

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 986 994**

51 Int. Cl.:

C12Q 1/66 (2006.01)

C07K 16/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.05.2016** **E 21209960 (0)**

97 Fecha y número de publicación de la concesión europea: **14.08.2024** **EP 3995587**

54 Título: **Proteínas de fusión recombinantes y bibliotecas de repertorios de células inmunitarias**

30 Prioridad:

09.06.2015 US 201562173224 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
13.11.2024

73 Titular/es:

**GIGAGEN, INC. (100.0%)
75 Shoreway Road, Suite A
San Carlos, CA 94070, US**

72 Inventor/es:

**JOHNSON, DAVID, SCOTT;
ADLER, ADAM y
MIZRAHI, RENA**

74 Agente/Representante:

ILLESCAS TABOADA, Manuel

Observaciones:

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

ES 2 986 994 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

Proteínas de fusión recombinantes y bibliotecas de repertorios de células inmunitarias

Campo

La invención está definida por las reivindicaciones. La presente divulgación se refiere, en general, a métodos y a composiciones para modificación por ingeniería de proteínas para su uso en aplicaciones biomédicas.

Antecedentes de la invención

Las compañías farmacéuticas usan cada vez más las terapias con anticuerpos para tratar enfermedades intratables como el cáncer (Carter 2006 Nature Reviews Immunology 6: 343-357). Sin embargo, el proceso de descubrimiento de fármacos de anticuerpos es costoso y tedioso, y se ha procedido mediante la identificación de un antígeno y luego el aislamiento y la producción de anticuerpos con actividad contra el antígeno. Además, la selección por afinidad y la expresión de un número limitado de anticuerpos a partir de esta selección han proporcionado un mecanismo de tratamiento más limitado que el proporcionado por el organismo. Sin embargo, no se ha logrado la generación artificial de un repertorio inmunitario representativo a partir de un individuo con receptores de células T o inmunoglobulina de cadena pesada y ligera apareada análoga.

Las personas que han estado expuestas a enfermedades producen naturalmente anticuerpos contra antígenos asociados con esa enfermedad. Por tanto, lo que se necesita son métodos mejorados de generación de alto rendimiento de proteínas de fusión recombinantes que comprendan dominios variables de cadena ligera y pesada para que sea posible usar repertorios inmunitarios naturales en el tratamiento y el descubrimiento y desarrollo farmacéutico.

Sumario de la invención

La presente invención está definida por las reivindicaciones. Por consiguiente, la presente invención se refiere a un método para preparar una pluralidad de constructos de expresión de inmunoglobulinas recombinantes, que comprenden:

proporcionar una pluralidad de polinucleótidos de fusión recombinantes que comprenden, cada uno, un primer polinucleótido que codifica para un primer dominio variable, un segundo polinucleótido que codifica para un segundo dominio variable, y un polinucleótido ligador que liga los polinucleótidos primero y segundo;

expresar una pluralidad de polipéptidos de fusión recombinantes codificados por dicha pluralidad de polinucleótidos de fusión recombinantes para presentar dicha pluralidad de polipéptidos de fusión recombinantes en una pluralidad de superficies;

enriquecer dicha pluralidad de polinucleótidos de fusión recombinantes para unirse a un antígeno exponiendo dicha pluralidad de superficies a dicho antígeno y seleccionar en base a la unión entre dicho antígeno y cada uno de dicha pluralidad de polipéptidos de fusión recombinantes;

circularizar mediante ensamblaje de Gibson uno o más de dichos polinucleótidos de fusión recombinantes enriquecidos; e

insertar un tercer polinucleótido que comprende una secuencia que codifica para un promotor y una secuencia que codifica para una región constante entre dichos polinucleótidos primero y segundo en cada uno de dichos polinucleótidos de fusión recombinantes enriquecidos circularizados, generando de ese modo una pluralidad de constructos de expresión de inmunoglobulinas recombinantes.

Según una realización preferida, el método comprende además las etapas anteriores de

proporcionar células inmunitarias primarias obtenidas de al menos un donante de mamífero; y encapsular dichas células inmunitarias primarias en una pluralidad de gotas con perlas que comprenden sondas de ácido polinucleico unidas con secuencias que son complementarias a ambas, polinucleótidos que comprenden secuencias que codifican para dominios variables de cadena pesada y polinucleótidos que comprenden secuencias que codifican para dominios variables de cadena ligera en una mezcla de reacción que lisa las células inmunitarias primarias y estimula la hibridación de ácido polinucleico, permitiendo así que las perlas capturen polinucleótidos que codifican para dominios variables de cadena pesada y polinucleótidos que codifican para dominios variables de cadena ligera.

Según una realización preferida adicional, el método comprende además las etapas anteriores de

insertar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en una pluralidad de células hospedadoras; y

expresar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en dicha pluralidad de células hospedadoras, generando de ese modo una biblioteca de inmunoglobulinas recombinantes, en la que cada inmunoglobulina recombinante comprende un par análogo de dominios variables de cadena pesada y de cadena ligera ligados de una célula individual aislada.

Según una realización preferida, dichos dominios variables primero y segundo son de una célula individual.

Según otra realización preferida adicional, dicho primer dominio variable es de una cadena pesada de inmunoglobulina y dicho segundo dominio variable es de una cadena ligera de inmunoglobulina.

Según una realización preferida adicional, dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes comprenden al menos 1.000, 10.000 ó 100.000 pares análogos únicos de secuencias codificantes de cadena pesada y de cadena ligera.

La presente invención también se refiere a un método para generar una biblioteca de inmunoglobulinas, que comprende:

generar una pluralidad de polinucleótidos de fusión recombinantes que comprenden, cada uno, un primer polinucleótido que codifica para un primer dominio variable y un segundo polinucleótido que codifica para un segundo dominio variable, en el que dicho primer dominio variable y dicho segundo dominio variable de dicha pluralidad de polinucleótidos de fusión recombinantes son un par análogo de una célula inmunitaria aislada individual, en el que dicha pluralidad de polinucleótidos de fusión recombinantes comprenden además un polinucleótido ligador que liga dichos polinucleótidos primero y segundo;

circularizar mediante ensamblaje de Gibson cada una de dicha pluralidad de los polinucleótidos de fusión recombinantes;

insertar un tercer polinucleótido que comprende una secuencia que codifica para un promotor y una secuencia que codifica para una región constante entre dichos polinucleótidos primero y segundo en cada uno de dichos polinucleótidos de fusión recombinantes circularizados, generando de ese modo una pluralidad de constructos de expresión de inmunoglobulinas recombinantes;

insertar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en una pluralidad de células hospedadoras; y

expresar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en dicha pluralidad de células hospedadoras, generando de ese modo una biblioteca de inmunoglobulinas recombinantes que comprende una pluralidad de inmunoglobulinas recombinantes, en el que cada una de dicha pluralidad de inmunoglobulinas recombinantes comprende un par análogo de dominios variables de cadena pesada y de cadena ligera de una célula individual aislada.

Según una realización preferida, dicha inserción de dicho tercer polinucleótido se realiza en paralelo para dicha pluralidad de constructos de polinucleótido circularizados.

Según una realización preferida adicional, al menos una inmunoglobulina recombinante de dicha biblioteca de inmunoglobulinas recombinantes se une específicamente a un epítipo de *Streptococcus pneumonia*, un epítipo de *Hemophilus influenza* o un epítipo de *Klebsiella pneumoniae*.

Según aún una realización preferida adicional, dicho primer dominio variable es de una cadena pesada de inmunoglobulina y dicho segundo dominio variable es de una cadena ligera de inmunoglobulina.

Breve descripción de los dibujos

Los objetos, características y ventajas anteriores y otros resultarán evidentes a partir de la siguiente descripción de implementaciones particulares de la divulgación, tal como se ilustra en los dibujos adjuntos en los que los mismos caracteres de referencia se refieren a las mismas partes a lo largo de las diferentes vistas. Los dibujos no están necesariamente a escala, sino que se hace hincapié en ilustrar los principios de diversas realizaciones de la invención.

Las figuras 1A-D muestran un método para generar un constructo ligado inicial usando sondas dirigidas a genes encapsuladas con ácidos nucleicos de células que expresan genes individuales dentro de gotas monodispersas.

La figura 2 muestra un constructo de expresión de scFv que comprende un polinucleótido que codifica para dominios variables de cadena pesada y de cadena ligera ligados de una célula individual.

La figura 3 muestra un método para generar un constructo de expresión de inmunoglobulina a partir de polinucleótidos que codifican para el dominio variable ligado inicial de una célula individual.

La figura 4 muestra un método para seleccionar bibliotecas de células que expresan fragmentos de anticuerpo con afinidad por un antígeno de interés.

La figura 5 muestra una inmunotransferencia de tipo Western que muestra la expresión de fragmentos de anticuerpo scFv en levadura a lo largo del tiempo. Los scFv se marcan con etiqueta con el marcador peptídico c-myc. Los carriles 1-6 se inducen usando el promotor AOX1 de tipo silvestre a las 72 h, 66 h, 47 h, 24 h y 18 h, respectivamente. Los carriles 7-12 son los mismos puntos de tiempo que los carriles 1-6, respectivamente, pero usando un promotor diferente. El carril 13 es un marcador de tamaño.

La figura 6 muestra una inmunotransferencia de tipo Western que muestra la expresión de anticuerpos de longitud completa en levadura. El carril 1 es un marcador de tamaño. Los carriles 2-5 son anticuerpos expresados en levadura, el carril 6 es anticuerpo expresado en CHO, todos en condiciones reducidas. Los carriles 7-10 son los mismos anticuerpos expresados en levadura, el carril 11 es el anticuerpo expresado en CHO, en condiciones no reducidas.

Descripción detallada

Los detalles de diversas realizaciones de la invención se exponen en la siguiente descripción. Otras características, objetos y ventajas de la invención resultarán evidentes a partir de la descripción y los dibujos, así como de las reivindicaciones.

Definiciones

Se entenderá que los siguientes términos, a menos que se indiquen de otro modo, tienen los siguientes significados:

Tal como se usa en el presente documento, el término "recombinante" se refiere a una biomolécula, por ejemplo, un gen o una proteína, que (1) se ha retirado de su entorno natural, (2) no está asociada con la totalidad o una parte de un polinucleótido en el que el gen se encuentra en la naturaleza, (3) está unido operativamente a un polinucleótido al que no se liga en la naturaleza, o (4) no se encuentra en la naturaleza. El término "recombinante" puede usarse en referencia a aislados de ADN clonados, análogos de polinucleótidos sintetizados químicamente o análogos de polinucleótidos que se sintetizan biológicamente por sistemas heterólogos, así como proteínas y/o ARNm codificados por tales ácidos nucleicos.

Tal como se usa en el presente documento, el término "ácido nucleico" se refiere a cualquier material que se compone de ADN o ARN. Los ácidos nucleicos pueden producirse de manera sintética o por células vivas.

Tal como se usa en el presente documento, el término "polinucleótido" se refiere a una cadena polimérica de nucleótidos. El término incluye moléculas de ADN (por ejemplo, ADNc o ADN genómico o sintético) y moléculas de ARN (por ejemplo, ARNm o ARN sintético), así como análogos de ADN o ARN que contienen análogos de nucleótidos no naturales, enlaces internucleosídicos no nativos, o ambos. El ácido nucleico puede estar en cualquier conformación topológica. Por ejemplo, el ácido nucleico puede ser monocatenario, bicatenario, tricatenario, cuádruplexado, parcialmente bicatenario, ramificado, en forma de horquilla, circular o en una conformación inmovilizada de candado.

A menos que se indique de otro modo, y como ejemplo de todas las secuencias descritas en el presente documento con el formato general "SEQ ID NO:", "ácido nucleico que comprende la SEQ ID NO: 1" se refiere a un ácido nucleico, al menos una parte del cual tiene (i) la secuencia de SEQ ID NO:1, o (ii) una secuencia complementaria a SEQ ID NO:1. La elección entre ambas está dictada por el contexto. Por ejemplo, si el ácido nucleico se usa como sonda, la elección entre ambas está dictada por el requisito de que la sonda sea complementaria a la diana deseado.

Tal como se usa en el presente documento, el término "proteína" se refiere a moléculas biológicas grandes, o macromoléculas, que consisten en una o más cadenas de residuos de aminoácido. Muchas proteínas son enzimas que catalizan reacciones bioquímicas y son vitales para el metabolismo. Las proteínas también tienen funciones estructurales o mecánicas, tales como actina y miosina en el músculo y las proteínas en el citoesqueleto, que forman un sistema de andamiaje que mantiene la forma celular. Otras proteínas son importantes en la señalización celular, las respuestas inmunitarias, la adhesión celular y el ciclo celular. Sin embargo, las proteínas pueden ser completamente artificiales o recombinantes, es decir, no existir de manera natural en un sistema biológico.

Tal como se usa en el presente documento, el término "polipéptido" se refiere tanto a proteínas que se producen de manera natural como a las que se producen de manera no natural, y a fragmentos, mutantes, derivados y análogos de las mismas. Un polipéptido puede ser monomérico o polimérico. Un polipéptido puede comprender varios dominios diferentes, cada uno de los cuales tiene una o más actividades distintas.

Tal como se usa en este documento, el término “antígeno” se refiere a una biomolécula que se une específicamente al anticuerpo respectivo. Un anticuerpo del repertorio diverso se une a una estructura antigénica específica por medio de su interacción de región variable (bucles CDR), siendo una analogía el ajuste entre una cerradura y una llave.

Tal como se usa en el presente documento, el término “anticuerpo” (es decir, “inmunoglobulina” (Ig)) se refiere a un polipéptido, al menos una parte del cual está codificada por al menos un gen de inmunoglobulina, o fragmento del mismo, y que puede unirse específicamente a una molécula diana deseada. El término incluye formas que se producen de manera natural, así como fragmentos y derivados. El anticuerpo reconoce una parte única de una diana foránea, denominada antígeno. Cada punta de la “Y” de un anticuerpo contiene un parátipo (una estructura análoga a una cerradura) que es específico para un epítipo particular (análogo de manera similar a una llave) en un antígeno, lo que permite que estas dos estructuras se unan con precisión. Mediante este mecanismo de unión, un anticuerpo puede marcar con una etiqueta un microbio o una célula infectada para que otras partes del sistema inmunitario lo ataquen, o puede neutralizar su diana directamente (por ejemplo, bloqueando una parte de un microbio que es esencial para su invasión y supervivencia). El anticuerpo consiste en cuatro cadenas de polipéptido; dos “cadenas pesadas” idénticas y dos “cadenas ligeras” idénticas conectadas por enlaces disulfuro. La cadena pesada es más larga que la cadena ligera y tiene dos regiones, la región constante y la región variable. La región constante es idéntica en todos los anticuerpos del mismo isotipo, pero difiere en los anticuerpos de isotipos diferentes. Una cadena ligera tiene dos dominios sucesivos: un dominio constante y un dominio variable. La longitud aproximada de una cadena ligera es de 211 a 217 aminoácidos. Cada anticuerpo contiene dos cadenas ligeras que son siempre idénticas; sólo está presente un tipo de cadena ligera, κ o λ , por anticuerpo en mamíferos.

Los fragmentos dentro del alcance del término “anticuerpo” incluyen los producidos mediante digestión con diversas proteasas, los producidos mediante escisión química y/o disociación química y los producidos de manera recombinante, siempre que el fragmento siga pudiendo unirse específicamente a una molécula diana. Entre tales fragmentos están Fab, Fab', Fv, F(ab')₂ y fragmentos Fv de cadena sencilla (scFv).

Los derivados dentro del alcance del término incluyen anticuerpos (o fragmentos de los mismos) que se han modificado en la secuencia, pero siguen siendo que pueden unirse específicamente a una molécula diana, incluyendo: anticuerpos quiméricos y humanizados interespecie; fusiones de anticuerpos; complejos de anticuerpos y fusiones de anticuerpos heteroméricos, tales como diacuerpos (anticuerpos biespecíficos), diacuerpos de cadena sencilla e intracuerpos (véase, por ejemplo, *Intracellular Antibodies: Research and Disease Applications* (1998) Marasco, ed., Springer-Verlag New York, Inc.)).

Los anticuerpos pueden producirse mediante cualquier técnica conocida, incluida la recogida a partir de cultivos celulares de linfocitos B nativos, la recogida a partir de cultivos de hibridomas, sistemas de expresión recombinante y la presentación en fagos. Los anticuerpos recombinantes también pueden producirse mediante los métodos que se dan a conocer en el presente documento.

Tal como se usa en el presente documento, el término “receptor de células B” o “BCR” se refiere a una proteína receptora transmembrana ubicada en la superficie exterior de las células B. El resto de unión del receptor de células B se compone de un anticuerpo unido a membrana que, como todos los anticuerpos, tiene un sitio de unión a antígeno único y determinado aleatoriamente que comprende dominios variables.

Tal como se usa en el presente documento, el término “región” se refiere a una parte contigua físicamente de la estructura primaria de una biomolécula. En el caso de proteínas, una región se define por una parte contigua de la secuencia de aminoácidos de esa proteína.

Tal como se usa en el presente documento, el término “dominio” se refiere a una estructura de una biomolécula que contribuye a una función conocida o sospechada de la biomolécula. Los dominios pueden ser coextensivos con regiones o partes de los mismos; los dominios también pueden incluir regiones distintas no contiguas de una biomolécula. Los ejemplos de dominios de proteína incluyen, pero no se limitan a, un dominio constante, un dominio variable, un dominio de cadena ligera y un dominio de cadena pesada.

Tal como se usa en el presente documento, el término “región variable” o “dominio variable” se refiere a la región de unión a antígeno que es muy variable. Esta variabilidad proporciona estructuras de punta o sitios de unión a antígeno ligeramente diferentes, lo que permite que existan millones de anticuerpos con estructuras de punta o sitios de unión a antígeno ligeramente diferentes. Cada una de estas variantes puede unirse a un antígeno diferente. Esta enorme diversidad de anticuerpos permite que el sistema inmunitario reconozca una variedad igualmente amplia de antígenos. La gran población diversa de anticuerpos se genera mediante combinaciones aleatorias de un conjunto de segmentos de genes que codifican para diferentes sitios de unión a antígeno (o parátipos), seguidas de mutaciones aleatorias en esta zona del gen del anticuerpo, lo que crea una mayor diversidad. Las moléculas de receptor de células T también tienen regiones variables o dominios variables.

Cada cadena pesada o ligera de anticuerpo tiene dos regiones, la región constante y la región variable. El “dominio

constante" es idéntico en todos los anticuerpos del mismo isotipo, pero difiere en anticuerpos de isotipos diferentes. Los isotipos de cadena pesada incluyen IgG1, IgG2, IgE e IgA. Los isotipos diferentes tienen funciones "efectoras" particulares en el sistema inmunitario, es decir, activan rutas biológicas particulares para la inmunidad. Las moléculas de receptor de células T también tienen dominios constantes.

Tal como se usa en el presente documento, el término "fragmento variable de cadena sencilla" (scFv) se refiere a un fragmento de anticuerpo de cadena sencilla que se compone de una cadena pesada y una cadena ligera ligadas por un ligador peptídico. En algunos casos, se expresan scFv en la superficie de una célula modificada por ingeniería, con el fin de seleccionar scFv particulares que se unen a un antígeno de interés.

Tal como se usa en el presente documento, el término "polipéptido de fusión" o "proteína de fusión" se refiere a un polipéptido que comprende un fragmento de polipéptido o proteína acoplado a secuencias de aminoácidos heterólogas. Las proteínas de fusión son útiles porque pueden construirse para contener dos o más elementos funcionales deseados de una o más proteínas. Una proteína de fusión comprende al menos 10 aminoácidos contiguos de un polipéptido de interés, al menos 20 ó 30 aminoácidos, al menos 40, 50 ó 60 aminoácidos, o al menos 75, 100 ó 125 aminoácidos. Las proteínas de fusión pueden producirse de manera recombinante construyendo una secuencia de ácido nucleico que codifica para el polipéptido o un fragmento del mismo en marco con una secuencia de ácido nucleico que codifica para una proteína o un péptido diferente y luego expresando la proteína de fusión. Alternativamente, una proteína de fusión puede producirse químicamente reticulando el polipéptido o un fragmento del mismo con otra proteína.

Tal como se usa en el presente documento, el término "modulador transcripcional" se refiere a cualquier secuencia de polinucleótido que modula la transcripción en una ubicación en cis o trans con respecto al modulador transcripcional. Normalmente, el mecanismo del modulador transcripcional es efectuar la unión de una proteína de factor de transcripción, o complejo de proteínas. Los ejemplos de moduladores transcripcionales incluyen, pero no se limitan a, promotores transcripcionales (es decir, "promotores") y potenciadores transcripcionales (es decir, "potenciadores"). Los promotores transcripcionales son secuencias en el extremo 5' de genes que modulan la expresión de genes. Los potenciadores transcripcionales son secuencias de polinucleótido, normalmente en cis con respecto al gen modulado, que efectúan la unión de factores de transcripción.

Tal como se usa en el presente documento, el término "modificador traduccional" o "modulador traduccional" se refiere a una secuencia de polinucleótido o de polipéptido que modula la traducción de un transcrito génico, tal como un sitio interno de entrada al ribosoma (IRES).

