

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 712 632**

51 Int. Cl.:

C12N 15/10 (2006.01)

C12N 15/90 (2006.01)

C12N 9/22 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.05.2015 PCT/GB2015/051287**

87 Fecha y número de publicación internacional: **05.11.2015 WO15166272**

96 Fecha de presentación y número de la solicitud europea: **01.05.2015 E 15723283 (6)**

97 Fecha y número de publicación de la concesión europea: **28.11.2018 EP 3137602**

54 Título: **Preparación de bibliotecas de variantes de proteínas expresadas en células eucariotas y su uso para seleccionar moléculas de unión**

30 Prioridad:

02.05.2014 GB 201407852

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

14.05.2019

73 Titular/es:

**IONTAS LIMITED (100.0%)
82b High Street, Sawston
Cambridge CB22 3HJ, GB**

72 Inventor/es:

**MCCAFFERTY, JOHN;
DYSON, MICHAEL y
PARTHIBAN, KOTHAI**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 712 632 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Preparación de bibliotecas de variantes de proteínas expresadas en células eucariotas y su uso para seleccionar moléculas de unión

5

Campo de la invención

Esta invención se refiere a métodos para producir bibliotecas de células eucariotas (por ejemplo, de mamíferos) para el cribado y/o la selección de moléculas de unión tales como anticuerpos. Las bibliotecas se pueden usar para contener y mostrar un repertorio diverso de aglutinantes, permitiendo que los aglutinantes se criben para seleccionar uno o más aglutinantes que tengan una propiedad deseada, tal como la especificidad para una molécula diana. La invención se refiere especialmente a métodos para introducir ADN donante que codifica los aglutinantes en células eucariotas para proporcionar una biblioteca de células en la que un número deseado de moléculas de ADN donante se integran fielmente en un locus o loci deseado en las células.

15

Introducción

Las técnicas de ingeniería de proteínas permiten la creación de grandes poblaciones diversas de moléculas relacionadas (por ejemplo, anticuerpos, proteínas, péptidos) a partir de las cuales se pueden aislar variantes individuales con propiedades de unión o catalíticas nuevas o mejoradas. La capacidad de construir grandes poblaciones de células eucariotas, particularmente células de mamíferos, donde cada célula expresa un anticuerpo individual, un péptido o una proteína modificada, tendrá un gran valor en la identificación de aglutinantes con propiedades deseadas.

20

El principio básico de la tecnología de presentación se basa en la unión de una molécula de unión a la información genética que codifica esa molécula. Las propiedades de unión de la molécula de unión se utilizan para aislar el gen que la codifica. Estos mismos principios subyacentes se aplican a todas las formas de tecnología de presentación, incluida la presentación de bacteriófagos, la presentación de bacterias, la presentación retroviral, la presentación de baculovirus, la presentación de ribosomas, la presentación de levaduras y la presentación en eucariotas superiores tales como las células de mamífero [1, 2, 3, 4].

25

30

La tecnología de presentación se ha ilustrado mejor mediante la presentación de anticuerpos en bacteriófagos filamentosos (presentación de fagos de anticuerpos) que durante los últimos 24 años ha proporcionado herramientas importantes para el descubrimiento y la creación de nuevas moléculas de unión, incluida la generación de anticuerpos terapéuticos humanos. El uso de moléculas de anticuerpo de presentación en fagos se presenta en la superficie de las partículas de bacteriófagos filamentosos mediante la clonación del gen que codifica un anticuerpo o fragmento de anticuerpo en marco con el gen que codifica una proteína de recubrimiento de fago. Los genes de anticuerpos se clonan inicialmente en *E. coli*, de tal forma que cada bacteria codifica un único anticuerpo. La generación de bacteriófagos a partir de bacterias usando métodos convencionales da como resultado la generación de partículas de bacteriófagos que presentan un fragmento de anticuerpo en su superficie y encapsulan el gen del anticuerpo codificante dentro del bacteriófago. La colección de bacterias o el bacteriófago derivado de ellas se conocen como una "biblioteca de anticuerpos". Usando la presentación en fagos de anticuerpos, los anticuerpos y sus genes asociados pueden enriquecerse dentro de la población exponiendo el bacteriófago presentador de anticuerpos a una molécula diana de interés.

35

40

45

Para permitir la recuperación de un bacteriófago que presenta un aglutinante que reconoce una de interés, la molécula diana debe inmovilizarse sobre la superficie de un recipiente de selección o debe ser recuperable de la solución mediante reactivos secundarios, por ejemplo, una proteína diana biotinilada, recuperada de la solución usando perlas recubiertas con estreptavidina. Después de la incubación de la biblioteca de bacteriófagos que presentan aglutinante con la molécula diana, se eliminan los fagos no unidos. Esto implica lavar la matriz a la que se une la diana (y el bacteriófago asociado) para eliminar el bacteriófago no unido. El bacteriófago unido con su gen de anticuerpo asociado puede recuperarse y/o infectarse en células bacterianas hospedadoras. Usando el enfoque descrito anteriormente, es posible enriquecer un subconjunto de clones de bacteriófagos capaces de unirse a una molécula diana de elección. Se ha demostrado que las bibliotecas de presentación de fagos proporcionan una rica fuente de diversidad de anticuerpos, proporcionando cientos de anticuerpos únicos para una sola diana [5,6,7].

50

55

Históricamente, los sistemas de presentación para aislar nuevas especificidades de unión de anticuerpos se han basado en sistemas procariotas y, en particular, en la presentación de Fv de una sola cadena (scFv) y, en menor medida, como Fab en bacteriófagos. La presentación de los aglutinantes en la superficie de las bacterias se ha descrito, pero no se ha utilizado ampliamente, y las aplicaciones se han limitado en gran medida a la presentación de péptidos o la presentación de fragmentos de anticuerpos pre-enriquecidos para aglutinantes a través de la inmunización [8]. A pesar del poder de los sistemas de presentación procariotas, incluida la presentación en fagos, existen limitaciones. Tras la selección por presentación en fagos o ribosomas, los genes que codifican las moléculas de unión individuales se identifican introduciendo la población de genes seleccionada en bacterias, colocando en placas las poblaciones bacterianas, recogiendo colonias, expresando las moléculas de unión en el sobrenadante o periplasma, e identificando los clones positivos en ensayos de unión tales como ensayos de inmunoabsorción ligada

60

65

a enzimas o fluorescencia (ELISA). Aunque las moléculas de unión están identificadas, este enfoque no resuelve la información sobre el grado de expresión y la afinidad de unión de los clones resultantes. Por lo tanto, aunque es posible generar miles de aglutinantes, la capacidad de cribar el resultado está limitada por la necesidad de recogida de colonias, manipulación de líquidos, etc., junto con información primaria limitada sobre el nivel de expresión relativa y la afinidad.

La presentación de las moléculas de unión en la superficie de las células eucariotas tiene el potencial de superar algunos de estos problemas. Junto con la citometría de flujo, la presentación eucariótica permite una selección rápida y de alto rendimiento. Es posible estudiar millones de clones celulares que expresan diferentes moléculas de unión en su superficie. La presentación de la superficie celular se ha ilustrado mejor para la presentación de fragmentos de anticuerpos formateados como scFv en la superficie de las células de levadura. Una modalidad comúnmente utilizada para la presentación de la superficie de la levadura aprovecha las proteínas de aglutinina de levadura (Aga1p y Aga2p). Según lo descrito por Chao *et al.* [9], los genes que codifican un repertorio de scFv se fusionan genéticamente con la subunidad Aga2p de aglutinina de levadura. La subunidad Aga2p se une entonces a la subunidad Aga1p presente en la pared celular a través de enlaces disulfuro. Las células de levadura que expresan una molécula de unión de diana específica pueden identificarse mediante citometría de flujo utilizando una molécula diana etiquetada directa o indirectamente. Por ejemplo, la diana biotinilada se puede añadir a las células y la unión a la superficie celular se puede detectar con estreptavidina-ficoeritrina. Dentro de una población, es posible, utilizando concentraciones de diana limitadas, distinguir aquellos clones que expresan moléculas de unión de mayor afinidad, ya que estos clones capturarán más moléculas diana y, por lo tanto, mostrarán una fluorescencia más brillante. Normalmente, cada célula de levadura mostrará de 10.000 a 100.000 copias de un solo scFv en la superficie de la célula. Para controlar la variación en la expresión de la superficie del scFv en diferentes células, Chao *et al.* utilizaron un anticuerpo anti-etiqueta marcado con fluorescencia para medir el nivel de expresión del anticuerpo en la superficie de cada célula, lo que permite la normalización de la variación en el nivel de expresión. Por lo tanto, este enfoque permite que las células de levadura que muestran moléculas de unión de alta afinidad se diferencien de aquellas células que expresan niveles altos de un anticuerpo de afinidad más baja. Por lo tanto, utilizando la clasificación de células activada por fluorescencia (FACS) es posible separar los clones celulares según la afinidad y/o el nivel de expresión de la molécula de unión codificada.

Los sistemas eucariotas también han demostrado ser más eficaces que los sistemas procariotas para la presentación de fragmentos de anticuerpos de múltiples cadenas y, en particular, con fragmentos más grandes, tales como IgG completas, FAb o fusiones de scFv con dominios Fc (fusiones scFv-Fc). Los métodos basados en perlas o basados en clasificación de flujo como se ha descrito anteriormente para las células de levadura también podrían usarse para seleccionar anticuerpos de bibliotecas de presentación basadas en eucariotas superiores, tales como las células de mamíferos. La capacidad de formatear bibliotecas de presentación y seleccionarlas directamente como IgG, Fab o como fusiones scFv-Fc en células de mamíferos sería una ventaja adicional sobre la presentación de levadura. La maquinaria de glucosilación, expresión y secreción de células bacterianas y de levadura es diferente de los eucariotas superiores que dan lugar a anticuerpos con diferentes modificaciones postraduccionales que las producidas en células de mamífero. Dado que la fabricación de anticuerpos para la investigación, la aplicación diagnóstica y terapéutica se realiza normalmente en células de mamífero, la presentación en células de mamífero (u otras células eucariotas superiores, tales como las líneas celulares de invertebrados, aves o plantas) podría dar una mejor indicación de posibles problemas o beneficios para la fabricación aguas abajo, por ejemplo, identificando clones con propiedades de expresión óptimas. Además, los anticuerpos descubiertos en el contexto de la presentación en eucariotas superiores y particularmente células de mamífero podrían aplicarse directamente en los ensayos de indicadores basados en células sin una purificación extensa y sin el efecto de complicación de los contaminantes de las bacterias y las células de levadura. Además, las bibliotecas de aglutinantes podrían expresarse directamente en células indicadoras eucariotas tales como células de mamífero para identificar clones que afectan directamente al fenotipo celular.

A pesar de las ventajas anteriores prometidas por las bibliotecas de presentación eucariota, sigue habiendo problemas significativos con la creación de bibliotecas de aglutinantes en células eucariotas, especialmente células eucariotas superiores. La introducción de un repertorio de genes exógenos ("transgenes") para la expresión en eucariotas superiores es más difícil que en la levadura y las bacterias. Las células de los eucariotas superiores son más difíciles de manipular y escalar, y las eficiencias de transformación son menores. Los tamaños de biblioteca típicos alcanzados son mucho más pequeños. Además, el ADN introducido se integra aleatoriamente dentro del genoma, lo que lleva a varios efectos de posición. Además, el ADN donante introducido en células de mamífero mediante métodos estándar de transfección o electroporación se integra como una matriz lineal con un número de copias variable del transgén transfectado. La introducción del ADN que codifica un repertorio de genes de anticuerpos, por lo tanto, tiene el potencial de introducir múltiples genes de anticuerpos en cada célula, lo que da como resultado la expresión de múltiples anticuerpos distintos por célula. Además, la presencia de múltiples genes de anticuerpos reducirá la expresión relativa de cualquier anticuerpo dado y conducirá al aislamiento de muchos genes de anticuerpos del pasajero reduciendo la tasa de enriquecimiento de clones específicos.

Aunque la presentación de una biblioteca de aglutinantes en la superficie de los eucariotas superiores es más desafiante, se han descrito previamente algunos ejemplos. En una publicación temprana en la que se usaron presentación en mamíferos de IgG derivadas de la inmunización humana, se requirieron 3 rondas de selección (que

implicaban transfección transitoria, clasificación celular, recuperación de ADN y re-transfección) para lograr un enriquecimiento de 450 veces de las células específicas de antígeno, con un promedio de 7,6 veces de enriquecimiento por ronda [10]. De manera similar, la expresión transitoria de las bibliotecas inmunizadas expresadas en vectores de replicación episómica también se ha descrito con anticuerpos formateados como scFv [11, 12] o IgG [13].

Se han descrito varios enfoques para introducir un número único o un número limitado de genes de anticuerpos en cada célula. Esto incluye la dilución del ADN o la mezcla con el ADN portador [13], pero este es un método relativamente incontrolado para administrar el número de copias de los genes introducidos y reducir la entrada de ADN tendrá un efecto perjudicial sobre el tamaño de la biblioteca. La introducción de genes de anticuerpos por vectores virales ha proporcionado otra solución para controlar la introducción de múltiples genes de anticuerpos por célula. De este modo, se ha generado una biblioteca de presentación de la superficie celular a partir de varios cientos de linfocitos B humanos generados por inmunización y enriquecidos adicionalmente por clasificación de flujo de linfocitos B específicos de antígeno [14]. Los genes de anticuerpos de este grupo enriquecido se formatearon como scFv, se clonaron en un sistema de expresión de alfavirus Sindbis y se introdujeron en células BHK utilizando una baja multiplicidad de infección.

Breous-Nystrom *et al.* [15] usaron la infección retroviral secuencial para introducir un repertorio limitado de genes de anticuerpos kappa de 91 V seguidos de un repertorio de genes de cadena pesada de 6 donantes sanos en una línea de linfocitos pre-B murinos (1624-5). El retrovirus infeccioso se generó utilizando el sistema V-Pack basado en el virus de la leucemia murina de Moloney (Stratagene). Con el fin de desviarse hacia inserciones de copia única, se eligió una multiplicidad de infección que condujo a la infección de aproximadamente el 5 % de las células. Una desventaja importante de estos enfoques es que la integración dentro del genoma es aleatoria, lo que conduce a una variación potencial en el nivel de transcripción basado en la actividad transcripcional del sitio de integración. Otra desventaja en todos estos casos es que la integración de los genes de anticuerpos se controla mediante una infección o transfección limitadas que afectan al tamaño de la biblioteca.

La integración de sitio específico de transgenes dirigidos por recombinasas se ha descrito previamente. Las recombinasas son enzimas que catalizan reacciones de intercambio entre moléculas de ADN que contienen secuencias de reconocimiento de enzimas específicas. Por ejemplo, la recombinasa Cre (derivada del sistema de recombinación de sitio específico de *E. coli*) o la recombinasa Fip (utilizando un sistema de recombinación de *Saccharomyces cerevisiae*) actúa en sus sitios específicos de reconocimiento loxP de 34 pb y el sitio de diana de recombinación de Fip (FRT) de 34 pb respectivamente [16]. Las recombinasas se han utilizado principalmente en ingeniería celular para catalizar la integración de sitio específico. Varios estudios a partir del trabajo de Chen Zhou [17, 18, US7.884.054] han descrito la integración de sitio específico mediada por recombinasa de genes de anticuerpos en el genoma de células de mamífero utilizando la recombinasa Fip en el sistema "Fip-In", (http://tools.lifetechnologies.com/content/sfs/manuals/flpinsystem_man.pdf). El sistema Fip-In utiliza diversas líneas celulares que previamente tenían un solo sitio FRT introducido en su genoma. Al expresar la enzima Fip recombinasa es posible dirigir la integración de los plásmidos de expresión, incorporando un sitio de recombinación de FRT, en este sitio de FRT preintegrado en las células diana.

Usando el sistema Fip-In, Zhou *et al.* [17] introdujeron un plásmido de expresión de anticuerpo entrante que contiene un sitio de FRT en la línea celular de ovario de hámster chino (CHO) que incorpora un sitio de FRT (células CHOF). Su trabajo describe la construcción de una biblioteca de presentación en la que se mutagenizaron 4 restos dentro de un anticuerpo anti-OX40 existente. La biblioteca se seleccionó utilizando FACS para identificar anticuerpos con afinidad anti-ligando en la superficie celular. El éxito general en la generación de anticuerpos mejorados se limitó al aislamiento de un único anticuerpo mejorado. La cantidad de clones de células de mamífero únicos logrados no se informó.

En un documento de seguimiento, Li *et al.* en 2012 [18] usaron linfocitos de un paciente con hepatitis B para construir una biblioteca de presentación de anticuerpos. Se produjeron bibliotecas separadas con los genes de cadena pesada y ligera obtenidos de un donante que se había inmunizado con HBsAg, que se informó individualmente que eran bibliotecas de tamaño $1,02 \times 10^6$ y $1,78 \times 10^5$, respectivamente. Después, se produjo una biblioteca secundaria que incluía tanto las cadenas pesadas como las ligeras que, según se informa, tenían un tamaño de $4,32 \times 10^5$. El análisis FACS indicó según se informa que aproximadamente el 40 % de las células mostraban anticuerpos de longitud completa detectables sobre la superficie celular. El cribado de FACS de la biblioteca identificó anticuerpos que se unían a HBsAg. De una muestra de 8 miembros de la biblioteca seleccionados que se unieron al antígeno, se encontró que seis tenían el mismo anticuerpo, por lo que en total se identificaron tres clones únicos anti-HBsAg.

El éxito bastante limitado de este trabajo puede deberse al hecho de que el sistema Fip-In está diseñado para una integración precisa en un número limitado de clones en lugar de una construcción de biblioteca grande. Por lo tanto, existe un conflicto potencial entre lograr la fidelidad de la integración y lograr el tamaño máximo de la biblioteca. El sistema Fip-In utiliza una recombinasa Fip mutante en el plásmido pOG44 que posee solo el 10 % de la actividad a 37 °C de la recombinasa Fip nativa [19]. Se ha identificado una variante de la recombinasa Fip (Fipe) con mejor termoestabilidad y mayor actividad que la de tipo silvestre [19, 20]. Esto se mejoró aún más mediante la optimización

del codón para crear Flp_o codificado dentro del plásmido cCAGGS- Flp_o (Genebridges Cat. A203). Sin embargo, según el manual de Flp-In:

"Al generar líneas celulares de expresión Flp-In™, es importante recordar que está seleccionando un evento de recombinación relativamente raro, ya que se desea que la recombinación e integración de su construcción pcDNA™5/FRT tenga lugar solo a través del sitio de FRT y durante un tiempo limitado. En este caso, el uso de una recombinasa Flp altamente ineficiente es beneficioso y puede disminuir la aparición de otros eventos de recombinación no deseados...

...Para aumentar la probabilidad de obtener integrantes individuales, se deberá disminuir la eficacia de la transfección limitando la cantidad de ADN plasmídico que se transfecta"

Esto se repite en Buchholz *et al.*, 1996 [19]:

"FLP puede ser particularmente útil para aplicaciones que no dependen de la eficiencia, sino que dependen de una regulación estricta".

En experimentos de modelo y usando "instrucciones descritas en el manual", Zhou *et al.* (2010) [17] demostraron de hecho que tuvieron lugar inserciones de copias únicas en >90 % de los clones. Sin embargo, en la construcción de bibliotecas se utilizaron cantidades relativamente altas de plásmido de expresión (2,5-3,2 µg por 10⁶ células) y un exceso de donante sobre el plásmido que codifica la recombinasa pOG44 [17, 18]. El sistema Flp-In recomienda utilizar una relación de al menos 9:1 a favor del plásmido que codifica la recombinasa frente al plásmido de expresión. Sin embargo, cuando se busca aumentar el tamaño de la biblioteca transfecando grandes cantidades de ADN, existe la posibilidad de una integración aleatoria del plásmido entrante [21]. En todos los estudios, no se informó la precisión de la integración y el número de integrantes por célula en las condiciones de "construcción de la biblioteca".

En la integración de genes dirigida por nucleasas, se utiliza una nucleasa de sitio específico para escindir el ADN celular en una ubicación específica. Previamente se ha demostrado que esto aumenta la velocidad de recombinación homóloga en al menos 40.000 veces y también permite la reparación mediante mecanismos de unión de extremos no homólogos. Esta mejora de la integración de sitio específico no se ha usado o contemplado previamente para resolver los problemas asociados con la creación de bibliotecas de aglutinantes.

El documento de patente US20100212035 describe métodos para la generación de roedores capaces de expresar anticuerpos exógenos dirigiéndose al locus de inmunoglobulina de un embrión de mamífero con una meganucleasa para dirigir la integración de un ADN donante. Se describe el potencial para crear bibliotecas de variantes de meganucleasas para crear nuevas especificidades de escisión del ADN, pero no se contempla el uso de meganucleasas para la generación de bibliotecas de aglutinantes.

El documento de patente WO 2013/190032 A1 describe la integración de genes en un locus específico (Fer1L4) previamente modificado con ADN exógeno (una célula hospedadora de "integración de sitio específico" SSI) para incorporar sitios de recombinasa, tales como sitios de loxP y FRT para la introducción génica de sitio específico mediada por recombinasa. No se describe la generación de bibliotecas dirigidas por nucleasas.

El documento de patente WO 2012/167192 A2 describe el direccionamiento de genes a un locus que luego puede seleccionarse para la amplificación. Se emplean métodos dirigidos por nucleasas para dirigirse al locus. No se describe la generación de bibliotecas dirigidas por nucleasas.

El documento de patente US 2009/0263900A1 describe moléculas de ADN que comprenden brazos de homología y su uso en métodos de recombinación homóloga. No se describe la generación de bibliotecas dirigidas por nucleasas.

El documento de patente WO 2011/100058 describe métodos para la integración de ácido nucleico en un genoma que evita la necesidad de brazos de homología largos y, en cambio, se basa en la microhomología o "extremos adhesivos" en el genoma y el donante para ayudar a la integración directa. No se describe la generación de bibliotecas dirigidas por nucleasas.

El documento de patente WO 2011/090804 describe métodos para la integración de múltiples genes o múltiples copias del mismo gen utilizando diferentes nucleasas con dedos de cinc (ZFN) en rondas secuenciales. No se describe la generación de bibliotecas dirigidas por nucleasas.

El documento de patente WO2014/039872 describe métodos para modificar células vegetales, que incorporan un "sitio de aterrizaje" en el que el ADN donante se integra mediante recombinación homóloga o unión de extremos no homólogos utilizando nucleasas de sitio dirigido. Las bibliotecas de cromosomas artificiales bacterianos (BAC) se utilizan para la clonación inicial del ADN donante. Las bibliotecas se mencionan en relación con los métodos de secuenciación de Illumina. No se describe la generación de bibliotecas dirigidas por nucleasas.

El documento de patente WO2007/047859 A2 describe métodos para modificar la especificidad de las meganucleasas y su uso para dirigirse a loci genómicos. Se describen bibliotecas de meganucleasas mutantes que pueden contener meganucleasas con nueva especificidad de nucleasa. No se describe la generación de bibliotecas dirigidas por nucleasas.

5 El documento de patente US2014/0113375 A1 describe un sistema de expresión transitoria para la generación de secuencias de ADN monocatenarias homólogas a una secuencia genómica diana, que puede ser transportada al núcleo para alterar la información genética de la secuencia genómica diana a través de vías de reparación de ADN o recombinación homóloga. Se sugiere que se pueda crear una "biblioteca" de mutaciones mediante la transcripción inversa de baja fidelidad del ADN introducido (no de biblioteca). No se describe la presentación y selección de moléculas de mamífero con actividad de unión.

15 El documento de patente US2012/0277120 describe métodos y composiciones para la integración simultánea de una pluralidad de ácidos nucleicos exógenos en una sola reacción de transformación que utiliza la maquinaria de recombinación homóloga nativa en levaduras, cuya recombinación se puede mejorar adicionalmente mediante la inducción de roturas bicatenarias dirigidas en el genoma de la célula hospedadora en los sitios de integración previstos. Los métodos pretenden superar la necesidad de múltiples rondas de modificación para integrar múltiples ensamblajes de ADN, por ejemplo, para la construcción de vías metabólicas funcionales en microbios industriales, tal como la levadura. No se describe la presentación o expresión de bibliotecas de moléculas de unión, el uso de eucariotas superiores y la selección de moléculas con actividad de unión.

20 Moehle *et al.* [21] informan que una rotura bicatenaria colocada con precisión inducida por las nucleasas con dedos de cinc (ZFN) modificadas por ingeniería genética puede estimular la integración de largos segmentos de ADN en una ubicación genómica predeterminada, dando como resultado la adición génica de sitio dirigido de alta eficiencia. Los autores utilizaron un donante de ADN extracromosómico que llevaba una etiqueta de 12 pb, un ORF de 900 pb, o una unidad de transcripción de promotor de 1,5 kb flanqueada por brazos de homología de locus específico y notificaron frecuencias de integración dirigida del 15 %, 6 % y el 5 %, respectivamente, en las 72 h de tratamiento, y sin selección para el evento deseado. Los autores informan que el evento de integración se produce de forma dirigida por homología y conduce a la reconstrucción precisa del genotipo del donante específico en el locus cromosómico endógeno, y por lo tanto, es probable que sea resultado de la reparación de hibridación de la rotura de la cadena dependiente de la síntesis utilizando el ADN donante como plantilla. Esta adición génica de sitio específico se produce sin un aumento medible en la tasa de integración aleatoria. Los autores también informan que las ZFN pueden conducir la adición de una secuencia de 8 kb que lleva tres unidades de transcripción de promotor distintas a un locus endógeno a una frecuencia del 6 %, también en ausencia de cualquier selección. Proc Natl Acad Sci USA. (2007) Feb 27; 104(9):3055-60.

35 Cristea *et al.* (2013) [22] informan que la inclusión de un sitio diana de la nucleasa en el plásmido donante seguido de la escisión de la nucleasa *in vivo* tanto del donante como del cromosoma da como resultado una integración eficiente de moléculas grandes de ADN de tamaño transgénico en la rotura bicatenaria cromosómica. Los autores informan que las células CHO, un tipo de célula recalcitrante a la integración basada en homología, son competentes en la captura de donantes transgénicos *in vivo* linealizados. También informan de una desactivación del gen FUT8 de hámster a través de la integración simultánea mediada por la nucleasa ZFN o TALE de un casete de anticuerpos. Cristea *et al.* declaran que sus resultados permiten la adición eficiente de transgenes dirigida a células y organismos que se comportan mal con los enfoques tradicionales impulsados por la homología.

45 Para comprender plenamente el potencial de la presentación de anticuerpos en células de mamíferos y otros eucariotas superiores, existe la necesidad de un sistema para crear bibliotecas grandes que combinen una integración precisa en un sitio predefinido con una eficiencia que permita la construcción de bibliotecas grandes.

50 Sumario de la invención

La invención se expone en las reivindicaciones adjuntas.

55 Se ha superado el problema de crear grandes bibliotecas de aglutinantes que incluyan uno o dos genes de aglutinantes por célula mediante el uso de la integración dirigida por nucleasas de poblaciones de genes que codifican aglutinantes. Por lo tanto, la invención permite la preparación de poblaciones de células eucariotas en las que un repertorio de codificación de aglutinantes se integra en un locus fijo en el genoma permitiendo la expresión de la molécula de unión codificada, creando de este modo una población de células que expresan diferentes aglutinantes.

60 La presente invención se refiere a métodos para producir bibliotecas de células eucariotas que codifican un repertorio de moléculas de unión ("aglutinantes") que reconocen una diana, en los que los métodos usan una nucleasa de sitio específico para la escisión dirigida del ADN celular para mejorar la integración de sitio específico de genes de aglutinantes a través de mecanismos de reparación celular endógenos. Los aglutinantes son anticuerpos, proteínas o péptidos. Las nucleasas de sitio específico permiten la introducción precisa del ADN donante que codifica las moléculas de aglutinante en uno o más loci definidos dentro del genoma eucariota u otro

ADN celular eucariota. La invención proporciona métodos para preparar poblaciones de células eucariotas en las que un repertorio de genes que codifican aglutinantes se integran en un locus deseado en el ADN celular (por ejemplo, un locus genómico) que permite la expresión de la molécula de unión codificada, creando de este modo una población de células que expresan diferentes aglutinantes.

5 La construcción de bibliotecas de aglutinantes dentro de células eucariotas de acuerdo con la presente invención tiene ventajas sobre los enfoques dirigidos por recombinasa para la incorporación de sitio dirigido de construcciones de expresión. La presente invención usa la escisión del ADN celular por nucleasas de sitio específico para resolver problemas previamente asociados con la construcción de grandes repertorios de genes de aglutinantes en células eucariotas y particularmente eucariotas superiores. Esta invención permite la creación eficiente de grandes poblaciones de clones celulares, cada uno de los cuales expresa aglutinantes individuales integrados en un locus fijo en el ADN celular. A partir de estas bibliotecas de clones celulares, es posible aislar genes que codifican proteínas y péptidos novedosos de unión o modificadores de la función.

15 En lugar del intercambio de ADN dirigido por recombinasa, el enfoque de la presente invención utiliza la escisión de sitio específico del ADN celular (por ejemplo, genómico) seguido del uso de mecanismos de reparación naturales para integrar el ADN donante que codifica el aglutinante. Después de la escisión del ADN celular en una secuencia reconocida por la nucleasa de sitio específico ("secuencia de reconocimiento"), las roturas en el ADN celular se reparan utilizando mecanismos tales como la recombinación homóloga o la unión de extremos no homólogos (NHEJ). La creación de roturas de sitio específico en el ADN celular mejora la incorporación de ADN donante exógeno permitiendo la construcción de grandes poblaciones de células con genes de aglutinantes integrados en un locus fijo.

25 Hasta la fecha, las nucleasas de sitio específico tales como meganucleasas, ZFN, nucleasas TALE y sistemas CRISPR/Cas se han dirigido hacia la creación eficiente de células con modificaciones para genes endógenos o para la introducción de genes indicadores para el estudio de la función celular. También hay casos en los que la orientación genómica dirigida por nucleasas se ha utilizado para integrar genes que codifican anticuerpos de un solo secretado para la producción de anticuerpos (mediante purificación a partir de un medio de cultivo) [21, 22].

30 La invención simplifica la construcción de grandes bibliotecas al tiempo que dirige la integración a un número único o limitado de loci genéticos definidos. La integración del ADN donante en uno o más loci fijos normaliza la transcripción en comparación con la integración aleatoria de números variables de transgenes, y permite la selección de clones de anticuerpos sobre la base de las propiedades de traslación y estabilidad del propio aglutinante. La integración fiel del ADN donante en una ubicación o ubicaciones predeterminadas en el ADN celular da como resultado niveles relativamente uniformes de transcripción de aglutinantes en la biblioteca, y la alta eficiencia de la introducción del ADN donante hace que las poblaciones de células creadas por los métodos de la invención sean particularmente útiles como bibliotecas para la presentación y selección de aglutinantes. Por lo tanto, los métodos de la invención producen bibliotecas de alta calidad de aglutinantes en células eucariotas, que pueden cribarse para identificar células que codifican y expresan un aglutinante específico para una diana de interés.

40 En diversos aspectos, la invención se refiere a métodos nuevos y mejorados para preparar bibliotecas de células eucariotas, a las propias bibliotecas, al aislamiento de los aglutinantes deseados, a la codificación de ácido nucleico y a las células de las bibliotecas, y a usos de las bibliotecas tales como para la expresión y cribado de moléculas de unión y para el cribado de los efectos de las moléculas de unión. Se describirán diversos métodos para producir bibliotecas *in vitro* y usar bibliotecas *in vitro* o *in vivo*.

50 La invención proporciona un método para producir una biblioteca de clones de células eucariotas que contienen ADN que codifica un repertorio diverso de aglutinantes, comprendiendo el método usar una nucleasa de sitio específico para dirigirse a la escisión del ADN de células eucariotas para mejorar la integración de sitio específico de los genes de aglutinantes en el ADN celular a través de mecanismos de reparación del ADN celular endógenos.

Un método para producir una biblioteca de clones de células eucariotas que contienen ADN que codifica un repertorio diverso de aglutinantes puede comprender:

55 proporcionar moléculas de ADN donante que codifican los aglutinantes, y células eucariotas, introducir el ADN donante en las células y proporcionar una nucleasa de sitio específico dentro de las células, en donde la nucleasa escinde el ADN celular para crear un sitio de integración en el que el ADN donante se integra en el ADN celular, teniendo lugar la integración a través de mecanismos de reparación del ADN, endógenos a las células.

60 Para los aglutinantes multiméricos que comprenden al menos una primera y segunda subunidades (es decir, cadenas polipeptídicas separadas, tales como los dominios VH y VL del anticuerpo presentados en un formato Fab o IgG), las múltiples subunidades pueden codificarse en la misma molécula de ADN donante. Sin embargo, puede ser deseable integrar las diferentes subunidades en loci separados, en cuyo caso las subunidades pueden proporcionarse en moléculas de ADN donante separadas. Estos podrían integrarse dentro del mismo ciclo de integración dirigida por nucleasas o podrían integrarse secuencialmente utilizando integración dirigida por nucleasas

para una o ambas etapas de integración.

Los métodos para producir bibliotecas de clones de células eucariotas que codifican aglutinantes multiméricos pueden comprender:

5 proporcionar células eucariotas que contienen ADN que codifica la primera subunidad, y proporcionar moléculas de ADN donante que codifican la segunda subunidad de aglutinante,
 introducir el ADN donante en las células y proporcionar una nucleasa de sitio específico dentro de las células, en
 10 donde la nucleasa escinde una secuencia de reconocimiento en el ADN celular para crear un sitio de integración en el que el ADN donante se integra en el ADN celular, teniendo lugar la integración a través de mecanismos de reparación del ADN, endógenos a las células, creando de este modo células recombinantes mientras que
 contienen ADN donante integrado en el ADN celular. Estas células recombinantes contendrán ADN que codifica la primera y segunda subunidades del aglutinante multimérico, y pueden cultivarse para expresar ambas
 15 subunidades. Los aglutinantes multiméricos se obtienen por expresión y ensamblaje de las subunidades codificadas por separado.

En el ejemplo anterior, la integración dirigida por nucleasas se usa para integrar ADN que codifica una segunda subunidad en células que ya contienen ADN que codifica una primera subunidad. La primera subunidad podría introducirse previamente utilizando las técnicas de la presente invención o cualquier otro método de integración de
 20 ADN adecuado. Un enfoque alternativo es usar la integración dirigida por nucleasas en un primer ciclo de introducción de ADN donante, para integrar una primera subunidad, seguido de la introducción de la segunda subunidad ya sea por el mismo enfoque o por cualquier otro método adecuado. Si el enfoque dirigido por nucleasas se usa en múltiples ciclos de integración, se pueden usar opcionalmente diferentes nucleasas de sitio específico para impulsar la integración del ADN donante dirigido por nucleasas en diferentes sitios de reconocimiento. Un
 25 método para generar la biblioteca puede comprender:

proporcionar las primeras moléculas de ADN donante que codifican la primera subunidad y proporcionar células eucariotas,
 30 introducir el primer ADN donante en las células y proporcionar una nucleasa de sitio específico dentro de las células, en donde la nucleasa escinde una secuencia de reconocimiento en el ADN celular para crear un sitio de integración en el que el ADN donante se integra en el ADN celular, teniendo lugar la integración a través de mecanismos de reparación del ADN endógenos a las células, creando de este modo un primer conjunto de células recombinantes que contienen un primer ADN donante integrado en el ADN celular,
 35 cultivar el primer conjunto de células recombinantes para producir un primer conjunto de clones que contienen ADN que codifica la primera subunidad,
 introducir las moléculas del segundo ADN donante que codifican la segunda subunidad en las células del primer conjunto de clones, en donde el segundo ADN donante se integra en el ADN celular del primer conjunto de clones, creando de este modo un segundo conjunto de células recombinantes que contienen el primer y segundo
 40 ADN donante integrado en el ADN celular, y
 cultivar el segundo conjunto de células recombinantes para producir un segundo conjunto de clones, conteniendo estos clones ADN que codifica la primera y segunda subunidades del aglutinante multimérico,
 proporcionar de este modo una biblioteca de clones de células eucariotas que contienen ADN donante que
 45 codifica el repertorio de aglutinantes multiméricos.

La integración de sitio específico del ADN donante en el ADN celular crea células recombinantes, que pueden cultivarse para producir clones. Por lo tanto, las células recombinantes individuales en las que se ha integrado el
 50 ADN donante se replican para generar poblaciones clonales de células ("clones"), derivándose cada clon de una célula recombinante original. Por lo tanto, el método genera una serie de clones correspondientes al número de células en las que se integró con éxito el ADN donante. La colección de clones forma una biblioteca que codifica el repertorio de aglutinantes (o, en una etapa intermedia donde las subunidades de aglutinante se integran en rondas separadas, los clones pueden codificar un conjunto de subunidades de aglutinante). Por lo tanto, los métodos de la invención pueden proporcionar una biblioteca de clones de células eucariotas que contienen ADN donante que
 55 codifica el repertorio de aglutinantes.

Los métodos de la invención pueden generar bibliotecas de clones que contienen ADN donante integrado en un locus fijo, o en múltiples loci fijos, en el ADN celular. Por "fijo" se entiende que el locus es el mismo entre las células. Las células utilizadas para la creación de la biblioteca, por lo tanto, pueden contener una secuencia de reconocimiento de nucleasas en un locus fijo, que representa un sitio de aterrizaje universal en el ADN celular en el que el ADN donante puede integrarse. La secuencia de reconocimiento para la nucleasa de sitio específico puede estar presente en una o más de una posición en el ADN celular.
 60

Las bibliotecas producidas de acuerdo con la presente invención pueden emplearse de diversas formas. Se puede cultivar una biblioteca para expresar los aglutinantes, produciendo de este modo un repertorio diverso de aglutinantes. Una biblioteca puede seleccionarse para una célula de un fenotipo deseado, en donde el fenotipo es resultado de la expresión de un aglutinante por una célula. La selección de fenotipos es posible en la que las células de la biblioteca se cultivan para expresar los aglutinantes, seguido de la detección de si el fenotipo deseado se
 65

muestra en los clones de la biblioteca. Las lecturas celulares pueden basarse en la alteración del comportamiento celular, tal como la expresión alterada de genes indicadores endógenos o exógenos, el estado de diferenciación, la proliferación, la supervivencia, el tamaño celular, el metabolismo o las interacciones alteradas con otras células. Cuando se detecta el fenotipo deseado, pueden recuperarse entonces las células de un clon que presenta el fenotipo deseado. Opcionalmente, el ADN que codifica el aglutinante se aísla entonces del clon recuperado, proporcionando ADN que codifica un aglutinante que produce el fenotipo deseado cuando se expresa en la célula.

Un propósito clave para el que se han usado bibliotecas de células eucariotas es en los métodos de selección de aglutinantes que reconocen una diana de interés. En dichos métodos, se cultiva una biblioteca para expresar los aglutinantes, y los aglutinantes se exponen a la diana para permitir el reconocimiento de la diana por uno o más aglutinantes afines, si están presentes, y detectar si la diana es reconocida por un aglutinante análogo. En dichos métodos, los aglutinantes pueden mostrarse sobre la superficie celular y los clones de la biblioteca que presentan los aglutinantes con las propiedades deseadas pueden aislarse. Por lo tanto, las células que incorporan genes que codifican aglutinantes con características funcionales o de unión deseadas podrían identificarse dentro de la biblioteca. Los genes se pueden recuperar y usar para la producción del aglutinante o se pueden usar para una modificación adicional para crear bibliotecas de derivados de aglutinantes para producir aglutinantes con propiedades mejoradas.

La presente invención ofrece ventajas sobre los enfoques anteriores para la construcción de bibliotecas en eucariotas superiores. Algunos estudios han utilizado la infección lentiviral para introducir genes de anticuerpos en células indicadoras de mamífero [106]. Esto tiene la ventaja de que se pueden generar grandes bibliotecas, pero no hay control sobre el sitio de integración y el número de copias se controla mediante una baja multiplicidad de infección (como se ha analizado anteriormente). En un enfoque alternativo, los genes de anticuerpos se introdujeron mediante recombinación homóloga, sin el beneficio de la integración dirigida por nucleasas y utilizando brazos de homología de 10 kb, pero la eficacia del direccionamiento fue relativamente baja, lo que significa que el tamaño potencial de la biblioteca era limitado [105]. En contraste, el uso de nucleasas dirigidas por secuencia conserva los beneficios de la integración dirigida a uno o unos pocos loci de elección mientras permite la construcción eficiente de grandes bibliotecas. La integración dirigida por nucleasas tiene la ventaja de que los transgenes se dirigen a un locus fijo o loci fijos dentro del ADN celular. Esto significa que la actividad del promotor que impulsa la transcripción de los genes de aglutinante en todos los clones será la misma y la funcionalidad de cada aglutinante será un reflejo de su potencia inherente, la eficiencia de la traducción y la estabilidad en lugar de deberse a la variación relacionada con el sitio de integración. El direccionamiento a un solo número o un número limitado de loci también permitirá un mejor control de la expresión si es necesario, por ejemplo, utilizando promotores inducibles.

A continuación, se describen con más detalle diversas características de la invención. Se observa que los encabezados utilizados a lo largo de esta memoria descriptiva son solo para ayudar a la navegación y no deben interpretarse como definitivos, y que las realizaciones descritas en diferentes secciones se pueden combinar según sea apropiado.

Descripción detallada

Células eucariotas

El potencial de las poblaciones de células eucariotas que expresan un repertorio diverso de aglutinantes se ilustra y se analiza en los Ejemplos en el presente documento en relación con la expresión de repertorios de anticuerpos en la superficie de células de mamífero. Los beneficios de la invención no se limitan a las células de mamífero e incluyen todos los eucariotas.

La levadura (por ejemplo, *Saccharomyces cerevisiae*) tiene un genoma más pequeño que las células de mamífero y la recombinación homóloga dirigida por brazos de homología (en ausencia de escisión dirigida por nucleasas) es una forma eficaz de introducir ADN extraño en comparación con los eucariotas superiores. Por lo tanto, un beneficio particular de la integración dirigida por nucleasas de la presente invención se refiere a la integración de genes de aglutinante en células eucariotas superiores con genomas más grandes donde la recombinación homóloga en ausencia de escisión de nucleasas es menos eficaz. La integración dirigida por nucleasas se ha utilizado en células de levadura para resolver el problema de la integración eficiente de múltiples genes en células de levadura individuales, por ejemplo, para la modificación por ingeniería de vías metabólicas (documento US2012/0277120), pero este trabajo no incorpora la introducción de bibliotecas de aglutinantes ni aborda los problemas de la construcción de bibliotecas en los eucariotas superiores.

Las bibliotecas de células eucariotas de acuerdo con la presente invención son preferiblemente células eucariotas superiores, definidas aquí como células con un genoma mayor que el de *Saccharomyces cerevisiae* que tiene un tamaño de genoma de 12×10^6 pares de bases (pb). Las células eucariotas superiores pueden tener, por ejemplo, un tamaño de genoma superior a 2×10^7 pares de bases. Esto incluye, por ejemplo, células de mamífero, aves, insectos o vegetales. Preferiblemente, las células son células de mamífero, por ejemplo, ratón o humano. Las células pueden ser células primarias o líneas celulares. Las células de ovario de hámster chino (CHO) se usan comúnmente para la expresión de anticuerpos y proteínas, pero se puede usar en la invención cualquier línea celular

estable alternativa. Las células HEK293 se usan en los Ejemplos en el presente documento. Hay métodos disponibles para la introducción eficiente de ADN extraño en las células primarias, lo que permite su uso (por ejemplo, mediante electroporación donde se han logrado eficiencias y viabilidades de hasta el 95 %) <http://www.maxcyte.com/technology/primary-cells-stem-cells.php>.

Las células de linaje de linfocitos T (por ejemplo, linfocitos T primarios o una línea de linfocitos T) o células de linaje de linfocitos B están entre los tipos de células preferidos. De particular interés son los linfocitos T primarios o las líneas celulares derivadas de linfocitos T para su uso en bibliotecas de TCR, incluidas las líneas celulares que carecen de la expresión de TCR [23, 24, 25]. Los ejemplos de células del linaje de linfocitos B incluyen linfocitos B, linfocitos pre-B o linfocitos pro-B y líneas celulares derivadas de cualquiera de estos.

La construcción de bibliotecas en linfocitos B primarios o líneas de linfocitos B sería de particular valor para la construcción de bibliotecas de anticuerpos. Breous-Nystrom *et al.* [15] han generado bibliotecas en una línea de linfocitos pre-B murinos (1624-5). La línea celular derivada de linfocitos B de pollo DT40 (ATCC CRL-2111) tiene una promesa particular para la construcción de bibliotecas de aglutinantes. DT40 es una línea celular pequeña con una tasa de división celular relativamente rápida. Los repertorios de aglutinantes podrían dirigirse a loci específicos utilizando ZFN, nucleasas TALE o CRISPR/Cas9 dirigidas a secuencias endógenas o dirigiéndose a sitios heterólogos preintegrados que podrían incluir sitios de reconocimiento de meganucleasa. Las células DT40 expresan anticuerpos y, por lo tanto, será ventajoso dirigirse a los genes de anticuerpos dentro del locus de anticuerpos con o sin alteración de los dominios variables de anticuerpos de pollo endógenos. Las células DT40 también se han utilizado como la base de un sistema *in vitro* para la generación de IgM de pollo denominado sistema de biblioteca de diversificación autónoma (sistema ADLib) que aprovecha la diversificación intrínseca que se produce en el locus del anticuerpo de pollo. Como resultado de esta diversificación endógena, es posible generar especificidades novedosas. El enfoque dirigido por nucleasas descrito aquí podría usarse en combinación con ADLib para combinar diversas bibliotecas de aglutinantes de fuentes heterólogas (por ejemplo, repertorios de regiones variables de anticuerpos humanos o andamiajes alternativos derivados sintéticamente) con el potencial de una diversificación adicional con el locus de IgG de pollo. Se podrían aplicar beneficios similares a las líneas de linfocitos B humanos como Nalm6 [26]. Otras líneas celulares de interés del linaje B incluyen líneas tales como la línea de linfocitos pre-B murinos 1624-5 y la línea de linfocitos pro-B Ba/F3. Ba/F3 depende de IL-3 [27] y su uso se analiza en otra parte del presente documento. Finalmente, se podrían usar varias líneas celulares humanas, incluidas las enumeradas en la "Enciclopedia de líneas celulares de cáncer" [28] o "Catálogo COSMIC de mutaciones somáticas en el cáncer" [29].

Normalmente, la biblioteca estará compuesta por un único tipo de células, producidas por la introducción de ADN donante en una población de células eucariotas clonales, por ejemplo, por introducción de ADN donante en células de una línea celular particular. La principal diferencia significativa entre los diferentes clones de la biblioteca será entonces debido a la integración del ADN donante.

Sistemas virales eucariotas

Las ventajas del sistema en la creación de bibliotecas de aglutinantes en células eucarióticas podrían aplicarse a sistemas de presentación viral basados en sistemas de expresión eucariotas, por ejemplo, presentación baculoviral o presentación retroviral [1, 2, 3, 4]. En este enfoque, cada célula codificará un aglutinante capaz de incorporarse en una partícula viral. En el caso de los sistemas retrovirales, el ARNm codificante se empaquetaría y el aglutinante codificado se presentaría en la superficie celular. En el caso de los sistemas baculovirales los genes que codifican el aglutinante deberían encapsularse en la partícula baculoviral para mantener una asociación entre el gen y la proteína codificada. Esto podría lograrse utilizando células hospedadoras que portan copias episómicas del genoma baculoviral. Como alternativa, las copias integradas podrían liberarse siguiendo la acción de una nucleasa específica (distinta de la utilizada para impulsar la integración de sitio específico). En el caso de las moléculas de aglutinante multiméricas, algunos socios podrían codificarse dentro del ADN celular con los genes para uno o más compañeros empaquetados dentro del virus.

Nucleasa de sitio específico

La invención implica el uso de una nucleasa de sitio específico para la escisión dirigida de ADN celular en la construcción de una biblioteca de células eucarióticas que contienen ADN que codifica un repertorio de aglutinantes, en donde la escisión del ADN mediada por nucleasas mejora la integración de sitio específico de los genes de aglutinante a través de mecanismos de reparación de ADN celular endógenos. La nucleasa de sitio específico escinde el ADN celular después de la unión específica a una secuencia de reconocimiento, creando de este modo un sitio de integración para el ADN donante. La nucleasa puede crear una rotura bicatenaria o una rotura monocatenaria (una mella). Las células utilizadas para la creación de la biblioteca pueden contener secuencias endógenas reconocidas por la nucleasa de sitio específico o la secuencia de reconocimiento se puede modificar en el ADN celular.

La nucleasa de sitio específico puede ser exógena a las células, es decir, no se produce naturalmente en las células del tipo elegido.

La nucleasa de sitio específico puede introducirse antes, después o simultáneamente con la introducción del ADN donante que codifica el aglutinante. Puede ser conveniente que el ADN donante codifique la nucleasa además del aglutinante, o en un ácido nucleico separado que se cotransfecta o se introduce de otro modo al mismo tiempo que el ADN donante. Los clones de una biblioteca pueden conservar opcionalmente el ácido nucleico que codifica la nucleasa de sitio específico, o dicho ácido nucleico puede transfectarse solo transitoriamente en las células.

Se puede usar cualquier nucleasa de sitio específico adecuada con la invención. Puede ser una enzima de origen natural o una variante modificada. Hay una serie de nucleasas conocidas que son especialmente adecuadas, tales como las que reconocen, o pueden diseñarse para reconocer, secuencias que se producen raramente en el ADN celular. La escisión de la nucleasa en solo uno o dos sitios es ventajosa, ya que esto debería garantizar que solo una o dos moléculas de ADN donante estén integradas por célula. La rareza de la secuencia reconocida por la nucleasa de sitio específico es más probable si la secuencia de reconocimiento es relativamente larga. La secuencia específicamente reconocida por la nucleasa puede ser, por ejemplo, una secuencia de al menos 10, 15, 20, 25 o 30 nucleótidos.

Los ejemplos de nucleasas adecuadas incluyen meganucleasas, nucleasas con dedos de cinc (ZFN), nucleasas TALE y nucleasas guiadas por ácido nucleico (por ejemplo, guiadas por ARN) tales como el sistema CRISPR/Cas. Cada una de estas produce roturas bicatenarias, aunque se conocen formas modificadas que generan roturas monocatenarias.

Las meganucleasas (también conocidas como endonucleasas de orientación) son nucleasas que se producen en todos los reinos de la vida y reconocen secuencias relativamente largas (12-40 pb). Dada la larga secuencia de reconocimiento, están ausentes o se producen con poca frecuencia en genomas eucariotas. Las meganucleasas se agrupan en 5 familias según la secuencia/estructura. (LAGLIDADG, GIY-YIG, HNH, caja de His-Cys y PD-(D/E)XK). La familia mejor estudiada es la familia LAGLIDADG que incluye la meganucleasa I-SceI bien caracterizada de *Saccharomyces cerevisiae*. I-SceI reconoce y escinde una secuencia de reconocimiento de 18 pb (5' TAGGGATAACAGGGTAAT) dejando un saliente 3' de 4 pb. Otro ejemplo comúnmente utilizado es I-Cre1, que se origina en el cloroplasto de las algas verdes unicelulares de *Chlamydomonas reinhardtii*, y reconoce una secuencia de 22 pb [30]. Se han creado varias variantes modificadas con secuencias de reconocimiento alteradas [31]. Las meganucleasas representan el primer ejemplo del uso de nucleasas de sitio específico en la ingeniería del genoma [49, 50]. Al igual que con los enfoques basados en recombinasa, el uso de I-Sce1 y otras meganucleasas requiere la inserción previa de un sitio de reconocimiento apropiado a direccionar dentro del genoma o la ingeniería de las meganucleasas para reconocer los sitios endógenos [30]. Mediante este enfoque, la eficacia de direccionamiento en las células HEK293 (según lo juzgado por la "reparación" dirigida por homología de un gen GFP defectuoso integrado) se logró en el 10-20 % de las células a través del uso de I-Sce1 [32].

Una clase preferida de meganucleasas para su uso en la presente invención son las endonucleasas. Estas incluyen I-Sce I, I-Chu I, I-Cre I, Csm I, PI-Sce I, PI-Tli I, PI-Mtu I, I-Ceu I, I-Sce II, I-Sce III, HO, Pi-Civ I, PI-Ctr I, PI-Aae I, PI-Bsu I, PI-Dha I, PI-Dra I, PI-Mav I, PI-Mch I, PI-Mfu PI-Mfl I, PI-Mga I, PI-Mgo I, PI-Min I, PI-Mka I, PI-Mle I, PI-Mma I, PI-Msh I, PI-Msm I, PI-Mth I, PI-Mtu PI-Mxe I, PI-Npu I, PI-Pfu I, PI-Rma I, PI-Spb I, PI-Ssp I, PI-Fac I, PI-Mja I, PI-Pho I, Pi-Tag I, PI-Thy I, PI-Tko I, I-Msol, y PI-Tsp I; preferiblemente, I-Sce I, I-Cre I, I-Chu I, I-Dmo I, I-Csm I, PI-Sce I, PI-Pfu I, PI-Tli I, PI-Mtu I, e I-Ceu I.

En los últimos años, se han desarrollado varios métodos que permiten el diseño de nuevas nucleasas de secuencia específica mediante la fusión de dominios de unión de ADN de secuencia específica a nucleasas no específicas para crear nucleasas diseñadas de secuencia específica dirigidas a través de dominios de unión de ADN a medida. La especificidad de unión puede dirigirse por dominios de unión modificados tales como dominios con dedos de cinc. Estos son pequeños dominios modulares, estabilizados por iones de cinc, que están involucrados en el reconocimiento molecular y se utilizan en la naturaleza para reconocer secuencias de ADN. Las matrices de dominios con dedos de cinc se han diseñado para la unión de secuencia específica y se han vinculado al dominio de escisión de ADN no específico de la enzima de restricción de tipo II FokI para crear nucleasas con dedos de cinc (ZFN). Las ZFN se pueden usar para crear una rotura bicatenaria en sitios específicos dentro del genoma. FokI es un dímero obligado y requiere que dos ZFN se unan en estrecha proximidad para efectuar la escisión. La especificidad de las nucleasas modificadas se ha mejorado y su toxicidad se ha reducido al crear dos variantes de FokI diferentes que están modificadas para formar solo heterodímeros entre sí [33]. Se ha demostrado que dichas ZFN heterodiméricas obligadas logran una integración dirigida por homología en el 5-18 % de las células diana sin la necesidad de selección de fármacos [21, 34, 35]. La incorporación de inserciones de hasta 8 kb con frecuencias de >5 % se ha demostrado en ausencia de selección.

Recientemente se ha demostrado que los salientes 5' monocatenarios creados por nucleasas tales como las ZFN ayudan a impulsar la integración eficiente de los transgenes a los sitios de escisión [45]. Esto se ha extendido para mostrar que la escisión *in vivo* del ADN donante (a través de la inclusión de un sitio de reconocimiento de nucleasa específico dentro del plásmido donante) aumenta la eficiencia en la integración no homóloga. El mecanismo no está completamente claro, pero es posible que la exposición reducida a las nucleasas celulares a través de la linealización *in vivo* pueda haber contribuido a la mejora [45]. También son posibles las correspondencias en los salientes 5' del ADN donante y aceptor, generados por la unión impulsada por las nucleasas. El examen de las

- 5 secuencias en las uniones, sin embargo, mostró la aparición de eliminaciones. Es posible que las uniones perfectamente emparejadas continúen actuando como sustrato para las nucleasas de sitio dirigido hasta que se produzca la eliminación de la secuencia de reconocimiento. Para superar este problema potencial, Maresca *et al.* [36] han invertido los sitios de reconocimiento de las ZFN izquierda y derecha dentro del ADN donante, de modo que la ligadura del ADN donante en el locus genómico conducirá a la duplicación de dos ZFN izquierdas en un flanco de la integración y la duplicación de dos ZFN derechos en el otro flanco. El uso de nucleasas heterodímeras obligadas (como se describe para Fok1) significa que ninguna de estas secuencias flanqueantes recién creadas puede escindirse por la nucleasa dirigida.
- 10 La capacidad de modificar dominios de unión a ADN de especificidad definida se ha simplificado adicionalmente por el descubrimiento en bacterias *Xanathomonas* de moléculas de efectores de tipo activadores de la transcripción (TALE). Estas moléculas TALE consisten en matrices de monómeros de 33-35 aminoácidos, y reconociendo cada monómero una base única dentro de una secuencia diana [37]. Esta relación modular 1:1 ha hecho que sea relativamente fácil diseñar moléculas TALE modificadas para unirse a cualquier diana de ADN de interés. Al acoplar estos TALE diseñados a Fok1, ha sido posible crear nuevas nucleasas TALE de secuencia específica. Las nucleasas TALE, también conocidas como TALEN, ahora se han diseñado para un gran número de sitios y muestran una alta tasa de éxito para una actividad de modificación génica eficiente [38]. En los ejemplos en el presente documento, se demostró la integración mejorada del ADN donante mediante el uso de nucleasas TALE. Se han desarrollado otras variaciones y mejoras de la tecnología de nucleasas TALE y se podrían utilizar para la generación de bibliotecas de aglutinantes a través de la integración dirigida por nucleasas. Estos incluían "mega-TALEN" donde un dominio de unión a nucleasa TALE se fusiona con una meganucleasa [39] y "TALEN compactas" donde se usa un único dominio de reconocimiento de nucleasa TALE para efectuar la escisión [40].
- 25 En los últimos años, se ha descrito otro sistema para dirigir las roturas bicatenarias o monocatenarias a secuencias específicas en el genoma. Este sistema, denominado sistema de "repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR) y asociado a CRISPR (Cas)", se basa en un mecanismo de defensa bacteriana [41]. El sistema CRISPR/Cas se dirige al ADN para su escisión a través de un ARN monocatenario complementario corto (ARN CRISPR o ARNcr) unido a una repetición palindrómica corta. En el sistema "Tipo II" comúnmente utilizado, el procesamiento del ARN de direccionamiento depende de la presencia de un ARNcr de trans-activación (ARNtracr) que tiene una secuencia complementaria a la repetición palindrómica. La hibridación del ARNtracr con la secuencia de repetición palindrómica desencadena el procesamiento. El ARN procesado activa el dominio Cas9 y dirige su actividad a la secuencia complementaria dentro del ADN. El sistema se ha simplificado para dirigir la escisión de Cas9 a partir de un único transcrito de ARN y se ha dirigido a muchas secuencias diferentes dentro del genoma [42, 43]. Este enfoque de la escisión del genoma tiene la ventaja de que se dirige a través de una secuencia corta de ARN, lo que hace que sea relativamente sencillo diseñar la especificidad de la escisión. Por lo tanto, hay varias formas diferentes de lograr la escisión de sitio específico del ADN genómico. Como se ha descrito anteriormente, esto aumenta la velocidad de integración de un plásmido donante a través de mecanismos de reparación de ADN celular endógeno.
- 40 El uso de meganucleasas, ZFN, nucleasa TALE o sistemas guiados por ácidos nucleicos, tales como los sistemas CRISPR/Cas9, permitirá el direccionamiento de loci endógenos dentro del genoma. En los Ejemplos en el presente documento, se ha demostrado el direccionamiento al locus AAVS, pero se podrían dirigir loci alternativos. Por ejemplo, el locus del gen de colágeno de tipo I se ha utilizado para la expresión eficiente de transgenes [44].
- 45 Como alternativa, podrían introducirse por adelantado sitios de reconocimiento heterólogos para nucleasas dirigidas, incluyendo meganucleasas, ZFN y nucleasas TALE para el posterior direccionamiento de bibliotecas. En los Ejemplos en el presente documento, se describe el uso de una nucleasa TALE que reconoce una secuencia dentro del locus AAVS para introducir mediante recombinación homóloga, una secuencia de reconocimiento de meganucleasa I-Sce1 y sitios heterólogos de reconocimiento de nucleasa TALE dentro del locus AAVS. El direccionamiento dirigido por nucleasas podría usarse para impulsar la inserción de secuencias diana por recombinación homóloga o NHEJ utilizando un vector de ADN o incluso oligonucleótidos bicatenarios [45]. Como alternativa, se podrían usar métodos de direccionamiento no específicos para introducir sitios de direccionamiento a través del uso de la integración dirigida por transposón [46] para introducir sitios de reconocimiento para nucleasas de sitio específico. Los sistemas basados en virus, tales como lentivirus, aplicados a bajo título también podrían usarse para introducir sitios de direccionamiento. La transfección de ADN junto con el cribado de la inserción de una sola copia también se ha utilizado para identificar sitios de integración únicos [17]. Dichos enfoques no específicos serían particularmente útiles en el caso de células que no tienen un sitio obvio para el direccionamiento o para genomas que no se han secuenciado o para genomas para los cuales no se dispone de nucleasas TALE, ZFN o sistemas Cas9/CRISPR. Una vez que se ha establecido una línea celular después de la inserción aleatoria de un sitio de reconocimiento de nucleasas, la línea celular se puede usar posteriormente para crear bibliotecas de aglutinantes donde todos los clones de la biblioteca contienen el transgén en el locus fijo utilizando la integración dirigida por nucleasas.
- 60 En los Ejemplos presentados, se usan tres plásmidos diferentes que incluyen pares de nucleasas TALE o ZFN en plásmidos individuales con un plásmido separado para el ADN donante. En el caso de la meganucleasa, la nucleasa de sitio específico está codificada por un solo gen y se introduce en un plásmido con el ADN donante presente en un

segundo plásmido. Por supuesto, se podrían usar combinaciones que incorporen dos o más de estos elementos en el mismo plásmido y esto podría mejorar la eficiencia del direccionamiento reduciendo el número de plásmidos a introducir. Además, puede ser posible preintegrar la nucleasa o nucleasas que también podrían ser inducibles para permitir el control temporal de la actividad de la nucleasa, como se ha demostrado para las transposasas [46]. Finalmente, la nucleasa podría introducirse como proteína recombinante o complejo de proteína:ARN (por ejemplo, en el caso de una nucleasa dirigida por ARN, tal como CRISPR:Cas9).

Locus

Una secuencia de reconocimiento para la nucleasa de sitio específico puede estar presente en el ADN genómico, o ADN episómico que se hereda de forma estable en las células. Por lo tanto, el ADN donante se puede integrar en un locus genómico o episómico en el ADN celular.

En su forma más simple, un único gen que codifica un aglutinante (gen de aglutinante) se dirige a un sitio único dentro del genoma eucariota. La identificación de una célula que demuestre una actividad de unión particular o un fenotipo celular permitirá el aislamiento directo del gen que codifica la propiedad deseada (por ejemplo, mediante PCR a partir de ARNm o ADN genómico). Esto se facilita utilizando una secuencia de reconocimiento única para la nucleasa de sitio específico, que se produce una vez en el ADN celular. Por lo tanto, las células utilizadas para la creación de la biblioteca pueden contener una secuencia de reconocimiento de nucleasas en un solo locus fijo, es decir, un locus idéntico en todas las células. Las bibliotecas producidas a partir de tales células contendrán ADN donante integrado en el locus fijo, es decir, que se producen en el mismo locus en el ADN celular de todos los clones en la biblioteca.

Opcionalmente, las secuencias de reconocimiento pueden tener lugar varias veces en el ADN celular, de modo que las células tienen más de un sitio de integración potencial para el ADN donante. Esta sería una situación típica para células diploides o poliploides en las que la secuencia de reconocimiento está presente en las posiciones correspondientes en un par de cromosomas, es decir, loci replicados. Las bibliotecas producidas a partir de tales células pueden contener ADN donante integrado en loci fijos replicados. Por ejemplo, las bibliotecas producidas a partir de células diploides pueden tener un ADN donante integrado en loci fijos duplicados y las bibliotecas producidas a partir de células triploides pueden tener un ADN donante integrado en loci fijos triplicados. Muchas células de mamíferos adecuadas son diploides, y los clones de las bibliotecas de células de mamífero de acuerdo con la invención pueden tener ADN donante integrado en loci fijos duplicados.

La secuencia reconocida por la nucleasa de sitio específico puede tener lugar en más de un locus independiente en el ADN celular. Por lo tanto, el ADN donante puede integrarse en múltiples loci independientes. Las bibliotecas de células diploides o poliploides pueden comprender ADN donante integrado en múltiples loci fijos independientes y/o en loci fijos replicados.

En células que contienen secuencias de reconocimiento en múltiples loci (ya sean loci replicados o independientes), cada locus representa un sitio de integración potencial para una molécula de ADN donante. La introducción del ADN donante en las células puede dar como resultado la integración en el número total de secuencias de reconocimiento de nucleasas presentes en la célula, o el ADN donante puede integrarse en algunos, pero no en todos estos sitios potenciales. Por ejemplo, cuando se produce una biblioteca a partir de células diploides que contienen secuencias de reconocimiento en el primer y segundo loci fijo (por ejemplo, loci fijo duplicado), la biblioteca resultante puede comprender clones en los que el ADN donante se integra en el primer locus fijo, clones en los que el ADN donante se integra en el segundo locus fijo, y los clones en los que el ADN donante se integra en el primer y segundo loci fijos.

Por lo tanto, los métodos para producir bibliotecas pueden implicar la escisión por nucleasa de sitio específico de múltiples loci fijos en una célula, y la integración del ADN donante en los múltiples loci fijos. Como se ha señalado anteriormente, en los casos en que hay múltiples copias de la misma secuencia de reconocimiento (por ejemplo, como tiene lugar cuando se dirigen a loci endógenos en células diploides o poliploides), es posible que se integren dos genes de aglutinante, particularmente cuando se utiliza un mecanismo eficiente de direccionamiento, siendo un solo gen específico para la diana. Esto se puede resolver durante el cribado posterior una vez que se hayan aislado los genes de aglutinante.

En algunos casos, puede ser deseable introducir más de un aglutinante por célula. Por ejemplo, se podrían generar aglutinantes bi-específicos a partir de dos anticuerpos diferentes que se juntan y estos pueden tener propiedades ausentes en los aglutinantes individuales [47]. Esto podría lograrse mediante la introducción de diferentes genes de anticuerpos en ambos alelos en loci fijos duplicados o dirigiéndose a diferentes poblaciones de anticuerpos en loci fijos independientes utilizando los métodos descritos en el presente documento. Además, un aglutinante puede estar compuesto por múltiples cadenas (por ejemplo, los dominios VH y VL del anticuerpo presentados dentro de un formato Fab o IgG). En este caso, puede ser conveniente integrar las diferentes subunidades en diferentes loci. Estos podrían integrarse dentro del mismo ciclo de integración dirigida por nucleasas, o podrían integrarse secuencialmente utilizando integración dirigida por nucleasas para una o ambas etapas de integración.

Introducción de ADN donante

Se han descrito numerosos métodos para introducir ADN donante en células eucariotas, incluyendo transfección, infección o electroporación. La transfección de un gran número de células es posible mediante métodos estándar que incluyen transfección mediada por polietilimina como se describe en el presente documento. Además, hay métodos disponibles para la electroporación altamente eficiente de 10^{10} células en 5 minutos, por ejemplo, <http://www.maxcyte.com>.

Se podrían crear bibliotecas combinatorias en las que se introducen en plásmidos diferentes miembros de pares de unión multiméricos (por ejemplo, genes VH y VL de genes de anticuerpos) o incluso partes diferentes de la misma molécula de aglutinante. La introducción de moléculas de ADN donante separadas que codifican aglutinantes separados o subunidades de aglutinante se puede realizar de manera simultánea o secuencial. Por ejemplo, se podría introducir una cadena ligera de anticuerpo por transfección o infección, las células crecerán y se seleccionarán si es necesario. Después, se podrán introducir otros componentes en una etapa posterior de infección o transfección. Una o ambas etapas podrían implicar la integración dirigida por nucleasas a loci genómicos específicos.

Integración de ADN donante

El ADN donante se integra en el ADN celular, formando ADN recombinante que tiene una secuencia de ADN contigua en la que se inserta el ADN donante en el sitio de integración. En la presente invención, la integración está mediada por los mecanismos de reparación de ADN natural que son endógenos a la célula. Por lo tanto, se puede permitir que la integración tenga lugar simplemente introduciendo el ADN donante en una célula, permitiendo que la nucleasa de sitio específico cree un sitio de integración, y permitiendo que el ADN donante se integre. Las células pueden mantenerse en cultivo durante un tiempo suficiente para que el ADN se integre. Esto generalmente dará como resultado una población mixta de células, incluyendo (i) células recombinantes en las que el ADN donante se ha integrado en el sitio de integración creado por la nucleasa de sitio específico, y opcionalmente (ii) células en las que el ADN donante se ha integrado en los sitios distintos del sitio de integración deseado y/u opcionalmente (iii) células en las que el ADN donante no se ha integrado. Por lo tanto, las células recombinantes deseadas y los clones resultantes de la biblioteca pueden proporcionarse en una población mixta de otras células eucariotas. Los métodos de selección descritos en otras partes en el presente documento se pueden usar para enriquecer las células de la biblioteca.

Los mecanismos de reparación de ADN endógenos en células eucariotas incluyen recombinación homóloga, unión de extremos no homólogos (NHEJ) y unión de extremos dirigida por microhomología. La eficiencia de la modificación del ADN mediante tales procesos puede aumentarse mediante la introducción de roturas bicatenarias (DSB) en el ADN y se han informado ganancias de eficiencia de 40.000 veces utilizando endonucleasas de corte raras (meganucleasas) tal como I-Sce1 [48, 49, 50].

A diferencia de la recombinación de sitio específico implicada en sistemas tales tal como el sistema Flp-In [16], la presente invención no requiere recombinasas exógenas o sitios de reconocimiento de recombinasas modificadas. Por lo tanto, opcionalmente, la presente invención no incluye una etapa de integración del ADN mediada por recombinasa en la creación de la biblioteca, y/u opcionalmente las células eucariotas en las que se introduce el ADN donante carecen de un sitio de recombinación para una recombinasa de sitio específico. Los mecanismos y aspectos prácticos de la inserción dirigida de ADN donante en el ADN celular por recombinasas y nucleasas son muy distintos. Según lo analizado por Jasin 1996 [50]:

"... la reacción catalizada por las recombinasas sitio específico es bastante distinta de la reparación celular de las DSB. Las recombinasas sitio específico, tal como cre, hacen sinapsis con dos sitios de reconocimiento y crean roturas monocatenarias dentro de los sitios, formando de este modo intermedios de Holliday. Los intermedios se resuelven para producir eliminaciones, inversiones e inserciones (cointegrantes), todos los cuales restauran los dos sitios de reconocimiento. La reacción es absolutamente precisa y, por lo tanto, reversible. Las roturas nunca se exponen a la maquinaria de reparación celular".

En contraste, la nucleasa de sitio específico actúa para crear roturas o mellas dentro del ADN celular (por ejemplo, genómico o episómico), que se exponen a y se reparan mediante mecanismos de reparación celular endógenos, tal como la recombinación homóloga o NHEJ. Los enfoques basados en la recombinasa tienen un requisito absoluto para la preintegración de sus sitios de reconocimiento, por lo que dichos métodos requieren la modificación por ingeniería del sitio de integración del "punto caliente" en el ADN celular como una etapa preliminar. Con la integración dirigida por nucleasas, es posible modificar nucleasas o dirigirse a través del ARN guía en el caso de CRISPR:Cas9 para reconocer loci endógenos, es decir, las secuencias de ácido nucleico que se producen de forma natural en el ADN celular. Finalmente, a nivel práctico, los enfoques dirigidos por nucleasas son más eficientes para la integración directa de transgenes en los niveles requeridos para hacer grandes bibliotecas de aglutinantes.

El mecanismo de reparación del ADN mediante el cual el ADN donante se integra en los métodos de la invención se puede predeterminar o sesgar en cierta medida mediante el diseño del ADN donante y/o la elección de la nucleasa de sitio específico.

La recombinación homóloga es un mecanismo natural utilizado por las células para reparar roturas bicatenarias utilizando una secuencia homóloga (por ejemplo, de otro alelo) como plantilla para la reparación. La recombinación homóloga se ha utilizado en la ingeniería celular para introducir inserciones (incluidos los transgenes), eliminaciones y mutaciones puntuales en el genoma. La recombinación homóloga se promueve proporcionando brazos de homología en el ADN donante. El enfoque original para modificar células eucariotas superiores usó normalmente brazos de homología de 5-10 kb dentro de un plásmido donante para aumentar la eficiencia de la integración dirigida en el sitio de interés. A pesar de esto, la recombinación homóloga impulsada puramente por brazos de homología largos, es menos eficiente que la recombinación dirigida por Flp y Cre, particularmente en eucariotas superiores con genomas grandes. La recombinación homóloga es particularmente adecuada para eucariotas tal como la levadura, que tiene un tamaño de genoma de solo $12,5 \times 10^6$ pb, donde es más eficaz en comparación con las eucariotas superiores con genomas más grandes, por ejemplo, células de mamífero con 3000×10^6 pb.

La recombinación homóloga también puede dirigirse a través de mellas [52] en el ADN genómico y esto también podría servir como una ruta para la integración dirigida por nucleasas en el ADN genómico. Se ha demostrado que dos vías distintas promueven la recombinación homóloga en el ADN mellado. Uno es esencialmente similar a la reparación en roturas bicatenarias, utilizando Rad51/Brca2, mientras que el otro se inhibe por Rad51/Brca2 y utiliza preferiblemente ADN monocatenario o ADN donante bicatenario mellado [51].

La unión de extremos no homólogos (NHEJ) es un mecanismo alternativo para reparar las roturas bicatenarias en el genoma donde los extremos del ADN se vuelven a ligar directamente sin la necesidad de una plantilla homóloga. La escisión dirigida por nucleasas del ADN genómico también puede mejorar la integración transgénica a través de mecanismos no basados en homología. Este enfoque con respecto a la reparación del ADN es menos preciso y puede conducir a inserciones o eliminaciones. No obstante, NHEJ proporciona un medio simple para integrar los exones en marco en el intrón o permite la integración de los casetes promotor:gen en el genoma. El uso de métodos no homólogos permite el uso de vectores donantes que carecen de brazos de homología, simplificando de este modo la construcción del ADN donante.

Se ha señalado que se utilizan regiones cortas de homología terminal para volver a unir los extremos del ADN y se planteó la hipótesis de que se podrían utilizar 4 pb de microhomología para dirigir la reparación en roturas bicatenarias, denominadas unión de extremos dirigida por microhomología [50].

ADN donante

El ADN donante usualmente será ADN circularizado, y se puede proporcionar como un plásmido o vector. El ADN lineal es otra posibilidad. Las moléculas de ADN donante pueden comprender regiones que no se integran en el ADN celular, además de una o más secuencias de ADN donante que se integran en el ADN celular. El ADN es normalmente bicatenario, aunque en algunos casos se puede usar ADN monocatenario. El ADN donante contiene uno o más transgenes que codifican un aglutinante, por ejemplo, puede comprender un casete promotor:gen.

En el formato más simple, se puede usar ADN de plásmido circular bicatenario para dirigir la recombinación homóloga. Esto requiere regiones de ADN que flanquean los transgenes que son homólogos a la secuencia de ADN que flanquea el sitio de escisión en el ADN genómico. El ADN plasmídico bicatenario linealizado o el producto de PCR o los genes sintéticos podrían usarse para dirigir tanto la recombinación homóloga como las vías de reparación de la NHEJ. Como alternativa al ADN bicatenario, es posible utilizar ADN monocatenario para dirigir la recombinación homóloga [52]. Un enfoque común para generar ADN monocatenario es incluir un origen de replicación monocatenario de un bacteriófago filamentoso en el plásmido.

Se han utilizado virus de ADN monocatenario tal como el virus adenoasociado (AAV) para impulsar una recombinación homóloga eficiente donde se ha demostrado que la eficiencia mejora en varios órdenes de magnitud [53, 54]. Los sistemas tales como los sistemas AAV podrían usarse junto con la escisión dirigida por nucleasas para la construcción de grandes bibliotecas de aglutinantes. Los beneficios de ambos sistemas podrían aplicarse al direccionamiento de bibliotecas de aglutinantes. El límite de empaquetamiento de los vectores de AAV es de 4,7 kb, pero el uso de la digestión con nucleasa del ADN genómico diana reducirá esto permitiendo que se incorporen construcciones transgénicas más grandes.

Una molécula de ADN donante puede codificar un único aglutinante o múltiples aglutinantes. Opcionalmente, pueden codificarse múltiples subunidades de un aglutinante por molécula de ADN donante. En algunas realizaciones, el ADN donante codifica una subunidad de un aglutinante multimérico.

