



(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 268 907**

(51) Int. Cl.:

**C12N 15/10** (2006.01)

**C12N 15/13** (2006.01)

**C12N 15/55** (2006.01)

**C12N 5/10** (2006.01)

**C07K 16/00** (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **99972677 .1**

(86) Fecha de presentación : **18.11.1999**

(87) Número de publicación de la solicitud: **1131421**

(87) Fecha de publicación de la solicitud: **12.09.2001**

(54) Título: **Métodos para la preparación de ácidos nucleicos y polipéptidos mediante la recombinación *in vivo*, y utilizaciones de las mismas.**

(30) Prioridad: **19.11.1998 IT MI98A2509**

(45) Fecha de publicación de la mención BOPI:  
**16.03.2007**

(45) Fecha de la publicación del folleto de la patente:  
**16.03.2007**

(73) Titular/es: **S.I.S.S.A. Scuola Internazionale  
Superiore di Studi Avanzati  
Via Beirut, 2-4  
34014 Trieste, IT**

(72) Inventor/es: **Bradbury, Andrew, Raymon, Morton y  
Sblattero, Daniele**

(74) Agente: **Durán Moya, Carlos**

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

Métodos para la preparación de ácidos nucleicos y polipéptidos mediante la recombinación *in vivo*, y utilidades de las mismas.

## Sector al que pertenece la invención

La presente invención se refiere a un procedimiento para la generación de diversidad genética a través de la preparación de bibliotecas de genes, y se refiere asimismo a la utilización de dichas bibliotecas de genes para la preparación de nuevos polipéptidos y además a los polipéptidos obtenidos de esta manera.

## Antecedentes de la invención

La creación de diversidad en poblaciones de polipéptidos se ha convertido en una herramienta importante en la derivatización de polipéptidos con características útiles. Habitualmente, estos enfoques implican métodos para crear poblaciones de moléculas diversas (por ejemplo, ácidos nucleicos, polipéptidos, etc.) acoplados a métodos para seleccionar o monitorizar los miembros de la población que poseen una o más de las características deseadas.

Se han utilizado una serie de enfoques para crear poblaciones de moléculas diversas. Entre estos métodos se incluyen mutagénesis aleatoria, mutagénesis recombinatoria, y mutagénesis direccional. En la mutagénesis aleatoria, se insertan habitualmente de manera aleatoria las mutaciones en una secuencia de ADN. Esto puede llevarse a cabo utilizando varios métodos diferentes, entre los que se incluyen mutagénesis química, cepas bacterianas mutantes y PCR que es proclive a error. En la mutagénesis recombinatoria, la recombinación sucede entre diferentes hebras de ADN de forma que las mutaciones, diferencias o genes que están presentes originariamente en diferentes cadenas de ADN se unen en la misma cadena. En la mutagénesis direccional, las mutaciones se insertan en una secuencia de ADN en los lugares en los que se conoce o se sospecha que la proteína codificada participa en determinados procesos, tales como actividades de unión o enzimáticas. Esto puede hacerse mediante mutagénesis direccional a lugar o PCR.

De estas tres formas de mutagénesis, la mutagénesis recombinatoria y la mutagénesis direccional parecen ser las más eficientes. Sin embargo, la mutagénesis direccional requiere un conocimiento estructural detallado de la proteína y/o ácido nucleico objetivo.

La mutagénesis recombinatoria, con el objetivo de preparar bibliotecas de genes solamente, se ha realizado *in vitro* utilizando mezcla (*shuffling*) de ADN (Stemmer (1994) *Proc Natl Acad Sci USA*, 91: 10747-10751). Mediante la utilización de mezcla de ADN, clonación y transfección, se han obtenido varias proteínas diferentes, entre las que se incluyen beta lactamasa, galactosidasa y anticuerpos con sus propiedades modificadas (de forma directa).

Mediante la utilización de genes estrechamente relacionados en la mezcla de ADN, en lugar de las mutaciones acumuladas durante el PCR, se han producido proteínas nuevas que tienen propiedades no encontradas en ninguna de las proteínas codificadas por los genes utilizados como sustratos para la mezcla. Entre los ejemplos se incluyen interferones y cefalosporinas (Cramer, y otros (1998) *Nature*, 391: 288-91).

Una aplicación particular en la que se crea la diversidad mediante la mezcla por combinación de dos genes diferentes, en vez de la recombinación en un único gen o grupo de genes es la creación de bibliotecas de anticuerpos expuestas sobre la superficie de un fago filamentoso o fagómico (exposición en fago) (Marks y otros (1991) *J. Mol. Biol.*, 222, 581-597).

El dominio de unión de los anticuerpos se codifica en las regiones V, con cada dominio de unión preparado a partir de un único dominio  $V_H$  y un único dominio  $V_L$ . Se ha comprobado que es posible separar los dominios V del resto de las inmunoglobulinas y crear nuevos polipéptidos de unión funcional. Los fragmentos Fv se preparan de  $V_H$  y  $V_L$ , y se mantienen unidos mediante fuerzas no covalentes. Sin embargo, estos son relativamente inestables, y las dos cadenas se separan fácilmente. La unión de las dos cadenas con un polipéptido conector crea el fragmento de cadena única Fv (scFv) que es mucho más estable.

El fragmento Fab comprende  $V_H$  y  $C_H1$  acoplados covalentemente mediante enlaces disulfuro a  $V_L$  y  $C_L$ . Este fragmento es estable, pero tiene la desventaja de que es relativamente más difícil de preparar en bacterias y tiene tendencia a agregarse. La mayoría de las bibliotecas preparadas sobre fagos filamentosos están en el formato scFv, aunque una pequeña cantidad se ha preparado en el formato Fab.

Aunque se han aislado anticuerpos contra un número elevado de antígenos utilizando bibliotecas de anticuerpos en fago (ver, por ejemplo, (Marks y otros (1991) *J. Mol. Biol.*, 222, 581-597; Griffiths, y otros (1994) *EMBO J.*, 13, 3245-3260), el procedimiento aún no se utiliza de forma amplia. Esto puede deberse a la dificultad de preparar bibliotecas grandes de anticuerpos en fago, que habitualmente se requiere la creación de una biblioteca de, como mínimo,  $10^9$  clones independientes a partir de una buena fuente de genes  $V_H$  y  $V_L$  diversos. En general, estas bibliotecas se preparan mediante la realización de un gran número de ligaciones y transfecciones, y una vez realizadas, se convierten en un recurso limitado, ya que no se puede llevar a cabo la amplificación sin una reducción potencial en la diversidad. En general, la afinidad de los anticuerpos aislados es proporcional al tamaño inicial de la biblioteca utilizada para la selección.

La recombinación específica de lugar se ha propuesto como un método alternativo a la utilización de la clonación para crear bibliotecas de genes grandes, y para la preparación de anticuerpos con nuevas especificidades. Para recombinar cadenas  $V_H$  y  $V_L$  en los formatos Fab (Griffiths y otros 1994) o scFv (Tsurushita y otros (1996) *Gene* 172: 59-63), se ha utilizado recombinasa cre mientras que en otro caso se ha utilizado recombinasa lambda para recombinar Fab (Geoffroy y otros, (1994) *Gene* 151: 109-113).

Aunque la recombinación entre anticuerpos únicos para reconstituir regiones de unión funcional se ha descrito en todos estos informes, una biblioteca grande de anticuerpos en fago solamente se ha preparado utilizando recombinasa cre para recombinar Fabs (es decir, recombinación específica de lugar) descrita en la Patente W093/19172 y (Griffiths, y otros (1994) *EMBO J.*, 13: 3245-3260).

En ambas publicaciones, se preparó una biblioteca de genes diversos utilizando recombinación específica de lugar entre un primer vector que contiene el gen  $V_H$  y un segundo vector que contiene el gen  $V_L$ , con el vector resultante recombinado que contiene ambos genes  $V_H$  y  $V_L$ . Estos sistemas utilizaron dos vectores diferentes para la recombinación.

La introducción de los vectores de diferente expresión que portan los genes que van a recombinarse en la célula, sin embargo, planteó serios problemas para la creación de bibliotecas diversas grandes. Habitualmente, los dos vectores tenían que ser diferentes (por ejemplo, un vector es un plásmido, el otro vector un fago). Esto es debido a que se había descrito que bacterias infectadas por un fago filamentosos son resistentes a una infección posterior por otro fago filamentosos (Boeke y otros (1982) *Mol Gen Genet* 186: 185-192). De este modo, uno de los dos vectores tuvo que someterse a transfección en lugar de infección y esto disminuyó drásticamente la eficiencia de transformación de la célula.

Adicionalmente, se creía que las dos construcciones requerían diferentes orígenes de replicación que pertenecían a grupos de incompatibilidad diferente y también genes de resistencia antibiótica diferente. Esto es debido a que los expertos en la materia creen que plásmidos que utilizan el mismo sistema de replicación (origen de replicación) no pueden coexistir de forma estable en una célula y se dice que son incompatibles (ver páginas 1.3-4 de (Sambrook y otros (1989) *Molecular cloning: a laboratory manual*. Editorial Cold Spring Harbor Laboratory)). Para cada plásmido y origen de replicación, se requiere además un gen de resistencia antibiótica diferente. Esto limitó el número total de plásmidos diferentes que pudieron mantenerse en una única célula. Consecuentemente, el número de sustratos de recombinación que pudieron estar presentes en una única célula se limitó al número de orígenes de replicación y genes de resistencia antibiótica diferentes que pudieron utilizarse. De este modo, el tamaño y la diversidad de una biblioteca que se pudiera preparar utilizando este enfoque estaban seriamente limitados.

La utilización del enfoque de las “dos construcciones” (plásmido + fago) tiene además otros problemas que limitan su utilización. En primer lugar, el sistema de recombinación de dos construcciones es reversible, y como resultado las bibliotecas producidas utilizando este sistema están “contaminadas” con el fago inicial (por ejemplo, que contiene cadenas  $V_H$  falsas (*dummy*)) y plásmidos (que contienen la biblioteca  $V_H+C_H1$  o la falsa  $V_H+C_H2$ ) a unos niveles relativamente altos. De este modo, la células huésped estaban sobrecargadas con construcciones “inútiles” y el sistema requiere fuentes sustanciales (por ejemplo, células huésped) para mantener niveles significativos de construcciones “útiles” (adecuadamente recombinadas).

Un segundo problema es que los plásmidos, a pesar de que no contienen un origen de replicación Ff, pueden sin embargo incorporarse en las partículas de fago con eficiencias variables. Como resultado, la biblioteca final contiene una mezcla de muchos elementos genéticos diferentes, y por lo tanto se reduce la diversidad funcional efectiva.

Un tercer problema fue que la recombinación tiene lugar entre una única molécula de ADN de fago y ADN de plásmido en una célula huésped y la recombinación entre múltiples plásmidos diferentes no fue posible. Como resultado, cada bacteria (célula huésped) produjo un único anticuerpo recombinado nuevo. Esto limita el tamaño de la biblioteca obtenida al número de bacterias utilizadas en la etapa de recombinación, que alcanza un límite práctico de  $5 \times 10^{12}$  para 10 litros. Para obtener bibliotecas más grandes, es necesario utilizar más bacterias.

Una cuarta limitación fue que los productos de recombinación en un sistema de este tipo son productos “finales”, y no pueden recombinarse de nuevo. Como resultado, no hay facilidad para experimentar una recombinación posterior: genes  $V_H$  y  $V_L$  ya recombinados o genes  $V_H$  recombinados y la biblioteca de genes  $V_L$ . Solamente genes  $V_L$  recombinados se pudieron recombinar con la biblioteca de genes  $V_H$  original.

### Características de la invención

Los métodos y construcciones de la presente invención superan estos y otros problemas, y permiten la creación de bibliotecas de ácidos nucleicos y/o polipéptidos grandes y altamente diversas. Los métodos explotan el descubrimiento de que la recombinación, particularmente la recombinación con recombinasa, puede tener lugar entre dos o más construcciones del mismo tipo (por ejemplo, que tienen el mismo origen de replicación y/o los mismos marcadores regulatorios o seleccionables). De este modo, los métodos implican la introducción de, como mínimo, dos miembros de una población inicial de moléculas de ácido nucleico en, como mínimo, una célula (por ejemplo, una bacteria, célula vegetal, de levadura, de insecto, de hongo, de vertebrado, célula mamífera, etc.). Las moléculas de ácido nucleico comprenden preferentemente una secuencia de ácidos nucleicos que es idéntica para cada molécula (miembro

de la biblioteca) y que incluye un origen de replicación (y opcionalmente otras secuencias regulatorias, y/o vector esqueleto, etc.); y una secuencia de ácidos nucleicos que varía entre los miembros de la población y comprende, como mínimo, un sustrato para la recombinación. Una vez en la célula, los miembros de la biblioteca se recombinan de manera que, como mínimo, un sustrato para la recombinación se recombina entre, como mínimo, dos miembros de la población para producir de este modo una población que comprende miembros de ácido nucleico recombinados. La recombinación puede llevarse a cabo mediante un mecanismo de recombinación endógeno a la célula (por ejemplo, recombinación general o recombinación mediada por recombinasa) o mediante una recombinasa exógena a la célula (por ejemplo, recombinasa proporcionada externamente o mediante transfección con un ácido nucleico que se expresa por recombinasa). Habitualmente, los miembros de la biblioteca tienen, como mínimo, dos lugares de reconocimiento de recombinasa, y preferentemente la recombinación se lleva a cabo en el lugar preseleccionado para la recombinación. En realizaciones preferentes, la recombinación está mediada por una recombinasa seleccionada del grupo que consiste en un miembro de la familia *hin* de las recombinasas, un miembro de la familia de la integrasa *lambda*, una recombinasa *flp*, una resolvasa, un transposón, y una recombinasa *Cre* y, en realizaciones más preferentes, entre las recombinasas se incluyen, pero no constituyen limitación, *Cre*, *hin*, *gin*, *pin*, *cin*, y *flp*. En una realización más preferente, la recombinación se da en un lugar *loxP* o en un lugar *flp*.

En ciertas realizaciones, un sustrato para la recombinación comprende (por ejemplo, está bordeado por) un par de lugares de reconocimiento de recombinación (por ejemplo, un primer y un segundo lugares de reconocimiento de recombinación) y estos lugares puede ser los mismos o diferentes. En realizaciones particularmente preferentes, la recombinación da como resultado el intercambio, de como mínimo, un sustrato de recombinación entre, como mínimo, dos miembros de la población de ácido nucleico. Un par de lugares de reconocimiento de recombinación preferentes son *loxP* y un mutante de *loxP* (por ejemplo, *loxP* 511 ó *loxP* *fas*).

Los miembros de la biblioteca de ácidos nucleicos pueden introducirse en la célula por cualquiera de los métodos convenientes, entre los que se incluyen biofísicos, transfección, o infección (por ejemplo, cuando se contienen en partículas infecciosas). Poblaciones preferentes de moléculas de ácido nucleico comprenden, como mínimo, 2 miembros diferentes, preferentemente, como mínimo, 10 miembros diferentes, más preferentemente, como mínimo, 50 miembros diferentes, y más preferentemente, como mínimo, 500 miembros diferentes por célula (es decir, transfectados y/o infectados dentro de una célula). Bibliotecas de ácidos nucleicos preferentes comprenden, como mínimo,  $10^5$ , preferentemente, como mínimo,  $10^7$ , más preferentemente, como mínimo,  $10^8$ ,  $10^9$ , o  $10^{10}$  miembros diferentes, y más preferentemente, como mínimo,  $10^{11}$ ,  $10^{12}$ ,  $10^{13}$ ,  $10^{14}$ ,  $10^{15}$  o más miembros diferentes. Habitualmente (aunque no está restringido) las bibliotecas antes de la recombinación comprenderán  $10^6$ - $10^8$  miembros diferentes y después de la recombinación comprenderán  $10^9$ - $10^{13}$  o más miembros. Entre las partículas infecciosas preferentes utilizadas en este método se incluyen, pero no constituyen limitación, fagos (por ejemplo, fagos filamentosos, tales como fagos de la familia *Ff*) y fagómidos (por ejemplo, fagómidos derivados de miembros de la familia de fagos *Ff*).

Adicionalmente, los métodos de la presente invención pueden implicar la transfección o infección de una o más células con miembros de la población de miembros de ácido nucleico recombinados, de manera que las células se infectan a una multiplicidad de infección (*moi*, de *multiplicity of infection*) menor que 1, aproximadamente; y el empaquetado de los miembros de la biblioteca de ácidos nucleicos en paquetes de exposición genética replicables, de manera que una proteína sobre la superficie de los paquetes de exposición replicables se codifica por un ácido nucleico empaquetado en el paquete de exposición, que es una secuencia de ácidos nucleicos que varía entre los miembros de la biblioteca de ácidos nucleicos.

En realizaciones particularmente preferentes, la secuencia variable de ácidos nucleicos que comprende el sustrato para la recombinación en los miembros de las bibliotecas de ácidos nucleicos comprende un casete de expresión, preferentemente un casete de expresión que comprende secuencias de ácidos nucleicos que codifican uno o más polipéptidos. En una realización más preferente, el ácido nucleico que codifica, como mínimo, uno de los polipéptidos está bordeado por pares de lugares de reconocimiento de recombinasa (por ejemplo, en los que los miembros del par de los lugares de reconocimiento son los mismos o diferentes, tal como se ha descrito anteriormente).

Los casetes de expresión preferentes expresan los polipéptidos sobre la superficie de un fago, un fagómido, o una bacteria. En una realización, entre las secuencias variables (por ejemplo, el casete de expresión) se pueden incluir una secuencia de ácidos nucleicos que codifica una primera cadena de polipéptido y una secuencia de ácidos nucleicos que codifica una segunda cadena de polipéptido en las que las cadenas de polipéptido son de un miembro de un par de unión específica (por ejemplo, un anticuerpo, un receptor, etc.), de manera que posteriormente a la recombinación la secuencia variable codifica proteínas de unión (o pares de proteínas de unión) que no están codificadas en la población inicial de ácidos nucleicos. En este caso, entre los mencionados primer y segundo polipéptidos preferentes se incluyen polipéptidos de anticuerpo (por ejemplo, fragmentos/dominios de anticuerpo). Entre los fragmentos de anticuerpo particularmente preferentes se incluyen, pero no constituyen limitación, una región  $V_H$ , una región  $V_L$ , una  $V_H$  CDR1, una  $V_H$  CDR2, una  $V_H$  CDR3, una  $V_L$  CDR1, una  $V_L$  CDR2, una  $V_L$  CDR3, una  $V_H$  unida a un  $C_H1$ , una  $V_L$  unida a un  $C_L$ , y similares. En el caso en el que los dos polipéptidos forman un scFv, el primer polipéptido es una región  $V_H$  y el segundo polipéptido es una región  $V_L$ . Habitualmente, como mínimo un lugar de reconocimiento de recombinasa estará presente en el ácido nucleico que codifica un conector que une las regiones  $V_H$  y  $V_L$ . Las regiones  $V_H$  y  $V_L$  pueden estar en el orden  $V_L$  seguida de  $V_H$  o viceversa. Estos lugares de recombinasa serán a menudo miembros de un par de lugares de reconocimiento de recombinasa que están bordeados por ácido nucleico que codifica un primer polipéptido, y los lugares de recombinasa que forman el par pueden ser los mismos o diferentes, tal como se ha explicado anteriormente. Ambos polipéptidos de cualquier combinación de fragmentos de anticuerpo que comprenden un anticuerpo, pueden

tener uno o más de los fragmentos sustituyentes bordeados por un par de lugares de reconocimiento de recombinasa. Miembros de la biblioteca de ácidos nucleicos particularmente preferentes codifican anticuerpos de cadena única (por ejemplo, scFv). Entre otras construcciones preferentes se incluyen, pero no constituyen limitación, un Fab, un diacuerpo, un dímero de  $V_H$ , y un dímero de  $V_L$ , y similares.

En otra realización, la presente invención da a conocer bibliotecas de ácidos nucleicos preparadas mediante los métodos de la presente invención.

La presente invención da a conocer además una biblioteca de ácidos nucleicos que comprende una población de moléculas de ácido nucleico que comprende dos o más ácidos nucleicos individuales, cada uno de los cuales consiste en una secuencia de ácidos nucleicos que es idéntica para cada molécula y que incluye un origen de replicación y, como mínimo, dos lugares de reconocimiento de recombinasa; y una secuencia de ácidos nucleicos que varía entre los miembros de dicha población en la que dicha secuencia de ácidos nucleicos que varía comprende un sustrato para recombinación, y en la que todos los miembros de dicha biblioteca tienen el mismo origen de replicación. Estas bibliotecas comprenden habitualmente, como mínimo, 2 miembros diferentes, preferentemente, como mínimo 10 miembros diferentes, más preferentemente, como mínimo, 50 miembros diferentes, y más preferentemente, como mínimo, 500 miembros diferentes por célula. Estas bibliotecas habitualmente comprenden, como mínimo,  $10^3$ , preferentemente, como mínimo,  $10^7$ , más preferentemente, como mínimo,  $10^8$ , y más preferentemente, como mínimo, de  $10^9$  a  $10^{15}$  o más miembros diferentes, tal como los descritos anteriormente. Los lugares de recombinasa y los miembros de ácido nucleico comprenden alguna o la mayoría de las características descritas en la presente descripción y, opcionalmente, pueden estar contenidos dentro de partículas infecciosas tal como se describe en la presente descripción. De este modo, se comprenderá que las bibliotecas pueden comprender una colección de fagos, o fagómidos o bacterias aislados que contienen los miembros de la biblioteca de ácidos nucleicos. Los ácidos nucleicos pueden ser vectores fagómico que codifican los anticuerpos y estar preparados para la subclonación dentro de un vector fago o los ácidos nucleicos pueden ser una colección de fagómidos que portan ya el anticuerpo subclonado que codifica ácidos nucleicos.

En aún otra realización, la presente invención da a conocer métodos de preparación de un polipéptido. Los métodos comprenden obtención de una biblioteca de ácidos nucleicos tal como las descritas en la presente descripción; seleccionar uno o más miembros de la biblioteca; y expresar los ácidos nucleicos de los miembros seleccionados. En una realización preferente, la selección puede implicar la expresión de proteínas codificadas por los miembros de la biblioteca de ácidos nucleicos; y la monitorización las proteínas expresadas para una o más de las propiedades deseadas (por ejemplo, unión específica a una o más objetivos preseleccionados, avidez de unión mínima para una o más objetivos preseleccionados, avidez de unión máxima para una o más objetivos preseleccionados, termoestabilidad a una temperatura determinada preseleccionada, actividad catalítica predefinida, una actividad enzimática predefinida bajo condiciones seleccionadas, una actividad biológica predefinida; estabilidad y/o actividad bajo condiciones oxidantes o reductoras o en disolventes no acuosos, etc.) y la selección de los miembros de la biblioteca que cumplen con los criterios de la monitorización. En realizaciones particularmente preferentes, la monitorización comprende monitorización para la unión específica a un objetivo preseleccionado y en estos casos, miembros de la biblioteca particularmente preferentes expresan anticuerpos (por ejemplo, sobre paquetes de exposición genética replicables (rgdps) o sobre la superficie de células infectadas o transfectadas con los miembros de la biblioteca de ácidos nucleicos). Entre los rgdps preferentes se incluyen fagos o fagómidos que expresan uno o más polipéptidos sustituidos para proteínas de superficie o para enlazarse covalentemente a éstas. Los miembros seleccionados pueden utilizarse para otro ciclo de recombinación y monitorización o para enriquecer o generar una biblioteca para posteriores ciclos de recombinación y monitorización.

Además, en varias otras realizaciones, la presente invención da a conocer células huéspedes que comprenden una biblioteca de ácidos nucleicos tal como las descritas en la presente descripción

## Definiciones

Tal como se utiliza en la presente descripción, un "anticuerpo" se refiere a una proteína que consiste en uno o más polipéptidos sustancialmente codificados por genes de inmunoglobulina o fragmentos de genes de inmunoglobulina. Entre los genes de inmunoglobulina reconocidos se incluyen los genes de región constante kappa, lambda, alfa, gamma, delta, epsilon y mu, así como un sinnúmero de genes de inmunoglobulina de región variable. Las cadenas ligeras se clasifican bien como kappa o lambda. Las cadenas pesadas se clasifican bien como gamma, mu, alfa, delta, o epsilon, que, a su vez, definen las clases de inmunoglobulina, IgG, IgM, IgA, IgD e IgE, respectivamente.

Se conoce que una unidad estructural típica de inmunoglobulina (anticuerpo) comprende un tetrámero. Cada tetrámero se compone de dos pares de cadenas de polipéptido idénticas, teniendo cada par una cadena "ligera" (25 KD aproximadamente) y una cadena "pesada" (50-70 KD aproximadamente). El extremo N-terminal de cada cadena define una región variable de, aproximadamente, 100 a 110 o más aminoácidos responsables principalmente del reconocimiento del antígeno. Los términos cadena ligera variable ( $V_L$ ) y cadena pesada variable ( $V_H$ ) se refieren a estas cadenas ligera y pesada, respectivamente.