Tal como se usa en el presente documento, el término "efector inmunitario" se refiere a una secuencia de péptido o proteína que se une selectivamente a otra proteína y regula de ese modo la actividad inmunitaria. La actividad inmunitaria se regula a través de un aumento o una disminución de la actividad enzimática, la expresión génica o la señalización celular, de tal manera que las células inmunitarias se activan, desactivan, se hace que se dividan o se hace que no se dividan. En algunos casos, se secretan moléculas efectoras, tales como citocinas y anticuerpos. En otros casos, las moléculas efectoras se fijan a la superficie de una célula.

Tal como se usa en el presente documento, el término "constructo de expresión de proteína" se refiere a cualquier secuencia de ácido polinucleico que puede usarse para expresar una proteína recombinante. Un constructo de expresión de proteína contiene, como mínimo, al menos una región que codifica para proteína, un sitio de iniciación de la transcripción y una secuencia promotora que modula la transcripción. Muchos constructos de expresión de proteína son circulares, tales como un plásmido o un fagémido. Muchos constructos de expresión de proteína albergan un origen de replicación (ORI) que permite la autorreplicación del constructo en el interior de una célula hospedadora. Otros constructos de expresión de proteína son lineales y no albergan un ORI. Algunos constructos de expresión de proteína son explícitamente para la expresión en una célula modificada por ingeniería, mientras que otros constructos de expresión de proteína son para generar proteínas *in vitro* sin el beneficio de la maquinaria celular. Los ejemplos de células modificadas por ingeniería incluyen células de bacterias, levaduras y mamíferos (por ejemplo, células de ovario de hámster chino o HeLa).

Tal como se usa en el presente documento, el término "vector de expresión" se refiere a una molécula de ácido nucleico que puede introducir un constructo de expresión de proteína en una célula hospedadora. Algunos vectores de expresión también se replican en el interior de células hospedadoras, lo que aumenta la expresión de proteína por el constructo de expresión de proteína. Un tipo de vector es un "plásmido", que se refiere a un bucle circular de ADN bicatenario en el que pueden ligarse segmentos de ADN adicionales. Otros vectores incluyen cósmidos, cromosomas artificiales bacterianos (BAC) y cromosomas artificiales de levadura (YAC), fósmodos, fagos y fagémidos. Otro tipo de vector es un vector viral, en el que pueden ligarse segmentos de ADN adicionales en el genoma viral (comentado con más detalle a continuación). Determinados vectores pueden dar replicación autónoma en una célula hospedadora en la que se introducen (por ejemplo, vectores que tienen un origen de replicación que funciona en la célula hospedadora). Otros vectores pueden integrarse en el genoma de una célula hospedadora tras la introducción en la célula hospedadora y, por tanto, se replican de ese modo junto con el genoma hospedador. Además, determinados vectores preferidos pueden dirigir la expresión de genes a los que están unidos operativamente. Tales vectores se denominan en el presente documento "vectores de expresión recombinantes" (o

simplemente “vectores de expresión”).

Tal como se usa en el presente documento, el término “sonda de polinucleótido” (o “sonda de ácido nucleico”) se refiere a cualquier secuencia de ácido nucleico que sea complementaria o se una en condiciones rigurosas a una secuencia de ácido nucleico diana. La sonda de polinucleótido puede usarse para detectar, capturar y/o amplificar esa secuencia de ácido nucleico diana. Las sondas de polinucleótido incluyen, pero no se limitan a, estructuras de origami de ADN que incluyen 10-5000 componentes de oligonucleótido individuales. Un ejemplo de una sonda de polinucleótido es un cebador usado en la amplificación por PCR.

Tal como se usa en el presente documento, el término “condiciones rigurosas” se refiere a las condiciones en las que un compuesto de la divulgación hibridará con su secuencia diana, pero con un número mínimo de otras secuencias. Las condiciones rigurosas dependen de la secuencia y serán diferentes en diferentes circunstancias y en el contexto de esta divulgación, las “condiciones rigurosas” en las cuales los compuestos oligoméricos hibridan con una secuencia diana están determinadas por la naturaleza y composición de los compuestos oligoméricos y los ensayos en los que se investigan.

Tal como se usa en el presente documento, el término “ácido nucleico fusionado” o “polinucleótido de fusión recombinante” se refiere a una fusión de dos secuencias de ácido nucleico en una única molécula de ácido nucleico contigua. La fusión puede lograrse a través de métodos moleculares que fusionan ácidos nucleicos, como la PCR de extensión por solapamiento o ligamiento.

Tal como se usa en el presente documento, el término “recipiente de reacción” se refiere a cualquier entidad que proporcione la separación física de una reacción en compartimentos independientes. El recipiente de reacción puede usarse, por ejemplo, para cribar o realizar reacciones de una célula, subpoblación celular o diana de ácido nucleico particular con exclusión de otros. Los recipientes de reacción pueden componerse de un compartimento de plástico, una cámara microfluídica o una gota, por ejemplo de una disolución de reacción acuosa.

Tal como se usa en el presente documento, el término “tensioactivo” se refiere a compuestos que reducen la tensión superficial (o tensión interfacial) entre dos líquidos o entre un líquido y un sólido. Los tensioactivos suelen ser compuestos orgánicos que son anfífilos, lo que significa que contienen grupos hidrófobos (sus colas) y grupos hidrófilos (sus cabezas). Por tanto, un tensioactivo contiene tanto un componente insoluble en agua (o soluble en aceite) como un componente soluble en agua. Los tensioactivos difundirán en el agua y se adsorberán en las interfases entre el aire y el agua o en la interfase entre el aceite y el agua, en el caso de que el agua se mezcle con aceite. El grupo hidrófobo insoluble en agua puede extenderse fuera de la fase acuosa en volumen, hacia el aire o hacia la fase de aceite, mientras que el grupo de cabeza soluble en agua permanece en la fase acuosa. Los tensioactivos pueden actuar como detergentes, agentes humectantes, emulsionantes, agentes espumantes y dispersantes.

Tal como se usa en el presente documento, el término “acuosa” se refiere a una disolución que contiene agua, normalmente como disolvente o medio.

Tal como se usa en el presente documento, el término “aceite” se refiere a cualquier sustancia química neutra y apolar que es un líquido viscoso a temperatura ambiente y que es tanto hidrófobo (inmiscible con agua, literalmente “temeroso del agua”) como lipófilo (miscible con otros aceites, literalmente “amante de la grasa”).

Tal como se usa en el presente documento, el término “gota” se refiere a una pequeña cantidad de líquido. Las gotas suelen ser esféricas, pero pueden componerse de lingotes cilíndricos que abarcan todo el diámetro de un canal microfluídico. Las gotas pueden formarse en aire, aceite o disoluciones acuosas, dependiendo de su composición de materia y el método de formación. Las gotas aparecen tanto en poblaciones monodispersas como polidispersas.

Tal como se usa en el presente documento, el término “monodisperso” se refiere a una propiedad de los componentes caracterizada por un tamaño uniforme o casi uniforme. Por ejemplo, las gotas monodispersas normalmente requieren una dispersión de tamaño <5% para >90% de las gotas en una mezcla. En muchos casos, las poblaciones de gotas monodispersas son más estables que las poblaciones de gotas que no son monodispersas, es decir, las poblaciones de gotas polidispersas. En algunas realizaciones, la generación de gotas monodispersas requiere algún tipo de dispositivo microfluídico controlado.

Tal como se usa en el presente documento, el término “célula” se refiere a la unidad más pequeña de un organismo que puede replicarse de manera independiente. Las células son normalmente microscópicas y tienen un citoplasma y un núcleo encerrados en una membrana, ya sea de un organismo unicelular o derivado de un organismo multicelular.

Tal como se usa en el presente documento, el término “población celular” o “subpoblación celular” se refiere a al menos dos células de clase o clasificación similar. Por ejemplo, una subpoblación celular pueden ser dos células separadas de una población celular encapsulando las dos células en una gota.

Tal como se usa en el presente documento, el término “lisis” se refiere al proceso de romper la membrana celular de una célula o células a través de medios físicos o químicos. La lisis puede lograrse a través de un tensioactivo químico tal como Triton X-100, un tampón de lisis alcalino, calor, corrientes eléctricas o perturbación física.

Tal como se usa en el presente documento, el término “origami de ADN” se refiere al plegado a escala nanométrica del ADN para crear formas bidimensionales y tridimensionales arbitrarias a escala nanométrica. La especificidad de las interacciones entre pares de bases complementarias hace del ADN un material de construcción útil, a través del diseño de sus secuencias de bases. El origami de ADN puede ser sondas de polinucleótido plegadas en partículas.

Tal como se usa en el presente documento, el término “partícula” se refiere a cualquier sustancia sólida que se añade a una disolución compleja de material biológico (es decir, una proteína o un ácido polinucleico) para capturar y luego aislar físicamente un material biológico diana. Las partículas pueden ser objetos tales como perlas microscópicas compuestas por materiales tales como látex, vidrio o sílice, cuyo tamaño oscila entre 0,1 micrómetros y 1 mm. Las partículas también pueden referirse al ADN doblado a escala nanométrica del origami de ADN.

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entienden habitualmente los expertos en la técnica a la que pertenece esta invención. Se describen a continuación materiales y métodos a modo de ejemplo. En caso de conflicto, prevalecerá la presente memoria descriptiva, incluidas las definiciones. Los materiales, métodos y ejemplos son sólo ilustrativos y no pretenden ser limitativos.

A lo largo de esta memoria descriptiva y las reivindicaciones, se entenderá que el término “comprenden” o variaciones tales como “comprende” o “que comprende”, implican la inclusión de un número entero o grupo de números enteros indicado, pero no la exclusión de cualquier otro número entero o grupo de números enteros.

Métodos para aislar células individuales y polinucleótidos diana

En algunas implementaciones, se aísla una célula individual o, alternativamente, una subpoblación de células para capturar ADN o ARN de cada célula o subpoblación de células. En una realización, las células proceden de una reserva heterogénea de células T o B. En algunas implementaciones, las células son células T o células B primarias. En algunas realizaciones, las células las proporciona una única persona.

En algunas implementaciones, se usa un dispositivo microfluídico para generar gotas de emulsión de células individuales. El dispositivo microfluídico expulsa células individuales en un tampón de reacción acuoso a una mezcla de aceite hidrófobo. El dispositivo puede crear miles de microgotas de emulsión por minuto. Después de crear las microgotas de emulsión, el dispositivo expulsa la mezcla de emulsión en un canal. La mezcla puede pipetearse o recogerse en un tubo de reacción convencional para termociclado.

Los dispositivos microfluídicos personalizados para el análisis de células individuales se fabrican de manera rutinaria en laboratorios académicos y comerciales (Kintses *et al.*, 2010 *Current Opinion in Chemical Biology* 14:548-555). Por ejemplo, pueden fabricarse chips con polidimetilsiloxano (PDMS), plástico, vidrio o cuarzo. En algunas realizaciones, el fluido se mueve a través de los chips mediante la acción de una bomba de presión o de jeringa. Las células individuales pueden manipularse incluso en chips microfluídicos programables usando un dispositivo de dielectroforesis personalizado (Hunt *et al.*, 2008 *Lab Chip* 8:81-87). En una implementación, se usa un chip de PDMS basado en presión que se compone de geometría de enfoque de flujo fabricado con tecnología de litografía blanda (Dolomite Microfluidics (Royston, R.U.)) (Anna *et al.*, 2003 *Applied Physics Letters* 82:364-366). El diseño de disolución madre puede generar normalmente 10.000 microgotas monodispersas acuosas en aceite por segundo en rangos de tamaño de 10-150 μm de diámetro. En algunas implementaciones, la fase hidrófoba consistirá en aceite fluorado que contiene una sal de amonio de carboxi-perfluoropoliéter, lo que garantiza condiciones óptimas para la biología molecular y disminuye la probabilidad de coalescencia de las gotas (Johnston *et al.*, 1996 *Science* 271: 624-626). Para medir la periodicidad del flujo de células y gotas, se graban imágenes a 50.000 fotogramas por segundo usando técnicas convencionales, tales como una cámara Phantom V7 o Fastec InLine (Abate *et al.*, 2009 *Lab Chip* 9:2628-31).

El sistema microfluídico puede optimizar el tamaño de microgota, la densidad celular de entrada, el diseño del chip y los parámetros de carga celular de modo que más del 98% de las gotas contengan una célula individual. Hay tres métodos habituales para lograr tales estadísticas: (i) dilución extrema de la disolución celular; (ii) selección fluorescente de gotas que contienen células individuales; y (iii) el dispositivo microfluídico usa una dilución celular extrema para controlar la tasa de coincidencias múltiples y la clasificación de células fluorescentes para reducir la tasa de negativos.

En algunas implementaciones, el flujo celular de entrada se alinea con la periodicidad de formación de gotas, de modo que más del 98% de las gotas contienen una célula individual (Edd *et al.*, 2008 *Lab Chip* 8:1262-1264; Abate *et al.*, 2009 *Lab Chip* 9:2628-31). En estos dispositivos microfluídicos, se fuerza una suspensión de células de alta densidad a través de un canal de alta relación de aspecto, de modo que el diámetro celular es una gran fracción de

la anchura del canal. El chip está diseñado con un microcanal rectangular de 27 μm x 52 μm que hace fluir las células en microgotas a $>10 \mu\text{l/min}$ (Edd *et al.*, 2008 Lab Chip 8:1262-1264). Se someten a prueba varias anchuras de canal de entrada y velocidades de flujo para llegar a una disolución óptima.

En determinadas implementaciones de la divulgación, se usan chips microfluídicos para aislar 10, 100, 1.000, 10.000, 100.000, 1 millón o mil millones de células individuales de una reserva heterogénea de células T o B. En algunas implementaciones, los métodos de la divulgación usan células individuales en recipientes de reacción, en lugar de gotas de emulsión. Los ejemplos de estos recipientes de reacción incluyen placas de 96 pocillos, tubos de 0,2 ml, tubos de 0,5 ml, tubos de 1,5 ml, placas de 384 pocillos, placas de 1536 pocillos, etc.

En algunas implementaciones de la divulgación, las células se encapsulan con reactivos de PCR. La mezcla de reactivos de PCR está diseñada específicamente para lisar las células, liberando así las dianas de ADN o ARN de interés. Los reactivos de PCR luego amplifican una pluralidad de dianas de ADN o ARN de interés. En algunas implementaciones de la divulgación, los productos de PCR son regiones variables de inmunoglobulina de cadena ligera y pesada ligadas. En algunas implementaciones de la divulgación, los productos de PCR son polinucleótidos que codifican para polipéptidos de scFv.

En alguna implementación de la divulgación, se encapsulan células individuales en gotas con perlas que comprenden sondas de ácido polinucleico unidas con secuencias que son complementarias a las dianas de ácido polinucleico de interés en las células individuales. En determinadas implementaciones, las sondas tienen una longitud de 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120 ó 130 nucleótidos. Las sondas son de ARN, ADN, receptor celular o inmunoglobulina. En algunas implementaciones, el ácido polinucleico es ácido nucleico bloqueado (LNA), ácido nucleico peptídico (PNA), ácido nucleico glicólico (GNA) o cualquier análogo de ácido nucleico. Las dianas de ácido polinucleico son o bien ADN o bien ARN. En algunas implementaciones de la divulgación, las sondas de ácido polinucleico seleccionan como diana la región constante de las sondas T que seleccionan como diana la región constante de IgK o IgG. La mezcla de reactivos presente en estas implementaciones que comprenden perlas está diseñada específicamente tanto para lisar las células como para fomentar la hibridación de poli(ácidos nucleicos), lo que permite que las perlas capturen dianas de ADN o ARN de interés.

En determinadas implementaciones, las sondas de ácido polinucleico para inmunoglobulina de cadena ligera y pesada se unen a perlas. En otras implementaciones, las sondas para la cadena pesada se unen a una reserva de perlas, y las sondas para la cadena ligera se unen a una segunda reserva de perlas, y luego ambas reservas de perlas se encapsulan en gotas con células individuales. En algunas implementaciones, las sondas de ácido polinucleico modificado en 5'-amino se unen a perlas de ácido carboxílico usando tampón de ácido 2-(N-morfolino)etanosulfónico (MES) (Kojima *et al.*, 2005, Nucleic Acids Research 33:e150). En otras implementaciones, las sondas de ácido polinucleico biotinilado se unen a perlas recubiertas con estreptavidina.

En algunas implementaciones, los métodos de la divulgación usan células individuales en contenedores de reacción, en lugar de gotas de emulsión. Los ejemplos de estos contenedores de reacción incluyen placas de 96 pocillos, tubos de 0,2 ml, tubos de 0,5 ml, tubos de 1,5 ml, placas de 384 pocillos, placas de 1536 pocillos, etc. También se usa una variedad de otros diseños de chips microfluídicos para aislar células individuales (Marcus *et al.*, 2006, Anal Chem 78:3084-3089).

En algunas implementaciones, la célula o subpoblación de células se añade a un contenedor de reacción junto con perlas que comprenden sondas de ácido polinucleico unidas y un tampón de lisis. El tampón de lisis lisa las células para permitir que las sondas de ácido polinucleico se unan a las dianas de ácido polinucleico de interés de la célula o células. Las perlas hibridadas con las dianas de ácido polinucleico se aíslan del tampón de lisis, como perlas individuales o subpoblaciones de perlas en recipientes de reacción, y se ponen en contacto con una mezcla de PCR para permitir la amplificación o la fusión de las dianas de ácido polinucleico.

En determinadas implementaciones de la divulgación, la fase acuosa de las emulsiones de gotas que contienen perlas y sus dianas unidas se recupera usando un disolvente tal como etil éter. Las perlas se aíslan en emulsiones con una mezcla de PCR de tal manera que, en promedio, las perlas individuales se aíslan en microgotas de emulsión individuales (DeKosky *et al.*, 2015, Nat Med 21:86-91). Pueden formarse emulsiones monodispersas en un chip microfluídico o pueden formarse emulsiones polidispersas usando una máquina tal como el sistema IKA Ultra-Turrax Tube Drive. Los reactivos de PCR amplifican una pluralidad de dianas de ADN o ARN de interés. En algunas implementaciones de la divulgación, los productos de PCR son regiones variables de inmunoglobulina de cadena ligera y pesada ligadas. En algunas implementaciones de la divulgación, los productos de PCR son scFv.

En determinadas implementaciones de la divulgación, las bibliotecas amplificadas de subunidades de regiones variables ligadas se convierten luego en bibliotecas de expresión de proteína usando métodos para el ligamiento o ensamblaje de Gibson. En determinadas implementaciones de la divulgación, las bibliotecas de expresión de proteína se expresan en un sistema de producción de proteína recombinante, tales como células de levadura o de mamífero.

Métodos para amplificar y ligar regiones variables

La PCR se usa para amplificar muchos tipos de secuencias incluyendo, pero sin limitarse a, SNP, repeticiones cortas en tándem (STR), dominios de proteína variables, regiones metiladas y regiones intergénicas. Se usan métodos de PCR de extensión por solapamiento para crear productos de amplicón de fusión de varios loci genómicos independientes en una reacción de un único tubo. Se dan a conocer métodos para amplificar y ligar regiones variables en Johnson *et al.*, 2005 *Genome Research* 15: 1315-24; patente estadounidense 7.749.697; publicación PCT WO 2012/083225; y publicación PCT WO 2013/096643.

En algunas implementaciones, se eligen en la célula al menos dos secuencias diana de ácido nucleico (por ejemplo, secuencias diana de ácido nucleico primera y segunda, o loci primero y segundo) y se designan como loci diana. Se diseñan cebadores directos e inversos para cada una de las dos secuencias diana de ácido nucleico, y los cebadores se usan para amplificar las secuencias diana. Se generan amplicones "minoritarios" amplificando las dos secuencias diana de ácido nucleico por separado y luego se fusionan mediante amplificación para crear un amplicón de fusión, también conocido como amplicón "mayoritario". En una implementación, un amplicón "minoritario" es una secuencia de ácido nucleico amplificada a partir de un loci genómico diana, y un amplicón "mayoritario" es un complejo de fusión generado a partir de secuencias amplificadas entre múltiples loci genómicos, por ejemplo, un nucleótido de fusión recombinante.

El método usa cebadores "internos" (es decir, el cebador inverso para el primer locus y el cebador directo para el segundo locus) que comprenden un dominio que hibrida con un amplicón minoritario y un segundo dominio que hibrida con un segundo amplicón minoritario. Los cebadores "internos" son un reactivo limitante, de tal manera que durante la fase exponencial de la PCR, los cebadores internos se agotan, lo que hace que los dominios solapantes en los amplicones minoritarios se apareen y creen amplicones mayoritarios.

Los cebadores de PCR se diseñan contra dianas de interés usando parámetros convencionales, es decir, temperatura de fusión (T_f) de aproximadamente 55-65°C, y con una longitud de 20-50 nucleótidos. Los cebadores se usan con condiciones de PCR convencionales, por ejemplo, Tris-HCl 1 mM pH 8,3, cloruro de potasio 5 mM, cloruro de magnesio 0,15 mM, cebadores 0,2-2 μ M, dNTP 200 μ M y una ADN polimerasa termoestable. Hay muchos kits comerciales disponibles para realizar PCR, tales como Platinum Taq (Life Technologies), AmpliTag Gold (Life Technologies), Titanium Taq (Clontech), polimerasa Phusion (Finnzymes), Hot-StartTaq Plus (Qiagen). Cualquier ADN polimerasa termoestable convencional puede usarse para esta etapa, tal como la Taq polimerasa o el fragmento de Stoffel.