Promotores y elementos genéticos para la selección

La transcripción del aglutinante a partir del ADN donante codificante generalmente se logrará colocando la secuencia que codifica el aglutinante bajo el control de un promotor y, opcionalmente, uno o más elementos potenciadores para la transcripción. Un promotor (y opcionalmente otros elementos de control genético) puede incluirse en la propia molécula de ADN donante. Como alternativa, la secuencia que codifica el aglutinante puede carecer de un promotor en el ADN donante y, en cambio, puede colocarse en un enlace operable con un promotor

en el ADN celular, por ejemplo, un promotor endógeno o un promotor exógeno preintegrado, como resultado de su inserción en el sitio de integración creado por la nucleasa de sitio específico.

El ADN donante puede comprender además una o más secuencias codificantes adicionales, tales como elementos genéticos que permiten la selección de células que contienen o expresan el ADN donante. Al igual que con la secuencia que codifica el aglutinante, analizada anteriormente, dichos elementos pueden asociarse con un promotor en el ADN donante o pueden colocarse bajo el control de un promotor como resultado de la integración del ADN donante en un locus fijo. La última disposición proporciona un medio conveniente para seleccionar específicamente aquellas células que han integrado el ADN donante en el sitio deseado, ya que estas células deben expresar el elemento genético para la selección. Éste puede ser, por ejemplo, un gen que confiere resistencia a un agente de selección negativo, tal como blasticidina o puromicina. Se pueden aplicar una o más etapas de selección para eliminar las células no deseadas, tales como las células que carecen del ADN donante o que no han integrado el ADN donante en la posición correcta. Como alternativa, se puede permitir que estas células permanezcan mezcladas con los clones de la biblioteca.

La expresión de un aglutinante anclado a membrana podría usarse como una forma de marcador seleccionable. Por ejemplo, si se introduce una biblioteca de genes de anticuerpos, formateados como fusiones IgG o scFv-Fc, después, las células que expresan el anticuerpo pueden seleccionarse usando reactivos secundarios que reconocen la Fc expresada en la superficie usando los métodos descritos en el presente documento. Tras la transfección inicial con el ADN donante que codifica el transgén bajo el control de un promotor exógeno, se producirá la expresión transitoria (y la expresión de la superficie celular) del aglutinante y será necesario esperar a que desaparezca la expresión transitoria (para lograr una integración dirigida de, por ejemplo, 1-2 genes de anticuerpo/célula).

Como alternativa, una construcción que codifica un elemento de enlace a membrana (por ejemplo, el dominio Fc del presente ejemplo fusionado con el dominio transmembrana del receptor PDGF) podría preintegrarse antes de introducir la biblioteca de aglutinantes. Si este elemento de enlace a membrana carece de un promotor o está codificado dentro de un exón que está fuera del marco con el exón anterior, la expresión de la superficie se verá entonces comprometida. La integración dirigida de una molécula donante entrante puede corregir entonces este defecto (por ejemplo, mediante el direccionamiento de un promotor o un exón "en marco" en el intrón que está aguas arriba del elemento de enlace defectuoso). Si el "exón de corrección" de marco también codifica un aglutinante, se producirá entonces una fusión entre el aglutinante y el elemento de enlace a membrana, lo que dará como resultado la expresión en la superficie de ambos. Por lo tanto, la integración correctamente dirigida dará como resultado la expresión en marco del elemento de enlace a membrana solo o como parte de una fusión con el aglutinante entrante. Además, si la biblioteca entrante de aglutinantes carece de un elemento de enlace a membrana y estos están integrados incorrectamente, no se seleccionarán. Por lo tanto, la expresión del propio aglutinante en la superficie celular se puede utilizar para seleccionar la población de células con una integración dirigida correctamente.

Número de clones y diversidad de bibliotecas

Las bibliotecas de presentación de levadura de 10^7 - 10^{10} se han construido previamente y han demostrado que producen aglutinantes en ausencia de inmunización o preselección de la población [9, 55, 56, 57]. Muchas de las bibliotecas de presentación de mamíferos publicadas anteriormente utilizaban genes de anticuerpos derivados de donantes inmunizados o incluso linfocitos B específicos de antígeno enriquecidos, dadas las limitaciones del tamaño de la biblioteca y la variabilidad cuando se usan células de eucariotas superiores. Gracias a la eficacia del direccionamiento génico descrito en la presente invención, se pueden construir bibliotecas grandes y sin tratar en eucariotas superiores tales como las células de mamífero, que coinciden con las descritas para eucariotas más simples tal como levadura.

Después de la integración del ADN donante en el ADN celular, las células recombinantes resultantes se cultivan para permitir su replicación, generando un clon de células de cada célula recombinante producida inicialmente. Por lo tanto, cada clon se deriva de una célula original en la que se integró el ADN donante en un sitio de integración creado por la nucleasa de sitio específico. Los métodos de acuerdo con la presente invención se asocian con una alta eficiencia y alta fidelidad de la integración del ADN donante, y una biblioteca de acuerdo con la presente invención puede contener al menos 100, 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 o 10^{10} clones.

Usando la integración dirigida por nucleasas es posible dirigirse al 10 % o más de células de mamífero transfectadas. También es práctica para cultivar y transformar $>10^{10}$ células (por ejemplo, a partir de 5 litros de células que crecen a 2×10^6 células/ml). La transfección de tal gran número de células puede hacerse usando métodos estándar que incluyen transfección mediada por polietilimina como se describe en el presente documento. Además, hay métodos disponibles para la electroporación altamente eficiente de 10^{10} células en 5 minutos, por ejemplo, <http://www.maxcyte.com>. Por lo tanto, utilizando el enfoque de la presente invención, es posible crear bibliotecas de más de 10^9 clones.

Cuando la población de moléculas de ADN donante que se usa para crear la biblioteca contiene múltiples copias de la misma secuencia, se pueden obtener dos o más clones que contienen ADN que codifica el mismo aglutinante.

También puede darse el caso de que un clon pueda contener ADN donante que codifique más de un aglutinante diferente, por ejemplo, si hay más de una secuencia de reconocimiento para la nucleasa de sitio específico, como se detalla en otra parte en el presente documento. Por lo tanto, la diversidad de la biblioteca, en términos del número de diferentes aglutinantes codificados o expresados, puede ser diferente del número de clones obtenidos.

Los clones en la biblioteca contienen preferiblemente ADN donante que codifica uno o dos miembros del repertorio de aglutinantes y/o preferiblemente expresan solo uno o dos miembros del repertorio de aglutinantes. Un número limitado de diferentes aglutinantes por célula es una ventaja cuando se trata de identificar el clon y/o el ADN que codifica un aglutinante particular identificado al cribar la biblioteca contra una diana determinada. Esto es más simple cuando los clones codifican un solo miembro del repertorio de aglutinantes. Sin embargo, también es sencillo identificar el ADN codificante relevante para un aglutinante deseado si un clon seleccionado de una biblioteca codifica un pequeño número de aglutinantes diferentes, por ejemplo, un clon puede codificar dos miembros del repertorio de aglutinantes. Como se analiza en otra parte en el presente documento, los clones que codifican uno o dos aglutinantes son particularmente convenientes de generar al seleccionar una secuencia de reconocimiento para la nucleasa de sitio específico que tiene lugar una vez por copia cromosómica en un genoma diploide, ya que las células diploides contienen loci fijos duplicados, uno en cada cromosoma copia, y el ADN donante puede integrarse en uno o ambos loci fijos. Por lo tanto, los clones de la biblioteca pueden expresar cada uno solo uno o dos miembros del repertorio de aglutinantes.

Los aglutinantes mostrados en la superficie de las células de la biblioteca pueden ser idénticos a (tienen la misma secuencia de aminoácidos que) otros aglutinantes presentados en la misma célula. La biblioteca puede consistir en clones de células que presentan cada uno un solo miembro del repertorio de aglutinantes, o en clones que presentan una pluralidad de miembros del repertorio de aglutinantes por célula. Como alternativa, una biblioteca puede comprender algunos clones que muestran un solo miembro del repertorio de aglutinantes, y algunos clones que presentan una pluralidad de miembros (por ejemplo, dos) del repertorio de aglutinantes.

Por consiguiente, una biblioteca de acuerdo con la presente invención puede comprender clones que codifican más de un miembro del repertorio de aglutinantes, en donde el ADN donante se integra en loci fijos duplicados o múltiples loci fijos independientes.

Como se señaló anteriormente, es más fácil identificar el ADN codificante correspondiente para un aglutinante si el clon correspondiente expresa solo un aglutinante. Normalmente, una molécula de ADN donante codificará un único aglutinante. El aglutinante puede ser multimérico, de modo que una molécula de ADN donante incluya múltiples genes o marcos de lectura abiertos correspondientes a las diversas subunidades del aglutinante multimérico.

Una biblioteca de acuerdo con la presente invención puede codificar al menos 100, 10^3 , 10^4 , 10^5 o 10^6 , 10^7 , 10^8 , 10^9 o 10^{10} aglutinantes diferentes. Cuando los aglutinantes son multiméricos, la diversidad puede proporcionarse por una o más subunidades del aglutinante. Los aglutinantes multiméricos pueden combinar una o más subunidades variables con una o más subunidades constantes, donde las subunidades constantes son iguales (o de una diversidad más limitada) en todos los clones de la biblioteca. En la generación de bibliotecas de aglutinantes multiméricos, la diversidad combinatoria es posible cuando un primer repertorio de subunidades aglutinantes puede emparejarse con cualquiera de un segundo repertorio de subunidades aglutinantes.

Aglutinantes

Un "aglutinante" de acuerdo con la presente invención es una molécula de unión, que representa un compañero de unión específica para otra molécula. Los ejemplos típicos de compañeros de unión específicos son anticuerpo-antígeno y receptor-ligando.

El repertorio de aglutinantes codificados por una biblioteca usualmente compartirá una estructura común y tendrá una o más regiones de diversidad. Por lo tanto, la biblioteca permite la selección de un miembro de una clase estructural deseada de moléculas, tal como un péptido o una molécula de anticuerpo scFv. Por ejemplo, los aglutinantes pueden ser polipéptidos que comparten una estructura común y que tienen una o más regiones de diversidad de secuencias de aminoácidos.

Esto se puede ilustrar considerando un repertorio de moléculas de anticuerpo. Estas pueden ser moléculas de anticuerpos de una clase estructural común, por ejemplo, IgG, Fab, scFv-Fc o scFv, que se diferencian en una o más regiones de su secuencia. Las moléculas de anticuerpos normalmente tienen una variabilidad de secuencia en sus regiones determinantes de complementariedad (CDR), que son las regiones involucradas principalmente en el reconocimiento de antígenos. Un repertorio de aglutinantes en la presente invención puede ser un repertorio de moléculas de anticuerpos que difieren en una o más CDR, por ejemplo, puede haber diversidad de secuencias en las seis CDR, o en una o más CDR particulares tal como la CDR3 de cadena pesada y/o CDR3 de cadena ligera.

Las moléculas de anticuerpo y otros aglutinantes se describen con más detalle en otra parte en el presente documento. Sin embargo, el potencial de la presente invención se extiende más allá de la presentación de anticuerpos para incluir la presentación de bibliotecas de péptidos o proteínas modificadas, incluidos receptores,

ligandos, dominios de proteínas individuales y andamiajes proteicos alternativos [58, 59]. La integración de sitio específico dirigida por nucleasas se puede utilizar para crear bibliotecas de otros tipos de aglutinantes previamente modificados usando otros sistemas de presentación. Muchos de estos implican dominios de unión monoméricos tales como DARPins y lipocalinas, anticuerpos y adhirones [58, 59, 152]. La presentación en eucariotas, particularmente en células de mamífero, también abre la posibilidad de aislar y modificar aglutinantes o dianas que involucren dianas multiméricas más complejas. Por ejemplo, los receptores de linfocitos T (TCR) se expresan en linfocitos T y han evolucionado para reconocer el péptido presentado en complejo con moléculas MHC en células presentadoras de antígeno. Las bibliotecas que codifican y expresan un repertorio de TCR pueden generarse, y pueden cribarse para identificar la unión a complejos peptídicos de MHC como se describe adicionalmente en otra parte en el presente documento.

Para los aglutinantes multiméricos, el ADN donante que codifica el aglutinante se puede proporcionar como una o más moléculas de ADN. Por ejemplo, cuando los dominios VH y VL de anticuerpos individuales van a expresarse por separado, estos pueden codificarse en moléculas separadas del ADN donante. El ADN donante se integra en el ADN celular en múltiples sitios de integración, por ejemplo, el gen de aglutinante para el VH en un locus y el gen de aglutinante para el VL en un segundo locus. Los métodos para introducir ADN donante que codifican subunidades de aglutinante separadas se describen con más detalle en otra parte en el presente documento. Como alternativa, ambas subunidades o partes de un aglutinante multimérico pueden codificarse en la misma molécula de ADN donante que se integra en un locus fijo.

Un aglutinante puede ser una molécula de anticuerpo o una proteína sin anticuerpo que comprende un sitio de unión a antígeno. Se puede proporcionar un sitio de unión a antígeno mediante la disposición de bucles peptídicos en andamiajes de proteínas diferentes de anticuerpos, tales como fibronectina o citocromo B, etc., o aleatorizando o mutando restos de aminoácidos de un bucle dentro de un andamiaje de proteínas para conferir la unión a una diana deseada [60, 61, 62]. Los andamiajes para imitadores de anticuerpos se divulgan en el documento de patente WO/0034784, en el que los inventores describen proteínas (imitadores de anticuerpos) que incluyen un dominio de tipo III de fibronectina que tiene al menos un bucle aleatorio. Puede proporcionarse un andamiaje adecuado en el que injertar uno o más bucles peptídicos, por ejemplo, un conjunto de bucles de VH CDR de anticuerpo, por cualquier miembro de dominio de la superfamilia del gen de inmunoglobulina. El andamiaje puede ser una proteína humana o no humana.

El uso de sitios de unión a antígeno en andamiajes de proteínas diferentes de anticuerpo se ha revisado previamente [63]. Son típicas las proteínas que tienen un esqueleto estable y uno o más bucles variables, en los que la secuencia de aminoácidos del bucle o bucles se muta específica o aleatoriamente para crear un sitio de unión a antígeno que tiene para unirse al antígeno diana. Dichas proteínas incluyen los dominios de unión a IgG de la proteína A de *S. aureus*, transferrina, tetranequina, fibronectina (por ejemplo, 10^o dominio de fibronectina de tipo III) y lipocalinas. Otros enfoques incluyen péptidos restringidos pequeños, por ejemplo, basados en "knottin" y andamiajes de ciclótidos [64]. Dado su pequeño tamaño y complejidad, particularmente en relación con la correcta formación del enlace disulfuro, puede haber ventajas en el uso de células eucariotas para la selección de nuevos aglutinantes basados en estos andamiajes. Dadas las funciones comunes de estos péptidos en la naturaleza, las bibliotecas de aglutinantes basadas en estos andamiajes pueden ser ventajosas para generar pequeños aglutinantes de alta afinidad con una aplicación particular en el bloqueo de canales iónicos y proteasas.

Además de las secuencias de anticuerpos y/o un sitio de unión a antígeno, un aglutinante puede comprender otros aminoácidos, por ejemplo, formando un péptido o polipéptido, tal como un dominio plegado, o impartir a la molécula otra característica funcional además de la capacidad de unirse al antígeno. Un aglutinante puede llevar un marcador detectable, o puede conjugarse con una toxina o un resto o enzima diana (por ejemplo, a través de un enlace peptídico o enlazador). Por ejemplo, un aglutinante puede comprender un sitio catalítico (por ejemplo, en un dominio enzimático), así como un sitio de unión a antígeno, en el que el sitio de unión a antígeno se une al antígeno y, por lo tanto, dirige el sitio catalítico al antígeno. El sitio catalítico puede inhibir la función biológica del antígeno, por ejemplo, mediante escisión.

Moléculas de anticuerpo

Las moléculas de anticuerpo son aglutinantes preferidos. Las moléculas de anticuerpos pueden ser anticuerpos completos o inmunoglobulinas (Ig), que tienen cuatro cadenas polipeptídicas, dos cadenas pesadas idénticas y dos cadenas ligeras idénticas. Las cadenas pesada y ligera forman pares, teniendo cada una un par de dominios VH-VL que contiene un sitio de unión a antígeno. Las cadenas pesada y ligera también comprenden regiones constantes: CL de cadena ligera y CH1, CH2, CH3 y CH4 de cadena pesada, y a veces, CH4 (el quinto dominio CH4 está presente en IgM e IgE humanas). Las dos cadenas pesadas están unidas por puentes disulfuro en una región de bisagra flexible. Una molécula de anticuerpo puede comprender un dominio VH y/o VL.

El formato nativo más común de una molécula de anticuerpo es una IgG que es un heterotetrámero que consiste en dos cadenas pesadas idénticas y dos cadenas ligeras idénticas. Las cadenas pesada y ligera están constituidas por dominios modulares con una estructura secundaria conservada que consiste en una lámina beta antiparalela de cuatro cadenas y una lámina beta anti-paralela de tres cadenas, que se estabiliza por un solo enlace disulfuro. Cada

cadena de anticuerpo tiene un dominio variable N terminal (VH) y 3 dominios de inmunoglobulina "constante" relativamente conservados (CH1, CH2, CH3), mientras que las cadenas ligeras tienen un dominio variable N terminal (VL) y un dominio constante (CL). Los enlaces disulfuro estabilizan los dominios individuales y forman enlaces covalentes para unir las cuatro cadenas en un complejo estable. El VL y el CL de la cadena ligera se asocian con el VH y CH1 de la cadena pesada, y estos elementos pueden expresarse solos para formar un fragmento Fab. Los dominios CH2 y CH3 (también denominados "dominio Fc") se asocian con otro par CH2:CH3 para dar una molécula en forma de Y tetramérica con los dominios variables de las cadenas pesada y ligera en las puntas de la "Y". Los dominios CH2 y CH3 son responsables de las interacciones con las células efectoras y los componentes del complemento dentro del sistema inmunológico. Los anticuerpos recombinantes se han expresado previamente en formato IgG o como Fab (que consiste en un dímero de VH:CH1 y una cadena ligera). Además, la construcción artificial denominada Fv monocatenario (scFv) podría usarse y consistir en ADN que codifica fragmentos VH y VL fusionados genéticamente con ADN que codifica un enlazador flexible.

Los aglutinantes pueden ser moléculas de anticuerpos humanos. Por lo tanto, cuando están presentes dominios constantes, estos son preferiblemente dominios constantes humanos.

Los aglutinantes pueden ser fragmentos de anticuerpos o formatos de moléculas de anticuerpos más pequeños, tales como moléculas de anticuerpos monocatenarias. Por ejemplo, las moléculas de anticuerpo pueden ser moléculas scFv, que consisten en un dominio VH y un dominio VL unidos por un péptido enlazador. En la molécula scFv, los dominios VH y VL forman un par VH-VL en el que las regiones determinantes de complementariedad de VH y VL se unen para formar un sitio de unión a antígeno.

Otros fragmentos de anticuerpo que comprenden un sitio de unión a antígeno de anticuerpo incluyen, pero sin limitación, (i) el fragmento Fab que consiste en los dominios VL, VH, CL y CH1; (ii) el fragmento Fd que consiste en los dominios VH y CH1; (iii) el fragmento Fv que consiste en los dominios VL y VH de un único anticuerpo; (iv) el fragmento dAb [65, 66, 67], que consiste en un dominio VH o VL; (v) regiones CDR aisladas; (vi) fragmentos F(ab')₂, un fragmento bivalente que comprende dos fragmentos Fab enlazados (vii) scFv, en donde un dominio VH y un dominio VL están unidos por un enlazador peptídico que permite que los dos dominios se asocien para formar un sitio de unión a antígeno [68, 69]; (viii) dímeros Fv monocatenarios biespecíficos (documento PCT/US92/09965) y (ix) "diacuerpos", fragmentos multivalentes o multiespecíficos constituidos por fusión génica (documento WO94/13804; [70]). Las moléculas de Fv, scFv o diacuerpo pueden estabilizarse mediante la incorporación de puentes disulfuro que unen los dominios VH y VL [71].

Se han diseñado diversas moléculas de anticuerpo diferentes que incluyen uno o más sitios de unión a antígeno del anticuerpo, incluyendo, por ejemplo, Fab₂, Fab₃, diacuerpos, triacuerpos, tetracuerpos y minicuerpos (proteínas inmunes pequeñas). Se han descrito moléculas de anticuerpo y los métodos para su construcción y uso [72].

Otros ejemplos de fragmentos de unión son Fab', que difiere de los fragmentos Fab por la adición de unos pocos restos en el extremo carboxilo del dominio CH1 de la cadena pesada, incluyendo una o más cisteínas de la región bisagra del anticuerpo, y Fab'-SH, que es un fragmento Fab' en el que el resto o restos de cisteína de los dominios constantes llevan un grupo tiol libre.

Un dAb (anticuerpo de dominio) es un pequeño fragmento monomérico de unión a antígeno de un anticuerpo, concretamente, la región variable de una cadena pesada o ligera de anticuerpo. Los dAb de VH se producen naturalmente en camélidos (por ejemplo, camello, llama) y pueden producirse inmunizando un camélido con un antígeno diana, aislando linfocitos B específicos de antígeno y clonando directamente genes dAb de linfocitos B individuales. Los dAb también pueden producirse en cultivos celulares. Su pequeño tamaño, buena solubilidad y estabilidad a la temperatura los hace particularmente útiles fisiológicamente y adecuados para la selección y la maduración de la afinidad. Los dAb VH de camélido se están desarrollando para su uso terapéutico con el nombre de "nanobodies™".

Las moléculas de anticuerpos sintéticos pueden crearse por la expresión de genes generados por medio de oligonucleótidos sintetizados y ensamblados dentro de vectores de expresión adecuados, por ejemplo, como se describe por Knappik *et al.* [73] o Krebs *et al.* [74].

Los anticuerpos biespecíficos o bifuncionales forman una segunda generación de anticuerpos monoclonales en los que dos regiones variables diferentes se combinan en la misma molécula [75]. Su uso se ha demostrado tanto en el campo diagnóstico como en el campo de la terapia a partir de su capacidad para reclutar nuevas funciones efectoras o dirigirse a varias moléculas en la superficie de las células tumorales. Cuando se usan anticuerpos biespecíficos, estos pueden ser anticuerpos biespecíficos convencionales, que pueden fabricarse de diversas maneras [76], por ejemplo, preparados químicamente o a partir de hibridomas híbridos, o pueden ser cualquiera de los fragmentos de anticuerpos biespecíficos mencionados anteriormente. Estos anticuerpos pueden obtenerse por métodos químicos [77, 78] o por métodos somáticos [79, 80], pero de la misma manera y preferiblemente mediante técnicas de ingeniería genética que permiten forzar la heterodimerización y, por lo tanto, facilitan el proceso de purificación del anticuerpo buscado [81]. Los ejemplos de anticuerpos biespecíficos incluyen aquellos de la tecnología BiTE™ en los que los dominios de unión de dos anticuerpos con especificidad diferente pueden usarse y unirse directamente a

través de péptidos flexibles cortos. Esto combina dos anticuerpos en una cadena polipeptídica corta única. Los diacuerpos y scFv pueden construirse sin una región Fc, utilizando solo dominios variables, lo que reduce potencialmente los efectos de la reacción anti-idiotípica.

5 Los anticuerpos biespecíficos se pueden construir como IgG completa, como Fab'2 biespecífico, como Fab'PEG, como diacuerpos o también como scFv biespecífico. Además, dos anticuerpos biespecíficos se pueden vincular utilizando métodos de rutina conocidos en la técnica para formar anticuerpos tetravalentes.

10 Los diacuerpos biespecíficos, a diferencia de los anticuerpos completos biespecíficos, también pueden ser particularmente útiles. Se pueden seleccionar fácilmente diacuerpos (y muchos otros polipéptidos, tales como fragmentos de anticuerpos) de especificidades de unión apropiadas. Si un brazo del diacuerpo se mantiene constante, por ejemplo, con una especificidad dirigida contra un antígeno de interés, entonces se puede crear una biblioteca donde se varíe el otro brazo y se seleccione un anticuerpo de especificidad apropiada. Los anticuerpos completos biespecíficos se pueden producir mediante métodos de ingeniería alternativos como se describe en
15 Ridgeway *et al.*, 1996 [82].

Se puede usar una biblioteca de acuerdo con la invención para seleccionar una molécula de anticuerpo que se une a uno o más antígenos de interés. La selección de las bibliotecas se describe en detalle a continuación. Tras la selección, la molécula de anticuerpo se puede diseñar entonces en un formato diferente y/o contener características adicionales. Por ejemplo, la molécula de anticuerpo seleccionada se puede convertir a un formato diferente, tal como uno de los formatos de anticuerpo descritos anteriormente. Las moléculas de anticuerpo seleccionadas, y las moléculas de anticuerpo que comprenden las CDR VH y/o VL de las moléculas de anticuerpo seleccionadas, son un aspecto de la presente invención. Las moléculas de anticuerpo y su ácido nucleico codificante se pueden proporcionar en forma aislada.

25 Los fragmentos de anticuerpo pueden obtenerse a partir de una molécula de anticuerpo por métodos tales como la digestión por enzimas, por ejemplo, pepsina o papaína y/o por escisión de los puentes disulfuro por reducción química. De otra manera, los fragmentos de anticuerpo pueden obtenerse mediante técnicas de recombinación genética bien conocidas por los expertos en la técnica o también mediante síntesis peptídica mediante, por ejemplo, sintetizadores peptídicos automáticos, o mediante la síntesis y expresión de ácidos nucleicos.

30 Es posible tomar anticuerpos monoclonales y otros anticuerpos y usar técnicas de tecnología de ADN recombinante para producir otros anticuerpos o moléculas quiméricas que se unen al antígeno diana. Dichas técnicas pueden implicar la introducción de ADN que codifica la región variable de la inmunoglobulina, o las CDR, de un anticuerpo contra las regiones constantes, o regiones constantes más regiones marco, de una inmunoglobulina diferente. Véanse, por ejemplo, los documentos EP-A-184187, GB 2188638A o EP-A-239400, y una gran cantidad de bibliografía posterior.

40 Las moléculas de anticuerpo pueden seleccionarse a partir de una biblioteca y luego modificarse, por ejemplo, la semivida *in vivo* de una molécula de anticuerpo puede aumentarse por modificación química, por ejemplo, PEGilación, o por incorporación en un liposoma.

Fuentes de genes aglutinantes

45 La ruta tradicional para la generación de anticuerpos monoclonales utiliza el sistema inmune de animales de laboratorio como ratones y conejos para generar una agrupación de anticuerpos de alta afinidad que después se aíslan mediante el uso de tecnología de hibridoma. La presente invención proporciona una ruta alternativa para identificar anticuerpos que surgen de la inmunización. Los genes VH y VL podrían amplificarse a partir de linfocitos B de animales inmunizados y clonarse en un vector apropiado para su introducción en bibliotecas eucariotas, seguido de la selección de estas bibliotecas. La presentación de fagos y la presentación de ribosomas permite construir bibliotecas muy grandes ($>10^9$ clones) que permiten el aislamiento de anticuerpos humanos sin inmunización. La presente invención también podría usarse junto con tales métodos. Tras las rondas de selección de presentación de fagos, la población seleccionada de aglutinantes podría introducirse en células eucariotas mediante una integración dirigida por nucleasas como se describe en el presente documento. Esto permitirá el uso inicial de bibliotecas muy
50 grandes basadas en otros sistemas (por ejemplo, la presentación de fagos) para enriquecer una población de aglutinantes al tiempo que permite su cribado eficiente utilizando células eucariotas como se ha descrito anteriormente. Por lo tanto, la invención puede combinar las mejores características de la presentación de fagos y la presentación eucariota para proporcionar un sistema de alto rendimiento con cribado y clasificación cuantitativos.

60 Usando la presentación de fagos y la presentación de levadura, se ha demostrado previamente que también es posible generar aglutinantes sin recurrir a la inmunización, siempre que se utilicen bibliotecas de presentación de tamaño suficiente. Por ejemplo, se generaron múltiples aglutinantes a partir de una biblioteca de anticuerpos no inmunes de $>10^7$ clones [83]. Esto, a su vez, permite la generación de aglutinantes para dianas que son difíciles por las rutas de inmunización tradicionales, por ejemplo, la generación de anticuerpos contra "autoantígenos" o epitopos que se conservan entre especies. Por ejemplo, los aglutinantes de reactividad cruzada humano/ratón pueden enriquecerse mediante selección secuencial en versiones humanas y luego de ratón de la misma diana. Dado que
65

no es posible inmunizar específicamente a los seres humanos contra la mayoría de las dianas de interés, esta instalación es particularmente importante para permitir la generación de anticuerpos humanos que se prefieren para los enfoques terapéuticos.

5 En ejemplos de presentación de mamíferos hasta la fecha, donde los tamaños y la calidad de las bibliotecas eran limitados, los aglutinantes solo se han generado utilizando repertorios que se enriquecieron previamente para los aglutinantes, por ejemplo, de la inmunización o de la ingeniería de aglutinantes preexistentes. La capacidad de crear grandes bibliotecas en células eucariotas y particularmente en eucariotas superiores crea la posibilidad de aislar los aglutinantes directamente de estas bibliotecas, comenzando con los aglutinantes no inmunes o los aglutinantes que
10 no se han seleccionado previamente dentro de otro sistema. Con la presente invención es posible generar aglutinantes a partir de fuentes no inmunes. Esto, a su vez, abre las posibilidades para el uso de genes de aglutinantes de múltiples fuentes. Los genes de aglutinantes podrían provenir de la PCR de fuentes naturales tal como genes de anticuerpos. Los genes de aglutinantes también podrían volver a clonarse a partir de bibliotecas existentes, tales como las bibliotecas de presentación de fagos de anticuerpos, y clonarse en un vector donante adecuado para la integración dirigida por nucleasas en las células diana. Los aglutinantes pueden ser de origen completa o parcialmente sintético. Además, se describen diversos tipos de aglutinantes en otra parte en el presente documento, por ejemplo, los genes de aglutinantes podrían codificar anticuerpos o podrían codificar andamiajes alternativos [58, 59], péptidos o proteínas modificadas o dominios de proteínas.

20 *Presentación de aglutinantes*

Para proporcionar un repertorio de aglutinantes para el cribado contra una diana de interés, la biblioteca puede cultivarse para expresar los aglutinantes en forma secretada soluble o en forma transmembrana. Para la presentación en la superficie celular es necesario conservar el aglutinante expresado sobre la superficie celular que
25 lo codifica. Los aglutinantes pueden comprender o estar unidos a un anclaje de membrana, tal como un dominio transmembrana, para la presentación extracelular del aglutinante en la superficie celular. Esto puede implicar la fusión directa del aglutinante a una señal de localización de membrana, tal como una secuencia de reconocimiento de GPI o a un dominio transmembrana, tal como el dominio transmembrana del receptor PDGF [84]. La retención de aglutinantes en la superficie celular también puede hacerse indirectamente por asociación con otra molécula retenida en la superficie celular expresada dentro de la misma célula. Esta molécula asociada podría ser parte de un aglutinante heterodimérico, tal como la cadena pesada del anticuerpo unido en asociación con un compañero de cadena ligera que no está directamente unido.

Aunque la inmovilización de la superficie celular facilita la selección del aglutinante, en muchas aplicaciones es necesario preparar un aglutinante secretado libre de células. Será posible combinar la unión a la membrana y la secreción soluble utilizando un método de recaptura para unir los aglutinantes secretados a los receptores de la superficie celular. Un enfoque consiste en formatear la biblioteca de aglutinantes como moléculas secretadas que pueden asociarse con una molécula unida a la membrana expresada dentro de la misma célula que puede funcionar para capturar un aglutinante secretado. Por ejemplo, en el caso de anticuerpos o moléculas de aglutinante fusionadas a dominios Fc de anticuerpos, un Fc unido a la membrana puede "muestrear" moléculas de aglutinante secretadas que se expresan en la misma célula, dando como resultado la presentación de una fracción monomérica de las moléculas de aglutinante que se expresan mientras que el resto se secreta en forma bivalente (documento US 8.551.715). Una alternativa es usar un dominio de unión a IgG anclado, tal como la proteína A.

45 Otros métodos para conservar anticuerpos secretados con las células que los producen se revisan en Kumar *et al.* (2012) [85] e incluyen la encapsulación de células dentro de microgotas, captura asistida por matriz, visualización de la superficie de captura por afinidad (ACSD), tecnología de secreción y captura (SECANT) y "captura en frío" [85]. En los ejemplos dados para ACSD y SECANT [85], la biotinylación se usa para facilitar la inmovilización de la estreptavidina o un anticuerpo de captura en la superficie celular. La molécula capturada a su vez captura anticuerpos secretados. En el ejemplo de SECANT *in vivo* se produce la biotinylación de la molécula secretada. Usando la técnica de "captura en frío", el anticuerpo secretado puede detectarse en células productoras usando anticuerpos dirigidos a la molécula secretada. Se ha propuesto que esto se debe a la asociación del anticuerpo secretado con el glucocáliz de la célula [86]. Como alternativa, se ha sugerido que el producto secretado queda atrapado por anticuerpos de tinción en la superficie celular antes de someterse a endocitosis [87]. Los métodos anteriores se han usado para identificar clones de alta expresión dentro de una población, pero potencialmente podrían adaptarse para la identificación de la especificidad de unión, siempre que la asociación tenga suficiente longevidad en la superficie celular.

Incluso cuando el aglutinante está directamente unido a la superficie celular, es posible generar un producto soluble. Por ejemplo, el gen que codifica el aglutinante seleccionado puede recuperarse y clonarse en un vector de expresión que carece de la secuencia unida a la membrana. Como alternativa, y como se demuestra en los Ejemplos, se puede usar una construcción de expresión en la que el dominio transmembrana está codificado dentro de un exón flanqueado por sitios de recombinación, por ejemplo, sitios de reconocimiento de ROX para la recombinasa Dre [88]. En este ejemplo, el exón que codifica el dominio transmembrana puede eliminarse mediante transfección con un gen que codifica la recombinasa Dre para cambiar la expresión a una forma secretada. Como se demuestra en el presente documento, el anticuerpo secretado se produjo mediante este método sin la necesidad de una acción de la

recombinasa. Esto es presumiblemente como resultado de un corte y empalme alternativo [89].

Se puede usar cualquiera de los métodos anteriores u otros enfoques adecuados para garantizar que los aglutinantes expresados por los clones de una biblioteca se presenten en la superficie de sus células de expresión.

Cribado para identificar aglutinantes con respecto a una diana de interés

Como se indicó, la biblioteca de células eucariotas se puede usar en un método de cribado de un aglutinante que reconoce una diana. Tal método puede comprender:

- proporcionar una biblioteca como se describe en el presente documento,
- cultivar células de la biblioteca para expresar los aglutinantes,
- exponer los aglutinantes a la diana, permitiendo el reconocimiento de la diana por uno o más aglutinantes análogos, si están presentes, y
- detectar si la diana es reconocida por un aglutinante análogo.

Las selecciones podrían realizarse utilizando un intervalo de clases de moléculas diana, por ejemplo, proteínas, ácido nucleico, carbohidrato, lípido, moléculas pequeñas. La diana se puede proporcionar en forma soluble. La diana puede marcarse para facilitar la detección, por ejemplo, puede llevar una etiqueta fluorescente o puede estar biotinilado. Las células que expresan un aglutinante de diana específica pueden aislarse utilizando una molécula diana marcada directa o indirectamente, donde el aglutinante captura la molécula marcada. Por ejemplo, las células que se unen, mediante la interacción aglutinante:diana, a una diana marcada con fluorescencia se pueden detectar y clasificar por citometría de flujo o FACS para aislar las células deseadas. Las selecciones que implican citometría requieren moléculas diana que están directamente marcadas con fluorescencia o que están marcadas con moléculas que pueden detectarse con reactivos secundarios, por ejemplo, la diana biotinilada se puede añadir a las células y la unión a la superficie celular se puede detectar con estreptavidina marcada con fluorescencia tal como estreptavidina-ficoeritrina. Una posibilidad adicional es inmovilizar la molécula diana o los reactivos secundarios que se unen a la diana en una superficie sólida, tales como perlas magnéticas o perlas de agarosa, para permitir el enriquecimiento de las células que se unen a la diana. Por ejemplo, las células que se unen, a través de la interacción aglutinante:diana, a una diana biotinilada se pueden aislar en un sustrato recubierto con estreptavidina, por ejemplo, perlas recubiertas con estreptavidina.

En las bibliotecas de cribado, es preferible sobremuestrear, es decir, cribar más clones que el número de clones independientes presentes dentro de la biblioteca para garantizar una representación eficaz de la biblioteca. La identificación de los aglutinantes de bibliotecas muy grandes proporcionadas por la presente invención se podría hacer por clasificación de flujo, pero esto tardaría varios días, particularmente si se realiza un sobremuestreo de la biblioteca. Como alternativa, las selecciones iniciales podrían basarse en el uso de antígeno recuperable, por ejemplo, antígeno biotinilado recuperado en perlas magnéticas recubiertas con estreptavidina. Por lo tanto, las perlas magnéticas recubiertas con estreptavidina podrían usarse para capturar células que se han unido al antígeno biotinilado. La selección con perlas magnéticas podría usarse como el único método de selección o esto podría hacerse junto con la citometría de flujo, donde se puede lograr una mejor resolución, por ejemplo, diferenciando entre un clon con niveles de expresión más altos y uno con una afinidad más alta [56, 57].

La naturaleza *in vitro* de los enfoques de tecnología de presentación hace que sea posible controlar la selección de una manera que no es posible mediante la inmunización, por ejemplo, seleccionando en un estado conformacional particular de una diana [90, 91]. Las dianas podrían marcarse a través de modificación química (fluoresceína, biotina) o mediante fusión genética (por ejemplo, proteína fusionada con una etiqueta de epitopo, tal como una etiqueta FLAG u otro dominio de proteína o una proteína completa). La etiqueta podría ser ácido nucleico (por ejemplo, ADN, ARN o ácidos nucleicos no biológicos) donde la etiqueta está parcialmente fusionada para dirigirse al ácido nucleico o podría unirse químicamente a otro tipo de molécula, tal como una proteína. Esto podría ser a través de la conjugación química o a través de la unión enzimática [92]. El ácido nucleico también podría fusionarse con una diana a través de un proceso de traducción, tal como la presentación de ribosomas. La "etiqueta" puede ser otra modificación que tenga lugar dentro de la célula (por ejemplo, glucosilación, fosforilación, ubiquitinilación, alquilación, PASilación, SUMO-lación y otras descritas en la base de datos postraduccional (db-PTM) en <http://dbptm.mbc.nctu.edu.tw/statistics.php>) que puede detectarse a través de reactivos secundarios. Esto produciría aglutinantes que se unen a una proteína diana desconocida sobre la base de una modificación particular.

Las dianas podrían detectarse utilizando aglutinantes existentes que se unen a esa molécula diana, por ejemplo, anticuerpos específicos de diana. El uso de los aglutinantes existentes para la detección tendrá la ventaja adicional de identificar los aglutinantes dentro de la biblioteca de aglutinantes que reconocen un epitopo distinto del aglutinante utilizado para la detección. De esta manera, los pares de aglutinantes podrían identificarse para su uso en aplicaciones tales como ELISA en sándwich. Cuando sea posible, se preferirá una molécula diana purificada. Como alternativa, la diana puede mostrarse en la superficie de una población de células diana y los aglutinantes se muestran en la superficie de las células de la biblioteca, comprendiendo el método exponer los aglutinantes a la diana poniendo las células de la biblioteca en contacto con las células diana. La recuperación de las células que expresan la diana (por ejemplo, el uso de células biotiniladas que expresan la diana) permitirá el enriquecimiento de

las células que les expresan aglutinantes. Este enfoque sería útil cuando se involucren interacciones de baja afinidad, ya que existe la posibilidad de un fuerte efecto de avidéz.

La molécula diana también podrían ser dianas recombinantes no purificadas o nativas no purificadas siempre que haya una molécula de detección disponible para identificar la unión celular (como se ha descrito anteriormente). Además, la unión de las moléculas diana a la célula que expresa el aglutinante podría detectarse indirectamente a través de la asociación de la molécula diana a otra molécula que se está detectando, por ejemplo, un lisado celular que contiene una molécula marcada podría incubarse con una biblioteca de aglutinantes para identificar aglutinantes no solo para la molécula marcada, sino también aglutinantes con respecto a sus proteínas de compañero asociadas. Esto daría como resultado un panel de anticuerpos contra estos compañeros que podrían usarse para detectar o identificar el compañero (por ejemplo, usando espectrometría de masas). El fraccionamiento celular podría usarse para enriquecer las dianas de ubicaciones subcelulares particulares. Como alternativa, la biotilación diferencial de las fracciones superficiales o citoplasmáticas se podría usar junto con los reactivos de detección de estreptavidina para la presentación eucariota [93, 94]. El uso de preparaciones diana solubilizadas con detergente es un enfoque particularmente útil para proteínas de membrana intactas, tales como GPCR y canales iónicos que, por lo demás, son difíciles de preparar. La presencia de detergentes puede tener un efecto perjudicial sobre las células eucariotas que presentan los aglutinantes que requieren la recuperación de genes de aglutinantes sin crecimiento adicional de las células seleccionadas.

Después de la detección del reconocimiento de dianas por un aglutinante análogo, pueden recuperarse las células de un clon que contiene el ADN que codifica el aglutinante análogo. El ADN que codifica el aglutinante puede aislarse entonces (por ejemplo, identificarse o amplificarse) del clon recuperado, obteniendo de este modo el ADN que codifica un aglutinante que reconoce la diana.

Los aglutinantes y dianas a modo de ejemplo se detallan en otra parte en el presente documento. Un ejemplo clásico es una biblioteca de moléculas de anticuerpos, que pueden cribarse para determinar su unión a un antígeno diana de interés. Otros ejemplos incluyen el cribado de una biblioteca de TCR frente a un complejo de MHC diana:péptido o el cribado de una biblioteca de complejos de MHC:péptido frente a un TCR diana.

TCR:MHC y otras interacciones del receptor

Los receptores de linfocitos T (TCR) se expresan en linfocitos T y han evolucionado para reconocer el péptido presentado en complejo con moléculas MHC en células presentadoras de antígeno. Los TCR son heterodímeros que consisten en el 95 % de los casos de heterodímeros alfa y beta y en el 5 % de los casos de heterodímeros gamma y delta. Ambas unidades monoméricas tienen un dominio de inmunoglobulina N terminal que tiene 3 regiones determinantes de complementariedad variables (CDR) implicadas en la conducción de la interacción con la diana. El TCR funcional está presente dentro de un complejo de otras subunidades y la señalización se mejora mediante la coestimulación con moléculas CD4 y CD8 (específicas para moléculas MHC de clase I y clase II, respectivamente). En las células presentadoras de antígenos, las proteínas se procesan y se presentan en la superficie celular en complejo con moléculas MHC que son parte de un complejo proteico multimérico. Los TCR que reconocen los péptidos que se originan en los "mismos" se eliminan durante el desarrollo y el sistema está preparado para el reconocimiento de los péptidos extraños presentados en las células presentadoras de antígenos para efectuar una respuesta inmune. El resultado del reconocimiento de un complejo de péptido:MHC depende de la identidad del linfocito T y la afinidad de esa interacción.

Sería valioso identificar los genes que codifican TCR o complejos de MHC:péptido que impulsan las interacciones involucradas en afecciones patológicas, por ejemplo, como ocurre en una enfermedad autoinmune. En el caso de una enfermedad autoinmune, la identificación de compañeros de interacción, por ejemplo, un TCR que impulsa una afección patógena podría allanar el camino ya sea para bloquear específicamente dichas interacciones o eliminar las células citotóxicas ofensoras. Sería deseable modificar TCR para una unión alterada, por ejemplo, mayor afinidad con las dianas de interés, por ejemplo, para redirigir los linfocitos T en el cáncer o mejorar el efecto de los linfocitos T existentes [95]. Como alternativa, el comportamiento de los linfocitos T reguladores o supresores podría alterarse como modalidad terapéutica, por ejemplo, para dirigir o mejorar la inmunoterapia del cáncer mediante la introducción de TCR específicos en los linfocitos T o mediante el uso de una proteína TCR expresada como entidades terapéuticas [96].

La presentación de bibliotecas de TCR en la superficie de células de levadura y células de mamífero se ha demostrado previamente. En el caso de las células de levadura fue necesario modificar el TCR y presentarlo en un formato monocatenario. Dado que la afinidad de interacción entre el TCR y el complejo péptido:MHC es baja, el componente soluble (por ejemplo, péptido:MHC en este caso) normalmente se presenta en un formato multimérico. La especificidad de TCR se ha modificado para péptidos en complejo con MHC de clase I [97] y MHC de clase II [98]. Los TCR también se han expresado en la superficie de linfocitos T mutantes de ratón (que carecen de cadenas alfa y beta de TCR) y se han aislado TCR variantes con propiedades de unión mejoradas [99]. Por ejemplo, Chervin *et al.* introdujeron TCR por infección retroviral y se generó un tamaño de biblioteca eficaz de 10^4 clones [100]. Usando la integración dirigida por nucleasas de los aglutinantes como se propone aquí, se podría adoptar un enfoque similar para los linfocitos T de ingeniería. Además de seleccionar TCR con propiedades de reconocimiento

alteradas, las bibliotecas de presentación podrían usarse para cribar bibliotecas de péptidos o de variantes de MHC para el reconocimiento por TCR. Por ejemplo, los complejos de péptido:MHC se han presentado en células de insecto y se han utilizado para hacer un mapa de epítomos de TCR presentados en formato multimérico [101].

5 Como se señaló, los métodos de cribado pueden implicar presentar el repertorio de aglutinantes en la superficie celular y sondear con una diana presentada como una molécula soluble, que puede ser una diana multimérica. Una alternativa, que puede ser especialmente útil con dianas multiméricas, es detectar directamente las interacciones célula:célula, donde el aglutinante y la diana se presentan en la superficie de diferentes células. Por ejemplo, si la activación de un TCR de interés condujo a la expresión de un gen indicador, esto podría usarse para identificar péptidos activadores o moléculas MHC activadas presentadas dentro de una biblioteca de péptidos:MHC. En este ejemplo particular, la célula indicadora no codifica el miembro de la biblioteca, pero podría utilizarse para identificar la célula que lo codifica. El enfoque podría potencialmente extenderse a un enfoque de "biblioteca frente a biblioteca". Por ejemplo, extendiendo el ejemplo descrito anteriormente, una biblioteca de TCR podría cribarse frente a una biblioteca de péptidos:MHC. Más ampliamente, el ejemplo de cribado de una biblioteca de aglutinantes presentados en una superficie celular utilizando un compañero de unión en otra célula podría extenderse a otros tipos de interacciones célula:célula, por ejemplo, la identificación de aglutinantes que inhiben o activan la señalización dentro de las rutas Notch o Wnt. Por lo tanto, la presente invención podría usarse en un sistema alternativo de cribado basado en células que incluye sistemas de reconocimiento basados en interacciones célula:célula.

20 Como ejemplo, los receptores de antígenos quiméricos ("CARS") representan una fusión entre un dominio de unión a anticuerpo (normalmente formateado como scFv) y un dominio de señalización. Estos se han introducido en los linfocitos T con el objetivo de redirigir el linfocito T *in vivo* para atacar a las células tumorales a través del reconocimiento de anticuerpos y la unión a antígenos específicos del tumor. Una serie de factores diferentes podrían afectar el éxito de esta estrategia, incluida la combinación de especificidad del anticuerpo, formato, afinidad del anticuerpo, longitud del enlazador, módulo de señalización fusionado, nivel de expresión en los linfocitos T, subtipo de linfocitos T e interacción del CAR con otras moléculas de señalización [102, 103]. La capacidad de crear grandes bibliotecas de CAR en linfocitos T primarios que incorporan individuales o combinaciones de las variables anteriores permitiría una búsqueda funcional para una construcción de CAR eficaz y óptima. Esta "búsqueda" funcional podría realizarse *in vitro* o *in vivo*. Por ejemplo, Alonso-Camino (2009) fusionó un CEA que reconoce scFv a la cadena ζ del complejo TCR:CD3 e introdujo esta construcción genética en una línea de células Jurkat humanas [104]. Tras la interacción con el CEA presente en células HeLa o en células tumorales, mostraron una regulación positiva del marcador de activación de linfocitos T CD69. Este enfoque podría usarse para identificar construcciones de fusión de CAR con propiedades de activación o inhibición apropiadas usando células cultivadas o primarias.

35 Avanzando una etapa más, la funcionalidad de las construcciones CAR podría evaluarse *in vivo*. Por ejemplo, una biblioteca de CAR construida en linfocitos T primarios de ratón podría introducirse en ratones portadores de tumor para identificar clones de linfocitos T estimulados para proliferar a través del encuentro con el tumor. Si es necesario, esta biblioteca de linfocitos T podría preseleccionarse basándose en la especificidad de unión a antígeno utilizando los métodos descritos anteriormente. En cualquier caso, la biblioteca entrante de aglutinantes podría usarse para reemplazar una molécula aglutinante existente (por ejemplo, MHC o TCR o dominio variable de anticuerpo).

Cribas de fenotipo

45 Aquí se describen diversos métodos para la selección de aglutinantes que modifican la señalización celular y el comportamiento celular.

Una biblioteca puede cribarse para detectar un fenotipo celular alterado como resultado de la acción del aglutinante en su diana.

50 Los anticuerpos que modifican la señalización celular mediante la unión a ligandos o receptores tienen un historial probado en el desarrollo de fármacos y la demanda de dichos anticuerpos terapéuticos continúan creciendo. Tales anticuerpos y otras clases de aglutinantes funcionales también tienen potencial para controlar el comportamiento celular *in vivo* e *in vitro*. Sin embargo, la capacidad de controlar y dirigir el comportamiento celular se basa en la disponibilidad de ligandos naturales que controlan vías de señalización específicas. Desafortunadamente, muchos ligandos naturales, tales como los que controlan la diferenciación de células madre (por ejemplo, miembros de las superfamilias FGF, TGF-beta, Wnt y Notch) a menudo presentan interacciones promiscuas y tienen una disponibilidad limitada debido a sus malos perfiles de expresión/estabilidad. Con su exquisita especificidad, los anticuerpos tienen un gran potencial para controlar el comportamiento celular.

60 La identificación de anticuerpos funcionales que modifican la señalización celular ha sido históricamente laboriosa e implica la recogida de clones, la expresión de anticuerpos, la caracterización según la secuencia y las propiedades de unión, la conversión en sistemas de expresión de mamífero y la adición a ensayos basados en células funcionales. El enfoque de presentación eucariota descrito en el presente documento reducirá este esfuerzo, pero todavía hay un requisito para la producción de anticuerpos y la adición a un cultivo de células indicadoras separadas. Por lo tanto, una alternativa preferida puede ser cribar directamente bibliotecas de aglutinantes que se

expresan en células eucariotas para determinar el efecto de la unión en la señalización celular o el comportamiento celular utilizando la propia célula de producción como una célula indicadora. Tras la introducción de genes de anticuerpos, se pueden identificar los clones dentro de la población resultante de células que muestran alteración en la expresión del gen indicador o fenotipos alterados.

Una serie de publicaciones recientes han descrito la construcción de bibliotecas de anticuerpos mediante la clonación de repertorios de genes de anticuerpos en células indicadoras [47, 105, 106]. Estos sistemas combinan la expresión y el informe dentro de una célula, y normalmente introducen una población de anticuerpos seleccionados contra una diana predefinida (por ejemplo, usando la presentación de fagos).

Se puede introducir una población de genes de anticuerpos en células indicadoras para producir una biblioteca mediante los métodos descritos en el presente documento, y se pueden identificar los clones dentro de la población con una alteración dirigida por anticuerpos en el fenotipo (por ejemplo, la expresión o supervivencia génica alteradas). Para que esta selección fenotípicamente dirigida funcione, existe el requisito de conservar un enlace entre el gen del anticuerpo presente dentro de la célula de expresión (genotipo) y la consecuencia de la expresión del anticuerpo (fenotipo). Esto se ha logrado previamente ya sea mediante la unión del anticuerpo a la superficie celular [47] como se describe para la presentación del anticuerpo o a través del uso de un medio semisólido para conservar anticuerpos secretados en la proximidad de las células productoras [105]. Como alternativa, los anticuerpos y otros aglutinantes pueden conservarse dentro de la célula [107]. Los aglutinantes conservados en la superficie celular o en el medio circundante pueden interactuar con un receptor endógeno o exógeno en la superficie celular que causa la activación del receptor. Esto, a su vez, puede causar un cambio en la expresión de un gen indicador o un cambio en el fenotipo de la célula. Como alternativa, el anticuerpo puede bloquear el receptor o el ligando para reducir la activación del receptor. El gen que codifica el aglutinante que causa el comportamiento celular modificado se puede recuperar entonces para la producción o la modificación adicional.

En una alternativa a este enfoque "dirigido a la diana", es posible introducir una población de anticuerpos "sin tratar" que no se ha preseleccionado para una diana particular [108]. El sistema de notificación celular se utiliza para identificar a los miembros de la población con comportamiento alterado. Dado que no hay un conocimiento previo de la diana, este enfoque no dirigido tiene un requisito particular para un gran repertorio de anticuerpos, ya que no es posible el enriquecimiento previo de la población de anticuerpos en la diana. Este enfoque se beneficiará del uso de la integración de transgén dirigida por nucleasas como se describe en la presente invención.

El enfoque de "selección funcional" podría usarse en otras aplicaciones que implican bibliotecas en células eucariotas, particularmente eucariotas superiores tales como células de mamífero. El anticuerpo podría fusionarse con un dominio de señalización de tal forma que la unión a la diana provoca la activación del receptor. Kawahara *et al.* han construido receptores quiméricos donde una fluoresceína que se dirige a scFv extracelular se fusionó con un dominio espaciador (el dominio D2 del receptor Epo) y diversos dominios del receptor de citocinas intracelulares que incluyen el receptor de trombopoyetina (Tpo), el receptor de eritropoyetina (Epo), gp130, receptor IL-2 y el receptor EGF [109, 110, 111]. Estos se introdujeron en una línea celular proB dependiente de IL-3 (BaF3) [27], donde se demostró que los receptores quiméricos muestran una activación dependiente de antígeno del receptor quimérico que conduce a un crecimiento independiente de IL-3. Este mismo enfoque se usó en experimentos modelo para demostrar la quimioatracción mediada por antígeno de las células BaF3 [110]. El enfoque se extendió más allá de las células de cultivo estable a las células primarias ilustradas por la supervivencia y el crecimiento de células madre hematopoyéticas sensibles a Tpo [112] o linfocitos T primarios dependientes de IL2, donde la estimulación normal por Tpo e IL-2 respectivamente se reemplazó por la estimulación dirigida por fluoresceína de receptores quiméricos scFv. Por lo tanto, un sistema basado en anticuerpo quimérico-quimeras de receptor se puede usar para dirigir la expresión de genes dependientes de la diana o los cambios fenotípicos en células indicadoras primarias o estables. Esta capacidad podría usarse para identificar los aglutinantes fusionados que impulsan una respuesta de señalización o aglutinantes que inhiben la respuesta.

En una modificación del enfoque anterior, los dominios VH y VL separados de un anticuerpo anti-lisozima se fusionaron con el dominio intracelular Epo [113]. Las células crecieron en respuesta a la adición de lisozima que indica una dimerización o estabilización inducida por antígeno de los compañeros de fusión VH y VL separados. Por lo tanto, tres componentes de interacción se unen para determinar una respuesta óptima en este sistema.

Aunque se describe aquí con referencia a moléculas de anticuerpo, los métodos anteriores también pueden adaptarse y realizarse con bibliotecas de otros aglutinantes.

La complementación de los fragmentos de proteínas representa un sistema alternativo para estudiar y seleccionar interacciones proteína:proteína en células de mamífero [114, 115]. Esto implica restaurar la función de las proteínas indicadoras divididas a través de las interacciones proteína:proteína. Las proteínas indicadoras que se han usado incluyen ubiquitina, inteína DNAE, beta-galactosidasa, dihidrofolato reductasa, GFP, luciferasa de luciérnaga, beta-lactamasa, TEV proteasa. Por ejemplo, un ejemplo reciente de este enfoque es el enfoque de híbrido de la membrana 2 de mamífero (MaMTH) donde la asociación de una fusión de proteína cebo:ubiquitina dividida:factor de transcripción con una proteína asociada:ubiquitina dividida restaura el reconocimiento de ubiquitina y libera el factor de transcripción para realizar la expresión del gen indicador [116]. Nuevamente, los aglutinantes que interfieren o

mejoran esta interacción podrían identificarse a través de la señalización perturbada.

Recuperación y reformateo de aglutinantes y ADN codificante

5 Después de la selección de un aglutinante o clon de interés de la biblioteca, un siguiente paso común será aislar (por ejemplo, identificar o amplificar) el ADN que codifica el aglutinante. Opcionalmente, puede desearse modificar el ácido nucleico que codifica el aglutinante, por ejemplo, para reestructurar el aglutinante y/o insertar la secuencia de codificación en un vector diferente.

10 Cuando el aglutinante es una molécula de anticuerpo, un método puede comprender aislar el ADN que codifica la molécula de anticuerpo de las células de un clon, amplificar el ADN que codifica al menos una región variable del anticuerpo, preferiblemente el dominio tanto VH como VL, e insertar el ADN en un vector para proporcionar un vector que codifica la molécula de anticuerpo. Una molécula de anticuerpo multimérico que lleva un dominio constante puede convertirse en una molécula de anticuerpo monocatenaria para la expresión en una forma secretada soluble.

15 Los anticuerpos pueden presentarse en diferentes formatos, pero sea cual sea el formato en el que se seleccione un anticuerpo, una vez que se aísla el gen del anticuerpo, es posible reconfigurarlo en varios formatos diferentes. Una vez que se han aislado los dominios VH o VL, se pueden volver a clonar en vectores de expresión que incluyen los dominios asociados requeridos (véase el Ejemplo 1 que muestra un casete de expresión de IgG de promotor dual).

20 Una etapa de reformateo puede comprender reformatear los aglutinantes compuestos por un par de subunidades (por ejemplo, moléculas scFv), en un formato de aglutinante molecular diferente (por ejemplo, Ig o Fab) en el que se mantiene el emparejamiento original de las subunidades. Dichos métodos se describen con más detalle en otra parte en el presente documento y se pueden usar para el reformateo monoclonal, oligoclonal o policlonal. El método se puede utilizar para convertir "en masa" una población de salida completa de cualquiera de las tecnologías de presentación comúnmente utilizadas, incluida la presentación de fagos, levaduras o ribosomas.

25

Presentación de scFv en la superficie de células de mamífero fusionadas a dominios Fc.

30 Aunque muchas bibliotecas de presentación de fagos de anticuerpos están formateadas para mostrar scFv, los sistemas de presentación eucariotas permitirán la presentación en formato Fab o IgG. Para aprovechar al máximo el potencial de expresión de IgG/Fab, en particular cuando se usan scFv de otros sistemas de presentación, será necesario tomar dominios VH y VL enlazados seleccionados dentro de un sistema de expresión bacteriano y expresarlos dentro de un sistema eucariota fusionado a dominios constantes apropiados. Aquí se describe un método para convertir poblaciones de scFv en inmunoglobulina (Ig) o un fragmento, formato de unión a antígeno (Fab) de tal manera que se mantengan los pares de cadenas VH y VL originales. En la presente invención, la conversión es posible utilizando clones individuales, mezclas oligoclonales o poblaciones completas formateadas como scFv mientras se conserva el emparejamiento original de las cadenas VH y VL. El método avanza a través de la generación de un ADN de "mini-círculo" no replicante intermedio que trae un nuevo fragmento de ADN "de relleno". El ADN circular se linealiza (por ejemplo, mediante digestión de restricción o PCR) que altera la posición relativa de los fragmentos VH y VL originales y coloca el ADN "de relleno" entre ellos. Tras la linealización, el producto puede clonarse en un vector de elección, por ejemplo, un vector de expresión de mamífero. De esta manera, todos los elementos, aparte de VH y VL, pueden reemplazarse. Los elementos para la expresión bacteriana pueden reemplazarse con elementos para la expresión de mamífero y fusión a compañeros alternativos. El proceso de conversión completo solo requiere una única etapa de transformación de las bacterias *E. coli* para generar una población de colonias bacterianas, cada una de las cuales alberga un plásmido que codifica un anticuerpo recombinante con formato Ig o Fab único. Extendiéndose más allá de la conversión de scFv en IgG/Fab, el método puede emplearse para reformatear cualquiera de los dos elementos de ADN unidos para clonarlos en un vector, de manera que, después del reformateo, cada elemento de ADN está rodeado por diferentes características de control de ADN mientras se mantiene el emparejamiento original. Se ha descrito un método anterior en el que se utilizan 2 etapas de clonación secuencial [117] para reemplazar estos elementos en contraste con el presente método, que se realiza a través de un intermedio circular no replicante intermedio.

55 Un método para reestructurar un aglutinante, o una población de aglutinantes, puede comprender convertir scFv en Ig o un fragmento del mismo, por ejemplo, Fab. El método puede comprender convertir el ácido nucleico que codifica scFv en ADN que codifica una inmunoglobulina (Ig) o un fragmento del mismo, tal como el formato Fab, de tal manera que se mantengan los pares de cadenas VH y VL variables originales. Preferiblemente, la conversión se realiza a través del intermedio de ADN circular que puede ser un ADN de "mini-círculo" no replicante. El método requiere una transformación única de *E. coli* para la generación directa de transformantes bacterianos que albergan plásmidos que codifican ADN de Ig o Fab.

60

El método se puede usar para el reformateo de clones monoclonal, oligoclonal o policlonal. El método se puede utilizar para convertir "en masa" una población de salida completa de cualquiera de las tecnologías de presentación comúnmente utilizadas, incluida la presentación de fagos, levaduras o ribosomas.

65

Más generalmente, este aspecto de la invención se refiere a un método que permite reformatear cualquiera de los dos elementos de ADN unidos en un vector donde los elementos de ADN se clonan bajo el control de promotores separados, o se separan por elementos de control alternativos, pero manteniendo el par de ADN original.

5 Los ejemplos 7 y 14 describen dichos métodos con más detalle.

El aislamiento, y la reestructuración opcional, de los aglutinantes que codifican el ADN pueden seguirse mediante la introducción de ese ADN en células adicionales para crear una biblioteca de derivados como se describe en otra parte en el presente documento, o el ADN que codifica uno o más aglutinantes particulares de interés se puede introducir en una célula hospedadora para la expresión. La célula hospedadora puede ser de un tipo diferente en comparación con las células de la biblioteca de la que se obtuvo. Generalmente el ADN se proporcionará en un vector. El ADN introducido en la célula hospedadora puede integrarse en el ADN celular de la célula hospedadora. Las células hospedadoras que expresan la molécula de anticuerpo soluble secretada pueden entonces seleccionarse.

15 Las células hospedadoras que codifican uno o más aglutinantes pueden proporcionarse en medio de cultivo y cultivarse para expresar el uno o más aglutinantes.

Bibliotecas de derivados

20 Después de la producción de una biblioteca por el método de la invención, uno o más clones de biblioteca pueden seleccionarse y usarse para producir una biblioteca adicional de segunda generación. Cuando se ha generado una biblioteca mediante la introducción de ADN en células eucariotas como se describe en el presente documento, la biblioteca puede cultivarse para expresar los aglutinantes, y se pueden recuperar uno o más clones que expresan aglutinantes de interés, por ejemplo, seleccionando aglutinantes contra una diana como se describe en otro lugar en el presente documento. Estos clones se pueden usar posteriormente para generar una biblioteca de derivados que contiene ADN que codifica un segundo repertorio de aglutinantes.

30 Para generar la biblioteca de derivados, el ADN donante de uno o más clones recuperados se muta para proporcionar el segundo repertorio de aglutinantes. Las mutaciones pueden ser la adición, sustitución o eliminación de uno o más nucleótidos. Cuando el aglutinante es un polipéptido, la mutación será cambiar la secuencia del aglutinante codificado mediante la adición, sustitución o eliminación de uno o más aminoácidos. La mutación puede centrarse en una o más regiones, tal como una o más CDR de una molécula de anticuerpo, proporcionando un repertorio de aglutinantes de una clase estructural común que difieren en una o más regiones de diversidad, como se describe en otra parte en el presente documento.

40 La generación de la biblioteca de derivados puede comprender aislar el ADN donante de uno o más clones recuperados, introducir la mutación en el ADN para proporcionar una población derivada de moléculas de ADN donante que codifican un segundo repertorio de aglutinantes, e introducir la población de derivados de moléculas de ADN donante en células para crear una biblioteca de derivados de células que contienen ADN que codifica el segundo repertorio de aglutinantes.

45 El aislamiento del ADN donante puede implicar obtener y/o identificar el ADN del clon. Dichos métodos pueden incluir la amplificación del ADN que codifica un aglutinante de un clon recuperado, por ejemplo, mediante PCR e introduciendo mutaciones. El ADN puede secuenciarse y el ADN mutado puede sintetizarse.

50 Como alternativa, la mutación puede introducirse en el ADN donante en uno o más clones recuperados induciendo la mutación del ADN dentro de los clones. Por lo tanto, la biblioteca de derivados puede crearse a partir de uno o más clones sin requerir el aislamiento del ADN, por ejemplo, a través de la mutación endógena en células DT40 aviares.

55 La presentación de anticuerpos se presta especialmente bien para la creación de bibliotecas de derivados. Una vez que se aíslan los genes de anticuerpos, es posible usar diversos enfoques de mutagénesis (por ejemplo, PCR propensa a errores, mutagénesis dirigida por oligonucleótidos, redistribución de cadenas) para crear bibliotecas de presentación de clones relacionados entre los cuales se pueden seleccionar variantes mejoradas. Por ejemplo, con la redistribución de cadenas el ADN que codifica la población del clon VH seleccionado, la mezcla o población oligoclonal se puede subclonar en un vector que codifica un formato de anticuerpo adecuado y que codifica un repertorio de cadenas VL adecuadamente formateado [118]. Como alternativa, y utilizando de nuevo el ejemplo de los VH, el clon, la oligomezcla o población de VH podría introducirse en una población de células eucariotas que codifican y expresan una población de compañeros de cadenas ligeras con el formato apropiado (por ejemplo, una cadena VL-CL para la asociación con una cadena pesada formateada con IgG o Fab). La población VH podría surgir a partir de cualquiera de las fuentes analizadas anteriormente, incluyendo los linfocitos B de animales inmunizados o genes scFv de poblaciones de fagos seleccionadas. En el último ejemplo, la clonación de VH seleccionados en un repertorio de cadenas ligeras podría combinar la redistribución de cadenas y el reformateo (por ejemplo, en formato IgG) en una sola etapa.

Una ventaja particular de la presentación en células eucariotas es la capacidad de controlar la rigurosidad de la etapa de selección/cribado. Al reducir la concentración de antígeno, las células que expresan los aglutinantes de mayor afinidad se pueden distinguir de los clones de menor afinidad dentro de la población. La visualización y la cuantificación del proceso de maduración de afinidad usando citometría de flujo es un beneficio importante de la presentación eucariota, ya que proporciona una indicación temprana de porcentajes positivos en bibliotecas sin tratar y permite una comparación directa entre la afinidad de los clones seleccionados y la población parental durante la clasificación. Después de la clasificación, la afinidad de los clones individuales se puede determinar mediante preincubación con un intervalo de concentraciones de antígenos y análisis en citometría de flujo o con un ensayo homogéneo de fluorescencia de tiempo resuelto (TRF) o usando resonancia de plasmón superficial (SPR) (Biacore).

Características y forma de la biblioteca

La presente invención permite la construcción de bibliotecas de células eucariotas que tienen muchas características ventajosas. La invención proporciona bibliotecas que tienen una o más de las siguientes características:
Diversidad. Una biblioteca puede codificar y/o expresar al menos 100, 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 o 10^9 aglutinantes diferentes.

Integración uniforme. Una biblioteca puede consistir en clones que contienen ADN donante integrado en un locus fijo, o en un número limitado de loci fijos en el ADN celular. Cada clon en la biblioteca, por lo tanto, contiene ADN donante en el locus fijo o al menos uno de los loci fijos. Preferiblemente, los clones contienen ADN donante integrado en uno o dos loci fijos en el ADN celular. Como se explica en otra parte en el presente documento, el sitio de integración está en una secuencia de reconocimiento para una nucleasa de sitio específico. La integración del ADN donante para producir ADN recombinante se describe en detalle en otra parte del presente documento y puede generar diferentes resultados dependiendo del número de sitios de integración. Cuando existe un único sitio de integración potencial en las células utilizadas para generar la biblioteca, la biblioteca será una biblioteca de clones que contienen ADN donante integrado en el locus fijo único. Todos los clones de la biblioteca, por lo tanto, contienen los genes de aglutinantes en la misma posición en el ADN celular. Como alternativa, cuando existen múltiples sitios de integración potenciales, la biblioteca puede ser una biblioteca de clones que contienen ADN donante integrado en múltiples y/o diferentes loci fijos. Preferiblemente, cada clon de una biblioteca contiene ADN donante integrado en un primer y/o segundo locus fijo. Por ejemplo, una biblioteca puede comprender clones en los que el ADN donante se integra en un primer locus fijo, clones en los que el ADN donante se integra en un segundo locus fijo, y los clones en los que el ADN donante se integra en el primer y segundo loci fijos. En realizaciones preferidas, solo hay uno o dos loci fijos en los clones en una biblioteca, aunque es posible integrar el ADN donante en múltiples loci si se desea para aplicaciones particulares. Por lo tanto, en algunas bibliotecas, cada clon puede contener ADN donante integrado en uno o más de varios loci fijos, por ejemplo, tres, cuatro, cinco o seis loci fijos.

Para bibliotecas que contienen subunidades de aglutinante integradas en sitios separados, los clones de la biblioteca pueden contener ADN que codifica una primera subunidad de aglutinante integrada en un primer locus fijo y ADN que codifica una segunda subunidad de aglutinante integrada en un segundo locus fijo, en donde los clones expresan aglutinantes multiméricos que comprenden la primera y segunda subunidades.

Transcripción uniforme. Los niveles relativos de transcripción de los aglutinantes entre diferentes clones de la biblioteca se mantienen dentro de límites controlados debido a que el ADN donante se integra en un número controlado de loci, y en el mismo locus en los diferentes clones (locus fijo). La transcripción relativamente uniforme de genes de aglutinantes conduce a niveles comparables de expresión de aglutinantes en o desde clones en una biblioteca. Los aglutinantes mostrados en la superficie de las células de la biblioteca pueden ser idénticos a (tienen la misma secuencia de aminoácidos que) otros aglutinantes presentados en la misma célula. La biblioteca puede consistir en clones de células que presentan cada uno un solo miembro del repertorio de aglutinantes, o en clones que presentan una pluralidad de miembros del repertorio de aglutinantes por célula. Como alternativa, una biblioteca puede comprender algunos clones que muestran un solo miembro del repertorio de aglutinantes, y algunos clones que presentan una pluralidad de miembros (por ejemplo, dos) del repertorio de aglutinantes. Preferiblemente, los clones de una biblioteca expresan uno o dos miembros del repertorio de aglutinantes.

Por ejemplo, una biblioteca de clones de células eucariotas de acuerdo con la presente invención puede expresar un repertorio de al menos 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 o 10^9 aglutinantes diferentes, por ejemplo, fragmentos de anticuerpos IgG, Fab, scFv o scFv-Fc, conteniendo cada célula ADN donante integrado en un locus fijo en el ADN celular. El ADN donante codifica el aglutinante y puede comprender además un elemento genético para la selección de células en las que el ADN donante se integra en el locus fijo. Las células de la biblioteca pueden contener ADN que codifica una nucleasa exógena de sitio específico.

Estas y otras características de las bibliotecas de acuerdo con la presente invención se describen adicionalmente en otra parte en el presente documento.

La presente invención se extiende a la biblioteca en forma pura, como una población de clones de biblioteca en ausencia de otras células eucariotas, o mezclada con otras células eucariotas. Otras células pueden ser células

eucariotas del mismo tipo (por ejemplo, la misma línea celular) o células diferentes. Se pueden obtener ventajas adicionales combinando dos o más bibliotecas de acuerdo con la presente invención, o combinando una biblioteca de acuerdo con la invención con una segunda biblioteca o una segunda población de células, ya sea para facilitar o ampliar el cribado o para otros usos como se describe en el presente documento o que serán evidentes para el experto en la técnica.

Una biblioteca de acuerdo con la invención, uno o más clones obtenidos de la biblioteca, o células hospedadoras en las que se ha introducido el ADN que codifica un aglutinante de la biblioteca, puede proporcionarse en un medio de cultivo celular. Las células pueden cultivarse y después concentrarse para formar un sedimento celular para un transporte o almacenamiento conveniente.

Las bibliotecas normalmente se proporcionarán *in vitro*. La biblioteca puede estar en un recipiente tal como un matraz de cultivo celular que contiene células de la biblioteca suspendidas en un medio de cultivo, o un recipiente que comprende un sedimento o suspensión concentrada de células eucariotas que comprenden la biblioteca. La biblioteca puede constituir al menos el 75 %, 80 %, 85 % o el 90 % de las células eucariotas en el recipiente.

Breve descripción de los dibujos

Las realizaciones de la invención se describirán ahora con más detalle, con referencia a los dibujos adjuntos, que son los siguientes:

Figura 1. Vector para la expresión de anticuerpos formateados con IgG.

- a. pDUAL D1.3, un vector de expresión de promotor dual para la secreción de IgG (en el esqueleto del vector "pCMV/myc/ER")
- b. pINT3-D1.3, un vector de expresión de promotor dual para la secreción de IgG (en el esqueleto del vector "pSF-CMV-f1-Pac1")
- c. Esqueleto del vector pCMV/myc/ER. El sitio ECoR1 precede al promotor CMV. Los sitios BstB1 y BstZ171 flanquean los sitios poli A de SV40.
- d. Esqueleto del vector pSF-CMV-f1-Pac1 (Oxford Genetics)
- e. gen sintético con exón que codifica la región transmembrana (TM) de PDGFR y exón que causa la secreción (sec). Las flechas sólidas representan los sitios de recombinación de Rox.

Figura 2. Secuencia de pD1 (SEQ ID NO: 1, 2, 3, 4 y 5): un casete de expresión del anticuerpo de promotor dual para la expresión superficial.
Características:

promotor pEF 13-1180
líder BM40 1193-1249
D1.3 VL humanizado 1250-1578
C kappa humano 1577-1891 BGH poli A 1916-2130
promotor CMV 2146-2734
Líder VH de ratón con intrón 32832-3414
D1.3 VH humanizado 3419-3769
CH1-CH3 de IgG2 humano optimizado

Figura 3. Construcción del plásmido donante de AAVS (pD2)

- a. Representación del locus AAVS humano. El Exón 1 y Exón 2 del locus AAVS (que codifica la proteína fosfatasa 1, subunidad reguladora 12C, PPP1R12C) se separan por un intrón de 4428 pb. El corte y empalme está en el marco "0", es decir, el corte y empalme se produce entre 2 codones intactos de cada exón. Las construcciones TALEN y CRISPR/Cas9 están disponibles para escindirse dentro de este intrón. Los bloques sombreados representan las regiones a la izquierda y derecha de este sitio de escisión que se utilizan dentro de construcciones de vectores para impulsar la recombinación homóloga en este locus (brazo de homología AAVS izquierdo "HA izquierdo" y brazo de homología AAVS derecho ("HA derecho").
- b. Representación del anticuerpo que codifica el plásmido donante pD2. Los brazos de homología izquierdo y derecho se muestran en los extremos de la representación de la construcción. Un gen blasticidina sintético está precedido por un aceptor de corte y empalme que crea una fusión "en marco" con el exón 1 de AAVS. También se muestra el casete de expresión de anticuerpos que consiste en una cadena ligera D1.3 y una cadena pesada D1.3 IgG2 impulsada por pEF y promotores de CMV respectivamente.
- c. Secuencia de la construcción donante pD1-huD1.3 (SEQ ID NO: 6, 7 y 8). Los brazos de homología AAVS se muestran subrayados y en negrita. Para mayor brevedad, el casete de anticuerpos (mostrado anteriormente en la Figura 2) no se muestra en detalle. Los sitios de restricción utilizados en el clon se muestran en negrita. La secuencia del esqueleto plasmídico se muestra al gen de resistencia a la ampicilina.

Figura 4. Expresión de IgG en la superficie celular.

5 a, c. El análisis se centró en las células viables utilizando dispersión directa y tinción en el canal FL3 (a, c). Se excluyeron las células positivas para la tinción en el canal FL3 (que representan células no viables que absorbieron 7-AAD). Todas las células se transfectaron con pD2-D1.3 en ausencia (a, b) o presencia (c, d) de las TALEN de AAVS y se tiñeron con anticuerpo anti-Fc.

10 Figura 5. Unión a anticuerpos del antígeno en la superficie celular. Las células viables se seleccionaron basándose en la dispersión directa y la oclusión de 7AAD (a). Las células se incubaron con lisozima de huevo de gallina marcada con fluorescencia (b).