Los anticuerpos existen como inmunoglobulinas intactas o como varios fragmentos bien caracterizados producidos por digestión con varias peptidasas. De este modo, por ejemplo, la pepsina digiere un anticuerpo por debajo de los enlaces disulfuro en la región bisagra para producir  $F(ab)_2$ , un dímero de Fab que en sí mismo es una cadena ligera unida a un  $V_H$ - $C_H1$  por un enlace disulfuro. El  $F(ab)_2$  puede reducirse en condiciones suaves para romper el enlace

disulfuro en la región bisagra para convertir de este modo el dímero (Fab)<sub>2</sub> en un monómero Fab'. Esencialmente, el monómero Fab' es un Fab con una parte de la región bisagra (ver, *Fundamental Immunology*, W. E. Paul, editores, Raven Press, N. Y. (1993), para una descripción más detallada de otros fragmentos de anticuerpo). Aunque varios fragmentos de anticuerpo se definen en términos de la digestión de un anticuerpo intacto, alguien con experiencia entenderá que estos fragmentos Fab' pueden sintetizarse de novo o bien químicamente o bien mediante la utilización de metodologías de ADN recombinante. De este modo, el término anticuerpo, tal como se utiliza en la presente descripción, incluye además fragmentos de anticuerpo o bien producidos por la modificación de anticuerpos enteros o sintetizados de novo utilizando metodologías de ADN recombinante. Entre los anticuerpos preferentes se incluyen anticuerpos de cadena única (anticuerpos que existen como una única cadena de polipéptido), más preferentemente anticuerpos de cadena única Fv (sFv o scFv) en los que una cadena pesada variable y una cadena ligera variable se unen entre sí (directamente o a través de un conector peptídico) para formar un polipéptido continuo. El anticuerpo de cadena única Fv es un heterodímero V<sub>H</sub>-V<sub>L</sub> enlazado covalentemente que puede expresarse a partir de ácidos nucleicos, entre los que se incluyen secuencias de codificación de V<sub>H</sub>- y V<sub>L</sub>- o bien unidas directamente o bien unidas mediante un conector que codifica péptidos. Huston, y otros (1988) *Proc. Nat. Acad. Sci. USA*, 85: 5879-5883. Aunque las V<sub>H</sub> y V<sub>L</sub> están conectadas entre sí como una cadena de polipéptido única, los dominios V<sub>H</sub> y V<sub>L</sub> se asocian de manera no covalente. Las primeras moléculas de anticuerpo funcionales a expresarse sobre la superficie del fago filamentoso fueron Fv de cadena única (scFv), sin embargo, también han tenido éxito estrategias alternativas de expresión. Por ejemplo, moléculas Fab pueden expresarse sobre fago si una de las cadenas (pesada o ligera) se fusiona a una proteína cápside g3 y la cadena complementaria se exporta al periplasma como una molécula soluble. La dos cadenas pueden codificarse sobre el mismo o sobre diferentes replicones; el punto importante es que las dos cadenas de anticuerpo en cada molécula Fab se ensamblen post-translacionalmente y el dímero se incorpore dentro de la partícula de fago a través del enlace de una de las cadenas a, por ejemplo, g3p (ver, por ejemplo, la Patente U.S.A. No. 5733743). Son conocidos para los técnicos en la materia anticuerpos scFv y varias otras estructuras que convierten las cadenas de polipéptido ligera y pesada agregadas naturalmente, pero químicamente separadas, de una región V de un anticuerpo en una molécula que se pliega en una estructura tridimensional sustancialmente similar a la estructura de un lugar de unión de antígeno (ver, por ejemplo, las Patentes U.S.A. Nos. 5.091.513, 5.132.405, y 4.956.778). Entre los anticuerpos particularmente preferentes se deben incluir todos los que han sido expresados sobre fago (por ejemplo, scFv, Fv, Fab y Fv enlazado por disulfuro (Reiter y otros (1995) *Protein Eng.* 8: 1323-1331).

Un "lugar de unión de antígeno" o "parte de unión" se refiere a la parte de una molécula de inmunoglobulina que participa en la unión al antígeno. El lugar de unión de antígeno está formado por residuos de aminoácido de regiones variables del extremo N-terminal ("V") de las cadenas pesada ("H") y ligera ("L"). Tres trozos altamente divergentes dentro de las regiones V de las cadenas pesada y ligera se refieren como "regiones hipervariables", que están interpuestas entre trozos adyacentes más conservados, conocidos como "regiones marco" o "FRs". De este modo, el término "FR" se refiere a secuencias de aminoácidos que se encuentran naturalmente entre, las regiones hipervariables en las inmunoglobulinas y adyacentes a las mismas. En una molécula de anticuerpo, las tres regiones hipervariables de una cadena ligera y las tres regiones hipervariables de una cadena pesada se disponen de forma relativa una a la otra en un espacio tridimensional para formar una "superficie" de unión de antígeno. Esta superficie media el reconocimiento y la unión del antígeno objetivo. Se hace referencia a las tres regiones hipervariables de cada una de las cadenas pesada y ligera como "regiones que determinan la complementariedad" o "CDR" y están caracterizadas, por ejemplo, según Kabat y otros, en *Sequences of proteins of immunological interest* 4ª edición, Departamento de Servicios Humanos y de la Salud, Servicios de Salud Públicos de los EE.UU, Bethesda, MD (1987).

Tal como se utiliza en la presente descripción, los términos "unión inmunológica" y "propiedades de unión inmunológica" se refieren a las interacciones no covalentes del tipo que tienen lugar entre una molécula de inmunoglobulina y un antígeno para el que la inmunoglobulina es específica. La fuerza o afinidad de las interacciones de unión inmunológica puede expresarse en términos de la constante de disociación (K<sub>d</sub>) de la interacción, en la que una K<sub>d</sub> más pequeña representa una mayor afinidad. Las propiedades de unión inmunológica de polipéptidos seleccionados se pueden cuantificar utilizando métodos bien conocidos en la técnica. Uno de estos métodos implica la medición de las velocidades de formación y disociación del complejo lugar de unión de antígeno/antígeno, en el que esas velocidades dependen de las concentraciones de los miembros del complejo, la afinidad de la interacción, y de parámetros geométricos que influyen a la velocidad en las dos direcciones por igual. De este modo, tanto la "constante de velocidad de activación (on) (k<sub>on</sub>) y la "constante de velocidad de desactivación (off) (k<sub>off</sub>) se pueden determinar mediante el cálculo de las concentraciones y las velocidades reales de asociación y disociación. La proporción k<sub>off</sub>/k<sub>on</sub> permite la cancelación de todos los parámetros no relacionados con la afinidad y, de este modo, es igual a la constante de disociación K<sub>d</sub> (ver, generalmente, Davies y otros (1990) *Ann. Rev. Biochem.*, 59: 439-473).

La frase "se une específicamente a una proteína" o "es específicamente inmunorreactivo a", cuando se hace referencia a un anticuerpo, se refiere a una reacción de unión que es determinativa de la presencia de la proteína en presencia de una población heterogénea de proteínas y otros componentes biológicos. De este modo, bajo las condiciones de inmunoensayo designadas, los anticuerpos especificados se unen a una proteína determinada y no se unen en una cantidad significativa a otras proteínas presentes en la muestra. La unión específica a una proteína bajo estas condiciones puede requerir un anticuerpo que se selecciona por su especificidad para una proteína determinada. Por ejemplo, los anticuerpos F5 o C1 se pueden incorporar a la proteína c-erbB-2 que une c-erbB-2 y no a otras proteínas presentes en una muestra de tejido. Pueden utilizarse una variedad de formatos de inmunoensayo para seleccionar anticuerpos específicamente inmunorreactivos a una proteína determinada. Por ejemplo, inmunoensayos de ELISA en fase sólida se utilizan de forma rutinaria para seleccionar anticuerpos monoclonales específicamente inmunorreactivos a una proteína. Ver Harlow y Lane, (1988) *Antibodies, A Laboratory Manual*, Publicaciones Cold Spring Harbor,

Nueva York, para una descripción de formatos de inmunoensayo y condiciones que pueden utilizarse para determinar la inmunorreactividad específica.

Los términos “polipéptido”, “péptido”, o “proteína” se utilizan en la presente descripción de forma intercambiable para designar una serie lineal de residuos de aminoácido conectados uno a otro mediante enlaces peptídicos entre los grupos amino alfa y carboxilo de residuos adyacentes. Los residuos de aminoácido están preferentemente en la forma isomérica natural “L”. Sin embargo, residuos en la forma isomérica “D” pueden sustituirse por cualquier residuo de L-aminoácido, mientras que el polipéptido mantenga la propiedad funcional deseada. Adicionalmente, entre los aminoácidos, además de los 20 aminoácidos “estándar”, se incluyen aminoácidos modificados e inusuales, entre los que se incluyen, pero no constituyen limitación, aquellos que aparecen en la lista 37 CFR 1.822 (b)(4). Adicionalmente, debe destacarse que una raya al principio o final de una secuencia de residuos de aminoácidos indica bien un enlace peptídico a una secuencia adicional de uno o más residuos de aminoácido o un enlace covalente a un grupo carboxilo o hidroxilo terminal.

El término “polipéptido de unión” se refiere a un polipéptido que se enlaza específicamente a una molécula objetivo (por ejemplo, una célula receptora) en una manera análoga a la unión de un anticuerpo a un antígeno. Polipéptidos de unión se distinguen de los anticuerpos en que los polipéptidos de unión no derivan en última instancia de genes de inmunoglobulina o fragmentos de genes de inmunoglobulina.

El término “ácido nucleico” se refiere a desoxirribonucleótidos o ribonucleótidos y polímeros de los mismos, o bien en forma de cadena única o bien cadena doble. A menos que se limite específicamente, el término abarca ácidos nucleicos que contienen análogos conocidos de nucleótidos naturales que tienen propiedades de unión similares al ácido nucleico de referencia y se metabolizan de una manera similar a los nucleótidos que se dan de forma natural. A menos que se indique de otra manera, una secuencia determinada de ácidos nucleicos abarca además implícitamente variantes de los mismos modificadas conservativamente (por ejemplo, sustituciones de codón degenerado) y secuencias complementarias, y también la secuencia indicada explícitamente. Específicamente, se pueden conseguir sustituciones de codón degenerado mediante la generación de secuencias en las que la tercera posición de uno o más codones seleccionados (o todos) se sustituye con residuos de base mixta y/o desoxiinosina (Batzner y otros (1991) *Nucleic Acid Res.* 19: 5081; Ohtsuka y otros (1985) *J. Biol. Chem.* 260: 2605-2608; y Cassol y otros (1992); Rossolini y otros, (1994) *Mol. Cell. Probes* 8: 91-98). El término ácido nucleico se utiliza de forma intercambiable con gen, ADNc, y ARNm codificado por un gen.

Los términos “aislado” o “biólogicamente puro” se refieren a materiales que están sustancialmente o esencialmente libres de componentes que normalmente lo acompañan tal como se encuentran en su estado natural. Sin embargo, no se pretende que el término “aislado” se refiere a los componentes presentes en un gel electroforético u otro medio de separación. Un componente aislado está libre de estos medios de separación y en una forma preparada para su utilización en otra aplicación o ya está en utilización en la nueva aplicación/medio.

El término “ácido nucleico heterólogo” se refiere a un ácido nucleico que no es nativo a la célula en la que se encuentra o cuyo origen último no es la célula o línea celular en la que se encuentra actualmente el “ácido nucleico heterólogo”.

El idiotipo representa el lugar de unión de antígeno altamente variable de un anticuerpo y es inmunogénico por sí mismo. Durante la generación de una respuesta inmune mediada por anticuerpo, un individuo desarrollará anticuerpos al antígeno así como anticuerpos anti-idiotipo, cuyo lugar de unión inmunogénico (idiotipo) mimetiza el antígeno. Además, los anticuerpos anti-idiotípicos pueden generarse mediante la inmunización con un anticuerpo, o fragmento del mismo.

Una “biblioteca de exposición en fago” se refiere a una colección de fagos (por ejemplo, fagos filamentosos) en la que el fago expresa una proteína externa (habitualmente heteróloga). La proteína externa es libre para interactuar con otras partes con las que el fago está en contacto, o unirse a las mismas. Cada fago que expone una proteína externa es un “miembro” de la biblioteca de exposición en fago.

El término “fago filamentoso” se refiere a una partícula vírica capaz de exponer un polipéptido heterogéneo sobre su superficie. Aunque alguien con experiencia en la técnica comprenderá que pueden utilizarse una variedad de bacteriófagos en la presente invención, en realizaciones preferentes, el vector es un bacteriófago filamentoso o se deriva de los mismos, tal como, por ejemplo, fl, fd, Pfl, M13, etc. El fago filamentoso puede contener un marcador seleccionable, tal como tetraciclina (por ejemplo, “fd-tet”). Varios sistemas de exposición en fago filamentoso se conocen bien por aquellos con experiencia en la técnica (ver, por ejemplo, Zacher y otros (1980) *Gene* 9: 127-140, Smith y otros (1985) *Science* 228: 1315-1317 (1985); y Parmley y Smith (1988) *Gene* 73: 305-318).

Una “señal de empaquetamiento viral” es una secuencia de ácidos nucleicos necesaria y suficiente para la incorporación directa de un ácido nucleico en una cápside viral.

Una célula de ensamblaje es una célula en la que un ácido nucleico puede empaquetarse dentro de una proteína de recubrimiento viral (cápside). Las células de ensamblaje se pueden infectar con una o más partículas de virus diferentes (por ejemplo, un fago normal o debilitado y un fago ayudante) que individualmente o en combinación direccionan el empaquetamiento de un ácido nucleico dentro de una cápside viral.

Las siguientes abreviaturas se utilizan en la presente descripción: AMP, ampicilina; CDR, región que determina la complementariedad; ELISA, ensayo de inmunosorbente unido a enzima; FACS, clasificador celular activado por fluorescencia; FR, región marco; Glu, glucosa; IMAC, cromatografía de afinidad al metal inmovilizado;  $k_{on}$ , constante de velocidad de asociación;  $k_{off}$ , constante de velocidad de disociación; PCR, reacción de cadena de polimerasa; RU, unidades de resonancia; scFv o sFv, fragmento Fv de cadena única; resonancia plasmónica de superficie;  $V_{\kappa}$ , región variable de la cadena ligera de la inmunoglobulina kappa;  $V_{\lambda}$ , región variable de la cadena ligera de la inmunoglobulina lambda;  $V_L$ , región variable de la cadena ligera de inmunoglobulina;  $V_H$ , región variable de la cadena pesada de inmunoglobulina; wt, tipo salvaje.

Una “recombinasa” se refiere a una proteína que media la recombinación entre dos o más ácidos nucleicos diferentes. Una recombinasa preferente “reconoce” una o más secuencias de nucleótidos determinadas (un lugar de reconocimiento de recombinasa) y extirpa el segmento de ácido nucleico en estos lugares. De ese modo, la utilización de una recombinasa y lugares de reconocimiento de recombinasa apropiados puede delinear la recombinación del sustrato (la secuencia de ácidos nucleicos sujeta a un evento de recombinación).

En una realización preferente, una “partícula infecciosa” se refiere a una molécula de ácido nucleico recubierta con “proteínas de recubrimiento” que es capaz de entrar en células con una alta eficiencia en virtud de interacciones específicas entre la o las proteínas de recubrimiento y un receptor sobre la superficie de la célula. Un ejemplo de una partícula infecciosa es un fago Ff o fagómido.

Una “proteína de recubrimiento” se refiere a una proteína encontrada sobre la superficie de una partícula infecciosa que “recubre” el ácido nucleico encerrado. Las proteínas de recubrimiento se pueden designar mayores o menores dependiendo del número de copias presentes sobre la partícula infecciosa. En el fago Ff, la proteína de recubrimiento mayor es p8 y está presente en 2800 copias aproximadamente, mientras que p3 es una proteína de recubrimiento menor presente en 3-5 copias.

El término “infección” se refiere al proceso por el que una partícula infecciosa entra en una célula en virtud de interacciones específicas entre una proteína o proteínas de recubrimiento encontradas sobre la superficie de la partícula infecciosa y un receptor sobre la superficie de la célula. En el caso de fagos Ff y fagómidos, la interacción tiene lugar entre la p3 sobre la superficie del fago y el pilus F sobre la bacteria.

Una “biblioteca primaria” o una “biblioteca primaria de genes” se refiere a una población de moléculas de ácido nucleico individuales cada una de las cuales está constituida de dos partes: una secuencia de ácidos nucleicos presente en cada molécula que es idéntica; y una secuencia de ácidos nucleicos variable, no necesariamente contiguos, que constituye el sustrato para la recombinación. Cada molécula de ácido nucleico es diferente a las otras moléculas de ácido nucleico que se encuentran en la biblioteca de genes. Las bibliotecas pueden tomar varias formas. De este modo, en una realización, la biblioteca es una colección de células que contienen los miembros de la biblioteca de exposición en fago, mientras que, en otra realización, la biblioteca consiste en una colección de fagos aislados, y en aún otra realización la biblioteca consiste en una biblioteca de ácidos nucleicos que codifican una biblioteca de exposición en fago. Los ácidos nucleicos pueden ser vectores fagómidos o plásmidos que codifican polipéptidos o anticuerpos y están preparados para la subclonación dentro de un vector fago o los ácidos nucleicos pueden ser una colección de fagómidos que ya portan los polipéptidos subclonados o los ácidos nucleicos que codifican los anticuerpos.

Una “biblioteca secundaria” o una “biblioteca secundaria de genes” se refiere a una colección de ácidos nucleicos que comprenden moléculas de ácido nucleico individuales derivadas de una biblioteca de genes primaria en la cual la diversidad se ha aumentado mediante el proceso de recombinación específica o general de lugar.

Una “biblioteca de partículas infecciosas” se refiere a una colección de partículas infecciosas, cada una de las cuales contiene una molécula de ácido nucleico que es diferente de las moléculas de ácido nucleico portadas por las otras partículas infecciosas en la biblioteca, siendo limitadas estas diferencias habitualmente a un gen específico que codifica una función específica, cuyas propiedades van a ser modificadas, permaneciendo idénticas todas las otras partes del genoma de la partícula infecciosa. Cuando el número de partículas infecciosas es mayor que la diversidad de la biblioteca, cada una de estas partículas puede estar presente en más de una copia.

Un “sustrato de recombinación” se refiere a una molécula de ácido nucleico que es capaz de recombinarse con otro sustrato de recombinación, o bien en virtud de la presencia de lugares de recombinación específicos (por ejemplo, loxP, en el que la recombinación está catalizada por una recombinasa específica, cre), o bien en virtud de similitudes en las secuencias de ácidos nucleicos que permiten que tenga lugar la recombinación celular general (recombinación general).

El término “recombinación específica” se refiere a la recombinación que tiene lugar en lugares específicos (por ejemplo, loxP) reconocidos por recombinasas (por ejemplo, recombinasa cre) que actúan en esos lugares. El término “multiplicidad de infección (MOI)” se refiere a la proporción entre el número de fagos o fagómidos (medidos como unidades formadoras de colonias o unidades formadoras por placa) añadidos a un cultivo de bacterias y el número de bacterias. Un MOI de 200 indica que los fagos o fagómidos están en un exceso de 200 veces con respecto a las bacterias.

Un “paquete de exposición genética replicable” o “rdgp” se refiere a una partícula biológica que tiene información genética que proporciona a la partícula la capacidad de replicarse bajo condiciones apropiadas. La partícula puede exponer sobre su superficie, como mínimo, una parte de un polipéptido. El polipéptido puede estar codificado por información genética nativa a la partícula y/o colocada de forma artificial dentro de la partícula. El polipéptido expuesto o una parte de él, puede ser cualquier miembro de un par de unión específica, por ejemplo, Fv de cadena única (scFv), dominios de cadena pesada o ligera basados en una molécula de inmunoglobulina, un enzima o receptor, etc. Un rdgp puede ser una célula, una bacteria, un fago, un fagómido, o cualquier otra partícula biológica con las características descritas anteriormente.

Un vector es una molécula de ácido nucleico que es capaz de replicarse por sí misma, y cualquier otra secuencia contenida dentro de éste, en una célula huésped (por ejemplo, bacteriana). Normalmente, los vectores contienen un origen de replicación, una resistencia antibiótica al gen y lugares de restricción enzimática que permiten la clonación de otros fragmentos de ADN.

Un “vector de exposición” se refiere a un vector utilizado para la creación de rgdps. Normalmente, los vectores de exposición, además de un origen de replicación y resistencia antibiótica al gen, contienen un gen que codifica una proteína, dentro del cual un polipéptido que es extraño a la partícula biológica puede expresarse, y que se expone sobre la superficie de la partícula biológica. Además, contienen preferentemente un origen de replicación adicional que permite que el vector ADN se incorpore dentro de la partícula biológica que expone el polipéptido extraño. pDAN5, un vector fagómido, es un ejemplo de un vector de expresión.

Un “fagómido” es una partícula infecciosa que comprende ADN plásmido que contiene el origen de replicación de un fago (por ejemplo, un fago filamentoso) que se recubre con proteínas de recubrimiento de fago (por ejemplo, proporcionadas por un fago ayudante). El término fagómido puede utilizarse además para aplicarlo al vector de ADN contenido dentro de la partícula infecciosa, un fagómido Ff se refiere a un fagómido que utiliza un origen de replicación derivado de la familia de fagos filamentosos Ff, y requiere un fago ayudante derivado de la misma familia Ff. pDAN5 es un ejemplo de un fagómido Ff.

Un “conector” se refiere a una secuencia de aminoácidos que une otras dos secuencias de aminoácidos. De este modo, por ejemplo, un conector puede unir las cadenas  $V_L$  y  $V_H$  de un anticuerpo permitiéndoles formar una estructura scFv correcta. En general, un conector es una secuencia de aminoácidos que enlaza covalentemente dos polipéptidos de manera que se forma un único polipéptido.

Un “par de unión específica” o “sbp” se refiere a un par de moléculas (siendo cada una miembros de un par de unión específica) que se derivan naturalmente o se producen sintéticamente, que se unen una a la otra específicamente. Entre los ejemplos de tipos de sbps se incluyen antígeno-anticuerpo, biotina-avidina, hormona-receptor hormonal, receptor-ligando, enzima-sustrato, IgG-proteína A.

Una “biblioteca de exposición en fago” se refiere a una colección de fagos (por ejemplo, fagos filamentosos) en la que el fago expresa una proteína externa (habitualmente heteróloga). La proteína externa es libre para interactuar con otra parte con la que el fago está en contacto o unirse a la misma. Cada fago que expone una proteína externa es un “miembro” de la biblioteca de exposición en fago.

El término “fago filamentoso” se refiere a una partícula viral capaz de expresar un polipéptido heterogéneo sobre su superficie. Aunque alguien con experiencia en la técnica entenderá que puede ser utilizada una variedad de bacteriófagos en la presente invención, en realizaciones preferentes el vector es un bacteriófago filamentoso, o se deriva del mismo, tal como, por ejemplo, fl, fd, Pfl, M13, etc. El fago filamentoso puede contener un marcador seleccionable tal como tetraciclina (por ejemplo, “fd-tet”). Varios sistemas de exposición en fago filamentoso son bien conocidos por aquellos con experiencia en la técnica (ver, por ejemplo, Zacher y otros (1980) *Gene* 9: 127-140, Smith y otros (1985) *Science* 228: 1315-1317 (1985); y Parmley y Smith (1988) *Gene* 73: 305-318).

Un “marcador seleccionable” se refiere a una secuencia de ácidos nucleicos que permite a los organismos que expresan una secuencia y/o que tienen una secuencia presente que se distingan de organismos que no expresan aquella secuencia y/o que no tienen aquella secuencia presente. Los marcadores seleccionables se conocen bien por aquellos con experiencia en la técnica. Entre algunos ejemplos se incluyen el gen hprt (Littlefield (1964) *Science* 145: 709-710), el gen tk del virus del herpes simplex (timidina quinasa) (Giphart-Gassler y otros (1989) *Mutat. Res.* 214: 223-232), el gen nDtII (Thomas y otros (1987) *Cell* 51: 503-512; Mansour y otros (1988) *Nature* 336: 348-352), u otros genes que confieren resistencia a los aminoácidos o a análogos de nucleósido, o antibióticos, etc.

Un “lugar loxP” se refiere a un lugar de reconocimiento de recombinasa que puede actuar como un lugar de reconocimiento de recombinasa para Cre. Un lugar loxP incluye secuencias loxP nativas así como lugares loxP modificados. Lugares loxP modificados se conocen bien por aquellos con experiencia en la técnica y entre éstos se incluyen, pero no constituyen limitación, loxP 511, fas, y otras mutaciones varias tales como las que han sido descritas en la literatura (ver, por ejemplo, Mack y otros, arriba; Hoess y otros (1986), arriba; Hoess y otros (1984) *Biochem* 81: 1026-29; Hoess y otros (1985) *Gene*, 40: 325-329; Abremski y otros (1986) *J. Biolog. Chem.*, 261: 391-396; Sauer (1996) *Nuc. Acids Res.* 24: 4608-13).

## Breve descripción de los dibujos

La figura 1 ilustra la generación de diversidad (por ejemplo, una biblioteca de anticuerpos) en el sistema de un vector según la presente invención. Se muestra la infección de una bacteria por dos fagómidos que contienen regiones  $V_H A/V_L A$  y  $V_H B/V_L B$ . Después de la recombinación, estarán presentes cuatro productos: los  $V_H A/V_L A$  y  $V_H B/V_L B$  originales, y los dos nuevos productos  $V_H A/V_L B$  y  $V_H B/V_L A$ . Todos los fagómidos, tanto antes como después de la recombinación, contienen anticuerpos funcionales (scFv). Además del número de células utilizadas, el tamaño de la biblioteca de genes final se determina también por el número de fagómidos que entran en una célula y por la extensión de la recombinación que tiene lugar.

La figura 2 ilustra la secuencia del conector loxP 511 scFv utilizado en los ejemplos. En particular, se indica el ADN y secuencias de aminoácidos deducidas del conector loxP 511 utilizado para crear mini-bibliotecas y para la clonación del scFv. Los lugares de restricción utilizados para la clonación de regiones V se indican debajo de la secuencia de ADN. La secuencia del lugar de reconocimiento loxP 511, implicado en la recombinación, está en cursiva, los lugares de restricción utilizados para la clonación están subrayados, y las repeticiones invertidas en la secuencia de reconocimiento del loxP 511 están doblemente subrayadas. Las regiones V del anticuerpo se clonaron en el orden  $V_L$  conector  $V_H$ .

La figura 3 ilustra el policonector pDAN5 descrito en la presente descripción. Se da la secuencia del policonector pDAN5 para la clonación de scFvs (se omite la secuencia de las regiones V y el gen 3). Se da la posición de las cadenas  $V_L$  y  $V_H$ , y los lugares de restricción utilizados para clonarlos se indican en negrita. Se indican los lugares loxP 511 y loxP tipo salvaje utilizados para inducir recombinación, con las repeticiones invertidas subrayadas. También se indican otras características importantes, entre las que se incluyen la secuencia líder, la etiqueta SV5 utilizada para reconocer el scFv, la etiqueta His para la purificación, el codón de detención ámbar y gen 3.

La figura 4 ilustra el mapa del pDAN5. En particular, se muestra un mapa del vector de exposición pDAN5 con el D1.3 scFv clonado. Los lugares utilizados para la clonación del gen V están en negrita.

La figura 5 ilustra el esquema del experimento de recombinación D1.3. Dos fagómidos que contienen  $V_L/X$ - $V_H/D1.3$  y  $V_L/D1.3$ - $V_H/Y$  (en los que X e Y representan anticuerpos desconocidos) se utilizan para infectar o bien bacterias que expresan cre o bien DH5 $\alpha$ F tipo salvaje (que no expresa recombinasa cre) a una alta multiplicidad de infección (MOI). Después de crecimiento durante 12 horas aproximadamente, se prepararon los fagómidos, se reinfectaron dentro de DH5 $\alpha$ F (es decir, no en células que expresan cre) para acoplar el genotipo y fenotipo, y se examinaron por PCR y ELISA para la presencia de  $V_L/D1.3$ - $V_H/D1.3$  funcional.

