

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 603 062**

51 Int. Cl.:

C12P 19/34 (2006.01)
C12N 15/09 (2006.01)
C12N 15/11 (2006.01)
C12N 15/63 (2006.01)
C07H 21/04 (2006.01)
C12N 15/85 (2006.01)
C12N 15/64 (2006.01)
C12N 15/66 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.05.2005 E 11009458 (8)**

97 Fecha y número de publicación de la concesión europea: **17.08.2016 EP 2484772**

54 Título: **Métodos de ensamblaje dinámico de vectores para la clonación de ADN de vectores plasmídicos**

30 Prioridad:

18.05.2004 US 572011 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.02.2017

73 Titular/es:

**INTREXON CORPORATION (100.0%)
1750 Kraft Drive, Suite 1400
Blacksburg, VA 24060 , US**

72 Inventor/es:

REED, THOMAS D.

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 603 062 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

Métodos de ensamblaje dinámico de vectores para la clonación de ADN de vectores plasmídicos

5 **Antecedentes de la invención**

La presente invención se refiere en general al campo de los vectores plasmídicos de clonación, y en particular a los métodos para ensamblar rápidamente transgenes con los vectores plasmídicos de clonación, concretamente a un método para sintetizar simultáneamente una matriz de transgenes. La base de la biología molecular es la tecnología de ADN recombinante, que se puede resumir aquí como la modificación y propagación de ácidos nucleicos con el fin de estudiar la estructura y función de ácidos nucleicos y sus productos proteicos. Los genes individuales, regiones génicas reguladoras, subgrupos de genes, y de hecho los cromosomas completos en los que están contenidos, están comprendidos por secuencias anti-paralelas de doble cadena de los nucleótidos adenina, timina, guanina y citosina, que se identifican convencionalmente por las iniciales A, T, G, y C, respectivamente. Estas secuencias de ADN, así como las secuencias de ADNc, que son copias de ADN de doble cadena derivadas de moléculas de ARNm (ARN mensajero), se pueden escindir en fragmentos distintos, aislarse, e insertarse en un vector tal como un plásmido bacteriano para estudiar los productos génicos. Un plásmido es una porción de ADN extra-cromosómica que originalmente se derivó de bacterias, y que se pueden modificar y re-introducir en una bacteria huésped con el fin de estudiar o producir un producto génico. El ADN de un plásmido es similar en todo al ADN cromosómico, ya que está compuesto de los mismos nucleótidos A, T, G y C que codifican genes y regiones reguladoras génicas, sin embargo, es una molécula relativamente pequeña que comprende menos de aproximadamente 30.000 pares de bases, o 30 kilobases (kb). Además, los pares de bases de nucleótidos de un plásmido de doble cadena forman una molécula circular continua, lo que también distingue el ADN del plásmido del ADN cromosómico.

Los plásmidos aumentan el intercambio rápido de material génico entre organismos bacterianos y permiten la rápida adaptación a los cambios en el entorno, tales como temperatura, suministro de alimento, u otros desafíos. Cualquier plásmido adquirido debe expresar un gen o genes que contribuyan a la supervivencia del huésped, ya que si no, será destruido o desechado por el organismo, debido a que el mantenimiento de plásmidos innecesarios sería un derroche de recursos. Una población clónica de células contiene un material génico idéntico, incluyendo cualquier plásmido que pueda albergar. El uso de un vector plasmídico de clonación con una inserción de ADN en dicha población clónica de células huésped amplificará la cantidad del ADN de interés que está disponible.

El ADN que se clona de esta manera se puede aislar luego y recuperarse para una modificación posterior en las etapas que sean necesarias para construir una construcción de ADN. Por lo tanto, se puede apreciar que los vectores plasmídicos de clonación son herramientas útiles en el estudio de la función génica, proporcionando la capacidad para producir rápidamente grandes cantidades de la inserción de ADN de interés.

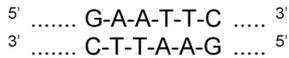
Aunque algunos elementos que se encuentran en los plásmidos son de origen natural, otros se han modificado para aumentar la utilidad de los plásmidos como vectores de ADN. Estos incluyen los genes de resistencia a antibióticos o química y un sitio de clonación múltiple (MCS) entre otros. Cada uno de estos elementos tiene un papel en la presente invención, así como en la técnica anterior. La descripción del papel que tiene cada elemento resaltarán las limitaciones de la técnica anterior y demostrará la utilidad de la presente invención.

Un gen surgido de un plásmido particularmente útil que puede adquirir un huésped es el que confiere resistencia a antibióticos. En la práctica diaria de la tecnología de ADN recombinante, los genes de resistencia a antibióticos se explotan como elementos de selección positiva o negativa para aumentar preferentemente el cultivo y la amplificación del plásmido deseado sobre la de otros plásmidos.

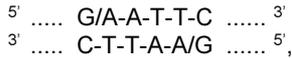
Con el fin de que un plásmido sea mantenido por una bacteria huésped, debe contener también un segmento de secuencias que dirijan al huésped para que duplique el plásmido. Secuencias conocidas como el elemento de origen de replicación (ORI) dirigen al huésped para que use sus enzimas celulares para producir copias del plásmido. Cuando dicha bacteria se divide, cada una de las células hijas mantendrá una copia o copias de cada uno de dicho plásmido. Ciertas cepas de bacterias *E. coli* se han derivado para maximizar esta duplicación, produciendo hasta 300 copias por bacteria. De esta manera, se puede aumentar el cultivo de un plásmido deseado.

