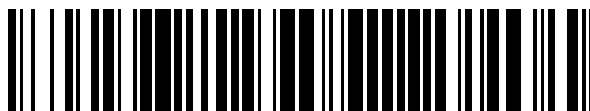


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 394 877**

51 Int. Cl.:

**C12N 15/90** (2006.01)

**C12N 15/82** (2006.01)

**C12N 15/66** (2006.01)

**A01K 67/027** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.08.2001 E 01962178 (8)**

97 Fecha y número de publicación de la solicitud europea: **21.05.2003 EP 1311661**

54 Título: **Recombinación homóloga mejorada mediada por proteínas de recombinación de lambda**

30 Prioridad:

**14.08.2000 US 225164 P**

**26.02.2001 US 271632 P**

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

**06.02.2013**

73 Titular/es:

**THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (100.0%)**

**The National Institutes of Health, Office of Technology Transfer, 6001 Executive Boulevard, Suite No. 325  
Rockville, MD 20852, US**

72 Inventor/es:

**COURT, DONALD, L.;  
YU, DAIGUAN;  
LEE, E-CHIANG;  
ELLIS, HILARY, M.;  
SWAMINATHAN, SRIVIDYA y  
SHARAN, SHYAM K.**

74 Agente/Representante:

**ARIAS SANZ, Juan**

**ES 2 394 877 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

Recombinación homóloga mejorada mediada por proteínas de recombinación de lambda

La presente divulgación se refiere a métodos para mejorar la recombinación homóloga en bacterias y células eucariotas empleando proteínas de recombinación, como las obtenidas del bacteriófago lambda. También se refiere a métodos para modificar el ADN genómico en cromosomas artificiales bacterianos (BAC) y para subclonar el ADN genómico de los BAC en plásmidos multicopia.

## ANTECEDENTES DE LA INVENCION

El uso concertado de endonucleasas de restricción y ligasas de ADN permite la recombinación *in vitro* de secuencias de ADN. El ADN recombinante generado mediante restricción y enlace puede ser amplificado en un microorganismo apropiado, como *E. coli*, y usado para diversos fines, incluyendo la terapia genética. No obstante, el planteamiento de restricción-enlace presenta dos limitaciones prácticas: en primer lugar, las moléculas de ADN solamente pueden ser combinadas de forma precisa, si se dispone de sitios de restricción apropiados; en segundo lugar, dado que los sitios de restricción útiles a menudo se repiten en un tramo largo de ADN, el tamaño de los fragmentos de ADN que se pueden manipular es limitado, normalmente inferior a unas 25 kilobases.

La recombinación homóloga, definida generalmente como un intercambio entre segmentos homólogos en cualquier punto a lo largo de dos moléculas de ADN, proporciona un método alternativo para modificar el ADN. Al generar ADN recombinante con una recombinación homóloga, un microorganismo, como *E. coli*, o una célula eucariota, como una levadura, o una célula de vertebrado, se transforma con ADN exógeno. El centro del ADN exógeno contiene el transgen deseado, mientras que cada uno de los flancos contiene un segmento de homología con el ADN de las células. El ADN exógeno se introduce en la célula con técnicas estándar, como la electroporación o la transfección mediada por fosfato cálcico, y se recombina en el ADN de la célula, por ejemplo con la ayuda de proteínas que promueven la recombinación en la célula.

Para generar ADN recombinante mediante recombinación homóloga, a menudo es beneficioso trabajar con segmentos lineales cortos de ADN. Por ejemplo, se puede introducir una mutación en un segmento lineal de ADN empleando técnicas de reacción en cadena de la polimerasa (PCR). En las circunstancias apropiadas, la mutación puede ser introducida posteriormente en el ADN celular mediante recombinación homóloga. Estos segmentos lineales cortos de ADN pueden transformar una levadura, aunque la manipulación posterior del ADN recombinante en la levadura es laboriosa. Por lo general, resulta más sencillo trabajar con bacterias, aunque los fragmentos lineales de ADN no transforman fácilmente las bacterias (debido en parte a la degradación provocada por las exonucleasas bacterianas). Por consiguiente, los elementos recombinantes son poco comunes, requieren cadenas de crecimiento limitado especiales (como las cadenas de RecBCD-mutante) y, por lo general, precisan miles de pares de bases de homología. Así pues, se necesitan métodos mejorados para promover la recombinación homóloga en las bacterias.

En las células eucariotas, la recombinación homóloga selectiva sienta la base para dirigirse a cualquier secuencia deseada de una molécula de ADN dúplex y modificarla esencialmente, como dirigirse a una secuencia de ADN de un cromosoma para la sustitución por otra secuencia. El planteamiento puede resultar útil para el tratamiento de enfermedades genéticas humanas.

La recombinación homóloga se ha utilizado para crear animales knock-out, mutantes y transgénicos, por lo que ha desempeñado un papel sumamente importante para el conocimiento de la función genética. Los animales transgénicos son organismos que contienen copias establemente integradas de genes o estructuras genéticas obtenidas de otra especie en el cromosoma del animal transgénico. Estos animales se pueden generar introduciendo estructuras de ADN clonado de los genes extraños en células totipotentes mediante diversos métodos, incluyendo la recombinación homóloga.

En la actualidad, los métodos para producir transgénicos se han aplicado en células madre embrionarias (ES) totipotentes con cigotos fertilizados. Las células ES presentan una ventaja en el sentido de que se pueden manipular *in vitro* grandes cantidades de células, antes de emplearlas para generar transgénicos. Alternativamente, también se puede introducir el ADN en oocitos fertilizados mediante microinyección en el pronúcleo o inyectarse en la línea germinal de organismos, incluyendo las especies *C. elegans* o *Drosophila*.

Se han desarrollado diversos métodos para detectar y/o seleccionar recombinantes de sitio específico diana entre el ADN del vector y la secuencia cromosómica homóloga diana (Capecchi, *Science* 244: 1288, 1989). Las células que presentan un fenotipo específico tras la recombinación, tal como ocurre con la alteración del gen de hipoxantina fosforibosiltransferasa (*hprt*), se pueden obtener mediante selección directa en el medio de crecimiento adecuado. Alternativamente, se puede incorporar un marcador seleccionable, como la resistencia a la neomicina, en un vector bajo control de un promotor, y se puede valorar el éxito de la transfección seleccionando células resistentes a G418 (Joyner *et al.*, *Nature* 338:153,1989). Se han descrito muchos otros procedimientos de selección (Jasin y Berg, *Genes and Development* 2: 1353, 1988; Doetschman *et al.* *Proc. Natl. Acad. Sci. (EE.UU.)* 85: 8583, 1988; Dorini *et al.*, *Science* 243: 1357, 1989; Itzhaki y Porter, *Nucl. Acids Res.* 19: 3835,1991). Lamentablemente, las secuencias exógenas

transferidas a las células eucariotas solamente se someten a recombinación homóloga a frecuencias muy bajas, incluso cuando hay presentes regiones de homología muy largas (Koller *et al.* Proc. Natl. Acad. Sci. (EE.UU.), 88:10730, 1991, y Snouwaert *et al.* Science 257:1083, 1992). Por tanto, es necesario transfectar, seleccionar y estudiar grandes números de células, a fin de generar un recombinante homólogo correctamente selectivo.

- 5 Muniyappa *et al.*, J. Biol. Chem., V261, nº 16, p7472-7478, 1986, se refiere al fago lambda. Datsenko *et al.*, PNAS, 97, pp 6640-6645, 2000, se refiere a la recombinación homóloga en lambda. Huang *et al.*, Gene, 251, pp 187-197, 2000, se refiere a un promotor de arabinosa. Yu *et al.*, PNAS, 97, pp5978-5983, 2000, se refiere a recombinación homóloga en lambda.

## RESUMEN DE LA DIVULGACIÓN

- 10 La presente divulgación proporciona métodos para clonar moléculas de ADN en células que tienen un ADN que codifica recombinasas de lambda operativamente unidas a un promotor susceptible de desrepresión, como por ejemplo el promotor PL del fago lambda. El promotor PL es activado, mediante un cambio de temperatura, permitiendo así la expresión de recombinasas lambda. Las recombinasas lambda promueven la recombinación homóloga entre los ácidos nucleicos de la célula. Los ácidos nucleicos que se someten a recombinación pueden ser intracromosómicos o  
15 extracromosómicos, por ejemplo en un cromosoma artificial bacteriano.

- La presente divulgación también proporciona métodos para inducir la recombinación homóloga empleando moléculas de ADN de cadena única, mediante la introducción en el ADN de la célula capaz de someterse a recombinación homóloga, y un polipéptido de unión a ADN de cadena única capaz de promover la recombinación homóloga. Estos polipéptidos de unión a ADN de cadena única incluyen el lambda Beta, RecT, P22 Erf y Rad52, así como variantes y  
20 fragmentos funcionales de los polipéptidos de unión a ADN de cadena única.

La presente divulgación también proporciona células bacterianas que promueven una recombinación homóloga eficiente. Estas células bacterianas contienen uno o más genes o promotores de un profago lambda defectuoso dentro del cromosoma bacteriano.

- 25 La divulgación también proporciona métodos para alterar los genes eucariotas, expresando recombinasas operativamente vinculadas a un promotor susceptible de desrepresión en las células bacterianas. Los genes eucariotas así modificados pueden ser empleados para modificar células eucariotas, por ejemplo para generar animales transgénicos o knockout.

## BREVE DESCRIPCIÓN DE LAS FIGURAS

- 30 La **Fig. 1** es una ilustración lineal del profago lambda defectuoso y está integrado en el cromosoma de *E. coli*. Los genes de profago están indicados con la línea continua y los genes de *E. coli* mediante líneas discontinuas. Las regiones de codificación para los genes de recombinación de lambda -exo, bet y gam- se encuentran aproximadamente en el centro del profago defectuoso. Los genes de lambda, cro a attR, sometidos a delección en el profago defectuoso, se encuentran entre paréntesis junto con el gen de *E. coli*, bioA, para indicar su delección.

- 35 La **Fig. 2A** es un diagrama esquemático que muestra la tecnología recombinante clásica y la Fig. 2B es un diagrama esquemático que muestra la "recombinería" empleando recombinación homóloga, tal y como se divulga en el presente documento.

- 40 La **Fig. 3** es un gráfico que muestra el efecto del tiempo de inducción de las proteínas de lambda Beta, Exo y Gam sobre la eficiencia de la recombinación. Se empleó una cassette de resistencia al cloranfenicol lineal para focalizar los genes del profago y se electroporó en las células, después de que éstas hubiesen sido calentadas hasta 42 °C durante el tiempo indicado. El tiempo de inducción (cambio de temperatura hasta 42 °C) se representa en el eje X, frente al número de recombinantes resistentes al cloranfenicol obtenidos del eje Y.

- 45 La **Fig. 4** es un gráfico que muestra el efecto de la cantidad lineal de ADN sobre la recombinación. La cadena DY330 de *E. Coli* fue inducida a 42 °C durante 15 minutos y se electroporó en las células el ADN que codifica una cassette de resistencia a cloranfenicol (<cat>) lineal a la concentración indicada. La concentración de ADN que codifica <cat> está trazada en el eje X, frente al número de recombinantes resistentes a cloranfenicol obtenidos del eje Y.

- 50 La **Fig. 5** es un gráfico que muestra el efecto de la longitud del brazo homólogo sobre la recombinación en células de *E. coli* inducida para la expresión de las proteínas de recombinación de lambda Beta, Exo y Gam. Se sintetizó una cassette de resistencia a cloranfenicol con los brazos homólogos de la longitud indicada en sus extremos 3' y 5'. La longitud del brazo homólogo se modificó de 0 a 1.000 pares de bases. La longitud de los brazos homólogos está trazada en el eje X, frente al número de recombinantes resistentes a cloranfenicol (escala logarítmica) del eje Y.

La **Fig. 6** es una ilustración lineal del profago lambda defectuoso modificado integrado en el cromosoma de *E. coli* de las células DY380, EL250 y EL350. Esta figura ilustra que los profagos defectuosos empleados para generar BAC contienen los genes X de c857 a int. P<sub>L</sub> y P<sub>R</sub> denotan los promotores izquierdo y derecho de lambda, respectivamente.

Los genes *gam* y *red*, *exo* y *bet* se encuentran bajo el control de  $P_L$ , que es sometido a represión por el represor sensible a la temperatura, *cl857*, a 32°C y a desrepresión a 42°C. *tet* sustituye al segmento de *cro*-*bioK* en las células DY380. La cassette *araC-P<sub>BAD</sub>flpe* o la cassette *araC-P<sub>BAD</sub>cre* sustituye al segmento de *cro*-*bioK* en las células EL250 o EL350, respectivamente. El promotor del operón *araBAD* ( $P_{BAD}$ ), que puede ser inducido mediante L-arabinosa, controla la expresión de los genes *flpe* o *ere*. Las líneas negras gruesas representan al profago, mientras que las finas representan la secuencia de *E. coli*. < > define los extremos de la región *cro*-*bioA* que fue sustituida por *tet*, *araC-P<sub>BAD</sub>flpe*, o *araC-P<sub>BAD</sub>ere*.

La Fig. 7 ilustra una estrategia para generar BAC. Esta figura ilustra la posición relativa del gen *Eno2* en el BAC de 250-kbp, 284H12, completamente secuenciado y los diferentes pasos seguidos para introducir *Cre* en el último exón de *Eno2*. En la cassette diana, los sitios de FRT están señalados por elipses, el gen *kan* por un rectángulo rojo y el gen de fusión GFP $\text{Cre}$  por un rectángulo azul. Los recuadros verdes representan los exones *Eno2*.

La Fig. 8 es un diagrama esquemático que ilustra el uso de la reparación de brecha en fragmentos del subclon de hasta 80 kbp de BAC. La Fig. 8A muestra en forma de flechas negras gruesas y cortas la ubicación de las homologías del extremo 5' en los cebadores de amplificación empleados para amplificar pBR322 para la subclonación mediante reparación de brecha. Cada cebador también contiene 20 segmentos de nt en su extremo 3' para cebar pBR322. Los sitios de clivaje *NoI* y *SaiI* fueron incluidos en estos cebadores para facilitar la liberación de los fragmentos subclonados del segmento principal del plásmido. También se muestra la ubicación de sitios de restricción *SpeI* cercanos a *Eno2* ("S"). Los sitios de restricción *SpeI* no están presentes en el vector pBR322 amplificado lineal. La Fig. 8B muestra un paso intermedio en la reparación de la brecha, emparejando una cassette diana pBR322 amplificada típica y el BAC *Eno2* modificado. Ap, gen de resistencia a *amp*; ORI, origen de replicación.

La Fig. 9 es un diagrama esquemático de un profago  $\lambda$  defectuoso. El profago defectuoso DY380 que expresa las funciones de  $\lambda$  *Exo*, *Beta*, y *Gam* se muestra con los genes bajo control del promotor  $P_L$  y el represor sensible a la temperatura *Cl857*. Se describen las ventajas y desventajas de los sistemas. Los genes que codifican *Cre* y *Flpe* están presentes en otros derivados de DY380 (EL250 y EL350) y sustituyen al gen *tet* como se muestra.

La Fig. 10 es un diagrama esquemático de la clonación mediante recombinación *in vivo*, empleando la reparación de brecha de un plásmido de vector lineal. El método de clonación *in vivo* emplea dos ADN lineales, un vector y un ADN diana, que presentan homologías entre sí en sus extremos. Ambos son electroporados en células competentes para permitir la recombinación y reparación de brecha del plásmido. El vector lineal se produce de forma muy similar al mostrado en la Fig. 8.

La Fig. 11 es un diagrama esquemático de recuperación *in vivo* del ADN de clones de BAC.

La recuperación de segmentos de hasta 80 kbp de BAC en vectores amplificados por PCR ha sido posible empleando técnicas de recombinación divulgadas en el presente. Aquí solamente el plásmido es linealizado y transformado en una célula de recombinación competente que contiene el BAC. La recombinación se produce entre homologías en el extremo del vector lineal y el BAC. Este método elimina la tecnología de clonación estándar del BAC y, lo que es más importante, el segmento cerrado nunca es replicado *in vitro*, reduciendo así la posibilidad de cambios extraños en la secuencia.

La Fig. 12 es un diagrama esquemático de un círculo de ADN de mini-lambda. Este elemento de ADN de lambda no es un plásmido y no presenta ninguna actividad del origen de replicación. Contiene el represor *cJ8S7* de lambda y el operón  $p_L$  que controla el represor. También contiene una cassette que codifica un marcador de fármacos, en este caso los genes *tet*. Este ADN cuando es transformado en la mayoría de las cadenas, incluyendo las cadenas de BAC, hace que la proteína *Int* permita la integración del ADN circular en el sitio de unión de  $\lambda$  en el cromosoma bacteriano y que el represor *cl857* permita la represión de  $p_L$ . El mini-lambda integrado es estable, pero capaz de ser inducido a 42 °C para activar la expresión de *Gam*, *Beta* y *Exo*, a fin de que la recombinación de la célula sea competente.

La Fig. 13 es un diagrama esquemático de la recombinación de ssADN en el genoma. Cuando el ssADN es electroporado en las células, tiene unida la proteína *Beta* y se recombina en el genoma o en un plásmido de BAC mediante homología. La evidencia sugiere que el ssADN unido a *Beta* se anilla en su complemento de ssADN en la horquilla de replicación. La cadena de ADN correspondiente a la generada retardando la síntesis de la cadena es sumamente recombinogénica, lo que sugiere que *Beta* simplemente anilla el ssADN a una brecha causada por la replicación de ADN.

## DESCRIPCIÓN DETALLADA DE REALIZACIONES CONCRETAS

En esta técnica existe la necesidad de métodos para alterar, de forma precisa y eficiente, secuencias genéticas endógenas predeterminadas mediante recombinación homóloga *in vivo*. Independientemente, también existe la necesidad en la técnica de una focalización genética de alta eficiencia, a fin de evitar complejos protocolos de selección *in vitro* o *in vivo*.

La presente invención se explica en las reivindicaciones adjuntas.

Los métodos divulgados en el presente satisfacen una o más de las anteriores necesidades, incluyendo métodos para la clonación de moléculas de ADN empleando la recombinación homóloga mediada por recombinasas de lambda y proteínas similares.

5 Uno de estos métodos de la invención emplea una célula que contiene ADN que codifica los productos genéticos funcionales Exo, Beta y Gam de lambda homólogos, operativamente unidos a un promotor susceptible de desrepresión (p. ej., el promotor PL). La desrepresión del promotor susceptible de desrepresión (p. ej., la inducción de la transcripción desde el promotor pL mediante inactivación de el) induce la expresión de exo, bet y gam, y en algunas realizaciones puede ser activado selectivamente para este fin. Un ácido nucleico (como un polinucleótido que es homólogo a una secuencia de ADN diana) capaz de someterse a recombinación homóloga es introducido en la célula, y las células en  
10 las que se ha producido la recombinación homóloga son seleccionadas o bien localizadas mediante detección directa de las células. En realizaciones concretas, el ácido nucleico introducido en la célula puede ser ADN de cadena doble o ADN con el extremo 5' protuberante.

15 En otras realizaciones concretas, al menos 1 de 5.000 células contiene ADN en el que se ha producido la recombinación homóloga. En otras realizaciones, al menos 1 de 1.000 o 1 de 500 células, 1 de 100 células o 1 de 20 células contiene ADN en el que se ha producido la recombinación homóloga.

20 La célula puede ser una célula eucariota o una célula procariota, como una célula bacteriana, por ejemplo una cadena de *E. coli*. El ADN que codifica las proteínas de recombinación de lambda y el promotor PL puede ser intracromosómico o extracromosómico. De forma similar, la secuencia de ADN diana puede ser intracromosómica o extracromosómica; por ejemplo, la secuencia de ADN diana se puede encontrar en un cromosoma de la célula o un plásmido (incluyendo derivados de colEI, pSCIOI, p15A y vectores transportadores que se replican tanto en las células bacterianas como en las eucariotas), un cromosoma bacteriano artificial, un cromosoma PI artificial, un cromosoma de levadura artificial, un cósmido o similares.

25 En otras realizaciones concretas, el ácido nucleico introducido en la célula puede ser ADN de cadena doble o ADN con el extremo 5' protuberante, y puede incluir un marcador seleccionable positivo o negativo. El ácido nucleico introducido puede alterar la función de una secuencia de ácido nucleico, como un gen en la célula, o añadir un gen al ADN de la célula. La célula puede ser tratada para mejorar la adsorción macromolecular, por ejemplo empleando electroporación, precipitado de ADN con fosfato cálcico, transferencia mediada por liposomas u otros métodos adecuados. En otras realizaciones concretas, el método puede producir una recombinación homóloga que altera la función de un gen en la célula o añade un gen a la célula.

30 En otras realizaciones concretas, la célula puede ser tratada para mejorar la adsorción macromolecular, por ejemplo con electroporación, coprecipitado de ADN con fosfato cálcico, transfección mediada por DEAE-dextrano, transfección mediada por polibreno, microinyección, fusión de liposomas, lipofección, fusión de protoplastos, transferencia mediada por adenovirus inactivado, transferencia mediada por HVJ-liposoma y biolística.

35 Otro de estos métodos que satisface una o más de las mencionadas necesidades incluye métodos que introducen en la célula un ADN capaz de someterse a recombinación homóloga y un polipéptido de unión a ssADN capaz de promover la recombinación homóloga. En realizaciones concretas, el ADN es ssADN o ADN con el extremo 3' protuberante.

40 El polipéptido de unión a ADN de cadena única (ssADN) es un tipo de polipéptido de unión a ADN que media en la recombinación homóloga de la reparación de la brecha de la cadena doble, uniéndose al ssADN o a un extremo 3' protuberante del dsADN y promoviendo la recombinación. Promueve la recombinación anillando el ssADN enlazado a su complemento en la célula. Algunos ejemplos de este polipéptido de unión a ssADN incluyen el lambda Beta, el *E. coli* RecT, el Erf del bacteriófago P22 y Rad52. El polipéptido de unión a ssADN puede ser introducido como un ácido nucleico. Por ejemplo, un ácido nucleico que expresa el polipéptido de unión a ssADN es introducido en una célula, como una célula eucariótica. La expresión del polipéptido de unión a ssADN de un ácido nucleico puede ser inducida, por ejemplo, mediante la activación de un promotor inducible. En realizaciones concretas, el ácido nucleico puede incluir también el lambda exo y gam, y el promotor inducible puede ser el promotor PL de lambda. En otras realizaciones, el  
45 polipéptido de unión a ssADN es introducido en la célula como un polipéptido.

50 La célula empleada en métodos divulgados en el presente puede ser una célula bacteriana, como una cadena de *E. coli*, o una célula eucariota, como una célula de mamífero, o una célula madre, o prácticamente cualquier otro tipo de célula eucariota. El ADN empleado en el método puede ser una única secuencia de oligonucleótido o pueden ser dos o más secuencias solapadas, por ejemplo con más de 10, o más de 20 pares de bases de solapamiento complementario en sus extremos 5' o 3' (en ejemplos específicos del caso del extremo 5', el ácido nucleico incluye secuencias de ácido nucleico exo y bet). El ADN puede comprender un marcador seleccionable y la recombinación homóloga con el ssADN puede conferir un fenotipo seleccionable a la célula. En realizaciones concretas, la célula puede ser tratada para mejorar la adsorción macromolecular, por ejemplo mediante electroporación, coprecipitado de ADN con fosfato cálcico, transferencia mediada por liposomas u otros métodos apropiados. El efecto de la recombinación homóloga puede consistir en alterar la función de un gen en la célula o añadir un gen a la célula.  
55

En ejemplos concretos, el ssADN se emplea en una cantidad de unos 0,01  $\mu\text{M}$  a unos 10 mM; o de unos 0,1  $\mu\text{M}$  a unos 1 mM; o de unos 1  $\mu\text{M}$  a unos 100  $\mu\text{M}$ . En otros ejemplos, el polipéptido de unión a ssADN se emplea en una cantidad de 0,001  $\mu\text{M}$  a unos 0,01  $\mu\text{M}$ , o de unos 0,01  $\mu\text{M}$  a unos 10 mM; o de unos 0,1  $\mu\text{M}$  a unos 1 mM; o de unos 1  $\mu\text{M}$  a unos 100  $\mu\text{M}$ .

5 También se divulgan células bacterianas que pueden ser útiles para la práctica de los métodos divulgados. Estas incluyen células bacterianas que contienen un profago lambda defectuoso del genotipo  $\lambda\text{cl857 } \Delta(\text{cro-bioA})$ . En ejemplos concretos, las células bacterianas pueden tener un marcador seleccionable, como un marcador de resistencia a antibióticos, en sentido ascendente del gen cl857. En algunos ejemplos concretos, las células bacterianas divulgadas incluyen un promotor inducible en sentido ascendente del gen cl857, que puede ser operativamente conectado a un gen que codifica una recombinasa, como flp, flpe o ere, o un gen que codifica fragmentos o variantes funcionales de estas recombinasas. En otros ejemplos concretos, las células bacterianas pueden contener un cromosoma artificial bacteriano, que puede tener un marcador seleccionable, loxp, y/o sitios de FRT. En realizaciones concretas, el marcador seleccionable en el cromosoma artificial bacteriano se puede eliminar mediante una recombinasa. En otros ejemplos concretos, el cromosoma artificial bacteriano puede tener al menos un exón o al menos un intrón de un gen de mamífero.

20 La divulgación también incluye métodos para alterar genes eucariotas mediante la expresión en una célula bacteriana de un gen intracromosómico que codifica una recombinasa operativamente unida a un promotor PL. La célula bacteriana también incluye un gen eucariota extracromosómico o un fragmento de gen (que tiene al menos un intrón o al menos un exón de un gen eucariota). Un ácido nucleico capaz de someterse a recombinación homóloga con el gen eucariota es introducido en la célula bacteriana y el ácido nucleico se somete a recombinación homóloga con el gen eucariota o el fragmento de gen. En una realización concreta, el ácido nucleico se somete a recombinación homóloga con una frecuencia de focalización de al menos 1 de 1.000.

25 En una realización, la recombinasa expresada es una recombinasa de reparación de roturas bicatenarias, como un producto genético de lambda Beta u otra proteína de unión a ADN de cadena única; el lambda Exo o un producto genético de lambda Gam. En otra realización, el gen eucariota extracromosómico o fragmento de gen puede estar ubicado en un cromosoma artificial bacteriano, cromosoma artificial de levadura, cromosoma artificial de PI, plásmido o cósmido. En otra realización más, el gen eucariota o fragmento de gen se obtiene de un organismo mamífero, como un ratón o un ser humano.

30 En varias otras realizaciones, el ácido nucleico que se somete a recombinación homóloga puede codificar una recombinasa, fragmentos o variantes funcionales de una recombinasa o una etiqueta de epítipo.

También se divulgan métodos para alterar el ADN intracromosómico de una célula eucariota. En estos métodos, se introduce un gen eucariota o fragmento de gen alterado en la célula eucariota. El gen eucariota o fragmento de gen introducido se ha alterado mediante recombinación homóloga, empleando los métodos de esta divulgación.

35 Por ejemplo, el ADN extracromosómico que incluye el gen eucariota o fragmento de gen es introducido en una célula bacteriana que tiene un gen intracromosómico que codifica una recombinasa operativamente unida a un promotor susceptible de desrepresión. Posteriormente la célula bacteriana es inducida a expresar la recombinasa. Se introduce una molécula de ácido nucleico capaz de someterse a recombinación homóloga con el gen eucariota o fragmento de gen en la célula bacteriana. El gen eucariota o fragmento de gen se somete a recombinación homóloga con el ácido nucleico, y entonces el gen eucariota o fragmento de gen alterado puede ser aislado e introducido en una célula eucariota.

40 En una realización, el gen eucariota o fragmento de gen introducido en la célula eucariota está ubicado en un cromosoma artificial bacteriano. El gen eucariota o fragmento de gen es capaz de someterse a recombinación homóloga con un gen diana de la célula, alterando así la secuencia de ácido nucleico del ADN intracromosómico de la célula eucariota. En ejemplos específicos, aunque no limitadores, la célula eucariota es una célula de mamífero, una célula madre embrionaria o un cigoto.

También se divulgan mamíferos mutantes en los que se han alterado uno o más de sus genes mediante recombinación homóloga con un cromosoma artificial bacteriano que porta un gen eucariota o fragmento de gen alterado a través de los métodos divulgados. La alteración del gen puede introducir una recombinasa en el mamífero mutante, como una recombinasa de sitio específico.

50 También se divulga en el presente un ADN de lambda movilizable que es aislado como un profago mini-lambda. El ADN de lambda movilizable puede ser transformado en cualquier cadena bacteriana de interés. El ADN de lambda se integra en el cromosoma bacteriano para generar un profago defectuoso que expresa la recombinasa.

55 La presente divulgación proporciona métodos para mejorar la eficiencia de la recombinación homóloga. La divulgación se comprenderá mejor consultando la siguiente explicación de términos utilizados y la descripción detallada de los métodos para la práctica de la invención.

A menos que se señale lo contrario, los términos técnicos se utilizan de acuerdo con el uso convencional. Las definiciones de términos comunes en biología molecular se pueden encontrar en Benjamin Lewin, *Genes V*, publicado por Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, publicado por Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); y Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, publicado por VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

### Explicación de términos empleados

**Extremo 3' protuberante:** dos secuencias de ácido nucleico que al anillarse son parcialmente de doble cadena y parcialmente de cadena única. El extremo o los extremos de cadena única se extienden más allá del segmento de doble cadena en dirección del extremo 5' al extremo 3'.

10 **Extremo 5' protuberante:** dos secuencias de ácido nucleico que al anillarse son parcialmente de doble cadena y parcialmente de cadena única. El extremo o los extremos de cadena única se extienden más allá del segmento de doble cadena en dirección del extremo 3' al extremo 5'.

**Arabinosa:** un azúcar simple de cinco carbonos metabolizado por *E. coli*. En una realización, se utiliza como sustancia química para inactivar la represión e inducir y activar la expresión del promotor pBAD.

15 **Sitio de unión (att):** un sitio específico para la recombinación presente en un fago o cromosoma. Un sitio de unión en lambda se denomina "attP", mientras que un sitio de unión de un cromosoma bacteriano es "attB." La recombinación mediada por integrasa de un sitio attP con un sitio attB permite la integración del profago X en el cromosoma bacteriano.

20 **Cromosoma artificial bacteriano (BAC):** los cromosomas artificiales bacterianos (BAC) han sido construidos para permitir la clonación de largos fragmentos de ADN en *E. coli*, tal y como se describe en O'Conner *et al.*, *Science* 244: 1307-12, 1989; Shizuya *et al.*, *Proc. Natl. Acad. Sci. USA* 89: 8794-7, 1992; Hosoda *et al.*, *Nucleic Acids Res.* 18: 3863-9, 1990; y Ausubel *et al.* Este sistema es capaz de propagar de forma estable ADN de mamífero más de 300 kb. En una realización, un BAC porta los sistemas de división y replicación F que garantizan un reducido número de copias y la segregación fiel del ADN del plásmido en las células hijas. Se pueden clonar grandes fragmentos genómicos en plásmidos de tipo F, lo que hace que resulten útiles para la creación de bibliotecas genómicas.

25 **Beta:** el polipéptido de unión a ssADN de lambda Beta de 28 kDa (y el ácido nucleico que codifica el lambda beta) implicado en la recombinación homóloga de la reparación de roturas bicatenarias. El ADN que codifica Beta (bet) y las cadenas de polipéptidos que presentan actividad de lambda Beta también se denominan "bet" en el presente. Véanse los ejemplos 1 y 14 y las referencias de los mismos para más información. La proteína de lambda Beta se une al ADN de cadena única y promueve la renaturalización de las regiones complementarias de cadena única del ADN. Véase Karakouxis *et al.*, *J. Mol. Biol.* 276: 721-733, 1.998; Li *et al.*, *J. Mol. Biol.* 276: 721-733, 1998; Passy *et al.* *PNAS* 96. 4279-4284, 1999.