La figura 6 ilustra un esquema para la creación de una biblioteca de genes secundaria mediante recombinación in vivo. Se infectaron bacterias que expresan cre por una biblioteca de fagómidos primaria a una MOI de 200:1. Después de la recombinación, se prepararon los fagómidos y se reinfectaron dentro de DH5 $\alpha$ F a una MOI menor que 1 para acoplar fenotipo y genotipo. Estos fagómidos, con el fenotipo y genotipo acoplados, se utilizaron para selecciones de bibliotecas. La diversidad creada a partir de la infección de una única bacteria con solamente cuatro fagómidos (con diferentes genes  $V_H$  y  $V_L$  sombreados diferentemente) se indica por simplicidad. Sin embargo, según la presente invención, una única bacteria se infecta con un número de fagómidos mucho mayor. Experimentos de impresión de huella dactilar de PCR (ver la figura 9) confirman estos datos.

La figura 7 ilustra la evaluación de la diversidad en bacterias que expresan recombinasa cre infectadas por múltiples fagómidos. Se aislaron colonias bacterianas individuales que expresan especificidades de fagómidos únicas (después de la reinfección a una MOI menor que 1) después de la recombinación como se describe en la figura 6. Se amplificaron los genes  $V_H$  y  $V_L$  a partir de estas colonias, utilizando cebadores específicos de  $V_H$  y  $V_L$  (Sblattero y Bradbury (1998) *Immunotechnology* 3: 271-278) y se determinó su impresión dactilar utilizando BstNI. Las diferentes impresiones dactilares del gen  $V_H$  se numeraron horizontalmente 1-17 para la célula 1, y 1-18 para la célula 2, mientras que las diferentes impresiones dactilares del  $V_L$  se numeraron verticalmente 1-12 para la célula 1 y 1-14 para la célula 2. Solamente hay una tabla, que representa una célula en la figura 7. Se analizaron las huellas  $V_H$  y  $V_L$  encontradas en 37 bacterias individuales para la célula 1 y 41 bacterias individuales para la célula 2 y se resumen en la tabla, indicando el número de veces que se encontró cada combinación  $V_H/V_L$  indicado. Estos datos dan una indicación del número y las diferentes combinaciones de  $V_H/V_L$  presentes en las células originales 1 y 2.

La figura 8 ilustra el esquema de recombinación general utilizado para crear genes resistentes a la cefotaxima. Bacterias tipo salvaje se infectaron por una biblioteca de fagómidos primaria beta lactamasa a una MOI mayor que 10:1. La recombinación tiene lugar, llevando mutaciones que estaban inicialmente en diferentes genes dentro del mismo gen. Solamente los fagómidos que contienen los genes recombinados de beta lactamasa son capaces de proporcionar resistencia a la cefotaxima y se llevará a cabo la selección. La biblioteca primaria utilizada tenía una diversidad de  $5 \times 10^5$ . Esto se representa en la figura, mediante dos fagómidos, cada uno de los cuales contiene una única mutación indicada por un asterisco o por un punto. Ninguna de estas mutaciones por sí sola es capaz de proporcionar resistencia a la cefotaxima, mientras que ambas juntas (lo que sucede después de la recombinación) pueden proporcionar resistencia a la cefotaxima. Los fagos seleccionados para el evento de recombinación que se produjo en el gen beta lactamasa portarán además una nueva combinación de mutaciones en el segundo gen.

La figura 9 ilustra un esquema de recombinación general que puede utilizarse para seleccionar genes recombinados mediante su unión a la recombinación de un marcador o gen selector. Una biblioteca de diferentes genes, o versiones mutantes de un único gen, se clonaría dentro de un vector que, en sí mismo, es una biblioteca de vectores que contiene mutaciones dentro de un marcador o gen selector (por ejemplo, beta lactamasa). Una biblioteca de este tipo podría además premonitorizarse para que contuviera solamente mutaciones beneficiosas conocidas a efectos de aumentar la incidencia de eventos de recombinación útiles, o incluso prepararse con dos únicas mutaciones que fueran efectivas solamente cuando se encuentran en el mismo gen. Una vez seleccionadas sobre el medio de selección (por ejemplo, cefotaxima si se utiliza beta lactamasa), serán seleccionados solamente aquellos vectores que han experimentado recombinación dentro del gen de beta lactamasa. Como la recombinación sucede habitualmente en dos lugares, esto seleccionará simultáneamente vectores que han experimentado recombinación en otra parte del vector, que frecuentemente estará dentro del gen que va a recombinarse. La figura ilustra la situación con dos vectores, cada uno de los cuales contiene mutaciones beneficiosas tanto dentro del gen selector (beta lactamasa) y el gen de interés. La recombinación dentro del gen de beta lactamasa, mediada por la maquinaria de recombinasa celular, lleva las dos mutaciones de beta lactamasa en cis y por lo tanto permite la selección por resistencia a la cefotaxima. Simultáneamente, hay recombinación dentro del gen de interés que lleva también mutaciones beneficiosas en cis y permite la identificación de este gen útil por selección o monitorización. En el esquema, las cajas abiertas representan fagómidos y los círculos abiertos representan el vector en su forma replicativa (plasmídica).

### Descripción detallada de la invención

La presente invención da a conocer nuevos métodos para crear bibliotecas de ácidos nucleicos y/o proteínas expresadas, extensas y altamente diversas. La creación de diversidad en poblaciones de ácidos nucleicos o polipéptidos se ha convertido en una herramienta importante en la derivatización de polipéptidos con características útiles. Estos enfoques implican habitualmente la producción de una población extensa y diversa de moléculas sujeto y, a continuación, la monitorización de esas moléculas para la propiedad o propiedades deseadas de interés. De este modo, estos métodos mimetizan el proceso de evolución y selección natural, y pueden utilizarse para evolucionar y seleccionar moléculas que tienen una amplia variedad de propiedades deseadas.

Estos métodos de selección han sido ampliamente utilizados en una variedad de contextos. De este modo, por ejemplo, la evolución direccionada y procedimientos de selección se han utilizado para obtener anticuerpos de cadena única humanos, no humanos o quiméricos que tienen una especificidad de unión y/o aidez determinadas (ver, por ejemplo, Stemmer y otros (1992) *Biotechniques* 14: 256-265; Marks, y otros (1992) *Bio Technology*, 10: 779-782, etc.). Estas aproximaciones se han utilizado también para obtener proteínas de unión determinadas (ver, por ejemplo, Clackson y Wells (1994) *Trends in Biotechnology* 12: 173-184), para obtener enzimas de ARN (Beaudry y otros (1992) 257: 635-641), para evolucionar y seleccionar proteínas que tienen actividades enzimáticas o catalíticas determinadas.

A diferencia de las aproximaciones de la técnica anterior, que utilizan aproximaciones direccionadas a lugar o mutagénesis aleatoria (por ejemplo, PCR proclive a error) para aumentar la diversidad de una biblioteca, los métodos de la presente invención utilizan procesos de recombinación. Entre éstos se incluyen los mecanismos de recombinación naturales (endógenos) que tienen lugar dentro de una célula (por ejemplo, entrecruzado, recombinación de homólogos, etc.) o procesos mediados por recombinasa en los que el proceso está mediado por una recombinasa endógena y/o exógena.

A diferencia de otras aproximaciones que explotan las recombinasas para mediar en la recombinación entre dos construcciones diferentes (por ejemplo, un vector y un ADN genómico, dos vectores diferentes, etc.), en una realización, los métodos de la presente invención utilizan recombinasas para mediar la recombinación entre dos o más construcciones idénticas; esto es, construcciones que son esencialmente idénticas salvo para el segmento o segmentos que van a recombinarse.

El descubrimiento sorprendente de la presente invención ha sido que, contrariamente a lo explicado en la técnica anterior, construcciones múltiples de la misma clase (por ejemplo, fagómidos, plásmidos, etc.) que son idénticas esencialmente en los elementos regulatorios (por ejemplo, origen de replicación y/o promotor o promotores y/o potenciadores y/o secuencias de terminación) y/o marcadores seleccionables (por ejemplo, genes de resistencia antibiótica) pueden coexistir en una célula huésped (por ejemplo, una célula bacteriana) durante un tiempo suficientemente largo para que pueda tener lugar la recombinación extensiva (por ejemplo, mediada por mecanismos celulares sin recombinasa, por recombinasas celulares, o por recombinasas exógenas) entre las construcciones y permite la creación de bibliotecas de ácidos nucleicos y/o péptidos, extensas y altamente diversas. Adicionalmente, ha sido un descubrimiento de la presente invención que la recombinación está particularmente potenciada en células infectadas a una alta multiplicidad de infección (por ejemplo, m.o.i. mayor que 5 aproximadamente, más preferentemente m.o.i. mayor que 20 aproximadamente, y más preferentemente m.o.i. mayor que aproximadamente 200 o más). Ha sido también un descubrimiento de la presente invención que las bibliotecas clonadas en vectores con un alto número de copias (por ejemplo, más de 300) tienen una copermanencia más larga que vectores con un bajo número de copias (por ejemplo, 10) permitiendo que la recombinación tenga lugar más extensivamente. De este modo, la transfección para introducir más de una copia de vector (preferentemente 200, 300, o incluso más de 500) dentro de cada célula debería dar como resultado la recombinación.

Debido a que la recombinación tiene lugar en una única clase de construcciones esencialmente idénticas (salvo para el o los ácidos nucleicos que van a ser recombinados) se superan los problemas asociados con los llamados sistemas de

dos construcciones descritos anteriormente. De este modo, por ejemplo, todas las construcciones introducidas dentro de una única célula según los métodos de la presente invención pueden tener orígenes idénticos de replicación y/o marcadores seleccionables, y/o secuencias de control. De este modo, solamente es necesaria la construcción de un único tipo de "vector de recombinación" y las diferentes secuencias que van a ser recombinadas se pueden insertar dentro del vector en uno o más lugares preseleccionados (por ejemplo, proporcionados por uno o más policonectores). Adicionalmente, pueden haber hasta 500, 1000 o más construcciones dentro de una célula, solamente limitadas por el número de copias del vector utilizado, y la recombinación puede tener lugar entre alguna o todas estas construcciones. De este modo, se puede producir un grado de diversidad extremadamente alto en una célula única y una población de células de tamaño modesto (por ejemplo,  $10^5$ - $10^7$  células) puede producir una biblioteca de diversidad remarcable (por ejemplo,  $10^{11}$ - $10^{18}$  o más miembros diferentes). El sistema no está limitado a un recombinante único por célula.

Adicionalmente, incluso después de la recombinación, las construcciones utilizadas en los métodos de la presente invención están inalteradas esencialmente, salvo por la recombinación. Por consiguiente, éstas se pueden utilizar para aumentar la diversidad en otras bibliotecas o en múltiples ciclos de recombinación y selección sin modificación esencial.

Estas ventajas son meramente ilustrativas. Otras propiedades y ventajas se harán evidentes para aquellos con experiencia a la luz de la presente descripción.

Las bibliotecas, una vez producidas, pueden utilizarse para modificar o mejorar propiedades determinadas. En cualquier biblioteca recombinada, una pequeña proporción de proteínas (por ejemplo, scFv) tendrá propiedades determinadas, tales como un plegamiento satisfactorio, resistencia a agentes reductores o elevada temperatura. Mediante la preselección de alguna de estas propiedades, se obtendrá un pequeño subconjunto de la biblioteca recombinada que está enriquecida en esa propiedad. La potencia del método de recombinación descrito en la presente descripción permite la creación de nuevas bibliotecas que pueden ser tan extensas como la biblioteca recombinada inicial, pero que son propensas a la propiedad para la que han sido preseleccionadas. Por ejemplo, en el caso de las bibliotecas scFv, la selección de la biblioteca recombinada en la proteína LA (que reconoce un epítipo conformacional de algunas regiones  $V_H$  y  $V_L$ ) preselecciona scFv con plegado adecuado, no susceptible a la degradación. Estas scFv están presentes en una proporción pequeña en la biblioteca recombinada final, pero realizando ciclos de recombinación adicionales, se puede crear una biblioteca de scFv muy diversa compuesta de scFv plegado adecuadamente. De forma similar, una pequeña proporción de scFv en la biblioteca recombinada puede ser resistente a agentes reductores tales como DTT. Mediante el tratamiento de la biblioteca con DTT y la selección de aquellos scFv que sobreviven (por ejemplo, mediante la selección de una etiqueta localizada en el extremo N-terminal, o mediante la utilización de la proteína LA), se pueden seleccionar los scFv que son resistentes a este agente. Un ciclo de recombinación adicional creará una biblioteca que está enriquecida en scFv que son resistentes al DTT, y de este modo pueden ser funcionales intracelularmente. Las bibliotecas de anticuerpos en fagos o fagómidos se hacen crecer habitualmente a temperaturas bajas (por ejemplo, 30°C). Esto es debido a que muchos anticuerpos son tóxicos o inestables a temperaturas más elevadas. Mediante el crecimiento de las bacterias de la biblioteca recombinada a temperaturas más elevadas (por ejemplo, 37°C o incluso 42°C) sólo se seleccionarán aquellas bacterias que contienen los scFv menos tóxicos y más estables. Mediante la realización de otro ciclo de selección, todos los scFv en la biblioteca se prepararán con aquellas regiones  $V_H$  y  $V_L$  que se preseleccionan por su no toxicidad o estabilidad. De forma similar, mediante la selección para la unión a un objetivo determinado, se derivará una biblioteca limitada direccionada hacia una especificidad determinada. Mediante la realización de la recombinación entre estos scFv, los cuales todos reconocen un antígeno determinado, se pueden seleccionar anticuerpos con una afinidad más elevada que no estaban presentes en la biblioteca original. Se pueden prever métodos similares de preselección, por ejemplo, resistencia a disolventes determinados, condiciones iónicas, temperaturas, actividad, etc.

#### *I. Creación eficiente de una biblioteca de ácidos nucleicos, extensa y altamente diversa*

En una realización preferente, la presente invención da a conocer métodos de preparación de una biblioteca de ácidos nucleicos, extensa y altamente diversa. Los métodos implican la transfección de, como mínimo, una célula, (más preferentemente una población de células), con más de una molécula de ácido nucleico por célula (preferentemente más de 2, más preferentemente más de 10 y más preferentemente más de 50 o 500). En los casos que se utiliza una partícula infectiva, la célula (o población de células) se infecta preferentemente a una multiplicidad de infección mayor de 1 (por ejemplo, m.o.i. mayor de 5, más preferentemente m.o.i. mayor de 20, y más preferentemente m.o.i. de aproximadamente 200 o más). En realizaciones preferentes, las moléculas de ácido nucleico son miembros de una biblioteca en la que cada molécula de ácido nucleico tiene el mismo (idéntico) origen de replicación y/o marcadores seleccionables (por ejemplo, gen o genes de resistencia antibiótica), y/o secuencias regulatorias. Cada una de las moléculas de ácido nucleico contiene también un sustrato para la recombinación que es diferente en cada molécula de ácido nucleico.

El método se entiende fácilmente en referencia a la figura 1. Por motivos de simplicidad, se ha mostrado esquemáticamente la recombinación solamente entre dos fagómidos, pero está claro que el número de vectores diferente que puede infectar una única célula es muy alto.

Los sustratos para la recombinación se clonan dentro de una única columna vertebral del vector, para constituir un fagómido que contiene el origen de replicación ColEI y Ff. Los marcadores seleccionables (por ejemplo, beta lactamasa o gen o genes de resistencia antibiótica) están presentes opcionalmente. Se crean dos fagómidos que portan los

genes para  $V_HA/V_LA$  y  $V_HB/V_LB$  respectivamente. En la recombinación siguiente, se obtienen 4 fagómidos diferentes: los dos fagómidos de partida ( $V_HA/V_LA$  y  $V_HB/V_LB$ ) y dos nuevos fagómidos ( $V_HA/V_LB$  y  $V_HB/V_LA$ ).

Como ya se ha indicado anteriormente, se ha mostrado un esquema simplificado en la figura 1. De hecho, una única célula bacteriana puede infectarse con un número de vectores muy alto. De este modo, en el ejemplo 3, una única célula se infecta con hasta 19 fagómidos. Esto da como resultado un potencial recombinatorial que es enorme. En el ejemplo 3, una única célula puede contener hasta 361 ( $19 \times 19$ ) miembros de la biblioteca diferentes. De forma similar, una célula que contiene 50 construcciones podría, después de la recombinación, contener hasta 2500 miembros de la biblioteca diferentes.

Los miembros de la biblioteca resultante pueden someterse a uno o más ciclos de monitorización de modo que la biblioteca se enriquece para miembros de la población que expresan ciertas propiedades deseadas. Esta biblioteca enriquecida puede combinarse con otra biblioteca para aumentar la diversidad de aquella biblioteca o puede someterse ella misma a otro ciclo de recombinación en células huéspedes, tal como se ha descrito anteriormente. El proceso de selección y recombinación de los miembros seleccionados de la biblioteca se puede repetir tantas veces como se desee hasta que se obtenga una biblioteca que contiene miembros que exhiben las propiedades deseadas.

Debe apreciarse que, en los casos en que los miembros de la biblioteca son capaces de empaquetarse (por ejemplo, en un paquete de expresión replicable), es posible que las partículas producidas de este modo puedan tener un genotipo que está desacoplado del fenotipo del paquete. De este modo, por ejemplo, en los casos que los miembros de la biblioteca contienen un sustrato para la recombinación que expresa una proteína sobre la superficie de un fago (por ejemplo, una proteína de recubrimiento pIII), debido a que la célula contiene muchos miembros de la biblioteca diferentes y la maquinaria de empaquetamiento es "independiente" de la secuencia de ácidos nucleicos que comprenden el miembro de la biblioteca, es posible que la proteína de recubrimiento en la que el miembro de la biblioteca (el ácido nucleico se empaqueta) no sea la proteína de recubrimiento codificada por ese ácido nucleico. En otras palabras, el fenotipo (proteína de recubrimiento) del paquete está "desacoplado" del genotipo (ácido nucleico) del paquete.

Debido a que los miembros de la biblioteca de la presente invención son fácilmente corregibles para utilizarse nuevamente en ciclos adicionales de recombinación/selección, el genotipo y fenotipo del paquete de expresión replicable puede reacoplarse fácilmente (por ejemplo, como se ilustra en la figura 6). Para conseguir esto, la biblioteca de ácidos nucleicos, o una fracción de la misma, se utiliza para retransfectar/reinfectar células huéspedes a una m.o.i. baja o bajo número de copias (es decir, número de copias o m.o.i. de aproximadamente 1). Esto produce una segunda biblioteca "derivada" que tiene, en promedio, un miembro de la biblioteca por célula. En este caso, cuando los miembros de ácido nucleico se reempaquetan, como hay solamente un miembro de la biblioteca (aunque posiblemente muchas copias de ese miembro único de la biblioteca) por célula, esencialmente todos las proteínas de recubrimiento de paquetes genéticos replicables (por ejemplo, proteínas de recubrimiento de fagómidos) no proporcionadas por una célula auxiliar están codificados por el miembro de la biblioteca. Cuando el miembro de la biblioteca de ácidos nucleicos se empaqueta, se empaquetará dentro de una proteína de recubrimiento que comprende proteínas codificadas por ese miembro de la biblioteca. En este caso, el fenotipo del paquete de exposición replicable se codifica por el ácido nucleico contenido dentro de ese paquete. En otras palabras, el genotipo y fenotipo del paquete están unidos.

La recombinación descrita anteriormente puede tener lugar aprovechando los sistemas de recombinación sin recombinasa presentes en la célula, y en este caso se hace referencia a la recombinación general, o ésta puede tener lugar en lugares específicos (por ejemplo, mediarse por recombinasa). La eficiencia de la recombinación general se puede aumentar utilizando cadenas bacterianas que contienen ciertas mutaciones, por ejemplo, *mutS* o *recO98* o con tratamientos físicos, por ejemplo, tratamiento con UV, mutagénesis química, etc.

Habitualmente, la mutación general tiene lugar con una eficiencia mucho más baja que la recombinación mediada por recombinasa. De este modo, en realizaciones preferentes que utilizan la recombinación general, a menudo es deseable incluir una etapa selección para seleccionar positivamente recombinantes anteriormente a etapas de monitorización y/o recombinación posteriores. Varias estrategias de selección son conocidas por aquellos con experiencia en la técnica. Sin embargo, una estrategia selección particularmente preferente se ilustra en el Ejemplo 4 y en la figura 9. Brevemente, uno o más marcadores seleccionables se acoplan a los segmentos que se desea recombinar. Marcadores seleccionables preferentes se seleccionan para que sean defectivos en ausencia de un evento de recombinación. Cuando la recombinación tiene lugar, la propiedad del marcador seleccionable se recupera y el miembro recombinado puede aislarse. En paralelo, en una alta proporción de casos, la recombinación también tendrá lugar en los segmentos que se desea recombinar. De este modo, a modo de ilustración, en el Ejemplo 4, la beta lactamasa tipo salvaje, que normalmente es incapaz de proporcionar resistencia significativa al antibiótico cefotaxima se convierte en un marcador con resistencia significativa a la cefotaxima cuando tienen lugar eventos de recombinación, y de este modo se proporciona un marcador efectivo para la selección de recombinantes.

En los casos en que se utiliza recombinación mediada por recombinasa, habitualmente la recombinación tiene lugar, como mínimo, en uno, más preferentemente, como mínimo, en dos o más lugares de reconocimiento de recombinasa. Habitualmente, los lugares de reconocimiento de recombinasa se organizarán dentro del miembro de la biblioteca de ácidos nucleicos, de manera que estos se sitúen al lado de una o más secuencias que sean sustratos de recombinación (es decir, secuencias que se desea recombinar). De este modo, por ejemplo, en el caso de una biblioteca de anticuerpos, los lugares de reconocimiento de recombinasa se colocarán de manera que estén bordeados por una o más regiones de anticuerpo (por ejemplo, una cadena pesada variable, una cadena ligera variable), etc. Es posible proporcionar muchos

lugares de reconocimiento de recombinasa diferentes en una única construcción y estos pueden ser lugares para un tipo de recombinasa o pueden utilizarse múltiples sistemas recombinasa.

La recombinación mediada por recombinasa puede medirse por una recombinasa endógena en la célula y/o por una recombinasa exógena. La recombinasa exógena puede ser una proteína de recombinasa exógena administrada a la célula, pero en una realización más preferente la recombinasa exógena se expresa mediante un ácido nucleico que ha sido transfectado dentro de la célula. Además, la célula puede ser una célula que ha sido modificada para expresar una recombinasa endógena en un nivel anormalmente alto.

## II. Construcciones y células huésped para la creación de bibliotecas diversas

### A) Células huésped

Virtualmente, cualquier célula capaz de ser transfectada o infectada con los miembros de las bibliotecas de ácidos nucleicos descritos en la presente descripción y capaz de soportar la recombinación (general o mediada por recombinasa) puede utilizarse en la presente invención. Pueden utilizarse células eucarióticas, entre las que se incluyen levaduras, células vegetales, de hongo, insecto, y de mamífero, así como células procarióticas tanto gram negativas como gram positivas, que pueden infectarse por partículas infecciosas que normalmente infectan estas células, o que han sido diseñadas o seleccionadas para infectar estas células. Las células utilizadas pueden ser normales, mutadas para inducir niveles de recombinación más elevados, o ideadas para expresar recombinasas específicas capaces de actuar en secuencias específicas de ácidos nucleicos presentes dentro de las moléculas de ácido nucleico contenidas en las partículas infecciosas. Entre las células particularmente preferentes se incluyen células bacterianas (por ejemplo, *E. coli*), células de levadura, y células de mamífero.

### B) Vectores para su utilización en la presente invención

Los miembros de las bibliotecas de ácidos nucleicos descritos en la presente descripción pueden estar virtualmente en cualquier forma conveniente, siendo formas preferentes las formas que facilitan la replicación y clonación. De este modo, mientras que, en algunas realizaciones, los miembros de la biblioteca de ácidos nucleicos pueden ser simplemente lugares de recombinación con lugares de reconocimiento de recombinasa asociados, los miembros de la biblioteca de ácidos nucleicos preferentes son vectores entre los que se incluyen, pero no constituyen limitación, plásmidos, cósmidos, y fagómidos, fagos filamentosos (M13, F1, fd, etc.), y similares. Aunque no son preferentes, en ciertas realizaciones, particularmente aquellas que utilizan segmentos de ácidos nucleicos extensos, pueden utilizarse vectores P1, YACs, y BACs.

De este modo, tal como se ha indicado anteriormente, la utilización de la presente invención no está limitada a *E. coli* y fagómidos. Para cinco métodos que utilizan vectores de infección, es adecuada cualquier partícula infecciosa que está en posición de infectar una célula de forma que múltiples copias (es decir, dos o más) del ácido nucleico pueden entrar en la célula y convertirse en sustratos para la recombinación. En una realización preferente, una partícula infecciosa puede considerarse que es una molécula de ácido nucleico cubierta con "moléculas de recubrimiento" (por ejemplo, proteínas) que permiten a la partícula entrar en las células con alta eficiencia, por ejemplo, en virtud de interacciones específicas entre las moléculas de recubrimiento y un receptor o marcador sobre la superficie celular. Estas partículas infecciosas pueden darse naturalmente, preparadas para que contengan sustratos de recombinación, o partículas infecciosas recombinantes que carecen de la información genética necesaria para replicarse por ellas mismas, pero son capaces de entrar específicamente en células, y preferentemente a una eficiencia alta. Un fagómido es un ejemplo de una partícula infecciosa de este tipo.

En realizaciones preferentes, los miembros de las bibliotecas de ácidos nucleicos de la presente invención tienen secuencias de ácidos nucleicos (dominios) que son esencialmente idénticas para cada miembro de la biblioteca y otras secuencias que varían entre los miembros de la biblioteca. Cuando se dice que las secuencias son idénticas para "cada uno" de los miembros de la biblioteca, se entenderá que las metodologías de ADN recombinante no son perfectas y que una cierta fracción de la biblioteca contendrá miembros defectivos y/o contaminados en los que las secuencias "idénticas" son diferentes. Estos miembros "anormales" se darán preferentemente a baja frecuencia, de manera que habitualmente, como mínimo, el 80%, preferentemente, como mínimo, el 90%, más preferentemente, como mínimo, el 95%, y más preferentemente, como mínimo, el 99% o, como mínimo, el 99,9% de los miembros de la biblioteca tienen secuencias idénticas.

Las secuencias "idénticas" habitualmente son más largas de un codón y, en realizaciones preferentes, son, como mínimo, 10, preferentemente, como mínimo, 20, más preferentemente, como mínimo, 50 nucleótidos contiguos en longitud. Habitualmente, las secuencias idénticas incluyen, como mínimo, un origen de replicación y/o uno o más marcadores seleccionables (por ejemplo, un gen de resistencia a la ampicilina). Entre otras características presentes opcionalmente en la secuencia "idéntica" se incluyen, pero no constituyen limitación, un líder de secreción, etiquetas (por ejemplo, SV5 y His), un gen o ADNc que codifica un fago o proteína de recubrimiento bacteriana (por ejemplo, gen 3), lugares cebadores de PCR, lugares de reconocimiento de recombinasa (por ejemplo, loxP, loxP511, etc.), vector "columna vertebral", y similares.