Otro elemento esencial en cualquier vector de clonación es una localización para la inserción de los materiales génicos de interés. Esto es un elemento sintético que se tiene que modificar en plásmidos de "tipo silvestre", que confiere de esta manera utilidad como vector de clonación. Cualquier vector plasmídico de clonación típico disponible en el mercado contiene al menos una de dichas regiones, que se conoce como sitio de clonación múltiple (MCS). Un MCS normalmente comprende secuencias de nucleótido que han sido escindidas por una única enzima endonucleasa, o una serie de enzimas endonucleasas, cada una de las cuales tiene una secuencia de reconocimiento y un patrón de escisión distintos. Las denominadas secuencias de reconocimiento de un sitio de endonucleasa de restricción (RE) codificado en la molécula de ADN comprende secuencias palindrómicas de doble cadena. Para algunas enzimas RE, son suficientes solo unos 4-6 nucleótidos para proporcionar un sitio de reconocimiento, mientras que algunas enzimas RE necesitan una secuencia de 8 o más nucleótidos. La enzima RE EcoRI, por ejemplo, reconoce la secuencia hexanucleotídica de doble cadena: $5' \text{G-A-A-T-T-C} 3'$, en la que $5'$ indica

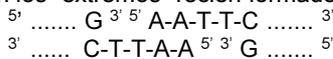
el final de la molécula que se considera el extremo "secuencia arriba", y 3' de igual manera indica el extremo "secuencia abajo". La cadena complementaria de la secuencia de reconocimiento debería ser su cadena anti-paralela, 3' G-A-A-T-T-C- 5'. Como cada sitio de endonucleasa es una secuencia de nucleótidos de doble cadena, un sitio de reconocimiento de 6 nucleótidos es, de hecho, de 6 pares de bases (pb). Por lo tanto el sitio de reconocimiento de doble cadena se puede representar con la molécula de doble cadena más grande en la cual se produce como:



Como muchas otras enzimas RE, EcoRI no escinde exactamente en el eje simétrico de la díada, sino en posiciones separadas cuatro nucleótidos en las dos cadenas de ADN entre los nucleótidos que se indica con una "P".



de manera que la molécula de ADN de doble cadena se escinde y tiene la configuración de nucleótidos resultante con los "extremos" recién formados.



Esta escisión escalonada da lugar a fragmentos de ADN que protruyen en el extremo 5'. Debido a que las parejas A-T y G-C se forman espontáneamente cuando se aproximan unas a otras, la protrusión de los extremos tales como estos se llaman extremos cohesivos o pegajosos. Cualquiera de estos extremos puede formar enlaces de hidrógeno con otros extremos complementarios escindidos con la misma enzima de restricción. Como cada ADN que contiene una secuencia de reconocimiento específica se cortará de la misma manera que el otro ADN que contienen la misma secuencia, los extremos escindidos serán complementarios. Por lo tanto, los extremos de cada molécula de ADN cortados con la misma enzima RE "se corresponden" entre ellos como la manera en que "encajan" las piezas adyacentes de un puzle y se pueden unir enzimáticamente entre ellos. Esta es la propiedad que permite la formación de moléculas de ADN recombinante, y permite la introducción de fragmentos de ADN ajeno en los plásmidos bacterianos, o en cualquier otra molécula de ADN.

Un principio general adicional que se tiene que considerar cuando se construyen moléculas de ADN recombinante es que todos los sitios de endonucleasa que existen en una molécula se cortarán con una enzima RE particular, no solo el sitio de interés. Cuanto mayor es una molécula de ADN, más probable es que un sitio de endonucleasa se repita. Asumiendo que cada uno de los sitios de endonucleasa se distribuyen aleatoriamente a lo largo de la molécula de ADN, un sitio tetranucleotídico estará presente, en promedio, una vez cada 4^4 (es decir, 256) nucleótidos o pb, mientras que un sitio hexanucleotídico estará presente una vez cada 4^6 (es decir, 4096) nucleótidos o pb, y los sitios octanucleotídico estarán presentes cada 4^8 (es decir, 114.688) nucleótidos o pb. Por lo tanto, se puede apreciar fácilmente que las secuencias de reconocimiento más cortas estarán presentes frecuentemente, mientras que las más largas existirán más raramente. Cuando se planea la construcción de un transgén u otra molécula de ADN recombinante, esto es un problema vital, ya que dicho proyecto necesita frecuentemente el ensamblaje de varias porciones de ADN de distintos tamaños. Cuanto más largas son estas porciones, más probable es que los sitios que se desean utilizar estén presentes en varias porciones de los componentes de ADN, haciendo la modificación difícil en el mejor de los casos.

Se hace referencia a los sitios de una enzima endonucleasa que existen frecuentemente como sitios comunes, y se hace referencia a las endonucleasas que escinden estos sitios como enzimas endonucleasas comunes. Se hace referencia a las enzimas de restricción con sitios de restricción equivalentes de más de 6 pb como enzimas de restricción raras y a sus sitios de restricción equivalentes como sitios de restricción raros. Sin embargo, hay algunos sitios de endonucleasa de 6 pb que están presentes con menos frecuencia de lo que se preveía estadísticamente, y también se hace referencia a estos sitios y a las endonucleasas que los escinden como raros. Por lo tanto las denominaciones "raro" y "común" no se refieren a la abundancia o disponibilidad relativas de cada enzima de restricción en particular, sino más bien a la frecuencia de la presencia de la secuencia de nucleótidos que produce su sitio de reconocimiento equivalente en cualquier molécula de ADN o fragmento aislado de una molécula de ADN, o cualquier gen de su secuencia de ADN.

Se ha aislado recientemente una segunda clase de enzimas endonucleasa, llamadas enzimas endonucleasas buscadoras (HE). Las enzimas HE tienen grandes sitios de reconocimiento asimétricos no palindrómicos (12-40 pares de bases). Los sitios de reconocimiento de HE son extremadamente raros. Por ejemplo, la HE conocida como I-SceI tiene un sitio de reconocimiento de 18 pb, ((5'...TAG-GGATAACAGGGTAAT...3'), que se prevé que está presente cada 7×10^{10} pb de la secuencia aleatoria. Esta tasa de presencia es equivalente a un solo sitio en un genoma del tamaño de 20 mamíferos. La rara naturaleza de los sitios de reconocimiento de HE aumenta enormemente la probabilidad de que un ingeniero genetista pueda cortar un producto transgénico final sin alterar la integridad del transgén si se incluyeran sitios de reconocimiento de HE en localizaciones apropiadas en un vector plasmídico de clonación.

