



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 358 994**

51 Int. Cl.:

**C12N 15/00** (2006.01)

**C12N 15/09** (2006.01)

**C12N 15/63** (2006.01)

**C12N 15/70** (2006.01)

**C12N 15/74** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03748973 .9**

96 Fecha de presentación : **24.07.2003**

97 Número de publicación de la solicitud: **1546322**

97 Fecha de publicación de la solicitud: **29.06.2005**

54

Título: **Vectores basados en transposones y métodos para la integración de ácidos nucleicos.**

30

Prioridad: **24.07.2002 US 398628 P**

45

Fecha de publicación de la mención BOPI:  
**17.05.2011**

45

Fecha de la publicación del folleto de la patente:  
**17.05.2011**

73

Titular/es: **MANOA BIOSCIENCES Inc.**  
**1717 Mott-Smith Drive 3213**  
**Honolulu, Hawai 96822, US**

72

Inventor/es: **Kaminski, Joseph, M.**

74

Agente: **Elzaburu Márquez, Alberto**

ES 2 358 994 T3

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

## DESCRIPCIÓN

Vectores basados en transposones y métodos para la integración de ácidos nucleicos

## I. Antecedentes de la invención

5 Las investigaciones han revelado tres componentes principales para el transporte eficaz de vectores virales y no virales a través de la membrana citoplásmica y en el núcleo de las células eucariotas. Éstos incluyen un ligando específico para la endocitosis mediada por receptor, un factor de disrupción endosomal y una señal de localización nuclear. Estos componentes se han empleado con éxito en vectores no virales (1-6). En vectores que carecen de o que no pueden interactuar con una señal de localización nuclear, la transfección eficaz sólo ocurrirá en aquellas células que se dividen activamente. Los tres requerimientos de ADN para la integración son (1) la secuencia de ADN, (2) una estructura de ADN local del anfitrión, y (3) las proteínas de unión al ADN endógenas asociadas [45]. Para que ocurra la integración, se requiere una enzima (p. ej., transposasa) para mediar el proceso. Esta enzima puede ser una transposasa o una recombinasa específica de sitio. Las recombinasas específicas de sitio permiten la recombinación y algunas no requieren cofactores permitiendo de esta manera una actividad fuera de su entorno normal. Por ejemplo, la recombinasa Cre, aunque se obtiene del fago P1 de *Escherichia coli*, actúa eficazmente en células de plantas, levaduras y de mamíferos (18). Las recombinasas selectivas de sitio tales como FLP, Cre y  $\beta$ -recombinasa llevan a cabo tanto la integración como la escisión eficazmente con los mismos sitios diana; sin embargo, la frecuencia neta de integración es baja (p. ej., 0,03% para Cre) (18-20).

20 Las limitaciones de los vectores virales tales como patogenicidad, gasto de la producción e inestabilidad sistémica han resultado obstáculos importantes en el uso de sistemas basados en virus. De hecho, la readministración de vectores basados en virus puede estimular respuestas inmunes que pueden resultar en efectos sistémicos potencialmente mortales y limitar la eficacia de la transferencia génica (64-65). Los vectores no virales (es decir, basados en lípidos, basados en polímeros, basados en lípidos-polímeros y polilisina) son medios sintéticos para encapsular el ADN transgénico hasta que alcanza la diana celular. Comparados con los vectores virales, los vectores no virales son más seguros de preparar; se disminuye el riesgo de complicaciones patológicas e inmunológicas. Los vectores no virales se han diseñado modificando la superficie del vector no viral para la terapia dirigida (7-12). Los liposomas típicamente se internalizan en los endosomas, que frecuentemente son dirigidos a los lisosomas, degradando así el plásmido. En estos vectores se han empleado factores de disrupción endosomal y las señales de localización nuclear. Sin embargo, los lipoplexos (ADN plasmídico y liposoma) están limitados principalmente a la transfección de células en división a no ser que esté presente o que interactúe con el vector un factor de localización nuclear (16). Además, la integración eficaz en el anfitrión no ocurre excepto en los plásmidos basados en transposones (17-20). No obstante, se ha demostrado que los liposomas son seguros en los ensayos de terapia génica en seres humanos (21-24).

35 Los transposones son móviles, respecto a que pueden moverse de una posición en el ADN a una segunda posición en el ADN en presencia de una transposasa. Existen dos componentes fundamentales en cualquier sistema de transposón móvil de tipo corte y pega, una fuente de una transposasa activa y las secuencias de ADN que son reconocidas y movilizadas por la transposasa. La movilización de las secuencias de ADN permite que también se movilice el ácido nucleico implicado entre las secuencias de ADN reconocidas.

La integrasa y retrotransposasa dependen de su propio dominio de unión al ADN o una interacción con un factor de direccionamiento de ADN del anfitrión para dirigir el complejo ADN-enzima (p. ej., transposón/transposasa) en yuxtaposición con el ADN del anfitrión para que ocurra la integración (25, 35-37).

40 Si el anfitrión no tiene este factor de direccionamiento o una secuencia de ADN específica del anfitrión reconocida por el complejo transposón/transposasa, la eficacia de la integración disminuye sustancialmente (25, 38). Por ejemplo, se ha mostrado que una proteína endógena humana específica, integrasa que interacciona 1, se une a la integrasa y estimula la integración in vitro y posiblemente in vivo mediante la unión y direccionamiento de la integrasa a los sitios hipersensibles de la ADNasa 1 (25). Alternativamente, el elemento semejante a retrovirus de levaduras Ty3 se inserta en los sitios de comienzo de la transcripción de genes transcritos por la ARN polimerasa III mediante su unión con este complejo (37). Además, algunas transposasas o integrasas requieren determinados sitios en el ADN del anfitrión para su actividad catalítica incluso si el complejo ADN-enzima se pone próximo al ADN del anfitrión. Por ejemplo, el transposón Tc1/mariner se integra en un dinucleótido TA (32).

50 Los elementos de ADN transponibles para manipulación genética han estado disponibles desde hace más de 15 años. Esta tecnología se ha aplicado tanto en bacterias como eucariotas para verificar si un fragmento de ADN clonado contiene o no el gen funcional completo de interés. Rubin y Spradling fueron los primeros en demostrar esto para los elementos P de *D. melanogaster*. Un fragmento de ADN que porta el gen *rosy* se insertó en las repeticiones terminales de un elemento P y se clonó en un plásmido. Este plásmido y otro que codificaba la transposasa se inyectaron en los embriones de una cepa M con una delección en el gen *rosy*. Aproximadamente el 50% de las moscas obtenidas de los

embriones inyectados poseían el fenotipo rosy, sugiriendo de esta manera que el gen rosy se había insertado en el cromosoma y había mantenido su función en distintos sitios en el genoma. Además, ninguno de los ADN plasmídicos que flanquean se había integrado en el genoma del anfitrión lo que sugiere que la escisión del plásmido sólo había tenido lugar en las repeticiones terminales (39).

5 Para que un ácido nucleico codificado en un vector se incorpore en el ADN diana debe producirse la integración. El modelo putativo de integración es similar en los retrovirus, transposones y retrotransposones semejantes a retrovirus. Por ejemplo, el dominio catalítico está conservado en las integrasas y las transposasas. Las reacciones in vitro han mostrado que la integrasa o transposasa son las únicas enzimas necesarias para la integración (25-28). Se ha mostrado  
10 que la integrasa y muchas transposasas en bacterias y eucariotas se unen específicamente al sitio att en los extremos de las repeticiones terminales. Requieren la presencia de CA en el extremo 3' tanto para el procesamiento como la escisión/ligación (29-30).

Los transposones tienen muchas aplicaciones en la manipulación genética de un genoma anfitrión, incluyendo la administración transgénica y la mutagénesis insercional. Sin embargo, la eficacia de la integración del transposón puede variar sustancialmente entre las líneas celulares, lo que sugiere la implicación de factores del anfitrión. Tomando como  
15 base los requerimientos para la integración de los elementos transponibles, parece que es necesario un factor de direccionamiento de ADN del anfitrión para una integración eficaz por yuxtaposición del complejo transposón-transposasa adyacente al ADN del anfitrión. El requerimiento de un factor de direccionamiento de ADN del anfitrión se ha establecido en retrovirus y retrotransposones semejantes a retrovirus. Por ejemplo, el elemento semejante a retrovirus de levadura Ty3 se inserta en los sitios de comienzo de la transcripción de genes transcritos por la ARN  
20 polimerasa III mediante su interacción con este complejo [82]. Alternativamente, la integrasa del virus de la inmunodeficiencia humana se une a la proteína integrasa endógena humana que interacciona 1 para estimular la integración in vitro y posiblemente in vivo [83, 25]. De hecho, las transposasas Tc1/mariner también tienen sitios de unión al ADN. Sin embargo, estos dominios de unión al ADN aparentemente no son selectivos de sitio (35), posiblemente carecen de sitios de reconocimiento fuerte en determinados genomas anfitriones, y pueden requerir otras  
25 proteínas del anfitrión para una integración eficaz por acoplamiento de transposón-transposasa con el ADN del anfitrión.

En muchos casos, el anfitrión no tiene el factor de acoplamiento requerido tal como una secuencia de ADN reconocida por la transposasa o un factor endógeno que yuxtapone el complejo transposón-transposasa al ADN del anfitrión. Así, la eficacia de la integración en estos anfitriones se reducirá de manera importante. Además, incluso si el complejo transposón-transposasa se acopla con el ADN del anfitrión, la integración puede no ocurrir porque el sitio del ADN tiene  
30 que ser permisivo. La presente invención supera los problemas asociados con los sistemas de vector no virales actualmente conocidos.

WO 2002/08286 describe proteínas de unión al ADN que comprenden dominios de dedo de cinc. Se requieren tres construcciones de ácido nucleico independientes para ensamblar una proteína con dedo de cinc con los dominios de  
35 dedo de cinc respectivos. WO 2001/30965 se refiere a métodos y composiciones para introducir un ácido nucleico en el genoma de al menos una célula, en el que se usa un transposón no quimérico Bella Durmiente que no comprende un dominio de unión al ADN.

## II. Resumen de la invención

De acuerdo con los propósitos de esta invención, como se plasma y describe ampliamente en la presente memoria, esta invención, en un aspecto, se refiere a composiciones que comprenden una construcción de ácido nucleico para la  
40 integración selectiva de sitio en el genoma de un sujeto y a métodos para su uso.

Las ventajas adicionales de la invención se mostrarán en parte en la descripción siguiente y en parte serán obvias a partir de la descripción o pueden aprenderse mediante la práctica de la invención. Las ventajas de la invención se llevarán a cabo y se lograrán mediante los elementos y combinaciones destacadas particularmente en las reivindicaciones adjuntas. Debe entenderse que tanto la descripción general precedente como la descripción detallada  
45 siguiente son sólo como ejemplo y explicación y no son restrictivas de la invención, según se reivindica.

## III. Descripción breve de los dibujos

Los dibujos adjuntos, que se incorporan en y constituyen una parte de esta especificación, ilustran varias realizaciones de la invención y conjuntamente con la descripción, sirven para explicar los principios de la invención. Los ejemplos específicos de la invención se observan en los Ejemplos 1-9.

50 La Figura 1 muestra un plásmido que codifica una transposasa selectiva de sitio. El diagrama proporciona una explicación general de un aspecto de la invención. IR=Repeticiones Invertidas; E,P=Amplificador, Promotor.

- La Figura 2 muestra dos plásmidos, uno que codifica el transgén y el otro la transposasa (o transposasa quimérica-factor de direccionamiento del anfitrión). El primer plásmido tiene un transgén flanqueado por repeticiones terminales invertidas (o semejantes) y también contiene una secuencia similar al anfitrión. En esta representación, la secuencia similar actúa principalmente para acoplar este plásmido al ADN complementario del anfitrión. El plásmido de la transposasa podría estar bajo un promotor inducible que permitiría la regulación temporal de la transposasa. La escisión del ADN fuera de las repeticiones terminales y la integración ocurrirían como se muestra en el diagrama de la Figura 1. IR=Repeticiones Invertidas; E,P=Amplificador, Promotor.
- La Figura 3 muestra dos plásmidos, el basado en el transposón [que codifica el transgén y contiene un sitio de unión a proteína (PBS)] y el otro para un polipéptido de fusión que contiene dos dominios de unión al ADN (o un dominio de unión al ADN y uno de unión a proteína). El polipéptido de fusión se unirá al PBS y dirigirá el plásmido con transposón al sitio de ADN del anfitrión reconocido por el dominio de unión al ADN de la proteína de fusión (o una proteína que está asociada con el ADN del anfitrión). La transposasa podría estar bajo un promotor inducible que permitiría la regulación temporal de la transposasa. La escisión del ADN fuera de las repeticiones terminales y la integración ocurrirían como se muestra en el diagrama de la Figura 1. IR=Repeticiones Invertidas; E,P=Amplificador, Promotor; PBS=Sitio de Unión a la Proteína.
- La Figura 4 muestra la construcción de ácido nucleico no viral contenida en una cubierta catiónica no específica con la transposasa unida a las repeticiones terminales invertidas que están flanqueando el transgén. Signos menos=carga negativa del ADN; Signos positivos=Carga positiva del empaquetamiento no viral (p. ej., bicapa lipídica compuesta por DOPE, un lípido catiónico y polietilén glicol-ceramida).
- La Figura 5 muestra la construcción de ácido nucleico no viral con un sistema de empaquetamiento que incluye un ligando específico, factor de disrupción endosomal y señal de localización nuclear. Signos menos=carga negativa del ADN; Signos positivos=Carga positiva del empaquetamiento no viral (p. ej., bicapa lipídica compuesta por DOPE, un lípido catiónico y polietilén glicol-ceramida); Círculos con tres vías=Ligando para el direccionamiento celular selectivo; Líneas serpenteantes= Factor de disrupción endosomal; Círculos sólidos=Señal de localización nuclear.
- La Figura 6 muestra la construcción de ácido nucleico no viral con dicho sistema de empaquetamiento de la figura 5; sin embargo, la transposasa contiene un factor de direccionamiento a ADN del anfitrión. Signos menos=carga negativa del ADN; Signos positivos=Carga positiva del empaquetamiento no viral (p. ej., bicapa lipídica compuesta por DOPE, un lípido catiónico y polietilén glicol-ceramida); Círculos con tres vías=Ligando para el direccionamiento celular selectivo; Líneas serpenteantes= Factor de disrupción endosomal; Círculos sólidos=Señal de localización nuclear; Triángulo en la transposasa=Dominio de unión al ADN.
- La Figura 7 muestra construcciones genéticas para evaluar la integración dirigida de maT en células de insecto. Se introducen cuatro plásmidos diferentes en las células de insecto, cada uno portando 1) un transposón maT modificado, con ITR intactas (flechas sólidas negras) y una ORF de transposasa interrumpida que contiene un gen marcador seleccionable y promotor inducible (triángulo) y 2) una transposasa quimérica con un dominio de unión al ADN LexA o Ga14 fusionado con cualquiera de los extremos 5' ó 3' de la secuencia de transposasa. Un plásmido diana adicional (no mostrado), que porta sitios diana LexA o Gal4 se coadministrará a las células o embriones y después de la inducción de la transposasa, los plásmidos diana recombinantes que portan el transposón modificado se examinarán para integraciones dirigidas.
- La Figura 8 muestra un esquema de la construcción de un transposón Mos1 quimérico con las secuencias de reconocimiento para los dominios de unión al ADN Ga14 y LexA y una señal de localización nuclear.
- La Figura 9 muestra la organización del dominio de gpNu1.
- La Figura 10 muestra la construcción quimérica de los dominios de unión al ADN de gpNu1 y el dominio catalítico de la integrasa.
- La Figura 11 muestra análisis en gel de agarosa de los fragmentos de ADN purificados. Carril M, marcadores de peso molecular. Carril 1, pKKT7(-H) linealizado con *EcoRI* y *HindIII*. Carril 2, producto de PCR de la secuencia de gpNulΔE85 digerido con *EcoRI* y *HpaI*. Carril 3, producto de PCR de la secuencia de gpNulΔP141 digerido con *EcoRI* y *HpaI*. Carril 4, producto de PCR de la secuencia de IntC170 digerido con *NotI* y *HindIII*. Obsérvese que la fluorescencia de este producto de PCR está apantallada por el colorante de carga en el gel.
- La Figura 12 muestra la estrategia de clonación para la construcción de pNu1ΔE185-IntC170 y gpNu1ΔP141-IntC170.
- La Figura 13 muestra una representación esquemática de la proteína de fusión MBP-Zif-Cre.

La Figura 14 muestra la sobreexpresión y purificación de la proteína de fusión MBP-Zif-Cre. Se indica la posición de la proteína de fusión expresada.

5 La Figura 15 muestra un análisis de la actividad biológica de las proteínas MBP-Zif-Cre y MBP-Cre. Se mezclaron 5 µg de ADN plasmídico con proteína de fusión purificada y se incubó a 37°C durante 15 min. Las reacciones se purificaron posteriormente como se describe y se digirieron alícuotas con la enzima de restricción AflIII.

La Figura 16 muestra una representación esquemática de la proteína MBP-Zif-Cre. La probabilidad de superficie está indicada así como la estructura de los sitios diana del ADN, que puede usarse para evaluar la actividad de recombinación de la proteína de fusión MBP-Zif-Cre.

10 La Figura 17 muestra una representación esquemática de la construcción belleza durmiente y una transposasa quimérica que usa belleza durmiente.

La Figura 18 muestra la introducción de un sitio de restricción único en el extremo N terminal de Belleza Durmiente.

La Figura 19 muestra la introducción de un dominio de unión al ADN (bien de zif268 o similar), junto con un conector peptídico flexible, en ese sitio de restricción.

#### IV. Descripción detallada

15 La presente invención puede entenderse más fácilmente por referencia a la descripción detallada siguiente de las realizaciones preferidas de la invención y los Ejemplos incluidos en ellas y a las Figuras y su descripción previa y siguiente.

20 Antes de revelar y describir los presentes compuestos, composiciones, artículos, dispositivos y/o métodos, debe entenderse que esta invención no está limitada a los métodos sintéticos específicos, métodos de biotecnología recombinantes específicos a no ser que se especifique otra cosa, o a reactivos particulares a no ser que se especifique otra cosa, ya que éstos pueden, por supuesto variar. También debe entenderse que la terminología usada en la presente memoria es sólo para el propósito de describir realizaciones particulares y no se pretende que sea limitante.

25 A lo largo de esta solicitud, se hace referencia a varias proteínas y ácidos nucleicos. Debe entenderse que cualquiera de los nombres usados para proteínas o ácidos nucleicos son nombres reconocidos en la técnica, de manera que la referencia al nombre constituye una descripción de la molécula en sí misma.

#### A. Definiciones

Tal y como se usa en la especificación y las reivindicaciones adjuntas, las formas singulares "un", "una" y "el o la" incluyen los referentes en plural a no ser que el contexto dicte claramente otra cosa. Así, por ejemplo, la referencia a "un vehículo farmacéutico" incluye mezclas de dos o más de dichos vehículos y semejantes.

30 Los intervalos pueden expresarse en la presente memoria como desde "aproximadamente" un valor particular y/o hasta "aproximadamente" otro valor particular. Cuando se expresa dicho intervalo, otra realización incluye desde el valor particular y/o hasta el otro valor particular. De manera similar, cuando los valores se expresan como aproximaciones, por el uso del "aproximadamente" anterior, se entenderá que el valor particular forma otra realización. Además se entenderá que los puntos finales de cada uno de los intervalos son ambos significativos en relación con el otro punto final e independientemente del otro punto final.

35 En esta especificación y en las reivindicaciones siguientes, se hará referencia a varios términos que se definirán para tener los significados siguientes:

"Opcional" u "opcionalmente" significa que el evento o circunstancia descrito posteriormente puede o no ocurrir, y que la descripción incluye casos en los que dicho evento o circunstancia ocurre y casos en los que no ocurre.

40 Por "tratar" se quiere decir que se observa y/o detecta una mejora en el estado patológico, es decir, trastorno genético, enfermedad autoinmune, cáncer, infección viral, infección bacteriana o infección parasítica después de la administración de una sustancia de la presente invención a un sujeto. El tratamiento puede variar desde un cambio positivo en un síntoma o síntomas de la enfermedad hasta la mejora completa del trastorno genético, enfermedad autoinmune, cáncer, infección viral, infección bacteriana o infección parasítica (p. ej., reducción de la gravedad o intensidad de la enfermedad, alteración de los parámetros clínicos indicativos de la condición del sujeto, alivio de las molestias o función incrementada o amplificada), como se detecta por técnicas conocidas en la técnica. Los métodos de la presente invención pueden utilizarse para tratar un trastorno genético, enfermedad autoinmune, cáncer, infección viral, infección bacteriana o infección parasítica establecidos. Un experto en la técnica reconocerá que el trastorno genético, enfermedad autoinmune, cáncer, infección viral, infección bacteriana o infección parasítica se refieren a condiciones caracterizadas

por la presencia de un patógeno extraño o crecimiento celular anormal. Los síntomas clínicos dependerán de la condición particular y son fácilmente reconocibles por los expertos en la técnica de tratar la condición específica.

5 Por "prevenir" se quiere decir que después de la administración de una sustancia de la presente invención a un sujeto, el sujeto no desarrolla los síntomas completos de la condición (p. ej., trastorno genético, enfermedad autoinmune, cáncer, infección viral, bacteriana o parasítica y/o no desarrolla el trastorno genético, enfermedad autoinmune, cáncer, infección viral, bacteriana o parasítica). Así, la condición se previene completamente o se previenen algunos de los síntomas o indicios reconocidos de la condición o se previene su manifestación completa.

10 Por "elementos transponibles" se quiere decir cualquier construcción genética que incluye pero no está limitada a cualquier gen, fragmento de gen o ácido nucleico que puede integrarse en una secuencia de ADN diana bajo el control de una enzima de integración.

Por "repetición terminal" se quiere decir cualquier secuencia repetitiva en una secuencia de ácidos nucleicos que incluye pero no está limitada a repeticiones invertidas y repeticiones directas.

15 Por "vector" se quiere decir cualquier composición capaz de administrar un ácido nucleico, péptido, polipéptido o proteína en un ácido nucleico, célula, tejido u organismo diana incluyendo pero sin estar limitado a plásmido, fago, transposones, retrotransposones, vector viral y vector retroviral.

Por "vector no viral" se quiere decir cualquier vector que no comprende un virus o retrovirus.

#### B. Composiciones

20 La invención proporciona composiciones que comprenden un ácido nucleico que comprende un transgén flanqueado por dos repeticiones terminales y un ácido nucleico que codifica una enzima de integración bajo el control de un elemento promotor, en las que la enzima de integración es una enzima de integración quimérica que comprende un dominio de unión al ADN específico del anfitrión y en las que el ácido nucleico que codifica el transgén y el ácido nucleico que codifica la enzima de integración quimérica son el mismo ácido nucleico y en las que la enzima de integración quimérica es una transposasa.

25 Tal y como se usa en la presente memoria, plásmidos son agentes que transportan los ácidos nucleicos descritos a la célula sin degradación y permiten la expresión dirigida por un promotor de los ácidos nucleicos que codifican las proteínas (p. ej., transgén y enzima de integración) en las células en las que se administran. En algunas realizaciones, los promotores y amplificadores se obtienen bien de un virus o un retrovirus.

También se describen composiciones de la invención, en las que el ácido nucleico está presente en un vector no viral.

También se describen composiciones de la invención, en las que el elemento promotor es un promotor/amplificador.

30 También se describen composiciones de la invención, en las que el promotor es un promotor específico de sitio.

35 Se ha mostrado que todos los elementos reguladores específicos pueden clonarse y usarse para construir vectores de expresión que se expresan selectivamente en tipos celulares específicos. El promotor específico de sitio puede seleccionarse al menos del grupo que consiste en el promotor de la proteína fibrilar ácida de la glía (GFAP), promotor de mielina básica (MBP), promotor MCK, promotor NSE, promotor de nestina, promotor de sinapsina, promotor de Insulina 2 (Ins2), promotor PSA, promotor de la albúmina, promotor TRP-1 y el promotor de la tirosinasa. También se describe un promotor específico de tejido de mama, tal como el promotor WAP, un promotor específico de tejido de ovario, tal como el promotor ACTB o un promotor específico de tejido óseo. Puede usarse cualquier promotor específico de tejido.

40 También se describen composiciones de la invención, en las que el promotor es inducible. El promotor inducible puede seleccionarse al menos del grupo que consiste en promotor del choque por calor humano, promotor Egr-1, promotor de tetraciclina y el promotor de la calicreína glandular humana 2 (hK2).

Como se necesitará que el elemento transponible se integre en el genoma del anfitrión, se necesita una enzima de integración. Las enzimas de integración pueden ser cualquier transposasa.

Las enzimas de integración quiméricas de la presente invención comprenden dos componentes:

45 Factor de acoplamiento al ADN (primer dominio) (p. ej., Dominio de Unión al ADN (DBD)) y un dominio de integración (enzimático) (segundo dominio). El factor de acoplamiento al ADN puede disponerse en cualquier lugar en relación con el dominio de integración (p. ej., internamente o en el extremo amino o carboxi terminal).

Además, una parte de la enzima de integración de tipo salvaje, por ejemplo, la parte que tiene el DBD de la enzima nativa, podría delecionarse y reemplazarse por un DBD que reconozca el ADN de la célula diana. Las proteínas quiméricas de la invención comprenden un primer dominio que une la proteína quimérica a un ácido nucleico diana y un segundo dominio que integra el ácido nucleico donante (transgén) en el ácido nucleico diana. Tal y como se emplea en la presente memoria, la expresión "proteína quimérica" se refiere a una proteína recombinante modificada por ingeniería genética en la que los dominios de ésta se obtienen de regiones codificadoras heterólogas (es decir, regiones codificadoras obtenidas de diferentes genes). Para construir una transposasa quimérica de la invención, pueden usarse métodos moleculares generales y específicamente los de Katz et al. (Patente EEUU No. 6.150.511).

Las proteínas enzimáticas de integración quiméricas de la invención se preparan por métodos de ADN recombinante, en los que las secuencias de ADN que codifican cada dominio están "unidas de manera operativa" entre sí de manera que después de la expresión, se genera una proteína de fusión que tiene las funciones de direccionamiento y transposasa descritas anteriormente. Tal y como se usa en la presente memoria, el término "unido de manera operativa" significa que los segmentos de ADN que codifican la proteína de fusión están ensamblados uno respecto al otro y respecto a un vector de expresión en el que están insertados, de manera tal que se expresa eficazmente una proteína de fusión funcional.

Tal y como se usa en la presente memoria, "primer dominio" se refiere al dominio en la proteína quimérica cuya función es unir la proteína quimérica a una secuencia de reconocimiento específica en un ácido nucleico diana. El primer dominio tiene una longitud de al menos 5 aminoácidos y puede estar localizado en cualquier lugar de la proteína quimérica, p. ej., internamente o en el extremo amino o carboxi terminal de ésta. El primer dominio puede ser un factor de acoplamiento a ADN, bien un "dominio de unión al ADN" o un "dominio de unión a proteínas" que es operativo para acoplar y/o asociar la proteína quimérica con una secuencia de reconocimiento en el ácido nucleico diana.

Por "factor de acoplamiento al ADN" se quiere decir cualquier secuencia de aminoácidos que se asocia con el ADN directamente o indirectamente. Así, cuando la asociación de la enzima de integración quimérica con el ácido nucleico diana ocurre por unión indirecta, se emplea un dominio de unión a proteínas como el factor de acoplamiento. Los dominios de unión a proteínas adecuados pueden obtenerse de factores de transcripción virales (p. ej., HSV-VP16 y adenovirus E1A) y factores de transcripción celulares. A lo largo de la presente descripción, los términos dominio de unión al ADN, factor de dirección al ADN y dominio de unión a proteínas se usan para hacer referencia a factores de acoplamiento al ADN. Se entiende que estos términos pueden usarse indistintamente a lo largo de la presente invención sin afectar el objetivo global de la invención.

Tal y como se usa en la presente memoria, el término "dominio de unión al ADN" engloba una secuencia peptídica mínima de una proteína de unión al ADN, hasta la longitud completa de una proteína de unión al ADN sin pérdida de función. Cuando se emplea un dominio de unión al ADN en la invención, la asociación de la enzima de integración quimérica con el ácido nucleico diana ocurre por interacción directa con el ácido nucleico del anfitrión. El dominio de unión al ADN pone el segundo dominio (es decir, el dominio de integración) muy próximo a una secuencia de reconocimiento específica en el ácido nucleico diana de manera que un ácido nucleico donante deseado puede integrarse en la secuencia de ácido nucleico diana.

Los dominios de unión al ADN se obtienen típicamente de proteínas de unión al ADN. Se sabe que dichos dominios de unión al ADN funcionan de manera heteróloga en combinación con otros dominios proteicos funcionales manteniendo la capacidad de unir la secuencia de reconocimiento del ADN natural (véase, p. ej., Brent y Ptashne, 1985, Cell, 43: 729-736).

Por ejemplo, se sabe que los receptores de hormonas tienen dominios de unión al ADN intercambiables que funcionan en proteínas quiméricas (véanse, p. ej., Pat. EEUU No. 4.981.784; y Evans, R., 1988, Science, 240: 889-895).

Las "proteínas de unión al ADN" utilizadas en la presente memoria pertenecen a una clase de proteínas muy conocida que son capaces de unirse directamente al ADN y realizar varias funciones, tales como facilitar el inicio de la transcripción o represión de la transcripción. Las proteínas de unión al ADN ejemplares para usarse en la presente memoria incluyen proteínas de control de la transcripción (p. ej., factores de transcripción y semejantes; Conaway y Conaway, 1994, "Transcription Mechanisms and Regulation", Raven Press Series on Molecular and Cellular Biology, Vol. 3, Raven Press, Ltd., Nueva York, N.Y.); enzimas de recombinación (p. ej., recombinasa *hin* y semejantes); y enzimas que modifican el ADN (p. ej., enzimas de restricción y semejantes).

Los factores de transcripción con proteínas de unión al ADN adecuados para usarse en la presente memoria incluyen p. ej., proteínas con homeodominio, proteínas con dedo de cinc, receptores de hormonas, proteínas hélice-giro-hélice, proteínas hélice-bucle-hélice, proteínas cremallera básicas (bZip), factores lazo beta y semejantes. Véase, por ejemplo, Harrison, S., "A Structural Taxonomy of DNA-binding Domains", Nature, 353: 715-719.

- 5 Las proteínas de unión al ADN con homeodominio adecuadas para usarse en la presente memoria incluyen, pero no están limitadas a, HOX, STF-1 (Leonard et al., 1993, *Mol. Endo.*, 7:1275-1283), Antp, Mat, alfa-2, INV (véase también Scott et al., (1989), *Biochem. Biophys. Acta*, 989:25-48). Leonard et al. han encontrado que un fragmento de 76 aminoácidos (correspondiente a los a.a. 140-215 descrito en Leonard et al., 1993, *Mol. Endo.*, 7:1275-1283) que contiene el homeodominio STF-1 se une al ADN tan fuerte como STF-1 de tipo salvaje.
- 10 Los dedos de cinc pueden manipularse para reconocer un amplio intervalo de secuencias. Como tales, estas enzimas tienen el potencial de dirigir la escisión hacia dianas elegidas arbitrariamente. Una rotura bicatenaria (DSB) en la diana cromosómica incrementa en gran medida la frecuencia de eventos de recombinación localizados. Las nucleasas con dedo de cinc (ZFN) tienen un dominio de reconocimiento del ADN compuesto por tres dedos de cinc Cys<sub>2</sub>His<sub>2</sub> unidos a un dominio de escisión del ADN no específico (Y. G. Kim et al. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156). Para actuar como una nucleasa, el dominio de escisión puede dimerizarse (J. Smith et al. (2000) *Nucleic Acids Res.*, 28, 3361). Esto puede conseguirse proporcionando sitios de unión para dos conjuntos de dedos de cinc muy próximos y en las orientaciones apropiadas (J. Smith et al. (2000) *Nucleic Acids Res.*, 28, 3361; M. Bibikova et al. (2001) *Mol. Cell Biol.* 21, 289). Las proteínas de unión al ADN con dedo de cinc adecuadas proporcionadas para usarse en la presente memoria incluyen pero no se limitan a Zif268, GLI y XFin. Estas proteínas pueden encontrarse a lo largo de la bibliografía a través de Klug y Rhodes (1987), *Trends Biochem. Sci.*, 12:464; Jacobs y Michaels (1990), *New Biol.*, 2:583; y Jacobs (1992), *EMBO J.*, 11:4507-4517.
- 20 Las proteínas de unión al ADN de receptores de hormonas ejemplares para usarse en la presente memoria incluyen pero no están limitadas a receptor de glucocorticoides, receptor de la hormona tiroidea y receptor de estrógenos y están descritas en la bibliografía (Pat. EEUU Nos. 4.981.784; 5.171.671; y 5.071.773).
- 25 Las proteínas de unión al ADN hélice-giro-hélice ejemplares para usarse en la presente memoria incluyen pero no están limitadas a represor lambda, represor cro, represor 434 y 434 cro. Se proporcionan estas proteínas de unión al ADN hélice-giro-hélice (Pabo y Sauer, 1984, *Annu. Rev. Biochem.*, 53:293-321).
- 30 Las proteínas de unión al ADN hélice-bucle-hélice ejemplares para usarse en la presente memoria incluyen pero no están limitadas a MRF4 (Block et al., 1992, *Mol. and Cell Biol.*, 12(6):2484-2492), CTF4 (Tsay et al., 1992, *NAR*, 20(10):2624), NSCL, PAL2 y USF. Véanse, para revisión, Wright (1992) *Current Opinion in Genetics and Development*, 2(2):243-248; Kadesch, T. (1992), *Immun. Today*, 13(1):31-36 y Garell y Campuzano (1991), *Bioessays*, 13(10):493-498.
- 35 Las proteínas cremallera básicas de unión al ADN ejemplares para usarse en la presente memoria incluyen pero no están limitadas a GCN4, fos y jun (véase, para revisión, Lamb y McKnight, 1991, *Trends Biochem. Sci.*, 16:417-422). Los factores lazo beta ejemplares proporcionados para usarse en la presente memoria incluyen Met-J, ARC y MNT.
- 40 Las enzimas de recombinación con proteínas de unión al ADN adecuadas para usarse en la presente memoria incluyen pero no están limitadas a la familia hin de recombinasas (p. ej., hin, gin, pin y cin; véase, Feng et al., 1994, *Science*, 263:348-355), la familia lambda de integrasas, recombinasa flp, transposones TN916 y la familia de resolvasas (p. ej., resolvasa TN21).
- 45 Las enzimas que modifican el ADN con proteínas de unión al ADN adecuadas para usarse en la presente memoria incluyen, por ejemplo, enzimas de restricción, enzimas de reparación del ADN y metilasas específicas de sitio. Para usarse en la presente invención, las enzimas de restricción se modifican usando métodos muy conocidos en la técnica para eliminar la función de digestión con restricción de la proteína a la vez que se mantiene la función de unión al ADN (véase, p. ej., King et al., 1989, *J. Biol. Chem.*, 264(20):11807-11815). Así, en la presente memoria puede emplearse cualquier enzima de restricción. La utilización de una enzima de restricción que reconoce una secuencia de ADN rara permite la unión de la proteína quimérica de la invención a relativamente pocos sitios en una cadena particular de ADN genómico.
- 50 La modificación de dominios de unión al ADN existentes para reconocer nuevas secuencias de reconocimiento diana también se contempla en la presente memoria. Se ha encontrado que los métodos de evolución in vitro pueden aplicarse para modificar y mejorar dominios de unión al ADN existentes. Devlin et al., 1990, *Science*, 249:404-406; y Scott y Smith, 1990, *Science*, 249:386-390 para enseñanzas en la modificación de dominios de unión al ADN existentes.
- El o los "dominios de unión a proteínas" adecuados para usarse como el "primer dominio" de la proteína quimérica de la invención se obtienen típicamente de proteínas capaces de unirse a otra proteína (p. ej., un factor de transcripción) que se une (acopla) bien directamente o indirectamente a la secuencia de ácido nucleico diana. Así, cuando un dominio de unión a proteínas se emplea como el primer dominio, la asociación de la proteína quimérica de la invención con el ácido nucleico diana ocurre por unión indirecta. Los dominios de unión a proteínas adecuados pueden obtenerse, por ejemplo, de factores de transcripción virales (p. ej., HSV-VP16, adenovirus E1A y semejantes), factores de transcripción celulares y semejantes usando métodos moleculares rutinarios.

- Además de los dominios de unión a proteínas fácilmente disponibles, pueden obtenerse dominios de unión a proteínas pequeños, p. ej., en el intervalo de aproximadamente 5-25 aminoácidos empleando los métodos descritos de "biblioteca de exposición en fagos" (Rebar y Pabo, 1994, Science, 263:671-673). Se ha encontrado que pueden aislarse péptidos cortos usando bibliotecas de exposición en fagos que se unen a una proteína seleccionada. Por ejemplo, se obtuvo un péptido a partir de una biblioteca que exponía hexámeros de aminoácidos aleatorios en la superficie de un fago que se unían específicamente a la avidina; este péptido no presentaba similitud con ninguno de los ligandos de la avidina conocidos (Devlin et al., 1990, Science, 249:404-406). Este método muy conocido se usa para crear dominios de unión a proteínas que se unen a proteínas que ya están unidas in vivo al ácido nucleico diana deseado.
- Las regiones microsátelite son secuencias repetitivas en el genoma. Mediante el direccionamiento de secuencias repetitivas ya sea a través de una enzima de integración quimérica o a través de secuencias homólogas se puede dirigir la integración a regiones no transcritas del genoma (es decir, eliminando el riesgo de mutagénesis insercional) y al tener más dianas se incrementa la eficacia de la integración, es decir, es mejor muchas dianas que una. Son regiones repetitivas no codificadoras en el genoma que permiten la integración como se describe en la presente memoria, seguido de la transcripción del transgén dirigida por el promotor proporcionado en la construcción.
- La enzima de integración quimérica de la invención comprende un dominio de integración (enzimático) (segundo dominio). El dominio de integración comprende o se obtiene de una enzima de integración. Las enzimas de integración pueden ser cualquiera.
- Se describen composiciones, en las que la enzima de integración es una transposasa. Se entiende y se contempla en la presente memoria que la transposasa de la composición no está limitada a ninguna transposasa y que puede seleccionarse de al menos el grupo que consiste en *Bella Durmiente* (SB), *Tn7*, *Tn5*, *mos 1*, *piggybac*, *Himar1*, *Hermes*, elemento To12, *Pokey*, *Minos*, elementos S, elemento P, ICES1, elementos Quetzal, *Tn916.maT*, *Tc1/mariner* y Tc3.
- Cuando la enzima de integración es una transposasa, se entiende que la transposasa de la composición no está limitada a ninguna transposasa y que puede seleccionarse de al menos el grupo que consiste en *Bella Durmiente* (SB), *Tn7*, *Tn5*, *Tn916*, *Tc1/mariner*, *Minos* y elementos S, elementos Quetzal, elementos Txr, *maT*, *mos1*, *piggybac*, *Himar1*, *Hermes*, elemento To12, *Pokey*, elemento P y Tc3. En la técnica pueden encontrarse transposasas adicionales, por ejemplo, Patente EEUU No. 6.225.121, Patente EEUU No. 6.218.185, Patente EEUU No. 5.792.924, Patente EEUU No. 5.719.055, Solicitud de Patente EEUU No. 20020028513 y Solicitud de Patente EEUU No. 20020016975. Como la parte principal aplicable de la invención sigue siendo la misma, las composiciones de la invención pueden incluir transposasas quiméricas construidas a partir de transposasas todavía sin identificar.
- También se describen enzimas de integración en las que la enzima es una integrasa. Por ejemplo, la enzima de integración puede ser una integrasa de bacteriófago. Dicha integrasa puede incluir cualquier integrasa de bacteriófago y puede incluir pero no está limitada a bacteriófago lamda ( $\lambda$ ) y bacteriófago mu ( $\mu$ ), así como Hong Kong 022 (Cheng Q., et al. Specificity determinants for bacteriophage Hong Kong 022 integrase: analysis of mutants with relaxed core-binding specificities. (2000) Mol. Microbiol. 36(2):424-36), HP1 (Hickman, A.B., et al. (1997). Molecular organization in sitespecific recombination: The catalytic domain of bacteriophage HP1 integrase at 2.7 Å resolution. Cell 89:227-237), P4 (Shoemaker, NB, et al. (1996). The Bacteroides mobilizable insertion element, NBU1, integrates into the 3' end of a Leu-tRNA gene and has an integrase that is a member of the lambda integrase family. J. Bacteriol. 178(12):3594-600), P1 (Li Y, y Austin S. (2002) The P1 plasmid in action: timelapse photomicroscopy reveals some unexpected aspects of plasmid partition. Plasmid. 48(3):174-8) y T7 (Rezende, L.F., et al. (2002) Essential Amino Acid Residues in the Single-stranded DNA-binding Protein of Bacteriophage T7. Identification of the Dimer Interface. J. Biol. Chem. 277, 50643-50653).
- La integrasa mantiene su actividad cuando se fusiona con otras proteínas. Esto se ha demostrado por el uso de proteínas de fusión represor-integrasa de lambda (40) y proteína de unión a maltosa integrasa (41). Además, las recombinasas quiméricas, factores de transcripción, oncogenes, etc. han mantenido su actividad cuando se fusionan con otros dominios proteicos (42). Sin embargo, todavía no se han demostrado los intentos de dirigir in vivo retrovirus selectivos de sitio que incluían secuencias que codifican proteínas de fusión de integrasa (43-45). Los elementos Tc1/mariner son promiscuos y se han usado con éxito como vectores de transgén de una especie a otra en moscas (49-53), mosquitos (54), bacterias (55), protozoos (56) y vertebrados.
- También se describen enzimas de integración en las que la enzima es una recombinasa. Por ejemplo, la recombinasa puede ser una recombinasa Cre, recombinasa Fip, recombinasa HIN o cualquier recombinasa adicional. Las recombinasas son muy conocidas en la técnica. Una lista extensa de recombinasas puede encontrarse en Nunes-Duby SE, et al. (1998) Nuc. Acids Res. 26(2):391-406, por sus enseñanzas sobre recombinasas y sus secuencias.
- También se describen enzimas de integración en las que la enzima es una retrotransposasa. Por ejemplo, la retrotransposasa puede ser una retrotransposasa Gate (Kogan GL, et al. (2003) The GATE retrotransposon in Drosophila

melanogaster: mobility in heterochromatin and aspects of its expression in germline tissues. *Mol Genet Genomics*. 269(2):234-42).

La enzima de integración quimérica de la invención puede tener el dominio de unión específico del anfitrión fusionado con el N terminal de la transposasa.

5 La enzima de integración quimérica de la invención puede tener el dominio de unión específico del anfitrión fusionado con el C terminal de la transposasa.

10 También se proporcionan composiciones que comprenden una construcción de ácido nucleico que codifica un transgén bajo el control de un elemento promotor flanqueado por dos repeticiones internas y un ácido nucleico que codifica una enzima de integración bajo el control de un elemento promotor. Algunas repeticiones internas (p. ej., algunos elementos nucleares cortos y largos intercalados) son permisivas para la integración selectiva de sitio (68-69) y permitirán la expresión del transgén incluso sin regiones de unión a la matriz nuclear flanqueando el transgén (66-67). Se han identificado proteínas que se unen selectivamente a los elementos repetidos intercalados (70-73) y se incorporan en la presente memoria por referencia. El desarrollo de proteínas de fusión que incorporan dominios de unión al ADN con secuencias de ADN conocidas, permisivas para la transcripción y repetitivas permite la integración dirigida como se ha descrito anteriormente.