En ciertas realizaciones preferentes, la secuencia "idéntica" incluye esencialmente toda secuencia diferente de la secuencia o secuencias que se desea recombinar y las secuencias "idénticas" son preferentemente secuencias del mis-

mo “tipo” de molécula (por ejemplo, fago, plásmido, fagómido, etc.). De este modo, los miembros de la “biblioteca” pueden verse como construcciones idénticas dentro de las que se insertan los sustratos “variables” (de recombinación).

Las secuencias del “sustrato de recombinación” variarán entre los miembros de la biblioteca. No es necesario que todos los miembros de la biblioteca sean diferentes a cada uno de los otros miembros de la biblioteca. Sin embargo, si una biblioteca es extensa y diversa aumenta la probabilidad de seleccionar una molécula que tiene las propiedades deseadas. Se puede ver a una biblioteca que tiene, en promedio  $X$  miembros de los que un  $y\%$  son diferentes (por ejemplo, una biblioteca que tiene  $10^{12}$  miembros de los que el 90% son diferentes contendrá, aproximadamente,  $10^{11}$  miembros diferentes) o alternatively medir el tamaño de la biblioteca como el número medio de miembros diferentes. De este modo, una biblioteca que tiene  $10^{12}$  miembros físicos el 90% de los cuales son diferentes puede considerarse como una biblioteca de, aproximadamente,  $10^{11}$  miembros.

El sustrato de recombinación puede ser una secuencia única o dos o más secuencias. En realizaciones particularmente preferentes, los sustratos de recombinación incluyen, como mínimo, dos secuencias. Las secuencias pueden ser contiguas o pueden contener secuencias interpuestas (por ejemplo, conectores y/o lugares de reconocimiento de recombinasa, etc).

Tal como se ha indicado anteriormente, las secuencias “idénticas” pueden incluir uno o más lugares de reconocimiento de recombinasa. Se ha observado que, cuando tiene lugar la recombinación (por ejemplo, en el sistema Cre-lox) a menudo el lugar de reconocimiento de recombinasa se corta y una parte del lugar se mueve con el ácido nucleico escindido. El lugar se reconstituye de forma efectiva cuando el fragmento se reinserta en el ADN objetivo para que, de forma efectiva, el lugar de reconocimiento de recombinasa permanezca inalterado y se considere “idéntico”.

Entre varios de los lugares de recombinación adecuados se incluyen, pero no constituyen limitación, loxP, mutantes de loxP y similares, los cuales se describen a continuación.

El sustrato de recombinación puede incluir uno o más genes o ADNc que codifican uno o más polipéptidos, o pueden ser secuencias de ácido nucleico con actividad biológica por sí mismas, tales como secuencias implicadas en la regulación genética, secuencias que codifican ribozimas, enzimas de ADN y similares.

Un miembro de la biblioteca particularmente preferente se ilustra por el vector pDAN5 (ejemplo 2). El vector pDAN5 comprende el origen de replicación, la resistencia a la ampicilina, un líder de secreción, las etiquetas (SV5 y His), el gen 3 y el origen de replicación Ff. En el ejemplo que se proporciona, hay dos secuencias “variables” que codifican las regiones  $V_H$  y  $V_L$  de un anticuerpo de cadena única, respectivamente. Un conector que comprende un lugar de reconocimiento de recombinasa (por ejemplo, loxP) se encuentra entre los segmentos  $V_H$  y  $V_L$ . Esta construcción no está limitada a los  $V_H$  y  $V_L$  determinados ilustrados en el ejemplo, sino que se puede sustituir virtualmente por cualquier otro  $V_H$  y/o  $V_L$ . Adicionalmente, no es necesario que los lugares de reconocimiento de recombinasa simplemente estén bordeados por  $V_H$  y  $V_L$ , sino que, por el contrario, puedan estar bordeados por cualquier fragmento o combinación de fragmentos de anticuerpo deseados. Adicionalmente, se pueden proporcionar múltiples lugares de reconocimiento de recombinasa, y se pueden seleccionar para que funcionen con las mismas o con múltiples recombinasas diferentes.

Tal como se ha indicado anteriormente, en ciertas realizaciones preferentes, los miembros de la biblioteca de ácidos nucleicos se diseñan para que sean capaces de infectar una célula huésped. Habitualmente, estos miembros de ácido nucleico se empaquetan en partículas infecciosas y, en estos casos, el componente “idéntico” (invariable) del miembro codifica, preferentemente, una o más secuencias esenciales para el empaquetamiento y/o la replicación y/o la supervivencia dentro de la célula. Los componentes invariables de un vector de este tipo pueden proporcionar todas las secuencias de ácido nucleico necesarias y suficientes para el empaquetamiento, replicación y supervivencia en la célula huésped, o éstos se pueden seleccionar de forma que la replicación, el empaquetamiento y la supervivencia se lleven a cabo con la ayuda de una construcción auxiliar (por ejemplo, un plásmido auxiliar).

### C) Bibliotecas de exposición

Los miembros de las bibliotecas de ácidos nucleicos de la presente invención se pueden diseñar de manera que cuando se empaquetan en un paquete genético replicable (por ejemplo, fago, bacteria, etc.) éstos exponen una proteína codificada sobre la superficie del paquete. En los casos en que la proteína codificada es una que está codificada por uno o más sustratos de recombinación, la colección de paquetes proporciona un sistema conveniente de “exposición” de la inmensa variedad de secuencias recombinadas disponibles. Las proteínas expuestas pueden someterse a una o más etapas de monitorización/selección (por ejemplo, por especificidad/avidez de unión, actividad catalítica, etc) y los miembros de la biblioteca seleccionados se pueden utilizar para generar bibliotecas posteriores para ciclos posteriores de enriquecimiento/selección. Tal como se explica a continuación, la biblioteca de “exposición” puede ser monovalente o polivalente.

#### 1) Bibliotecas de anticuerpos monovalentes y bibliotecas de polipéptidos

La capacidad de expresar polipéptidos y fragmentos de anticuerpo sobre la superficie de virus que infectan bacterias (bacteriófagos o fagos) o bacterias (ver, por ejemplo, Charbit y otros (1988) *Gene*, 70(1): 181-189) hace posible el aislamiento de un polipéptido o fragmento de anticuerpo de unión única de una biblioteca mayor de  $10^{10}$  clones no

enlazantes. Para expresar polipéptidos o fragmentos de anticuerpo sobre la superficie de un fago (exposición en fago), un polipéptido o un fragmento de anticuerpo se inserta dentro de un gen que codifica una proteína de superficie de un fago (por ejemplo, pIII) y la proteína de fusión fragmento-pIII se expone sobre la superficie del fago (McCafferty y otros (1990) *Nature*, 348: 552-554; Hoogenboom y otros (1991) *Nucleic Acids Res.* 19: 4133-4137). Dado que los fragmentos del anticuerpo sobre la superficie del fago son funcionales, los fagos que llevan polipéptidos de unión a antígenos o fragmentos de anticuerpo se pueden separar de los fagos no enlazantes mediante cromatografía de afinidad a antígeno (McCafferty y otros (1990) *Nature*, 348: 552-554). En función de la afinidad del fragmento y del anticuerpo, se obtienen factores de enriquecimiento de 20 veces-1.000.000 veces para una selección de afinidad de un único ciclo. Sin embargo, mediante la infección de bacterias con el fago eluído, se pueden hacer crecer más fagos y someter a otro ciclo de selección. De este modo, un enriquecimiento de 1000 veces se puede convertir en 1.000.000 de veces en dos ciclos de selección (McCafferty y otros (1990) *Nature*, 348: 552-554). Incluso cuando los enriquecimientos son bajos (Marks y otros (1991) *J. Mol. Biol.* 222: 581-597), múltiples ciclos de selección por afinidad pueden conducir al aislamiento de fagos raros. Dado que la selección de la biblioteca de anticuerpos de fago sobre antígenos da como resultado el enriquecimiento, la mayoría de los clones se unen al antígeno tras cuatro ciclos de selección. De este modo, sólo es necesario analizar un número relativamente pequeño de clones (varios cientos) en cuanto a su unión con el antígeno. Habitualmente, estas bibliotecas de exposición en fago, cuando se utilizan para exponer anticuerpos son monovalentes que exponen un anticuerpo único sobre la superficie de cada miembro de la biblioteca.

## 2) Bibliotecas de anticuerpos polivalentes

A diferencia de las bibliotecas en fago que exponen péptidos de forma multivalente, las bibliotecas de anticuerpos en fago habitualmente exponen cadenas Fv de cadena única (scFv) o fragmentos de anticuerpo Fab fusionados con pIII, como copias únicas sobre la superficie del fago utilizando un sistema de fagómido (Marks y otros (1991) *J. Mol. Biol.* 222: 581-597; Sheets y otros (1998) *Proc. Natl. Acad. Sci. USA* 95: 6157-6162.). Tal como se utiliza en la presente descripción, una biblioteca de anticuerpos de exposición en fago polivalente se refiere a una biblioteca en la cual cada miembro (por ejemplo, partículas de fago) expone en promedio dos o más dominios de unión, en los que cada dominio de unión incluye una región pesada variable y una región ligera variable. De forma más general, una biblioteca de exposición en fago multivalente expone, en promedio, dos o más fusiones pIII por partícula de fago. La exposición en fago polivalente se puede conseguir mediante la expresión de diacuerpos (es decir, una proteína formada por la fusión o conjugación de dos anticuerpos de cadena única, (por ejemplo, scFv)) o mediante la exposición de, en promedio, dos o más anticuerpos sobre cada partícula de fago. En contraste, una biblioteca monovalente expone, en promedio, un anticuerpo de cadena única por partícula viral.

### i) Expresión de diacuerpos

Los diacuerpos son dímeros de scFv en los que cada cadena comprende dominios de cadena pesada ( $V_H$ ) y ligera ( $V_L$ ) variables, conectados utilizando un conector (por ejemplo, un péptido conector) que es demasiado corto para permitir el emparejamiento entre los dominios de la misma cadena. Por consiguiente, el emparejamiento tiene lugar entre dominios complementarios de dos cadenas diferentes, creando un dímero estable no covalente con dos lugares de unión (Holliger y otros (1993) *Proc. Natl. Acad. Sci.* 90: 6444-6448). En una aproximación, determinados genes de diacuerpo se subclonan para expresión como fusiones pIII en el fagómido (ver, por ejemplo, Hoogenboom y otros (1991) *Nucleic Acid Res.* 19: 4133-4137). Esto genera fagómidos que expresan predominantemente un único scFv o diacuerpo-fusión pIII después de recuperarlos con un fago auxiliar (Marks y otros (1992) *J. Biol. Chem.* 267: 16007-16010). Los fagómidos de diacuerpos exponen un fragmento de anticuerpo bivalente resultante del emparejamiento intermolecular de una molécula scFv-fusión pIII y una molécula scFv nativa. Utilizando los conocimientos proporcionados por la presente descripción, alguien con experiencia en la técnica puede producir otros diacuerpos de forma rutinaria.

### ii) Expresión polivalente de anticuerpos de cadena única

Como una alternativa a la utilización de diacuerpos, se crean bibliotecas de anticuerpos de exposición en fago en las que cada partícula viral, en promedio, expresa 2 como mínimo, preferentemente 3 como mínimo, más preferentemente, 4 como mínimo, y más preferentemente 5 copias como mínimo de un anticuerpo de cadena única.

En principio, cada copia de pIII sobre el fago (y existe controversia sobre si hay 3 o 5 copias de pIII por fago) debería expresar un anticuerpo. Sin embargo, tiene lugar la proteólisis y el número que se expresa habitualmente es habitualmente menor. De este modo, bibliotecas de anticuerpos multivalentes preferentes se construyen en un vector fago y no en un vector fagómido. Esto significa que no es necesario añadir un fago auxiliar para hacer un fago. Un fago auxiliar lleva dentro de *E. coli* la pIII tipo salvaje, que compite con la fusión scFv-pIII. De este modo, en un vector fagómido, esta competición deja en promedio, solamente 1 (o menos) anticuerpo por fago.

Para producir bibliotecas de anticuerpos multivalentes, los anticuerpos de cadena única, habitualmente expresados en el fagómido, son subclonados a partir del vector fagómido dentro de un vector fago. No se requiere un fago auxiliar y no hay competición entre el pIII tipo salvaje y la fusión scFv-fusión pIII. De este modo, en promedio, el fago expresa dos o más fusiones pIII.

## IV. Sistemas de recombinasa para su utilización en la presente invención

## A) Recombinasas varias

- 5 En la presente invención, el intercambio de segmentos de ácidos nucleicos se consigue mediante la utilización de proteínas de recombinación, entre las que se incluyen recombinasas y cofactores asociados y proteínas. Varias proteínas de recombinación se describen en la técnica. Entre los ejemplos de estas recombinasas se incluyen, pero no constituyen limitación, la familia de las Cre, la familia de las integrasas, y la familia de las resolvasas.
- 10 Cre, una proteína de bacteriófago P1 (Abremski y Hoess (1984) *J. Biol. Chem.* 259 (3): 1509-1514) cataliza el intercambio (es decir, causa la recombinación) entre una secuencia de ADN de 34 pares de bases denominada lugares loxP (lugar de entrecruzamiento) (ver, por ejemplo, Hoess y otros, (1986) *Nucl. Acids Res.* 14 (5): 2287). Cre está disponible comercialmente (Novagen, No. de Catálogo 69247-1).
- 15 La recombinación mediada por Cre es reversible de forma libre. Cre opera en soluciones tampón simples con magnesio o espermidina como cofactor, como es bien conocido en la técnica. Los sustratos de ADN pueden ser o bien lineales o bien superenrollados. Han sido descritos varios lugares loxP mutantes (Hoess y otros, arriba). Uno de éstos, el loxP 511, se recombina con otro lugar loxP 511, pero no se recombinará con un lugar loxP.
- 20 La integrasa es una proteína de bacteriófago lambda que media la integración del genoma lambda dentro del cromosoma de *E. coli*. Las proteínas de integrasa recombinatorial de bacteriófago lambda promueven la recombinación irreversible entre sus lugares de sustrato att como parte de la formación o inducción de un estado lisogénico. La reversibilidad de las reacciones de recombinación da como resultado dos rutas independientes para la recombinación integrativa y escisiva. Cada ruta utiliza una única, aunque solapada, colección de 15 lugares de unión de proteína que comprenden lugares de ADN att. Las interacciones cooperativa y competitiva que implican cuatro proteínas (Int, Xis, IHF y FIS) determinan la dirección de la recombinación. La recombinación integrativa incluye a las proteínas Int e IHF y los lugares attP (240 pares de bases) y attB (25 pares de bases). La recombinación da como resultado la formación de dos nuevos lugares: attL y attR. La recombinación escisiva requiere de Int, IHF, y Xis, y lugares attL y attR para generar attP y attB. Bajo ciertas condiciones, la FIS estimula la recombinación escisiva. Además de estas reacciones normales, debe entenderse que los attP y attB, cuando están colocados sobre la misma molécula, pueden promover la recombinación escisiva para generar dos productos de escisión, uno con attL y otro con attR. De forma similar, la recombinación intermolecular entre moléculas que contienen attL y attR, en presencia de Int, IHF y Xis, puede dar como resultado la recombinación integrativa y la generación de attP y attB. Por lo tanto, mediante la colocación de combinaciones apropiadas de lugares att diseñados a los lados de los segmentos de ADN, en presencia de las proteínas de recombinación apropiadas, se puede direccionar la recombinación escisiva o integrativa, como también las reacciones inversas de cada una.
- 35 Cada uno de los lugares att contiene una secuencia núcleo de 15 pares de bases; la secuencia individual de elementos de significancia funcional se sitúa dentro, fuera, y a través de los límites de este núcleo común (Landy (1989) *Ann. Rev. Biochem.* 58: 913). La recombinación eficiente entre los diferentes lugares att requiere que la secuencia de la región central común sea idéntica en todos los miembros que se recombinan, sin embargo, se ha descubierto actualmente que la secuencia exacta es modificable. Por consiguiente, se ha descubierto actualmente que los derivados de los lugares att con cambios en el núcleo se recombinan, como mínimo, de forma tan eficiente como las secuencias con el núcleo nativo.
- 45 La integrasa actúa para recombinar los lugares attP sobre el bacteriófago lambda (aproximadamente 240 pares de bases) con el lugar attB sobre el genoma de *E. Coli* (aproximadamente 25 pares de bases) (Weisberg y Landy (1983) en *Lambda II*, pág. 211, Cold Spring Harbor Laboratory), para producir el genoma lambda integrado, bordeado por con los lugares attL (aproximadamente 100 pares de bases) y attR (aproximadamente 160 pares de bases). En ausencia de Xis (ver más abajo), esta reacción es esencialmente irreversible. La reacción de integración mediada por integrasa e IHF funciona *in vitro*, con una solución tampón simple que contiene espermidina. La integrasa se puede obtener tal como describe Nash (1983) *Meth. Enzym.*, 100: 210-216. La IHF se puede obtener tal como describen Filutowicz y otros (1994) *Gene* 147: 149-150.
- 50 En presencia de la proteína Xis lambda (escisión) la integrasa cataliza la reacción de attR y attL para formar attP y attB, es decir, promueve la reacción inversa a la descrita anteriormente. Esta reacción se puede aplicar también en la presente invención.
- 55 Se pueden utilizar también varios sistemas de recombinación a partir de varios organismos, basados en los conocimientos y orientaciones proporcionados en la presente descripción (ver, por ejemplo, Hoess y otros (1986) *Nucleic Acids Res.* 14(6):2287; Abremski y otros (1986) *J. Biol. Chem.* 261(1): 391; Campbell (1992) *Bacteriol.* 174(23): 7495; Qian y otros (1992) *J. Biol. Chem.* 267(11): 7794; Araki y otros (1992) *J. Mol. Biol.* 225(1): 25). Muchos de éstos pertenecen a la familia de integrasas de recombinasas (Argos y otros (1986) *EMBO J.* 5: 433-440). Tal como se ha indicado anteriormente, entre éstos posiblemente los estudiados más satisfactoriamente son los sistemas integrasa/att del bacteriófago lambda (Landy (1993) *Current Opinions in Genetics y Devel.* 3: 699-707), el sistema Cre/loxP del bacteriófago P1 (Hoess y Abremski (1990) en *Nucleic Acids and Molecular Biology*, vol. 4. Editores: Eckstein y Lilley, Berlin-Heidelberg: Springer-Verlag; págs. 90-109), y el sistema FLP/FRT de plásmido circular de *Saccharomyces cerevisiae* 2 mu (Broach y otros (1982) *Cell* 29: 227-234). Además, miembros de una segunda familia de

recombinasas específicas de lugar, la familia de las resolvasas (por ejemplo, gamma delta, resolvasa Tn3, Hin, Gin y Cin) se conocen y son adecuados para su utilización en la presente invención. Aunque los miembros de esta familia de recombinasas altamente relacionada están habitualmente limitados a reacciones intramoleculares (por ejemplo, inversiones y escisiones) y pueden requerir factores de codificación del huésped, se han aislado mutantes que eliminan algunos de los requerimientos de factores huésped (Maeser y Kahnmann (1991) *Mol. Gen. Genet.* 230: 170-176), así como también algunas de las limitaciones de recombinación intramolecular.

Otras recombinasas específicas de lugar similares a la lambda Int, y similares a la P1Cre pueden sustituirse por Int y Cre. Estas recombinasas son conocidas. En muchos casos, la purificación de estas otras recombinasas ha sido descrita en la técnica. En los casos en que éstas no son conocidas, se pueden utilizar extractos celulares o los enzimas pueden ser purificados parcialmente utilizando procedimientos descritos para Cre e Int.

Mientras que la Cre e Int se describen en detalle a modo de ejemplo, existen muchos sistemas de recombinasa relacionados y su aplicación a la presente invención descrita se proporciona también según la presente invención. Tal como se ha indicado anteriormente, la familia de integrasas de las recombinasas específicas de lugar se puede utilizar para proporcionar proteínas de recombinación alternativas y lugares de recombinación para la presente invención, como proteínas de recombinación específicas de lugar codificadas por bacteriófago lambda, phi 80, P22, P2, 186, P4 y P1. Mientras que el grupo de proteínas exhibe una diversidad de secuencias inesperadamente extensa, todas estas recombinasas pueden alinearse en sus mitades C-terminales, y esto proporciona un medio de identificación de nuevas recombinasas.

Una región de 40 residuos cercana al extremo C-terminal se conserva particularmente bien en todas las proteínas y es homóloga a una región cercana al extremo C-terminal de la proteína Flp del plásmido 2 mu de levadura. Tres posiciones están perfectamente conservadas en esta familia: histidina, arginina y tirosina se encuentran en las posiciones de alineamiento 396, 399 y 433 respectivas dentro de la región C-terminal bien conservada. Estos residuos contribuyen al lugar activo de esta familia de recombinasas, y sugiere que esa tirosina-433 forme un enlace covalente transitorio con el ADN durante la ruptura y reunión de la cadena (ver, por ejemplo, Argos y otros (1986) *EMBO J.* 5: 433-40).

De forma alternativa, IS231 y otros elementos transponibles de *Bacillus thuringiensis* podrían utilizarse como proteínas de recombinación y lugares de recombinación. El *Bacillus thuringiensis* es una bacteria entomopatogénica cuya toxicidad es debida a la presencia en las esporas de cristales de delta-endotoxina activos contra las plagas agrícolas y vectores de enfermedades humanas y animales. La mayoría de los genes que codifican las proteínas de estas toxinas son portados por un plásmido y están asociadas estructuralmente de forma general con secuencias de inserción (IS231, IS232, IS240, ISBT1 e ISBT2) y transposones (Tn4430 y Tn5401). Se ha demostrado que varios de estos elementos móviles son activos y participan en la movilidad del gen de los cristales, y de este modo contribuyen a la variación de la toxicidad bacteriana.

El análisis estructural de los elementos iso-IS231 indica que están relacionados con el IS 1151 de *Clostridium perfringens* y están relacionados de forma distante con el IS4 y el IS186 de *Escherichia coli*. Al igual que los otros miembros de la familia IS4, éstos contienen un motivo de transposasa-integrasa conservado, encontrado en otras familias de IS y retrovirus.

Adicionalmente, datos funcionales obtenidos del IS231A en *Escherichia coli* indican un modo de transposición no replicativo, con preferencia hacia objetivos específicos. Se han obtenido resultados similares también en *Bacillus subtilis* y *B. thuringiensis* (ver, por ejemplo, Mahillon y otros, (1994) *Genetica* 93: 13-26; Campbell (1992) *J. Bacteriol.* 7495-7499).

En una realización preferente, la presente invención utiliza recombinación conducida por Cre/Lox. Preferentemente, los lugares loxP y loxP 511 se utilizan como lugares Lox. Varias mutaciones de esta secuencia, tales como las que se han descrito en la literatura (ver, por ejemplo, Mack y otros, anteriormente; Hoess y otros (1986), anteriormente; Hoess y otros (1984) *Biochem* 81: 1026-29; Hoess y otros (1985) *Gene*, 40: 325-329; Abremski y otros (1986) *J. Biol. Chem.*, 261: 391-396) son también adecuadas. Pueden utilizarse también secuencias mutadas similares de loxP, que aún deben aislarse, en tanto que estas secuencias sean capaces de servir como lugares de recombinación para Cre.

La recombinasa expresada intracelularmente está presente de forma habitual en una concentración suficiente para conducir adecuadamente la recombinación en los métodos de la presente invención. Cuando se suministra una proteína de recombinasa exógena, la cantidad de recombinasa que conduce la reacción de recombinación se puede determinar utilizando ensayos conocidos. De forma específica, un ensayo de valoración se puede utilizar para determinar la cantidad apropiada de un enzima de recombinasa purificado, o la cantidad apropiada de un extracto.

#### B) Ingeniería/modificación de lugares de recombinación

Las recombinasas anteriores y los correspondientes lugares de recombinasa son adecuados para su utilización en la generación de bibliotecas diversas según los métodos de la presente invención. Sin embargo, lugares de recombinación tipo salvaje contienen frecuentemente secuencias que reducen la eficiencia o la especificidad de las reacciones de recombinación. Por ejemplo, codones múltiples de detención de traducción (*stop*) en los lugares de recombinación de attB, attR, attP, attL y loxP pueden darse en estructuras de lectura múltiples sobre las dos cadenas, de modo que se

reducen las eficiencias de recombinación, por ejemplo, en los casos en que la secuencia de codificación debe cruzar los lugares de recombinación (sólo una estructura de lectura está disponible sobre cada cadena de loxP y lugares attB) o se imposibilita esta recombinación (en attP, attR o attL).

- 5 Por consiguiente, la presente invención da a conocer además lugares de recombinación diseñados que superan estos problemas. Por ejemplo, se pueden diseñar los lugares att para que tengan una o múltiples mutaciones para potenciar la especificidad o eficiencia de la reacción de recombinación y las propiedades de los ADNs producto (por ejemplo, lugares att1, att2 y att3); para disminuir la reacción inversa (por ejemplo, eliminando el Pland H1 de attB). El examen de estos mutantes determina qué mutantes generan actividad recombinatoria suficiente para ser adecuados para la recombinación según la presente invención.

- 10 Las mutaciones pueden ser introducidas en los lugares de recombinación para potenciar la recombinación específica de lugar. Entre estas mutaciones se incluyen, pero no constituyen limitación: lugares de recombinación sin codones de detención de traducción, que permiten que las proteínas de fusión sean codificadas, lugares de recombinación reconocidos por las mismas proteínas pero que difieren en la secuencia de bases de forma que estos reaccionan extensiva o exclusivamente con sus miembros homólogos permitiendo que se contemplen múltiples reacciones. Se puede especificar qué reacciones determinadas tienen lugar por los miembros determinados que están presentes en la mezcla de reacción.

- 20 Existen procedimientos bien conocidos para la introducción de mutaciones específicas dentro de secuencias de ácido nucleico. Varios de éstos se describen en Ausubel y otros (1989-1996) en *Current Protocols in Molecular Biology*, Wiley Interscience, Nueva York. Las mutaciones se pueden diseñar en oligonucleótidos, que pueden ser utilizados para modificar secuencias clonadas existentes, o en reacciones de amplificación. La mutagénesis aleatoria se puede utilizar también si los métodos de selección adecuados están disponibles para aislar el ADN o ARN mutante deseado. La presencia de las mutaciones deseadas se puede confirmar mediante la secuenciación de ácidos nucleicos mediante métodos bien conocidos.