Como una molécula de ADN de cualquier organismo fuente se cortará de manera idéntica por una enzima endonucleasa, las porciones ajenas de ADN de cualquier especie se pueden cortar con una enzima endonucleasa, insertada en un vector plasmídico bacteriano que se escinda con la misma enzima endonucleasa y se amplifique en una célula huésped adecuada. Por ejemplo, si un gen humano se puede cortar en 2 lugares con la enzima RE conocida como EcoRI, el fragmento deseado con los extremos EcoRI se pueden aislar y mezclar con un plásmido que también se haya cortado con EcoRI en lo que se conoce comúnmente como una mezcla de unión. En condiciones apropiadas en la mezcla de unión, algunos de los fragmentos del gen humano aislado se combinarán con los extremos de las moléculas de plásmido. Estos extremos recién unidos se pueden unir entre ellos (ligarse) para re-circularizar enzimáticamente el plásmido que contiene ahora su nueva inserción. La mezcla de unión se introduce entonces en *E. coli* u otro huésped adecuado, y los plásmidos modificados recientemente se amplificarán según se divida la bacteria. De esta manera, se pueden obtener y recolectar un número de copias relativamente grande del gen humano a partir de la bacteria. Estas copias génicas se pueden alterar entonces adicionalmente con fines de investigación, análisis, o producción de su producto génico proteico.

La tecnología de ADN recombinante se incorpora frecuentemente en la generación de los denominados "transgenes". Los transgenes frecuentemente comprenden varios materiales génicos que se derivan de uno o más organismos donantes y se introducen en un organismo huésped. Normalmente, un transgén se construye utilizando un vector de clonación como el punto de inicio o "almacén" del proyecto, y se planea una serie de complejas etapas de clonación para ensamblar el producto final con ese vector. Los elementos de un transgén, que comprende las secuencias de nucleótido, incluyen, pero no se limitan a 1) un promotor regulador y/o elementos amplificadores, 2) un gen que se expresará como una molécula de ARNm, 3) elementos de ADN que proporcionan la estabilización del mensaje del ARNm, 4) secuencias de nucleótido que imitan las regiones génicas intrónicas de mamífero, y 5) las señales de ARNm que procesan dicha cola de poli-A añadida al final de los ARNm de origen natural. En algunos casos, un diseño experimental puede necesitar la adición de una señal de localización para proporcionar el transporte del producto génico a una localización subcelular particular.

Cada uno de los elementos de un transgén se puede derivar de un fragmento de una molécula de ADN mayor que se corta a partir de un genoma donante, o, en algunos casos se sintetiza en un laboratorio. Cuando la presente invención emplea endonucleasas para los métodos que se reivindican en el presente documento, se sabe que cada uno de los elementos más pequeños que comprenden, por ejemplo, las inserciones o módulos que se utilizan en los métodos del presente documento, se pueden crear por síntesis *de novo*, recombinación, y/o clonación por PCR de la protuberancia terminadora. Dicho método de síntesis de los elementos que componen un transgén incluye el método desvelado por Jarrell et al. en la Pat. de EE. UU. N° 6.358.712. Aunque Jarrell desvela un método para "soldar" entre ellos los elementos de transgén, solo los métodos de la presente invención desvelan una manera "des-soldar" y re-ensamblar los elementos una vez que se habían ensamblado. De acuerdo con un aspecto de la invención, cada porción se ensambla con las otras en un orden y orientación 5'-3' precisos en un vector plasmídico de clonación.

El promotor de cualquier gen se puede aislar como un fragmento de ADN y colocar en una molécula sintética, tal como un plásmido, para dirigir la expresión de un gen deseado, asumiendo que se proporcionan las condiciones necesarias para la estimulación del promotor de interés. Por ejemplo, se pueden aislar las secuencias promotoras del gen de insulina, colocarlas en un vector plasmídico de clonación junto con un gen indicador, y utilizarlas para estudiar las condiciones necesarias para la expresión del gen de insulina en un tipo celular apropiado. De manera alternativa, el promotor del gen de insulina se puede unir con la secuencia codificante de la proteína de cualquier gen de interés en un vector plasmídico de clonación, y utilizarse para dirigir la expresión del gen de interés en células que expresan insulina, asumiendo que están presentes todos los elementos necesarios en el transgén de ADN construido de esa manera.

Un gen indicador es un componente particularmente útil de algunos tipos de transgenes. Un gen indicador comprende secuencias de nucleótido que codifican una proteína que se expresará bajo la dirección de un promotor particular de interés al que está unido en un transgén, proporcionando una respuesta bioquímica de la actividad del promotor. Un gen indicador es normalmente fácil de detectar o medir respecto al fondo de proteínas celulares endógenas. Los genes indicadores que se utilizan habitualmente incluyen, pero no se limitan a, LacZ, proteína fluorescente verde, y luciferasa, y otros genes indicadores, muchos de los cuales son bien conocidos por los expertos en la técnica.

Los intrones, que son regiones no codificantes de genes de mamífero, no se encuentran en los genomas bacterianos, pero se necesitan para la formación apropiada de moléculas de ARNm en células de mamífero. Por lo tanto, cualquier construcción de ADN para su uso en sistemas mamíferos debe tener al menos un intrón. Los intrones se pueden aislar de cualquier gen de mamífero e insertarse en una construcción de ADN, junto con sus señales de corte y empalme apropiadas que permiten a las células de mamífero escindir el intrón y empalmar entre ellos los extremos restantes de ARNm.

Un elemento de estabilización del ARNm es una secuencia de ADN que es reconocido por las proteínas de unión que protegen algunos ARNm de la degradación. La inclusión de un elemento de estabilización del ARNm frecuentemente aumentará el nivel de expresión génica a partir del ARNm en algunos tipos de células de mamífero, y de esta manera pueden ser útiles en algunas construcciones de ADN o transgenes. Un elemento de estabilización

de ARNm se puede aislar del ADN o ARN de origen natural, o producirse sintéticamente para la inclusión en una construcción de ADN.

5 Una señal de localización es una secuencia de ADN que codifica una proteína de señal para el enrutamiento subcelular de una proteína de interés. Por ejemplo, una señal de localización nuclear dirigirá una proteína al núcleo; una señal de localización de membrana plasmática la dirigirá a la membrana plasmática, etc. Por lo tanto, una señal de localización se puede incorporar a una construcción de ADN para promover la translocación de su producto proteico hacia la localización subcelular deseada.

10 Se puede codificar una secuencia marcadora en una construcción de ADN de manera que el producto proteico tendrá unida una región única. Esta región única funciona como un marcador proteico que puede distinguirlo de su equivalente endógeno. De manera alternativa, puede funcionar como un identificador que puede detectarse por una amplia variedad de técnicas bien conocidas en la técnica, incluyendo, pero sin limitarse a, RT-PCR, inmunohistoquímica, o hibridación in situ.