Figura 6. Efecto de la escisión genómica dirigida por TALEN sobre la integración del gen de resistencia a la blasticidina. La figura muestra el número de colonias en las condiciones descritas.

15 Figura 7. Análisis de la integración del plásmido donante pD2-D1.3 en el locus AAVS. Después de la transfección, las células se seleccionaron en Blastidina y se preparó el ADN genómico. Las muestras 1-9 se beneficiaron de la adición de nucleasas TALE dirigidas por AAVS, las muestras 10-11 eran de clones que surgen de células transfectadas en ausencia de nucleasa TALE. Análisis de ADN genómico realizado por PCR como se describe en el texto. a. Verificación desde el extremo 5' del sitio de integración. B. Verificación desde el extremo 3' del sitio de integración.

20

Figura 8. Construcción del vector de expresión de scFv-Fc pD6

25 a. Los anticuerpos formateados como scFv se clonan en los sitios Nco1/Not1 para crear una fusión "en marco" con la región Fc humana de IgG2.
b. Secuencia de pD6 de los sitios Nco1 a Pme1 (SEQ ID NO: 9,10 y 11).

Figura 9. Selección de aglutinantes de la biblioteca scFv-Fc de la superficie celular (de poblaciones de fagos seleccionadas). Se muestra el análisis de citometría de flujo de células con:

30 a-c. Población sFv-Fc anti-CD229 integrada de 2 rondas de selección de presentación de fagos en CD229
d. f Población sFv-Fc anti-β-galactosidasa integrada de 1 ronda de selección de presentación de fagos en β-galactosidasa (células β-galR1)
e. Población sFv-Fc anti-p-galactosidasa integrada de 2 rondas de selección de presentación de fagos en β-galactosidasa
35 La muestra (a) muestra células sin teñir y el resto se tiñó con anti-Fc-ficoeritina humana (en FL2) y antígeno biotinilado apropiado 100 nM/estreptavidina-FITC (en FL1). Las células se analizaron después de 13 días (a, b, d, e). Los ejemplos c y f muestran las células teñidas después de 20 días y la región marcada muestra las células recogidas por citometría de flujo.
40 h. Las células β-galR1 seleccionadas por citometría de flujo (Figura 6f) se cultivaron durante 22 días y se volvieron a analizar para determinar la expresión de scFv-Fc y la unión a antígeno (utilizando un antígeno 100 nM). g. Se muestra el equivalente sin teñir.
j. Se muestran células β-galR1 sin clasificar de la población original (como en d) que se habían cultivado durante 42 días después de la transfección (j). Las células no marcadas de cada población se muestran con fines comparativos (g, i)

45

Figura 10. Presentación y clasificación de mamíferos de la biblioteca formateada por IgG.

50 Se seleccionó una población de anticuerpos en la β-galactosidasa utilizando 1 o 2 rondas de presentación de fagos, se reformatearon como IgG y se dirigieron a través la integración dirigida por nucleasas al locus AAVS de las células HEK293. Los paneles a, b muestran células derivadas de la población de fagos de la ronda 1, ya sea a, sin clasificar (después de un crecimiento de 38 días) o b, clasificadas por citometría de flujo y cultivadas durante 19 días. Los paneles c, d muestran células derivadas de la población de fagos de la ronda 2, ya sea c, sin clasificar (después de 38 días de crecimiento o d, clasificadas y cultivadas durante 19 días).

55 Figura 11. Construcción de una gran biblioteca de scFv-Fc sin tratamiento y selección de aglutinantes. Las células de la biblioteca de scFv-Fc sin tratamiento se tiñeron con antígeno biotinilado 500 nM y estreptavidina-FITC junto con el anticuerpo anti-Fc marcado con ficoeritina como anteriormente. La región muestra las células que se seleccionaron por clasificación de flujo. Las muestras se marcaron con biotinilaciones:

60 a. CD28
b. β-galactosidasa
c. Tiroglobulina
d. EphB4

65 Figura 12. Vector de direccionamiento para introducir un intrón que contiene sitios de "aterizaje múltiple" para la comparación de los métodos de integración.

- a. Plásmido de expresión de GFP intermedio (pD3)
 b. Vector de direccionamiento AAVS1_directed (pD4). El "sitio de aterrizaje" incorpora elementos para dirigir la integración que son FRT, lox 2272, meganucleasa I-Sce1 y GFP TALEN. Después de la integración de pD4 en el locus AAVS, están presentes múltiples sitios de recombinación o escisión de nucleasas dentro del genoma. El plásmido pD5 entrante (Figura 15) tiene brazos de homología izquierdo y derecho equivalentes a la secuencia presente en cada lado del "sitio de aterrizaje" para impulsar la inserción del anticuerpo por recombinación homóloga.

Figura 13. Secuencia de pD4 (SEQ ID NO: 12, 13, 14, 15, 16, 17 y 18).
 Las características de secuencia incluyen:

Homología izquierda de AAVS 19-822
 Sitio de FRT 832-879, Sitio de Lox 2272 884-917, sitio de meganucleasa I-Sce1 933-950
 Unión a TALEN izquierda I-Sce1 GFP 954-968, unión a TALEN derecha GFP 984-997
 T2A 1041-1103
 GFP 1104-1949
 promotor PGK 2178-2691
 Puromicina delta timidina quinasa 2706-4307
 loxP 4634-4667
 Homología derecha de AAVS 4692-5528

Figura 14. Verificación de la integración del intrón "sitio de aterrizaje múltiple" en el clon 6F. Después de la transfección, las células se seleccionaron en puromicina y se preparó el ADN genómico. Las muestras 1 representan a toda la población seleccionada. La muestra 2 representa el clon 6F, la muestra 3 es un clon transfectado en ausencia de TALEN y la muestra 4 son células HEK293 de tipo silvestre. Los cebadores y las condiciones descritas en el texto se utilizaron para verificar mediante PCR la integración correcta en los extremos 5' y 3' de la inserción genómica. La banda principal (tamaño correcto) se ve para el clon seleccionado (6F), así como para la población seleccionada.

Figura 15. Secuencia de plásmido donante para la integración en los sitios de TALEN Flp/GFP (pD5) (SEQ ID NO: 19, 20 y 21). Las características incluyen:

AAVS HA 13-233
 sitio de FRT 243-290
 Lox2272 295-328
 I-Sce1 344-361
 Resistencia a blasticidina 417-818
 Poli A 832-1070

Figura 16. Secuencia de la construcción de meganucleasa I-Sce1 (SEQ ID NO: 22 y 23)

Figura 17. Análisis de citometría de flujo que compara la integración dirigida por nucleasas utilizando meganucleasa I-Sce1 con recombinasas.

Las células del clon 6F se cotransfectaron con pD5-D1.3 y los plásmidos que codifican la nucleasa/recombinasa indicada. Las células se seleccionaron con blasticidina y se analizaron 13 días después de la transfección usando anticuerpos biotinilados anti-Fc humanos y estreptavidina-ficoeritrina. Se indica el porcentaje de células positivas (también se resume en la Tabla 5)

- a. Sin transfectar, b Donante solamente c. I-Sce1, d, TALEN eGFP, e. Cre, f. Flp recombinasa (codificada por el plásmido pOG44).

Figura 18. La integración dirigida por nucleasas impulsa la recombinación homóloga y la "unión de extremos no homólogos" (NHEJ).

- a. Representación de la estructura del plásmido pD5 utilizado para dirigirse al "sitio de aterrizaje" múltiple dentro del intrón de las "células de clones 6F que muestran la posición del cebador J48. El "sitio de aterrizaje" en este plásmido incorpora un sitio de FRT, un sitio lox2272 y un sitio de meganucleasa I-Sce1 (pero no un sitio de TALEN GFP).
 b. Representación del sitio de integración dentro del clon 6F (derivado de pD4) que muestra la posición del cebador J44. El "sitio de aterrizaje" incorpora elementos para dirigir la integración que son FRT, lox 2272, meganucleasa I-Sce1 y TALEN GFP.
 c. Representación del sitio de integración del clon 6F después de la recombinación homóloga de pD5, que muestra la posición de los cebadores J44 y J48.
 d. Representación del sitio de integración del clon 6F después de la recombinación NHEJ o Flp de pD5, que muestra la posición de los cebadores J44, J46 y el cebador inverso J44. La flecha de dos puntas indica el plásmido "extra" derivado del ADN incorporado por NHEJ o integración dirigida por Flp. Se observa en este

ejemplo que el ADN plasmídico entrante (pD5) tiene brazos de homología (que dirigen la recombinación homóloga pero no son necesarios para NHEJ). Estas secuencias se conservan después de la integración por NHEJ, lo que provoca una duplicación de la secuencia representada dentro de los brazos de homología con un par que procede del plásmido en este caso y el otro par que representa las secuencias genómicas endógenas. Por simplicidad, los brazos de homología codificados con plásmidos no se muestran, solo su secuencia equivalente dentro del genoma.

e. Los cebadores 44 y 48 se usaron como cebadores de PCR para las muestras i-iv donde se usó el ADN genómico de las células transfectadas con las siguientes nucleasas/integrasas:

i. Sce1, ii. TALEN (GFP), iii. Flp (pOG44), iv. Donante solamente. Los marcadores de peso molecular eran la escalera de 1 kb "GeneRuler (New England Biolabs). Los cebadores J44 y J48 revelan que se ha producido una recombinación homóloga produciendo una banda de 1928 pb (indicada por una flecha) en las muestras escindidas por nucleasas i y ii.

Los cebadores 44 y 46 se utilizaron para las muestras v-viii donde se usó el ADN genómico de las siguientes muestras. v. Sce1, vi. TALEN (GFP), vii. Flp (pOG44), viii. Donante solamente.

Los cebadores J44 y J46 revelan que la escisión del ADN donante y genómico por la meganucleasa I-Sce1 ha dado como resultado la NHEJ (muestra v.) que produce una banda de 1800 pb (indicada por la flecha). Como se esperaba, se logró una banda de tamaño similar mediante la integración mediada por Flp (vii). NHEJ no se ha producido con TALEN GFP ya que no había ningún sitio de escisión en el plásmido entrante.

Figura 19. Secreción de anticuerpos IgG en sobrenadante de cultivo.

a. Gel teñido con Coomassie de IgG purificada por proteína A de sobrenadantes de cultivo.

i. IgG purificada a partir de sobrenadante de células pD2-D1.3 sin transfección del gen de la recombinasa Dre.

ii. i. IgG purificada a partir de sobrenadante de células pD2-D1.3 transfectadas con el gen de la recombinasa Dre.

ELISA policlonal de anticuerpos secretados. Las células clasificadas del experimento mostrado en la Figura 9H (originalmente de células de la población de anticuerpos seleccionadas por 1 ronda de presentación de fagos) se cultivaron durante 7 días después de la clasificación y se recogió el sobrenadante del cultivo. Las placas de ELISA se recubrieron con β -galactosidasa (10 ug/ml) o BSA (10 ug/ml) durante una noche. Los sobrenadantes de cultivo se diluyeron hasta el 66 % después de mezclarlos con un volumen del 33 % de Marvel-PBS al 6 % (esto se ha descrito anteriormente como la muestra "pura"). El sobrenadante también se diluyó 1/10 en PBS y se mezcló con MPBS al 6 % de la misma manera. La detección de la fusión scFv-Fc unida se realizó utilizando IgG-Eu anti-humana (Perkin Elmer Cat. 1244-330).

Figura 20. Preparación de fragmentos de ADN para la conversión de poblaciones seleccionadas de scFv a formato con IgG. (a) Generación del inserto de ADN CL-pA-CMV-Sigp (como se muestra en la Figura 21b). Amplificación por PCR a partir del plásmido pD2 con los cebadores 2595 y 2597 y purificado en gel. Carril m, escalera de 1 kb Generuler (Thermo, SM031D), carril 1 inserto de ADN CL-pA-CMV-Sigp. (b) La generación del inserto de ADN scFv era como se describe en el Ejemplo 6. Carril m, escalera 1 kb Generuler (Thermo, SM031D), blanco del carril 1, scFv purificado por el carril 2, población de ScFv del resultado de la ronda 1 de β -galactosidasa del carril 3, población de scFv del resultado de la ronda 2 de β -galactosidasa del carril 4, población de scFv del resultado de la ronda 2 de CD229 del carril 5. (c) La purificación de ADN "mini-círculo" digerido por NheI y XhoI. Las ligaduras entre el ADN digerido de NcoI/NotI que codifica scFv (Figura 21, inserto a) y el ADN que codifica la cadena ligera (CL) constante, poli A (pA), promotor de CMV y péptido señal (Figura 21, inserto b) para formar un ADN "mini-círculo" (Figura 21c) se purificaron en una columna de centrifugación, se digirieron con NheI y XhoI y se purificaron con gel de agarosa al 1 %. Los carriles son m, escalera de 1 kb Generuler (Thermo, SM031D), resultado de la ronda 1 de β -galactosidasa del carril 1, resultado de la ronda 2 de β -galactosidasa del carril 2, resultado de la ronda 2 de CD229 del carril 3. El producto linealizado a 2,6 kb, indicado por flecha, se escindió y se purificó.

Figura 21. Representación esquemática del proceso de conversión de formato scFv a IgG. Un inserto de ADN (a) que codifica los dominios V_H y V_L de anticuerpo se liga con el fragmento de ADN (b) que codifica una cadena constante ligera (CL), una secuencia de poliadenilación (pA), un promotor de citomegalovirus (CMV) y un péptido señal (SigP). La unión de las moléculas de ADN a y b para crear un ADN "mini-círculo" no replicante c se facilita mediante una ligadura de "extremo adhesivo". Después de la ligadura, el "mini-círculo" c se linealiza con las enzimas de restricción NheI y XhoI. El producto linealizado d se purifica luego y se liga con el vector digerido e. El vector e incluye un promotor de pEF y una secuencia SigP aguas arriba del sitio NheI y codifica los dominios pesados (CH) constantes de anticuerpo 1 a 3 aguas abajo del sitio XhoI. El producto de la ligadura del inserto d con el vector e daría como resultado el plásmido f, que se puede usar para transformar bacterias y el crecimiento con un marcador seleccionable adecuado permitiría la producción y purificación de ADN plasmídico mediante métodos estándar. El plásmido f purificado se puede introducir en células de mamífero [134] para la expresión

heteróloga de anticuerpos Ig. Como alternativa, el ADN que codifica CH1-3 en el vector e, podría reemplazarse con el ADN que codifica un solo dominio CH1 para la expresión de Fab. V_H y V_L son cadenas pesada y ligera variables de anticuerpo respectivamente. El ADN que codifica un promotor del factor de elongación (pEF), se representan una cadena ligera (C_L) constante de anticuerpo y dominios pesados constantes 1 a 3 (CH1-3), una secuencia de poliadenilación (pA), un promotor de citomegalovirus (CMV) y un péptido señal (SigP).

Figura 22. Ejemplo adicional de preparación de fragmentos de ADN requeridos para la conversión de scFv en IgG (a) Los insertos de scFv generados como se describe en el Ejemplo 14 se separaron en un gel de TBE de agarosa al 1 %. Los carriles 1 y 14 son una escalera de ADN de 500 pb que comienza en 500 pb. Los carriles 2 a 13 son PCR de scFv. (b) La purificación del "minicírculo" linealizado d (Figura 21) se realizó mediante separación en un gel de TBE de agarosa al 1 %. De izquierda a derecha, el primer carril es una escalera de ADN (escalera de 1 kb, Lifetech, 15615-024) y los carriles restantes un "mini-círculo" linealizado d. (c) Como (b) excepto que el promotor de CMV se reemplaza por una secuencia P2A y la escalera de ADN empleada era la escalera de 1 kb Generuler (Thermo, SM031D).

Figura 23. Integración dirigida por nucleasas de genes de aglutinantes utilizando sistemas de electroporación de flujo. Una mezcla 50:50 de plásmidos pD6 que codifican un anticuerpo anti-FGFR1 o FGFR2 se electroporó usando un sistema de electroporación de flujo. Después de 13 días, las células de selección de blasticidina se marcaron con FGFR1-Fc marcado con Dyelight-633 (FGFR1-Dy633) o FGFR2-Fc marcado con Dyelight 488 (FGFR2-Dy488). Las manchas de puntos representan:

- a. Tinción única FGFR2-488 (muestra 1b de la Tabla 7)
- b. Tinción única FGFR1-633 (muestra 1b de la Tabla 7)
- c. Tinción doble FGFR1-633/FGFR2-488 (muestra 1b de la Tabla 7)
- d. Tinción doble FGFR1-633/FGFR2-488 (muestra 3 de la Tabla 7)

Figura 24. Recuperación de genes de anticuerpos después de la clasificación por flujo.

Se seleccionó una población de anticuerpos mediante una ronda de presentación de fagos y se creó una biblioteca de presentación de mamíferos mediante electroporación de flujo utilizando el sistema Maxcyte. Las células se clasificaron utilizando antígeno 1 nM o 10 nM y el ARNm se aisló directamente. Los genes de anticuerpos se recuperaron mediante PCR, se clonaron en un vector de expresión bacteriana y se determinó la proporción de positivos para ELISA ("resultado 1 nM", "resultado 10 nM"). Esto se comparó con el resultado de la ronda 1 original ("resultado del fago R1"). El gráfico muestra el perfil de las señales ELISA obtenidas con cada población.

Figura 25. Vector pINT20 para la expresión de los receptores de linfocitos T.

- a. Representación del plásmido del promotor dual pINT20 que muestra los brazos de homología de AAVS, gen seleccionable de puromicina (con la región alrededor del sitio del aceptor de corte y empalme que se muestra a continuación). La cadena alfa (que incluye alfa variable, constante alfa de ratón-CD3ζ) está flanqueada por los sitios de restricción Nhe1, Not1 y Acc65I y está bajo el control del promotor de pEF. La cadena beta (que incluye la variable beta, la constante beta de ratón-CD3ζ) está flanqueada por los sitios Nco1, Xho1 y hind3 y está bajo el control del promotor de CMV.
 - b. Secuencia en el aceptor de corte y empalme y comienzo del gen de puromicina (SEQ ID NO: 24 y 25)
 - c. Secuencia de la construcción de la cadena alfa del clon del receptor de linfocitos T c12/c2 que muestra los sitios de restricción Nhe1, Not1 y Acc65I (SEQ ID NO: 26 y 27).
 - d. Secuencia de la construcción de cadena beta del clon del receptor de linfocitos T c12/c2 que muestra los sitios de restricción Nco1, Xho1 y Hind 3 (SEQ ID NO: 28 y 29).
 - e. Secuencia de la construcción de cadena alfa del clon del receptor de linfocitos T 4JFH que muestra los sitios de restricción Nhe1/Not1 (SEQ ID NO: 30 y 31).
 - f. Secuencia de la construcción de cadena beta del clon del receptor de linfocitos T 4JFH que muestra los sitios de restricción Nco1/Xho1 (SEQ ID NO: 32 y 33).
 - g. Estrategia y cebador utilizados para mutar CDR3 de la cadena alfa de TCR c12/c2 (SEQ ID NO: 34, 35, 36 y 37).
 - g. Estrategia y cebador utilizados para mutar CDR3 de la cadena beta de TCR c12/c2 (SEQ ID NO: 38, 39, 40 y 41).
- (N = A, C, G, T; S = C O G; W = A O T)

Figura 26. Reconocimiento de complejos péptido:MHC por receptores de linfocitos T introducidos en células de mamífero mediante integración dirigida por nucleasas.

TCR1 es el péptido 1 que reconoce TCR c12/c2 (SLLMWITQV) en forma de complejo con HLA-A2 marcado con ficoeritrina (péptido 1). TCR2 es el péptido 2 que reconoce TCR 4JFH (ELAGIGLTV) en forma de complejo con HLA-A2 marcado con ficoeritrina (péptido 2). Las muestras a y c muestran células que expresan TCR1 expuestas al péptido 1 (a) o al péptido 2 (c). Las muestras e y f muestran células que expresan TCR2 expuestas al péptido 1 (a) o al péptido 2 (c).

Las muestras e y f muestran células HEK293 no transfectadas marcadas con el péptido 1 (e) o el péptido 2 (f).

g. El plásmido que codifica TCR1 se mezcló con un exceso de 100 veces del plásmido TCR2, introducido mediante integración dirigida por nucleasa en células HEK. El 1,15 % de las células se marcó con el péptido 1. El 1,15 % de las células eran positivas.

5 h. El plásmido que codificaba TCR2 se mezcló con un exceso de 100 veces del plásmido TCR1, se introdujo mediante integración dirigida por nucleasas y se marcó con el péptido 2. Se marcaron el 0,62 % de las células. Las células positivas se recogieron por clasificación de flujo y se recuperó el ARNm para el análisis del enriquecimiento de TCR específico.

10 Las muestras i-I ilustran la expresión de una biblioteca de linfocitos T en células HEK293. La biblioteca TCR se introdujo por electroporación Maxcyte y se seleccionó durante 11 días en puromicina

i. Muestra células marcadas con un anticuerpo anti-TCR marcado con APC (eje y).

j. Muestra las células marcadas con el péptido 1 marcado con ficoeritrina:MHC (eje x)

15 k. Muestra células no transfectadas marcadas tanto con anticuerpo anti-TCR como con el péptido 1:MHC

l. Muestra la biblioteca TCR1 de células transfectadas marcadas tanto con anticuerpos anti-TCR como con el péptido 1:MHC. Las muestras m-n ilustran la expresión de TCR en células Jurkat. TCR1 se administró por electroporación de Amaxa y se seleccionó durante 25 días en puromicina. El plásmido se transfectó en presencia (m) o ausencia (n) de nucleasa TALE y se incubó con un anticuerpo anti-TCR de cadena β marcado con APC. Las muestras o-r ilustran la activación del receptor de linfocitos T de las mismas células Jurkat transfectadas con TCR1. Todas las células están marcadas con anticuerpo anti-CD69 (eje y). La muestra o no se estimuló y la p se estimuló durante 24 horas con un anticuerpo anti-CD3. Las muestras q y r se incubaron durante 24 horas con 2 μ l y 6 μ l, respectivamente, de complejo MHC:péptido 1 marcado con PE. Todas las células también se expusieron al anticuerpo CD28.

25 Figura 27. Vectores pINT21 CAR1 y pINT21 CAR2 para la introducción de bibliotecas del receptor de antígeno quimérico (CAR) en células humanas.

Representación del plásmido de promotor único pINT21 que muestra los brazos de homología AAVS, gen seleccionable de puromicina, promotor de CMV que impulsa la fusión del aglutinante al dominio de señalización CD3 ζ . Los sitios Nco1 y Not 1 son demandados para clonar el aglutinante.

30 a. pINT21 CAR1 fusiona el aglutinante con el dominio de yuxtamembrana, transmembrana y señalización de CD3 ζ .

b. pINT21 CAR2 fusiona el aglutinante con el dominio de bisagra y transmembrana CD8, los dominios de activación 4-1BB y CD3 ζ

35 c. Secuencia de CD3 ζ en pINT 21_CAR1 (SEQ ID NO: 42, 43 y 44)

d. Secuencia de CD8, 4-1BB y CD3 ζ en pINT 21_CAR2 (SEQ ID NO: 45 y 46)

e. Secuencia de FMC63 H-L (anticuerpo anti CD19) (SEQ ID NO: 47 y 48)

40 Figura 28 Expresión de scFv y andamiaje alternativo dentro de la construcción del receptor de antígeno quimérico introducida en células humanas mediante integración mediada por nucleasas.

Las células HEK se transfectaron con anticuerpos anti-FGFR1 (b) o adherón lox1 (c) y se marcaron con FGFR1 y lox1 marcados respectivamente. Como control, los mismos antígenos se incubaron con células HEK293 no transfectadas (a y c, respectivamente).

45 Las poblaciones de bibliotecas de presentación de fagos seleccionadas en mesotelina y CD229 se introdujeron en células HEK mediante integración mediada por nucleasas (f y h respectivamente) y se seleccionaron en puromicina durante 11 días. Estas células o células HEK293 no transfectadas se incubaron con mesotelina marcada (e, f) o CD229 (g, h). e y h representan células HEK293 no transfectadas.

50 Figura 29. Secuencia de andamiajes de aglutinante alternativos para la presentación de mamíferos. Las bibliotecas de diferentes formatos de aglutinante pueden introducirse fácilmente mediante integración dirigida por nucleasas utilizando el vector descrito en el presente documento. Por ejemplo, las construcciones de Adhirón se prepararon con los sitios Nco1 y Not 1 flanqueantes para la introducción en construcciones CAR o construcciones de fusión Fc.

55 a. Secuencia de unión a lox1 Adhiron_lox1A (SEQ ID NO: 49 y 50)

b. Secuencia de unión a lox1 Adhiron_lox1B (SEQ ID NO: 51 y 52)

(Los bucles variables se muestran en negrita y subrayados en la secuencia de la proteína).

60 c. Posibles cebadores mutagénicos para la construcción de la biblioteca de aglutinantes dentro del bucle 1 (adhirón mut1) (SEQ ID NO: 53, 55 y 56) o el bucle 2 (adhirón mut2) (SEQ ID NO: 54, 57 y 58),

A continuación se muestra una representación de la región cubierta por los cebadores (cadena inferior) que muestra la traducción de proteínas. n representa el número variable de codones NNS que dan lugar a diferentes longitudes de bucle.

65 d. Secuencia de knottin MCoTI-II de unión a tripsina con sitios Nco1 y Not1 flanqueantes que permiten la expresión de knottin dentro de los vectores descritos en el presente documento. La secuencia del primer

bucle está subrayada SEQ ID NO: 59 y 60).

e. Estrategia para la creación de biblioteca de mutantes knottin. En este ejemplo, el bucle 1 se reemplaza por 10 aminoácidos aleatorizados. En este ejemplo, se introducen codones VNS (V = A, C o G) que proporcionan 24 codones que codifican 17 aminoácidos. Esta secuencia se puede introducir en un clon que codifica MCoTI-II usando métodos estándar (SEQ ID NO: 61, 62 y 63).

Figura 30. Ejemplo de secuencias para la inserción de genes de anticuerpo mediada por nucleasas por ligadura o unión de extremos mediada por microhomología (MMEJ).

a. Secuencia de pD7-Sce1 (nucleótidos 1-120) (SEQ ID NO: 64, 65). La secuencia es como pD6 (véase la Figura 3c, 8), excepto que el brazo izquierdo AAVS entre EcoR1 y Nsi1 ha sido reemplazado por el reconocimiento de meganucleasa I-Sce1 (en negrita). Además, el brazo derecho de AAVS entre Asc1 y Mlu1 se ha reemplazado por un inserto codificado por los cebadores 2723 y 2734 (no se muestra).

b. Secuencia de pD7-ObLiGaRe (nucleótidos 1-120) (SEQ ID NO: 66, 67). La secuencia es como pD6 (véase la Figura 3c, 8), excepto que el brazo izquierdo de AAVS entre EcoR1 y Nsi1 ha sido reemplazado por los sitios de reconocimiento del brazo derecho e izquierdo AAVS TALE. Además, el brazo derecho de AAVS entre Asc1 y Mlu1 se ha reemplazado por un inserto codificado por los cebadores 2723 y 2734 (no se muestra).

Figura 31. Integración dirigida por nucleasas de aglutinantes en el locus ROSA 26.

a. Muestra la secuencia del brazo de homología izquierdo hasta el comienzo del gen de puromicina que muestra los cebadores y los sitios de restricción mencionados en el ejemplo 22 (SEQ ID NO: 68).

b. Secuencia del brazo de homología derecho para la integración dirigida por nucleasas en el locus ROSA 26 que muestra los cebadores y los sitios de restricción mencionados en el ejemplo 22 (SEQ ID NO: 69).

Ejemplos

Ejemplo 1. Construcción de vectores para la expresión de anticuerpos formateados con IgG

Para efectuar selecciones genéticas de aglutinantes (por ejemplo, anticuerpo, proteína o péptido) es necesario introducir un gen que codifica este aglutinante e impulsar la expresión de este gen desde un promotor exógeno, o dirigiendo la integración del transgén aguas abajo de un promotor preexistente en el ADN celular, por ejemplo, un promotor endógeno. Los anticuerpos representan la clase de aglutinantes más utilizada comúnmente y pueden formatearse para su expresión en diferentes formas. En los ejemplos a continuación, se describe la expresión de un formato de gen único donde un scFv se fusiona con un dominio Fc (scFv-Fc). También se ilustra la expresión de anticuerpos formateados como moléculas de IgG2 humanas. Para expresar anticuerpos con formato IgG o FAb en células productoras tales como eucariotas superiores, es necesario expresar las cadenas pesadas y ligeras separadas. Esto se puede hacer introduciendo plásmidos separados que codifican cada cadena o introduciéndolos en un solo plásmido. Dentro de un único plásmido, las 2 cadenas pueden expresarse a partir de un ARNm único multicistrónico. La expresión de proteínas distintas de un solo mensaje requiere elementos tales como una secuencia de entrada interna del ribosoma (IRE) que permite iniciar la traducción en una ubicación secundaria aguas abajo. Como alternativa, se podrían usar elementos de secuencia que promuevan el estancamiento/reinicio de la traducción, tal como las secuencias virales 2A [119].

Como alternativa, pueden expresarse múltiples proteínas distintas a partir de un único plásmido usando múltiples promotores. Las Figuras 1a y 1b muestran la organización de 2 casetes de expresión similares dentro de diferentes esqueletos de vectores (pDUAL y pINT3) que se desarrollaron para la expresión de anticuerpos con formato IgG secretados. Estos casetes de expresión se crearon utilizando una combinación de síntesis génica y amplificación de la reacción en cadena de la polimerasa de elementos estándar tales como promotores y secuencias poliA. Los primeros plásmidos separados se crearon dentro de pCMV/myc/ER (Figura 1c, Life Technologies) para la expresión de la cadena pesada del anticuerpo (pBIOCAM1-NewNot) y la cadena ligera (pBIOCAM2-pEF). Los elementos de pBIOCAM2-pEF (incluidos el promotor de pEF, el gen de la cadena ligera y el sitio poli A) se clonaron en pBIOCAM1-NewNot para crear pDUAL. Los ejemplos mostrados incluyen los dominios VH y VL de un anticuerpo anti-lisozima humanizado llamado D1.3 [120] y se denominan pDUAL-D1.3 y pINT3-D1.3. Los elementos de pDUAL D1.3 representados en la Figura 1a están presentes entre el sitio poliA EcoR1 y BGH del esqueleto plasmídico de pCMV/myc/ER (Life Technologies Cat. V82320 Figura 1c).

De manera similar, se introdujeron casetes de cadena ligera y cadena pesada separados en pSF-pEF (Oxford Genetics OG43) y pSF-CMV-F1-Pac1 (Oxford Genetics OG111) respectivamente para crear pINT1 y pINT2. Estos se combinaron mediante la clonación del casete de cadena ligera (incluido el promotor de pEF, el gen de la cadena ligera y el sitio poli A) aguas arriba del promotor de CMV en pINT2, para crear pINT3. Los elementos de pINT3-D1.3 representados en la Figura 1b se clonan entre el primer Bgl2 y el Sbf1 representado dentro del plásmido pSF-CMV-F1-Pac1 (Figura 1d, Oxford Genetics OG111).

El promotor temprano inmediato de citomegalovirus (promotor de CMV) es un potente promotor y se usó para impulsar la expresión de cadenas pesadas. pDUAL D1.3 también incorpora un líder tripartita (TPL) de adenovirus 2 y un promotor tardío principal mejorado (enh MLP) inmediatamente aguas abajo del promotor de CMV [121]. La proteína alfa del factor de alargamiento 1 se expresa de forma ubicua y abundante en la mayoría de las células eucariotas y su promotor (promotor de pEF) se usa comúnmente para impulsar la expresión del transgén [122]. En pDUAL-D1.3 y pINT3-D1.3, el promotor de pEF se usa para impulsar la expresión de la cadena ligera del anticuerpo. Los sitios de poliadenilación que se originan en la hormona de crecimiento bovina (BGH poliA) están presentes al final de cada casete de expresión.

La secreción de las cadenas pesada y ligera separadas en el retículo endoplásmico (y en última instancia el sobrenadante de cultivo) está dirigida por 2 secuencias líderes diferentes. La secreción de la cadena ligera está dirigida por una secuencia líder BM40 [123]. A esto le siguen los sitios de clonación Nhe1 y Not1 que permiten la clonación en marco de los genes VL que, a su vez, se fusionan con un gen C kappa humano. La secreción de la cadena pesada está dirigida por un líder dividido por un intrón que se origina a partir de un gen VH de ratón (como se encuentra en pCMV/myc/ER). Al líder le siguen los sitios Nco1 y Xho1, lo que permite la clonación en marco de los genes VH del anticuerpo, seguido de un gen IgG2 optimizado por codones. Los genes VL y VH del anticuerpo D1.3 humanizado [120] se clonaron en los sitios Nhe1/Not1 y Nco1/Xho1 respectivamente dentro de pDUAL-D1.3 y pINT3-D1.3.

Las versiones ancladas a membrana de estos plásmidos se crearon para la presentación de mamíferos. El plásmido pD1 se creó al digerir pDUAL-D1.3 con Bsu36I (que corta el dominio CH3 del gen de la cadena pesada de IgG2) y con BstZ171, que corta el esqueleto después de la región poliA de SV40 del casete de resistencia a neomicina (Figura 1c). Por lo tanto, esto elimina la mayor parte del dominio CH3 y la totalidad del casete de expresión de neomicina. El dominio CH3 se reemplaza por un inserto sintético con los extremos compatibles Bsu36I y BstZ171 (representados en la Figura 1e). El inserto sintético fue diseñado para reemplazar el codón de terminación en el extremo del dominio CH3 del anticuerpo con un donante de corte y empalme e intrón que causa el corte y empalme del extremo CH3 con respecto a un exón que codifica el dominio transmembrana del receptor PDGF humano [84] los primeros 5 restos intracelulares, un codón de terminación y un donante de corte y empalme adicional. A esto le sigue un intrón y un aceptor de corte y empalme adicionales seguidos por un codón para un solo aminoácido y luego un codón de terminación (Figura 1e). Los 2 intrones sintéticos que flanquean el exón que codifica el dominio transmembrana se diseñaron con sitios de reconocimiento ROX ubicados dentro de ellos. Los sitios ROX son reconocidos por la recombinasa Dre que causa la recombinación entre el ADN que contiene estos sitios [88]. La inclusión de 2 sitios ROX que flanquean el exón que codifica el dominio transmembrana crea el potencial de eliminar este exón mediante la transfección de un gen que codifica la recombinasa Rox. Esto se anticiparía para crear un producto de anticuerpo secretado.

La Figura 2 muestra la secuencia del plásmido de expresión del anticuerpo del promotor dual resultante que expresa un anticuerpo anti-lisozima D1.3 humanizado (en lo sucesivo en el presente documento denominado pD1-D1.3 (SEQ ID NO: 1). La especificidad de unión a la lisozima se incorpora a través de la inclusión de secuencias VH y VL de D1.3 [120] entre los sitios de restricción Nco/Xho1 y Nhe1/Not1 respectivamente. La secuencia se muestra desde el sitio ECoR1 hasta el BstZ171. Las secuencias más allá de los sitios ECoR1 y BstZ171 son del esqueleto del vector como se representa en la Figura 1c.

Ejemplo 2. Construcción del vector (pD2) para dirigir un casete de anticuerpo al locus AAVS

La escisión dentro del genoma usando nucleasas de sitio específico facilita la inserción de ADN heterólogo a través de recombinación homóloga o unión de extremos no homólogos (NHEJ). Las células HEK293 humanas se escindieron con nucleasas dirigidas al primer intrón del gen de la subunidad reguladora 12C de la proteína fosfatasa 1 (PPP1R12C). Este locus se identificó como un sitio de integración común de virus adenoasociados y se conoce como el sitio AAVS (Figura 3a). El sitio AAVS se considera un locus "seguro" para la inserción y expresión de genes heterólogos en células humanas [124].

Después de la escisión de sitio específico dentro del genoma, es posible promover la integración de un casete que expresa proteínas utilizando recombinación homóloga. Para esto, es necesario flanquear el casete de expresión con regiones homólogas a las secuencias encontradas en cada lado del sitio de escisión genómica. Para dirigir la integración en el locus AAVS, una sección de 804 pb del locus AAVS 5' con respecto al sitio de escisión deseado, se amplificó por PCR para crear un sitio EcoR1 y Mfe 1 en los extremos 5' y 3' respectivamente. Este producto, que representa el brazo de homología izquierdo para dirigirse al casete de anticuerpos, se clonó en el sitio EcoR1 de pD1 recreando el sitio EcoR1 en el extremo 5'. Para el brazo de homología correcto, se amplificó por PCR una sección de 836 pb del locus AAVS, 3' del sitio de escisión, para crear sitios Bstz171 en cada extremo y se clonó en el Bstz171 de pD1. La construcción se representa en la Figura 3b y la secuencia de la construcción resultante (pD2) se muestra en la Figura 3c.

Durante la clonación del brazo de homología izquierdo AAVS, los sitios de restricción Nsi1 y Pac1 también se insertaron en el extremo 3'. Estos sitios se usaron posteriormente para clonar un intrón sintético seguido de un gen blasticidina con un sitio de poliA adjunto. El gen blasticidina carece de un promotor pero está precedido por un sitio

aceptor de corte y empalme que crea una fusión en marco con el exón aguas arriba del locus AAVS (Figura 3a, b). La integración en el locus AAVS causa la expresión del gen blasticidina sin promotor. La secuencia de esta construcción final, llamada pD2, se muestra en la Figura 3c.

5 En la Figura 2 se muestra la secuencia del casete de anticuerpos, que incluye el promotor de pEF, la cadena ligera D1.3, la región poli A, el promotor de CMV, la cadena pesada D1.3, los sitios de corte y empalme alternativos y el sitio poli A. Para evitar la duplicación de esta secuencia se representa en la Figura 3c como un bloque etiquetado como "CASETE DE EXPRESIÓN DE ANTICUERPO D1.3".

10 Ejemplo 3. Integración dirigida por AAVS TALEN de la construcción de IgG para la expresión de anticuerpos en la superficie celular y la unión a antígeno

Las células HEK293F (Life Technologies), cultivadas en medio Freestyle, se transfectaron con ADN pD2-D1.3 en presencia o ausencia de un par de vectores TALEN dirigidos por AAVS. Un par AAVS TALEN ("AAVS original") se ha descrito anteriormente [125] y reconoce la secuencia:

TALEN IZQUIERDA: 5' (T)CCCTCCACCCACAGT (SEQ ID NO: 70)
Espaciador 5' GGGGCCACTAGGGAC (SEQ ID NO: 71)
TALEN derecha: complemento de 5' AGGATTGGTGACAGAAAA (SEQ ID NO: 72)
20 (es decir, 5' TTTTCTGTCACCAATCCT (SEQ ID NO: 73)

Se identificó un par TALEN dirigido por AAVS alternativo, más eficiente y se usó en experimentos posteriores (pZT-AAVS1 L1 TALE-N y pZT-AAVS1 R1 TALE, Cat. N.º GE601A-1 System Biosciences). Este par, que reconoce el mismo sitio (pero no el primer resto "T" que se muestra entre paréntesis anteriormente), se conoce como el par TALEN "AAVS-SBI".

Las células se sembraron a $0,5 \times 10^6$ células/ml y se transfectaron al día siguiente a 10^6 células/ml usando ADN:polietilenimina (PolyPlus) añadida en una relación de 1: 2 (p/p). Las células se transfectaron con 0,6 µg/ml de pD2 y se cotransfectaron con pcDNA3.0 como control (0,6 µg/ml) o los plásmidos TALEN "AAVS original" izquierdo y derecho combinados (0,3 µg cada uno/ml). pD3 que expresa EGFP del promotor de CMV (véase más abajo) se incluyó en el experimento como control de transfección y mostró un 35 % de eficiencia de transfección. Las células se seleccionaron en cultivo en suspensión utilizando medio Freestyle (Life Technology) complementado con 5 µg/ml de blasticidina.

35 Para determinar si la expresión del anticuerpo había ocurrido en la superficie celular, las células se tiñeron con un anticuerpo anti-Fc humano de acuerdo con el siguiente protocolo:

1. 16 días después de la transfección, se centrifugaron $0,5-1 \times 10^6$ células de las poblaciones seleccionadas con blasticidina durante 2 minutos (200-300xg) a 4 °C.
- 40 2. Se lavaron las células con 1 ml de tampón de lavado (BSA al 0,1 % en PBS Gibco n.º 10010) y las células se centrifugaron durante 2 minutos (200-300xg) a 4 °C.
3. Se resuspendieron las células en 100 µl de tampón de tinción (BSA al 1 % en PBS) y se añadieron 5-10 ul de anticuerpos conjugados con fluorocromo. Los anticuerpos eran Fc de IgG anti-humana marcada con ficoeritrina (clon HP6017, Cat. N.º 409304, Biolegend) o IgG2a de ratón marcada con ficoeritrina, control de isotipo κ (Cat. N.º 400214, Biolegend). Se incubaron durante >30 min a 4 °C en la oscuridad.
- 45 4. Se lavaron dos veces con 1 ml de tampón de lavado y se resuspendieron en 500 ul de tampón de lavado.
5. Se añadieron 5 µl de solución de tinción de viabilidad celular (n.º 00-6993-50 eBioscience) que contenía 50 µg/ml de 7-amino-actinomicina D (7-AAD) para identificar las células muertas.
- 50 6. Las células se analizaron en un citómetro de flujo (Beckton Dickinson FACS II).

La Figura 4 muestra que había una población significativamente mayor de células que expresaban anticuerpos cuando pD2-D1.3 se transfecta en presencia de TALEN dirigido por AAVS con un 86 % de resultados positivos en comparación con pD2-D1.3 solo con un 1,5 % de positivos.

55 La funcionalidad del anticuerpo anti-lisozima expresado en la superficie se determinó evaluando la unión al antígeno marcado. La lisozima de huevo de gallina (Sigma: L6876) se marcó utilizando el sistema de conjugación Lightning-Link Rapid (Dylight 488, Innova Biosciences: 322-0010) de la siguiente manera:

- 60 1. Se añadieron 10 ul de reactivo LL-Rapid Modifier a 100 ul de lisozima (200 µg disueltos en 100 ul de PBS) y se mezclaron suavemente.
2. Se añadió la mezcla a la mezcla Lightning-Link® Rapid y se resuspendió suavemente pipeteando hacia arriba y hacia abajo.
3. Se incubó la mezcla durante 15-30 minutos en la oscuridad a temperatura ambiente.
- 65 4. Se añadieron 10 ul de reactivo LL-Rapid Quencher a la reacción y mezclaron suavemente.
5. Se almacenó a 4 °C. La concentración final de lisozima-Dy488 es de 1,6 µg/µl.
6. Se usaron 6 ul de lisozima-Dy488 (~10 ug) por tinción.

7. La tinción, el lavado y la citometría de flujo fueron como se ha descrito anteriormente.

El análisis muestra que el 86 % de las células transfectadas con pD2-huD1.3 se unen a HEL marcado (según lo juzgado por la puerta M1) en comparación con el 0,29 % de las células no transfectadas (Figura 5).

Ejemplo 4. Las nucleasas de sitio específico (TALEN dirigidas por AAVS) mejoran la integración del ADN donante

Las células transfectadas también se colocaron en placas y se seleccionaron con blastidina para determinar el número de células en las que se activó la expresión del gen blastidina sin promotor. 24 horas después de la transfección, las células se colocaron en placas de 0,25 x 10⁶ células/placa de petri de 10 cm (se trataron con cultivo tisular) y se cultivaron en suero bovino fetal al 10 % (10270-106, Gibco) y aminoácido no esencial de medio esencial mínimo al 1 % (MEM_NEAA N.º 11140-035 Life Technologies). Se añadieron 5 ug/ml de blastidina después de 24 horas más y el medio se cambió cada 2 días. Después de 9 días, las células que no recibieron el plásmido pD2 estaban todas muertas. Después de 12 días, las placas se tiñeron con azul de metileno al 2 % (en metanol al 50 %). La densidad de colonias fue demasiado alta para una cuantificación precisa, pero mostró un mayor número de colonias resistentes a la blastidina en presencia de TALEN AAVS, lo que sugiere una integración dirigida en el locus AAVS. Se introdujo una cantidad reducida de ADN para una cuantificación más precisa.

Las transfecciones se llevaron a cabo como se ha descrito anteriormente utilizando 50, 200 o 400 ng de PD2-D1.3/10⁶ células en presencia o ausencia de las TALEN AAVS (0,3 µg/ml de cada TALEN donde esté presente, Tabla 1A). La entrada total de ADN se ajustó a 1,2 ug de ADN por 10⁶ células con el plásmido de control pcDNA3.0. Después de 24 horas de transfección, se colocaron en placas 0,25 x 10⁶ células en una placa de 10 cm y se añadieron 7,5 µg/ml de blastidina después de 24 horas de la colocación en placas. 10 días después de la selección con blastidina, las colonias se tiñen con azul de metileno al 2 % (en metanol al 50 %). Los resultados se muestran en la Figura 6 y se resumen en la Tabla 1A. Esto muestra que la cotransfección de ADN que codifica TALEN dirigidas por AAVS aumenta el número de colonias resistentes a la blastidina alcanzadas en aproximadamente 10 veces.

Se realizó una comparación entre los pares TALEN "AAVS original" y "AAVS SBI" dirigidos al locus AAVS. La Tabla 1B muestra un aumento en el número de colonias resistentes a la blastidina utilizando el par TALEN "AAVS SBI".

Tabla 1. Cuantificación de colonias resistentes a blastidina de la transfección de pD2-D1.3

A.			
Plásmido enzimático	pD2-D1.3 donante (ng/10 ⁶ células)	Con AAVS TALEN	Sin AAVS TALEN
AAVS original	50	319	32
AAVS original	200	526	41
AAVS original	400	686	75
B.			
Plásmido enzimático	pD2-D1.3 donante (ng/10 ⁶ células)	Con AAVS TALEN	Sin AAVS TALEN (pcDNA3.0 de control)
AAVS original	300	1420	111
AAVS original	1000	1080	127
AAVS original	3000	560	70
AAVS-SBI	300	2800	111
AAVS-SBI	1000	1630	127
AAVS-SBI	3000	870	70

Aquí se ha comparado el efecto de la adición de nucleasas TALEN utilizando la expresión del anticuerpo de la superficie celular (Ejemplo 3) o la activación de un gen blastidina sin promotor (Ejemplo 4). El beneficio de la integración dirigida por nucleasas es más obvio cuando se mide la expresión de anticuerpos en comparación con el efecto sobre las colonias resistentes a la blastidina. Una explicación probable es que los niveles de expresión requeridos para efectuar la supervivencia en presencia de blastidina pueden ser significativamente menores que los niveles de expresión requeridos para detectar la expresión de IgG2 en la superficie. Por lo tanto, la falta de incorporación/corte y empalme del gen blastidina sin promotor podría conducir a un bajo nivel de expresión del gen de resistencia a la blastidina que causa un mayor trasfondo de colonias resistentes a la blastidina en ausencia de una expresión significativa de anticuerpos.

Ejemplo 5. Determinación de la precisión de la integración utilizando AAVS TALEN

Para investigar la precisión de la integración, se seleccionaron colonias del experimento del Ejemplo 4/Tabla 1A (de placas duplicadas, sin teñir), se usó ADN expandido y genómico de estas células como plantilla en la PCR. Para la preparación del ADN genómico, las células se recolectaron y se resuspendieron en 700 µl de tampón de lisis (Tris.Cl 10 mM, pH = 8,0, EDTA 50 mM, NaCl 200 mM, SDS al 0,5 %, complementado con 0,5 mg/ml de proteinasa K (agregado justo antes de la lisis). La resuspensión celular en el tampón de lisis se transfirió entonces a un tubo de microcentrifuga y se mantuvo a 60 °C durante aproximadamente 18 horas. Al día siguiente, se añadieron 700 µl de isopropanol al lisado para precipitar el ADN genómico. El tubo de microcentrifuga se centrifugó a 13.000 rpm durante 20 minutos. Luego, el sedimento de ADN genómico se lavó con etanol al 70 % y se centrifugó a 13.000 rpm durante 10 minutos más. Después de girar, el sobrenadante se separó cuidadosamente con cuidado de no tocar el sedimento de ADN genómico. El sedimento de ADN genómico se resuspendió entonces en 100 µl de tampón que contenía Tris 10 mM (pH 8,0) y EDTA 1 mM y se mantuvo a 60 °C durante 30 minutos manteniendo la tapa abierta para eliminar las trazas de etanol. A esta solución de 100 µl, se le añadió RNAsa A (concentración final de 20 µg/ml), y se incubó a 60 °C durante aproximadamente una hora. La concentración de ADN genómico se midió utilizando un espectrofotómetro nanodrop (Nanodrop).

Para identificar la integración correcta, se diseñaron cebadores de PCR que hibridan en el locus genómico de AAVS más allá de los brazos de homología izquierdo y derecho. Estos se emparejaron con cebadores específicos de inserción. En el extremo 5' los cebadores eran:
AAVS-Left-arm-junction-PCR-Forw (9625) 5' CCGGAACCTCTGCCCTCTAAC (SEQ ID NO: 74) BSD_Junction PCR-rev (9626): 5' TAGCCACAGAATAGTCTTCGGAG (SEQ ID NO: 75)

Estos dan un producto de 1,1 kb donde se produce la integración correcta. Los clones 8/9 que surgieron de la integración dirigida por AAVS dieron una banda del tamaño correcto (Figura 7a, b). 2 clones resistentes a la blasticidina derivados sin TALEN no dieron un producto (Figura 11a) indicativo de integración aleatoria. En el extremo 3' los cebadores eran:
Donor_plasmid_seq_PDGFRTM-2 Forw 5' ACACGCAGGAGGCCATCGTGG (SEQ ID NO: 76) AAVS1_right arm_junction_PCR_rev 5' TCCTGGGATACCCCGAAGAG (SEQ ID NO: 77)

Estos dan un producto de 1,5 kb con integración correcta. Los clones 7/9 que surgieron de la integración dirigida por AAVS dieron una banda del tamaño correcto. 2 clones resistentes a la blasticidina derivados sin TALEN no dieron un producto (Figura 11b). Por lo tanto, la mayoría de las células resistentes a la blasticidina surgen de la correcta integración en el locus AAVS, mientras que las colonias resistentes a la blasticidina que surgen en ausencia de TALEN no están integradas correctamente.

Ejemplo 6. Construcción de una biblioteca de presentación de scFv a partir de una población seleccionada de la presentación de fagos y selección a través de la presentación de mamíferos

Los anticuerpos solubles en formato scFv se han expresado previamente a partir del vector pBIOCAM5-3F, donde la expresión es impulsada por el promotor de CMV y el vector proporciona un compañero de fusión C-terminal, que consiste en Fc humano, His6 y 3xFLAG, al gen del anticuerpo [105, 126]. Esto se modificó para crear el vector pBIOCAM5newNot donde el sitio Not1 se incrustó dentro de la región Fc del anticuerpo (como se muestra en la Figura 8). Esto se usó como punto de partida para crear el vector pD6 (Figura 8) para la expresión de fusiones scFv-Fc atadas a la superficie celular. Los cebadores (2598 y 2619) se diseñaron para permitir la amplificación del casete de expresión del promotor de CMV-scFv-Fc de pBIOCAM5newNot. El cebador 2598 hibrida aguas arriba del promotor de CMV y coloca un sitio Pac1 (subrayado) al final.
2598: TTTTTTTTAAATTA GATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTC (SEQ ID NO: 78)

El cebador 2619 se hibrida cerca del final del dominio Fc e introduce un sitio donante de corte y empalme y un sitio Pme1 (subrayado) al comienzo del intrón.
2619: TTTTTTGTTTAAACTTACCTTGGATCCCTTGCCGGGGCTCAGGCTCAGGGAC (SEQ ID NO: 79)

El producto de PCR resultante es compatible con los sitios Pac1 y Pme1 de pD2 (Figura 3). La digestión de pD2 con Pac1 y Pme1 elimina: el promotor de pEF-líder--cadena ligera-promotor de CMV-líder-cadena pesada
Clonación de las inserciones de producto de PCR Pac1/Pme1: promotor de CMV-líder-sitios Nco1/Not1-Fc humano.

La clonación de esta manera posiciona el casete scFv-Fc de manera apropiada para el corte y el empalme del dominio trans-membrana aguas abajo descrito previamente para la presentación de IgG en la superficie celular en pD2. El vector final pD6 se muestra en la Figura 8, se muestra la secuencia de D6 de los sitios Nco1 a Pme1.

Las selecciones de presentación de fagos se realizaron utilizando la biblioteca de presentación de fagos de McCafferty [7] utilizando beta-galactosidasa (Rockland, Cat. B000-17) y CD229 (R y D Systems, Cat. 898-CD-050) como antígenos. Los métodos de selección y subclonación fueron esencialmente como se ha descrito anteriormente [6, 7, 118, 127]. Los genes scFv de poblaciones que surgen de una o dos rondas de selección en beta-galactosidasa

y dos rondas de selección en CD229 se recuperaron por PCR. Los cebadores M13Leadseq se hibridan dentro de la secuencia líder bacteriana que precede al gen scFv y Notmycseq se hibrida en la etiqueta myc que sigue al gen scFv en el vector de presentación de fagos [127].

M13Leadseq (SEQ ID NO: 80)

5 AAA TTA TTA TTC GCA ATT CCT TTG GTT GTT CCT

Notmycseq (SEQ ID NO: 81)

GGC CCC ATT CAG ATC CTC TTC TGA GAT GAG

10 El producto de PCR se digirió con Nco1 y Not1, el inserto digerido se purificó en gel. El producto digerido se ligó en los sitios Nco1 y Not 1 del plásmido de expresión bacteriano pSANG10-3F y el anticuerpo se expresó y se cribó como se describe [127]. Después de 2 rondas de selección en beta-galactosidasa y CD229, se encontró que 40/190 (21 %) y 35/190 (18 %) de los clones eran positivos por ELISA.

15 También se ligaron 550 ng de inserto de corte de Nco/Not en los sitios Nco1 y Not 1 de pD6 (2,4 µg) para crear una construcción que expresa una fusión entre el scFv y la región Fc de la IgG2 humana. El ADN ligado se transformó en células NEB5alfa electro-competentes (New England Biolabs, Cat. C2989) que generaron un tamaño de biblioteca de $2-3 \times 10^7$ clones para cada población. El ADN se preparó y se cotransfectó en 100 ml de células HEK293 cultivadas en medio Freestyle como se ha descrito anteriormente utilizando 0,3 µg de ADN donante (biblioteca pD6) por 10^6 células. Las células se cotransfectaron con 0,5 µg de TALEN "AAVS-SBI" (pZT-AAVS1 L1 TALE-N y pZT-AAVS1 R1 TALE, Cat. N.º GE601A-1 System Biosciences).

20 24 horas después de la transfección, el volumen del cultivo en masa se duplicó y 24 horas después se añadió blasticidina (10 µg/ml). El medio se refrescó cada 3-4 días y después de 6 días la concentración de blasticidina se aumentó a 20 µg/ml.

25 Para determinar el tamaño de la biblioteca, se colocaron 20.000 células en una placa petri de 10 cm (tratadas con cultivo tisular) 24 horas después de la transfección y se cultivaron en suero bovino fetal al 10 % (10270-106, Gibco) y aminoácido no esencial de medio esencial mínimo al 1 % (MEM_NEAA N.º 11140-035 Life Technologies). Se añadieron 10 µg/ml de blasticidina después de 24 horas más y el medio se cambió cada 2 días. Después de 8 días, las placas se tiñeron con azul de metileno al 2 % (en metanol al 50 %). Los resultados se muestran en la Tabla 2. Esto muestra que se obtuvieron bibliotecas de aproximadamente 3×10^6 clones (que representan el 3 % de las células transfectadas) para las 3 poblaciones.

Tabla 2. Determinación del tamaño de la biblioteca de scFv-Fc.

Muestra	Sin colonias/20.000 células	Sin colonias/ 10^6 células	Tamaño de la biblioteca
β-galactosidasa Rd1	546	27.300	$2,7 \times 10^6$
β-galactosidasa Rd2	654	32.700	$3,2 \times 10^6$
CD229 Rd2	556	27.800	$2,8 \times 10^6$

35 El protocolo para etiquetar y clasificar por flujo $10-20 \times 10^6$ células se muestra a continuación. El análisis inicial se realizó 13 días después de la transfección con solo 10^6 células/muestra y con volúmenes de incubación reducidos (volúmenes de reactivo que son 1/10 de los mostrados).

40 La Figura 9 muestra que, 13 días después de la transfección, al menos el 43-46 % de las células expresan la fusión scFv-Fc en la superficie celular y esto puede detectarse utilizando anticuerpos anti-Fc marcados con FITC o ficoeritrina. La unión de beta-galactosidasa biotinilada también se detecta dentro de esta población utilizando estreptavidina marcada con FITC o ficoeritrina. Usando estreptavidina-FITC, el 11,8 % y el 39 % de las células fueron positivas tanto para la expresión de anticuerpos como para la unión al antígeno utilizando bibliotecas derivadas de poblaciones de salida que surgen de 1 o 2 rondas de selección de presentación de fagos, respectivamente. Para CD229 derivado de 2 rondas de presentación de fagos, el 66 % de las células fueron positivas para scFv-Fc y el 24 % de estas fueron positivas para la unión a CD229 (15 % de la población total).

50 A los 20 días después de la transfección, las células se marcaron de acuerdo con el protocolo a continuación (utilizando antígeno biotinilado/estreptavidina marcada con ficoeritrina y Fc anti-humano marcado con FITC).

1. Se cosecharon, se lavaron y se ajustaron las células en $15-20 \times 10^6$ células por muestra. Se centrifugaron las células a 250 g durante 4', TA, se lavaron las células con 1 ml de PBS + BSA al 0,1 % (4 °C), las células se redujeron a 250 g durante 4', TA, se resuspendieron en 1 ml de PBS + BSA al 1 %
- 55 2. Se añadió el antígeno biotinilado a una concentración final de 100 nM y se incubó durante 30' a 4 °C
3. Las células se lavaron 2 veces con 1 ml de BSA al 0,1 % mediante centrifugación a 1500 rpm durante 5 minutos
4. Se añadieron:

60

10 µl de estreptavidina marcada con FITC (1 µg/ml, Sigma Cat. S3762) y 20 µl de Fc anti humano marcado con ficoeritrina (200 µg/ml, BioLegend Cat. 409304), o:

20 µl de estreptavidina marcada con ficoeritrina (200 µg/ml, Biolegend Cat 405203) y 20 µl de Fc antihumano marcado con FITC (200 µg/ml, Biolegend Cat. 409310) PBS+BSA al 1 %, durante 15 a 4 °C en la oscuridad

5 5. Las células se lavaron 2 veces con 1 ml de BSA al 0,1 % mediante centrifugación a 1500 rpm durante 5 minutos

6. Se resuspendieron en 500 µl de PBS enfriado con hielo + BSA al 1 %

7. Se añadieron 20 µl de 7AAD/vial para la tinción de viabilidad

10 Para la clasificación, las células se clasificaron en función del tamaño de la célula, la granularidad, el ancho de pulso y la viabilidad (a través de tinción con 7-AAD, dispersión directa y dispersión lateral. Los resultados se muestran en la Figura 9c y f. En total se clasificaron 10 millones de células y se recogieron el 3,1 % y el 7 % de células doblemente positivas para bibliotecas derivadas de poblaciones de salida que surgen de 2 rondas de selección de CD229 (CD229 R2) y 1 ronda de selección de β-galactosidasa (β-galR1) respectivamente.

15 Las células seleccionadas de las células derivadas de β-galR1 se cultivaron durante 20 días más y se volvieron a analizar (Figura 9h). Esto muestra que la mayoría de las células ahora expresan scFv-Fc y se unen a β-galactosidasa. Esta figura también muestra que la proporción de células dobles positivas dentro de la población no seleccionada no ha disminuido 42 días después de la transfección (Figura 9k).

El ADN genómico se preparó a partir de 150.000-10⁶ células clasificadas. El ADN genómico se preparó utilizando el método descrito anteriormente o utilizando un kit de minipreparación de ADN genómico de mamífero GenElute (Sigma G1N10).

25 Los genes scFv se amplificaron por PCR a partir del ADN genómico utilizando los siguientes cebadores:

2623 (SEQ ID NO: 82)

TAAAGTAGGCGGTCTTGAGACG

2624 (SEQ ID NO: 83)

30 GAAGGTGCTGTTGAACTGTTCC

35 Las reacciones de PCR se llevaron a cabo utilizando polimerasa Phusion (NEB Cat. M0532S) en un tampón del fabricante que contenía 0,3 µM de cada cebador y el 3 % de DMSO. Se usaron 100-1000 ng de ADN genómico como plantilla en una reacción de 50 µl. Se llevaron a cabo 30 ciclos a 98 °C durante 10 s, 55 °C durante 25 s, 72 °C durante 45 s. Esto dio un producto de 1,4 kB que se digirió con Nco1 y Not1. Se generó una banda de aproximadamente 750-800 pb y se purificó en gel antes de la clonación en pSANG10. El ADN ligado se transformó en células BL21 (células competentes Edge Bio Ultra BL21 (DE3), Cat. 45363). De esta manera, los fragmentos scFv derivados de la población clasificada pueden expresarse en bacterias como se ha descrito anteriormente [7, 127].

40 Como alternativa al aislamiento del gen del anticuerpo y la expresión en una combinación alternativa de vector/hospedador, es posible derivar el anticuerpo secretado directamente de las células seleccionadas tras la clonación de una sola célula o utilizando una población clasificada para generar una mezcla de anticuerpos policlonales. Para ilustrar este cultivo, el sobrenadante se tomó de células clasificadas (de células βgalR1) después de 7 días en cultivo. Se demostró que esto era positivo en ELISA utilizando placas recubiertas con βgalactosidasa (véanse el Ejemplo 13 y Figura 19b).

50 Ejemplo 7. Construcción y selección de una biblioteca de presentación de IgG de una población seleccionada de la presentación de fagos

Los fragmentos de ADN que codifican scFv, que representan los resultados de presentación de fagos de las rondas 1 y 2 de las selecciones contra 3-galactosidasa y CD229, se generaron como se describe en el Ejemplo 6. Las poblaciones scFv se convirtieron en el formato IgG de acuerdo con el Ejemplo 14 y se detallan en el siguiente método.

55 Un inserto de ADN que codifica el dominio constante (C_L) de la cadena ligera kappa humana, la secuencia de poliadenilación (pA), el promotor de CMV y el péptido señal de la cadena V_H murina (que se representan entre los sitios Not1 y Nco1 de pD2 que se muestran en la Figura 21b) se amplificó por PCR a partir del plásmido pD2 con los cebadores 2595 (GAGGGCTCTGGCAGCTAGC) (SEQ ID NO: 84) y 2597 (TCGAGACTGTGACGAGGCTG) (SEQ ID NO: 85). Las reacciones de PCR se llevaron a cabo utilizando la polimerasa de inicio en caliente KOD (Novagen Cat. 71086-4) en un tampón del fabricante que contenía 0,25 µM de cada cebador. Se usaron 10 ng de ADN plasmídico pD2 como plantilla en una reacción de 50 µl. Se realizaron 25 ciclos a 98 °C durante 10 s, 55 °C durante 25 s, 72 °C durante 40 s. Esto dio un producto de 1,8 kB que se digirió con Nco1 y Not1 y se purificó en gel (Figura 20a representada como inserto de C_L-pA-CMV-SigP en la Figura 21b).

65

Los fragmentos de ADN que codifican scFv, que representan los resultados de presentación de fagos de las rondas 1 y 2 de las selecciones contra β -galactosidasa y CD229, se generaron como se describe en el Ejemplo 6. La Figura 20b muestra poblaciones de scFv seleccionadas contra β -galactosidasa y CD229 separadas por electroforesis en gel de agarosa al 1 %.

Las ligaduras entre el inserto de scFv y el inserto de C_L -pA-CMV-SigP se realizaron incubando el inserto de scFv digerido con NcoI/NotI (1 μ g) con el inserto de C_L -pA-CMV-SigP digerido con NcoI/NotI (1 μ g) con ADN ligasa T4 (1,5 μ l, Roche, 10-481-220-001) en un tampón del fabricante en un volumen total de 40 μ l para formar el "minicírculo" representado en la Figura 21c. Las ligaduras se incubaron a 16 °C durante 16 horas, se purificaron mediante columna de centrifugación, se digirieron con NheI y XhoI y el producto de 2,6 kb (representado en la Figura 21d) se purificó por separación electroforética en gel de agarosa al 1 % (Figura 20c).

El inserto de ADN representado en la Figura 21d que codifica V_L - C_L -pA-CMV-SigP- V_H (0,5 μ g) se ligó con NheI/XhoI digerido, se purificó en gel el vector pD2 (0,7 μ g) (Figura 21e) con ADN ligasa T4 (1,5 μ l, Roche, 10-481-220-001) en un tampón del fabricante en un volumen total de 40 μ l para producir el vector de direccionamiento representado en la Figura 21f. Esto codifica poblaciones de anticuerpos formateados como IgG, que se originan a partir de selecciones de presentación de fagos de anticuerpos de primera o segunda ronda con respecto a β -galactosidasa o CD229. Las ligaduras se incubaron a 16 °C durante 16 horas, se purificaron mediante columna de centrifugación y se eluyeron con agua de calidad HPLC.

El ADN ligado se transformó en células NEB5alfa electro-competentes (New England Biolabs, Cat. C2989) que generaron un tamaño de biblioteca de 1-4 x 10⁵ clones para cada población. El ADN se preparó y se cotransfectó en 100 ml de células HEK293 cultivadas en medio Freestyle como se ha descrito anteriormente utilizando 0,3 μ g de ADN donante (biblioteca pD6) por 10⁶ células. Las células se cotransfectaron con 0,5 μ g de TALEN "AAVS-SBI" (pZT-AAVS1 L1 TALE-N y pZT-AAVS1 R1 TALE, Cat. N.º GE601A-1 System Biosciences).

24 horas después de la transfección, el volumen del cultivo en masa se duplicó y 24 horas después se añadió blasticidina (10 μ g/ml). El medio se refrescó cada 3-4 días y después de 6 días la concentración de blasticidina se aumentó a 20 μ g/ml.

Para determinar el tamaño de la biblioteca, se colocaron 250.000 células en una placa petri de 10 cm (tratadas con cultivo tisular) 24 horas después de la transfección y se cultivaron en suero bovino fetal al 10 % (10270-106, Gibco) y aminoácido no esencial de medio esencial mínimo al 1 % (MEM_NEAA N.º 11140-035 Life Technologies). Se añadieron 10 μ g/ml de blasticidina después de 24 horas más y el medio se cambió cada 2 días. Después de 8 días, las placas se tiñeron con azul de metileno al 2 % (en metanol al 50 %). Los resultados se muestran en la Tabla 3. Esto muestra que se obtuvieron bibliotecas de entre 5 x 10⁵ y 9 x 10⁵ clones (que representan del 0,5 % al 0,9 % de las células transfectadas) para las 3 poblaciones.

Tabla 3. Determinación del tamaño de las bibliotecas de presentación de mamíferos formateadas como IgG.

Muestra	N.º de colonias/0,25 x 10 ⁶ células	N.º de colonias/10 ⁶ células	Tamaño de la biblioteca (x10 ⁵)
β -galactosidasa Rd1	1337	5348	5,3
β -galactosidasa Rd2	1972	7888	7,9
CD229 Rd2	2175	8700	8,7

La mayor parte de la población de células transfectadas con los resultados de las rondas 1 o 2 de selección en β -galactosidasa se seleccionó en medio que contenía blasticidina como se ha descrito anteriormente. Después de 19 días, se marcaron 10-20 x 10⁶ células y se realizó la clasificación de flujo como se describe en el Ejemplo 6. Las células clasificadas se cultivaron durante 17 días y se volvieron a analizar mediante citometría de flujo (Figura 10). Esto mostró que la mayoría de las células eran ahora doblemente positivas para la expresión de IgG y la unión a β -galactosidasa.

El ADN genómico se preparó a partir de las células clasificadas y el ADN que codificaba el inserto de IgG se aisló mediante PCR. El inserto que codificaba IgG se amplificó utilizando polimerasa KOD (Merck, cat. n.º 71086-3), con una temperatura de hibridación de 60 °C y empleando 30 ciclos. El tampón proporcionado por el fabricante con DMSO al 5 % se usó con 0,3 μ M de los cebadores 2597 (SEQ ID NO: 54) y 2598 (SEQ ID NO: 47). El producto del tamaño deseado se purificó en gel. El producto purificado en gel se usó luego para la PCR anidada utilizando KOD polimerasa (Merck, cat. n.º 71086-3) en un tampón del fabricante con DMSO al 5 % usando 0,3 μ M de cebador 2625 (SEQ ID NO: 55) en combinación con el cebador 1999 (SEQ ID NO: 56) (para la muestra R1), o 2595 (SEQ ID NO: 53) (para 4R1 y 5R1), usando una temperatura de hibridación de 60 °C empleando 30 ciclos. Estos productos de PCR anidados se purificaron en gel y se sometieron a doble digestión con NheI-HF (NEB, Cat. n.º R3131S) y XhoI (NEB, cat. N.º R0146S) para ligarlos con pINT3 de doble digestión similar (Figura 1) para la expresión de aglutinantes solubles formateados como IgG. Las secuencias de cebador son:

2597: AGGGGTTTTATGCGATGGAGTT (SEQ ID NO: 85)

2598: GTTACAGGTGTAGGTCTGGGTG (SEQ ID NO: 78)

5 2625: CCTTGGTGCTGGCACTCGA (SEQ ID NO: 86)

1999: **AAAAGCAGGCTACCATGAGGGCCTGGATCTTCTTTCTCC** (SEQ ID NO: 87)

2595: GAGGGCTCTGGCAGCTAGC (SEQ ID NO: 84)

10

Ejemplo 8. Construcción y selección de una biblioteca de scFv sin tratar

Schofield *et al.* [7] describen la construcción de una biblioteca de presentación de fagos (biblioteca de McCafferty") en la que los genes de anticuerpos de los linfocitos B de varios donantes humanos se clonaron primero en una "biblioteca intermedia" antes de volver a clonarse en la biblioteca de presentación de fagos funcional final. Esta misma biblioteca intermedia y la misma metodología se usaron para generar una nueva biblioteca (biblioteca IONTAS) de 4×10^{10} clones. El ADN plasmídico se preparó a partir de esta biblioteca con el fin de garantizar una representación suficiente de la biblioteca dentro de la inoculación bacteriana. Se establecieron varias reacciones de PCR utilizando un total de 2 ug de plantilla de ADN. El producto de PCR se digirió con Nco1 y Not 1, se purificó en gel y se ligó como se describe en el Ejemplo 6. Se ligaron 9,3 ug de pD6 y 0,93 ug de inserto de PCR durante una noche, la reacción de ligadura se limpió mediante la extracción con fenol cloroformo y el ADN se sometió a electroporación en células DH5 alfa como se ha descrito anteriormente [7]. Como resultado, se creó una biblioteca de $2,4 \times 10^8$ clones dentro del vector de presentación de scFv-Fc. El ADN se preparó a partir de esta "biblioteca sin tratar" clonada en pD6 y transfectada en 1 litro de células HEK293F (Life Technologies) cultivadas en medio Freestyle (como se ha descrito anteriormente). 0,3 ug de ADN de la biblioteca pD6, 0,5 ug de cada par TALEN "AAVS-SBI". 24 horas después de la transfección, el volumen de cultivo se duplicó y 48 horas después de la transfección, la selección con blasticidina se inició como se ha descrito anteriormente. El tamaño de la biblioteca se determinó colocando alícuotas del cultivo en placas 24 horas después de la transformación y seleccionando blasticidina como se ha descrito anteriormente. Se creó una biblioteca de $0,9 \times 10^7$ clones.

20

Se biotinilaron varios antígenos usando el kit EZ-Link Sulfo-NHS-LC-Biotin (Pierce Cat. N.º 21327) de acuerdo con las instrucciones del fabricante. Los antígenos fueron tiroglobulina bovina (Calbiochem Cat. 609310), quimera CD28-Fc humana (R y D Systems, Cat. 342-CD-200) y quimera EphB4-Fc de ratón (R y D Systems, Cat. 446-B4-200). También se usó β -galactosidasa biotinilada (Rockland Cat. B000-17).

35

Las células transfectadas se seleccionaron en blasticidina en cultivo líquido como se ha descrito anteriormente durante 17 días. Las células se recogieron, se lavaron y se ajustaron a $15-20 \times 10^6$ células por muestra. Las células se prepararon como se describe y el antígeno biotinilado se añadió a una concentración de 500 nM. El etiquetado y la clasificación de flujo a como se ha descrito anteriormente. Usando células de control incubadas solo con el anticuerpo anti-Fc marcado con ficoeritrina, se creó una "puerta" que incluía el 0,05 % de estas células. Usando la misma puerta para las células marcadas, se incluyeron entre el 0,28 - 0,51 % de las células (Figura 11). Estas se recogieron y se cultivaron para permitir rondas adicionales de clasificación y amplificación de genes scFv de la biblioteca sin tratar.

40

Ejemplo 9. Creación de una línea celular con "sitios de aterrizaje" múltiples para comparar los enfoques dirigidos por nucleasas y recombinasas para la integración genómica

Para permitir la comparación de métodos de integración basados en escisión genómica o integración mediada por recombinasa, se construyó un vector de direccionamiento dirigido por AAVS (pD4) que introduce un intrón con múltiples "sitios de aterrizaje" (Ejemplo 3). Estos incluyen un sitio de FRT reconocido por la recombinasa F1p y un par de sitios lox2272/loxP reconocidos por la recombinasa Cre. Para permitir la escisión dirigida, pD4 también incluye una secuencia de GFP para la cual se ha diseñado un par TALEN [128] y un sitio de meganucleasa I-Sce 1 para permitir la integración dirigida por endonucleasas. Se construyó un plásmido donante entrante compatible (pD5) con sitios de reconocimiento apropiados, de modo que la integración dirigida por nucleasas o recombinasas provoca la activación de un gen blasticidina sin promotor y la integración de un casete de expresión de anticuerpos.

55

La organización del plásmido y la secuencia de pD4 se muestran en la Figura 12b. Primero se creó un plásmido pD3 intermedio que incluía un gen de GFP bajo el control del promotor de CMV seguido de una fusión del gen puromicina/timidina quinasa bajo el control del promotor de PGK (Figura 12a). Esto se creó al digerir pBIOCAM1-newNot con Sac1 (en el extremo del promotor de CMV) y BstB1 (entre el gen Neo y el sitio poli A, Figura 1a). Esto eliminó el casete de expresión de neomicina y permitió el reemplazo con un inserto sintético que incluía un gen de proteína verde fluorescente mejorada (EGFP) bajo el control del promotor de CMV. Esta construcción GFP se fusionó en el extremo C a los restos 422-461 de una secuencia PEST de ornitina descarboxilasa de ratón mutada. Esta secuencia PEST se incorpora en el plásmido pZsGreen1-DR (Clontech) y se ha demostrado que reduce la semivida de la GFP fusionada a 1 hora. Un casete que codifica un promotor de PGK, una fusión del gen puromicina/timidina quinasa (Puro deltaTK) y un casete de poliA se escindió del plásmido pFLEXIBLE [129] usando

60

65

Xmn1 y Fse1 y se clonó en los sitios Sma1 y Fse1 presentes en el inserto sintético original. El plásmido resultante (llamado pD3) codifica un gen GFP impulsado por CMV y un gen de resistencia a la puromicina impulsado por un promotor de PGK.

5 Para crear el vector de direccionamiento final pD4, se eliminó el promotor de CMV y se insertaron los brazos de homología de AAVS. Una sección de 850 pb del locus AAVS se amplificó por PCR para crear un brazo de homología izquierdo AAVS flanqueado por un EcoR1 en el extremo 5' y un Mre1 en el extremo 3'. Esto se clonó en el sitio ECoR1/Mre1 de pD3, eliminando de este modo el promotor de CMV. También se incorporó un sitio Nsi1 en el extremo 3' de este brazo de homología izquierdo AAVS. Los sitios Mre1 y Nsi1 vecinos se utilizaron para introducir un fragmento sintético que fusionaba un intrón con el gen EGFP como se muestra en la Figura 13. El intrón sintético que precede al gen EGFP incorpora:

15 un sitio de reconocimiento de FRT para la recombinasa Flp
un sitio de recombinación lox 2272
un sitio de meganucleasa I-Sce1
un sitio de reconocimiento GFP TALEN
una secuencia de estancamiento ribosómico T2A [130]

20 El brazo de homología derecho AAVS se generó mediante PCR para crear sitios Hpa1 y BstZ171 en los extremos 5' y 3'. Este fragmento se clonó en los sitios Hpa1 y BstZ171 de pD3. El plásmido pD4 resultante codifica un casete de resistencia a la puromicina ("Puro deltaTK") y se puede usar para introducir un "sitio de aterrizaje" en el locus AAVS que incorpora diversos sitios de nucleasa y recombinasa para comparación. La secuencia de pD4 se muestra en la Figura 13 (SEQ ID NO: 5, SEQ ID NO: 6 y SEQ ID NO: 7). Los brazos de direccionamiento izquierdo y derecho AAVS se muestran en detalle en la Figura 3 y, por lo tanto, se abrevian en la Figura 13.

