



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 334 577**

51 Int. Cl.:

C12N 15/90 (2006.01)

C12N 15/85 (2006.01)

C12N 5/10 (2006.01)

A61K 48/00 (2006.01)

A61K 31/70 (2006.01)

A01K 67/027 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06008807 .7**

96 Fecha de presentación : **29.08.2000**

97 Número de publicación de la solicitud: **1681355**

97 Fecha de publicación de la solicitud: **19.07.2006**

54

Título: **Recombinación de ADN específico de secuencia en células eucariotas.**

30

Prioridad: **30.08.1999 DE 199 41 186**

45

Fecha de publicación de la mención BOPI:
12.03.2010

45

Fecha de la publicación del folleto de la patente:
12.03.2010

73

Titular/es: **Peter Dröge**
100 Nanyang Crescent Meadows 11-01
Singapore 637819, SG

72

Inventor/es: **Dröge, Peter;**
Christ, Nicole y
Lorbach, Elke

74

Agente: **Carpintero López, Mario**

ES 2 334 577 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 de ADN específico de secuencia en células eucariotas.

5 La presente invención se refiere a un procedimiento para la recombinación específica de secuencia de ADN en células eucariotas *in vitro* que comprende la introducción en una célula de una primera secuencia de ADN que comprende una secuencia *attB* según SEQ ID N°:1 o su homólogo, o una secuencia *attP* según SEQ ID N°:2 o su homólogo, o una secuencia *attL* según SEQ ID N°:3 o su homólogo, o una secuencia *attR* según SEQ ID N°:4 o su homólogo, la introducción de una segunda secuencia de ADN en la célula que comprende una secuencia *attB* según SEQ ID N°:1 o su homólogo, o una secuencia *attP* según SEQ ID N°:2 o su homólogo, o una secuencia *attL* según SEQ ID N°:3 o su homólogo, o una secuencia *attR* según SEQ ID N°:4 o su homólogo, y la realización de la recombinación específica de secuencia mediante la acción de una integrasa Int del bacteriófago *lambda*, designando el término “homólogo” secuencias *attB*, *attP*, *attL* y *attR* que son idénticas al menos en el 70% en comparación con las secuencias de recombinación de origen natural. Una forma de realización preferida se refiere a un procedimiento que comprende además la realización de una segunda recombinación específica de secuencia de ADN mediante la acción de una Int y de un factor Xis.

La manipulación controlada del genoma eucariota es un procedimiento importante para estudiar la(s) función(es) de determinados genes en el organismo vivo. Además de esto desempeñan una función en el procedimiento de terapia génica en el campo de la medicina. La preparación de cepas animales transgénicas, la modificación de genes o recortes de genes (denominados “genes diana”) y la integración dirigida de ADN extraño en el genoma de eucariotas superiores son de especial importancia a este respecto. Mediante la caracterización y el uso de sistemas de recombinación específicos de secuencia pudieron mejorarse estas tecnologías en los últimos tiempos.

25 Las recombinasas de ADN específicas de secuencia conservativas se clasifican en dos familias.

Los miembros de la primera familia, la denominada familia de la “integrasa”, catalizan la separación y reconexión de hebras de ADN entre dos secuencias de nucleótidos definidas, que se designan en lo sucesivo secuencias de recombinación. Estas pueden encontrarse o bien sobre dos moléculas de ADN distintas o bien sobre una. Entonces se produce la recombinación inter- o intramolecular. En el último caso el resultado depende de la reacción de la disposición respectiva de las secuencias de recombinación unas respecto a otras. Si las secuencias de recombinación se encuentran en orientación inversa, es decir, opuestas unas respecto a otras, se produce la inversión del segmento de ADN que se encuentra entre ellas. Si las secuencias de recombinación se encuentran como repeticiones directas, es decir, en paralelo, sobre un sustrato de ADN, se produce la delección. En la recombinación intermolecular, es decir, cuando las dos secuencias de recombinación están situadas sobre dos moléculas de ADN distintas, puede producirse una fusión de las dos moléculas de ADN. Mientras que los miembros de la familia de la integrasa catalizan normalmente tanto la recombinación intra- como la intermolecular, las recombinasas de la segunda familia, las denominadas “invertasas/resolvasas”, sólo pueden efectuar la recombinación intramolecular.

40 Las recombinasas que actualmente se usan principalmente para la manipulación de genoma eucariota pertenecen a la familia de la integrasa. Éstas son la recombinasa Cre del bacteriófago *P1* y la recombinasa FLP de levaduras; véase Müller, U. (1999) Mech. Develop., 82, página 3. Las secuencias de recombinación a las que se une la recombinasa Cre se designan *loxP*. *loxP* es una secuencia de nucleótidos de 34 pb (pares de bases) de longitud que está constituida por dos secuencias de nucleótidos invertidos de 13 pb de longitud y un espaciador de 8 pb de longitud que se encuentra entre ellas; véase Hoess, R. y col. (1985) J. Mol. Biol., 181, página 351. Las secuencias de unión para FLP, designadas FRT, se construyen de forma similar, pero se diferencian de *loxP*; véase Kilby, J. y col. (1993) Trends Genet., 9, página 413. Por tanto, las secuencias de recombinación no pueden intercambiarse unas con otras, es decir, Cre no puede recombinar secuencias FRT y FLP tampoco secuencias *loxP*. Ambos sistemas de recombinación son activos en amplias distancias, es decir, el segmento de ADN que se va a invertir o deleccionar, flanqueado por dos secuencias *loxP* o FRT, puede ser de más de 10.000 pares de bases (kb) de longitud.

Con estos dos sistemas se produjo, por ejemplo, la recombinación específica del tejido en el modelo de ratón, la translocación cromosómica en plantas y animales, y una inducción controlada de la expresión génica; véase el artículo de revisión de Müller, U. (1999) Mech. Develop., 82, página 3. De este modo se deleccionó el gen de la ADN polimerasa β en determinados tejidos del ratón; véase Gu, H. y col. (1994) Science, 265, página 103. Otro ejemplo es la activación específica del oncogén del virus del tumor de ADN SV40 en las pupilas de los ojos del ratón, que condujo a la formación de tumor exclusivamente en estos tejidos. Además, la estrategia Cre-*loxP* también se usó en relación con los promotores inducibles. Para ello se reguló, por ejemplo, la expresión de la recombinasa con un promotor inducible por interferón, lo que condujo a la delección de un gen determinado en el hígado y no -o sólo mínimamente- en otros tejidos; véase Kühn, R. y col. (1995) Science, 269, página 1427.

Hasta ahora se han usado dos miembros de la familia de la invertasa/resolvasa para la manipulación del genoma eucariota. Un mutante del gen de la invertasa de bacteriófagos Mu puede catalizar sin cofactores la inversión de un fragmento de ADN en protoplastos vegetales. Sin embargo, se comprobó que este mutante es hiperrecombinante, es decir, también cataliza separaciones de hebras de ADN distintas de sus secuencias de recombinación naturales. Esto conduce a resultados de recombinación espontáneos parcialmente letales en el genoma de protoplastos vegetales. La β -recombinasa de *Streptococcus pyogenes* cataliza la recombinación entre dos secuencias de recombinación directa-

mente repetidas en cultivos celulares de ratón, lo que conduce al recorte del segmento. Sin embargo, además de la delección también se detectó la inversión, con lo que el uso controlado de este sistema es apto para la manipulación del genoma eucariota.

La manipulación del genoma eucariota con la recombinasa Cre o FLP también presenta desventajas considerables. En la delección, es decir, la recombinación de dos secuencias de recombinación *loxP* o *FRT* presentes repetidamente de forma directa en un genoma, se produce la pérdida irreversible del segmento de ADN que se encuentra entre ellas. Si sobre este segmento de ADN se encuentra un gen, en consecuencia éste se va perder para siempre en la célula y en el organismo. Por tanto, no es posible la reproducción del estado original para un nuevo análisis de la función del gen en estas células, por ejemplo, en un estadio de desarrollo tardío del organismo. La pérdida definitiva del segmento de ADN mediante delección puede evitarse mediante la inversión del segmento de ADN en cuestión. Con esto podría inactivarse un gen sin que llegue a perderse, y conectarse de nuevo en un momento posterior en el desarrollo o en el animal adulto mediante una expresión regulable temporalmente de la recombinasa mediante una recombinación inversa. Sin embargo, la desventaja en el uso de la recombinasa Cre o FLP en este procedimiento modificado se basa en que no puede regularse la inversión ya que las secuencias de recombinación no se cambian con el proceso de recombinación. En consecuencia se producen procesos de recombinación repetidos en los que la inactivación del gen en cuestión es provocada por la inversión del segmento de ADN en cuestión en sólo algunas, máximo el 50%, de las células diana. Se ha intentado resolver este problema al menos parcialmente construyendo secuencias *loxP* mutadas que ya no pueden usarse después de la recombinación única para otras reacciones. Sin embargo, a este respecto, la desventaja se encuentra en la unicidad de la reacción, es decir, después de la inactivación de un gen mediante inversión ya no puede realizarse una activación posterior mediante la recombinación inversa.

Otra desventaja de la recombinasa FLP es su reducida estabilidad al calor a 37°C, lo que limita considerablemente la eficiencia de la reacción de recombinación en eucariotas superiores, por ejemplo, del ratón con una temperatura corporal de aproximadamente 39°C. En este caso se construyeron mutantes FLP que presentaban una estabilidad al calor mayor que la de la recombinasa de tipo salvaje, sin embargo, todavía muestran una eficiencia de recombinación inferior en comparación con la recombinasa Cre.

Además, en el campo de la medicina se encuentra un uso de recombinasas específicas de secuencia, por ejemplo, en la terapia génica, en la que las recombinasas deben integrar de forma estable y específica un segmento de ADN deseado en el genoma de la célula diana humana respectiva. Tanto la Cre como la FLP pueden catalizar la recombinación intermolecular. Ambas recombinasas recombinan un ADN de plásmido, que lleva una copia de su secuencia de recombinación respectiva, con una secuencia de recombinación correspondiente insertada previamente de forma estable por recombinación homóloga en el genoma eucariota. Sin embargo, se desea que esta reacción pueda realizarse con una secuencia de recombinación de origen "natural" en el genoma eucariota. Debido a que *loxP* y *FRT* tienen 34 ó 54 nucleótidos de longitud, un origen de estas secuencias de recombinación como componente del genoma es extremadamente poco probable por motivos estadísticos. Incluso con la presencia resulta la desventaja de la reacción inversa previamente descrita, es decir, después de la integración satisfactoria tanto la recombinasa Cre como la recombinasa FLP pueden recortar de nuevo el segmento de ADN insertado mediante recombinación intramolecular.

Por tanto, un objetivo de la presente invención consiste en poner a disposición un sistema de recombinación simple y regulable y el medio de trabajo necesario. Además, un objetivo de la presente invención consiste en poner a disposición un sistema de recombinación y el medio de trabajo necesario que puede realizar una integración estable y específica de una secuencia de ADN seleccionada.

Los objetivos se alcanzan mediante el objeto definido en las reivindicaciones.

El procedimiento según la invención se explica mediante las siguientes figuras.

La Figura 1 muestra una representación esquemática de las reacciones de recombinación, integración y escisión catalizadas por la integrasa Int. Se representa un ADN de plásmido tensado superhelicoidal (arriba) que lleva una copia de la secuencia de recombinación *attP*. Ésta contiene cinco denominados puntos de unión al brazo para Int (P1, P2, P1', P2', P3'), dos puntos de unión de Int al "núcleo" (C y C'); marcados mediante flechas negras), tres puntos de unión para IHF (H1, H2, H'), dos puntos de unión para Xis (X1, X2) y la denominada región de "solapamiento" (rectángulo abierto) en la que tiene lugar el intercambio de la hebra de ADN propiamente. La secuencia asociada para *attP*, *attB* se representa debajo sobre un segmento de ADN lineal. Está constituido por dos puntos de unión del "núcleo" para Int (B y B'); marcados por flechas abiertas) y la región de "solapamiento". La recombinación entre *attB* y *attP* exige Int e IHF. Ésta conduce a la integración del plásmido en el segmento de ADN que lleva *attB*. Para este fin se forman dos nuevas secuencias de recombinación híbridas, *attL* y *attR*, que sirven de secuencias diana para la escisión. Esta reacción necesita, además de Int e IHF, otro cofactor, Xis, que está codificado por fagos *lambda*.