15 En el transgén flanqueado por las repeticiones terminales, las repeticiones terminales pueden obtenerse a partir de transposones conocidos. Los ejemplos de transposones incluyen, pero no están limitados a los siguientes:

20 Bella Durmiente (Izsvak Z, Ivics Z, y Plasterk RH. (2000) Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J. Mol. Biol.* 302:93-102), *mos1* (Bessereau JL, et al. (2001) Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line. *Nature*, 413(6851):70-4; Zhang L. et al. (2001) DNA-binding activity and subunit interaction of the mariner transposase. *Nucleic Acids Res.* 29(17):3566-75, *piggybac* (Tamura T, et al. Germline transformation of the silkworm *Bombyx mori* L. using a piggyBac transposon-derived vector. *Nat. Biotechnol.* 2000 Ene;18(1):81-4), Himar1 (Lampe DJ, et al. (1998) Factors affecting transposition of the Himar1 mariner transposon in vitro. *Genetics*. 149(1):179-87), Hermes, elemento Tol2, Pokey, Tn5 (Bhasin A, et al. (2000) Characterization of a Tn5 pre-cleavage synaptic complex. *J Mol Biol* 302:49-63), Tn7 (Kuduvalli PN, Rao JE, Craig NL. (2001) Target DNA structure plays a critical role in Tn7 transposition. *EMBO J* 20:924-932), Tn916 (Marra D, Scott JR. (1999) Regulation of excision of the conjugative transposon Tn916. *Mol Microbiol* 2:609-621), Tc1/mariner (Izsvak Z, Ivics Z, Hackett PB. (1995) Characterization of a Tc-1 like transposable element in zebrafish (*Danio rerio*). *Mol. Gen. Genet.* 247:312-322), elementos *Minos* y *S* (Franz G y Savakis C. (1991) Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. *Nucl. Acids Res.* 19:6646; Merriman PJ, Grimes CD, Ambroziak J, Hackett DA, Skinner P y Simmons MJ. (1995) S elements: a family of Tc1-like transposons in the genome of *Drosophila melanogaster*. *Genetics* 141:1425-1438), elementos Quetzal (Ke Z, Grossman GL, Cornel AJ, Collins FH. (1996) Quetzal: a transposon of the Tc1 family in the mosquito *Anopheles albimanus*. *Genetica* 98:141-147); elementos Txr (Lam WL, Seo P, Robison K, Virk S, y Gilbert W. (1996) Discovery of amphibian Tc1-like transposon families. *J Mol Biol* 257:359-366), subfamilias de transposón semejantes a Tc1 (Ivics Z, Izsvak Z, Minter A, Hackett PB. (1996) Identification of functional domains and evolution of Tc1-like transposable elements. *Proc. Natl. Acad Sci USA* 93:5008-5013), y Tc3 (Tu Z, Shao H. (2002) Intra- and inter-specific diversity of Tc-3 like transposons in nematodes and insects and implications for their evolution and transposition. *Gene* 282:133-142), ICEst1 (Burrus V et al. (2002) The ICEst1 element of *Streptococcus thermophilus* belongs to a large family of integrative and conjugative elements that exchange modules and change their specificity of integration. *Plasmid*. 48(2):77-97), *maT* y elemento P (Rubin GM y Spradling AC. (1983) Vectors for P element mediated gene transfer in *Drosophila*. *Nucleic Acids Res.* 11:6341-6351).

45 La translocación del transposón Bella Durmiente (SB) requiere la unión específica de la transposasa SB a repeticiones terminales invertidas (ITR) de aproximadamente 230 pb en cada extremo del transposón, que es seguido por una transferencia de corte y pega del transposón en la secuencia de ADN diana. Las ITR contienen dos repeticiones directas imperfectas (DR) de aproximadamente 32 pb. Las DR exteriores están en los extremos terminales del transposón mientras que las DR internas están localizadas en el interior del transposón, 165-166 pb desde las DR externas. Cui et al. (*J. Mol Biol* 318:1221-1235) investigaron los papeles de los elementos DR en la transposición. En el elemento de 1286 pb, las regiones esenciales están contenidas en los intervalos limitados por las coordenadas 229-586, 735-765 y 939-1066, que numeran los pares de bases desde el extremo terminal 5' del elemento. Estas regiones pueden contener secuencias que son necesarias para la unión de la transposasa o que se necesitan para mantener un espaciamiento apropiado entre los sitios de unión.

55 Los transposones están rodeados de repeticiones terminales invertidas que contienen sitios de unión para la transposasa. Los elementos del subgrupo IR/DR de la superfamilia Tc1/mariner tienen un par de sitios de unión a transposasa en los extremos de las repeticiones invertidas (IR) de 200-250 pb (Izsvak, et al. 1995). Los sitios de unión contienen repeticiones directas cortas (DR) de 15-20 pb. Esta estructura característica puede encontrarse en varios

elementos de especies evolutivamente distantes, tales como Minos y elementos S en moscas (Franz y Savakis, 1991; Merriman et al, (1995), elementos Quetzal en mosquitos (Ke et al, 1996), elementos T<sub>xr</sub> en ranas (Lam et al, 1996) y al menos tres subfamilias de transposones semejantes a Tc1 en peces (Ivics et al., 1996), incluyendo SB [Bella Durmiente].

5 Mientras que los transposones Tc1 requieren un sitio de unión para su transposasa en cada IR, la Bella Durmiente requiere dos sitios de unión de repeticiones directas (DR) en cada IR y, por lo tanto, se clasifica con Tc3 en un subgrupo IR/DR de la superfamilia Tc1/mariner (96, 97). La Bella Durmiente se transpone en sitios con dinucleótido TA y deja la huella característica de Tc1/mariner, es decir, duplicación del TA, después de la escisión. La construcción de ácido nucleico contiene el transgén que está flanqueado por secuencias IR/DR, que actúan como los sitios de unión para la transposasa. La transposasa catalíticamente activa puede expresarse a partir de un sistema de plásmido independiente (trans) o del mismo (cis). La transposasa se une a las IR/DR, cataliza la escisión del transgén flanqueado y media su integración en el genoma del anfitrión diana.

10 Tc3 de *Caenorhabditis elegans* es uno de los miembros fundadores de la familia Tc1 que incluye transposones de ADN en vertebrados, insectos, nematodos y hongos. Tu A, et al. (Gene 282:133-142) presentan la caracterización de varios transposones semejantes a Tc3 en *C. elegans*, *Caenorhabditis briggsae* y *Drosophila melanogaster*, que ha revelado altos niveles de diversidad inter e intra específica y sugiere además una amplia distribución de los transposones semejantes a Tc3. Estos transposones recién definidos y los descritos anteriormente Tc3 y Msq Tc3 forman un clado altamente divergente y distinto en la familia Tc1. La mayor parte de los transposones semejantes a Tc3 contienen dos sitios de unión putativos para sus transposasas. El primero está cerca del extremo y el segundo está a aproximadamente 164-184 pb del primer sitio. Existe una gran variación en la longitud (27-566 pb) y estructura de las repeticiones terminales invertidas (TIR) de los transposones semejantes a Tc3.

15 *Mos1* es un miembro de la familia *mariner/Tc1* de elementos transponibles identificados originalmente en *Drosophila mauritiana*. Tiene repeticiones terminales invertidas de 28 pb y como otros elementos de este tipo se transpone por un mecanismo de corte y pega, se inserta en los dinucleótidos TA y codifica una transposasa. Es la única proteína requerida para la transposición *in vitro*. Zhang y colaboradores (Nucleic Acids Res 29:3566-3575) han investigado las propiedades de unión al ADN de la transposasa *Mos1* y el papel de las interacciones transposasa-transposasa en la transposición. La transposasa purificada reconoce las repeticiones terminales invertidas de *Mos1* debido a un dominio de unión al ADN en los 120 aminoácidos N terminales. Esto requiere un resto hélice-giro-hélice putativo entre los restos 88 y 108. La unión se realiza preferentemente en el extremo del lado derecho, que se diferencia en cuatro posiciones de la repetición en el extremo del lado izquierdo. La escisión de *Mos1* por la transposasa también se realiza preferentemente en el extremo del lado derecho.

20 Tomando como base los requerimientos para la integración de los elementos transponibles, parece que es necesario un factor de direccionamiento del ADN anfitrión para la integración eficaz mediante la yuxtaposición del complejo transposón-transposasa adyacente al ADN del anfitrión. De hecho, las transposasas Tc1/mariner tienen dominios de unión al ADN. Sin embargo, estos dominios de unión al ADN aparentemente no son selectivos de sitio (35), posiblemente carecen de sitios de reconocimiento fuerte en determinados genomas de anfitriones y pueden requerir otras proteínas del anfitrión para la integración eficaz por acoplamiento del transposón-transposasa con el ADN del anfitrión.

25 La invención supera esta deficiencia proporcionando composiciones que comprenden una construcción de ácido nucleico que comprende además una enzima de integración quimérica (es decir, enzima de integración-dominio de unión al ADN del anfitrión) para eludir el requerimiento potencial de un factor o factores de direccionamiento al ADN del anfitrión para una integración eficaz y selectiva de sitio. Se entiende que la enzima de integración quimérica puede ser cualquier transposasa quimérica.

30 Así, se describen composiciones que comprenden un transgén flanqueado por repeticiones terminales de un elemento transponible, p. ej., Bella Durmiente, y una enzima quimérica requerida (p. ej., dominio de unión al ADN del anfitrión-transposasa) en un sistema de empaquetamiento no viral para la integración dirigida en el genoma del anfitrión. Es una realización de la presente invención que esta enzima quimérica sustituirá al dominio de unión al ADN nativo de la enzima de integración por uno que es específico del anfitrión y selectivo de sitio, eludiendo de esta manera el requerimiento de un factor de direccionamiento al ADN del anfitrión.

35 También se describen composiciones de la invención, en las que la transposasa es una transposasa quimérica que comprende un dominio de unión al ADN específico de anfitrión o específico de sitio.

40 Así, la presente invención se refiere a nuevas transposasas quiméricas y a los transposones que se usan para introducir secuencias de ácido nucleico en el ADN de una célula. Una transposasa es una enzima que es capaz de unirse al ADN en regiones del ADN denominadas repeticiones invertidas. Los transposones contienen típicamente al menos una, y preferiblemente dos, repeticiones invertidas que flanquean una secuencia de ácido nucleico implicada. La transposasa

se une a sitios de reconocimiento en la repeticiones invertidas y cataliza la incorporación del transposón en el ADN del anfitrión. La función de transposón está limitada frecuentemente a la especie de anfitrión. Incluso en aquellos transposones que no están limitados a su "anfitrión normal" la eficacia de la integración varía dramáticamente. Esta invención incrementa la eficacia de la integración modificando una transposasa para incluir un dominio del unión al ADN del anfitrión (ya sea o no para el propósito de selectividad de sitio) como se describe en la presente memoria. El nuevo dominio de unión al ADN de esta transposasa quimérica puede añadirse a las transposasas nativas o puede sustituir el dominio de unión al ADN de la transposasa nativa. Así, la transposasa quimérica de ADN del anfitrión [factor de direccionamiento], sitios de reconocimiento en el plásmido que reconocerán una proteína endógena (o una proteína recién introducida) que dirigirán el complejo a las proximidades del ADN del anfitrión, incorporando secuencias semejantes a las del anfitrión (p. ej., secuencias repetitivas) o una combinación de los papeles anteriores en la integración de transgenes selectiva de sitio y/o eficaz se proporciona por la presente invención.

Los vectores de transferencia génica para la terapia génica pueden clasificarse de manera amplia como vectores virales o vectores no virales. El uso de las construcciones de ácido nucleico que comprenden una enzima de integración quimérica proporciona una mejora importante y sorprendente sobre la transferencia génica mediada por ADN no viral. Hasta el momento, los vectores virales han sido el centro de los esfuerzos de la terapia génica, ya que se ha encontrado que son más eficaces para introducir y expresar genes en las células que los vectores no virales. Una vez que se han superado los problemas de eficacia de la técnica anterior, como se enseña en la presente memoria, existen varias ventajas de la transferencia génica no viral sobre la transferencia génica mediada por virus para el desarrollo de nuevas terapias génicas. Por ejemplo, la adaptación de virus como agentes para la terapia génica restringe el diseño genético a las limitaciones del genoma de ese virus en términos de tamaño, estructura y regulación de la expresión. Los vectores no virales se generan fundamentalmente a partir de materiales de partida sintéticos y, por lo tanto, se fabrican más fácilmente que los vectores virales. Es menos probable que los reactivos no virales sean inmunogénicos comparados con los agentes virales lo que hace posible la administración repetida. Los vectores no virales son más estables que los vectores virales y, por lo tanto, son más adecuados para la formulación y la aplicación farmacéutica que los vectores virales.

En realizaciones pasadas, los sistemas de transferencia génica no virales no se han equipado para estimular la integración de ácido nucleico en el ADN de una célula, incluyendo los cromosomas del anfitrión. Como resultado, las frecuencias de transferencia génica estable usando sistemas no virales ha sido muy baja; 0,1% como máximo en células de tejido cultivadas y mucho menor en células y tejidos y primarios. Los esfuerzos de la técnica anterior en vectores no virales basados en transposones han intentado proporcionar un sistema de transferencia génica con construcciones de ácido nucleico no virales que facilitan la integración y mejoran de manera importante la frecuencia de la transferencia génica estable. Sin embargo, la integración no es específica de sitio y no es eficaz de manera uniforme y puede variar de manera importante dependiendo de la línea celular anfitriona. Esta invención permite la integración selectiva de sitio en el genoma del anfitrión y proporciona la ventaja sorprendente de una integración eficaz en aquellos anfitriones que no tienen el factor de direccionamiento al ADN requerido como se menciona en la presente memoria.

En el sistema de transferencia génica de esta invención, la enzima de integración quimérica puede introducirse en la célula como una proteína o como un ácido nucleico que codifica la proteína. En una realización, el ácido nucleico que codifica la proteína es ARN y en otra, el ácido nucleico es ADN. Además, el ácido nucleico que codifica la proteína transposasa quimérica puede incorporarse en una célula a través de un vector viral, lípido catiónico u otros mecanismos de transfección estándar incluyendo electroporación o bombardeo de partículas usados para células eucariotas. Después de o simultáneamente a la introducción del ácido nucleico que codifica la transposasa quimérica, el fragmento de ácido nucleico de esta invención puede introducirse en la misma célula. Alternativamente, el ácido nucleico que codifica la transposasa quimérica puede ser el mismo ácido nucleico que incluye el transgén y las repeticiones terminales.

De manera similar, el fragmento de ácido nucleico puede introducirse en la célula como un fragmento lineal o como un fragmento circular. Preferiblemente, la secuencia de ácido nucleico comprende al menos una parte de un marco de lectura abierto para producir un producto que contiene aminoácidos funcional. En una realización preferida, la secuencia de ácido nucleico codifica al menos un péptido, polipéptido o proteína activo o funcional e incluye al menos un promotor seleccionado para dirigir la expresión del marco de lectura abierto o región codificadora de la secuencia de ácido nucleico. La proteína codificada por la secuencia de ácido nucleico puede ser cualquiera de varias proteínas recombinantes nuevas o conocidas en la técnica. En una realización, la proteína codificada por la secuencia de ácido nucleico es una proteína marcadora tal como la proteína verde fluorescente (GFP), cloranfenicol acetiltransferasa (CAT), hormonas del crecimiento, por ejemplo para estimular el crecimiento en un animal transgénico, beta-galactosidasa (lacZ), luciferasa (LUC) y factores de crecimiento semejantes a la insulina (IGF).

El sistema de transferencia génica de esta invención puede usarse fácilmente para producir animales transgénicos que portan un marcador particular o que expresan una proteína particular en una o más células del animal. Los métodos para producir animales transgénicos son conocidos en la técnica y la incorporación del sistema de transferencia génica de

esta invención en estas técnicas no requiere demasiada experimentación. Además, una revisión de la producción de proteínas biofarmacéuticas en la leche de animales lecheros transgénicos (véase Young et al., *BIO PHARM* (1997), 10, 34-38) y las referencias proporcionadas en ésta, detallan métodos y estrategias para producir proteínas recombinantes en la leche y se incorporan en la presente memoria en su totalidad respecto a las enseñanzas relacionadas con la producción de proteínas biofarmacéuticas. Los métodos y el sistema de transferencia génica de esta invención pueden incorporarse fácilmente en estas técnicas transgénicas sin demasiada experimentación a la vista de lo que se conoce en la técnica y particularmente a la vista de esta descripción.

También se describe un animal transgénico, en el que el animal transgénico actúa como un biorreactor, la proteína es un producto para aislamiento de una célula. Se conocen los animales transgénicos como biorreactores. La proteína puede producirse en grandes cantidades en la leche, orina, sangre o huevos. Se conocen promotores que estimulan la expresión en la leche, orina, sangre o huevos y éstos incluyen, pero no están limitados a, promotor de la caseína, promotor de la proteína urinaria de ratón, promotor de la beta-globina y promotor de la ovoalbúmina, respectivamente. Usando otros métodos para producir proteínas en una célula se han producido hormona de crecimiento recombinante, insulina recombinante y varias proteínas recombinantes más. Los ácidos nucleicos que codifican éstas u otras proteínas pueden incorporarse en el fragmento de ácido nucleico de esta invención e introducirse en una célula. La incorporación eficaz del fragmento de ácido nucleico en el ADN de una célula ocurre cuando está presente una transposasa quimérica como se describe en la presente memoria. Cuando la célula es parte de un tejido o parte de un animal transgénico, pueden obtenerse grandes cantidades de proteína recombinante. Existen varios métodos para producir animales transgénicos para investigación o para la producción de proteínas. Las referencias siguientes son respecto a enseñanzas de métodos para producir animales transgénicos (Hackett et al. (1993). *The molecular biology of transgenic fish*. En *Biochemistry and Molecular Biology of Fishes* (Hochachka y Mommseb, eds) Vol. 2, p. 207-240. Otros métodos para producir animales transgénicos incluyen las enseñanzas de M. Markkula et al., *Rev. Reprod.*, 1, 97-106 (1996); R.T. Wall et al., *J. Dairy Sci*, 80, 2213-2224 (1997); J.C. Dalton, et al., *Adv. Exp. Med. Biol.*, 411, 419-428 (1997); y H. Lubon et al., *Transfus. Med. Rev.*, 10, 131-143 (1996). Se obtuvieron peces zebra transgénicos como describen Hackett et al (Solicitud de Patente #20020016975). Los sistemas basados en transposones también se han ensayado a través de la introducción del ácido nucleico con una proteína marcadora en células madre embrionarias de ratón (ES) y se sabe que estas células pueden usarse para producir ratones transgénicos (A. Bradley et al., *Nature*, 309, 255-256 (1984)).

En general, existen dos métodos para conseguir preparaciones madre mejoradas de animales comercialmente importantes. El primero es cría clásica, que ha funcionado bien para animales terrestres pero tarda décadas en producir cambios importantes. Una revisión por Hackett et al. (1997) señala que mediante la cría controlada las velocidades de crecimiento en salmón coho (*Oncorhynchus kisutch*) se incrementaban 60% en cuatro generaciones y los pesos corporales de dos cepas de bagre de canal (*Ictalurus punctatus*) se incrementaron 21 a 29% en tres generaciones. El segundo método es ingeniería genética, un proceso selectivo por el cual se introducen genes en los cromosomas de animales o plantas para proporcionar a estos organismos un nuevo rasgo o característica, tal como crecimiento mejorado o mayor resistencia a enfermedades. Los resultados de la ingeniería genética han superado a los de la cría en algunos casos. En una única generación, se obtuvieron incrementos en el peso corporal de 58% en la carpa común (*Cyprinus carpio*) con genes extra de hormona de crecimiento I de trucha arcoiris, más del 1.000% en salmón con genes extra de la hormona de crecimiento de salmón y menos en trucha. La ventaja de la ingeniería genética en peces, por ejemplo, es que un organismo puede alterarse directamente en periodos muy cortos de tiempo si se ha identificado el gen apropiado (véase Hackett, 1997). La desventaja de la ingeniería genética en peces es que se ha identificado una pequeña parte de los muchos genes que están implicados en el crecimiento y desarrollo y no se conocen en profundidad las interacciones de sus productos proteicos. Los procedimientos para la manipulación genética hacen falta en muchos animales económicamente importantes. La presente invención proporciona un sistema eficaz para llevar a cabo mutagénesis insercional (direccionamiento génico) y procedimientos eficaces para producir animales transgénicos.

El sistema basado en transposones de esta invención tiene aplicaciones en muchas áreas de la biotecnología. El desarrollo de elementos transponibles para vectores en animales permite lo siguiente: 1) inserción eficaz de material genético en los cromosomas animales usando los métodos proporcionados en esta solicitud; 2) identificación, aislamiento y caracterización de genes implicados en el crecimiento y desarrollo a través del uso de transposones como mutágenos de inserción (p. ej., véase Kaiser et al., 1995, "Eukaryotic transposable elements as tools to study gene structure and function". En *Mobile Genetic Elements*, IRL Press, p. 69-100); 3) identificación, aislamiento y caracterización de secuencias reguladoras de la transcripción que controlan el crecimiento y desarrollo; 4) uso de construcciones marcadoras para análisis de loci de rasgos cuantitativos (QTL); y 5) identificación de loci genéticos de rasgos económicamente importantes, además de los de crecimiento y desarrollo, es decir, resistencia a enfermedades (p. ej., Anderson et al., 1996, *Mol. Mar. Biotech.*, 5, 105-113). En un ejemplo, el sistema de esta invención puede usarse para producir peces transgénicos estériles. Una camada madre con genes inactivados podría aparearse para producir crías estériles para la contención biológica o para maximizar las velocidades de crecimiento en peces criados en acuicultura.

En otro uso más del sistema de transferencia génica de esta invención, el fragmento de ácido nucleico incluye un gen para proporcionar una terapia génica a una célula. El gen se pone bajo el control de un promotor específico de tejido o de un promotor ubicuo o una o más regiones de control de la expresión para la expresión de un gen en una células que necesita este gen. Los ácidos nucleicos terapéuticos de interés incluyen genes que reemplazan genes defectuosos en la célula diana del anfitrión, tales como los responsables de condiciones patológicas basadas en defectos genéticos, genes que tienen utilidad terapéutica en el tratamiento del cáncer y semejantes. Se han ensayado varios genes para varias terapias génicas incluyendo, pero sin estar limitadas a, el gen del regulador transmembrana de la fibrosis quística (CFTR), adenosina desaminasa (ADA) para trastornos del sistema inmune, factor IX e interleuquina-2 (IL-2) para enfermedades de las células sanguíneas, alfa-1-antitripsina para enfermedades pulmonares y factores de necrosis tumoral (TNF) y proteínas de resistencia a múltiples fármacos (MDR) para terapias de cáncer. Otros genes terapéuticos específicos para usarse en el tratamiento de condiciones patológicas basadas en defectos genéticos incluyen los genes que codifican los productos siguientes: factor VIII, beta-globina, receptor de la proteína de baja densidad, purina nucleósido fosforilasa, esfingomielinasa, glucocerebrosidasa, regulador transmembrana de la fibrosis quística, CD-18, ornitina transcarbamilasa, arginosuccinato sintetasa, fenilalanina hidroxilasa, alfa-cetoácido deshidrogenasa de cadena ramificada, fumarilacetoacetato hidrolasa, glucosa 6-fosfatasa, alfa-L-fucosidasa, beta-glucuronidasa, alfa-L-iduronidasa, galactosa 1-fosfato uridiltransferasa y semejantes. Los genes terapéuticos para el cáncer que pueden administrarse a través de los presentes vectores incluyen: genes que incrementan la actividad antitumoral de los linfocitos, genes cuyo producto de expresión incrementa la inmunogenicidad de las células tumorales, genes supresores de tumores, genes de toxinas, genes suicidas, genes de resistencia a múltiples fármacos, secuencias antisentido, ARN de interferencia pequeños y semejantes. Debido a la longitud del ácido nucleico que pueden portar los presentes vectores, los presentes vectores pueden usarse no sólo para introducir un gen terapéutico de interés sino también cualquier elemento regulador de la expresión, tal como promotores, y semejantes, que pueden desearse para obtener la expresión temporal o espacial deseada del gen terapéutico. Estas secuencias génicas y varias secuencias génicas humanas o animales específicas incluyendo secuencias génicas para codificar proteínas marcadoras y varias proteínas recombinantes están disponibles en las bases de datos génicas conocidas tales como GenBank y semejantes.

La invención puede ser particularmente útil para la administración de vacunas. En este aspecto de la invención, el antígeno o inmunógeno puede expresarse de manera heteróloga (p. ej., por inserción recombinante de una secuencia de ácido nucleico que codifica el antígeno) o como un inmunógeno (incluyendo fragmentos antigénicos o inmunogénicos) en un vector viral. Alternativamente, el antígeno o inmunógeno puede expresarse en una vacuna de virus vivo atenuado pseudotipificada, por ejemplo. También se entiende que las construcciones de ácido nucleico descritas en la presente memoria pueden usarse para la administración de vacunas. Generalmente, los métodos pueden usarse para generar respuestas inmunes humorales y celulares, p. ej., a través de la expresión de proteínas heterólogas obtenidas de patógenos o fragmentos de éstas en células diana específicas.

Un problema superado por la presente invención es la integración no selectiva como se observa en la mayor parte de los sistemas de transposón (p. ej., Solicitud de Patente #20020016975) que crea el potencial de mutagénesis insercional de genes vitales (p. ej., disrupción de un antioncogén, dando lugar así potencialmente a carcinogénesis).

Las composiciones y métodos de la presente invención también son útiles para la introducción de una secuencia de ácido nucleico de interés en células de planta para producir plantas transgénicas. Tal y como se usa en la presente memoria, el término "planta transgénica" se refiere a la introducción de secuencias de ácido nucleico extrañas en el genoma nuclear, mitocondrial o de plástidos de una planta. Tal y como se usa en la presente memoria, el término "planta" se define como un organismo unicelular o multicelular capaz de realizar fotosíntesis. Éstos incluyen las algas procariotas y eucariotas (incluyendo algas cianofitas y azules-verdes), protistas eucariotas fotosintéticos, organismos fotosintéticos multicelulares no vasculares y vasculares, incluyendo angiospermas (monocotiledóneas y dicotiledóneas), gimnospermas, plantas con esporas o que se reproducen de manera vegetativa. También se incluyen hongos unicelulares y multicelulares.

La producción de una planta transgénica puede conseguirse modificando un elemento transponible aislado del tipo descrito en la presente memoria para incluir la secuencia de ácido nucleico de interés flanqueada por el extremo terminal del elemento transponible aislado. El elemento transponible modificado puede introducirse en una célula de planta en presencia de una proteína transposasa o una secuencia de ácido nucleico que codifica una transposasa o un virus que codifica una proteína transposasa (p. ej., plásmido auxiliar) usando técnicas muy conocidas en la técnica. Las técnicas ejemplares se discuten con detalle en Gelvin et al., "Plant Molecular Biology Manual", 2ª Ed., Kluwer Academic Publishers, Boston (1995). La transposasa (junto con la proteína de direccionamiento al ADN como se describe en la presente memoria) cataliza la transposición del elemento transponible modificado que contiene la secuencia de ácido nucleico de interés en el ADN genómico de la planta. La presente invención incrementa, por lo tanto, la eficacia de la integración.

Por ejemplo, para pastos como el maíz, los elementos del método basado en transposón pueden introducirse en una célula usando, por ejemplo, bombardeo con microproyectiles (véase p. ej. Sanford J.C. et al. Pat. EEUU No. 5.100.792

(1992). En este método, los elementos de las composiciones basadas en transposón se usan para recubrir partículas pequeñas que se introducen en el tejido diana (células) a través de penetración balística de alta velocidad. Las células transformadas se cultivan en condiciones apropiadas para la regeneración de las plantas, lo que resulta en la producción de plantas transgénicas. Las plantas transgénicas que portan una secuencia de ácido nucleico de interés se examinan para el fenotipo deseado usando varios métodos incluyendo, pero sin limitarse a, un marcador fenotípico apropiado, tal como resistencia a antibiótico o resistencia a herbicida, u observación visual del momento de la inducción floral comparado con las plantas naturales.

Además, el sistema de transferencia génica de esta invención puede usarse como parte de un proceso para trabajar con o para cribar una biblioteca de secuencias recombinantes, por ejemplo, para evaluar la función de las secuencias o para cribar la expresión de proteínas, o para evaluar el efecto de una proteína particular o una región de control de la expresión particular en un tipo de célula particular. En este ejemplo, puede incorporarse una biblioteca de secuencias recombinantes, tal como el producto de una biblioteca combinatoria o el producto de intercambio de genes, ambas técnicas ahora conocidas en la técnica, en el fragmento de ácido nucleico de esta invención para producir una biblioteca de fragmentos de ácido nucleico con distintas secuencias de ácido nucleico situadas entre secuencias repetidas invertidas constantes.

Una ventaja de este sistema es que no está limitado en un grado significativo por el tamaño de la secuencia de ácido nucleico implicada situada entre las repeticiones invertidas. Por ejemplo, la proteína SB se ha usado para incorporar transposones que varían de 1,3 kilobases (kb) hasta aproximadamente 5,0 kb y la transposasa mariner ha movilizado transposones de hasta aproximadamente 13 kb. No existe un límite conocido para el tamaño de la secuencia de ácido nucleico que puede incorporarse en el ADN de una célula usando la proteína SB.

El método de vectores basados en transposones tiene varias ventajas sobre las técnicas de recombinación usadas actualmente tales como el sistema Cre/LoxP. Por ejemplo, la introducción de secuencias de ácidos nucleicos de interés se lleva a cabo directamente por el transposón Minos. No se requieren componentes adicionales, tales como sitios diana. Además, usando el presente método, una única copia de una secuencia de ácido nucleico de interés puede integrarse y escindirse de manera precisa del material genético de una célula en cada etapa de integración.

Esta invención tiene ventajas significativas sobre los vectores actuales basados en transposones para la integración dirigida (véase por ejemplo, la Patente EEUU #5.958.775 Inventor: E. Wickstrom y Stephen Cleaver; Wickstrom E, et al. Gene (2000) 254:37-44), que describe los usos y limitaciones del sitio attTn7 o de secuencia similar que puede ser o no lo suficientemente similar en determinadas especies. La presente invención permite el potencial de incrementar la eficacia de integración selectiva de sitio mediante la inserción de secuencias semejantes a las del anfitrión como se describe en la presente memoria. Además, esta invención podría usarse para eludir el sitio o los sitios diana normales de la transposasa Tn7 mediante la sustitución de su factor de direccionamiento al ADN del anfitrión por otro. También, esta invención permite el potencial de utilizar la proteína de direccionamiento de Tn7 (es decir, TnsD) en un sistema más sencillo y más eficaz, p. ej. haciendo una transposasa quimérica Tn5-TnsD por los métodos recombinantes descritos en la presente memoria.

Lo que también ha constituido una limitación en el uso de terapias basadas en transposones es el método por el que el sistema de transferencia génica de esta invención se introduce en las células. Las estrategias mediadas por virus han limitado la longitud de la secuencia de ácido nucleico situada entre las repeticiones invertidas, según esta invención. Por el contrario, para el presente método basado en transposón no viral se usa la microinyección y existe una restricción muy pequeña respecto al tamaño de la secuencia implicada del fragmento de ácido nucleico de esta invención. De manera similar, las estrategias mediadas por lípidos descritas en la presente memoria para administrar los presentes ácidos nucleicos no tienen limitaciones sustanciales respecto al tamaño.

Existen varias combinaciones potenciales de mecanismos de administración para la parte de transposón que contiene el transgén de interés flanqueado por las repeticiones terminales invertidas (IR) y el gen que codifica la transposasa. Por ejemplo, tanto el transposón como el gen de la transposasa quimérica pueden estar contenidos conjuntamente en el mismo genoma viral recombinante (o plásmido); una única infección administra ambas partes del presente sistema de transposón de manera que la expresión de la transposasa dirige la escisión del transposón del genoma viral recombinante para la integración posterior en un cromosoma celular. En otro ejemplo, la transposasa quimérica y el transposón pueden administrarse independientemente por una combinación de sistemas virales y/o no virales tales como reactivos que contienen lípidos. En estos casos, el transposón quimérico y/o el gen de la transposasa pueden administrarse por un virus recombinante. En cada caso, el gen de la transposasa expresado dirige la liberación del transposón de su ADN transportador (genoma viral) para la integración específica de sitio en el ADN cromosómico.

Esta invención también se refiere a composiciones para usarse en el sistema de transferencia génica de esta invención. Así, la invención se refiere a la introducción de un fragmento de ácido nucleico que comprende una secuencia de ácido nucleico situada entre al menos dos repeticiones invertidas en una célula. En una realización preferida, la incorporación

5 eficaz del fragmento de ácido nucleico en el ADN de una célula ocurre cuando la célula también contiene una transposasa quimérica como se describe en la presente memoria. Como se ha discutido anteriormente, la transposasa quimérica puede proporcionarse a la célula como una transposasa quimérica o como ácido nucleico que codifica la transposasa quimérica. El ácido nucleico que codifica la transposasa quimérica puede tener la forma de ARN o ADN. La proteína puede introducirse en la célula sola o en un vector, tal como un plásmido o un vector viral. Además, el ácido nucleico que codifica la proteína transposasa quimérica puede incorporarse de manera estable o temporal en el genoma de la célula para facilitar la expresión temporal o prolongada de la transposasa quimérica en la célula. Además, pueden unirse de manera operativa promotores u otras regiones de control de la expresión al ácido nucleico que codifica la transposasa quimérica para regular la expresión de la proteína de una manera cuantitativa o específica de tejido. Muchas transposasas tienen una señal de localización nuclear (NLS). La NLS se requiere para el transporte al núcleo después de la traducción en el citosol en aquellas células que no se están dividiendo. Por ejemplo, la proteína SB contiene un dominio de unión al ADN, un dominio catalítico (que tiene actividad transposasa) y una señal NLS.

15 La construcción de ácido nucleico de esta invención se introduce en una o más células usando cualquiera de las distintas técnicas conocidas en la técnica tales como, pero sin limitarse a, microinyección, combinando el fragmento de ácido nucleico con vesículas lipídicas, tales como vesículas lipídicas catiónicas, bombardeo de partículas, electroporación, reactivos de condensación de ADN (p. ej., fosfato de calcio, polilisina o polietilimina) o incorporando el fragmento de ácido nucleico en un vector viral y poniendo en contacto el vector viral con la célula. Cuando se usa un vector viral, el vector viral puede incluir cualquiera de los distintos vectores virales conocidos en la técnica incluyendo vectores virales seleccionados del grupo que consiste en un vector retroviral, un vector de adenovirus o un vector viral adeno-asociado.

20 Se proporcionan vectores obtenidos del elemento P que incluyen al menos las secuencias de inserción del elemento P de *Drosophila* reconocidas por la transposasa del elemento P. Como tal, esta invención incluye un par del dominio de repetición invertida de 31 pares de bases del elemento P o el equivalente funcional de éste, es decir, un dominio reconocido por la transposasa quimérica codificada por el elemento P. La repetición invertida de 31 pares de bases se describe en Beall et al., "Drosophila P-element transposase is a novel site-specific endonuclease", *Genes Dev* (15 ago, 1997)11(16):2137-51. La secuencia de aminoácidos de la transposasa del elemento P se describe en Rio et al., *Cell* (17 ene, 1986 44:21-32).

30 Las Figuras 4-6 son esquemas de construcciones de ácido nucleico lineales en vehículos de administración no virales. Nota: el empaquetamiento de la construcción de ácido nucleico no viral como se ha explicado puede aplicarse obviamente a una construcción de plásmido. Los sistemas de empaquetamiento no virales (p. ej., basados en lípidos, basados en polímeros, basados en lípidos-polímeros y polilisina, entre otros) son muy conocidos para los expertos en el campo de la administración transgénica no viral. Son muy conocidas técnicas adicionales para aumentar la administración en el núcleo y se han empleado en vectores no virales. Los métodos para ensamblar in vitro un complejo transposón-transposasa se han descrito en la bibliografía (Lamberg, A, et al. (2002) Efficient insertion mutagenesis strategy for bacterial genomes involving electroporation of in vitro-assembled DNA transposition complexes of bacteriophage Mu. *Applied and Environmental Microbiology*).

35 Los ejemplos de ligandos específicos para el direccionamiento celular en el empaquetamiento de la invención son muy conocidos en la técnica. Las referencias siguientes son por sus enseñanzas sobre ligandos específicos: (1) Lestina, B.J., Sagnella, S.M., Xu, Z., Shive, M.S., Richter, N.J., Jayaseharan, J., Case, A.J., Kottke-Marchant, K., Anderson, J.M. y Marchant, R.E. (2002) Surface modification of liposomes for selective cell targeting in cardiovascular drug delivery. *J. Control Release* 78:35-247. (2) Moreira, J.N., Gaspar, R., y Allen, T.M. (2001) Targeting stealth liposomes in a murine model of human small cell lung cancer. *Biochim. Biophys. Acta.* 1515:167-176; (3) Xu, L., Tang, W.H., Huang, C.C., Alexander, W., Xiang, L.M., Pirollo, K.F., Rait, A., y Chang, E.H. (2001) Systemic p53 gene therapy of cancer with immunolipoplexes targeted by anti-transferrin receptor scFv. *Mol. Med.* 7:723-734; (4) Sudhan Shaik, M., Kanikkanan, N., y Singh, M. (2001) Conjugation of anti-My9 antibody to stealth monensin liposomes and the effect of conjugated liposomes on the cytotoxicity of immunotoxin. *J. Control Release* 76:285-295; (5) Li, X., Stuckert, P., Bosch, I., Marks, J.D., y Marasco, W.A. (2001) Single-chain antibody-mediated gene delivery into ErB2-positive human breast cancer cells. *Cancer Gene Ther.* 8:555-565; (6) Park, J.W., Kirpotin, D.B., Hong, K., Shalaby, R., Shao, Y., Nielsen, U.B., Marks, J.D., Papahadjopoulos, D., y Benz, C.C. (2001) Tumor targeting using anti-her2 immunoliposomes. *J. Control Release* 74:95-113.

50 Los ejemplos de factores de disrupción endosomal que se usan en el presente empaquetamiento del vector son muy conocidos en la técnica. Las referencias siguientes son por sus enseñanzas sobre factores de disrupción endosomal: (1) Farhood, H., Gao, X., Son, K., Yang, Y.Y., Lazo, J.S., Huang, L., Barsoum, J., Bottega, R., y Epanand, R.M. (1994) Cationic liposomes for direct gene transfer in therapy of cancer and other diseases. *Ann. NY Acad. Sci.* 716:23-25; (2) Tachibana R, Harashima H, Shono M, Azumano M, Niwa M, Futaki S, y Kiwada H. (1998) Intracellular regulation of macromolecules using pH-sensitive liposomes and nuclear localization signal: qualitative and quantitative evaluation of intracellular trafficking. *Biochem. Biophys. Res. Commun.* 251:538-544; (3) E1 Ouahabi A, Thiry M, Pector V, Fuks R, Ruyschaert

JM, y Vandenbranden M. (1997) The role of endosome destabilization activity in the gene transfer process mediated by cationic lipids. *FEBS Lett* 414:187-192.

5 Los factores de localización nuclear para usarse en la administración de las presentes construcciones de ácidos nucleicos son conocidos en la técnica. Las referencias siguientes son por sus enseñanzas sobre factores de localización nuclear: (1) Subramanian A, Ranganathan P, y Diamond SL (1999) Nuclear targeting peptide scaffolds for lipofection of nondividing mammalian cells. *Nat Biotechnol* 17:873-877; (2) Tachibana R, Harashima H, Shono M, Azumano M, Niwa M, Futaki S, y Kiwada H. (1998) Intracellular regulation of macromolecules using pH-sensitive liposomes and nuclear localization signal: qualitative and quantitative evaluation of intracellular trafficking. *Biochem. Biophys. Res. Commun.* 251:538-544; (3) Aronson AI y Hughes JA. (1998) Nuclear localization signal peptides enhance cationic liposome-mediated gene transfer. *J Drug Target* 5:163-169; (4) Boehm U, Heinlein M, Behrens U, y Kunze R. (1995) One of three nuclear localization signals of maize Activator (Ac) transposase overlaps the DNA-binding domain. *Plant J* 7:441-451.

10 También se describen composiciones de la invención, en las que la secuencia de ácido nucleico que codifica la enzima de integración quimérica está localizada fuera de las repeticiones terminales.

15 También se describen composiciones de la invención, en las que el ácido nucleico que codifica el transgén y el ácido nucleico que codifica la enzima de integración quimérica son el mismo ácido nucleico.

También se describen composiciones, en las que el transgén y la enzima de integración están codificados en ácidos nucleicos independientes.

También se describen composiciones, que comprenden además una secuencia homóloga que es homóloga al ADN del anfitrión.

20 También se describen composiciones, en las que la secuencia homóloga está localizada fuera de las repeticiones terminales.

También se describen composiciones, que comprenden además una secuencia de unión a proteína y un ácido nucleico independiente que codifica dos dominios de unión al ADN.

25 También se describen composiciones, que comprenden además una secuencia de unión a proteína y un ácido nucleico independiente que codifica un dominio de unión al ADN y un dominio de unión a proteínas.

También se describen composiciones de la invención, en las que el ácido nucleico presente en el vector no viral es al menos una proteína funcional.

30 También se describen composiciones de la invención, en las que el transgén codifica una molécula biológicamente activa. El transgén puede codificar múltiples y diferentes moléculas biológicamente activas. El transgén puede seleccionarse al menos del grupo que consiste en genes informadores (p. ej., luciferasa, cloranfenicol-acetil transferasa, GFP), oncogenes (p. ej., ras y c-myc) y antioncogenes (p. ej., p53 y retinoblastoma). Se están ensayando varios genes diferentes para terapia génica incluyendo CFTR para fibrosis quística, adenosina desaminasa (ADA) para trastornos inmunes, factor IX, factor VIII e interleuquina-2 (IL-2) para enfermedades de las células sanguíneas, alfa-1-antitripsina para enfermedades pulmonares y factor de necrosis tumoral, endostatina, cotransportador sodio/yodo, angiostatina y resistencia a múltiples fármacos (MDR) para terapias de cáncer. Otros ejemplos de genes incluyen, p. ej., bax, bak, E2F-1, BRCA-1, BRCA-2, bak, ras, p21, CDKN2A, pHyde, ligando FAS, ligando inductor de la apoptosis relacionado con TNF, DOC-2, E-cadherina, caspasas, clusterina, ATM, factor estimulante de las colonias de granulocitos y macrófagos, B7, factor alfa de necrosis tumoral, interleuquina 12, interleuquina 15, interferón gamma, interferón beta, MUC-1, PSA, WT1, WT2, myc, MDM2, DCC, VEGFB, VEGFC, VWF, NEFL, NEF3, TUBB, MAPT, SGNE1, RTN1, GAD1, PYGM, AMPD1, TNNT3, TNNT2, ACTC, MYH7, SFTPB, TPO, NGF, conexina 43.

35 Los compuestos descritos en la presente memoria también pueden usarse para el tratamiento de condiciones precancerígenas tales como displasias cervicales y anales, otras displasias, displasias graves, hiperplasias, hiperplasias atípicas y neoplasias.