35 Los fragmentos y variantes funcionales de Beta incluyen aquellas variantes que mantienen su capacidad para unirse a ssADN y que median en la función de recombinación del lambda Beta tal y como se describe en el presente y en las publicaciones mencionadas como referencia. Se conoce que el gen que codifica Beta puede ser considerablemente mutado sin alterar materialmente la función de unión a ssADN o la función de recombinación del lambda Beta. En primer lugar, se sabe que el código genético se degenera, por lo que diferentes codones codifican los mismos aminoácidos. En segundo lugar, aún cuando se introduzca la mutación de un aminoácido, esta mutación puede ser conservadora y no tener ningún impacto material sobre funciones esenciales del lambda Beta. Véase Stryer, *Biochemistry* 3ª Ed., (c) 1988. En tercer lugar, parte de la cadena polipeptídica del lambda Beta puede ser sometida a delección sin alterar ni eliminar su función de proteína de unión a ssADN ni su función de recombinación. En cuarto lugar, se pueden realizar inserciones o adiciones en la cadena polipeptídica del lambda Beta (por ejemplo, añadiendo etiquetas de epítopos) sin alterar ni eliminar sus funciones esenciales. Ausubel *et al.* 1997.

45 **Biolística:** inserción de ADN en las células incluyendo microproyectiles revestidos de ADN. También conocido como bombardeo de partículas o bombardeo de micropartículas, este planteamiento se describe y define más detalladamente en la Patente estadounidense nº 4.945.050.

**ADNc (ADN complementario):** Un tramo de ADN que carece de segmentos internos que no codifican (intrones) y de secuencias reguladoras que determinan la transcripción. El ADNc se puede sintetizar en el laboratorio mediante transcripción inversa del ADN mensajero extraído de las células.

50 **Cósmido:** vector de clonación construido artificialmente que contiene el gen *cos* del fago lambda. Los cósmidos pueden ser empaquetados en partículas del fago lambda para la infección en *E. coli*. Esto permite la clonación de fragmentos de ADN más largos (de hasta 45 kb) que pueden ser introducidos en huéspedes bacterianos de vectores de plásmidos.

**Cre:** La recombinasa Cre es una recombinasa de sitio específico. Reconoce y se une a sitios específicos llamados *loxP*. Dos sitios *loxP* se recombinan con una eficiencia cercana al 100% en presencia de Cre, permitiendo así que el ADN clonado entre esos dos sitios sea retirado de forma eficiente a través de la recombinación mediada por Cre.

**Promotor susceptible de desrepresión:** Cuando un represor está unido a un promotor susceptible de desrepresión, la transcripción se ve sustancialmente reducida en comparación con la transcripción de un promotor susceptible de desrepresión en ausencia del represor. Al regular la unión del represor, por ejemplo cambiando el entorno, el represor se libera del promotor susceptible de desrepresión y se incrementa la transcripción. En esta definición, un promotor susceptible de desrepresión no precisa de un activador para la transcripción. Un ejemplo específico, a título meramente enunciativo, es el promotor PL, que está regulado por el represor *cl*, pero que no está activado por un activador. (El promotor de la arabinosa no es un promotor susceptible de desrepresión simple, dado que la arabinosa inactiva el represor AraC y lo convierte en un activador.)

En una realización, el promotor susceptible de desrepresión es un promotor susceptible de desrepresión sensible a la temperatura. Por ejemplo, al aumentar la temperatura, el represor se libera del promotor y ya no se puede unir al promotor con una afinidad elevada, y se incrementa la transcripción del promotor. Un ejemplo específico, a título meramente enunciativo, es la inducción de la actividad del promotor PL aumentando la temperatura de la célula. La temperatura aumentada inactiva el represor sensible a la temperatura *cl*, permitiendo que los genes que están operativamente unidos al promotor PL se expresen a niveles incrementados. Cualquiera con conocimientos en la técnica podrá identificar fácilmente un promotor susceptible de desrepresión.

En una realización, un promotor susceptible de desrepresión está autorregulado. Un ejemplo específico, a título meramente enunciativo, de un promotor susceptible de desrepresión autorregulado es PL. Si solamente hay presente una copia de un gen que codifica *cl*, pero muchas copias del promotor PL, la expresión de *cl* se regula al alza, de forma que la transcripción es bloqueada por cualquiera de los promotores PL.

**Recombinación de la reparación de roturas bicatenarias:** Un tipo de recombinación homóloga ejemplificada por las proteínas de recombinación de *lambda* Exo, Beta y Gam, y compartida por muchos otros sistemas de recombinasa. Una rotura bicatenaria es el punto de inicio para la acción concertada de las proteínas de recombinación. Típicamente, una exonucleasa se degrada progresivamente desde los extremos 5' de estos sitios de rotura y el polipéptido de unión a ssADN se une a la cola restante de cadena única del extremo 3', protegiendo y preparando el ADN acoplado para la invasión de la cadena homóloga (Szostak *et al.*, Cell 33: 25-35, 1983; Little, J, Biol. Chem. 242: 679-686, 1967; Carter *et al.*, J. Biol. Chem. 246:2502-2512, 1971; Lindahl *et al.* Science 286:1897-1905, 1999). Algunos ejemplos de polipéptidos de unión a ssADN que se unen a ssADN y/o dsADN con el extremo 3' protuberante y que promueven la reparación de roturas bicatenarias incluyen el *lambda* Beta, RecT de *E. coli*, Erf del fago p22 y Rad52 en diversas células eucariotas, incluyendo las células de levadura y de mamíferos.

**Electrocompetente:** células capaces de adsorción macromolecular tras un tratamiento con electroporación.

**Electroporación:** un método para inducir o permitir que una célula adsorba macromoléculas, aplicando campos eléctricos a fin de permeabilizar de forma reversible las paredes celulares. Diversos métodos y aparatos empleados se definen y describen más detalladamente en: Patente estadounidense nº 4.695.547; Patente estadounidense nº 4.764.473; Patente estadounidense nº 4.882.28; y Patentes estadounidenses nº 4.946.793; 4.906.576; 4.923.814; y 4.849.089.

**Célula eucariota:** una célula que tiene un núcleo organizado rodeado por una membrana nuclear. Estas incluyen organismos inferiores, como las levaduras, mohos y similares, así como células de organismos multicelulares, como invertebrados, vertebrados y mamíferos. Incluyen diversos tipos de tejidos, tales como: célula endotelial, célula de músculo liso, célula epitelial, hepatocito, célula procedente de la cresta neural, célula tumoral, célula hematopoyética, célula inmunitaria, célula T, célula B, monocito, macrófago, célula dendrítica, fibroblasto, queratinocito, célula neuronal, célula glial, adipocito, mioblasto, miocito, condroblasto, condrocito, osteoblasto, osteocito, osteoclasto, célula secretora, célula endocrina, oocito y espermatozoido. Estos tipos de células se describen en textos de histología estándar, tales como McCormack, *Introduction to Histology*, (c) 1984 de J.P. Lippincott Co.; Wlieter *et al.*, eds., *Functional Histology*, 2ª Ed, (r.) 1987 de Churchill Livingstone; Fawcett *et al.*, eds., *Bloom y Fawcett: A Textbook of Histology*, (c) 1984 de William y Wilkins.

**Exo:** la exonucleasa de *lambda* (y el ácido nucleico que codifica la proteína de la exonucleasa) implicada en la recombinación homóloga de reparación de roturas bicatenarias. Véase el Ejemplo 1 y las referencias del mismo para más información.

**Gam:** una proteína *lambda* (y el ácido nucleico que codifica Gam) implicada en la recombinación homóloga de reparación de roturas bicatenarias. Se cree que inhibe la actividad de la nucleasa celular, como la codificada por el sistema *sbcC* y *recBCD* de *E. coli*. Véanse los ejemplos 1, 7 y 14, y las referencias de los mismos para más información. La función de Gam, cuando se expresa en la célula, es sumamente tóxica para la célula e impide el crecimiento. Por esta razón, siempre es necesario un estricto control sobre su expresión. Como se ha descrito en el presente, PL y *cl* 857 son capaces de regular la expresión de Gam.

Fragmentos y variantes funcionales de Exo y Gam: como se señala en el caso de Beta (véase "Fragmentos y variantes funcionales de Beta"), se reconoce que los genes que codifican Exo o Gam pueden ser considerablemente mutados sin



alterar de forma material su función, debido a la degeneración del código genético, sustituciones conservadoras de aminoácidos, deleciones o inserciones no críticas, etc. A menos que el contexto deje claro lo contrario, el término lambda Exo, Exo o exonucleasa de lambda incluirá la exonucleasa de lambda nativa, y todos los fragmentos y variantes de la exonucleasa de lambda.

5 **Extracromosómico:** no incorporado al cromosoma o los cromosomas de una célula. En el contexto de los ácidos nucleicos, extracromosómico indica un oligonucleótido de ADN que no está covalentemente incorporado al cromosoma o los cromosomas de una célula. Intracromosómico se refiere al material, como un oligonucleótido, que está incorporado al cromosoma o los cromosomas de una célula, como un oligonucleótido de ADN incorporado covalentemente al ADN cromosómico de una célula.

10 **Brazo homólogo:** nucleótidos en el extremo 5' o 3', o cerca de estos extremos, de un polinucleótido que son idénticos o similares en secuencia al ácido nucleico diana de una célula y capaces de mediar la recombinación homóloga con el ácido nucleico diana. Los brazos homólogos también se denominan brazos de homología. En una realización, un brazo de homología incluye al menos 20 bases de una secuencia homóloga a un ácido nucleico de interés. En otra realización, el brazo de homología incluye al menos 30 pares de bases de una secuencia homóloga a un ácido nucleico de interés.  
15 En otra realización más, un brazo de homología incluye al menos 40 pares de bases de una secuencia homóloga a un ácido nucleico de interés. En otra realización más, un brazo de homología incluye entre unos 50 y unos 100 pares de bases de una secuencia homóloga a un ácido nucleico de interés.

**Recombinación homóloga:** un intercambio de segmentos de polinucleótidos homólogos en cualquier lugar a lo largo de dos moléculas de ácido nucleico.

20 **Célula hospedadora:** una célula que se emplea en técnicas de laboratorio, como la clonación de ADN, para recibir moléculas de ácido nucleico exógeno. En una realización, se emplea una célula hospedadora para mantener o permitir la reproducción de un vector, o para facilitar la manipulación de moléculas de ácido nucleico *in vitro*. Una célula hospedadora puede ser una célula procariota o eucariota.

25 **Transferencia de genes mediada por HVJ:** un método de transferencia macromolecular a las células empleando el virus hemaglutinante del Japón y liposomas, tal y como se describe en Morishita *et al.*, J. Clin. Invest. 91: 2580-2585, 1993; Morishita *et al.*, J. Clin. Invest. 94: 978-984, 1994.

**Promotor inducible:** un promotor cuya actividad puede ser incrementada (o que puede ser sometido a desrepresión) por algún cambio en el entorno de la célula. Ejemplos de promotores inducibles abundan en la naturaleza y son muchos los cambios ambientales u hormonales que pueden activarlos o someterlos a represión.

30 **Aislado:** un componente biológico "aislado" (como un ácido nucleico o una proteína) ha sido sustancialmente separado o purificado de otros componentes biológicos de la célula del organismo en el que el componente se encuentra naturalmente presente, es decir otro ARN y ADN cromosómico y extracromosómico, y proteínas. De este modo, los ácidos nucleicos y proteínas que han sido "aislados" incluyen ácidos nucleicos y proteínas purificados mediante métodos de purificación estándar. El término también incluye ácidos nucleicos y proteínas preparados mediante  
35 expresión recombinante en una célula hospedadora, así como ácidos nucleicos sintetizados químicamente.

**Vector de plásmido lineal:** una secuencia de ADN (1) que contiene un origen de replicación de un plásmido bacteriano, (2) que tiene un extremo 5' y 3' libre, y (3) que es capaz de circularizar y replicarse como un plásmido bacteriano uniendo sus extremos 5' y 3' libres. Algunos ejemplos de vectores de plásmidos lineales incluyen el vector pBluescript linealizado y los vectores pBR322 linealizados aquí descritos.

40 **Lipofección:** El proceso de transferencia macromolecular a las células empleando liposomas. Véase la Patente estadounidense nº 5.651.981.

**Mini lambda:** Un derivado de lambda ( $\lambda$ ) en el que la mayoría de los genes líticos virales, incluyendo aquellos necesarios para la replicación y la lisis, están sometidos a deleción. Un mini-lambda mantiene las funciones Red (Beta, Exo y Gam) para la recombinación homóloga y mantiene las funciones de integración/excisión (p. ej., att, integrasa (int) y excisionasa (xis)) para insertar y separar su ADN del cromosoma.  
45

**Ácido nucleico:** un polímero de desoxirribonucleótido o ribonucleótido en forma de cadena doble o única, incluyendo análogos conocidos de nucleótidos naturales, a menos que se indique lo contrario.

**Oligonucleótido:** un ácido nucleico de cadena única que oscila entre 2 y unas 500 bases de longitud, por ejemplo polinucleótidos que contienen al menos 20 o 40 nucleótidos. Los oligonucleótidos suelen ser sintéticos pero también se  
50 pueden producir a partir de polinucleótidos naturalmente presentes.

**Operativamente unido:** Una primera secuencia de ácido nucleico está operativamente unida a una segunda secuencia de ácido nucleico cuando la primera secuencia de ácido nucleico mantiene una relación funcional con la segunda secuencia de ácido nucleico. Por ejemplo, un promotor está operativamente unido a una secuencia de codificación,

cuando el promotor afecta a la transcripción o expresión de la secuencia de codificación. Por lo general, las secuencias de ADN operativamente unidas se encuentran contiguas y, cuando es necesario para unir dos regiones de codificación de proteínas, en el mismo marco de lectura.

5 **Cromosoma artificial fagémido:** también denominado cromosoma artificial P1. Un tipo de cromosoma artificial que permite la clonación estable de fragmentos de ADN muy largos. Descrito más detalladamente en Shepherd, *et al.*, Proc. Natl. Acad. Sci. USA 92: 2629,1994; Iannou *et al.*, Nature Genetics, 6: 84-89,1994.

10 **Sistemas de recombinación basados en fagos:** bacterias como *E. coli* codifican sus propios sistemas de recombinación homóloga, que se emplean en la reparación de daños del ADN y para mantener un cromosoma funcional. La virus o fagos que habitan las bacterias a menudo portan sus propias funciones de recombinación. El fago  $\lambda$  porta el sistema de recombinación Red. Estos sistemas de fagos pueden funcionar con las funciones de recombinación bacterianas o independientemente de ellas.

15 **Promotor PL:** el principal promotor hacia la izquierda del bacteriófago lambda. Una vez que el ADN de lambda está incorporado al cromosoma bacteriano, la transcripción de este promotor se ve sustancialmente sometida a represión por el represor cl. Tras la inactivación del represor cl, por ejemplo mediante golpe de calor de un mutante sensible a la temperatura, la transcripción del promotor PL queda activada, conduciendo a la expresión de genes de lambda. Véase la Fig. 1; Sambrook *et al.* Ch. 2, Stryer, y Court y Oppenheim.

**Plásmido:** moléculas de ADN extracromosómico, que se replican de forma autónoma, distintas del genoma bacteriano normal y no esenciales para la supervivencia de la célula bacteriana en condiciones no selectivas.

20 **Polinucleótido:** una secuencia de ácido nucleico de doble cadena o de cadena única de cualquier longitud. Por tanto, un polinucleótido incluye moléculas que tienen 15, 50, 100, 200 nucleótidos de largo (oligonucleótidos) y también nucleótidos tan largos como un ADNc de longitud completa.

**Polipéptido:** cualquier cadena de aminoácidos, independientemente de la modificación de la longitud o postranslacional (p. ej., glicosilación o fosforilación).

25 **Procariota:** Célula u organismo que carece de un núcleo discreto estructuralmente rodeado por una membrana y de otros compartimentos subcelulares.

**Sondas y cebadores:** Una sonda de ácido nucleico comprende un ácido nucleico aislado unido a una etiqueta detectable o molécula informante. Las etiquetas típicas incluyen isótopos radiactivos, ligandos, agentes quimioluminiscentes y enzimas. Los métodos para etiquetar y guiar en la selección de las etiquetas apropiadas para diversos fines se recogen, por ejemplo, en Sambrook *et al.* (1989) y Ausubel *et al.* (1997).

30 Los cebadores son ácidos nucleicos cortos, preferiblemente oligonucleótidos de ADN de 15 nucleótidos de longitud o más. Los cebadores pueden ser anillados a una cadena de ADN diana complementaria mediante hibridación de ácido nucleico para formar un híbrido entre el cebador y la cadena de ADN diana. El extremo 3'-hidroxilo del cebador se puede extender a lo largo de la cadena de ADN diana mediante el uso de una enzima de ADN polimerasa. Los pares de cebadores (a cualquier lado de la secuencia de ácido nucleico diana) pueden ser empleados para la amplificación de una secuencia de ácido nucleico, p. ej., mediante la reacción en cadena de la polimerasa (PCR) u otros métodos de amplificación de ácido nucleico conocidos en el campo.

40 Los métodos para preparar y utilizar sondas y cebadores se describen, por ejemplo, en Sambrook *et al.* (1989), Ausubel *et al.* (1987). Los pares de cebadores de PCR se pueden obtener de una secuencia conocida, por ejemplo, utilizando programas informáticos diseñados para tal fin, como Primer (Versión 0.5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, MA). En las condiciones apropiadas, la especificidad de una sonda o cebador concreto aumenta con su longitud. Así, a fin de obtener una mayor especificidad, se pueden seleccionar sondas y cebadores que comprendan 20, 25, 30, 35, 40, 50 o más nucleótidos consecutivos de una secuencia genética o ADNc relacionado.

45 **Purificado:** El término purificado no exige una pureza absoluta, sino que ha de entenderse más bien como un término relativo. Así, por ejemplo, una preparación de lambda Beta o un polipéptido de unión a ssADN purificado es aquel en el que la proteína está más enriquecida que la proteína en su entorno natural dentro de una célula. Preferiblemente, una preparación de lambda Beta es purificada de forma que el polipéptido represente al menos el 50% del contenido total de proteína del preparado.

50 **RecA:** La proteína RecA es una proteína central que tiene una actividad como en la función de recombinación de *E. coli*. Se encuentran homólogos en todos los demás organismos. La proteína RecA permite que dos ADN homólogos se encuentren entre sí entre ADN no homólogos, para después intercambiar o transferir cadenas entre ellos. Este intercambio se produce mediante la unión de RecA a una región de cadena única de uno de los ADN y empleando esa cadena para buscar su homólogo del dsADN, uniéndose al dsADN y provocando que la cadena única se empareje con su complemento en el dsADN, desplazando en última instancia la cadena idéntica del dúplex. Esta transferencia de cadenas genera un intermedio clave en el proceso de recombinación mediado por RecA.

**Genes *recE* y *recT*, y el profago *Rac*:** *E. coli* y otras bacterias contienen en sus cromosomas remanentes de virus. Estos virus o profagos son en su mayoría defectuosos y pueden contener solamente algunos genes del virus original. En *E. coli*, un profago defectuoso se denomina *Rac*. Dos genes, *recE* y *recT* del profago *Rac*, codifican funciones de recombinación homóloga. Normalmente estos genes no se pronuncian, pero la mutación de *sbcA* activa su expresión constitutiva. Así, el mutante de *sbcA* está activo para la recombinación.

**Recombinasas:** proteínas que, cuando se incluyen con un polinucleótido diana exógeno, proporcionan un incremento mensurable de la frecuencia de recombinación entre dos o más oligonucleótidos que son al menos parcialmente homólogos.

**Marcadores de selección o marcadores seleccionables:** secuencias de ácido nucleico que, tras la expresión intracelular, son capaces de conferir un marcador de selección positivo o negativo o una característica fenotípica a la célula que expresa la secuencia. El término "marcador de selección" o "marcador seleccionable" incluye tanto los marcadores de selección positivos como negativos. Un "marcador de selección positiva" es una secuencia de ácido nucleico que permite la supervivencia de células que contienen el marcador de selección positiva en condiciones de crecimiento que matan o impiden el crecimiento de las células que carecen del marcador. Un ejemplo de marcador de selección positivo es una secuencia de ácido nucleico que promueve la expresión del gen de resistencia a la neomicina. Las células que no contienen el gen de resistencia a la neomicina son seleccionadas mediante la aplicación de G418, mientras que las células que expresan el gen de resistencia a la neomicina no se ven afectadas por G418 (selección positiva). Un "marcador de selección negativa" es una secuencia de ácido nucleico que mata, evita el crecimiento o selecciona de otro modo las células que contienen el marcador de selección negativa, normalmente tras la aplicación de un agente exógeno apropiado. Un ejemplo de un marcador de selección negativa es una secuencia de ácido nucleico que promueve la expresión del gen de la timidina quinasa del virus del herpes simple (HSV-TK). Las células que expresan HSV-TK son seleccionadas mediante la aplicación de ganciclovir (selección negativa), mientras que las células que no expresan el gen prácticamente no se ven afectadas por el ganciclovir. Los términos y métodos se definen y explican más detalladamente en la Patente estadounidense nº 5.464.764.

**Fenotipo seleccionable:** Una célula con un fenotipo seleccionable es aquel que expresa un marcador de selección positivo o negativo.

**Identidad de la secuencia:** la similitud entre dos secuencias de ácido nucleico o dos secuencias de aminoácidos se expresa en términos de la similitud entre las secuencias, también denominada identidad de la secuencia. A menudo la identidad de la secuencia se mide en forma de identidad porcentual (o similitud u homología); cuanto más elevado es el porcentaje, más similares son las dos secuencias.

Los métodos de alineación de secuencias para la comparación son bien conocidos en la técnica. Diversos programas y algoritmos de alineación se describen en: Smith y Waterman. *Adv. Appi: Math.* 2:482, 1981; Needleman y Wunscli, *J. Mol. Bio.* 48:443, 1970; Pearson y Lipman. *Methods in Molec. Biology* 24: 307-331, 1988; Biggins y Sharp. *Gene* 73:237-244, 1988; Higgins y Sharp, *CABIOS* 5:151-1.53, 1-989; Capet *et al.*, *Nucleic Acids Research* 16:10881-90, 1988; Huang *et al.*, *Computer Applications in Biosciences-X*:155-65,1992; y Pearson *et al.*, *Methods in Molecular Biology* 24:307-31,1.994. Allschul *et al.* (1994) presenta un análisis detallado de los métodos de alineación de la secuencia y cálculos de homología.

La herramienta de búsqueda de alineación local básica de NCBI (BLAST) (Allschul *et al.*, *Mol. Biol.* 215:403-410, 1990) está disponible en varias fuentes, incluyendo el Centro Nacional de Información Biológica (NCBI, Bethesda, MD) y en Internet, para su uso en relación con los programas de análisis de secuencias *blastp*, *blastn*, *blastx*, *tblastn* y *tblastx*. En el sitio web del NCBI, se puede acceder al mismo, así como a una descripción de cómo determinar la identidad de la secuencia empleando este programa.

Los homólogos de los productos genéticos de *lambda* Beta, *Exo* y *Gam* de acuerdo con la invención y las proteínas de unión a ssADN poseen al menos un 60% de identidad de la secuencia contada con respecto a la alineación de la longitud completa con la secuencia de aminoácidos de la proteína que se va a evaluar (es decir, *lambda* Beta, *Exo* o *Gam*, proteína de unión a ssADN como P22 Erf; *RecT*, y *Rad52*) empleando el NCB1 Blast 2.0, con Blastp "con huecos" (o gaps) establecido conforme a los parámetros por defecto. Para las comparaciones de secuencias de aminoácidos de más de 30 aminoácidos, la función de secuencias de Blast 2 se emplea utilizando la matriz BLOSUM62 por defecto configurada conforme a los parámetros por defecto (penalización de existencia de gap de 11 y penalización de gap por residuo de 1). Al alinear péptidos cortos (de menos de unos 30 aminoácidos), la alineación se deberá realizar empleando la función de secuencias de Blast 2, usando la matriz PAM30 establecida conforme a los parámetros por defecto (penalizaciones por gap abierto de 9 y por extensión de gap de 1). Las proteínas con una similitud incluso mayor a la secuencia de referencia registrarán mayores identidades porcentuales al analizarse con este método, como al menos un 70%, al menos un 75%, al menos un 80%, al menos un 90%, al menos un 95%, al menos un 98% o al menos un 99% de identidad de la secuencia. Cuando se compara menos de la secuencia entera para determinar la identidad de la secuencia, los homólogos típicamente poseerán al menos un 75% de identidad de la secuencia en ventanas cortas de 10-20 aminoácidos, y podrán poseer identidades de la secuencia de al menos el 85% o al menos el 90% o el 95%, dependiendo de su similitud con la secuencia de referencia. Los métodos para determinar la identidad de la secuencia en ventanas cortas se describen en el sitio web del NCBI.

Una persona con conocimientos en la técnica apreciará que estos rangos de identidad de la secuencia se proporcionan únicamente a título orientativo; es perfectamente posible obtener homólogos fuertemente significativos u otras variantes que no se enmarquen dentro de los rangos proporcionados.

5 **ADN de cadena única (ssADN) y ADN de doble cadena (dsADN):** ssADN es el ADN de una cadena de polinucleótidos única; las bases del ADN no están implicadas en el emparejamiento de bases de Watson-Crick con otra cadena de polinucleótidos. El dsADN implica dos o más cadenas de polinucleótidos complementarias, en el que las dos cadenas de polinucleótidos se someten al menos parcialmente al emparejamiento de bases de Watson-Crick entre sí. Cabe señalar que un segmento de ADN puede ser parcialmente ssADN y parcialmente dsADN, por ejemplo si existen huecos en una cadena de polinucleótidos de un segmento de dsADN o si tiene el extremo 5' o 3' protuberante. El ssADN y dsADN pueden contener análogos de nucleótidos, nucleótidos no naturalmente presentes o sintéticos, biotina, o etiquetas de epítomos o fluorescentes. El ADNss o dsADN pueden ser etiquetados; las etiquetas típicas incluyen isótopos radiactivos, ligandos, agentes quimioluminiscentes y enzimas.

10 **Recombinasa de sitio específico:** una recombinasa cuya actividad se limita al ADN de una secuencia específica. Algunos ejemplos incluyen las recombinasas Cre, FLP y FLPE. El lambda Int es específico de los sitios att. Las recombinasas Cre son de sitio específico para los sitios de recombinación **loxP** de una secuencia de ADN, mientras que las recombinasas FLP y FLPE son de sitio específico para los sitios de recombinación **FRT**. Un sitio de recombinación es una secuencia de ácido nucleico específicamente reconocida por una recombinasa. Por ejemplo, la recombinasa Cre se une específicamente a un sitio de recombinación **loxP** y de este modo induce la recombinación.

15 **Secuencia de ácido nucleico diana:** El segmento de ácido nucleico que es objetivo de la recombinación homóloga. Típicamente, se trata de un segmento de ADN cromosómico o extracromosómico en una célula. El ADN extracromosómico que contiene las secuencias de ácido nucleico diana puede incluir ADN episomal, ADN plásmido, cromosoma artificial bacteriano, cromosomas artificiales fagémidos, cromosomas artificiales de levadura, cósmidos y similares. Por lo general, la secuencia de ácido nucleico diana contiene un gen o fragmento de gen que mutará de alguna forma tras la recombinación homóloga. Algunos ejemplos de secuencias de ácido nucleico diana incluyen las secuencias de ADN que rodean a la mutación de tyr145 UAG ámbar de galK, tal y como se describe en Yu *et al.*, PNAS 97: 5798-5983,2000, y en el Ejemplo 3 de la presente solicitud; el segundo exón del gen *hox 1.1* de ratón, tal y como se describe en la Patente estadounidense nº 5.464.764; la mutación del gen hemoglobina S humano tal como se describe en el Ejemplo 15 de la presente solicitud.

20 **Frecuencia de focalización:** La frecuencia con la que una secuencia de ácido nucleico diana se somete a recombinación homóloga. Por ejemplo, el ADN extracromosómico es introducido en una célula eucariota. El ADN extracromosómico tiene secuencias capaces de someterse a recombinación homóloga con una secuencia de ADN intracromosómico diana. Después de introducir el ADN extracromosómico y de permitir que se produzca la recombinación homóloga, se puede determinar el número total de células y el número de células que contienen la secuencia de ADN diana alterada mediante recombinación homóloga. La frecuencia de focalización es el número de células que tiene la secuencia de ADN diana alterada, dividida por el número total de células. Por ejemplo, si hay un número total de un millón de células y 1.000 de estas células contienen la secuencia de ADN diana alterada, entonces la frecuencia de focalización será de 1 de cada 1.000 o de  $10^{-3}$ .

25 **Transformado:** Para los fines del presente, el término transformación abarca todas las técnicas por las que una molécula de ácido nucleico podría ser introducida en una célula, incluyendo la transfección con vectores virales, la transformación con vectores de plásmidos y la introducción de ADN (incluyendo ADN unido a la proteína Beta) mediante electroporación, lipofección y biolística.

30 **Sentido ascendente:** se refiere a secuencias de ácido nucleico que preceden a los codones que se transcriben a un ARN de interés. De forma similar, sentido descendente se refiere a secuencias de ácido nucleico que siguen a los codones que se transcriben a un ARN de interés.

35 **Variantes de secuencias de aminoácidos y de ácido nucleico:** La producción del lambda Beta, Exo o Gam, o de otro polipéptido de unión a ADNss, se puede realizar de diversas maneras. Las secuencias de ADN que codifican la proteína, o un fragmento de la proteína, pueden ser modificadas de forma que permitan que la proteína se exprese en células eucariotas, bacterias, insectos y/o plantas. A fin de conseguir esta expresión, la secuencia de ADN puede ser modificada y operativamente unida a otras secuencias reguladoras. El producto final, que contiene las secuencias reguladoras y el ácido nucleico que codifica la proteína terapéutica, está operativamente unido a un vector, permitiendo el mantenimiento estable en una célula. Este vector puede introducirse entonces en las células eucariotas, las bacterias, los insectos y/o la planta. Una vez en el interior de la célula, el vector permite que se produzca la proteína.

40 **Cualquiera con conocimientos en la técnica apreciará que el ADN se puede alterar de muchas maneras, sin afectar a la actividad biológica de la proteína codificada. Por ejemplo, se puede emplear la PCR para producir variaciones en la secuencia de ADN que codifica el lambda Beta, Exo o Gam, u otras proteínas de unión a ADNss. Estas variantes pueden ser variantes que están optimizadas para la preferencia por el codón en una célula hospedadora que se empleará para expresar la proteína, u otros cambios en la secuencia que facilitan la expresión.**

Se pueden producir dos tipos de variantes de secuencias de ADNc. En el primer tipo, la variación en la secuencia de ADNc no se manifiesta como un cambio en la secuencia de aminoácidos del polipéptido codificado. Estas variaciones táctas son simplemente un reflejo de la degeneración del código genético. En el segundo tipo, la variación de la secuencia de ADNc no produce un cambio en la secuencia de aminoácidos de la proteína codificada. En estos casos, la variante de la secuencia de ADNc produce una variante de la secuencia polipeptídica. A fin de preservar la identidad funcional e inmunológica del polipéptido codificado, estas sustituciones de aminoácidos son idealmente conservadoras en las regiones altamente conservadas. Las sustituciones conservadoras sustituyen un aminoácido por otro aminoácido similar por lo que respecta al tamaño, la hidrofobicidad, etc. Fuera de las regiones altamente conservadas, se pueden realizar sustituciones no conservadoras más fácilmente sin afectar a la función de la proteína. Algunos ejemplos de sustituciones conservadoras se muestran en la Tabla 1 siguiente.

**Tabla 1**

15

Residuo original	Sustitución conservadora
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu, Val
Leu	Ile; Val
Lys	Arg; Gln; Glu
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Las variaciones en la secuencia de ADNc que provocan cambios en los aminoácidos, sean conservadoras o no, deberán ser minimizadas, a fin de preservar la identidad funcional e inmunológica de la proteína codificada. La identidad inmunológica de la proteína puede ser evaluada determinando si es reconocida por un anticuerpo de la proteína; una variante que es reconocida por este anticuerpo es inmunológicamente conservada. Algunos ejemplos concretos de variantes de secuencias de ADNc introducen no más de 20 y menos de 10 sustituciones de aminoácidos, en el polipéptido codificado. Las variantes de secuencias de aminoácidos pueden ser, por ejemplo, al menos un 80, un 90 o incluso un 95% idénticas a las secuencias de aminoácidos nativas.