- 30 Se puede utilizar una variedad de métodos para diseñar una región nuclear de un lugar de recombinación dado, para obtener lugares mutados adecuados para su utilización en la presente invención. Entre éstos se incluyen, pero no constituyen limitación, la mutación de la secuencia del núcleo deseado (por ejemplo, mediante mutagénesis específica de lugar, PCR proclive a error, mutagénesis química, etc.) o mediante la recombinación de dos secuencias de ADN parenteral mediante mecanismos de recombinación específicos de lugar (por ejemplo, attL y attR para dar attB) u otros (por ejemplo, homólogos).

- 35 La funcionalidad de los lugares de recombinación mutante se puede demostrar de modo que dependa de la característica determinada que se desea. Por ejemplo, la falta de codones de detención de traducción en un lugar de recombinación se puede demostrar por la expresión de las proteínas de fusión apropiadas. La especificidad de recombinación entre miembros homólogos se puede demostrar por la introducción de moléculas apropiadas dentro de reacciones *in vitro*, y el ensayo para productos de recombinación tal como se describe en la presente descripción o se conoce en la técnica. Entre otras mutaciones deseadas en los lugares de recombinación se pueden incluir la presencia o ausencia de lugares de restricción, señales de inicio de traducción o transcripción, lugares de unión de proteínas, y otras funcionalidades conocidas de las secuencias de bases en los ácidos nucleicos. Los esquemas de selección genética para atributos funcionales particulares en los lugares de recombinación se pueden utilizar según etapas de métodos conocidos. Por ejemplo, se podría conseguir la modificación de lugares para obtener (a partir de un par de lugares que no interaccionan) miembros que sí interaccionan mediante la demanda de delección, a través de la recombinación entre los lugares, de una secuencia de ADN que codifica una sustancia tóxica. De forma similar, la selección de lugares que eliminan secuencias de interrupción de traducción, la presencia o ausencia de lugares de unión de proteína, etc., pueden ser concebidas fácilmente por aquellos con experiencia en la técnica.

#### 50 V. Selección de moléculas para la recombinación (mezclado)

- Virtualmente, cualquier ácido nucleico o cualquier molécula codificada por un ácido nucleico puede recombinarse ("mezclarse") y seleccionarse según los métodos de la presente invención. Se observa que, entre los ácidos nucleicos, se incluyen desoxiriboácidos nucleicos, riboácidos nucleicos y, en algunos casos, péptido-ácidos nucleicos. Entre los ácidos nucleicos se puede incluir, pero no constituyen limitación, ADNs genómicos (ADNg), ADNc, ARNm, ADNc transcritos inversos, productos de amplificación, y similares. De forma similar, entre los polipéptidos, se incluyen polipéptidos naturales y no naturales.

- 60 Entre las moléculas sujetas a recombinación y/o selección posterior según la métodos de la presente invención se incluyen, pero no constituyen limitación, ligandos polipeptídicos (por ejemplo, interleuquinas normales y/o modificadas, factores de crecimiento, y similares), enzimas, proteínas receptoras, marcadores y/o antígenos de la superficie celular, anticuerpos y/o fragmentos de anticuerpo, ácidos nucleicos que codifican ARN catalíticos (por ejemplo, ribozimas), ADN catalíticos, lugares de unión de ADN (por ejemplo, lugares de unión de factores de transcripción y/o recombinasas), y otros elementos regulatorios (por ejemplo, promotores, potenciadores, señales, etc.). Los elementos "funcionales" únicos se pueden recombinar en un único experimento, o grupos de elementos relacionados (por ejemplo, conjuntos de promotores y potenciadores en una ruta metabólica (por ejemplo, una poliquétido-sintasa) pueden recombinarse en una etapa única de recombinación.

En una realización particularmente preferente, los métodos de la presente invención se utilizan para recombinar fragmentos de anticuerpo para producir anticuerpos que tienen especificidades y/o avideces de unión determinadas. Los métodos de la presente invención pueden utilizarse para recombinar regiones  $V_H$  y  $V_L$  o fragmentos de las mismas. En realizaciones preferentes, la recombinación es entre dos o más elementos, entre los que se incluyen, pero no constituyen limitación  $V_H$ ,  $V_L$ ,  $V_H$  CDR1,  $V_H$  CDR2,  $V_H$  CDR3,  $V_L$  CDR1,  $V_L$  CDR2,  $V_L$  CDR3, cadenas individuales de un fragmento Fab (por ejemplo,  $V_H+C_H1$  y  $V_L+C_L$ ), dímeros de  $V_H$ , dímeros de  $V_L$ , y similares. De nuevo, tal como se ha indicado anteriormente, pueden tener lugar múltiples eventos de recombinación al mismo tiempo. De este modo, por ejemplo,  $V_H$  CDR1 y  $V_L$  CDR3 pueden recombinarse en el mismo experimento.

Las bibliotecas de anticuerpos para su utilización en los métodos de la presente invención se pueden obtener mediante métodos bien conocidos por aquellos con experiencia en la técnica (ver, por ejemplo, Marks y otros (1991) *J. Mol. Biol.* 222: 581-597). En una realización, estos métodos implican el aislamiento de repertorios naturales  $V_H$  y  $V_L$  presentes en las células (por ejemplo, linfocitos de sangre periférica humana) por PCR. Los repertorios de gen  $V$  se unen aleatoriamente utilizando PCR para crear un repertorio de genes scFv que se clona dentro de un vector (por ejemplo, un vector fago) para crear una biblioteca de anticuerpos en fago extensa (*Id.*).

Los ácidos nucleicos (por ejemplo, que codifican ligandos, antígenos, receptores, etc.) pueden prepararse utilizando técnicas de ADN recombinante o sintetizarse químicamente según métodos estándar bien conocidos por aquellos con experiencia en la técnica (ver, por ejemplo, Sambrook, y otros (1989) *Molecular Cloning: a Laboratory Manual* (2ª Ed., Vols. 1-3), Cold Spring Harbor Laboratory, Berger & Kimel (1987) *Methods in Enzymology*, vol. 152: *Guide to Molecular Cloning Techniques*, Academic Press, San Diego, Ausubel y otros (1987) *Current Protocols in Molecular Biology*, Greene Publishing y Wiley-Interscience, Nueva York, etc.).

Las diferentes moléculas descritas anteriormente, y otras, se someterán a recombinación y selección para obtener nuevas moléculas que tienen propiedades nuevas y/o alteradas tales como las descritas a continuación.

#### VI. Monitorización de polipéptidos recombinados

Las proteínas y/o ácidos nucleicos recombinados según los métodos de la presente invención se pueden monitorizar para una o más propiedades. A continuación, aquellos miembros de las bibliotecas que cumplan los criterios de monitorización pueden utilizarse, por ellos mismos, para ciclos de recombinación y selección posteriores o pueden utilizarse para enriquecer otras bibliotecas para ciclos de recombinación y selección posteriores.

Los miembros de las bibliotecas se pueden monitorizar virtualmente por cualquier propiedad o actividad. De este modo, por ejemplo, los miembros de ácido nucleico se pueden monitorizar por la capacidad de hibridarse con secuencias objetivos determinadas bajo condiciones de hibridación preseleccionadas, o éstos pueden monitorizarse por la actividad catalítica (por ejemplo, actividad de ARNasa), o éstos se pueden monitorizar por la actividad regulatoria alterada (por ejemplo, respuesta a determinados promotores, determinada especificidad tisular, etc.) y similares. De forma similar, las proteínas se pueden monitorizar por la capacidad para unirse específicamente a un objetivo determinado (por ejemplo, proteína, receptor, glicoproteína, grasa, etc. objetivos) a una avidez mínima seleccionada, o por la capacidad para catalizar una determinada reacción química bajo ciertas condiciones (por ejemplo, a una temperatura determinada, o pH, o en disolventes determinados (por ejemplo, disolventes orgánicos, etc.) o bajo condiciones reductoras o oxidantes o en presencia de proteasas, etc.), por su estabilidad bajo determinadas condiciones (por ejemplo, en presencia de ciertos desnaturalizantes, a un pH determinado, o temperaturas extremas, etc.).

De este modo, por ejemplo, en el caso de anticuerpos o proteínas de unión, la especificidad y/o avidez de unión se puede determinar en un BiaCore, un biosensor basado en la resonancia plasmónica superficial. Para esta técnica, el objetivo (por ejemplo, el antígeno, receptor, etc.) se acopla a un chip sensor derivatizado capaz de detectar cambios en masa. Cuando los miembros de las bibliotecas pasan sobre el chip sensor, algunos miembros de las bibliotecas se unen al objetivo inmovilizado dando como resultado un aumento en masa que es cuantificable. La medida de la velocidad de asociación como una función de la concentración de un miembro determinado puede utilizarse para calcular la constante de velocidad de asociación ( $k_{on}$ ). Después de la fase de asociación, una solución tampón se hace pasar sobre el chip y se determina la velocidad de disociación ( $k_{off}$ ) del anticuerpo. En ciertas realizaciones, habitualmente se mide  $k_{on}$  en el intervalo de  $1,0 \times 10^2$  a  $5,0 \times 10^6$  y  $k_{off}$  en el intervalo de  $1,0 \times 10^{-1}$  a  $1,0 \times 10^{-6}$ . La constante de equilibrio  $K_d$  se calcula a menudo como  $k_{off}/k_{on}$  y, de este modo, se mide habitualmente en el intervalo de  $10^{-5}$  a  $10^{-12}$ . Las afinidades medidas de esta manera se correlacionan bien con las afinidades medidas en solución por valoración por eliminación de fluorescencia.

Las actividades catalíticas o enzimáticas se pueden medir proporcionando un sustrato y un sistema tampón apropiado y determinando la velocidad de reacción (por ejemplo, monitorizando la velocidad de pérdida de material de partida o la velocidad de producción de producto de reacción). De este modo, por ejemplo, la actividad de una proteasa recombinada según los métodos de la presente invención se puede determinar proporcionando un sustrato polipeptídico apropiado que porta moléculas fluorescentes que se bloquean mutuamente la fluorescencia. A continuación, el miembro o los miembros de la biblioteca dividen el sustrato, las moléculas fluorescentes se separan, cesa el bloqueo de la fluorescencia mutua y una señal fluorescente es detectable (ver, por ejemplo, la Patente U.S.A. 5.714.342). Se pueden hacer medidas similares para actividad catalítica de ARN o ADN.

La actividad de elementos que controlan la expresión génica se puede ensayar utilizando genes informadores (es decir, genes o ADNc que codifican un producto detectable, por ejemplo, p-galactosidasa, proteína fluorescente verde, etc.) enlazados operativamente a estos elementos. La monitorización del nivel de expresión del gen o genes informadores bajo circunstancias determinadas (por ejemplo, en un tejido o célula determinado, en presencia de un inductor determinado, etc.) proporciona una medida de la actividad de los elementos de control.

Se observa que esta lista de criterios de monitorización y ensayos posibles es ilustrativa y no se pretende que constituya limitación. El ensayo particular apropiado se determinará por la naturaleza del miembro de la biblioteca y la actividad o propiedad que se desea evaluar. Ensayos para esencialmente cualquier actividad enzimática o catalítica, propiedades físicas o químicas, o especificidad y/o aidez de unión son conocidos por aquellos con experiencia en la técnica y se pueden llevar a cabo con la mayoría de la experimentación rutinaria.

Aunque la monitorización de las bibliotecas de la presente invención se puede llevar a cabo manualmente, se entenderá que muchos de estos ensayos pueden estar altamente automatizados. De hecho, en la actualidad, es habitual monitorizar bibliotecas combinatoriales extensas utilizando técnicas de "alta capacidad de procesamiento" que permiten que varios cientos de miles de ensayos se lleven a cabo en tan poco tiempo como una semana. Los ensayos de alta capacidad de procesamiento para la presencia, ausencia, o cuantificación de productos de ácidos nucleicos o proteínas determinados se conocen bien por aquellos con experiencia en la técnica. De forma similar, ensayos de unión y ensayos para actividades químicas determinadas son igualmente bien conocidos. De este modo, por ejemplo, la Patente U.S.A. 5.559.410 describe métodos de monitorización de alta capacidad de procesamiento para proteínas y la Patente U.S.A. 5.585.639 da a conocer métodos de monitorización de alta capacidad de procesamiento para unión de ácidos nucleicos (es decir, utilizando matrices de alta densidad), mientras que las Patentes U.S.A. 5.576.220 y 5.541.061 describen métodos de monitorización de alta capacidad de procesamiento para la unión ligando/anticuerpo.

Los sistemas de monitorización de alta capacidad de procesamiento pueden ser diseñados a medida para los ensayos particulares de interés. Alternativamente, varios sistemas generales de alta capacidad de procesamiento se pueden adaptar para una amplia variedad de ensayos. Sistemas de monitorización de alta capacidad de procesamiento están disponibles comercialmente (ver, por ejemplo, Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, etc.). Estos sistemas habitualmente automatizan procedimientos enteros que incluyen el pipeteo de toda la muestra y reactivos, la administración de líquido, las incubaciones calculadas, y las lecturas finales de microplaca en el detector o los detectores apropiados para el ensayo. Estos sistemas configurables proporcionan alta capacidad de procesamiento y una rápida puesta en marcha así como un alto grado de flexibilidad y personalización. Los fabricantes de estos sistemas proporcionan protocolos detallados para varios de estos aparatos con alta capacidad de procesamiento. De este modo, por ejemplo, Zymark Corp. proporciona boletines técnicos que describen sistemas de monitorización para detectar la modulación de la transcripción genética, unión de ligando, y similares.

## VII. *Cojuntos (kits)*

En una realización, la presente invención da a conocer conjuntos para su utilización en los métodos de la presente invención. Un conjunto preferente comprenderá habitualmente un contenedor que contiene una o más bibliotecas según la presente invención. La biblioteca puede tomar varias formas. De este modo, en una realización, la biblioteca es una colección de células que contienen miembros de una biblioteca de exposición en fago, mientras que, en otra realización, la biblioteca comprende una colección de fagos aislados. Otra biblioteca adicional comprende una biblioteca de ácidos nucleicos empaquetados (por ejemplo, una biblioteca de plásmidos). Los ácidos nucleicos pueden ser vectores fagómidos que codifican el sustrato o los sustratos de recombinación deseados, preparados para la subclonación dentro de un vector fago o los ácidos nucleicos pueden ser una colección de fagómidos que portan ya los ácidos nucleicos subclonados.

La biblioteca puede ser una biblioteca primaria que comprende una población de moléculas de ácido nucleico individuales que no ha sido sometida aún a recombinación tal como se describe en la presente descripción, o alternativamente, una biblioteca secundaria que comprende moléculas de ácido nucleico derivadas de una biblioteca de genes primaria, en la que la diversidad ha sido aumentada por el proceso de recombinación específica de lugar o bien general.

En otra realización, el conjunto puede comprender un contenedor que contiene vectores (por ejemplo, plásmidos, fagómidos, etc.) que contienen lugares de reconocimiento de recombinasa localizados de forma apropiada y, opcionalmente, que proporcionan policonectores u otros lugares para facilitar la clonación de uno o más sustratos de recombinación.

Adicionalmente, los conjuntos pueden comprender opcionalmente uno o más reactivos (por ejemplo, placas de microvaloración, soluciones tampón, células, moléculas informadoras, antibióticos, etc.) adecuados para la práctica de los métodos descritos en la presente descripción.

Adicionalmente, entre los conjuntos se pueden incluir materiales de instrucción que contienen directrices (es decir, protocolos) para la práctica de los métodos de la presente invención. Los materiales de instrucción preferentes proporcionan protocolos para generar bibliotecas de ácidos nucleicos extensas y diversas y/o para monitorizar los miembros de estas bibliotecas. Aunque los materiales de instrucción comprenden habitualmente materiales escritos o impresos,

no se limitan a éstos. Cualquier medio capaz de almacenar estas instrucciones y comunicarlas a un usuario final está contemplado por la presente invención. Entre estos medios se incluyen, pero no constituyen limitación, medios de almacenamiento electrónico (por ejemplo, discos magnéticos, cintas, cartuchos, chips), medios ópticos (por ejemplo, CD ROM), y similares. Entre estos medios se pueden incluir direcciones de sitios de Internet que proporcionan estos materiales de instrucción.

### Ejemplos

Los siguientes ejemplos se describen para ilustrar, pero no para que constituyan limitación de la invención reivindicada.

#### Ejemplo 1

*Las bacterias se pueden infectar por más de un fagómido*

Como una primera etapa hacia el desarrollo de un sistema de recombinación *in vivo*, el número de fagómidos diferentes que podrían infectar bacterias únicas fue evaluado utilizando cinco fagómidos diferentes que expresan diferentes resistencias a los antibióticos: resistencia a la ampicilina (Bluescript: Stratagene, La Jolla), resistencia a la kanamicina (pMPM-K1), resistencia a la tetraciclina (pMPM-T1), resistencia al cloranfenicol (pBSL 121) y resistencia a la gentamicina (pBSL 141). Los plásmidos pMPM son de Mayer (1995) *Gene* 163: 41-46 y los plásmidos pBSL de Alexeyev y otros (1995) *Gene* 160: 63-67. Mediante la infección de bacterias con partes alícuotas de estos fagómidos y la colocación en placas sobre antibióticos simples o dobles, y considerando el número de colonias presentes sobre las placas sin antibióticos como el 100%, puede observarse (Tabla 1) que la mayoría de las bacterias infectadas con dos fagómidos diferentes tienen dos resistencias diferentes.

TABLA 1

*Entrada y supervivencia de fagómidos en una célula que muestra resistencia a diferentes antibióticos*

|               |              |            |               |             |            |
|---------------|--------------|------------|---------------|-------------|------------|
| Tetraciclina  | 100          |            |               |             |            |
| Kanamicina    | 85           | 74         |               |             |            |
| Cloranfenicol | 70           | 63         | 64            |             |            |
| Gentamicina   | 73           | 33         | 39            | 52          |            |
| Ampicilina    | 94           | 81         |               |             | 92         |
|               | Tetraciclina | Kanamicina | Cloranfenicol | Gentamicina | Ampicilina |

Por ejemplo, el 94% de todas las bacterias son resistentes a la ampicilina y a la tetraciclina, y el 85% son resistentes a la tetraciclina y a la kanamicina. Como mínimo se estima que el 10-40% de las bacterias parece que ha tenido el potencial para mostrar resistencia a los cinco antibióticos. La resistencia a la gentamicina con otros antibióticos se encontró más infrecuentemente, pero esto es probablemente una característica de la resistencia a este antibiótico, ya que sólo el 52% de las bacterias pudieron proporcionar resistencia a la gentamicina, incluso cuando se utilizaron solas. La resistencia a algunos pares de antibióticos no se indica ya que los plásmidos utilizados no permitieron a los presentes inventores examinar estas combinaciones.

Estos datos muestran que la infección utilizando fagómidos es un método eficiente para introducir, como mínimo, cinco moléculas de ácido nucleico diferentes dentro de una célula.

#### Ejemplo 2

*La recombinación intracelular se puede utilizar para construir un anticuerpo funcional con sus componentes de cadena presentes en fagómidos separados*

A efectos de utilizar recombinasa cre para mezclar los genes de región variable de los anticuerpos en el formato scFv, se colocó un lugar lox entre los genes de cadena pesada y ligera. Esto implicó la utilización de un lugar lox traducido como conector proteico. Un examen de las seis posibles estructuras disponibles para el lugar loxP tipo salvaje y el lugar loxP 511 mutado (que no se recombinará con el loxP tipo salvaje (Hoess y Abremski (1985) *J Mol Biol*, 181: 351-362) identificó una translación del loxP 511 (ITSYNVYYTKL, Identificación de Secuencia No:1) que tenía sólo un único aminoácido básico (para reducir la posibilidad de proteólisis), carecía de codones de parada y fue la menos hidrofóbica. La secuencia del conector tal como se utilizó en los scFvs se da en las figuras 2 y 3. La capacidad de esta secuencia de actuar como conector de scFv así como dos conectores usados habitualmente, (gly<sub>4</sub>ser)<sub>3</sub> (Identificación de Secuencia No:2) (Bird y otros (1988) [errata publicada aparece en *Science* 28 de abril 1989; 244 (4903):409].

*Science* 242:423-426) y 220, una versión modificada del conector 218 (Whitlow (1993) *Protein Engineering*, 6: 989-995), se llevaron a cabo para crear pequeñas bibliotecas scFv para cada uno de estos conectores y se evaluaron los niveles de exposición mediante el desarrollo de transferencias western con el anticuerpo SV5 que reconoce la etiqueta entre el scFv clonado y la proteína del gen 3. El conector del loxP 511 mostró niveles de expresión tan buenos, o incluso mejores, que los otros dos conectores más ampliamente utilizados.

Sobre la base de este resultado, se diseñó y se construyó un vector de exposición en fagómido nuevo, pDAN5, (figuras 3, 4, Identificación de Secuencia No:3). (Nótese que la secuencia ilustrada es el vector pDAN5 D1.3 que tiene el scFv D1.3 clonado en sí. Los nucleótidos 102-426 son el gen  $V_L$  D1.3 y los nucleótidos 490-837 son el gen  $V_H$  D1.3. El vector pDAN5 básico (anticuerpo ausente) no incluye los nucleótidos 102-426 y 490-837, e incluirán secuencias policonectoras en su lugar). Este vector se examinó por la capacidad de exponer tres scFvs diferentes derivados de anticuerpos monoclonales caracterizados previamente. Los resultados muestran que los tres scFvs reconocen los antígenos apropiados cuando se exponen en este nuevo vector utilizando el conector loxP 511. Las señales de ELISA obtenidas utilizando este vector fueron similares aquéllas obtenidas utilizando vectores de exposición con el conector gly-ser.

Para hacer pDAN5, se clonó un nuevo policonector dentro de pUC119 utilizando EcoRI y HindIII mediante PCR de solapamiento de dos oligonucleótidos largos. Esto introdujo la secuencia bacteriana líder, un casete de policlonación que contiene los lugares de restricción indicados en negrita en la figura 3, la etiqueta SV5 (Hanke y otros (1992) *J Gen Virol*, 73: 653-660.), una etiqueta His<sub>6</sub> (Identificación de Secuencia No:4) y un codón de parada ámbar (ver las figuras 3 y 4). El gen 3 se clonó posteriormente dentro de su policonector mediante amplificación de PCR a partir de fdtet utilizando NotI y EcoRI. El extremo 5' del gen 3 maduro se insertó corriente abajo del codón de parada ámbar, y una secuencia lox de tipo salvaje se insertó en el extremo 3' del gen 3 después del codón de parada y antes del lugar EcoRI mediante PCR. El scFv D1.3 se ensambló (en el orden VL-VH) a partir de scFv D1.3 (orden VH-VL) amplificando el VH D1.3 con VHback-DAN y VHfor-2-DAN, Y VK D1.3 con VK2back-DAN y VK2for-DAN (Tabla 2).

TABLA 2

*Cebadores utilizados para la amplificación por PCR, además de aquellos descritos por Sblattero y Bradbury (1998) Immunotechnology 3:271-8*

| Nombre de oligo | Secuencia de oligonucleótido                                                     | No de SEC. |
|-----------------|----------------------------------------------------------------------------------|------------|
| VHback-DAN      | <u>TTA TCC TCG AGC GGT ACC</u> SAG GTS MAR CTG CAG<br>SAG TCW GG                 | 5          |
| VHfor2-DAN      | <u>GAT TGG TTT GCC GCT AGC</u> TGA GGA GAC GGT GAC<br>CGT GGT                    | 6          |
| VK2back-D       | <u>AGC AAG CGG CGC GCA TGC</u> CGA CAT CGA GCT<br>CAC CCA GTC TC                 | 7          |
| VK2for-DA       | <u>GAA GTT ATG GTC GAC CCT CCG</u> GAA CGT TTG ATC<br>TCG AGC TTG GTC CC         | 8          |
| VLbackPT1       | CGC TGG ATT GTT ATT ACT CGC AGC AAG CGG CGC<br>GCA TGC C                         | 9          |
| VLbackPT2       | TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA<br>TTA CTC                           | 10         |
| VHforPT1        | CCA GGC CCA GCA GTG GGT TTG GGA TTG GTT TGC<br>CGC TA                            | 11         |
| VHforPT2        | TGG TGA TGG TGA GTA CTA TCC AGG CCC AGC AGT<br>GGG TTT G                         | 12         |
| VLforPTL        | ACC GCT CGA GGA TAA CTT CGT ATA GTA TAC ATT<br>ATA CGA AGT TAT GGT CGA CCC TCC   | 13         |
| VHbackPTL       | GGA GGG TCG ACC ATA ACT TCG TAT AAT GTA TAC<br>TAT ACG AAG TTA TCC TCG AGC GGT A | 14         |

## ES 2 268 907 T3

Todos los genes V se purificaron por gel y se utilizaron aproximadamente 200 ng como plantillas para una amplificación posterior para añadir una región de solapamiento en el conector scFv, así como colas largas para facilitar la restricción de la digestión enzimática. Los cebadores VLbackPT1 y VLforPTL se utilizaron para amplificar los genes VL y VHforPT1 y VHbackPTL para los genes VH. Las bandas amplificadas se purificaron por gel tal como se describe a continuación.

El scFv D1.3 se ensambló mediante la mezcla de cantidades iguales (50-200 ng) de genes VH y VL y llevando a cabo el ensamblaje esencialmente tal como se describe en Krebber y otros (1997) *J. Immunol. Meth.* 201: 35-55): 8 ciclos de PCR sin cebadores seguido de 25 ciclos en presencia de VLbackPT2 y VHforPT2. Los parámetros de ciclación fueron 94°C durante 1 minuto (desnaturalización), 60°C durante 1 minuto (hibridación) y 72°C durante 1'30" (extensión).

El scFv amplificado se digirió con BssHII y NheI y se ligó en pDAN5 cortado con BssHII/NheI. La mezcla de ligado se separó por electroforesis en DH5 $\alpha$ f electrocompetente y se colocó en placas 2XTY ampicilina 100  $\mu$ g/ml / glucosa al 1%, los clones se confirmaron por secuenciación.

La capacidad para mezclar genes V de cadena pesada y ligera *in vivo* para crear anticuerpos funcionales se examinó utilizando el scFv derivado del mAb antilisoizima, D1.3. Se crearon dos scFvs que contenían D1.3 VH o bien D1.3 VL con cadenas complementarias irrelevantes (VL/D1.3-VH/X y VL/Y-VH/D1.3) mediante clonación por PCR. El reconocimiento de la lisozima por scFv D1.3 mostró que se requería la presencia de las cadenas pesada y ligera de D1.3; cadenas D1.3 únicas asociadas con cadenas complementarias irrelevantes no fueron funcionales. Un resumen del esquema utilizado se indica en la figura 5.