45 También se describen construcciones de ácido nucleico de la invención, en las que el transgén es un antígeno de un virus. El antígeno viral puede seleccionarse del grupo que consiste en virus Herpes simple tipo 1, virus Herpes simple tipo 2, Citomegalovirus, virus Epstein-Barr, virus de la Varicela-zoster, Virus del herpes humano 6, Virus del herpes humano 7, Virus del herpes humano 8, Virus de la viruela, Virus de la estomatitis vesicular, Virus de la hepatitis A, Virus de la hepatitis B, Virus de la hepatitis C, Virus de la hepatitis D, Virus de la hepatitis E, Rinovirus, Coronavirus, Virus de la influenza A, Virus de la influenza B, Virus del sarampión, Poliomavirus, Papilomavirus humano, Virus respiratorio sincitial, Adenovirus, Virus coxsackie, Virus del dengue, Virus de las paperas, Poliovirus, Virus de la rabia, Virus del sarcoma de Rous, Virus de la fiebre amarilla, Virus de Ébola, Virus de Marburg, Virus de la fiebre Lassa, Virus de la

50

- 5 Encefalitis Equina del Este, Virus de la Encefalitis Japonesa, Virus de la Encefalitis de St. Louis, Virus de la fiebre del Valle de Murray, Virus del Nilo Oeste, Virus de la fiebre del Valle Rift, Rotavirus A, Rotavirus B, Rotavirus C, Virus Sindbis, Virus de la Inmunodeficiencia de Simios, Virus tipo 1 de la Leucemia de célula T Humana, Hantavirus, Virus de la Rubeola, Virus de la Inmunodeficiencia de Simios, Virus tipo 1 de la Inmunodeficiencia Humana y Virus tipo 2 de la Inmunodeficiencia Humana.
- 10 También se describen construcciones de ácido nucleico de la invención, en las que el transgén es un antígeno de una bacteria. El antígeno bacteriano puede seleccionarse del grupo que consiste en *M. tuberculosis*, *M. bovis*, *M. bovis* cepa BCG, subcepas BCG, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subespecie *paratuberculosis*, *Nocardia asteroides*, otras especies de *Nocardia*, *Legionella pneumophila*, otras especies de *Legionella*, *Salmonella typhi*, otras especies de *Salmonella*, especies de *Shigella*, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, otras especies de *Pasteurella*, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, otras especies de *Brucella*, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetii*, otras especies de *Rickettsia*, especies de *Ehrlichia*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, especies de *Campylobacter*, *Neisseria meningitidis*, *Neisseria gonorrhoea*, *Pseudomonas aeruginosa*, otras especies de *Pseudomonas*, *Haemophilus influenzae*, *Haemophilus ducreyi*, otras especies de *Haemophilus*, *Clostridium tetani*, otras especies de *Clostridium*, *Yersinia enterocolitica* y otras especies de *Yersinia*.
- 15 También se describen construcciones de ácido nucleico de la invención, en las que el transgén es un antígeno de un parásito. El antígeno de parásito puede seleccionarse del grupo que consiste en *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, otras especies de *Plasmodium*, *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, otras especies de *Leishmania*, *Schistosoma mansoni*, otras especies de *Schistosoma* y *Entamoeba histolytica*.
- 20 También se describen construcciones de ácido nucleico de la invención, en las que el transgén es un antígeno tumoral. El antígeno tumoral puede seleccionarse de la lista que consiste en mucina de células epiteliales humana (Muc-1; una repetición en el núcleo de 20 aminoácidos para la glicoproteína Muc-1, presente en las células de cáncer de mama y las células de cáncer pancreático), el producto del oncogén Ha-ras, p53, antígeno carcino-embriionario (CEA), el producto del oncogén raf, gp100/pmel17, GD2, GD3, GM2, TF, sTn, MAGE-1, MAGE-3, BAGE, GAGE, tirosinasa, gp75, Melan-A/Mart-1, gp100, HER2/neu, EBV-LMP 1 y 2, HPV-F4, 6, 7, antígeno específico de la próstata (PSA), HPV-16, MUM, alfa-fetoproteína (AFP), CO17 1A, GA733, gp72, p53, el producto del oncogén ras, HPV E7, antígeno del tumor de Wilms 1, telomerasa y gangliósidos de melanoma.
- 25 Se describen componentes para usarse para preparar las composiciones descritas así como las composiciones en sí mismas para usarse en los métodos descritos en la presente memoria. Éstos y otros materiales se describen en la presente memoria y se entiende que cuando se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales que aunque no se describa explícitamente una referencia específica de cada una de las distintas combinaciones y permutaciones individuales y colectivas de estos compuestos, cada una se contempla específicamente y se describe en la presente memoria. Por ejemplo, si se describe y discute una transposasa quimérica particular y se discuten las distintas modificaciones que pueden hacerse a varias moléculas incluyendo la transposasa quimérica, se contempla específicamente cada una y todas las combinaciones y permutaciones de la transposasa quimérica y las modificaciones que son posibles a no ser que se indique específicamente lo contrario. Así, si se describe una clase de moléculas A, B y C así como una clase de moléculas D, E y F y se describe un ejemplo de una molécula de combinación, A-D, entonces incluso si cada una no se indica individualmente cada una se contempla individualmente y colectivamente lo que significa que se consideran descritas las combinaciones A-E, A-F, B-D, B-E, B-F, C-D, C-E y C-F. De la misma manera, también se describe cualquier subconjunto o combinación de éstas. Así, por ejemplo, el subgrupo de A-E, B-F y C-E se considerará descrito. Este concepto se aplica a todos los aspectos de esta solicitud incluyendo, pero sin limitarse a, las etapas de los métodos para preparar y usar las composiciones descritas. Así, si existen varias etapas adicionales que pueden llevarse a cabo, se entiende que cada una de estas etapas adicionales puede llevarse a cabo con cualquier realización específica o combinación de realizaciones de los métodos descritos.
- 30 Se describen componentes para usarse para preparar las composiciones descritas así como las composiciones en sí mismas para usarse en los métodos descritos en la presente memoria. Éstos y otros materiales se describen en la presente memoria y se entiende que cuando se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales que aunque no se describa explícitamente una referencia específica de cada una de las distintas combinaciones y permutaciones individuales y colectivas de estos compuestos, cada una se contempla específicamente y se describe en la presente memoria. Por ejemplo, si se describe y discute una transposasa quimérica particular y se discuten las distintas modificaciones que pueden hacerse a varias moléculas incluyendo la transposasa quimérica, se contempla específicamente cada una y todas las combinaciones y permutaciones de la transposasa quimérica y las modificaciones que son posibles a no ser que se indique específicamente lo contrario. Así, si se describe una clase de moléculas A, B y C así como una clase de moléculas D, E y F y se describe un ejemplo de una molécula de combinación, A-D, entonces incluso si cada una no se indica individualmente cada una se contempla individualmente y colectivamente lo que significa que se consideran descritas las combinaciones A-E, A-F, B-D, B-E, B-F, C-D, C-E y C-F. De la misma manera, también se describe cualquier subconjunto o combinación de éstas. Así, por ejemplo, el subgrupo de A-E, B-F y C-E se considerará descrito. Este concepto se aplica a todos los aspectos de esta solicitud incluyendo, pero sin limitarse a, las etapas de los métodos para preparar y usar las composiciones descritas. Así, si existen varias etapas adicionales que pueden llevarse a cabo, se entiende que cada una de estas etapas adicionales puede llevarse a cabo con cualquier realización específica o combinación de realizaciones de los métodos descritos.
- 35 Se describen componentes para usarse para preparar las composiciones descritas así como las composiciones en sí mismas para usarse en los métodos descritos en la presente memoria. Éstos y otros materiales se describen en la presente memoria y se entiende que cuando se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales que aunque no se describa explícitamente una referencia específica de cada una de las distintas combinaciones y permutaciones individuales y colectivas de estos compuestos, cada una se contempla específicamente y se describe en la presente memoria. Por ejemplo, si se describe y discute una transposasa quimérica particular y se discuten las distintas modificaciones que pueden hacerse a varias moléculas incluyendo la transposasa quimérica, se contempla específicamente cada una y todas las combinaciones y permutaciones de la transposasa quimérica y las modificaciones que son posibles a no ser que se indique específicamente lo contrario. Así, si se describe una clase de moléculas A, B y C así como una clase de moléculas D, E y F y se describe un ejemplo de una molécula de combinación, A-D, entonces incluso si cada una no se indica individualmente cada una se contempla individualmente y colectivamente lo que significa que se consideran descritas las combinaciones A-E, A-F, B-D, B-E, B-F, C-D, C-E y C-F. De la misma manera, también se describe cualquier subconjunto o combinación de éstas. Así, por ejemplo, el subgrupo de A-E, B-F y C-E se considerará descrito. Este concepto se aplica a todos los aspectos de esta solicitud incluyendo, pero sin limitarse a, las etapas de los métodos para preparar y usar las composiciones descritas. Así, si existen varias etapas adicionales que pueden llevarse a cabo, se entiende que cada una de estas etapas adicionales puede llevarse a cabo con cualquier realización específica o combinación de realizaciones de los métodos descritos.
- 40 Se describen componentes para usarse para preparar las composiciones descritas así como las composiciones en sí mismas para usarse en los métodos descritos en la presente memoria. Éstos y otros materiales se describen en la presente memoria y se entiende que cuando se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales que aunque no se describa explícitamente una referencia específica de cada una de las distintas combinaciones y permutaciones individuales y colectivas de estos compuestos, cada una se contempla específicamente y se describe en la presente memoria. Por ejemplo, si se describe y discute una transposasa quimérica particular y se discuten las distintas modificaciones que pueden hacerse a varias moléculas incluyendo la transposasa quimérica, se contempla específicamente cada una y todas las combinaciones y permutaciones de la transposasa quimérica y las modificaciones que son posibles a no ser que se indique específicamente lo contrario. Así, si se describe una clase de moléculas A, B y C así como una clase de moléculas D, E y F y se describe un ejemplo de una molécula de combinación, A-D, entonces incluso si cada una no se indica individualmente cada una se contempla individualmente y colectivamente lo que significa que se consideran descritas las combinaciones A-E, A-F, B-D, B-E, B-F, C-D, C-E y C-F. De la misma manera, también se describe cualquier subconjunto o combinación de éstas. Así, por ejemplo, el subgrupo de A-E, B-F y C-E se considerará descrito. Este concepto se aplica a todos los aspectos de esta solicitud incluyendo, pero sin limitarse a, las etapas de los métodos para preparar y usar las composiciones descritas. Así, si existen varias etapas adicionales que pueden llevarse a cabo, se entiende que cada una de estas etapas adicionales puede llevarse a cabo con cualquier realización específica o combinación de realizaciones de los métodos descritos.
- 45 Se describen componentes para usarse para preparar las composiciones descritas así como las composiciones en sí mismas para usarse en los métodos descritos en la presente memoria. Éstos y otros materiales se describen en la presente memoria y se entiende que cuando se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales que aunque no se describa explícitamente una referencia específica de cada una de las distintas combinaciones y permutaciones individuales y colectivas de estos compuestos, cada una se contempla específicamente y se describe en la presente memoria. Por ejemplo, si se describe y discute una transposasa quimérica particular y se discuten las distintas modificaciones que pueden hacerse a varias moléculas incluyendo la transposasa quimérica, se contempla específicamente cada una y todas las combinaciones y permutaciones de la transposasa quimérica y las modificaciones que son posibles a no ser que se indique específicamente lo contrario. Así, si se describe una clase de moléculas A, B y C así como una clase de moléculas D, E y F y se describe un ejemplo de una molécula de combinación, A-D, entonces incluso si cada una no se indica individualmente cada una se contempla individualmente y colectivamente lo que significa que se consideran descritas las combinaciones A-E, A-F, B-D, B-E, B-F, C-D, C-E y C-F. De la misma manera, también se describe cualquier subconjunto o combinación de éstas. Así, por ejemplo, el subgrupo de A-E, B-F y C-E se considerará descrito. Este concepto se aplica a todos los aspectos de esta solicitud incluyendo, pero sin limitarse a, las etapas de los métodos para preparar y usar las composiciones descritas. Así, si existen varias etapas adicionales que pueden llevarse a cabo, se entiende que cada una de estas etapas adicionales puede llevarse a cabo con cualquier realización específica o combinación de realizaciones de los métodos descritos.
- 50 También se describen métodos para acoplar el vector basado en transposón adyacente al ADN del anfitrión, utilizando secuencias repetitivas para la recombinación homóloga para estimular una integración selectiva de sitio eficaz, así como otros métodos selectivos de sitio no virales.
- También se describen métodos que emplean sitio o sitios de reconocimiento en el plásmido que pueden reconocer una proteína endógena (o una proteína recién introducida, p. ej., producida a partir de un gen localizado en el plásmido) que puede dirigir el complejo hacia la proximidad del ADN del anfitrión para la integración selectiva de sitio.

También se describen métodos para incorporar elementos repetitivos (p. ej., secuencias semejantes a Alu) en el plásmido basado en transposón. Se entiende que dichos métodos pueden incrementar el acoplamiento y al mismo tiempo permitir la recombinación homóloga (66-67) o la integración del transgén en el genoma del anfitrión.

5 La incorporación de elementos repetitivos (p. ej., secuencias semejantes a Alu) en el plásmido basado en transposón puede incrementar el acoplamiento y al mismo tiempo permitir la recombinación homóloga o la integración del transgén en el genoma del anfitrión.

También se describen métodos que emplean sitios de reconocimiento en el plásmido que pueden reconocer una proteína endógena (o una proteína recién introducida) que puede dirigir el complejo hacia la proximidad del ADN del anfitrión.

## 10 1. Administración de las composiciones de vector a las células

Existen varias composiciones y métodos que pueden usarse para administrar ácidos nucleicos a células, bien *in vitro* o *in vivo*. Por ejemplo, los ácidos nucleicos pueden administrarse mediante varios sistemas de administración directos tales como electroporación, lipofección, precipitación con fosfato de calcio, plásmidos, cósmidos o mediante la transferencia de material genético en células o vehículos tales como liposomas catiónicos. Los medios apropiados para la transfección, incluyendo los transfectantes químicos, o los métodos físico-mecánicos tales como electroporación y difusión directa del ADN, son descritos, por ejemplo, por Wolff, J.A. et al., *Science*, 247, 1465-1468, (1990); y Wolff, J.A. *Nature*, 352, 815-818, (1991). Dichos métodos son muy conocidos en la técnica y se pueden adaptar fácilmente para usarse con las composiciones y métodos descritos en la presente memoria. En determinados casos, los métodos se modificarán para funcionar específicamente con moléculas de ADN grandes. Además, estos métodos pueden usarse para el direccionamiento hacia determinadas enfermedades y poblaciones de células usando las características de direccionamiento del vehículo.

Las composiciones descritas pueden administrarse a las células diana de varias maneras. Por ejemplo, las composiciones pueden administrarse mediante electroporación o mediante lipofección o mediante precipitación con fosfato de calcio. El mecanismo de administración elegido dependerá en parte del tipo de célula diana y de si la administración ocurre por ejemplo *in vivo* o *in vitro*.

Así, las composiciones pueden comprender, además de los vectores con la construcción de ácido nucleico no virales, por ejemplo, lípidos tales como liposomas, tales como liposomas catiónicos (p. ej., DOTMA, DOPE, DC-colesterol) o liposomas aniónicos. Los liposomas pueden comprender además proteínas para facilitar el direccionamiento a una célula particular, si se desea. La administración de una composición que comprende un compuesto y un liposoma catiónico puede administrarse a la sangre aferente hacia un órgano diana o inhalarse en el tracto respiratorio hacia las células diana del tracto respiratorio. Respecto a los liposomas, véase, p. ej., Brigham et al. *Am. J. Resp. Cell. Mol. Biol.* 1:95-100 (1989); Felgner et al. *Proc. Natl. Acad. Sci USA* 84:7413-7417 (1987); Pat. EEUU No. 4.897.355. Además, el compuesto puede administrarse como un componente de una microcápsula que puede dirigirse hacia tipos celulares específicos, tales como macrófagos, o en el que la difusión del compuesto o administración del compuesto de la microcápsula se diseña para una velocidad o dosificación específica.

En los métodos descritos anteriormente que incluyen la administración y captación del ADN exógeno por las células de un sujeto (es decir, transducción o transfección génica), la administración de las composiciones a las células puede ser a través de varios mecanismos. Como ejemplo, la administración puede ser a través de un liposoma, usando preparaciones de liposoma disponibles comercialmente, tales como LIPOFECTINA, LIPOFECTAMINA (GIBCO-BRL, Inc., Gaithersburg, MD), SUPERFECT (Qiagen, Inc. Hilden, Alemania) y TRANSFECTAM (Promega Biotec, Inc., Madison, WI), así como otros liposomas desarrollados según los procedimientos estándar en la técnica. Además, el ácido nucleico o vector de esta invención puede administrarse *in vivo* por electroporación, tecnología que está disponible en Genetronics, Inc. (San Diego, CA) así como mediante un equipo SONOPORATION (ImaRx Pharmaceutical Corp., Tucson, AZ).

Los materiales pueden estar en disolución, suspensión (por ejemplo, incorporados en micropartículas, liposomas o células). Éstos pueden dirigirse hacia un tipo de célula particular mediante anticuerpos, receptores o ligandos de receptores. Las referencias siguientes son ejemplos del uso de esta tecnología para dirigir proteínas específicas hacia tejido tumoral (Senter, et al., *Bioconjugate Chem.*, 2:447-451, (1991); Bagshawe, K.D., *Br. J. Cancer*, 60:275-281, (1989); Bagshawe, et al., *Br. J. Cancer*, 58:700-703, (1988); Senter, et al., *Bioconjugate Chem.*, 4:3-9, (1993); Battelli, et al., *Cancer Immunol. Immunother.*, 35:421-425, (1992); Pietersz y McKenzie, *Immunol. Reviews*, 129:57-80, (1992); y Roffler, et al., *Biochem. Pharmacol.*, 42:2062-2065, (1991)). Estas técnicas pueden usarse para una variedad de otros tipos de células específicos. Los vehículos tales como liposomas durmientes y otros liposomas conjugados con anticuerpos (incluyendo direccionamiento de fármaco mediado por lípidos hacia carcinoma de colon), direccionamiento hacia ADN mediado por receptor a través de ligandos específicos de célula, direccionamiento hacia tumores dirigido por

5 linfocitos y direccionamiento retroviral terapéutico altamente específico de células de glioma murino *in vivo*. Las referencias siguientes son ejemplos del uso de esta tecnología para dirigir proteínas específicas hacia tejido tumoral (Hughes et al., *Cancer Research*, 49:6214-6220, (1989); y Litzinger y Huang, *Biochimica et Biophysica Acta*, 1104:179-187, (1992)). En general, los receptores están implicados en rutas de endocitosis, bien de forma constitutiva o inducida por ligandos. Estos receptores se agrupan en depresiones recubiertas con clatrina, entran en la célula a través de vesículas recubiertas con clatrina, pasan a través de un endosoma ácido en el que los receptores se separan y bien se reciclan hacia la superficie celular, se almacenan intracelularmente o se degradan en los lisosomas. Las rutas de internalización cumplen varias funciones, tales como captación de nutrientes, eliminación de proteínas activadas, aclaramiento de macromoléculas, entrada oportunista de virus y toxinas, disociación y degradación del ligando y regulación del nivel del receptor. Muchos receptores siguen más de una ruta intracelular, dependiendo del tipo celular, concentración del receptor, tipo de ligando, valencia del ligando y concentración del ligando. Los mecanismos moleculares y celulares de la endocitosis mediada por receptor se han revisado (Brown y Greene, *DNA and Cell Biology*, 10:6, 399-409 (1991)).

15 Los ácidos nucleicos que se administran a las células que se quieren integrar en el genoma de la célula anfitriona, contienen típicamente secuencias de integración. Estas secuencias son frecuentemente secuencias relacionadas con virus, particularmente cuando se usan sistemas basados en virus. Estos sistemas de integración virales también pueden incorporarse en ácidos nucleicos que se quieren administrar usando un sistema de administración no basado en ácido nucleico, tal como un liposoma, de manera que el ácido nucleico contenido en el sistema de administración pueda integrarse en el genoma del anfitrión.

20 Otras técnicas generales para la integración en el genoma del anfitrión incluyen, por ejemplo, sistemas diseñados para estimular la recombinación homóloga con el genoma del anfitrión. Estos sistemas se basan típicamente en la secuencia que flanquea el ácido nucleico que se quiere expresar que tiene una homología suficiente con una secuencia diana en el genoma de la célula del anfitrión como para que tenga lugar la recombinación entre el ácido nucleico del vector y el ácido nucleico diana, causando que el ácido nucleico administrado se integre en el genoma del anfitrión. Estos sistemas y los métodos necesarios para estimular la recombinación homóloga son conocidos para los expertos en la técnica.

25 Los 3 requerimientos para la administración selectiva de célula eficaz de un vector en el núcleo de una célula son un ligando (o receptor) para el direccionamiento celular selectivo, un factor de disrupción endosomal si el vector es captado a través de endocitosis mediada por receptor y una señal de localización nuclear. Éstos se han empleado en terapia génica y los métodos para la construcción e implementación son muy conocidos en la bibliografía.

30 Las modificaciones de la superficie de los liposomas para el direccionamiento celular selectivo se han descrito con detalle y se han empleado con éxito (Lestini, B.J., et al. (2002) Surface modification of liposomes for selective cell targeting in cardiovascular drug delivery. *J. Control Release* 78:235-247; Moreira, J.N., et al. (2001) Targeting stealth liposomes in a murine model of human small cell lung cancer. *Biochim. Biophys. Acta.* 1515:167-176; Xu, L., et al. (2001) Systemic p53 gene therapy of cancer with immunolipoplexes targeted by anti-transferrin receptor scFv. *Mol. Med.* 7:723-734. Sudhan Shaik, M., et al. (2001) Conjugation of anti-My9 antibody to stealth monensin liposomes and the effect of conjugated liposomes on the cytotoxicity of immunotoxin. *J. Control Release* 76:285-295; Li, X., et al. (2001) Single-chain antibody-mediated gene delivery into ErbB2-positive human breast cancer cells. *Cancer Gene Ther.* 8:555-565; Park, J.W., et al. (2001) Tumor targeting using anti-her2 immunoliposomes. *J. Control Release* 74:95-113). Por ejemplo, podría utilizarse un inmunolipoplex catiónico que incorpora un receptor anti-transferrina etiquetado biosintéticamente con lípidos como describen Xu y colaboradores.

45 Los factores de disrupción endosomal se han empleado en lípidos catiónicos y son muy conocidos para los expertos en la técnica (Tachibana R, et al. (1998) Intracellular regulation of macromolecules using pH-sensitive liposomes and nuclear localization signal: qualitative and quantitative evaluation of intracellular trafficking. *Biochem. Biophys. Res. Commun.* 251:538-544; E1 Ouahabi A, et al. (1997) The role of endosome destabilization activity in the gene transfer process mediated by cationic lipids. *FEBS Lett* 414:187-192). Por ejemplo, Tachibana y colaboradores utilizaron liposomas sensibles al pH con el fin de conseguir la disrupción endosomal y la liberación posterior en el citosol.

50 Los factores de localización nuclear también pueden incorporarse como se representa con un diagrama en el esquema (Figura 5 y 6) (Subramanian A, et al. (1999) Nuclear targeting peptide scaffolds for lipofection of nondividing mammalian cells. *Nat Biotechnol* 17:873-877; Aronsohn AI, et al. (1998) Nuclear localization signal peptides enhance cationic liposome-mediated gene transfer. *J Drug Target* 5:163-169; Boehm U, et al. (1995) One of three nuclear localization signals of maize Activator (Ac) transposase overlaps the DNA-binding domain. *Plant J* 7:441-451). Por ejemplo, Aronsohn y colaboradores construyeron un vehículo de administración no viral que consiste en un conglomerado de un péptido sintético de localización nuclear obtenido del virus SV40, un plásmido PGL3 que codifica una luciferasa y un liposoma DOTAP:DOPE de lípidos catiónicos.

## 55 2. Sistemas de expresión

5 Los ácidos nucleicos que se administran a las células contienen típicamente sistemas de control de la expresión. Por ejemplo, los genes insertados en sistemas no virales y virales contienen habitualmente promotores y/o amplificadores para ayudar en el control de la expresión del producto génico deseado. Un promotor es generalmente una secuencia o secuencias de ADN que funcionan cuando están en una localización relativamente fija respecto al sitio de comienzo de la transcripción. Un promotor contiene elementos centrales requeridos para la interacción básica de la ARN polimerasa y los factores de transcripción y puede contener elementos en 5' y elementos de respuesta.

#### a) Promotores y Amplificadores

10 Los promotores preferidos que controlan la transcripción desde vectores en células anfitrionas de mamífero pueden obtenerse a partir de varias fuentes, por ejemplo, los genomas de virus tales como: polioma, Virus 40 de Simio (SV40), adenovirus, retrovirus, virus de la hepatitis B y lo más preferiblemente citomegalovirus o a partir de promotores de mamífero heterólogos, p. ej., promotor de la beta actina. Los promotores tempranos y tardíos del virus SV40 se obtienen convenientemente como un fragmento de restricción de SV40 que también contiene el origen de replicación del virus SV40 (Fiers et al., Nature. 273:113 (1978)). El promotor temprano inmediato del citomegalovirus humano se obtiene convenientemente como un fragmento de restricción HindIII E (Greenway, P.J. et al., Gene 18:355-360 (1982)). Por supuesto, también son útiles en la presente memoria los promotores de la célula anfitriona o de especies relacionadas.

20 Amplificador se refiere generalmente a una secuencia de ADN que funciona a una distancia no fijada del sitio de comienzo de la transcripción y puede estar bien en 5' (Laimins, L. et al., Proc. Natl. Acad. Sci. 78:993 (1981)) ó 3' (Lusky, M.L. et al., Mol. Cell Bio. 3:1108 (1983)) respecto a la unidad de transcripción. Además, los amplificadores pueden estar en un intrón (Banerji, J.L. et al. Cell 33:729 (1983)) así como en la secuencia codificadora misma (Osborne, T.F., et al., Mol. Cell Bio. 4:1293 (1984)). Habitualmente tienen una longitud de entre 10 y 300 pb y funcionan en cis. La función de los amplificadores es incrementar la transcripción a partir de promotores cercanos. Frecuentemente, los amplificadores también contienen elementos de respuesta que median la regulación de la transcripción. Los promotores también pueden contener elementos de respuesta que median la regulación de la transcripción. Los amplificadores determinan frecuentemente la regulación de la expresión de un gen. Aunque actualmente se conocen muchas secuencias amplificadoras para genes de mamíferos (globina, elastasa, albúmina, fetoproteína e insulina), se usará típicamente un amplificador de un virus de una célula eucariota para la expresión general. Los ejemplos preferidos son el amplificador de SV40 en el lado tardío del origen de replicación (pb 100-270), el amplificador del promotor temprano de citomegalovirus, el amplificador de polioma en la lado tardío del origen de replicación y los amplificadores de adenovirus.

30 El promotor y/o amplificador puede activarse específicamente bien por luz o por eventos químicos específicos que desencadenan su función. Los sistemas pueden regularse por reactivos tales como tetraciclina y dexametasona. También existen maneras de amplificar la expresión génica del vector viral por exposición a irradiación, tal como irradiación gamma, o fármacos de quimioterapia alquilantes.

35 En determinadas realizaciones, la región promotora y/o amplificadora puede actuar como un promotor y/o amplificador constitutivo para maximizar la expresión de la región de la unidad de transcripción que se va a transcribir. En determinadas construcciones, la región promotora y/o amplificadora es activa en todos los tipos celulares eucariotas, incluso si sólo se expresa en un tipo de célula particular en un momento particular. Un promotor preferido de este tipo es el promotor CMV (650 bases). Otros promotores preferidos son los promotores SV40, citomegalovirus (promotor de longitud completa) y vector retroviral LTF.

40 Se ha mostrado que todos los elementos reguladores específicos pueden clonarse y usarse para construir vectores de expresión que se expresan selectivamente en tipos celulares específicos tales como células de melanoma. El promotor de la proteína fibrilar ácida de la glía (GFAP) se ha usado para expresar selectivamente genes en células de origen glial.

45 Los promotores adecuados para usarse en plantas también son muy conocidos en la técnica. Por ejemplo, los promotores constitutivos para la expresión de genes de plantas incluyen los promotores de la octopina sintasa, nopalina sintasa o manopina sintasa de *Agrobacterium*, el promotor del virus del mosaico de la coliflor (35S), el promotor del virus del mosaico de escrofularia (FMV) y el promotor del virus del mosaico del tabaco (TMV). Los ejemplos específicos de promotores regulados en plantas incluyen los promotores Kin1 y cor6.6 de baja temperatura (Wang, et al., Plant Mol. Biol. 28:605 (1995); Wang, et al., Plant Mol. Biol. 28:619-634 (1995)), el promotor inducible ABA (Marcotte et al., Plant Cell 1:969-976 (1989)), los promotores de choque térmico y el promotor inducible por frío de *B. napus* (White et al., Plant Physiol. 106:917 (1994)).

50 Los vectores de expresión usados en células anfitrionas eucariotas (células de levaduras, hongos, insectos, plantas, animales, humanas o nucleadas) también pueden contener secuencias necesarias para la finalización de la transcripción que pueden influir en la expresión del ARNm. Estas regiones se transcriben como segmentos poliadenilados en la parte no traducida del ARNm que codifica la proteína de factor de tejido. Las regiones 3' no traducidas también incluyen sitios de finalización de la transcripción. Se prefiere que la unidad de transcripción también contenga una región de

5 poliadenilación. Un beneficio de esta región es que incrementa la probabilidad de que la unidad transcrita se procese y transporte como ARNm. La identificación y uso de las señales de poliadenilación en las construcciones de expresión está bien establecida. Se prefiere usar señales de poliadenilación homólogas en las construcciones de transgén. En determinadas unidades de transcripción, la región de poliadenilación se obtiene de la señal de poliadenilación temprana de SV40 y consiste aproximadamente en 400 bases. También se prefiere que las unidades transcritas contengan otras secuencias estándar solas o en combinación con las secuencias anteriores para mejorar la expresión, o la estabilidad, de la construcción.

#### b) Marcadores

10 El vector puede incluir una secuencia de ácido nucleico que codifica un producto marcador. El término "gen marcador", tal y como se usa en la presente memoria, se refiere a una secuencia de ácido nucleico cuyo producto puede ensayarse fácilmente, por ejemplo, colorimétricamente como un producto de una reacción enzimática, tal como el gen lacZ que codifica la beta-galactosidasa. El gen marcador puede unirse de manera operativa a un promotor adecuado que está opcionalmente unido a una secuencia de ácido nucleico de interés de manera que la expresión del gen marcador puede usarse para ensayar la integración del transposón en el genoma de una célula y de esta manera la integración de la secuencia de ácido nucleico de interés en el genoma de la célula. Los ejemplos de moléculas marcadoras ampliamente usadas incluyen enzimas tales como beta-galactosidasa, beta-glucuronidasa, beta-glucosidasa; moléculas luminiscentes tales como la proteína verde fluorescente y la luciferasa de luciérnaga; y marcadores auxótrofos tales como His3p y Ura3p. (Véase, p. ej., Capítulo 9 en Ausubel, F.M., et al. *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., (1998)).

20 En algunas realizaciones, el marcador puede ser un marcador seleccionable. Los ejemplos de marcadores seleccionables adecuados para células de mamífero son dihidrofolato reductasa (DHFR), timidina quinasa, neomicina, análogo de la neomicina G418, hidromicina y puromicina. Cuando dichos marcadores seleccionables se transfieren con éxito a una célula anfitriona de mamífero, la célula anfitriona de mamífero transformada puede sobrevivir si se pone bajo presión selectiva. Existen dos categorías distintas ampliamente usadas de regímenes selectivos. La primera categoría se basa en el metabolismo de una célula y el uso de una línea celular mutante que carece de la capacidad de crecer independientemente de un medio suplementado. Dos ejemplos son: células CHO DHFR y células de ratón LTK. Estas células carecen de la capacidad de crecer sin la adición de nutrientes tales como la timidina o hipoxantina. Como estas células carecen de determinados genes necesarios para una ruta completa de síntesis de nucleótidos, no pueden sobrevivir a menos que se proporcionen los nucleótidos ausentes en un medio suplementado. Una alternativa a suplementar el medio es introducir un gen DHFR o TK intacto en las células que carecen de los genes respectivos, alterando así sus requerimientos de crecimiento. Las células individuales que no se transformaron con el gen DHFR o TK no serán capaces de sobrevivir en medio no suplementado.

35 La segunda categoría es la selección dominante que se refiere a un esquema de selección usado en cualquier tipo celular y no requiere el uso de una línea celular mutante. Estos esquemas usan típicamente un fármaco para parar el crecimiento de una célula anfitriona. Aquellas células que tienen un gen nuevo expresarán una proteína que confiere resistencia al fármaco y sobrevivirán a la selección. Los ejemplos de dicha selección dominante usan los fármacos neomicina, (Southern P. y Berg, P.J., *Molec. Appl. Genet.* 1:327 (1982)), ácido micofenólico, (Mulligan, R.C. y Berg, P. *Science* 209:1422 (1980)) o higromicina, (Sugden, B. et al., *Mol. Cell. Biol.* 5:410-413 (1985)). Los tres ejemplos emplean genes bacterianos bajo control eucariota para conferir resistencia al fármaco apropiado G418 o neomicina (genética), xgpt (ácido micofenólico) o higromicina, respectivamente. Otros incluyen el análogo de neomicina G418 y puramicina.

#### C. Métodos para usar las composiciones

45 El sistema de transposón de esta invención tiene aplicaciones en muchas áreas de la biotecnología. El desarrollo de elementos transponibles para vectores en animales permite lo siguiente: 1) inserción eficaz del material genético en los cromosomas del animal usando los métodos proporcionados en esta solicitud; 2) identificación, aislamiento y caracterización de los genes implicados en el crecimiento y desarrollo mediante el uso de transposones como mutágenos insercionales (p. ej., véase Kaiser et al., 1995, "Eukaryotic transposable elements as tools to study gene structure and function". En *Mobile Genetic Elements*, IRL Press, p. 69-100); 3) identificación aislamiento y caracterización de secuencias reguladoras de la transcripción que controlan el crecimiento y desarrollo; 4) uso de construcciones marcadoras para análisis de loci de rasgos cuantitativos (QTL); y 5) identificación de loci genéticos de rasgos económicamente importantes, además de aquellos para crecimiento y desarrollo, es decir, resistencia a enfermedades (p. ej., Anderson et al., 1996, *Mol. Mar. Biol. Biotech.*, 5, 105-113).

##### 1. Métodos de modificación génica y disrupción génica

Debido a su capacidad inherente para moverse de una localización cromosómica a otra en y entre genomas, los elementos transponibles se han explotado como vectores genéticos para la manipulación genética en varios organismos.

El etiquetado con transposones es una técnica en la que los transposones se movilizan para "saltar" en los genes, inactivándolos de esta manera por mutagénesis insercional. Estos métodos se discuten por Evans et al., TIG 1997 13, 370-374. En el proceso, los genes inactivados se "etiquetan" con el elemento transponible que puede usarse entonces para recuperar el alelo mutado. La capacidad de los proyectos del genoma humano y de otros genomas para adquirir datos de secuencia génica ha superado la capacidad de los científicos para atribuir funciones biológicas a los nuevos genes. Por lo tanto, la presente invención proporciona un método eficaz para introducir una etiqueta en el genoma de una célula. Cuando la etiqueta se inserta en una localización en la célula que interrumpe la expresión de una proteína que está asociada con un fenotipo particular, la expresión de un fenotipo alterado en una célula que contiene el ácido nucleico de esta invención permite la asociación de un fenotipo particular con un gen particular que ha sido interrumpido por el fragmento de ácido nucleico de esta invención. Aquí, el fragmento de ácido nucleico funciona como una etiqueta. Los cebadores diseñados para secuenciar el ADN genómico que flanquea el fragmento de ácido nucleico de esta invención pueden usarse para obtener información de la secuencia del gen interrumpido.

La construcción de ácido nucleico también puede usarse para el descubrimiento de genes. En un ejemplo, el fragmento de ácido nucleico en combinación con la transposasa quimérica o el ácido nucleico que codifica la transposasa quimérica se introduce en una célula. La construcción de ácido nucleico comprende preferiblemente una secuencia de ácido nucleico situada entre al menos dos repeticiones invertidas, en la que las repeticiones invertidas se unen a la proteína transposasa quimérica y en la que el fragmento de ácido nucleico se integra en el ADN de la célula en presencia de la proteína transposasa quimérica. En una realización preferida, la secuencia de ácido nucleico incluye una proteína marcadora, tal como GFP, y un sitio de reconocimiento de una endonucleasa de restricción, preferiblemente una secuencia de reconocimiento de 6 bases. Después de la integración, el ADN de la célula se aísla y se digiere con la endonucleasa de restricción. Cuando se usa una endonucleasa de restricción que emplea una secuencia de reconocimiento de 6 bases, el ADN de la célula se corta en fragmentos de aproximadamente 4.000 pb de media. Estos fragmentos pueden clonarse o pueden añadirse conectores a los extremos de los fragmentos digeridos para proporcionar secuencia complementaria para los cebadores de PCR. Cuando se añaden los conectores, las reacciones de PCR se usan para amplificar los fragmentos usando cebadores de los conectores y cebadores que se unen a las repeticiones directas de las repeticiones invertidas en el fragmento de ácido nucleico. Los fragmentos amplificados se secuencian y el ADN que flanquea las repeticiones directas se usa para buscar en bases de datos informatizadas tales como GenBank.

La invención puede usarse para el etiquetado dirigido a sitio. Por ejemplo, la incorporación de una secuencia génica similar a la del anfitrión (pero no funcional) en un plásmido basado en transposón permite el etiquetado de ese gen como se ha descrito anteriormente. Una aplicación de la invención es determinar la función de una proteína específica. Por ejemplo, puede insertarse ADNc (ARNm transcrito de manera inversa), ADN genómico o híbridos ARN/ADN (quimeraplasto) en un plásmido basado en transposón después de mutagénesis dirigida a sitio de manera que la región codificadora pueda inactivarse. Este ADNc o ADN genómico alterado puede insertarse en un plásmido basado en transposón como se describe en la presente memoria. El vector basado en transposón que contiene la secuencia semejante a la del anfitrión se acopla con el ADN del anfitrión mediante hibridación. La expresión de la transposasa y la integración posterior ocurre en la diana deseada. Otra realización de la invención es preparar una transposasa quimérica sin selectividad de sitio para los propósitos descritos anteriormente. Por ejemplo, si una transposasa dada en una determinada célula no tiene el factor de direccionamiento de ADN para esa célula la eficacia de la integración se reduce de manera importante. Si se proporciona la transposasa con un factor de direccionamiento de ADN requerido la integración se incrementa significativamente lo que resulta en una mejora obvia respecto a la transposasa "convencional".

También se describe un método para movilizar una secuencia de ácido nucleico en una célula. En este método la construcción de ácido nucleico de esta invención se incorpora en el ADN de una célula, como se proporciona en la discusión anterior. Se introduce en la célula una transposasa quimérica adicional o ácido nucleico que codifica la transposasa quimérica y la proteína es capaz de movilizar (es decir, mover) el fragmento de ácido nucleico desde una primera posición en el ADN de la célula hasta una segunda posición en el ADN de la célula. El ADN de la célula puede ser ADN genómico o ADN extracromosómico. El método permite el movimiento del fragmento de ácido nucleico desde una localización en el genoma hasta otra localización en el genoma o, por ejemplo, desde un plásmido en una célula al genoma de esa célula.

Las composiciones y métodos descritos pueden usarse para la disrupción y modificación génica dirigida en cualquier animal que pueda sufrir estos eventos. La modificación génica y la disrupción génica se refieren a métodos, técnicas y composiciones que rodean la eliminación o alteración selectiva de un gen o cadena de cromosomas en un animal, tal como un mamífero, de manera que propaga la modificación a través de la línea germinal del mamífero. En general, una célula se transforma con un vector que se diseña para recombinar de manera homóloga con una región de un cromosoma particular contenido en la célula, como por ejemplo, se describe en la presente memoria. Este evento de recombinación homóloga puede producir un cromosoma que tiene ADN exógeno introducido, por ejemplo en marco, con el ADN circundante. Este tipo de protocolo permite que se introduzcan mutaciones muy específicas, tales como

mutaciones puntuales, en el genoma contenido en la célula. Los métodos para llevar a cabo este tipo de recombinación homóloga se describen en la presente memoria.

5 Una de las características preferidas de llevar a cabo recombinación homóloga en células de mamífero es que las células deben ser capaces de ser cultivadas ya que los eventos de recombinación deseados ocurren con una frecuencia muy baja.

10 Una vez que se produce la célula mediante los métodos descritos en la presente memoria, puede producirse un animal a partir de esta célula mediante tecnología de células madre o tecnología de clonación. Por ejemplo, si la célula en la que se ha transfectado el ácido nucleico fuera una célula madre para el organismo, entonces esta célula, después de la transfección y cultivo, puede usarse para producir un organismo que contendrá la modificación o disrupción génica en células de la línea germinal, que a su vez pueden usarse para producir otro animal que posea la modificación o disrupción génica en todas sus células. En otros métodos para la producción de un animal que contiene la modificación o disrupción génica en todas sus células, pueden usarse tecnologías de clonación. Estas tecnologías toman generalmente el núcleo de la célula transfectada y bien mediante fusión o reemplazamiento fusionan el núcleo transfectado con un oocito que puede manipularse para producir un animal. La ventaja de los procedimientos que usan la clonación en lugar de la tecnología ES es que se pueden transfectar células distintas de las células ES. Por ejemplo, una célula fibroblasto, que es muy fácil de cultivar, puede usarse como la célula que se transfecta y en la que tiene lugar un evento de modificación o disrupción génica y las células obtenidas de esta célula pueden usarse para clonar un animal completo.

20 Para modificar un gen de interés, pueden clonarse ácidos nucleicos en un vector diseñado, por ejemplo, para recombinación homóloga. Este gen podría ser, por ejemplo, una secuencia reguladora heteróloga o sintética de un antioncogén (p. ej., p53 y retinoblastoma). Se ensayan varios genes distintos para terapia génica incluyendo CFTR para fibrosis quística, adenosina desaminasa (ADA) para trastornos inmunes, factor IX, factor VIII e interleuquina-2 (IL-2) para enfermedades de las células sanguíneas, alfa-1-antitripsina para enfermedades pulmonares y factor de necrosis tumoral, endostatina, cotransportador sodio/yodo, angiostatina y resistencia a múltiples fármacos (MDR) para terapias de cáncer. Otros ejemplos de genes incluyen, p. ej., bax, bak, E2F-1, BRCA-1, BRCA-2, bak, ras, p21, CDKN2A, pHyde, ligando FAS, ligando inductor de la apoptosis relacionado con TNF, DOC-2, E-cadherina, caspasas, clusterina, ATM, factor estimulante de las colonias de granulocitos y macrófagos, B7, factor alfa de necrosis tumoral, interleuquina 12, interleuquina 15, interferón gamma, interferón beta, MUC-1, PSA, WT1, WT2, myc, MDM2, DCC, VEGFB, VEGFC, VWF, NEFL, NEF3, TUBB, MAPT, SGNE1, RTN1, GAD1, PYGM, AMPD1, TNNT3, TNNT2, ACTC, MYH7, SFTPB, TPO, NGF, conexina 43.

## 30 **2. Métodos para llevar a cabo la administración génica**

La administración génica se lleva a cabo in vitro (p. ej., por electroporación u otras técnicas muy conocidas en la técnica) o in vivo. Las técnicas in vivo incluyen administración intravenosa, inyección directa en el sitio deseado o por inhalación.

## **3. Métodos para tratar enfermedades**

Se describen métodos para tratar un sujeto con una condición que comprende administrar el vector de la invención.

35 Las composiciones descritas pueden usarse para tratar cualquier enfermedad en la que ocurre la proliferación celular incontrolada como cánceres. Una lista no limitativa de los diferentes tipos de cánceres es como sigue: linfomas (Hodgkins y no Hodgkins), leucemias, carcinomas, carcinomas de tejidos sólidos, carcinomas de células escamosas, adenocarcinomas, sarcomas, gliomas, gliomas de alto grado, blastomas, neuroblastomas, plasmacitomas, histiocitomas, melanomas, adenomas, tumores hipóxicos, mielomas, linfomas o sarcomas relacionados con SIDA, cánceres metastásicos o cánceres en general.

40 Una lista representativa pero no limitativa de los cánceres que se pueden tratar usando las composiciones descritas es como sigue: linfoma, linfoma de células B, linfoma de células T, micosis fungoides, Enfermedad de Hodgkin, leucemia mieloide, cáncer de vejiga, cáncer de cerebro, cáncer del sistema nervioso, cáncer de cabeza y cuello, carcinoma de células escamosas de cabeza y cuello, cáncer de riñón, cánceres de pulmón tales como cáncer de pulmón de células pequeñas y cáncer de pulmón de células no pequeñas, neuroblastoma/glioblastoma, cáncer de ovarios, cáncer pancreático, cáncer de próstata, cáncer de piel, cáncer de hígado, melanoma, carcinomas de células escamosas de la boca, garganta, laringe y pulmón, cáncer de de colon, cáncer cervical, carcinoma cervical, cáncer de mama y cáncer epitelial, cáncer renal, cáncer genitourinario, cáncer pulmonar, carcinoma de esófago, carcinoma de cabeza y cuello, cáncer del intestino delgado, cánceres hematopoyéticos; cáncer testicular; cánceres de colon y rectal, cáncer prostático o cáncer pancreático.