La Figura 2A muestra una representación esquemática de los vectores de expresión de la integrasa y la Figura 2B una representación esquemática de un análisis Western. (A): el vector pKEXInt contiene el gen de la integrasa de tipo salvaje, el vector pKEXInt-h contiene el gen del mutante Int-h, y el vector pKEXInt-h/218 contiene el gen del mutante Int-h/218. El vector de control (pKEX) no contiene ningún gen de Int. Los genes respectivos para la integrasa de tipo salvaje (Int) y los dos mutantes (Int-h e Int-h/218) se representan como barras grises. Detrás de las regiones codificantes se encuentran las secuencias de señal para el procesamiento de ARN que deben garantizar una elevada estabilidad intracelular del ARNm respectivo (rectángulos punteados) y se designan SV-40, corte y empalme t-Ag y

señales poli A. La expresión de los genes de la integrasa se realiza mediante el promotor del citomegalovirus (CMV) humano. (B): después de introducirse el vector respectivo, como se muestra, en las líneas celulares indicadoras B2 y B3 se prepararon lisados celulares y se separaron proteínas en una electroforesis en gel de SDS-poliacrilamida según su peso molecular. La presencia de la proteína Int-h se hizo visible mediante anticuerpos de ratón policlonales, que se dirigen contra Int de tipo salvaje (carriles 2 y 4). La posición de Int en el gel está marcada con una flecha.

La Figura 3 muestra una representación esquemática de los vectores sustrato. (A): pGFP *attB/attP*. Se representa el vector linealizado con ApaLI. Las grandes flechas negras marcan la posición y la orientación de ambas secuencias de recombinación, *attB* y *attP*. Flanquean el gen GFP (proteína verde fluorescente) que está presente en orientación inversa respecto al promotor del CMV. pA designa la señal de poliA. Adicionalmente, sobre el vector se encuentra el gen de resistencia "neo" que se expresa con el promotor de SV40 temprano y hace posible la selección de líneas celulares indicadoras estables. También se marcan los puntos de corte de restricción para la enzima NcoI. La recombinación integrativa entre *attB* y *attP* conduce a la inversión del gen GFP y, por tanto, a su expresión. Las flechas pequeñas, abiertas y cerradas marcan la posición y la orientación del cebador de PCR individual y se designan p1 a p7. (B): pGFPattL/*attR*. El vector es en la constitución idéntico a pGFPattB/*attP*, sin embargo contiene *attL* y *attR* en lugar de *attB* y *attP*. El gen GFP se encuentra respecto al promotor del CMV en la orientación 3'-5'. Las cajas sombreadas designan la posición de la sonda (muestra), que se usó para el análisis Southern.

La Figura 4A a D muestra en una representación esquemática la detección de la recombinación integrativa en líneas celulares indicadoras mediante PCR después de la separación de las moléculas de ADN en geles de agarosa (1,2% en peso/volumen), en cuyo ADN se hizo visible mediante coloración con bromuro de etidio. (A): PCR de transcriptasa inversa (RT-PCR). Los vectores pKEX y pKEXInt-h (véase Figura 2A) se incorporaron por separado en la línea celular indicadora respectiva B1 a B3 mediante electroporación. El análisis por RT-PCR partiendo de poli A-ARNm aislado muestra el producto esperado mediante el par de cebadores p3/p4 (véase la Figura 3) sólo si las células se trataron con pKEXInt-h (carril 1, 3 y 5). Para el control del contenido en ARN se amplificó el gen de β -actina a partir de las mismas preparaciones de ARN. Carril M: patrón de longitud de ADN; carril O: RT-PCR de control sin molde de ARN; (B, C): análisis de PCR genómico. Se aisló ADN genómico a partir de las líneas celulares respectivas 72 horas después de la electroporación y se amplificó con pares de cebadores p3/p4 (véase la Figura 3) y p1/p2 (véase la Figura 3). La numeración y la designación de los carriles se corresponde con la de en la (Figura 4A). (D): ensayo en delección. Se amplificó ADN genómico aislado con el par de cebadores p5/p6 (véase la Figura 3). La posición del producto de PCR esperado (420 pb) mediante delección, en lugar de inversión, está marcada con una flecha. La numeración y la designación de los carriles es como en la Figura 4A.

La Figura 5A y B muestra en una representación esquemática la detección de la inversión en líneas celulares indicadoras mediante PCT e hibridación Southern después de la separación de las moléculas de ADN en geles de agarosa (1,2% en peso/volumen). (A): análisis por PCR. Se amplificó una fracción de ADN genómico que se aisló a partir de las líneas celulares B1, B2, B3 y BL60 tratadas con los vectores pKEX y pKEXInt-h con pares de cebadores p3/p4 y p5/p7 (véase la Figura 3). Los productos de PCR a los que se llegó con la inversión catalizada con integrasa del gen GFP se observan en los carriles 1, 3 y 5. Carril M: patrón de longitud de ADN; carril O: PCR de control sin ADN genómico. (B) Análisis Southern: la fracción restante del ADN analizado en la Figura 5A se incubó con la enzima de restricción NcoI, se separó mediante una electroforesis en gel de agarosa según peso molecular y a continuación se transfirió a una membrana de nitrocelulosa. Para la detección de la recombinación se hicieron visibles fragmentos de ADN que llevaban GFP mediante una sonda marcada radiactivamente (véase la Figura 3B). Carril 9: pGFPattB/*attP* no recombinado; carril 10: pGFPattB/*attP* recombinado.

La Figura 6A muestra una representación de secuencias de ácido nucleico que comprenden *attB* o *attH*. La Figura 6B muestra una representación de secuencias parciales de *attP* y *attP**. (A): comparación de secuencia entre *attB* y *attH*. Los puntos de unión del "núcleo" de Int B y B' en *attB* están marcados con una línea sobre las secuencias. Los puntos de unión del "núcleo" de Int H y H' en *attH* están marcados con una línea discontinua sobre las secuencias. Las secuencias de "solapamiento" están marcadas por rectángulos abiertos. Las diferencias de secuencia están marcadas con líneas dobles verticales. La numeración de los restos en las regiones núcleo y de solapamiento se refiere al centro del solapamiento definido por Landy y Ross ((1977), Science, 197, página 1147), que se designa con O. *attB* o *attH* es la secuencia de -9 a +11. (B): comparación de secuencia entre secuencias parciales de *attP* y *attP** que se corresponden con las secuencias de *attB* o *attH*. Las designaciones se seleccionan como en la Figura 6A.

La Figura 7 muestra en una representación esquemática la detección de la recombinación entre *attH* y *attP** sobre el vector pACH en *E. coli* después de la separación en una electroforesis en gel de agarosa. El vector sustrato pACH se cotransformó junto con los vectores de expresión procariota respectivos para Int, Int-h o Int-h/218 en la cepa de *E. coli* CSH26 o CSH26 delta IHF. El ADN de plásmido se aisló 36 horas después de la selección, se incubó con las enzimas de restricción HindIII y Aval, se separó mediante electroforesis en gel de agarosa y se hizo visible. La posición de los fragmentos de restricción formados mediante la inversión se indica con "invers". La posición del ADN no recombinante se marca con pACH. Carriles 1 y 12: patrones de longitud de ADN; carriles 2 y 3: vector de expresión o pACH-ADN no recombinado; carriles 4 a 7: ADN aislado de CSH26; carriles 8 a 11: ADN aislado de CSH26 delta IHF.

La Figura 8 muestra una representación esquemática de la estrategia para la integración del vector pEL13 en el locus de *attH* genómico y el principio del procedimiento de detección. El vector de integración pEL13 lleva un gen de resistencia (flecha caracterizada con "higr"), el gen para Int-h (flecha caracterizada con "int h") con el control del

promotor del CMV y una copia de *attP** (rectángulo abierto caracterizado con “*attP*/P*OP**”). Después de la incorporación del vector en células BL60 mediante electroporación se expresa Int-h (véase la Figura 2B). La recombinasa cataliza entonces la recombinación intermolecular entre *attP** y *attH* localizado cromosómico (rectángulo sombreado caracterizado con “*attH/HOH**”). Esto conduce a la integración del vector pEL13 en el genoma de las células BL60.

5 Las células que han captado el vector de forma estable pueden seleccionarse e identificarse mediante PCR con el par de cebadores *attX1/B2* (flechas caracterizadas con “*attX1*” y “*B2*”). EcoR V y SpH I designan puntos de corte de restricción para las enzimas de restricción correspondientes.

La Figura 9 muestra en una representación esquemática la detección de la recombinación intermolecular entre *attP** (pEL13) y *attH* en células BL60. El ADN genómico se aisló después de la electroporación de pEL13 y posterior selección durante varias semanas de 31 poblaciones celulares distintas y se amplificó con el par de cebadores *attX1/B2* (véase la Figura 8). Los productos de PCR se separaron mediante electroforesis en gel de agarosa y se hicieron visibles. La posición del producto esperado (295 pb) en los geles está marcada por la flecha. Los productos se analizaron a continuación mediante secuenciación de ADN. En el borde derecho de los geles se encuentra un patrón de longitud.

El término “transformación” o “transformar” usado en esta invención designa la introducción respectiva de una secuencia de ácido nucleico en una célula. La introducción puede ser, por ejemplo, una transfección o lipofección, o se realiza mediante el procedimiento del calcio, procedimiento de electrochoque o una inyección de oocitos. El término “transformación” o “transformar” también significa en esta invención la introducción de una secuencia de ácido nucleico viral que comprende, por ejemplo, la(s) secuencia(s) de recombinación y un gen o fragmento de gen terapéutico en la trayectoria seguida de forma natural por el virus respectivo. En este caso, la secuencia de ácido nucleico viral no debe presentarse como secuencia de ácido nucleico desnuda sino que puede estar envuelta en una envuelta de proteína viral. Por tanto, el término no sólo designa el procedimiento que normalmente se entiende con el término transformación o transformar.

El término usado en esta invención “derivado” designa secuencias de *attB* y *attP* y secuencias de *attL* y *attR* que presentan en comparación con las secuencias de recombinación de origen natural modificaciones en forma de una o varias, como máximo seis, especialmente dos, tres, cuatro o cinco sustituciones.

La expresión usada en esta invención “homólogo” u “homológamente” o “similar” en referencia a las secuencias de recombinación designa una secuencia de ácido nucleico que es idéntica a las secuencias de recombinación de origen natural aproximadamente en el 70%, preferiblemente aproximadamente en el 80%, con especial preferencia aproximadamente en el 85%, con más especial preferencia aproximadamente en el 90%, con más especial preferencia aproximadamente en el 95% y especialmente en el 99%.

El término usado en esta invención “vector” designa constructos de origen natural o sintético para la captación, multiplicación, expresión o transmisión de ácidos nucleicos, por ejemplo, plásmidos, fagémidos, cósmidos, cromosomas sintéticos, bacteriófagos, virus o retrovirus.

La integrasa (designada normalmente y en lo sucesivo “*Int*”) del bacteriófago *lambda* pertenece, como Cre y Flp, a la familia de la integrasa de las recombinasas de ADN específicas de secuencia conservativas. Int cataliza la recombinación integrativa entre dos secuencias de recombinación distintas, *attB* y *attP*. *attB* comprende 21 nucleótidos y originalmente se ha aislado del genoma de *E. coli*; véase Mizuuchi, M. y Mizuuchi, K. (1980) Proc. Natl. Acad. Sci. USA, 77, página 3220. Por el contrario, *attP* es con 243 nucleótidos considerablemente más larga y se origina de forma natural en el genoma del bacteriófago *lambda*; véase Landy, A. y Ross, W. (1977) Science, 197, página 1147. La recombinasa Int tiene en total siete puntos de unión en *attP* y dos en *attB*. La función biológica de Int es la integración específica de secuencia del genoma de fagos circular en el locus *attB* del cromosoma de *E. coli*. Para la recombinación integrativa, Int necesita un cofactor de proteína, el denominado “factor de integración del huésped” (“Integration Host Factor”) (designado normalmente y en lo sucesivo “IHF”), véase Kikuchi, Y. y Nash, H. (1978) J. Biol. Chem., 253, 7149. IHF se necesita para la construcción de un complejo de recombinación funcional con *attP*. Un segundo cofactor para la reacción de integración es el superenrollamiento de ADN negativo de *attP*. La recombinación entre *attB* y *attP* conduce finalmente a la formación de dos nuevas secuencias de recombinación, *attL* y *attR* que sirven como sustrato y secuencia de reconocimiento para una reacción de recombinación adicional, la reacción de escisión. Una amplia visión general de la integración de los bacteriófagos *lambda* se encuentra en, por ejemplo, Landy, A. (1989) Annu. Rev. Biochem., 58, página 913.

La recombinación por escisión (escisión) del genoma de fagos a partir del genoma de bacterias también está catalizada por la recombinasa Int. Para esto, además de Int e IHF se necesita sin falta otro cofactor que también esté codificado por bacteriófagos *lambda*. Se trata de la escisionasa (designada normalmente y en lo sucesivo Xis), que presenta dos puntos de unión en *attR*; véase Gottesman, M. y Weisberg, R. (1971). The Bacteriophage Lambda, Cold Spring Harbor Laboratory, página 113. Para la recombinación por escisión no se requiere, al contrario que para la recombinación integrativa, el superenrollamiento de ADN negativo de las secuencias de recombinación, sin embargo aumenta la eficiencia de la reacción de recombinación. Puede conseguirse una mejora adicional de la eficiencia de la reacción de escisión mediante un segundo cofactor, FIS (factor de estimulación de inversión, “Factor for Inversion Stimulation”), que puede actuar junto con Xis; véase Landy, A. (1989) Annu. Rev. Biochem., 58, página 913. La escisión es genéticamente la exacta reacción inversa de la integración, es decir, se generan de nuevo *attB* y *attP*. Una amplia visión general de la escisión de bacteriófagos *lambda* se encuentra, por ejemplo, en Landy, A. (1989) Annu. Rev. Biochem., 58, página 913.