25 El intrón introducido en pD4 contiene sitios de reconocimiento TALEN originados en GFP [128]. El par TALEN dirigido por eGFP (eGFP-TALEN-18-izquierdo y eGFP-TALEN-18-derecho) reconoce la secuencia que se muestra a continuación (todas las secuencias y los cebadores se presentan en una dirección de 5' a 3') donde las mayúsculas representan el sitio de reconocimiento de TALEN izquierdas y derechas y las minúsculas muestran la secuencia de espaciador. La TALEN derecha reconoce el complemento de la secuencia mostrada. El primer par de bases es equivalente a la posición menos 16 de la secuencia en relación con la secuencia ATG iniciadora de GFP (se muestra subrayada en el espaciador).
TCCACCGTGCCAccatggtagcaaggCGAGGAGCTGTTCA (SEQ ID NO: 88)

35 El plásmido pD4 también incorpora un sitio de meganucleasa I-Sce1 y un sitio de FRT reconocido por la recombinasa Flp. Finalmente, pD4 incorpora lox 2272 y loxP (que son mutuamente incompatibles) que flanquean los casetes de expresión de GFP y puromicina. La incorporación de estos mismos 2 sitios loxP que flanquean el plásmido donante (pD5 a continuación) brinda la oportunidad de sustituir el casete integrado (incluido el casete PGK puro delta TK) reemplazándolo con un casete entrante que impulsa la expresión de blastidina y anticuerpos a través del intercambio de casete mediado por recombinasa.

Creación de una línea celular por transfección de pD4.

45 Las células HEK293F se resuspendieron a 10⁶ células/ml y se añadió ADN:polietilenimina (PolyPlus) a una relación de 1:2 (p/p). Las células se transfectaron con 0,6 µg/ml de pD4 y se cotransfectaron con el par TALEN "AAVS original" o pcDNA3.0 como control (0,6 µg/ml). pD3 que expresa EGFP del promotor de CMV se incluyó en el experimento como control de transfección y mostró un 35 % de eficacia de transfección. Después de 24 horas, las células se colocaron en placas de 0,5 x 10⁶ células/placa de petri de 10 cm (se trataron con cultivo tisular) y se cultivaron en suero bovino fetal al 10 % (10270-106, Gibco) y aminoácido no esencial de medio esencial mínimo al 1 % (MEM_NEAA N.º 11140-035 Life Technologies). Se añadieron 5 µg/ml de puromicina después de 24 horas más y el medio se cambió cada 2 días. Después de 5 días, las células no transfectadas o las células transfectadas con pD3 solo estaban muertas. Después de 12 días, hubo aproximadamente 200 colonias en células transfectadas con pD4 solamente y aproximadamente 400 colonias en células transfectadas con pD4 y el par AAVS TALEN.

55 La población resistente a la puromicina que surge de la transfección con pD4 y el par AAVS TALEN se analizó para su correcta integración. Además, se seleccionó una única colonia de esta población (clon 6F) y se comparó con una colonia de la población resistente a la puromicina derivada de la transfección de pD4 en ausencia del par AAVS TALEN. Para identificar la integración correcta, se diseñaron cebadores de PCR que hibridan en el locus genómico de AAVS más allá de los brazos de homología izquierdo y derecho. Estos se emparejaron con cebadores específicos de inserción. En el extremo 5' los cebadores eran:

60 AAVS1_HA-L_Nested_Forw1 GTGCCCTTGCTGTGCCGCCGGA ACTCTGCCCTC (SEQ ID NO: 89)
EGFP_Synthetic_gene_Rev_Assembly TTCACGTCGCCGTCCAGCTCGAC (SEQ ID NO: 90)
Purotk_seq_fow2 TCCATACCGACGATCTGCGAC (SEQ ID NO: 91)

AAVS1_Right_arm_Junction_PCR_Rev TCCTGGGATACCCCGAAGAG (SEQ ID NO: 77)

La figura 14 muestra que el clon 6F y la población son correctos en los extremos izquierdo y derecho, pero el clon seleccionado de la población no dirigida por AAVS es negativo. Por lo tanto, el análisis de PCR indica que la precisión de la integración del casete donante es mayor cuando está dirigida por la escisión de AAVS TALEN del ADN genómico.

pD4 introduce un gen GFP en marco, sin promotor, sin promotor impulsado por el promotor de AAVS. La citometría de flujo de la población resistente a puomicina mostró una ausencia de expresión de GFP. Este fallo en la expresión podría deberse a la combinación de una semivida corta (del elemento de secuencia PEST de ornitina descarboxilasa murina) combinada con una expresión reducida derivada del uso del promotor de T2A. De hecho, se encontró que la adición del elemento T2A frente a un elemento de blasticidina sin promotor (como se describe para pD2) redujo el número de colonias resistentes a la blasticidina en 4 veces. A pesar de la ausencia de expresión de GFP, la integración de múltiples sitios de aterrizaje todavía ofrece una oportunidad para la comparación de la integración genómica dirigida por recombinasas frente a dirigida por la división de ADN.

Ejemplo 10. Construcción de un vector para insertar un casete de anticuerpo (pD5) en el sitio de "aterrizaje múltiple"

Después de la introducción del intrón de "sitio de aterrizaje múltiple" en el locus AAVS, es posible introducir un casete de anticuerpo a través de medios dirigidos por nucleasas o dirigidos por recombinasas. Para hacer esto, se creó un plásmido pD5 donante donde el casete de expresión está flanqueado por los brazos de homología izquierdo y derecho que son equivalentes a las secuencias que flanquean el sitio de escisión de GFP TALEN introducido en pD4. pD5 no incorpora un sitio de reconocimiento de GFP TALEN intacto y la integración se impulsa por recombinación homóloga. La integración dirigida por la homología del plásmido donante conducirá a la introducción de un gen blasticidina que carece de un promotor, pero está precedido por un sitio aceptor de corte y empalme que crea una fusión en marco con el exón aguas arriba del locus AAVS como se ha descrito anteriormente. La integración en el locus AAVS causará la expresión del gen blasticidina sin promotor. El casete insertado también codifica cadenas pesadas y ligeras de anticuerpos con formato IgG bajo el control de los promotores de pEF y CMV, respectivamente, como se ha descrito anteriormente. pD5 incorpora un sitio de meganucleasa I-Sce1 que puede conducir a la escisión del donante entrante que brinda una oportunidad para NHEJ (véase el Ejemplo 12). También se incorpora un sitio FRT en el plásmido donante pD5 que permite la incorporación dirigida por recombinasas del gen blasticidina sin promotor y los casetes de expresión de anticuerpos en el mismo locus. Como se ha analizado anteriormente, la recombinasa Cre actuará en los sitios loxP en el ADN donante y genómico para dirigir el intercambio de casetes mediado por recombinasas.

La secuencia de pD5 se muestra en la Figura 15. La secuencia en el extremo 5' del sitio GFP TALEN es del locus AAVS. Se generó una sección de 267 pb del locus AAVS aguas arriba del sitio de escisión de TALEN por PCR. Se utilizaron cebadores que crearon un sitio EcoR1 y un sitio Mfe 1 en los extremos 5' y 3' y el producto se clonó en el sitio ECoR1 de pD1-D1.3. El sitio EcoR1 se crea de nuevo en el extremo 5'. Durante la clonación del brazo de homología izquierdo, también se insertaron Nsi1 y Pac1 en el extremo 3'. Un ensamblaje de PCR creó un brazo de homología derecho, que incorporaba aproximadamente 700 pb equivalentes a la secuencia 3' de GFP TALEN. Los cebadores de PCR introducen los sitios BstZ171 en 5' y en el extremo 3' del fragmento ensamblado y esto se clonó en el sitio BstZ171 de pD1-D1.3. Los cebadores de PCR también introdujeron un sitio Hpa1 en el extremo 5'.

Se creó un fragmento de PCR que incluía el intrón (que incorpora sitios de reconocimiento para GFP TALEN, endonucleasa I-Sce1, recombinasa Flp y recombinasa Cre), una región de aceptor de corte y empalme, un gen de resistencia a la blasticidina y un sitio poli A (descrito anteriormente) con el sitio Nsi1 en el extremo 5' y un sitio Pac1 en el extremo 3'. Esto se clonó en el sitio Nsi1 y Pac1 del plásmido descrito anteriormente para crear pD5-D1.3 (secuencia mostrada en la Figura 15 y estructura plasmídica mostrada en la Figura 18a).

Ejemplo 11 Comparación de la integración dirigida por nucleasas y dirigida por Flp de un casete de expresión de anticuerpos

El sistema Flp-In que se ha utilizado previamente para la integración de casetes de expresión de anticuerpos mediada por recombinasa [18] utiliza una recombinasa Flp mutante (en el plásmido pOG44) que posee solo el 10 % de la actividad a 37 °C de la recombinasa Flp nativa [19]. Se ha identificado una variante de la recombinasa Flp (Flpe) con mejor termoestabilidad y actividad a 37 °C que la de tipo silvestre [19, 20]. Esto se mejoró aún más mediante la optimización por codón para crear Flpo [131] codificado dentro del plásmido cCAGGS-Flpo (Genebridges Cat. A203). Se comparó el efecto de ambas variantes de la recombinasa Flp (codificada en pOG44 y cCAGGS-Flpo). La recombinación dirigida por la recombinasa Cre también se examinó mediante la cotransfección de células con un plásmido que codifica la recombinasa Cre [132] (pCAGGS-Cre, Genebridges Cat. A204). En cada vector, la recombinasa se expresa bajo el control del promotor de la β -actina de pollo y un potenciador temprano inmediato de CMV. Se utiliza una secuencia de localización nuclear T de SV40 grande para la localización nuclear [20]. En los vectores originales (cCAGGS-Flpo y pCAGGS-Cre), la expresión de recombinasa se vinculó a un gen de resistencia a la puomicina mediante un sitio de entrada ribosomal interno (IRES) que se eliminó utilizando técnicas estándar de biología molecular.

Se llevó a cabo un experimento para comparar las eficiencias de la integración dirigida a la escisión genómica frente a la integración dirigida por recombinasas de un casete de anticuerpos. El resultado se evaluó de 2 maneras:

1. Medición del número de colonias resistentes a blasticidina que surgen de la integración de un gen blasticidina sin promotor
2. Evaluación del grado de expresión de anticuerpos alcanzado por los diferentes enfoques.

Como se describe en el Ejemplo 9, los sitios de reconocimiento para la recombinasa Cre (lox2272 y loxP) y la recombinasa Flp (FRT) se integraron previamente en el locus AAVS dentro del clon 6F. Además, los sitios de reconocimiento para un par de GFP TALEN y para la meganucleasa I-Sce1 también están presentes dentro del mismo intrón. El plásmido donante pD5-D1.3 transporta los mismos sitios de reconocimiento (aparte de GFP TALEN) dentro de un intrón aguas arriba de un gen blasticidina sin promotor. La correcta integración conducirá a la activación del gen blasticidina. pD5-D1.3 también codifica un gen de anticuerpo D1.3 con formato IgG que se expresará en la superficie celular.

La cotransfección de pD5-D1.3 con pOG44 o pCAGGS-Flpo (que codifica 2 variantes de la recombinasa Flp) debería dar como resultado la integración del plásmido pD5 completo en el sitio FRT del clon 6F. El plásmido donante pD5-D1.3 también tiene un sitio lox2272 dentro del intrón sintético aguas arriba del gen blasticidina y un sitio loxP en el extremo del casete de expresión de anticuerpo. Bajo la acción de la recombinasa Cre expresada a partir de pCAGGS-Cre, el intercambio de casetes mediado por la recombinasa debería dar como resultado la integración de los casetes de expresión de blasticidina y anticuerpo en los sitios lox2272 y loxP dentro del clon 6F.

La eficiencia de la integración de vectores utilizando enfoques dirigidos por recombinasa con enfoques dirigidos por escisión genómica se comparó utilizando un par TALEN (eGFP-TALEN-18-izquierdo y eGFP-TALEN-18-derecho) dirigidas hacia una secuencia de GFP (Reyon *et al.*, 2012). En el caso de las GFP TALEN, el elemento entre los brazos de homología izquierdo y derecho se integrará después de la escisión genómica por parte de las TALEN.

Para permitir la comparación con la meganucleasa I-Sce1, se construyó un gen de optimizado por codones que codificaba I-Sce1 (Figura 16). Este gen tiene una etiqueta de epítipo HA terminal N/señal de localización nuclear SV40 (NLS) en el extremo N y está flanqueado por los sitios Nco1 y Xba1 en los extremos 5' y 3'. El gen se clonó en el vector pSF-CMV-F1-Pac1 (Oxford Genetics OG111) donde la expresión se impulsa desde el promotor de CMV.

Las transfecciones se llevaron a cabo utilizando el clon 6F con un "sitio de aterrizaje múltiple" correctamente integrado. Las células se suspendieron a 10^6 /ml y se transfectaron con 50 ng de plásmido donante pD5-D1.3/células 10^6 junto con plásmidos que codifican enzimas (Tabla 4A).

Después de 24 horas, las células se colocaron en placas de $0,5 \times 10^6$ células/placa de petri de 10 cm (se trataron con cultivo tisular) y se cultivaron en suero bovino fetal al 10 % (10270-106, Gibco) y aminoácido no esencial de medio esencial mínimo al 1 % (MEM_NEAA N.º 11140-035 Life Technologies). Se añadieron 5 ug/ml de blasticidina después de 24 horas más y el medio se cambió cada 2 días. Después de 12 días, las placas se tiñeron con azul de metileno al 2 % (en metanol al 50 %) y se contó el número de colonias. En una comparación directa entre la recombinasa Flp, la recombinasa Cre y TALEN, el mayor número de colonias se obtuvo mediante el uso de GFP TALEN, donde hubo un aumento de 9 veces en comparación con "solo donante" (Tabla 4A). También parece que el uso del gen Flpo optimizado en realidad dio como resultado una reducción del número de colonias resistentes a la blasticidina en comparación con el control "solo donante", probablemente a través de la toxicidad de la recombinasa Flp de actividad mejorada. También hubo un aumento en el número de colonias usando pCAGGS-Cre en comparación con el control del donante solamente.

Se realizó un segundo experimento comparando GFP TALEN tanto con Flp mejorada (de cCAGGS- Flpo), así como la enzima Flp de baja actividad codificada en pOG44 a partir del sistema Flp-In (como se usa por Zhou *et al.* [17, 18, US7.884.054]). Se compararon con GFP TALEN y Cre recombinasa (Tabla 4B). Las células se transfectaron con las cantidades de ADN mostradas por millón de células. Se colocaron en placas $0,25 \times 10^6$ células y se determinó el número de colonias resistentes a blasticidina como se ha descrito anteriormente. Las células también se seleccionaron para la resistencia a la blasticidina en cultivo líquido durante 30 días antes de determinar la proporción de células que expresaban IgG de superficie (como se ha descrito anteriormente). La Tabla 4B muestra que TALEN fue superior a los otros enfoques en términos de número de colonias resistentes. Nuevamente, el uso de Flp optimizado en cCAGGS - Flpo en realidad causó una reducción en el número de colonias resistentes a la blasticidina en comparación con los controles "solo donantes". La recombinasa Cre condujo de nuevo a un aumento en el recuento de colonias de blasticidina en comparación con el control, mientras que el gen Flp dentro de pOG44 mostró solo un aumento marginal en comparación con el control.

Tabla 4. Comparación de los enfoques de integración dirigidos por nucleasas TALE y dirigidos por recombinasas.

A.				
Plásmido enzimático	Cantidad	pD5-D1.3 donante (ng/10 ⁶ células)	Colonias resistentes a blasticidina (colonias/10 ⁶ células)	
Par GFP TALEN	0,575 µg cada vez	50	152 (304)	
pCAGGS-Flp _o	1,15 µg	50	1 (2)	
pCAGGS-Cre	1,15 µg	50	57 (114)	
control (pCDNA3.0)	1,15 µg	50	17 (34)	
B.				
Muestra 2 µg cada vez por 10 ⁶ células	pD5-D1.3 Donante µg/10 ⁶ células	Colonias de blasticidina	Colonias de blasticidina por células 10 ⁶ (porcentaje de resistencia a blasticidina)	Porcentaje positivo en 1 flujo
Par GFP TALEN	0,6	270	1080 (1,1 %)	95,6
cCAGGS-Flp _o	0,6	35	140 (0,14 %)	Muy pocas células
pOG44	0,6	96	384 (0,38 %)	6,4
pCAGGS-Cre	0,6	180	720 (0,72 %)	4
Control (PCDNA3.0)	0,6	81	324 (0,32 %)	37,3
C.				
Muestra 2 µg cada vez por 10 ⁶ células	pD5-D1.3 Donante µg/10 ⁶ células	Colonias de blasticidina	Porcentaje positivo en 1 flujo	
GFP TALEN	2	210	95,3	
GFP TALEN	6	120	73,6	
cCAGGS-Flp _o	2	62	Muy pocas células	
cCAGGS-Flp _o	6	41	Muy pocas células	
pOG44	2	178	35,6	
pOG44	6	63	58,7	
pCAGGS-Cre	2	84	40,9	
pCAGGS-Cre	6	52	Muy pocas células	
Control (PCDNA3.0)	2	340	65,6	
Control (PCDNA3.0)	6	82	55,5	

Con la adición de más ADN donante (Tabla 4C) hubo un aumento en el número de colonias en niveles intermedios (2 µg/millón de células) y una disminución a niveles más altos en todos los ámbitos (6 µg/millón de células). Ninguna de las otras muestras logró niveles de presentación de anticuerpos observados con la integración dirigida por GFP TALEN.

5 Las células también se seleccionaron para resistencia a la blasticidina en cultivo de líquido y las células se tiñeron para la expresión de anticuerpos como se ha descrito anteriormente. La integración dirigida por TALEN proporcionó una proporción significativamente mayor de células positivas para anticuerpos en comparación con los otros enfoques. Las células transfectadas con cCAGGS-Flp_o y con altas concentraciones de pCAGGS-Cre no estaban sanas y no había números suficientes para llevar a cabo la citometría de flujo.

15 La comparación se amplió para incluir la endonucleasa I-Sce1. Se sintetizó un gen sintético que codificaba I-Sce1 (Figura 16) y se clonó en el sitio Nco1/Xba1 de pSF-CMV-f1-Pac1 (Oxford Genetics). Las células se suspendieron a 10⁶/ml y se transfectaron para cada ml de células (10⁶ células/ml) con 300 ng de plásmido donante pD5-D1.3 junto con plásmidos que codificaban enzimas (1 µg/10¹ células). Al día siguiente, se colocaron en placas 0,05 ml de células y se seleccionaron en blasticidina y se tiñeron después de 14 días como se describe. La Tabla 5 muestra que el mayor número de colonias resistentes a la blasticidina provino de la meganucleasa I-Sce1 seguida del par eGFP TALEN. Tanto la recombinasa Cre como Flp (codificada en pOG44) dieron números ligeramente más altos que el control "solo de donante". Como antes de la transfección con los plásmidos codificantes de Flpe, en realidad se redujeron los números de colonias en comparación con "solo de donante".

Tabla 5. Comparación de los enfoques de integración dirigidos por meganucleasas y dirigidos por recombinasas

Muestra 2 µg cada vez por 10 ⁶ células	pD5-D1.3 Donante µg/10 ⁶ células	Colonias de blasticidina	Colonias de blasticidina por células 10 ⁶ (porcentaje de resistencia a blasticidina)	Porcentaje positivo en flujo d7	Porcentaje positivo en flujo d13
Par GFP TALEN	0,6	90	1800	27,6	55 %
meganucleasa I-Sce1		150	3000	29,9	47 %
pOG44	0,6	60	1200	2,79	6,5 %
pCAGGS-Cre	0,6	56	1120	2,95	6,6 %
cCAGGS-Flp _o	0,6	4	80	5,9 (bajo n.º de células)	Muy pocas células
Control (PCDNA3.0)		40	800	3,3	4,9 %
(SBI AAVS en WT HEK		251 (x2)	10040 (1 %)	ND	ND

Después de la transfección, la mayor parte de las células se seleccionaron para la resistencia a la blasticidina en cultivo de líquido y después de 7 y 13 días se tiñeron con un anticuerpo marcado con ficoeritina anti-Fc como se ha descrito anteriormente. La Figura 17 (resumida en la Tabla 5) muestra que se logró una expresión de anticuerpos significativamente mayor para las células transfectadas con endonucleasa I-Sce1 y eGFP TALEN (47 % y 55 % respectivamente) en comparación con "solo donante" (4,9 %). En contraste, el porcentaje de células con anticuerpos positivos cuando las células se cotransfectaron con plásmidos que codifican la recombinasa Flp (pOG44) o la recombinasa Cre fue del 6,6 y el 6,5 % respectivamente. La proporción de células positivas a anticuerpos continúa aumentando con la selección continua en blasticidina y alcanza un 85-90 % de anticuerpos positivos en el caso de las muestras transfectadas con I-Sce1 y EGFP TALEN cuando se ensayan el día 19. Por lo tanto, las meganucleasas proporcionan un enfoque alternativo para realizar la integración dirigida por nucleasas de transgenes que codifican anticuerpos.

Ejemplo 12. La integración dirigida por nucleasas de un casete de anticuerpo puede ocurrir por recombinación homóloga y NHEJ

La eficiencia de la integración de los transgenes en el ADN celular puede mejorarse mediante la introducción de roturas bicatenarias (DSB). Los mecanismos de reparación del ADN endógeno en las células eucariotas incluyen la recombinación homóloga, unión de extremos no homólogos (NHEJ) y sus variantes. Todos proporcionan un medio para introducir genes que codifican aglutinantes dentro de una biblioteca. La recombinación homóloga proporciona una unión precisa entre las regiones de homología y el transgén insertado, pero requiere la provisión de regiones de homología en el plásmido donante. El ADN para la recombinación homóloga se puede proporcionar como ADN lineal o circular. Con NHEJ, los extremos del ADN se vuelven a ligar directamente sin la necesidad de una plantilla homóloga. Este enfoque con respecto a la reparación del ADN es menos preciso y puede conducir a inserciones o eliminaciones. No obstante, NHEJ proporciona un medio simple para integrar los exones en marco en el intrón o permite la integración de los casetes promotor:gen en el genoma. El uso de métodos no homólogos permite el uso de vectores donantes que carecen de brazos de homología, simplificando de este modo la construcción del ADN donante.

El clon 6F tiene sitios de reconocimiento de nucleasa GFP TALEN e I-Sce1 integrados en el genoma y estos se escindirán cuando se proporcionen estas nucleasas. El vector donante pD5 no tiene un sitio de reconocimiento de nucleasa GFP TALE sino que tiene brazos de homología que flanquean el sitio de escisión y, por lo tanto, se espera que se integre solo por recombinación homóloga. La escisión del ADN genómico en la meganucleasa I-Sce1 vecina también conducirá a la integración de los elementos pD5 mediante recombinación homóloga. Sin embargo, pD5 también tiene un sitio de meganucleasa I-Sce1 que se puede escindir in vivo cuando se proporciona I-Sce1. Esto creará un producto de ADN lineal que potencialmente puede integrarse por NHEJ. Como se ha descrito anteriormente, incluso puede haber ventajas de eficiencia al usar la escisión in vivo del ADN donante cuando se usa NHEJ.

La Figura 18a representa el ADN donante pD5-D1.3 entrante y la Figura 18b representa el locus genómico de las células del clon 6F que incorporan el sitio de "aterizaje múltiple". La Figura 18c representa la consecuencia de la recombinación homóloga entre pD5-D1.3 (Figura 18a) y el sitio de aterizaje múltiple del clon 6F (Figura 18b). La Figura 18d en contraste representa la consecuencia de NHEJ. En este caso, se incorpora ADN adicional del esqueleto del plásmido entrante (representado por una flecha doble). La recombinación mediada por Flp en el sitio de "aterizaje múltiple" conducirá a un producto similar. Con el fin de determinar qué ruta se está utilizando con las muestras descritas en el Ejemplo 11 (que se muestra en la Figura 17), se preparó ADN genómico a partir de la población seleccionada con blasticidina como se ha descrito anteriormente. Se diseñó un cebador de PCR inversa (J44) que hibrida con el promotor PGK integrado. Esto se usó junto con J48 que hibrida en el extremo de la proteína IgG. Los cebadores J44 y J48 fueron diseñados para revelar una recombinación homóloga que produce una banda

de 1928 pb cuando I-Sce1 es responsable de la integración (se indica por la flecha en la Figura 18e). (Este par de cebadores podría producir potencialmente una banda de 5131 pb cuando la NHEJ tuvo lugar, pero este producto más largo no fue visible en las PCR genómicas de este experimento).

- 5 El cebador J46 se diseñó para hibridar dentro del gen de β -lactamasa dentro del esqueleto del vector. Se anticipa que los cebadores J44 y J46 producen una banda de 1800 pb cuando ha tenido lugar la NHEJ. Se espera una banda de tamaño similar cuando la recombinasa FIp ha conducido a una integración mediada por recombinasas.

J44: AAAAGCGCCTCCCTACCCGGTAGAAT (SEQ ID NO: 92)

10 J46: GGCGACACGGAAATGTTGAATACTCAT (SEQ ID NO: 93)

J48: CACTACACCCAGAAGTCCCTGAGCCTG (SEQ ID NO: 94)

- 15 La Figura 18e revela claramente que la recombinación homóloga se produce solo con las muestras tratadas con GFP TALEN y meganucleasa I-Sce1 ((i e ii en comparación con iii e iv). Por el contrario, la NHEJ solo se produce cuando la escisión se realiza por I-Sce1 meganucleasa (Figura 18e v.) pero no GFP TALEN (Figura 18e vi). Como se esperaba, se encuentra una banda de tamaño similar en la muestra tratada con la recombinasa FIp (Figura 18e vii). Por lo tanto, este experimento revela que la integración dirigida por nucleasas de un casete de anticuerpo puede producirse tanto por recombinación homóloga como por NHEJ.

20 Ejemplo 13. Generación de fragmentos de anticuerpos secretados y unidos a membrana de la misma célula

- Como se ha descrito anteriormente, los vectores de presentación de mamíferos pD2 y pD5 se construyeron con un exón que codificaba un dominio transmembrana flanqueado por dos sitios de reconocimiento ROX reconocidos por recombinasa Dre [88]. Con el fin de determinar si era posible convertir de una forma unida a membrana a una forma secretada, la población resistente a la blastidina que surge de la transfección con el par pD2-D1.3/AAVS TALEN se volvió a transfectar con el plásmido que codificaba la recombinasa Dre (pCAGGs-Dre). Esto se basó en el plásmido pCAGGs-Dre-IRES puro [88] que impulsa el gen de la recombinasa Dre de un promotor CAGG (GeneBridges A205). El gen de resistencia a la puomicina se eliminó utilizando técnicas estándar de biología molecular. Después de 22 días de selección con blastidina, las células se ajustaron a $0,5 \times 10^6$ células/ml y se transfectaron como se ha descrito anteriormente con $0,5 \mu\text{g}$ de pCAGGs-Dre por 10^6 células. Después de 6 días, se recogieron los sobrenadantes, se purificó el anticuerpo usando la proteína A y las muestras se procesaron en un gel SDS-PAGE y se tiñeron con azul de Coomassie. La Figura 19a muestra que el anticuerpo secretado se encontró en el sobrenadante incluso sin transfección con el gen de la recombinasa Dre. Esto puede surgir de un corte y empalme alternativo donde se omite el exón que codifica el dominio transmembrana. Como alternativa, el anticuerpo en el sobrenadante de cultivo podría surgir de la escisión del anticuerpo unido a la membrana. La transfección de recombinasa Dre aumentó el nivel de anticuerpos secretados (Figura 19a).

- La producción de fusión scFv-Fc secretada también se demostró en el experimento descrito en el ejemplo 7 (Figura 9h). Las poblaciones de anticuerpo scFv seleccionadas por la ronda 1 de presentación de fagos en β -galactosidasa se introdujeron en el vector pD6 y se integraron en el locus AAVS de células HEK293 utilizando AAVS TALEN. Las células de unión a antígeno se clasificaron por clasificación de flujo y las células seleccionadas se cultivaron durante 7 días después de la clasificación sin un cambio de medio para permitir que el anticuerpo se acumulara. Las placas de ELISA se recubrieron con β -galactosidasa (10 $\mu\text{g/ml}$) o BSA (10 $\mu\text{g/ml}$) durante una noche. Los sobrenadantes de cultivo de los cultivos de 7 días se mezclaron con un volumen del 50 % de Marvel-PBS al 6 % y la muestra se analizó por triplicado. También se probó una dilución 1/10. La detección de la fusión scFv-Fc unida se realizó utilizando IgG-Eu anti-humana (Perkin Elmer Cat. 1244-330). La Figura 19b muestra que la unión del anticuerpo puede detectarse directamente a partir de sobrenadantes de cultivo puros o con una dilución 1/10. Esto ilustra que tanto la presentación en la superficie como la secreción de anticuerpos se pueden lograr dentro de las mismas células sin etapas adicionales. Será posible derivar el anticuerpo secretado directamente de las células seleccionadas, ya sea después de la clonación de una sola célula o utilizando una población clasificada como se muestra aquí para generar una mezcla de anticuerpos policlonales.

Ejemplo 14. Un método sencillo para la conversión de scFv a formato IgG o Fab

- 55 Se inventó un método novedoso para efectuar la conversión de anticuerpos formateados como scFv en un formato IgG como se describe en el Ejemplo 7. Esta conversión es un proceso necesario durante los proyectos de descubrimiento de fármacos de anticuerpos que emplean bibliotecas de presentación de fagos de anticuerpos scFv donde se requiere un anticuerpo formateado por IgG o Fab como formato final. Los métodos actuales son tediosos e implican la clonación individual de las cadenas pesada variable (V_H) y ligera variable (V_L) en vectores de expresión adecuados. Además, la conversión de una población de scFv "en masa" no es posible porque se pierde el enlace entre las cadenas V_H y V_L . Esto es un problema porque las cadenas V_H y V_L contribuyen a la especificidad de unión al antígeno. La incapacidad actual para convertir fácilmente las poblaciones de scFv en formato Ig o Fab limita la capacidad de cribar grandes cantidades de anticuerpos en el formato final que se utilizarán en la clínica. La capacidad de cribar anticuerpos recombinantes en formato Ig o Fab para determinar la unión a dianas, las cribas de indicador celular y las propiedades y funciones biofísicas, incluido el estado de agregación, es una etapa necesaria para elegir los fármacos de anticuerpos candidatos como candidatos clínicos. Cuanto mayor sea el número de

anticuerpos probados en esta fase, en formato IgG o Fab, mayor será la posibilidad de seleccionar el mejor candidato de fármaco para anticuerpos.

Aquí se describe un método para convertir las poblaciones de anticuerpos monocatenarios (scFv) al formato de inmunoglobulina (Ig) o Fab de tal manera que se mantengan los pares originales de cadena pesada variable (V_H) y de cadena ligera variable (V_L). El método permite convertir scFv monoclonales, oligoclonales o policlonales simultáneamente al formato Ig o Fab. Preferiblemente, el método procede a través de la generación de un ADN de "mini-círculo" no replicante. Preferiblemente, el proceso de conversión completo conlleva una transformación única de bacterias tales como *E. coli* para generar una población de colonias bacterianas que albergan cada una un plásmido que codifica un anticuerpo recombinante con formato Ig o Fab único. Esto es distinto de los métodos alternativos que requieren dos etapas separadas de clonación y transformación [117].

Más ampliamente, este aspecto de la invención se refiere a un método para convertir una construcción genética con 3 elementos genéticos enlazados A, B y C (representados por el enlazador V_H , y V_L respectivamente en el caso de scFv) en un formato donde el orden de los elementos flanqueantes (A y C) se invierte, en una sola etapa de clonación. El elemento intermedio podría conservarse, pero lo más útil es que el método permite el reemplazo de este elemento intermedio por un nuevo elemento D (para dar C-D-A). En el ejemplo de la conversión de un scFv en una IgG o Fab, entonces C es un dominio V_L de anticuerpo y A es un dominio V_H . En este ejemplo, el elemento D encapsula un dominio constante de cadena ligera, un sitio poli A, un promotor y una secuencia líder fusionados con el V_H (elemento A). En el proceso, el producto (C-D-A) se vuelve a clonar permitiendo que también se cambien las secuencias de flaqueo. En el ejemplo de conversión de scFv a IgG, el elemento V_L está precedido por un promotor y una secuencia líder y a V_H le sigue un dominio $C_{H1}-C_{H2}-C_{H3}$ en el caso de anticuerpos con formato IgG y un dominio C_{H1} en el caso de anticuerpos con formato Fab. El método podría aplicarse más ampliamente donde los elementos A y C (utilizando la nomenclatura anterior) podrían representar otros elementos genéticos, por ejemplo, en la construcción de proteínas con permutación circular donde los extremos N y C originales están fusionados y los nuevos extremos internos están modificados.

La Figura 21 ilustra el proceso de conversión esquemáticamente usando scFv en la conversión de IgG como ejemplo. Un inserto de ADN (a) que codifica los dominios V_H y V_L de anticuerpo se liga con el fragmento de ADN (b) que codifica una cadena constante ligera (CL), una secuencia de poliadenilación (pA), un promotor de citomegalovirus (CMV) y un péptido señal (SigP). El fragmento de ADN (b) también podría codificar cualquier promotor en lugar del promotor de CMV. Además, el casete de pA-CMV podría ser reemplazado por un sitio de entrada ribosomal interno (IRES) [119] o un péptido pequeño de tipo 2A "auto-escindido" [130, 133]. La unión de las moléculas de ADN (a) y (b) para crear un ADN "mini-círculo" no replicante (c) se facilita mediante una ligadura de "extremo adhesivo". En la Figura 21, los sitios NcoI y NotI se emplean porque se usaron en la creación de la biblioteca de presentación de fagos de McCafferty [7] sin embargo, cualquier sitio de restricción adecuado podría usarse para crear el "mini-círculo" no replicante c. Después de la ligadura, el "mini-círculo" c se linealiza con las enzimas de restricción NheI y XhoI, cuyos sitios de reconocimiento flanquean el enlazador entre los dominios V_H y V_L . NheI y XhoI se eligieron para ilustrar esta invención porque se usaron en la creación de la biblioteca de presentación de fagos de McCafferty [7], sin embargo, se podría usar cualquier sitio de restricción adecuado.

El producto linealizado d se purifica luego y se liga con el vector digerido (e). El vector (e) incluye un promotor de CMV o pEF y una secuencia de señal aguas arriba del sitio NheI y codifica los dominios pesados (C_H) constantes del anticuerpo 1 a 3 aguas abajo del sitio XhoI. El vector también codificaría un origen bacteriano o replicación y un marcador de resistencia a antibióticos (no mostrado) para permitir la selección y replicación del ADN plasmídico resultante en bacterias. El producto de la ligadura del inserto (d) con el vector (e) daría como resultado el plásmido f, que se puede usar para transformar bacterias y el crecimiento con un marcador seleccionable adecuado permitiría la producción y purificación de ADN plasmídico mediante métodos estándar. El plásmido f purificado se puede introducir en células de mamífero [134] para la expresión heteróloga de anticuerpos Ig. Como alternativa, el ADN que codifica C_{H1-3} en el vector (e), podría reemplazarse con el ADN que codifica un solo dominio C_{H1} para la expresión de Fab.

En la descripción detallada del método que se usa a continuación para ilustrar esta invención, el inserto b contiene un promotor de CMV o un péptido P2A que permite la expresión de cadenas ligeras y pesadas de anticuerpos separados de un ARN mensajero único (ARNm). El método no es obvio y se refinó después de varios intentos experimentales. Por ejemplo, inicialmente se intentó la linealización del ADN "mini-círculo" (c) mediante PCR. Sin embargo, esto dio lugar a la amplificación de productos secundarios homodiméricos, dando como resultado un bajo rendimiento del producto deseado (d). En contraste, la digestión directa del ADN "mini-círculo" (c) proporcionó suficiente material (d) para permitir que el método se implementara con éxito. En segundo lugar, en un intento por evitar un producto de homodímero no deseado, el inserto (a) se desfosforiló inicialmente. Sin embargo, esto requirió un control cuidadoso para evitar la digestión "final" que da como resultado que el producto carezca de los "extremos adhesivos" deseados para la ligadura. El método óptimo no incluye la desfosforilación para maximizar la proporción de producto competente en ligadura. Por último, se requirió un control cuidadoso de las relaciones utilizadas en la ligadura de los insertos de ADN (a) y (b) para maximizar el rendimiento del "mini-círculo" de ADN (c).

1. Preparación de inserciones scFv de PCR

A partir de una reserva de glicerol bacteriano, que albergaba ADN plasmídico que codificaba scFv, se raspó en 50 μ l de agua. Se diluyó esto 1 en 10. Se usaron 5 μ l de esto para la reacción por PCR que contenía el cebador directo pSANG10peIB (CGCTGCCAGCCGGCCATGG SEQ ID NO: 95) (2,5 μ l, 5 μ M), el cebador inverso 2097 (GATGGTATGATGATGTGCGGATGCG SEQ ID NO: 96), (2,5 μ l, 5 μ M), 10x tampón KOD (kit de inicio en caliente KOD de Merck, 71086-4), dNTPs (5 μ l, 2 mM), MgSO₄ (2 μ l, 25 mM), polimerasa de inicio en caliente KOD (2,5 unidades) en un volumen total de 50 μ l. Las condiciones de ciclado fueron 94 °C durante 2 min, después 25 ciclos de 94 °C durante 30 s, 54 °C durante 30 s, después 72 °C durante 1 min. La limpieza de la PCR se realizó mediante una columna de centrifugación (Qiagen o Fermentas) y las reacciones de la PCR se eluyeron en 90 μ l. La Figura 22a muestra 1 μ l de reacción de PCR cargada en un gel de agarosa al 1 % en TBE. El ADN scFv purificado (80 μ l, 8 μ g) se digirió mediante la adición del tampón 4 (New England Biolabs), BSA (0,1 mg/ml) y 40 unidades de NcoI-HF y NotI-HF en un volumen total de 100 μ l y se incubó durante 2 horas a 37 °C. Los insertos se purificaron con un kit de limpieza Qiagen PCR, se eluyeron en 30 μ l y la concentración de ADN se midió midiendo la absorbancia a 260 nm con un espectrofotómetro nanodrop (Thermo).

2. Ligadura de insertos de ADN (Figura 21a y b)

Se realiza una reacción de ligadura para producir el "mini-círculo" de ADN (Figura 21c). La reacción de ligadura contiene el inserto b (125 ng), el inserto scFv a (Figura 21) (125 ng), el tampón de ligadura 10x (kit de ADN ligasa T4 de Roche, 1,5 μ l), la ADN ligasa T4 (1 unidad) en un volumen total de 15 μ l. Se incubó durante 1-2 h a 21 °C. Se añadió agua (35 μ l) a la mezcla de ligadura y se purificó con un kit de limpieza Qiagen PCR y se eluyó en 30 μ l.

3. Digestión del ADN "mini-círculo" (Figura 21c) con XhoI/NheI

La reacción de ligadura purificada (28 μ l) se digiere mediante la adición del tampón 4 (New England Biolabs, 3,5 μ l), BSA (0,1 mg/ml) y 10 unidades de NcoI-HF y NotI-HF en un volumen total de 35 μ l y se incubó durante 2 horas a 37 °C. Esto se purifica entonces por separación en agarosa al 1 % en TBE (Figura 22b). Como alternativa, la Figura 22c muestra un "mini-círculo" linealizado que contiene una secuencia P2A en lugar de un promotor de CMV. La banda de ADN a 2,6 kb (Figura 22b) se escinde y se purifica con el kit de extracción de gel Qiagen y se eluye en 30 μ l.

4. Ligadura de ADN "mini-círculo" linealizado d con el vector pINT3 (corte XhoI/NheI) y transformación de DH5 α de *E. coli*.

Se estableció una ligadura estándar con el vector de corte pINT3 (50 ng), el "mini-círculo" linealizado d (20 ng), el tampón de ligasa 10x (Roche, 1,5 μ l) y 1 unidad de ADN ligasa T4 (NEB) en un volumen final de 15 μ l. La incubación se realizó a 21 °C durante 2 h. La transformación de las células químicamente competentes de DH5 alfa de *E. coli*, la eficiencia de subclonación (Invitrogen, cat. 18265017) fue de acuerdo con las instrucciones del fabricante. Se añadieron 80 μ l de células DH5a químicamente competentes a 6 μ l de la mezcla de ligadura, se colocaron en hielo durante 1 hora y se sometieron a choque térmico a 42 °C durante 1 min, hielo durante 2 min, y después se transfirieron a un tubo de polipropileno de 14 ml que contenía 900 μ l de medio SOC y se incubaron a 37 °C durante 1 hora y se colocaron en placas de amp LB.

Ejemplo 15: Construcción de grandes bibliotecas de presentación en células de mamífero mediante integración dirigida por nucleasas usando electroporación de flujo

La electroporación es una forma eficiente de introducir ADN, ARN y proteínas en las células y los sistemas de flujo de electroporación permiten la introducción eficiente de ADN en grandes cantidades de células de mamíferos. Por ejemplo, el "Sistema de transfección escalable MaxCyte STX" (Maxcyte) permite la electroporación de 10¹⁰ células en 30 minutos, lo que crea la posibilidad de transfectar hasta 10¹¹ células en un día. Las células y el ADN se mezclan y se pasan de un reservorio a una cámara de electroporación, se someten a electroporación, se bombean y el proceso se repite con una alícuota fresca de células y ADN. Se puede aplicar el mismo método para la introducción de ADN, ARN, proteína o mezclas de los mismos en células cultivadas (por ejemplo, células HEK293 humanas o células Jurkat) o células primarias, por ejemplo, linfocitos humanos [135]. La electroporación de flujo se ha utilizado para introducir eficazmente ADN, ARN y proteínas en un gran número de células primarias y cultivadas.

Aquí se ilustra el uso de un sistema de este tipo para introducir ADN donante, que codifica genes de anticuerpos, mediante co-transfección con ADN que codifica un par de nucleasas TALE dirigidas al locus AAVS humano de células HEK293 humanas y células Jurkat.

La distribución de las 2 especificidades de anticuerpos diferentes se determinó mediante citometría de flujo usando un antígeno marcado con fluorescencia. La generación de anticuerpos que reconocen los receptores de FGF FGFR1 o FGFR2 se ha descrito previamente [105]. Los clones α -FGFR1_A y α -FGFR2_A (descritos allí) se clonaron en pD6 como se describe en el ejemplo 6. Además, una población de anticuerpos scFv seleccionados de la biblioteca de presentación de fagos "McCafferty" [7] usando una ronda de presentación de fagos en β -galactosidasa (β -gal) también se clonó en este vector (como se describe en el ejemplo 6).

Las células HEK293 se centrifugaron y se resuspendieron en un volumen final de 10^8 células/ml en el tampón de electroporación del fabricante (tampón de electroporación Maxcyte, Thermo Fisher Scientific Cat. N.º NC0856428)). Se añadió una alícuota de 4×10^7 células (0,4 ml) a la cubeta de electroporación con 100 µg de ADN (es decir, $2,5 \mu\text{g}/10^6$ células). Las cantidades de los diferentes componentes utilizados se muestran a continuación. El ADN donante que codifica los anticuerpos α -FGFR1_A y α -FGFR2_A se proporcionó como una mezcla equimolar con la cantidad total por 10^6 células que se muestra en la Tabla 6 a continuación. El ADN que codifica las TALEN AAVS-SBI (pZT-AAVS1 L1 y pZT-AAVS R1 Systems Bioscience Cat. N.º GE601A-1) se usó como una mezcla equimolar con la cantidad total por 10^6 células que se muestra en la Tabla 6 a continuación. En muestras sin TALEN añadidas, el ADN de entrada se llevó a $2,5 \mu\text{g}/10^6$ células utilizando el plásmido de control pcDNA3.0.

El porcentaje de eficacia de transfección se calculó contando el número de colonias de blasticidina conseguidas para una entrada dada de células totales. La diferencia de veces en comparación con los controles negativos (es decir, sin adición de ADN de TALEN) se muestra entre paréntesis. Finalmente, el número de colonias transformadas que se puede lograr al procesar un ciclo completo del sistema Maxcyte, que implica la electroporación de 10^{10} células, se calcula en la última columna. Esto representa un ciclo único de aproximadamente 30 minutos, lo que brinda la posibilidad de ejecutar varios ciclos en un día. Por lo tanto, la producción diaria podría ser 5-10 veces mayor. Se pueden usar sistemas de fermentación y cultivo a gran escala tal como el sistema Wavebag (GE Healthcare) o el sistema Celltainer (Celltainer Biotech) para generar células para la transfección, y se pueden usar para cultivar las bibliotecas resultantes.

Tabla 6. Electroporación de células HEK293

Muestra	α FGFR1_A/ α FGFR2_A donante/ 10^6 células	ng ADN	ng ADN TALEN/ 10^6 células	% de transfección	N.º de clones por 10^{10} células
1	580		1920	5,1 (51x)	$5,1 \times 10^8$
1b	580		640	3,3 (33x)	$3,3 \times 10^8$
2	580		-	0,1 -	$0,1 \times 10^8$
3	194		1920	2,7 (89x)	$2,7 \times 10^8$
4	194		-	0,03	$0,03 \times 10^8$
5	1185		1315	5,8 (25x)	$5,8 \times 10^8$
6	1185		-	0,23	$0,23 \times 10^8$
7	1825		675	6,1 (11x)	$6,1 \times 10^8$
8	1825		-	0,57	$0,57 \times 10^8$
9	580 (FGFR1 solo)		1920	5,3	$5,3 \times 10^8$
10	580 (FGFR2 solo)		1920	5,3	$5,3 \times 10^8$
Muestra	ADN donante de β -galactosidasa/ 10^6 células		ng ADN TALEN/ 10^6 células	% de transfección	N.º de clones por 10^{10} células
11	580		1920	4,5	$4,5 \times 10^8$
12	580		-	0,21	$0,21 \times 10^8$
13	1185		1315	5,5	$5,5 \times 10^8$
14	1185		-	0,21	$0,21 \times 10^8$

Este ejemplo demuestra que es posible crear bibliotecas muy grandes de células con casetes de anticuerpos integrados. La eficiencia de transfección varió del 2,7 al 6,1 %. En el caso de la población seleccionada de β -galactosidasa (muestra 13), se puede crear una biblioteca de $5,5 \times 10^8$ clones en una sola sesión de electroporación de flujo. Con más de una sesión en un día, se puede generar una biblioteca de $2-5 \times 10^9$ clones.

Después de 13 días de selección con blasticidina (10 µg/ml), las células se marcaron con anticuerpo anti-Fc marcado con ficoeritrina (Biolegend, Cat. N.º 409304) como se ha descrito anteriormente. De la población de anticuerpos seleccionada en la β -galactosidasa, el 34-36 % de las células fueron positivas para la expresión de Fc y el 11-13 % fueron positivas para la unión del antígeno marcado con Dyelight-633 a una concentración de 10 nM.

Cuando se utilizaron clones de unión a FGFR, el 98-99 % de las células fueron positivas para la expresión de Fc. El uso de una mezcla de anticuerpos α -FGFR1_A y α -FGFR2_A brinda la oportunidad de examinar la proporción de células que contienen múltiples eventos de integración. Para una célula individual con un casete correctamente integrado (por ejemplo, α -FGFR1_A), hay aproximadamente una probabilidad de 50:50 de que una segunda integración sea de la especificidad alternativa (es decir, α -FGFR2_A). Si hay integraciones múltiples frecuentes, entonces la proporción de clones de doble positivo será alta, sin embargo, no se encontró que la proporción de clones de doble positivo sea alta, lo que ilustra la fidelidad del sistema de integración de bibliotecas dirigida por nucleasas en la generación de un gen de anticuerpo/por célula. Se confirmó la capacidad de los anticuerpos anti-

FGFR presentados en la superficie para unirse específicamente a su antígeno apropiado. La expresión del antígeno fue de un plásmido pTT3DestrCD4(d3+4)-His10 [134] que codificaba el ectodominio Fgfr1 de ratón (ENSMUSP00000063808). Esto se usó para transfectar células de suspensión HEK293 y se secretó Fgfr1-rCd4-His10 purificado por cromatografía de afinidad de metales inmovilizados como se ha descrito anteriormente [134]. El ectodominio Fgfr2 de ratón se amplificó por PCR a partir del clon IMAGE 9088089 utilizando cebadores:

2423 (TTTTTTCCATGGGCCGCGCCCTCCTTCAGTTTAGTTGAG) (SEQ ID NO: 97)

y

2437 (TTTTTTGCGGCCGCGGAAGCCGTGATCTCCTTCTCTCTC) (SEQ ID NO: 98),

digerido con NcoI/NotI y clonado en el plásmido de expresión pBIOCAM5 [126]. Fgfr2-Fc se expresó por transfección transitoria de células HEK293 como se ha descrito previamente [134] y se purificó por cromatografía de afinidad.

Las poblaciones transfectadas se probaron para determinar la unión dual utilizando ambos antígenos marcados y la proporción de dobles positivos fue baja. En este experimento, el equilibrio óptimo del tamaño de la biblioteca ($2,7 \times 10^8$ clones/por sesión de Maxcyte) con un bajo porcentaje de dobles positivos (3,5 %) se encontró utilizando 197 ng de ADN donante por 10^6 células (Figura 23).

Esta proporción de dobles positivos podría representar una mala incorporación de un segundo casete de anticuerpo, pero dada la eficacia de la integración de la biblioteca dirigida por nucleasas, también es posible que ambos alelos (el locus AAVS en este ejemplo) se puedan dirigir con aglutinantes entrantes en una proporción de células. La presencia de dos genes de anticuerpos diferentes dentro de una célula en sí misma no impide el aislamiento de los aglutinantes o sus genes codificantes, pero esto se puede evitar modificando primero la célula diana en un solo locus para introducir un único sitio de direccionamiento de nucleasas, por ejemplo, un sitio de meganucleasa SceI integrado previamente como se demostró en el ejemplo 9.

Ejemplo 16. Recuperación de genes que codifican aglutinantes de poblaciones de bibliotecas clasificadas

Las selecciones de presentación de fagos se realizaron en β -galactosidasa (como se describe en el ejemplo 6) y las poblaciones de anticuerpos de 1-2 rondas de selección se clonaron en el vector pD6 y se introdujeron en el locus AAVS de células HEK293 como se describe en el ejemplo 6. La β -galactosidasa se marcó usando Lightning Link Dyelight-633 (Innova Bioscience Cat. N.º 325-0010) de acuerdo con las instrucciones del fabricante. Las poblaciones de células transfectadas se seleccionaron durante 25 días en blasticidina (10 μ g/ml) y se marcaron con β -galactosidasa marcada con Dyelight-633 10 nM y anti-Fc marcado con ficoeritrina (Biolegend Cat. N.º 409304). Las células se incubaron con anticuerpos durante 30 minutos a 4 °C, se lavaron dos veces en PBS/BSA al 0,1 %, se resuspendieron en PBS/BSA al 0,1 % y se clasificaron las células de doble positivo usando un clasificador de flujo.

Las células clasificadas se expandieron y se llevó a cabo una segunda ronda de clasificación utilizando antígeno 10 nM. Las células se cultivaron y se aisló ADN genómico o ARNm de las poblaciones clasificadas y seleccionadas.

Cuando los aglutinantes que incluyen diferentes cadenas (por ejemplo, anticuerpos con formato IgG) están presentes en la misma secuencia genómica (por ejemplo, por introducción en el mismo plásmido) pero se transcriben en diferentes ARNm, puede ser óptimo recuperar los genes separados que codifican el aglutinante multimérico por amplificación a partir de ADN genómico. Como alternativa, los aglutinantes que incluyen múltiples cadenas de proteínas se pueden codificar en el mismo ARNm mediante el uso de elementos de "secuencia de entrada interna del ribosoma" (IRES) o secuencias tales como las secuencias virales P2A o T2A que promueven el estancamiento de la traducción/la escisión de proteínas [133, 136]. En este caso y en el caso de los aglutinantes codificados en una sola cadena de proteína, también será posible aislar los genes codificados del ARNm.

El ADN genómico se preparó usando el "kit de tejido y sangre DNeasy" (QIAGEN Cat. N.º 69504). El ARNm se preparó utilizando un "mini kit de ARN del aislado II" (Bioline Bio-52072). Para la amplificación del ADN genómico, se configuró una reacción de PCR utilizando la polimerasa Phusion con la mezcla "2xPhusion GC" de acuerdo con las instrucciones del fabricante. Los cebadores que flanquean los sitios de clonación Nco1 y Not1, por ejemplo, los cebadores 2622 (GAACAGGAACACGGAAGGTC) (SEQ ID NO: 99) y 2623 (TAAAGTAGGCGGTCTTGAGACG) (SEQ ID NO: 82) se utilizaron para amplificar el casete de anticuerpos (98 °C durante 10 s, 58 °C durante 20 s, 72 °C durante 90 s durante 35 ciclos).

Los genes que codifican genes scFv seleccionados se amplificaron a partir de ARNm. El ARN total se aisló de las células clasificadas utilizando el "mini kit de ARN del aislado II" (Bioline Cat. N.º Bio-52072). El ADNc se sintetizó a partir de 2 μ g de ARN utilizando la transcriptasa inversa Superscript II (Life Technologies, Cat. N.º 180064-022). Los genes scFv seleccionados se amplificaron entonces a partir del ADNc mediante PCR utilizando la ADN polimerasa de inicio en caliente KOD (Merck Millipore Cat. N.º 71086-3) utilizando cebadores que flanquean los sitios de clonación Nco1 y Not1. En este caso, los cebadores:

41679 ATGAGTTGGAGCTGTATCATCC (SEQ ID NO: 100) y

2621 GCATTCCACGGCGGCCGC (SEQ ID NO: 101)

se usaron para amplificar el casete de anticuerpos (95 °C durante 20 segundos, 60 °C durante 10 segundos, 70 °C durante 15 segundos, para 25 ciclos). Los productos de PCR se digirieron con Nco1 y Not 1 antes de la clonación en el sitio Nco/Not1 del vector de expresión del anticuerpo bacteriano pSANG10. La construcción de pSANG10, los métodos para la expresión bacteriana y el cribado por ELISA se describen en Martin et al 2006 [127].

El cribado por ELISA de la población de la ronda 1 de selección por presentación de fagos reveló que 0/90 de los clones eran positivos. Por el contrario, cuando esta misma población se sometió a presentación de mamíferos y la población del gen scFv se recuperó y se cribó, 27/90 clones (30 %) fueron positivos en ELISA. Esto ilustra que es posible llevar a cabo la presentación de mamíferos en una biblioteca y recuperar una población enriquecida de aglutinantes.

El Ejemplo 15 describe la introducción de la población seleccionada por la ronda 1 de presentación de fagos en β -galactosidasa en células HEK usando electroporación de flujo. La población de células de presentación de mamíferos se seleccionó en blastidina como anteriormente y se sometió a una clasificación por flujo usando 10 nM de β -galactosidasa marcada. Después de 9 días de crecimiento, se descubrió por citometría de flujo que el 75 % de las células eran positivas para la unión de β -galactosidasa utilizando β -galactosidasa 10 nM). Estas se clasificaron y se expandieron adicionalmente. Para ilustrar el potencial de impulsar la rigurosidad, el etiquetado se llevó a cabo utilizando concentraciones de antígeno de 1 nM o 10 nM. El 20,3 % y el 55,9 % de las células respectivamente se clasificaron de cada población. Después de la clasificación, el ARNm se preparó inmediatamente a partir de la población clasificada sin cultivo celular adicional. Tras la clonación, la expresión en bacterias y el cribado por ELISA (como anteriormente), se encontró que la tasa de éxito en ELISA aumentó con la rigurosidad durante la clasificación de flujo. Los clones que muestran el nivel de señal más alto provinieron de este grupo y también se mejoró el número de clones positivos (Figura 24). Esto ilustra la capacidad de controlar la rigurosidad de la selección dentro de las poblaciones de presentación, lo que se refleja en el mejor rendimiento de los anticuerpos resultantes.

Ejemplo 17. Incorporación de genes de receptores de linfocitos T (TCR) para la producción de bibliotecas de mamíferos mediante integración dirigida por nucleasas

Los métodos descritos en el presente documento tienen aplicación más allá de la presentación de anticuerpos. Para demostrar el potencial para cribar bibliotecas de receptores de linfocitos T usando integración dirigida por nucleasas, se construyó una construcción de vector (pINT20) que permitía la expresión de receptores de linfocitos T.

pINT20 (Figura 25a) es un vector promotor dual para dirigirse al locus AAVS humano. Tiene los brazos de homología izquierdo y derecho como se presentan en la Figura 3. El brazo de homología AAVS izquierdo está flanqueado por los sitios únicos AsiSI y Nsi1 y va seguido de un aceptor de corte y empalme y un gen de puromicina. La secuencia entre el extremo del brazo de homología izquierdo y el aceptor de corte y empalme es la misma que la descrita previamente y el gen de puromicina comienza con un ATG en marco con el exón aguas arriba ((Figura 25B y como también se muestra para el gen de blastidina en la Figura 3). La correcta integración dirigida por nucleasas conducirá a un corte y empalme en marco de un gen de puromicina que se corta y empalma a un exón aguas arriba endógeno que a su vez es impulsado por el promotor AAVS endógeno que permite la selección de clones con una integración correcta. El gen de resistencia a la puromicina va seguido de un sitio de poliadenilación SV40. El brazo derecho de homología AAVS está flanqueado por sitios únicos BstZ171 y Sbf1.

pINT20 está configurado con un promotor de pEF (de pSF-pEF, Oxford Genetics Cat. N.º OG43), que permite clonar genes de interés en sitios Nhe1/Kpn1. El sitio Nhe1 está precedido por una secuencia líder de secreción y el sitio Kpn1 va seguido de la señal de poliadenilación de la hormona de crecimiento bovina (bGH poli A) como se muestra previamente en la Figura 2. Aguas abajo se encuentra un promotor de CMV (de pSF-CMV-f1-Pac1, Oxford Genetics Cat. N.º OG111) que permite la clonación a través de Nco y Not (como se muestra en la Figura 2) o Hind3. El sitio Nco1 está precedido por una secuencia líder de secreción y el casete va seguido de un sitio bGH poliA.

Para ilustrar la presentación y el enriquecimiento de los receptores de linfocitos T (TCR), se usó un TCR que reconoce un marcador de cáncer descrito por Li *et al.* (2005) y más tarde por Zhao *et al.* (2007) [137, 138]. Este TCR denominado c12c2 reconoce el péptido SLLMWITQV (SEQ ID NO: 102) presente en HLA-A2 con una afinidad de 450 nM. Este péptido representa los restos 157-165 de NY-ESO-1 (NY-ESO-1 157-165). Esta es una variante de maduración por afinidad de un anticuerpo parental llamado 1G4 con una afinidad de 32 μ M.

Se utilizó un segundo TCR que reconoce otro marcador de cáncer. El TCR de MEL5 parental reconoce el péptido MART-1 26-35 ("antígeno de melanoma reconocido por los linfocitos T 1") presente en HLA-A2 con la secuencia peptídica ELAGIGILTV (SEQ ID NO: 103). Este TCR maduró por afinidad mediante la presentación en fago para dar el clon α 24/ β 17 con una afinidad de 0,5 nM descrita en Madura *et al.* (2013) [139]. La estructura del complejo entre TCR y el complejo MHC:péptido se resolvió (pdb clon 4JFH) y el clon se denomina en lo sucesivo "4JFH". El mismo TCR parental también se ha modificado basándose en el diseño de Pierce *et al.* (2014) [140] y la estructura del complejo se resolvió.

Según Debets *et al.*, la unión del dominio CD3 ζ como se muestra tiende a causar la asociación del gen heterólogo incluso en presencia de TCR nativos humanos [141, 142]. El elemento CD3 ζ utilizado está compuesto por un

dominio extracelular, un dominio transmembrana y un dominio citoplásmico completo. Además, la sustitución de dominios constantes humanos por dominios constantes de ratón en los genes heterólogos también tiende a dirigir su asociación sobre la asociación con dominios constantes humanos endógenos [143]. Finalmente, el uso de dominios constantes de ratón ofrece la opción de detectar las cadenas de TCR heterólogas contra un fondo de TCR humanos. Estos elementos se incorporaron en el diseño del casete de expresión de TCR.

Se diseñaron y se sintetizaron dos genes sintéticos dando lugar a construcciones génicas con la siguiente estructura:

10 $V\alpha$ de TCR humano-CD3 ζ constante de ratón α -humano
 $V\beta$ de TCR humano-CD3 ζ constante de ratón β -humano

La secuencia del gen sintético que incorpora las construcciones de cadena α y de cadena β que incorporan los dominios variables de c12/c2 se muestra en la Figura 25c y d. Estos se clonan en los sitios Nhe1/Kpn1 y los sitios Nco1/Hind3 de pINT20 respectivamente. La construcción que codifica este TCR se denomina pINT20-c12/c2. En el primer caso, el gen sintético se diseñó para incorporar un dominio $V\alpha$ C α (flanqueado por los sitios Nhe1 y Not 1) y el dominio $V\beta$ (flanqueado por los sitios Nco1/Xho1 que codifican el TCR c12/c2). Estos elementos pueden reemplazarse por TCR alternativos que utilizan estos sitios de restricción.

20 Se crearon dos genes sintéticos adicionales que codificaban los dominios $V\alpha$ y $V\beta$ de 4JFH [139] (Figura 25e y f). La construcción que codifica este TCR se denomina pINT20-4JFH.

Se transfectaron 10^7 células HEK293 usando 3 μ g de pINT20-c12/c2 y pINT20-4JFH como ADN donante (300 ng de ADN donante por 10^6 células) en las relaciones que se muestran en la Tabla 7. pINT20-c12/c2 se denomina TCR1 y pINT20-4JFH se denomina TCR2. Se añadieron 5 μ g de cada uno de TALEN pZT-AAVS1 L1 y pZT-AAVS R1 a 10^7 células (500 ng cada vez por 10^6 células) con la excepción de la muestra 6 en la que esto reemplazó por 10 μ g de ADN de control (pcDNA3.0). Se introdujo ADN por transfección con polietilenimina, como se ha descrito anteriormente.

30 Tabla 7

Muestra	Relación TCR1/TCR2
1	1:100
2	100:1
3	50:50
4	100 % de TCR1
5	100 % de TCR2
6	50:50 (sin nucleasa TALE)

Después de 12 días en selección (0,25 μ g/ml de puomicina), las células se marcaron con antígeno diana. Los complejos péptido:MHC reconocidos por los TCR descritos anteriormente se presentan como un pentámero marcado con ficoeritrina (ProImmune). c12/c2 reconoce el péptido SLLMWITQV presente en HLA-A2 que representa NY-ESO-1 157-165 (código de producto ProImmune 390). 4JFH (también conocido como α 24/ β 17) reconoce el péptido ELAGIGILTV presente en HLA-A2 que representa MelanA/MART 26-35 (código de producto ProImmune 082). En cada caso, el complejo MHC:péptido se marcó con ficoeritrina y se usó de acuerdo con las instrucciones del fabricante. La Figura 26 (a-d) muestra que cada TCR es específico para el complejo MHC:péptido esperado (a, d) y no se une al péptido no relacionado en el complejo (Figura 26b, c). El ADN que codifica cada TCR se mezcló con un exceso de 100 veces del ADN que codifica el otro (muestras 1-2, Tabla 7). Las células HEK293 se transfectaron y se seleccionaron en puomicina. La clasificación de las células positivas para el antígeno se realizó después de 14 días de selección con puomicina (Figura 26 g, h).

Para cuantificar el nivel de enriquecimiento dentro de las poblaciones de salida clasificadas por flujo, los genes TCR se recuperaron mediante amplificación por PCR y las cantidades relativas de cada especie de TCR se determinaron después de la clonación. El ARN total se aisló de la población clasificada. La síntesis de ADNc se realizó como se describe en el ejemplo 16. Los cebadores para amplificar la cadena alfa y beta de TCR fueron 1999/2782 y 41679/2789 respectivamente (Tabla 8). La amplificación por PCR empleó la polimerasa de inicio en caliente KOD utilizando el protocolo recomendado por el fabricante (EMD Millipore, 71086, EMD Millipore). Las condiciones de reacción de la PCR fueron 95 °C (2 min) y 25 ciclos de 95 °C (20 s), 60 °C (10 s), 70 °C (15 s) seguidos de 70 °C (5 min). Las cadenas alfa y beta de TCR amplificadas se digirieron con Nhe1/NotI o Nco1/XhoI y se subclonaron en vectores con sitios compatibles (en este caso, vectores de corte pBIOCAM1-Tr-N Nhe1/NotI o pBIOCAM2-Tr-N (Nco1/XhoI) respectivamente). La PCR de los clones individuales se amplificó con un cebador de TCR alfa específico para c2c12 (TCR1) (2781) o con cebadores TCR alfa específicos para 4JFH (TCR2) (4JFH- $V\alpha$ -F) y cebadores específicos de vector para analizar la identidad del clon de TCR. Después de clasificar la muestra 1 (Tabla 7) donde se empleó una relación de 1:100 de plásmido donante TCR1/TCR2 (Tabla 7) con enriquecimiento para clones

específicos de TCR1 utilizando el péptido SLLMWITQV presente en HLA-A2 (Proimmune, código de producto 390), la proporción de clones TCR1, según lo determinado por el PCR de colonias, aumentó a 11/15 (73 %). El enriquecimiento mediante la clasificación de la muestra 2 donde se empleó una relación de 100:1 de plásmido donante TCR1/TCR2 (Tabla 7), con el péptido ELAGIGILTV presente en HLA-A2 (código de producto Proimmune 082) dio como resultado un aumento de la proporción de clones TCR2 a 4/15 (27 %), determinado por el PCR de colonias.

Para demostrar la selección de bibliotecas utilizando integración dirigida por nucleasas, se creó una biblioteca de mutantes basada en c12/c2 clonando un repertorio de genes que codifican cadenas alfa de TCR mutantes junto con un repertorio de genes que codifican cadenas beta de TCR mutantes. Tal biblioteca podría crearse utilizando enfoques de mutagénesis dirigida a oligonucleótidos. Por ejemplo, métodos basados en la mutagénesis de Kunkel [144]. Como alternativa y a modo de ejemplo, se utilizó un enfoque de ensamblaje de PCR para crear una cadena alfa de TCR mutante (como un fragmento Nhe1/Kpn1) y una cadena beta de TCR mutante (como un fragmento Kpn/Hind 3, incluido el promotor de CMV) que se clona en el sitio Nhe1/Hind 3 de pINT20. Esto se hizo usando los cebadores que se muestran en la Tabla 8.

Tabla 8. Cebadores utilizados en la construcción de bibliotecas y recuperación de clones

4JFH-V α -F (SEQ ID NO: 104)	ACACACGCTAGCCAGAAAGAGGTGGAACAG
1999 (SEQ ID NO: 87)	AAAAGCAGGCTACCATGAGGGCCTGGATCTTCTTTCTCC
2770 (SEQ ID NO: 105)	CAAAGAACAGCTCGCCGTSNNCCCGASSNNGGAGCTGG CGCAAAAGTAC
2771 (SEQ ID NO: 106)	CTCGCCCGAAGGTGGGAATGTANGWTCCSNNSNNAAGTG GGCGCACGGCGCAC
2780 (SEQ ID NO: 107)	CTGGCAGCTAGCAAGCAGGAAG
2781 (SEQ ID NO: 108)	TACATCCCACCTTCGGGCGAG
2782 (SEQ ID NO: 109)	TTTTTTGCGGCCGCGGACAGGTTCTG
2783 (SEQ ID NO: 110)	CGTAAGCTGGTACCTTATTATCTAGGG
2785 (SEQ ID NO: 111)	CCCTAGATAATAAGGTACCAGCTTACG
2787 (SEQ ID NO: 112)	ACCGGCGAGCTGTTCTTTG
2788 (SEQ ID NO: 113)	AGTGACAAGCTTTTATTATCTGGGTG
2789 (SEQ ID NO: 114)	CAGGTCCTCGAGCACTGTC
41679 (SEQ ID NO: 100)	ATGAGTTGGAGCTGTATCATCC
(N = A, C, G, T. S = C O G, W = A O T)	

Se diseñó un oligonucleótido mutante (cebador 2771) que aleatorizaba 2 posiciones de aminoácidos dentro de CDR3 de la cadena alfa c12/c2 y se proporciona la opción de serina o treonina en otra posición (el cebador 2771 también está representado por la cadena inferior de Figura 25 g). El cebador 2771 se usó junto con el cebador 2780 para crear un repertorio de TCR alfa mutante que va desde el sitio de clonación Nhe1 incorporando la región de mutagénesis CDR3 con una secuencia invariante en el extremo. El cebador 2781 es complementario al extremo 5' invariante del cebador 2771. La PCR con los cebadores 2781 y 2783 proporcionó el resto del casete de TCR alfa-CD3 zeta hasta el sitio Kpn1. El ensamblaje de PCR de los 2 fragmentos de PCR se utiliza para crear el fragmento TCR alfa-CD3 zeta que se puede clonar en pINT20 después de la digestión con Nhe1 y Kpn1.

Se diseñó un segundo oligonucleótido mutante (cebador 2770) que asignó al azar 2 posiciones de aminoácidos dentro de CDR3 de la cadena beta c12/c2 y proporciona la opción de valina o leucina en otra posición (cebador 2770 representado por la cadena inferior de la Figura 25h). El cebador 2770 se usó junto con el cebador 2785 para crear un repertorio beta de TCR mutante del sitio de clonación Kpn1 que incorpora la región de mutagénesis CDR3 con una secuencia invariante en el extremo. El cebador 2787 es complementario al extremo 5' invariante del cebador 2700. La PCR con los cebadores 2787 y 2788 proporcionó el resto del casete de TCR beta-CD3 zeta hasta el sitio Hind 3. El ensamblaje de PCR de estos 2 fragmentos de PCR se utiliza para crear un fragmento de TCR beta-CD3 zeta que se puede clonar en pINT20. Se creó un repertorio completo que incorporaba mutaciones tanto en la CDR 3 de las cadenas alfa como beta mediante la clonación del fragmento Nhe1/Kpn, el fragmento Kpn1/Hind3 en pINT20 digerido con Nhe1/Hind3. Después de la ligadura, la biblioteca se clonó en células DH10B electrocompetentes. Se preparó ADN plasmídico y el ADN se transfectó en las células HEK293 junto con vectores que codifican la nucleasa TALE como se ha descrito anteriormente (una mezcla equimolar de pZT-AAVS1 L1 y pZT-AAVS R1 Systems

Bioscience Cat. N.º GE601A-1).

Después de la ligadura de las cadenas alfa y beta mutantes de la biblioteca mutante c12/c2 en pINT20, se electroporó el ADN en células DH10B, se preparó ADN plasmídico y la biblioteca se cotransfectó con nucleasa TALE que se dirigía al locus AAVS. La transfección se realizó mediante electroporación Maxcyte. El crecimiento y la selección fueron como se ha descrito anteriormente. La cuantificación del tamaño de la biblioteca, mediante la titulación de células transfectadas y la selección basada en placa en puromicina, indicó que se creó un tamaño de biblioteca de 5×10^5 . Después de la selección de puromicina durante 11 días, las células se marcaron con un anticuerpo marcado con APC específico para la cadena β del TCR de ratón (Life Technologies Cat. H57-957). La Figura 26i-j muestra que el 38 % de los clones en la población expresan un receptor de linfocitos T. De esta población con TCR positivo, el 13 % también se unía al péptido 1 (5 % de la población total). Mediante este enfoque, se pueden aislar clones con expresión mejorada o actividad de unión a péptido:MHC.

A partir de la Figura 26 se puede ver que cada receptor de linfocitos T reconoció solo su antígeno análogo. Además, cuando se utiliza una mezcla de las dos especificidades diferentes, es posible distinguir cada una de ellas a través del antígeno marcado. Este enfoque también permite que los clones de TCR con afinidad (o expresión) mejorada se distingan dentro de la biblioteca de TCR mutante al identificar un subconjunto de la biblioteca que fue etiquetado en mayor medida que el clon parental

Los genes del receptor de linfocitos T también se introdujeron en células Jurkat por electroporación. Las células Jurkat se centrifugaron y se resuspendieron en un volumen final de 10^8 células/ml en el tampón de electroporación del fabricante (tampón de electroporación Maxcyte, Thermo Fisher Scientific Cat. n.º NC0856428)). Se añadió una alícuota de 4×10^7 células (0,4 ml) a la cubeta de electroporación OC400 con 40 μg de ADN (es decir, 1 $\mu\text{g}/10^6$ células). El ADN consistió en una mezcla de ADN de plásmido donante (biblioteca de TCR pINT20-c12/c2 o pINT20-4JHF o pINT20-c12/c2, 9,2 μg) y una mezcla equimolar de ADN (30,8 μg en total) de ADN que codifica las TALEN AAVS-SBI (pZT-AAVS1 L1 y pZT-AAVS R1 Systems Bioscience Cat. N.º GE601A-1). En muestras sin TALEN añadidas, el ADN de entrada se llevó a 1 $\mu\text{g}/10^6$ células utilizando el plásmido de control pcDNA3.0. Un método alternativo para introducir genes de receptores de linfocitos T en células Jurkat utiliza el 4D-Nucleofector (Lonza). Aquí, el protocolo de transfección siguió las instrucciones del fabricante según el kit de línea celular SE (Lonza, Cat. PBC1-02250). En resumen, 2 μg de ADN, que consiste en una mezcla de ADN de plásmido donante (biblioteca de TCR pINT20-c12/c2 o pINT20-4JHF o pINT20-c12/c2, 0,46 μg) y una mezcla equimolar de ADN (1,54 μg en total) de ADN que codifica las TALEN AAVS-SBI (pZT-AAVS1 L1 y pZT-AAVS R1 Systems Bioscience Cat. N.º GE601A-1) se transfectó por 10^6 células Jurkat. La configuración del código de pulso fue CL120 y el programa de tipo de célula fue específico para las células Jurkat E6.1 (ATCC).

La Figura 26 demuestra que se logró la expresión de TCR y se logró el reconocimiento del péptido apropiado:molécula de MHC. Esto dependía del uso de la nucleasa TALE (comparar las Figuras 26m y n). La señalización a través del TCR introducido también se logró utilizando el péptido relevante:molécula MHC.

Se sembraron células Jurkat transfectadas con pINT20-c12/c2 o células Jurkat no transfectadas en una placa de 96 pocillos a una densidad de $1 \times 10^6/\text{ml}$, 200 μl por pocillo. Las células se estimularon con 2 μl o 6 μl por pocillo de péptido 1 marcado con PE-pentámero MHC (Prolimmune) o 2 $\mu\text{g}/\text{ml}$ de CD3 antihumano (BD Pharmingen, Cat. 555329) en presencia y ausencia de CD28 antihumano (BD Pharmingen Cat. 555725) a 2 $\mu\text{g}/\text{ml}$. Después de una incubación de 24 horas a 37 °C y CO_2 al 5 %, se detectó la activación de células Jurkat investigando la expresión de CD69. Las células se tiñeron con 50 μl de PBS + BSA al 1 % + 0,5 μl de CD69-APC antihumano (Invitrogen, Cat. MHCD6905) por pocillo durante 45 minutos a 4 °C.

La Figura 26 (muestra o y p) demuestra la regulación positiva de CD69 tras la estimulación con CD3. La figura también muestra el efecto de añadir 2 μl (q) o 6 μl (r) de péptido 1:MHC. Una población de células de doble positivo que se unen al péptido:MHC y expresan CD69 es obvia. Este ejemplo muestra células incubadas en presencia de CD28, pero se observó el mismo efecto en ausencia de CD28 (no mostrado).

Este ejemplo demuestra el potencial de la integración dirigida por nucleasas de bibliotecas de un tipo alternativo de aglutinante, es decir, receptores de linfocitos T. Se demuestra que es posible expresar y detectar la expresión de TCR en la superficie celular utilizando anticuerpos específicos. También se demuestra que estos receptores de linfocitos T reconocen específicamente sus dianas respectivas. También se ha construido una biblioteca de mutantes que permite la selección de aglutinantes mejorados. Finalmente, se ha demostrado que el cribado de bibliotecas basado en la activación de la señalización de TCR en los linfocitos T es posible. Aquí se ha utilizado una línea de linfocitos T humana cultivada. También es posible introducir ADN en linfocitos T primarios mediante electroporación Maxcyte. Los métodos para el aislamiento y la preparación de linfocitos T primarios se conocen por los expertos en la técnica (por ejemplo, Cribbs *et al.*, 2013, Oelke *et al.* 2003 [145, 146]. La exposición de linfocitos transfectados con TCR a péptidos multiméricos:MHC puede usarse entonces para lograr la activación a través de la exposición al multímeros de péptido:MHC [146] o a células presentadoras de antígenos cargadas con el péptido apropiado [146, 147]. La activación puede detectarse a través de la expresión de genes indicadores o a través de la regulación ascendente de genes endógenos tal como CD69 [104, 148].