50 También se describen métodos de la invención, en los que la condición es una infección viral. La infección viral puede seleccionarse de la lista de virus que consiste en Virus Herpes simple tipo 1, Virus Herpes simple tipo 2, Citomegalovirus, virus Epstein-Barr, virus de la Varicela-zoster, Virus del herpes humano 6, Virus del herpes humano 7,

5 Virus del herpes humano 8, Virus de la viruela, Virus de la estomatitis vesicular, Virus de la hepatitis A, Virus de la hepatitis B, Virus de la hepatitis C, Virus de la hepatitis D, Virus de la hepatitis E, Rinovirus, Coronavirus, Virus de la influenza A, Virus de la influenza B, Virus del sarampión, Poliomavirus, Papilomavirus humano, Virus respiratorio sincitial, Adenovirus, Virus coxsackie, Virus del dengue, Virus de las paperas, Poliovirus, Virus de la rabia, Virus del sarcoma de Rous, Virus de la fiebre amarilla, Virus de Ébola, Virus de Marburg, Virus de la fiebre Lassa, Virus de la Encefalitis Equina del Este, Virus de la Encefalitis Japonesa, Virus de la Encefalitis de St. Louis, Virus de la fiebre del Valle de Murray, Virus del Nilo Oeste, Virus de la fiebre del Valle Rift, Rotavirus A, Rotavirus B, Rotavirus C, Virus Sindbis, Virus de la Inmunodeficiencia de Simios, Virus tipo 1 de la Leucemia de célula T Humana, Hantavirus, Virus de la Rubeola, Virus de la Inmunodeficiencia de Simios, Virus tipo 1 de la Inmunodeficiencia Humana y Virus tipo 2 de la Inmunodeficiencia Humana.

15 También se describen métodos de la invención, en los que el transgén es un antígeno de un virus. El antígeno viral puede seleccionarse del grupo de virus que consiste en Virus Herpes simple tipo 1, Virus Herpes simple tipo 2, Citomegalovirus, virus Epstein-Barr, virus de la Varicela-zoster, Virus del herpes humano 6, Virus del herpes humano 7, Virus del herpes humano 8, Virus de la viruela, Virus de la estomatitis vesicular, Virus de la hepatitis A, Virus de la hepatitis B, Virus de la hepatitis C, Virus de la hepatitis D, Virus de la hepatitis E, Rinovirus, Coronavirus, Virus de la influenza A, Virus de la influenza B, Virus del sarampión, Poliomavirus, Papilomavirus humano, Virus respiratorio sincitial, Adenovirus, Virus coxsackie, Virus del dengue, Virus de las paperas, Poliovirus, Virus de la rabia, Virus del sarcoma de Rous, Virus de la fiebre amarilla, Virus de Ébola, Virus de Marburg, Virus de la fiebre Lassa, Virus de la Encefalitis Equina del Este, Virus de la Encefalitis Japonesa, Virus de la Encefalitis de St. Louis, Virus de la fiebre del Valle de Murray, Virus del Nilo Oeste, Virus de la fiebre del Valle Rift, Rotavirus A, Rotavirus B, Rotavirus C, Virus Sindbis, Virus de la Inmunodeficiencia de Simios, Virus tipo 1 de la Leucemia de célula T Humana, Hantavirus, Virus de la Rubeola, Virus de la Inmunodeficiencia de Simios, Virus tipo 1 de la Inmunodeficiencia Humana y Virus tipo 2 de la Inmunodeficiencia Humana.

25 También se describen métodos de la invención, en los que la condición es una infección bacteriana. La infección bacteriana puede seleccionarse de la lista de bacterias que consiste en *M. tuberculosis*, *M. bovis*, *M. bovis* cepa BCG, subcepas BCG, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subespecie *paratuberculosis*, *Nocardia asteroides*, otras especies de *Nocardia*, *Legionella pneumophila*, otras especies de *Legionella*, *Salmonella typhi*, otras especies de *Salmonella*, especies de *Shigella*, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, otras especies de *Pasteurella*, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, otras especies de *Brucella*, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, otras especies de *Rickettsia*, especies de *Ehrlichia*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyrogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, especies de *Campylobacter*, *Neisseria meningitidis*, *Neisseria gonorrhoea*, *Pseudomonas aeruginosa*, otras especies de *Pseudomonas*, *Haemophilus influenzae*, *Haemophilus ducreyi*, otras especies de *Haemophilus*, *Clostridium tetani*, otras especies de *Clostridium*, *Yersinia enterocolitica* y otras especies de *Yersinia*.

35 También se describen métodos de la invención, en los que el transgén es un antígeno de una bacteria. El antígeno bacteriano puede seleccionarse del grupo que consiste en *M. tuberculosis*, *M. bovis*, *M. bovis* cepa BCG, subcepas BCG, *M. avium*, *M. intracellulare*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. avium* subespecie *paratuberculosis*, *Nocardia asteroides*, otras especies de *Nocardia*, *Legionella pneumophila*, otras especies de *Legionella*, *Salmonella typhi*, otras especies de *Salmonella*, especies de *Shigella*, *Yersinia pestis*, *Pasteurella haemolytica*, *Pasteurella multocida*, otras especies de *Pasteurella*, *Actinobacillus pleuropneumoniae*, *Listeria monocytogenes*, *Listeria ivanovii*, *Brucella abortus*, otras especies de *Brucella*, *Cowdria ruminantium*, *Chlamydia pneumoniae*, *Chlamydia trachomatis*, *Chlamydia psittaci*, *Coxiella burnetti*, otras especies de *Rickettsia*, especies de *Ehrlichia*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus pyrogenes*, *Streptococcus agalactiae*, *Bacillus anthracis*, *Escherichia coli*, *Vibrio cholerae*, especies de *Campylobacter*, *Neisseria meningitidis*, *Neisseria gonorrhoea*, *Pseudomonas aeruginosa*, otras especies de *Pseudomonas*, *Haemophilus influenzae*, *Haemophilus ducreyi*, otras especies de *Haemophilus*, *Clostridium tetani*, otras especies de *Clostridium*, *Yersinia enterocolitica* y otras especies de *Yersinia*.

40 También se describen métodos de la invención, en los que la condición es una infección por parásitos. La infección por parásitos puede seleccionarse de la lista de parásitos que consiste en *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae*, otras especies de *Plasmodium*, *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, otras especies de *Leishmania*, *Schistosoma mansoni*, otras especies de *Schistosoma* y *Entamoeba histolytica*.

55 También se describen métodos de la invención, en los que el transgén es un antígeno de un parásito. El antígeno de parásito puede seleccionarse del grupo que consiste en *Toxoplasma gondii*, *Plasmodium falciparum*, *Plasmodium vivax*,

*Plasmodium malariae*, otras especies de *Plasmodium*, *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania major*, otras especies de *Leishmania*, *Schistosoma mansoni*, otras especies de *Schistosoma* y *Entamoeba histolytica*.

También se describen métodos de la invención, en los que la condición es cáncer.

5 Los vectores y composiciones que contienen vectores descritos pueden usarse para tratar cualquier enfermedad en la que ocurre proliferación celular incontrolada tal como cánceres. Una lista no limitativa de los diferentes tipos de cánceres es como sigue: linfomas (Hodgkins y no Hodgkins), leucemias, carcinomas, carcinomas de tejidos sólidos, carcinomas de células escamosas, adenocarcinomas, sarcomas, gliomas, gliomas de alto grado, blastomas, neuroblastomas, plasmacitomas, histiocitomas, melanomas, adenomas, tumores hipóxicos, mielomas, linfomas o sarcomas relacionados con SIDA, cánceres metastásicos o cánceres en general.

10 Una lista representativa pero no limitativa de los cánceres que se pueden tratar usando las composiciones descritas es como sigue: linfoma, linfoma de células B, linfoma de células T, micosis fungoides, Enfermedad de Hodgkin, leucemia mieloide, cáncer de vejiga, cáncer de cerebro, cáncer del sistema nervioso, cáncer de cabeza y cuello, carcinoma de células escamosas de cabeza y cuello, cáncer de riñón, cánceres de pulmón tales como cáncer de pulmón de células pequeñas y cáncer de pulmón de células no pequeñas, neuroblastoma/glioblastoma, cáncer de ovarios, cáncer  
15 pancreático, cáncer de próstata, cáncer de piel, cáncer de hígado, melanoma, carcinomas de células escamosas de la boca, garganta, laringe y pulmón, cáncer de de colon, cáncer cervical, carcinoma cervical, cáncer de mama y cáncer epitelial, cáncer renal, cáncer genitourinario, cáncer pulmonar, carcinoma de esófago, carcinoma de cabeza y cuello, cáncer del intestino delgado, cánceres hematopoyéticos; cáncer testicular; cánceres de colon y rectal, cáncer prostático o cáncer pancreático.

20 También se describen métodos de la invención, en los que el transgén es un antígeno tumoral. El antígeno tumoral puede seleccionarse de la lista que consiste en mucina de células epiteliales humana (Muc-1; una repetición en el núcleo de 20 aminoácidos para la glicoproteína Muc-1, presente en las células de cáncer de mama y las células de cáncer pancreático), el producto del oncogén Ha-ras, p53, antígeno carcino-embriionario (CEA), el producto del oncogén raf, gp100/pm17, GD2, GD3, GM2, TF, sTn, MAGE-1, MAGE-3, BAGE, GAGE, tirosinasa, gp75, Melan-A/Mart-1, gp100,  
25 HER2/neu, EBV-LMP 1 y 2, HPV-F4, 6, 7, antígeno específico de la próstata (PSA), HPV-16, MUM, alfa-fetoproteína (AFP), CO17 1A, GA733, gp72, p53, el producto del oncogén ras, HPV E7, antígeno del tumor de Wilms 1, telomerasa y gangliósidos de melanoma.

30 Se describen métodos para tratar una condición en un sujeto que comprende administrar al sujeto el vector de la invención, en el que la condición se debe a un gen mutado, desregulado, interrumpido o delecionado; autoinmunidad; o enfermedades inflamatorias.

35 Se describen métodos para tratar una condición en un sujeto, en el que la condición puede seleccionarse de la lista que consiste en fibrosis quística, asma, esclerosis múltiple, distrofia muscular, diabetes, tay-sachs, espina bifida, anemia de células falciforme, hemocromatosis hereditaria, parálisis cerebral, enfermedad de Parkinson, enfermedad de Lou Gehrigg, enfermedad de Alzheimer, lupus eritematoso sistémico, hemofilia, enfermedad de Addison, enfermedad de Huntington, y enfermedad de Cushing.

Se describen métodos para tratar una condición, en los que el transgén comprende un gen funcional para reemplazar un gen mutado asociado con un trastorno genético. También se describen métodos para tratar una condición, en los que el transgén puede seleccionarse de la lista de genes que consiste en regulador de la conductancia transmembrana de la fibrosis quística, HFE y HBB.

#### 40 **4. Vehículos farmacéuticos/. Administración de los productos farmacéuticos**

45 Como se ha descrito anteriormente, las composiciones también pueden administrarse *in vivo* en un vehículo farmacéuticamente aceptable. Por "farmacéuticamente aceptable" se quiere decir un material que no es indeseable biológicamente o de otra manera, es decir, el material puede administrarse a un sujeto, junto con el ácido nucleico o vector, sin causar ningún efecto biológico indeseable o sin interaccionar de una manera perjudicial con ninguno de los demás componentes de la composición farmacéutica en la que está contenido. El vehículo se seleccionará naturalmente para minimizar cualquier degradación del ingrediente activo y para minimizar cualquier efecto secundario adverso en el sujeto, como conocerá muy bien el experto en la técnica.

50 Las composiciones pueden administrarse por vía oral, parenteral (p. ej., intravenosa), por inyección intramuscular, por inyección intraperitoneal, transdérmica, extracorpórea, tópica o semejante, aunque se prefiere típicamente la administración intranasal tópica o administración por inhalante. Tal y como se usa en la presente memoria, "administración intranasal tópica" significa la administración de las composiciones en la nariz y vías nasales a través de uno o ambos orificios nasales y puede comprender la administración por un mecanismo de pulverización o mecanismo de gotas, o a través de la aerosolización del ácido nucleico o vector. El último puede ser eficaz cuando se quiere tratar

5 simultáneamente un gran número de animales. La administración de las composiciones por inhalante puede ser a través de la nariz o la boca mediante la administración por un mecanismo de pulverización o gotas. La administración también puede ser directamente a cualquier área del sistema respiratorio (p. ej., los pulmones) mediante intubación. La cantidad exacta de las composiciones requerida variará de un sujeto a otro, dependiendo de la especie, edad, peso y condición general del sujeto, la gravedad del trastorno alérgico que se está tratando, el ácido nucleico o vector particular usado, su modo de administración y semejantes. Así, no es posible especificar una cantidad exacta para cada composición. Sin embargo, el experto en la técnica puede determinar una cantidad apropiada usando sólo la experimentación rutinaria proporcionada en las enseñanzas de la presente memoria.

10 La administración parenteral de la composición, si se usa, se caracteriza generalmente por inyección. Los inyectables pueden prepararse en formas convencionales, bien como disoluciones líquidas o suspensiones, formas sólidas adecuadas para la disolución de la suspensión en líquido antes de la inyección o como emulsiones. Un método revisado más recientemente para la administración parenteral implica el uso de un sistema nuevo de liberación lenta o liberación sostenida de manera que se mantiene una dosificación constante. Véase, p. ej., la Patente EEUU No. 3.610.795.

15 Los materiales pueden estar en disolución, suspensión (por ejemplo, incorporados en micropartículas, liposomas o células). Éstos pueden estar dirigidos a un tipo celular particular mediante anticuerpos, receptores o ligandos de receptores. Las referencias siguientes son ejemplos del uso de esta tecnología para dirigir proteínas específicas a tejido tumoral (Senter, et al., *Bioconjugate Chem.*, 2:447-451, (1991); Bagshawe, K.D., *Br. J. Cancer*, 60:275-281, (1989); Bagshawe, et al., *Br. J. Cancer*, 58:700-703, (1988); Senter, et al., *Bioconjugate Chem.*, 4:3-9, (1993); Battelli, et al., *Cancer Immunol. Immunother.*, 35:421-425, (1992); Pietersz y McKenzie, *Immunol. Reviews*, 129:57-80, (1992); y Roffler, et al., *Biochem. Pharmacol.*, 42:2062-2065, (1991)). Los vehículos tales como liposomas durmientes y otros liposomas conjugados con anticuerpos (incluyendo direccionamiento de fármaco mediado por lípidos hacia carcinoma de colon), direccionamiento hacia ADN mediado por receptor a través de ligandos específicos de célula, direccionamiento hacia tumores dirigido por linfocitos y direccionamiento retroviral terapéutico altamente específico de células de glioma murino *in vivo*. Las referencias siguientes son ejemplos del uso de esta tecnología para dirigir proteínas específicas hacia tejido tumoral (Hughes et al., *Cancer Research*, 49:6214-6220, (1989); y Litzinger y Huang, *Biochimica et Biophysica Acta*, 1104:179-187, (1992)). En general, los receptores están implicados en rutas de endocitosis, bien de forma constitutiva o inducida por ligandos. Estos receptores se agrupan en depresiones recubiertas con clatrina, entran en la célula a través de vesículas recubiertas con clatrina, pasan a través de un endosoma ácido en el que los receptores se separan y bien se reciclan hacia la superficie celular, se almacenan intracelularmente o se degradan en los lisosomas.

20 Las rutas de internalización cumplen varias funciones, tales como captación de nutrientes, eliminación de proteínas activadas, aclaramiento de macromoléculas, entrada oportunista de virus y toxinas, disociación y degradación del ligando y regulación del nivel del receptor. Muchos receptores siguen más de una ruta intracelular, dependiendo del tipo celular, concentración del receptor, tipo de ligando, valencia del ligando y concentración del ligando. Los mecanismos moleculares y celulares de la endocitosis mediada por receptor se han revisado (Brown y Greene, *DNA and Cell Biology*, 10:6, 399-409 (1991)).

25

30

35

#### a) Vehículos Farmacéuticamente Aceptables

Las composiciones, incluyendo anticuerpos, pueden usarse terapéuticamente en combinación con un vehículo farmacéuticamente aceptable.

40 Los vehículos farmacéuticos son conocidos por los expertos en la técnica. En su mayoría éstos serán típicamente vehículos estándar para la administración de fármacos a los seres humanos, incluyendo disoluciones tales como agua estéril, disolución salina y disoluciones tamponadas a pH fisiológico. Las composiciones pueden administrarse por vía intramuscular o subcutánea. Otros compuestos se administrarán según los procedimientos estándar usados por los expertos en la técnica.

45 Las composiciones farmacéuticas pueden incluir vehículos, espesantes, diluyentes, tampones, conservantes, agentes superficiales activos y semejantes además de la molécula elegida. Las composiciones farmacéuticas también pueden incluir uno o más ingredientes activos tales como agentes antimicrobianos, agentes antiinflamatorios, anestésicos y semejantes.

50 La composición farmacéutica puede administrarse de varias maneras dependiendo de si se desea un tratamiento local o sistémico y del área que se va a tratar. La administración puede ser por vía tópica (incluyendo oftalmológica, vaginal, rectal, intranasal), oral, por inhalación o parenteral, por ejemplo, por goteo intravenoso, inyección subcutánea, intraperitoneal o intramuscular. Los anticuerpos descritos pueden administrarse por vía intravenosa, intraperitoneal, intramuscular, subcutánea, intracavital o transdérmica.

Las preparaciones para la administración parenteral incluyen disoluciones estériles acuosas y no acuosas, suspensiones y emulsiones. Los ejemplos de disolventes no acuosos son propilen glicol, polietilen glicol, aceites vegetales tal como

5 aceite de oliva y ésteres orgánicos inyectables tal como oleato de etilo. Los vehículos acuosos incluyen agua, disoluciones alcohólicas/acuosas, emulsiones o suspensiones, incluyendo disolución salina y medios tamponados. Los vehículos parenterales incluyen disolución de cloruro de sodio, dextrosa de Ringer, dextrosa y cloruro de sodio, Ringer con lactato o aceites fijados. Los vehículos intravenosos incluyen reponedores de fluidos o nutrientes, reponedores de electrolitos (tal como los basados en dextrosa de Ringer) y semejantes. También pueden estar presentes conservantes y otros aditivos tales como, por ejemplo, antimicrobianos, antioxidantes, agentes quelantes y gases inertes y semejantes.

Las formulaciones para la administración tópica pueden incluir pomadas, lociones, cremas, geles, gotas, supositorios, pulverizadores, líquidos y polvos. Los vehículos farmacéuticos convencionales, acuosos, en polvo o en bases oleosas, espesantes y semejantes pueden ser necesarios o deseables.

10 Las composiciones para la administración oral incluyen polvos o gránulos, suspensiones o disoluciones en agua o medios no acuosos, cápsulas, sobres o comprimidos. Pueden ser deseables espesantes, saporíferos, diluyentes, emulsionantes, agentes de dispersión o aglutinantes.

15 Algunas de las composiciones pueden administrarse potencialmente como una sal de adición de ácido o base farmacéuticamente aceptable, formada por la reacción con ácidos inorgánicos tales como ácido clorhídrico, ácido bromhídrico, ácido perclórico, ácido nítrico, ácido tiocianico, ácido sulfúrico y ácido fosfórico y ácidos orgánicos tales como ácido fórmico, ácido acético, ácido propiónico, ácido glicólico, ácido láctico, ácido pirúvico, ácido oxálico, ácido malónico, ácido succínico, ácido maleico y ácido fumárico o por la reacción con una base inorgánica tal como hidróxido de sodio, hidróxido de amonio, hidróxido de potasio y bases orgánicas tales como mono, di, trialkil y aril aminas y etanolaminas sustituidas.

## 20 **b) Usos Terapéuticos**

Los intervalos de dosificación para la administración de las composiciones son aquellos lo suficientemente amplios como para producir el efecto deseado en el que se vean influidos los síntomas del trastorno. La dosificación no debería ser tan amplia como para causar efectos secundarios adversos, tales como reacciones cruzadas no deseadas, reacciones anafilácticas y semejantes. Generalmente, la dosificación variará con la edad, condición, sexo y grado de la enfermedad en el paciente y puede determinarla un experto en la técnica. La dosificación puede ajustarse por el médico individual en el evento de cualquier contraindicación. La dosificación puede variar y puede administrarse en una o más administraciones de dosis diariamente durante uno o varios días.

30 Otros vectores que no tienen una función farmacéutica específica pero que pueden usarse para seguir los cambios en los cromosomas celulares o para la administración de herramientas de diagnóstico por ejemplo pueden administrarse de maneras similares a las descritas para los productos farmacéuticos.

Las construcciones de ácido nucleico de la invención también pueden usarse por ejemplo como herramientas para aislar y ensayar nuevos fármacos candidatos para varias enfermedades. También pueden usarse para el aislamiento y estudio continuo, por ejemplo, del ciclo celular. Su uso como dispositivos de administración de ADN exógeno puede extenderse a casi cualquier razón deseada por los expertos en la técnica.

## 35 **5. Similitudes de secuencia**

40 Se entiende que, como se discute en la presente memoria, el uso de los términos homología e identidad significa la misma cosa que similitud. Así, por ejemplo, si el uso de la palabra homología se usa entre dos secuencias no naturales se entiende que esto no indica necesariamente una relación evolutiva entre estas dos secuencias sino más bien se centra en la similitud o el grado de relación entre sus secuencias de ácido nucleico. Muchos de los métodos para determinar la homología entre dos moléculas relacionadas evolutivamente se aplican rutinariamente a cualquiera de dos o más ácidos nucleicos o proteínas para el propósito de medir la similitud de secuencia independientemente de si están o no relacionadas evolutivamente.

45 En general, se entiende que una manera de definir cualquier variante o derivado conocido o los que podrían surgir, de los genes y proteínas descritos en la presente memoria, es mediante la definición de las variantes y derivados en términos de homología respecto a secuencias conocidas específicas. Esta identidad de secuencias particulares descrita en la presente memoria también se discute en otros lugares en la presente memoria. En general, las variantes de los genes y proteínas descritos en la presente memoria tienen al menos aproximadamente 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 ó 99 de porcentaje de homología con la secuencia indicada o la secuencia nativa. Los expertos en la técnica entienden fácilmente cómo determinar la homología de dos proteínas o ácidos nucleicos, tales como genes. Por ejemplo, la homología puede calcularse después de alinear las dos secuencias de manera que la homología está en su nivel más alto.

Otra manera de calcular la homología puede llevarse a cabo mediante algoritmos publicados. La alineación óptima de secuencias para comparación puede realizarse por el algoritmo de homología local de Smith y Waterman Adv. Appl. Math. 2:482 (1981), por el algoritmo de alineación por homología de Needleman y Wunsch, J. Mol. Biol. 48:443 (1970), por la búsqueda de similitud por el método de Pearson y Lipman, Proc. Natl. Acad. Sci. U.S.A. 85:2444 (1988), por implementaciones informáticas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI) o por inspección.

Los mismos tipos de homología pueden obtenerse para ácidos nucleicos mediante por ejemplo los algoritmos descritos en Zuker, M. Science 144:48-52, 1989, Jaeger et al. Proc. Natl. Acad. Sci. USA 86:7706-7710, 1989, Jaeger et al. Methods Enzymol. 183:281-306, 1989 para al menos material relacionado con la alineación del ácido nucleico. Se entiende que cualquiera de los métodos puede usarse típicamente y que en determinados casos los resultados de estos diferentes métodos pueden ser distintos pero el experto en la técnica entiende que si se encuentra identidad con al menos uno de estos métodos, se dirá que las secuencias tienen la identidad indicada y se describirá en la presente memoria.

Por ejemplo, tal y como se usa en la presente memoria, una secuencia de la que se dice que tiene un porcentaje de homología particular respecto a otra secuencia se refiere a secuencias que tienen la homología indicada según se calcula por uno cualquiera o más de los métodos de cálculo descritos anteriormente. Por ejemplo, una primera secuencia tiene 80 por ciento de homología, como se define en la presente memoria, respecto a una segunda secuencia si se calcula que la primera secuencia tiene 80 por ciento de homología respecto a la segunda secuencia usando el método de cálculo de Zuker incluso si la primera secuencia no tiene 80 por ciento de homología respecto a la segunda secuencia según se calcula por cualquiera de los demás métodos de cálculo. Como otro ejemplo, una primera secuencia tiene 80 por ciento de homología, como se define en la presente memoria, respecto a una segunda secuencia si se calcula que la primera secuencia tiene 80 por ciento de homología respecto a la segunda secuencia usando tanto el método de cálculo de Zuker como el método de cálculo de Pearson y Lipman incluso si la primera secuencia no tiene un 80 por ciento de homología respecto a la segunda secuencia según se calcula por el método de cálculo de Smith y Waterman, el método de cálculo de Needleman y Wunsch, los métodos de cálculo de Jaeger o cualquiera de los demás métodos de cálculo. Como otro ejemplo más, una primera secuencia tiene 80 por ciento de homología, como se define en la presente memoria, respecto a una segunda secuencia si se calcula que la primera secuencia tiene 80 por ciento de homología respecto a la segunda secuencia usando cada uno de los métodos de cálculo (aunque, en la práctica, los diferentes métodos de cálculo resultarán frecuentemente en porcentajes de homología calculados diferentes).

## 6. Hibridación/hibridación selectiva

El término hibridación significa típicamente una interacción guiada por secuencia entre al menos dos moléculas de ácido nucleico, tales como un cebador o una sonda y un gen. La interacción guiada por secuencia significa una interacción que ocurre entre dos nucleótidos o análogos de nucleótido o derivados de nucleótido de manera específica de nucleótido. Por ejemplo, G interaccionando con C o A interaccionando con T son interacciones guiadas por secuencia. Las interacciones guiadas por secuencia ocurren típicamente en la cara Watson-Crick o cara Hoogsteen del nucleótido. La hibridación de dos ácidos nucleicos se ve influida por varias condiciones y parámetros conocidos para los expertos en la técnica. Por ejemplo, las concentraciones de sal, pH y temperatura de la reacción influirán todas en si dos moléculas de ácido nucleico hibridarán.

Los parámetros para la hibridación selectiva entre dos moléculas de ácido nucleico son muy conocidas para los expertos en la técnica. Por ejemplo, en algunas realizaciones las condiciones para la hibridación selectiva pueden definirse como condiciones de hibridación astringentes. Por ejemplo, la astringencia de la hibridación se controla tanto por la temperatura como la concentración de sal de la etapa de hibridación y lavado o de ambas. Por ejemplo, las condiciones de hibridación para conseguir la hibridación selectiva pueden implicar la hibridación en disolución con alta fuerza iónica (6X SSC ó 6X SSPE) a una temperatura que es aproximadamente 12-25<sup>o</sup>C por debajo de la T<sub>m</sub> (la temperatura de fusión a la que la mitad de las moléculas se disocian de sus parejas de hibridación) seguido de lavado a una combinación de temperatura y concentración de sal elegida de manera que la temperatura de lavado es aproximadamente 5<sup>o</sup>C a 20<sup>o</sup>C por debajo de la T<sub>m</sub>. Las condiciones de temperatura y de sal se determinan fácilmente de manera empírica en experimentos preliminares en los que se hibridan muestras de ADN de referencia inmovilizadas en filtros con un ácido nucleico marcado de interés y se lavan bajo condiciones de diferentes astringencias. Las temperaturas de hibridación son típicamente mayores para hibridaciones ADN-ARN y ARN-ARN. Las condiciones pueden usarse como se ha descrito anteriormente para conseguir la astringencia o como se muestra en la técnica (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2<sup>a</sup> Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, 1989; Kunkel et al. Methods Enzymol. 1987:154:367, 1987 para material al menos relacionado con la hibridación de ácidos nucleicos). Una condición de hibridación astringente preferida para una hibridación ADN:ADN puede ser a aproximadamente 68<sup>o</sup>C (en disolución acuosa) en 6X SSC ó 6X SSPE seguido de lavado a 68<sup>o</sup>C. La astringencia de la hibridación y el lavado, si se desea, puede reducirse consecuentemente al disminuir el grado de complementariedad deseado y, además, dependiendo de la riqueza en G-C o A-T de cualquier área en la que se busca

variabilidad. De la misma manera, la astringencia de la hibridación y el lavado, si se desea, puede incrementarse consecuentemente al aumentar la homología deseada y, además, dependiendo de la riqueza en G-C o A-T de cualquier área en la que se desea alta homología, todo como se conoce en la técnica.

5 Otra manera de definir la hibridación selectiva es observar la cantidad (porcentaje) de uno de los ácidos nucleicos unido al otro ácido nucleico. Por ejemplo, en algunas realizaciones las condiciones de hibridación selectiva serán cuando al menos aproximadamente 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 por cien del ácido nucleico limitante está unido al ácido nucleico no limitante. Típicamente, el cebador no limitante está, por ejemplo, en un exceso de 10 ó 100 ó 1.000 veces. Este tipo de ensayo puede llevarse a cabo bajo condiciones en las que tanto el cebador limitante como no limitante están, por ejemplo, 10 veces ó 100 veces ó 1.000 veces por debajo de su  $k_d$  o cuando sólo una de las moléculas de ácido nucleico está 10 veces ó 100 veces ó 1.000 veces o cuando una o ambas moléculas de ácido nucleico están por encima de su  $k_d$ .

10 Otra manera de definir la hibridación selectiva es observar el porcentaje de cebador que resulta enzimáticamente manipulado bajo condiciones en las que la hibridación se requiere para estimular la manipulación enzimática deseada. Por ejemplo, en algunas realizaciones las condiciones de hibridación selectiva serán cuando al menos aproximadamente 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 por cien del cebador está enzimáticamente manipulado bajo condiciones que estimulan la manipulación enzimática, por ejemplo si la manipulación enzimática es la extensión del ADN, entonces las condiciones de hibridación selectiva serán cuando al menos aproximadamente 60, 65, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 por cien de las moléculas de cebador se han extendido. Las condiciones preferidas también incluyen las sugeridas por el fabricante o las indicadas en la técnica como apropiadas para la enzima que lleva a cabo la manipulación.

15 Como con la homología, se entiende que existen varios métodos descritos en la presente memoria para determinar el nivel de hibridación entre dos moléculas de ácido nucleico. Se entiende que estos métodos y condiciones pueden proporcionar diferentes porcentajes de hibridación entre dos moléculas de ácido nucleico, pero a no ser que se indique otra cosa será suficiente con cumplir los parámetros de cualquiera de los métodos. Por ejemplo, si se requería el 80% de hibridación y siempre que la hibridación ocurra con los parámetros requeridos en uno cualquiera de estos métodos se considera descrita en la presente memoria.

20 Se entiende que los expertos en la técnica entienden que si una composición o método cumple con uno cualquiera de estos criterios para determinar la hibridación bien colectivamente o individualmente es una composición o método que se describe en la presente memoria.

### 7. Ácidos nucleicos

25 Existen varias moléculas descritas en la presente memoria que se basan en ácido nucleico, incluyendo por ejemplo los ácidos nucleicos que codifican, por ejemplo, una transposasa quimérica, así como varios ácidos nucleicos funcionales. Los ácidos nucleicos descritos están constituidos, por ejemplo, por nucleótidos, análogos de nucleótidos o sustitutos de nucleótidos. Los ejemplos no limitativos de éstas y otras moléculas se discuten en la presente memoria. Se entiende que, por ejemplo, cuando un vector se expresa en una célula, el ARNm expresado estará constituido típicamente por A, C, G y U. De la misma manera, se entiende que, por ejemplo, si se introduce una molécula antisentido en una célula o medio celular, por ejemplo, mediante administración exógena, es ventajoso que la molécula antisentido esté constituida por análogos de nucleótidos que reducen la degradación de la molécula antisentido en el medio celular.

#### 40 a) In vivo/ex vivo

Como se ha descrito anteriormente, las composiciones pueden administrarse en un vehículo farmacéuticamente aceptable y pueden administrarse a las células del sujeto in vivo y/o ex vivo mediante varios mecanismos muy conocidos en la técnica (p. ej., captación de ADN desnudo, fusión de liposomas, inyección intramuscular de ADN mediante una pistola de genes, endocitosis y semejantes).

45 Si se emplean métodos ex vivo, las células o tejidos pueden tomarse y mantenerse fuera del cuerpo según protocolos estándar muy conocidos en la técnica. Las composiciones pueden introducirse en las células mediante cualquier mecanismo de transferencia génica, tal como, por ejemplo, administración génica mediada por fosfato de calcio, electroporación, microinyección o proteoliposomas. Las células transducidas pueden volver a infundirse (p. ej., en un vehículo farmacéuticamente aceptable) o transplantarse de nuevo homotópicamente en el sujeto mediante métodos estándar para el tipo de célula o tejido. Se conocen métodos estándar para el trasplante o infusión de varias células en un sujeto.

### 8. Péptidos

### a) Variantes de proteínas

5 Como se discute en la presente memoria, existen numerosas variantes de las enzimas de integración quiméricas y que se conocen y se contemplan en la presente memoria. Además, existen derivados de las enzimas de integración quiméricas que también funcionan en los métodos y composiciones descritos. Las variantes y derivados de proteínas los  
 10 entienden muy bien los expertos en la técnica y pueden implicar modificaciones en la secuencia de aminoácidos. Por ejemplo, las modificaciones en la secuencia de aminoácidos pertenecen típicamente a una o más de tres clases: variantes por sustitución, inserción o delección. Las inserciones incluyen fusiones amino y/o carboxi terminales así como inserciones intrasecuencia de un único o múltiples restos de aminoácidos. Las inserciones serán habitualmente inserciones menores que las de fusiones amino o carboxi terminales, por ejemplo, del orden de uno a cuatro restos. Los  
 15 derivados de proteína por fusión inmunogénica, tales como los descritos en los ejemplos, se obtienen fusionando un polipéptido lo suficientemente grande para conferir inmunogenicidad a la secuencia diana por entrecruzamiento in vitro o por cultivo celular recombinante transformado con ADN que codifica la fusión. Las delecciones se caracterizan por la eliminación de uno o más restos de aminoácidos de la secuencia de la proteína. Típicamente, no se delecionan más de aproximadamente 2 a 6 restos en cualquier sitio en la molécula de la proteína. Estas variantes se preparan habitualmente por mutagénesis específica de sitio de nucleótidos del ADN que codifica la proteína, produciendo de esta  
 20 manera ADN que codifica la variante y expresando posteriormente el ADN en cultivo celular recombinante. Las técnicas para obtener mutaciones por sustitución en sitios predeterminados en el ADN que tienen una secuencia conocida son muy conocidas, por ejemplo mutagénesis por cebador M13 y mutagénesis por PCR. Las sustituciones de aminoácidos son típicamente de restos únicos, pero pueden ocurrir en varias localizaciones diferentes a la vez; las inserciones serán habitualmente del orden de aproximadamente 1 a 10 restos de aminoácidos; y las delecciones variarán de aproximadamente 1 a 30 restos. Las delecciones o inserciones se hacen preferiblemente en pares adyacentes, es decir, una delección de 2 restos o inserción de 2 restos. Las sustituciones, delecciones, inserciones o cualquier combinación de éstas pueden combinarse para conseguir una construcción final. Las mutaciones no deben poner la secuencia fuera del marco de lectura y preferiblemente no crearán regiones complementarias que podrían producir una estructura de ARNm  
 25 secundaria. Las variantes por sustitución son aquellas en las que al menos un resto se ha eliminado y un resto diferente se ha insertado en su lugar. Dichas sustituciones se hacen generalmente según las Tablas 1 y 2 siguientes y se refieren como sustituciones conservativas.

TABLA 1: Abreviaturas de Aminoácidos

| Aminoácido      | Abreviaturas |
|-----------------|--------------|
| alanina         | Ala; A       |
| alsoleucina     | Alle         |
| arginina        | Erg; R       |
| asparagina      | Asn; N       |
| ácido aspártico | Asp; D       |
| cisteína        | Cys; C       |
| ácido glutámico | Glu; E       |
| glutamina       | Gln; Q       |
| glicina         | Gly; G       |
| histidina       | His; H       |
| isoleucina      | Ile; I       |
| leucina         | Leu; L       |
| lisina          | Lys; K       |
| fenilalanina    | Phe; F       |
| prolina         | Pro; P       |

|                     |        |
|---------------------|--------|
| ácido piroglutámico | Glu    |
| serina              | Ser; S |
| treonina            | Thr; T |
| tirosina            | Tyr; Y |
| trptófano           | Trp; W |
| valina              | Val; V |

| TABLA 2: Sustituciones de Aminoácidos   |
|---|
| Sustituciones Conservativas Ejemplares Resto Original, otras se conocen en la técnica |
| Ala; Ser  |
| Arg; Lys, Gln   |
| Asn; Gln; His   |
| Asp; Glu  |
| Cys; Ser  |
| Gln; Asn, Lys   |
| Glu; Asp  |
| Gly; Pro  |
| His; Asn; Gln   |
| Ile; Leu; Val   |
| Leu; Ile; Val   |
| Lys; Arg; Gln   |
| Met; Leu; Ile   |
| Phe; Met; Leu; Tyr  |
| Ser; Thr  |
| Thr; Ser  |
| Trp; Tyr  |
| Tyr; Trp; Phe   |
| Val; Ile; Leu   |

5

Se hacen cambios sustanciales en la función o identidad inmunológica seleccionando las sustituciones que son menos conservativas que las de la Tabla 2, es decir, seleccionando restos que se diferencian más significativamente en su efecto para mantener (a) la estructura del núcleo del polipéptido en el área de la sustitución, por ejemplo, como una confirmación en lámina o hélice, (b) la carga o hidrofobicidad de la molécula en el sitio diana o (c) el tamaño de la cadena lateral. Las sustituciones que en general se espera que produzcan los cambios mayores en las propiedades de la proteína serán aquellas en las que (a) un resto hidrofílico, p. ej., seril o treonil, se sustituye por (o con) un resto hidrofóbico, p. ej., leucil, isoleucil, fenilalanil, valil o alanil; (b) una cisteína o prolina se sustituye por (o con) cualquier

otro resto; (c) un resto que tiene una cadena lateral electropositiva, p. ej., lisil, arginil o histidil, se sustituye por (o con) un resto electronegativo, p. ej., glutamil o aspartil; o (d) un resto que tiene una cadena lateral grande, p. ej., fenilalanina, se sustituye por (o con) uno que no tiene una cadena lateral, p. ej., glicina, en este caso, (e) incrementando el número de sitios para la sulfatación y/o glicosilación.

5 Por ejemplo, el reemplazamiento de un resto de aminoácido con otro que es biológicamente y/o químicamente similar es conocido para los expertos en la técnica como sustitución conservativa. Por ejemplo, una sustitución conservativa sería reemplazar un resto hidrofóbico por otro o un resto polar por otro. Las sustituciones incluyen combinaciones tales como, por ejemplo, Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; y Phe, Tyr. Dichas variaciones con  
10 sustituciones conservativas de cada secuencia descrita explícitamente están incluidas en los polipéptidos mosaico proporcionados en la presente memoria.

Puede emplearse mutagénesis sustitucional o delecional para insertar sitios para N-glicosilación (Asn-X-Thr/Ser) o O-glicosilación (Ser o Thr). Las deleciones de cisteína u otros restos lábiles también pueden ser deseables. Las deleciones o sustituciones de sitios potenciales de proteólisis, p. ej., Arg, se consigue por ejemplo delecionando uno de los restos básicos o sustituyendo uno por restos glutaminil o histidil.

15 Determinadas derivatizaciones posteriores a la traducción son el resultado de la acción de las células anfitrionas recombinantes en el polipéptido expresado. Los restos glutaminil y asparaginil se deamidán frecuentemente después de la traducción a los restos glutamil y asparil correspondientes. Alternativamente, estos restos se deamidán bajo condiciones ácidas suaves. Otras modificaciones posteriores a la traducción incluyen hidroxilación de prolina y lisina, fosforilación de grupos hidroxilo de restos seril o treonil, metilación de los grupos o-amino de las cadenas laterales de la  
20 lisina, arginina e histidina (T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco p 79-86 [1983]), acetilación del amino N-terminal y, en algunos casos, amidación del carboxilo C-terminal.

Se entiende que una manera para definir las variantes y derivados de las proteínas descritas en la presente memoria es definiendo las variantes y derivados en términos de homología/identidad con secuencias conocidas específicas. Se describen específicamente variantes de éstas y otras proteínas descritas en la presente memoria que tienen al menos  
25 70% ó 75% ó 80% ó 85% ó 90% ó 95% de homología con la secuencia indicada. Los expertos en la técnica entienden fácilmente cómo determinar la homología de dos proteínas. Por ejemplo, la homología puede calcularse después de alinear las dos secuencias de manera que la homología está a su nivel más alto.

Otra manera de calcular la homología puede llevarse a cabo por algoritmos publicados. La alineación óptima de secuencias para comparación puede realizarse por el algoritmo de homología local de Smith y Waterman Adv. Appl. Math. 2:482 (1981), por el algoritmo de alineación por homología de Needleman y Wunsch, J. Mol. Biol. 48:443 (1970), por la búsqueda de similitud por el método de Pearson y Lipman, Proc. Natl. Acad. Sci. U.S.A. 85:2444 (1988), por implementaciones informáticas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI) o por inspección.

30 Los mismos tipos de homología pueden obtenerse para ácidos nucleicos mediante por ejemplo los algoritmos descritos en Zuker, M. Science 244:48-52, 1989, Jaeger et al. Proc. Natl. Acad. Sci. USA 86:7706-7710, 1989, Jaeger et al. Methods Enzymol. 183:281-306, 1989 para al menos material relacionado con la alineación del ácido nucleico.

Se entiende que la descripción de mutaciones conservativas y homología pueden combinarse conjuntamente en cualquier combinación, como las realizaciones que tienen al menos 70% de homología con una secuencia particular en la que las variantes son mutaciones conservativas.

40 Como esta especificación discute varias proteínas y secuencias de proteínas, se entiende que los ácidos nucleicos que codifican esas secuencias de proteínas también están descritos. Esto incluirá todas las secuencias degeneradas relacionadas con una secuencia de proteína específica, es decir, todos los ácidos nucleicos que tienen una secuencia que codifica una secuencia de proteína particular así como todos los ácidos nucleicos, incluyendo ácidos nucleicos degenerados, que codifican las variantes y derivados descritos de las secuencias de proteínas. Así, aunque en la  
45 presente memoria no aparezca escrita cada secuencia de ácido nucleico particular, se entiende que cada una y todas las secuencias están de hecho indicadas y descritas en la presente memoria a través de la secuencia de proteína descrita. Por ejemplo, una de las muchas secuencias de ácido nucleico que puede codificar una transposasa química obtenida de la unión de una transposasa [p. ej., Tcl (Referencia No. NM\_061407, AI878683, AI878522, AI794017); elemento P (Rio et al., Cell (1986) 44:21-32; entre otros)] a un factor de direccionamiento a ADN [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300, dominio de unión al ADN Hin (No. de Referencia J03245), dominio de unión al ADN STF-1 (No. de Referencia S67435, correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) Mol. Endo. 7:1275-1283), entre otros]. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otro motor de  
50 búsqueda de nucleótidos o proteínas.

## 9. Kits

5 En la presente memoria se describen kits que están concebidos como reactivos que pueden usarse en la práctica de los métodos descritos en la presente memoria. Los kits pueden incluir cualquier reactivo o combinación de reactivos discutida en la presente memoria o que se entenderá que se requiere o es beneficiosa en la práctica de los métodos descritos. Por ejemplo, los kits podrían incluir cebadores para llevar a cabo las reacciones de amplificación discutidas en determinadas realizaciones de los métodos, así como los tampones y enzimas requeridos para usar los cebadores como se pretende.

#### 10 **10. Composiciones con funciones similares**

10 Se entiende que las composiciones descritas en la presente memoria tienen determinadas funciones, tales como dirigir un transposón a un ácido nucleico diana o unirse a un ácido nucleico diana. En la presente memoria se describen determinados requerimientos estructurales para llevar a cabo las funciones descritas y se entiende que existen varias estructuras que pueden llevar a cabo la misma función que están relacionadas con las estructuras descritas y que estas estructuras conseguirán finalmente el mismo resultado.

#### 15 **D. Métodos para preparar las composiciones**

15 Las composiciones descritas en la presente memoria y las composiciones necesarias para llevar a cabo los métodos descritos pueden prepararse usando cualquier método conocido por los expertos en la técnica para ese reactivo o compuesto particular a no ser que se indique específicamente otra cosa.