**Cromosoma artificial de levadura (YAC):** Un vector empleado para clonar fragmentos de ADN (de hasta 400 kb); se construye a partir de secuencias teloméricas, centroméricas y de origen de replicación necesarias para la replicación en células de levadura (véase Ausubel et al.).

*Uso del sistema de recombinación Red codificado por lambda en la recombinación mediada por un profago defectuoso*

El bacteriófago  $\lambda$  contiene un sistema de recombinación homólogo denominado Red, que es funcionalmente análogo al sistema de recombinación RecET de Rac. Al igual que RecET, la recombinación de Red precisa dos genes: redo, o exo, que es análogo a recE, y red $\beta$  (o bet), que es análogo a recT). Exo es una exonucleasa del extremo 5'-3' que actúa progresivamente sobre el dsADN lineal. Beta se une al ssADN protuberante creado por Exo y estimula el anillamiento a una cadena complementaria, pero no puede promover la invasión directa de la cadena ni intercambiarse por sí misma. Las funciones de recombinación de Exo y Beta son de nuevo asistidas por la Gam codificada por el fago  $\lambda$ , que inhibe la actividad de RecBCD de la célula hospedadora<sup>47</sup>. Los eventos de recombinación mediados por  $\lambda$  Red son entre 10 y 1.000 veces más eficientes que los observados en las cadenas recBC, sbcBC o recD. Dado que la recombinación

homóloga se incrementa drásticamente por la adición al hospedador de funciones de la proteína codificada por el fago, este procedimiento es generalmente aplicable a cualquier cadena de *E. coli* y también a otras especies de bacterias.

5 En el presente se divulga un sistema basado en el profago  $\lambda$  defectuoso para la recombinación mediada por Red (véase la Fig. 2). En este sistema, Gam, Beta y Exo están codificados por un profago lambda defectuoso, que está integrado en el cromosoma de *E. coli* de una célula bacteriana (p. ej., *E. coli*) (véase la Fig. 6 y la Fig. 9). La expresión de Gam, Beta y Exo se encuentra bajo el estricto control de un promotor susceptible de desrepresión. En el ejemplo mostrado, el promotor PL susceptible de desrepresión se encuentra bajo el control del represor  $\text{cI857 } \lambda$  sensible a la temperatura. A 32°C, cuando el represor está activo, la expresión del promotor PL y estos genes resulta indetectable. No obstante, cuando las células se pasan a 42 °C durante un breve período de 15 minutos, el represor es inactivado y los genes se expresan a niveles muy elevados. En contraste, los promotores que pueden ser activados, que están presentes en los plásmidos, resultan notablemente difíciles de controlar, y las funciones de Red y Gam se expresarían incluso en ausencia del inductor, como la arabinosa. La expresión a bajo nivel de Gam causa un defecto de RecBCD, una condición que resulta en la inestabilidad del plásmido y la pérdida de viabilidad de la célula.

15 La regulación estricta que proporciona el sistema del profago, unida al hecho de que el promotor  $\lambda$ , que impulsa la expresión de Gam y Red, es un promotor muy sólido, hacen posible conseguir frecuencias de recombinación que son al menos 50 veces superiores a las obtenidas con el sistema basado en un plásmido empleado anteriormente (véase Muyrers *et al*, Nucleic Acids Res 27: 1555-1557, 1999; Yu, D. *et al*. Proc Natl Acad Sci USA 97: 5978-5983, 2000), y varias órdenes de magnitud superior a las de las cadenas anteriormente descritas en las que se ha estudiado la recombinación lineal. El profago en sí es genéticamente estable, a diferencia de los plásmidos, y no confía en la presencia de selección de fármacos para el mantenimiento.

20 La FIG. 2 ilustra el diseño de cebadores para la amplificación de una cassette de recombinación de dsADN, y una estrategia para generar moléculas de ADN recombinante y sustitución de genes. Los pasos y ventajas de este sistema se explican a continuación.

25 La tecnología de ADN recombinante o ingeniería genética clásica ha confiado principalmente en la presencia de sitios de clivaje de enzimas de restricción para el clivaje juicioso del ADN y el uso de ligasa de ADN para enlazar covalentemente diferentes ADN al fin de generar los recombinantes deseados. La capacidad para practicar la ingeniería genética se ha simplificado mediante la reacción en cadena de la polimerasa (PCR), que permite incorporar sitios de restricción a los productos de la PCR lineal, permitiendo así un posicionamiento más preciso de estos sitios. No obstante, toda la tecnología de ingeniería genética falla cuando los vehículos de clonación y la diana contienen cientos de kilobases de ADN. Algunos ejemplos incluyen el cromosoma bacteriano y grandes clones de BAC genómico. Incluso los sitios de la enzima de restricción poco habituales se encuentran frecuentemente en estas moléculas de ADN de gran tamaño, lo que hace que el esfuerzo por utilizar sitios únicos resulte infructuoso. Por otra parte, la manipulación *in vitro* de ADN lineales de esta longitud también resulta extremadamente difícil. Así pues, una vez que la tecnología de clonación de BAC de gran tamaño resultó posible en *E. coli*, la modificación de los clones de BAC se convirtió en el principal problema. Inicialmente se adaptó una combinación de la tecnología de ingeniería genética con las técnicas de recombinación homóloga clásicas, para modificar los grandes clones genómicos. La recombinación homóloga clásica en *E. coli* depende de importantes (>500 bp) tramos de homología entre los ADN.

35 La Fig. 2A ilustra un protocolo de ingeniería genética típico para modificar una diana en un clon de BAC con una cassette y compara esa tecnología con la tecnología de recombinación recientemente desarrollada, que emplea funciones de recombinación de fagos especiales. En general, la ingeniería clásica exige seguir numerosos pasos y el producto final no se puede modificar con tanta precisión como con la nueva tecnología de recombinación. El avance crítico en la metodología de recombinación es el uso de funciones de recombinación de fagos que generan productos de recombinación empleando homología de 50 bp (o menos). Señala que las homología diana de las Fig. 2A y 2B están representadas por los recuadros de franjas. En el método señalado en la Fig. 2A, esos recuadros deben tener al menos 500 bp de largo, mientras que en el método señalado en la Fig. 2B, solamente necesitan tener entre unos 40 y unos 50 bp de largo.

Los pasos de ingeniería genética para generar recombinante de BAC incluyen:

- Escisión del ADN de la cassette con una enzima de restricción.
- 50 ■ Escisión de la diana del plásmido con una enzima de restricción.(Cabe señalar que en este caso el vector ha sido previamente modificado para que contenga fragmentos diana.)
- Unir la cassette al plásmido mediante la ligasa de ADN.
- Transformar el ADN en las células.
- Selección de clones resistentes a fármacos (drug<sup>R</sup>) y aislamiento del plásmido.
- Verificación de la cassette clonada.

- Transformación de la cassette clonada en la cadena de BAC.
- Diversos pasos de recombinación -> -> -> para introducir la cassette clonada en el BAC.

Los pasos de recombinación para generar recombinantes de BAC incluyen:

- Generación de dos cebadores (flechas blancas y negras).
- 5 ■ Generación de cassette amplificada por PCR con homologías en los flancos (en esta ilustración, los ejemplos de segmentos de homología representados con franjas tienen 50 bp de longitud).
- Inducción de funciones de recombinación de fagos en la cadena de BAC o introducción de ADN de BAC en la cadena que presenta funciones de recombinación.
  - Transformación de las células que contienen BAC y funciones de recombinación con la cassette de PCR.
- 10 ■ Generación de recombinante *in vivo*.
- Detección de recombinante:
    - Detección mediante selección o contraselección.
    - Detección mediante análisis directo (hibridación de la colonia),
    - Detección mediante la localización de una etiqueta en el ácido nucleico (p. ej., cuando el ADN incluye un aducto de ADN o un marcador como la biotina).
- 15

Como se ha divulgado en el presente, el profago  $\lambda$  defectuoso también ha sido transferido a la cadena hospedadora del BAC, DH10B, para que pueda emplearse para la ingeniería de BAC. La cadena de DH10B modificada, denominada DY380, puede ser transformada con ADN de BAC a eficiencias de  $10^{-6}$  a  $10^{-4}$ . La utilidad de células de DY380 para la ingeniería de BAC ha sido demostrada mediante la introducción de un BAC de ratón de 250 kbp que contiene el gen específico neuronal enolasa 2 (Eno2) en las células DY380 mediante electroporación y posteriormente modificando el BAC mediante la introducción de una cassette de focalización que expresa Cre en el extremo 3' del gen Eno2 empleando la recombinación de Red (véase el Ejemplo 20). La cassette de focalización fue amplificada por PCR a partir de un plásmido de muestra empleando cebadores de 63 nucleótidos (nt) quiméricos. Los 21 nt del extremo 3' de cada cebador eran homólogos a la cassette de focalización, mientras que los 42 nt del extremo 5' eran homólogos al último exón de Eno2 en el que se iba a focalizar la cassette (véase la Fig. 10). Las células DY380 eran entonces electroporadas con la cassette de focalización amplificada y se obtenían colonias correctamente focalizadas a una eficiencia cercana a  $10^{-4}$  tras la inducción de la expresión de Red; en las células no inducidas no se obtuvo ninguna colonia focalizada.

20

25

Como también se ha divulgado en el presente, el BAC modificado de longitud completa fue purificado e inyectado en cigotos de ratón y una línea transgénica de BAC establecida. También se establecieron como controles otras dos líneas transgénicas portadoras de un subclón más corto de 25 kbp del gen Eno2 modificado en pBR322. El subclón de 25 kbp porta toda la región de codificación de Eno2 modificado, así como 10 kbp de secuencia flanqueadora del extremo 5' y 5 kbp de secuencia flanqueadora del extremo 3'. La actividad del gen cre en las diferentes líneas transgénicas fue posteriormente valorada cruzando los ratones con los ratones informantes RGSA 26. Estos ratones portan un gen informante lacZ que puede ser activado mediante la recombinasa Cre. En los ratones que portan el transgen BAC de longitud completa, la actividad de Cre se detectó en todas las neuronas Eno2-positivas. Sin embargo, no todas las neuronas Eno2-positivas expresaban Cre en los ratones transgénicos que portaban el subclón de más corto de 25 kbp, y el patrón de expresión de Cre variaba entre las dos diferentes líneas de subclones de 25 kbp. Estos resultados son coherentes con estudios anteriores al demostrar que las secuencias reguladoras se pueden encontrar a cientos de kilobases de un gen y ponen de manifiesto la utilidad de la ingeniería de BAC para, en este caso, generar líneas de expresión de Cre, para emplearlas en experimentos de knockout condicionales.

30

35

40

También se han introducido genes *flpe* o *cre* inducibles por arabinosa en el profago defectuoso portado en la cadena DY380. *flpe* es un *flpe* genéticamente modificado que presenta una frecuencia de recombinación superior que el *flpe* original (Buchholz, et al., Nat biotechnol 16: 657-662, 1998, incorporado al presente como referencia). Las recombinasas de sitio específico *Flpe* y *Cre* son importantes herramientas empleadas para añadir o someter a delección segmentos de ADN (p. ej., cassettes de fármacos). La expresión de *Flpe* y *Cre* se puede inducir mediante la adición de arabinosa y emplearse para eliminar el marcador de selección del lugar focalizado. Esto será especialmente importante en los casos en los que el marcador de selección interfiera en la expresión del lugar focalizado. Es importante señalar que incluso la extracción del marcador seleccionable mediante recombinación de *Flpe* o *Cre* deja atrás el sitio *frt* o *loxP* como una cicatriz en el lugar focalizado.

45

50

*Modificación de BAC sin tener marcadores o "cicatrices" en el sitio diana y modificación genómica directa*

Se ha desarrollado un procedimiento de dos pasos para la focalización de BAC en el que se pueden introducir muchos tipos de mutaciones en los BAC, sin dejar atrás un marcador seleccionable, como un marcador de selección de fármacos en el lugar focalizado. En una realización, se utiliza un procedimiento de dos pasos. Esto se pone de manifiesto en el siguiente ejemplo concreto, de carácter no limitador, en el que una cassette de focalización generada por PCR que contiene un gen de fusión *sacB*-neo fue focalizado para un BAC u otro ADN. Las células que contienen la cassette *sacB*-neo focalizada para el ADN genómico del BAC fueron posteriormente transformadas con un segundo ADN de focalización para la misma región. Esta cassette fue diseñada para sustituir la cassette *sacB*-neo y en un ejemplo contenía secuencias genómicas cortas que portaban una mutación más tenue, como una pequeña inserción. Al colocar estas células recientemente transformadas en un medio con un 7% de sacarosa, se aplicó presión selectiva contra la expresión de *SacB*, que convierte la sacarosa en una bacteriotoxina (Muyrers *et al.*, EMBO Rep 1: 239-243: 2000). El crecimiento en placas de sacarosa seleccionó así a las células que han potencialmente sustituido a la cassette de focalización de *sacB*neo por la segunda cassette de focalización que contenía la pequeña inserción. Dado que las mutaciones espontáneas se producen en *sacB* para producir resistencia a la sacarosa a frecuencias que se aproximan a  $1$  de  $10^4$ , se identificaron recombinantes entre las colonias resistentes a la sacarosa como aquellos que también se habían vuelto sensibles a la neomicina. Como se ha divulgado en el presente, al combinar el poder de la recombinación Red con la selección/contraselección empleando *sacB*-neo, también se pueden generar otros tipos de cambios genéticos además de las inserciones, incluyendo deleciones y mutaciones puntuales, y estas mutaciones se pueden introducir prácticamente en cualquier molécula de ADN grande, como un BAC, PAC o el cromosoma de *E. coli*, sin ningún marcador seleccionable acompañante.

La alta frecuencia de recombinación generada por el sistema de profago defectuoso descrita en el presente también hace posible modificar un genoma bacteriano o un BAC en un único paso sin selección de fármacos ni contraselección. En un ejemplo concreto, de carácter no limitador, se introdujo una etiqueta de 24 bp en un BAC de 125 kbp directamente mediante recombinación, sin selección en el extremo 5' de la caja de SRY que contiene el gen 4 (*Sox4*) (véase el Ejemplo 21). Los recombinantes fueron encontrados estudiando células individuales del cultivo electroporado de BAC.

Dado que las homologías implicadas en la recombinación mediada por Red pueden ser muy cortas, la focalización de las cassettes también se puede conseguir simplemente anillando juntos dos oligonucleótidos de ssADN sintético complementarios. Como se describe en el presente, una cassette de focalización de 70 bp construida de esta manera se recombina con el cromosoma de *E. coli* para crear mutaciones puntuales a frecuencias que se aproximan a una de cada mil células electroporadas. Las mutaciones puntuales correspondientes a mutaciones que causan una enfermedad humana pueden introducirse así en cualquier gen humano o de mamífero portado en un BAC con facilidad y el efecto de esta mutación sobre la función genética puede ser sometido a ensayo en un transgénico que porta una mutación nula en el correspondiente gen de ratón.

*Clonación de ADN mediante reparación de huecos*

Los fragmentos pueden ser subclonados a partir de los BAC mediante recombinación mediada por Red, sin emplear enzimas de restricción ni ligasas de ADN. Así, cualquier región del BAC es susceptible de subclonación y la subclonación no depende de la colocación de sitios de enzimas de restricción apropiados. La subclonación confía en la reparación de huecos para recombinar los extremos libres de un vector de plásmido lineal con secuencias homólogas portadas en el BAC. Un ejemplo se muestra en las Figuras 8 y 10. El vector de plásmido lineal con, por ejemplo, un marcador seleccionable *amp* y un origen de replicación, porta los extremos recombinogénicos. El vector es generado, por ejemplo, mediante amplificación por PCR empleando dos cebadores quiméricos. El extremo 5' de cada cebador presenta homología con los extremos de la secuencia del BAC que se va a subclonar; el extremo 3' de cada cebador se emplea para cebar y amplificar el ADN del plásmido lineal. La recombinación genera un plásmido circular en el que el elemento insertado de ADN se recupera del BAC mediante una reparación de gap. Los plásmidos recombinantes circulares son seleccionados por su fenotipo de resistencia a los fármacos (p. ej., *Amp<sup>R</sup>*). Diferentes tamaños de fragmentos que pueden ser subclonados dependiendo del vector de clonación utilizado. Con un vector de alta copia como pBluescript, se subclonan fragmentos de hasta 25 kbp. No obstante, cuando se utiliza un vector de copia inferior, como pBR322, se pueden subclonar fragmentos de incluso unos 80 kbp. Se demostró que estos fragmentos más largos se expresaban con mayor precisión de forma específica para el tejido (al igual que el clon del BAC completo anteriormente mencionado).

*Un profago lambda movilizable*

Como se ha divulgado en el presente, las funciones de recombinación se expresaron desde su lugar nativo en el operón pL de un profago lambda utilizando los controles naturales del represor de  $\lambda$  (Fig. 9). No obstante, una limitación del sistema del profago defectuoso como se divulga en las Fig. 6 y 9 es que los BAC estudiados deben ser trasladados a células DY380 competentes en recombinación, antes de poder manipular los BAC. A fin de superar esta limitación, se ha generado un nuevo derivado del profago que es aislado como un ADN circular de mini-lambda que porta un marcador seleccionable (p. ej., un marcador resistente a los fármacos, como una cassette *tet<sup>R</sup>*) y que contiene los genes *exo*, *bet* y *gam*, bajo el control del represor *cl857* inducible por temperatura (Fig. 12). Este mini-lambda puede ser



transformado en cualquier célula bacteriana, como una célula DH10B que porta un BAC. El mini-lambda se integra entonces en el sitio de unión de lambda para generar el profago defectuoso. Este profago movilizable hace posible introducir el profago en bibliotecas de DH10B que contienen BAC y obvia la necesidad de transferir el BAC a las células DY380.

## 5 Recombinería empleando ssADN

La recombinería, o el uso de una recombinasa para mediar la recombinación utilizando brazos de homología suficiente como para inducir la recombinación, como los divulgados en el presente, se puede realizar empleando oligos como cassette de focalización. Como se describe en los Ejemplos siguientes (p. ej., véase el Ejemplo 3), en *E. coli*, un cambio en una única base ha sido sustituido en el gen *galK* y se ha extraído una inserción de 3,3 kbp del gen *galK* empleando oligos de cadena única. Los oligos de cadena única también se han utilizado para reparar cinco inserciones de Tn10 diferentes en distintos lugares del cromosoma de *E. coli*. La recombinería empleando oligos de cadena única es muy eficiente, dado que hasta un 6% de las células electroporadas son recombinantes. Mientras que Exo, Beta y Gam facilitan la recombinación de cassettes de dsDNAADN amplificado mediante PCR con homología en los flancos, solamente se necesita Beta para la recombinación de ssADN (véase la Fig. 13). La recombinación máxima se consigue con oligonucleótidos de unas 70 bases de longitud. También se pueden emplear oligonucleótidos de unas 40 a 60 bases de longitud para conseguir la recombinación, aunque a una frecuencia cinco veces inferior. En una realización, se utiliza un ssADN de unos 40 a unos 70 nucleótidos de longitud. En otra realización, se utiliza un ssDNAADN de unos 70 a unos 100 nucleótidos de longitud. En otra realización más, se utiliza un ssDNAADN de unos 70 a unos 1.000 nucleótidos de longitud. Resulta interesante señalar que la recombinación mediada por Beta es menos eficiente cuando las moléculas de ssADN tienen unas 1.000 bases de longitud. En otra realización más, el ssADN es etiquetado, como con un nucleótido biotinilado, un nucleótido metilado o un aducto de ADN.

La recombinación con cualquiera de los oligos de ADN complementarios o con ambos ha puesto de manifiesto que a pesar de que se puede utilizar cualquiera de las cadenas de forma eficiente para la recombinación, una cadena es más competente para la recombinación que la otra. Esta tendencia de las cadenas ha sido examinada en diferentes posiciones alrededor del cromosoma bacteriano, con el resultado de que la cadena preferible está correlacionada con la cadena discontinua de la replicación de ADN para cada sitio testado. Sin limitarse a la teoría, estos resultados han indicado que la tendencia de la cadena está asociada con la dirección de replicación a través de la región objeto de la focalización y que la recombinación de ssADN ocurre eficientemente cerca de la horquilla de replicación. El proceso de replicación de ADN resulta en regiones transitorias de ssADN que puede resultar accesible para el anillamiento mediado por beta del oligo de ssADN. A pesar de que la recombinación se produce en la cadena continua, la eficiencia aumentada de la recombinación de los oligos de la cadena discontinua puede reflejar una mayor frecuencia de regiones de cadena única durante la síntesis de la cadena discontinua frente a la de la cadena continua (Fig. 13). La ADN polimerasa y ADN ligasa podrían después completar la unión del oligo anillado a la cadena discontinua. Sin limitarse a la teoría, el aumento de la frecuencia de la recombinación de ssADN probablemente refleja el hecho de que la recombinación de ssADN ocurre a través de un mecanismo más sencillo que la recombinación de dsADN. La recombinación de ssADN puede exigir solamente el anillamiento de un oligo de cadena única a regiones de cadena única en el ADN objeto de la replicación. Por otra parte, la recombinación de ssADN también ocurre en la levadura con una tendencia de las cadenas que también pueden ser dependientes tras la replicación. No obstante, las funciones de la levadura necesarias para esta recombinación se desconoce, lo que hace que el descubrimiento de que solamente la función Beta del fago  $\lambda$  resulta necesaria en *E. coli* adquiera una importancia mucho mayor.

En un ejemplo específico, de carácter no limitador, se realizó una mutación puntual en el *Brca2* de ratón portado en un BAC empleando un oligo de 70 nt. La eficiencia de focalización fue varias veces mayor de la que se obtendría con dsADN creado con anillamiento de oligos y al menos 50 veces mayor de la que se obtendría con dsADN generado mediante PCR y con grandes regiones de ausencia de homología en su parte central. También se ha empleado un oligonucleótido de 140 nt para introducir una delección en marco de 29 aminoácidos en el exón 1 del gen *Brca2* y una delección de 1,93 kb en el segmento principal del vector de BAC (Swaminathan, S. *et al.* Genesis 29, 14-21, 2001, incorporado al presente como referencia). Finalmente, se ha empleado un oligonucleótido de 164 nt para introducir una etiqueta Flag de 24 bp en el extremo 5' de *Brca2*. La eficiencia de focalización para el oligo de 164 nt ( $7,7 \times 10^{-3}$ ) fue prácticamente la misma que la eficiencia de focalización para la producción de delecciones empleando oligos de 140 nt ( $8,3 \times 10^{-3}$  y  $5,4 \times 10^{-3}$ , respectivamente).

## EJEMPLOS

### EJEMPLO 1

#### Profago lambda modificado para la expresión definida de proteínas recombinantes

La genética molecular del bacteriófago lambda, incluyendo sus ciclos de crecimiento lítico y lisogénico, se describe en Sambrook *et al.*, Bacteriophage Lambda Vector, Capítulo 2 en Molecular Cloning: a Laboratory Manual, 2ª Ed., (c) 1989 {en adelante, Sambrook *et al.*, Cap. 2}; Stryer, Control of Gene Expression in Prokaryotes, Capítulo 32 en Biochemistry 3ª Ed., pp. 799-823, (c) 1988 (en adelante, Stryer); y Court y Oppenheim, pp. 251-277 en Hendrix *et al.* eds., Lambda IT,

Cold Spring Harbor Lab Press, (c) 1983 (en adelante, Court y Oppenheim). La secuencia completa de lambda es conocida (véase el número de acceso Genbank nº NC 001416).

5 El fago lambda cuenta con un sistema de recombinación homóloga bien caracterizado. Las roturas de doble cadena en el ADN son los sitios de inicio para esta recombinación (Thaler *et al.*, J. Mol. Biol. 195: 75-87, 1987). La exonucleasa de lambda (Exo) se degrada progresivamente desde los extremos 5' de estos sitios de rotura y el lambda Beta se une a la cola restante de cadena única del extremo 3', protegiendo y preparando el ADN acoplado para la invasión de la cadena homóloga (Little, J. Biol. Chem. 242: 679-686, 1.967; Carter *et al.*, J. Biol. Chem. 246: 2502-2512, 1971).

10 El sistema de recombinación de lambda que contiene exo y bet sin gam es eficiente en la sustitución genética empleando sustratos lineales con brazos de homología de más de 1.000 bp en una cadena que carece de la nucleasa RecBCD (Murphy, Journal of Bacteriology 180: 2063-2071, 1998). Para testar brazos de homología de menos de 100 bp de largo como sustratos para la recombinación mediada por lambda, se modificó un profago lambda para que expresase elevados niveles de funciones de recombinación del fago durante un período de tiempo definido.

15 La FIG. 1 ilustra el profago  $\lambda$  defectuoso en el cromosoma de *E. coli*. El profago defectuoso contiene los genes de  $\lambda$  de *cl* a *int*. El operón de PL está intacto y expresado bajo el control del represor *cl* de lambda sensible a la temperatura (alelo *cl857*). Una delección (línea de puntos) elimina el lado derecho del profago de *cro* a *attR* e incluyendo *bioA* (Patterson *et al.*, Gene 132: 83-87, 1993). En el cromosoma, los operones *nadA* y *gal* se encuentran a la izquierda del profago y los genes *bio* sin *bioA* se encuentran a la derecha. Los genes del profago  $\lambda$  se muestran en la línea continua y los genes del hospedador se muestran en la línea discontinua. PL y PR indican los promotores tempranos izquierdo y derecho de  $\lambda$ . *attL* y *attR* indican los sitios de unión izquierdo y derecho de  $\lambda$ . Los genes y las funciones de lambda se describen en Sambrook *et al.*, capítulo 2, Stryer, y Court y Oppenheim.

25 La ausencia del represor de *cro* permite que la expresión del operón PL esté plenamente sometida a desrepresión cuando se inactiva el represor *cl* sensible a la temperatura, a 42 °C. La delección de *cro* a *bioA* elimina la replicación y los genes líticos del profago. Las funciones que codifican estos genes líticos son tóxicas para la célula y provocan la muerte celular en un plazo de 7 minutos a partir de una inducción del profago normal. Las funciones presentes en el operón PL también son tóxicas, pero matan a las células solamente tras una inducción continua de 60 minutos (Greer, Virology 66: 589-604, 1975; Kourilsky *et al.*, Biochimie 56: 1517-1523, 1974). De este modo, el cambio de las células que contienen la estructura del operón PL de unas condiciones de represión a 32 °C a unas condiciones de inducción a 42 °C permite la expresión del operón PL. Un nuevo cambio de las células de 42 °C a 32 °C (o menos) en un plazo de 60 minutos restablece la represión y evita la muerte celular.

30 Como demostrarán los siguientes ejemplos, este profago lambda modificado ha producido ventajas inesperadas en la mediación de la recombinación homóloga. Estas ventajas inesperadas incluyen una eficiencia de recombinación sorprendentemente elevada, un control preciso de las funciones de recombinación, una recombinación efectiva con brazos de homología cortos y la capacidad para generar recombinantes homólogos con polinucleótidos distintos al dsADN largo. Sin ánimo de limitarse a única explicación de los efectos observados, es probable que estas ventajas imprevistas sean el resultado de la incorporación de los genes lambda Red del profago en su contexto nativo, limitando así el número de copias de genes de recombinación de lambda. Por otra parte, el uso del promotor PL confiere la capacidad de controlar de forma precisa el tiempo y la producción de grandes cantidades de expresión del gen de recombinación de lambda.

40 Este sistema no se limita a la expresión en *E. coli*, sino que también puede funcionar en otras bacterias, como la *Salmonella* entre otras. También puede funcionar en células eucariotas, como células de mamíferos o levaduras, con la selección de los promotores apropiados y con otras modificaciones del presente para permitir la expresión de los genes de recombinasa de lambda. En un ejemplo específico, de carácter no limitador, en otra bacteria, los genes entre *gam* y *N* (incluyendo *N* pero no *gam*) pueden ser sometidos a delección para eliminar terminadores de la transcripción.

45 A pesar de que el promotor PL se emplea para ilustrar la invención, se podrían utilizar otros promotores constitutivos, inducibles o susceptibles de desrepresión. Algunos ejemplos, a título meramente enunciativo, de promotores inducibles son los promotores inducibles por fármacos (p. ej., un promotor inducible por tetraciclina), los promotores inducibles por metales (p. ej., el promotor por metalotioneína), o promotores inducibles por hormonas (p. ej., un elemento que reacciona a los esteroides).

50 El promotor PL también se podría utilizar para impulsar la expresión, por ejemplo, de los genes de P22, como *erf*, o de RecE y RecT.

## EJEMPLO 2

**Cadenas bacterianas, expresión del operón PL, métodos de electroporación, identificación de recombinantes. Las cadenas bacterianas empleadas en este trabajo se recogen en la Tabla 2**

Tabla 2

CADENAS	GENOTIPO
WJW23	<i>his ilv rpsI Δ(argf-λαχ)Y169 ναδA::Tv10 γαλ490 λcl857 Δ(cro-βιοA)</i>
ZH1141	<i>W3110 Δ(argF-lac)U169 gal490 λN:λαχZ Δ(N-ιντ) χI857 Δ(cro-βιοA)</i>
BR367	<i>lacI<sup>9</sup> lacZ(M15) Δ(srl-ρεχA)301::Tv10</i>
7	
DY329	<i>W3110 Δ(argF-λαχ)Y169 ναδA::Tv10 γαλ490 λcl857 Δ(cro-βιοA)</i>
DY330	<i>W3110 Δ(argF-λαχ)Y169 γαλ490 λcl857 Δ(cro-βιοA)</i>
DY331	<i>W3110 Δ(argF-λαχ)Y169 Δsrl-ρεχA)301::Tv10 γαλ490 λcl857 Δ(cro-βιοA)</i>
DY378	<i>W3110 λcl857 Δ(cro-βιοA)</i>
W3110	"Wild-type"
HME5	<i>Δ(argF-λαχ) Y169 λcl857 Δ(cro-βιοA)</i>
HME6	<i>Δ(argF-λαχ)Y169 γαλK<sub>τψρ145YAG</sub> λcl857 Δ(cro-βιοA)</i>
HME9	<i>Δ(argF-λαχ)Y169 γαλK<sub>τψρ145YAG</sub> λcl857 Δ(cro-βιοA) τψρT<sub>ς</sub>&lt;&gt;χατ</i>
HME10	<i>Δ(argF-λαχ)Y169 γαλK<sub>τψρ145YAG</sub> λ cl857 Δ(cro-βιοA) τψρT<sub>ς</sub>&lt;&gt;χατ Δ(srl-ρεχA)301::Tv10</i>
HME31	<i>Δ(argF-λαχ)Y169 γαλK&lt;&gt;χατσαχB λcl857 Δ(cro-βιοA)</i>
HME40	<i>Δ(argF-λαχ)Y169 INγαλ[γαλM<sup>+</sup>K<sub>τψρ145YAG</sub>T<sup>+</sup>E<sup>+</sup>] λcl857 Δ(cro-βιοA)</i>
HME43	<i>Δ(argF-λαχ)Y169 γαλK<sub>τψρ145YAG</sub> λ (exo-int)&lt;&gt;cat Δ(gam-N) χI857 Δ(cro-βιοA)</i>
HME47	<i>galK□34&lt;&gt;kan λexo&lt;&gt;cat cl857 Δ(cro-bioA)</i>
DY411	<i>galK□34&lt;&gt;kan λcl857 Δ(cro-βιοA)</i>
DH10B	<i>P mcrA Δ(mrr-ησδPMΣ-μερBX) φ80dlacZΔM15 ΔlacX74 deoR reeA1 endA1 araD139 Δ(ara, leu) 7649 γαλU galK rspL nupG</i>
DY303	<i>DH10B [λcl857recA<sup>*</sup>]</i>
DY374	<i>W3110 gal490 nadA::Tn10 [λcl857 Δ(cro-bioA)]</i>
DY363	<i>W3110 ΔlacU169 gal490 [λcl857 (cro-bioA)Δtet<sup>R</sup>]</i>
DY380	<i>DH10B [λcl857 (cro-bioA)Δtet]</i>
EL11	<i>DH10B [λcl85 (cro-bioA)Δcat-sacB]</i>
EL250	<i>DH10B [λcl857 (ero-bioA)ΔaraC-P<sub>BAD</sub>flpe<sup>b</sup>]</i>
EL350	<i>DH10B [λcl857 (ero-bioA)ΔaraC-P<sub>BAD</sub>cre]</i>

<sup>a</sup> (cro-bioA)<>tet indica la sustitución de cro-bioA por tet. <sup>b</sup> P<sub>BAD</sub> representa el promotor de araBAD.