Un aspecto de la presente invención se refiere a un procedimiento para la recombinación específica de secuencia de ADN en células eucariotas que comprende a) la introducción en una célula de una primera secuencia de ADN que comprende una secuencia *attB* según SEQ ID N°:1 o su homólogo, o una secuencia *attP* según SEQ ID N°:2 o su homólogo, o una secuencia *attL* según SEQ ID N°:3 o su homólogo, o una secuencia *attR* según SEQ ID N°:4 o su homólogo, b) la introducción en una célula de una segunda secuencia de ADN que comprende una secuencia *attB* según SEQ ID N°:1 o su homólogo, o una secuencia *attP* según SEQ ID N°:2 o su homólogo, o una secuencia *attL* según SEQ ID N°:3 o su homólogo, o una secuencia *attR* según SEQ ID N°:4 o su homólogo, y c) la realización de la recombinación específica de secuencia mediante la acción de una integrasa Int del bacteriófago *lambda*, designando el término “homólogo” secuencias *attB*, *attP*, *attL* y *attR* que son idénticas al menos en el 70% en comparación con las secuencias de recombinación de origen natural. Se prefiere un procedimiento comprendiendo la primera primera secuencia de ADN una secuencia *attB* según SEQ ID N°:1 o su derivado, y comprendiendo la segunda secuencia de ADN una secuencia *attP* según SEQ ID N°:2 o su derivado. También se prefiere un procedimiento comprendiendo la primera secuencia de ADN una secuencia *attL* según SEQ ID N°:3 o su derivado, y comprendiendo la segunda secuencia de ADN una secuencia *attR* según SEQ ID N°:4 o su derivado, realizándose en la etapa c) la recombinación específica de secuencia mediante la acción de una Int y de un factor Xis.

Otro aspecto de la presente invención se refiere a un procedimiento para la recombinación específica de secuencia de ADN en una célula eucariota *in vitro* que contiene en su genoma la primera secuencia de ADN, bien presentándose naturalmente en la célula o introduciéndose con ayuda de recombinación de ADN, que comprende la etapa b) y c) como se definieron anteriormente.

En este caso, el procedimiento según la invención no puede realizarse sólo con la secuencia *attB* y/o *attP* o secuencia *attL* y/o *attR* de origen natural, sino también con secuencias *attB* y/o *attP* o secuencias *attL* y/o *attR* modificadas, por ejemplo, sustituidas. De este modo se observó, por ejemplo, la recombinación integrativa del bacteriófago *lambda* y *E. coli* entre secuencias homólogas *attP* y *attB* (mutantes de las secuencias de tipo salvaje) que presentan una o una combinación de las siguientes sustituciones en las siguientes posiciones en *attB*: G, T (en la posición -9); A, C, G (-8); C, A, T (-7); T, G, A (-6); C, A (-5); A (-4); G, A (-3); A, C, G (-2); A, C, G (-1); A, C, G (0); T, C, G (+1); A, C, G (+2); T, G, C (+3); A, G, T (+4); A, C, G (+5); G, T, (+6); G, T (+7); G, T, A (+8); C, G, A (+9); C, G, A (+10); T, A, C (+11) (Nash, H. (1981) Annu. Rev. Genet., 15, página 143; Nussinov, R. y Weisberg, R. (1986) J. Biomol. Struct. Dynamics, 3, página 1134) y/o presentan en *attP*: T (en la posición +1); C (+2) y A (+4), véase Nash, H. (1981) Annu. Rev. Genet., 15, página 143.

Por tanto, la presente invención se refiere a un procedimiento en el que la secuencia *attB* usada en comparación con la secuencia *attB* de origen natural según SEQ ID N°:1 y la secuencia *attP* usada en comparación con la secuencia *attP* de origen natural según SEQ ID N°:2 presentan una o varias sustituciones. Además, la presente invención se refiere a un procedimiento en el que la secuencia *attL* usada en comparación con la secuencia *attL* de origen natural según SEQ ID N°:3 y la secuencia *attR* usada en comparación con la secuencia *attR* de origen natural según SEQ ID N°:3 presentan una o varias sustituciones. Se prefiere un procedimiento en el que las secuencias de recombinación presentan una, dos, tres, cuatro o cinco sustituciones. Las sustituciones pueden presentarse tanto de la región de solapamiento (véase la Figura 6A, rectángulo abierto) como en la región núcleo (véase la Figura 6A, línea). También puede substituirse la región de solapamiento completa que comprende siete nucleótidos. Se prefiere especialmente un procedimiento en el que en la secuencia *attB* y *attP* se introducen sustituciones o en la región núcleo o en la región de solapamiento. Se prefiere la introducción de una sustitución en la región de solapamiento y la introducción simultánea de una o dos sustituciones en la región núcleo.

Para el procedimiento según la invención no se requiere que en una sustitución en *attB* se introduzca o invierta la sustitución correspondiente en *attP*, o en una sustitución en *attL* la sustitución correspondiente en *attR*. Una modificación mediante sustitución en secuencias de recombinación debe elegirse pudiendo realizarse la recombinación a pesar de la(s) modificación(es). Ejemplos de tales sustituciones se citan, por ejemplo, en las publicaciones de Nash, H. (1981), anteriormente y Nussinov, R. y Weisberg, R. (1986), anteriormente, y no deben considerarse como exhaustivas. Pueden introducirse otras modificaciones fácilmente, por ejemplo, mediante procedimientos de mutagénesis y se prueban en cuanto a su uso mediante recombinaciones de prueba.

Por tanto, la presente invención se refiere a un procedimiento en el que o la secuencia *attB* usada en comparación con la secuencia *attB* de origen natural según SEQ ID N°:1 y la secuencia *attP* usada en comparación con la secuencia *attP* de origen natural según SEQ ID N°:2, o bien la secuencia *attL* usada en comparación con la secuencia *attL* de origen natural según SEQ ID N°:3 o la secuencia *attR* usada en comparación con la secuencia *attR* de origen natural según SEQ ID N°:4 presentan una o varias sustituciones. Por tanto, una o varias sustituciones en una de las secuencias de recombinación no debe dar como resultado la sustitución correspondiente en la otra secuencia de recombinación.

En una forma de realización preferida del procedimiento según la invención, la secuencia *attB* comprende 21 nucleótidos y coincide con la secuencia originalmente aislada del genoma de *E. coli* (Mizuuchi, M. y Mizuuchi, K. (1980) Proc. Natl. Acad. Sci. USA, 77, página 3220) y la secuencia *attP* comprende 243 nucleótidos y coincide con la secuencia originalmente aislada del genoma del bacteriófago *lambda*; véase Landy, A. y Ross, W. (1977) Science, 197, página 1147.

En una forma de realización preferida adicional del procedimiento según la invención, la secuencia *attL* comprende 102 nucleótidos y la secuencia *attR* 162 nucleótidos y ambas coinciden con las secuencias originalmente aisladas del genoma de *E. coli*; véase Landy, A. (1989) Annu. Rev. Biochem., 58, página 913.

5 Para la realización del procedimiento según la invención, la primera secuencia de ADN comprende, además de la secuencia de recombinación, otras secuencias de ADN que permiten una integración en un lugar diana deseado en el genoma de la célula eucariota. Esta integración discurre por la recombinación homóloga, que está mediada por mecanismos de recombinación internos de la célula. Para ello, las otras secuencias de ADN deben ser homólogas con el ADN del lugar diana y tanto 3' como también 5' se deben encontrar en la secuencia *attB* o *attL*. El experto
10 conoce cuál debe ser el grado de homología y cómo de largas deben ser las secuencias 3' y 5' respectivas para que la recombinación homóloga discurra con una probabilidad suficiente; véase el artículo de revisión de Capecchi, M. (1989), Science, 244, página 1288.

15 La segunda secuencia de ADN con la secuencia de recombinación *attP* o *attR* puede también comprender secuencias de ADN que son necesarias para una integración en un lugar diana deseado por medio de la recombinación homóloga. Para el procedimiento según la invención, tanto la primera como la segunda secuencia de ADN pueden comprender las otras secuencias de ADN. Se prefiere un procedimiento en el que ambas secuencias de ADN comprenden las otras secuencias de ADN.

20 La introducción de la primera y segunda secuencia de ADN con o sin otras secuencias ADN puede realizarse tanto de forma sucesiva como también en una cotransformación, presentándose las secuencias de ADN sobre dos moléculas de ADN distintas. Se prefiere un procedimiento en el que la primera y la segunda secuencia de ADN, con o sin otras secuencias de ADN, estén presentes sobre una molécula de ADN individual y se introducen en las células eucariotas. Además, la primera secuencia de ADN pueden introducirse en una célula y la segunda secuencia de ADN
25 en otra célula, fusionándose las células a continuación. Por fusión debe entenderse en el sentido amplio el cruce de organismos y la fusión celular.

El procedimiento según la invención puede usarse, por ejemplo, para invertir en una recombinación intramolecular el segmento de ADN que se encuentra entre las secuencias de recombinación orientadas indirectamente. Además, el
30 procedimiento según la invención puede usarse para delecionar en una recombinación intramolecular el segmento de ADN que se encuentra entre las secuencias de recombinación orientadas directamente. Las secuencias de recombinación se presentan orientadas directamente cuando las secuencias se incorporan respectivamente en la dirección 5'-3' ó 3'-5'. Las secuencias de recombinación se presentan orientadas indirectamente cuando, por ejemplo, la secuencia *attB* se incorpora en la dirección 5'-3' y la secuencia *attP* en la dirección 3'-5'. Si las secuencias de recombinación se colocan mediante recombinación homóloga respectivamente en una secuencia de intrones 5' y 3' de un exón y la recombinación se realiza mediante la acción de la integrasa, entonces el exón gira en las secuencias de recombinación orientadas indirectamente (invierte) o deleciona en las secuencias de recombinación orientadas directamente. Mediante esto, el polipéptido codificado por el gen correspondiente puede perder su actividad o función o detener la transcripción mediante la inversión o delección, de modo que no se forma ningún transcrito (completo). De esta forma
40 puede estudiarse, por ejemplo, la función biológica del polipéptido codificado.

Sin embargo, la primera y/o segunda secuencia de ADN puede/pueden comprender otras secuencias de ácido nucleico que codifican uno o varios polipéptido(s) de interés. Así, por ejemplo, con las secuencias de recombinación puede introducirse una proteína de estructura, una enzima o una proteína regulatoria en el genoma que se expresa transitoriamente después de la recombinación intramolecular satisfactoria. El polipéptido introducido puede ser endógeno o exógeno. Además, puede introducirse una proteína marcadora. El experto sabe que esta enumeración de aplicaciones del procedimiento según la invención es sólo a modo de ejemplo y no es exhaustivo. Ejemplos de aplicaciones según la invención que se realizaron con las recombinasas Cre y FLP usadas anteriormente se encuentran, por ejemplo, en el artículo de revisión de Kilby, N. y col., (1993), Trenes Genet., 9, página 413.
50

El procedimiento según la invención puede usarse además para delecionar o para invertir sobre vectores segmentos de ADN sobre sustratos episomales en la recombinación intramolecular. Para ello podría usarse una reacción de delección, por ejemplo, para separar secuencias de envuelta de los denominados virus de la levadura. Este procedimiento tiene un amplio campo de aplicación en la producción industrial de vectores virales para aproximaciones de terapia génica; véase Hardy, S. y col., (1997), J. Virol., 71, página 1842.
55

La recombinación intermolecular conduce a la fusión de dos moléculas de ADN que poseen respectivamente una copia de *attB* y *attP* o *attL* y *attR*. Así puede introducirse, por ejemplo, *attB* en primer lugar mediante recombinación homóloga en un locus genómico conocido bien caracterizado de una célula. En esta secuencia *attB* genómica puede
60 integrarse a continuación un vector que lleva *attP* mediante recombinación intermolecular. Se prefiere en este procedimiento la expresión del mutante de integrasa Int-h/218 cuyo gen se encuentra sobre un segundo vector de ADN que se cotransfecta. Sobre el vector que lleva *attP* pueden encontrarse otras secuencias, por ejemplo, un gen para una proteína marcadora determinada, flanqueadas por secuencias *loxP/FRT*. Con esta aproximación puede conseguirse, por ejemplo, que en los análisis de expresión comparativos de distintos genes en un tipo de célula éstos no se influyan por influencias positivas o negativas del lugar de integración genómicos respectivo.
65

Para la realización del procedimiento según la invención debe actuar una integrasa sobre las secuencias de recombinación. En este caso, la integrasa o el gen de la integrasa y/o el factor Xis o el gen del factor Xis pueden estar presente

ya antes de la introducción de la primera y segunda secuencia de ADN en la célula eucariota, entre la introducción de la primera y segunda secuencia de ADN o después de la introducción de la primera y segunda secuencia de ADN. La integrasa usada para la recombinación específica de secuencia se expresa preferiblemente en la célula en la que se realiza la reacción. Para ello se introduce en las células una tercera secuencia de ADN que comprende un gen de la integrasa. Si la recombinación específica de secuencia se realiza con *attL/attR*, adicionalmente todavía puede introducirse un gen del factor Xis (cuarta secuencia de ADN) en las células. Se prefiere especialmente un procedimiento en el que la tercera y/o cuarta secuencia de ADN se integre en el genoma eucariota de la célula por recombinación homóloga o de forma aleatoria. Además, se prefiere un procedimiento en el que la tercera y/o cuarta secuencia de ADN comprenda secuencias de ADN regulatorias que provoquen una expresión espacial y/o temporal del gen de la integrasa y/o del gen del factor Xis.