Ejemplo 18. Presentación de bibliotecas de receptores de antígenos quiméricos en células de mamífero

La activación de los linfocitos T se produce normalmente a través de la interacción del receptor de linfocitos T (TCR) con el complejo péptido:MHC específico. Esto a su vez conduce a la señalización dirigida a través de CD3 y otras moléculas de señalización de linfocitos T. Como alternativa al reconocimiento de dianas dirigido por el TCR, se ha demostrado que las moléculas de unión alternativas, tales como los Fv monocatenarios, se pueden presentar en los linfocitos T como fusiones a las moléculas de señalización aguas abajo de una manera que dirige la activación de los linfocitos T a la molécula reconocida por el scFv (o aglutinante alternativo). De esta manera, la activación de los linfocitos T ya no está limitada al reconocimiento molecular dirigido a los complejos péptido:MHC por los TCR, sino que puede dirigirse a otras moléculas de la superficie celular. Este formato alternativo en el que una entidad de unión diferente de TCR se fusiona con un componente de señalización se denomina "receptor de antígeno quimérico" (CAR). En el caso de los linfocitos T, se ha demostrado que es un medio importante y valioso para redirigir la activación de los linfocitos T.

Para cualquier diana dada, aún no está claro cuáles serán las características de epítipo o afinidad óptimas para la incorporación en un CAR [103]. Las características del diseño de CAR, tal como la longitud del enlazador o la elección del dominio transmembrana, pueden afectar a su vez, lo que constituye un epítipo óptimo. La combinación de la densidad del antígeno en las células diana y no diana junto con la elección del dominio de señalización podría afectar a los requisitos de afinidad óptimos. La capacidad de presentar bibliotecas de receptores de antígenos quiméricos en linfocitos T brinda la oportunidad de identificar la especificidad de unión óptima, el formato de enlace, la longitud/secuencia del enlazador, las variantes del módulo de señalización fusionado, etc., ya sea solo o en combinación. Aquí, se demuestra la utilidad de la integración dirigida por nucleasas para la construcción de bibliotecas de receptores de antígenos quiméricos en células de mamífero. El vector pINT21 (Figura 27a) es un único vector del promotor de CMV para la expresión y secreción convenientes de aglutinantes tal como scFv flanqueados por los sitios de restricción Nco1/Not1 para permitir la expresión en marco con una secuencia líder aguas arriba y un compañero de fusión aguas abajo (como se muestra anteriormente en Figura 8). El casete de expresión CAR en pINT21 está flanqueado por los brazos de homología AAVS como se describe anteriormente en la Figura 3.

El vector pINT21-CAR1 (Figura 27a, c) fusiona aglutinantes tales como Fv monocatenarios al dominio transmembrana y al dominio intracelular de CD3 ζ (Figura 27c y como se describe para la expresión de TCR en la Figura 25). Este formato se suele denominar receptor de antígeno quimérico de "primera generación". Los dominios de señalización de otras moléculas coestimuladoras también se han utilizado para proporcionar señales adicionales y se ha demostrado que éstas proporcionan una señalización mejorada. Se han denominado receptores de antígenos quiméricos de segunda y tercera generación. Por ejemplo, pINT21-CAR2 (Figura 27b, d) fusiona el aglutinante (clonado convenientemente en este caso como un fragmento Nco1/Not 1 a un dominio de segunda generación previamente descrito (documento WO 2012/079000 A1) que consiste en:

El dominio bisagra y transmembrana de CD8
dominio de señalización 4-1BB
dominio de señalización CD3 ζ

A modo de ejemplo, se clonaron varios grupos de aglutinantes diferentes en los sitios Nco/Not1 de pINT20-CAR1 y pINT20-CAR2. Se ha utilizado previamente CD19 en varios estudios diferentes para dirigirse a estudios de neoplasias de linfocitos B en Sadelain *et al.* (2013) [103]. Se preparó un anticuerpo anti-CD19 previamente descrito (documento WO 2012/079000 A1) (denominado FMC63) como un gen scFv sintético en una configuración VH-enlazador-VL o una configuración VL-enlazador-VH (FMC63 H-L o FMC L-H respectivamente, la Figura 27E muestra la secuencia de FMC63 H-L. FMC63 L-H se configuró con los dominios variables en la configuración VL-enlazador-VH flanqueada por Nco1 y Not1 en los extremos 5' y 3' respectivamente.

Como controles, los scFv con especificidades de unión alternativas también se clonaron en pINT20. Estos incluyen anti-FGFR1_A [105] y un anticuerpo de control anti-desmina [7]. Además, los Adhirones [152] que reconocen lox1 (Figura 29a, b) se introdujeron como ejemplo de un formato alternativo de aglutinante configurado como una fusión de CAR (véase el ejemplo 19 para una descripción).

Para demostrar la creación de bibliotecas de aglutinantes presentes en un formato CAR, se clonaron poblaciones de anticuerpos con formato scFv seleccionados en mesotelina y CD229. La mesotelina es una glucoproteína de la superficie celular que se expresa altamente en varios cánceres, incluido el mesotelioma. Se están desarrollando varios formatos basados en anticuerpos y en ensayos clínicos que incluyen CAR dirigidos a la mesotelina [149]. CD229 representa otro posible antígeno asociado a un tumor que podría dirigirse por terapia inmunológica en cánceres tal como leucemia linfocítica crónica y el mieloma múltiple [150, 151].

Se creó una población de anticuerpos que reconocían tanto a la mesotelina como CD229 mediante la selección utilizando la biblioteca de presentación de fagos "McCafferty" como se describe en el ejemplo 6 y la referencia 7). Se llevaron a cabo dos rondas de selección y los genes codificantes de scFv se recuperaron utilizando los cebadores M13leadseq y Notmycseq (ejemplo 6). Los productos se cortaron con Nco1/Not 1, se purificaron en gel y se clonaron

en pINT21-CAR2. Estos se dirigieron al locus AAVS de las células HEK293 por escisión de nucleasa TALE para generar una biblioteca de $4,8 \times 10^5$ para CD229 y $6,4 \times 10^5$ para mesotelina (que representa un aumento de 30x y 53x en la biblioteca en comparación con muestras transfectadas en ausencia de nucleasa TALE).

5 Se introdujeron pINT20-CAR1 y pINT20-CAR2 en células HEK293 por transfección con PEI. Aquí, el ADN plasmídico donante (pINT20-CAR1 o pINT20-CAR2, 6 μg se mezcló con una mezcla equimolar de ADN (20 μg en total) de ADN que codifica las TALEN AAVS-SBI (pZT-AAVS1 L1 y pZT-AAVS R1 Sistemas Bioscience Cat. N.º GE601A-1) en medio Freestyle 293 (Lifetechn, Cat. 12338-026), se añadió PEI lineal (52 μl , 1 mg/ml, Polysciences Inc.) y se incubó a temperatura ambiente durante 10 minutos. Después, la mezcla se añadió a 20 ml de células HEK293 en suspensión (1×10^6 células/ml) en un matraz Erlenmeyer con ventilación de 125 ml. También se introdujeron pINT20-CAR1 y pINT20-CAR2 en células Jurkat por electroporación. Las células Jurkat se centrifugaron y se resuspendieron en un volumen final de 10^8 células/ml en el tampón de electroporación del fabricante (tampón de electroporación Maxcyte, Thermo Fisher Scientific Cat. n.º NC0856428)). Se añadió una alícuota de 4×10^7 células (0,4 ml) a la cubeta de electroporación OC400 con 40 μg de ADN (es decir, 1 $\mu\text{g}/10^6$ células). El ADN consistió en una mezcla de ADN de plásmido donante (pINT20-CAR1 y pINT20-CAR2, 9,2 μg) y una mezcla equimolar de ADN (30,8 μg en total) de ADN que codifica las TALEN AAVS-SBI (pZT-AAVS1 L1 y pZT-AAVS R1 Systems Bioscience Cat. N.º GE601A-1). En muestras sin TALEN añadidas, el ADN de entrada se llevó a 1 $\mu\text{g}/10^6$ células utilizando el plásmido de control pcDNA3.0.

20 El etiquetado fluorescente de los diversos antígenos se realizó usando el kit de conjugación Lightning-Link Rapid Dye-Light 633 (Innova Biosciences, cat. 325-0010). La preparación de FGFR1 y FGFR2 se describe en el ejemplo 15. Lox1 y CD229 eran de R and D Systems (Cat. N.º 1798-LX-050, y 898-CD050 respectivamente), CD19-Fc y la mesotelina eran de (AcroBiosystems Cat. N.º CD9-H5259 y MSN-H526x respectivamente).

25 La Figura 28b ilustra la presentación del anticuerpo anti-FGFR1 dirigido por nucleasa [105] dentro de una construcción CAR de segunda generación (pINT21-CAR2-FGFR1_A). La Figura 28d también ilustra la presentación de una molécula de andamiaje alternativa (una referencia de Adhirón [152]) como una fusión con un receptor de antígeno quimérico de segunda generación (pINT21-CAR2-lox1). La Figura 28f y g también ilustran clones positivos dentro de una biblioteca de scFv seleccionados en mesotelina o CD229.

30 En este ejemplo, se introdujeron CAR en células HEK y células Jurkat, pero esto también se podría hacer mediante la introducción de las construcciones en células primarias tales como los linfocitos T humanos (por ejemplo, como se describe por Sadelain *et al.* (2013) [103], por ejemplo, usando electroporación [135]. Las construcciones de expresión para la expresión de CAR en linfocitos pueden optimizarse aún más, por ejemplo, optimizando la estabilidad y la traducción del ARNm a través de la variación en las regiones no traducidas 5' y 3', longitud poli A, etc., como se ha descrito previamente [135]. La señalización de construcciones CAR introducidas en linfocitos T primarios o líneas celulares de linfocitos T puede inducirse por exposición a células que expresan antígeno diana o que usan antígeno multimérico, por ejemplo, antígeno inmovilizado en una superficie o presente en perlas [104, 148].

40 Ejemplo 19. Presentación de bibliotecas de andamiajes alternativos construidos en células de mamífero a través de integración dirigida por nucleasas

45 El método descrito para construir bibliotecas de aglutinantes puede emplearse más allá de la presentación de anticuerpos y receptores de linfocitos T. Se han descrito varios andamiajes alternativos que permiten la construcción de bibliotecas de variantes a partir de las cuales se han aislado nuevas especificidades de unión, por ejemplo, Tiede *et al.* (2014) [152] y referencias en los mismos. En el ejemplo descrito por Tiede *et al.* (2014) se utilizó un andamiaje estable y versátil basado en una secuencia consenso de fitocistatinas derivadas de plantas. Este andamiaje se conoce como Adhirón y la Figura 29a muestra un gen sintético que codifica un Adhirón que se seleccionó para unirse a lox1 (documento (WO 2014125290 A1). La Figura 29B muestra un aglutinante alternativo de lox1 (lox1B). Ambos se sintetizaron y se clonaron en el sitio Nco1/Not1 de pINT20_CAR2 para crear una fusión con el compañero aguas abajo.

55 Se puede construir una biblioteca mediante la aleatorización de restos de bucle (por ejemplo, mediante mutagénesis de Kunkel o ensamblaje de PCR como se ha descrito anteriormente en el ejemplo 17). A modo de ejemplo, la Figura 29c muestra el diseño de un oligonucleótido mutante útil para crear una biblioteca siguiendo el mismo enfoque descrito en el ejemplo 17. En este caso, la asignación al azar se logra mediante la introducción de un número variable de restos NNS, aunque podrían usarse estrategias alternativas conocidas por los expertos en la técnica.

60 Como otro ejemplo, la Figura 29d y e demuestra los medios para crear una biblioteca de aglutinantes mediante integración dirigida por nucleasa basada en un andamiaje knottin [156]. Los knottins son péptidos de aproximadamente 30 aminoácidos que están estabilizados por tres enlaces disulfuro, con uno enhebrado a través de los otros dos para crear una estructura "anudada". La Figura 29d muestra el knottin de unión a tripsina MCoTi-II con un sitio Nco1 en el extremo 5' y un sitio Not en el extremo 3' que permite la expresión en marco con los vectores descritos en el presente documento. Como ejemplo para la construcción de bibliotecas, los 6 aminoácidos del primer bucle (subrayados en la Figura 29d) se pueden mutar con un número variable de aminoácidos. La Figura 29e ilustra

una estrategia mutagénica que reemplaza los 6 aminoácidos del bucle 1 con 10 aminoácidos al azar utilizando los codones VNS (donde V = A, C o G y S = C o G). El codón VNS incluye 24 codones que codifican 17 aminoácidos que excluyen las cisteínas. Esta estrategia tiene fines ilustrativos y los expertos en la técnica conocerán estrategias mutagénicas alternativas.

5

Ejemplo 20. Introducción dirigida por nucleasas de bibliotecas de anticuerpos usando CRISPR/Cas9

La integración dirigida por nucleasas a través de CRISPR/Cas9 se demostró usando el "kit de vector de nucleasas CRISPR Geneart" (Lifetech A21175). En este sistema, un promotor de ARN polimerasa III U6 impulsa la expresión de un ARN CRISPR complementario (ARNrc) complementario que está vinculado a un ARNrc trans-activador (ARNtracr). El ARNrc y el ARNtracr juntos forman un ARN guía que dirige la especificidad de escisión de una proteína Cas9 codificada en el mismo "vector de nucleasa CRISPR GeneArt" (véanse las instrucciones del fabricante). El vector se proporciona como un plásmido linealizado en el que se clona un oligonucleótido bicatenario corto con salientes 3' apropiados. La especificidad de escisión se determina entonces por la secuencia del segmento clonado. Se diseñaron dos secuencias de direccionamiento diferentes para la escisión directa al locus AAVS humano descrito anteriormente.

Las secuencias fueron:

Inserto de ADN bicatenario CRISPR1:

20

5' GGGGCCACTAGGGACAGGATGTTTT (SEQ ID NO: 115)
3' GTGGCCCCCGGTGATCCCTGTCCTAC (SEQ ID NO: 116)

Inserto de ADN bicatenario CRISPR2:

5' GTCACCAATCCTGTCCCTAGGTTTT (SEQ ID NO: 117)
3' GTGGCCAGTGGTTAGGACAGGGATC (SEQ ID NO: 118)

Los ARN de guía resultantes se dirigen a la escisión dentro de la misma región del locus AAVS que las nucleasas TALE descritas anteriormente (estando CRISPR2 en la orientación inversa de CRISPR1). Por lo tanto, los brazos de homología AAVS utilizados previamente para la integración directa del casete de expresión se pueden usar para la integración dirigida por estos ARN guía CRISPR. El vector linealizado y los oligonucleótidos bicatenarios se ligaron y se transformaron en células DH10B electrocompetentes. La clonación del inserto correcto se confirmó mediante secuenciación y se preparó el ADN plasmídico. La construcción Cas9/CRISPR2 (que incluye el oligonucleótido CRISPR2) se transfectó junto con el ADN donante que codifica una biblioteca de β -galactosidasa seleccionada por una ronda 1 de selección de presentación de fagos (ejemplo 15). Se transfectaron en células HEK293 utilizando el sistema de electroporación Maxcyte con el ensamblaje OC-400. Se transfectaron 4×10^7 células con 23,2 μ g de ADN donante que representa una población de genes scFv seleccionados por una ronda de presentación de fagos en β -galactosidasa clonada en pD6. Las células se cotransfectaron con 77 μ g de plásmido Cas9-CRISPR2, o 77 μ g de plásmido TALEN (38,5 μ g de pZT-AAVS1 L1 y pZT-AAVS) o 77 μ g de plásmido de control.

35

Tabla 9

Muestra n.º	Nucleasa usada	Eficiencia de transfección %	Número de clones por 10^{10} células
1	Cas9/CRISPR 2	10,53 (29x)	10×10^8
2	AAVS TALEN	5,1 (14x)	$5,1 \times 10^8$
3	Ninguna (pCDNA3.0)	0,36	$0,36 \times 10^8$

La valoración del número de transformantes formados por la transfección de Cas9/CRISPR2 (mediante la medición de las colonias de resistencia a la blastidina) reveló que se generaron 1053 colonias resistentes a la blastidina a partir de la siembra de 10.000 células, lo que equivale a una eficiencia de transfección del 10,5 % (Tabla 9). En el caso de dirigirse por nucleasas TALE, se logró una eficiencia de transfección del 5,1 %. Por el contrario, en ausencia de la construcción Cas9/CRISPR2, solo se logró una eficiencia de transfección del 0,36 %.

45

Como alternativa a la transfección utilizando ADN plasmídico para introducir la proteína Cas9 y el ARN guía en las células, también es posible introducir directamente un complejo de nucleoproteínas formado por la proteína Cas9 (Toolgen, Inc.) y un ARN guía. El ARN guía fue preparado por Toolgen, Inc. utilizando la transcripción *in vitro* de un promotor T7 como una sola transcripción que incluía la secuencia TRACR (subrayada) precedida por una secuencia complementaria al ADN diana (en negrita) como se muestra a continuación.

50

ARN CRISPR 1 (SEQ ID NO: 119):

5'

GGGGGGCCACUAGGGACAGGAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCU
AGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

ARN CRISPR 2 (SEQ ID NO: 120):

5' **GGGUCACCAAUCCUGUCCCUAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGC**
 5 UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

Se introdujeron 6,6 µg de proteína Cs9, 4,6 µg de ARN y 10 µg de ADN donante (que codifica el anticuerpo anti-FGFR1_A en pD6) en 10⁷ células HEK293 mediante electroporación Maxcyte como se ha descrito anteriormente. Se lograron eficiencias de transfección del 2,2 % y del 2,9 % para ARN CRISPR 1 y CRISPR2, respectivamente, con un 0,7 % y un 0,8 % en ausencia del complejo añadido Cas9:proteína ARN.

Estos ARN de guía se dirigen a las mismas secuencias codificadas por CRISPR1 y CRISPR2. Como alternativa, el ARNcr y el ARNtracr se pueden producir por síntesis química (por ejemplo, GE Dharmacon).

15 Ejemplo 21: Inserción de genes de anticuerpo mediada por nucleasas por ligadura o unión de extremos mediada por microhomología (MMEJ)

Aunque la recombinación homóloga (HR) es útil para la inserción precisa de grandes fragmentos de ADN, requiere la construcción de grandes vectores de direccionamiento que incorporen brazos de homología largos. Esto puede hacer que la construcción de grandes bibliotecas sea más difícil debido a la reducción de la eficiencia de transformación de construcciones de ADN más grandes. Como alternativa, pueden producirse reacciones de ligadura sencillas entre el ADN cromosómico y el vector de direccionamiento, si se incorpora una secuencia de reconocimiento de nucleasa en el vector de direccionamiento. Las reacciones de ligadura pueden ser de "extremo adhesivo" empleando, por ejemplo, nucleasas con dedos de cinc (Orlando *et al.*, 2010) [45] o TALEN (Cristea *et al.*, 2013) [22] que pueden hacer roturas bicatenarias (DSB) que dejan salientes 5' o "extremos romos" que emplean la ribonucleoproteína CRISP/Cas9. Un ejemplo de integración del gen de nucleasa por ligadura utilizando meganucleasa I-SceI se mostró mediante la construcción del vector pD7-Sce1. pD7 se deriva de pD6 (Figura 8) pero los brazos de homología AAVS izquierdo y derecho se reemplazaron con oligonucleótidos bicatenarios cortos. El brazo izquierdo de homología AAVS de la serie de vectores pD está flanqueado por las enzimas de restricción EcoR1 y Nsi1 (véase la Figura 3). Para convertir pD6 en pD7-Sce1, se reemplazó por un inserto de oligonucleótido bicatenario formado por los cebadores 2778 y 2779 que codifican una secuencia de reconocimiento de meganucleasa I-SceI con "extremos adhesivos" compatibles con los "extremos adhesivos" formados por EcoRI/NsiI. El brazo de homología AAVS derecho está flanqueado por los sitios de restricción Asc1 y Mlu1 (Figura 3). El brazo de homología derecho se reemplazó por un inserto de oligonucleótido bicatenario con "extremos adhesivos" compatibles con los "extremos adhesivos" formados por la digestión con AscI/MluI y se forma por los cebadores 2723 y 2724.

Tabla 10. Cebadores para la construcción de pD7 y pINT19

2723 (SEQ ID NO: 121)	CGCGCCAGAAGTCTCACCAAGCCCA
2724 (SEQ ID NO: 122)	CGCGTGGGCTTGGTGAGACTTCTGG
2768 (SEQ ID NO: 123)	AATTCTCCCCTCCACCCCACAGTAGGGACAGTGGGGCCA GGATTGGTGACAGAAAATGCA
2769 (SEQ ID NO: 124)	TTTTCTGTCACCAATCCTGGCCCCACTGTCCCTACTGTGG GGTGGAGGGGAG
2778 (SEQ ID NO: 125)	AATTCTAGGGATAACAGGGTAATATGCA
2779 (SEQ ID NO: 126)	TATTACCCTGTTATCCCTAG
2808 (SEQ ID NO: 127)	AATTCTTTTCTGTCACCAATCCTGGGGCCACTAGGGACAC TGTGGGGTGGAGGGGATGCA
2809 (SEQ ID NO: 128)	TCCCCTCCACCCACAGTGTCCCTAGTGGCCCCAGGATTG GTGACAGAAAAGAATTG

Los anticuerpos que reconocen Fgfr1 y Fgfr2 (ejemplo 15) se clonaron en pD7 para crear pD7-Scel anti-Fgfr1 y pD7-Scel anti-Fgfr2 respectivamente. Estos se cotransfectaron con el plásmido de expresión I-Sce1 (ejemplo 11, Figura 16) en la línea celular del clon 6K HEK293 (véase el ejemplo 11) que contiene un sitio de reconocimiento I-Sce1 integrado.

La ligadura de las DSB en el cromosoma y el vector de direccionamiento generado por la nucleasa con dedos de cinc o las nucleasas TALE también se puede lograr. Al invertir los sitios de reconocimiento de la nucleasa con dedos de cinc o TALEN en el vector de direccionamiento, esto puede garantizar que el producto de la inserción ya no sea una diana para la escisión en un método denominado "recombinación activada por ligadura obligatoria" u ObLiGaRe (Maresca *et al.*, 2013) [153]. Los vectores pD7-ObLiGaRe se pueden generar de la misma manera que se ha descrito anteriormente para la creación de pD7-Sce1. En este caso, el brazo de homología izquierdo se reemplaza por un oligonucleótido que consiste en los cebadores 2808 y 2809 que codifican un sitio de reconocimiento TALEN invertido (mostrado en negrita) y una región espaciadora. El brazo de homología derecho se reemplaza con los cebadores 2723 y 2809 como se ha descrito anteriormente.

Una alternativa a las reacciones de ligadura simple entre las DSB en el cromosoma y el vector de direccionamiento, mediada por la unión de extremos no homólogos (NHEJ) es la unión de extremos mediada por microhomología (MMEJ). La MMEJ es un mecanismo de reparación de DSB que utiliza secuencias microhomólogas entre 5 y 25 pb para la unión de extremos propensa a errores (McVey y Lee, 2008) [154]. Se ha ideado una estrategia para la integración precisa de genes en la que la secuencia genómica y el vector de direccionamiento contienen la misma secuencia de reconocimiento del par de nucleasas TALE, pero una secuencia espaciadora de vector diferente en la que la mitad anterior y posterior están cambiadas. La secuencia genómica y el vector se pueden cortar con el mismo par TALEN y MMEJ tiene lugar a través de los extremos de ADN microhomólogo. El vector de direccionamiento integrado resultante ya no es una diana para la nucleasa TALE debido a que la región espaciadora acortada no es óptima para la escisión de nucleasa TALE (Nakade *et al.*, 2014) [155].

El vector de direccionamiento AAVS de MMEJ pD7-MMEJ se puede generar de la misma manera que se ha descrito anteriormente para la creación de pD7-Sce1. En este caso, el brazo de homología izquierdo se reemplaza por un oligonucleótido que consiste en los cebadores 2768 y 2769 que codifican el sitio de reconocimiento de TALEN (mostrado en negrita) y la región espaciadora conmutada (subrayada). El brazo de homología derecho se reemplaza con los cebadores 2723 y 2809 como se ha descrito anteriormente.

Ejemplo 22: Diseño de cebadores para la creación de un casete de promotor único (CMV) flanqueado por los brazos ROSA26 (pINT19-ROSA)

Este ejemplo pretende demostrar que los genes de anticuerpos o de moléculas de unión alternativas pueden integrarse en el genoma de las células de mamífero mediante métodos dirigidos por nucleasas, y los clones resultantes se pueden cribar para una función deseada ya sea mediante un cribado indicador o un cribado fenotípico. Un ejemplo de esto se demostró previamente cuando los genes de anticuerpos se integraron en el cromosoma de las células madre embrionarias de ratón (ES) y las colonias de ES individuales se cribaron para determinar su capacidad para mantener la pluripotencia cuando se sometieron a condiciones de diferenciación [105]. Se demostró que los genes de anticuerpos recuperados de colonias de ES que mantenían un fenotipo pluripotente bloqueaban la vía de señalización de FGFR1/FGF4. Un problema con este método descrito previamente es que la recombinación homóloga puede dar lugar a tamaños de biblioteca pequeños, lo que limita su capacidad para cribar directamente clones raros presentes en bibliotecas de moléculas de unión grandes. Los métodos de integración génica mediada por nucleasas para la integración de genes de anticuerpos y de moléculas de unión son más eficientes, dando como resultado una mayor generación de bibliotecas y, por lo tanto, es más probable que generen bibliotecas de células de mamífero capaces de identificar anticuerpos funcionales mediante el cribado fenotípico o de células indicadoras.

El vector de direccionamiento donante pINT19 está diseñado para integrar genes de anticuerpos en el locus ROSA26 de ratón mediante métodos dirigidos por nucleasas para el cribado funcional directo. pINT19 es un vector de promotor de CMV único para la expresión de fusión scFv-Fc. El casete de expresión está flanqueado por los brazos ROSA26 de ratón. Dado que el exón aguas arriba no está traducido, el gen de la puromicina está precedido por un aceptor de corte y empalme y más abajo tiene una secuencia KOZAK que conduce al gen de la puromicina.

El brazo de homología izquierdo AAVS y el gen de resistencia a puromicina de pINT18 se reemplazaron por un casete que codificaba la homología izquierda ROSA26, el aceptor de corte y empalme, la secuencia consenso kozak optimizada y el gen de resistencia a puromicina. El brazo de homología izquierdo ROSA26 se amplificó inicialmente a partir de pGATOR (Melidoni *et al.*, 2013 (105)) como dos fragmentos que desactivaron un sitio NotI interno. Los dos fragmentos, generados por los cebadores J60/2716 y 2715/2706 se combinaron en una PCR de ensamblaje con los cebadores J60 y 2706 y se digirieron con AsiI y NsiI. El aceptor de corte y empalme se amplificó a partir de pGATOR utilizando los cebadores 2709 y 2710 y el casete de resistencia a la puromicina se amplificó con los cebadores 2745 (que incluían una región homóloga al aceptor de corte y empalme y la secuencia consenso Kozak optimizada) y J59. La región del aceptor de corte y empalme y el casete de resistencia a puromicina se combinaron en la PCR de ensamblaje utilizando los cebadores 2709 y J59 y se digirieron con NsiI y BglII. La homología de

brazo izquierdo ROSA26 y los casetes de aceptor de corte y empalme-puromicina se ligaron con el vector pINT18 (AsiS1/Bgl2).

5 Para completar el vector de direccionamiento ROSA, se introdujo el brazo de homología ROSA26 derecho, aguas abajo del casete CMV-scFv-Fc, para reemplazar el brazo de homología derecho pINT18 AAVS. Esto se realizó mediante PCR del brazo de homología derecho ROSA, presente en pGATOR (Melidoni *et al.*, 2013) utilizando los cebadores J61 y J62 para amplificar un fragmento con BstZ171 en un extremo y Sbf1 en el otro. El cebador 61 se colocó para excluir un sitio Sbf1 endógeno de 65 pb por encima del sitio de escisión ROSA ZFN. La Figura 31 muestra las secuencias de los brazos de homología izquierdo y derecho ROSA26.

10

Tabla 11. Cebadores para el direccionamiento dirigido por nucleasas del locus ROSA 26 de ratón

J60 (SEQ ID NO: 129)	ACACACGGTACCGCGATCGCGCTGATTGG CTTCTTTTCCTC	AsiSI-rosa26-L-F
2706 (SEQ ID NO: 130)	TTTTTTATGCATTCTAGAAAGACTGGAGTT GCAGA	Nsil-rosa26-L-R
2715 (SEQ ID NO: 131)	GAGCGTCCGCCACCCCTC	ROSA-Left-NotI_knockout_F
2716 (SEQ ID NO: 132)	GAGGGTGGGCGGACGCTC	ROSA-Left-NotI_knockout_R
2709 (SEQ ID NO: 133)	TTTTTTATGCATTAAGGGATCTGTAGGGCG CAG	Splice_acceptor-F-Nsil
2710 (SEQ ID NO: 134)	GTGAATTCCTAGAGCGGCCTC	Splice_acceptor-R
2745 (SEQ ID NO: 135)	GAGGCCGCTCTAGGAATTCACGCCGCCAC CATGACCGAGTACAAGCCCAC	Overlap-Puro-F+kozak
J59 (SEQ ID NO: 136)	AAAAAAGATCTGTGTGTTTCGAATCAGGC ACCGGGCTTGCGGGTCAT	Bgl2-Puro-R
J61 (SEQ ID NO: 137)	tttttGTATACGGGAATTGAACAGGTGTA TTG	ROSA-Right_F-BstZ171
J62 (SEQ ID NO: 138)	TTTTTTCCTGCAGGAGGTTGGATTCTCAAT ACATCTATTGTTG	ROSA-Right_R-Sbf1
2701 (SEQ ID NO: 139)	GCCGACGTCTCGTCTGATGTTTT	
2702 (SEQ ID NO: 140)	ATCAGCGACGAGACGTCGCGCGGTG	
2703 (SEQ ID NO: 141)	CGCCCA T CTT CTAGAAAGACGTTTT	
2704 (SEQ ID NO: 142)	GTCTTTCTAGAAGATGGGCGCGGTG	

15 La integración de pINT19, anticuerpo codificante o moléculas de unión alternativas, en el locus ROSA26 de ratón se podría lograr mediante la introducción dirigida por nucleasas de bibliotecas de anticuerpos usando CRISPR/Cas9 como se describe en el Ejemplo 20. Aquí, la integración dirigida por nucleasas a través de CRISPR/Cas9 podría demostrarse utilizando el "kit de vector de nucleasa CRISPR Geneart" (Lifetech A21175). En este sistema, un promotor de ARN polimerasa III U6 impulsa la expresión de un ARN CRISPR complementario (ARNrc) complementario que está vinculado a un ARNrc trans-activador (ARNtracr). El ARNcr y el ARNtracr juntos forman un ARN guía que dirige la especificidad de escisión de una proteína Cas9 codificada en el mismo "vector de nucleasa CRISPR GeneArt" (véanse las instrucciones del fabricante). El vector se proporciona como un plásmido linealizado en el que se clona un oligonucleótido bicatenario corto con salientes 3' apropiados. La especificidad de escisión se determina entonces por la secuencia del segmento clonado. Se diseñaron 2 secuencias de direccionamiento diferentes para dirigir la escisión al ROSA26 de ratón codificado por los cebadores 2701/2702 y 2703/2704 (véase la

20

Tabla 10).

Como alternativa, se han descrito nucleasas con dedos de cinc que se escinden dentro del locus ROSA26 [34]. Podrían construirse en un vector de expresión apropiado como se describe para la meganucleasa Sce-I (Figura 16, ejemplo 11).

La integración mediada por nucleasas del plásmido donante pINT19 daría lugar a clones que expresan un anticuerpo secretado que podría unirse al receptor endógeno o al ligando dando como resultado un antagonismo [105,107] o un agonismo [47,106,108] de las vías de señalización del receptor. Para permitir el enlace entre el fenotipo celular y la actividad funcional del anticuerpo secretado, las células pueden colocarse en placas a una densidad baja en medios semisólidos, de modo que las colonias individuales se propagan y la expresión del anticuerpo se puede iniciar a través de un promotor inducible [105]. Como alternativa, podría emplearse un promotor constitutivo para la expresión génica del anticuerpo. Los medios semisólidos mantendrían una concentración local elevada del anticuerpo expresado de manera endógena, de modo que cualquier cambio fenotípico específico de una colonia que surja de una célula individual será causado por el anticuerpo único expresado a partir de ese clon particular. Si se empleara un indicador de respuesta rápida, tal como el promotor Rex o Nanog fusionado con un gen indicador, sería posible colocar en placas las células a una baja densidad en medios semisólidos, recolectarlas y luego cribarlas por citometría de flujo. Como alternativa, el codón de terminación aguas abajo del gen del anticuerpo en pINT19 podría reemplazarse por un dominio transmembrana para permitir la unión del anticuerpo a la superficie celular. El codón de terminación aguas abajo del gen del anticuerpo en pINT19 también podría reemplazarse por una secuencia señal de retención del retículo endoplásmico (ER) para permitir la retención de anticuerpos en el ER y la posible regulación negativa de un receptor diana expresado de manera endógena o cualquier proteína o péptido secretado. pINT19 está diseñado específicamente para dirigirse al locus ROSA26 de ratón y puede emplearse para el cribado fenotípico de anticuerpos o moléculas de unión alternativas en células de ES de ratón. Sin embargo, los métodos de integración de anticuerpos o genes de moléculas de aglutinantes dirigidos por nucleasas también podrían aplicarse a otros cribados funcionales, tal como los descritos utilizando el enfoque lentiviral [47,106,107,108].

Bibliografía

- 1 Russell, S. J., Hawkins, R. E., & Winter, G. (1993). Retroviral vectors displaying functional antibody fragments. *Nucleic Acids Res*, 21(5), 1081-1085.
- 2 Boublik, Y., Di Bonito, P., & Jones, I. M. (1995). Eukaryotic virus display: Engineering the major surface glycoprotein of the *Autographa californica* nuclear polyhedrosis virus (AcNPV) for the presentation of foreign proteins on the virus surface. *Nature Biotechnology*, 13(10), 1079-1084.
- 3 Mottershead, D. G., Alfthan, K., Ojala, K., Takkinen, K., & Oker-Blom, C. (2000). Baculoviral display of functional scFv and synthetic IgG-binding domains. *Biochemical and Biophysical Research Communications*, 275(1), 84-90.
- 4 Oker-Blom, C., Airene, K. J., & Grabherr, R. (2003). Baculovirus display strategies: Emerging tools for eukaryotic libraries and gene delivery. *Briefings in Functional Genomics and Proteomics*, 2(3), 244-253.
- 5 Edwards, B. M., Barash, S. C., Main, S. H., Choi, G. H., Minter, R., Ullrich, S., et al. (2003). The Remarkable Flexibility of the Human Antibody Repertoire; Isolation of Over One Thousand Different Antibodies to a Single Protein, BLYS. *Journal of Molecular Biology*, 334(1), 103-118. doi:10.1016/j.jmb-2003.09.054
- 6 Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., et al. (2010). Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display. *Protein Engineering, Design and Selection*, 23(4), 279-288. doi:10.1093/protein/gzq003
- 7 Schofield, D. J., Pope, A. R., Clementel, V., Buckell, J., Chapple, S., Clarke, K. F., et al. (2007). Application of phage display to high throughput antibody generation and characterization. *Genome Biol*, 8(11), R254.
- 8 Salema, V., Marin, E., Martfnez-Arteaga, R., Ruano-Gallego, D., Fraile, S., Margolles, Y., et al. (2013). Selection of Single Domain Antibodies from Immune Libraries Displayed on the Surface of *E. coli* Cells with Two β -Domains of Opposite Topologies. *PLoS ONE*, 8(9), e75126. doi:10.1371/journal.pone.0075126.s007
- 9 Chao, G., Lau, W. L., Hackel, B. J., Sazinsky, S. L., Lippow, S. M., & Wittrup, K. D. (2006). Isolating and engineering human antibodies using yeast surface display. *Nat Protoc*, 1(2), 755-768.
- 10 Higuchi, K., Araki, T., Matsuzaki, O., Sato, A., Kanno, K., Kitaguchi, N., & Ito, H. (1997). Cell display library for gene cloning of variable regions of human antibodies to hepatitis B surface antigen. *J Immunol Methods*, 202(2), 193-204.
- 11 Ho, M., Nagata, S., & Pastan, I. (2006). Isolation of anti-CD22 Fv with high affinity by Fv display on human cells. *Proc Natl Acad Sci U S A*, 103(25), 9637-9642. doi:10.1073/pnas.0603653103
- 12 Ho, M., & Pastan, I. (2009). Display and selection of scFv antibodies on HEK-293T cells. *Methods Mol Biol*, 562, 99-113. doi:10.1007/978-1-60327-302-2_8
- 13 Akamatsu, Y., Pakabunto, K., Xu, Z., Zhang, Y., & Tsurushita, N. (2007). Whole IgG surface display on mammalian cells: Application to isolation of neutralizing chicken monoclonal anti-IL-12 antibodies. *Journal of Immunological Methods*, 327(1-2), 40-52. doi:10.1016/j.jim.2007.07.007
- 14 Beerli, R. R., Bauer, M., Buser, R. B., Gwerder, M., Muntwiler, S., Maurer, P., et al. Isolation of human monoclonal antibodies by mammalian cell display. *Proc Natl Acad Sci U S A*. 2008 Sep 23;105(38):14336-41. doi: 10.1073/pnas.0805942105

- 15 Breous-Nystrom, E., Schultze, K., Meier, M., Flueck, L., Holzer, C., Boll, M., et al. (2013). Retrocyte Display[®] technology: Generation and screening of a high diversity cellular antibody library. *Methods*, 1-11. doi:10.1016/j.ymeth.2013.09.003
- 5 16 Grindley, Whiteson & Rice. Mechanisms of Site-Specific Recombination. *Annu Rev Biochem* 2006 75:567-605
- 17 Zhou, C., Jacobsen, F. W., Cai, L., Chen, Q., & Shen, W. D. (2010). Development of a novel mammalian cell surface antibody display platform. *mAbs*, 2(5), 508-518
- 18 Li, C. Z., Liang, Z. K., Chen, Z. R., Lou, H. B., Zhou, Y., Zhang, Z. H., et al. (2012). Identification of HBsAg-specific antibodies from a mammalian cell displayed full-length human antibody library of healthy immunized donor. *Cellular and Molecular Immunology*, 9(2), 184-190.
- 10 19 Buchholz, F., Ringrose, L., Angrand, P. O., Rossi, F., & Stewart, A. F. (1996). Different thermostabilities of FLP and Cre recombinases: Implications for applied site-specific recombination. *Nucleic Acids Res*, 24(21), 4256-4262.
- 20 Schaft, J., Ashery-Padan, R., Van Hoeven, F. D., Gruss, P., & Francis Stewart, A. (2001). Efficient FLP recombination in mouse ES cells and oocytes. *Genesis*, 31(1), 6-10.
- 15 21 Moehle, E. A., Moehle, E. A., Rock, J. M., Rock, J. M., Lee, Y.-L., Lee, Y. L., et al. (2007). Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A*, 104(9), 3055-3060. doi:10.1073/pnas.0611478104
- 22 Cristea, S., Freyvert, Y., Santiago, Y., Holmes, M. C., Urnov, F. D., Gregory, P. D., & Cost, G. J. (2013). In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnology and Bioengineering*, 110(3), 871-880. doi:10.1002/bit.24733
- 20 23 Letourneur, F., & Malissen, B. (1989). Derivation of a T cell hybridoma variant deprived of functional T cell receptor alpha and beta chain transcripts reveals a nonfunctional alpha-mRNA of BW5147 origin. *European Journal of Immunology*, 19(12), 2269-2274. doi:10.1002/eji.1830191214
- 24 Kanayama, N., Todo, K., Reth, M., & Ohmori, H. (2005). Reversible switching of immunoglobulin hypermutation machinery in a chicken B cell line. *Biochem Biophys Res Commun*, 327(1), 70-75. doi:10.1016/j.bbrc.2004.11.143
- 25 25 Lin, W., Kurosawa, K., Murayama, A., Kagaya, E., & Ohta, K. (2011). B-cell display-based one-step method to generate chimeric human IgG monoclonal antibodies. *Nucleic Acids Research*, 39(3), e14-e14. doi:10.1093/nar/gkq1122
- 30 26 Adachi, N., So, S., Iizumi, S., Nomura, Y., Murai, K., Yamakawa, C., et al. (2006). The human pre-B cell line Nalm-6 is highly proficient in gene targeting by homologous recombination. *DNA and Cell Biology*, 25(1), 19-24. doi:10.1089/dna.2006.25.19
- 27 Palacios, R., & Steinmetz, M. (1985). IL3-Dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo. *Cell*, 41(3), 727-734
- 35 28 Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, 483(7391), 603-607. doi:10.1038/nature11003
- 29 Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Research*, 39(Database issue), D945-50. doi:10.1093/nar/gkq929
- 40 30 Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., & Pâques, F. (2011). Meganucleases and other tools for targeted genome engineering: Perspectives and challenges for gene therapy. *Current Gene Therapy*, 11(1), 11-27
- 45 31 Epinat, J. C., Silva, G. H., Pâques, F., Smith, J., & Duchateau, P. (2013). Engineered meganucleases for genome engineering purposes. *Topics in Current Genetics* (Vol. 23, pp. 147-185).
- 32 Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D. J., & Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nature Biotechnology*, 25(7), 786-793. doi:10.1038/nbt1317
- 50 33 Doyon, Y., Vo, T. D., Mendel, M. C., Greenberg, S. G., Wang, J., Xia, D. F., et al. (2011). Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods*, 8(1), 74-79. doi:10.1038/nmeth.1539
- 34 Perez-Pinera, P., Ousterout, D. G., Brown, M. T., & Gersbach, C. A. (2012). Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. *Nucleic Acids Research*, 40(8), 3741-3752. doi:10.1093/nar/gkr1214
- 55 35 Urnov, F. D., Miller, J. C., Lee, Y.-L., Beausejour, C. M., Rock, J. M., Augustus, S., et al. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 435(7042), 646-651. doi:10.1038/nature03556
- 36 Maresca, M., Lin, V. G., Guo, N., & Yang, Y. (2013). Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Research*, 23(3), 539-546. doi:10.1101/gr.145441.112
- 60 37 Bogdanove, A. J., & Voytas, D. F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science*, 333(6051), 1843-1846. doi:10.1126/science.1204094
- 38 Reyon, D., Tsai, S. Q., Khgayer, C., Foden, J. A., Sander, J. D., & Joung, J. K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30(5), 460-465. doi:10.1038/nbt.2170
- 65 39 Boissel, S., Jarjour, J., Astrakhan, A., Adey, A., Gouble, A., Duchateau, P., et al. (2013). megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Research*.

doi:10.1093/nar/gkt1224

- 40 Beurdeley, M., Bietz, F., Li, J., Thomas, S., Stoddard, T., Juillerat, A., et al. (2013). Compact designer TALENs for efficient genome engineering. *Nature Communications*, 4, 1762. doi:10.1038/ncomms2782
- 41 Sampson, T. R., & Weiss, D. S. (2014). Exploiting CRISPR/Cas systems for biotechnology. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, 36(1), 34-38. doi:10.1002/bies.201300135
- 5 42 Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, 343(6166), 84-87. doi:10.1126/science.1247005
- 43 Wang, T., Wei, J. J., Sabatini, D. M., & Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, 343(6166), 80-84. doi:10.1126/science.1246981
- 10 44 Beard, C., Hochedlinger, K., Plath, K., Wutz, A., & Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis*, 44(1), 23-28.
- 45 Orlando, S. J., Santiago, Y., DeKolver, R. C., Freyvert, Y., Boydston, E. A., Moehle, E. A., et al. (2010). Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Research*, 38(15), e152. doi:10.1093/nar/gkq512
- 15 46 Cadinanos, J; Bradley, A. (2007) *Nucleic Acids Res.* 35(12):e87
- 47 Zhang, H., Yea, K., Xie, J., Ruiz, D., Wilson, I. A., & Lerner, R. A. (2013). Selecting agonists from single cells infected with combinatorial antibody libraries. *Chemistry & Biology*, 20(5), 734-741. doi:10.1016/j.chembiol.2013.04.012
- 20 48 Porteus, M. H., & Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. *Science*, 300(5620), 763. doi:10.1126/science.1078395
- 49 Rouet, P., Smih, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, 14(12), 8096-8106.
- 50 Jasin, M. (1996). Genetic manipulation of genomes with rare-cutting endonucleases. *Trends in Genetics*, 12(6), 224-228
- 25 51 Davis, L., & Maizels, N. (2011). DNA nicks promote efficient and safe targeted gene correction. *PLoS ONE*, 6(9), e23981. doi:10.1371/journal.pone.0023981
- 52 Fujioka, K., Aratani, Y., Kusano, K., & Koyama, H. (1993). Targeted recombination with single-stranded DNA vectors in mammalian cells. *Nucleic Acids Res*, 21(3), 407-412
- 30 53 Khan, I. F., Hirata, R. K., & Russell, D. W. (2011). AAV-mediated gene targeting methods for human cells. *Nat Protoc*, 6(4), 482-501. doi:10.1038/nprot.2011.301
- 54 Deyle, D. R., & Russell, D. W. (2009). Adeno-associated virus vector integration. *Current Opinion in Molecular Therapeutics*, 11(4), 442-447
- 55 Benatui, L., Perez, J. M., Belk, J., & Hsieh, C. M. (2010). An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Engineering, Design and Selection*, 23(4), 155-159.
- 35 56 Feldhaus, M. J., Siegel, R. W., Opresko, L. K., Coleman, J. R., Feldhaus, J. M., Yeung, Y. A., et al. (2003). Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat Biotechnol*.
- 57 Zhao, A., Nunez-Cruz, S., Li, C., Coukos, G., Siegel, D. L., & Scholler, N. (2011). Rapid isolation of high-affinity human antibodies against the tumor vascular marker Endosialin/TEM1, using a paired yeast-display/secretory scFv library platform. *Journal of Immunological Methods*, 363(2), 221-232. doi:10.1016/j.jim.2010.09.001
- 40 58 Skerra, A. (2007). Alternative non-antibody scaffolds for molecular recognition. *Curr Opin Biotechnol*, 18(4), 295-304. doi:10.1016/j.copbio.2007.04.010
- 59 Gebauer, M., & Skerra, A. (2009). Engineered protein scaffolds as next-generation antibody therapeutics. *Current Opinion in Chemical Biology*, 13(3), 245-255. doi:10.1016/j.cbpa.2009.04.627
- 45 60 Haan & Maggos (2004) *BioCentury*, 12(5): A1-A6
- 61 Koide et al. (1998) *Journal of Molecular Biology*, 284: 1141-1151.
- 62 Nygren et al. (1997) *Current Opinion in Structural Biology*, 7: 463-469
- 63 Wess, L. In: *BioCentury, The Bernstein Report on BioBusiness*, 12(42), A1-A7, 2004
- 50 64 Chang, H.-J., Hsu, H.-J., Chang, C.-F., Peng, H.-P., Sun, Y.-K., Yu, H.-M., et al. (2009). Molecular Evolution of Cystine-Stabilized Miniproteins as Stable Proteinaceous Binders. *Structure*, 17(4), 620-631. doi:10.1016/j.str.2009.01.011
- 65 65 Ward, E.S. et al., *Nature* 341, 544-546 (1989)
- 66 McCafferty et al *Nature*, 348, 552-554 (1990)
- 55 67 Holt et al *Trends in Biotechnology* 21, 484-490 (2003)
- 68 Bird et al, *Science*, 242, 423-426, (1988)
- 69 Huston et al, *PNAS USA*, 85, 5879-5883, (1988)
- 70 Holliger, P. et al, *PNAS USA* 90 6444-6448, (1993)
- 71 Reiter, Y. et al, *Nature Biotech*, 14, 1239-1245, (1996)
- 60 72 Holliger & Hudson, *Nature Biotechnology* 23(9): 1126-1136 (2005)
- 73 Knappik et al. *J. Mol. Biol.* 296, 57-86 (2000)
- 74 Krebs et al. *Journal of Immunological Methods* 254, 67-84 (2001)
- 75 Holliger and Bohlen *Cancer and metastasis rev.* 18: 411-419 (1999)
- 76 Holliger, P. and Winter G. *Current Opinion Biotechnol* 4, 446-449 (1993)
- 65 77 Glennie M J et al., *J. Immunol.* 139, 2367-2375 (1987)
- 78 Repp R. et al., *J. Hemat.* 377-382 (1995)

- 79 Staerz U. D. and Bevan M. J. PNAS 83 (1986)
- 80 Suresh M. R. et al., Method Enzymol. 121: 210-228 (1986)
- 81 Merchand et al., Nature Biotech. 16:677-681 (1998)
- 82 Ridgeway, J. B. B. et al, Protein Eng., 9, 616-621, (1996)
- 5 83 Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D., & Winter, G. (1991). Bypassing immunization. Human antibodies from V-gene libraries displayed on phage. J Mol Biol, 222(3), 581-597.
- 84 Gronwald, R. G. K., Grant, F. J., Haldeman, B. A., Hart, C. E., O'Hara, P. J., Hagen, F. S., et al. (1988). Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor: Evidence for more than one receptor class (Vol. 85, pp. 3435-3439). Presented at the Proceedings of the National Academy of Sciences of the United States of America.
- 10 85 Kumar, N., & Borth, N. (2012). Flow-cytometry and cell sorting: an efficient approach to investigate productivity and cell physiology in mammalian cell factories. Methods, 56(3), 366-374. doi:10.1016/j.ymeth.2012.03.004
- 86 Brezinsky, S. C. G., Chiang, G. G., Szilvasi, A., Mohan, S., Shapiro, R. I., MacLean, A., et al. (2003). A simple method for enriching populations of transfected CHO cells for cells of higher specific productivity. J Immunol Methods, 277(1-2), 141-155.
- 15 87 Pichler, J., Hesse, F., Wieser, M., Kunert, R., Galosy, S. S., Mott, J. E., & Borth, N. (2009). A study on the temperature dependency and time course of the cold capture antibody secretion assay. Journal of Biotechnology, 141(1-2), 80-83. doi:10.1016/j.jbiotec.2009.03.001
- 88 Anastassiadis, K., Fu, J., Patsch, C., Hu, S., Weidlich, S., Duerschke, K., et al. (2009). Dre recombinase, like Cre, is a highly efficient site-specific recombinase in E. coli, mammalian cells and mice. Disease Models & Mechanisms, 2(9-10), 508-515. doi:10.1242/dmm.003087
- 20 89 Horlick, R. A., Macomber, J. L., Bowers, P. M., Neben, T. Y., Tomlinson, G. L., Krapf, I. P., et al. (2013). Simultaneous surface display and secretion of proteins from mammalian cells facilitate efficient in vitro selection and maturation of antibodies. J Biol Chem, 288(27), 19861-19869. doi:10.1074/jbc.M113.452482
- 25 90 Biffi, G., Tannahill, D., McCafferty, J., & Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. Nature Chemistry, 5(3), 182-186. doi:10.1038/nchem.1548
- 91 Gao, J., Sidhu, S. S., & Wells, J. A. (2009). Two-state selection of conformation-specific antibodies. Proc Natl Acad Sci U S A, 106(9), 3071-3076. doi:10.1073/pnas.0812952106
- 30 92 Gu, G. J., Friedman, M., Jost, C., Johnsson, K., Kamali-Moghaddam, M., Plückthun, A., et al. (2013). Protein tag-mediated conjugation of oligonucleotides to recombinant affinity binders for proximity ligation. N Biotechnol, 30(2), 144-152. doi:10.1016/j.nbt.2012.05.005
- 93 Cho, Y. K., & Shusta, E. V. (2010). Antibody library screens using detergent-solubilized mammalian cell lysates as antigen sources. Protein Eng Des Sel, 23(7), 567-577. doi:10.1093/protein/gzq029
- 35 94 Tillotson, B. J., Cho, Y. K., & Shusta, E. V. (2013). Cells and cell lysates: a direct approach for engineering antibodies against membrane proteins using yeast surface display. Methods, 60(1), 27-37. doi:10.1016/j.ymeth.2012.03.010
- 95 Kunert, A., Straetemans, T., Govers, C., Lamers, C., Mathijssen, R., Sleijfer, S., & Debets, R. (2013). TCR-Engineered T Cells Meet New Challenges to Treat Solid Tumors: Choice of Antigen, T Cell Fitness, and Sensitization of Tumor Milieu. Frontiers in Immunology, 4, 363. doi:10.3389/fimmu.2013.00363
- 40 96 Liddy, N., Bossi, G., Adams, K. J., Lissina, A., Mahon, T. M., Hassan, N. J., et al. (2012). Monoclonal TCR-redirectioned tumor cell killing. Nature Medicine, 18(6), 980-987. doi:10.1038/nm.2764
- 97 Holler, P. D., Holman, P. O., Shusta, E. V., O'Herrin, S., Wittrup, K. D., & Kranz, D. M. (2000). In vitro evolution of a T cell receptor with high affinity for peptide/MHC. Proc Natl Acad Sci U S A, 97(10), 5387-5392. doi:10.1073/pnas.080078297
- 45 98 Weber, K. S., Donermeyer, D. L., Allen, P. M., & Kranz, D. M. (2005). Class II-restricted T cell receptor engineered in vitro for higher affinity retains peptide specificity and function. Proc Natl Acad Sci U S A, 102(52), 19033-19038. doi:10.1073/pnas.0507554102
- 99 Kessels, H. W., van Den Boom, M. D., Spits, H., Hooijberg, E., & Schumacher, T. N. (2000). Changing T cell specificity by retroviral T cell receptor display. Proc Natl Acad Sci U S A, 97(26), 14578-14583. doi:10.1073/pnas.97.26.14578
- 50 100 Chervin, A. S., Aggen, D. H., Raseman, J. M., & Kranz, D. M. (2008). Engineering higher affinity T cell receptors using a T cell display system. Journal of Immunological Methods, 339(2), 175-184.
- 101 Crawford, F., Jordan, K. R., Stadinski, B., Wang, Y., Huseby, E., Marrack, P., et al. (2006). Use of baculovirus MHC/peptide display libraries to characterize T-cell receptor ligands. Immunological Reviews, 210, 156-170. doi:10.1111/j.0105-2896.2006.00365.x
- 55 102 Hinrichs, C. S., & Restifo, N. P. (2013). Reassessing target antigens for adoptive T-cell therapy. Nat Biotechnol, 31(11), 999-1008. doi:10.1038/nbt.2725
- 103 Sadelain, M., Brentjens, R., & Rivière, I. (2013). The basic principles of chimeric antigen receptor design. Cancer Discovery, 3(4), 388-398. doi:10.1158/2159-8290.CD-12-0548
- 60 104 Alonso-Camino, V., Sánchez-Martín, D., Compte, M., Sanz, L., & Alvarez-Vallina, L. (2009). Lymphocyte display: a novel antibody selection platform based on T cell activation. PLoS ONE, 4(9), e7174. doi:10.1371/journal.pone.0007174
- 105 Melidoni, A. N., Dyson, M. R., Wormald, S., & McCafferty, J. (2013). Selecting antagonistic antibodies that control differentiation through inducible expression in embryonic stem cells. Proceedings of the National Academy of Sciences, 110(44), 17802-17807. doi:10.1073/pnas.1312062110
- 65 106 Zhang, H., Wilson, I. A., & Lerner, R. A. (2012). Selection of antibodies that regulate phenotype from

- intracellular combinatorial antibody libraries. *Proceedings of the National Academy of Sciences*, 109(39), 15728-15733. doi:10.1073/pnas.1214275109
- 107 Xie, J., Yea, K., Zhang, H., Moldt, B., He, L., Zhu, J., & Lerner, R. A. (2014). Prevention of cell death by antibodies selected from intracellular combinatorial libraries. *Chemistry & Biology*, 21(2), 274-283.
- 5 108 Yea, K., Zhang, H., Xie, J., Jones, T. M., Yang, G., Song, B. D., & Lerner, R. A. (2013). Converting stem cells to dendritic cells by agonist antibodies from unbiased morphogenic selections (Vol. 110, pp. 14966-14971). Presented at the Proceedings of the National Academy of Sciences of the United States of America. doi:10.1073/pnas.1313671110
- 109 Kawahara, M., Kimura, H., Ueda, H., & Nagamune, T. (2004). Selection of genetically modified cell population using hapten-specific antibody/receptor chimera. *Biochem Biophys Res Commun*, 315(1), 132-138. doi:10.1016/j.bbrc.2004.01.030
- 110 Kawahara, M., Shimo, Y., Sogo, T., Hitomi, A., Ueda, H., & Nagamune, T. (2008). Antigen-mediated migration of murine pro-B Ba/F3 cells via an antibody/receptor chimera. *Journal of Biotechnology*, 133(1), 154-161. doi:10.1016/j.jbiotec.2007.09.009
- 15 111 Sogo, T., Kawahara, M., Ueda, H., Otsu, M., Onodera, M., Nakauchi, H., & Nagamune, T. (2009). T cell growth control using hapten-specific antibody/interleukin-2 receptor chimera. *Cytokine*, 46(1), 127-136. doi:10.1016/j.cyto.2008.12.020
- 112 Kawahara, M., Chen, J., Sogo, T., Teng, J., Otsu, M., Onodera, M., et al. (2011). Growth promotion of genetically modified hematopoietic progenitors using an antibody/c-Mpl chimera. *Cytokine*, 55(3), 402-408. doi:10.1016/j.cyto.2011.05.024
- 20 113 Ueda, H., Kawahara, M., Aburatani, T., Tsumoto, K., Todokoro, K., Suzuki, E., et al. (2000). Cell-growth control by monomeric antigen: the cell surface expression of lysozyme-specific Ig V-domains fused to truncated Epo receptor. *J Immunol Methods*, 241(1-2), 159-170.
- 114 Kerppola, T. K. (2009). Visualization of molecular interactions using bimolecular fluorescence complementation analysis: Characteristics of protein fragment complementation. *Chemical Society Reviews*, 38(10), 2876-2886.
- 25 115 Michnick, S. W., Ear, P. H., Manderson, E. N., Remy, I., & Stefan, E. (2007). Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nature Reviews Drug Discovery*, 6(7), 569-582. doi:10.1038/nrd231
- 30 116 Petschnigg, J., Groisman, B., Kotlyar, M., Taipale, M., Zheng, Y., Kurat, C. F., et al. (2014). The mammalian-membrane two-hybrid assay (MaMTH) for probing membrane-protein interactions in human cells. *Nat Methods*. doi:10.1038/nmeth.2895
- 117 Renaut, L., Monnet, C., Dubreuil, O., Zaki, O., Crozet, F., Bouayadi, K., et al. (2012). Affinity maturation of antibodies: Optimized methods to generate high-quality scfv libraries and isolate igg candidates by high-throughput screening. *Methods in Molecular Biology* (Vol. 907, pp. 451-461)
- 35 118 Dyson, M. R., Zheng, Y., Zhang, C., Colwill, K., Pershad, K., Kay, B. K., et al. (2011). Mapping protein interactions by combining antibody affinity maturation and mass spectrometry. *Anal Biochem*, 417(1), 25-35. doi:10.1016/j.ab.2011.05.005
- 119 de Felipe P (2002) Polycistronic viral vectors. *Curr Gene Ther* 2: 355-378. doi: 10.2174/1566523023347742.
- 40 120 Foote, J., & Winter, G. (1992). Antibody framework residues affecting the conformation of the hypervariable loops. *J Mol Biol*, 224(2), 487-499.
- 121 Massie, B., Dionne, J., Lamarche, N., Fleurent, J., & Langelier, Y. (1995). Improved adenovirus vector provides herpes simplex virus ribonucleotide reductase R1 and R2 subunits very efficiently. *Nature Biotechnology*, 13(6), 602-608.
- 45 122 Kim, D. W., Uetsuki, T., Kaziro, Y., Yamaguchi, N., & Sugano, S. (1990). Use of the human elongation factor 1 alpha promoter as a versatile and efficient expression system. *Gene*, 91(2), 217-223.
- 123 Holden, P., Keene, D. R., Lunstrum, G. P., Bächinger, H. P., & Horton, W. A. (2005). Secretion of cartilage oligomeric matrix protein is affected by the signal peptide. *J Biol Chem*, 280(17), 17172-17179.
- 50 124 Sadelain, M., Papapetrou, E. P., & Bushman, F. D. (2011). Safe harbours for the integration of new DNA in the human genome. *Nature Reviews Cancer*, 12(1), 51-58.
- 125 Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G., & Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. *Nat Protoc*, 7(1), 171-192. doi:10.1038/nprot.2011.431
- 126 Falk, R., Falk, A., Dyson, M. R., Melidoni, A. N., Parthiban, K., Young, J. L., et al. (2012). Generation of anti-Notch antibodies and their application in blocking Notch signalling in neural stem cells. *Methods*, 58(1), 69-78. doi:10.1016/j.ymeth.2012.07.008
- 55 127 Martin, C. D., Rojas, G., Mitchell, J. N., Vincent, K. J., Wu, J., McCafferty, J., & Schofield, D. J. (2006). A simple vector system to improve performance and utilisation of recombinant antibodies. *BMC Biotechnology*, 6, 46.
- 128 Reyon, D., Tsai, S. Q., Khgayer, C., Foden, J. A., Sander, J. D., & Joung, J. K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30(5), 460-465. doi:10.1038/nbt.2170
- 60 129 Van Der Weyden, L., Adams, D. J., Harris, L. W., Tannahill, D., Arends, M. J., & Bradley, A. (2005). Null and conditional Semaphorin 3B alleles using a flexible puroΔtk LoxP/FRT vector. *Genesis*, 41(4), 171-178
- 130 de Felipe, P., & Ryan, M. D. (2004). Targeting of proteins derived from self-processing polyproteins containing multiple signal sequences. *Traffic*, 5(8), 616-626.
- 65 131 Raymond, C. S., & Soriano, P. (2007). High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS ONE*, 2(1), e162. doi:10.1371/journal.pone.0000162

- 132 Kranz, A., Fu, J., Duerschke, K., Weidlich, S., Naumann, R., Stewart, A. F., & Anastassiadis, K. (2010). An improved Flp deleter mouse in C57Bl/6 based on Flpo recombinase. *Genesis*, 48(8), 512-520. doi:10.1002/dvg.20641
- 5 133 Szymczak AL, Vignali DA (2005) Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther* 5: 627-638. doi: 10.1517/14712598.5.5.627
- 134 Chapple, S.D., Crofts, A.M., Shadbolt, S.P., McCafferty, J., and Dyson, M.R. (2006). Multiplexed expression and screening for recombinant protein production in mammalian cells. *BMC biotechnology* 6, 49.
135. Zhao, Y. et al. Multiple injections of electroporated autologous T cells expressing a chimeric antigen receptor mediate regression of human disseminated tumor. *Cancer Research* 70, 9053-9061 (2010).
- 10 136 Szymczak, A. L. et al. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nature biotechnology* 22, 589-594 (2004).
137. Li, Y. et al. Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat. Biotechnol.* 23, 349-354 (2005).
138. Zhao, Y. et al. High-affinity TCRs generated by phage display provide CD4+ T cells with the ability to recognize and kill tumor cell lines. *J. Immunol.* 179, 5845-5854 (2007).
- 15 139. Madura, F. et al. T-cell receptor specificity maintained by altered thermodynamics. *The Journal of biological chemistry* 288, 18766-18775 (2013).
140. Pierce, B. G. et al. Computational design of the affinity and specificity of a therapeutic T cell receptor. *PLoS Comput Biol* 10, e1003478 (2014).
- 20 141. Sebestyén, Z. et al. Human TCR that incorporate CD3zeta induce highly preferred pairing between TCRalpha and beta chains following gene transfer. *J. Immunol.* 180, 7736-7746 (2008).
142. Roszik, J. et al. T-cell synapse formation depends on antigen recognition but not CD3 interaction: studies with TCRalpha, a candidate transgene for TCR gene therapy. *Eur. J. Immunol.* 41, 1288-1297 (2011).
- 25 143. Cohen, C. J., Zhao, Y., Zheng, Z., Rosenberg, S. A. & Morgan, R. A. Enhanced antitumor activity of murine-human hybrid T-cell receptor (TCR) in human lymphocytes is associated with improved pairing and TCR/CD3 stability. *Cancer Research* 66, 8878-8886 (2006).
144. Huovinen, T. et al. Primer extension mutagenesis powered by selective rolling circle amplification. *PLoS ONE* 7, e31817 (2012).
- 30 145. Cribbs, A. P., Kennedy, A., Gregory, B. & Brennan, F. M. Simplified production and concentration of lentiviral vectors to achieve high transduction in primary human T cells. *BMC biotechnology* 13, 98 (2013).
146. Oelke, M. et al. Ex vivo induction and expansion of antigen-specific cytotoxic T cells by HLA-Ig-coated artificial antigen-presenting cells. *Nat Med* 9, 619-624 (2003).
147. Wölfel, M. & Greenberg, P. D. Antigen-specific activation and cytokine-facilitated expansion of naive, human CD8+ T cells. *Nature Protocols* 9, 950-966 (2014).
- 35 148. Lipowska-Bhalla, G., Gilham, D. E., Hawkins, R. E. & Rothwell, D. G. Isolation of tumor antigen-specific single-chain variable fragment fragments using a chimeric antigen receptor bicistronic retroviral vector in a Mammalian screening protocol. *Hum Gene Ther Methods* 24, 381-391 (2013).
149. Kelly, R. J., Sharon, E., Pastan, I. & Hassan, R. Mesothelin-targeted agents in clinical trials and in preclinical development. *Mol. Cancer Ther.* 11, 517-525 (2012).
- 40 150. Atanackovic, D. et al. Surface molecule CD229 as a novel target for the diagnosis and treatment of multiple myeloma. *Haematologica* 96, 1512-1520 (2011).
151. Bund, D., Mayr, C., Kofler, D. M., Hallek, M. & Wendtner, C.-M. Human Ly9 (CD229) as novel tumor-associated antigen (TAA) in chronic lymphocytic leukemia (B-CLL) recognized by autologous CD8+ T cells. *Exp. Hematol.* 34, 860-869 (2006).
- 45 152. Tiede, C. et al. Adhiron: a stable and versatile peptide display scaffold for molecular recognition applications. *Protein Eng. Des. Sel.* 27, 145-155 (2014).
153. Maresca, M., Lin, V. G., Guo, N. & Yang, Y. Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Res.* 23, 539-546 (2013).
- 50 154. McVey, M. & Lee, S. E. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* 24, 529-538 (2008).
155. Nakade, S. et al. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat Commun* 5, 5560 (2014).
- 55 156. Chiche, L. et al. Squash inhibitors: From structural motifs to macrocyclic knottins. *Current Protein & Peptide Science* 5, 341-349.