- 5 La cadena DY329 se construyó mediante transducción de ZH1141 con el fago PI cultivado en WJW23, seleccionando la resistencia a tetraciclina de nadA::Tn10 (Tet<sup>R</sup>) a 32°C y después detectando la presencia de un profago lambda defectuoso que causa un crecimiento celular sensible a la temperatura a 42 °C. Una transducción de PI similar se empleó para crear otras cadenas descritas en la Tabla 2 empleando medios, métodos y selecciones estándar (Sambrook *et al.*, Molecular Cloning: a Laboratory Manual, 2<sup>a</sup> Ed., (c) 1989; Miller, Experiments in Molecular Genetics, Cold Spring Harbor Lab Press (c) 1972). El símbolo <> se emplea para indicar una sustitución generada mediante recombinación homóloga. El símbolo >> indica una inserción generada mediante recombinación homóloga. Una delección en el punto de inserción se indica en paréntesis después del gen insertado. El operón gal completo de HME40 está invertido (IN).

- 15 Para inducir la expresión a partir del operón PL y preparar células competentes en electroporación, los cultivos de una noche a 32 °C de colonias aisladas se diluyeron 50 veces en medio LB y se cultivaron a 32°C con agitación hasta una OD<sub>600</sub> de aproximadamente 0,4-0,8. La inducción se realizó en un cultivo de 10 ml en un matraz cónico con deflector (50 ml) colocando el matraz en un baño de agua a 42 °C con agitación (200 revoluciones/min.) durante 15 minutos. Inmediatamente después de la inducción de 15 minutos, el matraz se hizo girar en agua con hielo para que se enfriara durante 10 minutos. Un cultivo de control no inducido, mantenido a 32 °C todo el tiempo, también se colocó en el agua con hielo. Los cultivos de 10 ml enfriados fueron centrifugados durante 8 minutos a 5.500 x g a 4°C. Cada gránulo celular fue suspendido en 1 ml de agua estéril enfriada con hielo, transferido a un tubo de microcentrifugado de plástico de 1,5 ml y centrifugado durante 20 segundos a 4 °C a máxima velocidad en una microcentrifugadora. Después de lavar los gránulos celulares dos veces más como se ha descrito, las células fueron suspendidas en 100 μl de agua estéril enfriada con hielo. Este volumen de células competentes es suficiente para dos reacciones de electroporación estándar (~10<sup>8</sup> células por reacción). Se pueden preparar cultivos mayores para un mayor número de reacciones o para el almacenamiento de células electrocompetentes a -80 °C con glicerol al 12%. Las células competentes frescas presentan mayores eficiencias de recombinación. Para transformar células mediante electroporación, el ADN donante lineal purificado (1 a 10 μl) se mezcló con células competentes hasta un volumen final de 50 μl sobre hielo y posteriormente se introdujo en una cubeta de electroporación previamente enfriada (0,1 cm). La cantidad de ADN donante empleado

por reacción (normalmente de 1 a 100 ng) está indicada para experimentos relevantes. La electroporación se realizó empleando un Bio-Rad Gene Pulser a 1.8 kV, 25 µF con controlador de impulso de 200 ohms. Se han empleado dos protocolos intercambiamente para permitir la segregación del recombinante de los cromosomas precursores dentro de las células electroporadas. En ambos protocolos, las células electroporadas eran diluidas inmediatamente con 1 ml de medio LB. En uno, las células se incubaron de 1 a 2 horas a 32 °C antes de seleccionar los recombinantes. En el otro, las células fueron inmediatamente diluidas y extendidas sobre filtros de nitrocelulosa estériles (100 mm) en agar LB. Tras una incubación de 2 horas a 32 °C, los filtros fueron transferidos a las correspondientes placas de agar, necesarias para seleccionar los recombinantes. También se extendieron partes alícuotas directamente sobre el agar LB y se incubaron a 32 °C para determinar y examinar el total de células viables tras la electroporación. Para la selección de la resistencia a los fármacos, cada ml de medio LB contenía 10 µg de cloranfenicol, 12.5 µg de tetraciclina, 20 µg de canamicina, 30 µg de ampicilina o 50g de espectinomocina.

A pesar de que los recombinantes fueron verificados por más de un método, la detección primaria fue de un fenotipo alterado causado por el gen diana modificado. La disrupción o mutación del gen *galK* fue confirmada por la presencia de colonias blancas en el agar indicador de galactosa de MacConkey; la disrupción del gen *mc* de la endorribonucleasa RNaseHI fue confirmada por la incapacidad del fago tipo lambda para lisogenizar (Court, pp. 71-116 en Belasco *et al.*, eds., Control of Messenger RNA Stability, (c) 1993, Academic Press, New York), y la delección de *gam*, *kiI*, y *cIII* en el operón PL se valoró como una capacidad del lisógeno de λ para sobrevivir a un cultivo a 42°C (Court y Oppenheim; Greer, Virology 66: 589-604, 1975). El análisis por PCR se empleó para conformar la estructura alterada causada por la sustitución de un gen. Los análisis de la hibridación de Southern de los ADN precursor y recombinante confirmaron cambios estructurales y el ADN de los clones recombinantes puede ser amplificado mediante PCR y secuenciado.

Además de la electroporación, cualquier método apropiado para la transferencia molecular a células bacterianas resultaría efectivo para la práctica de los métodos divulgados en el presente. Por ejemplo, estos métodos pueden incluir la exposición a cationes divalentes, DMSO y similares, descritos en diversas publicaciones de laboratorios estándar, tales como Sambrook *et al.* (véanse concretamente las páginas 1.74-1.84) y Ausubel *et al.*, eds., Current Protocols In Molecular Biology, John Wiley & Sons (c) 1998 (en adelante, Ausubel *et al.*), incorporado al presente en su totalidad.

### EJEMPLO 3

#### Recombinación homóloga con fragmentos de ADN lineales cortos

El sistema de recombinación descrito en los Ejemplos 1 y 2 fue empleado para generar una mutación de un único bp en el gen *galK* bacteriano.

El gen *galK* codifica una galactoquinasa que fosforila galactosa y sus derivados. La galactoquinasa de *galK* fosforila 2-desoxigalactosa para generar 2-desoxigalactosa fosfato (2DGP). A pesar de que la 2-desoxigalactosa no fosforilada no tiene ningún impacto sobre el crecimiento celular, 2DGP es un azúcar-fosfato no metabolizado que inhibe el crecimiento celular. Por tanto, las células que contienen un gen *galK* de tipo salvaje tienen un crecimiento limitado en 2-desoxigalactosa, un fenotipo denominado en el presente Gal+. Por lo contrario, los mutantes defectuosos en *galK* crecen bien en presencia de 2-desoxigalactosa (Dog) y tienen un fenotipo denominado en el presente DogR- (Adhya, páginas 1503-1512 en Escherechia coli and Salmonella typhimurium: Cellular and Molecular Biology, Neihardt *et al.* eds., American Society of Microbiology, 1987). El fenotipo DogR- permite una rápida selección de células que albergan eventos de recombinación con éxito.

Se sintetizaron y anillaron entre sí oligonucleótidos complementarios de 70 pares de bases. El ADN anillado era homólogo a un segmento de codificación interno del gen *galK* bacteriano, salvo por que un codón UAU (TYR-145) se había cambiado por un codón UAG ámbar. Un evento de recombinación homóloga entre este fragmento de ADN de 70 pares de bases y el gen *galK* introduce un codón de parada prematuro en el gen *galK*, lo que se denomina en el presente mutación ámbar de *galK*. Esta mutación produce un producto genético de *galK* truncado que carece de función.

El fragmento de ADN de 70 bp fue transferido mediante electroporación a las células de *galK*+ (HME5) que habían sido inducidas para la expresión del operón PL de lambda mediante cultivo a 42 °C durante 15 minutos. Después de la electroporación con el ADN mutante (10 ng) o electroporación simulada sin ADN, las células se extendieron en un medio de agar con un 0,4% de glicerol como mínimo, que contenía un 0,2% de 2-desoxigalactosa. Los mutantes resistentes espontáneos aparecieron con frecuencia (10<sup>-4</sup>) en ausencia del ADN mutante. A pesar de ello, la adición del ADN mutante mejoró la frecuencia de mutantes resistentes de forma drástica, generando un mutante por cada 500 células electroporadas.

Para determinar que la inducción de temperatura fue necesaria, otro lote de células que no había sido inducido para la función de recombinación se testó de la misma manera. En este tratamiento, no se observó ningún efecto discernible del ADN mutante añadido. Esto indicó que tanto la inducción de la expresión del operón PL como la adición de ADN mutante fueron necesarias para una supervivencia mejorada. Sin ánimo de limitarse a una única explicación de los efectos observados, se cree que las funciones de lambda expresadas permitieron la recombinación eficiente de este ADN mutante lineal corto con el gen *galK* cromosómico.

Las colonias que sobrevivieron al tratamiento con 2-desoxigalactosa fueron estudiadas para detectar su fenotipo de Gal en placas indicadoras y todas las colonias testadas tenían el fenotipo Gal previsto para un mutante ámbar de *galK*. Para testar específicamente si la mutación ámbar de *galK* estaba presente, se testaron cuatro colonias de Gal independientes, transduciendo cultivos de cada una de ellas con un fago lambdaimm21 que porta el tRNA<sub>tyr</sub> supresor del alelo supF. Los cuatro mutantes testados fueron suprimidos con un fenotipo Gal+. Finalmente, la presencia de la mutación ámbar en *galK* fue verificada mediante amplificación por PCR y análisis de la secuencia del segmento del gen *galK* del cromosoma.

Este ejemplo demuestra que la expresión controlada de genes de recombinación de lambda por parte de un profago lambda defectuoso promueve de forma sorprendente la recombinación homóloga eficiente en las células bacterianas, incluso en los segmentos lineales cortos de ADN con brazos de homología muy cortos.

Es importante señalar que la elevada frecuencia de recombinación indica que los recombinantes fueron identificados sin necesidad de aplicar métodos de selección positivos o negativos. Las sondas de hibridación de ADN están por tanto diseñadas para detectar mutaciones puntuales, deleciones, inserciones u otras modificaciones del ADN celular. Las técnicas estándar de hibridación de colonias o hibridación *in situ* pueden emplearse para detectar células en las que se ha producido la recombinación. Alternativamente, se emplean métodos de enriquecimiento para la detección de la mutación, particularmente para la detección de mutaciones puntuales. Estos métodos de enriquecimiento se describen en Gockc *et al.*, Annals of the New York Academy of Sciences 906: 31 -38, 2000. Un ejemplo de enriquecimiento adecuado es el ensayo de TaqMan (TaqMAMA) descrito por Cha *et al.*, PCR Applications and Methods 2: 14-20, 1992.

El hecho de que los fragmentos de ADN con brazos de homología cortos sean capaces de recombinarse *in vivo* abre un amplio abanico de nuevas posibilidades para generar ADN recombinante. Se eliminan varios pasos normalmente implicados en la generación de moléculas de ADN recombinante. No son necesarias digestiones de enzimas de restricción para generar fragmentos de ADN y las reacciones de la ligasa del ADN no son necesarias para unir diferentes fragmentos de ADN en enlaces nuevos. La célula genera el recombinante completo unido de forma precisa a través de recombinación homóloga.

La eficiencia de la recombinación se acerca al 0,1% de las células supervivientes de una electroporación estándar. A esta eficiencia, las colonias no seleccionadas podrían ser analizadas para detectar ADN recombinante utilizando técnicas de hibridación de colonias, eliminando así la necesidad de pasos de selección. De este modo, este protocolo de recombinación hace que el ADN del plásmido y el cromosoma bacteriano sean susceptibles prácticamente de cualquier tipo de cambio deseado. Esto incluye mutagénesis dirigida de un gen, un segmento de un gen o incluso una base.

#### EJEMPLO 4

##### Preparación de cassettes de ADN lineal de más de 1.000 bp de longitud

Se utilizaron las condiciones de la reacción en cadena de la polimerasa (PCR) estándar para amplificar fragmentos de ADN lineales con el sistema de PCR Expand<sup>TM</sup> High Fidelity de Boehringcr Mannheim. La cassette resistente a cloranfenicol (Cm<sup>R</sup>) *cut* fue amplificada a partir de pPCR-Script Cam (Stratagene) con los cebadores 5'TG'GACGGAAGATCACTTCG (SEQ ID NO: 1) y VACCAGCAATAGACATAAGCG (SEQ ID NO: 2). La cassette resistente a tetraciclina (Tet<sup>R</sup>) *tet* fue amplificada a partir de Tn10 con los cebadores 5'CTCTTGGGTTATCAAGAGGG (SEQ ID NO: 3) y 5'ACTCGACATCTTGGTTACCG (SEQ ID NO: 4). La cassette resistente a ampicilina (Ap<sup>R</sup>) *amp* fue amplificada a partir de pBluescript (Stratagene) con los cebadores 5'CATTCAAATATGTATCCGCTC (SEQ ID NO: 5) y 5'AGAGTTGGTAGCTCTTGATC (SEQ ID NO: 6). La cassette resistente a kanamicina *kan* fue amplificada a partir de Tn5 con los cebadores TATGGACAGCAAGCGAACCG (SEQ ID NO: 7) y 5TCAGAAGAACTCGTCAAGAAG (SEQ ID NO: 8). Los productos de la PCR fueron purificados empleando equipos de purificación para PCR de Qiagen y concentrados, cuando resultó necesario, mediante precipitado con etanol. Los ADN lineales amplificados fueron suspendidos en agua estéril o solución tampón de TE (10mM Tris-Cl pH7.5; 1mM EDTA) y cuantificados mediante espectroscopia. El ADN en agua se almacenó a -20 °C. Los inventores evitaron programas de purificación de productos de la PCR a partir de geles, en los que el ADN es sometido a radiación ultravioleta.

A fin de diseñar cebadores para la amplificación de una cassette de recombinación, los oligonucleótidos recombinantes fueron sintetizados químicamente con los 30 a 50 nucleótidos del extremo 5' idénticos a secuencias de la secuencia de ácido nucleico diana y con los 20 nucleótidos del extremo 3' homólogos a los extremos de la cassette a introducir. Se genera una cassette por PCR que está flanqueada por las 30 a 50 homología de bases presentes en la diana.

Las células que portan el ADN diana, sea en el cromosoma o en un plásmido, son inducidas para la función de Exo, Beta y Gam. Estas células se vuelven competentes para la electroporación y se mezclan con la cassette amplificada. Tras la electroporación, la recombinación ocurre entre las secuencias homólogas de la cassette lineal y la diana sustituye el segmento diana por la cassette.

Los segmentos de homología de *galK* de 50 nt (rectángulos) empleados para el experimento descrito en la Tabla 3 son:

5'GTTTGC GCGCAGTCAGCGATATCCATTTTCGCGAATCCGGAGTGTAAGAA (SEQ ID NO: 9) y

5"TTCATATTGTTTCAGCGACAGCTTGCTGTACGGCAGGCACCAGCTCTTCCG (SEQ ID NO: 10)

En una realización, la cassette es un marcador de resistencia a fármacos, pero puede ser cualquier ADN si la secuencia diana de los pasos posteriores puede ser contraseleccionada. La transcripción de la cassette del marcador ha sido orientada arbitrariamente en la misma dirección que la región diana que va a ser sustituida. Los cebadores contienen dos partes: un extremo 5' homólogo a las regiones de los flancos del ADN diana y un extremo 3' que ceba el ADN de la cassette para la replicación. La PCR empleando estos cebadores y una plantilla de ADN que contiene la cassette del marcador genera un producto de ADN lineal con la cassette flanqueada por la homología diana. Cabe señalar que si la transformación con el ADN de la plantilla va a generar el fenotipo seleccionado (por ejemplo, la plantilla es un plásmido), entonces la plantilla es eliminada. El ADN de una plantilla de plásmido puede ser destruido mediante tratamiento con DpnI tras la PCR; DpnI corta el ADN de la plantilla de GATC metilado, dejando el ADN no metilado recién replicado intacto. Una vez que se ha generado una cassette lineal, puede ser almacenada y empleada como plantilla para las PCR posteriores.

### EJEMPLO 5

#### 15 Sustitución de genes mediante recombinación homóloga focalizada

Habiendo demostrado en el Ejemplo 3 que el ADN lineal de 70 bp puede dirigir las mutaciones a una diana específica, se construyó un ADN sintético con segmentos de ADN de *galK* de 50 bp a los flancos de la cassette de *cat* (resistencia a cloranfenicol, o Cm<sup>R</sup>) para focalizarse en la sustitución de un gen *galK* por *cat*.

La cassette de *cat* lineal con ADN de *galK* en los flancos se realizó mediante PCR, empleando cebadores sintetizados químicamente, tal y como se describe en el Ejemplo 4.

Los datos de estos experimentos se presentan en la Tabla 3. Las células competentes DY330 fueron electroporadas con 100 ng de la cassette de *cat* focalizada para sustituir *galK* (*galK*<>*cat*) o los genes del profago *cIII kil gam* (*cIII kil gam*<>*cat*; véase la FIG. 1 para obtener el mapa de genes del profago). Los recombinantes totales por electroporación se muestran en la columna que se encuentra más a la derecha, "Recombinantes Cm<sup>R</sup>". La cassette de *cat* fue transferida por electroporación a las células *galK*<sup>+</sup> que habían sido inducidas por calor para la expresión del operón PL (15 minutos de cambio de temperatura a 42 °C, como se ha indicado con "15" en la columna central de la Tabla 3) o no inducidas (mantenidas a 32 °C, tal y como se ha indicado con "0" en la columna central de la Tabla 3). Véase también el Ejemplo 2 para obtener una descripción de la inducción y otros métodos. Tras la electroporación, los recombinantes Cm<sup>R</sup> fueron seleccionados a 32°C y posteriormente cuantificados.

Como se muestra en la Tabla 3, las colonias de Cm<sup>R</sup> solamente se encontraron en el cultivo inducido por calor. Las 50 colonias Cm<sup>R</sup> testadas tenían un fenotipo Gal<sup>-</sup> en el agar indicador de galactosa de MacConkey, indicando la presencia de la sustitución *galK*<>*cat*. El símbolo <> indica una sustitución generada mediante técnicas de recombinación homóloga, por ejemplo, *galK*<>*cat* indica que el gen *galK* bacteriano está sustituido por *cat* empleando técnicas de recombinación homóloga.

35 **Tabla 3**

Sitio diana*	42°C, min	Recombinantes Cm <sup>R</sup>
<i>GalK</i>	0	<1
<i>GalK</i>	15	2.5 x10 <sup>4</sup>
<i>CIII kil gam</i>	0	<1
<i>CIII kil gam</i>	15	5.0 x10 <sup>4</sup>

En experimentos similares empleando los mismos brazos homólogos de 50 bp, el *galK* ha sido intercambiado por cassettes de *kan*, *amp*, y *tet* seleccionando Km<sup>R</sup>, Amp<sup>R</sup>, y Tet<sup>R</sup>, para crear sustituciones *galK*<>*kan*, *galK*<>*amp*, y *galK*<>*tet*, respectivamente.

Para testar si este planteamiento también funciona en otras posiciones del cromosoma bacteriano, se creó una cassette de *cat* lineal flanqueada por segmentos de ADN de 50 bp encontrados inmediatamente en sentido ascendente y descendente del gen *mc* que codifica RNasell. Se cree que el gen *mc* no es esencial (Takiff *et al.*, Journal of Bacteriology 171: 2581-20 2590, 1989); por tanto, se testó si una sustitución exacta de la región de codificación de *cat* por la región de codificación de *mc* (de AUG al codón 224) se podía realizar empleando las técnicas de recombinación aquí descritas. En esta estructura, *cat* se transcribía desde el promotor de *mc*, y el cebador del extremo 5' empleado para generar *cat* comenzaba en el codón de inicio de *cat*.

Siguiendo los procedimientos empleados para *galK* descritos en los Ejemplos 2 y 3, se encontraron colonias de Cm<sup>R</sup>, pero solamente en el cultivo inducido. Las colonias de Cm<sup>R</sup> testadas tenían un fenotipo mutante de Rnc, tal y como se describe en el Ejemplo 2.

5 Se generaron otros dos recombinantes de *rnc*<>*cat*. Uno sustituía a la secuencia desde el inicio de AUG hasta el codón 126 de *rnc* y el otro desde el inicio de AUG hasta el codón 192 de *rnc*. Estos dos recombinantes generan fusiones de genes *cat::rnc* con un fenotipo mutante de *rnc*. Se seleccionaron diferentes conjuntos de cebadores para detectar sin ambigüedades los alelos recombinantes y/o de tipo salvaje. Este procedimiento de PCR sigue las directrices establecidas por los investigadores de levaduras al caracterizar las sustituciones cromosómicas en la levadura (Winzler *et al.*, Science 285: 901-906, 1989). El análisis por PCR de los recombinantes verificó la pérdida del gen *rnc*+ y las estructuras previstas de las tres sustituciones de genes *rnc*<>*cat*.  
10

15 En este ejemplo, varios genes diferentes del cromosoma bacteriano y de los plásmidos han sido sustituidos por marcadores de resistencia a fármacos. No obstante, también resulta posible crear recombinantes en los que el producto deseado no incluye un marcador seleccionable. Los genes han sido fusionados con cassettes que codifican etiquetas específicas, como la proteína fluorescente verde. Las etiquetas de fusión se pueden colocar de forma precisa en el gen que se pretende modificar, por ejemplo mediante cualquiera de las estrategias siguientes. En una, la cassette no seleccionada se une a un marcador de fármacos seleccionable y ambos son recombinados en el lugar elegido seleccionando la resistencia a fármacos. En otra, la cassette es recombinada en su ubicación, sustituyéndola por un marcador de selección negativo como *sacB* (Blomfield *et al.*, Mol Microbiol 5, 1447-57, 1991). Esta estrategia permite la clonación de cualquier ADN. En una tercera estrategia, los recombinantes son examinados de forma no selectiva mediante hibridación de ADN con sondas específicas para la cassette.  
20

25 En estos experimentos, habitualmente se obtuvo el producto de recombinación deseado. No obstante, algunos productos de recombinación fueron inesperados. En dos casos, se intentó desactivar genes esenciales y sorprendentemente resultó posible seleccionar algunos recombinantes poco comunes. Éstos resultaron ser diploides para la región del gen focalizado, dado que portaban el tipo salvaje y la copia mutante del gen, tal y como se determinó mediante un análisis por PCR. Se sabe que las regiones diploides poco frecuentes del cromosoma bacteriano ocurren de forma espontánea en los cultivos de células a una frecuencia aproximada del 0,1% (Haack y Roth, Genetics 14, 1245-1252, 1995). Dado que el objeto de la focalización fue un gen esencial, estos diploides poco comunes fueron seleccionados. Esto solamente fue posible debido a la gran eficiencia de la recombinación.

30 Este ejemplo demuestra que los métodos de esta invención se pueden emplear para promover una sustitución eficiente de los genes mediante recombinación homóloga. Las sustituciones de genes se realizaron en todo el cromosoma bacteriano. En el Ejemplo 7, se demuestra que los métodos se pueden emplear para modificar ácidos nucleicos extracromosómicos, como plásmidos, cromosomas artificiales bacterianos, cósmidos, fagémidos y similares.

## EJEMPLO 6

35 **El tiempo de inducción, la cantidad de ADN y la longitud del brazo de homología afectan a la eficiencia de la focalización**

40 *Tiempo de inducción.* La FIG. 3 muestra el efecto del tiempo de inducción sobre la recombinación. La cadena DY330 se cultivó a 32°C a una OD<sub>600</sub>=0.4 a 0.8, fue inducida por calor a 42 °C durante los tiempos indicados y después se hizo electrocompetente (véase el Ejemplo 2 para obtener una descripción de los métodos). Se empleó una cassette de *cat* lineal (10 ng) para focalizarse en los genes *cIII kil gam* del profago. Se trazaron los recombinantes Cm<sup>R</sup> totales frente al tiempo de inducción.

45 La inducción de la expresión del operón PL durante solo cinco minutos mejoró la actividad de recombinación. La FIG. 3 revela que a los 7,5 minutos de inducción por calor se alcanza una eficiencia máxima. Este nivel máximo se mantiene para los tiempos de inducción de 7,5 a 17,5 minutos, con alguna reducción apareciendo en los tiempos superiores a 17,5 minutos. La expresión del operón PL durante más de 60 minutos provoca la muerte cerebral.

50 Las células que albergan el profago lambda defectuoso se pueden cultivar a temperaturas diferentes a 32 °C. Por lo general, resulta recomendable cultivar las células a temperaturas superiores a 37 °C, porque estas temperaturas provocan una inactivación parcial del represor *cl* y una expresión más suave del promotor PL. En general, también es recomendable cultivar células a temperaturas inferiores a 20°C, debido al crecimiento ralentizado. Una persona con conocimientos en el campo reconocerá también que existe un alto grado de flexibilidad con respecto al tiempo y la temperatura de inducción. Por ejemplo, la expresión de genes de recombinación de lambda por parte del operón PL podría inducirse a temperaturas de tan solo 38 °C, generalmente permitiendo tiempos de inducción superiores. El límite para la expresión de la proteína en *E. coli* es de unos 45 °C.

55 Cantidad de ADN donante. La FIG. 4 muestra el efecto de la cantidad de la cassette de ADN lineal sobre la recombinación. La cadena DY330 se cultivó a 32°C hasta una OD<sub>600</sub>=0.4 a 0.8, fue inducida a 42 °C durante 15 minutos

y después se hizo electrocompetente. Se emplearon diferentes cantidades (1, 10, 100, 300, 1.000 ng) de una cassette de *cat* lineal (1 kbp de longitud) para focalizarse en los genes *cIII kil gam* del profago. Se trazaron los recombinantes  $Cm^R$  totales frente a la cantidad de ADN a 42 °C.

5 La Fig. 4 muestra que la eficiencia de focalización aumentó en una relación prácticamente lineal con el incremento de la concentración del ADN donante, en el rango de  $10^8$  (1 ng) a  $10^{10}$  (100 ng) moléculas por electroporación. Un nivel de saturación de ADN lineal se alcanza a  $3 \times 10^{10}$  moléculas que producen  $7.5 \times 10^4$  recombinantes por  $\sim 2 \times 10^8$  células electroporadas. Así, los métodos de la presente invención se pueden practicar con una amplia gama de concentraciones de oligonucleótidos.

10 *Longitud de homología.* La FIG. 5 muestra el efecto de la longitud del brazo homólogo sobre la recombinación. Las cadenas DY330 (*recA+*)(círculos negros) y DY331 (*recA-*)(círculos blancos) fueron cultivadas a 32°C, inducidas a 42°C durante 15 minutos y posteriormente se volvieron electrocompetentes. Se empleó una cassette de *cat* lineal (100 ng) para focalizarse en los genes *cIII kil gam* del profago. La longitud del brazo homólogo de la cassette se varió de 0 a 1.000 bp. Los cebadores que contenían las homología de 0 a 50 bp fueron sintetizados químicamente como se describe (Fig. 2). La cassette que contenía brazos homólogos de 1.000 bp se realizó mediante PCR empleando  
15 cebadores a 1.000 bp de distancia a cada extremo de una alteración (*cIII kil gam*)<>*cat* existente en la célula. Se trazaron los recombinantes  $Cm^R$  totales frente a la longitud del brazo homólogo.

20 Se realizaron varios pares de cebadores para amplificar la cassette de *cat* a fin de focalizar el cromosoma y cada par se diseñó con una longitud diferente de homología en los flancos. La longitud del segmento de homología en los cebadores variaba por incrementos de bases de 10 a 50 bases. Se realizó un conjunto encadenado de cassettes de *cat* lineales con los cebadores. Otra cassette de *cat* lineal se construyó flanqueada por 1.000 bp de homología. Este conjunto de ADN lineales fue posteriormente testado para determinar la eficiencia de focalización de recombinación, tal y como se muestra en la FIG. 5. No se encontró ningún recombinante con 10 bp de homología y en cada uno de los tres experimentos con 20 bp de homología se encontraron menos de 10 recombinantes. De 20 bp a 40 bp de homología, la recombinación homóloga aumentó en cuatro órdenes de magnitud. De 40 bp a 1.000 bp de homología, la recombinación  
25 aumentó solamente 10 veces.

Estos datos indican que los métodos divulgados en el presente se pueden practicar con brazos de homología sorprendentemente cortos, de tan solo 20-40 residuos. No obstante, los brazos de homología de 30 o más residuos aumentan la eficiencia.

## EJEMPLO 7

### 30 Sustitución de genes en los plásmidos y BAC: clonación *in vivo*

Para determinar si este método se podría emplear para modificar ADN de plásmido, se siguieron los procedimientos descritos en los Ejemplos 2 y 3 para modificar el plásmido pGB2, un derivado de pSCIOI (Bernardi *et al.*, Nucleic Acids Res. 12: 9415-9426, 1984). Una cassette de *cat* fue sintetizada *in vitro* y recombinada *in vivo* con pGB2 para sustituir el gen de resistencia a la espectinomicina por *cat*, confiriendo  $Cm^R$  a la célula portadora del plásmido recombinante.

35 El mismo experimento, realizado con derivados de pBR322, generó recombinantes, pero se unieron en tándem a plásmidos no recombinantes como dímeros y multímeros superiores. La inducción de la expresión de Gam de nuestro profago inactiva la nucleasa RedBCD. En ausencia de RecBCD, los derivados de pBR322 se replican mediante un modo de círculo rodante (Feiss *et al.*, Gene 17:123-130, 1982), y el plásmido se convierte de monómeros en multímeros. Esto es específico para los replicones de tipo pBR322, dado que el tipo pGB2 no formó multímeros.

40 Para generar recombinantes simples de derivados de pBR322, el protocolo fue modificado mediante coelectroporación de la cadena DY331 de *recA* con ADN de plásmido circular (0,1 ng) y una cassette de fármaco lineal. Los monómeros del plásmido recombinante fueron fácilmente seleccionados y aislados.

Además de los plásmidos, el método también es adecuado para focalizar genes en cromosomas artificiales bacterianos, cromosomas artificiales fagémidos, cromosomas artificiales de levaduras, cósmidos y similares. La recombinación  
45 homóloga entre las secuencias de ácido nucleico diana de los BAC y los oligonucleótidos sintéticos (como fragmentos de dsADN o fragmentos de PCR) se realiza en células bacterianas que portan el profago lambda defectuoso mostrado en la FIG 1 y descrito en el Ejemplo 1. Los oligonucleótidos sintéticos (como los fragmentos de PCR o fragmentos de dsADN anillados cortos) son electroporados en células bacterianas, tal y como se ha descrito en el Ejemplo 2. El análisis para detectar los eventos de recombinación fructuosos se realiza mediante amplificación por PCR selectiva, empleando cebadores específicos para la secuencia introducida, o bien mediante planteamientos de amplificación selectiva tales como el ensayo de TaqMan (TaqMAMA) descrito por Cha *et al.*, PCR Applications and Methods 2: 14-20, 1992, o por hibridación directa empleando sondas específicas. Empleando este planteamiento, se han observado frecuencias de recombinación de hasta 1:500, independientemente de la cadena focalizada. Así, este sistema resulta extremadamente útil para la manipulación y la rápida detección de recombinantes en vectores de BAC. Las eficiencias  
50 inesperadamente elevadas eliminan la necesidad de introducir marcadores seleccionables o de modificar estos marcadores en el BAC.



**EJEMPLO 8**

**Necesidad de RecA para la recombinación focalizada**

5 La necesidad de RecA se testó en la recombinación homóloga focalizada repitiendo el experimento descrito en la Fig. 5, pero utilizando una cadena de *recA*. Sorprendentemente, la eficiencia de la recombinación se vio reducida solamente unas 10 veces en el mutante de *recA*- para las longitudes de brazo testadas (Fig. 5). De este modo, la función de RecA no es necesaria y las cadenas de *recA*- median una recombinación homóloga eficiente con fragmentos de ADN lineales que tienen brazos de homología de 30 bp o más.