La expresión espacial significa en este caso que la recombinasa o el factor Xis sólo se expresa en un determinado tipo de célula, por ejemplo, células hepáticas, células renales, células nerviosas o células del sistema inmunitario, mediante el uso de promotores específicos del tipo celular y sólo cataliza la recombinación en estas células. Puede conseguirse una expresión temporal en la regulación de la expresión de la integrasa/factor Xis mediante promotores que son activos a partir de o en un estadio de desarrollo determinado o en un momento determinado en el organismo adulto. Además, la expresión temporal puede producirse mediante el uso de promotores inducibles, por ejemplo, mediante promotores dependientes de la interferón o tetraciclina; véase el artículo de revisión de Müller, U. (1999) *Mech. Develop.*, 82, página 3.

La integrasa usada en el procedimiento según la invención puede ser tanto la integrasa de tipo salvaje como una integrasa modificada del bacteriófago *lambda*. Debido a que la integrasa de tipo salvaje puede realizar la reacción de recombinación sólo con ayuda de un cofactor, concretamente IHF, para el procedimiento según la invención se prefiere el uso de una integrasa modificada. Si se usa la integrasa de tipo salvaje para el procedimiento según la invención, para la reacción de recombinación se necesita adicionalmente IHF. La integrasa modificada está modificada de forma que puede realizar la reacción de recombinación sin IHF. La preparación de polipéptidos modificados y la selección en cuanto a la actividad deseada son estado de la técnica y se realizan de forma sencilla; véase Erlich, H. (1989) *PCR Technology*. Stockton Press. Las integrasas de bacteriófagos *lambda* preferidas son dos mutantes de Int que se designan Int-h e Int-h/218; véase Miller y col. (1980) *Cell*, 20 página 721; Christ, N. y Dröge, P (1999) *J. Mol. Biol.*, 288, página 825. Int-h contiene un resto de lisina en lugar del resto de glutamato en la posición 174 frente a Int de tipo salvaje. Int-h/218 contiene otro resto de lisina en lugar de un resto de glutamato en la posición 218 y se preparó mediante mutagénesis de PCR del gen Int-h. Estos mutantes pueden catalizar la recombinación tanto entre *attB/attP* como también entre *attL/attR* sin los cofactores IHF, Xis y el superenrollamiento negativo en *E. coli* e *in vitro*, es decir, con sustratos purificados en el tubo de ensayo. En células eucariotas, los mutantes necesitan para la recombinación entre *attL/attR* sólo el cofactor Xis. Puede producirse una mejora adicional de la eficiencia de la recombinación entre *attL/attR* mediante otro cofactor, por ejemplo FIS. Se prefiere el mutante Int-h/218 ya que puede catalizar la reacción integrativa independiente del cofactor con una elevada eficiencia; véase Christ, N. y Dröge, P. (1999) *J. Mol. Biol.*, 288, página 825.

El procedimiento según la invención puede realizarse en todas las células eucariotas. Las células pueden presentarse, por ejemplo, en un cultivo celular y comprenden todos los tipos de células vegetales o animales. Por ejemplo, las células pueden ser oocitos, células madre embrionarias no humanas, células madre hematopoyéticas o cualquier tipo de células diferenciadas. Se prefiere un procedimiento en el que la célula eucariota es una célula de mamífero. Se prefiere especialmente un procedimiento en el que la célula de mamífero es una célula humana, célula de mono, ratón, rata, conejo, hámster, cabra, vaca, oveja o cerdo.

Una forma de realización preferida de la presente invención se refiere además a un procedimiento en el que dado el caso se realiza una segunda recombinación específica de secuencia de ADN mediante la acción de una integrasa de bacteriófagos *lambda* y de un factor Xis. La segunda recombinación necesita las secuencias *attL* y *attR* generadas mediante una primera recombinación de *attB* y *attP* o sus derivados. Por tanto, la segunda recombinación específica de secuencia se limita a un procedimiento que en la primera recombinación específica de secuencia usa secuencias *attB* y *attP* o sus derivados. Tanto la integrasa de tipo salvaje como los mutantes de Int pueden catalizar sin la adición de otros cofactores sólo la denominada recombinación integrativa, es decir, recombinan *attB* con *attP* y no *attL* con *attR*, si éstas están integradas de forma estable en el genoma de las células. En este caso, la integrasa de tipo salvaje necesita para la denominada recombinación por escisión los factores IHF, Xis y el superenrollamiento negativo. Los mutantes de Int Int-h e Int-h/218 necesitan para la recombinación por escisión sólo el factor Xis. Por tanto, es posible dejar transcurrir de forma controlada dos reacciones de recombinación una tras otra, si están presentes en las células otros factores para la segunda reacción de recombinación, concretamente la reacción de escisión. En combinación con otros sistemas de recombinación ya usados, en este documento pueden determinarse nuevas estrategias para la manipulación controlada de genomas eucariotas superiores. Esto es posible ya que distintos sistemas de recombinación usan exclusivamente sus propias secuencias de recombinación.

Así, por ejemplo, el sistema Int se usa para integrar secuencias *loxP* y/o *FRT* de forma específica en un locus genómico de un genoma eucariota y a continuación para activar o inactivar un gen mediante expresión controlada de Cre y/o Flp. Además, puede usarse el sistema Int para eliminar secuencias *loxP/FRT* tras uso, es decir, recombinación mediante la recombinasa correspondiente a partir de un genoma.

Se prefiere un procedimiento en el que además se introduce otra secuencia de ADN que comprende un gen del factor Xis en las células. Se prefiere especialmente un procedimiento en el que la otra secuencia de ADN comprende además una secuencia de ADN regulatoria que provoca una expresión espacial y/o temporal del gen del factor Xis.

5 Así, por ejemplo, después de la recombinación intramolecular integrativa satisfactoria (inversión) mediante Int, lo que ha llevado a la activación/inactivación de un gen en un tipo de célula determinado, este gen se inactiva o activa de nuevo en un momento posterior mediante la expresión espacial y/o temporal inducible de Xis con la expresión simultánea de Int.

10 La invención se refiere además al uso de una secuencia *attB* según SEQ ID N°:1 o su derivado y a una secuencia *attP* según SEQ ID N°:2 o su derivado, o a una secuencia *attL* según SEQ ID N°:3 o su derivado y a una secuencia *attR* según SEQ ID N°:4 o su derivado en una recombinación específica de secuencia de ADN en células eucariotas. La célula eucariota puede presentarse también en la conexión celular de un organismo, por ejemplo, de un mamífero que no contiene integrasa o factor Xis en sus células. Este organismo puede usarse para el cruce con otros organismos que por su parte contienen la integrasa o el factor Xis en sus células, de modo que generan descendencia en cuyas células se realiza entonces la recombinación específica de secuencia. Por tanto, la invención también se refiere al uso de una integrasa o un gen de la integrasa y un factor Xis o un gen del factor Xis en una recombinación específica de secuencia en células eucariotas.

20 Los inventores han determinado una secuencia del genoma humano (en lo sucesivo designada *attH*) que presenta una homología de aproximadamente el 85% con *attB*. *attH* puede usarse como secuencia de recombinación para la integración de un ADN extraño en el genoma humano. Para ello, la segunda secuencia de recombinación *attP* puede modificarse correspondientemente para que la integrasa pueda realizar la reacción de recombinación con alta eficiencia. Los inventores podrían mostrar que *attH* con una de las versiones modificadas por los inventores de *attP*, que se designa en lo sucesivo *attP** y se enumera SEQ ID N°:5, puede recombinarse mediante Int-h en *E. coli*. Experimentos con células humanas mostraron que *attH* también se recombina como constituyente del genoma humano con *attP**, cuando Int-H es sintetizado transitoriamente por las células.

30 De esto resulta la posibilidad de integrar un ADN extraño circular que contiene una secuencia de recombinación *attP* de forma estable y específica en el locus de origen natural *attH* del genoma humano. En este caso, *attH* es sólo un ejemplo de una secuencia de recombinación de origen natural en el genoma humano. En el marco del proyecto Genoma Humano pueden determinarse otras secuencias que presentan una homología con *attB*. Éstas pueden usarse asimismo para la integración de un ADN extraño en el genoma humano. En función de la secuencia homóloga con *attB* presente en el genoma humano, se selecciona una secuencia de recombinación *attP* correspondiente en el ADN extraño circular. Se prefiere un ADN extraño circular que contiene la secuencia de ácido nucleico de la secuencia *attP* de origen natural. Se prefiere especialmente un derivado de la secuencia *attP* de origen natural que presenta como máximo seis, preferiblemente de una a cinco, especialmente tres sustituciones. Se prefiere especialmente un ADN extraño circular que comprende la secuencia de ácido nucleico *attP** según SEQ ID N°:5 que presenta una homología de aproximadamente el 95% con *attP*.

40 En este caso, la integrasa puede introducirse a las células bien como polipéptido o mediante un vector de expresión. El gen de la integrasa puede presentarse además como secuencia de ácido nucleico expresable sobre la molécula de ADN que comprende la secuencia *attP* modificada o la natural o la secuencia *attP**.

45 El ADN extraño circular que contiene la secuencia *attP* natural o su derivado u homólogo, especialmente la secuencia *attP** según SEQ ID N°:5, comprende igualmente el gen o fragmento de gen terapéutico que va a introducirse en el genoma. Los genes terapéuticos pueden ser, por ejemplo, el gen CFTR, el gen ADA, el gen del receptor de LDL, el gen de la β -globina, el gen del factor VIII o el gen del factor IX, el gen de alfa-1-antitripsina o el gen de distrofina. El ADN extraño circular puede ser, por ejemplo, un vector viral ya usado en terapias génicas somáticas. El vector puede ser igualmente específico de la célula, de modo que sólo transfecta las células deseadas para la terapia génica, por ejemplo, las células del epitelio del pulmón, las células madre de la médula ósea, los linfocitos T, los linfocitos B, las células hepáticas, las células renales, las células nerviosas, las células del músculo esquelético, las células madre hematopoyéticas o los fibroblastos. Es evidente para el experto que este listado sólo representa una selección de los genes terapéuticos y células diana y pueden usarse igualmente otros genes y células diana para una terapia génica. Los fragmentos de gen comprenden, por ejemplo, delecciones de genes terapéuticos, exones individuales, secuencias de ácido nucleico antisentido o ribozimas. Los fragmentos de gen pueden comprender además recortes de un gen que contienen las repeticiones de trinucleótidos de un gen, por ejemplo, del gen X frágil.

55 En el uso de la integrasa de tipo salvaje debe estar presente IHF para la recombinación. Se prefiere el uso de una integrasa modificada en la que la recombinación puede realizarse sin IHF. Se prefiere especialmente el uso de Int-h o Int-h/218.

60 Por tanto, la presente invención describe la secuencia *attP* de origen natural o su derivado o su homólogo. La invención describe especialmente la secuencia de ácido nucleico *attP** según SEQ ID N°:5. La presente invención *da a conocer* además un vector que comprende la secuencia *attP* de origen natural o su derivado, especialmente la secuencia de ácido nucleico *attP** según SEQ ID N°:5, y otra secuencia de ácido nucleico que comprende un gen terapéutico o su fragmento de gen. Se prefiere un vector en el que el gen terapéutico comprende un gen CFTR, ADA, receptor de LDL, alfa o beta globina, alfa 1-antitripsina, factor VIII o factor IX o su fragmento. El vector puede comprender asimismo elementos de ADN regulatorios que controlan la expresión del gen terapéutico o su fragmento de gen.