15 Con un transgén complejo, o con uno que incluya regiones particularmente grandes de ADN, hay un aumento de probabilidades de que haya múltiples sitios de reconocimiento de endonucleasas en estas porciones de ADN. Recordando que las secuencias de reconocimiento que codifican uno cualquiera de los sitios hexanucleotídicos están presentes cada 4096 pb (4⁶). Si se van a ensamblar una secuencia promotora que tenga 3000 pb y un gen de interés de 1500 pb en un vector de clonación de 3000 pb, es muy probable estadísticamente que muchos sitios de 6 o menos nucleótidos no sean útiles, mientras que cualquiera de los sitios utilizables debe estar presente solo en dos de las porciones. Además, los sitios deben estar presentes en las áreas apropiadas de las moléculas apropiadas que se van a ensamblar. Además, la mayoría de los proyectos de clonación necesitarán tener añadidos elementos de ADN adicionales, incrementando de esta manera la complejidad de la molécula en crecimiento y la probabilidad de repeticiones inoportunas de cualquier sitio de restricción particular. Como cualquier enzima de restricción cortará todos sus sitios en una molécula, si se repite un sitio de enzima de restricción endonucleasa, todos los sitios inoportunos se cortarán junto con los sitios deseados, alterando la integridad de la molécula. Por lo tanto, cada etapa de clonación debe ser planeada cuidadosamente de manera de no alterar el crecimiento de la molécula cortándola con una enzima endonucleasa que ya se ha utilizado para incorporar un elemento precedente. Y finalmente, cuando un investigador desee introducir un transgén completado en un organismo mamífero, la construcción del transgén completamente ensamblada frecuentemente debe estar alineada con un sitio único de reconocimiento en al menos un extremo del transgén, necesitando por tanto otro único sitio de reconocimiento que se encuentre en cualquier lugar de la construcción. Como la mayoría de las construcciones de ADN se diseñan para un único fin, se piensa poco en cualquier modificación futura que pueda ser necesaria, lo que aumenta adicionalmente la dificultad para cambios experimentales futuros.

Tradicionalmente, el diseño y construcción de un transgén consume cantidades significativas de tiempo y energía por varias razones, que incluyen las siguientes:

40 1. Hay una amplia variedad de enzimas endonucleasas disponibles que generarán una matriz de extremos, sin embargo, la mayoría de estas no son compatibles entre ellas. Muchas enzimas endonucleasas, tales como EcoRI, genera fragmentos de ADN con extremos 5' protuberantes cohesivos o "colas"; otras (por ejemplo, PstI) generan fragmentos con colas protuberantes 3', mientras que otras más (por ejemplo Ball) escinden en el eje de simetría para producir fragmentos con extremos truncados. Algunos de estos serán compatibles con los extremos formados por escisión con otras enzimas endonucleasas, pero la mayoría de las útiles no. Los extremos que se pueden generar con cada aislamiento de fragmento de ADN se deben tener en consideración cuidadosamente al diseñar una construcción de ADN.

45 2. Los fragmentos de ADN necesarios para ensamblar una construcción de ADN o transgén primero se deben aislar de sus genomas fuente, colocarse en los vectores plasmídicos de clonación, y amplificarse para obtener cantidades útiles. Las etapas se pueden llevar a cabo utilizando cualquiera de los vectores de clonación disponibles en el mercado o alterados individualmente. Cada uno de los vectores plasmídicos disponibles en el mercado es, en su mayor parte, desarrollados independientemente, y por lo tanto contienen diferentes secuencias y sitios de endonucleasa para los fragmentos de ADN de genes o elementos génicos de interés. Los genes por lo tanto se deben hacer a medida individualmente para adaptar cada uno de estos vectores como se necesita para cualquier grupo de experimentos determinado. Se necesitará frecuentemente que los mismos fragmentos de ADN se alteren adicionalmente para experimentos posteriores o clonación con otras combinaciones para nuevas construcciones de ADN o transgenes. Como cada construcción de ADN o transgén se hace a medida para una aplicación particular sin pensar o saber cómo se usará después, frecuentemente se debe "retro-ajustar" para aplicaciones posteriores.

60 3. Además, la secuencia de ADN de un determinado gen o elemento génico varía y puede contener sitios de endonucleasa internos que la hacen incompatible con vectores disponibles actualmente, complicando la modificación de esta manera. Esto es especialmente cierto cuando se ensamblan varios fragmentos de ADN en una construcción de ADN o transgén.

65 Por lo tanto, existe la necesidad de un sistema que permita al usuario ensamblar rápidamente una cantidad de fragmentos de ADN en una molécula, a pesar de la redundancia de sitios de endonucleasas que se encuentra en los

extremos y en los fragmentos de ADN. Dicho sistema, puede también proporcionar un medio simple para alterar rápidamente los extremos de los fragmentos de manera que se añadan otras secuencias de endonucleasas a las mismas. La inclusión de sitios sencillos o pares opuestos de sitios HE aumentaría la probabilidad de que tengan sitios únicos de clonación. Un sistema que permitiera también sustituciones fáciles o la eliminación de uno o más de los fragmentos añadirían un nivel de versatilidad que no está disponible actualmente a los usuarios. Por lo tanto, un sistema "modular" es decir, un sistema que permita insertar o eliminar fragmentos de ADN o "inserciones" en o fuera de regiones de "casete" flanqueadas por sitios raros de endonucleasa en el vector de clonación, sería especialmente útil y bienvenido en el campo de la tecnología de ADN recombinante.

10 El documento WO 00/56901 A2 desvela elementos de expresión lineales y circulares para explorar la función génica, los efectos biológicos de la función génica, antígenos, y función de promotor. El documento WO 00/56901 A2 también proporciona métodos para ensayar la producción o regulación de la expresión de un segmento de ácido nucleico lineal que comprende un promotor y una fase de lectura abierta, comprendiendo el método la obtención de un segmento de ácido nucleico lineal que comprende un promotor y una fase de lectura abierta; colocando el segmento de ácido nucleico lineal en condiciones que dan lugar a la expresión de un polipéptido a partir de la fase de lectura abierta; y ensayando la producción y regulación de la expresión del polipéptido.