10 No se esperaba este resultado. Se sabe que el sistema de recombinación de  $\lambda$  funciona en células que carecen de la función de RecA bacteriana (Brooks y Clark, J. Virol. 1: 283-293, 1967). No obstante, la recombinación en mutantes de *recA* en condiciones empleadas por otros se reduce más de 50 veces con respecto a los niveles observados en células *recA*<sup>+</sup> (Stahl *et al.*, Genetics 77: 395-408, 1974; Murphy, J. Bacterid. 180: 2063-2071, 1998).

**EJEMPLO 9**

**Los genes de lambda promueven la recombinación focalizada de dsADN en E. coli de tipo salvaje**

15 Para determinar qué genes de lambda promueven la recombinación focalizada de dsADN, se generó un conjunto de deleciones de sustitución en el operón PL del profago, empleando cassettes de *cat* y *amp*. En la columna central de la Tabla 4, los paréntesis indican las deleciones realizadas por el evento de recombinación dentro del profago (véase la Fig. 1 para obtener un mapa lineal de los genes del profago). Cada una de estas deleciones realizadas se verificó estructuralmente mediante análisis por PCR y se testó para determinar la recombinación focalizada de una cassette del gen *tet* en *galk*.

20 Las células electrocompetentes de las cadenas indicadas en la Tabla 4 fueron inducidas por calor durante 15 minutos a 42 °C y electroporadas con 10 ng de *galk*<>*tet* lineal. Los resultados se presentan como número total de recombinantes Tet<sup>R</sup> por electroporación en la columna derecha de la Tabla 4.

Tabla 4

Cadena	Profago	Recombinante
DY330	tipo salvaje	4.100
DY392	( <i>hin-int</i> )<> <i>amp</i>	2.000
DY351	( <i>sieB-kil</i> )<> <i>cat</i>	4.400
DY386	( <i>hin-ini</i> )<> <i>amp</i> ( <i>sieB-kify</i> )<> <i>cat</i>	1.650
DY349	( <i>gam</i> )<> <i>cat</i>	0
DY360	( <i>bet</i> )<> <i>cat</i>	0
DY359	( <i>Exo</i> )<> <i>cat</i>	0

25 La Tabla 4 muestra que solamente las deleciones de *exo*, *bet* y *gam* afectaron a la recombinación focalizada de *galk*<>*tet*. La deleción de cualquiera de estos tres genes eliminó la recombinación, mientras que la deleción de todos los demás genes en el operón pL tuvo poco o ningún efecto.

30 Para demostrar que la sustitución *gam*<>*cat* no fue polar en *bet* y *exo*, el gen *gam* fue expresado en trans y se demostró que complementaba el defecto. Así, a pesar de que en los estudios aquí descritos se empleó el operón PL completo, solamente las funciones de *exo*, *bet* y *gam* son necesarias para la recombinación con cassettes de dsADN de más de 100-200 bp producidas mediante PCR.

**EJEMPLO 10**

**Recombinación eficiente con ADN de cadena única**

35 Para evaluar la recombinación entre ADN de cadena única exógeno y el cromosoma de *E. coli*, se realizaron los experimentos descritos en este Ejemplo. Por otra parte, los experimentos también se refieren al papel de los genes de recombinación de *recA* bacteriana en la mediación de los efectos observados.

40 La expresión del operón PL se indujo por calor, las células se hicieron competentes por electroporación, y los oligonucleótidos de 70-mer se electroporaron en las células, tal y como se describe en el Ejemplo 2. Tanto las cadenas de HME9 como de HME10 contienen la mutación ámbar de *galk*, y por tanto son de fenotipo Gal<sup>-</sup>. Difieren entre sí en que HME9 es *recA*<sup>+</sup>, mientras que HME10 es *recA*<sup>-</sup>. Por tanto, en este experimento se comparan las eficiencias de recombinación en las células de *recA*<sup>+</sup> y *recA*<sup>-</sup>.

Los oligonucleótidos de cadena única de 70-mer empleados en este experimento fueron diseñados para restaurar la actividad del gen *galK* de tipo salvaje (en adelante, *galK*<sup>+</sup>) tras una recombinación con éxito, produciendo así un fenotipo *Gal*<sup>+</sup> (es decir, capacidad para crecer en medios mínimos con galactosa como única fuente de carbono). Los 70-mer correspondientes a la cadena de ADN sin plantilla de transcripción del *galK* fueron

5 5'AAGTCGCGGTCGGAACCGTATTGCAGCAGCTTTATCATCTGCCGCTGGAC GGCGCACAAATCGCGCTTAA (SEQ ID NO 11).

El ADN de cadena única de 70-mer de la SEQ ID NO 11 o de su complemento fue electroporado en células. Alternativamente, los dos 70-mers fueron primero anillados entre sí y posteriormente electroporados en células como ADN de cadena doble. Se identificó un evento de recombinación con éxito mediante la restauración del fenotipo *Gal*<sup>+</sup>.

10 La Tabla 5 indica el número de recombinantes de *galK*<sup>+</sup> por célula viable,  $\times 10^{-4}$ .

La Tabla 5 presenta el número de recombinantes observados  $\times 10^{-4}$  tras la electroporación de las cadenas de HME9 o HME10 con ADN en las formas indicadas. La recombinación eficiente se observó con el ADN de cadena doble, de forma similar a la previamente descrita en los Ejemplos 3 y 5. Sorprendentemente, el ADN de cadena única fue más o menos igual de eficiente o incluso más que el ADN de cadena doble en la producción de recombinación homóloga, independientemente de la cadena utilizada. En este experimento, las eficiencias de recombinación para el ADN de cadena doble fue aproximadamente de una de cada 3.800 células ( $2.6 - 2.7 \times 10^{-4}$ ), mientras que la eficiencia de recombinación del ADN de cadena única inverso fue entre 3 y 7 veces superior. En otras cadenas (p.ej., DY374), se ha observado la recombinación de un ADN lineal de cadena única tan frecuente como un *nadA*<sup>+</sup> recombinante por 20 células viables. Por otra parte, la recombinación eficiente se observó tanto en las células de *recA*<sup>+</sup> como en las de *recA*<sup>-</sup>, estableciendo que los eventos de recombinación no precisaban de productos genéticos de *recA* bacteriana.

15  
20

**Tabla 5**

Cadena utilizada	ADN utilizado en la electroporación		
	dsADN (aprox. 1 µg)	cc-ssADN (aprox. 0,7 µg)	cw-ssADN (aprox. 0,6 µg)
HME9 <i>galK</i> <sup>am</sup> λC1857Δ( <i>cro-bio</i> )	2,6	18,5	2,2
HME10 <i>galK</i> <sup>am</sup> λC1857Δ( <i>cro-bio</i> ) <i>recA</i> <sup>-</sup>	2,7	9	0,35

cc=directa (cadena con la orientación del extremo 5' al 3' con respecto a la transcripción) cw=inversa (cadena con la orientación del extremo 3' al extremo 5')

#### EJEMPLO 11

##### 25 Generación eficiente de grandes deleciones mediante recombinación con ADN de cadena única

El Ejemplo 10 demuestra una recombinación mediada por lambda eficiente utilizando ssADN para generar un cambio de una sola base en el cromosoma de *E. coli*. Este ejemplo demuestra una alta eficiencia similar cuando el planteamiento se utiliza para generar grandes deleciones.

Utilizando los métodos descritos en los Ejemplos 2, 3 y 10, la expresión del operón de lambda fue inducida por calor, las células se volvieron competentes por electroporación y el ADN de cadena única de 70-mer de la SEQ ID NO: 11 o su complemento fue electroporado en las células. El ADN fue electroporado en dos cadenas: una que contenía la mutación ámbar de *galK* y una en la que el gen *galK* fue interrumpido mediante una cassette de *cat-sacB* insertada con precisión en la posición de la mutación ámbar en *galK*. Por lo demás las cadenas eran genéticamente idénticas. Se identificó un evento de recombinación con éxito mediante la restauración del fenotipo *Gal*<sup>+</sup>. La Tabla 6 indica el número de recombinantes de *galK*<sup>+</sup> por célula viable,  $\times 10^{-4}$ .

30  
35

Los datos de la Tabla 6 demuestran que la eficiencia de la recombinación es similar empleando el mismo oligonucleótido, independientemente de que la recombinación de ssADN mediada por lambda vaya a ser utilizada para generar el cambio de una única base o una deleción de gran tamaño, la eliminación de la cassette *cat-sacB* de 3264 bp. En ambos casos, el método es altamente eficiente para generar recombinantes.

Tabla 6

Cadena utilizada	ADN utilizado en la electroporación	
	cc-ssADN (200 ng)	cw-ssADN (200 ng)
HME6 <i>galK<sup>am</sup>λCI857Δ(cro-bio)</i>	15	3
HME31 <i>galK&lt;&gt;catsacB λCI857Δ(cro-bio)</i>	10	0,5

La longitud que resulta razonable para sintetizar químicamente limita la longitud de los oligonucleótidos de cadena única empleados para la recombinación. En el presente se demuestra que dos oligonucleótidos que presentan una región de solapamiento complementaria en sus extremos 3', cuando son coelectroporados en células DY411, pueden anillarse y generar recombinantes con ADN cromosómico o extracromosómico (Figura 14). Esta recombinación necesita la inducción del operón PL y de las funciones Gam, Beta y Exo. Se creó mediante recombinación una mutación de *galK* en la que la cassette *kan* fue introducida en *galK* a fin de someter a delección 34 bp del gen *galK*. Se sintetizaron dos oligos que tenían 70 bases de largo, con 34 bases de la región sometida a delección en sus extremos 3', que eran complementarias y podían actuar para anillar los dos oligos juntos. El extremo 5' de cada oligo contenía 36 bases de homología a cada lado de la delección de 34 bp causada por *kan*. Cada uno de los oligos no podía generar recombinantes *gal+* por sí solo, pero juntos generaron hasta 10<sup>5</sup> recombinantes por 10<sup>8</sup> células electroporadas. Los oligos con la misma secuencia pero acortados en su extremo 3' para solaparse solamente en 2 bases no produjeron recombinantes. Sin embargo, los solapamientos de 10 o más bases generaron recombinantes. No fue necesario un preanillamiento y los dos oligos se pueden mezclar y utilizar directamente para la electroporación.

Si los extremos de los solapamientos se rellenan con ADN polimerasa, se genera un dúplex de 104 bp. Este dsADN sólo genera un número ligeramente mayor de recombinantes que el ADN con extremos prominentes de 10 a 34 bases. De este modo, los oligonucleótidos con un solapamiento múltiple (en más de 10 bases) de cualquier número par pueden ser empleados para producir recombinantes, en los que los oligonucleótidos más externos tienen solapamientos en el extremo 5'. Los oligonucleótidos del extremo también presentan 30-50 bases de homología con la región de focalización. El uso de múltiples oligonucleótidos de solapamiento permite la producción de largos sustratos de recombinación sin emplear la PCR. El oligonucleótido o los oligonucleótidos centrales pueden ser cualquier cassette que se tenga previsto emplear para la recombinación de dsADN. Esta recombinación con oligonucleótidos de solapamiento que tienen los extremos 5' protuberantes es sumamente eficiente con Exo, Beta, y Gam, pero puede ser recombinada solamente por Beta (sin Exo y Gam) en la célula. Esto simplifica mucho el procedimiento de recombinación (dado que solamente se necesita Beta). A pesar de que el ADN dúplex de 104 bp se recombina de forma más eficiente si están presentes Exo, Beta y Gam, la recombinación también ocurre en ausencia de Gam, aunque con una eficiencia menor (el dúplex necesita tanto Exo como Beta para la recombinación).

Se pueden generar oligonucleótidos sintéticos de solapamiento similares con protuberancias en el extremo 3' de 34 bases, que pueden ser coelectroporadas en las células. Éstos también son recombinados en las dianas definidas por homología en los extremos. Una vez más, solamente se necesita Beta para esta recombinación. En este caso, Exo no es necesario ni estimula la recombinación. En una realización, se pueden solapar múltiples oligonucleótidos como anteriormente para abarcar una distancia mayor. Siempre que los oligonucleótidos más externos tengan extremos 3' protuberantes, la recombinación será independiente de Exo. Las eficiencias del presente sistema permiten la detección de recombinantes en este caso.

Los Ejemplos 10 y 11, en su conjunto, documentan que los métodos divulgados en el presente se pueden practicar con oligonucleótidos de ssADN. Este sorprendente resultado permite la recombinación homóloga de alta eficiencia con ADN sintético de cadena única o doble.

El presente sistema permite ampliar el límite del tamaño de los oligonucleótidos sintéticos de forma drástica, al solapar los oligonucleótidos. Por otra parte, el sistema permite realizar la recombinación de estos ADN solamente con Beta, o Exo con Beta, pero sin Gam. La recombinación sin necesidad de Gam es importante, porque Gam es la función tóxica que representaba un factor de limitación en los métodos anteriormente descritos. Dado que el presente sistema solamente necesita Beta, se puede emplear un promotor constitutivo.

**EJEMPLO 12**

**Efecto de la longitud de ssADN sobre la eficiencia de la recombinación**

En los Ejemplos 10 y 11, la recombinación mediada por lambda se utilizó para incorporar de manera eficiente ssADN de 70-mer en el cromosoma de *E. coli*. En este ejemplo, se investigó el efecto de la longitud de los oligonucleótidos sobre la eficiencia de la recombinación.

5 Empleando los métodos descritos en los Ejemplos 2, 3, 5, 10 y 11, la expresión del operón de lambda fue inducida por calor, las células se volvieron competentes por electroporación y los oligonucleótidos del ssADN (200 ng cada uno) fueron electroporados en células de la cadena HME9 de *E. coli*. Los oligonucleótidos de ssADN electroporados incluían los 70-mer de la SEQ ID NO: 11, unos 60-mer construidos eliminando los 5 últimos nucleótidos de los extremos 5' y 3' de la SEQ ID NO: 11, y unos 50-mer, 40-mer, 30-mer, o 20-mer construidos eliminando los 10, 15, 20 o 25 últimos nucleótidos, respectivamente, de los extremos 5' y 3' de la SEQ ID NO: 11. Al igual que en el Ejemplo 10, los oligonucleótidos del ssADN empleados en este experimento estaban todos diseñados para restaurar el gen *galK+* tras la recombinación con éxito, confiriendo así a la célula un fenotipo Gal+. La Tabla 7 indica el número de recombinantes *galK+* por célula viable, x 10<sup>-4</sup>.

**Tabla 7**

15 **Longitud del oligonucleótido**

	0	20	30	40	50	60	70
Eficiencia (x10 <sup>-4</sup> )	0,004	0,01	0,47	4	4	6	22

20 Como demuestran los datos de la Tabla 7, la eficiencia de la recombinación aumenta cuando se incrementa la longitud del ssADN. La eficiencia de la recombinación era baja cuando el ssADN empleado era de 20-mer, pero aumentaba considerablemente cuando era de 30-mer. La eficiencia era prácticamente óptima con 40-mer y se incrementaba hasta 1 de cada 450 células viables a 70-mer. Por tanto, ejemplos concretos de la invención emplean moléculas de ADN de cadena única que tienen al menos 40 nucleótidos de longitud.

25 Sin ánimo de limitarse a una única explicación de los efectos observados, los inventores consideran actualmente que la relación observada entre la longitud y la eficiencia puede reflejar los datos publicados que indican que la proteína Beta de lambda se une de forma estable a secuencias de ADN de 36 bases o más de longitud, pero que no se une de forma tan eficiente a los oligonucleótidos más cortos (Mytbili *et al*, Gene 182: 81-87, 1996).

**EJEMPLO 13**

**La proteína Beta de lambda media una recombinación eficiente con ssADN**

30 Para determinar si la proteína Beta de lambda era suficiente para mediar la recombinación entre ssADN exógeno y el cromosoma de *E. coli*, se investigó la eficiencia de la recombinación en una cadena que expresaba Beta de lambda, pero no Exo ni Gam.

Para estos experimentos, se empleó la cadena HME43. Su genotipo es idéntico al de la cadena HME6, salvo por el hecho de que el profago lambda contiene deleciones genéticas adicionales, desde *int* hasta *exo* y desde *gam* hasta *N* (véase la FIG. 1). Por otra parte, el gen *cat* que confiere el fenotipo CmR se inserta entre *attL* y *bet*.

35 Empleando los métodos descritos en los Ejemplos 2, 3, y 10-12, la expresión del operón de lambda modificado fue inducida por calor, las células se volvieron competentes por electroporación y el ssADN de 70-mer de la SEQ ID NO: 11 (200 ng) fue electroporado en las células. Empleando este procedimiento, la cadena HME43 expresa la proteína Beta de lambda, pero no expresa *gam*, *exo* ni ningún gen codificado por el profago. Se identificó un evento de recombinación con éxito mediante la restauración del fenotipo Gal+. La Tabla 8 indica el número de recombinantes *galK+* por célula viable, x 10<sup>-4</sup>.

40 **Tabla 8**

Cadena	Modificaciones del profago	Eficiencia de recombinación (x 10 <sup>-4</sup> )
HME43	(int-aro)<->cat, (gam-N)<->Δ	7,7

5 A diferencia del Ejemplo 9, que empleaba ADN de doble cadena generado por PCR, los datos presentados en este ejemplo establecen que Beta de lambda es suficiente por sí sola para mediar una recombinación eficiente entre ssADN y el cromosoma de *E. coli*. Por otra parte, se pueden emplear dos o más oligonucleótidos de solapamiento, si tienen un extremo 3' protuberante y más de unos 10 bp de solapamiento. Los oligonucleótidos de solapamiento con un extremo 5' protuberante también promueven la recombinación homóloga con Beta solamente. No obstante, en caso de extremo 5' protuberante, *exo* y *gam* (o una exonucleasa similar y la función de inhibición de RecBCD) parecen mejorar la eficiencia máxima.

10 Una modificación del método consiste en poner el ADN que codifica otros polipéptidos de unión a ssADN bajo el control del promotor PL. Por ejemplo, la cadena HME43 es también modificada para someter *bet* a delección e insertar ADN que codifica P22 Erf, RecT, o Rad52. La expresión del polipéptido de unión a ssADN es inducida por un cambio de temperatura, como para la inducción de la expresión *bet* de lambda. *Exo* y *Gam*, o proteínas con una función similar, también se puede poner bajo el control del promotor PL. Por otra parte, se pueden emplear otros promotores inducibles o constitutivos.

## 15 EJEMPLO 14

### La combinación *ex vivo* de ssADN con Beta de lambda media una recombinación homóloga eficiente

El ADN de cadena única puede ser combinado con la proteína Beta de lambda antes de la electroporación en las células y mediar una recombinación eficiente entre el ssADN y el ADN hospedador.

20 Las proteínas Beta de lambda se pueden preparar con técnicas conocidas en el campo (Karakousis et al., J. Mol. Biol. 276: 721-731, 1998), y preincubarse a 37 °C con oligonucleótidos de cadena única de 20-mer o más de longitud. En este ejemplo, los oligonucleótidos de ssADN de la SEQ ID NO: 11, 60-mer construidos eliminando los 5 últimos nucleótidos de los extremos 5' y 3' de la SEQ ID NO 11, y 50-mer construidos eliminando los 10 últimos nucleótidos de los extremos 5' y 3' de la SEQ ID NO 11. Típicamente, la concentración de proteína Beta de lambda es de unos 2,5 µM y la concentración de ADN de unos 5 µM, aunque el método es efectivo con un amplio rango de concentraciones de proteína y ADN (por ejemplo, de 0,1 µM a 10 mM de proteína, y 0,01 µM a 10 mM de ssADN). Alternativamente, la proteína Beta y el ssADN pueden ser coelectroporados en las células sin premezclado o preincubación.

30 El ADN y la proteína son electroporados en *E. coli* empleando métodos descritos en los Ejemplos 2 y 3. En este ejemplo, se emplea la cadena BMB 43, aunque se pueden utilizar muchas otras cadenas. La expresión del operón de lambda modificado en un conjunto de células es inducida por calor y un segundo conjunto de células se mantiene a 32 °C. Ambos conjuntos de células se vuelven competentes por electroporación y el ssADN de 70-mer de la SEQ ID NO 11 (200 ng) es electroporado tanto en células inducidas por calor como en células no inducidas. Empleando este procedimiento, la cadena HME43 expresa la proteína Beta de lambda tras un cambio de temperatura a 42 °C, pero no expresa Beta de *bet* ni ningún otro gen codificado por el profago en ausencia de un cambio de temperatura. Se identifica un evento de recombinación con éxito mediante la restauración del fenotipo Gal+.

35 En este experimento, se observa una recombinación de alta eficiencia tanto en las células inducidas por calor como en las no inducidas. Por otra parte, se cree que se observó una recombinación de alta eficiencia más o menos igual al seguir estas técnicas en cadenas de *E. coli* que no contienen ningún gen del profago lambda.

40 Este planteamiento puede ser modificado sustituyendo otros polipéptidos de unión a ssADN por Beta de lambda, como p22 Erf, RecT y Rad52. La secuencia de ácido nucleico diana puede estar en el cromosoma bacteriano, o en ADN exógeno como un cromosoma artificial bacteriano, un cromosoma artificial fagémido, plásmido, cósmido o similares. Por otra parte, no existe ningún requisito concreto de una especie bacteriana específica; estos polipéptidos de unión a ADN de cadena única mediarán una recombinación eficiente en un amplio rango de bacterias. En efecto, estos polipéptidos mediarán una recombinación eficiente también en células eucariotas, como en el Ejemplo 15.

## EJEMPLO 15

### 45 La proteína Beta de lambda media una recombinación homóloga eficiente en células eucariotas

50 El planteamiento *ex vivo* descrito en el Ejemplo 14 puede emplearse para focalizar genes de células eucariotas para la recombinación homóloga. En células eucariotas, la transfección del ssADN con la proteína Beta de lambda puede realizarse mediante electroporación, como en los Ejemplos 2, 3 y 14, o a través de los métodos de Chang *et al.*, Biochimica et Biophysica Acta, 153-160 (1992), Keating y Toneguzzo, Bone Marrow Purging and Processing, 491-498 (1990), u otros protocolos de electroporación conocidos en la técnica. Por otra parte, en la técnica se conocen diversos medios para los métodos de transferencia macromolecular, incluyendo el coprecipitado de ADN con fosfato cálcico (Ausubel *et al.*), la transfección mediada por DEAE-dextrano (Matthews et al, Experimental Hematology. 21: 697-702, 1993), la transfección mediada por polibreno (Costello *et al.*, Gene Therapy 5: 596-604, 2000), la microinyección (Davis *et al.*, Blood 95: 437-44, 2000), la fusión de liposomas y lipofección (Veit *et al.*, Cardiovascular Research. 43:

808-22, 1999), la fusión de protoplastos (Schafner, Proc. Natl. Acad. Sci. USA 77: 2163, 1980), la transferencia mediada por adenovirus inactivado (Wagner *et al.*, Proc Natl Acad Sci USA 89:6099-6103, 1992), la transferencia mediada por el virus hemaglutinante del Japón (HVJ) (Morisluta *et al.*, Journal of Clinical Investigation 93: 1458-1464, 1994), biolítica (bombardeo de partículas) y similares. Cualquiera de estos planteamientos de transferencia macromolecular resulta adecuado. El diseño de moléculas de dsADN para facilitar la recombinación homóloga con genes eucariotas es bien conocido en la técnica (por ejemplo, tal y como se describe en Mansour, Nature 336: 348-352, 1988; Shescly, PNAS 88: 4294-4298, 1991; Capocchi, M. R., Trends in Genetics 5: 70-76, 1989; Patente estadounidense nº 6.063.630).

Las células que se pretenden transfectar con ADN exógeno son combinadas con una estructura de ADN que comprende el ADN exógeno, secuencias de ADN de focalización y, opcionalmente, ADN que codifica uno o más marcadores seleccionables. La combinación resultante es tratada de tal manera que la estructura de ADN se introduce en las células. Esto se consigue sometiendo la combinación a electroporación, microinyección u otro método de introducción de ADN en las células de vertebrados. Una vez en la célula, el ssADN exógeno se integra en el ADN celular mediante recombinación homóloga entre las secuencias de ADN de la estructura de ADN y las secuencias de ADN del ADN celular.

Por ejemplo, el ácido nucleico diana es el gen beta-globina de células madre hematopoyéticas de una paciente con anemia drepanocítica (Beutler, Disorders of Hemoglobin, Ch. 107 in Harrison's Principles of Internal Medicine, 14<sup>th</sup> ed. © 1998. El gen beta-globina de drepanocitos contiene una mutación puntual que sustituye un Val por un Glu en la posición seis de la cadena de polipéptidos, resultando en una hemoglobina anómala que tiende a una polimerización inapropiada. Los métodos de la presente invención se pueden emplear para corregir la mutación.

Las células madre hematopoyéticas de un paciente con drepanocitos son aisladas, cultivadas y expandidas *ex vivo* como se conoce en la técnica (Brugger Seminars in Hematology 37(1 Suppl 2): 42-49, 2000; Dao *et al.*, Blood 92:4612-21, 1998; Aglierta *et al.*, Haematologica 83: 824-4S, 1998; Emerson, Blood 87: 3082-8, 1996). Un oligonucleótido de ADN de 60-mer de la SEQ ID NO 12 (A ACAGACACC ATGGTGCACC TGACTCCTGA GGAG AAGTCT GCCGTTACTG CCCTGTGGGG) es sintetizado y parcialmente purificado mediante técnicas estándar (Pflücker *et al.* Acta Biochimica Polonica. 43:37-44, 1996; Anderson *et al.*, Applied Biochemistry & Biotechnology 54: 19-42, 1995.)

Tras el cultivo y la expansión *ex vivo*, alrededor de  $10^6$  células madre hematopoyéticas son suspendidas en 0,4 mL de PBS con un 0,1% de glucosa, unos 10  $\mu$ M de proteína Beta de lambda purificada y aproximadamente 1  $\mu$ g de oligonucleótidos del ssADN de la SEQ ID NO: 12. La suspensión celular es electroporada en una cubeta de 1-mL a 280V y 250  $\mu$ F con un Gene Pulser (Bio-Rnd Laboratories Inc., Hercules, California, USA). A continuación, las células se colocan en placas y son cultivadas. Los recombinantes homólogos que contienen la mutación son identificados y aislados clonalmente, nuevamente expandidos *ex vivo*, y pueden ser devueltos al paciente o cultivados para otro estudio *in vitro*.

Aquellos versados en la técnica reconocerán que una amplia gama de concentraciones de ssADN y del polipéptido de unión a ssADN resultarán efectivas, como en el Ejemplo 14. Por ejemplo, tanto el ssADN como la proteína de unión a ssADN pueden estar presentes en concentraciones que oscilan entre 0,001  $\mu$ M y 100 mM; o entre 0.1 mM y 1  $\mu$ M; o entre 1  $\mu$ M y 100  $\mu$ M. La longitud del oligonucleótido se puede variar de acuerdo con los parámetros presentados en el Ejemplo 12. No existe ningún límite máximo concreto para la longitud del oligonucleótido. Por otra parte, se pueden incluir dos o más oligonucleótidos que tienen extremos 5' complementarios, creando así extremos 3' protuberantes que son sustratos efectivos para los polipéptidos de unión a ssADN como Beta de lambda. Por otra parte, RecT, P22 Erf, Rad52 y otros polipéptidos de unión a ssADN de reparación de roturas bicatenarias pueden ser sustituidos por Beta de lambda. Las condiciones de cultivo y electroporación son fácilmente variables sin reducir de forma material la recombinación homóloga. Por otra parte, el ácido nucleico puede ser introducido en la célula mediante cualquier método de transferencia macromolecular adecuado.

Se pueden emplear otros tipos de células madre para corregir defectos genéticos concretos asociados con las células obtenidas de estas células madre. Estas otras células madre incluyen células madre epiteliales, hepáticas, pulmonares, musculares, endoteliales, mesenquimales, neurales y óseas.

Alternativamente, determinados estados patológicos se pueden tratar modificando el genoma de las células de forma que no corrija un defecto genético en sí, sino que proporcione la suplementación del producto genético de un gen defectuoso. Por ejemplo, las células endoteliales son preferibles como diana para la terapia genética humana orientada a tratar trastornos que afectan a factores normalmente presentes en la circulación sistémica. En modelos de estudio con perros y cerdos, se ha demostrado que las células endoteliales forman cultivos primarios, que se pueden transformar con ADN en cultivo, y que son capaces de expresar un transgen tras su reimplantación en injertos arteriales en el organismo hospedador. Wilson *et al.*, Science, 244: 1344, 1989; Nabel *et al.*, Science, 244: 1342, 1989. Dado que las células endoteliales forman parte integral del injerto, estas células transformadas se pueden emplear para producir proteínas que se secretarán en el sistema circulatorio y, de este modo, servirán como agentes terapéuticos en el tratamiento de trastornos genéticos que afectan a factores de la circulación. Algunos ejemplos de estas enfermedades

incluyen la diabetes por deficiencia de insulina, la deficiencia de alfa-1-antitripsina y hemofilia. Las células epiteliales, miocitos y hepatocitos también son tipos de células útiles para la producción terapéutica de proteínas.

5 El método también resulta útil para la inactivación o modificación de genes en células madre embrionarias (ES). Estas células han sido manipuladas para introducir transgenes. Las células ES se obtienen de embriones en fase de preimplantación cultivados *in vitro*. Evans *et al.*, Nature, 292, 154-156, 1981; Bradley *et al.*, Nature, 309, 255-258, 1984; Gossler, *et al.*, Proc. Natl. Acad. Sci. U.S.A., 83, 9065-9069, 1986; Robertson, *et al.*, Nature, 322, 445-448, 1986; Patente estadounidense nº 5.464.764. Los oligonucleótidos diseñados para focalizar segmentos de genes específicos de la célula ES se combinan con la proteína Beta de lambda u otro polipéptido de unión a ssADN y se introducen en  
10 células ES mediante electroporación u otros métodos de transformación. Los oligonucleótidos pueden ser diseñados como una serie de segmentos de solapamiento con el extremo 3' protuberante. Estas células ES transformadas pueden ser posteriormente combinadas con blastocitos de un animal no humano. A continuación, las células ES colonizan el embrión y pueden contribuir a la línea germinal del animal quimérico resultante (Jaenisch, Science, 240, 1468-1474, 1988).

15 Por ejemplo, las secuencias que codifican un gen de resistencia a la neomicina de un marcador de selección positivo son sintetizadas como una serie de oligonucleótidos de 70-mer de solapamiento, 20 pares de bases de solapamiento y extremos 3' protuberantes. Los oligonucleótidos terminales del extremo 3' están diseñados para insertarse en el segundo exón del gen *hox 1.1* de ratón, tal como se describe en la Patente estadounidense nº 5.464.764. Dado que los oligonucleótidos de solapamiento se combinan para codificar un gen de resistencia a neomicina sin promotor, solamente aquellos que se incorporen con éxito al segundo exón de *hox 1.1* de ratón focalizado expresarán el producto genético  
20 neo y tendrán el fenotipo de resistencia a la neomicina. La focalización está diseñada para proporcionar al gen de resistencia a neomicina sintético un promotor operable y un inicio de traslación obtenido del gen *hox 1.1* de ratón. El ADN de focalización también está diseñado de forma que sea poco probable que las incorporaciones aleatorias en otras partes del genoma de la célula ES estén operativamente unidas a cualquier promotor para permitir la transcripción y traslación.

25 La serie de oligonucleótidos de solapamiento con extremos 3' protuberantes (unos 200 nanogramos cada uno) se combinan con 10 µM de proteína Beta de lambda y se introducen en células ES por electroporación, empleando el Promega Biotech X-Cell 2000. Las células de crecimiento rápido son tripsinizadas, lavadas en DMEM, contadas y resuspendidas en una solución tampón que contiene 20 mM de HEPES (pH 7.0), 137 mM de NaCl, 5 mM de KCl, 0,7 mM de Na<sub>2</sub>HPO<sub>4</sub>, 6 mM de dextrosa, y 0,1 mM de beta-mercaptoetanol. Justo antes de la electroporación, los oligonucleótidos y la proteína Beta de lambda son añadidos a 10<sup>7</sup> ES células en cubetas de 1 ml cada una. Las células y el ADN son expuestos a dos pulsos secuenciales de 625V/cm a temperatura ambiente, se dejan en la solución tampón durante 10 minutos y, a continuación, se ponen en placas en medios no selectivos sobre un soporte biológico de células.