En una implementación, se usa un conjunto de sondas de ácido nucleico (o cebadores) para amplificar una primera secuencia de ácido nucleico diana y una segunda secuencia de ácido nucleico diana para formar un complejo de fusión (figuras 1A-D). Tal como se muestra en la figura 1A, la primera sonda incluye una secuencia que es complementaria a una primera secuencia de ácido nucleico diana (por ejemplo, el extremo 5' de la primera secuencia de ácido nucleico diana). La segunda sonda incluye una secuencia que es complementaria a la primera secuencia de ácido nucleico diana (por ejemplo, el extremo 3' de la primera secuencia de ácido nucleico diana) y una segunda secuencia que es complementaria a una secuencia exógena. En algunas implementaciones, la secuencia exógena es una secuencia de ácido nucleico no humana y no es complementaria a ninguna de las secuencias de ácido nucleico diana. Por ejemplo, la secuencia exógena puede ser una secuencia de polinucleótido que codifica para una secuencia de polipéptido rica en aminoácidos Ser y Gly, que liga regiones variables de cadena ligera y pesada en un scFv (véase, por ejemplo, el documento PCT/US1992/001478). Las sondas primera y segunda son el cebador directo y el cebador inverso para la primera secuencia de ácido nucleico diana.

Tal como se muestra en la figura 1A, la tercera sonda incluye una secuencia que es complementaria a la parte de la segunda sonda que es complementaria a la secuencia exógena y una secuencia que es complementaria a la segunda secuencia de ácido nucleico diana (por ejemplo, el extremo 5' de la segunda secuencia de ácido nucleico diana). La cuarta sonda incluye una secuencia que es complementaria a la segunda secuencia de ácido nucleico diana (por ejemplo, el extremo 3' de la segunda secuencia de ácido nucleico diana). La tercera sonda y la cuarta sonda son los cebadores directo e inverso para la segunda secuencia de ácido nucleico diana.

Las sondas segunda y tercera también se denominan cebadores "internos" de la reacción (es decir, el cebador inverso para el primer locus y el cebador directo para el segundo locus) y tienen una concentración limitante (por ejemplo, 0,01 μ M para los cebadores internos y 0,1 μ M para todos los demás cebadores). Esto impulsará la amplificación del amplicón mayoritario preferentemente con respecto a los amplicones minoritarios. Las sondas primera y cuarta se denominan cebadores "externos".

Tal como se muestra en las figuras 1B y 1C, las secuencias de ácido nucleico primera y segunda se amplifican de manera independiente, de tal manera que la primera secuencia de ácido nucleico se amplifica usando la primera sonda y la segunda sonda, y la segunda secuencia de ácido nucleico se amplifica usando la tercera sonda y la cuarta sonda. A continuación, se genera un complejo de fusión hibridando las regiones de secuencias complementarias de las secuencias de ácido nucleico primera y segunda amplificadas y amplificando las secuencias hibridadas usando las sondas primera y cuarta (figura 1D). Esto se denomina amplificación por PCR de extensión

por solapamiento.

Durante la amplificación por PCR de extensión por solapamiento, las regiones de secuencias complementarias de las secuencias de ácido nucleico primera y segunda amplificadas actúan como cebadores para la extensión en ambas hebras y en cada sentido por moléculas de ADN polimerasa. En ciclos de PCR posteriores, los cebadores externos ceban la secuencia fusionada total de tal manera que la ADN polimerasa duplique el complejo fusionado. Este método produce una pluralidad de complejos de fusión. En algunas implementaciones, los complejos de fusión comprenden una región variable de cadena pesada y de cadena ligera de la misma célula o subpoblación de células. En algunas implementaciones, el complejo de fusión se usa como inserto de scFv.

Métodos para diseñar constructos de ácido polinucleico para expresión de proteínas

En algunas implementaciones, se ligan y se amplifican regiones variables para inmunoglobulina de cadena ligera y pesada, o se amplifican y se ligan, para formar una pluralidad de constructos de ácido polinucleico que comprenden una región variable de Ig de cadena pesada y ligera ligada por un polinucleótido ligador, para formar, por ejemplo, un polinucleótido que codifica para scFv. (figura 2). En otras implementaciones, se ligan y se amplifican regiones variables para el receptor de células T alfa y beta, o se amplifican y se ligan, para formar una pluralidad de constructos de ácido polinucleico que comprenden un receptor de células T alfa y beta conectados por un polinucleótido ligador. En determinadas implementaciones, las secuencias variables del receptor de células T o las secuencias de Ig se amplifican a partir de genes que se producen de manera natural, especialmente a partir de células individuales o poblaciones clonales de células. En otras realizaciones, las secuencias de receptor de células T variables o las secuencias de Ig se generan artificialmente usando metodologías de síntesis génica (Kosuri & Church, 2014, Nat Methods 11: 499-507). En otras implementaciones, la biblioteca de complejos ligados se genera a partir de secuencias completamente artificiales, es decir, una biblioteca de gran diversidad (por ejemplo, $\sim 10^{12}$ secuencias únicas) de secuencias de ADN aleatorizadas. En determinadas realizaciones, las regiones variables ligadas forman un scFv que puede expresarse en células procariontes o eucariotas hospedadoras como proteína de superficie o secretada recombinante. En algunas implementaciones, los scFv recombinantes se expresan usando la presentación ribosómica (Hanes *et al.*, 1997, PNAS 94: 4937-42) o la presentación de ARNm (Matheakis *et al.*, 1994, Affymax Research Institute 91:9022-6). En otras implementaciones de la presente divulgación, los constructos de polinucleótido de subunidad de región variable ligada nunca se expresan como proteína y, en su lugar, sirven como constructo de polinucleótido intermedio o biblioteca de constructos para la generación posterior de un constructo de polinucleótido o una biblioteca de constructos de polinucleótido que codifican para una proteína de longitud completa, o biblioteca de proteínas. En determinadas implementaciones de la divulgación, las bibliotecas se componen de una, decenas, centenas, miles, millones o miles de millones de secuencias únicas.

En algunas implementaciones de la divulgación, se circulariza el constructo de polinucleótido que comprende los polinucleótidos que codifican para el dominio variable de una célula individual aislada conectados por un polinucleótido ligador en un único polinucleótido (véase la figura 3, paneles superior y central). Por ejemplo, si el constructo original forma parte de una biblioteca de más de un constructo, entonces la circularización es beneficiosa para retener cualquier apareamiento entre las subunidades de polinucleótido amplificadas para formar la secuencia de polinucleótido ligada inicial. En una implementación, se crea una biblioteca de scFv a partir de miles de células B individuales. A menudo es deseable insertar en el polinucleótido ligador de scFv o reemplazar al menos una parte del polinucleótido ligador de scFv por un promotor y otros elementos de secuencia que son importantes para convertir la biblioteca de scFv en una biblioteca de anticuerpos de longitud completa usando los métodos de modificación por ingeniería de ADN descritos más adelante. Si el scFv lineal se escindiera en masa en el ligador, la biblioteca resultante ya no retendría los apareamientos nativos de región variable de cadena pesada y ligera de las células aisladas originales. Por tanto, es importante circularizar el constructo antes de la inserción de un promotor u otros elementos de secuencia en el constructo entre las secuencias que codifican para el dominio variable para mantener el apareamiento de las células aisladas originales (figura 3).

En algunas implementaciones, la biblioteca de scFv se circulariza usando primero endonucleasas de restricción para digerir un plásmido y los constructos de scFv, seguido del ligamiento de los constructos de scFv en el plásmido. En otras implementaciones, se usa un constructo de ADN lineal en lugar de un plásmido. Las enzimas de restricción cortan el ADN en o cerca de secuencias de nucleótidos de reconocimiento específicas conocidas como sitios de restricción. Se usan de manera rutinaria endonucleasas de restricción tales como EcoRI y NotI en biología molecular y están disponibles comercialmente de proveedores tales como New England Biolabs. En algunas realizaciones de la divulgación, la ADN ligasa se usa luego para ligar los "extremos cohesivos" de un plásmido con los extremos cohesivos de la biblioteca de scFv. En otra implementación de la divulgación, se usa una ADN ligasa de T4 para circularizar el scFv directamente, o insertando la biblioteca de scFv en un ADN lineal con extremos romos. En otras implementaciones de la divulgación, se usa el ensamblaje de Gibson para circularizar la biblioteca de scFv (Gibson *et al.*, 2008, Science 319: 1215-20). En el proceso de ensamblaje de Gibson, al menos dos constructos de ADN con extremos solapantes se mezclan en un tubo de reacción. Se usa una enzima 5'-exonucleasa para recortar los extremos 5' de los constructos. Luego, los constructos se aparean y se usa una ADN polimerasa para extender los extremos 3' del ADN. Finalmente, una ADN ligasa sella las mellas restantes en el ADN bicatenario. En otra implementación de la divulgación, la modificación por ingeniería de ADN se logra mediante recombinación homóloga

o unión de extremos no homólogos.

En algunas implementaciones de la divulgación, un plásmido circularizado se transforma en bacterias. En algunas implementaciones de la divulgación, el plásmido confiere resistencia a un antibiótico, y se usa la selección de las bacterias transformadas con medios que contienen antibiótico para generar una pluralidad de clones bacterianos que contienen el plásmido. Después de usar dichos métodos de modificación por ingeniería de ADN para alterar el ADN, se usa la secuenciación de ADN para verificar el constructo o la biblioteca de constructos. La secuenciación del ADN se realiza usando métodos paralelos masiva de proveedores tales como Illumina, o usando métodos de secuenciación de clones individuales tales como la secuenciación de Sanger de proveedores tales como Applied Biosystems.

En determinadas implementaciones, los métodos de modificación por ingeniería de ADN, tales como las endonucleasas de restricción y el ensamblaje de Gibson, tal como se describieron anteriormente, también se usan para insertar determinadas secuencias de ácido polinucleico en el constructo de ácido polinucleico circularizado ligado. En determinadas implementaciones, la secuencia insertada reemplaza parte o la totalidad del polinucleótido ligado que conecta las dos secuencias de polinucleótido que codifican para el dominio variable. En determinadas implementaciones de la divulgación, el ácido polinucleico insertado es necesario o beneficioso para la transcripción del constructo recombinante (por ejemplo, un promotor o potenciador) en células recombinantes (véase la figura 3, panel inferior). En otras implementaciones de la divulgación, el ácido polinucleico insertado codifica para segmentos de proteína importantes para una modalidad terapéutica, por ejemplo, secuencias efectoras inmunitarias. Estos segmentos que codifican para proteínas se insertan en el constructo inicial en marco, de tal manera que el constructo de polinucleótido resultante produce una proteína o proteínas en marco con el constructo original. En determinadas implementaciones de la divulgación, los componentes insertados se amplifican a partir de ADN genómico o ARNm. En otras implementaciones de la divulgación, los componentes insertados se sintetizan *in vitro* usando oligonucleótidos de ADN como material de partida. En determinadas implementaciones de la divulgación, las bibliotecas de scFv se componen de decenas, cientos, miles, millones o miles de millones de secuencias únicas, y los métodos de modificación por ingeniería de ADN se usan para insertar secuencias de ácido polinucleico en el constructo de ácido polinucleico ligado inicial en masa en un único tubo de reacción. Después de usar dichos métodos de modificación por ingeniería de ADN para alterar el ADN, se usa la secuenciación de ADN para verificar el constructo o la biblioteca de constructos. La secuenciación del ADN se realiza usando métodos paralelos masiva de proveedores tales como Illumina, o usando métodos de secuenciación de clones individuales tales como la secuenciación de Sanger de proveedores tales como Applied Biosystems.

En algunas implementaciones, el método de generación de la biblioteca de inmunoglobulinas o receptores de células T se realiza sin selección por afinidad. Los métodos de alto rendimiento descritos en el presente documento pueden usarse para generar una biblioteca de proteínas de unión a antígeno (por ejemplo, inmunoglobulinas) que comprende al menos 1.000, 10.000, 100.000 o más proteínas de unión a antígeno únicas, cada una de las cuales tiene secuencias variables apareadas a partir de una célula inmunitaria aislada individual (es decir, un par análogo). La selección por afinidad también puede usarse como una etapa adicional para generar una biblioteca de proteínas de unión a antígeno refinada.

Métodos para la expresión de proteína recombinante

Los scFv recombinantes se expresan de manera rutinaria en fagos, en un proceso denominado presentación en fago (Smith, 1985, Science 228:1315-17; McCafferty *et al.*, 1990, Nature 348:552-554). La presentación en fago es una técnica de laboratorio habitual para el estudio de las interacciones proteína-proteína, proteína-péptido y proteína-ADN que usa bacteriófagos (virus que infectan bacterias) para conectar las proteínas con la información genética que las codifica. Las aplicaciones de la tecnología de presentación en fago incluyen la determinación de las parejas de interacción de una proteína (que se usaría como el "cebo" de fago inmovilizado con una biblioteca de ADN que consiste en todas las secuencias codificantes de una célula, un tejido u organismo) de modo que pueda determinarse la función o el mecanismo de la función de esa proteína. La presentación en fago también es un método usado ampliamente para la evolución de proteínas *in vitro* (también denominada modificación por ingeniería de proteínas). Como tal, la presentación en fago es una herramienta útil en el descubrimiento de fármacos. Se usa para hallar nuevos ligandos (inhibidores de enzimas, agonistas y antagonistas de receptores) para proteínas diana (Lunder *et al.*, 2005, J Lipid Research 46:1512-1516; Bratkovic *et al.*, 2005, Biochem Biophys Res Commun 332: 897-903). La técnica también se usa para determinar antígenos tumorales (para su uso en diagnóstico y direccionamiento terapéutico) y en la búsqueda de interacciones proteína-ADN usando bibliotecas de ADN construidas especialmente con segmentos aleatorizados (Hufton *et al.*, 1999, J. Immunol Methods 231: 39-51; Gommans *et al.*, 2005, J Mol Biol 354: 507-519). Las bibliotecas de fagos modificados por ingeniería se componen de cientos, miles, millones o miles de millones de secuencias de scFv únicas. La gran diversidad de bibliotecas permite la detección de scFv con afinidad por un antígeno de interés de modo paralelo masiva.

Los scFv recombinantes se expresan de manera rutinaria en la superficie de células eucariotas tales como células de mamífero y levaduras. La ventaja de la presentación a la superficie celular es el uso de clasificación y análisis de citometría de flujo cuantitativa para identificar interacciones de alta afinidad y normalizar la expresión de proteínas de anticuerpo. Como mínimo, los constructos de expresión de polinucleótido para la expresión en superficie de scFv

incluyen secuencias de polinucleótido para un promotor transcripcional, una secuencia de región variable de cadena pesada, una secuencia ligadora de polipéptido y una secuencia de región variable de cadena ligera. La presentación en levadura es similar a la presentación en fago en el sentido de que un scFv recombinante se modifica por ingeniería en un constructo de expresión de polinucleótido y luego se transporta a la superficie de la levadura, usando una señal de tráfico de péptidos tal como Aga2 (Boder y Wittrup, 1997, Nat Biotech, 15:553-57). Las cepas de levadura usadas habitualmente para la expresión de proteína recombinante incluyen *Pichia pastoris* y *Saccharomyces cerevisiae*. En algunas implementaciones, la presentación en levadura se usa para el estudio de interacciones proteína-proteína, proteína-péptido y proteína-ADN. Las bibliotecas de levadura modificada por ingeniería se componen de cientos, miles, millones o miles de millones de secuencias de scFv únicas. La gran diversidad de bibliotecas permite el cribado de scFv con afinidad por un antígeno de interés de manera paralela masiva (Dangaj *et al.*, 2013, Cancer Res 73:4820-4829). En otras implementaciones, se usan células de mamífero para la expresión en superficie de scFv en lugar de levadura (Ho *et al.*, 2006, PNAS 103:25). En determinadas implementaciones, la expresión en la superficie de células de mamífero se produce al fusionar el scFv con la proteína CCR5, o el factor de crecimiento derivado de plaquetas (PDGF) (Urban *et al.*, 2005, Nucleic Acids Research 33:e35; Wolkowicz *et al.*, 2005, J Biol Chem 280:15195-15201). Las células de mamífero para la expresión de proteína recombinante incluyen células de ovario de hámster chino (CHO) (Anderson *et al.*, 2004, Curr Opin Biotechnol 15:456-462). El ADN recombinante se introduce en el genoma de la célula de mamífero usando un retrovirus, o se introduce de manera transitoria en las células de mamífero usando un plásmido de autorreplicación.

En determinadas implementaciones de la secuencia, se convierten complejos de polinucleótido que se componen de lg de cadena ligera y pesada ligadas (por ejemplo, scFv) en proteínas de anticuerpo de longitud completa para aplicaciones posteriores, tales como la terapia con anticuerpos. Tales aplicaciones requieren partes adicionales de las secuencias de anticuerpos que no están presentes en scFv y que serían difíciles de amplificar a partir de células individuales. Además, en determinadas implementaciones, debido a que los anticuerpos se componen de productos proteicos de dos genes (lg de cadena ligera y pesada), un constructo de expresión de anticuerpo de longitud completa requiere dos promotores, es decir, un promotor para cada cadena pesada y ligera (véase, por ejemplo, la figura 3, panel inferior). En otras implementaciones, el ligador de scFv se reemplaza por un sitio interno de entrada al ribosoma (IRES), que permite la expresión independiente de lg de cadena ligera y pesada. Los sistemas libres de células como la presentación ribosómica pueden producir bibliotecas de anticuerpos de gran diversidad (~de 10^{13} a 10^{14} anticuerpos únicos) (Hanes & Pluckthun, 1997, PNAS 94:4937-4942). Aunque los sistemas libres de células son útiles para muchas aplicaciones, los problemas con el plegamiento de proteínas, las modificaciones postraduccionales y el uso de codones limitan la utilidad de tales métodos para producir anticuerpos terapéuticos totalmente funcionales. En determinadas implementaciones de la divulgación, los complejos de polinucleótido que se componen de lg de cadena pesada y ligera ligadas se convierten en constructos de expresión de longitud completa usando los métodos descritos anteriormente. En otras implementaciones de la divulgación, los complejos de polinucleótido que se componen de subunidades de receptor de células T ligadas se convierten en constructos de expresión de longitud completa usando los métodos anteriores. En determinadas implementaciones de la divulgación, los constructos de expresión de longitud completa se usan para inducir la expresión de proteína recombinante en células tales como células de bacterias, levadura o mamífero. En determinadas implementaciones, la biblioteca de constructos de expresión de longitud completa se compone de decenas, cientos, miles, millones o miles de millones de proteínas diferentes. En determinadas implementaciones de la divulgación, la biblioteca completa de constructos de longitud completa se introduce en masa en células productoras de proteína recombinante, para producir una biblioteca celular que se compone de decenas, cientos, miles, millones o miles de millones de clones diferentes que pueden usarse para producir una biblioteca de proteínas que se compone de decenas, cientos, miles, millones o miles de millones de proteínas diferentes.

Los protocolos para la expresión de anticuerpos de longitud completa en células de mamífero se conocen bien, y los primeros anticuerpos monoclonales comerciales producidos en CHO llegaron al mercado en 1997 (Rituxan). Como mínimo, un constructo de expresión de polinucleótido de anticuerpo de longitud completa se compone de secuencias de polinucleótido que codifican para una proteína de cadena pesada de longitud completa, un promotor para la cadena pesada, una proteína de cadena ligera de longitud completa y un promotor para la cadena ligera. Los promotores para la expresión de anticuerpos incluyen el citomegalovirus humano (hCMV). Las células de mamífero se transfectan con los constructos de expresión de anticuerpos usando métodos tales como electroporación, precipitación con fosfato de calcio, lipofección y transfección retroviral. Estos métodos se usan para introducir una biblioteca de constructos de longitud completa en células productoras de proteína recombinante en masa, para producir una biblioteca celular que se compone de decenas, cientos, miles, millones o miles de millones de clones diferentes que pueden usarse para producir una biblioteca de proteínas que se compone de decenas, cientos, miles, millones o miles de millones de proteínas diferentes. Los vectores de expresión pueden incluir marcadores de resistencia frente a compuestos tales como la higromicina, que pueden usarse para seleccionar frente a células que no se han transfectado satisfactoriamente con el vector de expresión recombinante. Los cultivos de CHO se mantienen normalmente en matraces de agitación a 37°C en un 8% de CO₂, usando medios disponibles comercialmente de proveedores tales como GIBCO y HyClone. En determinadas implementaciones, las células hospedadoras son deficientes en enzimas metabólicas tales como la dihidrofolato reductasa (DHFR). Las células que son deficientes en DHFR son auxótrofos que requieren prolina, por lo que la transfección de células CHO deficientes en DHFR con vectores que contienen DHFR y luego el crecimiento de las células transfectadas en un medio deficiente en prolina puede ayudar a seleccionar clones que expresen los anticuerpos de longitud completa.