20 Los documentos WO 00/26386 A1 y US 6.096.523 A desvelan un sistema de transformación de un vector que incluye un grupo de vectores lanzadera y un vector de acoplamiento de manera que se puedan clonar segmentos de ADN individuales, incluyendo grandes segmentos de ADN codificante en vectores lanzadera individuales en la orientación adecuada y combinarse en un orden y orientación deseados, en un vector de acoplamiento único. El vector de acoplamiento tiene un sitio múltiple de clonación que tiene una pluralidad de sitios de corte por endonucleasas de restricción raros, estando presentes dichos sitios como un locus único en el vector. Cada vector lanzadera comprende al menos dos sitios de corte por endonucleasas raros que se corresponden con uno o más sitios de corte por endonucleasa raros en el vector de acoplamiento en la orientación congruente entre ellos y comprende además al menos otro sitio de restricción entre los sitios de restricción de corte raro.

30 El documento US 5.919.667 A desvela el ensamblaje modular de vectores retrovíricos que contienen repeticiones terminales largas (LTRS) que hacen posible una expresión de alto nivel y modulable por ligando de un producto génico deseado, incluso tras periodos prolongados de quiescencia celular.

35 El documento GB 2 393 441 A desvela un método para producir un vector recombinante multi-génico y un método para producir vectores recombinantes multi-génicos utilizando un vector receptor y al menos dos vectores donantes en los que un sistema de recombinación de ADN permite dos o más rondas de ensamblaje de ADN haciendo posible que se produzca el intercambio de ADN entre dicho receptor y los donantes, permitiendo de esta manera la inserción secuencial de ADN.

40 El documento US 2004/0253732 A1 desvela vectores de clonación para la recombinación homóloga que presentan un casete que comprende un polienlazador que hace posible que se integre un fragmento de ADN genómico, estando dicho polienlazador flanqueado por sitios de restricción raros, y estando dicho casete flanqueado por sitios de restricción muy raros.

Sumario de la invención

45 En general, la presente invención se refiere a un método para ensamblar rápidamente transgenes utilizando vectores plasmídicos de clonación. En particular la presente invención proporciona un método para sintetizar simultáneamente una matriz de transgenes, que comprende las etapas de:

50 a. proporcionar un vector plasmídico de clonación primario que comprende un armazón, comprendiendo el armazón

(i) al menos un primer punto de acoplamiento, estando cada punto de acoplamiento fijado en el armazón y que comprende al menos un sitio de restricción raro de más de 6 nucleótidos para una enzima de restricción rara no variable, y

55 (ii) un único sitio de endonucleasa buscadora en una orientación directa localizado secuencia arriba desde el extremo 5' del primer punto de acoplamiento y un único sitio de endonucleasa buscadora en una orientación inversa localizado secuencia abajo desde el extremo 3' del segundo punto de acoplamiento;

60 b. Escindir el primer punto de acoplamiento con una primera enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción raros de más de 6 nucleótidos del primer punto de acoplamiento;

c. escindir el segundo punto de acoplamiento con una segunda enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción del segundo punto de acoplamiento;

d. proporcionar al menos una inserción de Promotor comprendiendo la secuencia de Promotor de interés, un extremo 5' que es compatible con el extremo 3' del primer punto de acoplamiento;

65 e. proporcionar al menos una inserción de Expresión comprendiendo la secuencia de Expresión de interés, un extremo 5' que es compatible con el extremo 3' de la al menos una inserción de Promotor para formar un sitio de

restricción raro de más de 6 nucleótidos para una tercera enzima de restricción rara no variable;

f. proporcionar al menos una inserción Reguladora comprendiendo la secuencia Reguladora de interés, un extremo 5' que es compatible con el extremo 3' de la al menos una inserción de Expresión para formar un sitio de restricción raro de más de 6 nucleótidos para una cuarta enzima de restricción rara no variable, y un extremo 3' que es compatible con el extremo 5' del segundo punto de acoplamiento que se escindió en la etapa 'c', en el que las primera, segunda, tercera y cuarta enzimas de restricción raras no variables son diferentes;

g. colocar al menos dos tipos diferentes de al menos una de las inserciones de Promotor, de Expresión y Reguladora, al menos una de cada una de las inserciones restantes y el vector plasmídico de clonación escindido en una mezcla de reacción apropiada para producir la unión simultánea, auto-orientación y colocación secuencial de cada una de las inserciones de Promotor, Expresión y Reguladoras entre el primer y segundo punto de acoplamiento del armazón, creando de esta manera una matriz de plásmidos que tienen diferentes combinaciones de inserciones de Promotor, Expresión y Reguladoras en su armazón;

m. escindir el armazón de cada uno de los sitios únicos de endonucleasa buscadora con una única enzima de restricción endonucleasa buscadora;

n. purificar la parte escindida que contiene las inserciones; y

o. insertar la parte escindida en un genoma huésped de interés.

La invención proporciona un método que incorpora múltiples fragmentos de ADN, también conocidos como "inserciones" o "módulos", de manera que cada una de las secuencias de nucleótidos de Promotor, de Expresión, y Reguladora 3', en un vector plasmídico de clonación en una única etapa, mejor que tener que introducir cada inserción de manera secuencial. Dicho método se llama "Ensamblaje Dinámico de Vector" en el presente documento.

En una realización de la presente descripción que comprende la presente invención, se proporciona un método para construir un transgén, comprendiendo el método las etapas de proporcionar un vector plasmídico de clonación con un armazón capaz de aceptar una disposición secuencial de inserciones, proporcionando al menos una primera inserción y una segunda inserción que se van a incluir en el transgén, y transfiriendo tanto la primera inserción como la segunda inserción al armazón en una única reacción.

En otra realización de la presente descripción que comprende la invención, se proporciona un método para producir un transgén, comprendiendo el método las etapas de: proporcionar un vector plasmídico de clonación que comprende un primer y segundo punto de acoplamiento; introducir primeras secuencias de nucleótidos que se van a incluir en el transgén en un primer vector lanzadera; introducir segundas secuencias de nucleótidos que se van a introducir en el transgén en un segundo vector lanzadera; y transferir simultáneamente las primeras secuencias de nucleótidos, y las segundas secuencias de nucleótidos desde los vectores lanzadera al vector plasmídico de clonación entre el primer y el segundo punto de acoplamiento.