LISTADO DE SECUENCIAS

- <110> Iontas Limited
- 60 <120> Preparación de bibliotecas de variantes de proteínas expresadas en células eucariotas y uso para seleccionar moléculas de unión
- <130> SMW/CP7119530
- 65 <150> GB1407852.1

ES 2 712 632 T3

<151> 02-05-2014

<160> 142

5 <170> PatentIn versión 3.5

<210> 1

<211> 5396

<212> ADN

10 <213> Secuencia artificial

<220>

<223> Casete de expresión del anticuerpo de promotor dual pD1 para la expresión superficial

15 <400> 1

```

ggtagcgaat tccgtgaggc tccggtgcc gtcagtgggc agagcgcaca tcgcccacag      60
tccccgagaa gttgggggga ggggtcggca attgaaccgg tgcctagaga aggtggcgcg      120
gggtaaaactg gaaagtgat gtcgtgtact ggctccgcct ttttcccgag ggtgggggag      180
aacggtatat aagtgcagta gtcgccgtga acgttctttt tcgcaacggg tttgcccca      240
gaacacaggt aagtgccgtg tgtggttccc gcggcctgg cctctttacg ggttatggcc      300
cttgctgccc ttgaattact tccacctggc tccagtacgt gattcttgat cccgagctgg      360
agccaggggc gggccttgcg ctttaggagc cccttcgcct cgtgcttgag ttgaggcctg      420
gcctggggcg tggggccgcc gcgtgcgaat ctggtggcac cttcgcgcct gtctcgtgc      480
tttcgataag tctctagcca tttaaaattt ttgatgacct gctgcgacgc tttttttctg      540
gcaagatagt cttgtaaagt cgggccagga tctgcacact ggtatttcgg tttttgggcc      600
cgcgccgggc gacggggccc gtgcgtccca gcgcacatgt tcggcgaggc ggggcctgcg      660
agcgcggcca ccgagaatcg gacgggggta gtctcaagct ggccggcctg ctctggtgcc      720
tggcctcgcg ccgccgtgta tcgccccgcc ctgggaggca aggctggccc ggtcggcacc      780
agttgctgta gcggaaagat ggccgcttcc cggccctgct ccagggggct caaaatggag      840
gacgcggcgc tcgggagagc gggcgggtga gtcaccaca caaaggaaa gggcctttcc      900
gtcctcagcc gtcgcttcat gtgactccac ggagtaccgg gcgccgtcca ggcacctcga      960
ttagttctgg agcttttggg gtacgtcgtc tttaggttgg ggggaggggt tttatgcgat     1020
ggagtttccc cacactgagt gggtaggagac tgaagttagg ccagcttggc acttgatgta     1080
attctcgttg gaatttgccc tttttgagtt tggatcttgg ttcattctca agcctcagac     1140

```

ES 2 712 632 T3

agtggttcaa agtttttttc ttccatttca ggtgtcgtga gacgtggcca ccatgagggc 1200
 ctggatcttc tttctccttt gcctggccgg gagggtctg gcagctagcg acatccagat 1260
 gaccagagc ccaagcagcc tgagcgccag cgtgggtgac agagtgacca tcacctgtag 1320
 agccagcggt aacatccaca actacctggc ttggtaccag cagaagccag gtaaggctcc 1380
 aaagctgctg atctactaca ccaccacct ggctgacggt gtgccaagca gattcagcgg 1440
 tagcggtagc ggtaccgact acaccttac catcagcagc ctccagccag aggacatcgc 1500
 cacctactac tgccagcact tctggagcac cccaaggacg ttcggccaag ggaccaaggt 1560
 ggaatcaaa cgtaccgagg ccgccccttc cgtgttcate ttccctccct ccgacgagca 1620
 gctgaagtcc ggcaccgctt ctgtggtgtg cctgctgaac aacttctacc ctcgggaggc 1680
 caaggtgcag tggaaaggtg acaacgccct gcagtccggc aactcccagg aatccgtcac 1740
 cgagcaggac tccaaggact ctacctact cctgtcctcc accctgacct tgtccaaggc 1800
 cgactacgag aagcacaagc tgtacgcctg cgaagtgacc caccagggcc tgtcctctcc 1860
 cgtgaccaag tccttcaacc ggggcgagtg ctaataaaag cttacgacgt gatcagcctc 1920
 gactgtgcct tctagttgcc agccatctgt tgtttgcccc tccccgtgc cttccttgac 1980
 cctggaaggt gccactocca ctgtccttcc ctaataaaat gaggaaattg catcgattg 2040
 tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga 2100
 ttgggaagac aatagcaggc atgctgggga acattgatta ttgactagtt attaatagta 2160
 atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgta cataacttac 2220
 ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac 2280
 gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt 2340
 acggtaaact gcccacttgg cagtacatca agtgtatcat atgccaagtc cggcccctat 2400
 tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttacggga 2460
 cttcctact tggcagtaca tctacgtatt agtcatcgct attacatag tgatgcggtt 2520
 ttggcagtac accaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca 2580
 cccattgac gtcaatggga gtttgttttg gcacaaaat caacgggact ttccaaaatg 2640
 tcgtaataac cccgccccgt tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta 2700
 tataagcaga gctcgtttag tgaaccgtca gatcctcact ctcttccgca tcgctgtctg 2760
 cgagggccag ctggtgggct cgcggttag gacaaactct tcgcggtctt tccagtactc 2820
 ttggatcgga aaccgctcg cctccgaacg gtactccgcc accgagggac ctgagcgagt 2880
 ccgatcgac cggatcggaa aacctctcgt gaaaggcgtc taaccagtca cagtcgcaag 2940
 gtaggctgag caccgtggcg ggccgagcg ggtggcggtc ggggttgttt ctggcggagg 3000

ES 2 712 632 T3

tgctgctgat gatgtaatta aagtaggcgg tcttgagacg gcggatggtc gaggtgaggt	3060
gtggcaggct tgagatccag ctgttggggg gagtactccc tctcaaaagc gggcattact	3120
tctgcgctaa gattgtcagt ttccaaaaac gaggaggatt tgatattcac ctggcccgat	3180
ctggccatac acttgagtga caatgacatc cactttgcct ttctctccac aggtgtccac	3240
tcccaggtcc aagtttgtgg aaattaatac gacgtggcca ccatgagttg gagctgtatc	3300
atcctcttct tggtagcaac agctacaggt aaggggttaa cagtagcagg cttgaggtct	3360
ggacatatat atgggtgaca atgacatcca ctttgccttt ctctccacag gcgccatggc	3420
ccaggtccaa ctgcaggaga gcggtccagg tcttgtgaga cctagccaga ccctgagcct	3480
gacctgcacc gtgtctggca gcaccttcag cggctatggt gtaaactggg tgagacagcc	3540
acctggacga ggtcttgagt ggattggaat gatttggggg gatggaaaca cagactataa	3600
ttcagctctc aaatccagag tgacaatgct ggtagacacc agcaagaacc agttcagcct	3660
gagactcagc agcgtgacag ccgccgacac cgcggtctat tattgtgcaa gagagagaga	3720
ttataggctt gactactggg gtcaaggcag cctcgtcaca gtctcgagtg ccagcaccaa	3780
gggccccagc gtgttccctc tggccccctg tagcagaagc accagcgaga gcacagccgc	3840
cctgggctgc ctggtcaagg actacttccc cgagcccgtg accgtgtcct ggaactctgg	3900
cgctctgacc agcggcgtgc acaccttcc agccgtgctg cagagcagcg gcctgtacag	3960
cctgagcagc gtggtcaccg tgcccagcag caacttcggc acccagacct acacctgtaa	4020
cgtggaccac aagcccagca acaccaaggt ggacaagacc gtggaacgga agtgctgctg	4080
ggaatgcccc ccctgtcccc ctctccagt ggctggacct tccgtgttcc tgttcccccc	4140
aaagcccaag gacaccctga tgatcagccg gacccccgaa gtgacctgcg tgggtggtgga	4200
cgtgtcccac gaggaccccc aggtgcagtt caattggtac gtggacggcg tggaaagtga	4260
caacgccaaag accaagccca gagaggaaca gttcaacagc accttcoggg tgggtgtccgt	4320
gctgaccgtg gtgcaccagg actggctgaa cggcaaagag tacaagtgcg ccgtctccaa	4380
caagggcctg cctgccccca tcgagaaaac catcagcaag accaagggcc agcctcgcga	4440
gcctcaggtg tacacactgc cccccagccg ggaagagatg accaagaacc aggtgtccct	4500
gacctgcctc gtgaagggtc tctaccccag cgatatcgcc gtggaatggg agagcaacgg	4560
ccagcccag aacaactaca agaccacccc ccccatgctg gacagcgacg gctcattctt	4620
cctgtacagc aagctgacag tggacaagag ccggtggcag cagggcaacg tgttcagctg	4680
cagcgtgatg cacgaggccc tgcacaacca ctacaccag aagtccctga gcctgagccc	4740
cggaaggga tccaaggtaa gtttaaacat atatataact ttaaataatt ggcattattt	4800
aaagttacta ctaactaacc ctgattattt aaatcttcag gaacaaaaac tcatctcaga	4860
agaggatctg aatgctgtgg gccaggacac gcaggaggtc atcgtggtgc cacactcctt	4920

ES 2 712 632 T3

gccctttaag gtggtggtga tctcagccat cctggccctg gtggtgctca ccatcatctc 4980
 ccttatcatc ctcatcatgc tttggcagaa gaagccacgt tagtaacagg taagagtgta 5040
 actttaaata atgccaatta tttaaagtta ctgactctct ctgcttacga cgcttcttct 5100
 ttttttttct ctgcaggggt agtaatcagc ctgcactgtg ccttctagtt gccagccatc 5160
 tgttgtttgc ccctccccg tgccttcctt gaccctggaa ggtgccactc ccactgtcct 5220
 ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg 5280
 ggggtggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcatgctgg 5340
 ggatggcccg ggcatgataa cttcgtataa tgtatgctat acgaagtat gtatac 5396

<210> 2
 <211> 233
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> Líder BM40, secuencia D1.3 VL humanizada y C kappa humana

10

<400> 2

Met	Arg	Ala	Trp	Ile	Phe	Phe	Leu	Leu	Cys	Leu	Ala	Gly	Arg	Ala	Leu
1				5					10					15	
Ala	Ala	Ser	Asp	Ile	Gln	Met	Thr	Gln	Ser	Pro	Ser	Ser	Leu	Ser	Ala
			20					25					30		
Ser	Val	Gly	Asp	Arg	Val	Thr	Ile	Thr	Cys	Arg	Ala	Ser	Gly	Asn	Ile
		35					40					45			
His	Asn	Tyr	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Lys	Ala	Pro	Lys
	50					55					60				
Leu	Leu	Ile	Tyr	Tyr	Thr	Thr	Thr	Leu	Ala	Asp	Gly	Val	Pro	Ser	Arg
65					70					75					80
Phe	Ser	Gly	Ser	Gly	Ser	Gly	Thr	Asp	Tyr	Thr	Phe	Thr	Ile	Ser	Ser
				85					90					95	
Leu	Gln	Pro	Glu	Asp	Ile	Ala	Thr	Tyr	Tyr	Cys	Gln	His	Phe	Trp	Ser
			100					105						110	
Thr	Pro	Arg	Thr	Phe	Gly	Gln	Gly	Thr	Lys	Val	Glu	Ile	Lys	Arg	Thr
		115					120					125			
Ala	Ala	Ala	Pro	Ser	Val	Phe	Ile	Phe	Pro	Pro	Ser	Asp	Glu	Gln	Leu
	130					135					140				

ES 2 712 632 T3

Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro
145 150 155 160

Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly
165 170 175

Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr
180 185 190

Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His
195 200 205

Lys Leu Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val
210 215 220

Thr Lys Ser Phe Asn Arg Gly Glu Cys
225 230

5 <210> 3
<211> 15
<212> PRT
<213> Secuencia artificial

10 <220>
<223> Secuencia líder 1
<400> 3

Met Ser Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr
1 5 10 15

15 <210> 4
<211> 449
<212> PRT
<213> Secuencia artificial

20 <220>
<223> Secuencia líder 2 y D1.3 VH, dominio CH1, Bisagra, dominios CH2 y CH3
<400> 4

Gly Ala Met Ala Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val
1 5 10 15

Arg Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Ser Thr
20 25 30

Phe Ser Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Arg Gly
35 40 45

25 Leu Glu Trp Ile Gly Met Ile Trp Gly Asp Gly Asn Thr Asp Tyr Asn

ES 2 712 632 T3

50	55	60																		
Ser 65	Ala	Leu	Lys	Ser	Arg 70	Val	Thr	Met	Leu	Val 75	Asp	Thr	Ser	Lys	Asn 80					
Gln	Phe	Ser	Leu	Arg 85	Leu	Ser	Ser	Val	Thr 90	Ala	Ala	Asp	Thr	Ala 95	Val					
Tyr	Tyr	Cys	Ala 100	Arg	Glu	Arg	Asp	Tyr 105	Arg	Leu	Asp	Tyr	Trp 110	Gly	Gln					
Gly	Ser	Leu 115	Val	Thr	Val	Ser	Ser	Ala 120	Ser	Thr	Lys	Gly 125	Pro	Ser	Val					
Phe 130	Pro	Leu	Ala	Pro	Cys	Ser	Arg 135	Ser	Thr	Ser	Glu 140	Ser	Thr	Ala	Ala					
Leu 145	Gly	Cys	Leu	Val	Lys 150	Asp	Tyr	Phe	Pro	Glu 155	Pro	Val	Thr	Val	Ser 160					
Trp	Asn	Ser	Gly	Ala 165	Leu	Thr	Ser	Gly	Val 170	His	Thr	Phe	Pro	Ala 175	Val					
Leu	Gln	Ser	Ser 180	Gly	Leu	Tyr	Ser	Leu 185	Ser	Ser	Val	Val	Thr 190	Val	Pro					
Ser	Ser	Asn 195	Phe	Gly	Thr	Gln	Thr 200	Tyr	Thr	Cys	Asn	Val 205	Asp	His	Lys					
Pro 210	Ser	Asn	Thr	Lys	Val	Asp 215	Lys	Thr	Val	Glu	Arg 220	Lys	Cys	Cys	Val					
Glu 225	Cys	Pro	Pro	Cys	Pro 230	Ala	Pro	Pro	Val	Ala 235	Gly	Pro	Ser	Val	Phe 240					
Leu	Phe	Pro	Pro	Lys 245	Pro	Lys	Asp	Thr	Leu 250	Met	Ile	Ser	Arg	Thr 255	Pro					
Glu	Val	Thr	Cys 260	Val	Val	Val	Asp	Val 265	Ser	His	Glu	Asp	Pro 270	Glu	Val					
Gln	Phe	Asn 275	Trp	Tyr	Val	Asp	Gly 280	Val	Glu	Val	His	Asn 285	Ala	Lys	Thr					
Lys 290	Pro	Arg	Glu	Glu	Gln	Phe 295	Asn	Ser	Thr	Phe	Arg 300	Val	Val	Ser	Val					

ES 2 712 632 T3

Leu Thr Val Val His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
305 310 315 320

Ala Val Ser Asn Lys Gly Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
325 330 335

Lys Thr Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
340 345 350

Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
355 360 365

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
370 375 380

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Met Leu Asp Ser Asp
385 390 395 400

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
405 410 415

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
420 425 430

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Ser
435 440 445

Lys

5 <210> 5
<211> 60
<212> PRT
<213> Artificial

10 <220>
<223> Epítopo Myc, espaciador PDGFR y región transmembrana (TM) PDGFR
<400> 5

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ala Val Gly Gln Asp
1 5 10 15

Thr Gln Glu Val Ile Val Val Pro His Ser Leu Pro Phe Lys Val Val
20 25 30

Val Ile Ser Ala Ile Leu Ala Leu Val Val Leu Thr Ile Ile Ser Leu
35 40 45

Ile Ile Leu Ile Met Leu Trp Gln Lys Lys Pro Arg

15 50 55 60

20 <210> 6
<211> 1542
<212> ADN
<213> Secuencia artificial

<220>

ES 2 712 632 T3

<223> fragmento de construcción donante pD1-huD1.3 que contiene brazo de homología izquierdo AAVS1, sitio de aceptor de corte y empalme, gen blasticidina, SV40 poliA

<400> 6

5

```

ggtaccgaat tcgccctttg ctttctctga cctgcattct ctcccctggg cctgtgcccg      60
tttctgtctg cagcttgtgg cctgggtcac ctctacggct ggcccagatc cttccctgcc      120
gcctccttca ggttccgtct tccctcaactc cctcttcccc ttgctctctg ctgtgttggc      180
gcccgaaggat gctctttccg gagcaattcc ttctcggcgc tgcaccacgt gatgtcctct      240
gagcggatcc tcccctgtgc tgggtcctct ccgggcatct ctccctccctc acccaacccc      300
atgccgtctt cactcgtctg gttccctttt ccttctcctt ctggggcctg tgccatctct      360
cgtttcttag gatggccttc tccgacggat gtctcccttg cgtcccgcct ccccttcttg      420
taggcctgca tcatcaccgt ttttctggac aaccccaaag taccctgtct ccctggcttt      480
agccacctct ccatcctctt gctttctttg cctggacacc ccgttctcct gtggattcgg      540
gtcacctctc actcctttca tttgggcagc tcccctaccc cccttacctc tctagtctgt      600
gcaagctctt ccagccccct gtcattggcat cttccagggg tccgagagct cagctagtct      660
tcttctcca acccgggccc ctatgtccac ttcaggacag catgtttgct gcctccaggg      720
atcctgtgtc cccgagctgg gaccacctta tattcccagg gccggttaat gtggctctgg      780
ttctgggtac ttttatctgt cccctccacc ccacagtggg gcaagatgca tcttctgacc      840
tcttctcttc ctcccacagc gcatggcaaa acctctgagc caggaagaaa gcacactgat      900
tgaaagagca accgctaact tcaacagcat ccccatctcc gaagactatt ctgtggctag      960
tgccgctctg tccagcgacg ggagaatctt caccgggtgtg aacgtctacc actttacagg     1020
cggaccatgc gcagagctgg tggctctggg gactgcagcc gctgcagccg ctggtaatct     1080
gacctgtatc gtggccattg gcaacgaaaa taggggcatac ctgtcccatc gcggcaggtg     1140
tcggcaggtg ctgctggatc tgcacacctg catcaaggca attgtcaaag actctgatgg     1200
acagcctacc gccgtcggta tccgtgaact gctgcctagc ggctatgtct gggagggata     1260
atgagcttgg cttogaaatg accgaccaag cgacgcccaa cctgccatca cgagatttgg     1320
attccaccgc cgccttctat gaaaggttgg gcttcgggaat cgttttccgg gacgccggct     1380
ggatgatcct ccagcgcggg gatctcatgc tggagtctt cgtccacccc aacttgttta     1440
ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat     1500
tttttctact gcattctagt tgtggtttaa ttaagtcaat tc                               1542

```

10

<210> 7
 <211> 924
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <223> Fragmento pD1-huD1.3 que contiene el brazo de homología derecho AAVS1

<400> 7

ES 2 712 632 T3

```

tggcccgggc atgataactt cgtataatgt atgctatacg aagttatgta tacggcgcgc      60
ccactagggg caggattggg gacagaaaag ccccatcctt aggcctcctc cttcctagtc      120
tcctgatatt gggctaaacc cccacctcct gttaggcaga ttccttatct ggtgacacac      180
ccccatttcc tggagccatc tctctccttg ccagaacctc taaggtttgc ttacgatgga      240
gccagagagg atcctgggag ggagagcttg gcagggggtg ggaggggaag gggggatgcg      300
tgacctgccc ggttctcagt ggcaccctg cgctaccctc tcccagaacc tgagctgctc      360
tgacgcggct gtctgggtgc tttcactgat cctgggtgctg cagcttcctt aacttccca      420
agaggagaag cagtttggaa aaacaaaatc agaataagtt ggtcctgagt tctaactttg      480
gctcttcacc tttctagtcc ccaatttata ttgttcctcc gtgcgtcagt tttacctgtg      540
agataaggcc agtagccagc cccgtcctgg cagggtctgtg gtgaggaggg ggggtgccgt      600
gtggaaaact ccctttgtga gaatggtgcg tcctaggtgt tcaccaggtc gtggccgcct      660
ctactccctt tctctttctc catccttctt tccttaaaga gtccccagtg ctatctggga      720
catattcctc cgcccagagc aggggtcccgc ttccctaagg ccctgctctg ggcttctggg      780
tttgagtctt tggcaagccc aggagaggcg ctcaggcttc cctgtcccc ttctcgtcc      840
accatctcat gccctggct ctccctcccc ttccctacag gggttcctgg ctctgctctg      900
acgcgtgtat actcgatcct tccg                                           924

```

<210> 8
 <211> 132
 <212> PRT
 <213> Secuencia artificial

5

<220>
 <223> Gen de resistencia a blasticidina sintético

10

<400> 8

Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala
 1 5 10 15

Thr Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr Ser Val Ala
 20 25 30

ES 2 712 632 T3

Ser Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly Val Asn Val
 35 40 45

Tyr His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val Leu Gly Thr
 50 55 60

Ala Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val Ala Ile Gly
 65 70 75 80

Asn Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys Arg Gln Val
 85 90 95

Leu Leu Asp Leu His Pro Gly Ile Lys Ala Ile Val Lys Asp Ser Asp
 100 105 110

Gly Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro Ser Gly Tyr
 115 120 125

Val Trp Glu Gly
 130

5 <210> 9
 <211> 734
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Vector pD6 para la expresión scFv-Fc, fragmento del sitio Nco1 al sitio PmeI que contiene un fragmento (últimos 4 aminoácidos) de la secuencia líder, el dominio CH2 completo y el dominio CH3

<400> 9

ggcgccatgg	cccaggtcgc	ggccgccgtg	gaatgcccc	cctgtcccgc	tcctccagtg	60
gctggacctt	ccgtgttctt	gttcccccca	aagcccaagg	acaccctgat	gatcagccgg	120
acccccgaag	tgacctgcgt	ggtggtggac	gtgtcccacg	aggaccccga	ggtgcagttc	180
aattggtacg	tggacggcgt	ggaagtgcac	aacgccaaga	ccaagcccag	agaggaacag	240
ttcaacagca	ccttccgggt	ggtgtccctg	ctgaccgtgg	tgcaccagga	ctggctgaac	300
ggcaaagagt	acaagtgcgc	cgtctccaac	aagggcctgc	ctgcccccat	cgagaaaacc	360
atcagcaaga	ccaagggcca	gcctcgcgag	cctcaggtgt	acacactgcc	ccccagccgg	420
gaagagatga	ccaagaacca	ggtgtccctg	acctgcctcg	tgaagggctt	ctaccccagc	480
gatatcgccg	tggaatggga	gagcaacggc	cagcccgaga	acaactacaa	gaccaccccc	540
cccatgctgg	acagcgacg	ctcattcttc	ctgtacagca	agctgacagt	ggacaagagc	600
cgggtggcagc	agggcaacgt	gttcagctgc	agcgtgatgc	acgaggccct	gcacaaccac	660
tacaccacaga	agtcocctgag	cctgagcccc	ggcaagggat	ccaaggtaa	tttaaacata	720
tatataactt	taaa					734

20 <210> 10
 <211> 4
 <212> PRT
 <213> Secuencia artificial

<220>

ES 2 712 632 T3

<223> Vector pD6, fragmento al final de la secuencia líder

<400> 10

5 Gly Ala Met Ala
1

<210> 11

<211> 219

<212> PRT

10 <213> Secuencia artificial

<220>

<223> Vector pD6 dominios CH2 y CH3

15 <400> 11

Ala Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
1 5 10 15

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
20 25 30

Val Asp Val Ser His Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val
35 40 45

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
50 55 60

Phe Asn Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Val His Gln
65 70 75 80

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Ala Val Ser Asn Lys Gly
85 90 95

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly Gln Pro
100 105 110

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
115 120 125

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
130 135 140

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
145 150 155 160

Lys Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
165 170 175

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
180 185 190

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
195 200 205

Ser Leu Ser Leu Ser Pro Gly Lys Gly Ser Lys
210 215

20

ES 2 712 632 T3

<210> 12
 <211> 60
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Fragmento de la secuencia del vector pD4 anterior y que incluye el inicio del brazo de homología izquierdo AAVS1

10

<400> 12
 ggtaccgaat tcgcccttg ctttctcga cctgcattct ctcccctggg cctgtgccgc 60

<210> 13
 <211> 3933
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <223> Fragmento de la secuencia del vector pD4 que incluye el final del AAVS1 LHA, sitio FRT, sitio Lox2272, sitio de meganucleasa 1-Sce1, sitio de unión a GFP izquierdo TALEN, sitio de unión a GFP derecho TALEN, T2A, GFP, promotor PGK, puromicina delta TK, sitio loxP e inicio de AAVS1 RHA

20

<400> 13

	ttctgggtac ttttatctgt cccctccacc ccacagtggg gcaagatgca tgaagttcct	60
	attccgaagt tcctattctc tagaaagtat aggaacttcg accataactt cgtataaagt	120
	atcctatacg aagttatgcg atcgctcgcg cgtagggata acagggtaat aagtccaccg	180
	gtcgccacca tggtgagcaa gggcgaggag ctgttcactt ctgacctctt ctcttcctcc	240
	cacagggcct agagagatct ggcagcggag agggcagagg aagtcttcta acatgcggtg	300
	acgtggagga gaatcccgga ccgatggtga gcaagggaga agaactcttc accggggtgg	360
	tgcccatcct ggtcgagctg gacggcgacg tgaacggcca caagttcagc gtgtccggcg	420
	agggcgaggg cgatgccacc tacggcaagc tgaccctgaa gttcatctgc accaccggca	480
25	agctgcccgt gccctggccc accctcgtga ccaccctgac ctacggcgtg cagtgttca	540

ES 2 712 632 T3

gccgctaccc cgaccacatg aagcagcacg acttcttcaa gtccgccatg cccgaaggct 600
 acgtccagga gcgcaccatc ttcttcaagg acgacggcaa ctacaagacc cgcgccgagg 660
 tgaagttcga gggcgacacc ctggtgaacc gcatcgagct gaaggcctc gacttcaagg 720
 aggacggcaa catcctgggg cacaagctgg agtacaacta caacagccac aacgtctata 780
 tcatggccga caagcagaag aacggcatca aggtgaactt caagatccgc cacaacatcg 840
 aggacggcag cgtgcagctc gccgaccact accagcagaa ccccccatc ggcgacggcc 900
 ccgtgctgct gcccgacaac cactacctga gcaccagtc cgcctgagc aaagacccca 960
 acgagaagcg cgatcacatg gtctgctgg agttcgtgac cgcgcgggg atcactcacg 1020
 gcatggacga gctgtacaag aagctgagcc acggcttccc gccggcgggt gcggcgcagg 1080
 atgatggcac gctgcccatg tcttgtgcc aggagagcgg gatggaccgt caccctgcag 1140
 cctgtgcttc tgetaggatc aatgtgtagg tgatcagcct cgactgtgcc ttctagtgc 1200
 cagccatctg ttgtttgcc ctccccctg ccttccttga ccctggaagg tgccactccc 1260
 actgtccttt cctaataaaa tgaggaaatt gcatcgcat gtctgagtag gtgtcattct 1320
 attctggggg gtgggggtgg gcaggacagc aagggggagg attggaaga caatagcagg 1380
 catgctgggg atggcccaat tctaccgggt aggggaggcg cttttccca ggcagtctgg 1440
 agcatgcgct ttagcagccc cgctgggcac ttggcgctac acaagtggcc tctggcctcg 1500
 cacacattcc acatccaccg gtaggcgcca accggctccg ttctttggtg gcccttcgc 1560
 gccaccttct actcctccc tagtcaggaa gttcccccc gccccgcagc tcgctcgtg 1620
 caggacgtga caaatggaag tagcacgtct cactagtctc gtgcagatgg acagcaccgc 1680
 tgagcaatgg aagcgggtag gcctttgggg cagcggccaa tagcagcttt gctccttcgc 1740
 tttctgggct cagaggctgg gaaggggtgg gtccgggggc gggctcaggg gcgggctcag 1800
 gggcggggcg ggcgcccga ggtcctccgg aggccggca ttctgcacgc ttcaaaagcg 1860
 cagctctgcc gcgctgttct cctcttcctc atctccgggc ctttcgacct gcagccaacg 1920
 ccaccatggg gaccgagtac aagcccacgg tgcgcctcgc caccgcgac gacgtcccc 1980
 gggccgtacg caccctcgc gcccggttcg ccgactacct cgcacgcgc cacaccgtcg 2040
 acccggaccg ccacatcgag cgggtcaccg agctgcaaga actcttcctc acgcgcgtcg 2100
 ggctcgacat cggcaagtg tgggtcgcgg acgacggcgc cgcggtggcg gtctggacca 2160
 cgccggagag cgtcgaagcg gggcggtgt tcgccgagat cggccgcgc atggccgagt 2220
 tgagcggttc ccggtggcc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc 2280
 ccaaggagcc cgcgtggtc ctggccaccg tcggcgtctc gcccgaccac cagggcaagg 2340
 gtctgggcag cgcgctcgtg ctccccggag tggagggcgc cgagcgcgc ggggtgccg 2400
 ccttctgga gacctccgc ccccgcaacc tccccctta cgagcggctc ggcttcaccg 2460

ES 2 712 632 T3

tcaccgccga cgtcgagggtg cccgaaggac cgcgcacctg gtgcatgacc cgcaagcccg 2520
 gtgccggatc catgcccacg ctactgcggg tttatataga cggtcctcac gggatgggga 2580
 aaaccaccac cacgcaactg ctggtggccc tgggttcgcg cgacgatatc gtctacgtac 2640
 ccgagccgat gacttactgg caggtgctgg gggcttcgga gacaatcgcg aacatctaca 2700
 ccacacaaca ccgctcgcac cagggtgaga tatcggccgg ggacgcggcg gtggtaatga 2760
 caagcgccca gataacaatg ggcattgcctt atgccgtgac cgacgccgtt ctggctcctc 2820
 atatcggggg ggaggctggg agctcacatg ccccgcccc ggccctcacc ctcatcttgc 2880
 accgccatcc catgcgcccc ctctgtgctt acccggccgc gcgatacctt atgggcagca 2940
 tgacccccca ggcctgtctg gcgttcgtgg ccctcatccc gccgacctg cccgcgacaa 3000
 acatcgtggt gggggccctt ccggaggaca gacacatcga ccgctggcc aaacgccagc 3060
 gccccggcga gcggcttgac ctggctatgc tggcccgat tcgccgcgtt tacgggtgc 3120
 ttgccaatac ggtgcggtat ctgcagggcg gcgggtcgtg gcgggaggat tggggacagc 3180
 tttcggggac ggcctgccc cccagggtg ccgagcccc gagcaacgcg ggcccacgac 3240
 cccatatacg ggacacgtta tttaccctgt ttcgggcccc cgagttgctg gccccaacg 3300
 gcgacctgta caacgtgttt gcctggcctt tggactcctt ggccaaacgc ctccgtccca 3360
 tgcactcctt tctctggat tacgaccaat cgcgcccg ctgccgggac gccctgctgc 3420
 aactaacctc cgggatggtc cagaccacg tcaccacccc cggctccata ccgacgatct 3480
 gcgacctggc gcgcactgtt gcccgggaga tgggggaggc taactgagct cttagactcg 3540
 ctgatcagcc tcgactgtgc cttctagtgg ccagccatct gttgtttgcc cctccccgt 3600
 gccttcttg accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat 3660
 tgcatacgat tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag 3720
 caagggggag gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggc 3780
 ttctgaggcg gaaagaacca gctggggctc gagatccact agttctagcc tcgaggctag 3840
 agcggccggc cctataactt cgtataatgt atgctatacg aagttatcag gtaagttaac 3900
 agggcgcgcc cactagggac aggattggtg aca 3933

5 <210> 14
 <211> 20
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Fragmento de pD4 que incluye el final del brazo de homología derecho (RHA) AAVS1

<400> 14
 tctgctctga cgcgtgtata 20

15 <210> 15
 <211> 53
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> Fragmento de pD4 que comprende T2A y el inicio de la secuencia EGFP

<400> 15

ES 2 712 632 T3

Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
1 5 10 15

Glu Asn Pro Gly Pro Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly
20 25 30

Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys
35 40 45

Phe Ser Val Ser Gly
50

5 <210> 16
<211> 69
<212> PRT
<213> Secuencia artificial

10 <220>
<223> Fragmento de pD4 que comprende la parte terminal de la secuencia EGFP y la secuencia PEST de Mu
ornitina descarboxilasa
<400> 16

Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala
1 5 10 15

Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys Lys Leu Ser His Gly
20 25 30

Phe Pro Pro Ala Val Ala Ala Gln Asp Asp Gly Thr Leu Pro Met Ser
35 40 45

Cys Ala Gln Glu Ser Gly Met Asp Arg His Pro Ala Ala Cys Ala Ser
50 55 60

15 Ala Arg Ile Asn Val
65

20 <210> 17
<211> 18
<212> PRT
<213> Secuencia artificial

<220>
<223> Fragmento inicial de la secuencia purodelta TK codificada por pD4

25 <400> 17

Met Gly Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp
1 5 10 15

Val Pro

30 <210> 18
<211> 15
<212> PRT
<213> Secuencia artificial

35 <220>
<223> Fragmento terminal de la secuencia puromicina delta TK codificada por pD4

ES 2 712 632 T3

```

caggcatgct ggggatggcc cgggcatgat aacttcgtat aatgtatgct atacgaagtt      60
atgtatacgg cgcgcccag caagggcgag gagctgttca cttctgacct cttctcttcc      120
tcccacctga gcctagagag atctggcagc ggagagggca gaggaagtct tctaaccatgc      180
ggtgacgtgg aggagaatcc cggaccgtga gtgagcaagg gagaagaact cttcaccggg      240
gtggtgcccc tcctggtcga gctggacggc gacgtgaacg gccacaagtt cagcgtgtcc      300
ggcgagggcg agggcgatgc cacctacggc aagctgacct tgaagttcat ctgcaccacc      360
ggcaagctgc ccgtgccctg gccaccctc gtgaccacc tgaacctacgg cgtgcagtgc      420
ttcagccgct accccgacca catgaagcag cagcacttct tcaagtccgc catgcccgaa      480
ggctacgtcc aggagcgcac catcttcttc aaggacgacg gcaactaaa gaccgcgccc      540
gaggtgaagt tcgagggcga caccctggtg aaccgcatcg agctgaaggg catcgacttc      600
aaggaggacg gcaacatcct ggggcacaag ctggagtaca actacaacag ccacaacgtc      660
tatatcatgg ccgacaagca gaagaacggc atcaagggtga acttcaagat ccgccacaac      720
atcgaggacg gcagcgtgca gctcgccgac cactaccagc agaacacccc catcgccgac      780
ggccccgtgc tgctgcccga caaccactac ctgagcacc agtccgccct gagcaaagac      840
cccaacgaga agcgcgatca catggtcctg ctggagttcg tgaccgccgc cgggatcact      900
cacggcatgg acgagcctga cgcgtgtata c      931

```

<210> 21

<211> 131

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de proteína de resistencia a blasticidina codificada por plásmido donante para la integración de los sitios FLP/GFP TALEN (pD5)

<400> 21

ES 2 712 632 T3

Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala Thr
 1 5 10 15

Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr Ser Val Ala Ser
 20 25 30

Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly Val Asn Val Tyr
 35 40 45

His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val Leu Gly Thr Ala
 50 55 60

Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val Ala Ile Gly Asn
 65 70 75 80

Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys Arg Gln Val Leu
 85 90 95

Leu Asp Leu His Pro Gly Ile Lys Ala Ile Val Lys Asp Ser Asp Gly
 100 105 110

Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro Ser Gly Tyr Val
 115 120 125

Trp Glu Gly
 130

<210> 22
 <211> 813
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia de la construcción de meganucleasa I-Sce-1

<400> 22

ccatgggcta tccttacgat gtccttgatt acgccaacag tcctggtatc cctggtatgg 60
 gtcctaaaaa gaagcgaaaa gtgggtagac tggaacccgg catgaagaac attaagaaaa 120
 atcaggtgat gaacctggga cctaattcca agctgctgaa agagtacaag tctcagctga 180
 tcgaactgaa cattgagcag tttgaagcag ggatcggtct gattctgggg gacgcctaca 240
 tccggagcag ggatgagggc aagacttatt gcatgcagtt cgaatggaag aataaggcct 300
 acatggacca cgtgtgtctg ctgtatgatc agtgggtcct gtctcccct cacaagaaag 360

5

10

ES 2 712 632 T3

```

agagagtgaa ccatctgggc aatctggtca ttacttgggg agcacagacc ttcaagcatc      420
aggcctttaa caaactggct aacctgttca tcgtgaacaa caagaaaacc atccctaaca      480
atctggtcga aaactacctg acaccaatga gtctggccta ttggttcatg gacgatggcg      540
gaaaatggga ctacaacaag aacagcacia acaaagcat cgtgctgaat acccagtct      600
tcacatttga ggaagtggag tatctggtca agggcctgcg gaacaaattc cagctgaact      660
gctacgtgaa gatcaacaag aacaagccaa tcactacat cgattctatg agttacctga      720
tcttttataa cctgattaag ccatacctga tccccagat gatgtataaa ctgcccaata      780
caatcagctc cgagactttc ctgaaggtct aga                                  813
    
```

<210> 23

<211> 269

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia codificada por la construcción de meganucleasa I-SceI: etiqueta HA, NLS, e I-SceI

<400> 23

```

Met Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Asn Ser Pro Gly Ile
 1           5           10          15

Pro Gly Met Gly Pro Lys Lys Lys Arg Lys Val Gly Arg Leu Glu Pro
 20           25           30

Gly Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn
 35           40           45

Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile
 50           55           60

Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile
 65           70           75           80

Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys
 85           90           95

Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val
 100          105          110

Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu
 115          120          125

Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys
 130          135          140
    
```


ES 2 712 632 T3

Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn Asn
145 150 155 160

Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met
165 170 175

Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys Ser
180 185 190

Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr Leu
195 200 205

Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile
210 215 220

Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile
225 230 235 240

Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr Lys
245 250 255

Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys Val
260 265

<210> 24
<211> 67
5 <212> ADN
<213> Secuencia artificial

<220>
10 <223> Fragmento del vector pINT20 que muestra el sitio Nsi1, sitio de aceptor de corte y empalme e inicio de la secuencia del gen de resistencia a puromicina

<400> 24
atgcatcttc tgacctcttc tcttcctccc acagggcatg accgagtaca agcccacggt 60
gcgccctc 67

15 <210> 25
<211> 10
<212> PRT
<213> Secuencia artificial

20 <220>
<223> Fragmento inicial de la proteína de resistencia a puromicina codificada por el vector pINT20

25 <400> 25
Met Thr Glu Tyr Lys Pro Thr Val Arg Leu
1 5 10

30 <210> 26
<211> 1125
<212> ADN
<213> Secuencia artificial

35 <220>
<223> Secuencia de la construcción de cadena alfa del clon del receptor de linfocitos T c12/c2 que contiene los sitios de restricción Nhe1, Not1 y Acc651

<400> 26

ES 2 712 632 T3

```

gctagcaagc aggaagtgac tcagatccca gccgctctga gcgtgcctga gggagaaaac      60
ctggtcctga attgcagttt caccgactca gccatctata acctgcagtg gtttcgccag      120
gatccaggca agggactgac ctccctgctg ctgattcaga gctcccagag ggaacagaca      180
tctggcagac tgaatgctag tctggacaaa tctagtggac ggtctaccct gtacatcgca      240
gccagccagc ctggagattc cgcaacatat ctgtgcgccg tgcgccact tacaggcgga      300
agctacattc ccaccttcgg gcgaggtaca agcctgatcg tgcaccaga catccagaat      360
ccggagcccg ccgtatacca gctgaaggac cccagaagcc aggacagcac cctgtgcctg      420
ttcaccgact tcgacagcca gatcaactg cccaagacaa tggaaagcgg caccttcac      480
accgacaaga ccgtgctgga catgaaggct atggacagca agagcaacgg cgccattgcc      540
tggccaacc agaccagctt cacatgccag gacatcttca aagagacaaa cgccacctac      600
cccagcagcg acgtgccctg tgatgccacc ctgaccgaga agtccttca gacagacatg      660
aacctgaact tccagaacct gtccgcggcc gcaggcctgc tggatcccaa gctgtgctac      720
ctgctggacg ggatcctgtt catctacggt gtgactctga ctgccctgtt cctgcgagtc      780
aaattttctc ggagtgccga cgctcctgca taccagcagg ggcagaacca gctgtataac      840
gagctgaatc tgggtcggag ggaggaatat gacgtgctgg ataagagacg cggcagggat      900
ccagaaatgg ggggcaagcc ccagcgacgg aaaaacctc aggagggact gtataatgaa      960
ctgcagaagg acaaaatggc cgaggcttac tctgaaattg ggatgaaggg cgagaggaga     1020
cgcgcaaaag gacacgatgg cctgtaccag ggactgagca ctgctaccaa ggacacatat     1080
gatgctctgc atatgcaggc actgcccctc agataataag gtacc                          1125

```

5 <210> 27
 <211> 369
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> Secuencia de la cadena alfa del clon del receptor de linfocitos T c12/c2

<400> 27

```

Lys Gln Glu Val Thr Gln Ile Pro Ala Ala Leu Ser Val Pro Glu Gly
1           5           10           15

```

15

ES 2 712 632 T3

Glu Asn Leu Val Leu Asn Cys Ser Phe Thr Asp Ser Ala Ile Tyr Asn
 20 25 30
 Leu Gln Trp Phe Arg Gln Asp Pro Gly Lys Gly Leu Thr Ser Leu Leu
 35 40 45
 Leu Ile Gln Ser Ser Gln Arg Glu Gln Thr Ser Gly Arg Leu Asn Ala
 50 55 60
 Ser Leu Asp Lys Ser Ser Gly Arg Ser Thr Leu Tyr Ile Ala Ala Ser
 65 70 75 80
 Gln Pro Gly Asp Ser Ala Thr Tyr Leu Cys Ala Val Arg Pro Leu Thr
 85 90 95
 Gly Gly Ser Tyr Ile Pro Thr Phe Gly Arg Gly Thr Ser Leu Ile Val
 100 105 110
 His Pro Asp Ile Gln Asn Pro Glu Pro Ala Val Tyr Gln Leu Lys Asp
 115 120 125
 Pro Arg Ser Gln Asp Ser Thr Leu Cys Leu Phe Thr Asp Phe Asp Ser
 130 135 140
 Gln Ile Asn Val Pro Lys Thr Met Glu Ser Gly Thr Phe Ile Thr Asp
 145 150 155 160
 Lys Thr Val Leu Asp Met Lys Ala Met Asp Ser Lys Ser Asn Gly Ala
 165 170 175
 Ile Ala Trp Ser Asn Gln Thr Ser Phe Thr Cys Gln Asp Ile Phe Lys
 180 185 190
 Glu Thr Asn Ala Thr Tyr Pro Ser Ser Asp Val Pro Cys Asp Ala Thr
 195 200 205
 Leu Thr Glu Lys Ser Phe Glu Thr Asp Met Asn Leu Asn Phe Gln Asn
 210 215 220
 Leu Ser Ala Ala Ala Gly Leu Leu Asp Pro Lys Leu Cys Tyr Leu Leu
 225 230 235 240
 Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Leu Thr Ala Leu Phe Leu
 245 250 255
 Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln Gly
 260 265 270

ES 2 712 632 T3

Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
 275 280 285

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
 290 295 300

Pro Gln Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln
 305 310 315 320

Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu
 325 330 335

Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr
 340 345 350

Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro
 355 360 365

Arg

<210> 28
 <211> 1226
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia de la construcción de cadena veta del clon del receptor de linfocitos T c12/c2 que contiene los sitios de restricción Nco1, Xho1 y Hind3

<400> 28

```

ccatggccaa cgctggagtg actcagacct ctaagttcca ggtcctgaaa actgggcaga      60
gtatgaccct gcagtgcgca caggacatga atcacgagta catgtcatgg tatcggcagg      120
atccagggat gggctctgagg ctgatccatt acagcgtggg cgctggaact accgaccagg      180
gcgaggtgcc caacggatat aatgtctcaa gaagcaccac agaagatttc cactgcgac      240
tgctgagcgc cgctcctagc cagacatccg tgtacttttg cgccagctcc aatgtcggga      300
acaccggcga gctgttcttt ggggaaggtt cccgcctgac agtgctcgag gacctgagaa      360
acgtgacccc cccaaggtg tccctgttcg agcctagcaa ggccgagatc gccaacaagc      420
agaaagccac cctcgtgtgc ctggccagag gcttcttccc cgaccacgtg gaactgtctt      480
ggtaggtcaa cggcaaagag gtgcacagcg gctgtgccac cgatccccag gcctacaaa      540
agagcaacta cagctactgc ctgagcagca gactgcgggt gtccgccacc ttctggcaca      600
acccccgaa cacttcaga tgccaggtgc agtttcacgg cctgagcga gaggacaagt      660
ggcccagggg cagccctaag cccgtgacct agaatatctc tgccgaagcc tggggcagag      720
    
```

5

10

15

ES 2 712 632 T3

```

ccgactgtgg cattaccagc gccagctacc agcagggcgt gctgtctgcc accatcctgt      780
acgaggtcgc gagcggactg ctggacccaa agctgtgcta cctgctggat gggatcctgt      840
tcatctacgg tgtgattctg acagccctgt tcctgcgagt caagttcagc cggagcgcgg      900
acgcaccagc ataccagcag gggcagaatc agctgtataa cgagctgaat ctgggtcgga      960
gggaggaata cgacgtgctg gataagagac gcggcagggg tcccgaaatg ggcggaaagc     1020
ctcagcgacg gaaaaaccca caggagggac tgtacaatga actgcagaag gacaaaaatg     1080
ctgaggcata ttctgaaatc ggcataaagg gagagaggag acgcgggcaa ggacacgatg     1140
ggctgtacca ggtctgagt acagccacta aggacaccta tgatgccctg catatgcagg     1200
ctctgccacc cagataataa aagctt                                           1226

```

<210> 29

<211> 404

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia de la construcción de cadena veta del clon del receptor de linfocitos T c12/c2

<400> 29

```

Met Ala Asn Ala Gly Val Thr Gln Thr Pro Lys Phe Gln Val Leu Lys
 1                               5                               10 15

Thr Gly Gln Ser Met Thr Leu Gln Cys Ala Gln Asp Met Asn His Glu
 20                               25                               30

Tyr Met Ser Trp Tyr Arg Gln Asp Pro Gly Met Gly Leu Arg Leu Ile
 35                               40                               45

His Tyr Ser Val Gly Ala Gly Thr Thr Asp Gln Gly Glu Val Pro Asn
 50                               55                               60

Gly Tyr Asn Val Ser Arg Ser Thr Thr Glu Asp Phe Pro Leu Arg Leu
 65                               70                               75 80

Leu Ser Ala Ala Pro Ser Gln Thr Ser Val Tyr Phe Cys Ala Ser Ser
 85                               90                               95

Asn Val Gly Asn Thr Gly Glu Leu Phe Phe Gly Glu Gly Ser Arg Leu
 100                              105                              110

Thr Val Leu Glu Asp Leu Arg Asn Val Thr Pro Pro Lys Val Ser Leu
 115                              120                              125

Phe Glu Pro Ser Lys Ala Glu Ile Ala Asn Lys Gln Lys Ala Thr Leu
 130                              135                              140

```

ES 2 712 632 T3

Val Cys Leu Ala Arg Gly Phe Phe Pro Asp His Val Glu Leu Ser Trp
 145 150 155 160

Trp Val Asn Gly Lys Glu Val His Ser Gly Val Ser Thr Asp Pro Gln
 165 170 175

Ala Tyr Lys Glu Ser Asn Tyr Ser Tyr Cys Leu Ser Ser Arg Leu Arg
 180 185 190

Val Ser Ala Thr Phe Trp His Asn Pro Arg Asn His Phe Arg Cys Gln
 195 200 205

Val Gln Phe His Gly Leu Ser Glu Glu Asp Lys Trp Pro Glu Gly Ser
 210 215 220

Pro Lys Pro Val Thr Gln Asn Ile Ser Ala Glu Ala Trp Gly Arg Ala
 225 230 235 240

Asp Cys Gly Ile Thr Ser Ala Ser Tyr Gln Gln Gly Val Leu Ser Ala
 245 250 255

Thr Ile Leu Tyr Glu Val Ala Ser Gly Leu Leu Asp Pro Lys Leu Cys
 260 265 270

Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Leu Thr Ala
 275 280 285

Leu Phe Leu Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr
 290 295 300

Gln Gln Gly Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg
 305 310 315 320

Glu Glu Tyr Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met
 325 330 335

Gly Gly Lys Pro Gln Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn
 340 345 350

Glu Leu Gln Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met
 355 360 365

Lys Gly Glu Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly
 370 375 380

Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala
 385 390 395 400

Leu Pro Pro Arg

- 5 <210> 30
- <211> 677
- <212> ADN
- <213> Secuencia artificial

ES 2 712 632 T3

<220>

<223> Secuencia de la construcción de cadena alfa del clon del receptor de linfocitos T 4JFH con sitios de restricciones Nhe1 inicial y Not1 terminal

5 <400> 30

```

gctagccaga aagaggtgga acagaacagc ggcctctga gcgtgccaga aggcgctatc      60
gccagcctga actgcaccta cagctttctg ggcagccaga gcttcttctg gtacagacag      120
tacagcggca agagccccga gctgatcatg ttcacctaca gagagggcga caaagaggac      180
ggcagattca ccgcccagct gaacaaggcc agccagcacg tgtccctgct gatcagagac      240
agccagccta gcgacagcgc cacctacctg tgcgccgtga atgatggcgg cagactgacc      300
tttggcgacg gcaccaccct gaccgtgaag cctgacatcc agaatccgga gcccgccgta      360
taccagctga aggaccccag aagccaggac agcaccctgt gcctgttcac cgacttcgac      420
agccagatca acgtgcccaa gacaatggaa agcggcacct tcatcaccga caagaccgtg      480
ctggacatga aggctatgga cagcaagagc aacggcgcca ttgcctggtc caaccagacc      540
agcttcacat gccaggacat cttcaaagag acaaacgcc cctaccccag cagcgacgtg      600
ccctgtgatg ccaccctgac cgagaagtcc ttcgagacag acatgaacct gaacttccag      660
aacctgtccg cggccgc                                                    677
    
```

<210> 31

10 <211> 221

<212> PRT

<213> Secuencia artificial

<220>

15 <223> Secuencia de la construcción de cadena alfa del clon del receptor de linfocitos T 4JFH

<400> 31

```

Gln Lys Glu Val Glu Gln Asn Ser Gly Pro Leu Ser Val Pro Glu Gly
 1           5           10           15

Ala Ile Ala Ser Leu Asn Cys Thr Tyr Ser Phe Leu Gly Ser Gln Ser
      20           25           30

Phe Phe Trp Tyr Arg Gln Tyr Ser Gly Lys Ser Pro Glu Leu Ile Met
      35           40           45
    
```

20

ES 2 712 632 T3

Phe Thr Tyr Arg Glu Gly Asp Lys Glu Asp Gly Arg Phe Thr Ala Gln
 50 55 60

Leu Asn Lys Ala Ser Gln His Val Ser Leu Leu Ile Arg Asp Ser Gln
 65 70 75 80

Pro Ser Asp Ser Ala Thr Tyr Leu Cys Ala Val Asn Asp Gly Gly Arg
 85 90 95

Leu Thr Phe Gly Asp Gly Thr Thr Leu Thr Val Lys Pro Asp Ile Gln
 100 105 110

Asn Pro Glu Pro Ala Val Tyr Gln Leu Lys Asp Pro Arg Ser Gln Asp
 115 120 125

Ser Thr Leu Cys Leu Phe Thr Asp Phe Asp Ser Gln Ile Asn Val Pro
 130 135 140

Lys Thr Met Glu Ser Gly Thr Phe Ile Thr Asp Lys Thr Val Leu Asp
 145 150 155 160

Met Lys Ala Met Asp Ser Lys Ser Asn Gly Ala Ile Ala Trp Ser Asn
 165 170 175

Gln Thr Ser Phe Thr Cys Gln Asp Ile Phe Lys Glu Thr Asn Ala Thr
 180 185 190

Tyr Pro Ser Ser Asp Val Pro Cys Asp Ala Thr Leu Thr Glu Lys Ser
 195 200 205

Phe Glu Thr Asp Met Asn Leu Asn Phe Gln Asn Leu Ser
 210 215 220

<210> 32
 <211> 353
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia de la construcción de cadena beta del clon del receptor de linfocitos T 4JFH con sitio de restricción Nco1 inicial y sitio de restricción Xho1 terminal

<400> 32

ccatggccag ccagaccatc catcagtggc ctgccaccct ggtgcagcct gtgggatctc 60
 ctctgagcct ggaatgcacc gtggaaggca ccagcaaccc caacctgtac tggtagacagc 120
 aggccgctgg cagaggcccc cagctgctgt tttactgggg cccctttggc cagatcagca 180
 gcgaggtgcc ccagaacctg agcgccagca gaccccagga cgggcagttt atcctgagca 240
 gcaagaagct gctgctgagc gacagcggct tctacctgtg cgcttgagc gagacaggcc 300
 tgggcatggg cggatggcag tttggcgagg gcagcagact gacagtgctc gag 353

<210> 33
 <211> 117
 <212> PRT

ES 2 712 632 T3

<213> Secuencia artificial

<220>

<223> Secuencia de la construcción de cadena beta del clon del receptor de linfocitos T 4JFH

5

<400> 33

```

Met Ala Ser Gln Thr Ile His Gln Trp Pro Ala Thr Leu Val Gln Pro
 1           5           10           15

Val Gly Ser Pro Leu Ser Leu Glu Cys Thr Val Glu Gly Thr Ser Asn
          20           25           30

Pro Asn Leu Tyr Trp Tyr Arg Gln Ala Ala Gly Arg Gly Pro Gln Leu
          35           40           45

Leu Phe Tyr Trp Gly Pro Phe Gly Gln Ile Ser Ser Glu Val Pro Gln
          50           55           60

Asn Leu Ser Ala Ser Arg Pro Gln Asp Arg Gln Phe Ile Leu Ser Ser
65           70           75           80

Lys Lys Leu Leu Leu Ser Asp Ser Gly Phe Tyr Leu Cys Ala Trp Ser
          85           90           95

Glu Thr Gly Leu Gly Met Gly Gly Trp Gln Phe Gly Glu Gly Ser Arg
          100          105          110

Leu Thr Val Leu Glu
          115
    
```

10

<210> 34

<211> 53

<212> ADN

<213> Secuencia artificial

15

<220>

<223> Cebador para la mutación de CDR3 de la cadena alfa de TCR c12/c2

<220>

<221> misc_feature

20

<222> (20)..(21)

<223> n es a, c, g, o t

<220>

<221> misc_feature

25

<222> (23)..(24)

<223> n es a, c, g, o t

<220>

<221> misc_feature

30

<222> (31)..(31)

<223> n es a, c, g, o t

<400> 34

cacgcgccac gcggtgaan nsnscctwc natgtaaggg tggaagcccg ctc 53

35

<210> 35

<211> 52

<212> ADN

<213> Secuencia artificial

ES 2 712 632 T3

<220>
 <223> Cebador usado para mutar CDR3 de cadena alfa de TCR c12/c2

5 <220>
 <221> misc_feature
 <222> (23)..(23)
 <223> n es a, c, g, o t

10 <220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n es a, c, g, o t

15 <220>
 <221> misc_feature
 <222> (33)..(34)
 <223> n es a, c, g, o t

20 <400> 35
 ctcgcccgaa ggtgggaatg tangwtccsn nsnnaagtgg gcgcacggcg ca 52

25 <210> 36
 <211> 17
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <223> Estrategia usada para mutar CDR3 de cadena alfa de TCR c12/c2

35 <220>
 <221> misc_feature
 <222> (7)..(8)
 <223> Xaa puede ser cualquier aminoácido de origen natural

40 <400> 36

Cys	Ala	Val	Arg	Pro	Leu	Xaa	Xaa	Gly	Ser	Tyr	Ile	Pro	Thr	Phe	Gly
1			5					10						15	

Arg

45 <210> 37
 <211> 17
 <212> PRT
 <213> Secuencia artificial

50 <220>
 <223> Estrategia usada para mutar CDR3 de cadena alfa de TCR c12/c2

55 <220>
 <221> misc_feature
 <222> (7)..(8)
 <223> Xaa puede ser cualquier aminoácido de origen natural

60 <400> 37

Cys	Ala	Val	Arg	Pro	Leu	Xaa	Xaa	Gly	Thr	Tyr	Ile	Pro	Thr	Phe	Gly
1			5					10						15	

Arg

65 <210> 38
 <211> 50
 <212> ADN

ES 2 712 632 T3

<213> Secuencia artificial

<220>
<223> Cebador usado para mutar CDR3 de cadena beta de TCR c12/c2

5

<220>
<221> misc_feature
<222> (20)..(21)
<223> n es a, c, g, o t

10

<220>
<221> misc_feature
<222> (29)..(30)
<223> n es a, c, g, o t

15

<400> 38
gtactttgc gccagctccn nsstcgggnn saccgcgag ctgttcttg 50

<210> 39
<211> 49
<212> ADN
<213> Secuencia artificial

20

<220>
<223> Cebador usado para mutar CDR3 de cadena beta de TCR c12/c2

25

<220>
<221> misc_feature
<222> (21)..(22)
<223> n es a, c, g, o t

30

<220>
<221> misc_feature
<222> (30)..(31)
<223> n es a, c, g, o t

35

<400> 39
caaagaacag ctgccggts nccccgassn nggagctggc gcaaaagta 49

40

<210> 40
<211> 16
<212> PRT
<213> Secuencia artificial

45

<220>
<223> Estrategia usada para mutar CDR3 de cadena beta de TCR c12/c2

<400> 40

50

	Tyr	Phe	Cys	Ala	Ser	Ser	Asn	Val	Gly	Asn	Thr	Gly	Glu	Leu	Phe	Phe
	1				5					10					15	

<210> 41
<211> 16
<212> PRT
<213> Secuencia artificial

55

<220>
<223> Estrategia usada para mutar CDR3 de cadena beta de TCR c12/c2

60

<400> 41

	Tyr	Phe	Cys	Ala	Ser	Ser	Asn	Leu	Gly	Asn	Thr	Gly	Glu	Leu	Phe	Phe
	1				5					10					15	

<210> 42

ES 2 712 632 T3

<211> 462
<212> ADN
<213> Secuencia artificial

5 <220>
<223> Secuencia de CD3 zeta en pINT21 CAR1

<400> 42

```
ggcgccatgg cccaggctgc ggccgcaagc ggactgctgg acccaaagct gtgctacctg      60
ctggatggga tcctgttcat ctacggtgtg attctgacag ccctgttcct gcgagtcaag      120
ttcagccgga gcgccgacgc accagcatac cagcaggggc agaatcagct gtataacgag      180
ctgaatctgg gtcggaggga ggaatacgac gtgctggata agagacgcgg cagggatccc      240
gaaatgggcg gaaagcctca gcgacggaaa aaccacagc agggactgta caatgaactg      300
cagaaggaca aaatggctga ggcatattct gaaatcggca tgaaggaga gaggagacgc      360
ggcaaaggac acgatgggct gtaccagggt ctgagtacag ccactaagga cacctatgat      420
10 gccctgcata tgcaggctct gccaccaga taataaaagc tt                               462
```

<210> 43
<211> 4
<212> PRT
15 <213> Secuencia artificial

<220>
<223> Fragmento codificado por el vector pINT21 CAR1

20 <400> 43

Gly Ala Met Ala
1

25 <210> 44
<211> 141
<212> PRT
<213> Secuencia artificial

30 <220>
<223> Secuencia CDR3 zeta codificada por pINT21 CAR1

<400> 44

ES 2 712 632 T3

Ser Gly Leu Leu Asp Pro Lys Leu Cys Tyr Leu Leu Asp Gly Ile Leu
 1 5 10 15

Phe Ile Tyr Gly Val Ile Leu Thr Ala Leu Phe Leu Arg Val Lys Phe
 20 25 30

Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln Gly Gln Asn Gln Leu
 35 40 45

Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp
 50 55 60

Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Gln Arg Arg
 65 70 75 80

Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met
 85 90 95

Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg Gly
 100 105 110

Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp
 115 120 125

Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
 130 135 140

<210> 45
 <211> 705
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia pINT21 CAR2 que codifica CD3 zeta

<400> 45

gccatggccc aggtcgcggc cgcaacaaca accccagccc ccagacctcc taccctgccc 60
 cctacaattg ccagccagcc tctgagcctg aggcccgagg cttgtagacc agctgctggc 120
 ggagccgtgc acaccagagg actggatttc gcctgcgaca tctacatctg ggcccctctg 180
 gccggcacat gtggcgtgct gctgctgagc ctctgatca ccctgtactg caagcggggc 240
 agaaagaaac tgctgtacat ctttaagcag cccttcatgc ggcccgtgca gaccacccag 300
 gaagaggacg gctgctcctg cagattcccc gaggaagaag aaggcggctg cgagctgaga 360
 gtgaagtcca gcagatccgc cgacgccctt gcctacaagc agggccagaa ccagctgtac 420
 aacgagctga acctgggcag acgggaagag tacgacgtgc tggacaagcg gagaggccgg 480
 gaccagaga tgggggaaa gcccagaaga aagaaccccc aggaaggcct gtataacgaa 540
 ctgcagaaag acaaaatggc cgaggcctac agcgagatcg gaatgaaggc cgagcggaga 600
 agaggcaagg ggcacgatgg cctgtaccag ggcctgagca ccgccaccaa ggacacctat 660
 gacgcctgac acatgcaggc cctgccccct agataataaa agctt 705

<210> 46

ES 2 712 632 T3

<211> 223
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> Secuencia codificada por pINT21 CAR2: Región bisagra/transmembrana CD8, 4-1BB y CD3 zeta
 <400> 46

```

    Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala
    1          5          10          15

    Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly
    20          25          30

    Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile
    35          40          45

    Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val
    50          55          60

    Ile Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe
    65          70          75          80

    Lys Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly
    85          90          95

    Cys Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg
    100         105         110

    Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln
    115         120         125

    Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp
    130         135         140

    Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro
    145         150         155         160

    Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp
    165         170         175

    Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg
    180         185         190

    Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr
    195         200         205

    Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
    210         215         220
    
```

10
 15 <210> 47
 <211> 750
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Secuencia de FMC63 H-L (anticuerpo anti-CD19)

ES 2 712 632 T3

<400> 47

```
gccatggccg aagtgaaact gcaggagtct ggaccgggcc tggaggcccc atctcagtct      60
ctgagcgtga cctgtaccgt gtccggcgtg tccctgcctg actatggcgt gtcctggatc     120
agacagcccc ccagaaaggg cctggaatgg ctgggagtga tctggggcag cgaaccacc      180
tactacaaca gcgcctgaa gtcccggctg accatcatca aggacaactc caagagccag     240
gtgttcctga agatgaacag cctgcagacc gacgacaccg ccatctacta ctgcgccaa      300
cactactact acggcggcag ctacgctatg gactactggg gccagggcac ctcggtcacc     360
gtctcgagtg gtggaggcgg ttcaggcggg ggtggctctg gcggtggcgc tagcgacatc     420
cagatgacct agaccaccag cagcctgagc gccagcctgg gcgatagagt gaccatcagc     480
tgcaagacca gccaggacat cagcaagtac ctgaactggg atcagcagaa acccgacggc     540
accgtgaagc tgctgatcta ccacaccagc agactgcaca gcggcgtgcc cagcagattt     600
tccggctctg gcagcggcac cgactacagc ctgaccatct ccaacctgga acaggaagat     660
atcgctacct acttctgtca gcaaggcaac accctgccct acaccttcgg cggagggacc     720
aagctggaga tcaaactgac cgcgcccgca                                     750
```

5

<210> 48

<211> 250

10

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia de FMC63 H-L (anticuerpo anti-CD19)

15

<400> 48

ES 2 712 632 T3

Ala Met Ala Glu Val Lys Leu Gln Glu Ser Gly Pro Gly Leu Val Ala
 1 5 10 15

Pro Ser Gln Ser Leu Ser Val Thr Cys Thr Val Ser Gly Val Ser Leu
 20 25 30

Pro Asp Tyr Gly Val Ser Trp Ile Arg Gln Pro Pro Arg Lys Gly Leu
 35 40 45

Glu Trp Leu Gly Val Ile Trp Gly Ser Glu Thr Thr Tyr Tyr Asn Ser
 50 55 60

Ala Leu Lys Ser Arg Leu Thr Ile Ile Lys Asp Asn Ser Lys Ser Gln
 65 70 75 80

Val Phe Leu Lys Met Asn Ser Leu Gln Thr Asp Asp Thr Ala Ile Tyr
 85 90 95

Tyr Cys Ala Lys His Tyr Tyr Tyr Gly Gly Ser Tyr Ala Met Asp Tyr
 100 105 110

Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Gly Gly Gly Gly Ser
 115 120 125

Gly Gly Gly Gly Ser Gly Gly Gly Ala Ser Asp Ile Gln Met Thr Gln
 130 135 140

Thr Thr Ser Ser Leu Ser Ala Ser Leu Gly Asp Arg Val Thr Ile Ser
 145 150 155 160

Cys Arg Ala Ser Gln Asp Ile Ser Lys Tyr Leu Asn Trp Tyr Gln Gln
 165 170 175

Lys Pro Asp Gly Thr Val Lys Leu Leu Ile Tyr His Thr Ser Arg Leu
 180 185 190

His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
 195 200 205

Tyr Ser Leu Thr Ile Ser Asn Leu Glu Gln Glu Asp Ile Ala Thr Tyr
 210 215 220

Phe Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gly Gly Thr
 225 230 235 240

Lys Leu Glu Ile Lys Arg Thr Ala Ala Ala
 245 250

5 <210> 49
 <211> 327
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Secuencia de unión a lox1 Adhiron lox1A

<400> 49

ES 2 712 632 T3

```

gccatggctg ctacaggcgt gcgggctgtg cccggcaatg agaacagcct ggaaatcgag      60
gaactggcca gattcgccgt ggacgagcac aacaagaaag agaacgccct gctggaattc      120
gtgcgggtcg tgaaggccaa agagcagtgg agcgaggccg acaacgactg gcacaccatg      180
tactacctga ccctggaagc caaggacggc ggcaagaaga agctgtacga ggccaaagtg      240
tgggtcaagc tggacctgga aacctggcag cacttcaact tcaaagagct ccaggaattc      300
aagcccgtgg gcgacgctgc ggccgcg                                           327

```

5 <210> 50
 <211> 109
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> Secuencia de unión a Lox1 Adhiron lox1A

<400> 50

```

Ala Met Ala Ala Thr Gly Val Arg Ala Val Pro Gly Asn Glu Asn Ser
 1                               5                               10          15

Leu Glu Ile Glu Glu Leu Ala Arg Phe Ala Val Asp Glu His Asn Lys
                20                               25          30

Lys Glu Asn Ala Leu Leu Glu Phe Val Arg Val Val Lys Ala Lys Glu
                35                               40          45

Gln Trp Ser Glu Ala Asp Asn Asp Trp His Thr Met Tyr Tyr Leu Thr
                50                               55          60

Leu Glu Ala Lys Asp Gly Gly Lys Lys Lys Leu Tyr Glu Ala Lys Val
65                70                               75          80

Trp Val Lys Leu Asp Leu Glu Thr Trp Gln His Phe Asn Phe Lys Glu
                85                               90          95

Leu Gln Glu Phe Lys Pro Val Gly Asp Ala Ala Ala Ala
                100                              105

```

15 <210> 51
 <211> 333
 <212> ADN
 20 <213> Secuencia artificial

<220>
 <223> Secuencia de unión a lox1 Adhiron lox1B

25 <400> 51

ES 2 712 632 T3

```

gccatggctg ctacagggcgt gcgggctgtg cccggcaatg agaacagcct ggaaatcgag      60
gaactggcca gattcgccgt ggacgagcac aacaagaaag agaacgccct gctggaattc      120
gtgcgggctcg tgaaggccaa agagcaggaa cagcccatcg gcgagcacc cgtgaacgac      180
accatgtact acctgaccct ggaagccaag gacggcggca agaagaagct gtacgaggcc      240
aaagtgtggg tcaagcgggtg gctgcggttc accgagatct acaacttcaa agagctccag      300
gaattcaagc ccgtgggcga cgctgcggcc gcg      333

```

5 <210> 52
 <211> 111
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> Secuencia de unión a lox1 Adhiron lox1B

<400> 52

```

Ala Met Ala Ala Thr Gly Val Arg Ala Val Pro Gly Asn Glu Asn Ser
 1                               5                               10          15

Leu Glu Ile Glu Glu Leu Ala Arg Phe Ala Val Asp Glu His Asn Lys
                20                               25          30

Lys Glu Asn Ala Leu Leu Glu Phe Val Arg Val Val Lys Ala Lys Glu
                35                               40          45

Gln Glu Gln Pro Ile Gly Glu His Pro Val Asn Asp Thr Met Tyr Tyr
 50                               55          60

Leu Thr Leu Glu Ala Lys Asp Gly Gly Lys Lys Lys Leu Tyr Glu Ala
 65                               70          75          80

Lys Val Trp Val Lys Arg Trp Leu Arg Phe Thr Glu Ile Tyr Asn Phe
                85                               90          95

Lys Glu Leu Gln Glu Phe Lys Pro Val Gly Asp Ala Ala Ala Ala
 100                               105          110

```

15 <210> 53
 <211> 54
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Cebador Adhiron_mut1

25 <220>
 <221> misc_feature
 <222> (23)..(24)
 <223> n es a, c, g, o t

30 <220>
 <221> misc_feature
 <222> (26)..(27)
 <223> n es a, c, g, o t

35 <220>
 <221> misc_feature
 <222> (29)..(30)

ES 2 712 632 T3

<223> n es a, c, g, o t

<220>
<221> misc_feature
5 <222> (32)..(33)
<223> n es a, c, g, o t

<400> 53
10 cagggtcagg tagtacctg tsnnsnnsnn snnctgctct ttggcctca cgac 54

<210> 54
<211> 48
<212> ADN
<213> Secuencia artificial

15 <220>
<223> Cebador Adhiron_mut2

<220>
20 <221> misc_feature
<222> (20)..(21)
<223> n es a, c, g, o t

<220>
25 <221> misc_feature
<222> (23)..(24)
<223> n es a, c, g, o t

<220>
30 <221> misc_feature
<222> (26)..(27)
<223> n es a, c, g, o t

<220>
35 <221> misc_feature
<222> (29)..(30)
<223> n es a, c, g, o t

<400> 54
40 ctggagctct tgaagtsn nsnnsnnsnn ctgaccac acttggc 48

<210> 55
<211> 54
<212> ADN
45 <213> Secuencia artificial

<220>
<223> Cebador mutagénico para la construcción de biblioteca de aglutinantes en el bucle 1 (adhiron mut 1)

50 <220>
<221> misc_feature
<222> (22)..(23)
<223> n es a, c, g, o t

<220>
55 <221> misc_feature
<222> (25)..(26)
<223> n es a, c, g, o t

<220>
60 <221> misc_feature
<222> (28)..(29)
<223> n es a, c, g, o t

<220>
65 <221> misc_feature

ES 2 712 632 T3

<222> (31)..(32)
 <223> n es a, c, g, o t

<400> 55
 5 gtcgtgaagg ccaaagagca gnnsnnsnns nnsacatgt actacctgac cctg 54

<210> 56
 <211> 18
 <212> PRT
 10 <213> Secuencia artificial

<220>
 <223> Secuencia de variantes de la biblioteca del bucle Adhiron mut1

15 <220>
 <221> misc_feature
 <222> (8)..(11)
 <223> Xaa puede ser cualquier aminoácido de origen natural

20 <400> 56

Val	Val	Lys	Ala	Lys	Glu	Gln	Xaa	Xaa	Xaa	Xaa	Thr	Met	Tyr	Tyr	Leu
1				5					10					15	

Thr Leu

<210> 57
 <211> 48
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Cebador mutagénico para la construcción de la biblioteca de aglutinantes en el bucle Adhiron mut2

<220>
 <221> misc_feature
 <222> (19)..(20)
 35 <223> n es a, c, g, o t

<220>
 <221> misc_feature
 <222> (22)..(23)
 40 <223> n es a, c, g, o t

<220>
 <221> misc_feature
 <222> (25)..(26)
 45 <223> n es a, c, g, o t

<220>
 <221> misc_feature
 <222> (28)..(29)
 50 <223> n es a, c, g, o t

<400> 57
 gccaaagtgt gggtaaggn snnsnnsnns aactcaaag agctccag 48

55 <210> 58
 <211> 16
 <212> PRT
 <213> Secuencia artificial

60 <220>
 <223> Secuencia de variantes de la biblioteca del bucle Adhiron mut 2

ES 2 712 632 T3

<220>
 <221> misc_feature
 <222> (7)..(10)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 5
 <400> 58
 Ala Lys Val Trp Val Lys Xaa Xaa Xaa Xaa Asn Phe Lys Glu Leu Gln
 1 5 10 15
 10
 <210> 59
 <211> 105
 <212> ADN
 <213> Secuencia artificial
 15
 <220>
 <223> Secuencia Knottin MCoTI-II de unión a tripsina
 <400> 59
 gccatggccg gtgtgtgccc caagatcttg aaaaagtgcc gccgtgacag cgattgtccc 60
 20
 ggcgcctgca tctgccgcg caatggctat tgcggagcgg ccgca 105
 <210> 60
 <211> 32
 <212> PRT
 25
 <213> Secuencia artificial
 <220>
 <223> Secuencia Knottin MCo-TI-II de unión a tripsina
 30
 <400> 60
 Ala Met Ala Gly Val Cys Pro Lys Ile Leu Lys Lys Cys Arg Arg Asp
 1 5 10 15
 Ser Asp Cys Pro Gly Ala Cys Ile Cys Arg Gly Asn Gly Tyr Cys Gly
 20 25 30
 35
 <210> 61
 <211> 99
 <212> ADN
 <213> Secuencia artificial
 40
 <220>
 <223> Cebador para la creación de una biblioteca de mutantes Knottin
 <220>
 <221> misc_feature
 <222> (11)..(11)
 45
 <223> n es a, c, g, o t
 <220>
 <221> misc_feature
 <222> (14)..(14)
 50
 <223> n es a, c, g, o t
 <220>
 <221> misc_feature
 <222> (17)..(17)
 55
 <223> n es a, c, g, o t
 <220>
 <221> misc_feature

<222> (20)..(20)
 <223> n es a, c, g, o t

5
 <220>
 <221> misc_feature
 <222> (23)..(23)
 <223> n es a, c, g, o t

10
 <220>
 <221> misc_feature
 <222> (26)..(26)
 <223> n es a, c, g, o t

15
 <220>
 <221> misc_feature
 <222> (29)..(29)
 <223> n es a, c, g, o t

20
 <220>
 <221> misc_feature
 <222> (32)..(32)
 <223> n es a, c, g, o t

25
 <220>
 <221> misc_feature
 <222> (35)..(35)
 <223> n es a, c, g, o t

30
 <220>
 <221> misc_feature
 <222> (38)..(38)
 <223> n es a, c, g, o t

35
 <400> 61
 ggtgtgtgcv nsvnsvnsvn svnsvnsvns vnsvnsvns gccgccgtga cagcgattgt 60
 cccggcgccct gcatctgccg cggcaatggc tattgcgga 99

40
 <210> 62
 <211> 99
 <212> ADN
 <213> Secuencia artificial

45
 <220>
 <223> Secuencia variada en la biblioteca de mutantes knottin

50
 <220>
 <221> misc_feature
 <222> (62)..(62)
 <223> n es a, c, g, o t

55
 <220>
 <221> misc_feature
 <222> (65)..(65)
 <223> n es a, c, g, o t

60
 <220>
 <221> misc_feature
 <222> (68)..(68)
 <223> n es a, c, g, o t

60
 <220>
 <221> misc_feature
 <222> (71)..(71)
 <223> n es a, c, g, o t

ES 2 712 632 T3

<220>
 <221> misc_feature
 <222> (74)..(74)
 <223> n es a, c, g, o t
 5

<220>
 <221> misc_feature
 <222> (77)..(77)
 <223> n es a, c, g, o t
 10

<220>
 <221> misc_feature
 <222> (80)..(80)
 <223> n es a, c, g, o t
 15

<220>
 <221> misc_feature
 <222> (83)..(83)
 <223> n es a, c, g, o t
 20

<220>
 <221> misc_feature
 <222> (86)..(86)
 <223> n es a, c, g, o t
 25

<220>
 <221> misc_feature
 <222> (89)..(89)
 <223> n es a, c, g, o t
 30

<400> 62
 tccgcaatag ccattgccgc ggcagatgca ggcgccggga caatcgctgt cacggcggca 60
 snbsnbsnbs nbsnbsnbsn bsnbsnbsnb gcacacacc 99

<210> 63
 <211> 33
 <212> PRT
 <213> Secuencia artificial
 40

<220>
 <223> Secuencia Knottin variada en la biblioteca
 <220>
 <221> misc_feature
 <222> (4)..(13)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 <400> 63
 Gly Val Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Arg Arg
 1 5 10 15
 Asp Ser Asp Cys Pro Gly Ala Cys Ile Cys Arg Gly Asn Gly Tyr Cys
 20 25 30
 Gly
 50

<210> 64
 <211> 121
 <212> ADN
 <213> Secuencia artificial
 55

<220>

ES 2 712 632 T3

<223> Secuencia pD7-Sce1 nucleótidos 1 - 120

<400> 64

ggtaccgaat tctagggata acagggtaat atgcatcttc tgacctcttc tcttcctccc 60
acagggcatg gcaaaacctc tgagccagga agaaagcaca ctgattgaaa gagcaaccgc 120

5 t 121

<210> 65

<211> 18

<212> PRT

10 <213> Secuencia artificial

<220>

<223> Fragmento del gen de resistencia a blasticidina codificado por pD7-Sce1

15 <400> 65

Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala
1 5 10 15

Thr Ala

<210> 66

20 <211> 120

<212> ADN

<213> Secuencia artificial

<220>

25 <223> Fragmento de secuencia pD7-ObLiGaRe

<400> 66

ggtaccgaat tcttttctgt caccaatcct ggggccacta gggacactgt ggggtggagg 60
aatgcatct tctgacctct tctcttcctc ccacagggca tggcaaaacc tctgagccag 120

30

<210> 67

<211> 7

<212> PRT

<213> Secuencia artificial

35

<220>

<223> Fragmento del gen de resistencia a Blasticidina codificado por pD7-ObLiGaRe

<400> 67

40

Met Ala Lys Pro Leu Ser Gln
1 5

<210> 68

45 <211> 1061

<212> ADN

<213> Secuencia artificial

<220>

50 <223> Secuencia del brazo de homología izquierdo ROSA 26

<400> 68

ES 2 712 632 T3

gcgatcgcgc tgattggcct cttttcctcc cgccgtgtgt gaaaacacaa atggcgtggt 60
 ttggttggcg taaggcgcct gtcagttaac ggcagccgga gtgcgcagcc gccggcagcc 120
 tcgctctgcc cactgggtgg ggcgggaggt aggtgggggt aggcgagctg gacgtgcggg 180
 cgcggtcggc ctctggcggg gcgggggagg ggagggaggg tcagcgaag tagctcgcgc 240
 gcgagcggcc gccaccctc cccttcctct gggggagtgc ttttaccgc cgccggccgg 300
 gcctcgtcgt ctgattggct ctcggggccc agaaaactgg cccttgccat tggctcgtgt 360
 tcgtgcaagt tgagtccatc cgccggccag cggggcgggc gaggaggcgc tcccaggttc 420
 cggccctccc ctcgggcccc cgccgcagag tctggccgcg cgcccctgcg caactgtgca 480
 ggaagcgcgc gctgggggcg gggacgggca gtagggctga gcggctgcg ggcggtgca 540
 agcacgtttc cgacttgagt tgcctcaaga gggcgtgct gagccagacc tccatcgcgc 600
 actccgggga gtggaggaa ggagcgaggg ctcagttggg ctgttttga ggcaggaagc 660
 acttgctctc ccaaagtcgc tctgagttgt tatcagtaag ggagctgcag tggagtaggc 720
 ggggagaagg ccgcaccctt ctccggaggg gggaggggag tgttgcaata cttttctggg 780
 agttctctgc tgcctcctgg cttctgagga ccgccctggg cctgggagaa tcccttcccc 840
 ctctccctc gtgatctgca actccagtct ttctagaatg cattaaggga tctgtagggc 900
 gcagtagtcc agggtttcct tgatgatgtc atacttatcc tgtccctttt tttccacag 960
 ctgcggttg aggacaaact cttcgcggtc tttccagtgg ggatcgacgg tatcgtagag 1020
 tcgagccgc tctaggaatt cacgccgcca ccatgaccga g 1061

<210> 69
 <211> 703
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Secuencia del brazo de homología derecho ROSA 26

10

<400> 69

gtatacggga attgaacag tgtaaaattg gagggacaag acttcccaca gattttcggg 60
 tttgtcggga agttttttaa taggggcaaa taaggaaaat gggaggatag gtagtcatct 120
 ggggttttat gcagcaaac tacaggttat tattgcttgt gatccgcctc ggagtatttt 180
 ccatcgaggg agattaaaga catgctcacc cgagttttat actctcctgc ttgagatcct 240
 tactacagta tgaaattaca gtgtcgcgag ttagactatg taagcagaat ttaatcatt 300
 tttaaagagc ccagtacttc atatccattt ctcccgctcc ttctgcagcc ttatcaaaag 360
 gtattttaga aactcattt tagccccatt ttcatttatt atactggctt atccaacccc 420
 tagacagagc attggcattt tccctttcct gatcttagaa gtctgatgac tcatgaaacc 480
 agacagatta gttacataca ccacaaatcg aggtgtgagc tggggcctca aactgcagt 540
 tctttataa ctcccttagta cactttttgt tgatcctttg ccttgatcct taattttcag 600
 tgtctatcac ctctcccgtc aggtggtggt ccacatttgg gcctattctc agtccagggg 660
 gttttacaac aatagatgta ttgagaatcc aacctcctgc agg 703

15

ES 2 712 632 T3

<210> 70
<211> 18
<212> ADN
<213> Secuencia artificial
5
<220>
<223> Secuencia de reconocimiento TALEN izquierda
<400> 70
10 tcccctccac cccacagt 18
<210> 71
<211> 15
<212> ADN
15 <213> Secuencia artificial
<220>
<223> Espaciador TALEN
20 <400> 71
ggggccacta gggac 15
<210> 72
<211> 18
25 <212> ADN
<213> Secuencia artificial
<220>
<223> Secuencia diana TALEN derecha
30 <400> 72
aggattggtg acagaaaa 18
<210> 73
<211> 18
35 <212> ADN
<213> Secuencia artificial
<220>
40 <223> Secuencia TALEN derecha
<400> 73
tttctgtca ccaatcct 18
45 <210> 74
<211> 20
<212> ADN
<213> Secuencia artificial
50 <220>
<223> Cebador AAVS-Left-arm-junction-PCR-Forw (9625)
<400> 74
55 ccggaactct gcctctaac 20
<210> 75
<211> 23
<212> ADN
60 <213> Secuencia artificial
<220>
<223> Cebador BSD_Junction PCR-rev (9626)
<400> 75
65 tagccacaga atagtctcg gag 23

ES 2 712 632 T3

<210> 76
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Cebador Donor_plasmid_seq_PDGFRTM-2 Forw
 <400> 76
 10 acacgcagga ggccatcgtg g 21
 <210> 77
 <211> 20
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Cebador AAVS1_right arm_junction_PCR_rev
 20 <400> 77
 tcctgggata ccccgaagag 20
 <210> 78
 <211> 54
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2598
 30 <400> 78
 tttttttaa ttaagattat tgactagtta ttaatagtaa tcaattacgg ggtc 54
 <210> 79
 <211> 52
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2619
 40 <400> 79
 tttttgttt aaacttacct tggatccctt gccggggctc aggctcaggg ac 52
 <210> 80
 <211> 33
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador M13Leadseq
 50 <400> 80
 55 aaattattat tcgcaattcc ttggtgtt cct 33
 <210> 81
 <211> 30
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> Cebador Notmycseq
 <400> 81
 65 ggccccattc agatcctctt ctgagatgag 30

ES 2 712 632 T3

<210> 82
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Cebador 2623
 <400> 82
 10 taaagtaggc ggtcttgaga cg 22
 <210> 83
 <211> 22
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Cebador 2624
 20 <400> 83
 gaaggtgctg tgaactgtt cc 22
 <210> 84
 <211> 19
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2595
 30 <400> 84
 gagggctctg gcagctagc 19
 <210> 85
 <211> 20
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2597
 40 <400> 85
 tcgagactgt gacgaggctg 20
 <210> 86
 <211> 19
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2625
 <400> 86
 50 ccttggtgct ggcactcga 19
 <210> 87
 <211> 40
 55 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> Cebador 1999
 <400> 87
 65 aaaaagcagg ctacatgag gcctggatc ttcttctcc 40

ES 2 712 632 T3

<210> 88
 <211> 44
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Secuencia de reconocimiento TALEN
 <400> 88
 10 tccaccggtc gccacatgg tgagcaaggg cgaggagctg ttca 44
 <210> 89
 <211> 33
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Cebador AAVS1_HA-L_Nested_Forw1
 20 <400> 89
 gtgcccttgc tgtgccgccg gaactctgcc ctc 33
 <210> 90
 <211> 23
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador EGFP_Synthetic_gene_Rev_Assembly
 30 <400> 90
 ttcacgtcgc cgtccagctc gac 23
 <210> 91
 <211> 21
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador Purotk_seq_fow2
 40 <400> 91
 tccataccga cgatctgcga c 21
 <210> 92
 <211> 27
 45 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador J44
 50 <400> 92
 aaaagcgct cccctaccg gtagaat 27
 55 <210> 93
 <211> 27
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> Cebador J46
 <400> 93
 65 ggcgacacgg aaatgtgaa tactcat 27