Se hicieron crecer 10 ml de *E. coli* BS1365, que expresa recombinasa cre constitutivamente (BS1365: BS591 F' kan (BS591: recA1 endA1 gyrA96 thi-1 D lacU169 supE44 hsdR17 lamdalm434 nin5 XI-cre (Sauer y Henderson (1988) *Gene* 70: 331-341)) a OD550 0,5 a 37°C en 2XTY con kanamicina 100  $\mu$ g/ml/ glucosa al 1%.

Se añadieron cantidades iguales de fagómido que contiene los dos genes scFv (VL/D1.3-VH/X y VL/Y-VH/D1.3) a las bacterias a un MOI de 20:1 ( $5 \times 10^{10}$  de cada fagómido añadido a  $5 \times 10^9$  de bacterias). Esto se dejó durante 30 minutos a 37°C sin agitación para permitir que tuviera lugar la infección.

Se añadió ampicilina a 100  $\mu$ g/ml y las bacterias se hicieron crecer durante, como mínimo, 12 horas a 30°C. La recombinación tuvo lugar durante este periodo.

Después del crecimiento durante 12 horas, las bacterias se diluyeron 1:20 en 10 ml del mismo medio de crecimiento, se hicieron crecer a OD550 0,5 y se añadió fago auxiliar M13 K07 a un MOI de 20:1.

Esto se dejó durante 30 minutos a 37°C sin agitación para permitir que tuviera lugar la infección.

El cultivo se hizo crecer durante 6-18 horas a 30°C, se centrifugó a 4000 rpm durante 15 minutos y se sacó el sobrenadante. Esta etapa prepara el fagómido que no expresa scFv, con el fenotipo o el genotipo (habitualmente) acoplado.

Se hicieron crecer 10 ml de DH5 $\alpha$ F a OD550 0,5 en 2XTY a 37°C.

Los fagómidos preparados en la etapa 6 se añadieron al DH5 $\alpha$ F a una MOI menor de 1, se dejaron durante 30 minutos a 37°C y se colocaron sobre placas 2XTY con ampicilina 100  $\mu$ g/ml/glucosa al 1%. Esta etapa acopla el fenotipo al genotipo, en ausencia de esta etapa, el scFv expresado puede no corresponder necesariamente al gen scFv dentro del fagómido.

El fagómido que expresa scFv se preparó a partir de 20 colonias individuales mediante crecimiento a OD550 0,5 en 10 ml de 2XTY ampicilina 100  $\mu$ g/ml/glucosa al 1%, infección con M13K07 a una MOI 20:1 durante 30 minutos a 37°C, centrifugación y resuspensión de las bacterias en 10 ml de 2XTY ampicilina 100  $\mu$ g/ml. Los cultivos se hicieron crecer durante 6-18 horas a 30°C.

Si la recombinación fue exitosa, cada bacteria debería contener cuatro genes scFv diferentes ( $V_L/D1.3-V_H/D1.3$ ;  $V_L/D1.3-V_H/X$ ;  $V_L/Y-V_H/D1.3$  y  $V_L/Y-V_H/X$ ) (ver la figura 5). En presencia de recombinasa cre, el 25% (16/64) de las colonias bacterianas obtenidas en las etapas 8/9 mostró que habían experimentado recombinación por PCR. Adicionalmente, se mostró que el 25% del fagómido preparado en la etapa 9 podía reconocer lisozima por ELISA. En bacterias que no expresan cre, no se encontró recombinación y no se identificó D1.3 funcional, bien por PCR o por ELISA. Esto mostró que la recombinación se indujo tal como se había predicho en aquellas células que expresan cre, en los lugares loxP encontrados entre  $V_L$  y  $V_H$  y en el extremo del gen 3, siendo el resultado neto un intercambio de  $V_H$  (y de gen 3 enlazado, que es constante) entre fagómidos diferentes.

## Ejemplo 3

*Creación de una biblioteca de genes extensa de anticuerpos en pDAN5 por medio de recombinación intracelular*A) *Creación de una biblioteca de genes scFv primaria por técnicas de clonación estándar*

Se creó una biblioteca de genes scFv en fagómido primaria que comprende  $7 \times 10^7$  clones independientes, en pDAN5 por clonación de scFv ensamblado por PCR derivado de genes de  $V\mu$ ,  $V\lambda$ , y  $V\kappa$  ( $V\mu$ ,  $V\lambda$ , y  $V\kappa$ ) de sangre periférica de la siguiente forma:

1. Se prepararon linfocitos de sangre periférica humana por centrifugación por gradiente de densidad sobre Ficoll Hypaque (Pharmacia).
2. El ARN total se preparó a partir de estos linfocitos mediante tiocianato de ácido guanidínico, extracción con cloroformo y fenol y precipitación en isopropanol (Chomczynski y Sacks (1987) *Anal. Biochem.* 162: 156-159).
3. El ADNc se preparó utilizando Transcriptasa H-Inversa SuperScript II RNasa (Gibco BRL) con hexámeros aleatorios comenzando con 1-5  $\mu$ g de ARN total en un volumen final de 20  $\mu$ l siguiendo las instrucciones proporcionadas con el SuperScript.
4. Los genes  $V_H$  de IgM se amplificaron en primer lugar a partir de 0,5  $\mu$ l de la reacción de ADNc, utilizando IgMfor y los cebadores  $V_H$ back descritos en (Sblattero y Bradbury, (1998) *Immunotechnology* 3: 271-8). Los volúmenes de reacción fueron de 20  $\mu$ l, utilizando 0,5  $\mu$ l de reacción de ADNc, 10 pmol de cada cebador, 200  $\mu$ M de dNTPs, 2  $\mu$ l de solución tampón de PCR 10X, y 0,5  $\mu$ l (2,5 U) de polimerasa de ADN Taq (Perkin Elmer): Los parámetros de ciclación fueron 94°C durante 1 minuto (desnaturalización), 55°C durante 1 minuto (hibridación) y 72°C durante 1' (extensión) durante treinta ciclos. La totalidad de los 20  $\mu$ l se cargaron en un gel de agarosa al 2% (FMC) y se purificaron por gel utilizando el conjunto de purificación Qiagen (Qiagen). Posteriormente, los genes  $V_H$  se reamplificaron utilizando la mezcla de cebadores  $V_H$ for con los cebadores  $V_H$ back individuales descritos en (Sblattero y Bradbury, 1998) en volúmenes de 50  $\mu$ l utilizando 1  $\mu$ l de  $V_H$  purificado (los otros parámetros tal como anteriormente).
5. Los genes  $V\lambda$  y  $V\kappa$  fueron amplificados de forma similar (utilizando cebadores VLback individuales con la mezcla de VL para cebadores) a partir del ADNc derivado previamente.
6. La totalidad de los genes V fueron purificados por gel y se utilizaron 200 ng aproximadamente, como plantillas para la amplificación para añadir una región de solapamiento en el conector scFv así como colas largas para facilitar la restricción de la digestión enzimática. Se utilizaron los cebadores VLbackPTI y VLforPTL para amplificar los genes VL y los VHforPTI y VHbackPTL para los genes VH. Las bandas amplificadas se purificaron por gel tal como anteriormente.
7. Se ensambló la biblioteca de scFv mediante la mezcla de cantidades iguales (200-500 ng) de genes VH y VL y llevando a cabo el ensamblaje esencialmente tal como se describe en (Kreber y otros, 1997): 8 ciclos de PCR sin cebadores seguido de 25 ciclos en la presencia de VLbackPT2 y VHforPT2. Los parámetros de ciclación fueron 94°C durante 1 minuto (desnaturalización), 60°C durante 1 minuto (hibridación) y 72°C durante 1'30" (extensión).
8. El scFv amplificado se digirió con BssHII y NheI y se ligó dentro de pDAN5-D1.3 cortado con BssHII/NheI. La mezcla de ligado se separó por electroforesis en DH5af electrocompetente y se colocó sobre placas 2XTY con ampicilina 100  $\mu$ g/ml/glucosa al 1% para obtener una biblioteca primaria que consiste en  $7 \times 10^7$  clones independientes.
9. Las colonias se rascaron en 2XTY con glicerol al 10% y se congelaron en partes alícuotas de 1 ml.

B) *Recombinación intracelular para crear una biblioteca de genes secundaria extensa*

Para crear la biblioteca extensa secundaria se siguió el esquema ilustrado en la figura 6. El protocolo detallado es tal como sigue:

1. 100  $\mu$ l de bacterias de la biblioteca primaria se diluyeron en 100 ml de 2XTY con ampicilina 100  $\mu$ g/ml/glucosa al 1% y se hicieron crecer a un OD<sub>550</sub> 0,5 y se añadió fago auxiliar M13 K07 a una MOI de 20:1.
2. Esto se dejó durante 30 minutos a 37°C sin agitación para permitir que tuviera lugar la infección.
3. Se añadió kanamicina a 100  $\mu$ g/ml y el cultivo se hizo crecer durante 5-18 horas a 30°C, se centrifugó a 4000 rpm durante 15 minutos y se sacó el sobrenadante. En esta etapa se prepara el fagómido que no expresa scFv, pero que contiene genes de scFv.

## ES 2 268 907 T3

4. Se hizo crecer 20 ml de *E. coli* BS1365, que expresa recombinasa cre constitutivamente, a un OD550 0,5 a 37°C en 2XTY con kanamicina 100 µg/ml/glucosa al 1%.
5. Los fagómidos preparados en la etapa 3 se añadieron a una MOI de 200:1 (2 x 10<sup>12</sup> fagómidos se añadieron a 1 x 10<sup>11</sup> bacterias). Esto se dejó durante 30 minutos a 37°C sin agitación para permitir que tuviera lugar la infección. Esto dio como resultado la infección de la bacteria por más de un fagómido.
6. Se añadió ampicilina a 100 mg/ml y se hicieron crecer las bacterias durante, como mínimo, 12 horas a 30°C. La recombinación tiene lugar durante este periodo.
7. Después del crecimiento durante aproximadamente 12 horas, se añadieron las bacterias a 380 ml del mismo medio de crecimiento (dilución 1/20), se hicieron crecer a OD550 0,5 a 37°C y se añadió fago auxiliar M13 K07 a una MOI de 20:1. Las bacterias se hicieron crecer durante 6-18 horas adicionales, y los fagómidos se prepararon tal como se ha descrito anteriormente. Esto produce una biblioteca diversa de fagómidos de anticuerpo recombinados genéticamente, pero que no expresa proteína (que tendría de todos modos el fenotipo y el genotipo desacoplado). Esto se soluciona mediante la reinfección de DH5αF a una MOI menor que 1.
8. Se hizo crecer 1 litro de DH5αF<sup>+</sup> a OD550 0,5 en 2XTY/glucosa al 1% a 37°C, se añadieron 5 x 10<sup>11</sup> fagómidos (MOI menor que 1) y se dejó durante 30 minutos para permitir que tuviera lugar la infección. Se añadieron M13K07 a una MOI de 20:1, el cultivo se dejó durante 30 minutos a 37°C para permitir que tuviera lugar la infección, y a continuación, se centrifugó a 4000 rpm durante 15 minutos. Las bacterias se resuspendieron en 1 litro de 2XTY con 100 µg/ml de ampicilina/100 µg/ml de kanamicina y se hicieron crecer durante 6-18 horas a 30°C. El cultivo se centrifugó (4000 rpm 30 minutos) y los fagómidos se recolectaron del sobrenadante.
9. Se añadieron 150 ml de NaCl 2,5 M PEG 8000 al 40% al sobrenadante (esto precipita los fagómidos). La torta se recuperó por centrifugación (4000 rpm 15 minutos), se resuspendió en 50 ml de PBS y se centrifugó (9000 rpm 15 minutos) para eliminar las bacterias y fagómidos agregados.
10. La etapa 9 se repitió mediante la adición de 10 ml de NaCl 2,5 M PEG 8000 al 40% a 50 ml de la preparación de fagómido. El precipitado final en PEG se resuspendió en 20 ml y los fagómidos se purificaron posteriormente por centrifugación por densidad en cloruro césico tal como se describe en (Smith, y Scott (1993) *Meth. Enzymol.*, 217: 228-157) y se resuspendieron en 20 ml. Esto constituyó la biblioteca final de fagómidos de anticuerpo que se utilizó para las selecciones.

### C) Evaluación de la diversidad en células individuales

Los experimentos descritos en los ejemplos 1 y 2 anteriores, indicaron que, como mínimo, cinco fagómidos diferentes podían entrar en una única bacteria, y que dos fagómidos que entraban en una única bacteria eran capaces de recombinarse uno con otro hasta el equilibrio, dando como resultado cuatro fagómidos diferentes. Sin embargo, estos no dieron una indicación de cuanto tiempo estos fagómidos sobrevivirían después del crecimiento en un cultivo líquido o sólido. Para clarificar este punto, y caracterizar la diversidad potencial de la biblioteca de genes producida, se produjeron fagómidos a partir de una única colonia de bacterias cre que contienen fagómidos que han experimentado recombinación. Estos fagómidos se utilizaron para obtener colonias bacterianas únicas conteniendo cada una solamente un tipo de fagómido. El análisis de estas colonias diferentes por PCR permitió la identificación de las diferentes cadenas V<sub>H</sub> y V<sub>L</sub> presentes en la colonia cre original (ver la figura 6 para detalles).

Después del crecimiento durante, como mínimo, 12 horas en bacterias que expresan cre, el cultivo se extrajo de las placas para aislar las colonias individuales. Si la recombinación ha sido exitosa, y están presentes aún todos los fagómidos, cada una de estas colonias debería contener muchos fagómidos con múltiples genes V recombinados. Con el objeto de identificarlos, en primer lugar, se prepararon fagómidos a partir de una única colonia (estos deberían comprender todas las combinaciones de scFv diferentes que se presentan dentro de la célula original de partida), y a continuación, estos se utilizaron para preparar colonias que contienen fagómidos únicos. De este modo, es posible aislar la mayoría de los genes scFv únicos encontrados en la colonia de partida.

Brevemente, esto se llevó a cabo de la manera siguiente:

1. Después de infección a una alta MOI, y el crecimiento durante toda una noche, bacterias que expresan cre se sacaron de las placas sobre 2XTY con 100 µg/ml de kanamicina /100 µg/ml de ampicilina/glucosa al 1% para aislar colonias individuales.
2. Se prepararon los fagómidos a partir de estas colonias individuales mediante el crecimiento de estas colonias a un OD550 0,5 en 1 ml de 2XTY con 100 µg/ml de kanamicina /100 µg/ml de ampicilina/glucosa al 1% a 37°C, infección con M13K07 y utilizando las técnicas descritas anteriormente para producir fagómidos.

3. Para derivar colonias a partir de estos fagómidos, se hicieron crecer bacterias DH5 $\alpha$ F a un OD550 0,5 en 2XTY a 37°C y se infectaron con el fagómido obtenido en la etapa 2 a una MOI menor de 1. Estas colonias contienen fagómidos únicos, y la población completa de las colonias representa la diversidad de los fagómidos contenidos dentro de las bacterias únicas de partida en crecimiento en la etapa 2.

4. Las cadenas V<sub>H</sub> y V<sub>L</sub> individuales presentes en cada fagómido se identificaron mediante amplificación por PCR y huella dactilar genética con BstNI o por secuenciación.

Se extrajeron dos células para análisis y los resultados para ambas se muestran en la figura 7. Ambas células dieron resultados muy similares, con 17-18 genes V<sub>H</sub> y 12-14 genes V<sub>L</sub> identificados por célula.

Estos genes V diferentes están presentes en 28-29 combinaciones diferentes, con muchos de los genes V diferentes encontrados en más de una combinación de scFv. En el caso de la primera célula ilustrada, pudieron identificarse cuatro combinaciones de dos pares V<sub>H</sub>/V<sub>L</sub>, indicando que la recombinación ha sido extensiva. En ambas células, alrededor del 50% de los pares V<sub>H</sub>/V<sub>L</sub> identificados están presentes en copias únicas, sugiriendo que no se ha identificado, en la pequeña muestra analizada (37-41 scFv) toda la dotación de genes V<sub>H</sub> y V<sub>L</sub> o sus combinaciones, y que la diversidad verdadera es más elevada que la identificada.

El grado de diversidad creado en una única célula fue difícil de evaluar. Es muy probable que los 18 genes V<sub>H</sub> diferentes recuperados de una única colonia correspondan a 18 genes V<sub>L</sub> diferentes. La diversidad potencial identificada en esta pequeña muestra es 324 (18<sup>2</sup>). Es probable que no todos los genes V diferentes presentes fueran identificados, y es conocido que los genes V puede tener huellas dactilares genéticas idénticas pero secuencias diferentes, sugiriendo que la diversidad creada en una única célula puede exceder esto, dando un máximo estimado que se aproxima a los 500-700 números de copias de pUC basados en plásmidos (Sambrook y otros, 1989, arriba). El más bajo de los estimados es el que se observó sobre la base del análisis de huella dactilar genética: 28-29 combinaciones scFv diferentes por célula.

El grado de recombinación fue extensivo, con muchos grupos diferentes de tres de cuatro combinaciones identificadas. Una cadena ligera se encontró con nueve cadenas pesadas diferentes, y una cadena pesada con cinco cadenas ligeras diferentes. El hecho de que tantos plásmidos diferentes puedan coexistir de forma estable en una única célula es, de alguna manera, sorprendente, y está en desacuerdo con el dogma que afirma que solamente un plásmido por grupo de complementación es capaz de sobrevivir en una célula bacteriana. Sin embargo, probablemente una vez que un número de genes V extenso se han introducido dentro de una célula, la recombinación en curso mantendrá tanto la diversidad como prevendrá la presión selectiva de eliminar genes V individuales. Esto da un intervalo potencial de diversidad creada por célula única de 29-500 scFvs diferentes.

En la biblioteca que se crea en la presente invención, 10<sup>10</sup> (20 ml) bacterias que expresan la recombinasa cre se infectaron con 2 x 10<sup>12</sup> fagos (una MOI de 200). Sobre la base de la diversidad observada realmente (cada bacteria produce 29 fagos diferentes), el tamaño final de la biblioteca será de 2,9 x 10<sup>11</sup> (29 x 10<sup>10</sup>). Mientras que, si la totalidad de las 324 combinaciones de genes V posibles están presentes, la biblioteca será aproximadamente diez veces más elevada (3,2 x 10<sup>12</sup>). Sin embargo, el tamaño de la biblioteca que puede utilizarse prácticamente, está limitado por la etapa reinfección a baja MOI cuando el fenotipo y el genotipo están acoplados. Mediante la utilización de un litro para llevar a cabo esta etapa, el número de bacterias en un litro no puede exceder 5 x 10<sup>11</sup>, aunque utilizando volúmenes más grandes se pueden crear fácilmente bibliotecas más extensas.

#### D) Examen de la biblioteca de anticuerpos en fago extensa

La biblioteca se examinó por selección sobre varios antígenos diferentes (ver Tabla 3) y los anticuerpos se aislaron contra todos en dos ciclos. En cada caso, como mínimo, se obtuvieron 3 scFvs diferentes por antígeno con un promedio de 6 por antígeno.

# ES 2 268 907 T3

TABLA 3

*Antígenos para los que se han seleccionado anticuerpos*

|    |                                          |                                                  |                                  |               |
|----|------------------------------------------|--------------------------------------------------|----------------------------------|---------------|
| 5  | Antígeno                                 | % positivo después de<br>2-3 turnos de selección | Número de scFV<br>independientes | Afinidad (nM) |
|    | Humano                                   |                                                  |                                  |               |
| 10 | Albúmina de suero                        | 96                                               | 7                                |               |
|    | Ciclina D                                | 50                                               | 4                                |               |
| 15 | CdK2                                     | 96                                               | 9                                | 29, 59, 82    |
|    | Rad52                                    | 80                                               | 11                               | 14            |
|    | Endonucleasa Flap                        | 50                                               |                                  |               |
| 20 | Ku70/80                                  | 25                                               |                                  |               |
|    | cdc25A                                   | 25                                               | 3                                |               |
| 25 | cdc25c                                   | 13                                               |                                  |               |
|    | Dominio de unión de ADN PARP             | 38                                               |                                  |               |
|    | PARP 85 KDa                              | 68                                               |                                  |               |
| 30 | Rodococos                                |                                                  |                                  |               |
|    | Pigs10:ferrodoxina piruvato oxireductasa | 38                                               |                                  |               |
| 35 | Pigs 12B:fosfoglicerol deshidrogenasa    | 50                                               | 4                                |               |
|    | Miscelánea                               |                                                  |                                  |               |
| 40 | Bucle HIV1                               | 75                                               | 6                                |               |

La selección se llevó a cabo de la manera siguiente:

- Un inmutotubo (Nunc) se recubrió con 1 ml del antígeno de interés a 10  $\mu\text{g/ml}$  toda una noche a temperatura ambiente. Después del recubrimiento el tubo se lavó con PBS tres veces y se bloqueó llenando el inmutotubo con Marvel al 2%/PBS (MPBS). Simultáneamente, 1 ml de biblioteca (que contiene  $5 \times 10^{12}$ - $10^{13}$  fagómidos) se bloqueó mediante la adición de Marvel al 10%/PBS a una concentración final del 2%.
- Después del bloqueo, el inmutotubo se lavó 3 veces con PBS y se añadió la biblioteca y se incubó 1 hora en reposo y 1 hora en rotación lenta de extremo a extremo.
- Después de la selección, el tubo se lavó diez veces con PBS y diez veces con Tween 20 al 0,1% /PBS.
- Los fagómidos de unión se eluyeron con 1 ml de trietanolamina 100 mM durante 8 minutos a temperatura ambiente, y inmediatamente se transfirió a un nuevo tubo y se neutralizó con Tris 1 M pH 7-7,5.
- Los fagómidos eluidos se añadieron a DH5 $\alpha$ F a OD550 0,5 y se dejó durante 30 minutos a 37°C sin agitación.
- Las bacterias infectadas se colocaron sobre placas 2XTY / ampicilina 100  $\mu\text{g/ml}$  / glucosa al 1% y se hicieron crecer durante 12-18 horas a 30°C.
- Las bacterias se extrajeron en 10 ml 2XTY, se mezclaron profundamente y se diluyeron en 50 ml de 2XTY / ampicilina 100  $\mu\text{g/ml}$  / glucosa al 1% a un OD550 0,05, y los fagómidos se prepararon por crecimiento a 37°C a un OD550 0,5 infectando con M13K07 y siguiendo el protocolo descrito anteriormente, que incluye dos precipitaciones con PEG/NaCl dando como resultado un volumen final de 1 ml de fagómido.

8. Se utilizaron  $10^{12}$  de estos fagómidos como inicio para el siguiente ciclo de selección que se llevó a cabo exactamente tal como se ha descrito anteriormente, excepto que en la etapa 3, el lavado se llevó a cabo veinte veces con PBS y diez veces con Tween 20 al 0,1% /PBS seguido de un lavado largo de 30-60 minutos en PBS a temperatura ambiente.

9. El fago eluido se sacó de las placas a densidad baja para aislar colonias individuales, se prepararon fagómidos a partir de éstos y se evaluaron con ELISA, tal como se describe en (Marks y otros, 1991).

#### Ejemplo 4

##### *Recombinación aleatoria in vivo sin recombinasas específicas*

La lactamasa beta de tipo salvaje es incapaz de dar resistencia al antibiótico cefotaxima excepto a una concentración de solamente  $0,025 \mu\text{g/ml}$ , que no es mayor que la proporcionada por las bacterias que no poseen esta resistencia al gen. Se ha mostrado que la resistencia a la cefotaxima se puede producir por la mutación del gen beta lactamasa en varios residuos clave, haciéndose esto por recombinación *in vitro*, pero no por solamente PCR proclive a error (Stemmer 1994).

Con el objeto de demostrar la recombinación aleatoria *in vivo* sin recombinasas específicas (es decir, utilizando la maquinaria de recombinación celular endógena), el gen para la beta lactamasa (que codifica la resistencia a la ampicilina) se amplificó a partir del vector pBSL167 (Alexeyev y otros (1995) *Gene* 160: 63-67) utilizando condiciones de PCR proclive a error. Después de la transfección, se obtuvo una biblioteca primaria de  $3 \times 10^5$  de clones independientes. Se obtuvieron fagómidos a partir de éstos y se utilizaron para infectar bacterias a diferentes proporciones fagómido:bacteria. Estas bacterias se hicieron crecer, como mínimo, durante 12 horas y los fagómidos se prepararon a partir de éstas. Se encontró que aquellos fagómidos preparados a partir de bacterias previamente infectadas a proporciones altas de fagómido:bacteria (mayores que 10:1) fueron capaces de dar lugar a muchas colonias bacterianas con resistencia a  $0,25 \mu\text{g/ml}$  de cefotaxima, mientras que aquellos fagómidos preparados a partir de bacterias previamente infectadas a proporciones bajas de fagómido:bacteria (menores que 1: 1) fueron capaces de dar muchas menos colonias resistentes a  $0,25 \mu\text{g/ml}$  de cefotaxima. La recombinación solamente puede tener lugar en los casos en que más de un fagómido está presente dentro de una única bacteria, y la recombinación útil solamente puede tener lugar si, como mínimo, dos fagómidos con mutaciones complementarias (mutaciones que si se encuentran sobre el mismo gen de beta lactamasa, a pesar de que sean diferentes, son capaces de dar lugar a una mayor resistencia a la cefotaxima) están presentes en la misma bacteria. Como la diferencia entre aquellos fagómidos capaces de dar lugar a más colonias resistentes y aquellos capaces de dar lugar a pocas colonias resistentes es solamente la multiplicidad de infección, está claro que la recombinación ha tenido lugar en el primer caso, pero no en el último.