La presente descripción que comprende la invención también proporciona un método para producir un transgén, que comprende las etapas de: proporcionar un vector plasmídico de clonación que comprende un primer y segundo puntos de acoplamiento; introducir secuencias de nucleótidos de Promotor que se van a incluir en el transgén en un vector lanzadera de Promotor; introducir secuencias de nucleótido de Expresión que se van a incluir en el transgén en un vector lanzadera de Expresión; introducir secuencias de nucleótidos Reguladoras que se van a incluir en el transgén en un vector lanzadera Regulador; y transferir simultáneamente las secuencias de nucleótidos de Promotor, de Expresión y Reguladoras de los vectores lanzadera de Promotor, de Expresión y Regulador al vector plasmídico de clonación, entre el primer y el segundo puntos de acoplamiento.

En otra realización de la presente descripción que comprende la invención, se proporciona un método para sintetizar simultáneamente una matriz de transgenes, comprendiendo el método las etapas de: proporcionar un vector plasmídico de clonación primario que comprende un primer y un segundo punto de acoplamiento; introducir al menos una secuencia de nucleótidos de Promotor que se va a incluir en el transgén en un vector lanzadera de Promotor correspondiente; introducir al menos una secuencia de nucleótidos de Expresión que se va a incluir en el transgén en un vector lanzadera de Expresión correspondiente; introducir al menos una secuencia de nucleótidos Reguladora que se va a incluir en el transgén en una vector lanzadera Regulador correspondiente; y transferir simultáneamente las secuencias de nucleótidos de Promotor, de Expresión y Reguladora desde los vectores lanzadera de Promotor, de Expresión y Regulador al vector plasmídico de clonación, entre el primer y segundo puntos de acoplamiento, en el que se transfieren al menos dos combinaciones de un módulo Promotor, un módulo de Expresión, y un módulo Regulador en dos moléculas distintas de vector de clonación primario.

En otra realización más de la presente descripción que comprende la invención, se proporciona un método para producir un vector plasmídico de clonación modular para la síntesis de un transgén u otra construcción de ADN complicada, comprendiendo el método las etapas de: proporcionar el vector plasmídico de clonación que comprende un armazón, comprendiendo el armazón un primer y un segundo punto de acoplamiento, estando cada punto de acoplamiento fijado en el armazón y comprendiendo al menos un sitio de endonucleasa rara no variable para una enzima endonucleasa; escindir el primer punto de acoplamiento con una primera enzima endonucleasa que se corresponde con al menos un sitio de restricción raro no variable del primer punto de acoplamiento, dejando el primer punto de acoplamiento escindido con un extremo 3'; escindir el segundo punto de acoplamiento con una

segunda enzima nucleasa que se corresponde con el al menos un sitio de endonucleasa rara no variable del segundo punto de acoplamiento, dejado el segundo punto de acoplamiento escindido con un extremo 5'; proporcionar al menos una primera y una segunda inserción, comprendiendo cada inserción un extremo 5', una secuencia de nucleótido de interés y un extremo 3', en el que el extremo 5' de la primera inserción es compatible con el extremo 3' del primer punto de acoplamiento escindido, el extremo 3' de la segunda inserción es compatible con el extremo 5' del segundo punto de acoplamiento escindido, siendo el extremo 3' de la primera inserción compatible con el extremo 5' de la segunda inserción para formar un tercer sitio de endonucleasa rara no variable para una tercera enzima endonucleasa; y colocar las inserciones y el vector plasmídico de clonación escindido en una mezcla de reacción apropiada para producir la unión simultánea y la auto-orientación de la primera y segunda inserciones entre el primer y el segundo punto de acoplamiento en el armazón, volviendo a formar el primer y segundo punto de acoplamiento, y formando el vector plasmídico de acoplamiento.

En otra realización de la presente descripción que comprende la invención, se proporciona un método para sintetizar un transgén u otra construcción de ADN complicada, comprendiendo el método las etapas de: proporcionar un primer vector plasmídico de clonación que comprende un armazón, comprendiendo el armazón al menos un primer punto de acoplamiento y un segundo punto de acoplamiento, estando cada punto de acoplamiento fijado en el armazón y comprendiendo al menos un sitio de restricción raro para una enzima de restricción rara no variable; escindir el primer punto de acoplamiento con una primera enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción raros del primer punto de acoplamiento, dejando el armazón escindido con un extremo 3'; escindir el segundo punto de acoplamiento con una segunda enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción del segundo punto de acoplamiento, dejando el armazón escindido con un extremo 5'; proporcionar una inserción Promotora en la que hay una secuencia Promotora de interés, con un extremo 5' que es compatible con el extremo 3' del primer punto de acoplamiento, y un extremo 3'; proporcionar una inserción de Expresión que comprende una secuencia de Expresión de interés, con un extremo 5' que es compatible con el extremo 3' de la inserción de Promotor para formar un sitio de restricción raro para una tercera enzima de restricción rara no variable, y un extremo 3'; proporcionar una inserción Reguladora que comprende una secuencia Reguladora de interés, con un extremo 5' que es compatible con el extremo 3' de la inserción de Expresión para formar un sitio de restricción raro para una cuarta enzima de restricción rara no variable, y un extremo 3' que es compatible con el extremo 5' del segundo punto de acoplamiento escindido en la etapa 'c'; y colocar las inserciones de Promotor, de Expresión y Reguladora y el vector plasmídico de clonación escindido en una mezcla de reacción adecuada para producir la unión simultánea, auto-orientación y la colocación secuencial de las inserciones de Promotor, de Expresión y Reguladora entre el primer y el segundo puntos de acoplamiento, volviendo a formar el primer y segundo punto de acoplamiento, y formando un vector plasmídico de clonación modular primario.