ES 2 712 632 T3

<210> 94
 <211> 27
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Cebador J48
 <400> 94
 10 cactacaccc agaagtcct gagcctg 27
 <210> 95
 <211> 20
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Cebador pSANG10pelB
 20 <400> 95
 cgctgcccag cggccatgg 20
 <210> 96
 <211> 26
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2097
 30 <400> 96
 gatggtgatg atgatgtcgc gatgcg 26
 <210> 97
 <211> 38
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 40 <223> Cebador 2423
 <400> 97
 tttttccat gggccggccc tcctcagtt tagttgag 38
 45 <210> 98
 <211> 39
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> Cebador 2437
 <400> 98
 55 tttttgcgg cgcggaagc cgtgatctcc ttctctc 39
 <210> 99
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> Cebador 2622
 <400> 99
 65 gaacaggaac acggaagtc 20

ES 2 712 632 T3

<210> 100
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Cebador 41679
 <400> 100
 10 atgagttgga gctgtatcat cc 22
 <210> 101
 <211> 18
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Cebador 2621
 20 <400> 101
 gcattccacg gcggccgc 18
 <210> 102
 <211> 9
 25 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Péptido SLLMWITQV (NY-ESO-1 157-165) reconocido por c12c2
 30 <400> 102
 Ser Leu Leu Met Trp Ile Thr Gln Val
 1 5
 35 <210> 103
 <211> 10
 <212> PRT
 <213> Secuencia artificial
 40 <220>
 <223> Péptido MART-1 26-35 reconocido por MEL5 TCR
 <400> 103
 Glu Leu Ala Gly Ile Gly Ile Leu Thr Val
 45 1 5 10
 <210> 104
 <211> 30
 <212> ADN
 50 <213> Secuencia artificial
 <220>
 <223> Cebador 4JFH-Valpha-F
 55 <400> 104
 acacacgcta gccagaaaga ggtggaacag 30
 <210> 105
 <211> 50
 60 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2770

ES 2 712 632 T3

<220>
 <221> misc_feature
 <222> (21)..(22)
 <223> n es a, c, g, o t
 5

<220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n es a, c, g, o t
 10

<400> 105
 caaagaacag ctcgccggtg nccccgassn nggagctggc gcaaaagtac 50

<210> 106
 <211> 53
 <212> ADN
 <213> Secuencia artificial
 15

<220>
 <223> Cebador 2771
 20

<220>
 <221> misc_feature
 <222> (23)..(23)
 <223> n es a, c, g, o t
 25

<220>
 <221> misc_feature
 <222> (30)..(31)
 <223> n es a, c, g, o t
 30

<220>
 <221> misc_feature
 <222> (33)..(34)
 <223> n es a, c, g, o t
 35

<400> 106
 ctcgccgcaa ggtgggaatg tangwtccsn nsnnaagtgg gcgcacggcg cac 53

<210> 107
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 40

<220>
 <223> Cebador 2780
 45

<400> 107
 ctggcagcta gcaagcagga ag 22
 50

<210> 108
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 55

<220>
 <223> Cebador 2781
 60

<400> 108
 tacattccca cctcgggag ag 22
 60

<210> 109
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 65

<220>
 <223> Cebador 2782

 <400> 109
 5 tttttgcgg ccgcgacag gttctg 26

 <210> 110
 <211> 27
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> Cebador 2783

 <400> 110
 15 cgtaagctgg tacctatta tctaggg 27

 <210> 111
 <211> 27
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador 2785

 <400> 111
 25 ccctagataa taaggtagca gcttagc 27

 <210> 112
 <211> 19
 30 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador 2787

 <400> 112
 accggcgagc tgttcttg 19

 <210> 113
 <211> 26
 40 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Cebador 2788

 <400> 113
 50 agtgacaagc tttattatc tgggtg 26

 <210> 114
 <211> 19
 <212> ADN
 55 <213> Secuencia artificial

 <220>
 <223> Cebador 2789

 <400> 114
 60 caggtcctcg agcactgc 19

 <210> 115
 <211> 25
 <212> ADN
 65 <213> Secuencia artificial

ES 2 712 632 T3

```

<220>
<223> Cadena de inserto de ADN bicatenario CRISPR1

<400> 115
5  ggggccacta gggacaggat gtttt    25

<210> 116
<211> 26
<212> ADN
10 <213> Secuencia artificial

<220>
<223> Cadena de inserto de ADN bicatenario CRISPR1

15 <400> 116
    catcctgtcc ctagtggcc ccggtg    26

<210> 117
<211> 25
20 <212> ADN
    <213> Secuencia artificial

<220>
<223> Cadena de inserto de ADN bicatenario CRISPR2

25 <400> 117
    gtcaccaatc ctgtccctag gtttt    25

<210> 118
<211> 25
30 <212> ADN
    <213> Secuencia artificial

<220>
<223> Cadena de inserto de ADN bicatenario CRISPR2

35 <400> 118
    ctagggacag gattggtgac cgggtg    25

<210> 119
<211> 102
40 <212> ARN
    <213> Secuencia artificial

45 <220>
<223> ARN CRISPR 1

<400> 119
    gggggggccac uagggacagg auguuuuaga gcuagaaaua gcaaguuaaa auaaggcuag    60
50    uccguuauca acuugaaaaa guggcaccga gucggugcuu uu    102

<210> 120
<211> 102
55 <212> ARN
    <213> Secuencia artificial

<220>
<223> ARN CRISPR 2

60 <400> 120
    gggucaccaa uccugucccu agguuuuaga gcuagaaaua gcaaguuaaa auaaggcuag    60
    uccguuauca acuugaaaaa guggcaccga gucggugcuu uu    102

```

ES 2 712 632 T3

5 <210> 121
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2723

10 <400> 121
cgcgccagaa gtctcaccaa gccca 25

15 <210> 122
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2724

20 <400> 122
cgcgtgggct tggtagact tctgg 25

25 <210> 123
<211> 60
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2768

30 <400> 123
aattctccc tccacccac agtagggaca gtggggccag gattggtgac agaaaatgca 60

35 <210> 124
<211> 52
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2769

40 <400> 124
tttctgtca ccaatcctgg cccactgtc cctactgtgg ggtggagggg ag 52

45 <210> 125
<211> 28
<212> ADN
<213> Secuencia artificial

50 <220>
<223> Cebador 2778

55 <400> 125
aattctaggg ataacagggt aatatgca 28

60 <210> 126
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2779

65 <400> 126
tattaccctg ttatccctag 20

ES 2 712 632 T3

<210> 127
<211> 60
<212> ADN
<213> Secuencia artificial
5
<220>
<223> Cebador 2808
<400> 127
10 aattcttttc tgcaccaat cctggggcca ctagggacac tgtggggtgg aggggatgca 60
<210> 128
<211> 57
<212> ADN
15 <213> Secuencia artificial
<220>
<223> Cebador 2809
20 <400> 128
tccctccac cccacagtgt ccctagtggc cccaggattg gtgacagaaa agaattg 57
<210> 129
<211> 41
25 <212> ADN
<213> Secuencia artificial
<220>
<223> Cebador J60
30 <400> 129
acacacggta cgcgcatcgc gctgattggc ttctttcct c 41
<210> 130
<211> 35
35 <212> ADN
<213> Secuencia artificial
<220>
40 <223> Cebador 2706
<400> 130
tttttatgc attctagaaa gactggagtt gcaga 35
45 <210> 131
<211> 18
<212> ADN
<213> Secuencia artificial
50 <220>
<223> Cebador 2715
<400> 131
55 gagcgtccgc ccaccctc 18
<210> 132
<211> 18
<212> ADN
60 <213> Secuencia artificial
<220>
<223> Cebador 2716
<400> 132
65 gaggtgggc ggacgctc 18

ES 2 712 632 T3

<210> 133
 <211> 33
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Cebador 2709
 <400> 133
 10 tttttatgc attaagggat ctgtagggcg cag 33
 <210> 134
 <211> 21
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Cebador 2710
 20 <400> 134
 gtgaattcct agagcggcct c 21
 <210> 135
 <211> 50
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador 2745
 30 <400> 135
 gaggcgcctc taggaattca cgccgccacc atgaccgagt acaagcccac 50
 <210> 136
 <211> 48
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador J59
 40 <400> 136
 aaaaaaagat ctgtgtgtt cgaatcaggc accgggcttg cgggtcat 48
 45 <210> 137
 <211> 36
 <212> ADN
 <213> Secuencia artificial
 50 <220>
 <223> Cebador J61
 <400> 137
 55 tttttgtat acggaattg aacagggtga aaattg 36
 <210> 138
 <211> 43
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> Cebador J62
 <400> 138
 65 tttttcctg caggagggtg gattctcaat acatctattg ttg 43

ES 2 712 632 T3

5
<210> 139
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2701

10 <400> 139
gccgacgtct cgtcgctgat gtttt 25

15 <210> 140
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2702

20 <400> 140
atcagcgacg agacgtcggc cgggtg 25

25 <210> 141
<211> 25
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador 2703

30 <400> 141
cgcccatctt ctagaagac gtttt 25

35 <210> 142
<211> 25
<212> ADN
<213> Secuencia artificial

40 <220>
<223> Cebador 2704

<400> 142
gtcttctag aagatgggcg cgggtg 25

REIVINDICACIONES

1. Un método para producir una biblioteca de clones de células eucariotas que contienen ADN que codifica un repertorio diverso de aglutinantes que reconocen dianas, que comprende
 5 proporcionar moléculas de ADN donante que codifican los aglutinantes, y células eucariotas, en donde los aglutinantes son anticuerpos, proteínas o péptidos,
 introducir el ADN donante en las células y proporcionar una nucleasa de sitio específico dentro de las células, en donde la nucleasa escinde una secuencia de reconocimiento en el ADN celular para crear un sitio de integración en el que el ADN donante se integra en el ADN celular, teniendo lugar la integración a través de mecanismos de
 10 reparación del ADN endógenos a las células, creando de este modo células recombinantes que contienen ADN donante integrado en el ADN celular, y
 cultivar las células recombinantes para producir clones,
 proporcionar de este modo una biblioteca de clones de células eucariotas que contienen ADN donante que codifica el repertorio de aglutinantes.
 15
2. Un método de acuerdo con la reivindicación 1, en el que los aglutinantes son moléculas de anticuerpo; receptores de linfocitos T; y/o en el que los aglutinantes son multiméricos, que comprenden al menos una primera y una segunda subunidades.
- 20 3. Un método de acuerdo con la reivindicación 2, en el que las moléculas de anticuerpo son inmunoglobulinas de longitud completa, IgG, Fab, scFv-Fc o scFv.
4. Un método para producir una biblioteca de clones de células eucariotas que contienen ADN que codifica un repertorio diverso de aglutinantes multiméricos que reconocen dianas, comprendiendo cada aglutinante una primera y una segunda subunidades, en donde el método comprende proporcionar células eucariotas que contienen ADN que codifica la primera subunidad y proporcionar moléculas de ADN donante que codifican la segunda subunidad aglutinante, en donde los aglutinantes son anticuerpos, proteínas o péptidos,
 25 introducir el ADN donante en las células y proporcionar una nucleasa de sitio específico dentro de las células, en donde la nucleasa escinde una secuencia de reconocimiento en el ADN celular para crear un sitio de integración en el que el ADN donante se integra en el ADN celular, teniendo lugar la integración a través de mecanismos de
 30 reparación del ADN endógenos a las células, creando de este modo células recombinantes que contienen ADN donante integrado en el ADN celular, y
 cultivar las células recombinantes para producir clones que contienen ADN que codifica la primera y la segunda subunidades del aglutinante multimérico,
 35 proporcionar de este modo una biblioteca de clones de células eucariotas que contienen ADN donante que codifica el repertorio de aglutinantes multiméricos.
5. Un método para producir una biblioteca de clones de células eucariotas que contienen ADN que codifica un repertorio diverso de aglutinantes multiméricos que reconocen dianas, comprendiendo cada aglutinante al menos una primera y una segunda subunidades, en donde los aglutinantes son anticuerpos, proteínas o péptidos, en donde el método comprende
 40 proporcionar las primeras moléculas de ADN donante que codifican la primera subunidad y proporcionar células eucariotas,
 introducir el primer ADN donante en las células y proporcionar una nucleasa de sitio específico dentro de las células,
 45 en donde la nucleasa escinde una secuencia de reconocimiento en el ADN celular para crear un sitio de integración en el que el ADN donante se integra en el ADN celular, teniendo lugar la integración a través de mecanismos de
 reparación del ADN endógenos a las células, creando de este modo un primer conjunto de células recombinantes que contienen un primer ADN donante integrado en el ADN celular,
 50 cultivar el primer conjunto de células recombinantes para producir un primer conjunto de clones que contienen ADN que codifica la primera subunidad,
 introducir las moléculas del segundo ADN donante que codifican la segunda subunidad en las células del primer conjunto de clones, en donde el segundo ADN donante se integra en el ADN celular del primer conjunto de clones,
 creando de este modo un segundo conjunto de células recombinantes que contienen el primer y el segundo ADN donantes integrados en el ADN celular, y
 55 cultivar el segundo conjunto de células recombinantes para producir un segundo conjunto de clones, conteniendo estos clones ADN que codifica la primera y la segunda subunidades del aglutinante multimérico,
 proporcionar de este modo una biblioteca de clones de células eucariotas que contienen ADN donante que codifica el repertorio de aglutinantes multiméricos.
- 60 6. Un método de acuerdo con la reivindicación 5, en el que las moléculas del segundo ADN donante se integran mediante un método que comprende proporcionar una nucleasa de sitio específico dentro de las células, en donde la nucleasa escinde una secuencia de reconocimiento en el ADN celular para crear un sitio de integración en el que el ADN donante se integra en el ADN celular, teniendo lugar la integración a través de mecanismos de reparación del ADN endógenos a las células.
 65

7. Un método de acuerdo con cualquiera de las reivindicaciones 3 a 6 en el que los aglutinantes multiméricos son moléculas de anticuerpo que comprenden un dominio variable de cadena pesada (VH) y un dominio variable de cadena ligera (VL) como subunidades separadas.
- 5 8. Un método de acuerdo con la reivindicación 7, en el que los aglutinantes multiméricos son inmunoglobulinas completas.
9. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que las células son células eucariotas superiores con un tamaño de genoma superior a 2×10^7 pares de bases.
- 10 10. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que las células son de mamífero.
11. Un método de acuerdo con la reivindicación 10, en el que las células son células HEK293, células de ovario de hámster chino (CHO), células de linaje de linfocitos T o células de linaje de linfocitos B o cualquiera de las líneas celulares enumeradas en la "Cancer Cell Line Encyclopedia" o el "COSMIC catalogue of somatic mutations in cancer".
- 15 12. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la secuencia de reconocimiento está en el ADN genómico de las células.
- 20 13. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la secuencia de reconocimiento para la nucleasa de sitio específico tiene lugar solo una o dos veces en el ADN celular.
- 25 14. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la nucleasa de sitio específico escinde el ADN celular para crear una rotura bicatenaria que sirve como un sitio de integración.
- 30 15. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la nucleasa es una meganucleasa, una nucleasa con dedos de cinc (ZFN), una nucleasa TALE, o en el que la escisión del ADN está dirigida por el sistema CRISPR/Cas.
- 35 16. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el ADN donante comprende un elemento genético para la selección de células en las que se integra el ADN donante.
17. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la integración del ADN donante en el ADN celular coloca la expresión del aglutinante y/o la expresión de un elemento de selección genética bajo el control de un promotor presente dentro del ADN celular.
- 40 18. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la biblioteca contiene al menos 100, 10^3 , 10^4 , 10^5 o 10^6 clones, derivándose cada clon de una célula recombinante individual producida por integración de ADN donante.
- 45 19. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la biblioteca codifica al menos 100, 10^3 , 10^4 , 10^5 o 10^6 aglutinantes diferentes.
- 50 20. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que cada clon contiene ADN donante integrado que codifica solo uno o dos miembros del repertorio de aglutinantes.
21. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que las células eucariotas son diploides y contienen una secuencia de reconocimiento para la nucleasa de sitio específico en loci fijados duplicados en el ADN celular.
22. Un método de acuerdo con cualquiera de las reivindicaciones 1 a 19, en el que cada clon contiene ADN donante integrado que codifica un solo miembro del repertorio de aglutinantes.
- 55 23. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que cada una de las moléculas de ADN donante codifican un único aglutinante o subunidad de aglutinante.
- 60 24. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que los aglutinantes se muestran en la superficie de la célula.
25. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que los aglutinantes se secretan de las células.
- 65 26. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que el ADN donante se obtiene de rondas de selección de presentación de fagos.

27. Un método de acuerdo con cualquiera de las reivindicaciones anteriores, que comprende además:

cultivar la biblioteca para expresar los aglutinantes,
recuperar uno o más clones que expresan un aglutinante de interés, y
5 generar una biblioteca de derivados a partir del uno o más clones recuperados, en donde la biblioteca de derivados contiene ADN que codifica un segundo repertorio de aglutinantes.

28. Un método de acuerdo con la reivindicación 27, en el que generar la biblioteca de derivados comprende aislar el ADN donante de uno o más clones recuperados, introducir la mutación en el ADN para proporcionar una población derivada de moléculas de ADN donante que codifican un segundo repertorio de aglutinantes, e introducir la población de derivados de moléculas de ADN donante en células para crear una biblioteca de derivados de células que contienen ADN que codifica el segundo repertorio de aglutinantes.

29. Un método de acuerdo con la reivindicación 27, en el que generar la biblioteca de derivados comprende introducir la mutación en el ADN donante en uno o más clones recuperados mediante la inducción de la mutación del ADN dentro de los clones.

30. Un método para producir un repertorio diverso de aglutinantes, que comprende producir una biblioteca mediante un método de acuerdo con cualquiera de las reivindicaciones 1 a 29 y cultivar las células de la biblioteca para expresar los aglutinantes.

31. Un método de cribado de un aglutinante que reconoce una diana, que comprende:

producir una biblioteca por el método de cualquiera de las reivindicaciones 1 a 29,
cultivar células de la biblioteca para expresar los aglutinantes,
exponer los aglutinantes a la diana, permitiendo el reconocimiento de la diana por uno o más aglutinantes afines, si están presentes, y
detectar si la diana es reconocida por un aglutinante afín.

32. Un método de acuerdo con la reivindicación 31, en el que los aglutinantes son moléculas de anticuerpo y la diana es un antígeno, o en el que los aglutinantes son TCR y la diana es un complejo MHC:péptido.

33. Un método de acuerdo con la reivindicación 31 o la reivindicación 32, que comprende además detectar el reconocimiento de dianas por un aglutinante afín y recuperar células de un clon que contiene ADN que codifica el aglutinante afín.

34. Un método de acuerdo con la reivindicación 33, que comprende además aislar el ADN que codifica el aglutinante del clon recuperado, obteniendo de este modo el ADN que codifica un aglutinante que reconoce la diana.

35. Un método de acuerdo con la reivindicación 33 o la reivindicación 34, que comprende introducir una mutación o convertir el ADN en un ADN modificado que codifica un aglutinante reestructurado.

36. Un método de acuerdo con la reivindicación 35, en el que el aglutinante es un scFv y el método comprende convertir el ADN que codifica el scFv en el ADN que codifica una Ig o un fragmento de la misma mientras se mantienen los pares de cadenas VH y VL variables originales.

37. Un método de acuerdo con las reivindicaciones 34, 35 o 36, que comprende además introducir el ADN en una célula hospedadora y opcionalmente que comprende además cultivar las células y concentrar las células para proporcionar un sedimento celular o una suspensión celular concentrada.

38. Un método de acuerdo con las reivindicaciones 33 o 37, que comprende además cultivar las células para expresar el aglutinante y purificar el aglutinante.

39. Uso de una nucleasa de sitio específico para la escisión dirigida de ADN celular en la construcción de una biblioteca de células eucariotas que contienen ADN que codifica un repertorio de aglutinantes, mediante un método de acuerdo con una cualquiera de las reivindicaciones 1 a 38, en donde la escisión del ADN mediada por nucleasas mejora la integración de sitio específico de los genes de aglutinante a través de mecanismos de reparación de ADN celular endógenos.

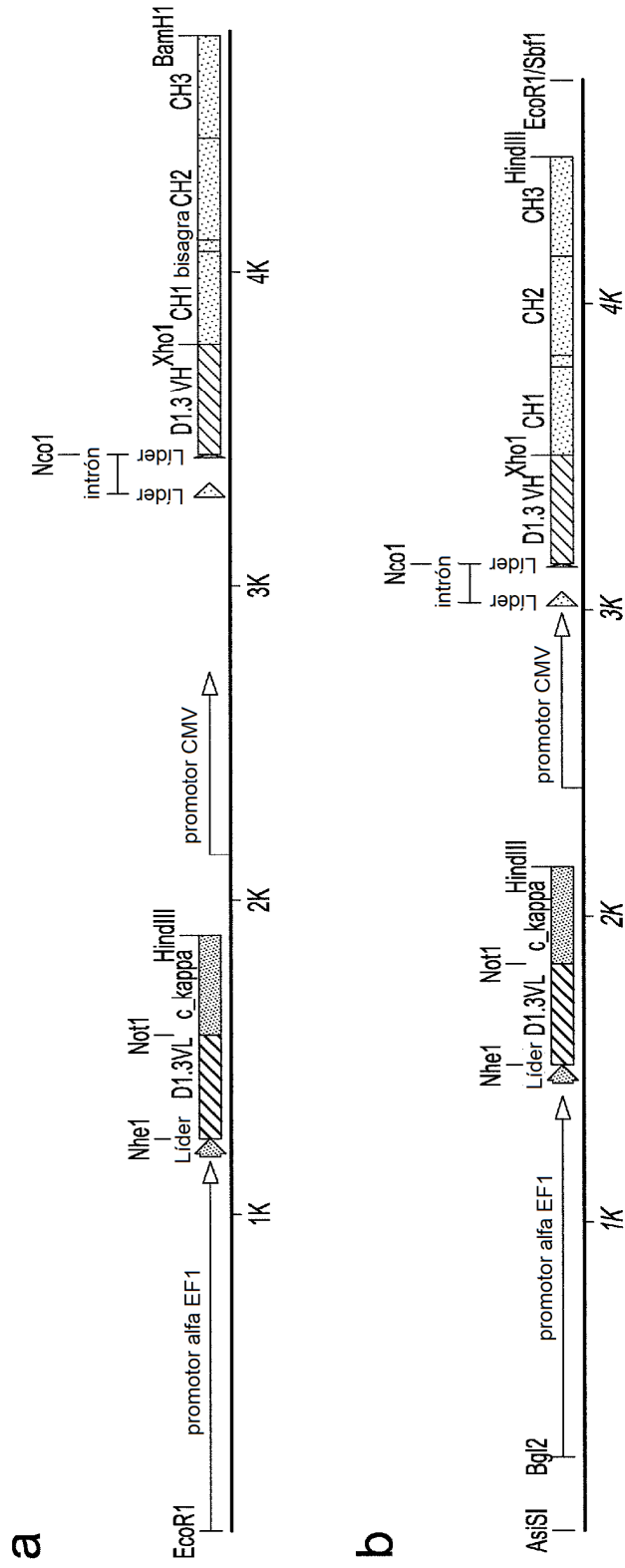


Fig. 1

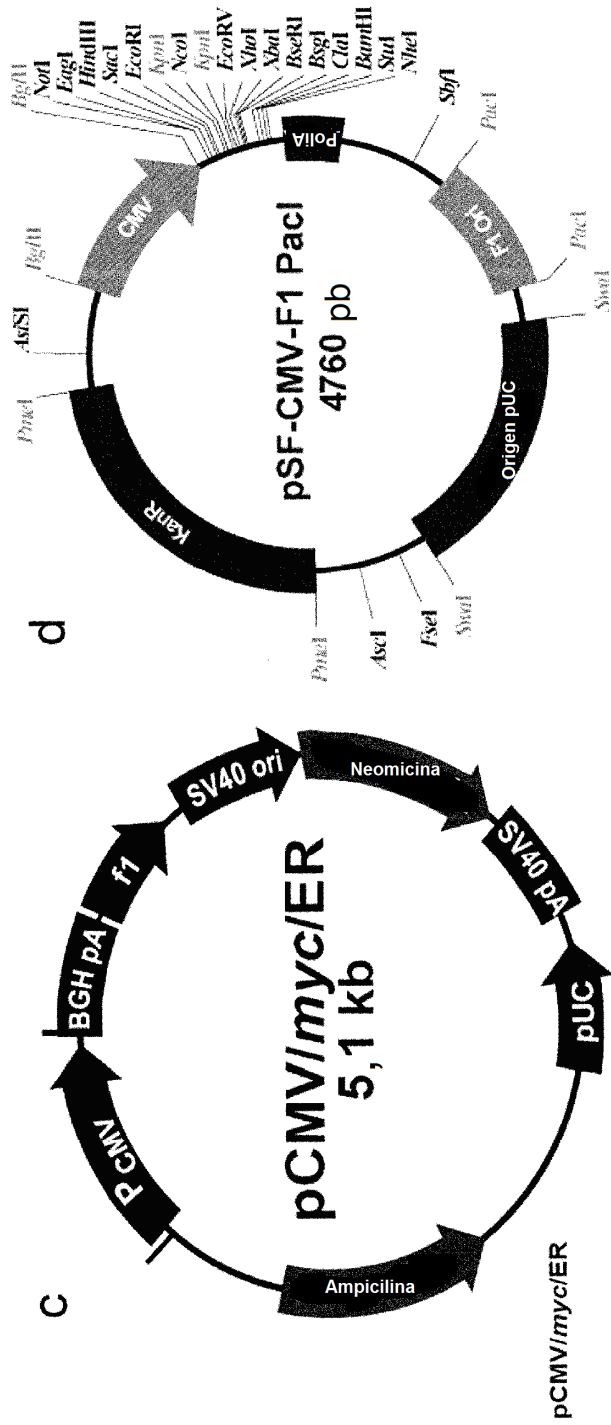


Fig. 1
(continuación)

ES 2 712 632 T3

>"promotor pEF
"Acc65I" >"EcoR1" |
| 10 | 20 | 30 | 40 | 50 | 60
GGT ACC GAA TTC CGT GAG GCT CCG GIG CCC GTC AGT GGG CAG AGC GCA CAT CGC CCA CAG
TCC CCG AGA AGT TGG GGG GAG GGG TCG GCA ATT GAA CCG GTG CCT AGA GAA GGT GGC GCG
GGG TAA ACT GGG AAA GTG ATG TCG TGT ACT GGC TCC GCC TTT TTC CCG AGG GTG GGG GAG
AAC CGT ATA TAA GTG CAG TAG TCG CCG TGA ACC TTC TTT TTC GCA ACG GGT TTG CCG CCA
GAA CAC AGG TAA GTG CCG TGT GTG GTT CCC GCG GGC CTG GCC TCT TTA CCG GTT ATG GCC
CTT GCG TGC CTT GAA TTA CTT CCA CCT GGC TCC AGT ACG TGA TTC TTG ATC CCG AGC TGG
AGC CAG GGG CCG GCC TTG CCG TTT AGG AGC CCC TTC GCC TCG TGC TTG AGT TGA GGC CTG
GCC TGG GCG CTG GGG CCG CCG CGT CCG AAT CTG GTG GCA CCT TCG CCG CTG TCT CGC TGC
TTT CGA TAA GTC TCT AGC CAT TTA AAA TTT TTG ATG ACC TGC TGC GAC GCT TTT TTT CTG
GCA AGA TAG TCT TGT AAA TGC GGG CCA GGA TCT GCA CAC TGG TAT TTC GGT TTT TGG GCC
CGC GGC CCG CGA CCG GGC CCG TGC GTC CCA GCG CAC ATG TTC GGC GAG GCG GGG OCT GCG
AGC GCG GCC ACC GAG AAT CCG ACG GGG GTA GTC TCA AGC TGG CCG GCC TGC TCT GGT GCC
TGG CGT CGC GCC GCC GCG TAT CCG CCC GCC CTG GGC GGC AAG GCT GGC CCG GTC GCC ACC
AGT TGC GTG AGC GGA AAG ATG GCC GCT TCC CCG CCC TGC TCC AGG GGG CTC AAA ATG GAG
GAC GCG GCG CTC GGG AGA GCG GGC GGG TGA GTC ACC CAC ACA AAG GAA AAG GGC CTT TCC
GTC CTC AGC CGT CCG TTC ATG TGA CTC CAC GGA GTA CCG GGC GCC GTC CAG GCA CCT CGA
TTA GTT CTG GAG CTT TTG GAG TAC GTC GTC TTT AGG TTG GGG GGA GGG GTT TTA TGC GAT
GGA GTT TCC CCA CAC TGA GTG GGT GGA SAC TGA AGT TAG GCC AGC TTG GCA CTT GAT GTA
ATT CTC GTT GGA ATT TGC CCT TTT TGA GTT TGG ATC TTG GTT CAT TCT CAA GCC TCA GAC
AGT GGT TCA AAG TTT TTT TCT TCC ATT TCA GGT GTC GTG AGA CGT GGCCACC ATG AGG GCC
M R A>
_ _ a _ a _ a _ >
 >"Nhe1" |
1210 1220 1230 1240 1250 1260
TGG ATC TTC TTT CTC CTT TGC CTG GCC GGG AGG GCT CTG GCA GCT AGC GAC ATC CAG ATG
W I F F L L C L A G R A L A A S> D I Q M>
_ _ a _ a _ a _ a _ a _ " lider BM40 " _ _ a _ a _ a _ a _ a _ a _ < _ "D1.3 VL" _ _ >
1270 1280 1290 1300 1310 1320
ACC CAG AGC CCA AGC AGC CTG AGC GCC AGC GTG GGT GAC AGA GTG ACC ATC ACC TGT AGA
T Q S P S S L S A S V G D R V I I T C R>
_ _ b _ b _ b _ b _ b _ b _ b _ "D1.3 VL" _ _ b _ b _ b _ b _ b _ b _ b _ b _ >
1330 1340 1350 1360 1370 1380
GCC AGC GGT AAC ATC CAC AAC TAC CTG GCT TGG TAC CAG CAG AAG CCA GGT AAG GCT CCA
A S G N I H N Y L A W Y Q Q K P G K A P>
_ _ b _ b _ b _ b _ b _ b _ b _ "D1.3 VL" _ _ b _ b _ b _ b _ b _ b _ b _ b _ >
1390 1400 1410 1420 1430 1440
AAG CTG CTG ATC TAC TAC ACC ACC ACC CTG GCT GAC GGT GTG CCA AGC AGA TTC AGC GGT
K L L I Y Y T T T L A D G V P S R F S G>
_ _ b _ b _ b _ b _ b _ b _ b _ "D1.3 VL" _ _ b _ b _ b _ b _ b _ b _ b _ b _ >
1450 1460 1470 1480 1490 1500
AGC GGT AGC GGT ACC GAC TAC ACC TTC ACC ATC AGC AGC CTC CAG CCA GAG GAC ATC GCC
S G S G T D Y T F T I S S L Q P E D I A>
_ _ b _ b _ b _ b _ b _ b _ b _ "D1.3 VL" _ _ b _ b _ b _ b _ b _ b _ b _ b _ >
1510 1520 1530 1540 1550 1560
ACC TAC TAC TGC CAG CAC TTC TGG AGC ACC CCA AGG ACG TTC GGC CAA GGG ACC AAG GTG
T Y Y C Q H F W S T P R T F G Q G T K V>
_ _ b _ b _ b _ b _ b _ b _ b _ "D1.3 VL" _ _ b _ b _ b _ b _ b _ b _ b _ b _ >

Fig. 2

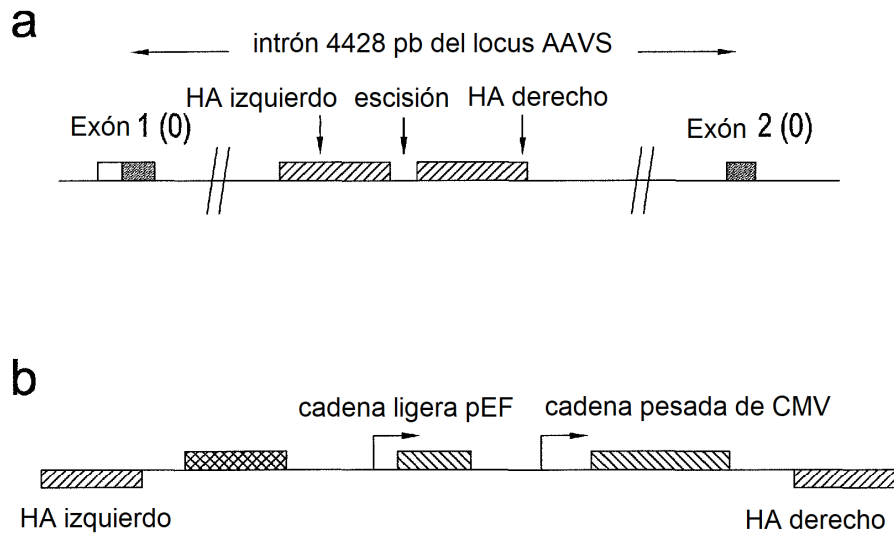


Fig. 3

ES 2 712 632 T3

← CASETE DE EXPRESIÓN DE ANTICUERPO D 1.3 →

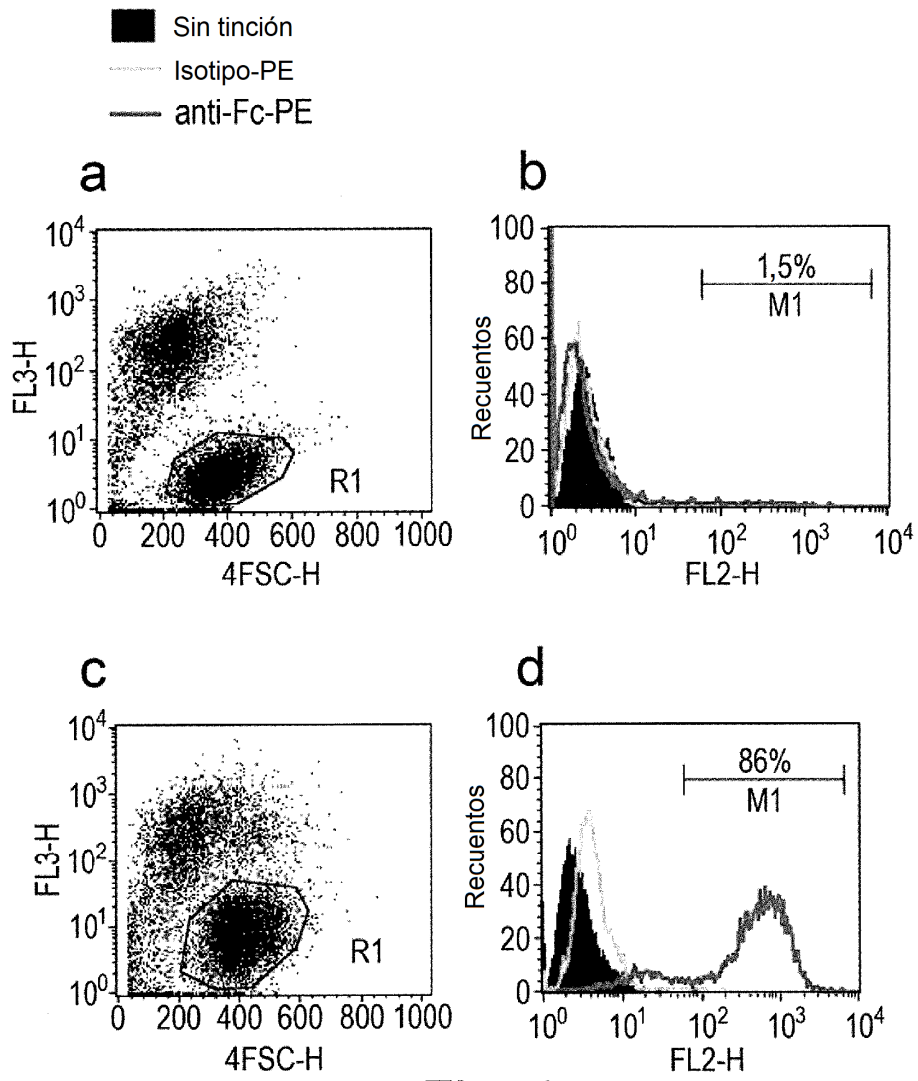
6887 sitio loxP BstZ171 AscI AAVS1_right_HA

TGGCCCGGGCAIG ATAACTTCGTATAATGTATGCTATACGAAGTTATG TATAC GGCGCGCCC ACTAGGGACAGGATTGGTGACA

6960 6970 6980 6990 7000 7010
GAA AAG CCC CAT CCT TAG GCC TCC TCC TTC CTA GTC TCC TGA TAT TGG GTC TAA CCC CCA
CCT CCT GPT AGG CAG ATT CCT TAT CTG GTG ACA CAC CCC CAT TTC CTG GAG CCA TCT CTC
TCC TTG CCA GAA CCT CTA AGG TTT GCT TAC GAT GGA GCC AGA GAG GAT CCT GGG AGG GAG
AGC TTG GCA GGG GGT GGG AGG GAA GGG GGG GAT GCG TGA CCT GCC CGG TTC TCA GTG GCC
ACC CTG CGC TAC CCT CTC CCA GAA CCT GAG CTG CTC TGA CGC GGC TGT CTG GTG CGT TTC
ACT GAT CCT GGT GCT GCA GCT TCC TTA CAC TTC CCA AGA GGA GAA GCA GTT TGG AAA AAC
AAA ATC AGA ATA AGT TGG TCC TGA GTT CTA ACT TTG GCT CTT CAC CTT TCT AGT CCC CAA
TTT ATA TTG TTC CTC CGT GCG TCA GTT TTA CCT GTG AGA TAA GGC CAG TAG CCA GCC CCG
TCC TGG CAG GGC TGT GGT GAG GAG GGG GGT GTC CGT GTG GAA AAC TCC CTT TGT GAG AAT
GGT GCG TCC TAG GTG TTC ACC AGG TCG TGG CCG CCT CTA CTC CCT TTC TCT TTC TCC ATC
CTT CTT TCC TTA AAG AGT CCC CAG TGC TAT CTG GGA CAT ATT CCT CCG CCC AGA GCA GGG
TCC CGC TTC CCT AAG GCC CTG CTC TGG GCT TCT GGG TTT GAG TCC TTG GCA AGC CCA GGA
GAG GCG CTC AGG CTT CCC TGT CCC CCT TCC TCG TCC ACC ATC TCA TGC CCC TGG CTC TCC

7740 7750 AAVS1 terminal Brazo de homología derecho **MluI** **BstZ1** 7790
TGC CCC TTC CCT ACA GGG GTT CCT GGC TCT GCT CTG AGG CGT GTATAC TCG ATC TTT CCG

Fig. 3 (continuación)



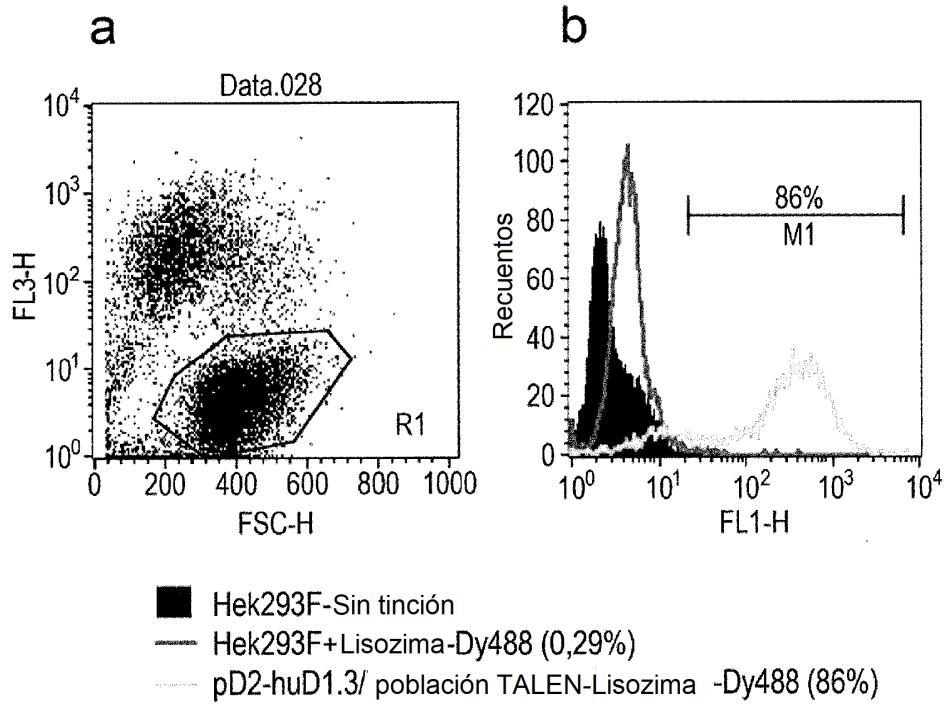


Fig. 5

ADN pD2/
10⁶ células

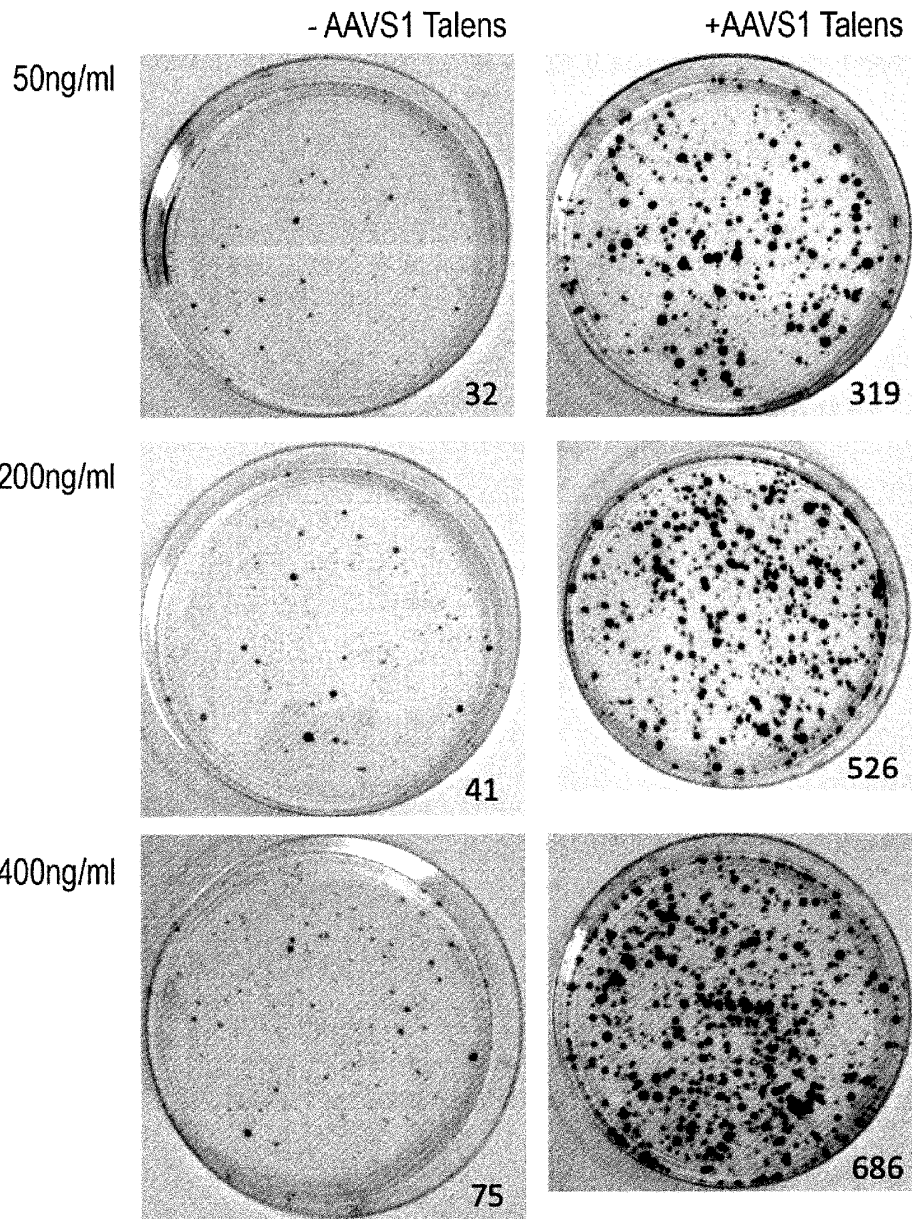
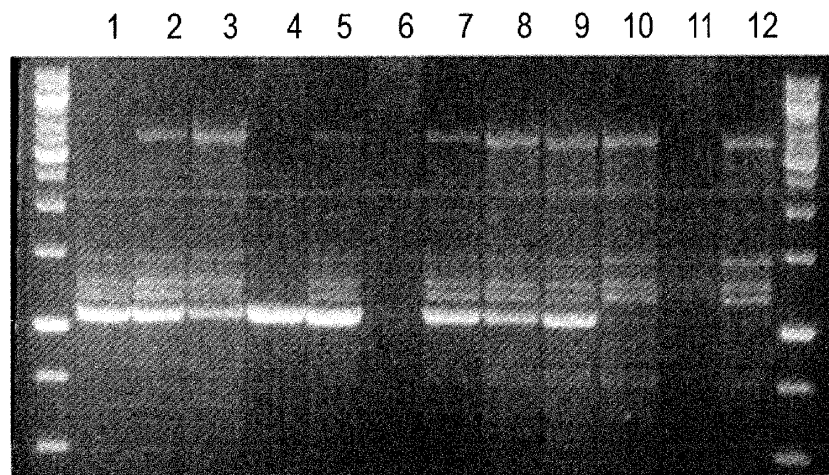


Fig. 6

a



b

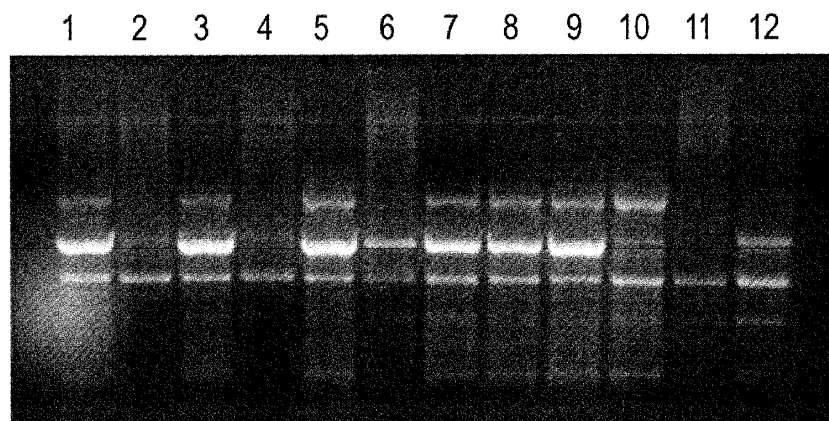


Fig. 7

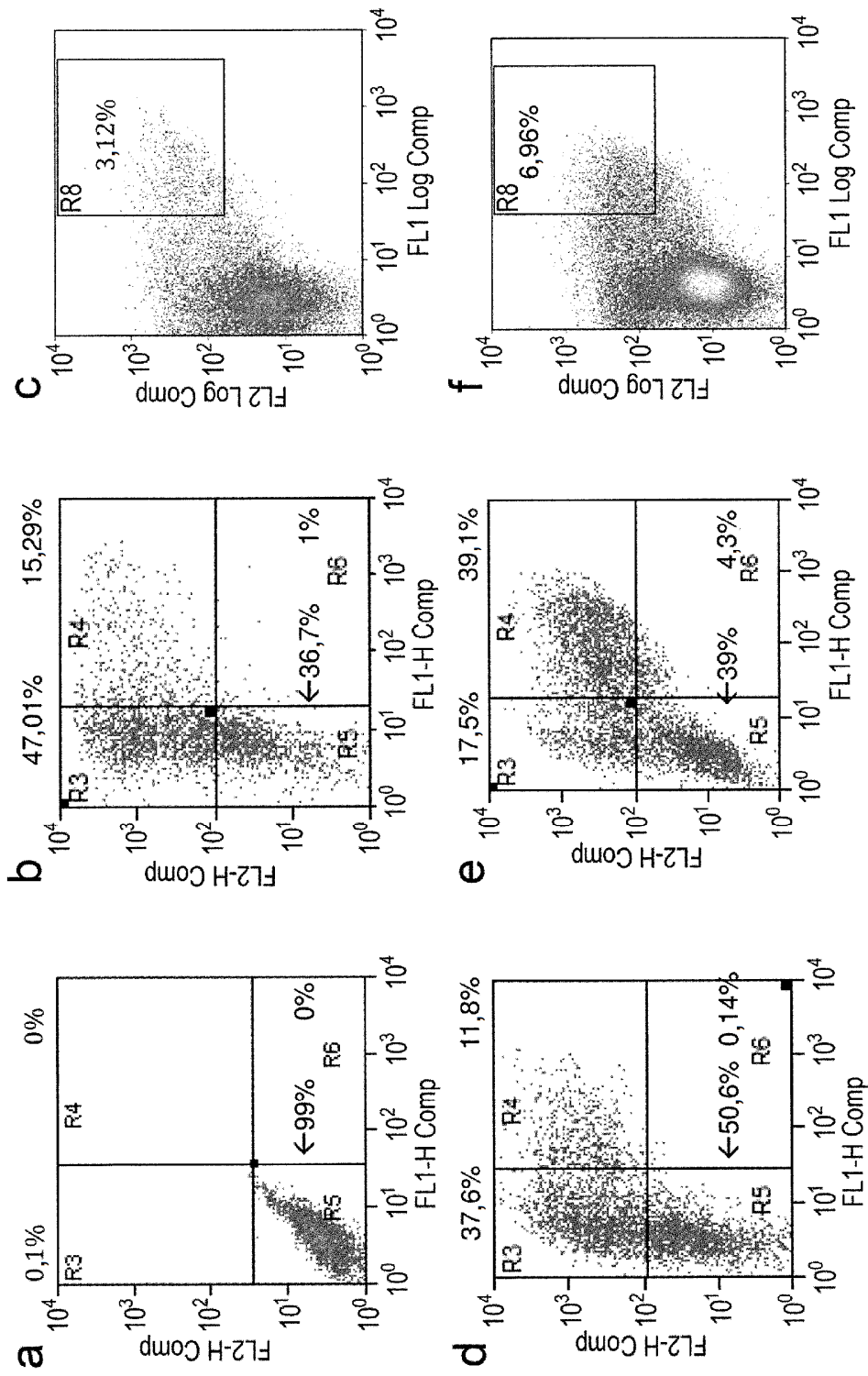


Fig. 9

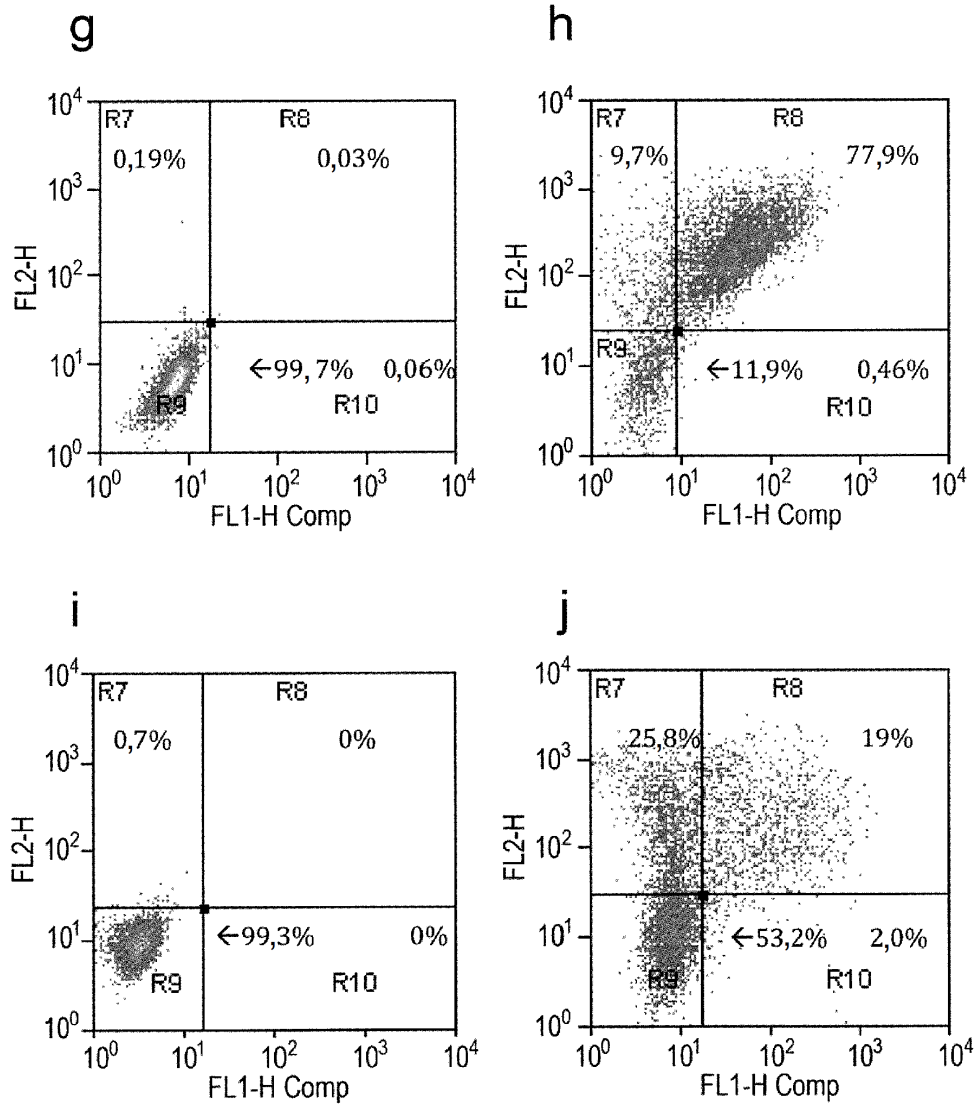


Fig. 9
(continuación)

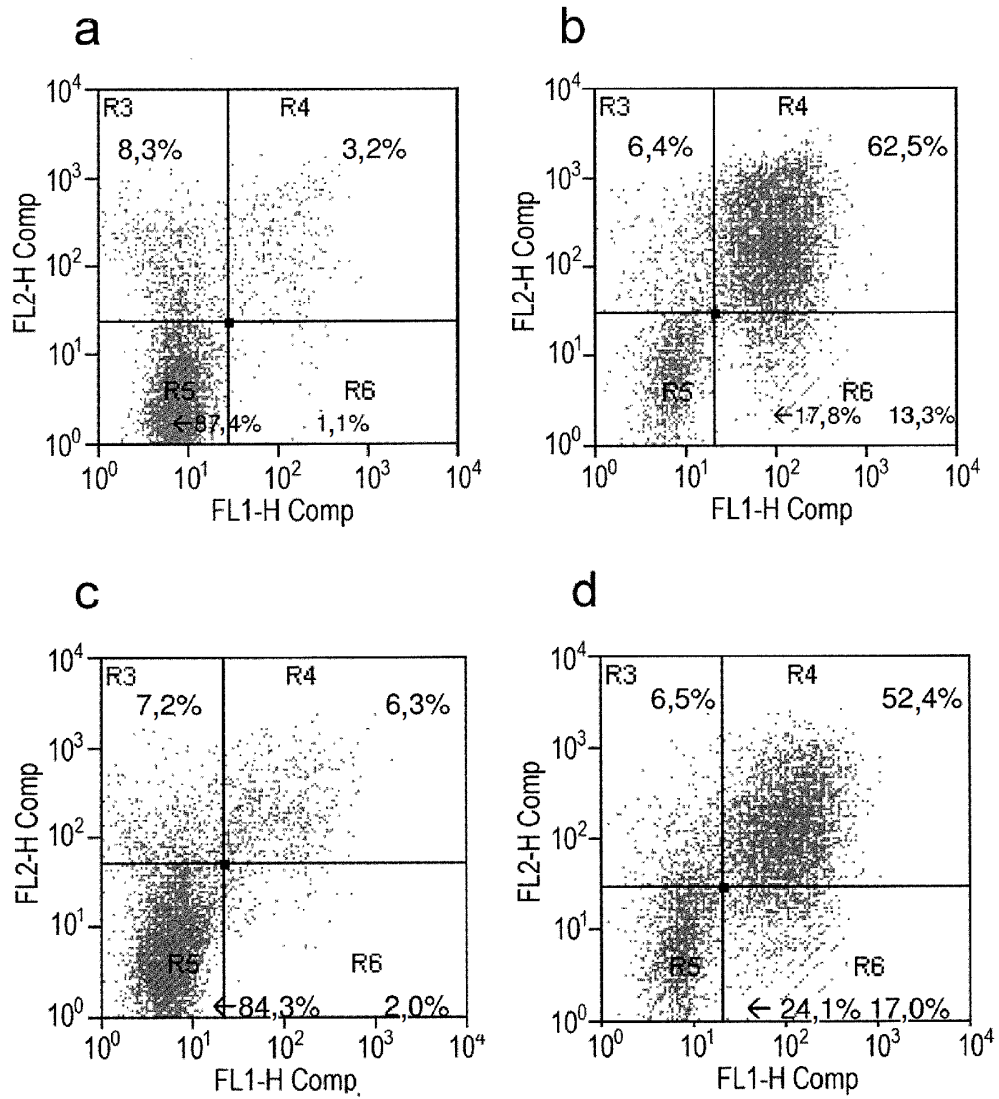


Fig. 10

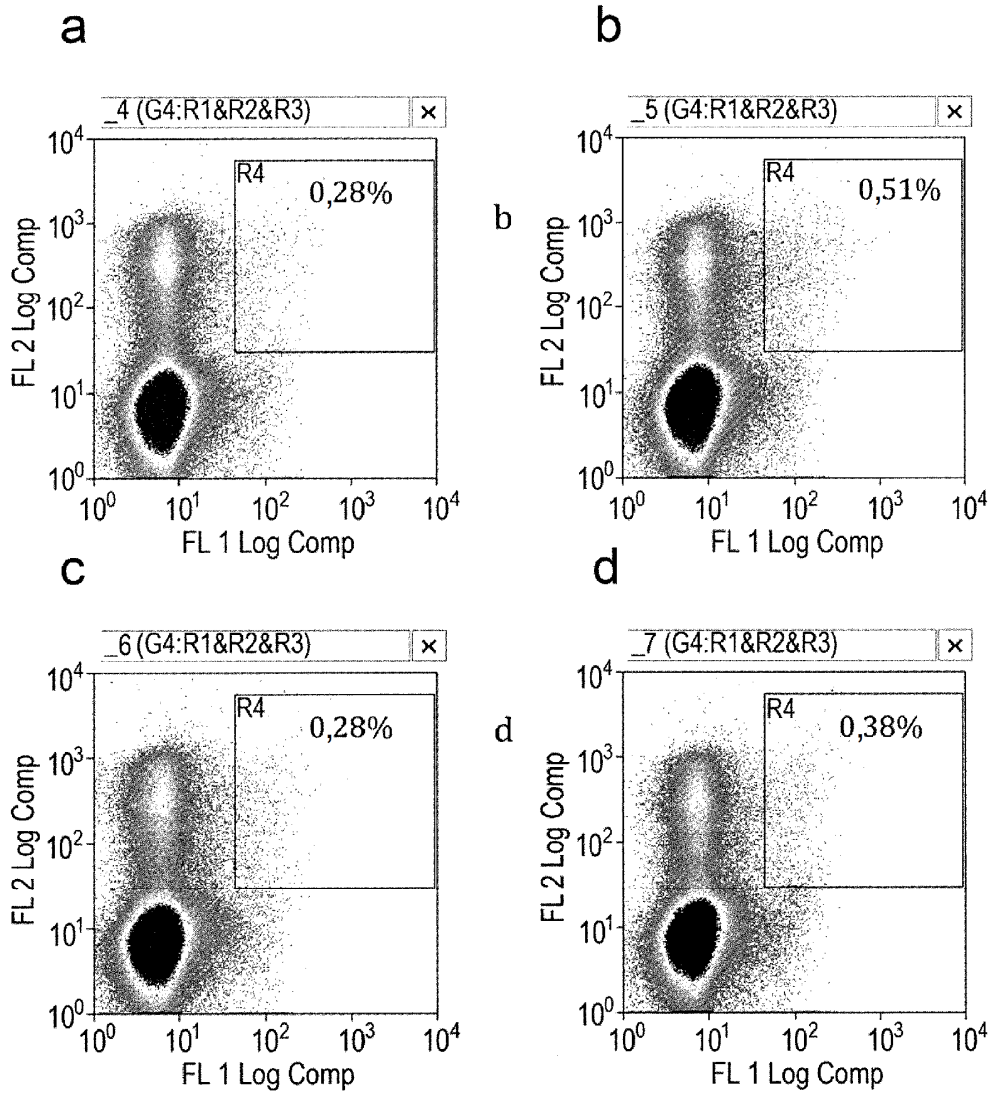


Fig. 11

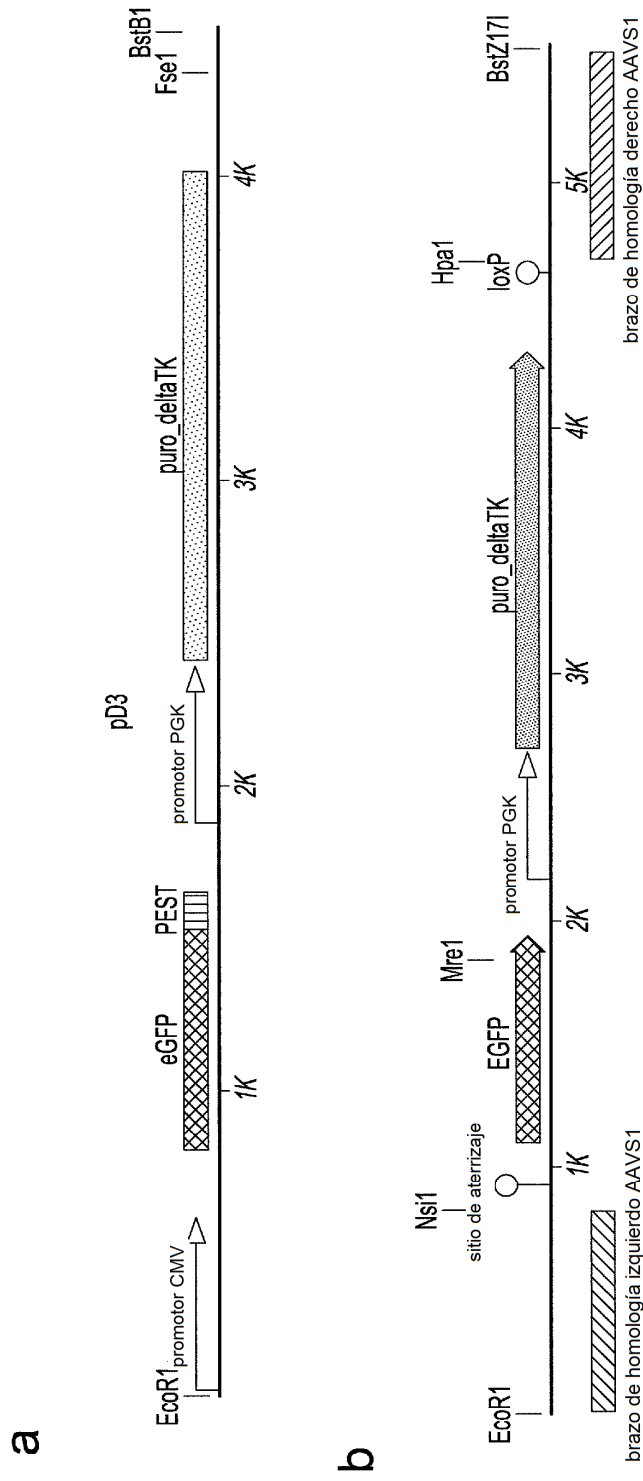


Fig. 12

ES 2 712 632 T3

```

Acc65I"      EcoRI"      AAVS1_left_homology_arm"→
|           | 10         20         30         40         50         60
GGT ACC GAA TTC GCC CTT TGC TTT CTC TGA CCT GCA TTC TCT CCC CTG GGC CTG TGC CGC
822
←---AAVS left homology arm---→TTC TGG GTA CTT TTA TCT GTC CCC TCC ACC CCA CAG TGG GGC
      NsiI      832      sitio de FRT
AAG AIG CAT      GAAGTTCCTAITCCGAAGTTCCTATTCTCTAGAAAAGTATAGGAAC TTC GACC
      lox_2272      I-SceI_meganuclease
ATAACTTCGTATAAAGTATCCTATACGAAGTTAI      GCGATCGCTCCGCGG      TAGGGATAACAGGGTAATAAG
TALEN izquierdo eGFP      espaciador      eGFP_right_TALEN
TCCACCGGTGCGCA      CCATGGTGAGCAAGGG      CGAGGAGCTGTTC      CTTCTGACCTCTTCTCTTCTCTCC
>"Splice_acceptor_site"
| 1030      1040|      1050      1060      1070      1080
CACAG GGC CTA GAG AGA ICT GGC AGC GGA GAG GGC AGA GGA AGT CTT CTA ACA TGC GGT GAG CTG
      G S G E G R G S L L T C G D V>
      _b_b_b_b_b_b_"T2A"__b_b_b_b_b_b_>
      1090      1100      1110      1120      1130
GAG GAG AAT CCC GGA CCG ATG GTG AGC AAG GGA GAA GAA CTC TTC ACC GGG GTG
      E E N P G P M V S K G E E L F T G V>
      _b_b_"T2A"__b_II__c_c_c_c_c_c_"EGFP"__c_c_c_c_c_>
1140      1150      1160      1170      1180      1190
GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTG AAC GGC CAC AAG TTC AGC GTG TCC GGC
      V P I L V E L D G D V N G H K F S V S G>
      _c_c_c_c_c_c_c_"EGFP"__c_c_c_c_c_c_c_>
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC
AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC
AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC
TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG
GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT
ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC
GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC
1740      1750      1760      1770      1780      1790
AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGC ATC ACT CAC
      N E K R D H M V L L E F V T A A G I T H>
      _c_c_c_c_c_c_c_"EGFP"__c_c_c_c_c_c_c_>
1800      1810      1820      1830      1840 MreI 1850
GGC ATG GAC GAG CTG TAC AAG AAG CTG AGC CAC GGC TTC CCG CCG GCG GTG GCG GCG CAG
      G M D E L Y K K L S H G F P P A V A A Q>
      <-----EGFP-----> II> SECUENCIA PEST MU ORNITINA DESCARBOXILASA
1860      1870      1880      1890      1900      1910
GAT GAT GGC ACG CTG CCC ATG TCT TGT GCC CAG GAG AGC GGG ATG GAC CGT CAC CCT GCA
      D D G T L P M S C A Q E S G M D R H P A>
      <-----SECUENCIA PEST MU ORNITINA DESCARBOXILASA----->
      >"bGH_polyA"
1920      1930      1940      1950|      1960      1970      1980
GCC TGT GCT TCT GCT AGG ATC AAT GTG TAG G TGA TCA GCC TCG ACT GTG CCT TCI AGT TGC
      A C A S A R I N V *>
SECUENCIA PEST MU ORNITINA DESCARBOXILASA >
      1990      2000      2010      2020      2030      2040
CAG CCA TCT GTT GTT TGC CCC TCC CCC GTG CCT TCC TTG ACC CTG GAA GGT GCC ACT CCC
      2050      2060      2070      2080      2090      2100
ACT GTC CTT TCC TAA TAA AAT GAG GAA AIT GCA TCG CAT TGT CTG AGT AGG TGT CAT TCT
      2110      2120      2130      2140      2150      2160
AIT CTG GGG GGT GGG CTG GGG CAG GAC AGC AAG GGG GAG GAT TGG GAA GAC AAT AGC AGG

```

Fig. 13

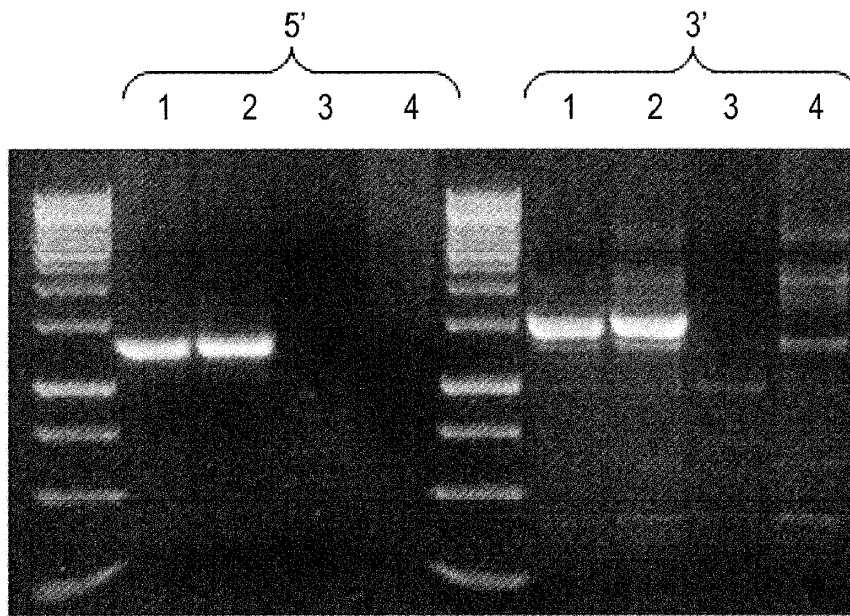


Fig. 14

ES 2 712 632 T3

```

>"Acc65I" >"EcoRI" | _____AAVS1_Left_homology_arm_____>>
| | 10 | 20 | 30 | 40 | 50 | 60
GCT ACC GAA TTC CTA GCT CTT CCA GCC CCC TGT CAT GGC ATC TTC CAG GGG TCC GAG AGC
70 80 90 100 110 120
TCA GCT AGT CTT CTT CCT CCA ACC CGG GCC CCT ATG TCC ACT TCA GGA CAG CAT GTT TGC
130 140 150 160 170 180
TGC CTC CAG GGA TCC TGT GTC CCC GAG CTG GGA CCA CCT TAT ATT CCC AGG GCC GGT TAA
190 200 210 | 220 230 240
TGT GGC TCT GGT TCT GGG TAC TTT TAT CTG TCC CCT CCA CCC CAC AGT GGG GC

NsiI FRT_site
AAGATGCAT GAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAAGTTC GACC

Lox 2272 meganuclease I-SceI
ATAACTTCGTATAAAGTATCCCTATACGAAGTTAT GCGATCGCTCGCGCG TAGGGATAACAGGGTAAT AAG

>"EGFP_left_TALEN" >"Splice_acceptor"
| 370 380 390 400 410
TCCACCGGTCGCC ACCATGGT CT TCT GAC CTC TTC TCT TCC TCC CA CAG GGC ATG
G M

420 430 440 450 460 470
GCA AAA CCT CTG AGC CAG GAA GAA AGC ACA CTG ATT GAA AGA GCA ACC GCT ACT ATC AAC
A K P L S Q E E S T L I E R A T A T I N>
_a_a_a_a_a_a_a_a_a_a_ "BLASTICIDINA " _a_a_a_a_a_a_a_a_a_a_>

480 490 500 510 520 530
AGC ATC CCC ATC TCC GAA GAC TAT TCT GTG GCT AGT GCC GCT CTG TCC AGC GAC GGG AGA
S I P I S E D Y S V A S A A L S S D G R>
_a_a_a_a_a_a_a_a_a_a_ "BLASTICIDINA " _a_a_a_a_a_a_a_a_a_a_>

540 550 560 570 580 590
ATC TTC ACC GGT GTG AAC GTC TAC CAC TTT ACA GGC GGA CCA TGC GCA GAG CTG GTG GTC
I F T G V N V Y H F T G G P C A E L V V>
_a_a_a_a_a_a_a_a_a_a_ "BLASTICIDINA " _a_a_a_a_a_a_a_a_a_a_>

600 610 620 630 640 650
CTG GGG ACT GCA GCC GCT GCA GCC GCT GGT AAT CTG ACC TGT ATC GTG GCC ATT GGC AAC
L G T A A A A A A G N L T C I V A I G N>
_a_a_a_a_a_a_a_a_a_a_ "BLASTICIDINA " _a_a_a_a_a_a_a_a_a_a_>

660 670 680 690 700 710
GAA AAT AGG GGC ATC CTG TCC CCA TGC GGC AGG TGT CGG CAG GTG CTG CTG GAT CTG CAT
E N R G I L S P C G R C R Q V L L D L H>
_a_a_a_a_a_a_a_a_a_a_ "BLASTICIDINA " _a_a_a_a_a_a_a_a_a_a_>

720 730 740 750 760 770
CCT GGC ATC AAG GCA ATT GTC AAA GAC TCT GAT GGA CAG CCT ACC GCC GTC GGT ATC CGT
P G I K A I V K D S D G Q P T A V G I R>
_a_a_a_a_a_a_a_a_a_a_ "BLASTICIDINA " _a_a_a_a_a_a_a_a_a_a_>

```

Fig. 15

ES 2 712 632 T3

```

Nco1 10      20      30      40      50      60
CC ATG GGC TAT CCT TAC GAT GTC CCT GAT TAC GCC AAC AGT CCT GGT ATC CCT GGT ATG
M  G  Y  P  Y  D  V  P  D  Y  A  N  S  P  G  I  P  G  M>
   _b_b_b_"HA TAG"_b_b_b_>
      70      80      90      100     110     120
GGT CCI AAA AAG AAG CGA AAA GTG GGT AGA CTG GAA CCC GGC ATG AAG AAC ATT AAG AAA
G  P  K  K  K  R  K  V  G  R  L  E  P  G  M  K  N  I  K  K>
   _c_c_"NLS"_c_c_c_>           _d_d_"I-SCE1"_d_d_>
      130     140     150     160     170     180
AAT CAG GTG ATG AAC CTG GGA CCT AAT TCC AAG CTG CTG AAA GAG TAC AAG TCT CAG CTG
N  Q  V  M  N  L  G  P  N  S  K  L  L  K  E  Y  K  S  Q  L>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      190     200     210     220     230     240
ATC GAA CTG AAC ATT GAG CAG TTT GAA GCA GGC ATC GGT CTG ATT CTG GGC GAC GCC TAC
I  E  L  N  I  E  Q  F  E  A  G  I  G  L  I  L  G  D  A  Y>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      250     260     270     280     290     300
ATC CGG AGC AGG GAT GAG GGC AAG ACT TAT TGC ATG CAG TTC GAA TGG AAG AAT AAG GCC
I  R  S  R  D  E  G  K  T  Y  C  M  Q  F  E  W  K  N  K  A>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      310     320     330     340     350     360
TAC ATG GAC CAC GTG TGT CTG CTG TAT GAT CAG TGG GTC CTG TCT CCC CCT CAC AAG AAA
Y  M  D  H  V  C  L  L  Y  D  Q  W  V  L  S  P  P  H  K  K>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      370     380     390     400     410     420
GAG AGA GTG AAC CAT CTG GGC AAT CTG GTC ATT ACT TGG GGA GCA CAG ACC TTC AAG CAT
E  R  V  N  H  L  G  N  L  V  I  T  W  G  A  Q  T  F  K  H>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      430     440     450     460     470     480
CAG GCC TTT AAC AAA CTG GCT AAC CTG TTC ATC GTG AAC AAC AAG AAA ACC ATC CCT AAC
Q  A  F  N  K  L  A  N  L  F  I  V  N  N  K  K  T  I  P  N>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      490     500     510     520     530     540
AAT CTG GTC GAA AAC TAC CTG ACA CCA ATG ACT CTG GCC TAT TGG TTC ATG GAC GAT GGC
N  L  V  E  N  Y  L  T  P  M  S  L  A  Y  W  F  M  D  D  G>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      550     560     570     580     590     600
GGA AAA TGG GAC TAC AAC AAG AGC ACA AAC AAA AGC ATC GTG CTG AAT ACC CAG TCC
G  K  W  D  Y  N  K  N  S  T  N  K  S  I  V  L  N  T  Q  S>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      610     620     630     640     650     660
TTC ACA TTT GAG GAA GTG GAG TAT CTG GTC AAG GGC CTG CGG AAC AAA TTC CAG CTG AAC
P  T  F  E  E  V  E  Y  L  V  K  G  L  R  N  K  F  Q  L  N>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      670     680     690     700     710     720
TGC TAC GTG AAG ATC AAC AAG AAC AAG CCA ATC ATC TAC ATC GAT TCT ATG AGT TAC CTG
C  Y  V  K  I  N  K  N  K  P  I  I  Y  I  D  S  M  S  Y  L>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      730     740     750     760     770     780
ATC TTT TAT AAC CTG ATT AAG CCA TAC CTG ATC CCC CAG ATG ATG TAT AAA CTG CCC AAT
I  F  Y  N  L  I  K  P  Y  L  I  P  Q  M  M  Y  K  L  P  N>
   _d_d_d_d_d_d_d_d_"I-SCE1"_d_d_d_d_d_d_d_d_d_>
      790     800
ACA ATC AGC TCC GAG ACT TTC CTG AAG GTCTAGA
T  I  S  S  E  T  F  L  K  V>
   _d_d_d_"I-SCE1"_d_d_d_>
Xba1