Por tanto, tanto la selección negativa (por ejemplo, higromicina) como la selección positiva (por ejemplo, deficiencia de DHFR rescatada) pueden usarse para generar grandes bibliotecas de clones de células de mamífero que expresan anticuerpos de longitud completa. En otras realizaciones, se generan transfectantes estables usando recombinación específica del sitio, por ejemplo, usando el sistema de modificación por ingeniería Cre/loxP (Kameyama *et al.*, 2010, *Biotechnol Bioeng* 105:1106-1114; Wiberg *et al.*, 2006, *Biotech Bioeng* 94:396-405). Tales métodos permiten niveles de expresión de proteína más predecibles en grandes bibliotecas de clones de células de mamífero. En otras implementaciones, se usa un sistema de expresión cromosómica artificial (ACE) para expresar proteínas recombinantes. El sistema ACE consiste en un cromosoma artificial basado en mamíferos conocido como plataforma ACE, un vector de direccionamiento de ACE (ATV) y una integrasa λ mutante (integrasa ACE) para la recombinación dirigida (Kennard *et al.*, 2009, *Biotechnol Bioeng* 104:540-553). La plataforma ACE consiste principalmente en genes ribosómicos repetidos en tándem y secuencias satélite repetitivas, que forman la heterocromatina pericentromérica. También tiene centrómeros y telómeros naturales para permitir la replicación del ADN sin necesidad de integración en el genoma de la célula hospedadora, lo que reduce la probabilidad de aberración cromosómica y heterogeneidad clonal.

En otras implementaciones de la divulgación, se usa levadura para producir proteína recombinante. Las cepas de levadura usadas habitualmente para la producción de proteína recombinante incluyen *Pichia pastoris* y *Saccharomyces cerevisiae*. En determinadas implementaciones, la producción de proteínas requiere modificaciones postraduccionales de proteínas que no se producen de manera natural en la levadura de tipo natural. En tales casos, es útil usar levaduras modificadas por ingeniería con glico, por ejemplo, la tecnología GlycoSwitch, que es una familia de cepas de *Pichia* que se modifican por ingeniería para tener una glicosilación postraducciona que es más "similar a la humana", por ejemplo, como Gal(2)GlcNAc(2)Man(3)GlcNAc(2)N-glicanos (Jacobs *et al.*, 2009, *Nat Protocol*, 4:58-70). La levadura se cultiva de manera rutinaria en medios tales como YPD (extracto de levadura al 1%, peptona al 2%, dextrosa al 2%), normalmente a 30°C. Muchos vectores diferentes están disponibles comercialmente, tal como la serie de vectores pPICZ de Life Technologies. Normalmente, estos vectores contienen resistencia a una sustancia química que puede usarse para seleccionar negativamente la levadura no transformada, tal como zeocina o kanomicina. Los constructos de polinucleótido o bibliotecas de constructos de polinucleótido se introducen de manera rutinaria en células de levadura usando un electroporador, aunque también se usan métodos de generación de esferoplastos, LiCl y polietilenglicol. En determinadas realizaciones, se usa el promotor AOX1 para inducir la expresión de proteínas (Cregg *et al.*, 2011, *Methods in Enzymology* 463:169-187). La secreción de proteína recombinante se dirige de manera rutinaria usando señales peptídicas tales como alfa-MF en el extremo NH₂-terminal de la proteína recombinante. En muchas implementaciones, es posible modificar por ingeniería levaduras para generar y secretar proteína recombinante a niveles de producción de hasta 10 g/l.

En determinadas implementaciones de la divulgación, es deseable expresar las proteínas en células T primarias, particularmente para células T modificadas con receptor de antígeno quimérico (CAR). En algunas implementaciones de la divulgación, el constructo de expresión de proteína se somete a transcripción *in vitro* para producir un ARNm. Estos ARNm se introducen luego en células T primarias mediante electroporación. En otras implementaciones de la divulgación, los constructos de expresión de proteína son retrovirus, que se transfectan en células T primarias, incorporando el constructo de expresión de proteína en el genoma.

Descubrimiento de fármacos de anticuerpos monoclonales

Las compañías farmacéuticas usan cada vez más las terapias con anticuerpos para tratar enfermedades intratables tales como cáncer (Carter 2006 *Nature Reviews Immunology* 6:343-357). Sin embargo, el proceso de descubrimiento de fármacos de anticuerpos es costoso y tedioso, y requiere la identificación de un antígeno y luego el aislamiento y la producción de anticuerpos monoclonales con actividad contra el antígeno. Las personas que han estado expuestas a una enfermedad producen anticuerpos contra los antígenos asociados con esa enfermedad, por lo que es posible extraer anticuerpos de los repertorios inmunitarios de los pacientes que podrían usarse para el desarrollo farmacéutico. Sin embargo, un anticuerpo monoclonal funcional requiere inmunoglobulinas de cadena pesada y ligera.

Determinadas implementaciones de la divulgación requieren una gran biblioteca de constructos de polinucleótido ligados que se componen de regiones variables de Ig de cadena ligera y pesada. En determinadas realizaciones, la biblioteca se genera a partir de cientos, miles, millones o miles de millones de células B individuales. El aislamiento de células B puede realizarse usando microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. La biblioteca está enriquecida previamente para una diana particular de interés o se enriquece a través del cribado por afinidad como scFv expresado en superficie (figura 4). Luego, la biblioteca enriquecida se convierte en una biblioteca de anticuerpos de longitud completa modificando por ingeniería primero los constructos de polinucleótido ligados en constructos de polinucleótido que codifican para anticuerpos de longitud completa, y luego cribando la afinidad o actividad de los anticuerpos de longitud completa contra un antígeno particular. La conversión en anticuerpos de longitud completa es particularmente útil, porque muchos scFv con afinidad contra un antígeno no tienen afinidad contra antígenos cuando se convierten en anticuerpos de longitud completa. En algunas implementaciones, la biblioteca de ADN de scFv se genera a partir de células B primarias aisladas de donantes humanos. En determinadas implementaciones, el enriquecimiento se produce al exponer a los donantes humanos a una vacuna que se compone de antígenos de un patógeno particular, de modo que las células B de los donantes

humanos se enriquecen en anticuerpos contra esos antígenos. En otras implementaciones, se aíslan células B de donantes con una enfermedad clínica particular, tal como una enfermedad autoinmunitaria o cáncer. En otras realizaciones, la biblioteca de complejos de Ig ligados se genera a partir de secuencias completamente artificiales, es decir, una biblioteca de gran diversidad de 10^{12} secuencias de ADN aleatorizadas.

En algunas implementaciones de la divulgación, se inmunizan ratones con una proteína u otra clase de antígeno de interés. Se aíslan células B individuales y se producen *in vitro* complejos ligados de regiones variables de cadena pesada y ligera apareadas. El aislamiento de células B puede realizarse usando microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. En algunas implementaciones, estas bibliotecas se expresan en la superficie de las células modificadas por ingeniería como scFv, y luego las células modificadas por ingeniería se clasifican por la unión a variantes de secuencia del antígeno de interés. En algunas implementaciones, se extrae el ARN o el ADN de las células modificadas por ingeniería clasificadas y se secuencian el ARN o el ADN para determinar el contenido de inmunoglobulina de las células seleccionadas. En algunas implementaciones, los complejos de Ig ligados amplificados a partir de células modificadas por ingeniería seleccionadas por antígeno se clonan luego en masa en un constructo de ADN circular, tal como un vector de plásmido, para producir una biblioteca de cientos, miles o millones de complejos ligados circularizados. En determinadas implementaciones, esta biblioteca seleccionada de vectores de plásmido se enriquece en complejos de Ig con afinidad hacia un antígeno de interés.

Para estudiar la función de estos complejos de Ig como anticuerpos de longitud completa, la secuencia ligadora entre las secuencias de polinucleótido de región variable de cadena pesada y ligera se reemplaza luego por una secuencia de polinucleótido que codifica para las subunidades de proteína de inmunoglobulina requeridas para la expresión del anticuerpo de longitud completa. En algunas implementaciones de esta descripción, la secuencia de polinucleótido insertada también incluye un promotor transcripcional que dirige la expresión de una de las cadenas de Ig. En algunas implementaciones de la divulgación, la biblioteca de constructos de expresión de proteína se introduce luego en una población de células hospedadoras de línea germinal no humana para producir una biblioteca de células modificadas por ingeniería que expresan una biblioteca de cientos, miles o millones de proteínas recombinantes. Luego, se analiza la biblioteca de anticuerpos de longitud completa para descubrir anticuerpos monoclonales que pueden ser de utilidad terapéutica. En determinadas implementaciones, los anticuerpos monoclonales funcionales se descubren primero aislando subpoblaciones de células modificadas por ingeniería y luego seleccionando reservas para determinar la afinidad contra un antígeno particular. Las reservas de modificación por ingeniería que muestran actividad contra un antígeno se dividen luego en células individuales y se criban de nuevo para determinar la afinidad contra un único antígeno. El aislamiento celular puede realizarse usando microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. De esta manera se descubren anticuerpos monoclonales con afinidad por antígenos de interés.

Terapia con anticuerpos policlonales

La inmunoglobulina intravenosa (IgIV) es un conjunto de proteínas aisladas del plasma de miles de donantes. La Administración de Alimentos y Fármacos de los EE. UU. (FDA) aprobó la terapia con IgIV para seis indicaciones, incluyendo la púrpura trombocitopénica idiopática (inmunitaria) (PTI), la vasculitis de Kawasaki, la leucemia linfocítica crónica (LLC) de células B y las inmunodeficiencias primarias (Orange *et al.*, 2006). Aunque se desconoce el mecanismo para la modulación autoinmunitaria, la mayoría de las IgIV se usan como terapia de reemplazo para pacientes con deficiencia de anticuerpos (Hartung *et al.*, 2009). Las ventas de IgIV ascienden a 7 mil millones de dólares en todo el mundo y crecen entre un 8% y un 10% por año, debido al envejecimiento de la población y a las modalidades no autorizadas en constante expansión (Taylor & Shapiro, 2013).

Los métodos actuales para la producción de IgIV amenazan la expansión continua de la terapia de IgIV debido al riesgo de la cadena de suministro, las impurezas y la variabilidad entre lotes. La producción de IgIV depende en gran medida del suministro limitado de sueros humanos y requiere inversiones en costosas instalaciones de purificación a gran escala. Más del 90% del suministro mundial está en manos de sólo 3 empresas. En 2006, la demanda de IgIV superó la oferta en un 4%, lo que obligó a los médicos a racionar la oferta rechazando a pacientes y administrando menores dosis (McGinnity, 2007). Debido a que la IgIV se purifica a partir de sueros primarios, las impurezas de proteínas y el espectro de contaminación viral son un problema continuado. Octapharma sufrió recientemente una retirada voluntaria masiva de su producto de IgIV (Octagam 5%) debido a complicaciones resultantes de la contaminación por el factor de coagulación XIa (Roemisch *et al.*, 2011). La IgIV depende de la unión a antígeno a través de su medio de región variable policlonal. Sin embargo, debido a la gran diversidad de repertorios inmunitarios en las poblaciones de donantes, las preparaciones siempre tienen un contenido de región variable diferente. Un estudio de títulos de anticuerpos anti-VHA de preparaciones de IgIV de 30 reservas diferentes de más de 60.000 donantes mostró una alta variabilidad con un CV del 33% entre reservas (Simon & Späth, 2003).

En una implementación de la divulgación, se produce IgIV en células recombinantes en lugar de extraerse del plasma del donante. En esta implementación, se recogen células B primarias de miles de donantes humanos. Se aíslan células individuales y se producen *in vitro* complejos ligados de regiones variables de cadena pesada y ligera apareadas. El aislamiento de células B puede realizarse usando microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. En algunas realizaciones, los complejos ligados se clonan en

masa en un constructo de ADN circular, tal como un vector de plásmido, para producir una biblioteca de cientos, miles o millones de complejos ligados circularizados. La secuencia ligadora entre las secuencias de polinucleótido de región variable de cadena pesada y ligera se reemplaza luego por un constructo de ligador que incluye un promotor transcripcional y cualquier parte requerida de Ig de cadena pesada o ligera. En algunas implementaciones de la divulgación, la biblioteca de constructos de expresión de proteína se introduce luego en una población de células, para producir una biblioteca de células modificadas por ingeniería que expresan una biblioteca de cientos, miles o millones de proteínas de anticuerpo. En algunas realizaciones, estas proteínas de anticuerpo son sustancialmente equivalentes a los anticuerpos producidos por las células B primarias originales. Por tanto, la reserva de proteínas de anticuerpo se usa como reemplazo recombinante de la IgIV. En algunas implementaciones, se usa la secuenciación en paralelo masiva de ADN para determinar la diversidad de las células B, la biblioteca inicial de Ig de cadena ligera y pesada apareadas, la biblioteca de constructos de expresión de proteína y/o las células hospedadoras modificadas por ingeniería. La secuenciación de ADN puede ser útil como etapa de control de calidad/garantía de calidad para la producción de bancos de células y bibliotecas de proteínas.

En otra implementación de la divulgación, los pacientes se seleccionan por la presencia de una afección médica particular, tal como la exposición a un patógeno particular. Estos pacientes actúan como donantes de células B que se enriquecen para la producción de anticuerpos contra ese patógeno particular. Los productores de IgIV convencionales ya comercializan productos hiperinmunes de IgIV convencionales con mayor actividad contra patógenos tales como la hepatitis B, la rabia, la toxina tetánica, la varicela-zóster y el citomegalovirus (CMV). En una implementación de la divulgación, a los donantes de células B se les inyecta una vacuna contra el neumococo. Se extraen células B de estos donantes y luego se producen grandes bibliotecas de constructos de expresión de proteína de polinucleótido a partir de cientos, miles o millones de células B individuales. El aislamiento de células B puede realizarse usando microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. En determinadas implementaciones, estos constructos de expresión de proteína se introducen en masa en células modificadas por ingeniería y luego se producen bibliotecas de proteínas a partir de las células modificadas por ingeniería. Estas bibliotecas de proteínas pueden contener cientos, miles o millones de anticuerpos individuales, según la diversidad de las células B de entrada iniciales. En algunas implementaciones, las bibliotecas de proteínas resultantes se usan como agentes terapéuticos policlonales dirigidos contra patógenos particulares, tales como el CMV. En otras implementaciones de la divulgación, se modifican por ingeniería células para expresar scFv de superficie usando una biblioteca de secuencias ligadas generadas a partir de células B primarias. Luego, las células que expresan scFv se exponen a un antígeno de interés, tal como el antígeno CMV, para seleccionar positivamente scFv con afinidad por dicho antígeno. La biblioteca de células enriquecidas se usa luego para generar una biblioteca de anticuerpos terapéuticos de longitud completa tal como se describió anteriormente. En algunas realizaciones, se usa la secuenciación de ADN paralela masiva para determinar la diversidad de las células B, la biblioteca inicial de Ig de cadena pesada y ligera apareadas, la biblioteca de constructos de expresión de proteína y/o las células hospedadoras modificadas por ingeniería. La secuenciación de ADN puede ser útil como etapa de control de calidad/garantía de calidad para la producción de bancos de células y bibliotecas de proteínas.

Alcance

En las reivindicaciones, los artículos como tales “un(o)”, “una” y “el/la” pueden significar uno o más de uno, a menos que se indique lo contrario o resulte evidente de otro modo del contexto. Las reivindicaciones o descripciones que incluyen “o” entre uno o más elementos de un grupo se consideran satisfechas si uno, más de uno o todos los elementos del grupo están presentes, se emplean en o son relevantes para un determinado producto o procedimiento, a menos que se indique de lo contrario o resulte evidente de otro modo por el contexto. También se observa que el término “que comprende” pretende ser abierto y permite, pero no requiere la inclusión de elementos o etapas adicionales. Cuando se usa el término “que comprende” en el presente documento, el término “que consiste en” también se incluye y da a conocer.

Cuando se dan rangos, se incluyen puntos finales. En caso de declaraciones contradictorias de una fuente citada y la presente solicitud, prevalecerá la declaración en la presente solicitud.

Los títulos de las secciones y tablas no pretenden ser limitativos.

Ejemplos

A continuación, se muestran ejemplos de realizaciones específicas para llevar a cabo la presente invención. Los ejemplos se ofrecen únicamente con fines ilustrativos y no pretenden limitar el alcance de la presente invención en modo alguno.

La práctica de la presente invención empleará, a menos que se indique de otro modo, métodos convencionales de química de proteínas, bioquímica, técnicas de ADN recombinante y farmacología, dentro de la experiencia en la técnica. Tales técnicas se explican completamente en la bibliografía. Véase, por ejemplo, TE Creighton, *Proteins: Structures and Molecular Properties* (W.H. Freeman and Company, 1993); A.L. Lehninger, *Biochemistry* (Worth Publishers, Inc., adición actual); Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual* (2ª edición, 1989); *Methods In Enzymology* (S. Colowick y N. Kaplan eds., Academic Press, Inc.); Remington's *Pharmaceutical*

Sciences, 18ª edición (Easton, Pensilvania: Mack Publishing Company, 1990); Carey y Sundberg Advanced Organic Chemistry 3ª ed. (Plenum Press) Vol. A y B (1992).

Ejemplo 1: Generación de bibliotecas de scFv

Ahora se comentarán los métodos de la invención en relación con la generación de bibliotecas de scFv.

Preparación de perlas

Se preparó tampón LiCl 2x de la siguiente manera: para 250 ml de tampón LiCl 2x, se añadieron 10 ml de Tris 1 M (pH 7,5), 31,25 ml de LiCl 8 M y 1 ml de EDTA 500 mM a 180,25 ml de agua de calidad molecular. El tampón de lisis 2x se preparó de la siguiente manera: por cada 1 ml de tampón de lisis 2x, se mezclaron 890 µl de disolución madre de LiCl 2x, 90 µl de agua de calidad molecular, 10 µl de DTT 1 M y 10 µl de Tween 20 para producir una disolución de 1 ml de tampón de lisis 2x.

Se sintetizaron sondas de IgK e IgG biotiniladas, que se unen a las regiones constantes de IgK e IgG, respectivamente, con las siguientes secuencias: SEQ ID NO: 6-7. Cada una de las sondas se añadió a 100 µl de tampón de lisis 2x hasta una concentración final de 10 µM.

Se agitó suavemente una disolución 1 µM de perlas magnéticas de estreptavidina de New England Biolabs (NEB S1420S) a temperatura ambiente durante 30 minutos. Se colocaron 200 µl de la disolución de perlas magnéticas de estreptavidina en un tubo de 1,5 ml. Se usó un imán potente para retirar el sobrenadante de las perlas. Luego, se lavaron las perlas con 600 µl de tampón de lisis 2x, seguido de la retirada del sobrenadante de las perlas usando un imán potente. Luego se resuspendieron las perlas en 60 µl de las sondas de IgK e IgG 10 µM. Se incubó la mezcla de perlas y sondas a temperatura ambiente durante 10 minutos. Después de la incubación, se expuso la mezcla a un imán para retirar el sobrenadante. Luego, se lavaron las perlas dos veces con 200 µl de tampón de lisis 2x, seguido de la retirada del sobrenadante con un imán después de cada lavado. Después del lavado, se resuspendieron las perlas en 500 µl de tampón de lisis 2x (dilución 1:2,5). Se añadieron 5 µl de un inhibidor de ARNasa a las perlas (concentración final de ARNasa del 1%) y se mezcló suavemente la disolución.

Captura de transcritos de ARN de células en perlas en emulsiones

Se montó un sistema microfluídico con tres bombas de presión (Dolomite microfluidics) y se conectó a una fuente de presión de al menos 6000 mbar. Se llenó una cámara de bomba de presión con una disolución de aceite mineral que comprendía Span-80 al 4,5%, Tween 80 al 0,4% y Triton X-100 al 0,05%. Se llenó la segunda cámara de presión con solución salina tamponada con fosfato de Dulbecco (DPBS). Se llenó la tercera cámara de la bomba de presión con agua. Se conectó cada una de las tres bombas de presión a un chip de gotas microfluídico (Dolomite microfluidics) que comprendía entradas para las fases oleosa y acuosa, una unión de enfoque de flujo para la generación de gotas y canales recubiertos con un material hidrófobo. Se purificaron células B primarias usando un kit de selección negativa pan-B (Stem Cells Inc.) y se mezclaron con un clon de células de ovario de hámster chino (CHO) que expresa adalumimab con una prevalencia del 0,1% en DPBS, y luego se cargaron en la segunda cámara de presión y se cargó la mezcla de perlas descrita anteriormente en la tercera cámara de presión. La línea CHO actúa como control positivo, expresando la secuencia de anticuerpo monoclonal adalumimab publicada previamente (<http://www.drugbank.ca/drugs/DB00051>). Luego, se inicializaron las tres bombas a una presión máxima de alrededor del 50%. Se normalizó la formación de gotas y se recogieron las emulsiones que contenían gotas con la mezcla de perlas/células en tubos para PCR. Luego, se incubaron los tubos a 50°C durante 30 minutos. Después de la incubación, se realizó la extracción con acetato de etilo. Se añadió un volumen 2:1 de acetato de etilo a cada tubo, se transfirió a un tubo de 1,5 ml, seguido de agitación en vórtex y centrifugación a máxima velocidad durante 1 minuto. Después de la centrifugación, se retiró el sobrenadante. Se repitió el proceso con un volumen 1:1 de acetato de etilo hasta que se hubo roto suficiente emulsión. Después de la extracción con acetato de etilo, se colocó el tubo en un imán potente durante 1-2 minutos y se retiró y se desechó todo el sobrenadante.