35 Después de dos días de crecimiento no selectivo, las células son tripsinizadas y se vuelven a colocar en placas sobre un medio de G418 (250 µg/ml). La selección positiva se aplica durante tres días. Debido a la elevada eficiencia de la recombinación mediada por Beta de lambda, la necesidad de una nueva selección (por ejemplo, selección negativa introduciendo un gen de timidina quinasa y la selección con ganciclovir) se puede obviar. Convenientemente transformadas, las células resistentes a G418 son cultivadas en medio no selectivo durante 2-5 días antes de la inyección en blastocistos, conforme al método de Bradley, en Teratocarcinomas and embryonic stem cells, a practical approach, editado por E. J. Robertson, URL Press, Oxford (1987), p. 125.  
40

Los blastocistos que contienen las células ES focalizadas son implantados en hembras pseudo-preñadas y se deja que se desarrollen a término. La descendencia quimérica es identificada mediante marcadores de manchas y los machos que muestran quimerismo son seleccionados para descendencia de reproducción. La descendencia que porta el alelo mutante se puede identificar por las manchas y la presencia del alelo mutante se puede reconfirmar a través de un análisis de ADN de la cola (tail-blot).  
45

De este modo, el método simplifica notablemente la construcción de knockouts genéticos y modificaciones genéticas en células ES. Además de su posible relevancia para la terapia genética de plantas, animales y seres humanos, el método simplificará la producción de animales transgénicos que presentan knockouts genéticos o modificaciones genéticas.

50 Como se describe para las células madre y en el Ejemplo 14, una amplia gama de concentraciones del ssADN y del polipéptido de unión a ssADN resultarán efectivas. Por ejemplo, tanto el ssADN como la proteína de unión a ssADN pueden estar presentes en cantidades que oscilan entre 0,001 µM y 100 mM; o entre 0,1 µM y 1 µM; o entre 1 µM y 100 µM. La longitud del oligonucleótido puede variar de acuerdo con los parámetros presentados en el Ejemplo 12. No existe ningún límite máximo concreto para la longitud del oligonucleótido. Por otra parte, se pueden incluir dos o más oligonucleótidos que tienen extremos 5' complementarios (por ejemplo, con extremos 5' complementarios de 10, 20, 30, 40 bp), creando así extremos 3' protuberantes que son sustratos efectivos para los polipéptidos de unión a ssADN como Beta de lambda. Por otra parte, RecT, P22 Erf, Rad52 y otros polipéptidos de unión a ssADN de reparación de roturas bicatenarias pueden ser sustituidos por Beta de lambda, en los mismos rangos que los descritos para Beta de lambda. Aquellos con conocimientos en la técnica reconocerán que las condiciones de cultivo y electroporación son fácilmente  
55

variables sin reducir de forma material la recombinación homóloga. Por otra parte, el ácido nucleico puede ser introducido en la célula mediante cualquier método de transferencia macromolecular adecuado.

## EJEMPLO 16

### Recombinación homóloga en plantas

- 5 Los métodos divulgados en el presente documento también son aplicables a la manipulación de células de plantas y en última instancia al genoma de toda la planta. Se han documentado diversas plantas transgénicas, incluyendo dicotiledóneas herbáceas, dicotiledóneas leñosas y monocotiledóneas. Para obtener un resumen, véase Gasser, *et al.* (1989), *Science*, 244,1293-1299. Se han desarrollado diversas técnicas de transferencia genética diferentes para producir estas plantas transgénicas y células vegetales transformadas. Una técnica empleó *Agrobacterium tumefaciens* como sistema de transferencia genética. Rogers, *et al.* (1986), *Methods Enzymol.*, 118, 627-640. Una transformación estrechamente relacionada utiliza la bacteria *Agrobacterium rhizogenes*. En cada uno de estos sistemas, se puede construir un vector de transformación vegetal Ti o Ri que contiene regiones fronterizas que definen la secuencia de ADN a insertar en el genoma vegetal. Estos sistemas se habían utilizado anteriormente para integrar de forma aleatoria ADN exógeno en genomas vegetales.
- 10
- 15 Preferiblemente, el ADN diseñado para la recombinación homóloga con una secuencia de ADN diana en las plantas se combina con la proteína Beta de lambda u otra proteína de ssADN y se transfiere directamente a los protoplastos de la planta mediante métodos análogos a los anteriormente empleados para introducir transgenes en los protoplastos. Las concentraciones de ADN y de las proteínas de unión a ssADN son las que se describen en el Ejemplo 15. Véase, p. ej., Paszkowski, *et al.* (1984), *EMBO J.*, 3,2717-2722; Hain, *et al.* (1985), *Mol. Gen. Genet.*, 199, 161-168; Shillito, *et al.* (1985), *Bio/Technology*, 3, 1099-1103; y Negrutiu, *et al.* (1987), *Plant Mol. Bio.*, 8, 363-373. Alternativamente, el vector de PNS se introduce en un liposoma que puede ser fusionado al protoplasto de una planta (véase, p. ej., Deshayes, *et al.* (1985), *EMBO J.*, 4, 2731-2738) o se inserta directamente en el protoplasto de la planta mediante microinyección intranuclear (véase, p. ej., Crossway, *et al.* (1986), *Mol. Gen. Genet.*, 202,179-185, y Reich, *et al.* (1986), *Bio/Technology*, 4, 1001-1004). La microinyección es el método preferible para la transfección de protoplastos. El ADN y las proteínas de unión a ssADN también pueden ser microinyectadas en inflorescencias meristémicas. De la Pena *et al.* (1987), *Nature*, 325, 274-276. Finalmente, los explantes de tejidos pueden ser transfectados mediante un microproyectil de alta velocidad revestido del ADN y de las proteínas de unión al ssADN análogo a los métodos empleados para la inserción de transgenes. Véase, p. ej., Vasil (1988), *Bio/Technology*, 6, 397; Klein, *et al.* (1987), *Nature*, 327, 70; Klein, *et al.* (1988), *Proc. Natl. Acad. Sci. 5 U.S.A.*, 85, 8502; McCabe, *et al.* (1988), *Bio/Technology*, 6, 923; y Klein, *et al.*, *Genetic Engineering*, Vol 11, J. K. Setlow editor (Academic Press, N.Y., 1989). Estos explantes transformados se pueden utilizar para regenerar por ejemplo diversos cultivos en serie. Vasil (1988), *Bio/Technology*, 6, 397.
- 20
- 25
- 30

Una vez que el ADN y la proteína de unión a ssADN han sido insertados en la célula vegetal mediante cualquiera de los métodos anteriores, la recombinación homóloga focaliza el oligonucleótido en el sitio apropiado del genoma de la planta. Como en ejemplos anteriores, el oligonucleótido puede ser una serie de ssADN de solapamiento con extremos 5' o 3' protuberantes. Dependiendo de la metodología empleada para la transfección, la selección se realiza en cultivos tisulares del protoplasto transformado o en la célula vegetal. En algunos casos, las células que se pueden someter a un cultivo tisular pueden ser obtenidas de una planta transformada de la generación F0 o de una generación posterior.

35

La composición de aminoácidos de diversas proteínas de reserva en el trigo y el maíz, por ejemplo, que se sabe que son deficientes en lisina y triptófano, también puede ser modificada. Los vectores de PNS pueden ser fácilmente diseñados para alterar codones específicos dentro de estas proteínas de reserva, para codificar lisina y/o triptófano, aumentando así el valor nutricional de estos cultivos. Por ejemplo, la proteína zein del maíz (Pederson *et al.* (1982), *Cell*, 29, 1015) puede ser modificada para que tenga un mayor contenido de lisina y triptófano, mediante los vectores y métodos de la invención.

40

## EJEMPLO 17

### 45 Creación de cadenas hospedadoras bacterianas mejoradas para la recombinación mediada por lambda en los BAC

#### *Transferencia de genes de recombinación de lambda a las células DH10B*

Para facilitar el uso de la recombinación mediada por lambda con los BAC, se ha creado un sistema de recombinación mediado por fago para una recombinación eficiente utilizando BAC. La cadena DH10B, a diferencia de la mayoría de las demás cadenas de *E. coli*, se transforma eficientemente con ADN de BAC y contiene muchas de las bibliotecas genómicas de BAC; se consideró una buena cadena hospedadora para la posterior modificación.

50

Dado que DH10B es defectuosa en recA, no se pueden emplear los cruces genéticos estándar para colocar el profago lambda defectuoso empleado para la recombinación mediada por lambda en la cadena DH10B de *E. coli*. Para superar este problema, la DH10B fue primero convertida a recA+, a continuación se cruzaron los genes de recombinación de lambda y la cadena se convirtió de nuevo a recA-, pero ya portando los genes de lambda.

55



Para convertir la DH10B en *recA+*, un fago de transducción de  $\lambda$  que porta el gen *recA+* fue utilizado para lisogenizar DH10B creando el derivado DY303. En la cadena DY330 empleada para la recombinación mediada por  $\lambda$ , el gen *tet* que confiere resistencia a la tetraciclina se insertó mediante recombinación homóloga donde existe deleción de *cro-bioA*, creando la cadena DY363. Un lisado de PI realizado en DY363 se utilizó para infectar DY303 y mediante genética bacteriana estándar el gen *tet* de la deleción de *cro-bioA* se cruzó en DY303. Esto somete a deleción un largo segmento del ADN de  $\lambda$  del lisógeno, incluyendo el gen *recA+*. Este nuevo derivado de DH10B y DY303 se denomina DY380, es *RecA-* y porta el marcador seleccionable *tet* sustituido por el segmento *cro-bioA*. Se observó que las células DY380 eran transformadas con el ADN de BAC a eficiencias de  $10^6$  a  $10^4$ .

*Creación de derivados de DY380 que contienen los genes flpe o ere inducibles por arabinosa*

La focalización de BAC a menudo hace uso de un marcador seleccionable para introducir la cassette de focalización en el lugar focalizado. No obstante, el marcador seleccionable puede interferir en la función subsiguiente del lugar focalizado. Para resolver este problema, los inventores señalaron que un marcador seleccionable flanqueado por sitios *FRT* o *loxP* puede ser eliminado mediante recombinasas *Flp* o *Cre*, respectivamente. De este modo, los inventores han creado dos nuevas cadenas, EL250 y EL350, sustituyendo en última instancia el gen *tet* de DY380 por *araC* y poniendo los genes *flpe* y *cre* bajo un promotor inducible por arabinosa. Los genotipos de DY380, EL250, y EL350 se muestran esquemáticamente en la FIG. 6. A pesar de que en este ejemplo se utilizó el promotor inducible por arabinosa  $P_{BAD}$ , básicamente se puede utilizar cualquier promotor inducible para activar la expresión de *flpe* y *cre*.

En el profago de DY380, *tet* se encuentra entre *cl857* y *bioA*. En el profago de EL250, *flpe* sustituye a *tet* (*flpe* es un *flp* modificado genéticamente que tiene una eficiencia de recombinación superior que el gen *flp* original; Buchholz *et al.*, Nature Biotechnology 16: 657-662, 1998). De este modo, como se ilustra en la FIG. 6, tanto EL250 como EL350 tienen funciones de recombinación homóloga inducibles por calor (los genes *red* de  $\lambda$ ) y funciones de recombinación de sitio específico inducibles por arabinosa (*flpe* o *cre*). Esta doble regulación permite tanto la focalización selectiva mediante recombinación como la posterior eliminación del marcador de selección del lugar focalizado mediante recombinación de sitio específico.

*Planteamiento mejorado para introducir profagos lambda defectuosos en las bacterias: círculos de minilambda*

Se ha desarrollado un método para introducir el sistema de recombinación mediado por  $\lambda$  de forma directa en prácticamente cualquier cadena de *E. coli*, incluyendo derivados de DH10B defectuosos en *recA*. Estos derivados pueden portar BAC, PAC u otros vectores.

La cadena DY330 porta la deleción de genes del profago de *cro* a *bioA*. Este segmento sometido a deleción de  $\lambda$  y *bioA* fue sustituido para crear un derivado que contiene un lisógeno de copia única de  $\lambda$ cl857 totalmente normal. Los lisógenos de este tipo se pueden inducir a 42 °C para expresar funciones de  $\lambda$  incluyendo las funciones de recombinación de *Red*. Dado que el  $\lambda$  porta todos los genes de replicación y líticos, la inducción durante más de 6 minutos provoca la muerte de las células que portan el lisógeno. No obstante, al inducir durante menos de 6 minutos, por ejemplo 4 minutos, las funciones de recombinación solo se activan parcialmente, pero las células sobreviven cuando se devuelven al cultivo a 32 °C. Empleando, por ejemplo, un tiempo de inducción de 4 minutos, se pueden generar recombinantes mediados por  $\lambda$  entre el ADN electroporado lineal y el cromosoma que incluye el ADN del profago. Así, el fago  $\lambda$  en sí puede ser empleado para lisogenizar y generar recombinantes en cadenas de BAC. No obstante, la eficiencia de recombinación sería baja, debido al reducido tiempo de inducción.

Las cassettes por PCR que contienen 5 genes para diferentes marcadores de resistencia a fármacos fueron amplificadas [*cat*, *kan*, *amp*, *tet*, *spec* (*strep*)] con homología a los flancos, a fin de sustituir genes del profago desde *cro* hasta *ea59* por los respectivos marcadores de fármacos, seleccionando con ese fármaco recombinantes resistentes a 32 °C. Un segmento de ADN del profago contiguo desde la posición de la base 38.044 del mapa de  $\lambda$  en *cro* hasta la posición de la base 25.737 del mapa de  $\lambda$  en *sib* es sustituido por las cassettes del fármaco (véase Court y Oppenheim). Esta deleción elimina todos los genes de replicación y lisis del profago, creando un profago defectuoso similar al de la DY330 original. La diferencia es que este profago tiene los dos sitios de unión *attL* y *attR* en el término del profago, mientras que DY330 tiene de *attR* a *bioA* sometido a deleción como parte de la deleción de *cro-bioA*.

Este conjunto de cadenas (con las respectivas cassettes del fármaco) puede ser inducido durante tiempos más prolongados que el  $\lambda$  completo sin matar las células, proporcionando así una actividad de recombinación homóloga máxima como con la DY330. El operón PL de estos profagos incluye los genes *int* y *xis*. La inducción activa su expresión y, dado que tanto *attL* como *attR* se encuentran presentes, causa la escisión de sitio específico del profago como un círculo de ADN que porta su marcador de fármacos asociado. Las células que se someten a inducción durante 15 minutos pueden perder el profago original. Esto ocurre en aproximadamente el 50% de las células. El otro 50% todavía cuenta con el profago. El 50% con el profago es probable que ocurra mediante la reintegración del ADN circular en el sitio del *attB* vacío mediante recombinación de sitio específico mediada por *Int*.

El ADN del profago defectuoso puede ser aislado y purificado a partir de estos lisógenos, si después de una inducción de 15 minutos, las células son lisadas y el ADN es aislado mediante protocolos de purificación de plásmidos, es decir en columnas Qiagen. El ADN del fago circular con sus marcadores de fármacos puede ser purificado. Este ADN no puede

5 replicarse tras la retransfección en cadenas de *E. coli*, pero puede expresar su operón PL y función Int para permitir la integración del ADN circular mediante recombinación de sitio específico entre el *attP* del ADN circular y el *attB* del cromosoma bacteriano. Solamente las funciones IHF del hospedador e Int son necesarias para la recombinación de sitio específico. Estos ADN integrados son estables, son inmunes y pueden ser seleccionados por el marcador de fármacos que porta cada uno.

Dado que RecA no es necesario para la recombinación mediada por Int de sitio específico, los derivados de DH10B pueden ser utilizados para la transformación y para la integración del ADN de fago defectuoso circular, seleccionado su marcador de fármacos apropiado.

10 El profago mini-lambda defectuoso también puede ser inducido como parte de un dilisógeno en el que hay un fago lambda CI857 completo presente. El lisado del fago creado por esta inducción de 90 minutos a 42 °C en L-Broth genera partículas de fago lambda normal, así como partículas que contienen el ADN del profago mini-lambda defectuoso (en terminología  $\lambda$ , partículas doCL). La infección de estos lisados en las células (p. ej., DH10B) permite la inyección de ADN del profago mini-lambda, la recombinación de sitio específico y la selección del marcador de fármacos portado por ese ADN.

## 15 EJEMPLO 18

### Una estrategia mejorada y reactivos mejorados para la ingeniería de BAC

20 Para testar el sistema del profago del Ejemplo 17 en ingeniería de BAC, se investigó la eficiencia de la recombinación de BAC en células EL250. En los experimentos descritos en este ejemplo, una cassette seleccionable fue focalizada para un lugar específico de neurona de ratón en un BAC de 250 kb. El BAC fue entonces modificado para mejorar su utilidad en posteriores estudios genéticos con ratones. Estos experimentos validaron una estrategia mejorada y proporcionaron reactivos mejorados para la ingeniería de BAC, empleando el sistema de recombinación de lambda.

25 El gen *Eno2* está ubicado en medio de 284H12, un BAC completamente secuenciado (obtenido de Research Genetics; Ansari-Lari *et al.*, Genome Research 8: 29-40, 1998). El gen *Eno2* fue focalizado, dado que se trata de un gen específico neural y se expresa en la mayoría de las neuronas maduras (Marangos y Schmechel, Annual Review of Neuroscience 10: 269-295, 1987). Al inactivar *eno2* y sustituirlo por una cassette que contenía *cre*, se creó una línea transgénica de BAC que expresa Cre en todas las neuronas maduras (descrito en el Ejemplo 20). Esta línea transgénica de BAC resulta útil para posteriores estudios de knockout condicionales. Los inventores emplearon un planteamiento de BAC en parte porque los BAC son lo suficientemente largos como para contener todas las secuencias reguladoras importantes necesarias para una buena regulación de la expresión genética.

30 A continuación se describe la construcción de la línea transgénica de BAC con expresión de Cre específica neuronal.

#### Generación de la cassette de focalización y de células EL250 que contienen BAC

35 La cassette de focalización IRES-eGFPcre-FRT-kan-FRT fue amplificada por PCR a partir de pICGN21, que se construyó subclonando un fragmento de 1,9 kbp digerido con *HinIII/AccI* y ocupado con *FRT-kan-FRT* de pFRTneo en el sitio de clonación ocupado y digerido con *NoTI/BclI* de pIRESeGC. La cassette *IRES-eGFPcre-FRT-kan-FRT* fue amplificada utilizando cebadores de 63 nt quiméricos. Los 21 nt del extremo 3' de cada cebador eran homólogos a la cassette de focalización empleada para la amplificación, mientras que los 42 nt del extremo 5' eran homólogos al último exón de *Eno2* en el que se iba a focalizar la cassette mediante recombinación. Los cebadores fueron diseñados para focalizar con precisión la cassette en sentido descendente del codón de parada de *Eno2* y en sentido ascendente de su sitio de polyA.

40 El BAC 284H12 que contenía *Eno2* fue electroporado en células EL250 y se seleccionaron seis colonias resistentes a cloranfenicol ( $\text{Cm}^R$ ). La digestión de ADN de BAC de seis colonias  $\text{Cm}^R$  con *EcoKL* o *HindIII* demostró que una tenía un patrón de digestión anómalo. No obstante, en otros experimentos de electroporación de BAC que implicaban el análisis de más de 76 colonias adicionales no se identificó ningún BAC anómalo. Estos resultados indican que las reorganizaciones de BAC durante la electroporación son raras. Los experimentos posteriores se realizaron con colonias 45 EL250 resistentes a  $\text{Cm}^R$  que contenían BAC con unos patrones de digestión de *EcoRL* y/o *HindIII* apropiados.

#### Generación y aislamiento de un BAC con un lugar *eno2* alterado

50 A continuación, el BAC 284H12 BAC fue modificado para alterar el lugar *eno2* con la cassette de focalización *IRES-eGFPcre-FRT-kan-FRT*. Los métodos empleados en estos experimentos fueron similares a los descritos extensamente en el presente y en Yu *et al.* (Proc. Natl. Acad. Sci. U.S.A. 97: 5978-5983, 2000). El planteamiento se ilustra esquemáticamente en la FIG.7

Las células EL250 que portan el BAC 284H12 se pasaron a 42 °C durante 15 minutos para inducir la expresión de lambda Exo, Beta, y Gam. Las células fueron entonces electroporadas con 300 ng de la cassette *IRES-eGFPcre-FRT-kan-FRT* amplificada y se seleccionaron las colonias resistentes a la kanamicina ( $\text{Km}^R$ ). Un fenotipo resistente a la

kanamicina indicaba que la cassette de focalización se integró con éxito en el BAC 284H12 (ilustrado como "Focalización," FIG. 7, en el centro). Se obtuvieron aproximadamente 5.200 colonias  $Km^R$  de  $10^8$  células electroporadas para una eficiencia de focalización aproximada de  $10^{-5}$ . No se obtuvo ninguna colonia de las células de control no inducidas por calor. Por tanto, la expresión de recombinasa de lambda fue necesaria para una recombinación eficiente.

5 Se analizaron 24 colonias resistentes a la kanamicina con PCR de célula completa, utilizando cebadores que flanqueaban el lugar focalizado. Los resultados de la PCR indicaron que todas fueron correctamente focalizadas. Sin embargo, la secuenciación de la región focalizada de seis colonias demostró que tres portaban mutaciones puntuales. Para determinar si estas mutaciones puntuales se introdujeron durante la amplificación por PCR o durante la recombinación homóloga, se repitió la focalización. Sin embargo, en esta ocasión la cassette *IRES-eGFPcre-FRT-kan-FRT* amplificada por PCR fue subclonada en el sitio Smal de pBluescript mediante el enlace del extremo romo antes de la focalización y se identificaron los plásmidos que portaban cassettes amplificadas de tipo salvaje mediante secuenciación de ADN. Estas cassettes fueron entonces liberadas del plásmido mediante digestión de BanHI y se utilizaron para la focalización. Utilizando este método de dos pasos, los 12 BAC focalizados que fueron posteriormente secuenciados contenían cassettes *IRES-eGFPcre-FRT-kan-FRT* de tipo salvaje. Estos resultados indican que las mutaciones puntuales fueron introducidas por los cebadores empleados o durante la amplificación por PCR de la cassette de focalización, y no durante la focalización.

#### Eliminación del marcador de resistencia a kanamicina

A continuación, el marcador seleccionable *kan* fue eliminado del BAC para evitar su posible interferencia en la expresión de Cre. Este proceso se inició mediante tratamiento con arabinosa, que induce a las células EL250 a expresar la recombinasa Flpe. El proceso se ilustra en la Fig. 7, línea inferior ("Salida de kan")

Los cultivos nocturnos de colonias  $Km^R$  únicas fueron diluidos 50 veces en medio LB y cultivadas hasta  $OD_{600} = 0,5$ . La expresión de Flpe de las células EL250 fue entonces inducida por incubación de los cultivos con 0,1% de L-arabinosa durante 1 hora. Las células bacterianas fueron posteriormente diluidas 10 veces en medio LB, cultivadas durante otra hora más y extendidas en placas de cloranfenicol (12,5 ug/ml). Al día siguiente, se recogieron 100 colonias CmR y se volvieron a poner en placas de kan (25 ug/ml) para testar la pérdida de resistencia a la kanamicina. La resistencia al cloranfenicol indica que la célula conservaba el BAC, mientras que la sensibilidad a la kanamicina indica que *kan* se ha eliminado con éxito del BAC. Todas las colonias eran  $Km^S$  y contenían un único sitio FRT en el lugar focalizado.

Sin limitarse a la teoría, es probable que la eficiencia de recombinación sorprendentemente elevada refleje el estrecho control de la expresión de *Flpe* que proporciona el promotor  $P_{BAD}$  de copia única y el gen *flpe*, y el hecho de que los sitios FRT se encuentran ubicados en *cis*, en lugar de en *trans* entre sí.

#### Eliminación de un sitio *loxP* indeseable en el segmento principal del vector BAC

Un sitio *loxP* incluido en el segmento principal del vector BAC (pBeloBAC1 1; Shizuya *et al.*, Proc. Natl. Acad. Sci. U.S.A. 89: 8794-8797, 1992) fue eliminado mediante una ronda final de focalización genética.

Para facilitar la eliminación de este sitio *loxP* indeseable, se construyó un nuevo plásmido, pTamp, que contiene un gen *amp* flanqueado por 920 bp de la secuencia pBeloBAC11 ubicada en el extremo 5' del sitio *loxP* y 370 bp de la secuencia pBeloBAC11 ubicada en el extremo 3' del sitio *loxP*. Este elemento insertado *amp* puede ser eliminado de pTamp mediante digestión de BamHI y utilizado para sustituir el sitio *loxP* en el transgen de BAC mediante focalización genética. Esta reacción de focalización es muy eficiente debido a la gran cantidad de homología entre la cassette de *amp* y el vector pBeloBAC11 (56.200 colonias por  $10^8$  células electroporadas).

Tras la eliminación del sitio *loxP* indeseado, el BAC 284H12 modificado se utilizó en los estudios con ratones transgénicos descritos en el Ejemplo 20.

## EJEMPLO 19

### Subclonación por reparación de gap

Este sistema de recombinación mediado por  $\lambda$  también se puede emplear para subclonar fragmentos de BAC, sin utilizar enzimas de restricción ni ligasas de ADN. Esta forma de subclonación confía en la reparación de huecos para recombinar los extremos libres de un vector de plásmido lineal con secuencias homólogas portadas en el BAC (FIG. 8). El método se puede adaptar fácilmente a otras formas de ADN intramolecular y extracromosómico, como plásmidos, cromosomas artificiales de levadura, cromosomas artificiales de PI y cósmidos. Este nuevo método combina la recombinación mediada por lambda con la reparación de huecos para permitir la recombinación de segmentos de ADN muy largos en un vector extracromosómico.

El vector del plásmido lineal con un marcador seleccionable *amp* y un origen de replicación porta los extremos recombinogénicos (FIG. 8B). El vector es generado mediante amplificación por PCR empleando dos cebadores quiméricos. Los 45-52 nt del extremo 5' de cada cebador son homólogos a los dos extremos de la secuencia de BAC

que se va a subclonar, mientras que los 20 nt del extremo 3' son homólogos al ADN del plásmido. La recombinación genera un plásmido circular en el que el elemento insertado de ADN se recupera del ADN del BAC mediante una reparación de gap. Los plásmidos circulares son seleccionados por su Amp<sup>R</sup>.

5 Para determinar el fragmento de tamaño máximo que se puede subclonar de los BAC utilizando este método, se generaron varios pares de cebadores diferentes en los que los segmentos de homología estaban ubicados a 25 kb, 48 kb, 60 kb, u 80 kb de distancia del ADN del BAC de *Eno2* (Fig. 8A). Los sitios de restricción de corte poco frecuentes *NotI* y *SalI* también se incorporaron a estos cebadores, de forma que los fragmentos subclonados pudieron ser liberados de los clones recombinantes intactos. Utilizando pBluescript como vector de clonación, resultó posible subclonar el fragmento de 25 kb. No obstante, los intentos de subclonar fragmentos más largos fracasaron. Como posible explicación de este resultado, se planteó la hipótesis de que los subclones que contenían fragmentos más largos en un vector de alta copia como pBluescript eran tóxicos para la célula.

10 Para determinar si la hipótesis era correcta, un vector con un número de copias menor (pBR322, con su elemento de control del número de copias intacto) fue utilizado como vector de clonación. Con un vector pBR322 se pudieron subclonar fragmentos de hasta 80 kb de largo. No todos los subclones obtenidos mediante reparación de gap tenían los elementos insertados correctos (tal y como se determinó mediante un análisis del patrón de la enzima de restricción). Algunos subclones carecían de elementos insertados mientras que otros contenían elementos insertados con patrones de restricción aberrantes. A fin de confirmar que se ha subclonado el elemento insertado correcto, a la hora de utilizar la subclonación mediante reparación de gap se puede emplear un método de exploración de subclones para garantizar que los subclones seleccionados contienen el elemento insertado deseado. Estos métodos incluyen el mapeo de la restricción, la secuenciación, el análisis por PCR, el análisis Southern, etc., y otros métodos bien conocidos por aquellos con conocimientos en la técnica.

15 La capacidad para subclonar grandes fragmentos de ADN genómico mediante reparación de huecos debería facilitar muchos estudios de investigación del genoma que resultaban difíciles o imposibles de realizar anteriormente. Por ejemplo, la reparación de huecos para la clonación en vectores se puede utilizar con muchos vectores diferentes empleados para la expresión de proteína en bacterias, plantas y células de animales; mutagénesis; clonación; transcripción, etc. Los vectores de focalización o estructuras transgénicas se pueden subclonar con facilidad, y prácticamente cualquier región del BAC modificado se puede ahora incluir en el subclon deseado.

20 La recombinación mediada por lambda combinada con la reparación de huecos hace posible subclonar fragmentos de mezclas complejas sin purificar primero el ADN que se va a subclonar. Esto facilita en gran medida el proceso de subclonación y permite una subclonación de alta producción de decenas de miles de genes o moléculas de ADN en los segmentos principales de muchos vectores diferentes. Esto facilitará en gran medida los estudios diseñados para determinar la función de los genes descubiertos en los proyectos de secuenciación a gran escala. Por ejemplo, los clones de ADNc para genes de función desconocida pueden ser subclonados en muchos vectores de expresión diferentes y la función de estos estudios de genes en ensayos basados en células *in vitro* o en el animal completo. Este tipo de subclonación no confía en la amplificación por PCR, que puede introducir mutaciones indeseadas en las secuencias subclonadas.

25 La subclonación por reparación de huecos también facilita la identificación de regiones de control de lugares u otros elementos reguladores que se pueden ubicar a cierta distancia del gen. Muchos de estos elementos potenciales están siendo identificados en la actualidad mediante técnicas como la secuenciación comparativa del genoma. Algunos ejemplos incluyen islas de patogenicidad, orígenes replicativos y elementos de segregación. La capacidad para modificar con precisión estas secuencias reguladoras en los BAC, combinada con la capacidad para incluirlas o excluirlas durante el proceso de subclonación, permitirá analizar detenidamente la función de estas secuencias en el animal completo o *in vitro* a un nivel que anteriormente no resultaba posible.

## EJEMPLO 20

### 45 Producción de ratones transgénicos empleando BAC

Los Ejemplos 17 y 18 describen la estructura de un BAC modificado que se considera que contiene todas las secuencias reguladoras necesarias para la expresión de Cre específica neural en los ratones transgénicos. Para investigar esta hipótesis, el BAC modificado descrito en el Ejemplo 18 fue inyectado en cigotos F<sub>2</sub> (C3H/HeN-Mtv<sup>-</sup> X C57BL/6Ncr). A continuación, se estableció una línea transgénica de BAC que porta aproximadamente dos copias del transgen.

50 Además de la línea transgénica de BAC, también se establecieron dos líneas transgénicas que portan 20 subclones de 25-kbp del BAC. Los subclones de 25 kbp portan toda la región de codificación de *Eno2* modificado, así como 10 kbp de secuencia flanqueadora del extremo 5' y 5 kbp de secuencia flanqueadora del extremo 3', respectivamente. Una línea transgénica, 25kbp-1, porta aproximadamente cuatro copias del transgen, mientras que la segunda, 25kbp-2, porta aproximadamente cinco copias del transgen. Así, la expresión de Cre en la línea transgénica de BAC se podría comparar con la expresión de Cre en las líneas transgénicas que portan el subclon.

Los ratones transgénicos se cruzaron con los ratones informantes ROSA26, que contienen un informante lacZ que se puede activar mediante recombinasa de Cre (Soriano, Nature Genetics 221:70-71, 1999). Posteriormente se analizaron heterocigotos dobles mediante tinción de X-gal a las cuatro semanas de edad.