Ejemplos

1. Preparación de los vectores de expresión y sustrato

1.1. Vectores de expresión

Los vectores de expresión eucariotas para Int de tipo salvaje (pKEXInt), Int-h (pKEXInt-h), Int-h/218 (pKEXInt-h/218) y pEL13 son derivados de pKEX-2-X (Rittner y col. (1991) *Methods Mol. Cell. Biol.*, 2, página 176). El vector contiene el elemento promotor/potenciador del citomegalovirus humano (CMV) y corte y empalme de ARN y secuencias de señal de poliadenilación del antígeno tumoral del virus 40 de los simios pequeños (SV40). Los genes Int se clonaron mediante PCR con los siguientes cebadores: (3343) 5'-GCTCTAGACCACCATGGGAAGAAGGCGAAGTCA-3', que se encuentra en el extremo 5' del gen Int y (3289) 5'-AAGGAAAGCGGCCGCTCATTATTTGATTTCAATTTTGTCC-3', que se encuentra en el extremo 3'. La amplificación se realizó después de una primera etapa de desnaturalización a 95°C (4 min) con 30 ciclos de desnaturalización (95°C, 45 s), formación de cebador (55°C, 45 s), síntesis de ADN (27°C, 2 min) y una última etapa de síntesis durante 4 min a 72°C. El fragmento de PCR resultante se usó en el vector pKEX-2-XR con XbaI y NotI. Int-h se generó del vector pHN16 como molde (Lange-Gustafson, B. y Nash, H. (1989) *J. Biol. Chem.*, 259, página 12724). Int de tipo salvaje e Int-h/218 se generaron con pTrcInt o pTrcInt-h/218 como moldes (Christ, N. y Dröge, P. (1999) *J. Mol. Biol.*, 288, página 825). pEL13 contiene, además del gen Int-h, adicionalmente una copia de *attP**.

*AttP** se construyó partiendo de *attP* mediante mutagénesis por PCR. Se usaron para ello los siguientes oligonucleótidos:

- (03) 5'-GTTCAGCTTTTTGATACTAAGTTG-3',
- (04) 5'-CAACTTAGTATCAAAAAG CTGAAC-3',
- (PC) 5'-TTGATAGCTCTTCCGCTTTCTGTTACAGGTCATAATACC-3' y
- (PD) 5'-ACGGTTGCTCTTCCAGCCAGGGAGTGGGACAAAATTGA-3'.

La amplificación se realizó según una primera etapa de desnaturalización a 95°C (4 min) con 30 ciclos de desnaturalización (95°C, 45 s), formación de cebador (57°C, 1 min 30 s), síntesis de ADN (72°C, 30 s) y una última etapa de síntesis durante 4 min a 72°C. El producto de PCR se incubó con la enzima de restricción SapI y se ligó con el pKEXInt-h recortado con SapI. El plásmido de control pKEX no contiene el gen Int.

1.2 Vectores sustrato

Los vectores sustrato son derivados de pEGFP (Clontech). Los casetes de recombinación se encuentran bajo el control del promotor del CMV, que garantiza una expresión constructiva fuerte. Se construyó pGFP*attB*/*attP* recortando el gen GFP (proteína de fluorescencia verde) en primer lugar de pEGFP mediante AgeI y BamHI. La secuencia *attB* de tipo salvaje se usó como oligonucleótido de doble hebra en el vector recortado con AgeI mediante los siguientes oligonucleótidos:

- (B10B) 5'-CCGGTTGAAGCCTGCTTTTTTATACTAACTTGAGCGAACGC-3 y
- (BOBI) 5'-AATTGCGTTCGCTCAAGTTAGTATAAAAAAGCAGGCTTCAA-3'.

La secuencia *attP* de tipo salvaje se amplificó mediante PCR del vector pAB3 (Dröge, P. y Cozzarelli, N. (1989) *Proc. Natl. Acad. Sci.*, 86, página 6062) usando los siguientes cebadores:

- (p7) 5'-TCCCCCGGGAGGGAGTGGGACAAAATTGA-3 y
- (p6) 5'-GGGGATCCTCTGTTACAGGTCATAATAC-3'.

La amplificación se realizó después de una primera etapa de desnaturalización a 95°C (4 min) con 30 ciclos de desnaturalización (95°C, 45 s), formación de cebador (54°C, 30 s), síntesis de ADN (72°C, 30 s) y una última etapa de síntesis durante 4 min a 72°C. El fragmento de PCR que contenía *attP* se recortó con XmaI y BamHI y se ligó con un fragmento de restricción que contenía el gen GFP. Este fragmento de restricción de GFP se preparó con AgeI y EcoRI a partir de pEGFP. El producto de la ligadura se usó en el vector que contenía *attB* recortado con MfeI/BamHI. El vector sustrato resultante contiene el gen GFP en la orientación invertida respecto al promotor del CMV. Su funcionalidad se ensayó en recombinación integrativa en *E. coli* mediante Int de tipo salvaje.

pGFP *attL/attR* es idéntico a pGFP *attB/attP* hasta las secuencias de recombinación. Se construyó el vector re-combinando en primer lugar pGFP *attB/attP* en *E. coli*, lo que conduce a la formación de *attL* y *attR*. El gen GFP orientado correctamente respecto al promotor del CMV se recortó mediante una reacción de restricción parcial con BsiEI y HindIII. Para usar el gen GFP en la orientación invertida respecto al promotor del CMV, se amplificó el gen en primer lugar mediante PCR usando los siguientes cebadores:

(p2) 5'-AATCCGCGGTCGGAGCTCGAGATCTGAGTCC-3' y

(p3) 5'-AATCCCAAGCTTCCACCATGGTGAGCAAGGG-3' (Fig. 3).

La amplificación se realizó después de una primera etapa de desnaturalización a 95°C (4 min) con 30 ciclos de desnaturalización (95°C, 45 s), formación de cebador (56°C, 45 s), síntesis de ADN (72°C, 1 min) y una última etapa de síntesis durante 4 min a 72°C. El fragmento de la PCR se recortó a continuación con HindIII y BsiEI y se integró en el vector recortado parcialmente. Este contiene *attL* y *attR* en la orientación invertida. Por tanto, pGFP *attL/attR* muestra la misma estructura global que pGFP *attB/attP*, con excepción de la presencia de *attL/attR* en lugar de *attB/attP*.

El homólogo de *attB* humano, *attH*, se amplificó por el ADN humano purificado mediante PCR usando los siguientes cebadores:

(B3) 5'-GCTCTAGATTAGCAGAAATTCTTTTGTG-3' y

(B2) 5'-AACTGCAGTAAAAAGCATGCTCATCACCCC-3'.

La amplificación se realizó después de una primera etapa de desnaturalización a 95°C (5 min) con 30 ciclos de desnaturalización (95°C, 45 s), formación de cebador (42°C, 1,45 min), síntesis de ADN (72°C, 1,45 min) y una última etapa de síntesis durante 10 min a 72°C. Las secuencias de cebador para la generación de *attH* se tomaron de un EST (número de registro: N31218; base de datos de EMBL). Se confirmó la secuencia incompleta de *attH*, como está presente en la base de datos, y se completó mediante secuenciación de ADN del producto de PCR aislado (192 pb). A continuación se recortó el fragmento con XbaI y PstI y se usó en el vector tratado correspondiente pACYC187 (New England Biolabs). *attP** se preparó mediante mutagénesis dirigida por PCR como se describió (Christ, N. y Dröge, P. (1999) J. Mol. Biol., 288, página 825) y se usó en el vector que contenía *attH* en orientación invertida respecto a *attH*. Esta construcción dio lugar al vector de prueba pACH.

Se aislaron los ADN de plásmido de la cepa de *E. coli* DH5 α (Hanahan, D. (1983) J. Mol. Biol., 166, página 557) mediante cromatografía de afinidad (Qiagen, Alemania). Se controlaron los vectores de expresión y sustrato, así como todos los constructos generados por PCR, mediante el sistema de secuenciación 373^a DNA basado en fluorescencia (Applied Biosystems).

Las reacciones de PCR se realizaron con el "Master Mix Kit" (Qiagen, Alemania) y los productos se analizaron por electroforesis en gel de agarosa (0,8% peso/volumen) en tampón de TBE.

2. Cultivo celular y construcción de las líneas celulares indicadoras

Los análisis de expresión y recombinación transitorios se realizaron con una línea celular del linfoma de Burkitt humano (BL60; (Wolf, J. y col., (1990) Cancer Res., 50, página 3095)). Las células BL60 se cultivaron en medio RPMI1640 (Life Technologies, Inc.), que se enriqueció con suero bovino fetal al 10%, y contenía L-glutamina 2 mM, estreptomycin (0,1 mg/ml) y penicilina (100 unidades/ml).

Las líneas celulares indicadoras BL60, que habían integrado bien pGFP *attB/attP* o pGFP *attL/attR* de forma estable en el genoma, se construyeron como sigue: se linealizaron aproximadamente 20 μ g de cada vector con ApaLI, se purificaron mediante extracciones con fenol/cloroformo, se precipitaron con etanol y se incorporaron en aproximadamente 2×10^7 células mediante electroporación a 260 V y 960 mF usando un "pulsador de genes Bio-Rad". Se seleccionaron líneas celulares estables con G418/genticina (300 μ g/ml) y a continuación se caracterizaron mediante PCR, secuenciación de ADN y análisis Southern.

3. Análisis de recombinación *in vivo*

Para realizar la recombinación intramolecular *in vivo* se transfectaron aproximadamente 2×10^7 células de la línea celular indicadora BL60 respectiva con 40 μ g de cada vector de expresión circular mediante electroporación, como se describe en el ejemplo 2. Después de 72 horas, las células se recogieron mediante centrifugación y se aisló el ADN genómico de la mitad de las células según las indicaciones del fabricante (Qiaamp Blood Kit; Qiagen, Alemania) mediante cromatografía de afinidad. De la segunda mitad de las células se aisló bien ARN (Rneasy Kit, Qiagen, Alemania) o se preparó un lisado celular para el análisis Western (véase el ejemplo 4).

ES 2 334 577 T3

Los análisis de recombinación con pACH se realizaron como ya se describieron previamente (Christ, N. y Dröge, P. (1999) J. Mol. Biol., 288, página 825) en *E. coli*. Para ello se usaron las recombinasas Int, Int-h e Int-h/218. La recombinación esperada de pACH conduce a la inversión y se detectó mediante análisis de restricción con HindIII y Aval.

La recombinación intermolecular para la integración de pEL13 en la *attH* genómicamente localizada de células BL60 se llevó a cabo como sigue: se transfectaron 2×10^7 células con $20 \mu\text{g}$ de pEL13 circular mediante electroporación, como se describió anteriormente. Después de 48 horas se cultivaron las células en placas con una concentración de 1×10^6 células/ml de medio de selección ($200 \mu\text{g/ml}$ de higromicina B) y se incubaron durante 6-8 semanas. De una parte de la población de células supervivientes se preparó en este momento ADN genómico según las indicaciones del fabricante (Qiaamp Blood Kit; Qiagen, Alemania).

Para comprobar la recombinación integrativa y por escisión intramolecular se amplificaron $0,4 \mu\text{g}$ de ADN genómico mediante PCR. Para ello se usaron de 20-50 pmol de los siguientes cebadores:

- (p1) 5'-GGCAAACCGGTTGAAGCCTGCTTTT-3';
- (p2) 5'-AATCCGCGGTCGGAGCTCGAGATCTGAGTCC-3';
- (p3) 5'-AATCCCAAGCTTCCACCATGGTGAGCAAGGG-3';
- (p4) 5'-AACCTCTACAAATGTGGTATGG-3';
- (p5) 5'-TACCATGGTGATGCGGTTTTG-3';
- (p6) 5'-GGGGATCCTCTGTTACAGGTCCTAATAC;
- (p7) 5'-TCCCCCGGGAGGGAGTGGAACAAAATTGA-3';

La amplificación se realizó después de una primera desnaturalización a 95°C (5 min) con 30 ciclos de desnaturalización (95°C , 45 s), formación de cebador (57°C , 45 s), síntesis de ADN (72°C , 1,5 min) y una última etapa de síntesis durante 4 min a 72°C .

La recombinación integrativa intermolecular de pEL13 se detectó como sigue. Se incubaron aproximadamente 400 ng de ADN genómico de poblaciones celulares supervivientes con los siguientes oligonucleótidos como cebadores para PCR:

- (attx1) 5'-AGTAGGAATTCAGTTGATTCATAGTGAAGTGC-3' y
- (B2) 5'-AACTGCAGTAAAAAGCATGCTCATCACCCC-3'.

La amplificación tuvo lugar después de una primera desnaturalización a 95°C (4 min) con 30 ciclos de desnaturalización (95°C , 45 s), formación de cebador (52°C , 45 s), síntesis de ADN (72°C , 45 s) y una última etapa de síntesis durante 4 min a 72°C .

La PCR de transcriptasa inversa (RT-PCR) se realizó con $4 \mu\text{g}$ de ARN aislado. En primer lugar se sintetizaron los ADNc según las indicaciones del fabricante usando cebadores oligo-dT (First Strand Synthesis Kit; Pharmacia). Una cuarta parte de estos ADNc se usó como molde para la posterior PCR usando los cebadores p3 y p4. Las condiciones de PCR fueron como se describen para p1-p7. Para probar la delección en lugar de inversión, el ADN genómico aislado se amplificó con los cebadores p5 y p6. Se analizaron transcritos de beta-actina partiendo de los ADNc usando los cebadores

- (AS) 5'-TAAAACGCAGCTCAGTAACAGTCCG-3' y
- (S) 5'-TGGAATCCTGTGGCATCCATGAAAC-3'

Las condiciones de PCR fueron como se describen para p1-p7.