Amplificación de complejos ligados de perlas en emulsiones para formar polinucleótidos que codifican para scFv

Las perlas que comprendían sondas unidas a los transcritos de ARN recogidos anteriormente se expusieron luego a RT-PCR de extensión por solapamiento para amplificar los transcritos de ARN.

Se transfirieron 25 µl de las perlas con transcritos de ARN unidos a un tubo de 1,5 ml y se colocaron sobre un imán, lo que permitió la retirada del sobrenadante. Luego, se lavaron las perlas con 100 µl de agua seguido por la retirada del sobrenadante. Luego, se aislaron perlas individuales o subpoblaciones de perlas en cámaras de reacción para amplificar el ARN de una célula individual o subpoblación de células. Para lograr esto, se resuspendieron las perlas en 250 µl de mezcla fría de RT-PCR (kits y enzimas de NEB, Thermo Fisher y Qiagen). La mezcla de RT-PCR consistía en una concentración final de: tampón de reacción 1x, cebador de IgK V externo 1 µM (SEQ ID NO: 8), cebador de IgK C interno 0,2 µM (SEQ ID NO: 9), cebador de IgG V interno 0,2 µM (SEQ ID NO: 10), cebador de IgG C externo 1 µM (SEQ ID NO: 11), ET SSB 4 ng/µl, inhibidor de ARNasa al 2% y transcriptasa inversa al 4% y

polimerasa termoestable. Se formó una emulsión que comprendía las perlas y la mezcla de RT-PCR usando un dispositivo de generación de emulsión (sistemas Tube Drive ULTRA-TURRAX de IKA con tubos DT-20). Se colocó un tubo DT-20 en el dispositivo generador de emulsión y se añadieron al tubo 750 μ E de una mezcla de aceite que comprendía Span-80 al 4,5%, Tween 80 al 0,4%, Triton X-100 al 0,05% en aceite mineral. Luego, se añadieron gota a gota 250 μ l de perlas frías y mezcla de RT-PCR a la parte superior de la fase de aceite en el tubo DT-20. Luego se formaron emulsiones de la disolución en el tubo DT-20 usando el dispositivo de generación de emulsión. Después de preparar las emulsiones, se dividió la mezcla de emulsión en alícuotas de 100 μ l en tubos para PCR.

Luego se realizó RT-PCR en la mezcla de emulsión usando las siguientes condiciones de termociclado secuencialmente:

55°C	30 min	
94°C	3 min	
94°C	30 s	17 ciclos; -0,5°C apareamiento por ciclo
De 65°C a 57°C	3 min	
68°C	1 min	
94°C	30 s	
57°C	3 min	26 ciclos [43 ciclos en total]
68°C	1 min	
68°C	2 min	
4°C	∞	

Después de la PCR, se añadieron 100 μ l de acetato de etilo a cada tubo y se mezcló la disolución con una pipeta. Luego, se transfirieron las emulsiones rotas a un tubo de 1,5 ml y se agitaron en vórtex pulsado para mezclar. Luego, se centrifugó la disolución a máxima velocidad durante 1 minuto y se retiró la fase superior (sobrenadante). Luego, se colocó el tubo sobre un imán y se transfirió la fase inferior (fase acuosa) que comprendía el producto de amplificación a un tubo nuevo de 2 ml sin perlas. Luego, se usó un kit de purificación QIAquick PCR de Qiagen para extraer el ADN amplificado de la fase acuosa. El ADN amplificado aislado incluye una biblioteca de insertos de scFv (polinucleótido que codifica para scFv) (es decir, SEQ ID NO: 13, de la secuencia de control de adalumimab) que se usará en la generación de constructos de expresión.

Generación de constructos de expresión

Se modificó el vector pPIC9 (Life Technologies) para que incluyese una parte de la secuencia de IgG1 humana (SEQ ID NO: 12). Se determinó la concentración de insertos de scFv generados anteriormente. Se combinaron 0,02 pmol del vector pPIC9_IgG1 con 0,04 pmol de insertos de scFv aislados (por ejemplo, ejemplo SEQ ID NO: 13) en agua hasta un volumen final de 5 μ l. Se añadieron 5 μ l de mezcla maestra de ensamblaje de Gibson 2X (New England Biolabs) a la disolución de inserción de vector/scFv. Luego, se incubaron las muestras a 50°C durante 60 minutos para generar un constructo circularizado del inserto de scFv en el vector pPIC9_IgG1.

Luego, se añadió un promotor AOX1 de tipo natural al constructo circularizado para inducir la expresión del inserto de scFv de la siguiente manera: se combinaron en agua 0,02 pmol del constructo circularizado que comprendía el inserto de scFv y el vector pPIC9_IgG1 hasta un volumen final de 5 μ l. Se añadieron 5 μ E de mezcla maestra de ensamblaje de Gibson 2X (New England Biolabs) a la disolución de inserción de vector/scFv. Luego, se incubaron las muestras a 50°C durante 60 minutos para generar un constructo circularizado del inserto de scFv en el vector pPIC9_IgG1.

Luego, se linealizó el constructo circularizado con endonucleasa de restricción XhoI de NEB para generar ADN de plásmido de scFv linealizado.

Expresión de proteínas en levadura

Se descongelaron células de levadura competentes congeladas en hielo. Se transfirieron 40 μ l de células de levadura a un tubo que contenía 10 μ l de ADN de plásmido de scFv linealizado 0,1 μ g/ μ l. Se incubó en hielo la mezcla de ADN y células de levadura durante 5 minutos. Luego, se sometió la mezcla a electroporación a 1,5 kV, 200 Omega y 25 μ F. Luego, se añadió a la mezcla 1 ml de medio de recuperación (50% de sorbitol 1 M, 50% de extracto de levadura de peptona y dextrosa (YPD)) y se agitó el tubo a 200 rpm durante 1 hora a 30°C. Luego, la levadura se sembró en placas de selección de agar RDB menos His y se incubó durante la noche a 30°C. Luego, se seleccionó una colonia de levadura transformada de la placa y se añadió a 50 ml de medio BMMY. Se agitó la levadura a 200 rpm durante la noche (al menos 20 horas) a 30°C, hasta alcanzar una DO600 de entre 5 y 10. Luego, se centrifugó el cultivo a 2000xg durante 5 minutos y se retiró el sobrenadante. Luego, se resuspendieron las células de levadura en 50 ml de medio BMMY y se cultivaron durante 24 horas a 30°C mientras se agitaban a 200 rpm. Después de 24 horas, se añadió metanol al cultivo de levadura hasta una concentración final de 1% en volumen por volumen. Luego, se centrifugó el cultivo de levadura a 2000 xg durante 5 minutos para recoger la biblioteca de scFv

en el sobrenadante. Se recogió el sobrenadante, se filtró y se almacenó a 4°C para su posterior purificación y análisis.

Se sometió una parte del sobrenadante recogido de la levadura a una inmunotransferencia de tipo Western tal como se muestra en la figura 5 para mostrar la biblioteca de scFv. Los scFv se marcan con etiqueta con el marcador peptídico c-myc. Los scFv se marcan con etiqueta con el marcador peptídico c-myc. Los carriles 1-6 se inducen usando el promotor AOX1 de tipo natural a las 72 h, 66 h, 47 h, 24 h y 18 h, respectivamente. Los carriles 7-12 son los mismos puntos de tiempo que los carriles 1-6, respectivamente, pero usando un promotor diferente. El carril 14 es un marcador de tamaño.

Ejemplo 2: Generación de anticuerpos

Generación de constructos de expresión

Se linealizó una biblioteca de scFv (que incluye, por ejemplo, SEQ ID NO: 1) por PCR usando métodos de PCR convencionales y cebadores directos e inversos (por ejemplo, SEQ ID NO: 4, 5). Un proveedor de síntesis de genes (IDT) sintetizó un inserto que contenía un segundo promotor AOX1. Luego, se añadió un inserto al constructo circularizado para inducir la expresión del anticuerpo de longitud completa, de modo que el vector contendría ahora dos promotores AOX1, es decir, uno para cada una de la inmunoglobulina de cadena pesada y de cadena ligera. La nueva biblioteca se diseñó de la siguiente manera: se combinaron 0,02 pmol del constructo linealizado que comprendía el vector pPIC9_IgG1_scFv con 0,04 pmol del inserto promotor (SEQ ID NO: 3) en agua hasta un volumen final de 5 µl. Se añadieron 5 µl de mezcla maestra de ensamblaje de Gibson 2X (New England Biolabs) a la disolución de promotor / vector de scFv. Luego, se incubaron las muestras a 50°C durante 60 minutos para generar un constructo circularizado del inserto promotor en el vector pPIC9_IgG1_scFv.

Expresión de proteínas en levadura

Se descongelaron células de levadura competentes congeladas en hielo. Se transfirieron 40 µl de células de levadura a un tubo que contenía 10 µl de ADN de anticuerpo de longitud completa linealizado 0,1 µg/µl de antes. Se incubó en hielo la mezcla de ADN y células de levadura durante 5 minutos. Luego, se sometió la mezcla a electroporación a 1,5 kV, 200 Ohm y 25 µF. Luego, se añadió a la mezcla 1 ml de medio de recuperación (50% de sorbitol 1 M, 50% de YPD) y se agitó el tubo a 200 rpm durante 1 hora a 30°C. Luego, se sembró la levadura en placas de selección de agar RDB menos His y se incubó. Luego, se seleccionó una colonia de levadura transformada de la placa y se añadió a 50 ml de medio BMGY. Se agitó la levadura a 200 rpm durante la noche (al menos 20 horas) a 30°C, hasta alcanzar una DO600 de entre 5 y 10. Luego, se centrifugó el cultivo a 2000xg durante 5 minutos y se retiró el sobrenadante. Luego, se resuspendieron las células de levadura en 50 ml de medio BMMY y se hicieron crecer durante 24 horas a 30°C mientras se agitaban a 200 rpm. Después de 24 horas, se añadió metanol al cultivo de levadura hasta una concentración final del 1% en peso por volumen. Luego, el cultivo de levadura se centrifugó a 2000 xg durante 5 minutos para recoger la biblioteca de anticuerpos recombinantes en el sobrenadante. Se recogió el sobrenadante, se filtró y se almacenó a 4°C para su posterior purificación y análisis.

Se sometió una parte del sobrenadante recogido de la levadura a una inmunotransferencia de tipo Western tal como muestra la biblioteca de anticuerpos recombinantes. La figura 6 muestra la inmunotransferencia de tipo Western resultante, que muestra la expresión de anticuerpos completos en levadura. El carril 1 es un marcador de tamaño. Los carriles 2-5 son anticuerpos expresados en levadura, el carril 6 es anticuerpo expresado en células CHO, todos en condiciones reducidas. Los carriles 7-10 son los mismos anticuerpos expresados en levadura, el carril 11 es el anticuerpo expresado en células CHO, en condiciones no reducidas.

Ejemplo 3: Descubrimiento de fármacos de anticuerpos monoclonales

Se inmunizan ratones con una proteína u otra clase de antígeno de interés. Se aíslan células B individuales de los ratones y se producen *in vitro* complejos ligados de regiones variables de cadena pesada y ligera apareadas. El aislamiento de células B se realiza mediante microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. Las bibliotecas de constructos de fusión se insertan y se expresan en células hospedadoras en la superficie como scFv. Se criban las células modificadas por ingeniería para determinar si se unen a variantes de secuencia del antígeno de interés. Se extrae ARN o ADN de células modificadas por ingeniería con afinidad de unión por el antígeno de interés, y se secuencian el ARN o el ADN para determinar el contenido de inmunoglobulina de las células seleccionadas. Los complejos de Ig ligados amplificados a partir de células modificadas por ingeniería seleccionadas genéticamente por antígeno se clonan en masa en vectores plasmídicos para producir una biblioteca de vectores plasmídicos que comprenden el constructo de fusión recombinante.

Para estudiar la función de estos complejos de Ig como anticuerpos de longitud completa, la secuencia ligadora entre las secuencias de polinucleótido de región variable de cadena ligera y pesada se reemplaza por una secuencia de polinucleótido que codifica para las subunidades de proteína de inmunoglobulina requeridas para la expresión de un anticuerpo de longitud completa que comprende regiones variables de cadena pesada y ligera apareadas producidas *in vitro*. La secuencia de polinucleótido insertada también incluye un promotor transcripcional que dirige

la expresión de una de las cadenas de Ig. La biblioteca de constructos de expresión de proteína se introduce en una población de células hospedadoras para producir una biblioteca de células modificadas por ingeniería que expresan una biblioteca de anticuerpos de longitud completa recombinantes. La biblioteca de anticuerpos de longitud completa recombinantes se analiza para descubrir anticuerpos monoclonales de utilidad terapéutica. Para realizar este análisis, se aíslan las subpoblaciones de células modificadas por ingeniería y se criban reservas aisladas para determinar su afinidad contra un antígeno particular. Las reservas de células modificadas que muestran actividad contra un antígeno se dividen en células individuales y se criban de nuevo para detectar afinidad contra un antígeno individual. El aislamiento celular se realiza mediante microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. Se descubren anticuerpos recombinantes con afinidad por antígenos de interés.

Ejemplo 4: Terapia con anticuerpos policlonales

En este ejemplo, se produce IgIV en células recombinantes en lugar de extraerse del plasma de un donante. Se recogen células B primarias de miles de donantes humanos. Se aíslan células individuales de cada donante y se producen *in vitro* complejos ligados de regiones variables de cadena pesada y ligera apareadas de cada célula. El aislamiento de células B se realiza usando microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. Los complejos de proteínas fusionadas resultantes se clonan en masa en vectores plasmídicos para producir una biblioteca de vectores plasmídicos que codifican para proteínas fusionadas. La secuencia ligadora entre las secuencias de polinucleótido de región variable de cadena pesada y ligera se reemplaza por un constructo de ligador que incluye un promotor transcripcional y cualquier parte requerida de Ig de cadena pesada o ligera, es decir, parte de una región constante que no estaba incluida en el constructo de expresión de scFv original. La biblioteca de constructos de expresión de proteínas fusionadas se introduce en una población de células para producir una biblioteca de células modificadas por ingeniería que expresan una biblioteca de proteínas de anticuerpo recombinantes. Las proteínas de anticuerpo recombinantes comprenden dominios de inmunoglobulina ligeros y pesados variables de las células B primarias de origen. La reserva de proteínas de anticuerpo puede usarse como reemplazo recombinante de la IgIV.

La diversidad de las células B, la biblioteca inicial de Ig de cadena ligera y pesada apareadas, la biblioteca de constructos de expresión de proteína y/o las células hospedadoras modificadas por ingeniería se determinarán usando secuenciación de ADN paralela masiva.

Ejemplo 5: Terapia de neumococos con anticuerpos policlonales

A los donantes de células B se les inyecta una vacuna contra el neumococo. Se aíslan células B de estos donantes. Luego, se producen grandes bibliotecas de constructos de expresión de proteína de polinucleótido a partir de cada una de las células B individuales. El aislamiento de células B se realiza mediante microfluídica de gotas o aislamiento en contenedores físicos tales como placas de 96 pocillos. Los constructos de expresión de proteína se introducen en masa en células modificadas y luego se producen bibliotecas de proteínas a partir de las células modificadas por ingeniería. Estas bibliotecas de proteínas contienen cientos, miles o millones de anticuerpos individuales, según la diversidad de las células B de entrada iniciales. Las bibliotecas de proteínas resultantes se usan como agentes terapéuticos policlonales dirigidos contra patógenos particulares, tales como el CMV.

Lista de secuencias informal

Secuencia 1. Secuencia de scFv en vector pD912

CTTCAAGTAATGTCTTGTTCCTTTTGTGTCAGTGGTGAGCCATTTTGACTTCGTGAAAGTTTCTTT
 AGAATAGTTGTTTCCAGAGGCCAAACATTCCACCCGTAGTAAAGTGCAAGCGTAGGAAGACCAAG
 ACTGGCATAAATCAGGTATAAGTGTGAGCACTGGCAGGTGATCTTCTGAAAGTTTCTACTAGCA
 GATAAGATCCAGTAGTCATGCATATGGCAACAATGTACCGTGTGGATCTAAGAACGCGTCCTACT
 AACCTTCGCATTTCGTTGGTCCAGTTTGTGTTATCGATCAACGTGACAAGGTTGTCGATTCCGCG
 TAAGCATGCATACCCAAGGACGCCTGTTGCAATTCCAAGTGAGCCAGTTCCAACAATCTTTGTAA
 TATTAGAGCACTTCATTGTGTTGCGCTTGAAAGTAAAATGCGAACAAATTAAGAGATAATCTCGA
 AACCGCGACTTCAAACGCCAATATGATGTGCGGCACACAATAAGCGTTCATATCCGCTGGGTGAC
 TTTCTCGCTTTAAAAAATTATCCGAAAAAATTTCTAGAGTGTGTTACTTTATACTTCCGGCTC
 GTATAATACGACAAGGTGTAAGGAGGACTAAACCATGGCTAAACTCACCTCTGCTGTTCCAGTCC
 TGAAGTGTGCTGATGTTGCTGGTGTGTTGAGTTCTGGACTGATAGACTCGGTTTCTCCCGTGAC
 TTCGTAGAGGACGACTTTGCCGGTGTGTTGACGTGACGACGTTACCCTGTTTCATCTCCGCGATTCA
 GGACCAGGTTGTGCCAGACAACACTCTGGCATGGGTATGGGTTCGTGGTCTGGACGAACTGTACG
 CTGAGTGGTCTGAGGTCGTGTCTACCAACTTCCGTGATGCATCTGGTCCAGCTATGACCGAGATC
 GGTGAACAGCCCTGGGGTCTGAGTTTGCAGTGGTGTGATCCAGCTGGTAACTGCGTGCATTTCTGT
 CGCAGAAGAACAGGACTAACAATTGACACCTTACGATTATTTAGAGAGTATTTATTAGTTTTATT
 GTATGTATACGGATGTTTTATTATCTATTTATGCCCTTATATTCTGTAACATATCCAAAAGTCCTA
 TCTTATCAAGCCAGCAATCTATGTCCGCGAACGTCAACTAAAAATAAGCTTTTTATGCTGTTCTC
 TCTTTTTTCCCTTCGGTATAATTATACCTTGCATCCACAGATTCTCCTGCCAAATTTTGCATAA
 TCCTTTTACAACATGGCTATATGGGAGCACTTAGCGCCCTCCAAAACCCATATTGCCTACGCATGT
 ATAGGTGTTTTTTCCACAATATTTTCTCTGTGCTCTCTTTTTTATTAAAGAGAAGCTCTATATCGG
 AGAAGCTTCTGTGGCCGTTATATTCCGCCCTTATCGTGGGACCACATTGCCTGAATTGGTTTGCCC
 CGGAAGATTGGGGAACTTGGATCTGATTACCTTAGCTGCAGGTACCACTGAGCGTCAGACCCCG
 TAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACA
 AAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAA
 GGTAACGGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCC
 ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCT
 GCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGC
 GCAGCGGTCCGGCTGAACGGGGGGTTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCG
 AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGAC
 AGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGC
 CTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCT
 CGTCAGGGGGGCGGAGCCTATGGAACAAACGCCAGCAACGCGGCCCTTTTACGGTTCCCTGGCCTTT
 TGCTGGCCCTTTTGCTCACATGTTCTTTCTGCGGTACCCAGATCCAATTCCCGCTTTGACTGCCT
 GAAATCTCCATCGCCTACAATGATGACATTTGGATTTGGTTGACTCATGTTGGTATTGTGAAATA
 GACGCAGATCGGGAACACTGAAAAATACACAGTTATTATTCAATTTAAATAACATCCAAAGACGAA
 AGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCATTCTCACACATAAGTGCCAAAC
 GCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAACGCAGGACCTCCACTCCTCTTCTCCT
 CAACACCCACTTTTTGCCATCGAAAAACCAGCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTC
 CAATTCCCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCTGTCTATCCTGGCCCCCTGGC