5 Se examinaron varios tejidos diferentes para comprobar la expresión de X-gal, incluyendo el cerebro, la médula espinal, el ojo, el pulmón, el corazón, el intestino, el músculo, el hígado, el bazo y el riñón. Las células teñidas de azul se encontraron solamente en tejido neural en las tres líneas transgénicas, lo que indica que tanto el BAC como el subclon de 25 kbp contienen los elementos reguladores necesarios para la expresión específica neural. No obstante, el patrón de actividad de Cre fue diferente en las tres líneas. Los cortes con vibrátomo del cerebro de los ratones transgénicos BAC mostraron células teñidas de azul en toda la materia gris, pero no en la materia blanca, lo que es indicativo de la actividad de Cre en la mayoría de las neuronas pero no en las células gliales. Por lo contrario, la tinción de X-gal en los ratones transgénicos de 25kbp-1 y 25kbp-2 estuvo presente solamente en un subconjunto de neuronas y la expresión era variable entre las dos líneas diferentes.

15 Un aumento de mayor potencia del cerebelo de los ratones transgénicos BAC demostró que Cre estaba expresado prácticamente en todas las células neuronales. Esto incluía las células de Purkinje en la capa de Purkinje, células granulares y células de Golgi en la capa granular, células en cesta y células en estrella en la capa molecular y neuronas de núcleos cerebelosos profundos. Por contra, en la línea 25kbp-1, Cre solamente estaba expresado en un subconjunto de células Golgi, además de en algunas células en la capa granular y de Purkinje. Las células gliales de la materia blanca también expresaban Cre, lo que es indicativo de una expresión más suave. En la línea 25kbp-2, la expresión de Cre se limitaba a la materia gris e incluía diversos tipos de células neuronales, incluyendo la mayoría de las células en cesta, células en estrella, células de Purkinje y neuronas de núcleos cerebelosos profundos. Por contra, eran pocas las células granulares y células de Golgi en la capa granular que expresaban Cre.

20 Un aumento a mayor potencia del hipocampo y la corteza demostró unos resultados similares. En el hipocampo de los ratones transgénicos BAC, prácticamente todas las neuronas de la región del Cornu Ammonis (CA) y el dentate gyrus (DG) expresaban Cre. Lo mismo sucedía en la corteza, donde las seis capas de la corteza que contenían neuronas (capas II-VI) expresaban Cre. Por contra, el hipocampo de los ratones transgénicos 25kbp-1 mostraba una expresión de Cre reducida en el DG (Fig. 4E) y en las capas II y VI de la corteza. Los ratones transgénicos 25kbp-2 mostraron unos niveles de expresión de Cre incluso inferiores en el DG. Las regiones CA1 y CA2 de CA tampoco expresaban Cre. La expresión de Cre también se vio notablemente reducida en la corteza, concentrándose en las capas II y III la mayor parte de la reducción.

30 La actividad de Cre en la médula espinal, el ganglio de la raíz dorsal (DRG) y la retina de los ratones transgénicos también fue examinada, a fin de determinar si Cre se había expresado en las neuronas maduras dentro del sistema nervioso periférico. De forma similar a lo que se observó en el sistema nervioso central, Cre se expresó en la mayoría de las neuronas periféricas maduras de los ratones transgénicos BAC, mientras que un menor número de neuronas periféricas expresaba Cre en las dos líneas de ratones transgénicos 25kbp (datos no mostrados).

35 Para determinar si Cre se expresaba en todas las neuronas positivas en la proteína Eno2, una sección del cerebro de un animal transgénico BAC fue sometida a inmunotinción con un anticuerpo anti-Eno2, seguida de tinción con X-gal, para detectar la actividad de Cre. Prácticamente todas las neuronas Eno2-positivas mostraban actividad de Cre. Así, la expresión de Cre en los animales transgénicos BAC estaba correlacionada directamente con la actividad nativa del promotor-facilitador de Eno2 de ratón.

40 La presente solicitud, en particular los Ejemplos 17-20, describe un sistema de recombinación altamente eficiente para manipular ADN de BAC en *E. coli*. El sistema de recombinación emplea un profago lambda defectuoso para proporcionar funciones que protegen y recombinan la cassette de focalización de ADN lineal electroporado con la secuencia de BAC. Dado que las funciones de recombinación se expresan por un profago defectuoso y no por un plásmido, las funciones de recombinación no se pierden durante el cultivo celular, tal como suele ocurrir con los sistemas basados en plásmidos. Otra ventaja de este sistema de profago es que los genes de recombinación *gam* y *red* de lambda se encuentran bajo el control del represor  $\lambda$  sensible a la temperatura, que proporciona un control mucho más estricto de la expresión de *gam* y *red* que la que se puede obtener en los plásmidos. Esta estricta regulación, combinada con el sólido promotor PL de lambda, que impulsa la expresión de *gam* y *red* hasta niveles muy elevados, hace posible conseguir frecuencias de recombinación que son sorprendentemente eficientes (al menos 50-100 veces superiores que las obtenidas con sistemas basados en plásmidos; Narayanan, et al., Gene Therapy 6: 446: 442-447, 1999; Muiers *et al.*, Nucleic Acids Research 27: 1555-1557, 1999). El estricto control evita la expresión de cualquier función de recombinación, salvo por la inducción por temperatura durante 15 minutos.

55 La capacidad de manipular de forma precisa grandes fragmentos de ADN genómico, independiente de la ubicación de los sitios de la enzima de restricción apropiados, tiene muchas aplicaciones para la genómica funcional, tanto en el ratón como en otros organismos. Como se muestra en el presente, se puede introducir Cre en las regiones de codificación de genes contenidos en BAC, facilitando la generación de líneas transgénicas que expresan Cre, para su uso en estudios de knockout condicionales o en estudios de expresión genética condicionales. Los genes también pueden ser un epítipo etiquetado y microinyectado en la línea germinal de los ratones que portan una mutación en el gen. Si el transgen etiquetado con el epítipo rescata el fenotipo mutante, la proteína etiquetada con el epítipo es funcional y la

etiqueta del epítipo puede servir como marcador para la expresión del gen. Del mismo modo, un gen incluido en un BAC puede ser sustituido por otro gen y la función de la mutación “knock-in” puede ser ensayada en los ratones transgénicos.

5 Este sistema de recombinación también facilita la generación de vectores de focalización condicionales complicados. A pesar de que la generación de estos vectores a menudo solía ocupar varios meses, ahora se puede hacer solamente en unas pocas semanas. La capacidad para expresar reversiblemente recombinasas *Cre* o *Flpe* en *E. coli* acelera este proceso todavía más. Por otra parte, como se ha demostrado en el Ejemplo 18, ahora se puede introducir un marcador seleccionable flanqueado por sitios *loxP* o *FRT* en un intrón de un gen y posteriormente retirarlo mediante expresión transitoria de *Cre* o *Flpe*, dejando detrás un único sitio *loxP* o *FRT* en el intrón.

## 10 EJEMPLO 21

### Recombinación de BAC sin selección de fármacos

15 La alta eficiencia de la recombinación descrita en el Ejemplo 18 y en otras partes de estos ejemplos sugería que la focalización podría resultar posible sin selección de fármacos. La focalización directa sin selección de fármacos ofrecería una serie de ventajas significativas. En particular, facilitaría experimentos genómicos en los que la presencia de una cicatriz de un marcador seleccionable, o incluso de un *FRT* o *loxP*, podría resultar indeseable.

20 Para demostrar que la focalización se puede conseguir sin selección de fármacos, una etiqueta FLAG de 24 bp fue focalizada para el extremo 5' del gen que contiene la caja de SRY4 (*Sox4*) en un BAC de 125 kb. Para estos experimentos, se generó una cassette de focalización de 114 bp en la que dos brazos homólogos de 45 bp del gen *Sox4* flanqueaban la secuencia de FLAG de 24 bp. Este fragmento de ADN se creó sintetizando dos oligonucleótidos de 79 bp que se solapaban en sus extremos 3' en 44 bp. Estos solapamientos fueron anillados y rellenados con Taq polimerasa.

25 La expresión de genes de recombinasa de lambda por parte del profago defectuoso fue inducida por calor en las células DY380 portadoras del BAC de *Sox4*. A continuación, la cassette etiquetada con FLAG fue introducida en las células mediante electroporación. A continuación, las células se extendieron en placas LB hasta una densidad de 2.000 células por placa. Las colonias que contenían la etiqueta FLAG fueron posteriormente identificadas mediante hibridación de colonias, empleando una sonda de oligonucleótidos específica de FLAG de 30 bp (etiqueta FLAG de 24 bp y 3 bp a cada lado que eran homólogos al sitio focalizado de *Sox4*).

30 Entre 3.800 colonias estudiadas de células no inducidas, no se identificó ninguna colonia FLAG-positiva. Por contra, siete colonias FLAG-positivas se identificaron en 4.210 colonias obtenidas de células inducidas para una frecuencia de focalización total de  $1,7 \times 10^{-3}$ . La amplificación por PCR y la secuenciación directa demostraron que cada una de las siete colonias FLAG-positivas estuvo correctamente focalizada.

35 Como se ha demostrado de forma inequívoca en este ejemplo, la eficiencia de recombinación sorprendentemente elevada que ofrece este sistema de recombinación hace posible manipular el BAC u otro ADN sin selección de fármacos. Las mutaciones puntuales, deleciones o inserciones se pueden practicar ahora en cualquier gen de un BAC, en ausencia de un desconcertante marcador de selección de fármacos incorporado o de un sitio *loxP* o *FRT*. En los casos en los que el gen es mutado en la enfermedad humana, se pueden realizar las mutaciones que causan la enfermedad exacta en el BAC, para analizar el efecto de las mismas en ratones transgénicos.

## EJEMPLO 22

### Materiales y métodos empleados en los Ejemplos 17-21

40 *Cadenas bacterianas* Todas las cadenas empleadas salvo DH10B se mantuvieron a 32 °C debido al profago inducible por temperatura. DY303 se construyó infectando células DH10B (Gibco) con un fago  $\lambda$  que portaba *recA* ( $\lambda$ cl857 *recA*<sup>+</sup>) (cortesía de F. W. Stahl) y se seleccionaron lisógenos. La cadena EL11 se construyó sustituyendo el gen *tet* de DY380 por una cassette que contenía los genes *cat* y *sacB* seleccionando Cm<sup>R</sup>. Las células EL11 son Tet<sup>s</sup>, Cm<sup>R</sup> y sensibles a la sacarosa al 2%. La cadena EL250 se construyó sustituyendo la cassette *cat-sacB* de las células EL11 por *araC* y el gen de la recombinasa de *flpe* impulsado por el promotor de arabinosa (*P*<sub>BAD</sub>*flpe*), seleccionando en presencia de sacarosa. Las células EL250 son resistentes a la sacarosa al 2%. La cadena EL350 se construyó de manera similar, salvo por *cre* en lugar de *flpe*.

50 *Construcción de plásmidos.* La cassette de focalización IRES-eGFPcre-FRT-kan-FRT fue amplificada por PCR a partir de pICGN21, que se construyó subclonando un fragmento de 1,9 kbp digerido con *HinIII/AccI* y ocupado con *FRT-kan-FRT* de pFRTneo en el sitio de clonación ocupado y digerido con *NotI/BclI* de pIRESeGC. El FRTneo se construyó amplificando el gen *kan* junto con el promotor de lactamasa Beta de pEGFP-CI (Clontech) con los cebadores 5'CTGC AAGGCGATTAAGTTGGGTAACGCC AGGGTTTTTCGTC AGGTGGCACTTTTCGGG (SEQ ID NO: 13) y 5'CTCAGAAGAAGTTCGTCAAGAAGG (SEQ ID NO: 14). El fragmento amplificado fue después focalizado entre los sitios *FRT* de pNeo $\beta$ -gal (Stratagene). El pIRESeGC se generó insertando el fragmento eGFPcre ocupado y digerido con

*NheI/MluI* de 2 kbp de pEGC en el sitio de clonación ocupado y digerido con 5amHI de 3,5 kbp de pNTRlacZPGKneoloxP (Arango *et al.*, 1999). El pEGC se generó subclonando un fragmento por PCR de EcoRV/KpnI de 1,05 kbp que contenía el gen *cre* de pGKm<sub>cre</sub> (cortesía de P. Soriano) en el sitio EcoKV/KpnI de pEGFP-CI. Este fragmento por PCR se generó amplificando el gen *cre* de pGKm<sub>cre</sub> con los cebadores

5 5'GTAGGTACCTCGAGAATCGCCATCTTCCAGCAGGC (SEQ ID NO: 15) y 5'TCGAATTTTCTGCATCCAATTTACT GACCGTACACC (SEQ ID NO: 16), que contenían sitios de clivaje de *EcoRI* y *KpnI*, respectivamente, en sus extremos 5'.

Para construir el vector pTamp, el a/wp-focalizado pBeloBAC1 1 fue primero generado substituyendo el sitio *loxP* de BeloBAC11 (Shizuya *et al.*, 1992) por el gen *amp* amplificado por PCR de pEGFP (Clontech). Los cebadores empleados para la amplificación son 5' GCAA GTGTGTCGCTGTGCGA CGA GCTCGCGA GCTCGGA CA TGA GGTGTCTT AG A CGTCAGGTGGCAC (SEQ ID NO: 17) y 5'CATAGTTAAG CCAGCCCCGACACCCGCCAACACCCGCTGACGCGA ACCTCACGT TAAGGGATTTGGTC (SEQ ID NO: 18), que son homólogos al gen *amp* de pEGFP (normal) y a las secuencias que flanquean el sitio *loxP* en pBe10BAC11 (cursiva). Un fragmento por PCR de 2,4 kbp amplificado a partir de pBe10Bac1 focalizado en *amp* con los cebadores 5'GCAGGATCCAGTTTGTCTCCTGGAGCGACA (SEQ ID NO: 19) y 5TGCAGGTGCGACTCTAGAGGATC (SEQ ID NO:20) fue posteriormente clonado en el sitio ocupado y *XhoI/XbaI* de pCS (Stratagene) para crear el vector pTamp. La cassette de *amp* de 2,4 kbp que contenía un gen *amp* junto con 920 bp del extremo 5' y 370 bp del extremo 3', la secuencia del vector pBe10BAC11 que flanquea el sitio de *loxP* puede ser liberada mediante digestión de *BamHI* y empleada directamente para substituir el sitio de *loxP* en cualquier BAC derivado de pBeloBAC1 1 con *amp*.

20 El vector pK04 que contiene la cassette de focalización *cat-sacB* es un derivado de pK03 (Link *et al.*, 1997) en el que se han sometido a delección 605 bp entre *cat* y *sacB*.

La cassette de focalización *araC-P<sub>BAD</sub>flpe* fue amplificada a partir de pBADflpe, que se construyó subclonando un fragmento *PstI/KpnI* de 1,4 kb de pOGFlpe (Buchholz *et al.*, 1998) en pBAD/MycHis-A (Invitrogen). La cassette de focalización *araC-P<sub>BAD</sub>cre* fue amplificada a partir de pBADcre, que se construyó introduciendo un fragmento HindIII/NcoI de 1,2 kb de pGKm<sub>cre</sub> en pBAD/His-C (Invitrogen).

25

*Cebadores de amplificación para la focalización o reparación de huecos de ADN de cassettes.* Para todos los cebadores recogidos a continuación, los nucleótidos en cursiva son homólogos a la secuencia focalizada, mientras que los que aparecen en letra normal son homólogos a las cassettes de amplificación. La cassette Tet<sup>R</sup> empleada para la focalización de *cro-bio* en DY330 fue amplificada de Tn10 con los cebadores:5'TGGCGGTGA TAA TGGTTGCA TGTA CTAA GGA GGTTGTA 7/GCTCTTGGTT ATC AA GAGGG (SEQ ID NO:21) y 5 'GGCGCTGCAAAAA TTCTTTGTGCGAA CA GGGTGTCTGGA 7XACTCG AC ATCTTGGT TACCG (SEQ ID NO:22). La cassette *cat-sacB* empleada para substituir el gen *tet* en DY363 fue amplificada de pK04 con los cebadores:5'TGGCGGTGA TAA TGGTTGCA TGTA CTAA GGA GGTTGTA 7/GCTGTGACGGAAGATCACTTCG (SEQ ID NO:23) y 51GGCGCTGCAAAAA TTCTTTGTGCGAA CA GGGTGTCTGGA rCCTGAGGTTCTT ATGG CTCTTG (SEQ ID NO:24).

35 Las cassettes *araC-P<sub>BAD</sub>flpe* y *araC-P<sub>BAD</sub>cre* empleadas para la sustitución de *cat-sacB* en EL11 fueron amplificadas de pBADflpe y pBADcre con los cebadores:5'TGGCGGTGA TAA TGGTTGCA TGTA CTAA GGA GGTTGTA TGAAGCGCATGCATA ATGTGC (SEQ ID NO:25) y 5'GGCGCTGCAAAAA TTCTTTGTGCGAA CA GGGTGTCTGGA TXCTGTCTACTC AG GAGAGCGTTC (SEQ ID NO:26). La cassette *IRE5-eGFP<sub>cre</sub>-FRT-kan -FRT* empleada para la focalización del lugar *Eno2* fue amplificada de pIGCN21 con los

40 cebadores:5'CGCTTCGCGGGACATAATTTCCGAAATCCCAGTGTGCTGTGAGCCAAGCTATCGA ATTCCGCC (SEQ ID NO:27) y 5'GA GGCTCCA GGA GAA TGA GA TGTTCCCGCGTTCA GGCAA GCGCTATTCC AGAAG TAGTGAGGA (SEQ ID NO:28). Los oligonucleótidos empleados para focalizar la cassette flag en el extremo 5' del gen *Sox4* fueron anillados y extendidos con polimerasa, empleando los cebadores: 5 'GCGAGCGTGTGA GCGCGCGTGGGCGCCCGGCAA GCCGGGGCCA TGG ATT AC A AGGATGACGACGATAAGGTACAACAGA (SEQ ID NO: 29) y 5'GGCCA GCA GA GCCTCAGTGTCTCCGCGTTGTTGGTCTGTTGTA CTTATCGTCCG TCATCCTTGTAATCCATGGCCCCC (SEQ ID NO:30).

El derivado de pBR322 lineal empleado para subclonar el fragmento de 25 kbp del lugar *Eno2* modificado fue amplificado con los cebadores:5'CTCTCCA TGCCTGTCTGGGTGA GGGTGGCCAGGGGCGA TGGCTA TGA GA GA GG TCGACTTCTTAGACGTCAGGTGGCAC (SEQ ID NO:31) (*Eno2*-C-L1) y GCAA TGCA GA GAA GCCTTGTA CTGGGA TGA CA GA GA CGGA GGGGAA GA GGA GGCG GCCGCGATACGCGAGCGAACGTGA (SEQ ID NO:32) (*Eno2*-C-R1/2). Los cebadores de amplificación para los otros experimentos fueron: fragmento modificado de 48 kbp, 5'GA CTTCTA TGA CCTGTA CGGA GGGGA GAA GTTTGCACGTGACA GA GCTGGTCGT CCACTTCTTAGACGTCAGGTGGCAC (SEQ ID NO:33) (*Eno2*-C-L2/3/4) y *Eno2*-C-R1/2; fragmento modificado de 60 kbp, *Eno2*-C-L2/3/4 y 5'GCCCCATACACGTAATGTACATA GAATCACACAGCATCA CTTCTA TGGA TGCGGC GGCCGCGATACGCGAGCGAACGTGA (SEQ ID NO:34) (*Eno2*-C-R3); fragmento modificado de 80 kbp, *Eno2*-C-L2/3/4 y

55 5'CATCCA GTA GAA CTTGGGA GTGAAGCTA GA GCCAA GGCCATCTAA GTGA CAGGCG GCCGCGATACGCGAGCGAACGTGA (SEQ ID NO:35) (*Eno2*-C-RA). Estos cebadores contenían regiones del extremo 5' homólogas a la secuencia diana y regiones del extremo 3' homólogas a pBR322. Los productos de la PCR fueron purificados utilizando un equipo de extracción de gel Qiaex U (Qiagen) y digeridos con *DpnI* para eliminar la plantilla contaminada.

60

5 *Preparación de células electrocompetentes y generación de recombinantes.* Para la modificación de BAC, los cultivos nocturnos que contenían el BAC fueron cultivados a partir de colonias únicas y diluidos 50 veces en medio LB, y cultivados hasta una  $OD_{600} = 0,5-0,7$ . A continuación se indujeron cultivos de 10 ml para la expresión de Beta, Exo, y Gam, pasando las células a 42°C durante 15 min, para después enfriarlas sobre hielo durante 20 minutos. Las células fueron posteriormente centrifugadas durante 5 minutos a 5.500 g a 4 °C y lavadas en tres ocasiones con 1,5 ml de agua estéril enfriada con hielo. Las células se resuspendieron después en 50 µl de agua estéril enfriada con hielo y electroporadas. Para la transformación de BAC, el paso de inducción se omitió.

10 La transformación de las células se realizó mediante electroporación de 100-300 ng de ADN lineal en 50 µl de células competentes enfriadas con hielo en cubetas (0,1 cm) empleando un generador de impulsos para genes Bio-Rad a 1,75 kV, 25 µF con el controlador de impulsos establecido a 200 ohms. Después de la electroporación, se añadió 1 ml de medio LB. Las células fueron incubadas a 32 °C durante 1,5 horas con agitación y se extendieron sobre el medio de agar selectivo o no selectivo apropiado.

15 *Producción de ratones transgénicos.* El BAC modificado y los ADN del subclon de p25-kbp fueron purificados empleando gradientes de cloruro de cesio, tal y como se describe (Antoch *et al.*, 1997). El ADN del subclon de 25 kbp fue linealizado mediante digestión con Noil antes de la microinyección. El ADN de BAC (1 µg/ml) y el ADN del subclon de 25-kbp (2 µg/ml) fueron microinyectados en el pronúcleo de cigotos F<sub>2</sub> (C3H/HeN-Mtv X C57BL/6Ncr). Los fundadores transgénicos fueron posteriormente identificados mediante análisis Southern utilizando una sonda de cre o mediante PCR, empleando los cebadores 5' CTGCTGG AAG ATGGCG ATTCTCG (SEQ ID NO:36) y 5' AACAGCAGGAGCGGTGAGTC (SEQ ID NO.37) que flanquean la unión de inserción del extremo 3'.

20 *Análisis histoquímico de la expresión de β-galactosidasa.* Los ratones de cuatro a cinco semanas de vida fueron sacrificados en CO<sub>2</sub> y perfundidos con paraformaldehído al 4% en PBS (pH 7.3). Los cerebros, las médulas espinales y los ojos fueron extraídos y fijados durante tres horas. Los cortes con vibrátomo (20 µm) de cerebros se montaron en portaobjetos y se emplearon directamente para la tinción con X-gal o para la inmunocitoquímica. Para las médulas espinales y los ojos, se emplearon cortes con criostat (20 µm) que se realizaron crioprotegiendo los tejidos en sacarosa al 30% en PBS durante una noche y poniendo los tejidos en compuesto de congelación (OCT, Sakura). Antes de la tinción con X-gal, las muestras de los portaobjetos fueron fijadas con glutaraldehído al 0,25% en PBS y se lavaron brevemente con una solución de enjuague (0,1 solución tampón de fosfato, pH7.3, 0,1% ácido desoxicólico, 0,2% NP40 y 2mM MgCl<sub>2</sub>). La tinción con X-gal se realizó incubando muestras en solución tampón de tinción (2,5 mg/ml X-gal, 5mM ferricianuro de potasio y 5 mM ferrocianuro de potasio en solución tampón de tinción) durante 2 horas a 37°C, seguida de contrateñido con eosina al 0,25% (Fisher).

35 *Inmunocitoquímica.* La inmunotinción se realizó empleando el equipo ABC Vectastain (Vector Labs) sobre cortes con vibrátomo de 20 µm. Las secciones se bloquearon con PBS (pH7.3 que contenía 0,2% Triton X-100, 1,5% albúmina de suero bovino y 5% suero de cabra normal) a temperatura ambiente durante dos horas y fueron incubadas con el anticuerpo de *Eno2* primario, un antisuero de ratón anti-*Eno2* policlonal (Chemicon) a una dilución de 1:100 en solución PBS. Tras la incubación con un anticuerpo biotinilado secundario y el reactivo ABC, se hizo reaccionar la peroxidasa con 0,05% diaminobencidina tetrahidrocloruro (DAB) y 0,003% peróxido de hidrógeno.

40 A menos que se defina lo contrario, todos los términos técnicos y científicos empleados en el presente tienen el significado que entendería normalmente cualquiera con conocimientos en la técnica al que pertenece la invención. A pesar de que se pueden emplear métodos y materiales similares o equivalentes a los descritos en el presente para la práctica o prueba de la presente invención, a continuación se describen métodos y materiales adecuados. Por otra parte, los materiales, métodos y ejemplos se ofrecen a título ilustrativo solamente y no tienen un carácter limitador.



LISTADO DE SECUENCIAS

<110> EL GOBIERNO DE LOS ESTADOS UNIDOS DE AMÉRICA, REPRESENTADO POR EL SECRETARIO DEL DEPARTAMENTO DE SANIDAD Y SERVICIOS HUMANOS, INSTITUTOS NACIONALES DE SANIDAD

5 LEE, E-CHIANG  
 COPELAND, NEAL G.  
 JENKINS, NANCY A.  
 COURT, DONALD L.  
 YU, DAIGUAN  
 10 ELLIS, HILLARY M.

<120> RECOMBINACIÓN HOMÓLOGA MEJORADA MEDIADA POR PROTEÍNAS DE RECOMBINACIÓN DE LAMBDA

<130> 4239-59371

<150> US 60/271,632

<151> 2001-02-26

<150> US 60/225,164

<151> 2000-08-14

15 <160> 37

<170> Versión de patente 3.1

20 <210> 1  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial

25 <220>  
 <223> Cebador

<400> 1  
 tgtgacggaa gatcacttcg 20

30 <210> 2  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Cebador

35 <400> 2  
 accagcaata gacataagcg 20

40 <210> 3  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Cebador

45 <400> 3  
 ctcttggggtt atcaagaggg 20

<210> 4  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial

	<220>		
	<223> Cebador		
	<400> 4		
5	actcgacatc ttggttaccg		20
	<210> 5		
	<211> 21		
	<212> ADN		
10	<213> Secuencia artificial		
	<220>		
	<223> Cebador		
	<400> 5		
	cattcaaata tgtatccgct c		21
	<210> 6		
15	<211> 20		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
20	<223> Cebador		
	<400> 6		
	agagttggta gctcttgatc		20
	<210> 7		
25	<211> 20		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
30	<223> Cebador		
	<400> 7		
	tatggacagc aagcgaaccg		20
	<210> 8		
35	<211> 21		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
40	<223> Cebador		
	<400> 8		
	tcagaagaac tcgtcaagaa g		21
	<210> 9		
45	<211> 50		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Segmento de homología		
	<400> 9		
50	gtttgcgcgc agtcagcgat atccattttc gcgaaatccgg agtgtaagaa		50
	<210> 10		
	<211> 50		
	<212> ADN		
	<213> Secuencia artificial		
55	<220>		

ES 2 394 877 T3

	<223> Segmento de homología	
	<400> 10	
	ttcatattgt tcagcgacag cttgctgtac ggcaggcacc agctcttccg	50
	<210> 11	
	<211> 70	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Segmento de homología	
	<400> 11	
	aagtcgcbgt cggaaccgta ttgcagcagc tttatcatct gccgctggac ggcgcacaaa	60
	tcgcbcttaa	70
10	<210> 12	
	<211> 60	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Segmento de homología	
	<400> 12	
	aacagacacc atgggtgcacc tgactcctga ggagaagtct gccgttactg ccctgtgggg	60
	<210> 13	
20	<211> 56	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
25	<223> Cebador	
	<400> 13	
	ctgcaaggcg attaagttgg gtaacgccag ggttttcgtc aggtggcact ttcggg	56
30	<210> 14	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> Cebador	
	<400> 14	
	ctcagaagaa ctcgtcaaga agg	23
40	<210> 15	
	<211> 35	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
45	<223> Cebador	
	<400> 15	
	gtaggtacct cgagaatcgc catcttccag caggc	35
	<210> 16	
	<211> 36	
	<212> ADN	
50	<213> Secuencia artificial	
	<220>	

ES 2 394 877 T3

	<223> Cebador	
	<400> 16	
	tcgaattttc tgcacccaat ttactgaccg tacacc	36
5	<210> 17 <211> 67 <212> ADN <213> Secuencia artificial	
10	<220> <223> Cebador	
	<400> 17	
	gcaagtgtgt cgctgctgac gagctcgcga gctcggacat gaggttgtct tagacgtcag	60
	gtggcac	67
15	<210> 18 <211> 69 <212> ADN <213> Secuencia artificial	
20	<220> <223> Cebador	
	<400> 18	
	catagttaag ccagccccga caccgcgcaa caccgcgtga cgcgaacctc acgttaaggg	60
	atthttggtc	69
25	<210> 19 <211> 29 <212> ADN <213> Secuencia artificial	
30	<220> <223> Cebador	
	<400> 19	
	gcaggatcca gtttgctcct ggagcgaca	29
35	<210> 20 <211> 22 <212> ADN <213> Secuencia artificial	
40	<220> <223> Cebador	
	<400> 20	
	tgcaggtcga ctctagagga tc	22
45	<210> 21 <211> 60 <212> ADN <213> Secuencia artificial	
50	<220> <223> Cebador	
	<400> 21	
	tggcggatgat aatgggttgca tgtactaagg aggttgatg ctcttgggtt atcaagaggg	60
	<210> 22 <211> 60 <212> ADN <213> Secuencia artificial	

ES 2 394 877 T3

<220>  
 <223> Cebador

<400> 22  
 ggcgctgcaa aaattctttg tcgaacaggg tgtctggatc actcgacatc ttggttaccg 60

5  
 <210> 23  
 <211> 61  
 <212> ADN  
 <213> Secuencia artificial

10  
 <220>  
 <223> Cebador  
 <400> 23  
 tggcggatgat aatggttgca tgtactaagg aggttgatg ctgtgacgga agatcacttc 60

g 61

15  
 <210> 24  
 <211> 61  
 <212> ADN  
 <213> Secuencia artificial

20  
 <220>  
 <223> Cebador  
 <400> 24  
 ggcgctgcaa aaattctttg tcgaacaggg tgtctggatc ctgaggttct tatggctctt 60

g 61

25  
 <210> 25  
 <211> 60  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Cebador  
 <400> 25  
 tggcggatgat aatggttgca tgtactaagg aggttgatg aagcggcatg cataatgtgc 60

30  
 <210> 26  
 <211> 66  
 <212> ADN  
 <213> Secuencia artificial

35  
 <220>  
 <223> Cebador  
 <400> 26  
 ggcgctgcaa aaattctttg tcgaacaggg tgtctggatc ctgtgtccta ctcaggagag 60

cgttc 65

40  
 <210> 27  
 <211> 63  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Cebador  
 <400> 27  
 cgcttcgagg gacataattt ccgaaatccc agtgtgctgt gagccaagct atcgaattcc 60

gcc 63

45  
 <210> 28  
 <211> 63

ES 2 394 877 T3

<212> ADN  
 <213> Secuencia artificial

5 <220>  
 <223> Cebador  
 <400> 28  
 gaggctccag gagaatgaga tgttcccgcg ttcaggcaag cgctattcca gaagtagtga 60  
 gga 63

10 <210> 29  
 <211> 79  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Cebador  
 <400> 29  
 gcgagcgtgt gagcgcgcgt gggcgcccgg caagccgggg ccatggatta caaggatgac 60  
 gacgataagg tacaacaga 79

15 <210> 30  
 <211> 79  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> Cebador  
 <400> 30  
 ggccagcaga gcctcagtgt tctccgcggt gttggtctgt tgtaccttat cgtcgtcacc 60  
 cttgtaatcc atggcccc 79

25 <210> 31  
 <211> 80  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Cebador  
 <400> 31  
 ctctccatgc ctgtctgggt gaggggtggcc caggggcgat ggctatgaga gaggtcgact 60  
 tcttagacgt caggtggcac 80

30 <210> 32  
 <211> 79  
 <212> ADN  
 <213> Secuencia artificial

35 <220>  
 <223> Cebador  
 <400> 32  
 gcaatgcaga gaagccttgt actgggatga cagagacgga ggggaagagg aggcggccgc 60  
 gatacgcgag cgaacgtga 79