El protocolo seguido es tal como se describe a continuación:

1. Se amplificó el gen de la beta lactamasa bajo condiciones de PCR proclive a error con dos cebadores que hibridizan al policonector en el plásmido pBSL167.
2. La banda amplificada se cortó con XsaI y KpnI y se reclonó atrás dentro del pBSL167 cortado con los mismos enzimas para dar lugar a una biblioteca primaria de  $3 \times 10^5$  colonias sobre placas de cloranfenicol ( $34 \mu\text{g/ml}$ ). Se proporciona resistencia al cloranfenicol por el gen de la cloranfenicol acetil transferasa encontrado en el esqueleto del vector.
3. Las bacterias se extrajeron en 2XTY, se diluyeron a OD550 0,1 en 10 ml de 2XTY con cloranfenicol  $50 \mu\text{g/ml}$ , se hicieron crecer a OD550 0,5 a  $37^\circ\text{C}$  y se añadieron  $10^{11}$  fagos auxiliares. Después de 30 minutos de incubación a  $37^\circ\text{C}$ , se añadió Kanamicina a  $25 \mu\text{g/ml}$ .
4. Las bacterias se hicieron crecer, como mínimo, durante 12 horas a  $37^\circ\text{C}$ , y se prepararon fagómidos a partir del cultivo. El cultivo se centrifugó (3500 rpm 20 minutos), el sobrenadante se trató con 2,5 ml de NaCl 2,5 M y PEG 8000 al 40% (esto precipita el fagómido). La partícula se recuperó por centrifugación (3000 rpm 20 minutos), se resuspendió en 1 ml de PBS y se centrifugó (13000 rpm 10 minutos) para eliminar las bacterias y fagómidos agregados. Esto constituye la biblioteca de fagómidos primaria.
5. Varias partes alícuotas de 10 ml de bacteria DH5 $\alpha$ F' se hicieron crecer a un OD550 0,5 en 2XTY a  $37^\circ\text{C}$ . Se añadieron partes alícuotas de la biblioteca de fagómidos primaria a estas bacterias a diferentes proporciones fagómido:bacteria, que varían desde 0,1:1 hasta 100:1. Las bacterias se dejaron durante 30 minutos a  $37^\circ\text{C}$  para permitir que tuviera lugar la infección, se añadió cloranfenicol y las bacterias se hicieron crecer a  $37^\circ\text{C}$  durante como mínimo 12 horas.
6. Después del crecimiento, se prepararon fagómidos a partir del sobrenadante bacteriano tal como se describe en las etapas 3-4 anteriores. Éstos constituyen la biblioteca secundaria o recombinada.
7. Para evaluar la capacidad de la biblioteca secundaria o recombinada para conferir una mayor resistencia a la cefotaxima, se hicieron crecer bacterias Dh5 $\alpha$ F' a OD550 0,5 en 2XTY a  $37^\circ\text{C}$ , y se añadieron fagómidos de las diferentes bibliotecas secundarias (producidas a diferentes proporciones fagómido:bacteria) a 1 ml

## ES 2 268 907 T3

de DH5 $\alpha$ F' (es decir, a una proporción menor de fagómido:bacteria que 1 para asegurar que cada bacteria se infecta, en promedio, por solamente un fagómido). El cultivo se dejó durante 30 minutos a 37°C, y se colocó sobre placas de cefotaxima de diferentes concentraciones.

- 5        8. Las bibliotecas secundarias creadas a partir de bacterias infectadas a proporciones fagómido:bacteria mayores que 10:1 dieron lugar a 100-800 colonias sobre 0,25  $\mu$ g/ml de cefotaxima, mientras que aquellas bibliotecas creadas a partir de bacterias infectadas a proporciones menores de fagómido:bacteria que 1 no dieron ninguna colonia resistente.

10       Se entiende que los ejemplos y realizaciones descritos en la presente descripción son solamente para propósitos ilustrativos, y que varias modificaciones o cambios a la luz de los mismos queda sugeridos a personas con experiencia en la técnica.

# REIVINDICACIONES

1. Método para la preparación de una biblioteca de ácidos nucleicos, comprendiendo dicho método la introducción de, como mínimo, dos miembros de una población inicial de moléculas de ácidos nucleicos dentro de, como mínimo, una célula, comprendiendo dicha población de moléculas de ácidos nucleicos dos o más ácidos nucleicos individuales, cada uno de los cuales comprende una secuencia de ácidos nucleicos que es idéntica para cada molécula y que incluye el mismo origen de replicación; y una secuencia de ácidos nucleicos que varía entre los miembros de dicha población y que comprende un sustrato para la recombinación, dando como resultado dicha introducción la recombinación de dichos sustratos para la recombinación entre, como mínimo, dos miembros de la población, para producir de este modo una población que comprende miembros de ácidos nucleicos recombinados.

2. Método, según la reivindicación 1, en el que dicha recombinación se lleva a cabo por un mecanismo de recombinación endógeno a dicha célula.

3. Método, según la reivindicación 1, en el que dicha recombinación está mediada por una recombinasa exógena.

4. Método, según la reivindicación 1, en el que dicha recombinación está mediada por una recombinasa endógena.

5. Método, según la reivindicación 1, en el que dicha recombinación se da en un lugar preseleccionado para la recombinación.

6. Método, según la reivindicación 5, en el que el lugar preseleccionado para la recombinación es un lugar de reconocimiento de recombinasa y la recombinación está mediada por una recombinasa expresada por dicha célula.

7. Método, según la reivindicación 1, en el que dicha recombinación está mediada por una recombinasa seleccionada del grupo que consiste en un miembro de la familia de recombinasas *hin*, un miembro de la familia de integrasas *lambda*, una recombinasa *flp*, una resolvasa, un transposón, y una recombinasa *Cre*.

8. Método, según la reivindicación 7, en el que dicha recombinasa se selecciona del grupo que consiste en *Cre*, *hin*, *gin*, *pin*, *cin*, y *flp*.

9. Método, según la reivindicación 5, en el que dicho lugar preseleccionado para la recombinación es un lugar *loxP*.

10. Método, según la reivindicación 1, en el que dicho sustrato para la recombinación comprende un primer lugar de reconocimiento de recombinasa y un segundo lugar de reconocimiento de recombinasa diferente del primer lugar de reconocimiento de recombinasa.

11. Método, según la reivindicación 10, en el que la recombinación resultada en el intercambio, entre dos miembros de dicha población de ácidos nucleicos, de los ácidos nucleicos que están bordeados por el primer y segundo lugares de reconocimiento de recombinasa.

12. Método, según la reivindicación 10, en el que el primer lugar de reconocimiento de recombinasa es un lugar *loxP* y el segundo lugar de reconocimiento de recombinasa es un lugar mutante de *loxP*.

13. Método, según la reivindicación 12, en el que el lugar mutante de *loxP* es *loxP 511*.

14. Método, según la reivindicación 1, en el que dicha célula se selecciona del grupo que consiste en una célula bacteriana, una célula de levadura, una célula de insecto, y una célula de mamífero.

15. Método, según la reivindicación 14, en el que dicha célula bacteriana es una célula de *Escherichia coli*.

16. Método, según la reivindicación 1, en el que dichos miembros de una población de moléculas de ácido nucleico se introducen dentro de la célula por transfección.

17. Método, según la reivindicación 1, en el que dicha población de moléculas de ácido nucleico comprenden, como mínimo, 10 miembros diferentes.

18. Método, según la reivindicación 1, en el que dichos miembros de una población de moléculas de ácido nucleico están contenidos dentro de partículas infecciosas y se introducen dentro de las células a través de infección con dichas partículas infecciosas.

19. Método, según la reivindicación 18, en el que dichas partículas infecciosas son fagos.

20. Método, según la reivindicación 19, en el que dichas partículas infecciosas son fagos filamentosos.

21. Método, según la reivindicación 20, en el que las partículas infecciosas son fagos filamentosos de la familia *Ff*.

22. Método, según la reivindicación 18, en el que dichas partículas infecciosas son fagómidos que contienen ADN fagomidoico.

23. Método, según la reivindicación 18, en el que dichas partículas infecciosas son fagómidos derivados de fagos filamentosos de la familia Ff.

24. Método, según la reivindicación 1, en el que dicho método comprende adicionalmente:

transfección o infección de una o más células con miembros de dicha población de miembros de ácidos nucleicos recombinados de modo que dichas células están infectadas a una multiplicidad de infección (moi) menor que, aproximadamente, 1.

25. Método, según la reivindicación 24, en el que dicho método comprende adicionalmente el empaquetamiento de los miembros de dicha biblioteca de ácidos nucleicos en paquetes de exposición genética replicables de modo que una proteína sobre la superficie del paquete de exposición replicable se codifica por un ácido nucleico empaquetado dentro del paquete de exposición, que es una secuencia de ácidos nucleicos que varía entre los miembros de la biblioteca de ácidos nucleicos.

26. Método, según la reivindicación 1, en el que la secuencia de ácidos nucleicos variable que comprende el sustrato para la recombinación comprende un casete de expresión.

27. Método, según la reivindicación 26, en el que dicho casete de expresión comprende secuencias de ácidos nucleicos que codifican uno o más polipéptidos.

28. Método, según la reivindicación 26, en el que dicho casete de expresión comprende secuencias de ácidos nucleicos que codifican uno o más polipéptidos y el ácido nucleico que codifica, como mínimo, uno de dichos polipéptidos tiene situado al lado un par de lugares de reconocimiento de recombinasa.

29. Método, según la reivindicación 27, en el que dichos polipéptidos se expresan sobre la superficie de un fago, un fagómido, o una bacteria.

30. Método, según la reivindicación 27, en el que dicha secuencia variable incluye un ácido nucleico que codifica una primera cadena de polipéptido y una segunda cadena de polipéptido a partir de un miembro par de unión específica de modo que la recombinación siguiente de dicha secuencia variable codifica proteínas de unión que no están presentes en la población inicial de ácidos nucleicos.

31. Método según la reivindicación 30, en el que dicho primer y dicho segundo polipéptido son polipéptidos de anticuerpo.

32. Método, según la reivindicación 31, en el que dicho primer y segundo polipéptido se seleccionan del grupo que comprende una región  $V_H$ , una región  $V_L$ , una  $V_H$  CDR<sub>1</sub>, una  $V_H$  CDR<sub>2</sub>, una  $V_H$  CDR<sub>3</sub>, una  $V_L$  CDR<sub>1</sub>, una  $V_L$  CDR<sub>2</sub>, una  $V_L$  CDR<sub>3</sub>, una  $V_H$  unida a un  $C_H1$ , y una  $V_L$  unida a un  $C_L$ .

33. Método, según la reivindicación 32, en el que el primer polipéptido es una región  $V_H$  y el segundo polipéptido es una región  $V_L$ .

34. Método, según la reivindicación 30, en el que un par de lugares de reconocimiento de recombinasa están bordeados por el ácido nucleico que codifica un primer polipéptido y dicho par de lugares de reconocimiento de recombinasa comprenden un primer lugar de reconocimiento de recombinasa y un segundo lugar de reconocimiento de recombinasa diferente.

35. Método, según la reivindicación 34, en el que el primer lugar de reconocimiento de recombinasa es un lugar LoxP y el segundo lugar de reconocimiento de recombinasa es un lugar LoxP 511.

36. Método, según la reivindicación 30, en el que el primer polipéptido está bordeado por un par de lugares de reconocimiento de recombinasa y los lugares de reconocimiento son diferentes el uno del otro.

37. Método, según la reivindicación 30, en el que el primer polipéptido y el segundo polipéptido está bordeado cada uno de ellos por un par de lugares de reconocimiento de recombinasa y los lugares de reconocimiento dentro de cada par son diferentes el uno del otro.

38. Método, según la reivindicación 35, en el que dichos lugares loxP se seleccionan del grupo que comprende loxP, loxP 511, y loxP fas.

39. Método, según la reivindicación 38, en el que los miembros de dicha biblioteca codifican un anticuerpo de cadena única.

40. Método, según la reivindicación 38, en el que dichos fragmentos de anticuerpo son scFv.

41. Método, según la reivindicación 38, en el que los miembros de dicha biblioteca codifican una fracción seleccionada del grupo que comprende un Fab, un Fv, un diacuerpo, un dímero V<sub>H</sub>, y un dímero V<sub>L</sub>.

42. Método, según la reivindicación 38, en el que los miembros de dicha biblioteca codifican un anticuerpo en el que las regiones V del anticuerpo están enlazadas por un polipéptido conector que comprende un lugar de reconocimiento de recombinasa.

43. Método, según la reivindicación 42, en el que dicho lugar de reconocimiento de recombinasa se selecciona del grupo que comprende un loxP, un mutante de loxP, un lugar de reconocimiento para una recombinasa de la familia hin, un lugar de reconocimiento para una integrasa lambda, un lugar de reconocimiento para una recombinasa flp, un lugar de reconocimiento para una resolvasa, y un lugar de reconocimiento para un transposón.

44. Método, según la reivindicación 1, en el que la parte variable del ácido nucleico comprende adicionalmente un marcador seleccionable, de manera que dicho marcador seleccionable debe recombinarse con un segundo marcador seleccionable para volverse activo.

45. Método, según la reivindicación 44, en el que dicho marcador seleccionable es inactivo sin recombinación debido a mutaciones.

46. Método, según la reivindicación 44, en el que dicho marcador seleccionable es inactivo sin recombinación debido a que es un marcador seleccionable incompleto.

47. Método, según la reivindicación 44, en el que dicho marcador seleccionable se localiza de modo que está unido al sustrato de recombinación y se cotransfiere con un gen de interés en dicho sustrato de recombinación.

48. Biblioteca de ácidos nucleicos que comprende una población de moléculas de ácidos nucleicos que comprende dos o más ácidos nucleicos individuales, cada uno de los cuales comprende una secuencia de ácidos nucleicos que es idéntica para cada molécula y que incluye el mismo origen de replicación y, como mínimo, dos lugares de reconocimiento de recombinasa; y una secuencia de ácidos nucleicos que varía entre los miembros de dicha población en la que dicha secuencia de ácidos nucleicos que varía comprende un sustrato para la recombinación, y en la que todos los miembros de dicha biblioteca tienen el mismo origen de replicación.

49. Biblioteca de ácidos nucleicos, según la reivindicación 48, en la que dicha biblioteca comprende, como mínimo, 10 miembros diferentes en una única célula.

50. Biblioteca de ácidos nucleicos, según la reivindicación 49, en la que dicha biblioteca comprende, como mínimo, 100 miembros diferentes en una única célula.

51. Biblioteca de ácidos nucleicos, según la reivindicación 48, en la que dichos lugares de reconocimiento de recombinasa comprenden lugares reconocidos por una recombinasa seleccionada del grupo que comprende un miembro de la familia de recombinasas hin, un miembro de la familia de integrasas lambda, una resolvasa, un transposón, y una recombinasa Cre.

52. Biblioteca de ácidos nucleicos, según la reivindicación 51, en la que dicha recombinasa se selecciona del grupo que comprende Cre, hin, gin, pin, cin, y flp.

53. Biblioteca de ácidos nucleicos, según la reivindicación 52, en la que dicho lugar de reconocimiento de recombinasa es un lugar LoxP o un lugar mutante de LoxP.

54. Biblioteca de ácidos nucleicos, según la reivindicación 53, en la que dichos miembros de una población de moléculas de ácido nucleico están contenidos dentro de partículas infecciosas.

55. Biblioteca de ácidos nucleicos, según la reivindicación 54, en la que dichas partículas infecciosas son fagos.

56. Biblioteca de ácidos nucleicos, según la reivindicación 55, en la que dichas partículas infecciosas son fagos filamentosos.

57. Biblioteca de ácidos nucleicos, según la reivindicación 56, en la que dichas partículas infecciosas son fagos filamentosos de la familia Ff.

58. Biblioteca de ácidos nucleicos, según la reivindicación 54, en la que dichas partículas infecciosas son fagómidos que contienen ADN fagómidoico.

59. Biblioteca de ácidos nucleicos, según la reivindicación 58, en la que dichas partículas infecciosas son fagómidos derivados de fago filamentosos de la familia Ff.

60. Biblioteca de ácidos nucleicos, según la reivindicación 48, en la que la secuencia variable de ácidos nucleicos que comprende el sustrato para la recombinación comprende un casete de expresión.

61. Biblioteca de ácidos nucleicos, según la reivindicación 60, en la que dicho casete de expresión comprende secuencias de ácidos nucleicos que codifican uno o más polipéptidos.

62. Biblioteca de ácidos nucleicos, según la reivindicación 61, en la que dichos polipéptidos se expresan sobre la superficie de un fago o fagómido.

63. Biblioteca de ácidos nucleicos, según la reivindicación 48, en la que dicha secuencia variable incluye ácidos nucleicos que codifican una primera cadena de polipéptido y una segunda cadena de polipéptido de un miembro par de unión específica de modo que después de la recombinación se producen ácidos nucleicos de unión que codifican proteínas de unión que no están presentes en la población inicial de ácidos nucleicos.

64. Biblioteca de ácidos nucleicos, según la reivindicación 63, en la que dicho primer y dicho segundo polipéptidos son polipéptidos de anticuerpo.

65. Biblioteca de ácidos nucleicos, según la reivindicación 64, en la que dicho primer y segundo polipéptidos se seleccionan del grupo que comprende una región  $V_H$ , una región  $V_L$ , una  $V_H$  CDR1, una  $V_H$  CDR2, una  $V_H$  CDR3, una  $V_L$  CDR1, una  $V_L$  CDR2, una  $V_L$  CDR3, una  $V_H$  unida a una  $C_H1$ , y una  $V_L$  unida a una  $C_L$ .

66. Biblioteca de ácidos nucleicos, según la reivindicación 65, en la que el primer polipéptido es una región  $V_H$  y el segundo polipéptido es una región  $V_L$ .

67. Biblioteca de ácidos nucleicos, según las reivindicaciones 63-66, en la que un par de lugares de reconocimiento de recombinasa bordean el ácido nucleico que codifica un primer polipéptido y el par de lugares de reconocimiento de recombinasa comprende un primer lugar de reconocimiento de recombinasa y un segundo lugar de reconocimiento de recombinasa diferente.

68. Biblioteca de ácidos nucleicos, según la reivindicación 67, en la que el primer lugar de reconocimiento de recombinasa es un lugar LoxP y el segundo lugar de reconocimiento de recombinasa es un lugar LoxP 511.

69. Biblioteca de ácidos nucleicos, según la reivindicación 48, en la que los miembros de dicha biblioteca codifican un anticuerpo de cadena única.

70. Biblioteca de ácidos nucleicos, según las reivindicaciones 63-66, en la que el primer y el segundo polipéptidos se expresan sobre la superficie de un fago o bacteria.

71. Biblioteca de ácidos nucleicos, según la reivindicación 70, en la que dichos polipéptidos son scFv, en los que  $V_H$  y  $V_L$  se unen por un conector de polipéptido codificado por un ácido nucleico que comprende un loxP, un mutante de loxP, un lugar de reconocimiento para una recombinasa de la familia hin, un lugar de reconocimiento para una integrasa lambda, un lugar de reconocimiento para una recombinasa flp, un lugar de reconocimiento para una resolvasa, y un lugar de reconocimiento para un transposón.

72. Biblioteca de ácidos nucleicos, según las reivindicaciones 48-66, en la que dichos ácidos nucleicos expresan polipéptidos rgdp.

73. Método para la preparación de un polipéptido, comprendiendo dicho método:

- a) disponer una biblioteca de ácidos nucleicos según las reivindicaciones 48-72;
- b) seleccionar uno o más miembros de dicha biblioteca; y
- c) expresar los ácidos nucleicos de uno o más miembros seleccionados.

74. Método, según la reivindicación 73, en el que dicha selección comprende:

- i) expresión de proteínas codificadas por los miembros de dicha biblioteca de ácidos nucleicos; y
- ii) monitorización de las proteínas expresadas para una o más propiedades seleccionadas del grupo que comprende unión específica a uno o más objetivos preseleccionados, una avidez de unión mínima a uno o más objetivos preseleccionados, una avidez de unión máxima a uno o más objetivos preseleccionados, termoestabilidad a una determinada temperatura preseleccionada, una actividad catalítica predefinida, una actividad enzimática predefinida bajo condiciones seleccionadas, una actividad biológica predefinida; y
- iii) selección de los miembros de la biblioteca que cumplen los criterios de monitorización.

75. Método, según la reivindicación 74, en el que dicha monitorización comprende la monitorización para la unión específica a un objetivo preseleccionado.

## ES 2 268 907 T3

76. Método, según la reivindicación 74, en el que las proteínas expresadas comprenden anticuerpos de cadena única.

5 77. Método, según la reivindicación 74, en el que los polipéptidos se expresan sobre la superficie de células infectadas o transfectadas con los miembros de la biblioteca de ácidos nucleicos.

78. Método, según la reivindicación 74, en el que los polipéptidos se expresan sobre la superficie de paquetes de exposición genética replicables (rgdp).

10 79. Método, según la reivindicación 74, en el que dichos miembros seleccionados en la etapa (iii) se utilizan para enriquecer o generar una biblioteca según el método de la reivindicación 1.

80. Método, según la reivindicación 78, en el que dicho paquete de exposición genética replicable (rgdp) es un fagómido que expresa uno o más polipéptidos unidos a proteínas de superficie.

15 81. Célula huésped, que comprende una biblioteca de ácidos nucleicos según las reivindicaciones 48-72.

82. Célula huésped, según la reivindicación 81, en la que dicha célula se selecciona del grupo que comprende una célula bacteriana, una célula vegetal, una célula de mamífero, una célula de insecto, y una célula de levadura.

20

25

30

35

40

45

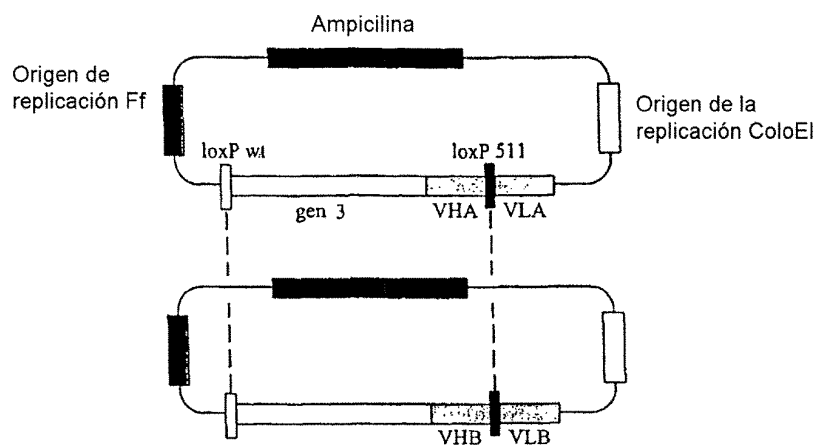
50

55

60

65

Antes de la recombinación



Después de la recombinación

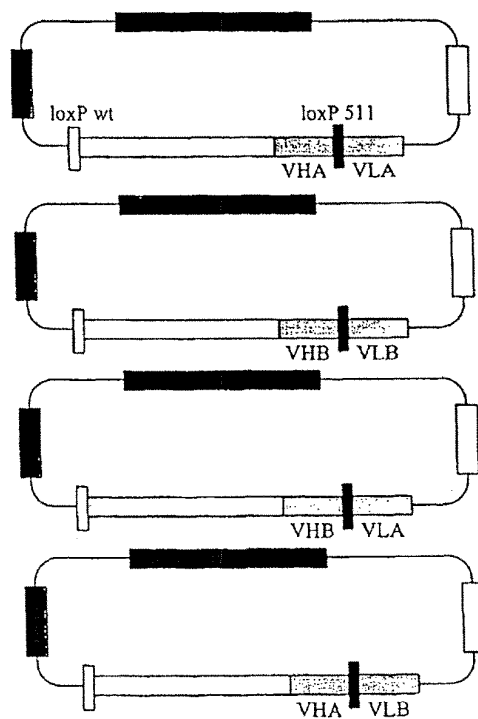


**Fig. 1** Todos los fagómidos después de la recombinación tienen anticuerpos funcionales

S G G S T I T S Y N V Y Y T K L S S S G T  
TCCGGAGGGTCGACC ATAACITCGTATAATGTATACTATACGAAGTTAT CCTCGAGCGGTACC  
BspEI Sali lox511 XhoI KpnI

Fig. 2

Lider bacteriano  
M K Y L L P T A A G L L L L A A S G A

H A  
AAGCTTGCCAAATCTTATTCAAGGACACAGTCATA  
ATGAAATACCTATTGCTACGGCAGCGCTGGATTGTATTACTCGCAGCAAGCGCGCGCATGCC  
HindIII Shine Delgano  
BssHII

... Cadena VL... TCCGGAGGGTCGACC ATAACTTCGTATAATGTATACTATACGAAGTTAT CCTCGAGCGGTACC..... Cadena VH....  
BspEI Sali Lugar loxp 511 XhoI KpnI

Etiqueta del epigono sv5  
A S G K P I P N P L L G L D S T H H H H H (Q)  
GCTAGC GGCAAACCAATCCCAACCCACTGCTGGGCTGGAT AGTACT CACCATCACCATCACCAT TAG  
NheI ScaI

Proteina del gen 3  
A A A T V E S C ..... R N K E s Parada  
GCGGCCGC TACTGTTGAAAGTTGT.....CGTAATAGGAGTCT TAA GCATGC ATAACTTCGTATAATGTATGCTATACGAAGTTAT  
GAATTC NotI  
EcoRI Lugar loxp