GAGGTTTCATGTTTGTGTTTATTTCCGAATGCAACAAGCTCCGCATTACACCCGAACATCACTCCAGA
 TGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCCAAATGGCCCAAACTGACAGTT
 TAAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAACCTAAGTTTGGT
 TCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGGCATACCGTTTGTCT
 TGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTCTCTATC
 GCTTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAATGGGGGAAACACCCGCTTTTTTGGATGATT
 ATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTT
 CATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATATATAAACAGAAGGAAGCTGCC
 CTGTCTTAAACCTTTTTTTTTTATCATCATTATTAGCTTACTTTCATAATTGCGACTGGTTCCAAT
 TGACAAGCTTTTGATTTTAAACGACTTTTAAACGACAACCTTGAGAAGATCAAAAAACAATAATTAT
 TGAAAGAATTCCGAAACGATGAGATTCCCATCTATTTTACCCGCTGTCTTGTTGCTGCCTCCTC
 TGCATTGGCTGCCCTGTTAACACTACCACTGAAGACGAGACTGCTCAAATTCAGCTGAAGCAG
 TTATCGGTTACTCTGACCTTGAGGGTGATTTGACGTCGCTGTTTTGCCTTTCTCTAACTCCACT
 AACAAACGGTTTGTTGTTTCATTAAACACCACTATCGCTTCCATTGCTGCTAAGGAAGAGGGTGTCTC
 TCTCGAGAAAAGAGAGGCCGAAGCTGACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCAT
 CTGTAGGGGACAGAGTCACCATCACTTGTGCGGGCAAGTCAGGGCATCAGAAATTACTTAGCCTGG
 TATCAGCAAAAACCAGGGAAAGCCCCTAAGCTCCTGATCTATGCTGCATCCACTTTGCAATCAGG
 GGTCCCATCTCGGTTTCAGTGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTAC
 AGCCTGAAGATGTTGCAACTTATTACTGTCAAAGGTATAACCGTGCACCGTATACTTTTGGCCAG
 GGGACCAAGGTGGAAATCAAACGAACTGTGGCTGCACCATCTGTGCGCGGATCCTCTAGGTCAAG
 TTCCAGCGGCGGCGGTGGCAGCGGAGGCGGCGGTGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCT
 TGGTACAGCCCGGCAGGTCCCTGAGACTCTCCTGTGCGGCCCTTGATTACCTTTGATGATTAT
 GCCATGCACTGGGTCCGGCAAGCTCCAGGGAAGGGCCTGGAATGGGTCTCAGCTATCACTTGGA
 TAGTGGTCACATAGACTATGCGGACTCTGTGGAGGGCCGATTACCATCTCCAGAGACAACGCCA
 AGAACTCCCTGTATCTGCAAATGAACAGTCTGAGAGCTGAGGATACGGCCGTATATTACTGTGCG
 AAAGTCTCGTACCTTAGCACCGCGTCTCCTTACTATTTGGGGCCAAGGTACCCTGGTCACCGT
 CTCGAGTGCCTCCACCAAGGGCCCATCGGTCTTCGAACAGAAGCTCATCTCAGAAGAGGATCTGT
 AAAGGGGCGGCCGCTCAAGAGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTT
 GATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCATTTTGTTTCTTCTCGTAC
 GAGCTTGCTCCTGATCAGCCTATCTCGCAGCAGATGAATATCTTGTGGTAGGGGTTTGGGAAAT
 CATTCGAGTTTGATGTTTTTCTTGGTATTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAA
 ACCTTCGTTTGTGCGGATC

Secuencia 2. Secuencia 1 de scFv convertida en secuencia de anticuerpo de longitud completa en vector pD912

CTTCAGTAATGTCTTGTTTCTTTTGTGTCAGTGGTGAGCCATTTTGACTTCGTGAAAGTTTCTTT
 AGAATAGTTGTTTCCAGAGGCCAAACATTCCACCCGTAGTAAAGTGCAAGCGTAGGAAGACCAAG
 ACTGGCATAAATCAGGTATAAGTGTGAGCACTGGCAGGTGATCTTCTGAAAGTTTCTACTAGCA
 GATAAGATCCAGTAGTCATGCATATGGCAACAATGTACCGTGTGGATCTAAGAACGCGTCCTACT
 AACCTTCGCATTGCTTGGTCCAGTTTGTGTTATCGATCAACGTGACAAGGTTGTCGATTCCGCG
 TAAGCATGCATACCCAAGGACGCCTGTTGCAATTCCAAGTGAGCCAGTTCCAACAATCTTTGTAA
 TATTAGAGCACTTCATTGTGTTGCGCTTGAAAGTAAAATGCGAACAAATTAAGAGATAATCTCGA
 AACCGCGACTTCAAACGCCAATATGATGTGCGGCACACAATAAGCGTTCATATCCGCTGGGTGAC
 TTTCTCGCTTTAAAAAATTATCCGAAAAAATTTCTAGAGTGTTGTTACTTTATACTTCCGGCTC

GTATAATACGACAAGGTGTAAGGAGGACTAAACCATGGCTAAACTCACCTCTGCTGTTCCAGTCC
 TGA CTGCTCGTGATGTTGCTGGTGCTGTTGAGTTCTGGACTGATAGACTCGGTTTCTCCCGTGAC
 TTCGTAGAGGACGACTTTGCCGGTGTTGTACGTGACGACGTTACCCTGTTTCATCTCCGAGTTCA
 GGACCAGGTTGTGCCAGACAACACTCTGGCATGGGTATGGGTTCGTGGTCTGGACGAACTGTACG
 CTGAGTGGTCTGAGGTCGTGTCTACCAACTTCCGTGATGCATCTGGTCCAGCTATGACCGAGATC
 GGTGAACAGCCCTGGGGTCTGAGTTTGCCTGCGTGATCCAGCTGGTAACGCGTGCATTTTCGT
 CGCAGAAGAACAGGACTAACAATTGACACCTTACGATTATTTAGAGAGTATTTATTAGTTTTATT
 GTATGTATACGGATGTTTTATTATCTATTTATGCCCTTATATTCTGTAACATATCCAAAAGTCCTA
 TCTTATCAAGCCAGCAATCTATGTCCGCGAACGTCAACTAAAAATAAGCTTTTTTATGCTGTTCTC
 TCTTTTTTTCCCTTCGGTATAATTATACCTTGCATCCACAGATTCTCCTGCCAAATTTTGCATAA
 TCCTTTTACAACATGGCTATATGGGAGCACTTAGCGCCCTCCAAAACCCATATTGCCTACGCATGT
 ATAGGTGTTTTTTCCACAATATTTTCTCTGTGCTCTCTTTTTTATTAAAGAGAAGCTCTATATCGG
 AGAAGCTTCTGTGGCCGTTATATTCCGCCCTTATCGTGGGACCACATTGCCTGAATTGGTTTGCCC
 CGGAAGATTGGGGAACTTGGATCTGATTACCTTAGCTGCAGGTACCACTGAGCGTCAGACCCCG
 TAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACA
 AAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAA
 GGTAAC TGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCC
 ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCT
 GCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGC
 GCAGCGGTCTGGGCTGAACGGGGGGTTCTGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCG
 AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGAC
 AGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGC
 CTGGTATCTTTATAGTCCTGTCTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCT
 CGTCAGGGGGGCGGAGCCTATGGA AAAACGCCAGCAACGCGGCCCTTTTACGGTTCTTGGCCTTT
 TGCTGGCCCTTTTGCTCACATGTTCTTTCTGCGGTACCCAGATCCAATTCCCGCTTTGACTGCCT
 GAAATCTCCATCGCCTACAATGATGACATTTGGATTTGGTTGACTCATGTTGGTATTGTGAAATA
 GACGCAGATCGGGAACACTGAAAAATACACAGTTATTATTTCATTTAAATAACATCCAAAGACGAA
 AGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCATTCTCACACATAAGTGCCAAAC
 GCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAAACGCAGGACCTCCACTCCTCTTCTCCT
 CAACACCCACTTTTTGCCATCGAAAAACCAGCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTC
 CAATTCCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCTGTCTATCCTGGCCCCCTGGC
 GAGGTT CATGTTTGTTTATTTCCGAATGCAACAAGCTCCGCATTACACCCGAACATCACTCCAGA
 TGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCCAAATGGCCCAAACTGACAGTT
 TAAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAAC TAAGTTTGGT
 TCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGGCATAACGTTTGTCT
 TGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTCTCTATC
 GCTTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAAATGGGGAAACACCCGCTTTTTTGATGATT
 ATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTT
 CATGATCAAAAATTTAACTGTTCTAACCCTACTTGACAGCAATATATAAACAGAAGGAAGCTGCC
 CTGTCTTAAACCTTTTTTTTTTATCATCATTATTAGCTTACTTTCATAATTGCGACTGGTTCCAAT
 TGACAAGCTTTTGATTTTAACGACTTTTAACGACAACCTTGAGAAGATCAAAAAACAATAATTAT
 TGAAAGAAATCCGAAACGATGAGATTCCCATCTATTTTACCGCTGTCTTGTTTCGCTGCCTCCTC
 TGCATTGGCTGCCCTGTTAACACTACCACTGAAGACGAGACTGCTCAAATTCAGCTGAAGCAG

TTATCGGTTACTCTGACCTTGAGGGTGATTTGACGTCGCTGTTTTGCCTTTCTCTAACTCCACT
 AACAAACGGTTTGTGTTCATTAACACCACTATCGCTTCCATTGCTGCTAAGGAAGAGGGTGTCTC
 TCTCGAGAAAAGAGAGGCCGAAGCTGACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCAT
 CTGTAGGGGACAGAGTCACCATCACTTGTCTGGGCAAGTCAGGGCATCAGAAATTACTTAGCCTGG
 TATCAGCAAAAACCAGGGAAAGCCCCTAAGCTCCTGATCTATGCTGCATCCACTTTGCAATCAGG
 GGTCCCATCTCGGTTTCAGTGGCAGTGGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTAC
 AGCCTGAAGATGTTGCAACTTATTACTGTCAAAGGTATAACCGTGCACCGTATACTTTTGCCAG
 GGGACCAAGGTGGAAATCAAACGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGA
 TGAGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGG
 CCAAAGTACAGTGGAAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAGTGTACAGAG
 CAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCCTGACGCTGAGCAAAGCAGACTACGA
 GAAACACAAAGTCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAGAGCT
 TCAACAGGGGAGAGTGTAAAGGGGCGGCCGCTCAAGAGGATGTCAGAATGCCATTTGCCTGAGA
 GATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATTTTTTTTGTCA
 TTTTGTCTTCTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCAGATGAATATCTTGTG
 GTAGGGGTTTGGGAAAATCATTTCAGTTTGTATGTTTTTCTTGGTATTTCCCACTCCTCTTCAGAG
 TACAGAAGATTAAGTGAACCTTCGTTTGTGCGTGTCTTTCTGCGGTACCCAGATCCAATTCC
 CGCTTTGACTGCCTGAAATCTCCATCGCCTACAATGATGACATTTGGATTTGGTTGACTCATGTT
 GGTATTGTGAAATAGACGCAGATCGGGAACACTGAAAAATACACAGTTATTATTTCATTTAAATAA
 CATCCAAAGACGAAAGGTTGAATGAACCTTTTTGCCATCCGACATCCACAGGTCCATTCTCACA
 CATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAACGCAGGACCTCC
 ACTCCTCTTCTCCTCAACACCCACTTTTTGCCATCGAAAAACCAGCCAGTTATTGGGCTTGATTG
 GAGCTCGCTCATTTCCAATTCCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCTGTCTATC
 CTGGCCCCCTGGCGAGGTTTCATGTTTGTGTTATTTCCGAATGCAACAAGCTCCGCATTACACCCG
 AACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAAATAGTTTCATGTTCCCCAAATGGCC
 CAAAACCTGACAGTTTAAACGCTGTCTTGAACCTAATATGACAAAAGCGTGATCTCATCCAAGAT
 GAACTAAGTTTGGTTTCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGG
 CATAACGTTTGTCTTGTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATGCTTAGC
 GCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAATGGGGAAACACCCG
 CTTTTTGGATGATTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCT
 GATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATATATAAAC
 AGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTTATCATCATTATTAGCTTACTTTTCATAATTG
 CGACTGGTTCCAATTGACAAGCTTTTGATTTTAAACGACTTTTAAACGACAACCTTGAGAAGATCAAA
 AAACAACCTAATTATTGAAAGAATTCCGAAACGATGAGATTCCCATCTATTTTCACCGCTGTCTTG
 TTCGCTGCCTCCTCTGCATTGGCTGCCCCTGTTAACACTACCACTGAAGACGAGACTGCTCAAAAT
 TCCAGCTGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTTGACGTCGCTGTTTTGCCTT
 TCTCTAACTCCACTAACAACGGTTTGTGTTTCATTAACACCACTATCGCTTCCATTGCTGCTAAG
 GAAGAGGGTGTCTCTCTCGAGAAAAGAGAGGCCGAAGCTGAGGTGCAGCTGGTGGAGTCTGGGGG
 AGGCTTGGTACAGCCCGCAGGTCCCTGAGACTCTCCTGTGCGGCCCTCTGGATTACCTTTGATG
 ATTATGCCATGCACCTGGGTCCGGCAAGCTCCAGGGAAGGGCCTGGAATGGGTCTCAGCTATCACT
 TGGAATAGTGGTCACATAGACTATGCGGACTCTGTGGAGGGCCGATTACCATCTCCAGAGACAA
 CGCCAAGAACTCCCTGTATCTGCAAATGAACAGTCTGAGAGCTGAGGATACGGCCGTATATTACT
 GTGCGAAAAGTCTCGTACCTTAGCACCGCGTCCCTTGGCTATTGGGGCCAAGGTACCCTGGTC

ACCGTCTCGAGTGCCTCCACCAAGGGCCCATCGGTCTTCCCCCTGGCACCCCTCCTCCAAGAGCAC
 CTCTGGGGGACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGACGGTGT
 CGTGGAATCAGGCGCCCTGACCAGCGGCGTGACACCTTCCCGGTGTCTTACAGTCCTCAGGA
 CTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTG
 CAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACA
 AAATCACACATGCCCACCGTGCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCTCTTC
 CCCCCAAAACCCAAGGACACCCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGA
 CGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATG
 CCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTC
 CTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGC
 CCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGC
 CCCCATCCCGGGATGAGCTGACCAAGAACCAGGTACGCTGACCTGCCTGGTCAAAGGCTTCTAT
 CCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGAGAACAACTACAAGACCACGCC
 TCCCGTGTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGT
 GGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
 AAGAGCCTCTCCCTGTCTCCGGGTAAATAAAGGGGCGGCCGCTCAAGAGGATGTGAGAATGCCAT
 TTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGATT
 TTTTTTGTCATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCAGATGA
 ATATCTTGTGGTAGGGGTTTGGGAAAATCATTCGAGTTTGATGTTTTTCTTGGTATTTCCCACTC
 CTCTTCAGAGTACAGAAGATTAAGTGAAACCTTCGTTTGTGCGGATC

Secuencia 3. Secuencia insertada en la secuencia 1 para crear la secuencia 2

TTCATCTTCCCGCCATCTGATGAGCAGTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAA
 TAACTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACT
 CCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCCTGACG
 CTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCCTGCGAAGTACCCATCAGGGCCTGAG
 CTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGTTAAAGGGGCGGCCGCTCAAGAGGATGTC
 AGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATTTGTAACCTATATA
 GTATAGGATTTTTTTTTGTCATTTTGTTTCTTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCG
 CAGCAGATGAATATCTTGTGGTAGGGGTTTGGGAAAATCATTCGAGTTTGATGTTTTTCTTGGTA
 TTTCCCACTCCTCTTCAGAGTACAGAAGATTAAGTGAAACCTTCGTTTGTGCGTGTTCTTTCTTG
 CGGTACCCAGATCCAATTCCCGCTTTGACTGCCTGAAATCTCCATCGCCTACAATGATGACATTT
 GGATTTGGTTGACTCATGTTGGTATTGTGAAATAGACGCAGATCGGGAACACTGAAAAATACACA
 GTTATTATTCAATTTAAATAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTTGCCATCCGACA
 TCCACAGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACC
 GTTGCAAACGCAGGACCTCCACTCCTCTTCTCCTCAACACCCACTTTTGCCATCGAAAAACCAGC
 CCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCTACTAACACCATG
 ACTTTATTAGCCTGTCTATCCTGGCCCCCCTGGCGAGGTTTCATGTTTGTGTTATTTCCGAATGCAA
 CAAGCTCCGCATTACACCCGAACATCACTCCAGATGAGGGCTTCTGAGTGTGGGGTCAAATAGT
 TTCATGTTCCCCAAATGGCCCCAAACTGACAGTTTAAACGCTGTCTTGGAACCTAATATGACAAA
 AGCGTGATCTCATCCAAGATGAACTAAGTTTGGTTTCGTTGAAATGCTAACGGCCAGTTGGTCAAA
 AAGAAACTTCCAAAAGTCGGCATACCGTTTGTCTTGTGTTGGTATTGATTGACGAATGCTCAAAAA
 TAATCTCATTAATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGAAA

CGCAAATGGGGAAACACCCGCTTTTTGGATGATTATGCATTGTCTCCACATTGTATGCTTCCAAG
ATTCTGGTGGGAATACTGCTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTA
CTTGACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTATCATCATT
TTAGCTTACTTTTCATAATTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAAC
GACAACTTGAGAAGATCAAAAAACAATAATTATTGAAAGAATTCCGAAACGATGAGATTCCCAT
CTATTTTCACCGCTGTCTTGTTCGCTGCCTCCTCTGCATTGGCTGCCCTGTTAACACTACCACT
GAAGACGAGACTGCTCAAATTCAGCTGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTT
CGACGTCGCTGTTTTGCCTTTCTCTAACTCCACTAACAACGGTTTGTTGTTTCATTAACACCACTA
TCGCTTCCATTGCTGCTAAGGAAGAGGGTGTCTCTCTCGAGAAAAGAGAGGCCGAAGCT

Secuencia 4. Cebador directo usado para linealizar la secuencia 1 para ensamblaje de Gibson en la secuencia 2.

GTCTCTCTCGAGAAAAGAGAGGCCGAAGCT SAGGTGCAGCTGGTGGAG

Secuencia 5. Cebador inverso usado para linealizar la secuencia 1 para ensamblaje de Gibson en la secuencia 2.

CAACTGCTCATCAGATGGCGGGAAGATGAA GACAGATGGTGCAGCCACAGT

Secuencia 6. Sonda para capturar Ig humana de cadena pesada

CTGCCACCTGCTCTTGTCCACGGTGAGCTTGCTGT

Secuencia 7. Sonda para capturar Ig humana de cadena ligera

TGATGGGTGACTTCGAGGCGTAGAGTTTGTGTTT

Secuencia 8. Cebador V de IgK exterior

GGACTGGACATCCAGWTGACCCAGTCT

Secuencia 9. Cebador C de IgK interior

GCCGCCGCTGGAACCTTGACCTAGAGGATCCGCC GACAGATGGTGCAGCCACAGT

Secuencia 10. Cebador V de IgG interior

AGGTCAAGTTCCAGCGGCGGCGGTGGCAGCGGAGGCGGCGGT SAGGTGCAGCTGGTGGAG

Secuencia 11. Cebador C de IgG exterior

CCRYGGCTTTGTCTTGGCAT

Secuencia 12, pIC9_IgG1

AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCA
TTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAACGCA
GGACCTCCACTCCTCTTCTCCTCAACACCCACTTTTGCCATCGAAAAACCAGCCAGTTATTGGG
CTTGATTGGAGCTCGCTCATTCCAATTCTTCTATTAGGCTACTAACACCATGACTTTATTAGCC
TGTCTATCCTGGCCCCCTGGCGAGGTTTCATGTTTGTATTATTCGAATGCAACAAGCTCCGCAT
TACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCC
AAATGGCCCCAAACTGACAGTTTAAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCA
TCCAAGATGAACTAAGTTTGGTTTCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCA
AAAGTCGCCATACCGTTTGTCTTGTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAA

TGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAATGGGGA
 AACACCCCGCTTTTTGGATGATTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGA
 ATACTGCTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAAT
 ATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTATCATCATTATTAGCTTACTTT
 CATAATTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGACAACCTTGAGA
 AGATCAAAAAACAATAATTATTTCGAAGGATCCAAACGATGAGATTTCTTCAATTTTTACTGCA
 GTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGC
 ACAAATTCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTT
 TGCCATTTTCCAACAGCACAAATAACGGGTATTGTTTATAAATACTACTATTGCCAGCATTGCT
 GCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTCCCCTGGCACCCCTCTCCAA
 GAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTGA
 CGGTGTCGTGGAACCTCAGGCGCCCTGACCAGCGGCGTGCACACCTTCCCGGCTGTCTACAGTCC
 TCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTA
 CATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTT
 GTGACAAAACCTCACACATGCCACCGTGCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTC
 CTCTTCCCCCCCCAAAACCCAAGGACACCCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGT
 GGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGC
 ATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTC
 ACCGTCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCT
 CCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACA
 CCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTGAGCCTGACCTGCCTGGTCAAAGGC
 TTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGAGAACAACTACAAGAC
 CACGCTCCCGTGTGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGA
 GCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTAC
 ACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATAATTCGCCTTAGACATGACTGTTTCTCAGTT
 CAAGTTGGGCACCTTACGAGAAGACCGGTCTTGCTAGATTCTAATCAAGAGGATGTCAGAATGCCA
 TTTGCCGTGAGAGATGCAGGCTTCATTTTTTGATACTTTTTTATTTGTAACCTATATAGTATAGGAT
 TTTTTTTGTCATTTTGTCTTCTCTCGTACGAGCTTGCTCCTGATCAGCCTATCTCGCAGCTGATG
 AATATCTTGTGGTAGGGGTTTGGGAAAATCATTCGAGTTTGATGTTTTTCTTGGTATTTCCCACT
 CCTCTTCAGAGTACAGAAGATTAAGTGAGAAGTTTCGTTTGTGCAAGCTTATCGATAAGCTTTAAT
 GCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAATGCG
 CTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTGGTTATGCCGGTACT
 GCCGGGCCCTCTTGCGGGATATCGTCCATTCCGACAGCATCGCCAGTCACTATGGCGTGTGCTAG
 CGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTCTCTCGGAGCACTGTCCGACCGCTTTGGC
 CGCCGCCAGTCCCTGCTCGCTTTCGCTACTTGAGCCACTATCGACTACGCGATCATGGCGACCAC
 ACCCGTCTGTGGATCTATCGAATCTAAATGTAAGTTAAATCTCTAAATAATTAAATAAGTCCC
 AGTTTCTCCATACGAACCTTAACAGCATTGCGGTGAGCATCTAGACCTTCAACAGCAGCCAGATC
 CATCACTGCTTGGCCAATATGTTTCAGTCCCTCAGGAGTTACGTCTTGTGAAGTGATGAACTTCT
 GGAAGGTTGCAGTGTTAACTCCGCTGTATTGACGGGCATATCCGTACGTTGGCAAAGTGTTGGTTG
 GTACCGGAGGAGTAATCTCCACAACCTCTCTGGAGAGTAGGCACCAACAAACACAGATCCAGCGTG
 TTGTACTTGATCAACATAAGAAGAAGCATTCTCGATTTGCAGGATCAAGTGTTTCAGGAGCGTACT
 GATTGGACATTTCCAAAGCCTGCTCGTAGGTTGCAACCGATAGGGTTGTAGAGTGTGCAATACAC
 TTGCGTACAATTTCAACCCTTGGCAACTGCACAGCTTGGTTGTGAACAGCATCTTCAATTCTGGC