#### 20 **1. Síntesis de ácidos nucleicos**

20 Por ejemplo, los ácidos nucleicos, tales como los oligonucleótidos que se van a usar como cebadores pueden prepararse usando métodos de síntesis química estándar o pueden producirse usando métodos enzimáticos o cualquier otro método conocido. Dichos métodos pueden variar desde digestión enzimática estándar seguida de aislamiento de los fragmentos de nucleótidos (véase, por ejemplo, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2ª Edición (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989) Capítulos 5 y 6) para métodos puramente sintéticos, por ejemplo, por el método de la cianoetil fosforamidita usando un Sistema Sintetizador de ADN Milligen o Beckman 1Plus (por ejemplo, sintetizador automático Modelo 8700 de Milligen-Bioscience, Burlington, MA o Modelo ABI 380B). Los métodos sintéticos útiles para preparar oligonucleótidos también se describen por Ikuta et al., *Ann. Rev. Biochem.* 53:323-356 (1984), (métodos de fosfotriéster y fosfitotriéster), y Narang et al., *Methods Enzymol.*, 65:610-620 (1980), (método de fosfotriéster). Las moléculas de ácido nucleico proteína pueden prepararse usando métodos conocidos tales como los descritos por Nielsen et al., *Biocong. Chem.* 5:3-7 (1994).

#### 30 **2. Síntesis de péptidos**

30 Un método para producir las proteínas descritas es unir dos o más péptidos o polipéptidos entre sí por técnicas de química de proteínas. Por ejemplo, los péptidos o polipéptidos pueden sintetizarse químicamente usando equipos de laboratorio actualmente disponibles usando química Fmoc (9-fluorenilmetiloxycarbonilo) o Boc (*tert*-butiloxycarbonilo). (Applied Biosystems, Inc., Foster City, CA). Un experto en la técnica apreciará fácilmente que un péptido o polipéptido correspondiente a las proteínas descritas puede sintetizarse, por ejemplo, por reacciones químicas estándar. Por ejemplo, un péptido o polipéptido puede sintetizarse y no escindirse de su resina de síntesis mientras que el otro fragmento de un péptido o proteína puede sintetizarse y escindirse posteriormente de la resina, exponiendo de esta manera un grupo terminal que está bloqueado funcionalmente en el otro fragmento. Por reacciones de condensación de péptidos, estos dos fragmentos pueden unirse covalentemente a través de un enlace peptídico en sus extremos carboxilo y amino terminales, respectivamente, para formar un anticuerpo, o fragmento de éste. (Grant GA (1992) *Synthetic Peptides: A User Guide*. W.H. Freeman and Co., N.Y. (1992); Bodansky M y Trost B., Ed. (1993) *Principles of Peptide Synthesis*. Springer-Verlag Inc., NY (al menos para el material relacionado con la síntesis de péptidos). Alternativamente, el péptido o polipéptido se sintetiza independientemente *in vivo* como se describe en la presente memoria. Una vez aislados, estos péptidos o polipéptidos independientes pueden unirse para formar un péptido o fragmento de éste a través de reacciones de condensación de péptidos similares.

45 Por ejemplo, la ligación enzimática de segmentos peptídicos clonados o sintéticos permite que se unan fragmentos peptídicos relativamente cortos para producir fragmentos peptídicos más largos, polipéptidos o dominios proteicos completos (Abrahmsen L et al., *Biochemistry*, 30:4151 (1991)). Alternativamente, puede utilizarse la ligación química nativa de péptidos sintéticos para construir sintéticamente péptidos o polipéptidos grandes a partir de fragmentos peptídicos más cortos. Este método consiste en una reacción química de dos etapas (Dawson et al. *Synthesis of Proteins by Native Chemical Ligation*. *Science*, 266:776-779 (1994)). La primera etapa es la reacción quimioselectiva de un péptido-tioéster sintético no protegido con otro segmento peptídico no protegido que contiene un resto Cys amino terminal para proporcionar un intermedio unido a tioéster como el producto covalente inicial. Sin cambiar las condiciones de la reacción, este intermedio sufre una reacción intramolecular espontánea rápida para formar un enlace peptídico nativo en el sitio de la ligación (Baggiolini M et al. (1992) *FEBS Lett.* 307:97-101; Clark-Lewis I et al., *J. Biol. Chem.*,

50

269:16075 (1994); Clark-Lewis I et al., *Biochemistry*, 30:3128 (1991); Rajarathnam K et al., *Biochemistry* 33:6623-30 (1994)).

5 Alternativamente, los segmentos peptídicos no protegidos se unen químicamente cuando el enlace formado entre los segmentos peptídicos como resultado de la ligación química es un enlace no natural (no peptídico) (Schnolzer, M et al. *Science*, 256:221 (1992)). Esta técnica se ha usado para sintetizar análogos de dominios de proteínas así como grandes cantidades de proteínas relativamente puras con actividad biológica completa (deLisle Milton RC et al., *Techniques in Protein Chemistry IV*. Academic Press, Nueva York, p. 257-267 (1992)).

### 3. Proceso para preparar las composiciones

10 Se describen procesos para preparar las composiciones así como para preparar los intermedios que dan lugar a las composiciones. Por ejemplo, se describen ácidos nucleicos para la construcción de una transposasa quimérica obtenida de unir una transposasa [p. ej., Tcl (Referencia No. NM\_061407, AI878683, AI878522, AI794017); elemento P (Rio et al., *Cell* (1986) 44:21-32; y entre otros)] a un factor de direccionamiento a ADN [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300, dominio de unión al ADN Hin (No. de Referencia J03245), dominio de unión al ADN STF-1 (No. de Referencia S67435, correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) *Mol. Endo.* 7:1275-1283), entre otros]. Las secuencias de éstas y otras transposasas conocidas pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otro motor de búsqueda de nucleótidos o proteínas. Existen varios métodos que pueden usarse para preparar estas composiciones, tales como métodos químicos sintéticos y métodos de biología molecular estándar. Se entiende que los métodos para preparar éstas y las demás composiciones descritas se describen específicamente.

20 Se describen moléculas de ácido nucleico producidas por el proceso que comprende unir de una manera operativa un ácido nucleico que comprende la secuencia mostrada en una transposasa quimérica obtenida de unir una transposasa [p. ej., Tcl (Referencia No. NM\_061407, AI878683, AI878522, AI794017); elemento P (Rio et al., *Cell* (1986) 44:21-32; y entre otros listados en la presente memoria. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas]] a un factor de direccionamiento a ADN [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300, dominio de unión al ADN Hin (No. de Referencia J03245), dominio de unión al ADN STF-1 (No. de Referencia S67435, correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) *Mol. Endo.* 7:1275-1283), y entre otros listados en la presente memoria. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas]] y una secuencia que controla la expresión del ácido nucleico.

30 También se describen moléculas de ácido nucleico producidas por el proceso que comprende unir de una manera operativa una molécula de ácido nucleico que comprende una secuencia que tiene 80% de identidad con una secuencia mostrada en una transposasa quimérica obtenida de unir una transposasa [p. ej., Tcl (Referencia No. NM\_061407, AI878683, AI878522, AI794017); elemento P (Rio et al., *Cell* (1986) 44:21-32; y entre otros listados en la presente memoria. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas]] a un factor de direccionamiento a ADN [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300, dominio de unión al ADN Hin (No. de Referencia J03245), dominio de unión al ADN STF-1 (No. de Referencia S67435, correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) *Mol. Endo.* 7:1275-1283), y entre otros listados en la presente memoria. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas]] y una secuencia que controla la expresión del ácido nucleico.

40 Se describen moléculas de ácido nucleico producidas por el proceso que comprende unir de una manera operativa una molécula de ácido nucleico que comprende una secuencia que hibrida bajo condiciones de hibridación estrictas con una secuencia de una transposasa mostrada en una transposasa quimérica obtenida de unir una transposasa [p. ej., Tcl (Referencia No. NM\_061407, AI878683, AI878522, AI794017); elemento P (Rio et al., *Cell* (1986) 44:21-32; y entre otros listados en la presente memoria. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas]] a un factor de direccionamiento a ADN [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300, dominio de unión al ADN Hin (No. de Referencia J03245), dominio de unión al ADN STF-1 (No. de Referencia S67435, correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) *Mol. Endo.* 7:1275-1283), y entre otros listados en la presente memoria. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas]] y una secuencia que controla la expresión del ácido nucleico.

55 Se describen moléculas de ácido nucleico producidas por el proceso que comprende unir de una manera operativa una molécula de ácido nucleico que comprende una secuencia que codifica un polipéptido de fusión que contiene dos dominios de unión al ADN (o un dominio de unión al ADN y un dominio de unión a proteínas) [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300, dominio de unión al ADN Hin (No. de Referencia J03245), unido al dominio de unión al ADN STF-1 (No. de Referencia S67435, correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) *Mol. Endo.*

7:1275-1283), y entre otros listados en la presente memoria que pueden combinarse. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas] y una secuencia que controla la expresión de la molécula de ácido nucleico.

5 Se describen moléculas de ácido nucleico producidas por el proceso que comprende unir de una manera operativa una molécula de ácido nucleico que comprende una secuencia que codifica un polipéptido de fusión que contiene dos dominios de unión al ADN (o un dominio de unión al ADN y un dominio de unión a proteínas) [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300, dominio de unión al ADN Hin (No. de Referencia J03245), unido al dominio de unión al ADN STF-1 (No. de Referencia S67435, correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) Mol. Endo. 7:1275-1283), y entre otros listados en la presente memoria que pueden combinarse. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas] que tienen 80% de identidad con un péptido y una secuencia que controla la expresión de la molécula de ácido nucleico.

10 Se describen células producidas por el proceso de transformar la célula con cualquiera de los ácidos nucleicos descritos. Se describen células producidas por el proceso de transformar la célula con cualquiera de los ácidos nucleicos no naturales descritos

15 Se describe cualquiera de los péptidos descritos producidos por el proceso de expresar cualquiera de los ácidos nucleicos descritos. Se describe cualquiera de los péptidos no naturales descritos producidos por el proceso de expresar cualquiera de los ácidos nucleicos descritos. Se describe cualquiera de los péptidos descritos producidos por el proceso de expresar cualquiera de los ácidos nucleicos no naturales descritos.

20 Se describen animales producidos por el proceso de transfectar una célula del animal con cualquiera de las moléculas de ácido nucleico descritas en la presente memoria. Se describen animales producidos por el proceso de transfectar una célula del animal con cualquiera de las moléculas de ácido nucleico descritas en la presente memoria, en los que el animal es un mamífero. También se describen animales producidos por el proceso de transfectar una célula del animal con cualquiera de las moléculas de ácido nucleico descritas en la presente memoria, en los que el mamífero es ratón, rata, conejo, vaca, oveja, cerdo o primate.

25 También se describen animales producidos por el proceso de añadir al animal cualquiera de las células descritas en la presente memoria.

### E. Ejemplos

30 Los ejemplos siguientes se muestran para proporcionar a los expertos en la técnica una exposición y descripción completas de cómo los compuestos, composiciones, artículos, dispositivos y/o métodos reivindicados en la presente memoria se preparan y evalúan y se pretende que sean puramente ejemplares de la invención y no se pretende que limiten el alcance de lo que los inventores consideran su invención. Se han hecho esfuerzos para asegurar la precisión respecto a los números (p. ej., cantidades, temperatura, etc.) pero se deben tener en cuenta algunos errores y desviaciones. A no ser que se indique otra cosa, las partes son partes en peso, la temperatura es en °C o es a temperatura ambiente y la presión es atmosférica o cerca de la atmosférica.

### 35 1. Ejemplo 1

#### a) Preparación de las Construcciones de Ácido Nucleico que Codifican la Invención

##### Proteínas Quiméricas

40 Las transposasas quiméricas [p. ej., Tc1 (Referencia No. NM\_061407, AI878683, AI878522, AI794017); elemento P (Rio et al., Cell (1986) 44:21-32; entre otros)] que contienen el dominio de unión al ADN en el extremo "amino terminal" o "carboxilo terminal" se construyen usando PCR de fusión (véanse, p. ej., Vallete et al., 1989, NAR, 17:723-733; y Yon y Fried, 1989, NAR, 17:4895). La región codificadora de la transposasa construida como se describe y el dominio de unión al ADN (p. ej., región codificadora zif268) construido como se describe se amplifican independientemente por PCR. Los cebadores se diseñan empleando métodos muy conocidos para contener una región de superposición que codifica la unión de fusión deseada. Los productos de PCR de dos reacciones independientes se purifican, se mezclan y se someten a una segunda reacción de PCR usando cebadores dirigidos a cualquiera de los sitios de la región de superposición. En el primer ciclo de la segunda ronda, las cadenas de los dos productos de la reacción pueden desnaturalizarse e hibridar para permitir la extensión por la polimerasa. En el ciclo siguiente, la cadena resultante puede amplificarse como en PCR normal.

50 Así, se pueden fusionar de manera precisa dos secuencias no relacionadas: el plásmido basado en transposón [que codifica el transgén, la transposasa y que contiene un sitio de unión a proteínas (p. ej., operadores  $\lambda$ ) y un segundo plásmido que comprende un polipéptido de fusión que contiene dos dominios de unión al ADN (o un dominio de unión al

ADN y un dominio de unión a proteínas) [p. ej., LexA DBD (No. de Registro J01643-V0029-V00300) unido al dominio de unión al ADN STF-1 (No. de Referencia S67435; correspondiente a los a.a 140-215 descrito en Leonard et al. (1993) Mol. Endo. 7:1275-1283), y entre otros listados en la presente memoria que pueden combinarse]. Las secuencias pueden obtenerse en la Base de Datos de Nucleótidos Entrez o GenBank u otros motores de búsqueda de nucleótidos o proteínas] se construyen por los métodos utilizados en el Ejemplo 1 (véanse, p. ej., Vallete et al., 1989, NAR, 17:723-733; y Yon y Fried, 1989, NAR, 17:4895). Los plásmidos se transfectan en células obtenidas del páncreas.

Cultivo Celular y Transfecciones. Las células obtenidas del páncreas se cultivan en OptiMEM (GIBCO/BRL) suplementado con 5% suero y 50 µg/ml penicilina/estreptomicina a 37°C y 5% CO<sub>2</sub>. El ensayo de transposición en cultivo de tejido se lleva a cabo de una manera similar a la descrita previamente (Ivics Z, et al. (1997) Cell 91:1-20).

Después de la cotransfección, se compara el número de colonias resistentes a G418 con el número obtenido después de la cotransfección de un vector basado en transposón control sin PBS. Se ha mostrado previamente que el incremento en el número de colonias resistentes está causado por la integración mediada por transposasa del transposón (Schouten GJ, et al. (1998) Nucleic Acids Res 26:3687-3693 e Ivics Z, et al. (1997) Cell 91:1-20) y, por lo tanto, es indicativo de la eficacia de la transposición.

Los vectores de esta invención se producen por métodos estándar de escisión por enzimas de restricción, ligación y clonación molecular. El protocolo general para construir los presentes vectores incluye las etapas siguientes. En primer lugar, los fragmentos de ácido nucleico purificados que contienen las secuencias de nucleótidos componentes deseadas así como secuencias extrañas se escinden con endonucleasas de restricción a partir de las fuentes iniciales. Los fragmentos que contienen las secuencias de nucleótidos deseadas se separan de los fragmentos no deseados de diferente tamaño usando métodos de separación convencionales, p. ej., por electroforesis en gel de agarosa. Los fragmentos deseados se escinden del gel y se ligan entre sí en la configuración apropiada de manera que se produce un ácido nucleico circular o plásmido que contiene las secuencias deseadas, p. ej., las secuencias correspondientes a los distintos elementos de los presentes vectores, como se ha descrito anteriormente. Cuando se desee, las moléculas circulares así construidas se amplifican en un anfitrión procariota, p. ej., E. coli. Los procedimientos de escisión, construcción del plásmido, transformación celular y producción del plásmido implicados en estas etapas son muy conocidos para un experto en la técnica y las enzimas requeridas para la restricción y ligación están disponibles comercialmente (Véanse, por ejemplo, R. Wu, Ed., Methods in Enzymology, Vol. 68, Academic Press, N.Y. (1979); T. Maniatis, E.F. Fritsch y J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1982); Catálogo 1982-83, New England Biolabs, Inc.; Catálogo 1982-83, Bethesda Research Laboratories, Inc.).

## 2. Ejemplo 2

Se proporcionan transposasas quiméricas que comprenden transposasas conocidas (p. ej., Bella Durmiente, Tn7, Tn916, Tc1/mariner, Tc3, maT, y otras listadas en la presente memoria) que contienen el dominio de unión al ADN (DBD) lexA fusionado de manera precisa en el extremo N o C terminal. Los ejemplos de transposasas no quiméricas conocidas pueden encontrarse en la bibliografía: Bella Durmiente (Izsvak Z, Ivics Z, y Plasterk RH. (2000) Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J. Mol. Biol. 302:93-102), Tn5 (Bhasin A, et al. (2000) Characterization of a Tn5 pre-cleavage synaptic complex. J Mol Biol 302:49-63), Tn7 (Kuduvalli PN, Rao JE, Craig NL. (2001) Target DNA structure plays a critical role in Tn7 transposition. EMBO J 20:924-932), Tn916 (Marra D, Scott JR. (1999) Regulation of excision of the conjugative transposon Tn916. Mol Microbiol 2:609-621), Tc1/mariner (Izsvak Z, Ivics Z, Hackett PB. (1995) Characterization of a Tc-1 like transposable element in zebrafish (Danio rerio). Mol. Gen. Genet. 247:312-322), elementos *Minos* y *S* (Franz G y Savakis C. (1991) Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. Nucl. Acids Res. 19:6646; Merriman PJ, Grimes CD, Ambroziak J, Hackett DA, Skinner P y Simmons MJ. (1995) S elements: a family of Tc1-like transposons in the genome of *Drosophila melanogaster*. Genetics 141:1425-1438), elementos Quetzal (Ke Z, Grossman GL, Cornel AJ, Collins FH. (1996) Quetzal: a transposon of the Tc1 family in the mosquito *Anopheles albimanus*. Genetica 98:141-147); elementos Tor (Lam WL, Seo P, Robison K, Virk S, y Gilbert W. (1996) Discovery of amphibian Tc1-like transposon families. J Mol Biol 257:359-366), subfamilias de transposón semejantes a Tc1 (Ivics Z, Izsvak Z, Minter A, Hackett PB. (1996) Identification of functional domains and evolution of Tc1-like transposable elements. Proc. Natl. Acad Sci USA 93:5008-5013), y Tc3 (Tu Z, Shao H. (2002) Intra- and inter-specific diversity of Tc-3 like transposons in nematodes and insects and implications for their evolution and transposition. Gene 282:133-142), elemento P (Rubin GM y Spradling AC. (1983) Vectors for P element mediated gene transfer in *Drosophila*. Nucleic Acids Res. 11:6341-6351. Estas transposasas de fusión incrementan la integración dirigida en el genoma del anfitrión.

### a) Construcción y purificación de proteínas de fusión LexA-transposasa.

El DBD LexA (No. de Registro J01643-V0029-V00300) se fusiona directamente al extremo N o C terminal de estas transposasas [p. ej., Tc1 (Referencia No. NM\_061407, AI878683, AI878522, AI794017); elemento P (Rio et al., Cell

(1986) 44:21-32; entre otros)]. Las técnicas para construir ácidos nucleicos que codifican proteínas de fusión y la purificación de sus productos son muy conocidas en la bibliografía y para los expertos en la técnica.

5 Las construcciones de transposón lineales y circulares se combinan con su transposasa quimérica respectiva y se añaden a extractos celulares de diferentes líneas celulares (p. ej., células HeLa) que contienen un plásmido diana. Un plásmido diana que contiene la secuencia de unión LexA proporciona el sustrato de ADN para ensayar la integración selectiva de sitio. Dependiendo del vector transposón (p. ej., Tc1/mariner) que se use se hacen modificaciones adicionales a la diana tal como incrementar el número de sitios TA a lo largo del plásmido (incluyendo regiones en, adyacentes a y lejos de su sitio diana putativo) para determinar el efecto en la eficacia de la integración. Los métodos para determinar el sitio de la integración se han descrito previamente en la bibliografía (40, 84). La transposasa quimérica puede ensayarse así para su nivel de integración y especificidad comparado con su transposasa respectiva sin un dominio de unión al ADN fusionado.

### 3. Ejemplo 3: Transposición dirigida del transposón *maT*

#### a) Evaluación de la integración dirigida de *maT* en células de insecto.

15 *maT* es un miembro de la superfamilia *Tc1/mariner* de transposones. Como es característico en los elementos semejantes a mariner, *maT* tiene una triada catalítica DDD. Las ITR de *maT* se parecen más a las de *Tc1* que a las de mariner y las indicaciones estructurales muestran que el dominio N terminal es único y distinto de mariner o *Tc1*. Además, el dominio de unión al ADN se parece más a los factores de transcripción *Pax/paired* y la transposasa *Tc3* que a las transposasas *Tel/mariner*.

20 Se evalúa la capacidad de una transposasa *maT* modificada, quimérica para estimular la integración del transposón en los sitios de unión Gal4 o LexA. Se transfectan líneas celulares de insectos y embriones de insectos con dos a tres plásmidos. El primer plásmido, referido como el plásmido donante, contiene un transposón *maT* modificado que tiene sus repeticiones terminales invertidas y dominios de unión a transposasa intactos, pero su gen transposasa se ha reemplazado o interrumpido por un gen marcador seleccionable (un gen de resistencia a antibiótico). El segundo plásmido expresa tanto las proteínas de unión al ADN como el gen transposasa *maT* fusionados en los extremos 5' ó 3' con los dominios de unión al ADN heterólogos para Gal4 o LexA (o ambos). Los elementos esenciales de estos dos plásmidos también se combinan en un único plásmido, para facilitar la cotransfección de estos elementos en las células (véase la figura 7). El gen transposasa se sitúa bajo el control del promotor del gen *hsp70* de *Drosophila*. El plásmido final, el plásmido diana, contiene los sitios diana para Gal4 o LexA. Después de la administración de los plásmidos a las células o embriones, se induce la expresión de la transposasa por un choque de calor, lo que estimula la transposición del transposón modificado. La presencia de plásmidos diana recombinantes que contienen el transposón *maT* marcado se determina en células/embriones tratados y los sitios de integración se evalúan para la integración específica de sitio. Para asegurar que las integraciones son dependientes de transposón, se llevan a cabo transfecciones control que carecen del plásmido transposasa.

#### b) Evaluación de la integración dirigida de *maT* en células humanas.

35 Se transfectan líneas celulares humanas (HeLa) con los mismos dos o tres plásmidos descritos anteriormente, excepto en que la transposasa está bajo el control de un promotor CMV y el transposón contiene el marcador seleccionable de resistencia a G418 bajo el control del promotor SV40. Los plásmidos diana recombinantes se recuperan para evaluar la proximidad de la integración del transposón respecto a un sitio diana LexA o Gal4.

#### c) Evaluación de la integración dirigida de *maT* en *Drosophila*.

40 En experimentos que demuestran que la integración del transposón *maT* puede dirigirse preferentemente a los sitios Gal4 o LexA en las líneas celulares, los plásmidos donante y transposasa se administran a líneas transgénicas de *Drosophila* que contienen sitios diana Gal4 para evaluar la integración dirigida en un organismo intacto.

### 4. Ejemplo 4: Uso de los dominios de unión al ADN Gal4 y LexA para conseguir la integración dirigida de los transposones *Mosl* y *PiggyBac*.

45 Las secuencias de reconocimiento de ADN Gal4 y LexA se han amplificado por PCR y clonado en el plásmido diana pGDV1. Los transposones *Mosl* y *piggyBac* se fusionan a NLS y se han incorporado secuencias conectoras en los oligonucleótidos y se usaron en las reacciones de amplificación por PCR para producir los dominios de unión al ADN requeridos tanto para Gal4 como para LexA. Se construyen fusiones dobles de los dominios de unión al ADN Gal4-LexA y plásmidos diana modificados pGDV1 que contienen secuencias de reconocimiento para los dominios de unión al ADN de Gal4 y LexA usando esta misma técnica. Los plásmidos independientes se cortan con enzimas de restricción y se ligan formando la construcción completa como se muestra en la Figura 8.

Los ensayos de transposición interplásmido se llevan a cabo usando los auxiliares dominio de unión al ADN-transposasa, transposones donantes modificados y plásmidos diana modificados pGDV1 para determinar a) un incremento general en la frecuencia de transposición y b) un direccionamiento específico de la integración del transposón en o cerca de las secuencias de reconocimiento.

- 5 Los experimentos de transformación genética se llevan a cabo en cepas de *Drosophila melanogaster* que contienen secuencias de reconocimiento Gal4 o LexA, usando el auxiliar modificado y los transposones donantes para determinar a) un incremento general en la frecuencia de transformación de la línea germinal y b) un direccionamiento específico de las integraciones en las copias genómicas de las secuencias de reconocimiento.

### 5. Ejemplo 5: Transposasas quiméricas

#### 10 a) Plásmidos.

Los fragmentos de PCR de las ORF que codifican las proteínas transposasa de Tc1, Tc3, *Himar1* y *Mos1* se clonaron en el fragmento *NotI* tratado con Klenow de 3,8 kb de pCMV $\beta$  (CLONTECH), lo que resulta, respectivamente, en pRP1341, pRP1342, pRP1389 y pRP1353.

- 15 Los fragmentos de PCR de las ORF que codifican las proteínas transposasa quiméricas de Tc1, Tc3, *Himar1* y *Mos1* fusionados con el dominio de unión al ADN del dominio de unión al ADN repetido ALU de 66 kD (Luk'yanov, et al. (1999) *Biochemistry* (Moscú) 64:25-33), los dominios de unión al ADN de dos proteínas de 120 kD y 35 kD aisladas de las proteínas celulares de HeLa que se unen específicamente al elemento Alu (Chiang Y y Vishwanatha JK (1996) *Molecular and Cellular Biochemistry* 155:131-138, 1996) u otros contenidos en la presente memoria y se diseñan por métodos proporcionados por Kim, JS, et al. (Solicitud de Patente No. 20020061512) en el dominio C terminal de la transposasa, 20 se clonaron en el fragmento *NotI* tratado con Klenow de 3,8 kb de pCMV $\beta$  (CLONTECH), lo que resulta, respectivamente, en los grupos de plásmidos pRP1341\*, pRP1342\*, pRP1389\* y pRP1353\*. Los plásmidos molde fueron, respectivamente, pRP470 (Vos J, et al (1993) *Genes Development* 7:1244-1253), pRP716 (Van Luenen, HGAM et al (1993) *EMBO J.* 12:2513-2520), pMar27fH (Lampe DJ, et al (1999) *PNAS* 96:1142S-11433) y *pMos1* (Medhora M, et al (1991) *Genetics* 128:311-318). Las mutaciones en las ORF de las transposasas quiméricas Tc3, *Mos1* y *Himar1* se introducen bien por mutagénesis dirigida a sitio usando cebadores mutagénicos o por un método de PCR-ligación-PCR 25 (Ali SA *BioTechniques* 18:746-750). Se construyen los vectores de expresión de citomegalovirus (CMV) siguientes: grupos de plásmidos pRP2301\* (Tc3 N225D/DBD), pRP2302\* (Tc3 V41E N225D/DBD), pRP1390\* (*Mos1* F344L/DBD), pRP1398\* (*Himar1* H267R/DBD), pRP1399\* (*Himar1* Q131RE137K/DBD) y pRP2300\* (*Himar1* Q131RE137K H267R/DBD). También se ensayan los plásmidos que contienen mutagénesis dirigida a sitio de las transposasas de Tc3 30 (pRP2301y pRP2302), *Mos1* (pRP1390), *Himar1* (pRP1398, pRP1399 y pRP2300) como se describe a continuación (Fisher SE (2001) *PNAS* 98:6759-6764).

- Un casete virus de simio 40 (SV40)-resistencia a G418 (un fragmento *Bam*HI-*Eco*RI con extremos romos de 1,6 kb de pRc/CMV (Invitrogen)) se clona en Tc1 [en los sitios romos Styl de pRP1212 (Ketting RF, et al (1997) *Nucleic Acids Res* 25:4041-4047)], lo que resulta en pRP1349, en Tc3 [en los sitios con extremos romos *Bsp*EI y *Nco*I de Tc3 en pRP790 35 (Fisher SE, et al (1999) *Mol Gen Genet* 262:268-274)], lo que resulta en pRP1351, en *Himar1* [el transposón *Himar1* se corta de pMarKan (Lampe DJ, et al (1999) *PNAS* 96:11428-11433) usando *NotI* y *Eco*RI y se clona en el sitio *Sma*I de pUC19, el casete SV40-resistencia a G418 se clona en los sitios *Hin*clI y *Bst*EII], lo que resulta en pRP1347 y en *Mos1* [el transposón *Mos1* se amplificó por PCR de *pMos1* (*pMos1* (Medhora M, et al (1991) *Genetics* 128:311-318) y se clonó en el sitio *Sma*I de pUC19; el casete SV40-resistencia a G418 se clona en el sitio *Nru*I], lo que resulta en pRP1388.

#### 40 b) Cultivo Celular y Transfecciones.

Las células humanas HeLa se cultivan en OptiMEM (GIBCO/BRL) suplementado con 5% suero y 50  $\mu$ g/ml penicilina/estreptomicina a 37°C y 5% CO<sub>2</sub>. Los ensayos de transposición en cultivo de tejido se llevan a cabo de una manera similar a la descrita previamente (Ivics Z, et al. (1997) *Cell* 91:1-20).

#### c) Eficacias de los Elementos Tc1/*marinier* Comparado en Células Humanas HeLa.

- 45 Se comparan las eficacias de la transposición de los transposones de *C. elegans* Tc1 y Tc3 y los transposones de insecto *Himar1* y *Mos1* en células humanas HeLa. Todos los ORF de transposasa se clonaron en sitios de restricción idénticos en un vector de expresión CMV. Los transposones correspondientes están todos interrumpidos por un casete SV40-resistencia a G418. El vector de expresión transposasa y el vector transposón correspondiente se cotransfectaron en células humanas HeLa. Después de la cotransfección, se compara el número de colonias resistentes a G418 con el 50 número obtenido después de la cotransfección de un vector de expresión control junto con el vector transposón. Se ha mostrado previamente que el incremento en el número de colonias resistentes está causado por la integración mediada por transposasa del transposón (Schouten GJ, et al. (1998) *Nucleic Acids Res* 26:3687-3693 e Ivics Z, et al. (1997) *Cell* 91:1-20) y, por lo tanto, es indicativo de la eficacia de la transposición.

## 6. Ejemplo 6

### a) Transposasas quiméricas que contienen el dominio de unión al ADN (DBD) $\lambda$ R fusionado de manera precisa en el extremo N o C terminal.

5 El DBD  $\lambda$ R se fusiona directamente en el extremo N o C terminal de una transposasa [p. ej., Tc1 (Referencia No. NM\_061407, A1878683, A1878522, A1794017); elemento P (Rio et al., Cell (1986) 44:21-32; entre otros)]. Las técnicas para construir ácidos nucleicos que codifican proteínas de fusión y la purificación de sus productos son muy conocidas en la bibliografía y para los expertos en la técnica.

10 Una construcción de transposón lineal y circular se combina con su transposasa quimérica respectiva y se añade a extractos celulares de diferentes líneas celulares (p. ej. células HeLa) que contienen el plásmido diana. El sustrato de ADN para ensayar la integración selectiva de sitio de este vector es un plásmido que contiene los operadores  $\lambda$ . Dependiendo del vector transposón (p. ej., Tc1/mariner) que se está usando se hacen modificaciones adicionales a la diana tal como incrementar el número de sitios TA a lo largo del plásmido (incluyendo regiones en, adyacentes a y lejos de su sitio diana putativo) para determinar el efecto en la eficacia de la integración. La transposasa quimérica puede evaluarse para su nivel de integración y especificidad comparado con su transposasa respectiva sin un dominio de unión al ADN fusionado.

## 7. Ejemplo 7

### a) Vector transposón que incluye una transposasa quimérica que contiene dominio de unión al ADN (DBD) $\lambda$ R fusionado de manera precisa en el extremo N o C terminal.

20 El DBD  $\lambda$ R se fusiona directamente en el extremo N o C terminal de una transposasa [p. ej., Tc1 (Referencia No. NM\_061407, A1878683, A1878522, A1794017); elemento P (Rio et al., Cell (1986) 44:21-32; entre otros)]. Las técnicas para construir ácidos nucleicos que codifican proteínas de fusión y la purificación de sus productos son muy conocidas en la bibliografía y para los expertos en la técnica.

25 El vector es un plásmido que contiene un transgén (p. ej. gen de resistencia a antibiótico, p53 o factor VIII) flanqueado por las repeticiones terminales de un transposón que también contiene una transposasa quimérica (p. ej., Bella Durmiente) que contiene el dominio de unión al ADN (DBD)  $\lambda$ R fusionado de manera precisa en el extremo N o C terminal. El sustrato de ADN para ensayar la integración selectiva de sitio es un plásmido que contiene los operadores I y el gen de resistencia a antibiótico que se había introducido previamente en una línea celular (p. ej. células HeLa). Dependiendo de la construcción de transposón usada en el vector que se está usando, se hacen modificaciones adicionales al plásmido diana tales como intercalar sitios TA a lo largo del plásmido (incluyendo regiones en, adyacentes a o lejos de su sitio diana putativo) como se requerirá en la familia de transposones Tc1/mariner. Después de la transfección, ocurrirá la escisión e integración del transgén flanqueado por las repeticiones terminales como se ha descrito previamente en la Figura 1. Esta transposasa quimérica puede evaluarse para su nivel de integración y especificidad comparado con su transposasa respectiva sin un dominio de unión al ADN fusionado.

## 8. Ejemplo 8

### 35 a) Modificación de Dominios de Unión al ADN Existentes para Reconocer Nuevas Secuencias de Reconocimiento Diana

#### (1) Modificación de Dedos de Cinc

40 La modificación de dominios de unión al ADN existentes permite el reconocimiento de muchas secuencias nuevas. Muchas proteínas con dedos de cinc consisten en varias repeticiones en tándem de dedos. En algunas de estas proteínas cada dedo reconoce tres bases de ADN adyacentes (Desjarlais y Berg, 1993, PNAS, USA, 90:2256-2260; y Pavletich y Pabo, 1991, Science, 252:809). Se ha encontrado que estos dedos de cinc pueden "mezclarse y emparejarse" para rendir nuevas especificidades de unión al ADN. Por ejemplo, se han preparado varias proteínas con dedos de cinc que contienen los mismos tres dedos de cinc pero en un orden reorganizado (Desjarlais y Berg, 1993, PNAS, USA, 90:2256-2260). Las proteínas modificadas reconocen los mismos sitios de 9 pares de bases, pero con las unidades de tres pares de bases reorganizadas en el orden predicho. Así, se crean nuevas especificidades mediante la reorganización del orden de los dedos de cinc en proteínas con múltiples dedos. Además, los métodos de selección in vivo para identificar dominios de dedos de cinc que reconocen cualquier sitio diana dado y otras secuencias de aminoácidos de dominios de dedos de cinc que reconocen un sitio particular están contenidas en la Solicitud de Patente No. 20020061512.

50 Además, se han conferido nuevas especificidades de unión al ADN a dedos Cys2-His2 individuales (Rebar y Pabo, 1994, Science, 263:671-673). Por ejemplo, se han usado bibliotecas de exposición en fagos para aislar nuevos dedos que

reconocen una secuencia preseleccionada de una biblioteca que contiene secuencias de reconocimiento aleatorizadas. La proteína con dedo zif-268, que contiene tres dedos de cinc, se aleatorizó en los aminoácidos de un dedo implicado en el contacto con ADN y se expresó en la superficie de un bacteriófago. Se aislaron los fagos capaces de unirse a un sitio de ADN mutante aplicando la biblioteca de exposición en fago a una placa recubierta con el sitio de ADN de interés. Después de lavar, los fagos unidos se eluyeron con alta concentración de sal y se crecieron. La selección se repitió varias veces más, usando como material de partida para la siguiente ronda el fago recuperado de la ronda previa. Al final de este procedimiento, los fagos que quedaban codificaban derivados zif268 con altas afinidades para las nuevas secuencias de ADN. Los expertos en la técnica reconocerán que la combinación de la capacidad de transformar dedos individuales para reconocer nuevos sitios con la capacidad de mezclar y emparejar diferentes dedos permite el diseño de proteínas con dedos de cinc que reconocen cualquier secuencia de ADN.

### (2) Modificación del Dominio de Unión al ADN Zif268 para Reconocer un Elemento L1

El elemento de ADN L1 es un sitio de reconocimiento de dominio de unión atractivo porque se sabe que las interrupciones en dichas secuencias no son dañinas. La proteína zif268 puede modificarse para unirse a la secuencia 5'GGGGCAGGG3', que se encuentra cerca del extremo 3' de los elementos L1 (Hattori et al., 1985, NAR, 13:7813-7827). Siguiendo las reglas de Desjarlais y Berg (Desjarlais y Berg, 1993, PNAS, USA, 90:2256-2260) y usando un nuevo elemento de reconocimiento del trabajo de Rebar y Pabo (Rebar y Pabo, 1994, Science, 263:671-673), se diseña una nueva especificidad para reconocer esta secuencia L1. Específicamente, el cambio de los aminoácidos de reconocimiento críticos en los tres dedos (13, 16 y 19 en cada dedo según la numeración de Desjarlais y Berg) para contener la secuencia RHR (dedo 1), QGS (dedo 2) y RHR (dedo 3) produce una proteína que reconoce la secuencia L1 deseada.

### (3) Modificación del Dominio de Unión al ADN Hin para Reconocer un Elemento L1

Cuando se desea un dominio de unión al ADN relativamente pequeño (p. ej. aproximadamente 50 aminoácidos), puede emplearse evolución dirigida in vitro para modificar el dominio de unión al ADN Hin de 52 aminoácidos para reconocer la secuencia L1 u otras secuencias deseadas. Por ejemplo, el ADN que codifica el dominio de unión al ADN Hin puede clonarse en el vector fUSE2 (Parmley y Smith, 1988, Gene, 73:305-318) y las selecciones se realizan usando los métodos de Rebar y Pabo descritos anteriormente.

## 9. Ejemplo 9

### a) Preparación de un Dominio de Unión a Proteínas que se Une Indirectamente al Ácido Nucleico Diana por Unión a STF-1.

Para producir un primer dominio de unión a proteínas de una proteína quimérica que se une a STF-1, se usa el método de biblioteca de exposición en fago para crear un péptido de aproximadamente 19 aminoácidos que se une fuertemente a STF-1. El péptido de dominio de unión a proteínas STF-1 se fusiona con transposasa para formar una proteína quimérica (como se describe en la presente memoria). La proteína quimérica transposasa-dominio de unión a proteínas STF-1 se ensaya (como se describe en la presente memoria) para determinar si dirige la integración a ADN unido a STF-1 in vitro. Esta proteína quimérica también se incorpora en un plásmido basado en transposón como se describe en la presente memoria y se ensaya intracelularmente en células que expresan STF-1 (p. ej., células obtenidas de páncreas).

## 10. Ejemplo 10

Utilizando secuencias del anfitrión similares (es decir, que no son funcionales) flanqueadas por las repeticiones terminales en un sistema basado en transposón, se puede tomar como diana un gen e interrumpirlo con una eficacia mayor que con los métodos convencionales. El potencial de esta tecnología se extiende obviamente a desarrollar modelos de inactivación génica, determinar funciones de genes, etc. Aquí utilizamos un plásmido basado en transposón que contiene las repeticiones terminales invertidas de un transposón y las secuencias del anfitrión homólogas entre y/o fuera de las repeticiones terminales. Una transposasa contenida en otro plásmido (o su ARNm) se microinyecta en las células (p. ej., embriones de ratón) como describen Dupuy y colaboradores (Dupuy AJ et al. Mammalian germ-line transgenesis by transposition. PNAS 99:4495-4499). Los métodos para determinar el sitio de integración se han descrito previamente en la bibliografía (40, 84 como referencias actualmente). El plásmido basado en transposón con secuencias homólogas se compara con un control sin secuencias homólogas.

## 11. Ejemplo 11

La replicación de ADN viral en muchos virus con ADN bicatenario ocurre a través de un mecanismo de círculo rodante que rinde concatémeros lineales del genoma viral (Furth ME y Wickner SH (1983) Lambda II, 145-155). El ensamblaje de un virus infeccioso requiere la escisión de un único genoma del concatémero y la inserción simultánea del ADN en una cápside preformada (Casjen, S y Hendix, R (1988) The Bacteriophages, 15-92; Black LW (1989) Annu. Rev. Microbiol.

43, 267-292). Las enzimas terminasas son comunes en estos virus y son las responsables del "empaquetamiento" del ADN viral (Black LW (1989) Annu. Rev. Microbiol. 43, 267-292; Fujisawa, H y Morita, M (1997) Genes to Cells 2, 537-545; Catalano, CE (2000) Cellular and Molecular Life Sciences, 57, 128-148). La subunidad pequeña de la terminasa del bacteriófago lambda, gpNu1, es la responsable del ensamblaje específico de sitio de la holoenzima en *cos*, el sitio de inicio del empaquetamiento del genoma lambda (Catalano CE (2000) Cellular and Molecular Life Sciences, 57, 128-148). Se han demostrado interacciones de unión específicas entre gpNu1 y los "elementos R" repetidos en *cos*. La organización de los dominios de gpNu1 se describe en la Figura 9. Los  $\approx 40$  restos C terminales están implicados en las interacciones proteína-proteína con la subunidad grande gpA, mientras que los restos  $\approx 100-140$  definen un dominio de autoensamblaje hidrofóbico de la proteína. Los  $\approx 55$  restos N terminales de la proteína definen el dominio de unión al ADN (DBD) mínimo de la proteína mientras que los restos  $\approx 55-100$  forman una espiral helicoidal extendida que conecta el DBD y el dominio de autoasociación de la proteína.

El análisis de las construcciones de delección de la proteína demostraron que el DBD retiene las interacciones de unión al ADN específicas de *cos*; sin embargo, la delección del dominio de autoasociación disminuye la afinidad de unión al ADN tres órdenes de magnitud (Yang, Q et al. (1999) Biochemistry 38, 465-477; Yang, Q et al. (1999) Biochemistry 38, 14238-14247).

La proteína integrasa del bacteriófago lambda se requiere para la integración específica de sitio del ADN viral en el cromosoma bacteriano durante la lisogenia (Landy, A (1989) Ann. Rev. Biochem. 58, 913-949). La proteína se une de manera específica de sitio a sitios "*att*" tanto en el genoma bacteriano (*attB*) como viral (*attP*). Cada sitio *att* contiene un par invertido de sitios de unión de "tipo núcleo" (9 pb cada uno) separados por una región de "superposición" de 7 pb. Elementos de unión al ADN "brazo" adicionales flanquean los sitios de unión de tipo núcleo. Se ha definido una organización de los dominios de la integrasa para el fago lambda, como sigue. Un dominio amino terminal (restos 1-64) se une con alta afinidad a los sitios de tipo brazo de *att*, mientras un dominio carboxi terminal (C65, restos 65-356) se une con baja afinidad a los sitios de tipo núcleo. Este dominio también posee una actividad topoisomerasa independiente de secuencia. Se ha identificado un dominio catalítico de la proteína que comprende los restos  $\approx 170-356$  (IntC 170) pero este dominio no forma complejos estables con ADN que contiene *att*.

Una construcción que comprende los 85 restos N terminales de gpNu1 (gpNu1 $\Delta$ E85) constituye un dominio de unión al ADN altamente estable de la proteína (Yang, Q et al. (1999) Biochemistry 38, 465-477; Yang, Q et al. (1999) Biochemistry 38, 14238-14247; Los estudios estructurales y biofísicos de éstas y construcciones relacionadas han demostrado que mientras los restos 1-55 forman un dominio globular completamente plegado, los restos  $\approx 55-85$  forman una estructura de hélice extendida y flexible (Bain, 2001 #686); de Beer, T et al. (2002) Mol. Cell 9, 981-991). Esta proteína es así ideal para la construcción de una proteína quimérica que proporciona un dominio de unión al ADN N terminal específico de sitio, unido mediante la hélice flexible al dominio catalítico de la integrasa (IntC170). La construcción que comprende gpNu1 $\Delta$ E85 e IntC170 se refiere como gpNu1 $\Delta$ E85-IntC170 (Figura 10).