Se realizaron análisis Southern esencialmente según el protocolo de Sambrook, J. (1989) Molecular Cloning (2ª edición) Cold Spring Harbor Laboratory Press. Se fragmentaron aproximadamente 10 g de ADN genómico con NcoI, se separó separaron mediante electroforesis en gel de agarosa (0,8% peso/volumen) en tampón TBE y se transfirió

durante la noche a una membrana de nailon. La sonda GFP para la detección de la recombinación se preparó mediante PCR usando los cebadores p2 y p3. El marcado radiactivo se realizó según las indicaciones del fabricante (Megaprime; Amersham) usando dATP y dCTP marcados con ^{32}P (Amersham).

5 4. Análisis Western

Se prepararon lisados celulares de células transfectadas transitoriamente mediante cocción (5 min) de las células en tampón de ensayo (New England Biolabs). Las proteínas se separaron en un gel de SDS-poliacrilamida al 12,5% según peso molecular y se transfirieron durante la noche a una membrana de nitrocelulosa (Immobilon P, Millipore).
 10 La membrana se trató con disolución de bloqueo al 1% (BM Chemiluminescence Western Blotting Kit; Boehringer Mannheim, Alemania) y se incubó con anticuerpo de ratón policlonal contra Int de tipo salvaje a una dilución de 1:50.000 (anticuerpos de A. Landy, EE.UU.). Se usaron los anticuerpos secundarios acoplados con peroxidasa para hacer visible la posición de la integrasa en el gel (BM Chemiluminescence Western Blotting Kit; Boehringer Mannheim, Alemania). Como control se usaron extractos celulares de *E. coli* que contenían Int de tipo salvaje.

15 5. Resultados

5.1. Síntesis de Int-h en células BL60

Para probar si la recombinación de Int-h puede catalizarse en células humanas fue necesario demostrar en primer lugar que la recombinasa puede sintetizarse de las células. Para tal fin se usó el vector de expresión eucariota, pKEXInt-h que contenía el gen Int-h bajo el control del promotor del CMV. Después de la introducción de pKEXInt-h en dos líneas celulares indicadoras BL60 distintas, B2 y B3, los transcritos completa y correctamente modificados específicos para el gen Int-h pudieron detectarse mediante análisis por RT-PCR. Se estudiaron lisados celulares 72 horas después de la electroporación con pKEXInt-h mediante un análisis Western. La detección de la recombinasa se realizó mediante anticuerpos policlonales de ratón que están dirigidos contra Int de tipo salvaje. Como control se incorporó pKEX en las células.

Los resultados mostraron que estaba presente una proteína con el peso molecular esperado en las células que se trataron con pKEXInt-h en la electroporación. Esta proteína no era detectable cuando se usó el vector de control pKEX.

5.2. Recombinación intramolecular integrativa catalizada por Int-h en células humanas

El análisis Western mostró que partiendo del vector pKEXInt-h se sintetiza la proteína Int-h de las dos líneas celulares indicadoras. Estas células contienen un vector sustrato, pGFPattB/attP, como ADN extraño integrado de forma estable en su genoma. Las dos secuencias de recombinación para la recombinación integrativa, attB y attP, se encuentran en orientación invertida una respecto a otra y flanquean el gen para GFP. El gen GFP se encuentra propiamente en orientación invertida respecto al promotor del CMV, que está aguas arriba de attB. La recombinación entre attB y attP mediante Int-h condujo a la inversión del gen GFP y por tanto a su expresión. Se construyeron en total tres líneas celulares indicadoras (B1-B3). El análisis Southern de su ADN genómico mostró que B1 y B3 han integrado varias copias de pGFPattB/attP como repeticiones directas en el genoma. La línea celular B2 contiene por el contrario sólo una copia. Las secuencias integradas se verificaron mediante PCR y secuenciación posterior.

Para probar la recombinación entre attB y attP se introdujeron pKEXInt-h y pKEX por separado en estas líneas celulares. Las células se recogieron 72 horas después de la electroporación, se aisló el ARN de una parte de las células y se investigó en cuanto a la expresión de GFP mediante RT-PCR usando el par de cebadores p3/p4. Estos cebadores amplificaron un fragmento de ADN de 0,99 kb de longitud sólo cuando el gen GFP se invirtió a consecuencia de la recombinación. Los resultados mostraron que el producto podía detectarse en las tres líneas celulares. No pudo detectarse ningún producto cuando se introdujo pKEX en las células. Los análisis de secuencia de ADN de los productos de PCR aislados confirman que se transcribió la región codificante del gen GFP y que pudo detectarse attR en el transcripto en lugar de attP. Como control para el contenido en ARN en las seis preparaciones celulares, así como para la síntesis satisfactoria de la primera hebra de ADN mediante la transcriptasa inversa, se analizó el transcripto de β -actina endógeno mediante PCR. Los resultados mostraron que el transcripto se presentaba en cantidades casi iguales.

La recombinación también se detectó mediante PCR directa de ADN genómico. Los resultados mostraron que los productos esperados sólo pudieron detectarse mediante los pares de cebadores p3/p4 (0,99 kb) y p1/p2 (0,92 kb) cuando pKEXInt-h se introdujo en las células. Los análisis de estos productos mediante secuenciación de ADN confirmaron que attR y attL estaban presentes en el genoma y que el gen GFP se invirtió mediante la recombinación. Estos experimentos se repitieron un total de tres veces, pudiendo detectarse la recombinación entre attB y attP mediante RT-PCR y/o PCR en las tres líneas celulares. Una detección de la delección del gen GFP mediante PCR con el par de cebadores p5/p6 resultó negativa. Sólo pudo amplificarse el fragmento de 1,3 kb esperado, que resultó del vector integrado.

La señal más fuerte que mostró la inversión entre attB y attP en la PCR se obtuvo repetida con la línea celular B3 en otro experimento. Se fragmentó el ADN genómico mediante NcoI y se investigó con un análisis Southern. Para ello se usó el gen GFP como sonda. Los resultados mostraron que el fragmento de restricción de ADN genómico esperado mediante la inversión entre attB y attP podía detectarse en la línea celular B3.

Para probar si el Int de tipo salvaje y el mutante Int-h/218 podían asimismo catalizar la recombinación integrativa intramolecular, los vectores pKEXInt-h, pKEXInt-h/218, pKEXInt y como control pKEX se incorporaron en la línea celular indicadora B3 mediante electroporación en un experimento adicional, como se describe en el ejemplo 2. Después de 72 horas, el ADN genómico se aisló y se probó en cuanto a recombinación mediante PCR con los pares de cebadores p5/p7 y p3/p4, como se describe en el ejemplo 3. Los resultados mostraron que ambos mutantes de Int podían catalizar la recombinación entre *attB* y *attP*, sin embargo la Int de tipo salvaje era inactiva.

5.3. No se detectaba la recombinación por escisión entre *attL* y *attR*

Debido a que Int-h también podía catalizar la recombinación por escisión entre *attL* y *attR* en ausencia de los cofactores IHF y Xis, se construyeron tres líneas celulares indicadoras BL60 que habían integrado el vector pGFP*attL/attR* de forma estable en el genoma. Estas líneas celulares contienen el gen GFP de nuevo en orientación invertida respecto al promotor del CMV, flanqueado no obstante por *attL* y *attR* en lugar de por *attB* y *attP*. Los análisis de recombinación se realizaron como se describe en el ejemplo 3 con pKEXInt-h como vector de expresión para la recombinasa, sin embargo mostraron que ni la inversión ni la delección podían detectarse entre *attL* y *attR* con RT-PCR o PCR.

5.4. Identificación y caracterización de una secuencia de nucleótidos similar a *attB* de origen natural en el genoma humano

Como se muestra en el ejemplo 3, ambos mutantes de recombinasa Int catalizaron la recombinación integrativa intramolecular en células humanas. Una de las dos secuencias de recombinación que participan en esta reacción, *attB*, tiene de 21 pares de bases (pb) de longitud y es constituyente natural del genoma de *E. coli*. Pudo mostrarse que Int-h tolera algunas desviaciones en la secuencia de la denominada región de reconocimiento “núcleo” de *attB* para una recombinación con *attP* (Nash (1981) Annu. Rev. Genet., 15, página 143). Por tanto, a partir de consideraciones estadísticas es posible la presencia de una secuencia funcional homóloga a *attB* en el genoma humano. En una búsqueda en bases de datos, los inventores pudieron identificar una secuencia aún incompleta como parte de una “etiqueta de secuencia expresada” (“expressed sequence tag”) (EST). Esta secuencia se aisló y clonó a continuación mediante PCR a partir de ADN humano. El análisis de secuencia de ADN completó la secuencia y un análisis Southern adicional de ADN genómico de células BL60 mostró que esta parte de la secuencia es de un gen humano aún desconocido que está presente en una copia en el genoma.

Esta secuencia que se designa en lo sucesivo *attH* se desvía de la secuencia *attB* de tipo salvaje en tres posiciones. Dos de los nucleótidos se encuentran en la región de reconocimiento de Int “núcleo” izquierda (B) y una es parte de la denominada región de “solapamiento”. Debido a que la identidad de la región de “solapamiento” entre dos secuencias de recombinación es una condición para una recombinación eficiente con Int-h, el nucleótido correspondiente en la posición 0 en el “solapamiento” de *attP* se cambió de timidina a guanina, lo que condujo a *attP**. *attH* y *attP** se usaron como secuencias invertidas en un vector (pACH) y se probaron en cuanto a recombinación en *E. coli*. Los resultados mostraron que Int-h e Int-h/218 catalizan la inversión entre *attH* y *attP** en presencia de IHF. Los análisis de secuencia de ADN de los productos de recombinación aislados confirmaron que la recombinación entre *attH* y *attP** transcurría según el mecanismo esperado. Por el contrario, la Int de tipo salvaje puede recombinar *attH/attP** también en presencia de IHF sólo de forma muy ineficiente. *attH* es por tanto una posible secuencia de integración para la integración catalizada por Int-h de ADN extraño que contiene una copia de *attP**.

5.5. Recombinación integrativa intermolecular entre *attH* y *attP** en células humanas

Para probar si *attH* como constituyente natural del genoma humano puede recombinarse con *attP** en una reacción intermolecular se construyó pEL13. Este vector contenía, además del gen Int-h con el control del promotor de CMV, una copia de *attP** y, como marcador de selección, el gen de resistencia a higromicina. Int-h pudo sintetizarse después de la introducción de pEL13 en células BL60 y catalizó la recombinación intermolecular entre *attP** como constituyente de pEL13 y *attH* genómico.

pEL13 se introdujo en células BL60 mediante electroporación como se describe en el ejemplo 2 y éstas se llevaron y diluyeron después de 72 horas a presión selectiva. Se estudiaron las poblaciones celulares supervivientes después de 6-8 semanas mediante PCR con el par de cebadores *attx1/B2* en cuanto a los acontecimientos de recombinación. Los resultados mostraron que en 13 de las 31 poblaciones celulares supervivientes podía detectarse una integración en *attH*. Los análisis de secuencia de ADN de los productos de PCR aislados de las distintas aproximaciones confirmaron su identidad como productos de la recombinación.

REIVINDICACIONES

1. Procedimiento para la recombinación específica de secuencia de ADN en una célula eucariota *in vitro* que comprende

a) la introducción en una célula de una primera secuencia de ADN que comprende:

una secuencia *attB* según SEQ ID N°:1 o su homólogo, o

una secuencia *attP* según SEQ ID N°:2 o su homólogo, o

una secuencia *attL* según SEQ ID N°:3 o su homólogo, o

una secuencia *attR* según SEQ ID N°:4 o su homólogo,

b) la introducción de una segunda secuencia de ADN en la célula que comprende:

una secuencia *attB* según SEQ ID N°:1 o su homólogo, o

una secuencia *attP* según SEQ ID N°:2 o su homólogo, o

una secuencia *attL* según SEQ ID N°:3 o su homólogo, o

una secuencia *attR* según SEQ ID N°:4 o su homólogo, y

c) la realización de la recombinación específica de secuencia mediante la acción de una integrasa Int del bacteriófago *lambda*,

designando el término “homólogo” secuencias *attB*, *attP*, *attL* y *attR* que son idénticas al menos en el 70% en comparación con las secuencias de recombinación de origen natural.

2. Procedimiento según la reivindicación 1, designando el término “homólogo” secuencias *attB*, *attP*, *attL* y *attR* que son idénticas al menos en el 80%, al menos en el 85%, al menos en el 90%, al menos en el 95% o al menos en el 99% en comparación con las secuencias de recombinación de origen natural.