AAGCTCCTTGTCTGTCATATCGACAGCCAACAGAATCACCTGGGAATCAATACCATGTTTCAGCTT
GAGACAGAAGGTCTGAGGCAACGAAATCTGGATCAGCGTATTTATCAGCAATAACTAGAACTTCA
GAAGGCCCAGCAGGCATGTCAATACTACACAGGGCTGATGTGTCATTTTGAACCATCATCTTGGC
AGCAGTAACGAACTGGTTTTCTGGACCAAATATTTTGTACACTTAGGAACAGTTTCTGTTCCGT
AAGCCATAGCAGCTACTGCCTGGGCGCCTCCTGCTAGCACGATACACTTAGCACCAACCTTGTGG
GCAACGTAGATGACTTCTGGGGTAAGGGTACCATCCTTCTTAGGTGGAGATGCAAAAACAATTTT
TTTGCAACCAGCAACTTTGGCAGGAACACCCAGCATCAGGGAAGTGGAAGGCAGAATTGCGGTTC
CACCAGGAATATAGAGGCCAACTTTCTCAATAGGTCTTGCAAAACGAGAGCAGACTACACCAGGG
CAAGTCTCAACTTGCAACGTCTCCGTTAGTTGAGCTTCATGGAATTTCTGACGTTATCTATAGA
GAGATCAATGGCTCTCTTAACGTTATCTGGCAATTGCATAAGTTCCTCTGGGAAAGGAGCTTCTA
ACACAGGTGTCTTCAAAGCGACTCCATCAAACCTGGCAGTTAGTTCTAAAAGGGCTTTGTACCA
TTTTGACGAACATTGTGACAATTGGTTTTGACTAATTCCATAATCTGTTCCGTTTTCTGGATAGG
ACGACGAAGGGCATCTTCAATTTCTTGTGAGGAGGCCTTAGAAACGTCAATTTTGCACAATTCAA
TACGACCTTCAGAAGGGACTTCTTTAGGTTTGGATTCTTCTTTAGGTTGTTCTTGGTGTATCCT
GGCTTGGCATCTCCTTTCTTCTAGTGACCTTTAGGGACTTCATATCCAGGTTTCTCTCCACCTC
GTCCAACGTACACCGTACTTGGCACATCTAACTAATGCAAAAATAAAATAAGTCAGCACATTCCC
AGGCTATATCTTCTTGGATTTAGCTTCTGCAAGTTTCATCAGCTTCCTCCCTAATTTTAGCGTTC
AACAAAACCTCGTCGTCAAATAACCGTTTGGTATAAGAACCTTCTGGAGCATTGCTCTTACGATC
CCACAAGGTGGCTTCCATGGCTCTAAGACCCTTTGATTGGCCAAAACAGGAAGTGCGTTCCAAGT
GACAGAAAACCAACACCTGTTTGTTCACCACAAATTTCAAGCAGTCTCCATCACAATCCAATTG
ATACCCAGCAACTTTTGAGTTGCTCCAGATGTAGCACCTTTATACCACAAACCGTGACGACGAGA
TTGGTAGACTCCAGTTTGTGTCCTTATAGCCTCCGGAATAGACTTTTTGGACGAGTACACCAGGC
CCAACGAGTAATTAGAAGAGTCAGCCACCAAAGTAGTGAATAGACCATCGGGGCGGTGAGTAGTC
AAAGACGCCAACAAAATTTCACTGACAGGGAACTTTTGACATCTTCAGAAAGTTGCTATTAGT
AGTCAATTGCCGAGCATCAATAATGGGGATTATACCAGAAGCAACAGTGGAAGTCACATCTACCA
ACTTTGCGGTCTCAGAAAAAGCATAAACAGTTCTACTACCGCCATTAGTGAACTTTTCAAATCG
CCCAGTGGAGAAGAAAAAGGCACAGCGATACTAGCATTAGCGGGCAAGGATGCAACTTTATCAAC
CAGGGTCTTATAGATAACCCTAGCGCCTGGGATCATCCTTTGGACAACCTTTTCTGCCAAATCTA
GGTCCAAAATCACTTCATTGATACCATTATTGTACAACCTTGAGCAAGTTGTGATCAGCTCCTCA
AATTGGTCCTCTGTAACGGATGACTCAACTTGCACATTAACTTGAAGCTCAGTCGATTGAGTGAA
CTTGATCAGGTTGTGTCAGCTGGTCAGCAGCATAGGGAAACACGGCTTTTCTACCAAACCTCAAGG
AATTATCAAACCTCTGCAACACTTGCGTATGCAGGTAGCAAGGGAAATGTCATACTTGAAGTCGGA
CAGTGAGTGTAGTCTTGAGAAATTCTGAAGCCGTATTTTATTATCAGTGAGTCAGTCATCAGGA
GATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCT
ATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTT
GGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGTCATGCACC
ATTCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGT
CGCATAAGGGAGAGCGTCGAGTATCTATGATTGGAAGTATGGGAATGGTGATACCCGCATTCTTC
AGTGTCTTGAGGTCTCCTATCAGATTATGCCCAACTAAAGCAACCGGAGGAGGAGATTTTCATGGT
AAATTTCTCTGACTTTTGGTCATCAGTAGACTCGAAGTGTGAGACTATCTCGGTTATGACAGCAG
AAATGTCTTCTTGGAGACAGTAAATGAAGTCCCAACATAAAGAAATCCTTGTTATCAGGAACA
AACTTCTTGTTCGAACTTTTTCGGTGCCTTGAAGTATAAAATGTAGAGTGGATATGTCGGGTAG
GAATGGAGCGGGCAAATGCTTACCTTCTGGACCTTCAAGAGGTATGTAGGGTTTGTAGATACTGA

TGCCAACTTCAGTGACAACGTTGCTATTTTCGTTCAAACCATTCGGAATCCAGAGAAATCAAAGTT
 GTTTGTCTACTATTGATCCAAGCCAGTGCGGTCTTGAAACTGACAATAGTGTGCTCGTGTTTTGA
 GGTCACTCTTTGTATGAATAAATCTAGTCTTTGATCTAAATAATCTTGACGAGCCAAGGCGATAAA
 TACCCAAATCTAAACTCTTTTAAACGTTAAAGGACAAGTATGTCTGCCTGTATTAAACCCCA
 AATCAGCTCGTAGTCTGATCCTCATCAACTTGAGGGGCACTATCTTGTTTTAGAGAAATTTGCGG
 AGATGCGATATCGAGAAAAAGGTACGCTGATTTTAAACGTGAAATTTATCTCAAGATCTCTGCCT
 CGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTT
 GTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGT
 CGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAAGTATGCGGCA
 TCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAG
 AAAATACCGCATCAGGCGCTCTTCCGCTTCTCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCCGC
 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC
 GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCT
 GCGGTTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT
 GCGGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGCGCT
 TTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTG
 TGCACGAACCCCCCGTTAGCCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC
 CCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTA
 TGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTAT
 TTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGC
 AAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAA
 AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC
 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAA
 TGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAAT
 CAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCTG
 TGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
 CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAG
 TGGTCTTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTA
 GTTCGCCAGTTAATAGTTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCG
 TCGTTTGGTATGGCTTCATTACGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCAT
 GTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAG
 TGTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCTATGCCATCCGTAAGATGC
 TTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTG
 CTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA
 TTGGAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATG
 TAACCCACTCGTGCACCCAAGTATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGC
 AAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCA
 TACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATA
 TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACC
 TGACGCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCT
 TTCGTCTTCAAGAATTAATTCTCATGTTTGACAGCTTATCATCGATAAGCTGACTCATGTTGGTA
 TTGTGAAAATAGACGCAGATCGGGAACACTGAAAAATAACAGTTATTATTCG

Secuencia 13, scFv de adalumimab

GACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGGGACAGAGTCACCATCAC
 TTGTCGGGCAAGTCAGGGCATCAGAAATTACTTAGCCTGGTATCAGCAAAAACCAGGGAAAGCCC
 CTAAGCTCCTGATCTATGCTGCATCCACTTTGCAATCAGGGGTCCCATCTCGGTTTCAGTGGCAGT
 GGATCTGGGACAGATTTCACTCTCACCATCAGCAGCCTACAGCCTGAAGATGTTGCAACTTATTA
 CTGTCAAAGGTATAACCGTGCACCGTATACTTTTGGCCAGGGGACCAAGGTGGAAATCAAACGAA
 CTGTGGCTGCACCATCTGTGCGCGGATCCTCTAGGTCAAGTTCCAGCGGCGGCGGTGGCAGCGGA
 GCGGCGGTGAGGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCCGGCAGGTCCCTGAG
 ACTCTCCTGTGCGGCCTCTGGATTACCTTTGATGATTATGCCATGCACTGGGTCCGGCAAGCTC
 CAGGGAAGGGCCTGGAATGGGTCTCAGCTATCACTTGGAATAGTGGTCACATAGACTATGCGGAC
 TCTGTGGAGGGCCGATTACCATCTCCAGAGACAACGCCAAGAACTCCCTGTATCTGCAAATGAA
 CAGTCTGAGAGCTGAGGATACGGCCGTATATTACTGTGCGAAAAGTCTCGTACCTTAGCACCGCGT
 CCTCCCTTGACTATTGGGGCCAAGGTACCCTGGTCACCGTCTCGAGTGCCTCCACCAAGGGCCCA
 TCGGTCTTC

REIVINDICACIONES

1. Un método para preparar una pluralidad de constructos de expresión de inmunoglobulinas recombinantes, que comprenden:

proporcionar una pluralidad de polinucleótidos de fusión recombinantes que comprenden, cada uno, un primer polinucleótido que codifica para un primer dominio variable, un segundo polinucleótido que codifica para un segundo dominio variable, y un polinucleótido ligador que liga los polinucleótidos primero y segundo;

expresar una pluralidad de polipéptidos de fusión recombinantes codificados por dicha pluralidad de polinucleótidos de fusión recombinantes para presentar dicha pluralidad de polipéptidos de fusión recombinantes en una pluralidad de superficies;

enriquecer dicha pluralidad de polinucleótidos de fusión recombinantes para unirse a un antígeno exponiendo dicha pluralidad de superficies a dicho antígeno y seleccionar en base a la unión entre dicho antígeno y cada uno de dicha pluralidad de polipéptidos de fusión recombinantes;

circularizar mediante ensamblaje de Gibson uno o más de dichos polinucleótidos de fusión recombinantes enriquecidos; e

insertar un tercer polinucleótido que comprende una secuencia que codifica para un promotor y una secuencia que codifica para una región constante entre dichos polinucleótidos primero y segundo en cada uno de dichos polinucleótidos de fusión recombinantes enriquecidos circularizados, generando de ese modo una pluralidad de constructos de expresión de inmunoglobulinas recombinantes.
2. El método según la reivindicación 1, que comprende además las etapas anteriores de:

proporcionar células inmunitarias primarias obtenidas de al menos un donante de mamífero; y

encapsular dichas células inmunitarias primarias en una pluralidad de gotas con perlas que comprenden sondas de ácido polinucleico unidas con secuencias que son complementarias a ambas, polinucleótidos que comprenden secuencias que codifican para dominios variables de cadena pesada y polinucleótidos que comprenden secuencias que codifican para dominios variables de cadena ligera en una mezcla de reacción que lisa las células inmunitarias primarias y estimula la hibridación de ácido polinucleico, permitiendo así que las perlas capturen polinucleótidos que codifican para dominios variables de cadena pesada y polinucleótidos que codifican para dominios variables de cadena ligera.
3. El método según la reivindicación 1 ó 2, que comprende además:

insertar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en una pluralidad de células hospedadoras; y expresar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en dicha pluralidad de células hospedadoras, generando de ese modo una biblioteca de inmunoglobulinas recombinantes, en la que cada inmunoglobulina recombinante comprende un par análogo de dominios variables de cadena pesada y de cadena ligera de una célula individual aislada.
4. El método según una de las reivindicaciones 1 a 3, en el que dichos dominios variables primero y segundo son de una célula individual.
5. El método según la reivindicación 1, en el que dicho primer dominio variable es de una cadena pesada de inmunoglobulina y dicho segundo dominio variable es de una cadena ligera de inmunoglobulina.
6. El método según la reivindicación 1, en el que dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes comprenden al menos 1.000, 10.000 ó 100.000 pares análogos únicos de secuencias codificantes de cadena pesada y de cadena ligera.
7. Un método para generar una biblioteca de inmunoglobulinas, que comprende:

generar una pluralidad de polinucleótidos de fusión recombinantes que comprenden, cada uno, un primer polinucleótido que codifica para un primer dominio variable y un segundo polinucleótido que codifica para un segundo dominio variable, en el que dicho primer dominio variable y dicho segundo dominio variable de dicha pluralidad de polinucleótidos de fusión recombinantes son un par análogo de una célula inmunitaria aislada individual, en el que cada uno de dicha pluralidad de polinucleótidos de fusión recombinantes comprende además un polinucleótido ligador que liga dichos polinucleótidos primero y segundo;

circularizar mediante ensamblaje de Gibson cada una de dicha pluralidad de los polinucleótidos de fusión

recombinantes;

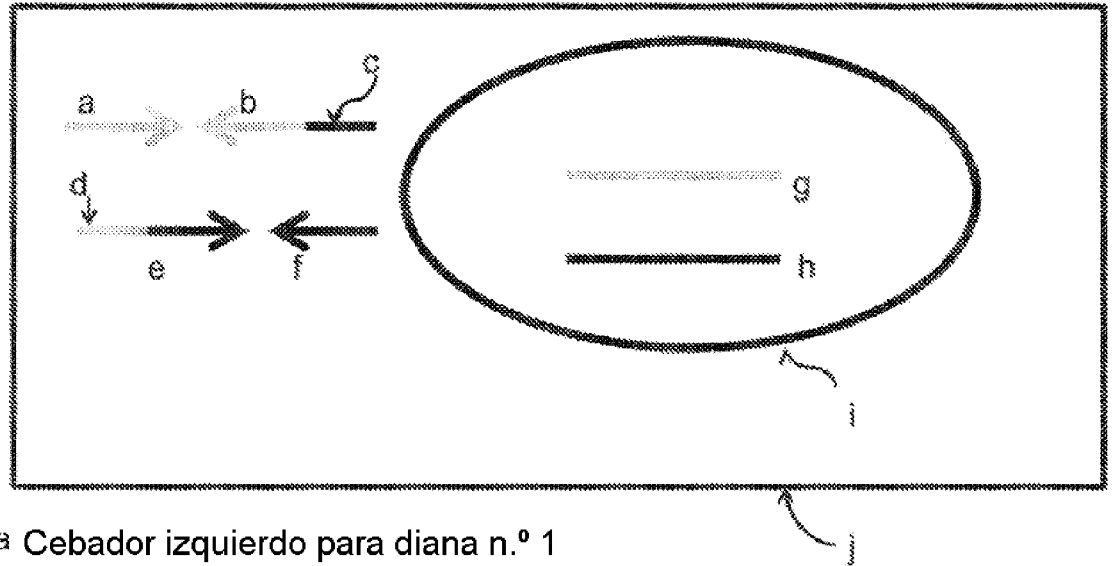
insertar un tercer polinucleótido que comprende una secuencia que codifica para un promotor y una secuencia que codifica para una región constante entre dichos polinucleótidos primero y segundo en cada uno de dichos polinucleótidos de fusión recombinantes circularizados,

generando de ese modo una pluralidad de constructos de expresión de inmunoglobulinas recombinantes;

insertar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en una pluralidad de células hospedadoras; y

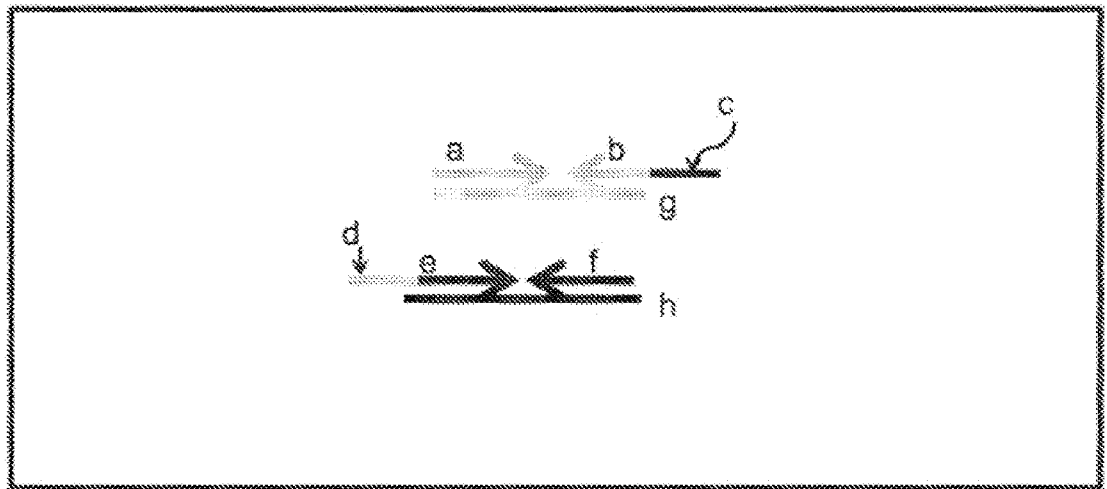
expresar dicha pluralidad de constructos de expresión de inmunoglobulinas recombinantes en dicha pluralidad de células hospedadoras, generando de ese modo una biblioteca de inmunoglobulinas recombinantes que comprende una pluralidad de inmunoglobulinas recombinantes, en el que cada una de dicha pluralidad de inmunoglobulinas recombinantes comprende un par análogo de dominios variables de cadena pesada y de cadena ligera de una célula individual aislada.

8. El método según la reivindicación 7, en el que dicha inserción de dicho tercer polinucleótido se realiza en paralelo para dicha pluralidad de constructos de polinucleótido circularizados.
9. El método según la reivindicación 7, en el que al menos una inmunoglobulina recombinante de dicha biblioteca de inmunoglobulinas recombinantes se une específicamente a un epítipo de *Streptococcus pneumoniae*, un epítipo de *Hemophilus influenza* o un epítipo de *Klebsiella pneumoniae*.
10. El método según la reivindicación 7, en el que dicho primer dominio variable es de una cadena pesada de inmunoglobulina y dicho segundo dominio variable es de una cadena ligera de inmunoglobulina.