Mientras gpNu1 $\Delta$ E85 se une con especificidad a *cos*-ADN, la afinidad de esta construcción para el ADN viral es relativamente baja. Así, se construye gpNu1 $\Delta$ P141-IntC170, una quimera que consiste en los 141 restos N terminales de gpNu1 unidos a IntC170 (Figura 10). Se ha demostrado que gpNu1 $\Delta$ P141 se une a ADN que contiene *cos* con una afinidad igual a la de la proteína de longitud completa (Yang, Q et al. (1999) Biochemistry 38, 14238-14247). Así, la construcción gpNu1 $\Delta$ P141-IntC170, puede proporcionar una proteína quimérica con alta especificidad y afinidad para los sustratos de ADN que contienen *cos*.

Las dos construcciones quiméricas dirigen específicamente el dominio catalítico IntC170 a sustratos de ADN que contienen *cos*. El dominio IntC170 aislado posee una actividad topoisomerasa que no es específica para ninguna secuencia de ADN. Así, las proteínas quiméricas poseen una actividad topoisomerasa que se incrementa significativamente en presencia de la secuencia *cos* del ADN del fago lambda.

La construcción de las proteínas quiméricas gpNu1 $\Delta$ E85-IntC170 y gpNu1 $\Delta$ P141-IntC170 usando tecnología PCR revela que la primera quimera posee el dominio de unión al ADN gpNu1 altamente soluble y la última quimera contiene además el dominio de autoasociación gpNu1 que proporciona una unión de alta afinidad. Se usó ADN de lambda maduro como molde para amplificar las secuencias de ADN viral que codifican la secuencia N terminal de gpNu1 que se extienden de Met1 a Glu85 y Met1 a Pro141, respectivamente. Los cebadores usados se listan en la Tabla 3. Obsérvese que la amplificación del ADN usando estos cebadores proporciona secuencias de restricción *EcoRI* y *HpaI* en los extremos 5' y 3' del producto de PCR, respectivamente. De manera similar, se sintetizaron cebadores que permiten la amplificación de la secuencia C terminal de la integrasa de lambda que se extiende de Ala170 a Lys356 (C170, Tabla 3). La amplificación de IntC170 usando estos cebadores proporcionó secuencias de restricción *NotI* y *HindIII* en los extremos 5' y 3' del producto de PCR, respectivamente.

| Cebador            | Secuencia   | Producto |
|--------------------|---|----------|
| gpNu1 Directo      | TTCTCC- <i>GAATTC</i> -ATG-GAA-GTC-AAC-AAA-AAG-C ( <i>EcoRI</i> )       | -        |
| gpNul-E85 Inverso  | TCCTTC- <i>GTTAAC</i> -TTC-GTA-CTC-AAT-AGT-TCC-T ( <i>HpaI</i> )        | 258 pb   |
| gpNul-P141 Inverso | TCCTTC- <i>GTTAAC</i> -CGG-AAA-ACG-CCG-CTG-C ( <i>HpaI</i> )            | 426 pb   |
| Integrasa Directa  | AAGAAT- <i>GCGGCCGC</i> -GCA-GCA-AAA-TCA-GAG-GTA ( <i>NotI</i> )        | -        |
| Integrasa Inversa  | ATTAAT- <i>AAGCTT</i> -TTA-TTT-GAT-TTC-AAT-TTT-GTC-C ( <i>HindIII</i> ) | 590 pb   |

**Tabla 3. Cebadores usados para la amplificación por PCR.** La secuencia en *itálica* indica las secuencias de reconocimiento relevantes de la enzima de restricción.

5 La Figura 11 muestra que la amplificación de todas las secuencias apropiadas se ha logrado con éxito. Los productos de PCR se han purificado, digerido con las endonucleasas de restricción apropiadas y purificado de nuevo por electroforesis en gel de agarosa. Estos productos de PCR se clonan en el plásmido pKKT7(-H) usando el protocolo presentado en la Figura 12.

10 Para construir los vectores de expresión de proteínas como se indica en la Figura 12, se transforman células de *E. coli* DH5 $\alpha$  con la mezcla de ligación y se aísla el ADN plasmídico de las colonias resistentes a ampicilina. Los plásmidos se analizan por análisis de digestión con restricción y secuenciación de ADN para verificar la presencia de los insertos apropiados. Una vez que se ha verificado la secuencia del vector, los plásmidos se usan para transformar células de *E. coli* BL21(DE3) y las proteínas quiméricas se expresan. Los estudios iniciales examinarán la solubilidad y estabilidad de la proteína.

## 12. Ejemplo 12: Transposición de transposasa-dominio de unión al ADN en líneas celulares

### a) Transformación de líneas celulares con resistencia a neomicina

15 Para seleccionar la supervivencia en presencia de G418, se transfectan líneas celulares con un vector basado en transposón con resistencia a neomicina  $\pm$  transposasa nativa, para medir la frecuencia de la transposición y la recombinación heteróloga. La frecuencia de la recombinación homóloga se mide usando un vector que incluye secuencias homólogas al sitio genómico humano pretendido. A continuación, las líneas celulares se transfectan con vectores transposones con resistencia a neomicina más transposasa quimérica. Se entiende que una frecuencia de supervivencia de siete días mayor que la línea base es una evidencia putativa de transposición más allá de los valores de la línea base de recombinación homóloga o heteróloga. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio.

### b) Transformación de líneas celulares con beta-galactosidasa

25 Se transfectan líneas celulares con un vector basado en transposón CMV/beta-galactosidasa  $\pm$  transposasa quimérica o la transposasa nativa. Se entiende que la actividad específica beta-galactosidasa a los siete días en los extractos celulares es una evidencia putativa de transposición. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio.

### c) Transformación de líneas celulares con luciferasa

30 Se transfectan líneas celulares con un vector basado en transposón CMV/luciferasa  $\pm$  transposasa quimérica o la transposasa nativa. Se entiende que la actividad específica luciferasa a los siete días en los extractos celulares es una evidencia putativa de transposición. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio.

### d) Transformación de líneas celulares con proteína verde fluorescente

Se transfectan líneas celulares con una mezcla de un vector MMTV/EGFP inducible con dexametasona  $\pm$  transposasa quimérica o la transposasa nativa. Se entiende que la intensidad de fluorescencia de EGFP inducible en las células vivas a los siete días es una evidencia putativa de transposición. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio.

### 5 13. Ejemplo 13: Transposición con integración dirigida por transposasa-TnsD en líneas celulares

#### a) Transformación de líneas celulares con resistencia a neomicina

10 Para seleccionar la supervivencia en presencia de G418, se transfectan líneas celulares con un vector Tn7 con resistencia a neomicina, para medir la frecuencia de la recombinación heteróloga en la línea base. La frecuencia de la recombinación homóloga se mide usando un vector que incluye secuencias homólogas al sitio genómico humano pretendido. A continuación, las líneas celulares se transfectan con vectores Tn7 con resistencia a neomicina más las proteínas de transposición TnsA, TnsB, TnsC y TnsD de Tn7. Se entiende que una frecuencia de supervivencia de siete días mayor que la línea base es una evidencia putativa de transposición más allá de los niveles de la línea base de recombinación homóloga o heteróloga. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio. La secuencia del ADN del dominio de reconocimiento de TnsD se altera para optimizar el reconocimiento de la secuencia diana humana parecida y se determina el dominio de unión al ADN mínimo de TnsD.

#### 15 b) Transformación de líneas celulares con beta-galactosidasa

20 Después de optimizar la afinidad de TnsD para su diana humana y de purificar la proteína de segunda generación, se transfectan líneas celulares con una mezcla de un vector CMV/beta-galactosidasa  $\pm$  proteínas de transposición purificadas TnsA, TnsB, TnsC y TnsD. Se entiende que la actividad específica beta-galactosidasa a los siete días en los extractos celulares es una evidencia putativa de transposición. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio.

#### c) Transformación de líneas celulares con luciferasa

25 Se construye un panel de vectores de expresión para producir quimeras dominio de unión al ADN de TnsD-transposasa. La quimera dominio de unión al ADN de TnsD-transposasa más eficaz se trunca, particularmente en el dominio TnsD para determinar la quimera activa con el menor tamaño. De manera similar, el espaciamiento entre los dominios TnsD y transposasa se varía para determinar el espaciamiento más eficaz para la inserción en el sitio deseado. Para la transposasa, puede usarse TnsB (es decir, el dominio catalítico de Tn7), maT, transposasa Tn5, integrasa del fago  $\lambda$ , integrasa del fago  $\mu$  o integrasa de VIH. TnsD puede estar orientado 3' ó 5' de la transposasa. Después, se transfectan líneas celulares con una mezcla de un vector CMV/luciferasa  $\pm$  dominio de unión al ADN TnsD-transposasa purificado.

30 Se entiende que la actividad específica luciferasa a los siete días en los extractos celulares es una evidencia putativa de transposición. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio.

#### d) Transformación de líneas celulares con proteína verde fluorescente

35 Se transfectan líneas celulares con una mezcla de un vector MMTV/EGFP inducible con dexametasona  $\pm$  proteínas TnsD-transposasa purificadas. Se entiende que la intensidad de fluorescencia de EGFP inducible en las células vivas a los siete días es una evidencia putativa de transposición. Los sitios de inserción se determinan para ensayar la hipótesis de transposición específica de sitio.

### 14. Ejemplo 14: Diseño y análisis de una recombinasa con especificidad de sitio alterada

#### a) Resultados

40 La modificación precisa de los genomas de mamíferos tiene gran importancia en la terapia génica (J.M. Kaminsky, et al., (2002) *Faseb J.* 16:1242-1247). Actualmente, la modificación específica de sitio de los genomas eucariotas se basa en la recombinación homóloga, que es demasiado ineficaz para ser útil en los métodos de terapia génica. Por lo tanto, se ha alterado la especificidad de sitio de la recombinasa específica de sitio más eficaz, concretamente Cre. Cre es un producto génico de 34 kDa del bacteriófago P1 e interacciona con sitios denominados loxP como parte del ciclo biológico del bacteriófago (K. Abremski, et al., (1983) *Cell* 32:1301-1311). No requiere ningún cofactor y cataliza los intercambios bicatenarios de ADN in vitro y en células de cualquier origen. La superficie de la proteína con la que interacciona Cre con el sitio loxP está distribuida a lo largo de una gran parte de la proteína (F. Guo, et al., (1997) *Nature* 389:40-46) y los intentos de alterar la especificidad de sitio por mutagénesis de los aminoácidos implicados en la interacción ADN proteína no han tenido éxito.

50 Por lo tanto, se intentó aumentar en lugar de reemplazar la capacidad de unión al ADN de Cre añadiendo el dominio de unión al ADN (DBD) del factor de transcripción con dedo de Cinc humano Zif268 (H.A. Greisman, et al., (1997) *Science*

275:657-661). El DBD de zif268 se fusionó con el extremo N terminal de la recombinasa Cre de manera que se genera un marco de lectura abierto contiguo (Fig. 13). Para facilitar la purificación y la estabilización de la proteína, se añadió un fragmento del marco de lectura abierto de la proteína de unión a maltosa (MBP) de *E. coli* al extremo N terminal de la proteína (Fig. 13). La proteína de fusión MBP-Cre muestra una estabilidad mejorada a la vez que retiene una actividad catalítica completa *in vitro* e *in vivo* (A.F. Kolb y S.G. Siddell, (1996) *Gene* 183:53-60).

La proteína MBP-Zif-Cre se purificó mediante una columna de amilosa (Fig. 14) y se analizó por su capacidad de catalizar la recombinación específica de sitio entre sitios loxP. Este análisis demuestra que la adición del DBD de Zif268 a la proteína MBP-Cre no anula su actividad catalítica (Fig. 15). Se generaron sitios de unión quiméricos que consisten en sitios loxP y sitios de unión Zif268, que están separados por segmentos espaciadores de longitud variada (Fig. 15). Los dominios de unión estaban orientados de manera que eran compatibles con los dominios de unión al ADN yuxtapuestos de la proteína MBP-Zif-Cre. Los sitios de unión quiméricos pueden analizarse para afinidad hacia la proteína de fusión y para su capacidad de recombinarse por la proteína de fusión. La construcción completa se muestra en la Figura 16.

#### b) Métodos

El plásmido pMALc2-Cre que codifica la proteína de fusión MBP-Cre se ha descrito anteriormente. El dominio de unión al ADN de Zif268 se ha escindido del plásmido pB-Zif268 como un fragmento SpeI/PstI de 283 pb y se ligó con el plásmido pMALc2-Cre digerido con PstI/XbaI. El plásmido resultante pMALc2-Zif-Cre se transformó en células *E. coli* TB1. Con el fin de producir la proteína MBP-Zif-Cre, se crecieron bacterias hasta una DO600 de 0,5 en un volumen total de 100 ml y se indujeron con una concentración final de 0,3 mM de IPTG. Las células se lisaron por sonicación en un tampón que contiene 20 mM Tris-HCl pH 7,4, 200 mM NaCl y 1 mM EDTA. El lisado se aclaró por centrifugación y la proteína se purificó mediante una columna de amilosa como se ha descrito (A.F. Kolb y S.G. Siddell, (1996) *Gene* 183:53-60).

#### 15. Ejemplo 15

Las composiciones descritas en la presente memoria pueden utilizar integrasa obtenida del bacteriófago Mu y otros elementos que comprenden un complejo donante escindido (CDC) activo y que comprende además un mecanismo de direccionamiento por el que la integración de un casete transponible Mu puede ser dirigido a un sitio diana predeterminado en el genoma de un organismo anfitrión. Estos vectores de integración comprenden un casete transponible Mu y una muA quimérica del bacteriófago. Los métodos de la invención utilizan los vectores de integración de la invención para insertar el casete transponible Mu en un sitio diana del genoma de un organismo. Esta inserción ocurre en ausencia de la proteína auxiliar MuB. Los métodos son útiles para modular la actividad de genes conocidos y para la integración dirigida de secuencias de nucleótidos de interés en una localización específica del genoma de un organismo. De acuerdo con esto, los métodos también pueden usarse para crear interrupciones e inactivaciones génicas.

Estos vectores de integración comprenden un complejo donante escindido (CDC) Mu y una "transposasa quimérica" que proporciona la transposición del casete transponible Mu de una manera específica de sitio y en ausencia de la proteína auxiliar MuB (Suzuki, Hideki; et al Solicitud de Patente de los Estados Unidos 20020132350).

Los complejos donantes escindidos (CDC) activos pueden obtenerse usando una reacción de transposición *in vitro* y un plásmido mini-Mu como donante del transposón. Por "plásmido mini-Mu" se pretende un plásmido que comprende un casete transponible Mu flanqueado por un dominio de ADN que no es del plásmido MU. Dichos plásmidos mini-Mu pueden construirse usando técnicas de biología molecular muy conocidas en la técnica. Véanse particularmente Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2ª Ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.); y Ausubel et al., eds. (1995) *Current Protocols in Molecular Biology* (Greene Publishing and Wiley Interscience, Nueva York).

Las composiciones de la invención comprenden nuevos vectores de integración que se obtienen de CDC del bacteriófago atenuado Mu, un transposón de clase III bacteriano de *Escherichia coli*. Este transposón presenta una frecuencia de transposición extremadamente alta (Toussaint y Rsibois (1983) en *Mobile Genetic Elements*, ed. Shapiro (Academic Press, Nueva York), p. 105-158). El bacteriófago Mu con su genoma de aproximadamente 37 kb es relativamente grande comparado con otros transposones. Mu codifica dos productos génicos que están implicados en el proceso de transposición: transposasa MuA, una proteína multidominio de 70 kDa y 663 aminoácidos y MuB, una proteína auxiliar de aproximadamente 33 kDa. Este elemento transponible tiene secuencias de reconocimiento MuA en el extremo izquierdo y en el extremo derecho (denominadas "L" y "R", respectivamente) que flanquean el casete transponible Mu, la región del transposón que se integra finalmente en el sitio diana. A diferencia de otros transposones conocidos en la técnica, estos extremos no son secuencias repetidas invertidas. El casete transponible Mu, cuando es necesario, puede incluir una secuencia amplificadora de la transposición (también referida en la presente memoria como

la secuencia activadora interna, o "IAS") localizada aproximadamente 950 pares de bases hacia el interior desde la secuencia de reconocimiento del extremo izquierdo.

Las secuencias de reconocimiento del extremo izquierdo y derecho del transposón Mu engloban cada una tres sitios de unión a la transposasa MuA de "tipo extremo" de 22 pares de bases, designados attL1 ("L1"), attL2 ("L2") y attL3 ("L3"); y attR1 ("R1"), attR2 ("R2") y attR3 ("R3"), que se numeran desde los extremos del casete transponible Mu hacia en el interior (véase la Fig. 1). Existen dos sitios de dinucleótidos de escisión del ADN fuera del casete transponible Mu, situados a 6 pb de distancia de los sitios de unión MuA más extremos L1 y R1. La secuencia amplificadora de la transposición de Mu también se une a la transposasa MuA, pero en un dominio diferente de la proteína que el usado para unirse a las secuencias de reconocimiento del extremo izquierdo y derecho. La transposasa MuA interacciona con las secuencias de reconocimiento del extremo izquierdo y derecho flanqueantes y la secuencia amplificadora de la transposición para llevar a cabo la inserción del casete transponible Mu en una secuencia de ADN diana.

La transposición es una característica esencial del ciclo vital del bacteriófago Mu. La integración del ADN infeccioso de Mu en el cromosoma de un anfitrión para formar un lisógeno estable ocurre por inserción simple no replicativa (Liebart et al. (1982) Proc. Natl. Acad. Sci. USA 79:4362-4366; Harshey (1984) Nature 311:580-581. Durante el crecimiento lítico, Mu genera múltiples copias de su genoma por ciclos repetidos de transposición replicativa (Ljungquist y Bukhari (1977) Proc. Natl. Acad. Sci. USA 74:3143-3147) a través de una ruta cointegrada (Chaconas et al. (1981) J. Mol. Biol. 150:341-359). Ambos tipos de transposición están facilitados por la transposasa MuA y la proteína auxiliar MuB. Las proteínas codificadas por E. coli tales como la proteína semejante a histona ("HU") y el factor de integración del anfitrión (IHF) ayudan en los cambios conformacionales tempranos que finalmente dan lugar a la transferencia del casete transponible Mu en la secuencia de ADN diana del anfitrión.

Los detalles de la transposición de Mu se han aclarado usando una reacción de transposición in vitro (Mizuuchi (1983) Cell 35:785-794; Mizuuchi (1984) Cell 39:395-404; Craigie y Mizuuchi (1985) Cell 41:867-876; Craigie et al. (1985) Proc. Natl. Acad. Sci. USA 82:750-7574; revisado por Chaconas et al. (1996) Curr. Biol. 6:817-820; Craigie (1996) Cell 85:137-140; Lavoie y Chaconas (1995) Curr. Topics Microbiol. Immunol. 201:83-99; y Mizuuchi (1992) Annu. Rev. Biochem. 61:1011-1051). En esta reacción in vitro, por ejemplo, el donante del transposón es un plásmido mini-Mu y otra molécula de ADN, comúnmente forma replicativa de ADN .phi.X174, sirve como la diana de la transposición. El plásmido mini-Mu se construye de manera que comprende dos dominios de ADN. El primero de estos dominios de ADN es un casete transponible Mu, que está flanqueado por el segundo dominio de ADN, referido en la presente memoria como el dominio de ADN del plásmido no perteneciente a Mu.

Usando un sistema in vitro, se ha mostrado que normalmente la transposasa MuA existe en su estado monomérico inerte que no reconoce los sitios de escisión del ADN adyacentes a las secuencias de reconocimiento del extremo izquierdo y extremo derecho del casete transponible Mu. En presencia de HU, IHF y iones metálicos divalentes, particularmente  $Mg^{2+}$ , la transposasa MuA se une inicialmente a la secuencia amplificadora de la transposición de Mu y a las secuencias de reconocimiento del extremo izquierdo y derecho. Después de esta unión, el plásmido mini-Mu sufre una serie de cambios conformacionales que resultan finalmente en la formación del complejo donante escindido (CDC).

En la transposición normal del bacteriófago Mu, el núcleo estructural y funcional de CDC es una unidad tetramérica de moléculas MuA (Lavoie et al. (1991) EMBO J. 10:3051-3059; Mizuuchi (1992) Annu. Rev. Biochem. 61:1011-1051; Baker et al. (1993) Cell 74:723-733, referido de aquí en adelante en la presente memoria como el núcleo tetramérico de MuA. Los tres sitios de unión de tipo final de la transposasa MuA designados attL1, attR1 y attR2 se consideran los sitios de unión del núcleo, ya que se unen establemente con el tetrámero MuA. Cuando la proteína MuA interacciona con los otros tres sitios de unión de tipo final de la transposasa MuA (attL2, attL3 y attR3) se une débilmente. Estas moléculas de MuA unidas débilmente pueden eliminarse bien con heparina, alta concentración de sal (0,5 M NaCl) o exceso de ADN competidor de extremo de Mu (Kuo et al. (1991) EMBO J. 10:1585-1591; Lavoie et al. (1991) EMBO J. 10:3051-3059; Mizuuchi et al. (1991) Proc. Natl. Acad. Sci. USA 88:9031-9035). Los tres sitios L1, L2 y L3 se consideran sitios auxiliares, ya que son dispensables individualmente y no se requieren para la reacción de transferencia de cadena intermolecular (Allison y Chaconas (1992) J. Biol. Chem. 267:19963-19970; Lavoie et al. (1991) EMBO J. 10:3051-3059; y Mizuuchi et al. (1991) Proc. Natl. Acad. Sci. USA 88:9031-9035). Sin embargo, los sitios R1, R2 y R3 pueden intercambiarse con los sitios L1, L2 y L3 para usarse en la construcción de plásmidos y para preparar los complejos donante escindidos activos de esta invención.

En el sistema in vitro, así como en células bacterianas, la proteína codificada por Mu MuB se une al ADN diana de una manera no específica en presencia de ATP. De acuerdo con esto, en el sistema in vitro, MuB se une a la molécula de ADN diana, mientras que in vivo se une al ADN del anfitrión. La forma unida a ADN de MuB tiene una fuerte afinidad por el CDC de Mu y así cuando está presente, MuB introduce el CDC en la molécula diana o genoma del anfitrión al que MuB se ha unido. Debido a la unión no específica de MuB, la introducción del CDC ocurre con poca preferencia de

diana. MuB también estimula las actividades de degradación del ADN y unión del ADN de MuA (Adzuma y Mizuuchi (1988) *Cell* 53:257-266; Baker et al. (1991) *Cell* 65:1003-1013; Maxwell et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:699-703; Surette y Chaconas (1991) *J. Biol. Chem.* 266:17306-17313; Surette et al. (1991) *J. Biol. Chem.* 266:3118-3124; y Wu y Chaconas (1992) *J. Biol. Chem.* 267:9552-9558; y Wu y Chaconas (1994) *J. Biol. Chem.* 269:28829-28833). Así, las moléculas de ADN unidas a MuB son dianas preferentes de la transposición de Mu. En ausencia de MuB, la introducción del CDC a un sitio de ADN diana también ocurre pero está limitada principalmente a las reacciones intramoleculares que tienen lugar en regiones adyacentes fuera del ADN de Mu.

La transferencia real del casete transponible Mu del CDC en el sitio de ADN diana está mediada por la transposasa MuA quimérica unida en el CDC. Aunque la invención no está sujeta a ninguna teoría o mecanismo de acción, se cree que los extremos 3' OH expuestos del CDC actúan como nucleófilos, atacando el enlace fosfodiéster del núcleo del ADN diana. Este ataque de un grupo fosfato por el grupo 3' OH expuesto forma un enlace entre los extremos 3' del ADN de Mu y los extremos 5' del ADN diana. Este proceso se refiere como una transferencia de cadenas y resulta en la formación de un complejo de transferencia de cadena (STC). Este complejo de nucleoproteína estable está implicado tanto en la cointegración como en la inserción simple (véase, generalmente, Haren et al. (1999) *Ann. Rev. Microbiol* 53:245-281). Los cointegrados se hacen por replicación de la parte de casete transponible Mu del STC, usando los extremos 3' libres del ADN diana como cebadores para la síntesis de ADN liderada por cadena. Los insertos simples se forman a partir del STC por degradación del dominio de ADN del plásmido no perteneciente a mini-Mu que flanquea la parte de casete transponible Mu de la molécula donante, seguido de reparación de huecos.

Los vectores de integración de la presente invención comprenden complejos de donante escindido (CDC) "activos" del bacteriófago Mu con la transposasa quimérica muA de manera que la inserción del casete transponible Mu en el genoma de un organismo anfitrión ocurre de una manera específica de sitio y en ausencia de la proteína auxiliar MuB. Esta integración puede ocurrir en ausencia de la expresión in vivo de la transposasa quimérica MuA porque CDC activo tiene unido el núcleo tetramérico intacto de MuA quimérica. Estos nuevos vectores de integración permiten la inserción del casete transponible Mu completo en un sitio diana predeterminado en el genoma de cualquier organismo anfitrión y así puede referirse como "CDC dirigidos". Por "sitio diana predeterminado" se pretende una localización deseada en el genoma del organismo anfitrión para la inserción del casete transponible Mu. Las localizaciones deseadas en el genoma incluyen, por ejemplo, localizaciones en las secuencias de ADN cromosómico, secuencias episomales (p. ej., plásmidos replicables o intermedios de replicación viral) y secuencias de ADN de cloroplastos y mitocondrias. Por "predeterminado" se pretende que el sitio diana puede seleccionarse por el investigador tomando como base información de secuencia conocida o predicha.

Los complejos donantes escindidos (CDC) activos pueden obtenerse usando una reacción de transposición in vitro y un plásmido mini-Mu como donante del transposón. Por "plásmido mini-Mu" se pretende un plásmido que comprende un casete transponible Mu flanqueado por un dominio de ADN que no es del plásmido MU. Dichos plásmidos mini-Mu pueden construirse usando técnicas de biología molecular muy conocidas en la técnica. Véanse particularmente Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2ª Ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.); y Ausubel et al., eds. (1995) *Current Protocols in Molecular Biology* (Greene Publishing and Wiley Interscience, Nueva York).

Puede usarse cualquier plásmido o plásmido mini-Mu para obtener los CDC, siempre que comprenda los elementos necesarios en el casete transponible Mu para la formación de un CDC activo. Por "CDC activo" se pretende un CDC que es capaz de llevar a cabo la transferencia de cadena intermolecular o intramolecular en una reacción de transposición in vitro. Dichos CDC activos, cuando se modifican para obtener los vectores de integración de la presente invención, respaldarán la transferencia de cadena intermolecular in vivo. Los elementos necesarios para la formación de CDC activos dependerán de las condiciones de reacción usadas durante la formación de CDC in vitro (véase, por ejemplo, Baker y Mizuuchi (1992) *Genes and Develop.* 6:2221-2232; Wu y Chaconas (1997) *J. Mol. Biol.* 267:132-141). Sin embargo, es posible obtener un CDC activo usando un casete transponible Mu cuyos extremos se definen bien por las secuencias de reconocimiento MuA izquierda o derecha. Además, si se usan casetes preescindidos, es posible obtener la integración en el genoma (es decir, un CDC activo) que retiene menos del conjunto completo de los tres sitios de unión de la secuencia o secuencias de reconocimiento MuA izquierda o derecha.

Así, también se describe que un CDC activo se obtiene usando un plásmido mini-Mu de tipo salvaje. Por "plásmido mini-Mu de tipo salvaje" se pretende el plásmido mini-Mu que tiene un casete transponible Mu que comprende las secuencias de reconocimiento completas Mu del extremo izquierdo y derecho en su orientación natural (es decir, invertida); estas secuencias de reconocimiento flanquean una secuencia de nucleótidos interna que comprende la secuencia amplificadora Mu de la transposición. Por "secuencias de reconocimiento completas Mu de los extremos izquierdo y derecho" se pretende cada una de las secuencias de reconocimiento terminales que comprende los tres sitios de unión a transposasa MuA de tipo extremo naturales de 22 pares de bases. Así, la secuencia de reconocimiento del extremo izquierdo comprende los sitios de unión a transposasa MuA de tipo terminal attL1, attL2, attL3, mientras que la secuencia de reconocimiento del extremo derecho comprende los sitios de unión a transposasa MuA de tipo terminal attR1, attR2 y

attR3. Cuando están presentes, las secuencias de reconocimiento terminales completas permiten la formación de un CDC activo que tiene la transposasa quimérica MuA unida de manera estable a los sitios de unión del núcleo attL1, attR1 y attR2 para formar el núcleo tetramérico de MuA y los monómeros de transposasa quimérica MuA unidos débilmente a los sitios de unión de transposasa de tipo extremo MuA auxiliares attL2, attL3 y attR3. Las secuencias de los pares de bases para las secuencias de reconocimiento completas Mu de los extremos izquierdo y derecho y del amplificador de la transposición Mu se conocen en la técnica. Véase Kahmann y Kamp (1979) *Nature* 280:247-250 y Allet (1978) *Nature* 274:553-558 para las secuencias de reconocimiento Mu del extremo izquierdo y derecho; obsérvese, sin embargo, que estas dos referencias contienen errores de secuenciación. La secuencia correcta se encuentra en No. de Registro GenBank AF083977 (secuencia del bacteriófago Mu, proporcionada por Grimaud (*Virology* 217:200-210 (1996) y Morgan et al., envío directo (13 ago, 1998)). Véase también, Mizuuchi y Mizuuchi (1989) *Cell* 58:399-408 para la secuencia amplificadora de la transposición Mu. Sin embargo, un experto en la técnica se dará cuenta que la secuencia de nucleótidos exacta de estas secuencias de reconocimiento pueden variar ligeramente y que no existe un requerimiento de secuencia exacto para los dominios de unión individuales. Así, por ejemplo, la secuencia de reconocimiento del extremo izquierdo comprende tres sitios de unión de transposasa MuA de tipo terminal que residen en los nucleótidos 1-180 del No. de Registro de GenBank AF083977 y la secuencia de reconocimiento del extremo derecho comprende tres sitios de unión de transposasa MuA de tipo terminal que residen en los nucleótidos 36641-36662 del No. de Registro de GenBank AF083977. En una realización de la invención, los sitios de unión de transposasa MuA en la secuencia de reconocimiento del extremo izquierdo están representados por los nucleótidos 6-27 (attL1), 111-132 (attL2) y 151-172 (attL3), respectivamente, del No. de Registro de GenBank AF083977; y los sitios de unión de transposasa MuA en la secuencia de reconocimiento del extremo derecho están representados por los nucleótidos 36691-36712 (attR1), 36669-36690 (attR2) y 36641-36662 (attR3), respectivamente, del No. de Registro de GenBank AF083977. Un experto en la técnica se dará cuenta de que pueden emplearse variaciones de estas secuencias en la invención siempre que se consiga el resultado deseado. Así, pueden emplearse secuencias que tienen al menos 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad de secuencia con las secuencias Mu nativas.

El uso de un plásmido mini-Mu de tipo salvaje para formar un CDC activo permite que la reacción de transposición in vitro se lleve a cabo bajo condiciones de reacción estándar. Para condiciones de reacción estándar, véanse, Mizuuchi et al (1992) *Cell* 70:303-311 y Surette y Chaconas (1992) *Cell* 68:1101-1108. Cuando se usa un plásmido mini-Mu de tipo salvaje en la reacción de transposición in vitro bajo condiciones estándar, el plásmido mini-Mu debe estar superenrollado negativamente para formar un CDC activo. Sin embargo, este requerimiento de superenrollamiento bajo condiciones de reacción estándar puede evitarse bajo otras condiciones de reacción, por ejemplo, incluyendo DMSO en la mezcla de reacción. Véase Baker y Mizuuchi (1992) *Genes and Develop.* 6:2221-2232.

También se describe que un CDC activo se obtiene usando un derivado de plásmido mini-Mu. Por "derivado de plásmido mini-Mu" se pretende un plásmido mini-Mu que tiene un casete transponible Mu que carece de una o más de las características del casete transponible Mu encontrado en un plásmido mini-Mu de tipo salvaje. Por "características" se pretende lo siguiente: (1) una secuencia de reconocimiento del extremo izquierdo completa, (2) una secuencia de reconocimiento del extremo derecho completa, (3) las secuencias de reconocimiento del extremo izquierdo y derecho en su orientación natural (es decir, invertidas), y (4) una secuencia amplificadora de la transposición Mu en la secuencia de nucleótidos interna que está flanqueada por las secuencias de reconocimiento del extremo izquierdo y derecho. Así, por ejemplo, un derivado de plásmido mini-Mu que carece de una secuencia de reconocimiento del extremo izquierdo o derecho completas carece de uno o más de los sitios de unión de transposasa de tipo terminal MuA en su casete transponible Mu.

Cuando se usa un derivado de plásmido mini-Mu para obtener un CDC activo, las condiciones de reacción requeridas en una reacción de transposición in vitro dependerán de qué característica del plásmido mini-Mu de tipo salvaje está ausente del casete transponible Mu. Así, cuando la única característica ausente es el sitio de unión de transposasa de tipo terminal MuA auxiliar attR3, las condiciones de reacción estándar rendirán un CDC activo que respalda la transferencia de cadena intermolecular (Baker y Mizuuchi (1992) *Genes and Develop.* 6:2221-2232).

Otros derivados de plásmidos mini-Mu que tienen características adicionales delecionadas del casete transponible Mu pueden usarse para obtener un CDC activo variando las condiciones de reacción in vitro. Por ejemplo, cuando se incluye dimetilsulfóxido (DMSO) en la reacción de transposición bajo condiciones de reacción estándar, los plásmidos mini-Mu que carecen del amplificador de la transposición Mu, que portan sólo una secuencia de reconocimiento del extremo izquierdo o derecho Mu completa, que portan sólo un único sitio de unión de transposasa de tipo terminal MuA adyacente a un sitio de escisión de ADN con o sin el amplificador de la transposición Mu, o que tienen las secuencias de reconocimiento del extremo izquierdo y derecho en orientación directa (en lugar de orientación invertida) pueden usarse para formar un CDC que es activo en las etapas de escisión del ADN y transferencia de cadena requeridas para la transposición intermolecular. Véase Baker y Mizuuchi (1992) *Genes and Develop.* 6:2221-2232. También se describe que el sitio de escisión del ADN puede ser un sitio que es reconocido y escindido por la proteína quimérica MuA o puede ser un sitio que es un sitio de reconocimiento de una enzima de restricción; así, los sitios de escisión del ADN usados

pueden ser nativos respecto a la secuencia de ADN en la que están localizados o pueden haberse obtenido por ingeniería o haberse añadido artificialmente a la secuencia en la que están localizados.

De acuerdo con esto, puede usarse cualquier plásmido o plásmido mini-Mu que rinde un CDC activo como la base para obtener las construcciones de ácidos nucleicos de integración de la invención. Los ejemplos de plásmidos mini-Mu de tipo salvaje que pueden usarse incluyen, pero no están limitados a, el basado en pBR322 pBLO7 (7,2 kb; Lavoie (1993) en Structural Aspects of the Mu Transposome (Universidad de Western Ontario, London, Canadá); el basado en pUC19 pBL03 (6,5 kb; Lavoie y Chaconas (1993) Genes Dev. 7:2510-2519; pMK:586 (Mizuuchi et al. (1991) Proc. Natl. Acad. Sci. USA 88:9031-9035); pMK108 (Mizuuchi (1983) Cell 35:785-794; Craigie y Mizuuchi (1986) Cell 45:793-800; pCL222 (Chaconas et al. (1981) Gene 13:37-46); y el basado en pBR322 pGG215 (7,1 kb; Surette et al. (1987) Cell 49:253-262). Los ejemplos de derivados de plásmidos mini-Mu que tienen uno o más sitios de unión MuA y/o la secuencia amplificadora de la transposición incluyen, pero no están limitados a, pBL05 (sitios de unión de transposasa MuA attR3 delecionado de pBL03; Allison y Chaconas (1992) J. Biol. Chem. 267:19963-19970); pMK426 (que porta dos secuencias de reconocimiento del extremo derecho Mu; Craigie y Mizuuchi (1987) Cell 51:493-501); pMK412 (pMK108 con la secuencia amplificadora de la transposición Mu eliminada; Mizuuchi y Mizuuchi (1989) Cell 58:399-408); y pMK395 (mini-Mu con una orientación incorrecta relativa de las dos secuencias terminales Mu; Craigie y Mizuuchi (1986) Cell 45:793-800; y otros descritos en Mizuuchi y Mizuuchi (1989) Cell 58:399-408. También son adecuados para la formación de un CDC mutante activo los derivados de pUC19 que portan sitios de unión MuA específicos, tales como los derivados descritos por Baker y Mizuuchi et al. (1992) Genes and Develop. 6:2221-2232. Todas las referencias anteriores que describen dichos plásmidos mini-Mu.

Cuando se desea la producción in vitro de CDC activos, el plásmido mini-Mu resultante se somete a las etapas iniciales de la reacción de transposición in vitro para formar un complejo donante escindido (CDC) activo. Los métodos para producir CDC activos son muy conocidos en la técnica. Véanse particularmente Craigie et al. (1985) Proc. Natl. Acad. Sci. USA 82:7570-7574; Wu y Chaconas (1997) J. Mol. Biol. 267:132-141. La reacción de transposición puede llevarse a cabo bajo condiciones de reacción estándar (Craigie et al. (1985) Proc. Natl. Acad. Sci. USA 82:7570-7574) o bajo condiciones de reacción modificadas (tales como con la adición de DMSO o glicerol; véase, por ejemplo, Mizuuchi y Mizuuchi (1989) Cell 58:399-408) para obtener un CDC activo.

Los CDC activos pueden obtenerse in vivo (es decir, en la célula anfitriona) en la que se introduce la MuA quimérica o se expresa en una célula en la que también está presente el ADN de un plásmido mini-Mu u otro plásmido capaz de formar un CDC activo. En algunas realizaciones, por ejemplo, la formación de CDC activos de ADN de un plásmido mini-Mu integrado previamente en el genoma del organismo anfitrión podría resultar en la deleción de la mayor parte del ADN integrado previamente y podría resultar también en la reintegración del CDC activo recién formado en una localización diferente del genoma del anfitrión.

Por ejemplo, cuando se desea la producción in vitro de CDC activos, se incuba un plásmido mini-Mu de interés con la proteína transposasa quimérica MuA purificada y la proteína HU de *E. coli* o variantes o fragmentos biológicamente activos de éstas como se define a continuación, en presencia de un ión metálico divalente tal como  $Mg^{2+}$  o  $Mn^{2+}$  (Mizuuchi et al. 1992 Cell 70:303-311). Cuando el casete transponible Mu comprende una secuencia amplificadora de la transposición Mu, la proteína IHF purificada de *E. coli* o una variante de ésta también se incluye en la reacción de incubación. Después de la formación de los CDC, la reacción se termina por la adición de EDTA (véase Wu y Chaconas (1997) J. Mol. Biol. 267:132-141) para obtener CDC activo estable. Las reorganizaciones adicionales espontáneas de los CDC también pueden inhibirse por incubación a 0 grados centígrados (véase Surette et al. (1987) Cell 49:253-262)). Cuando el CDC se ha obtenido de un plásmido mini-Mu de tipo salvaje, las moléculas de transposasa MuA unidas débilmente pueden eliminarse para obtener una versión recortada del CDC activo (Wu y Chaconas (1997) J. Mol. Biol. 267:132-141). Este CDC recortado activo puede usarse para preparar los vectores de integración de la invención. Sin embargo, cuando el CDC activo comprende las moléculas de transposasa MuA unidas débilmente a los sitios de unión auxiliares attL2, attL3 y attR3, la transferencia de cadena intermolecular ocurre cuatro veces más rápido que con el CDC recortado (Wu y Chaconas (1997), supra). Así, cuando se va a usar un CDC recortado, puede coadministrarse proteína quimérica MuA adicional a la célula del anfitrión para estimular la transferencia de cadena intermolecular. La MuA quimérica adicional puede coadministrarse directamente usando una técnica tal como microinyección o bombardeo de partículas o puede coadministrarse indirectamente por la administración de un vector de expresión que comprende la secuencia codificadora de MuA quimérica unida de manera operativa a elementos reguladores que estimulan la expresión en la célula anfitriona. Como la MuA quimérica debe importarse en el núcleo, dicha construcción de ADN comprenderá además una secuencia que codifica una señal de localización nuclear, tal como NLS de SV40, fusionada en marco con la secuencia codificadora de MuA quimérica. Además de la MuA quimérica, otras proteínas o compuestos pueden ser útiles para conseguir los resultados deseados de frecuencia incrementada de integración no aleatoria de los CDC y dichas proteínas o compuestos también pueden coadministrarse a la célula anfitriona con los vectores de la presente invención.

Así, un plásmido mini-Mu de interés y la transposasa quimérica MuA, HU, y proteínas IHF, o variantes o fragmentos biológicamente activos de éstos, pueden usarse en una reacción in vitro bajo condiciones de reacción estándar o modificadas para obtener un CDC activo estable que es capaz de transposición intermolecular. Durante la formación de este CDC, se ha introducido una mella en cada extremo del casete transponible Mu, exponiendo los grupos 3'-OH, relajando el dominio de ADN que no es del plásmido Mu. Este CDC estable puede modificarse en el dominio de ADN que no es del plásmido Mu para obtener nuevos vectores de integración de la invención.

Así, los nuevos vectores de integración de la invención pueden obtenerse usando plásmidos mini-Mu y cualesquiera otras proteínas necesarias o útiles, tales como, por ejemplo, la transposasa MuA nativa, las proteínas bacterianas HU, IHF y una proteína semejante a RecA o variantes o fragmentos biológicamente activos de éstas. Dichas proteínas pueden producirse in vivo por el genoma del anfitrión, por ejemplo como resultado de ingeniería genética previa del genoma o las proteínas pueden introducirse junto con los vectores de integración durante o después de la transformación del genoma del anfitrión con los vectores de integración. Dicha introducción puede ser directa o indirecta (por ejemplo, por cotransformación de un vector de integración con otra secuencia de ADN que codifica la transposasa MuA nativa). Así, pueden formarse CDC activos en la célula anfitriona en la que los elementos y secuencias apropiadas existen en el interior de la célula.

Cuando se van a usar proteínas purificadas, los métodos para obtener estas proteínas nativas purificadas o variantes o fragmentos biológicamente activos de éstas son conocidos en la técnica. Véanse, por ejemplo, Craigie y Mizuuchi (1985) J. Biol. Chem. 260:1832-1835 (clonación del gen MuA y purificación de MuA); Craigie et al. (1985) Proc. Natl. Acad. Sci. USA 82:7570-7574, Rouviere-Yaniv y Gros (1975) Proc. Natl. Acad. Sci. USA 72:3428-3432, Dixon y Komberg (1984) Proc. Natl. Acad. Sci. USA 81:424-428; y Surette et al. Cell 49:253-226 (purificación de HU); Wu y Chaconas (1994) J. Biol. Chem. 269:28829-28833, y las referencias citadas en éstas (MuA, HU e IHF); Yang et al. (1995) EMBO J 14:2374-2384 (MuA nativa y variantes de ésta, y HU); y Shibita et al. (1982) J. Biol. Chem. 257:370, Shibita et al. (1983) Methods Enzymol. 100:197, Cox et al. (1981) J. Biol. Chem. 256(9):4676, y Cox et al. (1981) Proc. Natl. Acad. Sci. USA 78:3433 (RecA purificada). Los métodos para construir enzimas de integración quiméricas se describen en la presente memoria.

Por "fragmento" se pretende una parte de la secuencia de aminoácidos y por tanto la proteína codificada de esta manera. Por ejemplo, una parte biológicamente activa de MuA, HU, IHF o proteína semejante a RecA pueden prepararse aislando una parte de sus secuencias codificadoras respectivas, expresando la parte codificada de la proteína respectiva (p. ej., por expresión recombinante in vitro) y evaluando la actividad de la parte codificada de la proteína respectiva. Las secuencias codificadoras de estas proteínas son conocidas en la técnica. Véanse, por ejemplo, Grimaud (1996) Virology 217(1):200-210 para la secuencia de nucleótidos para el bacteriófago Mu (No. de Registro de GenBank AF083977), que identifica la secuencia codificadora de la transposasa MuA (No. de Registro de GenBank AAF01083); Miller (1984) Cold Spring Harb. Symp. Quant. Biol. 49:691-698 para la secuencia codificadora de la subunidad alfa de IHF (No. de Registro de GenBank P06984) y Flamm y Weisberg (1985) J. Mol. Biol. 183(2):117-128 para la secuencia codificadora de la subunidad beta de IHF (No. de Registro de GenBank P08756); No. de Registro de GenBank U82664, nucleótidos 40901-41173, que codifican la proteína HU (No. de Registro de GenBank AAB40196); y Keener et al. (1984) J. Bacteriol. 160(1):153-160 y las referencias citadas en otro lugar en la presente memoria para secuencias codificadoras de proteínas semejantes a RecA.