3. Procedimiento para la recombinación específica de secuencia de ADN en una célula eucariota *in vitro* que contiene en su genoma la primera secuencia de ADN, bien presentándose de forma natural en la célula o introduciéndose previamente con ayuda de recombinación de ADN, que comprende las etapas b) y c) como se definieron en la reivindicación 1 ó 2.

4. Procedimiento según la reivindicación 1 a 3, en el que la primera secuencia de ADN comprende una secuencia *attB* según SEQ ID N°:1 o su homólogo, y la segunda secuencia de ADN comprende una secuencia *attP* según SEQ ID N°:2 o su homólogo, definiéndose el término “homólogo” como en la reivindicación 1.

5. Procedimiento según la reivindicación 1 a 3, en el que la primera secuencia de ADN comprende una secuencia *attL* según SEQ ID N°:3 o su homólogo, y la segunda secuencia de ADN comprende una secuencia *attR* según SEQ ID N°:4 o su homólogo, estando presente en la etapa c) adicionalmente un factor Xis, definiéndose el término “homólogo” como en la reivindicación 1.

6. Procedimiento según una de las reivindicaciones 1 a 5, en el que adicionalmente se introduce en la célula una tercera o una tercera y cuarta secuencia de ADN que comprenden un gen de Int o un gen de Int y un gen del factor Xis.

7. Procedimiento según la reivindicación 6, en el que la tercera o la tercera y/o cuarta secuencia de ADN comprenden además una secuencia de ADN reguladora que provoca una expresión espacial y/o temporal del gen de Int y/o del gen del factor Xis.

8. Procedimiento según una de las reivindicaciones 1 a 7, en el que Int es una integrasa modificada.

9. Procedimiento según la reivindicación 8, en el que la Int modificada es Int-h o Int-h/218.

10. Procedimiento según una de las reivindicaciones 1 a 9, en el que en la etapa c) participa adicionalmente un “factor de integración del huésped” (IHF).

11. Procedimiento según una de las reivindicaciones 1 a 10, en el que la primera y/o segunda secuencia de ADN comprende además secuencias de ADN que provocan una integración de la primera y/o segunda secuencia de ADN en el genoma de células eucariotas mediante recombinación homóloga.

ES 2 334 577 T3

12. Procedimiento según una de las reivindicaciones 1 a 11, en el que la primera y/o segunda secuencia de ADN comprende además una secuencia de ácido nucleico que codifica un polipéptido de interés.

5 13. Procedimiento según la reivindicación 12, en el que el polipéptido de interés es una proteína estructural, una enzima endógena o exógena, una proteína reguladora o una proteína marcadora.

14. Procedimiento según una de las reivindicaciones 1, 2 y 4 a 13, en el que la primera y segunda secuencia de ADN se introducen en la célula eucariota en la misma molécula de ADN.

10 15. Procedimiento según una de las reivindicaciones 1 a 14, en el que la célula eucariota es una célula de mamífero.

16. Procedimiento según la reivindicación 15, en el que la célula de mamífero es una célula humana, una célula de mono, ratón, rata, conejo, hámster, cabra, vaca, oveja o cerdo.

15 17. Procedimiento según una de las reivindicaciones 1 a 4 y 6 a 16, que comprende además

20 d) la realización de una segunda recombinación específica de secuencia de ADN después de una primera recombinación específica de secuencia según las etapas a) a c), en el que la primera secuencia de ADN comprende *attB* según SEQ ID N°:1 o su homólogo y la segunda secuencia de ADN *attP* según SEQ ID N°:2 o su homólogo o a la inversa, mediante la acción de una *Int* y de un factor *Xis*.

25 18. Procedimiento según la reivindicación 17, en el que en las células se introduce además una secuencia de ADN adicional que comprende un gen del factor *Xis*.

19. Procedimiento según la reivindicación 18, en el que la secuencia de ADN adicional comprende además una secuencia de ADN reguladora que provoca una expresión espacial y/o temporal del gen del factor *Xis*.

30 20. Uso de una secuencia *attB* según SEQ ID N°:1 o su homólogo y una secuencia *aP* según SEQ ID N°:2 o su homólogo, o una secuencia *attL* según SEQ ID N°:3 o su homólogo y una secuencia *attR* según SEQ ID N°:4 o su homólogo en una recombinación específica de secuencia de ADN en células eucariotas *in vitro*, designando el término “homólogo” secuencias *attB*, *attP*, *attL* y *attR* que son idénticas al menos en el 70% en comparación con las secuencias de recombinación de origen natural.

35 21. Uso según la reivindicación 20, designando el término “homólogo” secuencias *attB*, *attP*, *attL* y *attR* que son idénticas al menos en el 80%, al menos en el 85%, al menos en el 90%, al menos en el 95% o al menos en el 99% en comparación con las secuencias de recombinación de origen natural.

40 22. Célula eucariota que contiene secuencias de *attB*, *attP*, *attL* y *attR* u homólogos de las mismas que puede obtenerse sometiendo la célula eucariota mencionada en la reivindicación 1 y 3 al procedimiento según una de las reivindicaciones 1 a 19.

23. Organismo transgénico no humano que contiene al menos una célula según la reivindicación 22.

45 24. Organismo según la reivindicación 23, en el que este organismo es un ratón, una rata, un conejo o un hámster.

50

55

60

65

Fig. 1

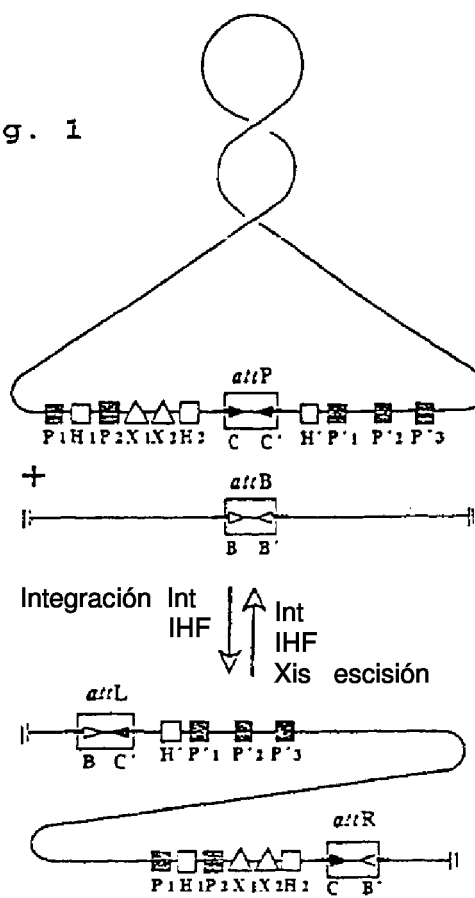
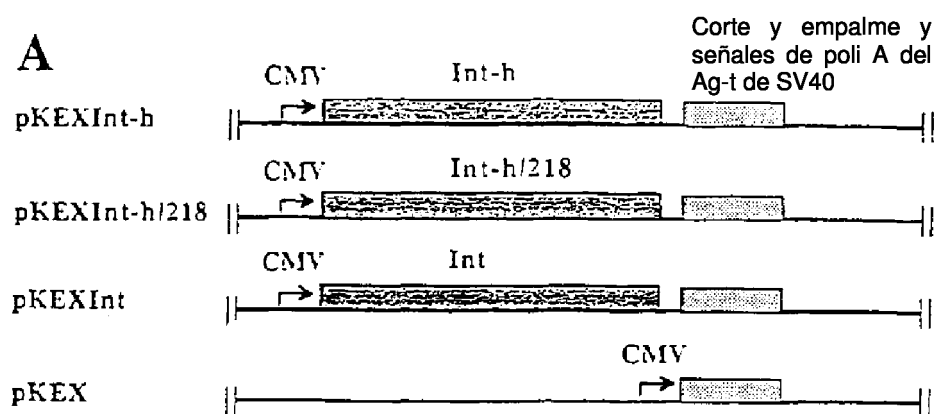