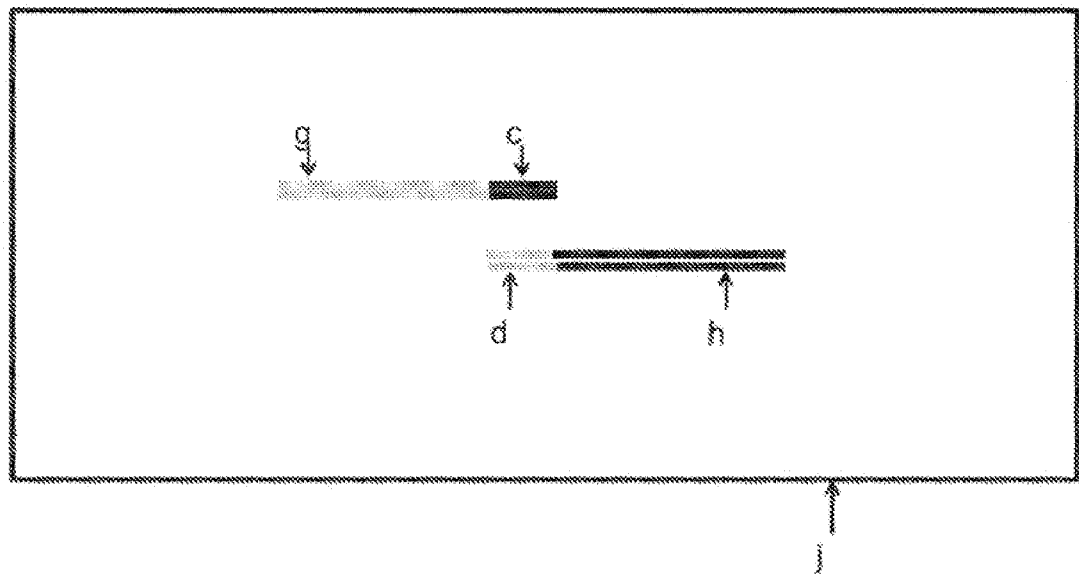
- a Cebador izquierdo para diana n.º 1
- b Cebador derecho para diana n.º 1
- c Región de cebador complementaria a d
- d Región de cebador complementaria a c
- e Cebador izquierdo para diana n.º 2
- f Cebador derecho para diana n.º 2
- g Diana de poli(ácido nucleico) n.º 1
- h Diana de poli(ácido nucleico) n.º 2
- i Membrana celular
- j Contenedor de reacción físico o gota de emulsión

Figura 1A



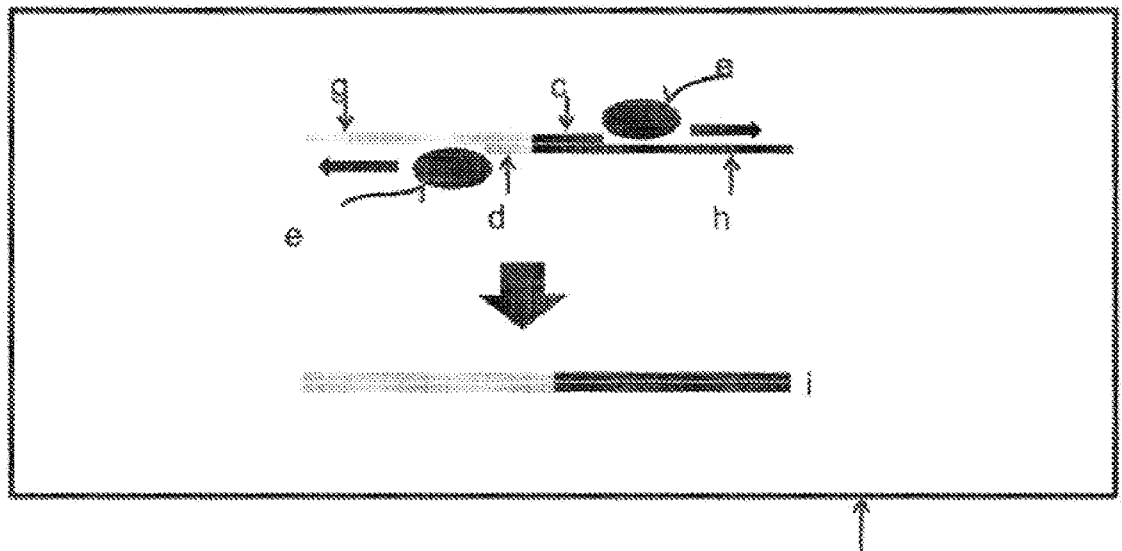
- a Cebador izquierdo para diana n.º 1
- b Cebador derecho para diana n.º 1
- c Región de cebador complementaria a d
- d Región de cebador complementaria a c
- e Cebador izquierdo para diana n.º 2
- f Cebador derecho para diana n.º 2
- g Diana de poli(ácido nucleico) n.º 1
- h Diana de poli(ácido nucleico) n.º 2
- j Contenedor de reacción físico o gota de emulsión

Figura 1B



- c Región producto complementaria a d
- d Región producto complementaria a c
- g Diana de poli(ácido nucleico) n.º 1
- h Diana de poli(ácido nucleico) n.º 2
- j Contenedor de reacción físico o gota de emulsión

Figura 1C



- c Región producto complementaria a d
- d Región producto complementaria a c
- e Polimerasa
- g Diana de poli(ácido nucleico) n.º 1
- h Diana de poli(ácido nucleico) n.º 2
- i Producto fusionado entre dianas de poli(ácido nucleico) n.º 1 y n.º 2
- j Contenedor de reacción físico o gota de emulsión

Figura 1D

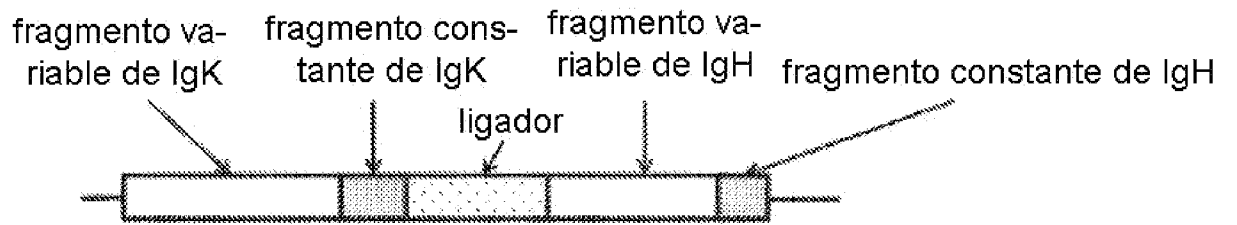


Figura 2

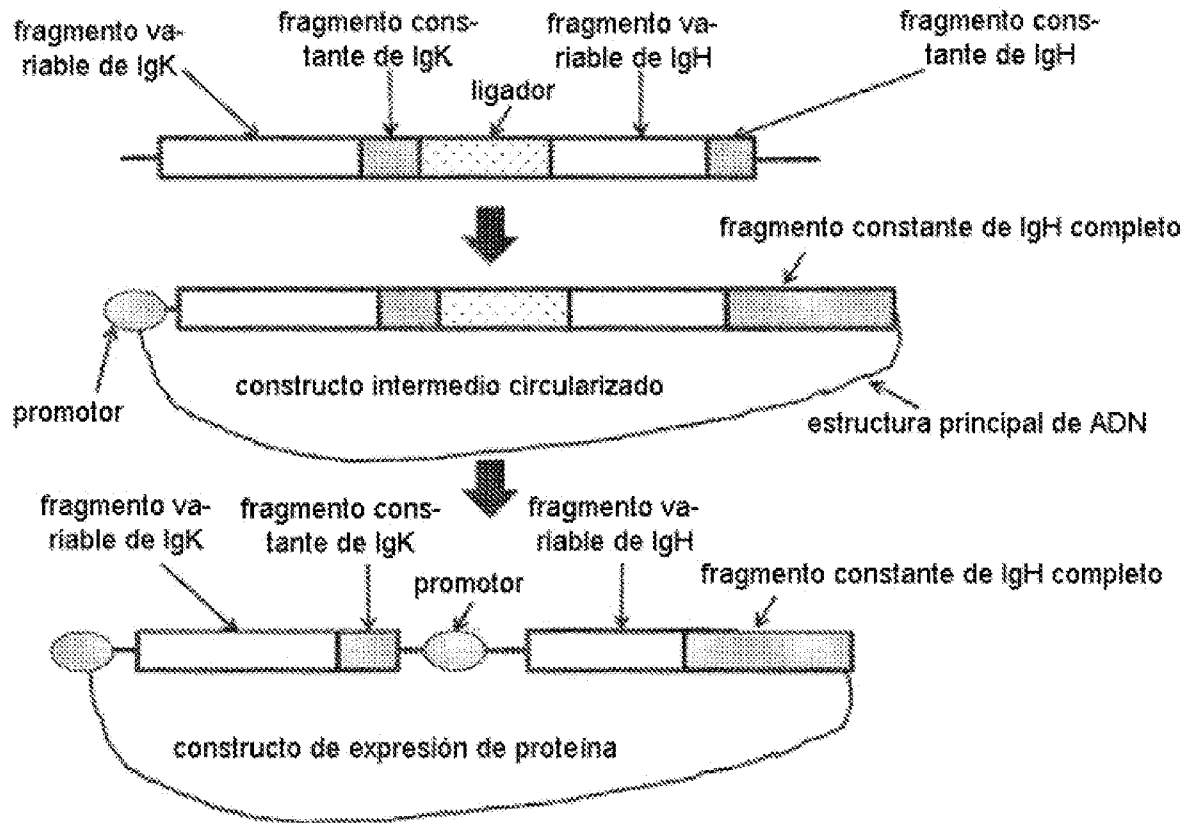


Figura 3

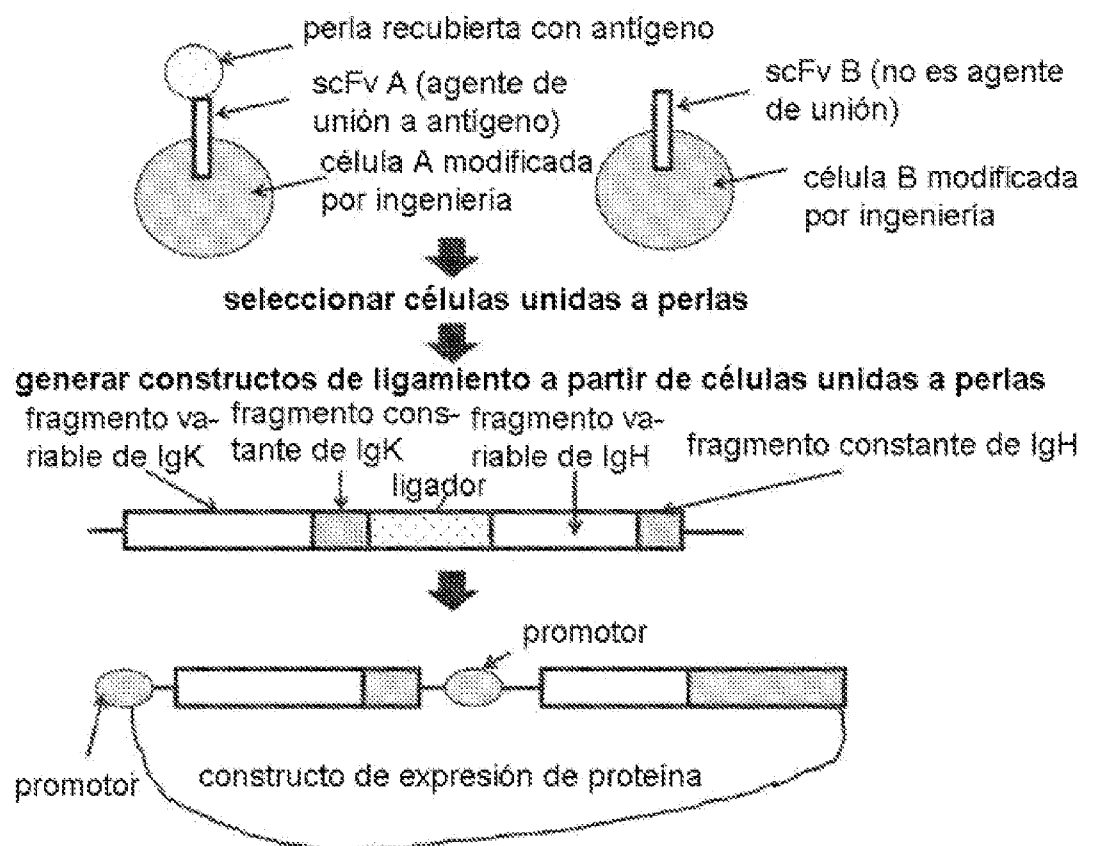


Figura 4

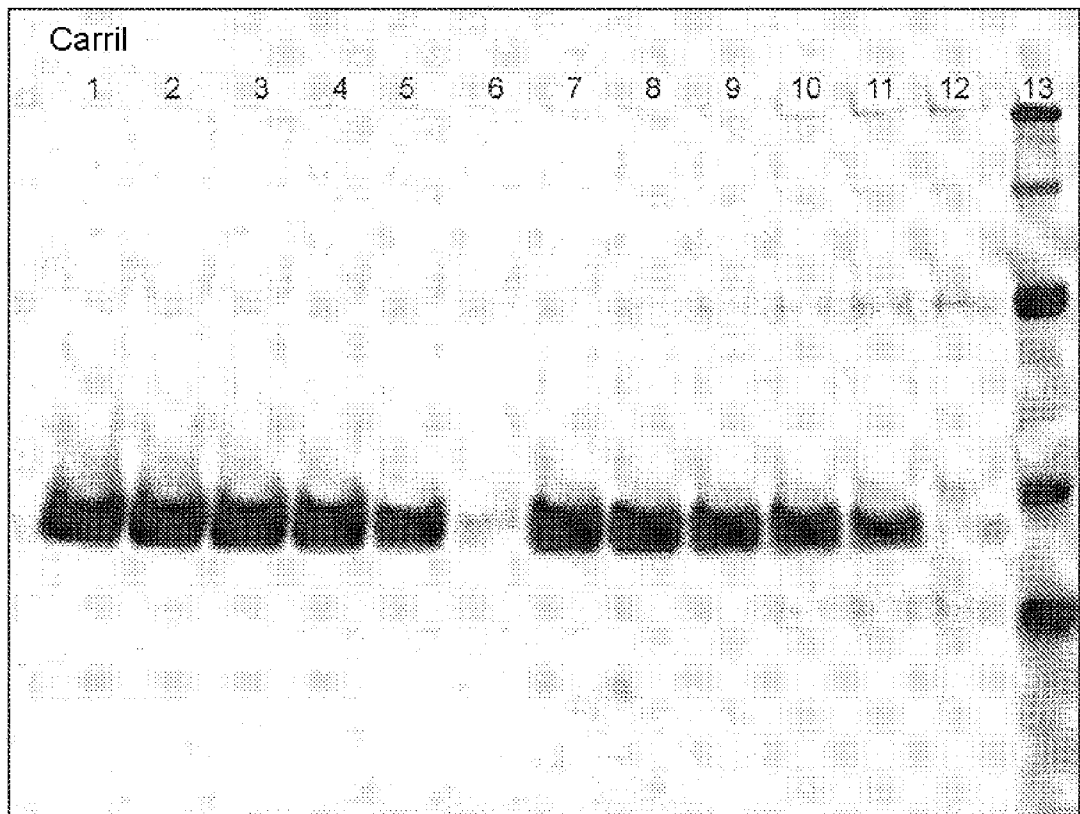


FIG. 5

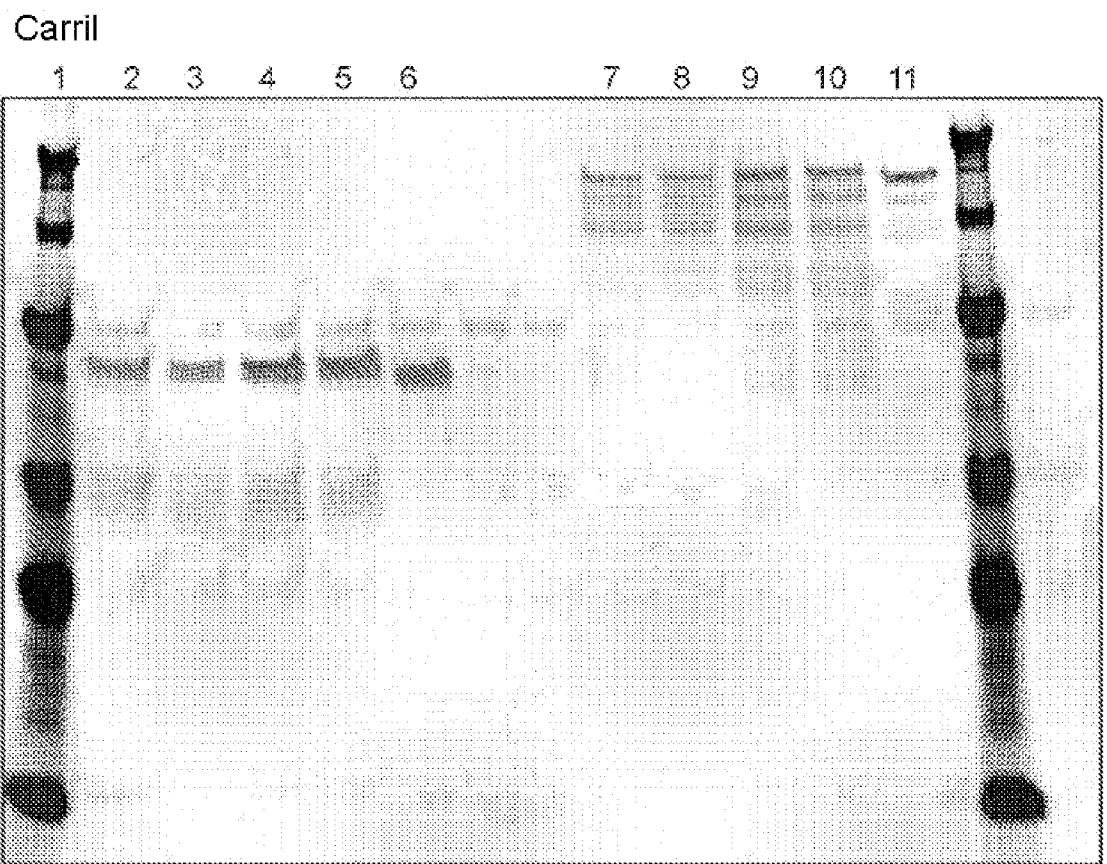


FIG. 6

REFERENCIAS CITADAS EN LA DESCRIPCION

La lista de referencias citadas por el solicitante es, únicamente, para conveniencia del lector. No forma parte del documento de patente europea. Si bien se ha tenido gran cuidado al recopilar las referencias, no pueden excluirse errores u omisiones y la OEP declina toda responsabilidad a este respecto.

Documentos de patente citados en la descripción

- US 7749697 B [0068]
- WO 2012083225 A [0068]
- WO 2013096643 A [0068]
- US 1992001478 W [0072]

Literatura no patente citada en la descripción

- CARTER. *Nature Reviews Immunology*, 2006, vol. 6, 343-357 [0002] [0089]
- Intracellular Antibodies: Research and Disease Applications. Springer-Verlag New York, Inc, 1998 [0025]
- KINTSES et al. *Current Opinion in Chemical Biology*, 2010, vol. 14, 548-555 [0057]
- HUNT et al. *Lab Chip*, 2008, vol. 8, 81-87 [0057]
- ANNA et al. *Applied Physics Letters*, 2003, vol. 82, 364-366 [0057]
- JOHNSTON et al. *Science*, 1996, vol. 271, 624-626 [0057]
- ABATE et al. *Lab Chip*, 2009, vol. 9, 2628-31 [0057] [0059]
- EDD et al. *Lab Chip*, 2008, vol. 8, 1262-1264 [0059]
- KOJIMA et al. *Nucleic Acids Research*, 2005, vol. 33, e150 [0063]
- MARCUS et al. *Anal Chem*, 2006, vol. 78, 3084-3089 [0064]
- DEKOSKY et al. *Nat Med*, 2015, vol. 21, 86-91 [0066]
- JOHNSON et al. *Genome Research*, 2005, vol. 15, 1315-24 [0066]
- KOSURI ; CHURCH. *Nat Methods*, 2014, vol. 11, 499-507 [0077]
- HANES et al. *PNAS*, 1997, vol. 94, 4937-42 [0077]
- MATTHEAKIS et al. *Affymax Research Institute*, 1994, vol. 91, 9022-6 [0077]
- GIBSON et al. *Science*, 2008, vol. 319, 1215-20 [0079]
- SMITH. *Science*, 1985, vol. 228, 1315-17 [0083]
- MCCAFFERTY et al. *Nature*, 1990, vol. 348, 552-554 [0083]
- LUNDER et al. *J Lipid Research*, 2005, vol. 46, 1512-1516 [0083]
- BRATKOVIC et al. *Biochem Biophys Res Commun*, 2005, vol. 332, 897-903 [0083]
- HUFTON et al. *J. Immunol Methods*, 1999, vol. 231, 39-51 [0083]
- GOMMANS et al. *J Mol Biol*, 2005, vol. 354, 507-519 [0083]
- BODER ; WITTRUP. *Nat Biotech*, 1997, vol. 15, 553-57 [0084]
- DANGA et al. *Cancer Res*, 2013, vol. 73, 4820-4829 [0084]
- HO et al. *PNAS*, 2006, vol. 103, 25 [0084]
- URBAN et al. *Nucleic Acids Research*, 2005, vol. 33, e35 [0084]
- WOLKOWICZ et al. *J Biol Chem*, 2005, vol. 280, 15195-15201 [0084]
- ANDERSON et al. *Curr Opin Biotechnol*, 2004, vol. 15, 456-462 [0084]
- HANES ; PLUCKTHUN. *PNAS*, 1997, vol. 94, 4937-4942 [0085]
- KAMEYAMA et al. *Biotechnol Bioeng*, 2010, vol. 105, 1106-1114 [0086]
- WIBERG et al. *Biotech Bioeng*, 2006, vol. 94, 396-405 [0086]
- KENNARD et al. *Biotechnol Bioeng*, 2009, vol. 104, 540-553 [0086]
- JACOBS et al. *Nat Protoc*, 2009, vol. 4, 58-70 [0087]
- CREGG et al. *Methods in Enzymology*, 2011, vol. 463, 169-187 [0087]
- T.E. CREIGHTON. *Proteins: Structures and Molecular Properties*. W.H. Freeman and Company, 1983 [0101]
- A.L. LEHNINGER. *Biochemistry*. Worth Publishers, Inc, [0101]
- SAMBROOK et al. *Molecular Cloning: A Laboratory Manual*. 1989 [0101]
- *Methods In Enzymology*. Academic Press, Inc, [0101]
- *Remington's Pharmaceutical Sciences*. Mack Publishing Company, 1990 [0101]
- CAREY ; SUNDBERG. *Advanced Organic Chemistry*. Plenum Press, 1992, vol. A, B [0101]