En otra realización más de la presente descripción que comprende la invención, se proporciona un método para sintetizar simultáneamente una matriz de transgenes u otras construcciones de ADN complicadas, comprendiendo el método las etapas de: proporcionar al menos un vector plasmídico de clonación primario que comprende un armazón en el cual se pueden insertar inserciones que tienen un extremo 5', una secuencia de nucleótidos de interés y un extremo 3', siendo el armazón operativo para aceptar una disposición secuencial de inserciones de Promotor, de Expresión, y reguladoras y que comprenden al menos un primer y un segundo punto de acoplamiento, estando cada punto de acoplamiento fijado en el armazón y comprendiendo al menos un sitio de restricción para una enzima de restricción rara no variable; escindir el primer punto de acoplamiento con una primera enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción del primer punto de acoplamiento; escindir el segundo punto de acoplamiento con una segunda enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción del segundo punto de acoplamiento; proporcionar al menos una inserción de Promotor en la que se ha insertado una secuencia de nucleótido de Promotor, el extremo 5' de al menos una inserción de Promotor compatible con el extremo 3' del primer punto de acoplamiento que se escindió en la etapa 'b'; proporcionar al menos una inserción de Expresión en la que se ha insertado una secuencia de nucleótidos de Expresión, siendo el extremo 5' de la al menos una inserción de Expresión compatible con el extremo 3' de la al menos una inserción de Promotor para formar un sitio de restricción para una tercera enzima de restricción rara no variable; proporcionar al menos una inserción Reguladora en la que se ha insertado una secuencia de nucleótidos Reguladora, siendo el extremo 5' de la al menos una inserción Reguladora compatible con el extremo 3' de la al menos una inserción de Expresión para formar un sitio de restricción para una cuarta enzima de restricción rara no variable, siendo el extremo 3' de la al menos una inserción Reguladora compatible con el extremo 5' del segundo punto de acoplamiento que se escindió en la etapa 'c'; y a continuación colocar al menos dos tipos diferentes de al menos una de las inserciones de Promotor, de Expresión y Reguladora, al menos una de cada una de las inserciones restantes, y el vector plasmídico de clonación escindido en una mezcla de reacción apropiada para producir la unión simultánea, auto-orientación y colocación secuencial de una de cada inserción de Promotor, de Expresión y Reguladora entre el primer y el segundo puntos de acoplamiento, creando de esta manera una matriz de plásmidos que tienen diferentes combinaciones de inserciones de Promotor, de Expresión y Reguladora en su armazón.

Se apreciará una comprensión adicional más completa de la naturaleza y ventajas de la presente invención con respecto a la descripción de los siguientes dibujos y detallada.

Breve descripción de los dibujos

Los dibujos adjuntos ilustran realizaciones de la presente descripción que comprende la invención y, junto con la descripción general que se ha dado anteriormente, y la descripción detallada que se da posteriormente, sirven para explicar los principios de la invención.

La FIG. 1 es un mapa lineal del concepto de módulo que representa un vector lanzadera P que se puede insertar en una estación de acoplamiento PE3, que se puede insertar en una estación de acoplamiento Primaria.

La FIG. 2 es una ilustración que representa el ensamblaje de un armazón del vector capacitado para las relaciones entre los sitios de restricción con los vectores lanzadera, tales como módulos de Promotor, de Expresión y Reguladores 3', y los puntos de acoplamiento de un vector plasmídico de clonación primario.

La FIG. 3 es una ilustración que representa el armazón del vector ensamblado de la FIG. 2.

Descripción detallada de la invención

Como se utiliza en el presente documento, la expresión "dominio de modificación de la cromatina" (CMD) se refiere a secuencias de nucleótidos que interactúan con varias proteínas asociadas a la tinción principal y/o a la alteración de la estructura de la cromatina. Como se utiliza en el presente documento, el término "clonación" se refiere al proceso para unir una molécula de ADN en un plásmido y transferirla a una célula huésped apropiada para su duplicación durante la propagación del huésped.

Como se utiliza en el presente documento, las expresiones "vector de clonación" y "vector plasmídico de clonación" se utilizan de manera intercambiable para referirse a una molécula de ADN circular que contienen como mínimo un Origen de Replicación, un medio para la selección positiva de las células huésped que albergan el plásmido tal como un gen de resistencia a los antibióticos; y un sitio de clonación múltiple.

Como se utiliza en el presente documento, el término "común" en relación a sitios de endonucleasa se refiere a cualquier sitio de endonucleasa que relativamente se presenta frecuentemente en un genoma.

Como se utiliza en el presente documento, la frase "compatible con" se refiere a un final o extremo, sea 5' o 3', de una cadena de ADN que puede formar enlaces hidrógeno con cualquier otro extremo complementario que se ha escindido con la misma enzima de restricción o se ha creado por algún otro método. Como cualquier ADN que contenga una secuencia de reconocimiento específica para una enzima de restricción se cortará de la misma manera que otro ADN que contenga la misma secuencia, estos extremos escindidos serán complementarios y por lo tanto compatibles. Por lo tanto, los extremos de cualquier molécula de ADN cortada con la misma enzima de restricción "se complementan" entre ellos de la misma manera que las piezas de un puzle "encajan", y se pueden unir entre ellas enzimáticamente. Los extremos compatibles formarán un sitio de restricción para una enzima de restricción particular cuando se combinan entre ellos.

Como se utiliza en el presente documento, la expresión "síntesis *de novo*" se refiere al proceso de síntesis de moléculas de ADN de doble cadena de cualquier longitud uniendo complementariamente protuberancias compatibles de moléculas de ADN de cadena sencilla que representan sub-secuencias de la molécula de ADN total deseada.

Como se utiliza en el presente documento, la expresión "construcción de ADN" se refiere a una molécula de ADN sintetizada por etapas de clonación consecutivas en un vector plasmídico de clonación, y se utiliza comúnmente para dirigir la expresión génica en una célula huésped apropiada tal como células cultivadas *in vitro*, o un ratón transgénico *in vivo*. También puede hacer referencia a un transgén que se utiliza para producir dicho ratón como una construcción de ADN, especialmente durante el periodo de tiempo en el que el transgén se va a diseñar y sintetizar.

Como se utiliza en el presente documento, un "fragmento de ADN" se refiere a una molécula de ADN aislada, que incluye pero no se limita a una secuencia codificante de proteína, gen indicador, promotor, amplificador, intrón, exón, cola poli-A, sitio de clonación múltiple, señal de localización nuclear, o señal de estabilización del ARNm, o cualquier otra molécula de ADN de origen natural o sintética, o cualquier parte de la misma. De manera alternativa, un fragmento de ADN puede ser de origen completamente sintético, producido *in vitro*. Además, un fragmento de ADN puede comprender cualquier combinación de fragmentos aislados de origen natural y/o sintéticos.