Fig. 3

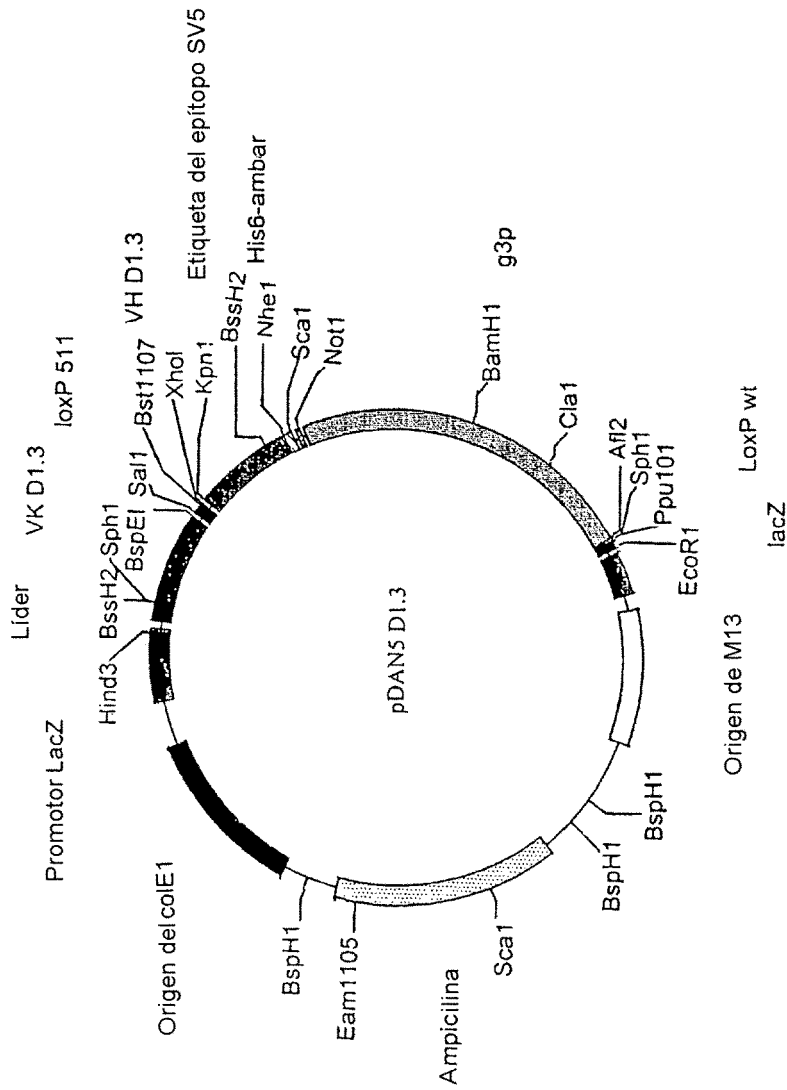


Fig. 4

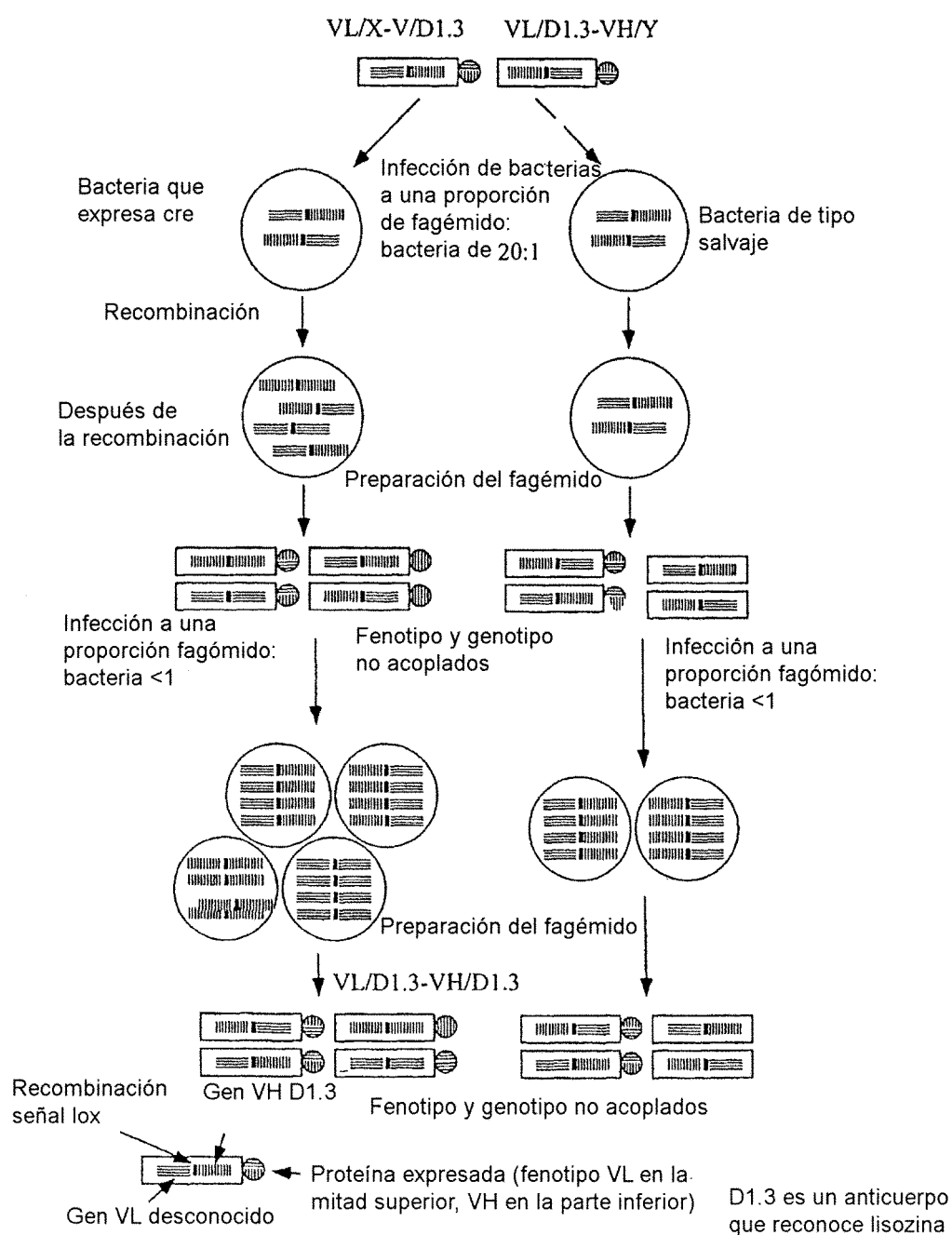


Fig.5

Biblioteca primaria ( $7 \times 10^7$ ) representada por 4 fagómeros diferentes

Biblioteca recombinada ( $\sim 10^{12}$ ) representada por 16 fagómeros diferentes

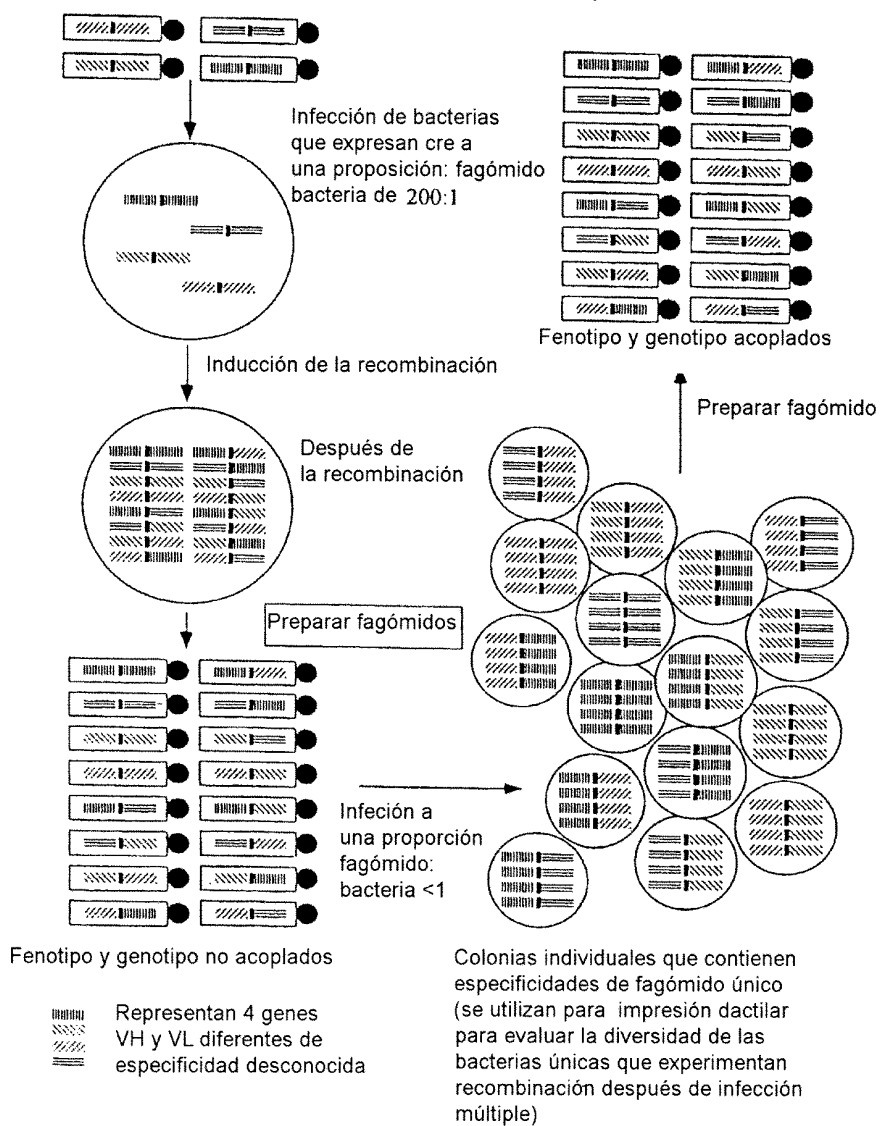


Fig. 6

|    |    | VH |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
|----|----|----|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|
|    |    | 1  | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| VL | 1  | 1  |   |   |   |   |   |   | 1 |   |    |    |    |    |    |    |    |    |
|    | 2  |    |   |   |   |   |   |   | 1 |   |    |    |    |    |    |    |    |    |
|    | 3  |    |   | 1 |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
|    | 4  |    |   |   | 1 |   |   |   |   |   |    | 1  |    |    |    |    |    |    |
|    | 5  |    |   |   |   | 1 |   |   |   |   |    |    |    |    |    |    |    |    |
|    | 6  |    |   |   |   |   |   |   |   |   |    | 1  |    |    |    |    |    |    |
|    | 7  |    |   |   |   |   |   |   |   |   |    |    | 1  |    |    |    |    |    |
|    | 8  |    |   |   |   |   |   |   |   |   |    |    |    |    | 1  |    |    |    |
|    | 9  | 2  |   | 2 |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
|    | 10 | 3  |   |   | 1 |   |   |   |   |   |    | 1  |    |    |    |    |    |    |
|    | 11 | 2  | 1 |   |   | 1 |   |   |   | 1 |    |    |    |    |    |    |    |    |
|    | 12 | 2  |   | 1 |   | 1 | 1 | 1 |   |   |    |    |    | 2  |    | 1  |    | 1  |

Fig. 7

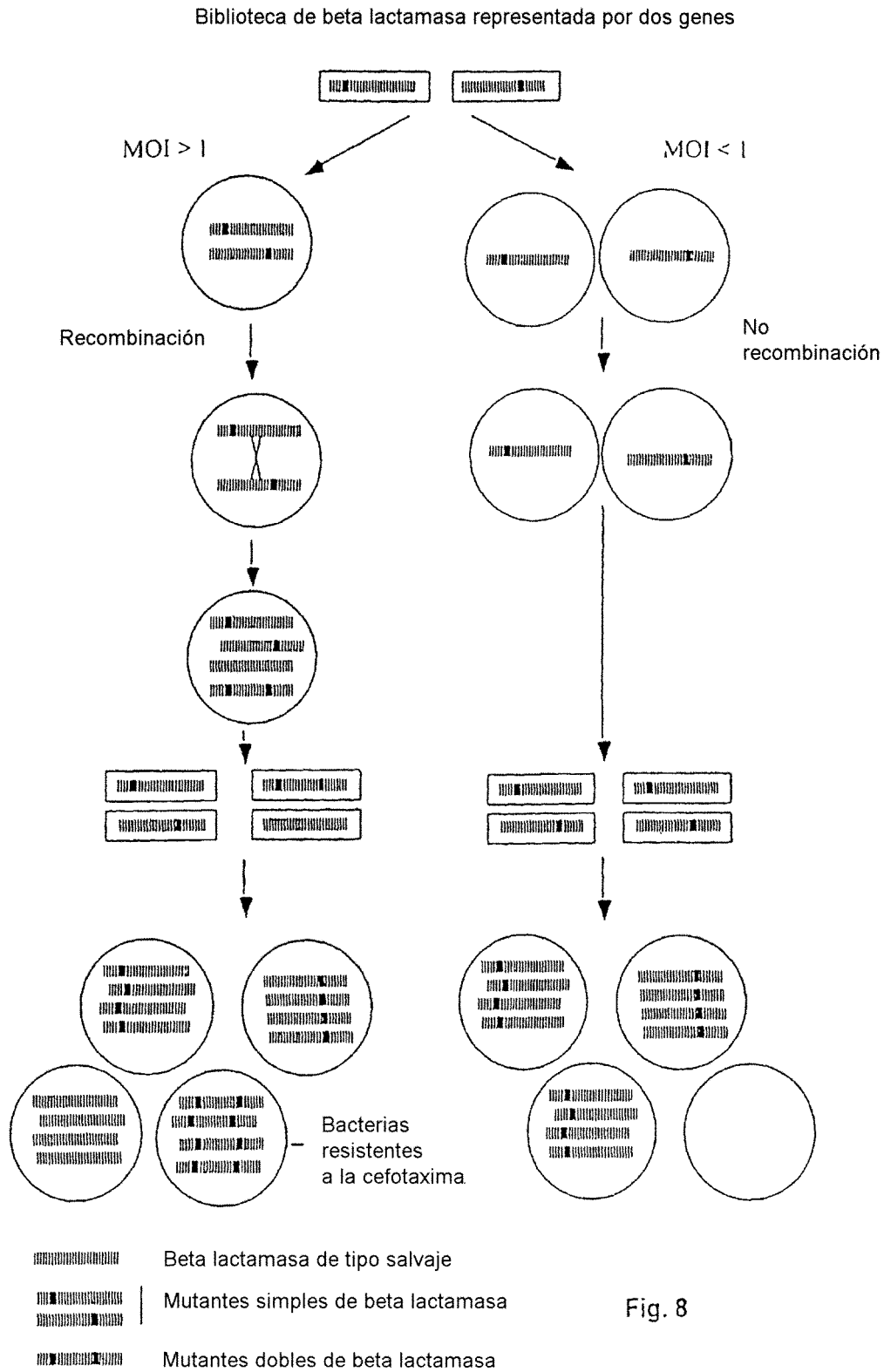


Fig. 8

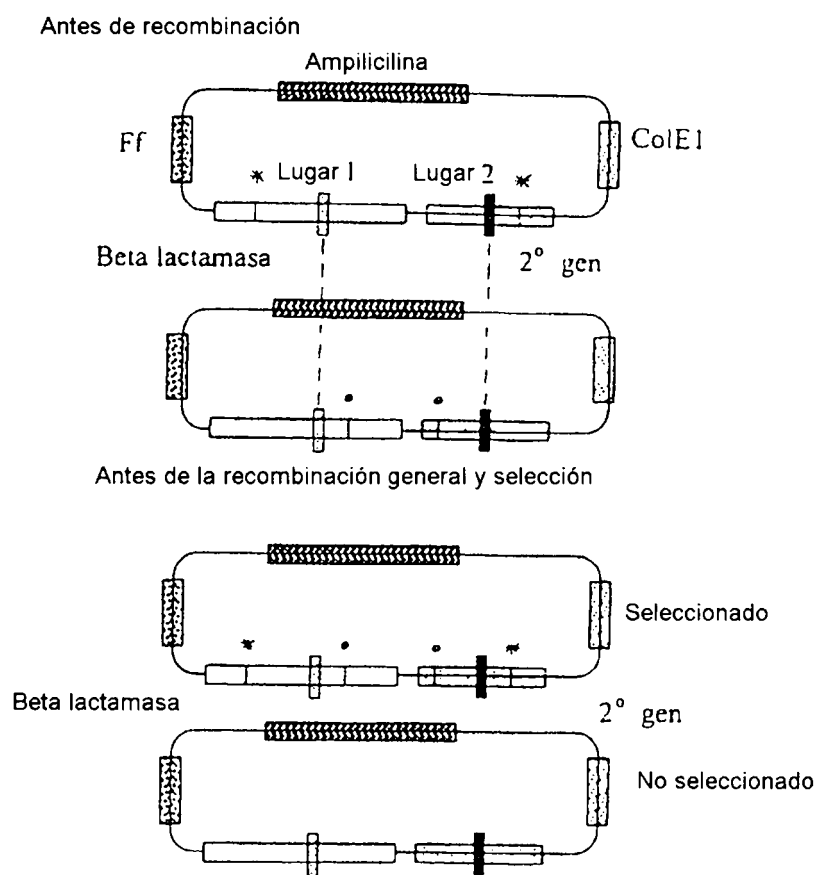


Fig. 9

## LISTA DE SECUENCIAS

## Identificación de Secuencia N° 1

5 ITSYNVYYTKL

## Identificación de Secuencia N° 2

10 GGGGS GGGGS GGGGS

## Identificación de Secuencia N° 3. pDAN5 D1.3

15 CTAGCGGCAG ACCAATCCCA AACCCACTGC TGGGCTTGGG TAGTACTCAC CATCACCATC ACCATTAGGC GGCGGCTACT  
 GTTGAAGTT GTTTAGCAAA ACCTCATACA GAAATTCAT TTAATAACGT CTGGAAGAC GACAAAACCT TAGATCGTTA  
 CGTAACTAT GAGGCTGTC TGTGGAATGC TACAGGCGTT GTGGTTTGTG CTGGTGACGA AACTCAGTGT TACGGTACAT  
 GGGTTCCTAT TGGGCTTGCT ATCCCTGAAA ATGAGGGTGG TGGCTCTGAG GGTGGCGGTT CTGAGGGTGG CGGTTCTGAG  
 GGTGGCGGTA CTAAACCTCC TGAGTACGGT GATACACCTA TTCCGGGCTA TACTTATATC AACCTCTCG ACGGCACTTA  
 TCCGCTTGGT ACTGAGCAAA ACCCGCTTAA TCCTAATCCT TCTCTGAGG AGTCTCAGCC TCTTAATATC TTCATGTTTC  
 AGAATAATAG GTTCCGAAAT AGGCAGGGTG CATTAACTGT TTATACGGGC ACTGTTACTC AAGGCACTGA CCCCGTTAAA  
 ACTTATTACC AGTACACTCC TGTATCATCA AAAGCCATGT ATGACGCTTA CTGGAACGGT AAATTCAGAG ACTGCGCTTT  
 CCATTCTGGC TTTAATGAGG ATCCATTCTGT TTGTGAATAT CAAGGCCAAT CGTCTGACCT GCCTCAACCT CCGTCAATG  
 20 CTGGCGGCGG CTCTGGTGGT GGTCTGGTG GCGGCTCTGA GGGTGGCGGC TCTGAGGGTG GCGGTTCTGA GGGTGGCGGC  
 TCTGAGGGTG GCGGTTCCGG TGGCGGCTCC GGTTCGGTG ATTTTGATTA TGAAAAATG GCAACGCTA ATAAGGGGGC  
 TATGACCGAA TGTGCGCCTT ATGTCTTTGG CGCTGGTAAA CCATATGAAT TTTCTATTGA TTGTGACAAA ATAAACTTAT  
 CTGCTATCGA TGGTTTCATT GGTGACGTTT CCGGCTTGG TAATGGTAAT GGTGCTACTG GTGATTTTGC TGGCTCTAAT  
 TCCCAATGG CTCAAGTCGG TGACGGTGAT AATTCACCTT TAATGAATAA TTTCCGTCAG TATTACCTT CTTTGCCTCA  
 25 GTCGGTTGAA TGTGCGCCTT ATGTCTTTGG CGCTGGTAAA CCATATGAAT TTTCTATTGA TTGTGACAAA ATAAACTTAT  
 TCCGTTGGTGT CTTTGGCTTT CTTTATATG TTGCCACCTT TATGTATGTA TTTTCGACGT TTGCTAACAT ACTGCGTAAT  
 AAGGAGTCTT AAGCATGCAT AACTTCGTAT AATGTATGCT ATACGAAGTT ATGAATTCAC TGGCCGTCGT TTTACAACGT  
 CGTGACTGGG AAAACCTGG CGTTACCCAA CTTAATCGCC TTGCAGCACA TCCCGCTTTC GCCAGCTGGC GTAATAGCGA  
 AGAGGCGCCG ACCGATCGCC CTTCCCAACA GTTGGCGAGC CTGAATGGCG AATGGCGCCT GATGCGGTAT TTTCTCCTTA  
 CGCATCTGTG CGGTATTTCA CACCGCATAC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG  
 30 GTGTGGTGGT TACGGCGAGC GTGACCGCTA CACTTGGCAG CGCCCTAGCG CCGGCTCCTT TCGCTTTCTT CCGTTCCTTT  
 CTCGCCACGT TCGCCGGCTT TCCCGCTCAA GCTCTAAATC GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA  
 CCTCGACCCC AAAAACTTG ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA  
 CGTTGGAGTG CACGTTCTTT AATAGTGGAC TCTTGTTCCT AACTGGAACA ACACCTCAACC CTATCTCGGG CTATTCTTTT  
 GATTTATAAG GGATTTTGGC GATTTGGGCC TATTGGTTAA AAAATGAGCT GATTAAACAA AAATTTAAGC CGAATTTTAA  
 CARAATATTA ACGTTTACAA TTTTATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATAGTTAAGC CAGCCCGGAC  
 35 ACCCGCCAAC ACCCGCTGAC GCGCCCTGAC GGGCTTGTCT GCTCCCGGCA TCCGCTTACA GACAAGCTGT GACCGTCTCC  
 GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA AACGCGCGAG ACGAAAGGGC CTCGTGATAC GCCTATTTTT  
 ATAGGTTAAT GTCATGATAA TAATGGTTTC TTAGACGCTA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCTTATT  
 GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT ATCCGCTTCA GAGACAATAA CCCTGATAAA TGCTTCAATA ATATTGAAAA  
 AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTGCGCCTTA TTCCCTTTTT TGCGGGCATTT TGCCCTCCTG TTTTGTCTCA  
 CCCAGAAACG CTGGTGAAAG TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAAGCT GATCTCAACA  
 GCGGTAAGAT CTTGAGAGT TTTCCGCCG AAGAAGCTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT ATGTGGCGCG  
 40 GTATTATCCC GTATTGACCG CGGGCAAGAG CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC  
 ACCAGTACA GAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA  
 CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GAGGCTAACC GCTTTTTTGC ACAACATGGG GATCATGTA  
 ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCTA TACCAACAGA CGAGCGGTGAC ACCACGATGC CTGTAGCAAT  
 GGCAACAACG TTGGCGAAAC TATTAATCG CGAAGTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG  
 CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG  
 45 CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG  
 TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTGAGACC  
 AAGTTTACTC ATATATACTT TAGATTGATT TAAAACCTTCA TTTTAAATTT AAAAGGATCT AGGTGAAGAT CTTTGTGAT  
 AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC  
 TTGAGATCCT TTTTCTGCG GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG  
 ATCAAGAGCT ACCAATCTTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATAACCA ATACTGTCCCT TCTAGTGTAG  
 CCGTAGTTAG GCCACCACTT CAAGAATCT GTAGCACCAG CTACATACCT CGCTCTGCTA ATCTGTGTAC CAGTGGCTGC  
 50 TGCCAGTGGC GATAAGTCGT GTCTTACCG GTTGGACTCA AGACGATAGT TACCGGATAA GCGCAGCGG TCGGGCTGAA  
 CGGGGGGTTG GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA GCATTGAGAA  
 AGCGCCACCG TTCCCGAAGG GAGAAAGCGG GACAGGTATC CCGTAAGCGG CAGGGTCCGA ACAGGAGAGC GCACGAGGGA  
 GCTTCCAGGG GGAACCGCCT GGTATCTTTA TAGTCTGTCT GGGTTTCGCC ACCTCTGACT TGAGCGTCTG TTTTGTGAT  
 GCTCGTCAGG GGGGCGGAGC CTATGAAAAA ACGCCAGCAA CGCGGCTTTT TTACGGTTTC TGGCCTTTTG TGGCCTTTT  
 GCTCACATGT TCTTCTCTG GTTATCCCTT GATTCTGTGG ATAACCGTAT TACCGGCTTT GAGTGAGCTG ATACCGCTCG  
 55 CCGAGCCGA ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAACCG CCTCTCCCGG

CGCGTTGGCC GATTTCATTA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT  
 GTGAGTTAGC TCACCTATTA GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG  
 60 ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCTTG CCAAAATCTA TTTCAAGGAG ACAGTCATAA  
 TGAATACCTT ATGCGCTACG GCAGCGCGTG GATTGTTATT ACTCGCAGCA AGCGGCGCGC ATGCGGACAT TCAGATGACC  
 CAGTCTCCAG CTTCCCTTTT TCGCTCTGTG GGAGAACTG TCACCATCAC ATGTCGAGCA AGTGGGAATA TTCACATTTA  
 TTTAGCATGG TATCAGCAGA AACAGGAAA ATCTCTCTAG TCCTCTGGTC ATTATACAAC AACCTTAGCA GATGGTGTG  
 CATCAAGGTT CAGTGGCAGT GGATCAGGAA CACAATATTC TCTCAAGATC AACAGCCTGC AACCTGAAGA TTTTGGGAGT  
 TATTACTGTC AACATTTTTG GAGTACTCCT CGGACGTTCT GTGGAGGGAC CAAGCTGGAG CTGAACCGTT CCGGAGGGTC  
 GACCATTAAT TCGTATAATG TATACTATAC GAAGTTATCC TCGAGCGGTA CCGAGGTGAA GCTGGTGGAG TCAGGACCTG  
 65 GCCTGTGGG GCGCTCACAG AGCCTGTCCA TCACATGCAC CGTCTCAGGG TTCTCATTA CCGGCTATGG TGTAAACTGG  
 GTTCGCCAGC CTCACGAAA GGGTCTGGAG TGGCTGGGAA TGATTGGGG TGATGAAAAC ACAGACTATA ATTACGCTCT  
 CAAATCCAGA CTGAGCATCA GCAAGGACAA CTCCAAGAGC CAAGTTTCT TAAAAATGAA CAGTCTGAC ACTGATGACA  
 CAGCCGTCTA CTACTGCGG CGAGAGAGAG ATTATAGGCT TGACTACTGG GGCAAGGCA CCACGGTCAC CGTCTCTCA

## ES 2 268 907 T3

Identificación de Secuencia N° 4.His6

HHHHHH

5 Identificación de Secuencia N° 5. VHback-DAN

TTA TCC TCG AGC GGT ACC SAG GTS MAR CTG CAG SAG TCW GG

10 Identificación de Secuencia N° 6. VHfor2-DAN

GAT TGG TTT GCC GCT AGC TGA GGA GAC GGT GAC CGT GGT

15 Identificación de Secuencia N° 7. VK2back-DAN

AGC AAG CGG CGC GCA TGC CGA CAT CGA GCT CAC CCA GTC TC

Identificación de Secuencia N° 8. VK2for-DAN

20 GAA GTT ATG GTC GAC CCT CCG GAA CGT TTG ATC TCG AGC TTG  
GTC CC

25 Identificación de Secuencia N° 9. VLbackPT1

CGC TGG ATT GTT ATT ACT CGC AGC AAG CGG CGC GCA TGC C

Identificación de Secuencia N° 10. VLbackPT2

30 TAC CTA TTG CCT ACG GCA GCC GCT GGA TTG TTA TTA CTC

Identificación de Secuencia N° 11. VHforPT1

35 CCA GGC CCA GCA GTG GGT TTG GGA TTG GTT TGC CGC TA

Identificación de Secuencia N° 12. VHforPT1

40 TGG TGA TGG TGA GTA CTA TCC AGG CCC AGC AGT GGG TTT G

Identificación de Secuencia N° 13. VLforPTL

45 ACC GCT CGA GGA TAA CTT CGT ATA GTA TAC ATT ATA CGA AGT  
TAT GGT CGA CCC TCC

Identificación de Secuencia N° 14. VHbackPTL

50 GGA GGG TCG ACC ATA ACT TCG TAT AAT GTA TAC TAT ACG AAG  
TTA TCC TCG AGC GGT A

55

60

65