40 <210> 33  
 <211> 80  
 <212> ADN  
 <213> Secuencia artificial

45 <220>  
 <223> Cebador

ES 2 394 877 T3

	<400> 33		
	gacttctatg acctgtacgg aggggagaag tttgcgacgt gacagagctg gtcgtegact	60	
	tcttagacgt caggtggcac	80	
	<210> 34		
	<211> 81		
5	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Cebador		
	<400> 34		
	gccccataca cgtaaagtga catagaatca cacagcatca cttctatgga tgcggcgcc	60	
	gcgatacgcg agcgaactg a	81	
10	<210> 35		
	<211> 79		
	<212> ADN		
	<213> Secuencia artificial		
15	<220>		
	<223> Cebador		
	<400> 35		
	catccagtag aacttgggag tgaagctaga gccaaaggcca tctaagtgac aggcggccgc	60	
	gatacgcgag cgaactga	79	
	<210> 36		
	<211> 23		
20	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Cebador		
	<400> 36		
25	ctgctggaag atggcgattc tcg	23	
	<210> 37		
	<211> 20		
	<212> ADN		
30	<213> Secuencia artificial		
	<220>		
	<223> Cebador		
	<400> 37		
	aacagcagga gcggtgagtc	20	
35			

**REIVINDICACIONES**

1. Un método para inducir la recombinación homóloga en una célula que comprende un ácido nucleico diana, comprendiendo el método :
- 5 la introducción en una célula de un primer ácido nucleico de interés que comprende una secuencia homóloga al ácido nucleico diana de longitud suficiente como para inducir la recombinación homóloga y donde el primer ácido nucleico de interés es un ADN de cadena única o un ADN que comprende un extremo 3' protuberante, donde la célula comprende un ácido nucleico que codifica un polipéptido de unión a ADN de cadena única que funciona en la recombinación homóloga de reparación de roturas bicatenarias, operativamente unido al promotor PL, y
- 10 la inducción de una expresión suficiente del ácido nucleico que codifica el polipéptido de unión a ADN de cadena única para promover la recombinación homóloga del primer ácido nucleico de interés con el ácido nucleico diana; induciendo así la recombinación homóloga en la célula.
2. El método de la reivindicación 1, donde el polipéptido de unión a ADN de cadena única es Beta de lambda y la célula comprende también un ácido nucleico que codifica Gam y una proteína seleccionada del grupo compuesto por Exo, F12A21.16 de Arabidopsis, orf del virus del cólera, gen YqaJ de B. subtilis, fago A118 gp47 de Listeria, gen 34.1 del fago SPP1 de B. subtilis, y orf del virus de la peste porcina africana operativamente unido al promotor PL susceptible de desrepresión.
- 15 3. El método de la reivindicación 1, donde el método comprende también la introducción de una segunda secuencia de ADN de cadena única, donde la segunda secuencia de ADN de cadena única comprende más de 10 bp de solapamiento complementario en el extremo 5' o 3' del primer ácido nucleico de interés.
- 20 4. El método de la reivindicación 1, donde la célula es una célula bacteriana o una célula eucariota.
5. El método de la reivindicación 4, donde la célula eucariota es una célula de mamífero.
6. El método de la reivindicación 5, donde la célula de mamífero es una célula madre.
7. El método de la reivindicación 1, donde el primer ácido nucleico de interés codifica un marcador seleccionable.
- 25 8. Un método para realizar la recombinación homóloga de una secuencia de nucleótidos de interés en una célula hospedadora, que comprende introducir una secuencia de ácido nucleico que codifica los productos genéticos de Exo, Beta y Gam de lambda, o un homólogo de los mismos que posea al menos un 60% de identidad de secuencia, operativamente unido al promotor PL;
- 30 introducir el polinucleótido que comprende una secuencia de nucleótidos homóloga a la secuencia de nucleótidos de interés en las células hospedadoras, donde el nucleótido de interés es un ADN de cadena única o un ADN que comprende un extremo 3' protuberante;
- activar el promotor PL en las células, induciendo así la expresión de los productos genéticos de Exo, Beta y Gam de lambda, u homólogos de los mismos;
- 35 seleccionar una célula de la población de células hospedadoras en las que se ha producido la recombinación homóloga entre el ácido nucleico de interés y una secuencia de ácido nucleico diana.
9. El método de la reivindicación 8, donde al menos una célula de 100.000 de la población de células hospedadoras contiene ADN en el que se ha ocurrido recombinación homóloga.
10. El método de la reivindicación 8, donde al menos una célula de 1.000 de la población de células hospedadoras contiene ADN en el que se ha ocurrido recombinación homóloga.
- 40 11. El método de la reivindicación 10, donde al menos una célula de 100 de la población de células hospedadoras contiene ADN en el que se ha ocurrido recombinación homóloga.
12. El método de la reivindicación 8, donde la secuencia de nucleótidos de interés es intracromosómica o extracromosómica.
- 45 13. El método de la reivindicación 12, donde la secuencia de nucleótidos de interés es extracromosómica y está ubicada en un cromosoma artificial bacteriano, un cromosoma artificial de P1, un plásmido, un cósmido o un cromosoma artificial de levadura.



14. Un método para la clonación de una molécula de ácido nucleico, que es un ADN de cadena única o un ADN que comprende un extremo 3' protuberante, en una célula, que comprende:
- 5 introducir en una célula hospedadora que comprende una secuencia de ácido nucleico que codifica los productos genéticos de Exo, Beta y Gam de lambda, o un homólogo de los mismos que posea al menos un 60% de identidad de la secuencia, operativamente unido al promotor PL, un vector que comprende un número suficiente de nucleótidos homólogos a un ácido nucleico diana como para conseguir la recombinación homóloga;
- activar el promotor PL en la célula hospedadora, induciendo así la expresión de los productos genéticos de Exo, Beta y Gam de lambda, o la variante funcional u homóloga de los mismos;
- 10 seleccionar una célula hospedadora en la que la recombinación homóloga se ha producido entre el vector y una secuencia de ácido nucleico diana, clonando así la molécula de ácido nucleico.
15. Un método para alterar una secuencia de ácido nucleico eucariota, que comprende:
- introducir un ácido nucleico de interés en una célula bacteriana que comprende un elemento extracromosómico que comprende al menos un intrón o al menos un exón de una secuencia de ácido nucleico eucariota;
- 15 donde el ácido nucleico de interés comprende un número de nucleótidos homólogos a la secuencia de ácido nucleico eucariota suficiente como para mediar la recombinación homóloga y es un ADN de cadena única o un ADN con el extremo 3' protuberante;
- y donde la célula bacteriana comprende el promotor PL operativamente unido a una secuencia de ácido nucleico que codifica una recombinasa; y
- activar la expresión de la recombinasa a partir del promotor PL;
- 20 donde la expresión de la recombinasa provoca que el ácido nucleico de interés se someta a recombinación homóloga con la secuencia de ácido nucleico eucariota.
16. El método de la reivindicación 15, donde la recombinasa es una proteína de unión a ADN de cadena única.
17. El método de la reivindicación 15, donde el ácido nucleico que codifica la recombinasa codifica el producto genético de Beta de lambda.
- 25 18. El método de la reivindicación 15, donde el ADN que codifica la recombinasa a expresar codifica también el producto genético de Exo de lambda o el producto genético de Gam de lambda.
19. El método de la reivindicación 15, donde el elemento extracromosómico es un cromosoma artificial bacteriano, un cromosoma artificial de levadura, un cromosoma artificial de P1, un plásmido o un cósmido.
- 30 20. El método de la reivindicación 15, donde la secuencia de ácido nucleico eucariota es una secuencia de ácido nucleico de mamífero.
21. El método de la reivindicación 15, donde el ácido nucleico de interés codifica una etiqueta del epítipo.
22. Un método de subclonación de una secuencia de ADN, que comprende:
- 35 proporcionar la célula bacteriana que comprende una secuencia de ácido nucleico que codifica los productos genéticos Exo, Beta y Gam de lambda, o un homólogo de los mismos que posea al menos un 60% de identidad de la secuencia, operativamente unido al promotor PL, y que comprende también una molécula de ADN con una secuencia de ácido nucleico diana ubicada extracromosómicamente;
- introducir en la célula bacteriana un vector de plásmido lineal que comprende un origen de replicación y un marcador seleccionable, y comprende una secuencia de nucleótidos que es un ADN de cadena única o un ADN que tiene un extremo 3' protuberante capaz de someterse a recombinación homóloga con la secuencia de ácido nucleico diana en el ADN;
- 40 inducir la expresión del ácido nucleico que codifica los productos genéticos de Exo, Beta y Gam de lambda, induciendo así la recombinación homóloga entre el vector de plásmido lineal y la secuencia de ácido nucleico diana; donde el ácido nucleico diana es insertado en el vector de plásmido lineal; y
- 45 circularizar el vector de plásmido lineal, subclonando así el ADN.
23. El método de la reivindicación 22 donde el ADN comprende al menos 20 kb del ADN extracromosómico.

24. El método de la reivindicación 22 donde el ADN comprende al menos 40 kb del ADN extracromosómico.
25. El método de la reivindicación 22 donde el ADN comprende al menos 80 kb del ADN extracromosómico.
26. El método de la reivindicación 22, donde el ADN extracromosómico es un cromosoma artificial bacteriano.

FIG. 1

*nadA gal attL int xis ea8.5 ea22 hin exo bet gam kil clll ssb ral sieB N rexB rexA cl857 (cro-attR-bioA)*  
 -----|----- PL PR Δ

FIG. 3

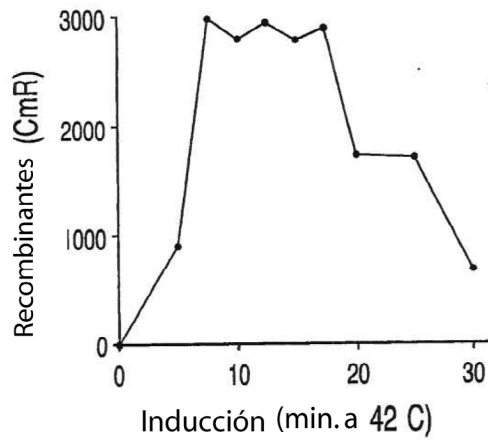


FIG. 2

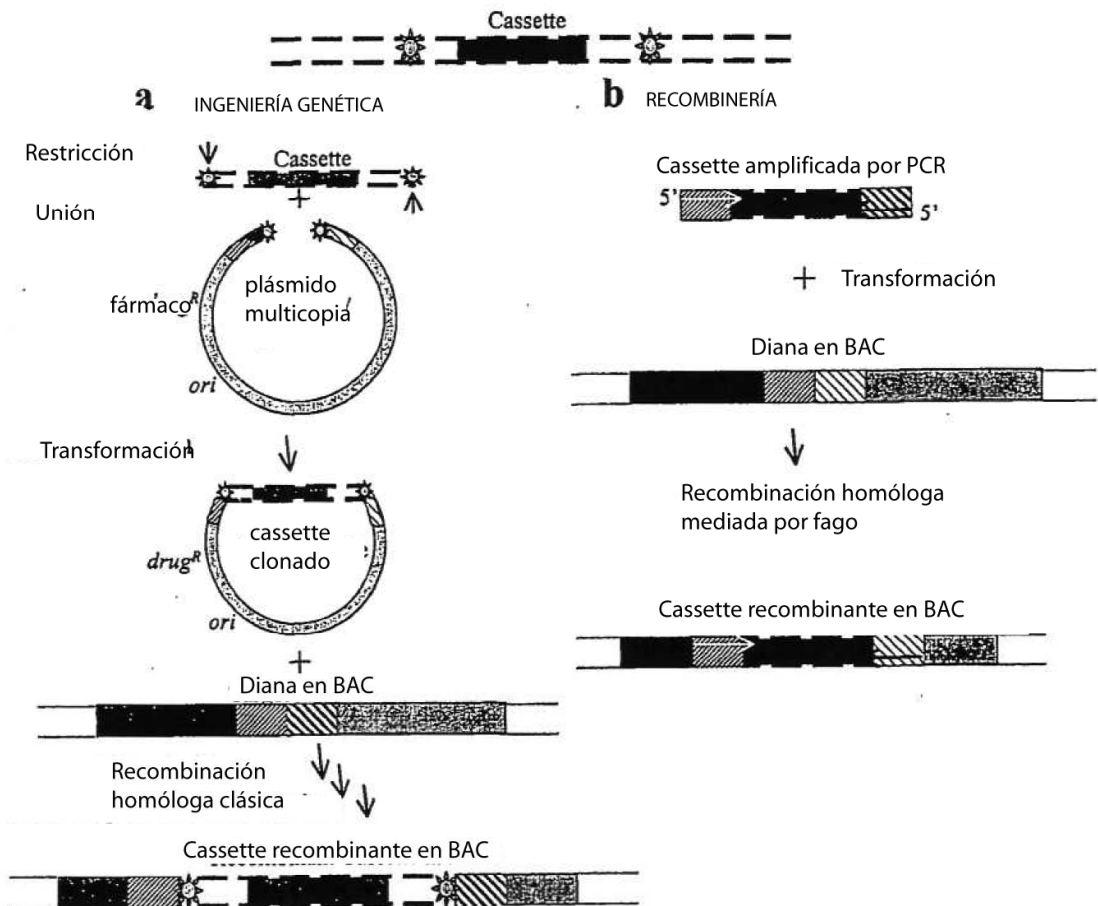


FIG. 4

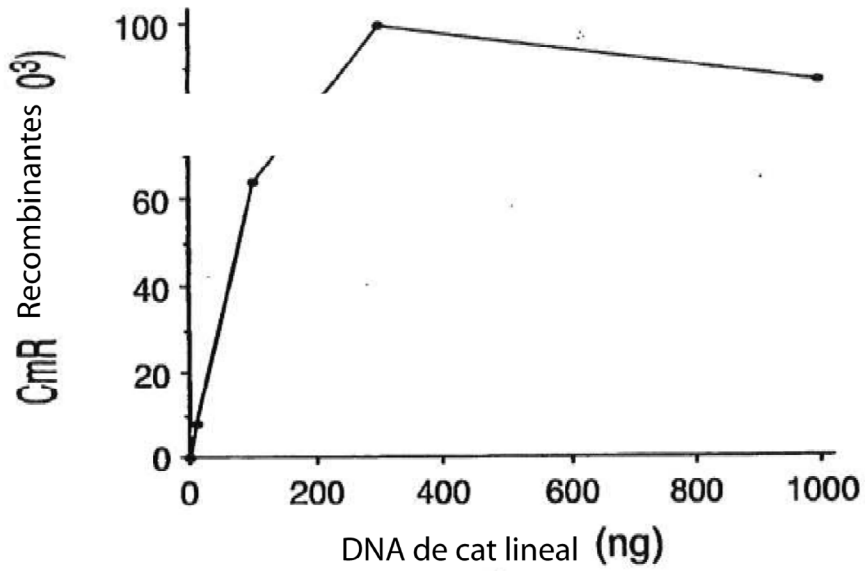


FIG. 5

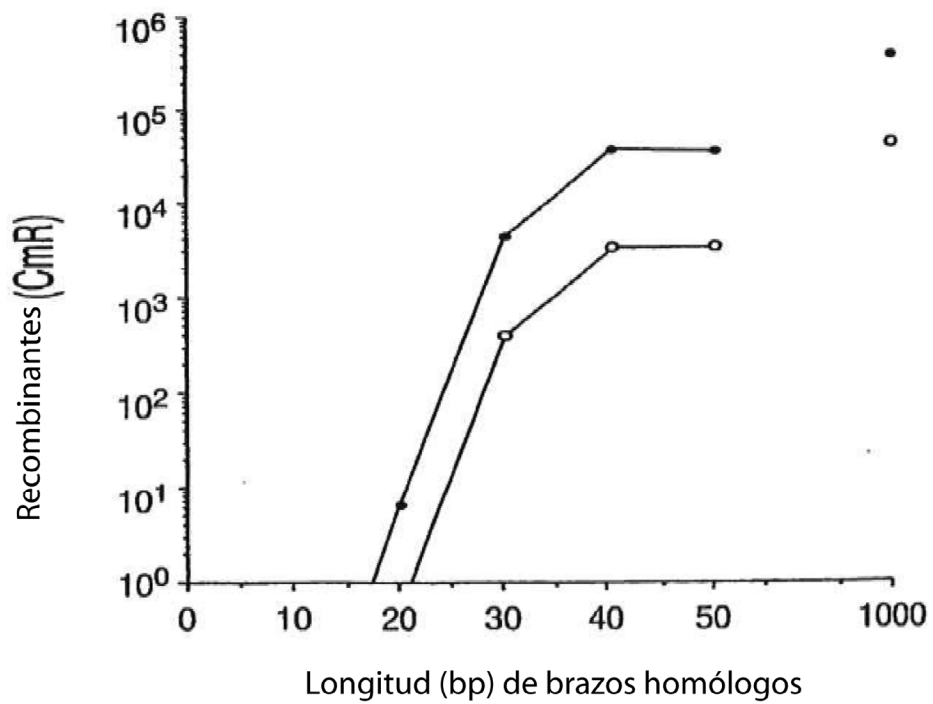


FIG. 6

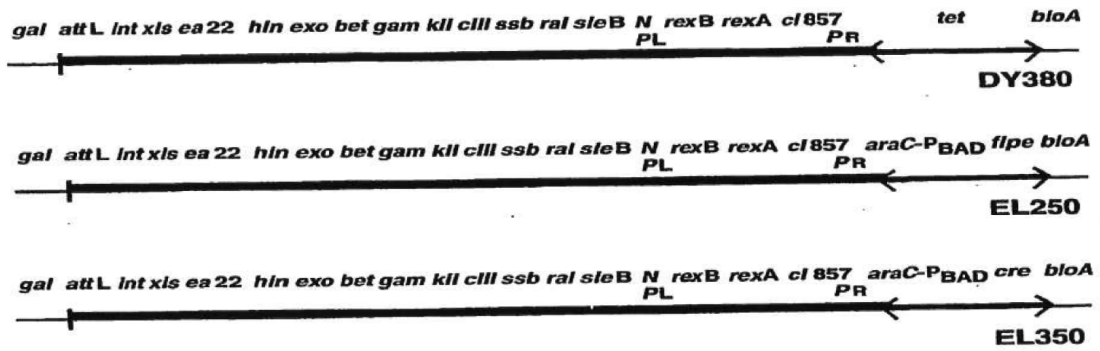


FIG. 7

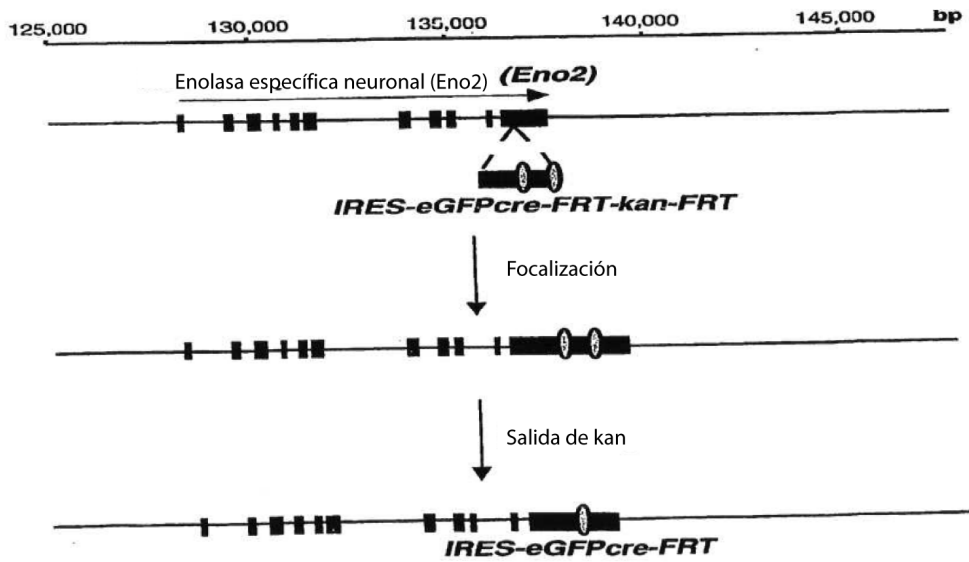


FIG. 8

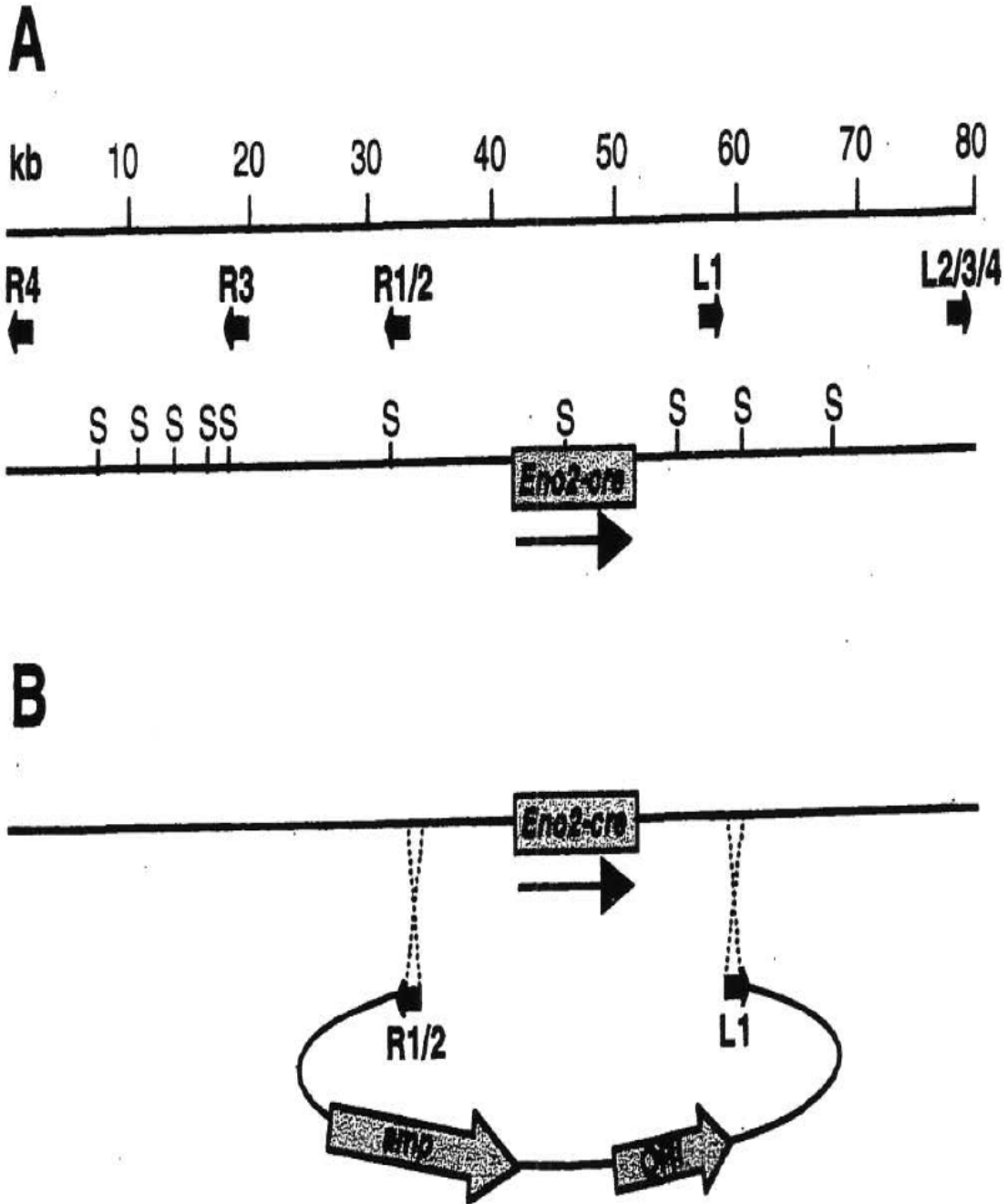


FIG. 9

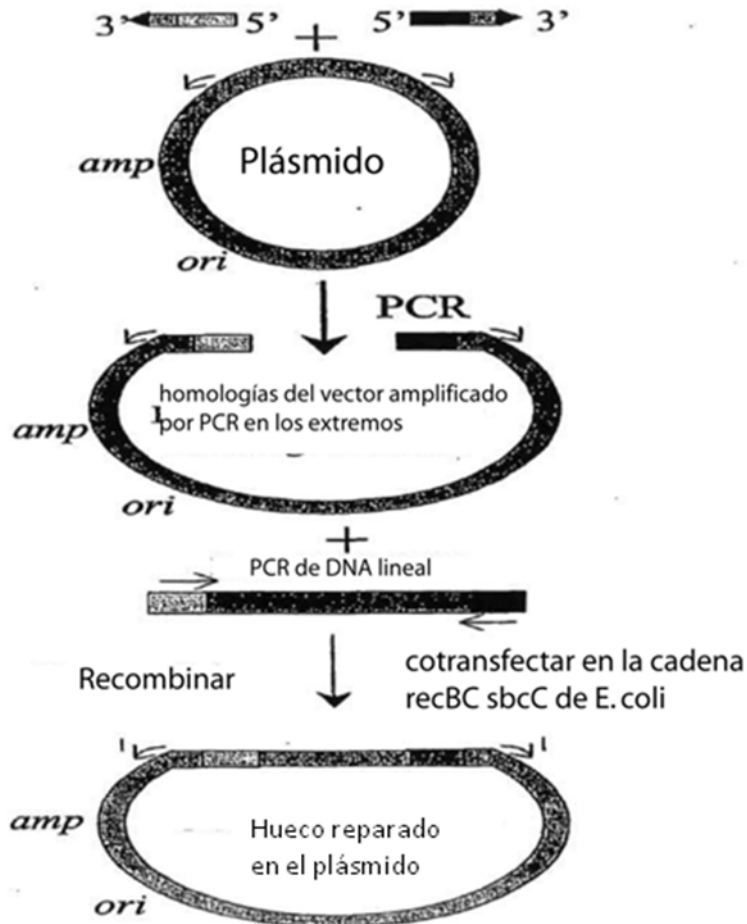




FIG. 10

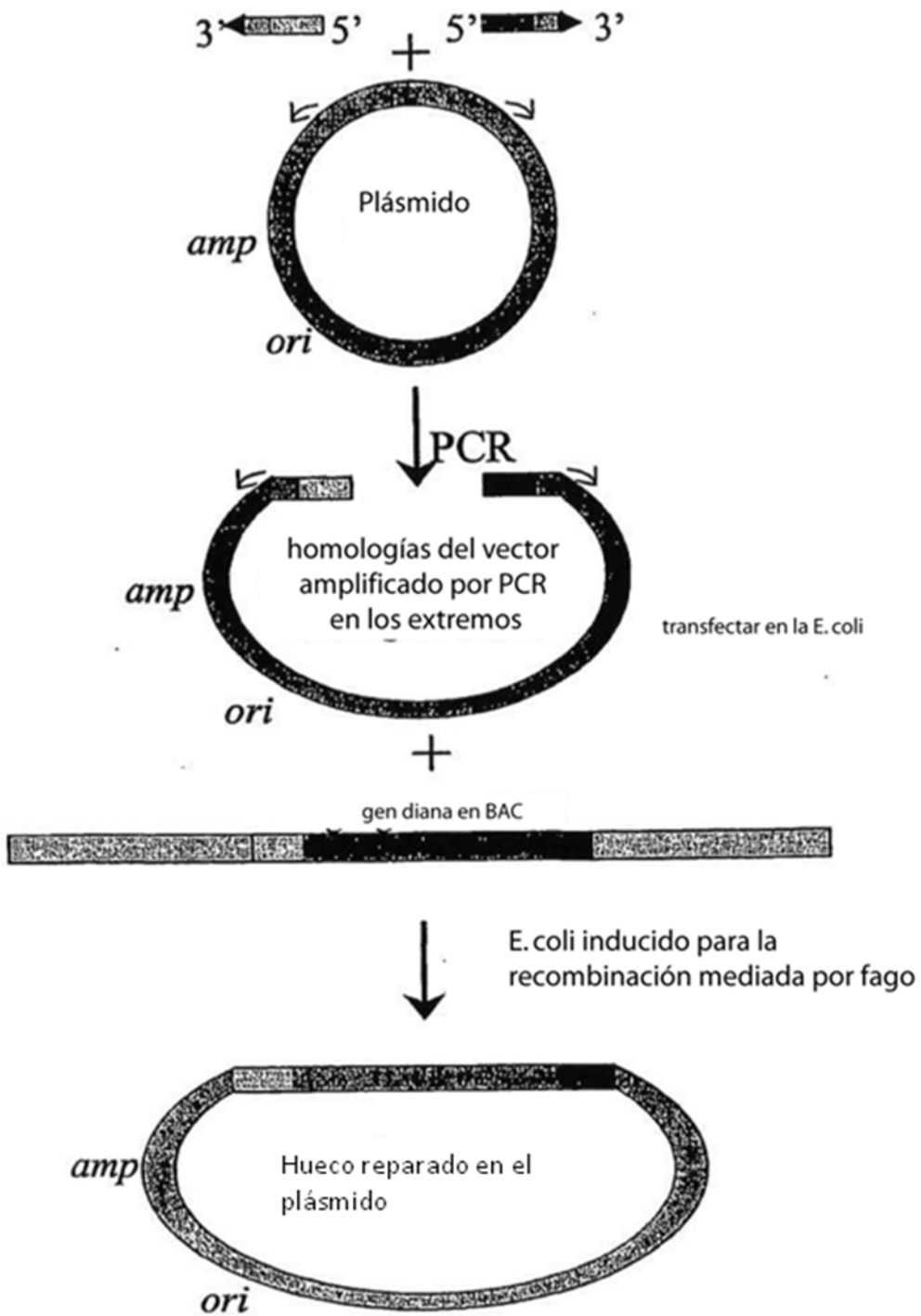


FIG. 11

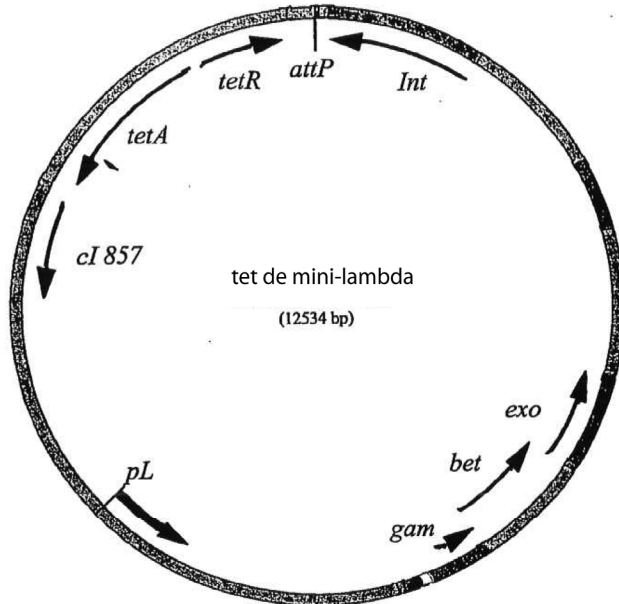
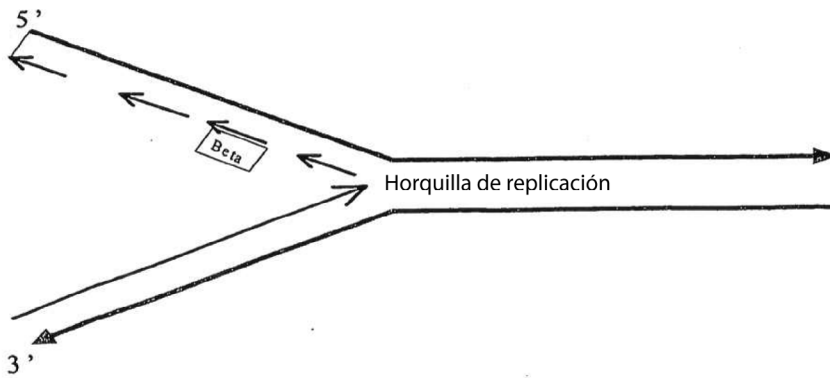


FIG. 12



# FIG. 13

Recombinación con extremo 5' protuberante  
El efecto de la longitud del solapamiento