Fig. 2



B

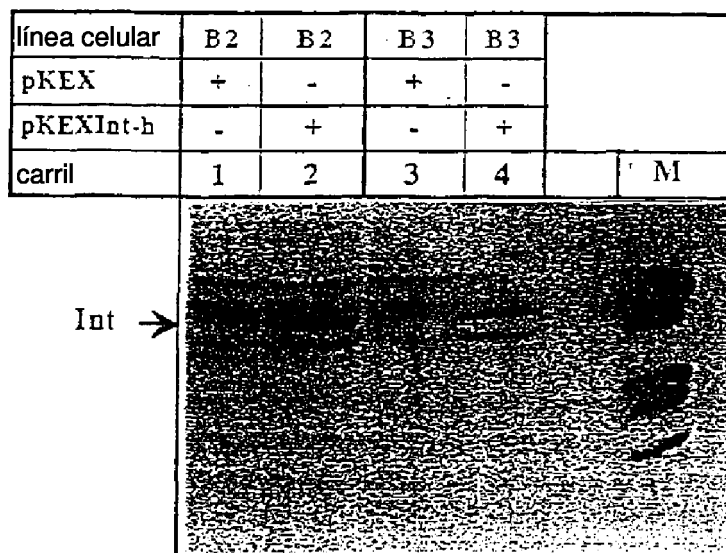


Fig. 3

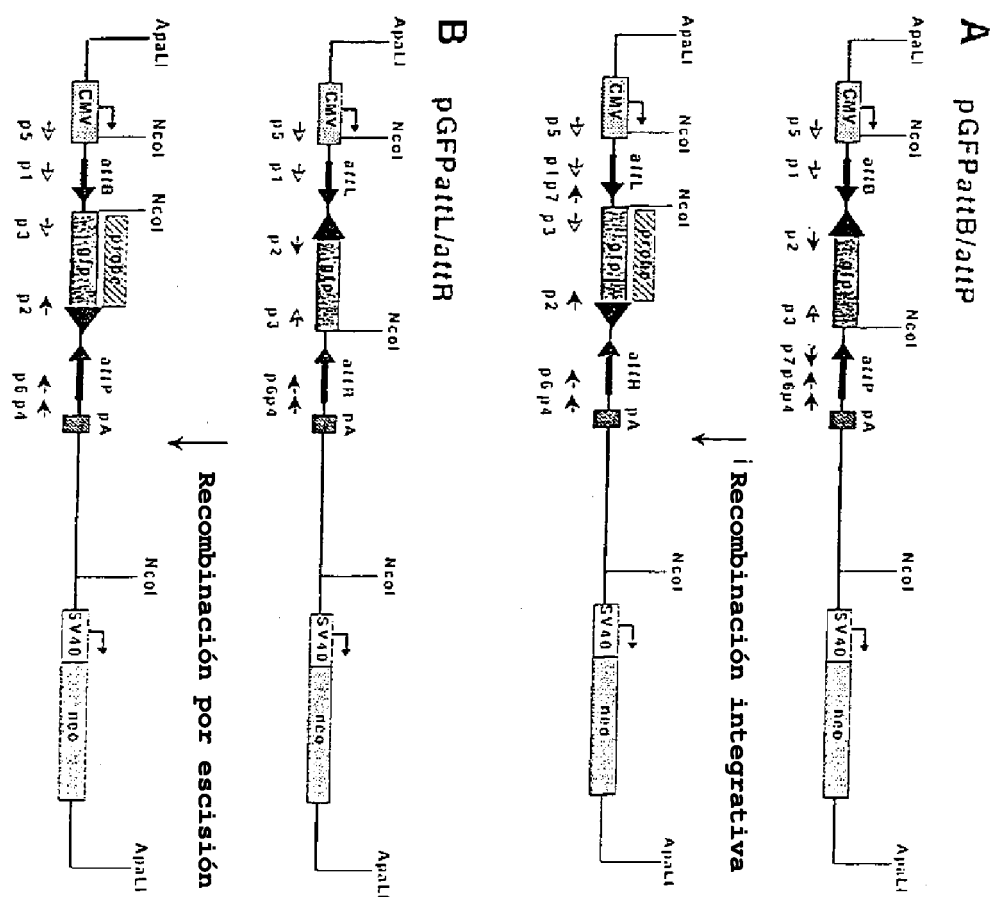


Fig. 4

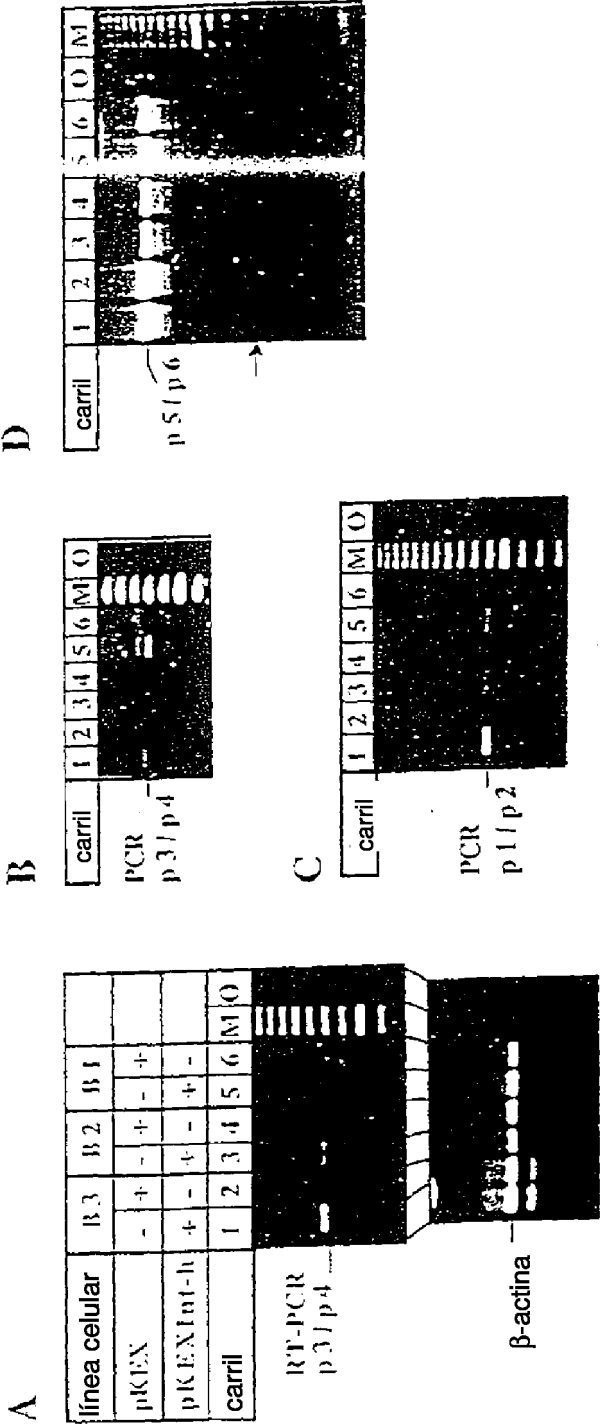
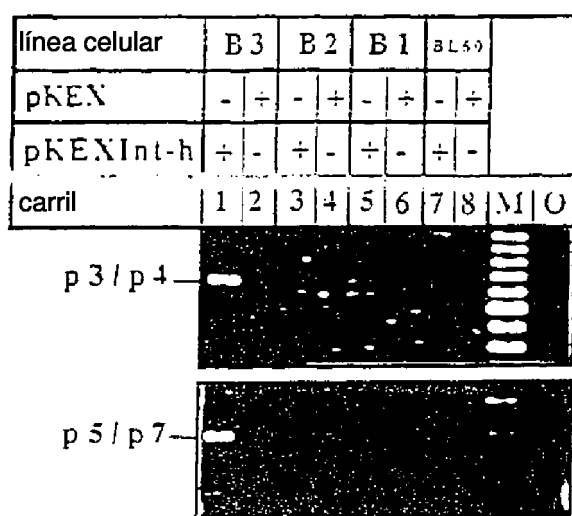


Fig. 5

A



B

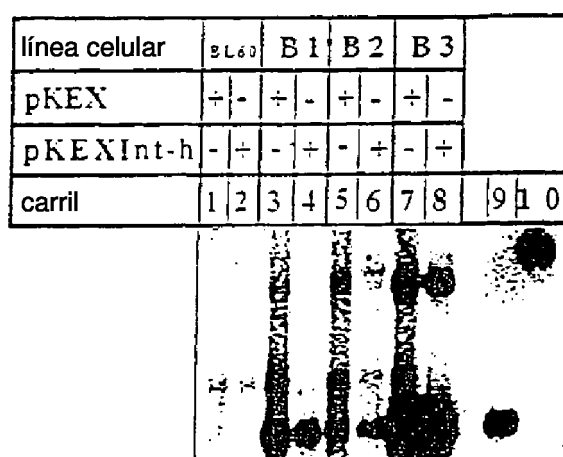


Fig. 6

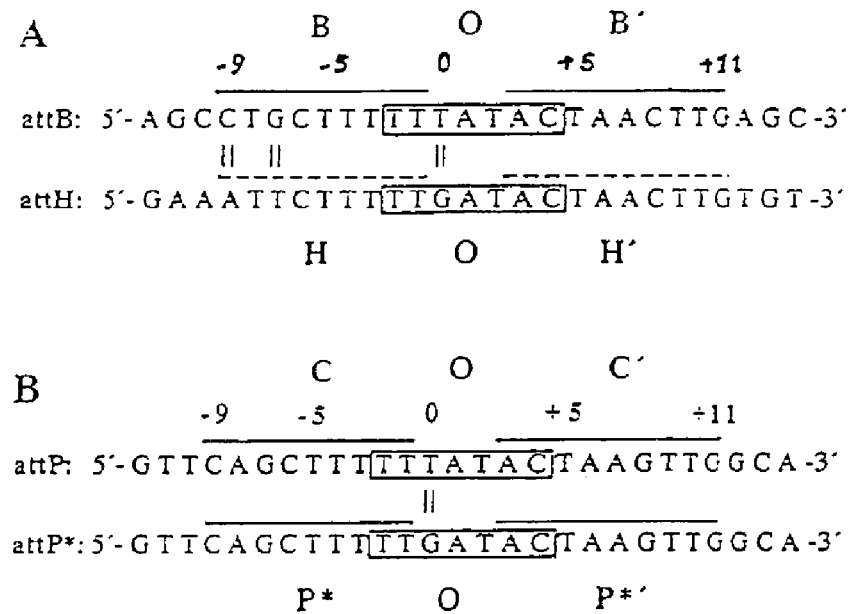


Fig. 7

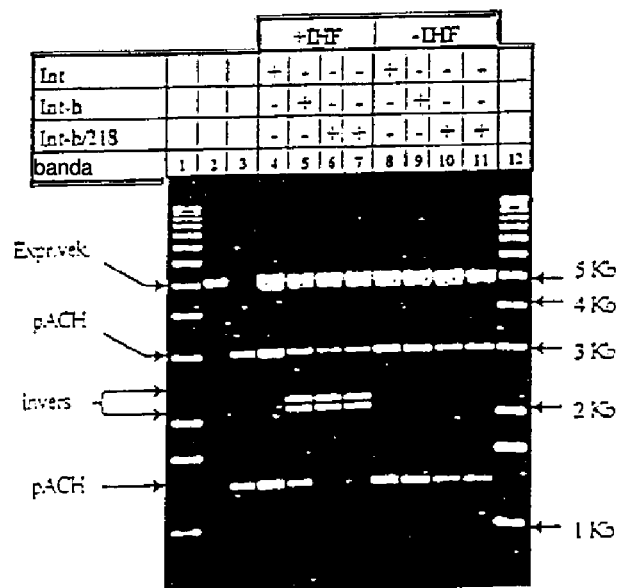


Fig. 8

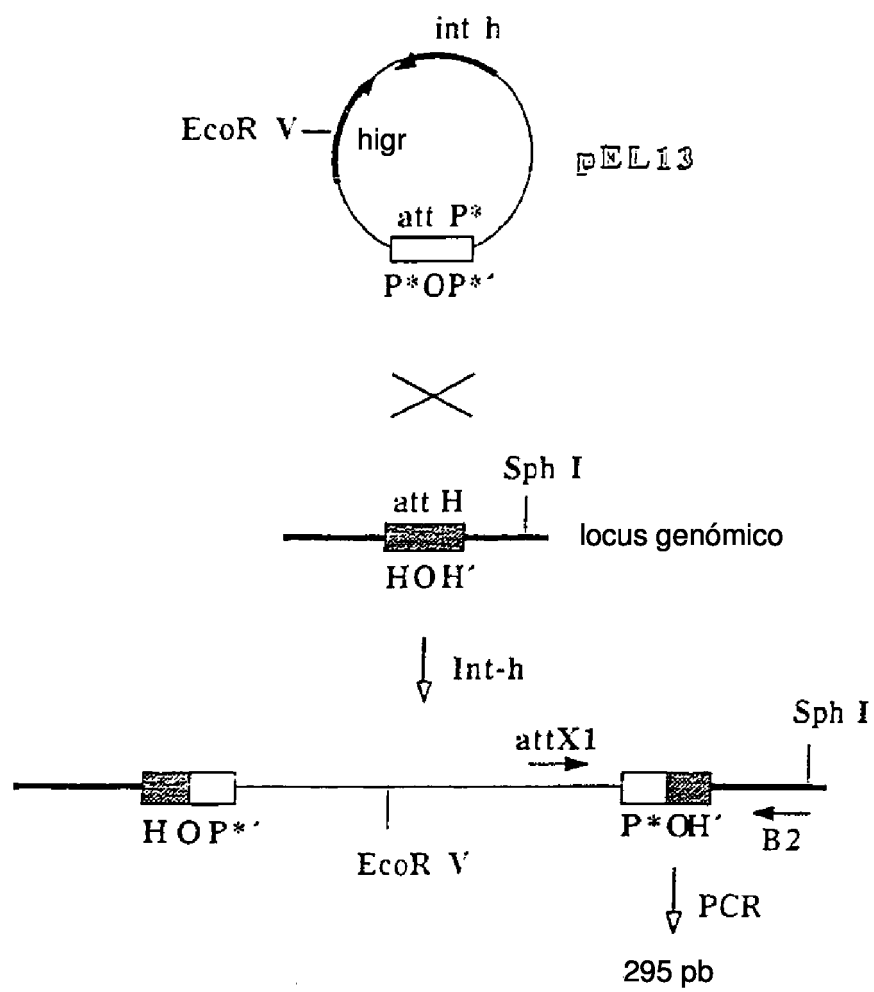
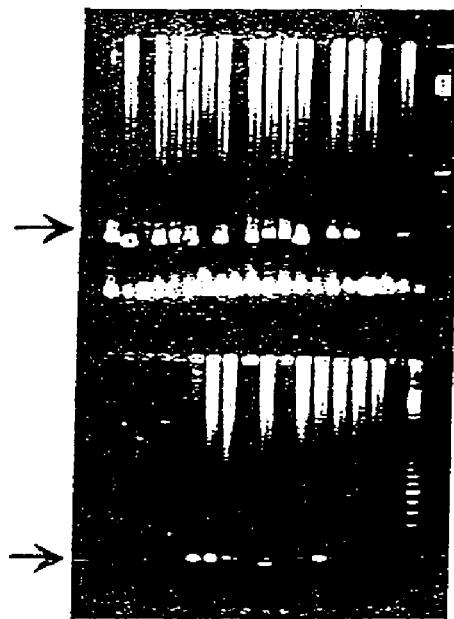


Fig. 9



ES 2 334 577 T3

LISTA DE SECUENCIAS

- <110> Dröge Dr., Peter
5 <120> Recombinación de ADN específico de secuencia en células eucariotas
<130> DRO-001 PCT/EP T1
<140> desconocido
<141> 2000-08-29
10 <160> 5
<170>
<210> 1
<211> 21
15 <212> ADN
<213> *Escherichia coli*

20 <400> 1

ctgctttttt atactaactt g

21

- <210> 2
25 <211> 243
<212> DNA
<213> Bacteriófago lambda

30 <400> 2

35 tctgttacag gtcactaata ccatctaagt agttgattca tagtgactgc atatgttggtg 60
ttttacagta ttatgtagtc tggttttttat gcaaaaatcta atttaatatata ttgatatttta 120
tatcatttta cgtttctcgt tcagcttttt tatactaagt tggcattata aaaaagcatt 180
gcttatcaat ttgttgcaac gaacagggtca ctatcagtca aaataaaaac attattttgat 240
ttc 243

- 40 <210> 3
<211> 102
<212> ADN
45 <213> *Escherichia coli*

<400> 3

50 ctgctttttt atactaagtt ggcattataa aaaagcattg cttatcaatt tggttgcaacg 60
aacagggtcac tatcagtcac aataaaaatca ttatttgatt tc 102

- 55 <210> 4
<211> 162
<212> ADN
<213> *Escherichia coli*

60 <400> 4

65 tctgttacag gtcactaata ccatctaagt agttgattca tagtgactgc atatgttggtg 60
ttttacagta ttatgtagtc tggttttttat gcaaaaatcta atttaatatata ttgatatttta 120
tatcatttta cgtttctcgt tcagcttttt tatactaact tg 162

ES 2 334 577 T3

<210> 5

<211> 243

<212> DNA

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: oligonucleótido

10 <400> 5

```

      tctgttacag gtcactaata ccatctaagt agttgattca tagtgactgc atatgttgtg 60
      ttttacagta ttatgtagtc tgttttttat gcaaaatcta atttaataata ttgatattta 120
15   tatcatttta cgtttctcgt tcagcttttt gatactaagt tggcattata aaaaagcatt 180
      gcttatcaat ttgttgcaac gaacaggtca ctatcagtea aaataaaatc attatttgat 240
      ttc                                     243
```

20

25

30

35

40

45

50

55

60

65