```

Fig. 16

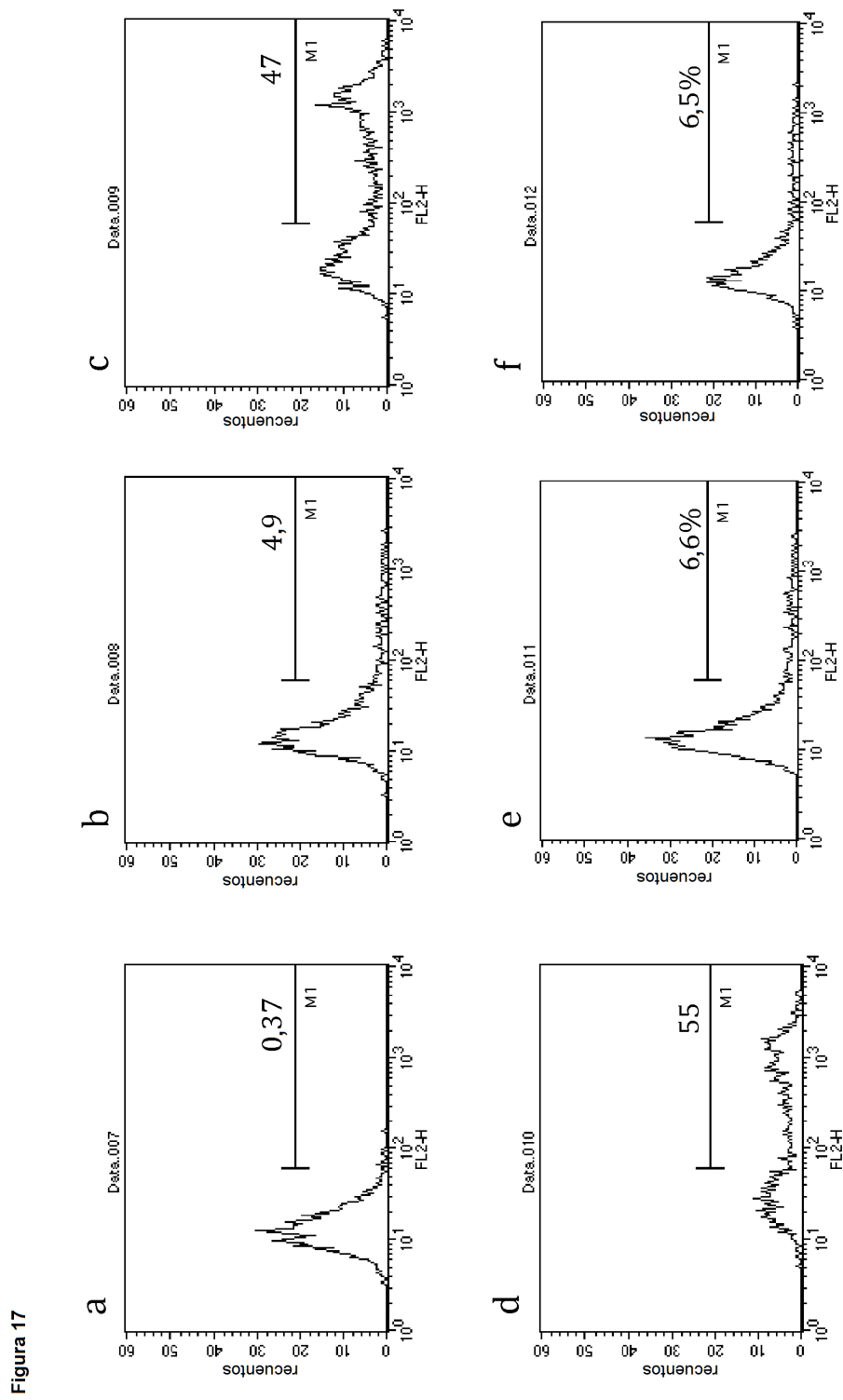


Figura 17

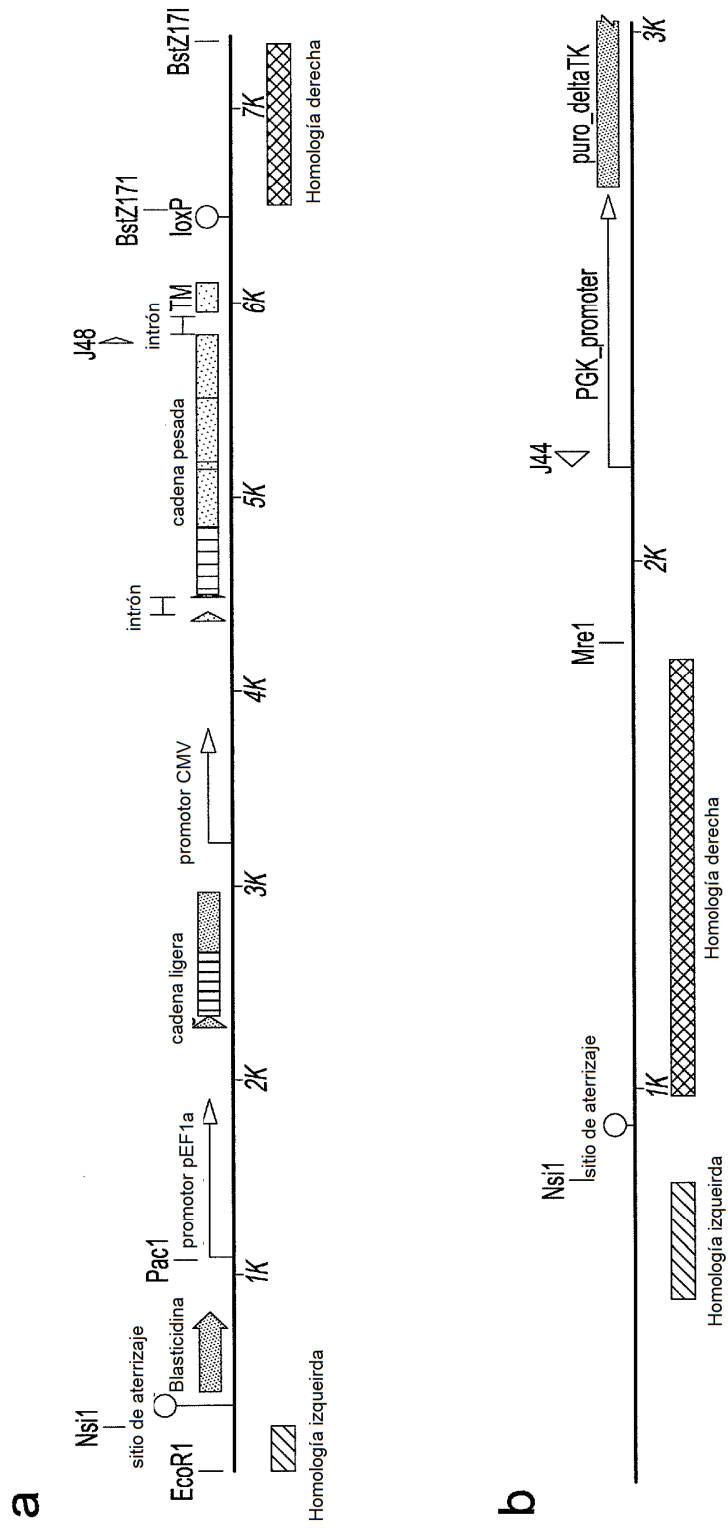


Fig. 18

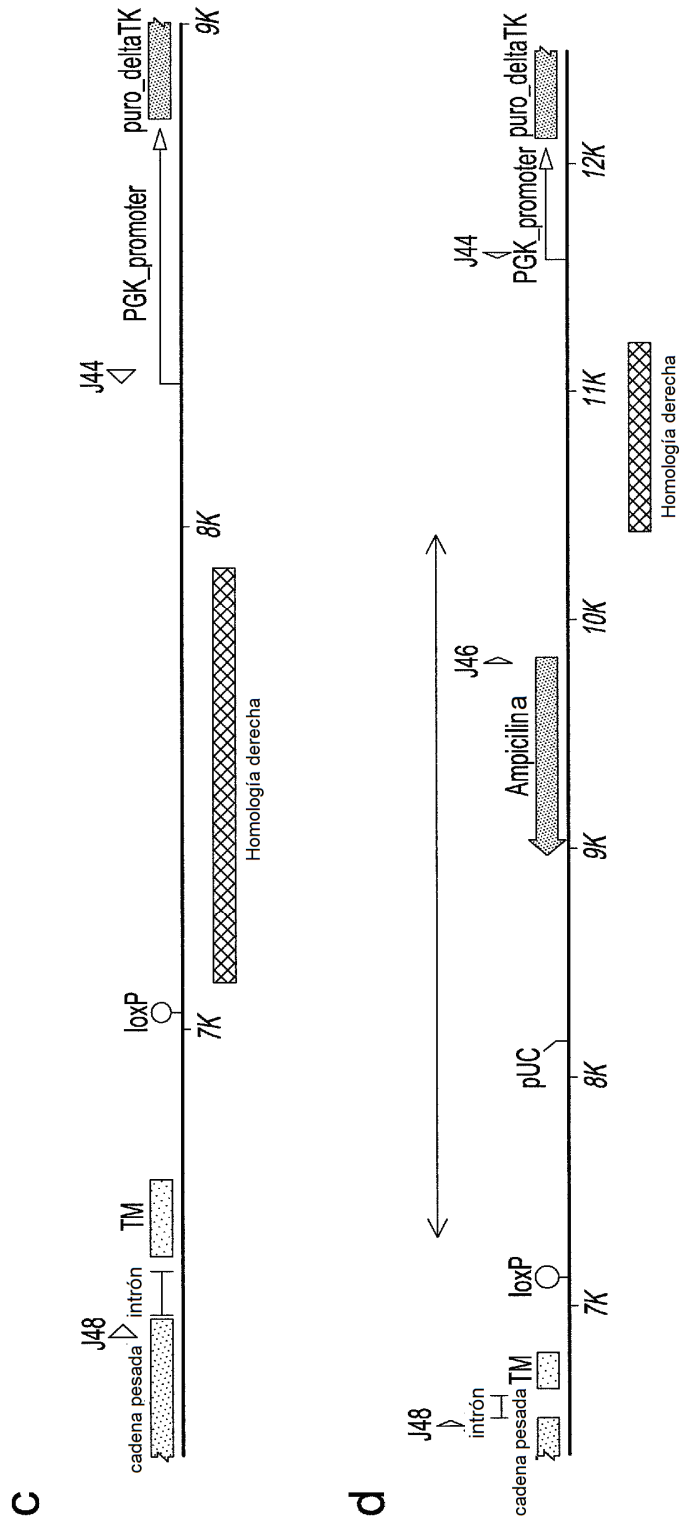


Fig. 18
 (continuación)

e

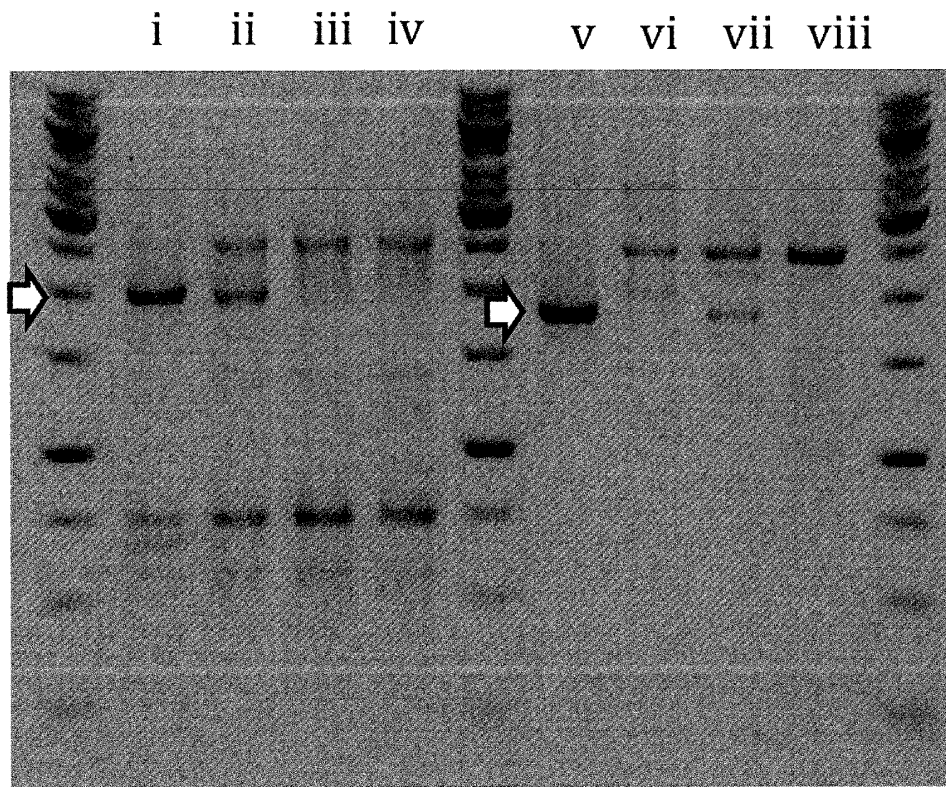


Fig. 18
(continuación)

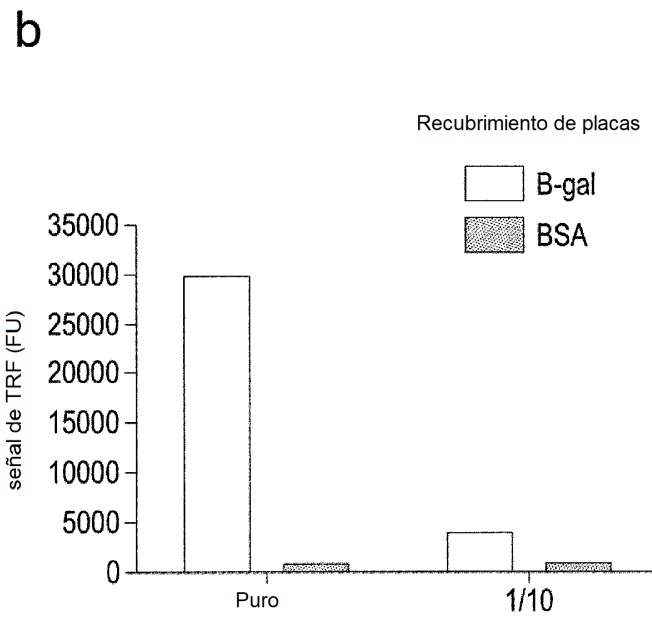
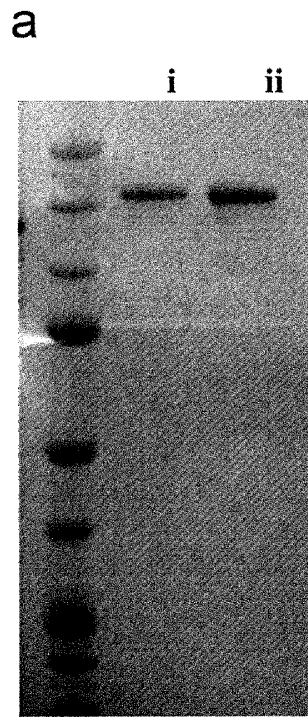


Fig. 19

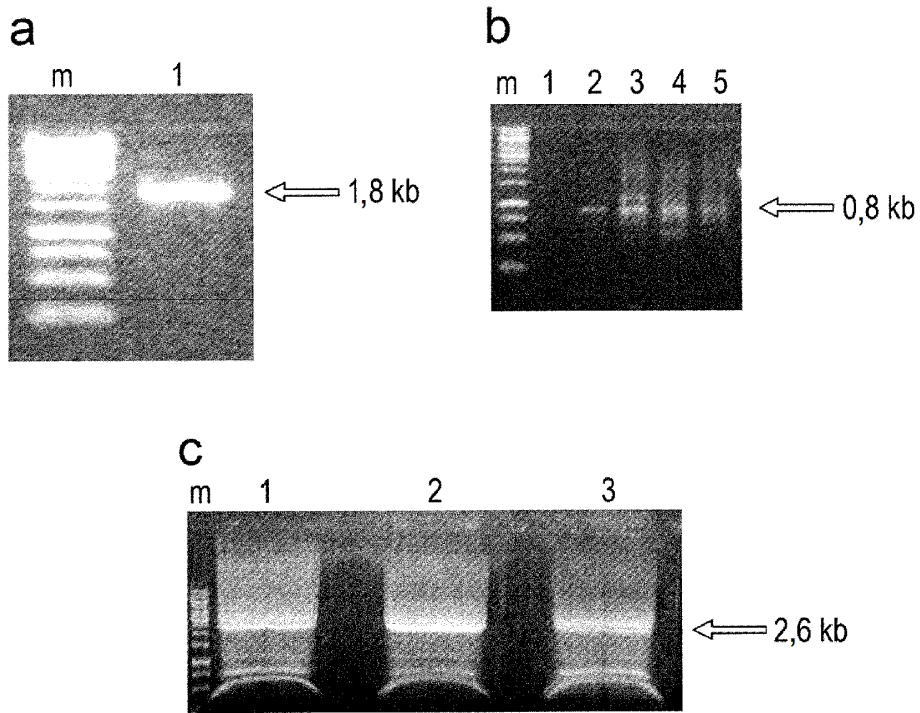


Fig. 20

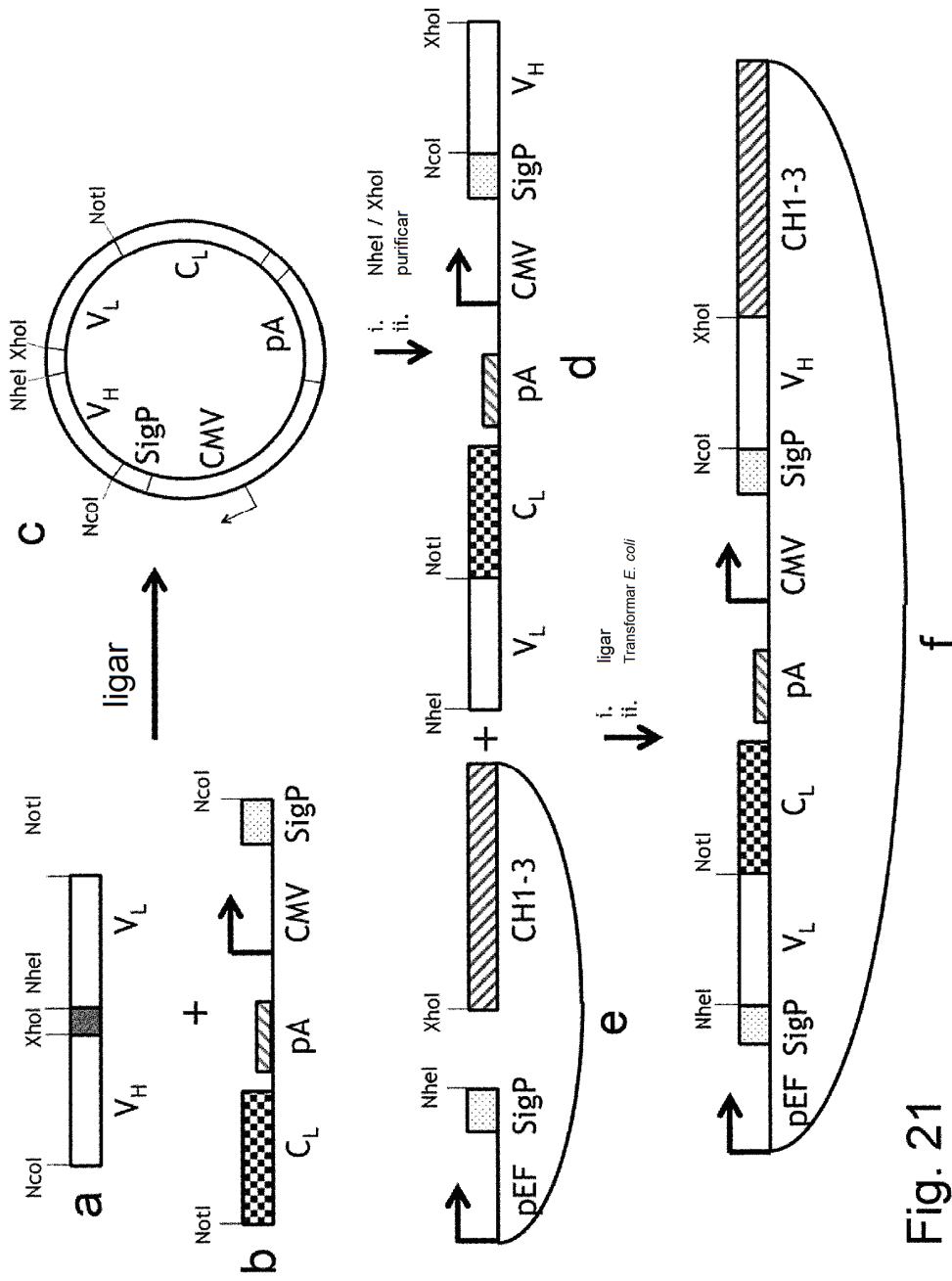


Fig. 21

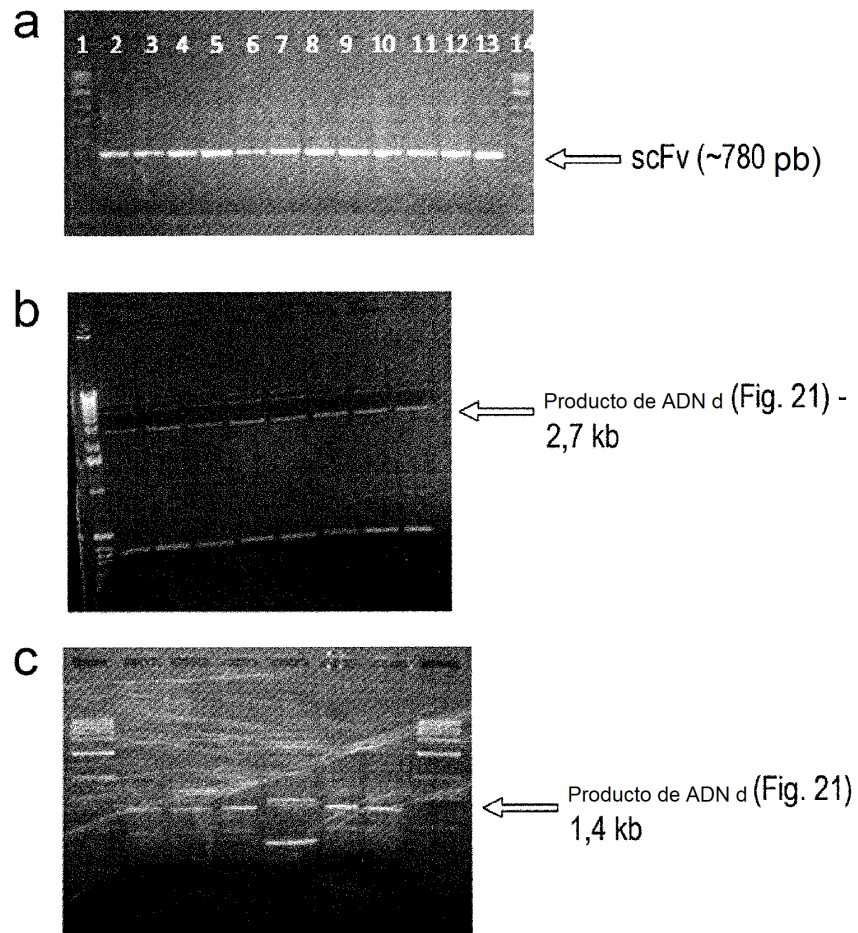


Fig. 22

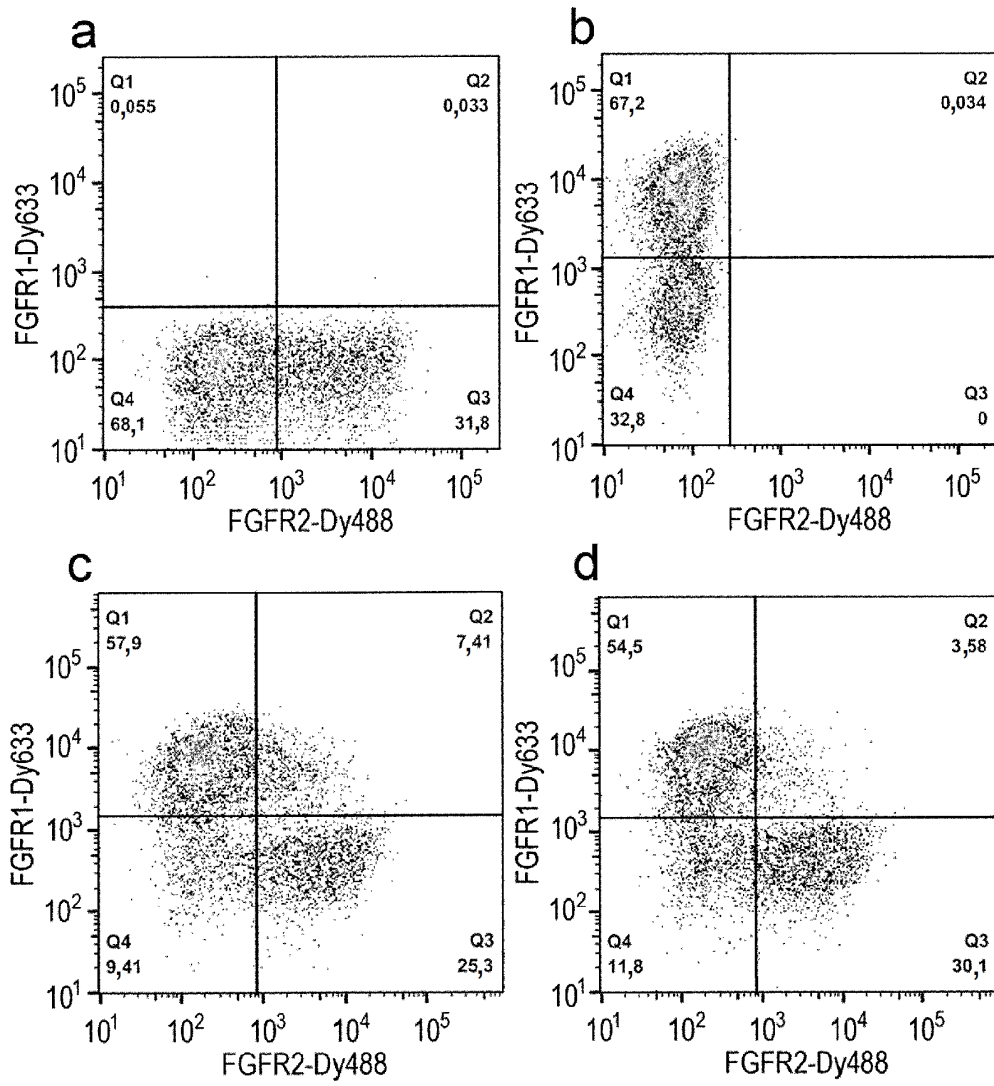


Fig. 23

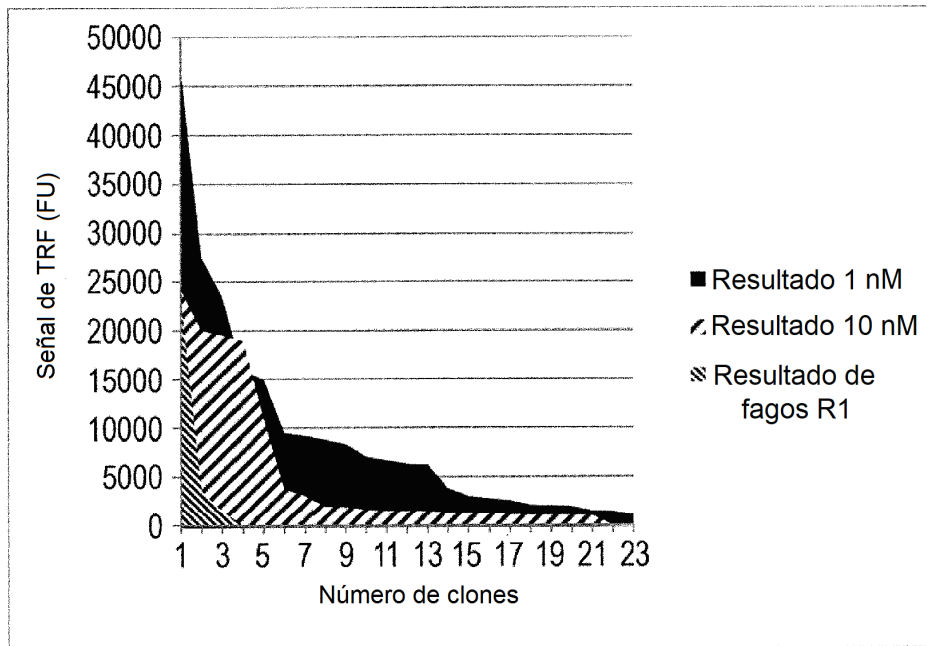


Fig. 24

C

gctagcaagcaggaagtgactcagatcccagccgctctgagcgtgcctgagggagaaaaac
Nhe1 K Q E V T Q I P A A L S V P E G E N
ctggctcctgaattgcagtttcaccgactcagccatctataacctgcagtggtttcgccag
L V L N C S F T D S A I Y N L Q W F R Q
gatccaggcaagggactgacctccctgctgctgattcagagctcccagagggaacagaca
D P G K G L T S L L L I Q S S Q R E Q T
tctggcagactgaatgctagtctggacaaatctagtggacggtctaccctgtacatcgca
S G R L N A S L D K S S G R S T L Y I A
gccagccagcctggagattccgcaacatatctgtgcccgtgcccacttacaggcgga
A S Q P G D S A T Y L C A V R P L T G G
agctacattcccaccttcgggagaggtacaagcctgatcgtgcacccagacatccagaat
S Y I P T F G R G T S L I V H P D I Q N
ccggagcccgcgtataaccagctgaaggaccccagaagccaggacagcaccctgtgcctg
P E P A V Y Q L K D P R S Q D S T L C L
ttcaccgacttcgacagccagatcaacgtgcccaagacaatggaaagcggcaccttcac
F T D F D S Q I N V P K T M E S G T F I
accgacaagaccgtgctggacatgaaggctatggacagcaagagcaacggcgccattggc
T D K T V L D M K A M D S K S N G A I A
tggtccaaccagaccagcttcacatgccaggacatcttcaaagagacaaaacgccacctac
W S N Q T S F T C Q D I F K E T N A T Y
cccagcagcagcgtgccctgtgatgccaccctgaccgagaagtcccttcgagacagacatg
P S S D V P C D A T L T E K S F E T D M
Aacctgaacttcagaacctgtccgcccgcaggcctgctggatcccaagctgtgtctac

Not1

N L N F Q N L S A A A G L L D P K L C Y
ctgctggacgggatcctgttcatctacggtgtgatcctgactgccctgttccctgcgagtc
L L D G I L F I Y G V I L T A L F L R V
aaatthttctcggagtgccgacgctcctgcataaccagcaggggcagaaccagctgtataac
K F S R S A D A P A Y Q Q G Q N Q L Y N
gagctgaatctgggtcggagggaggaatatgacgtgctggataagagacgcggcagggat
E L N L G R R E E Y D V L D K R R G R D
ccagaaatggggggcaagcccagcgcggaaccctcaggagggactgtataatgaa
P E M G G K P Q R R K N P Q E G L Y N E
ctgcagaaggacaaaatggccgaggttactctgaaattgggatgaagggcgagaggaga
L Q K D K M A E A Y S E I G M K G E R R
cgcggaaggacacgatggcctgtaccagggactgagcactgctaccaaggacacatat
R G K G H D G L Y Q G L S T A T K D T Y
gatgctctgcatatgcaggcactgcccctagataataaaggtacc
D A L H M Q A L P P R - - **Acc65I**

Fig. 25

d

ccatggccaacgctggagtgactcagaccctaagttccaggctcctgaaaactgggcag
Nco M A N A G V T Q T P K F Q V L K T G Q
 agtatgacctgcagtgccacaggacatgaatcacgagtacatgtcatggtatcggcag
 S M T L Q C A Q D M N H E Y M S W Y R Q
 gatccagggatgggtctgaggctgatccattacagcgtgggcgctggaactaccgaccag
 D P G M G L R L I H Y S V G A G T T D Q
 ggcgagtgcccaacggatataatgtctcaagaagcaccacagaagatttcccactgcga
 G E V P N G Y N V S R S T T E D F P L R
 ctgctgagcgcgctcctagccagacatccgtgtacttttgcgccagctccaatgtcggg
 L L S A A P S Q T S V Y F C A S S N V G
 aacaccggcgagctgttctttggggaaggttccgcctgacagtgctcgaagcctgaga
 N T G E L F F G E G S R L T V L E D L R

Xho1

aacgtgaccccccaaggtgtccctgttcgagcctagcaaggccgagatcgccaacaag
 N V T P P K V S L F E P S K A E I A N K
 cagaaagccacctcgtgtgcctggccagaggcttcttccccgaccacgtggaactgtct
 Q K A T L V C L A R G F F P D H V E L S
 tgggtgggtcaacggcaaagaggtgcacagcggcgtgtccaccgatccccaggcctacaaa
 W W V N G K E V H S G V S T D P Q A Y K
 gagagcaactacagctactgcctgagcagcagactgcggtgtccgccaccttctggcac
 E S N Y S Y C L S S R L R V S A T F W H
 aacccccggaaccttcagatgccaggtgcagtttcacggcctgagcgaagaggacaag
 N P R N H F R C Q V Q F H G L S E E D K
 tggccccgagggcagccctaagcccgtgacccagaatatctctgccgaagcctggggcaga
 W P E G S P K P V T Q N I S A E A W G R
 gccgactgtggcattaccagcgcagctaccagcagggcgtgctgtctgccaccatcctg
 A D C G I T S A S Y Q Q G V L S A T I L
 tacgaggtcgcgagcggactgctggacccaaagctgtgctacctgctggatgggatcctg
 Y E V A S G L L D P K L C Y L L D G I L
 ttcatctacggtgtgattctgacagccctgttccctgagcgtcaagttcagccggagcgc
 F I Y G V I L T A L F L R V K F S R S A
 gacgcaccagcataccagcaggggcagaatcagctgtataacgagctgaatctgggtcgg
 D A P A Y Q Q G Q N Q L Y N E L N L G R
 agggaggaatacagcgtgctggataagagacgcggcagggatccccgaaatgggcggaaag
 R E E Y D V L D K R R G R D P E M G G K
 cctcagcagcggaaaaacccacaggaggactgtacaatgaactgcagaaggacaaaatg
 P Q R R K N P Q E G L Y N E L Q K D K M
 gctgaggcatattctgaaatcggcatgaagggagagaggagacgcggcaaggacacgat
 A E A Y S E I G M K G E R R R G K G H D
 gggctgtaccaggtctgagtagcagccactaaggacacctatgatgccctgcatatgcag
 G L Y Q G L S T A T K D T Y D A L H M Q
 gctctgccaccagataataaaagctt

A L P P R - - **Hind 3**

Fig. 25 (continuación)

e

gctagccagaaagaggtggaacagaacagcgccctctgagcgtgccagaaggcgtatc
NheI Q K E V E Q N S G P L S V P E G A I
gccagcctgaactgcacctacagctttctgggcagccagagcttcttctggtagacag
A S L N C T Y S F L G S Q S F F W Y R Q
tacagcggcaagagccccgagctgatcatgttcacctacagagagggcgacaaagaggac
Y S G K S P E L I M F T Y R E G D K E D
ggcagattcaccgcccagctgaacaaggccagccagcagctgtccctgctgatcagagac
G R F T A Q L N K A S Q H V S L L I R D
agccagcctagcgacagcgccacctacctgtgcgccgtgaatgatggcggcagactgacc
S Q P S D S A T Y L C A V N D G G R L T
tttggcgagcggcaccacctgaccgtgaagcctgacatccagaatccggagcccgcgta
F G D G T T L T V K P D I Q N P E P A V
taccagctgaaggaccccagaagccaggacagcaccctgtgcctgttcaccgacttcgac
Y Q L K D P R S Q D S T L C L F T D F D
agccagatcaacgtgcccaagacaatggaaagcggcaccttcatcaccgacaagaccgtg
S Q I N V P K T M E S G T F I T D K T V
ctggacatgaaggctatggacagcaagagcaacggcgccattgcctggccaaccagacc
L D M K A M D S K S N G A I A W S N Q T
agcttcacatgccaggacatcttcaaagagacaaacgccacctaccccagcagcgagctg
S F T C Q D I F K E T N A T Y P S S D V
ccctgtgatgccaccctgaccgagaagtcccttcgagacagacatgaacctgaactccag
P C D A T L T E K S F E T D M N L N F Q
aacctgtccgcccgcg
N L S **NotI**

f

NcoI

ccatggccagccagaccatccatcagtgccctgccaccctggcagcctgtgggatct
M A S Q T I H Q W P A T L V Q P V G S
cctctgagcctggaatgcaccgtggaaggcaccagcaaccccaacctgtactggtacaga
P L S L E C T V E G T S N P N L Y W Y R
cagcgcgctggcagagggccccagctgctgtttactggggcccctttggccagatcagc
Q A A G R G P Q L L F Y W G P F G Q I S
agcgaggtgcccagaacctgagcgcagcagaccccaggaccggcagtttatcctgagc
S E V P Q N L S A S R P Q D R Q F I L S
agcaagaagctgctgctgagcgacagcggcttctacctgtgcgcttggagcgagacagcc
S K K L L L S D S G F Y L C A W S E T G
ctgggcatggcggatggcagtttggcgagggcagcagactgacagtgctcagc
L G M G G W Q F G E G S R L T V L E
XhoI

Fig. 25 (continuación)

g

```

5' C ACG CGG CAC GCG GGT GAA NNS NNS CCT WCN ATG TAA GGG TGG AAG CCC GCTC
   C ACG CGG CAC GCG GGT GAA NNS NNS CCT WGN ATG TAA GGG TGG AAG CCC GCTC 5'
     C  A  V  R  P  L  X  X  G  S/T Y  I  P  T  F  G  R
-----CDR3"-----
    
```

h

```

5' G TAC TTT TGC GCC AGC TCC NNS STC GGG NNS ACC GGC GAG CTG TTC TTG
   C ATG AAA ACG CGG TCG AGG NNS SAG CCC NNS TGG CCG CTC GAC AAG AAAC 5'
     Y  F  C  A  S  S  N  V/L G  N  T  G  E  L  F  F
-----CDR3-----
    
```

N=A,C,G o T

S= C O G

W= A o T

Fig. 25 (continuación)

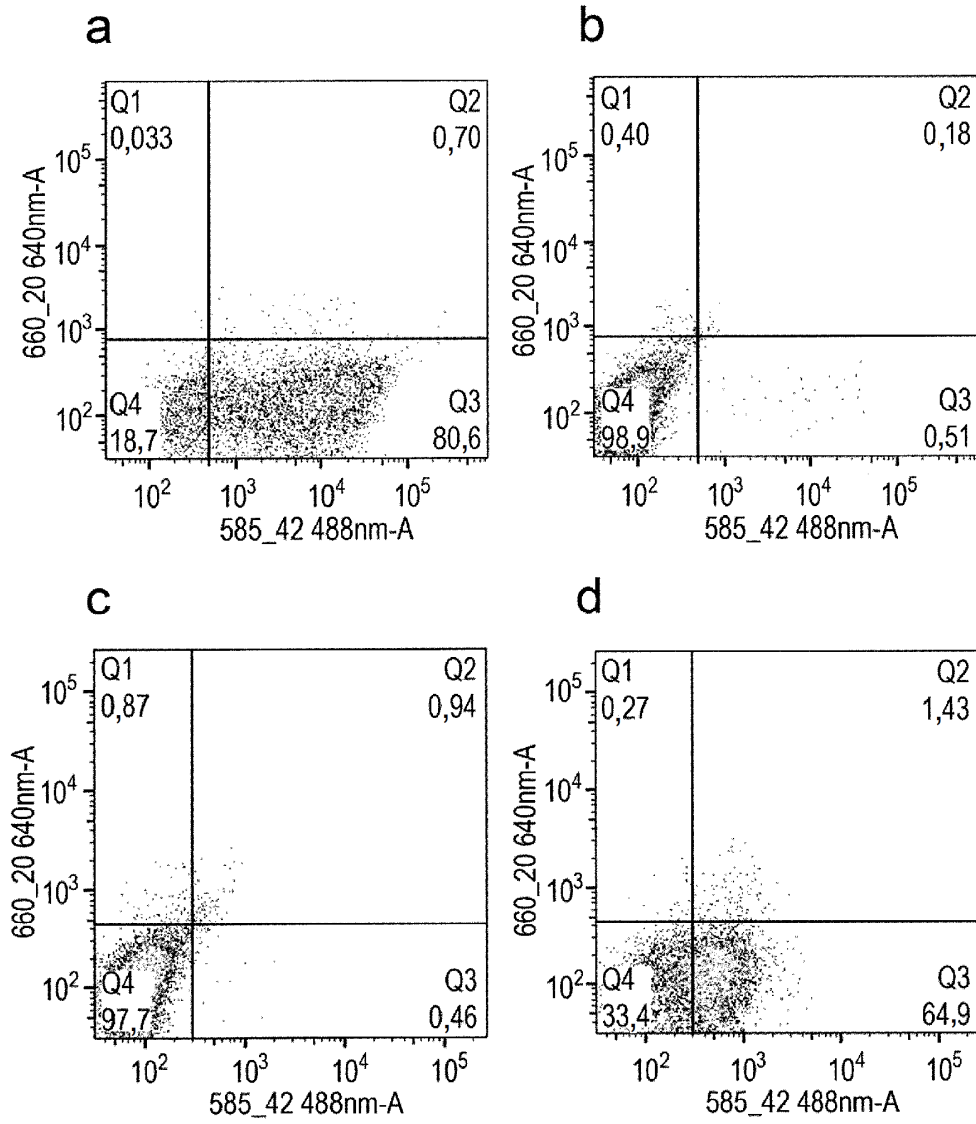


Fig. 26

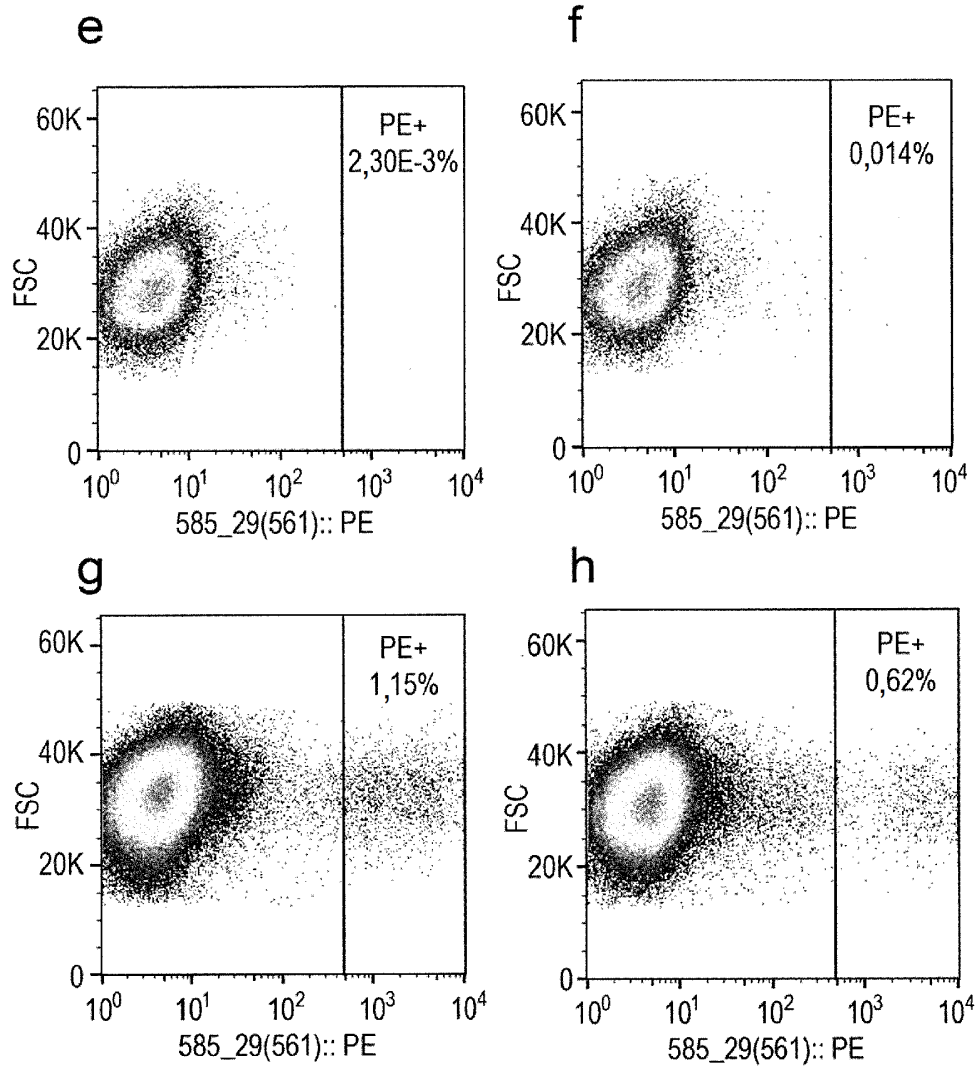


Fig. 26
(continuación)

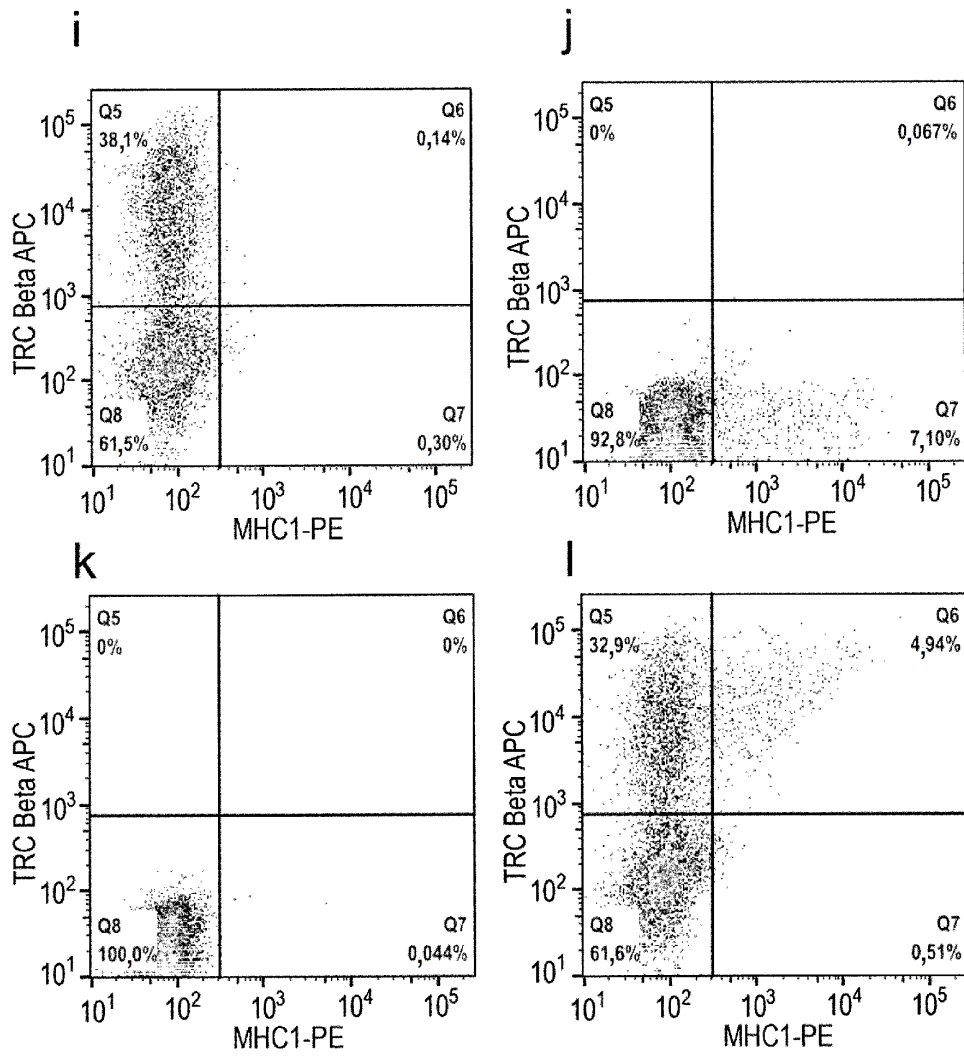


Fig. 26
(continuación)

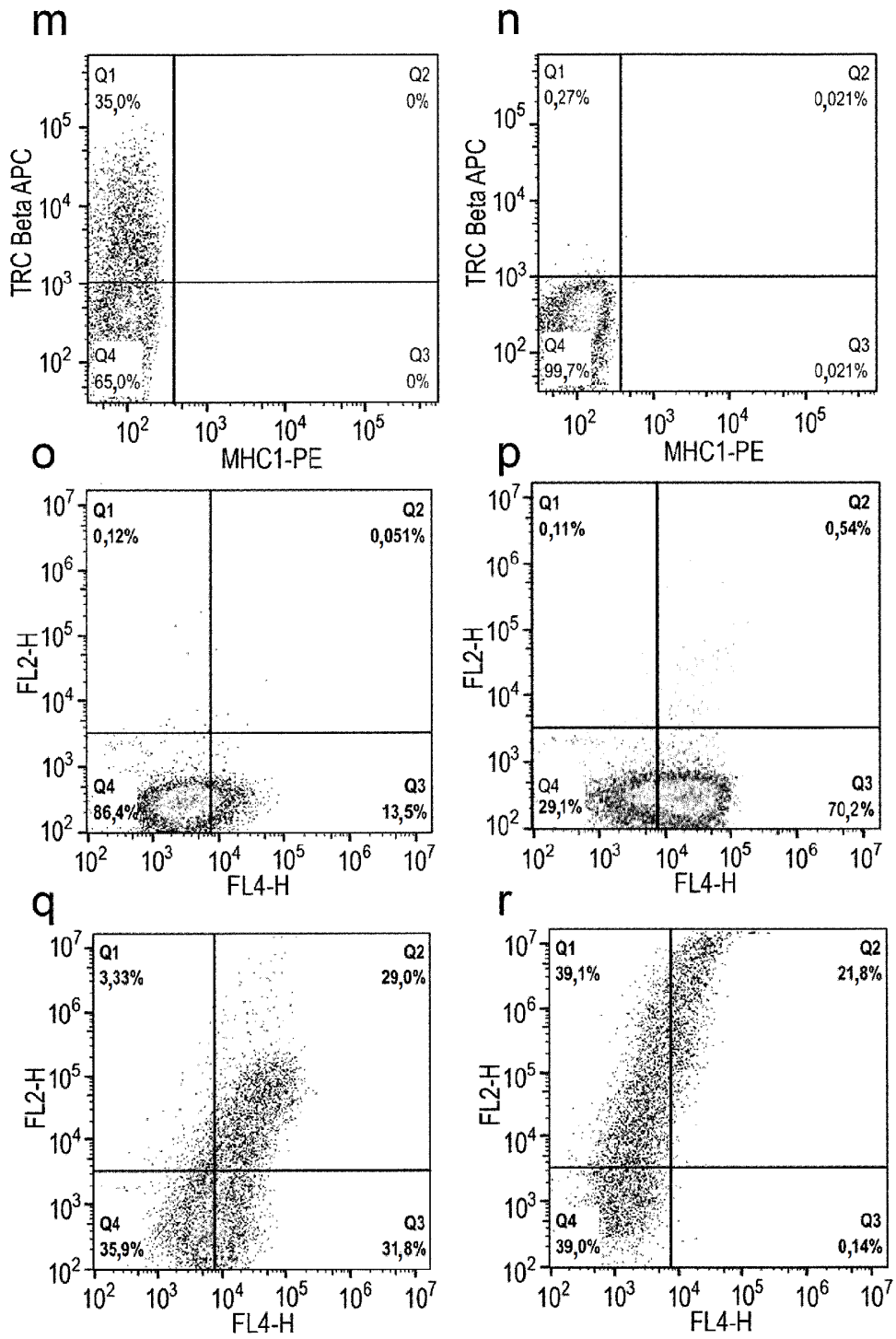


Fig. 26 (continuación)

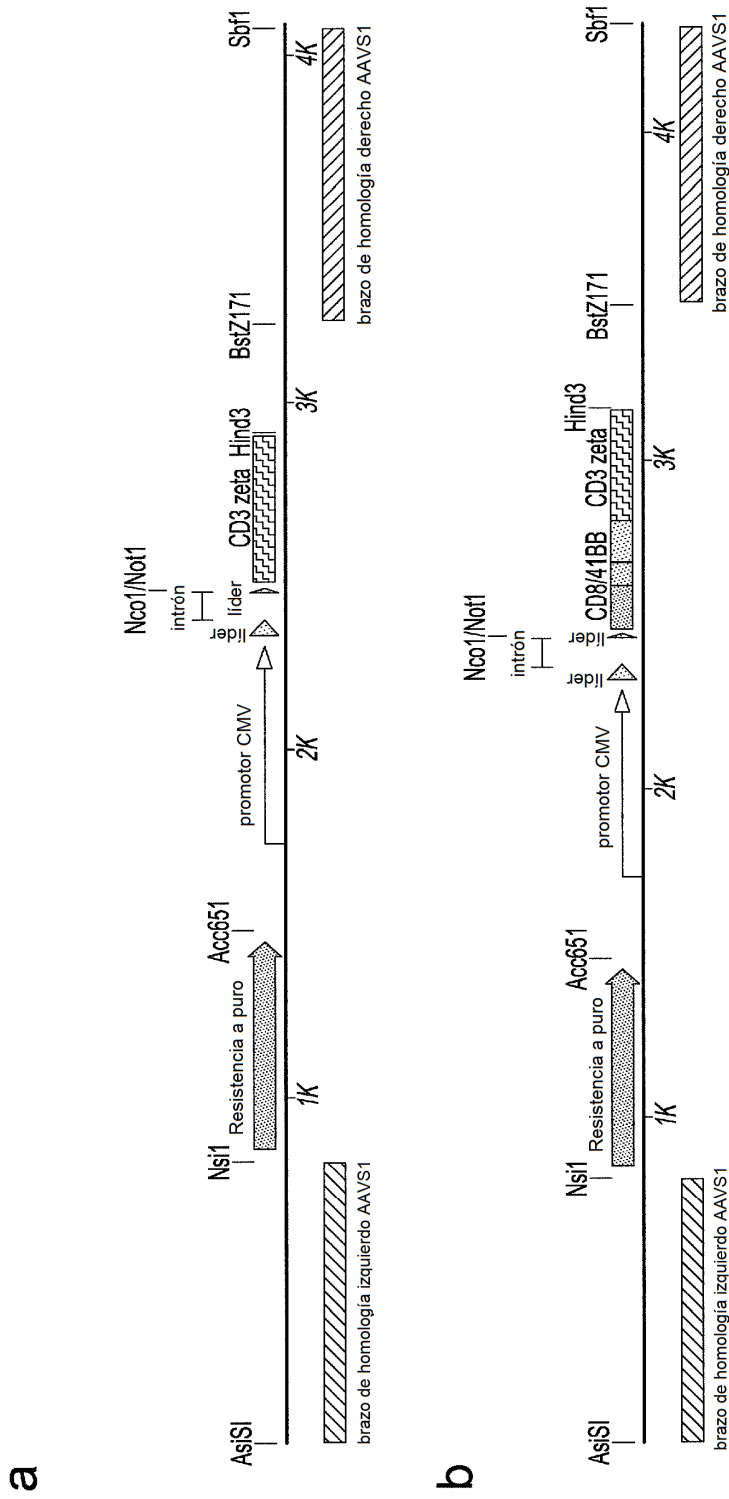


Fig. 27

C

líder	Nco1	Not1
ggcgccatggcccaggtc	<u>cgccggccgcaagcgg</u> actgctggacccaaagctgtgctacctg	
G A M A	S G L L D P K L C Y L	
ctggatgggatcctgtt	catctacgggtgtgattctgacagccctgttcctg	cgagtcaag
L D G I L F I Y G V I L T A L F L R V K		
ttcagccggagcgcggacgcaccagcataccagcaggggcagaatcagctgtataacgag		
F S R S A D A P A Y Q Q G Q N Q L Y N E		
ctgaatctgggtcggagggaggaatacgcgtgctggataagagacgcggcagggatccc		
L N L G R R E E Y D V L D K R R G R D P		
gaaatgggcggaagcctcagcgcagggaaaaacccacaggagggactgtacaatgaactg		
E M G G K P Q R R K N P Q E G L Y N E L		
cagaaggacaaaatggctgagggcatattctgaaatcggcatgaagggagagaggagacgc		
Q K D K M A E A Y S E I G M K G E R R R		
ggcaaaggacacgatgggctgtaccagggctgagtagcagccactaaggacacctatgat		
G K G H D G L Y Q G L S T A T K D T Y D		
	Hind3	
gccctgcatatgcaggctctgccacccagataataaaagctt		
A L H M Q A L P P R - -		

Fig. 27 (continuación)

d

Lider **Nco1** **Not1** I--- bisagra CD8/Transmembrana →
gccatggcc caggtcggggccgca acaacaaccccagccccagacctcctaccct
 A M A T T T P A P R P P T P
 gccctacaattgccagccagcctctgagcctgaggcccgaggctttagaccagctgct
 A P T I A S Q P L S L R P E A C R P A A
 ggcggagccgtgcacaccagaggactggatttcgctgcgacatctacatctgggccct
 G G A V H T R G L D F A C D I Y I W A P

←--- bisagra CD8/Transmembrana ---II 4-1BB→
 ctggccggcacatgtggcgtgctgctgctgagcctcgtgatcacctgtactgc aagcgg
 L A G T C G V L L L S L V I T L Y C K R
 ggcagaaagaaactgctgtacatctttaagcagccctcatgcggcccggtgcagaccacc
 G R K K L L Y I F K Q P F M R P V Q T T

←-- 4-1BB--I
 caggaagaggacggctgctcctgcagattccccgaggaagaagaaggc ggctgcgagctg
 Q E E D G C S C R F P E E E E G G C E L

I---CD3ζ→
 agagtgaagttcagcagatccgccgacgccctgcctacaagcagggccagaaccagctg
 R V K F S R S A D A P A Y K Q G Q N Q L
 tacaacgagctgaacctgggcagacgggaaggtacgacgtgctggacaagcggagagggc
 Y N E L N L G R R E E Y D V L D K R R G
 cgggaccagagatggcggaagcccagaagaagaacccccaggaaggcctgtataac
 R D P E M G G K P R R K N P Q E G L Y N
 gaactgcagaaagacaaaatggccgaggcctacagcgagatcggaatgaagggcgagcgg
 E L Q K D K M A E A Y S E I G M K G E R
 agaagaggcaaggggcacgatggcctgtaccagggcctgagcaccgccaccaaggacacc
 R R G K G H D G L Y Q G L S T A T K D T
 tatgacgccctgcacatgcaggccctgccccctagataataaaagctt
 Y D A L H M Q A L P P R - - **Hind 3**

Fig. 27 (continuación)

e

Nco1

g**ccatgg**ccgaagtgaaactgcaggagtctggacccggcctggtggcccatctcagtct
 A M A E V K L Q E S G P G L V A P S Q S
 ctgagcgtgacctgtaccgtgtccggcgtgtccctgcctgactatggcgtgtcctggatc
 L S V T C T V S G V S L P D Y G V S W I
 agacagccccccagaaagggcctggaatggctgggagtgatctggggcagcgaaaccacc
 R Q P P R K G L E W L G V I W G S E T T
 tactacaacagcgcctgaagtcccggctgaccatcatcaaggacaactccaagagccag
 Y Y N S A L K S R L T I I K D N S K S Q
 gtgttcctgaagatgaacagcctgcagaccgacgacaccgccatctactactgcgccaag
 V F L K M N S L Q T D D T A I Y Y C A K
 cactactactacggcggcagctacgctatggactactggggccagggcacctcggtcacc
 H Y Y Y G G S Y A M D Y W G Q G T S V T
 gtctcgagtggaggcgggttcaggcggaggtggctctggcgggtggcgctagcgacatc
 V S S G G G G S G G G G S G G G A S D I
 cagatgaccagaccaccagcagcctgagcggcagcctgggcgatagagtgaccatcagc
 Q M T Q T T S S L S A S L G D R V T I S
 tgcagagccagccaggacatcagcaagtacctgaactggtatcagcagaaaccogacggc
 C R A S Q D I S K Y L N W Y Q Q K P D G
 accgtgaagctgctgatctaccacaccagcagactgcacagcggcgtgcccagcagattt
 T V K L L I Y H T S R L H S G V P S R F
 tccggctctggcagcggcaccgactacagcctgaccatctccaacctggaacaggaagat
 S G S G S G T D Y S L T I S N L E Q E D
 atcgctaccttctgtcagcaaggcaacaccctgcctacaccttcggcggaggacc
 I A T Y F C Q Q G N T L P Y T F G G G T
 aagctggagatcaaacgtaccg**ggccgca**
 K L E I K R T A A A

Not1

Fig. 27 (continuación)

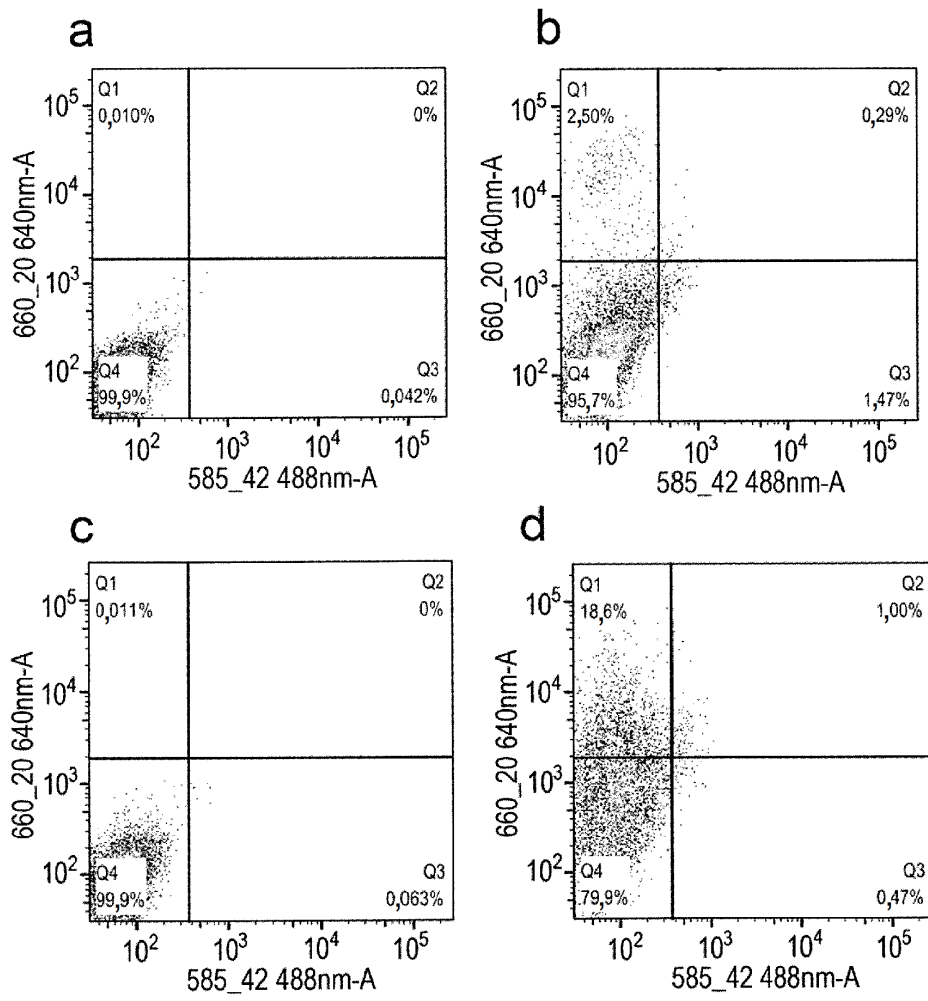


Fig. 28

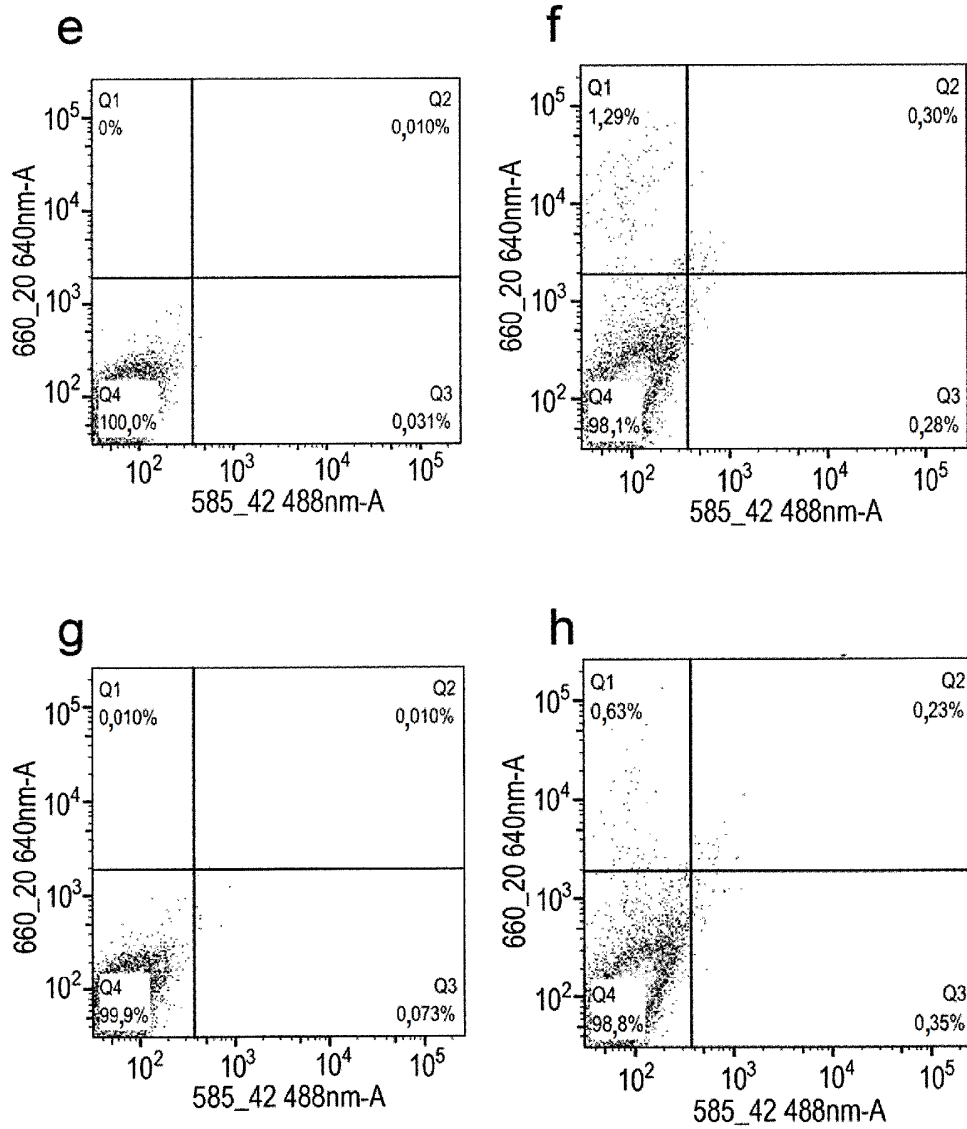


Fig. 28
(continuación)

a

Nco1

gccatggctgctacagggcgtgcgggctgtgcccggaatgagaacagcctggaaatcgag
 A M A A T G V R A V P G N E N S L E I E
 gaactggccagattcgccgtggacgagcacaacaagaagagaacgccctgctggaattc
 E L A R F A V D E H N K K E N A L L E F
 gtgcggtcgtgaaggccaagagcagtgagcgaggccgacaacgactggcacaccatg
 V R V V K A K E Q W S E A D N D W H T M
 tactacctgaccctggaagccaaggacggcggaagaagaagctgtacgaggccaaagtg
 Y Y L T L E A K D G G K K K L Y E A K V
 tgggtcaagctggacctggaacctggcagcacttcaacttcaaagagctccaggaattc
 W V K L D L E T W Q H F N F K E L Q E F
 aagcccgtggcgacgctgcgccgcg
 K P V G D A A A A

Not1

b

Nco1

gccatggctgctacagggcgtgcgggctgtgcccggaatgagaacagcctggaaatcgag
 A M A A T G V R A V P G N E N S L E I E
 gaactggccagattcgccgtggacgagcacaacaagaagagaacgccctgctggaattc
 E L A R F A V D E H N K K E N A L L E F
 gtgcggtcgtgaaggccaagagcaggaacagcccatcggcgagcaccctggaacgac
 V R V V K A K E Q E Q P I G E H P V N D
 accatgtactacctgaccctggaagccaaggacggcggaagaagaagctgtacgaggcc
 T M Y Y L T L E A K D G G K K K L Y E A
 aaagtgtgggtcaagcgggtgctgcggtcaccgagatctacaacttcaaagagctccag
 K V W V K R W L R F T E I Y N F K E L Q
 gaattcaagcccgtggcgacgctgcgccgcg
 E F K P V G D A A A A

Not1

Fig. 29

C

Adhiron_mut1

5' CAGGGTCAGGTAGTACATGGTSNNSNNSNN(SNN)_n CTGCTCTTTGGCCTTCACGAC

Adhiron_mut2

CTGGAGCTCTTTGAAGTTSNNSNNSNN(SNN)_n CTTGACCCACACTTTGGC

GTC GTG AAG GCC AAA GAG CAG (NNS)_n NNS NNS NNS ACC ATG TAC TAC CTG ACC CTG
 CAG CAC TTC CGG TTT CTC GTC (NNS)_n NNS NNS NNS TGG TAC ATG ATG GAC TGG GAC
 V V K A K E Q X X X X T M Y Y L T L

GCC AAA GTG TGG GTC AAG (NNS)_n NNS NNS NNS AAC TTC AAA GAG CTC CAG
 CGG TTT CAC ACC CAG TTC (NNS)_n NNS NNS NNS TTG AAG TTT CTC GAG GTC
 A K V W V K X X X X N F K E L Q

Fig. 29 (continuación)

d

Nco1

gccatggccggtgtgtgccccaagatottgaaaaagtgccgccgtgacagegattg
 A M A G V C P K I L K K C R R D S D C
cccggcgctgcatctgccgcggcaatggctattgcggacgcccgca
 P G A C I C R G N G Y C G **Not 1**

e

GGT GTG TGC VNS VNS VNS VNS VNS VNS VNS VNS VNS VNS TGC CGC CGT
 CCA CAC ACG BNS BNS BNS BNS BNS BNS BNS BNS BNS BNS ACG GCG GCA
 G V C X X X X X X X X X X X X C R R

GAC AGC GAT TGT CCC GGC GCC TGC ATC TGC CGC GGC AAT GGC TAT TGC GGA
 CTG TCG CTA ACA GGG CCG CGG ACG TAG ACG GCG CCG TTA CCG ATA ACG CCT
 D S D C P G A C I C R G N G Y C G

Fig. 29
 (continuación)

ES 2 712 632 T3

a

```

EcoRI                               NsiI
| 10           20           30           | 40           50           60
G GTA CCG AAT TCT AGG GAT AAC AGG GTA ATA TGC ATC TTC TGA CCT CTT CTC TTC CTC CCA

              70           80           90           100           110           120
CAG GGC ATG GCA AAA CCT CTG AGC CAG GAA GAA AGC ACA CTG ATT GAA AGA GCA ACC GCT
M  A  K  P  L  S  Q  E  E  S  T  L  I  E  R  A  T  A  >
_____ "BLASTICIDINA " _____>

```

b

```

EcoRI                               NsiI
| 10           20           30           40           50           60
GGT ACC GAA TTC TTT TCT GTC ACC AAT CCT GGG GCC ACT AGG GAC ACT GTG GGG TGG AGG
              AAVS_TALE_Right_binding           AAVS_TALE_Left_binding

              NsiI
              | 70           80           90           100           110           120
AAA TGC ATC TTC TGA CCT CTT CTC TTC CTC CCA CAG GGC ATG GCA AAA CCT CTG AGC CAG
              M  A  K  P  L  S  Q  >
              _____ "BLASTICIDINA " _____>

```

Fig. 30

a

Cebador J60>

GCGATCGCGCTGATTGGCTTCTTTTCCCTCCCGCCGTGTGTGAAAAACACAAATGGCGTGT
 GGTGGCGTAAGGCGCTGTCAAGTAAACGGCAGCCGGAGTGCAGCAGCCCGGCAGCCTCGC
 TCTGCCACTGGGTGGGGCGGGAGGTAGGTGGGGTGAGGCGAGCTGGACGTGCGGGCGCGGT
 CGGCCTCTGGCGGGGCGGGGAGGGGAGGGGAGGGTCAAGGAAAGTAGCTCGCGCGAGCGG
 CCGCCACCCTCCCCTTCTCTGGGGAGTCGTTTTACCCGCCGCCGGCCGGGCTCGTCGT
 CTGATTGGCTCTCGGGGCCAGAAAAGTGGCCCTTGCCATTGGCTCGTGTTCGTGCAAGTTG
 AGTCCATCCGCCCGCAGCGGGGCGGCGAGGAGGCGCTCCCAGGTTCCGGCCCTCCCCTCG
 GCCCCGCGCCGAGAGTCTGGCCGCGCGCCCTGCGCAACGTGGCAGGAAGCGCGCGCTGGG
 GCGGGGACGGGCAGTAGGGCTGAGCGGCTGCGGGGCGGGTGCAAGCACGTTTCCGACTTGA
 GTTGCCCAAGAGGGGCGTGCTGAGCCAGACCTCCATCGCGCACTCCGGGGAGTGGAGGGAA
 GGAGCGAGGGCTCAGTTGGGCTGTTTTGGAGGCAGGAAGCACTTGCTCTCCCAAAGTCGCTC
 TGAGTTGTTATCAGTAAGGGAGCTGCAGTGGAGTAGGCGGGGAGAAGGCCCGCACCCTTCTCC
 GGAGGGGGGAGGGGAGTGTGCAATACCTTTCTGGGAGTTCTCTGCTGCCTCCTGGCTTCTG
 AGGACCGCCCTGGGCTGGGAGAATCCCTTCCCCCTCTTCCCTCGTGATCTGCAACTCCAGT
CTTTCTAGAATGCATTAAGGGATCTGTAGGGCGCAGTAGTCCAGGGTTTCCTTGATGATGTC
 < 2706 Nsi1 2709→

ATACTTATCCTGTCCCTTTTTTTTTCCACAGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTT
 TCCAGTGGGGATCGACGGTATCGTAGAGTCGAGGCCGCTCTAGGAATTCACGCCGCCACC

← Cebador 2710

ATG ACC GAG

M T E

Puromicina →

b

BstZ171 Cebador J61→

GTATACGGGAATTGAACAGGTGTAATAATGGAGGGACAAGACTTCCCACAGATTTTCGGTTT
 TGTCGGGAAGTTTTTAATAGGGGCAAATAAGGAAAATGGGAGGATAGGTAGTCATCTGGGG
 TTTTATGCAGCAAACTACAGGTTATTATTGCTTGTGATCCGCTCGGAGTATTTTCCATCG
 AGGTAGATTAAAGACATGCTCACCCGAGTTTTATACTCTCCTGCTTGAGATCCTTACTACAG
 TATGAAATTACAGTGTGCGGAGTTAGACTATGTAAGCAGAATTTAATCATTTTTAAAGAGC
 CCAGTACTTCATATCCATTTCTCCCGCTCCTTCTGCAGCCTTATCAAAAGGTATTTTAGAAC
 ACTCATTTTAGCCCCATTTTCATTTATTATACTGGCTTATCCAACCCCTAGACAGAGCATTG
 GCATTTTCCCTTTCCCTGATCTTAGAAGTCTGATGACTCATGAAACCAGACAGATTAGTTACA
 TACACCACAAATCGAGGCTGTAGCTGGGGCCTCAACACTGCAGTTCTTTTATAACTCCTTAG
 TACACTTTTTGTTGATCCTTTGCCTTGATCCTTAATTTTCAGTGTCTATCACCTCTCCCGTC
 AGGTGGTGTTCACATTTGGGCCTATTCTCAGTCCAGGGAGTTTTACAACAATAGATGTATT
GAGAATCCAACCTCCTGCAGG

← Cebador J62 Sbf1

Fig. 31