#### 16. Ejemplo 16

El desarrollo de una transposasa quimérica usa técnicas muy conocidas en la técnica de la biología molecular. Por ejemplo, un dominio de unión al ADN de una fuente exógena puede introducirse en la transposasa Bella durmiente (Figura 17). La estrategia implica dos etapas de clonación; ambas están basadas en PCR e implican mutagénesis dirigida por cebador. La primera etapa introduce un único sitio de restricción en el extremo N terminal de Bella Durmiente (Figura 18). La segunda etapa introduce un dominio de unión al ADN (bien de zif268 o similar), junto con un conector peptídico flexible, en ese sitio de restricción (Figura 19). Este conector peptídico flexible puede ayudar a que la proteína adopte la conformación necesaria para unir ADN y catalice el evento de transposición simultáneamente. Los clones deben secuenciarse después de cada etapa, para seleccionar un vector que no contenga mutaciones. Con esta estrategia, también se retienen sitios únicos de restricción alrededor del casete de expresión (incluyendo el promotor y la señal de poli-A) en el vector final, para permitir la subclonación de éste en otros vectores (p. ej. un plásmido suicida). Los productos secundarios de esta estrategia son pCMV-antisentidoSB\* y Bella Durmiente con el dominio de unión al ADN-conector peptídico flexible fusionados en una orientación antisentido.

#### 17. Ejemplo 17: Vector de expresión para la proteína de fusión Transposasa-dominio de unión al ADN

Se eligieron dos dominios de unión al ADN bien caracterizados: los de Leu3p y Uga3p. Tanto Leu3p como Uga3p son activadores transcripcionales en *S. cerevisiae*. Se unen a secuencias de ADN específicas que se encuentran en los genes diana. Sus dominios de unión al ADN consisten en un dedo de cinc y un dominio de dimerización que permite la unión homodimérica al ADN. Uga3p y Leu3p están altamente relacionados, pero con distintas dianas de ADN (Noel y

Turcotte (1998) J. Biol. Chem. 273:17463), por ejemplo, el dominio de unión al ADN purificado de Leu3p (aminoácidos 1 a 147) se une in vitro a la secuencia de ADN SEQ ID No: 1 TCCGGCCGGAACCGGCTTT (Hellauer, et al., (1996) Mol. Cell. Biol. 16:6096).

5 Las secuencias correspondientes al dominio de unión al ADN de Leu3p se amplificaron por PCR usando el ADN genómico de *S. cerevisiae* como molde y los oligos siguientes:

Y10

CG GGATCCCACCTATGGAAGGAAGATCAGATTT SEQ ID No: 2

Y11

AGATTACTCGAGTCAAAGTGTTTTGTATGATGTCG SEQ ID No: 3

10 El producto de PCR se cortó con BamHI y XhoI y se subclonó en el plásmido pGRTEMP2. Después se cortó con las enzimas compatibles BglII y Sall. El marco de lectura abierto de la transposasa contiene un sitio BglII justo antes del codón de parada y un sitio Sall justo en 3' del codón de parada (AAGATCTGATCCGTCGAC SEQ ID No: 4 con el codón de parada subrayado). La proteína de fusión debería ser Tn5-Pro-Pro-Leu3p (a.a. 1-147). El plásmido que codifica una fusión de transposasa Uga3p (a.a. 1-124) se obtuvo de manera similar excepto en que los oligos usados para la PCR fueron:

U11

CGGGATCCCACCTATGAATTATGGCGTGGAGAA SEQ ID No: 5

U12

AGATTACTCGAGTCAGTTGTACAGCTGCAATCCCA SEQ ID No: 6

20 Se obtuvieron plásmidos del tamaño apropiado. Las construcciones que se secuenciaron (usando el cebador GGAAGCCCTGCAAAGTAAA SEQ ID No: 7) tenía secuencias Leu3p o Uga3p insertadas en pGRTEMP2. Sin embargo, todas las construcciones secuenciadas tenían mutaciones que resultaron en desplazamientos de marco en la región codificadora de Leu3p o Uga3p. Como tales, no se pudieron producir proteínas de fusión de longitud completa por las bacterias transformadas con los plásmidos.

## 25 18. Ejemplo 18

En la presente memoria se describen ejemplos de construcciones de la presente invención preparadas según la enseñanza de la presente memoria. La terminología usada tiene el significado siguiente: TR = repetición terminal; E/P = Amplificador/Promotor; Transgén = Ácido nucleico insertado en el ADN diana. Todas las construcciones tienen la fórmula básica de repetición terminal amplificador/promotor-transgén-repetición terminal-amplificador/promotor-transposasa quimérica (transposasa/factor de acoplamiento a ADN), en las que la transposasa se representará por el nombre del transposón correspondiente (p. ej., Bella Durmiente y TcI/mariner) y el factor de acoplamiento a ADN está representado por LexA, STF-1, Zif268 o cualquier otro factor de acoplamiento descrito en la presente memoria. Además, las repeticiones terminales representan la repetición correspondiente a la transposasa identificada. Todos los genes representan el ácido nucleico que codifica la proteína identificada.

35 TR-E/P)-Transgén-TR-E/P-----integrasa lambda / Gal4-----

TR-E/P-Transgén-TR-E/P-----Gal4 / integrasa lambda-----

TR-E/P-Transgén-TR-E/P-----integrasa lambda / LexA-----

TR-E/P-Transgén-TR-E/P-----LexA / integrasa lambda-----

TR-E/P-Transgén-TR-E/P-----piggybac / Gal4-----

40 TR-E/P-Transgén-TR-E/P-----Gal4 / piggybac-----

TR-E/P-Transgén-TR-E/P-----piggybac / LexA-----

TR-E/P-Transgén-TR-E/P-----LexA / piggybac-----

TR-E/P-Transgén-TR-E/P-----mos1 / Gal4 -----

|    |  |
|----|--|
|    | TR-E/P-Transgén-TR-E/P-----Gal4 / mos1-----                                    |
|    | TR-E/P-Transgén-TR-E/P-----mos1 / LexA-----                                    |
|    | TR-E/P-Transgén-TR-E/P-----LexA / mos1-----                                    |
|    | TR-E/P-Transgén-TR-E/P-----Cre / Zif268-----                                   |
| 5  | TR-E/P-Transgén-TR-E/P-----Zif268 / Cre -----                                  |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / LexA -----                        |
|    | TR-E/P-Transgén-TR-E/P-----LexA / bella durmiente -----                        |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / STF-1-----                        |
|    | TR-E/P-Transgén-TR-E/P-----STF-1 / bella durmiente-----                        |
| 10 | TR-E/P-Transgén-TR-E/P-----bella durmiente / Zif268-----                       |
|    | TR-E/P-Transgén-TR-E/P-----Zif268 / bella durmiente-----                       |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / XFin-----                         |
|    | TR-E/P-Transgén-TR-E/P-----XFin / bella durmiente-----                         |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / receptor de glucocorticoides----- |
| 15 | TR-E/P-Transgén-TR-E/P-----receptor de glucocorticoides / bella durmiente----- |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / 434-cro -----                     |
|    | TR-E/P-Transgén-TR-E/P-----434-cro / bella durmiente -----                     |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / MRF4 -----                        |
|    | TR-E/P-Transgén-TR-E/P-----MRF4 / bella durmiente -----                        |
| 20 | TR-E/P-Transgén-TR-E/P-----bella durmiente / PAL2 -----                        |
|    | TR-E/P-Transgén-TR-E/P-----PAL2 / bella durmiente -----                        |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / GCN4 -----                        |
|    | TR-E/P-Transgén-TR-E/P-----GCN4 / bella durmiente -----                        |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / jun -----                         |
| 25 | TR-E/P-Transgén-TR-E/P-----jun / bella durmiente -----                         |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / pin -----                         |
|    | TR-E/P-Transgén-TR-E/P-----pin / bella durmiente -----                         |
|    | TR-E/P-Transgén-TR-E/P-----bella durmiente / HSV-VP16 -----                    |
|    | TR-E/P-Transgén-TR-E/P-----HSV-VP16 / bella durmiente -----                    |
| 30 | TR-E/P-Transgén-TR-E/P-----bella durmiente / HOX -----                         |
|    | TR-E/P-Transgén-TR-E/P-----HOX / bella durmiente -----                         |
|    | TR-E/P-Transgén-TR-Egr-1 -----bella durmiente / LexA -----                     |
|    | TR-E/P-Transgén-TR-Egr-1 -----LexA / bella durmiente -----                     |
|    | TR-E/P-Transgén-TR-Egr-1 -----bella durmiente / STF-1-----                     |

|    |                          |  |
|----|--------------------------|--|
|    | TR-E/P-Transgén-TR-Egr-1 | STF-1 / bella durmiente                        |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / Zif268                       |
|    | TR-E/P-Transgén-TR-Egr-1 | Zif268 / bella durmiente                       |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / XFin                         |
| 5  | TR-E/P-Transgén-TR-Egr-1 | XFin / bella durmiente                         |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / receptor de glucocorticoides |
|    | TR-E/P-Transgén-TR-Egr-1 | receptor de glucocorticoides / bella durmiente |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / 434-cro                      |
|    | TR-E/P-Transgén-TR-Egr-1 | 434-cro / bella durmiente                      |
| 10 | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / MRF4                         |
|    | TR-E/P-Transgén-TR-Egr-1 | MRF4 / bella durmiente                         |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / PAL2                         |
|    | TR-E/P-Transgén-TR-Egr-1 | PAL2 / bella durmiente                         |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / GCN4                         |
| 15 | TR-E/P-Transgén-TR-Egr-1 | GCN4 / bella durmiente                         |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / jun                          |
|    | TR-E/P-Transgén-TR-Egr-1 | jun / bella durmiente                          |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / pin                          |
|    | TR-E/P-Transgén-TR-Egr-1 | pin / bella durmiente                          |
| 20 | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / HSV-VP16                     |
|    | TR-E/P-Transgén-TR-Egr-1 | HSV-VP16 / bella durmiente                     |
|    | TR-E/P-Transgén-TR-Egr-1 | bella durmiente / HOX                          |
|    | TR-E/P-Transgén-TR-Egr-1 | HOX / bella durmiente                          |
|    | TR-E/P-Transgén-TR-GFAP  | bella durmiente / LexA                         |
| 25 | TR-E/P-Transgén-TR-GFAP  | LexA / bella durmiente                         |
|    | TR-E/P-Transgén-TR-GFAP  | bella durmiente / STF-1                        |
|    | TR-E/P-Transgén-TR-GFAP  | STF-1 / bella durmiente                        |
|    | TR-E/P-Transgén-TR-GFAP  | bella durmiente / Zif268                       |
|    | TR-E/P-Transgén-TR-GFAP  | Zif268 / bella durmiente                       |
| 30 | TR-E/P-Transgén-TR-GFAP  | bella durmiente / XFin                         |
|    | TR-E/P-Transgén-TR-GFAP  | XFin / bella durmiente                         |
|    | TR-E/P-Transgén-TR-GFAP  | bella durmiente / receptor de glucocorticoides |
|    | TR-E/P-Transgén-TR-GFAP  | receptor de glucocorticoides / bella durmiente |
|    | TR-E/P-Transgén-TR-GFAP  | bella durmiente / 434-cro                      |

|    |  |
|----|--|
|    | TR-E/P-Transgén-TR-GFAP-----434-cro / bella durmiente -----                |
|    | TR-E/P-Transgén-TR-GFAP-----bella durmiente / MRF4 -----                   |
|    | TR-E/P-Transgén-TR-GFAP-----MRF4 / bella durmiente -----                   |
|    | TR-E/P-Transgén-TR-GFAP-----bella durmiente / PAL2 -----                   |
| 5  | TR-E/P-Transgén-TR-GFAP-----PAL2 / bella durmiente -----                   |
|    | TR-E/P-Transgén-TR-GFAP-----bella durmiente / GCN4 -----                   |
|    | TR-E/P-Transgén-TR-GFAP-----GCN4 / bella durmiente -----                   |
|    | TR-E/P-Transgén-TR-GFAP-----bella durmiente / jun -----                    |
|    | TR-E/P-Transgén-TR-GFAP-----jun / bella durmiente -----                    |
| 10 | TR-E/P-Transgén-TR-GFAP-----bella durmiente / pin -----                    |
|    | TR-E/P-Transgén-TR-GFAP-----pin / bella durmiente -----                    |
|    | TR-E/P-Transgén-TR-GFAP-----bella durmiente / HSV-VP16 -----               |
|    | TR-E/P-Transgén-TR-GFAP-----HSV-VP16 / bella durmiente -----               |
|    | TR-E/P-Transgén-TR-GFAP-----bella durmiente / HOX -----                    |
| 15 | TR-E/P-Transgén-TR-GFAP-----HOX / bella durmiente -----                    |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / LexA -----                        |
|    | TR-E/P-Transgén-TR-E/P-----LexA / Tc1/mariner -----                        |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / STF-1-----                        |
|    | TR-E/P-Transgén-TR-E/P-----STF-1 / Tc1/mariner-----                        |
| 20 | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / Zif268-----                       |
|    | TR-E/P-Transgén-TR-E/P-----Zif268 / Tc1/mariner-----                       |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / XFin-----                         |
|    | TR-E/P-Transgén-TR-E/P-----XFin / Tc1/mariner-----                         |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / receptor de glucocorticoides----- |
| 25 | TR-E/P-Transgén-TR-E/P-----receptor de glucocorticoides / Tc1/mariner----- |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / 434-cro -----                     |
|    | TR-E/P-Transgén-TR-E/P-----434-cro / Tc1/mariner -----                     |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / MRF4 -----                        |
|    | TR-E/P-Transgén-TR-E/P-----MRF4 / Tc1/mariner -----                        |
| 30 | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / PAL2 -----                        |
|    | TR-E/P-Transgén-TR-E/P-----PAL2 / Tc1/mariner -----                        |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / GCN4 -----                        |
|    | TR-E/P-Transgén-TR-E/P-----GCN4 / Tc1/mariner -----                        |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / jun -----                         |

|    |   |
|----|---|
|    | TR-E/P-Transgén-TR-E/P-----jun / Tc1/mariner -----                            |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / pin -----                            |
|    | TR-E/P-Transgén-TR-E/P-----pin / Tc1/mariner -----                            |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / HSV-VP16 -----                       |
| 5  | TR-E/P-Transgén-TR-E/P-----HSV-VP16 / Tc1/mariner -----                       |
|    | TR-E/P-Transgén-TR-E/P-----Tc1/mariner / HOX -----                            |
|    | TR-E/P-Transgén-TR-E/P-----HOX / Tc1/mariner -----                            |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / LexA -----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----LexA / Tc1/mariner -----                        |
| 10 | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / STF-1-----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----STF-1 / Tc1/mariner-----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / Zif268-----                       |
|    | TR-E/P-Transgén-TR-Egr-1 -----Zif268 / Tc1/mariner-----                       |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / XFin-----                         |
| 15 | TR-E/P-Transgén-TR-Egr-1 -----XFin / Tc1/mariner-----                         |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / receptor de glucocorticoides----- |
|    | TR-E/P-Transgén-TR-Egr-1 -----receptor de glucocorticoides / Tc1/mariner----- |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / 434-cro -----                     |
|    | TR-E/P-Transgén-TR-Egr-1 -----434-cro / Tc1/mariner -----                     |
| 20 | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / MRF4 -----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----MRF4 / Tc1/mariner -----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / PAL2 -----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----PAL2 / Tc1/mariner -----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / GCN4 -----                        |
| 25 | TR-E/P-Transgén-TR-Egr-1 -----GCN4 / Tc1/mariner -----                        |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / jun -----                         |
|    | TR-E/P-Transgén-TR-Egr-1 -----jun / Tc1/mariner -----                         |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / pin -----                         |
|    | TR-E/P-Transgén-TR-Egr-1 -----pin / Tc1/mariner -----                         |
| 30 | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / HSV-VP16 -----                    |
|    | TR-E/P-Transgén-TR-Egr-1 -----HSV-VP16 / Tc1/mariner -----                    |
|    | TR-E/P-Transgén-TR-Egr-1 -----Tc1/mariner / HOX -----                         |
|    | TR-E/P-Transgén-TR-Egr-1 -----HOX / Tc1/mariner -----                         |
|    | TR-E/P-Transgén-TR-GFAP -----Tc1/mariner / LexA -----                         |

- TR-E/P-Transgén-TR-GFAP -----LexA / Tc1/mariner -----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / STF-1-----
- TR-E/P-Transgén-TR-GFAP-----STF-1 / Tc1/mariner-----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / Zif268-----
- 5 TR-E/P-Transgén-TR-GFAP-----Zif268 / Tc1/mariner-----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / XFin-----
- TR-E/P-Transgén-TR-GFAP-----XFin / Tc1/mariner-----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / receptor de glucocorticoides-----
- TR-E/P-Transgén-TR-GFAP-----receptor de glucocorticoides / Tc1/mariner-----
- 10 TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / 434-cro -----
- TR-E/P-Transgén-TR-GFAP-----434-cro / Tc1/mariner -----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / MRF4 -----
- TR-E/P-Transgén-TR-GFAP-----MRF4 / Tc1/mariner -----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / PAL2 -----
- 15 TR-E/P-Transgén-TR-GFAP-----PAL2 / Tc1/mariner -----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / GCN4 -----
- TR-E/P-Transgén-TR-GFAP-----GCN4 / Tc1/mariner -----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / jun -----
- TR-E/P-Transgén-TR-GFAP-----jun / Tc1/mariner -----
- 20 TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / pin -----
- TR-E/P-Transgén-TR-GFAP-----pin / Tc1/mariner -----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / HSV-VP16 -----
- TR-E/P-Transgén-TR-GFAP-----HSV-VP16 / Tc1/mariner -----
- TR-E/P-Transgén-TR-GFAP-----Tc1/mariner / HOX -----
- 25 TR-E/P-Transgén-TR-GFAP-----HOX / Tc1/mariner -----

Aunque la invención se ha descrito con detalle con referencia a determinadas realizaciones preferidas de ésta, deberá entenderse que las modificaciones y variaciones están dentro del espíritu y alcance de lo que se describe y reivindica.

**F. Referencias**

1. Wu, G.Y., Zhan, P., Sze, L.L., Rosenberg, A.R. y Wu, C.H. (1994) Incorporation of adenovirus into a ligandbased DNA carrier system results in retention of original receptor specificity and enhances targeted gene expression. *J. Biol. Chem.* 269:11542-11546.
- 5 2. Chowdhury, N.R., Wu, C.H., Wu, G.Y., Yemeni, P.C., Bommineni, V.R. y Chowdhury, J.R. (1993) Fate of DNA targeted to the liver by asialoglycoprotein receptor-mediated endocytosis in vivo. *J. Biol. Chem.* 268: 11265-11271.
3. Farhood, H., Gao, X., Son, K., Yang, Y.Y., Lazo, J.S., Huang, L., Barsoum, J., Bottega, R. y Eband, R.M. (1994) Cationic liposomes for direct gene transfer in therapy of cancer and other diseases. *Ann. NY Acad. Sci.* 716:23-35.
- 10 4. Curiel, D.T. (1994) High efficiency gene transfer mediated by adenovirus-polylysine-DNA complexes. *Ann. NY Acad. of Sci.* 716:36-58.
- 5 5. Cotten, M., Wagner, E., Zatloukal, K. y Birnstiel ML. (1993) Chicken adenovirus (CELO virus) particles augment receptor-mediated DNA delivery to mammalian cells and yield exceptional levels of stable transformants. *J. Virology* 67: 3777-3785.
- 15 6. Schagen, F.H., Rademaker, H.J., Cramer, S.J., van Ormondt, H., van der Eb, A.J., van de Putte, P. y Hoeben, R.C. (2000) Towards integrating vectors for gene therapy: expression of functional bacteriophage MuA and MuB proteins in mammalian cells. *Nucleic Acids Res.* 28: E104.
7. Lestina, B.J., Sagnella, S.M., Xu, Z., Shive, M.S., Richter, N.J., Jayaseharan, J., Case, A.J., Kottke-Marchant, K., Anderson, J.M. y Marchant, R.E. (2002) Surface modification of liposomes for selective cell targeting in cardiovascular drug delivery. *J. Control Release* 78:235-247.
- 20 8. Moreira, J.N., Gaspar, R. y Allen, T.M. (2001) Targeting stealth liposomes in a murine model of human small cell lung cancer. *Biochim. Biophys. Acta.* 1515:167-176.
9. Xu, L., Tang, W.H., Huang, C.C., Alexander, W., Xiang, L.M., Pirolo, K.F., Rait, A. y Chang, E.H. (2001) Systemic p53 gene therapy of cancer with immunolipoplexes targeted by anti-transferrin receptor scFv. *Mol. Med.* 7:723-734.
- 25 10. Sudhan Shaik, M., Kanikkannan, N. y Singh, M. (2001) Conjugation of anti-My9 antibody to stealth monensin liposomes and the effect of conjugated liposomes on the cytotoxicity of immunotoxin. *J. Control Release* 76:285-295.
11. Li, X., Stuckert, P., Bosch, I., Marks, J.D. y Marasco, W.A. (2001) Single-chain antibody-mediated gene delivery into ErbB2-positive human breast cancer cells. *Cancer Gene Ther.* 8:555-565.
12. Park, J.W., Kirpotin, D.B., Hong, K., Shalaby, R., Shao, Y., Nielsen, U.B., Marks, J.D., Papahadjopoulos, D. y Benz, C.C. (2001) Tumor targeting using anti-her2 immunoliposomes. *J. Control Release* 74:95-113.
- 30 13. Lee, S.E., Jin, R.J., Lee, S.G., Yoon, S.J., Park, M.S., Heo, D.S. y Choi, H. (2000) Development of a new plasmid vector with PSA-promoter and enhancer expressing tissuespecificity in prostate carcinoma cell lines. *Anticancer Res.* 20:417-422.
14. Gottschalk, S., Cristiano, R.J., Smith, L.C. y Woo, S.L. (1993) Folate receptor mediated DNA delivery into tumor cells: potosomal disruption results in enhanced gene expression. *Gene Ther* 1:185-191.
- 35 15. Boulikas, T. (1997) Gene therapy of prostate cancer: p53, suicidal genes, and other targets. *Anticancer Res.* 17:1471-1505.
16. Kaneda, Y., Iwai, K. y Uchida, T. (1989) Increased expression of DNA cointroduced with nuclear protein in adult rat liver. *Science* 243:375-378.

17. Izsvak, Z., Ivics, Z. y Plasterk, R.H. (2000) Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J. Mol. Biol.* 302:93-102.
18. Sauer, B. (1994) Site-specific recombination: developments and applications. *Curr. Opin. Biotechnol.* 5:521-527.
- 5 19. Diaz, V., Rojo, F., Martinez-A, C., Alonso, J.C. y Bernard, A. (1999) The prokaryotic beta-recombinase catalyzes site-specific recombination in mammalian cells. *J. Biol. Chem.* 274:6634-6640.
20. O'Gorman, S., Fox, D.T. y Wahl, G.M. (1991) Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* 251:1351-1355.
- 10 21. Waters, J.S., Webb, A., Cunningham, D., Clare, P.A., Raynaud, F., di Stefano, F. y Cotter, F.E. (2000) Phase I clinical and pharmacokinetic study of bcl-2 antisense oligonucleotide therapy in patients with non-Hodgkin's lymphoma. *J. Clin. Oncol.* 18:1812-1823.
22. Nabel, E.G., Yang, Z., Muller, D., Chang, A.E., Gao, X., Huang, L., Cho, K.J. y Nabel, G.J. (1994) Safety and toxicity of catheter gene delivery to the pulmonary vasculature in a patient with metastatic melanoma. *Hum. Gene Ther.* 5:1089-1094.
- 15 23. Nabel, G.J., Nabel, E.G., Yang, Z.Y., Fox, B.A., Plautz, G.E., Gao, X., Huang, L., Shu, S., Gordon, D. y Chang, A.E. (1993) Direct gene transfer with DNA-liposome complexes in melanoma: expression, biologic activity, and lack of toxicity in humans. *Proc. Natl. Acad. Sci. U. S. A.* 90:11307-11311.
24. Caplen, N.J., Alton, E.W., Middleton, P.G., Dorin, J.R., Stevenson B.J., Gao X., Durham S.R., Jeffrey P.K., Hodson M.E., Coutelle C. et al. (1995) Liposome-mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis. *Nat. Med.* 1:39-46.
- 20 25. Kalpana, G.V., Marmon, S., Wang, W., Crabtree, G.R. y Goff, S.P. (1994) Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor snf5. *Science* 266:2002-2006.
26. Bushman FD, Fujiwara T. y Cragie R. (1995) Retroviral DNA integration directed by HIV integration protein in vitro. *Science* 249:1555-1558.
- 25 27. Lampe, D.J., Churchill, M.E. y Robertson, H.M. (1996) A purified mariner transposase is sufficient to mediate transposition in vitro. *EMBO J.* 15:5470-5479.
28. Vos, J.C., De Baere, I. y Plasterk, R.H. (1996) Transposase is the only nematode protein required for in vitro transposition of Tc1. *Genes Dev.* 10:755-761.
29. Skalka, A.M. (1993) Retroviral DNA integration: lessons for transposon shuffling. *Gene* 135: 175-182.
- 30 30. Ellison, V. y Brown, P.O. (1994). A stable complex between integrase and viral DNA ends mediates human immunodeficiency virus integration in vitro. *Proc. Natl. Acad. U.S.A.* 91: 7316-7320.
31. Engelman, A. (1999) In vivo analysis of retroviral integrase structure and function. *Adv. Virus Res.* 52:411-426.
32. Fischer, S.E., Wienholds, E. y Plasterk, R.H. (2001) Regulated transposition of a fish transposon in the mouse germ line. *Proc. Natl. Acad. Sci. U.S.A* 98:6759-6764.
- 35 33. Flint, S.J., Enquist, L.W., Krug, R.M., Racaniello, V.R. y Skalka, A.M. (2000) Reverse transcription and integration: hallmarks of the reovirus. In *Principles of Virology: Molecular Biology, Pathogenesis, and Control*. pp. 199-234, ASM Press, Washington, DC.
34. Goldhaber-Gordon, I., Williams, T.L. y Baker, T.A. (2001) DNA recognition sites activate MuA transposase to perform transposition of non-Mu DNA. *J. Biol. Chem.* 277:7694-7702.

35. Yant, S.R., Meuse, L., Chin, W., Ivics, Z., Izsvak, Z. y Kay, M.A. (2000) Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat. Genet.* 25:35-41.
36. Bushman, F. (1995) Targeting retroviral integration. *Science* 267:1443-1444.
- 5 37. Kirchner, J., Connolly, C.M. y Sandmeyer, S.B. (1995) Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. *Science* 267:1488-1491.
38. Thyagarajan, B., Olivares, E.C., Hollis, R.P., Ginsburg, D.S. y Calos, M.P. (2001) Sitespecific genomic integration in mammalian cells mediated by phage  $\phi$ C31 integrase. *Mol. Cell Biol.* 21:3926-3934.
39. Rubin, G.M. y Spradling, A.C. (1983) Vectors for P element mediated gene transfer in *Drosophila*. *Nucleic Acids Res.* 11:6341-6351.
- 10 40. Bushman, F.D. (1994) Tethering human immunodeficiency virus 1 integrase to a DNA site directs integration to nearby sequences. *Proc. Natl. Acad. Sci. U.S.A.* 91:9233-9237.
41. Kulkosky, J., Katz, R., Merkel, G. y Skalka, A.M. (1995) Activities and substrate specificity of the evolutionarily conserved central domain of retroviral integrase. *Virology* 206:44.8-456.
- 15 42. Logie, C. y Stewart, F.A. (1995) Ligand-regulated site-specific recombination. *Proc. Natl. Acad. Sci. U.S.A.* 92:5940-5944.
43. Katz, R.A., Merkel, G. y Skalka, A.M. (1996) Targeting of retroviral integrase by fusion to a heterologous DNA binding domain: in vitro activities and incorporation of a fusion protein into viral particles. *Virology* 217:178-190.
44. Bushman FD. y Miller MD. (1997) Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *J. Virol.* 71:458-464.
- 20 45. Holmes-Son, M.L., Appa, R.S. y Chow, S.A. (2001) Molecular genetics and target site specificity of retroviral integration. *Adv. Genet.* 43:33-69.
46. Holmes-Son, M.L. y Chow, S.A. (2000) Integrase-LexA fusion proteins incorporated into human immunodeficiency virus type I that contains a catalytically inactive integrase gene are functional to mediate integration. *J. Virol.* 74:11548-11556.
- 25 47. Kaminski, J.M., Nguyen, K., Buyyounouski, M. y Pollack, A. (2002) Prostate cancer gene therapy and the role of radiation. *Cancer Treat. Rev.* 28:49-64.
48. Yant, S.R., Meuse, L., Chiu, W., Ivics, Z., Izsvak, Z. y Kay, M.A. (2000) Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nature Genetics* 25:35-41.
- 30 49. Loukeris, T.G., Livadaras, I., Arca, B., Zabalou, S. y Savakis, C. (1995) Gene transfer into the medfly, *Ceratitis capitata*, with a *Drosophila hydei* transposable element. *Science* 270:2002-2005.
50. Loukeris, T.G., Area, B., Livadaras, I., Dialektaki, G. y Savakis, C. (1995) Introduction of the transposable element Minos into the germline of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 92:9485-9489.
51. Lidholm, D.A., Lohe, A.R. y Hartl, D.L. (1993) The transposable element mariner mediates germline transformation in *Drosophila melanogaster*. *Genetics* 134:859-868.
- 35 52. Lohe, A.R. y Hartl, D.L. (1996) Germline transformation of *Drosophila virilis* with the transposable element mariner. *Genetics* 143:365-374.

53. Coates, C.J., Turney, C.L., Frommer, M. y O'Brochta, D.A. (1997) Interplasmid transposition of the mariner transposable element in non-drosophilid insects. *Mol. Gen. Genet* 253:728-733.
54. Coates, C.J., Jasinskiene, N., Miyashiro, L. y James, A.A. (1998) Mariner transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc. Natl. Acad. Sci. U.S.A.* 95:3748-3751.
- 5 55. Rubin, E.J., Akerley, B.J., Novik, V.N., Lampe, D.J., Husson, R.N. y Mekalanos, J.J. (1999) In vivo transposition of mariner-based elements in enteric bacteria and mycobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 96:1645-1650.
56. Gueiros-Filho, F.J. y Beverly, S.M. (1997) Trans-kingdom transposition of the *Drosophila* element mariner within the protozoan *Leishmania*. *Science* 276:1716-1719.
- 10 57. Sherman, A., Dawson, A., Mather, C., Gilhooley, H., Li, Y., Mitchell, R., Finnegan, D. y Sang, H. (1998) Transposition of the *Drosophila* element mariner into the chicken germ line. *Nat. Biotech.* 16:1050-1053.
58. Raz, E., van Luenen, H.G., Schaerringer, B., Plasterk, R.H., Driever, W. (1998) Transposition of the nematode *Caenorhabditis elegans* Tc3 element in the zebrafish *Danio rerio*. *Curr. Biol* 8:82-88.
59. Fadool, J.M., Hartl, D.L. y Dowling, J.E. (1998) Transposition of the mariner element from *Drosophila mauritiana* in zebrafish. *Proc. Natl. Acad. Sci. U.S.A.* 95:5182-5186.
- 15 60. Luo, G., Ivics, Z., Izsvak, Z. y Bradley, A. (1998) Chromosomal transposition of a Tc1 /mariner-like element in a mouse embryonic stem cell. *Proc. Natl. Acad. Sci. U.S.A.* 95:10769-10773.
61. Ivics, Z., Hackett, P.B., Plasterk, R.H. y Izsvak, Z. (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 91:501-510.
- 20 62. Schouten, G.J., van Luenen, H.G., Verra, N.C., Valerio, D. y Plasterk, R.H. (1998) Transposon Tc1 of the nematode *Caenorhabditis elegans* jumps in human cells. *Nucleic Acids Res.* 26:3013-3017.
63. Zhang, L., Sankar, U., Lampe, D.J., Robertson, H.M. y Graham, F.L. (1998) The Himar1 mariner transposase cloned in a recombinant adenovirus vector is functional in mammalian cells. *Nucleic Acids Res.* 16:3687-3693.
- 25 64. Kay, M.A., Meuse, L., Gown, A.M., Linsley, P., Hollenbaugh, D., Aruffo, A., Ochs, H.D. y Wilson, C.B. (1997) Transient immunomodulation with anti-CD40 ligand antibody and CTLA4lg enhances persistence and secondary adenovirus-mediated gene transfer into mouse liver. *Proc. Natl. Acad. Sci. U.S.A.* 94:4684-4691.
65. Hernandez, Y.J., Wang, J., Kearns, W.G., Loiler, S., Poirier, A. y Flotte, T.R. (1999) Latent adeno-associated virus infection elicits humoral but not cell-mediated immune responses in a nonhuman primate model. *J. Virol.* 73:8549-8558.
66. Koduri, R.K., Miller, J.T. y Thammana P. (2001) An efficient homologous recombination vector pTV(I) contains a hot spot for increased recombinant protein expression in Chinese hamster ovary cells. *Gene* 280:87-95.
- 30 67. Kang, Y.K., Park, J.S., Lee, C.S., Yeom, Y., Chung, A.S. y Lee, K.K. (1999) Efficient integration of short interspersed element-flanked foreign DNA via homologous recombination. *J. Biol. Chem.* 274:36585-36591.
68. Stevens, S.W. y Griffith, J.D. (1994) Human immunodeficiency virus type 1 may preferentially integrate into chromatin occupied by L1Hs repetitive elements. *Proc. Natl. Acad. Sci. U.S.A.* 91:5557-5561.
- 35 69. Stevens, S.W. y Griffith, J.D. (1996) Sequence analysis of the human DNA flanking sites of Human Immunodeficiency Virus Type 1 integration. *J. Virol.* 70:6459-6462.
70. Cox, G.S., Gutkin, D.W., Haas, M.J. y Cosgrove, D.E. (1998) Isolation of an Alu repetitive DNA binding protein and effect of CpG methylation on binding to its recognition sequence. *Biochim. Biophys. Acta.* 1493:302-318.

71. Chesnokov, I.N. y Schmid, C.W. (1995) Specific Alu binding protein from human sperm chromatin prevents DNA methylation. *J. Biol. Chem.* 270:18539-18542.
- 5 72. de Belle, I., Cai, S. y Kohwi-Shigematsu, T. (1998) The genomic sequences bound to special AT-rich sequence-binding protein 1 (SATB 1) in vivo in Jurkat T cells are tightly associated with the nuclear matrix at the bases of the chromatin loops. *J. Cell. Biol.* 141:335-348.
73. Chiang, Y. y Vishwanatha, J.K. Characterization of the HeLa cell 35 kDa Alu-element binding protein. *Mol. Cell Biochem.* 155:131-138.
74. Jurka, J., Walichiewicz, J. y Milosavljevic, A. (1992) Prototypic sequences for human repetitive DNA. *J. Mol. Evol.* 35:286-291.
- 10 75. Zhang, L., Dawson, A. y Finnegan, D.J. (2001) DNA-binding activity and subunit interaction of the mariner transposase. *Nucleic Acids Res.* 29:3566-3575.
76. Vega MA. (1991) Prospects for homologous recombination in human gene therapy *Human Genet.* 87:245-253.
77. Kren, B.T., Bandyopahyay, P. y Steer, C.L. (1998) In vivo site-directed mutagenesis of the factor IX gene by chimeric RNA/DNA oligonucleotides. *Nat. Med.* 4:285-290.
- 15 78. Strauss, M. (1998) The site-specific correction of genetic defects. *Nat. Med.* 4:274-275.
79. Van der Steege, G., Schuilenga-Hut, P.H., Buys, C.H., Scheffer, H., Pas, H.H. y Jonkman, M.F. (2001) Persistent failures in gene repair. *Nat. Biotechnol.* 19:305-306.
80. Albuquerque-Silva, J., Vassart, G., Lavinha, J. y Abramowicz, M.J. (2001) Chimeroplasty validation. *Nat. Biotechnol.* 19:1011.
- 20 81. Groth, A.C., Olivares, E.C., Thyagarajan, B. y Calos, M.P. (2000) A phage integrase directs efficient sitespecific integration in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 97:5995-6000.
82. Kirchner, J., Connolly, C. M. y Sandmeyer, S. B. *Science* 267, 1488-1491 (1995).
83. Morozov, A., Yung, E. y Kalpana, G. V. *Proc. Natl. Acad. Sci. USA* 95, 1120-1125 (1998).
- 25 84. Goulaouic H y Chow SA. (1996) Directed integration of viral DNA mediated by fusion proteins consisting of human immunodeficiency virus type 1 integrase and *E. coli* LexA protein. *J Virol* 70:37-46.
85. Manome Y, Kunieda T, et al. (1998) Transgene expression in malignant glioma using a replication-defective adenoviral vector containing the Egr-1 promoter: activation by ionizing radiation or uptake of radioactive iododeoxyuridine. *Human Gene Ther* 9:1409-1417.
- 30 86. Borrelli MJ, et al. (2001) Heat-activated transgene expression from adenovirus vectors infected into human prostate cancer cells. *Cancer Res* 61:1113-1121.
87. Xie X, Zhao X, et al (2001) Robust prostate-specific expression for targeted gene therapy based on the human kallikrein 2 promoter. *Human Gene Ther* 12:549-561.
88. Furth PA, et al. (1994) Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter. *Proc. Natl. Acad. Sci. USA* 91:9302-9306.
- 35 89. Izsvak Z, Ivics Z, Hackett PB. (1995) Characterization of a Tc-1 like transposable element in zebrafish (*Danio rerio*). *Mol. Gen. Genet.* 247:312-322.

90. Franz G y Savakis C. (1991) Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. *Nucl. Acids Res.* 19:6646.
91. Mernman PJ, Grimes CD, Ambroziak J, Hackett DA, Skinner P. y Simmons MJ. (1995) S elements: a family of Tc1-like transposons in the genome of *Drosophila melanogaster*. *Genetics* 141:1425-1438.
- 5 92. Ke Z, Grossman GL, Cornel AJ, Collins FH. (1996) Quetzal: a transposon of the Tc1 family in the mosquito *Anopheles albimanus*. *Genetica* 98:141-147.
93. Lam WL, Seo P, Robison K, Virk S. y Gilbert W. (1996) Discovery of amphibian Tc1-like transposon families. *J Mol Biol* 257:359-366.
- 10 94. Ivics Z, Izsvak Z, Minter A, Hackett PB. (1996) Identification of functional domains and evolution of Tc1-like transposable elements. *Proc. Natl. Acad Sci USA* 93: 5008-5013.
95. Richardson PD, Augustin LB, Kren BT y Steer CJ. (2002) Gene repair and transposon-mediated gene therapy. *Stem Cells* 20:105-118.
96. Plasterk RH, Izsvak Z, Ivics Z. (1999) Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet* 15:326-332.
- 15 97. Izsvak Z, Ivics Z. y Plasterk RH. (2000) Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J. Mol. Biol.* 302:93-102.
98. Furth, M. E. y Wicker, S. H. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W. y Weisberg, R. A., Eds.) pp 145-155, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
99. Casjen, S. y Hendrix, R. (1988) in *The Bacteriophages* (Calendar, R., Ed.) pp 15-92, Plenum, New York, N.Y.
- 20 100. Black.L. W. (1989) *Ann. Rev. Microbiol.* 43, 267-292.
101. Fujisawa, H. y Morita, M. (1997) *Genes to Cells* 2, 537-545.
102. Catalano, C. E. (2000) *Cellular and Molecular Life Sciences* 57, 128-148.
103. Yang, Q., de Beer, T., Woods, L., Meyer, J., Manning, M., Overduin, M. y Catalano, C. E. (1999) *Biochemistry* 38, 465-477.
- 25 104. Yang, Q., Berton, N., Manning, M. C. y Catalano, C. E. (1999) *Biochemistry* 38, 14238-14247.
105. Landy, A. (1989) *Ann. Rev. Biochem.* 58, 913-949.
106. de Beer, T., Meyer, J., Ortega, M., Yang, Q., Maes, L., Duffy, C., Berton, N., Sippy, J., Overduin, M., Feiss, M. y Catalano, C. E. (2002) *Mol. Cell* 9, 981-991.
- 30 107. J.M. Kaminski, M.R. Huber, J.B. Summers, M.B. Ward, Design of a nonviral vector for site-selective, efficient integration into the human genome, *Faseb J.* 16 (2002) 1242-1247.
108. K. Abremski, R. Hoess, N. Sternberg, Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination, *Cell* 32 (1983)1301-1311.
109. F. Guo, D.N. Gopaul, G.D. van Duyne, Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse, *Nature* 389 (1997) 40-46.

110. H.A. Greisman, C.O. Pabo, A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites, *Science* 275 (1997) 657-661.

111. A.F. Kolb, S.G. Siddell, Genomic targeting with an MBP-Cre fusion protein, *Gene* 183 (1996) 53-60.

5

**G. Secuencias**

SEQ ID No: 1 TCCGGCCGGAACCGGCTTT  
 SEQ ID No: 2 CG GGATCCACCTATGGAAGGAAGATCAGATTT  
 SEQ ID No: 3 AGATTACTCGAGTCAAAGTGTTTTGTATGATCTCG  
 SEQ ID No: 4 AAGATCTGATCCGTCGAC  
 SEQ ID No: 5 CGGGATCCACCTATGAATTATGGCGTGGAGAA  
 SEQ ID No: 6 AGATTACTCGAGTCAGTTGTACAGCTGCAATCCCA  
 SEQ ID No: 7 GGAAGCCCTGCAAAGTAAA

10

15

**LISTADO DE SECUENCIAS**

<110> Kaminski, Joseph

20

<120> Vectores basados en transposones y métodos para la integración de ácidos nucleicos

<130> 11000.0001P1

25

<150> 60/398,628

<151> 2002-07-24

<160> 7

30

<170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 19

<212> DNA

<213> Secuencia Artificial

35

<220>

<223> Descripción de la Secuencia Artificial:/nota = constructo sintético

40

<400> 1

tccggccgga accggcttt

19

<210> 2

<211> 33

<212> DNA

45

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial:/nota = constructo sintético

50

<400> 2

cgggatccca cctatggaag gaagatcaga ttt

33

<210> 3

<211> 35

<212> DNA

55

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial:/nota = constructo sintético

<400> 3  
 5 agattactcg agtcaaagtg tttgtatga tctcg 35

<210> 4  
 <211> 18  
 <212> DNA  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de la Secuencia Artificial:/nota = constructo sintético

<400> 4  
 15 aagatctgat ccgctgac 18

<210> 5  
 <211> 33  
 <212> DNA  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de la Secuencia Artificial:/nota = constructo sintético

<400> 5  
 25 cgggatccca cctatgaatt atggcgtgga gaa 33

<210> 6  
 <211> 35  
 <212> DNA  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de la Secuencia Artificial:/nota = constructo sintético

<400> 6  
 35 agattactcg agtcagttgt acagctgcaa tccca 35

<210> 7  
 <211> 19  
 <212> DNA  
 <213> Secuencia Artificial

40 <220>  
 <223> Descripción de la Secuencia Artificial:/nota = constructo sintético

<400> 7  
 45 ggaagccctg caaagtaaa 19

50

## REIVINDICACIONES

- 5 1. Una composición que comprende una construcción de ácido nucleico que comprende un transgén flanqueado por dos repeticiones terminales y un ácido nucleico que codifica una enzima de integración bajo el control de un elemento promotor, en la que la enzima de integración es una enzima de integración quimérica que comprende un dominio de unión al ADN específico de anfitrión y en la que el ácido nucleico que codifica el transgén y el ácido nucleico que codifica la enzima de integración quimérica son el mismo ácido nucleico, y en la que la enzima de integración quimérica es una transposasa.
2. La composición de la reivindicación 1, en la que el elemento promotor es un promotor/amplificador.
- 10 3. La composición de la reivindicación 1, en la que el promotor es un promotor específico de sitio.
- 15 4. La composición de la reivindicación 3, en la que el promotor específico de sitio puede seleccionarse de al menos el grupo que consiste en el promotor de la proteína fibrilar ácida de la glía (GFAP), promotor de mielina básica (MBP), promotor MCK, promotor NSE, promotor de nestina, promotor de sinapsina, promotor de Insulina 2 (Ins2), promotor PSA, promotor de la albúmina, promotor TRP-1, el promotor de tirosinasa, el promotor EIIA, un promotor específico de tejido de mama, tal como el promotor WAP, un promotor específico de tejido de ovarios, tal como el promotor ACTB o un promotor específico de tejido óseo.
5. La composición de la reivindicación 1, en la que el promotor es inducible.
- 20 6. La composición de la reivindicación 5, en la que el promotor inducible puede seleccionarse de al menos uno del grupo que consiste en el promotor del choque por calor humano, promotor Egr-1, promotor de tetraciclina, sistema recombinasa cre-lox y el promotor de la calicreína glandular humana 2 (hK2).
7. La composición de la reivindicación 1, en la que la transposasa quimérica puede seleccionarse de al menos una del grupo que consiste en Bella Durmiente (SB), mosl, piggyBac, Himarl, Hermes, elemento To 12, Pokey, Tn7, Tn916, maT, Tc1/mariner y Tc3.
- 25 8. La composición de la reivindicación 1, en la que el dominio de unión al ADN específico del anfitrión de la enzima de integración quimérica se fusiona con el extremo N terminal de la transposasa.
9. La composición de la reivindicación 1, en la que el dominio de unión al ADN específico del anfitrión de la enzima de integración quimérica se fusiona con el extremo C terminal de la transposasa.
- 30 10. La composición de la reivindicación 1, en la que la secuencia de ácido nucleico que codifica la enzima de integración quimérica está localizada fuera de las repeticiones terminales.
- 35 11. La composición de una cualquiera de las reivindicaciones precedentes, en la que el dominio de unión al ADN específico del anfitrión se selecciona del grupo que consiste en proteínas con homeodominio, proteínas con dedo de cinc, receptores de hormonas, proteínas hélice-giro-hélice, proteínas hélice-bucle-hélice, proteínas cremallera básicas (bZip) y factores lazo beta.
12. Una construcción de ácido nucleico que comprende i) un transgén flanqueado por dos repeticiones terminales, ii) una enzima de integración quimérica en la forma de una transposasa bajo el control de un elemento promotor capaz de dirigir la expresión de dicha transposasa, y iii) un dominio de unión al ADN específico del anfitrión.
13. Una composición según una cualquiera de las reivindicaciones 1-11 o una construcción de ácido nucleico según la reivindicación 12 para usarse como un medicamento.

Figura 1

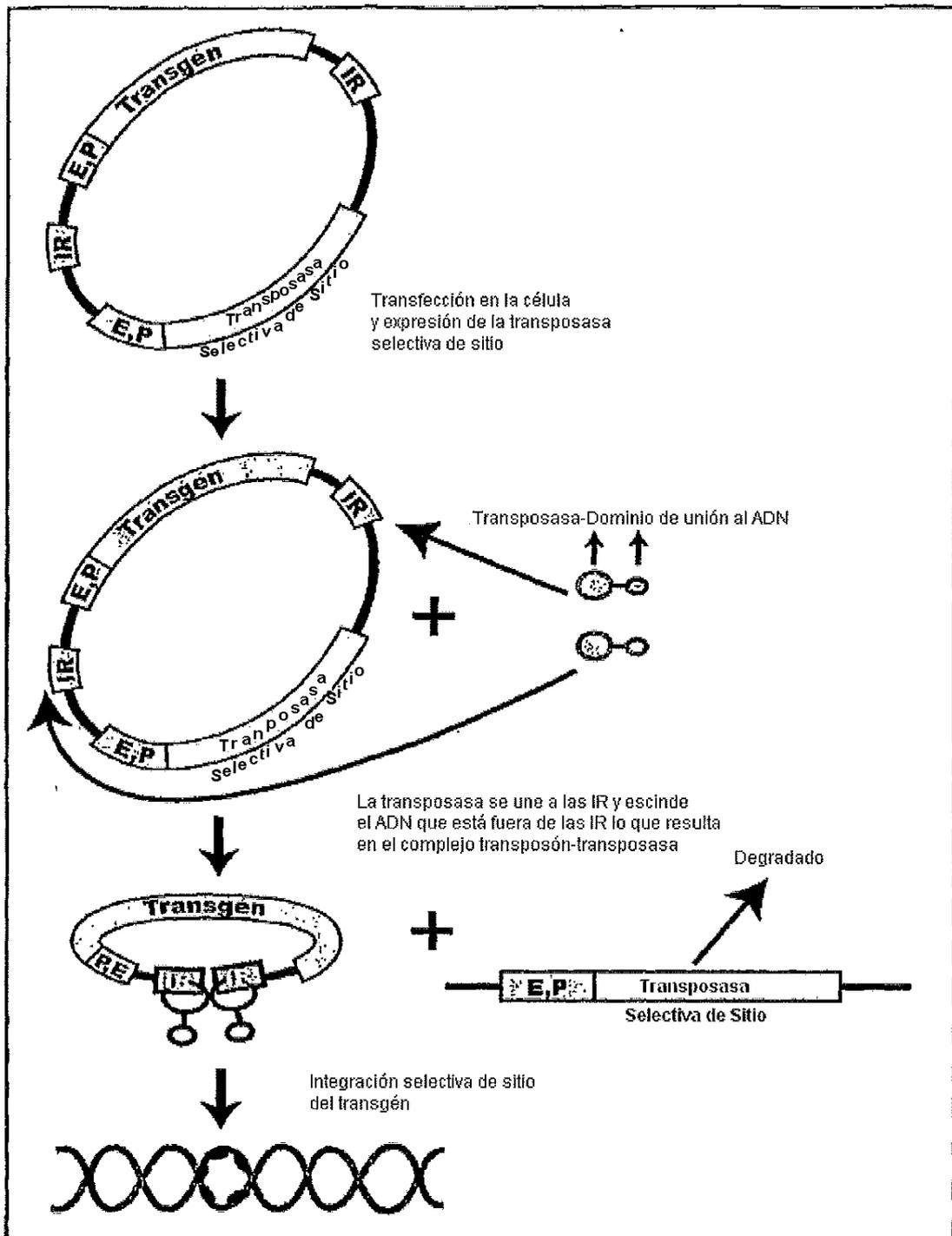
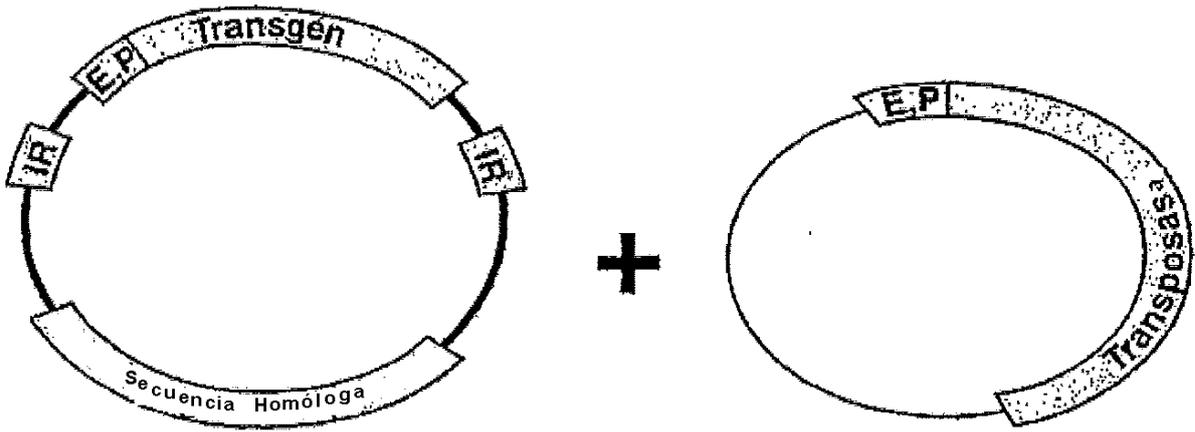


Figura 2



5

Figura 3

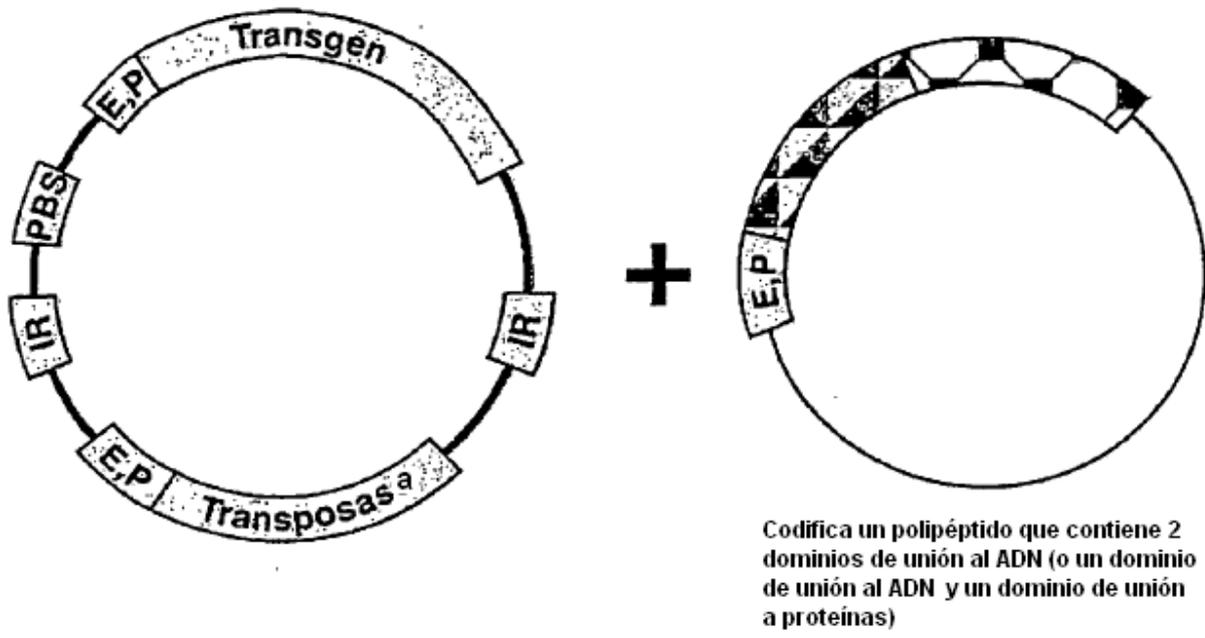


Figura 4

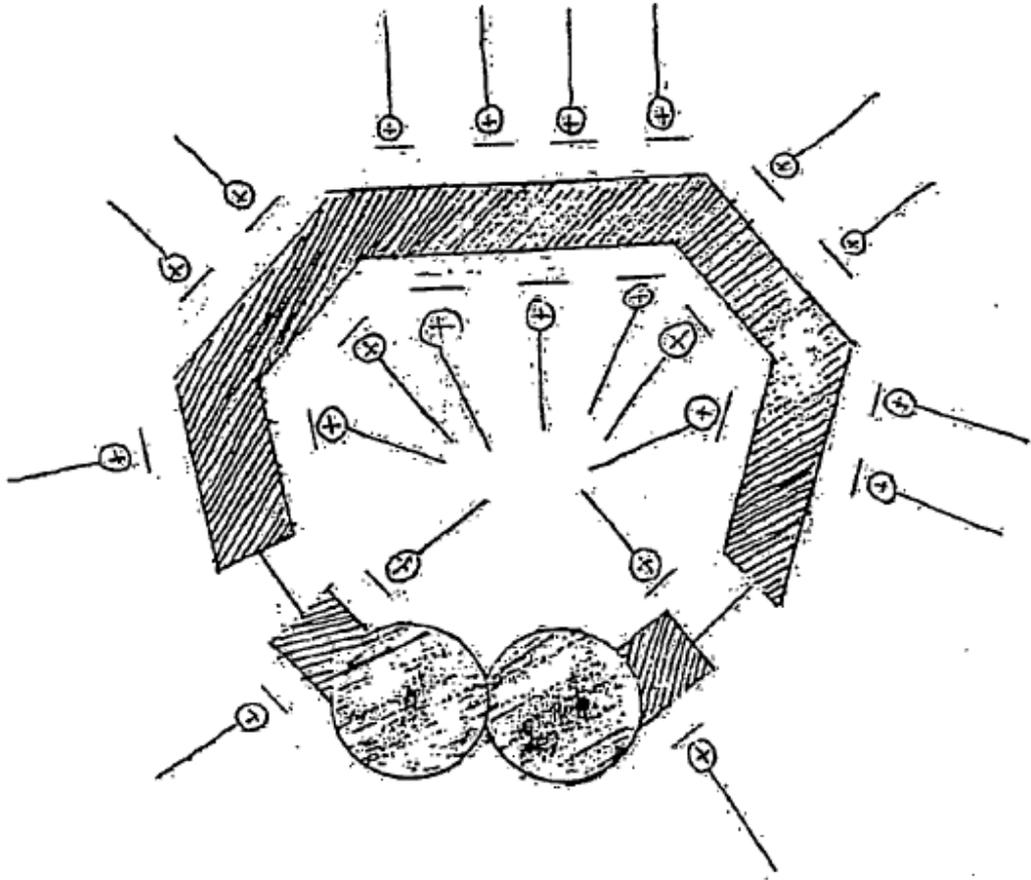


Figura 5

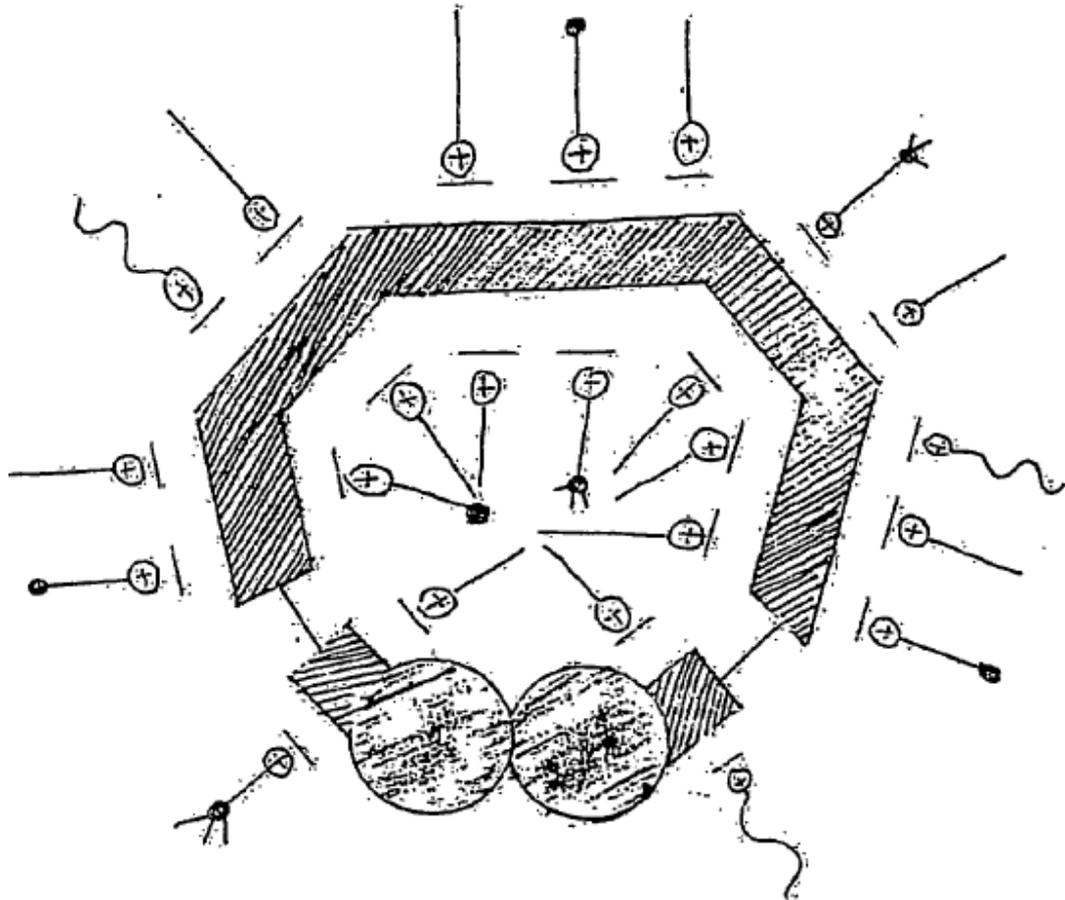


Figura 6

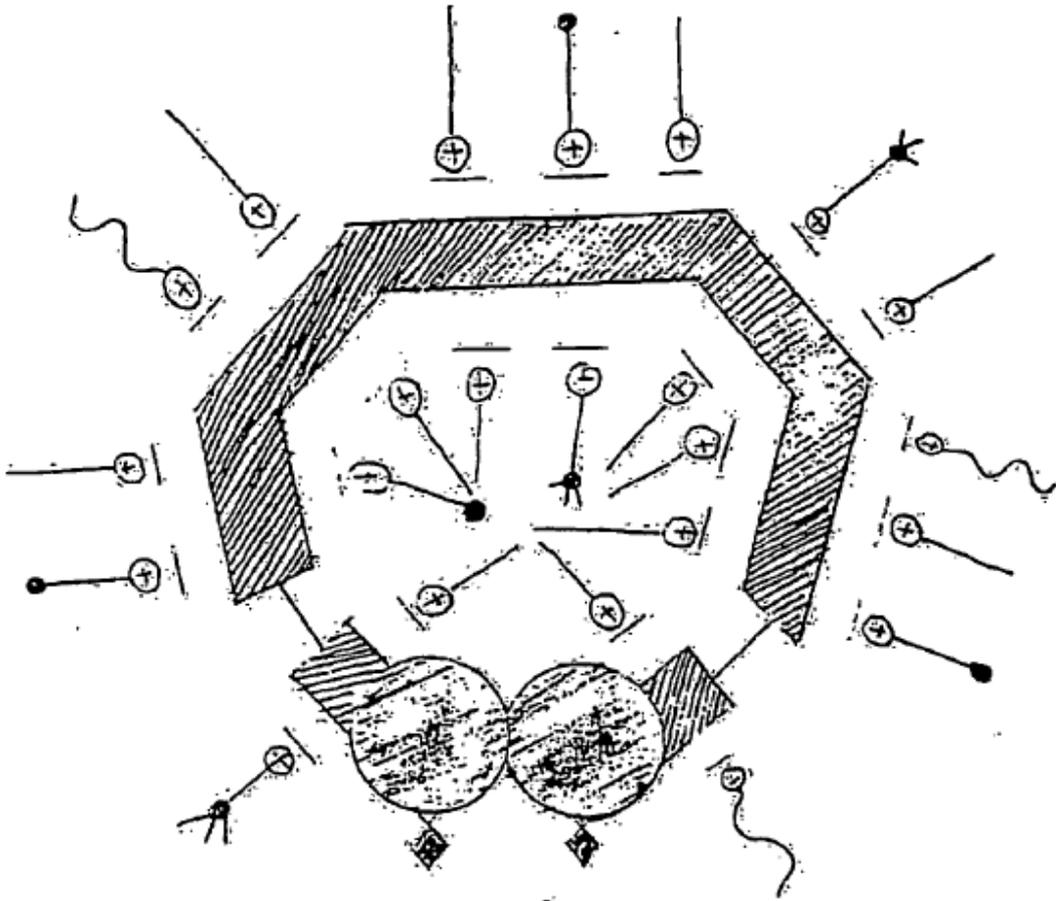
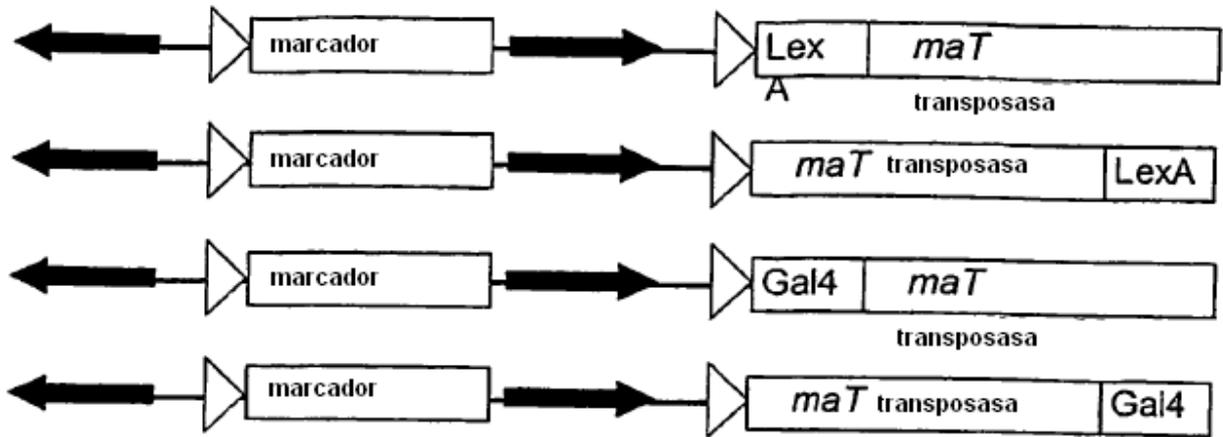
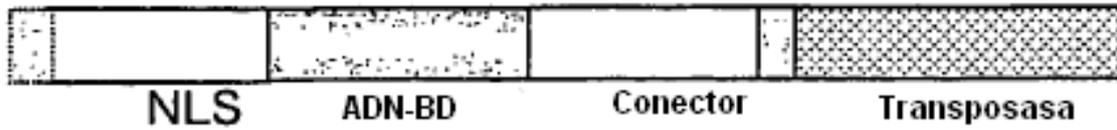


Figura 7



5

Figura 8



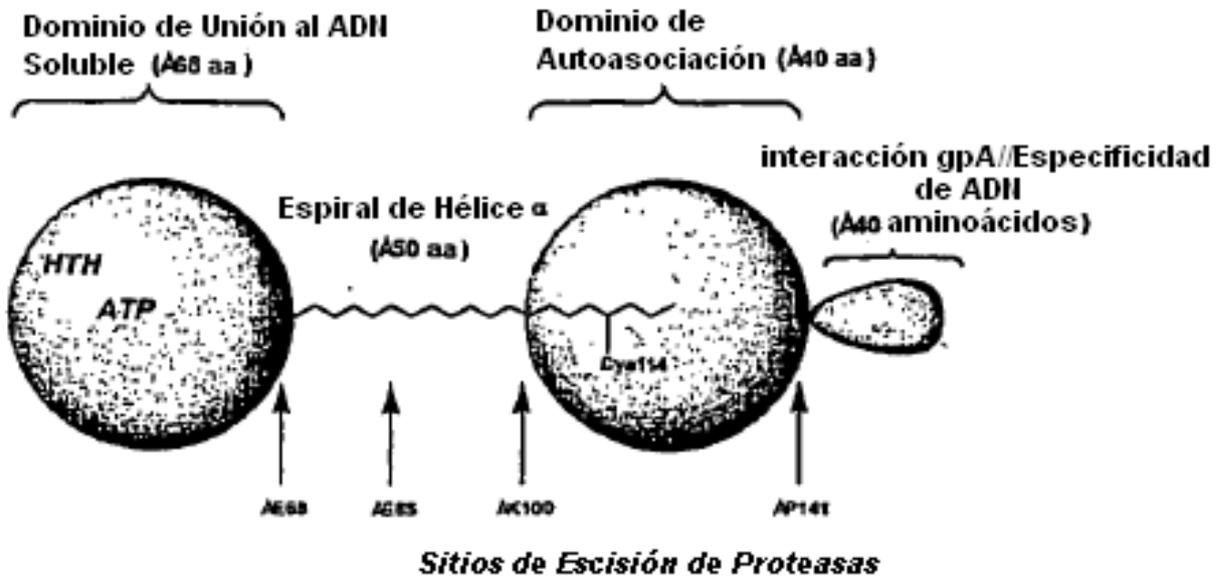
**NLS - Señal de Localización Nuclear**

**ADN-BD - Dominio de unión al ADN de la proteína Gal4 o LexA**

**Conector - Cadena flexible de aminoácidos**

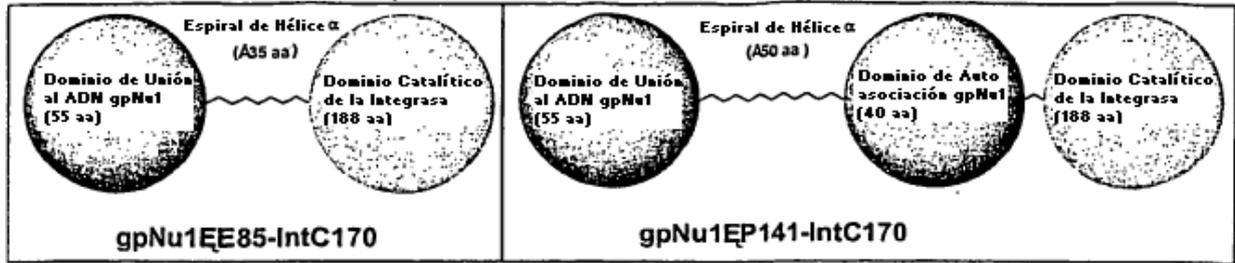
**Transposasa - Transposasa *Mos1* o *piggyBac***

Figura 9



5

Figura 10



5

Figura 11

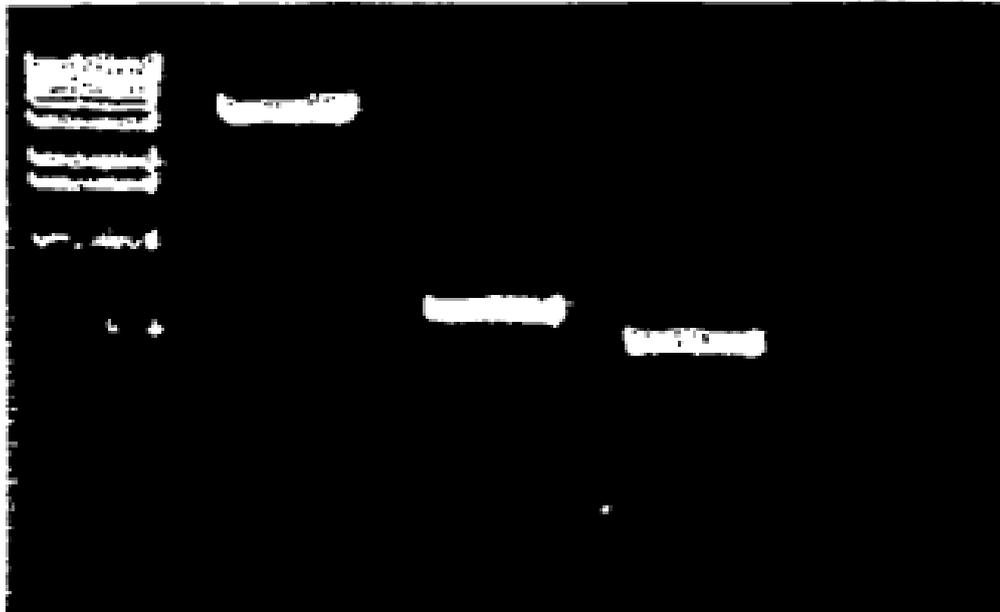


Figura 12

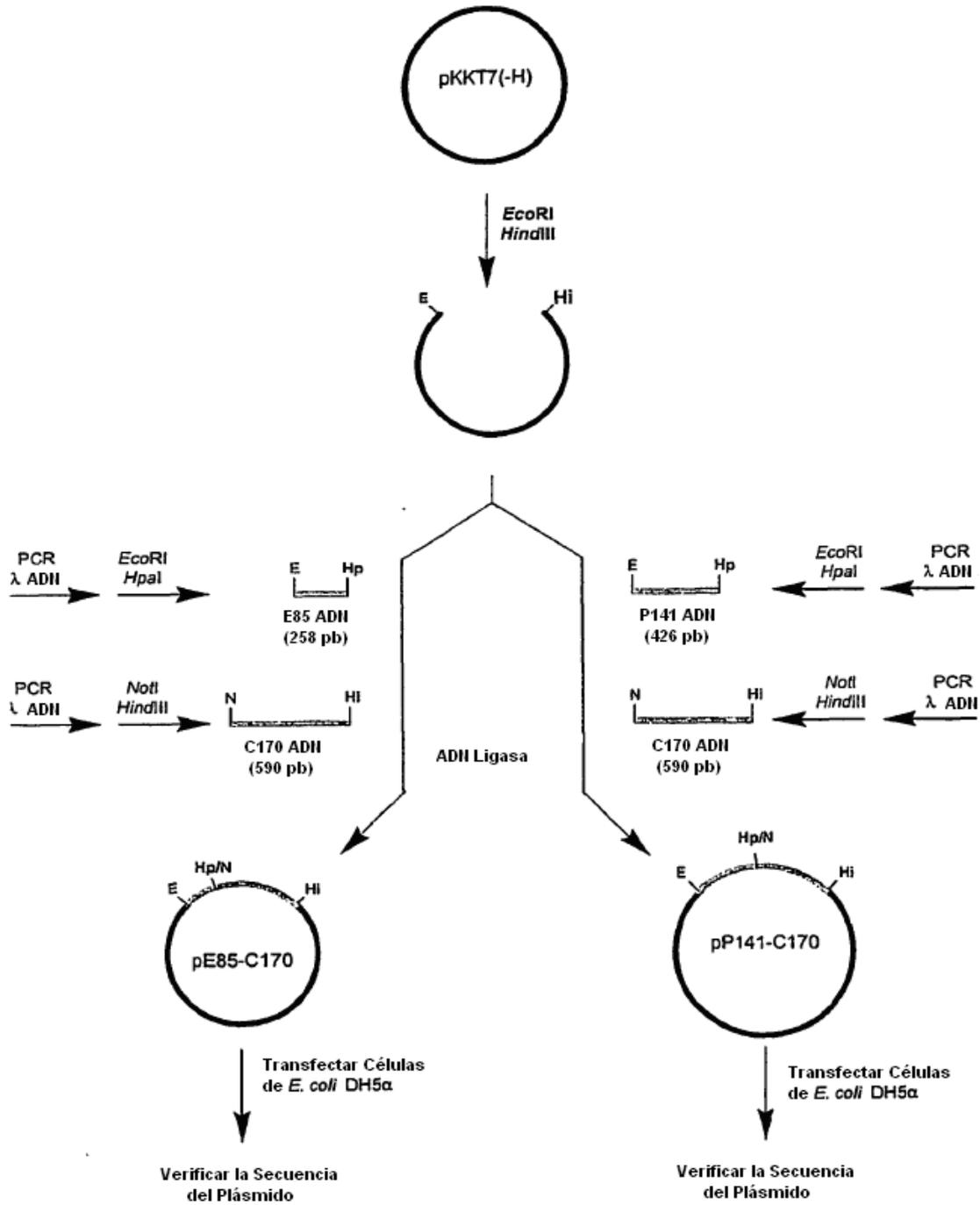
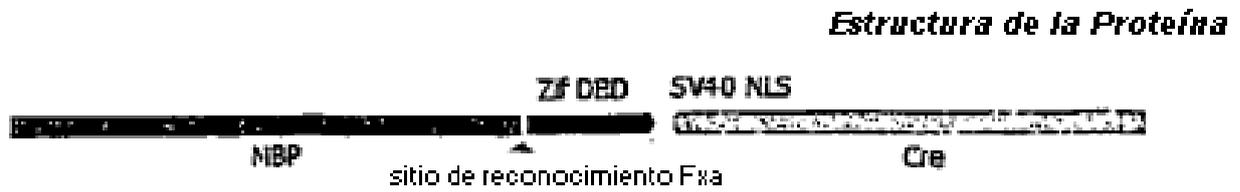


Figura 13



*componentes de la proteína:*

|          |   |
|----------|---|
| MBP      | proteína de unión a maltosa de E. coli                      |
| DBD Zif  | dominio de unión al ADN de Zif                              |
| Fxa.     | señal de reconocimiento para la proteasa factor Xa          |
| SV40 NLS | señal de localización nuclear del antígeno T grande de SV40 |
| Cre      | proteína Cre del bacteriófago P1                            |

Figura 14

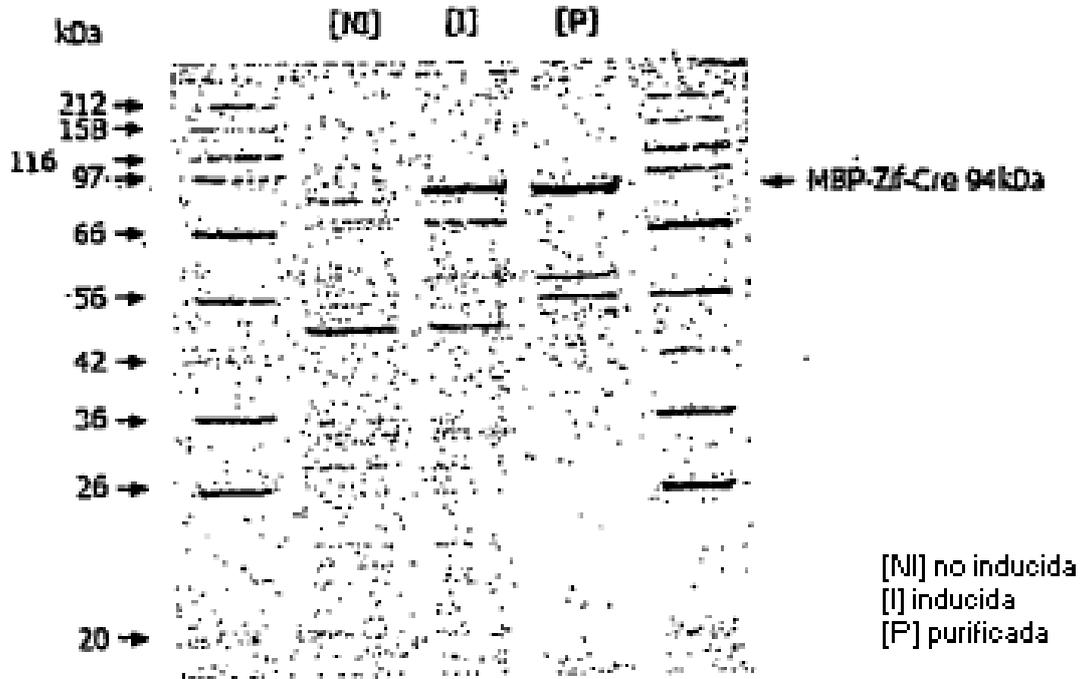


Figura 15

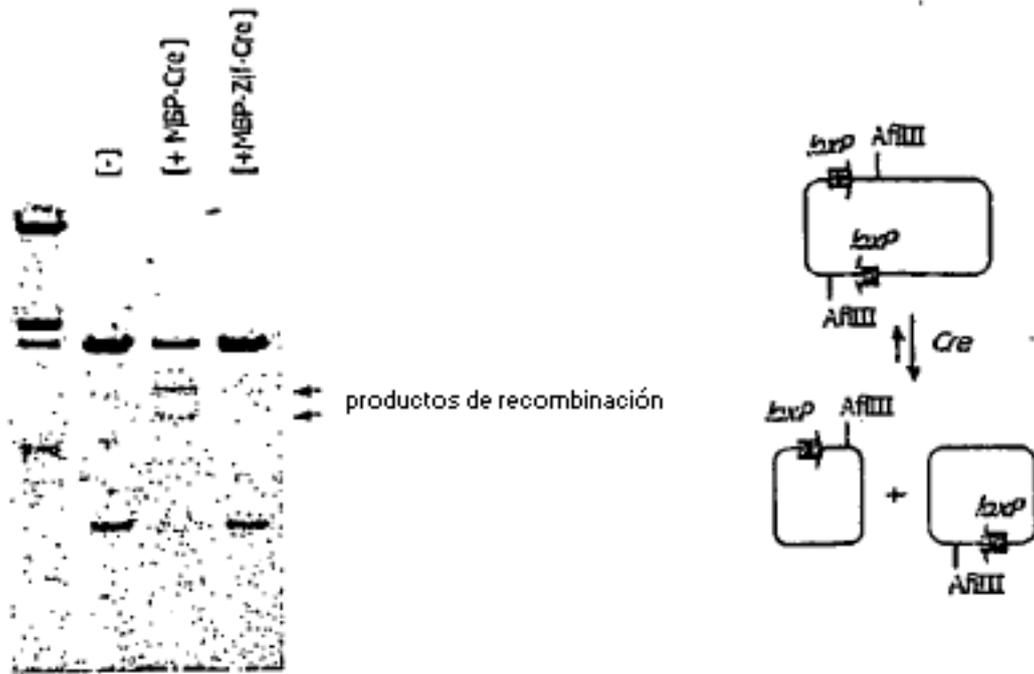


Figura 16



Figura 17

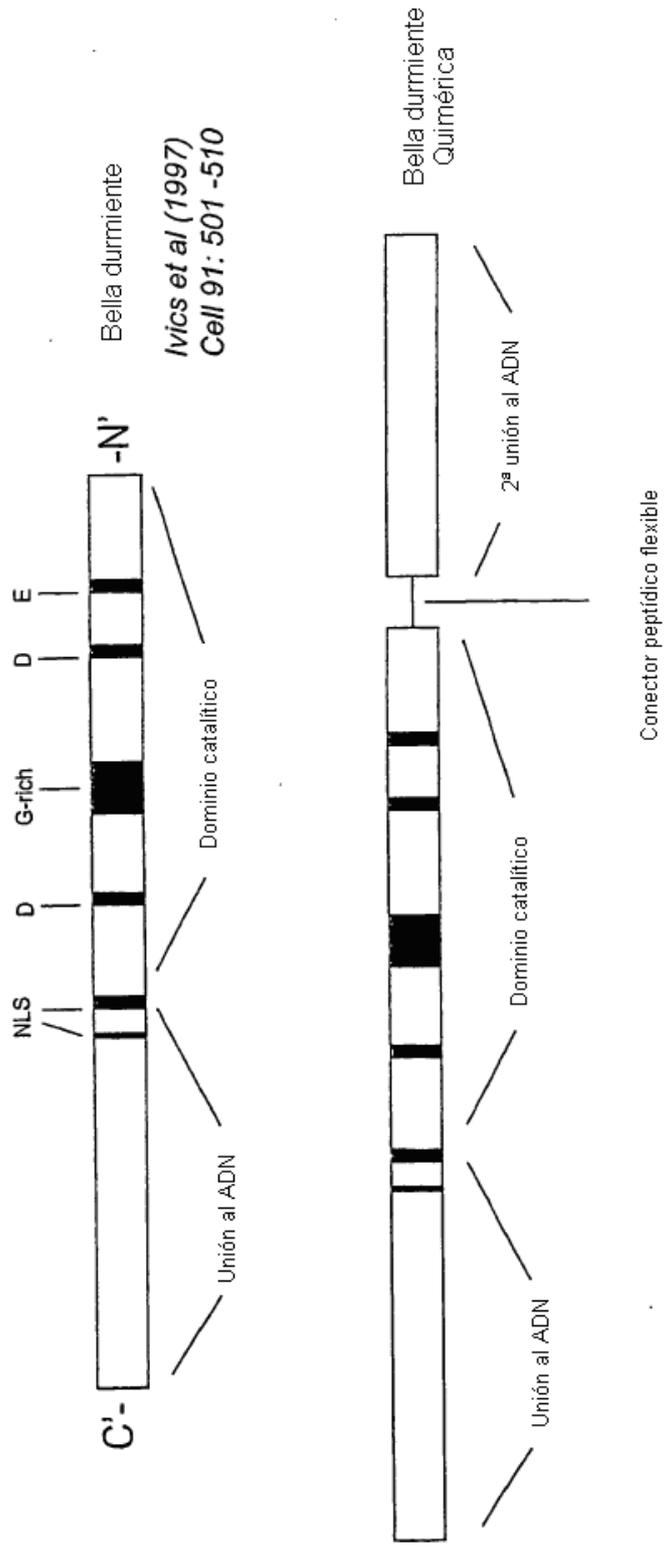


Figura 18

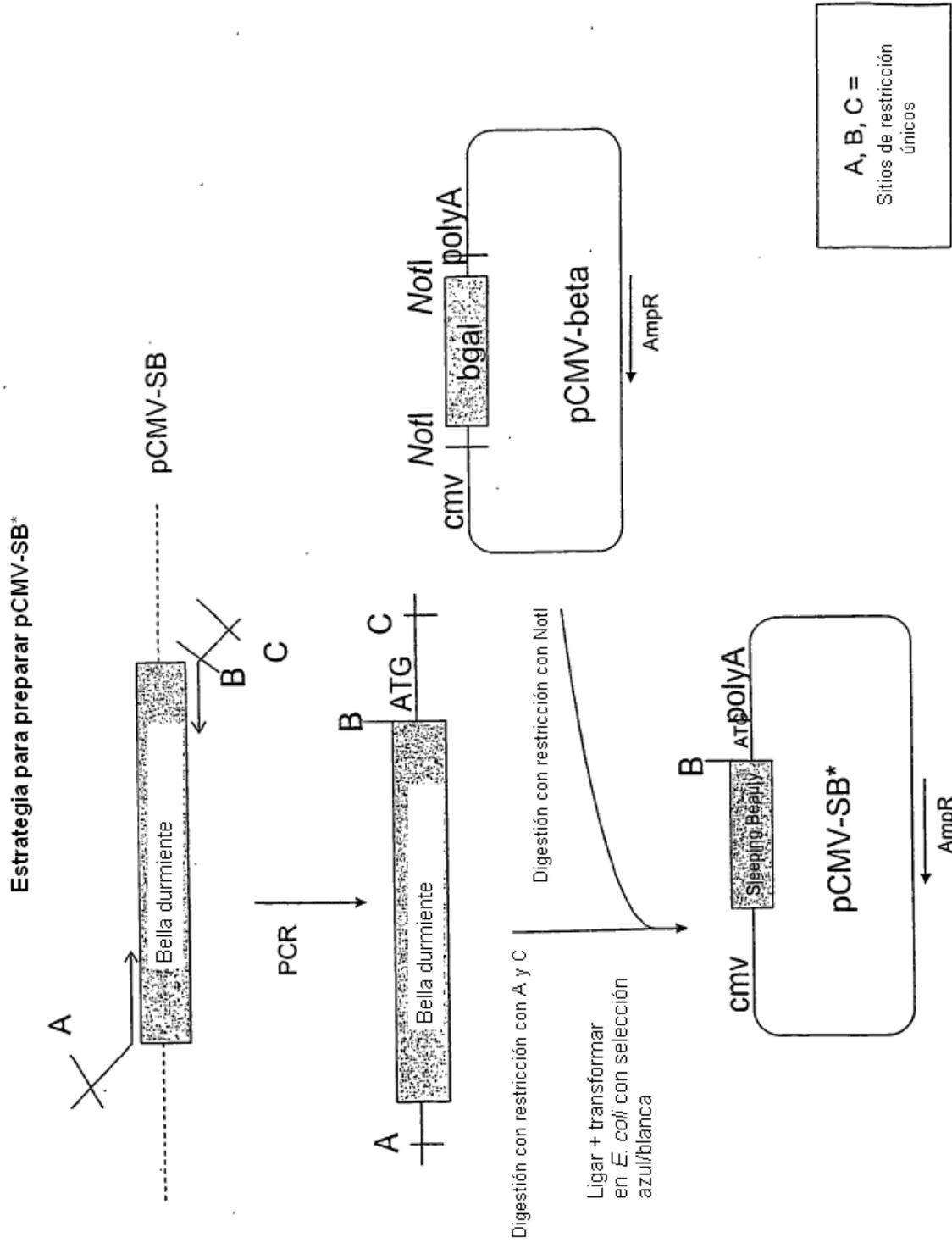


Figura 19

Estrategia para preparar pSB-L-ZF