Como se utiliza en el presente documento, la expresión "plásmido de acoplamiento" se refiere a un vector plasmídico de clonación especializado que se utiliza en la invención para ensamblar fragmentos de ADN en una construcción de ADN. Como se utiliza en el presente documento, las expresiones "endonucleasa" o "enzima endonucleasa" se refieren a un miembro o miembros de una clasificación de moléculas catalíticas que se unen a un sitio de reconocimiento codificado en una molécula de ADN y escinde la molécula de ADN en una localización precisa en o cerca de la secuencia.

Como se utiliza en el presente documento, las expresiones "sitio de reconocimiento de endonucleasa", "sitio de reconocimiento", "sitio equivalente" o "secuencias equivalentes" se refiere al tramo mínimo de nucleótidos necesario

para que una enzima de restricción se una y escinda una molécula de ADN o gen. Como se utiliza en el presente documento, la expresión “región amplificadora” se refiere a una secuencia de nucleótido que no es necesaria para la expresión de un gen diana, pero que aumentará el nivel de expresión génica en condiciones apropiadas.

5 Como se utiliza en el presente documento, la expresión “gen selector del gen de expresión del huésped” (GEH-S) se refiere a un elemento génico que puede conferir a un organismo huésped un rasgo que se puede seleccionar, seguir, o detectar por sensores ópticos, amplificación por PCR, ensayos bioquímicos, o por ensayos de supervivencia de células/organismos (resistencia o toxicidad de las células u organismos cuando se tratan con un antibiótico o agente químico apropiado).

10 Como se utiliza en el presente documento, la expresión “gen promotor” o “promotor” se refiere a una secuencia de nucleótido necesaria para la expresión de un gen, o una parte del promotor de longitud completa. Como se utiliza en el presente documento, los términos “inserción” y “módulo” son esencialmente intercambiables, con la única pequeña distinción de que una “inserción” se inserta en el vector, y una vez que se inserta se llama más comúnmente “módulo”. Un módulo se puede retirar entonces del vector. También, se utiliza comúnmente el término inserción para un módulo aislado que se utiliza como una inserción en un vector modular receptor.

15 Como se utiliza en el presente documento, el término “intrón” se refiere a las secuencias de nucleótido de una región no codificante de proteína en un gen de una célula de mamífero que se encuentra entre dos regiones codificantes de proteína o exones.

20 Como se utiliza en el presente documento, la expresión “señal de localización” (LOC) se refiere a secuencias de nucleótido que codifican una señal para el enrutamiento de una proteína de interés. Como se utiliza en el presente documento, la expresión “sitio de clonación múltiple” (MCS) se refiere a secuencias de nucleótidos que comprenden al menos un único sitio de endonucleasa, y, más normalmente, un agrupamiento de sitios únicos de endonucleasa, con el fin de clonar fragmentos de ADN en un vector plasmídico de clonación.

25 Como se utiliza en el presente documento, la expresión “elemento de estabilización del ARNm” se refiere a una secuencia de ADN que es reconocida por proteínas de unión concebidas para proteger algunos ARNm de la degradación.

30 Como se utiliza en el presente documento, la expresión “Origen de Replicación” (ORI) se refiere a secuencia de nucleótidos que dirigen la replicación o la duplicación de un plásmido en una célula huésped.

35 Como se utiliza en el presente documento, la frase “tecnología de PCR de clonación de la protuberancia terminadora” se refiere al proceso de amplificación de módulos génicos utilizando la reacción en cadena de polimerasa en conjunción con cebadores ADN de cadena sencilla con protuberancias nucleotídicas 5' que pueden servir como sitios de unión con protuberancias de ADN complementarias.

40 Como se utiliza en el presente documento, la expresión “cola poli-A” se refiere a una secuencia de nucleótidos adenina (A) que se encuentran comúnmente al final de las moléculas de ARN mensajero (ARNm). Una señal de cola poli-A se incorpora en los extremos 3' de las construcciones de ADN o transgenes para facilitar la expresión del gen de interés.

45 Como se utiliza en el presente documento, la expresión “sitio cebador” se refiere a secuencias de nucleótido que funcionan como matrices de ADN en los que se pueden hibridar oligonucleótidos ADN de cadena sencilla con el fin de iniciar la secuenciación de ADN, la amplificación por PC, y/o la transcripción de ARN. Como se utiliza en el presente documento, el término “pUC19” se refiere a un vector plasmídico de clonación bien conocido por los expertos en la técnica, y se puede encontrar en la base de datos Genbank con el n° de registro L09137.

50 Como se utiliza en el presente documento, la expresión “secuencias aleatorias de nucleótidos” se refiere a cualquier combinación de secuencias de nucleótidos que no duplican las secuencias que codifican otros elementos especificados como componentes de la misma molécula. El número de nucleótidos necesarios en las secuencias aleatorias depende de las necesidades de las enzimas endonucleasas que flanquean las secuencias aleatorias. La mayoría de las endonucleasas necesitan un mínimo de 2-4 secuencias aleatorias adicionales para estabilizar la unión de ADN. Se prefiere que el número de secuencias aleatorias sea un múltiplo de 3, que se corresponde con el número de nucleótidos que forman un codón. El número mínimo preferido de secuencias aleatorias por lo tanto es 6, sin embargo, se pueden utilizar menos o más nucleótidos.

55 Como se utiliza en el presente documento, el término “raro” en relación con sitios de endonucleasa se refiere a un sitio de endonucleasa que se presenta relativamente infrecuentemente en un genoma. Como se utiliza en el presente documento, la expresión “brazo de recombinación” se refiere a secuencias de nucleótidos que facilitan la recombinación homóloga entre el ADN transgénico y el ADN genómico. La recombinación satisfactoria necesita la presencia de un brazo de recombinación izquierdo (LRA) y un brazo de recombinación derecho (RRA) que flanquean una región de ADN transgénico que se va a incorporar en un genoma huésped por medio de recombinación homóloga.

60 Como se utiliza en el presente documento, el término “raro” en relación con sitios de endonucleasa se refiere a un sitio de endonucleasa que se presenta relativamente infrecuentemente en un genoma. Como se utiliza en el presente documento, la expresión “brazo de recombinación” se refiere a secuencias de nucleótidos que facilitan la recombinación homóloga entre el ADN transgénico y el ADN genómico. La recombinación satisfactoria necesita la presencia de un brazo de recombinación izquierdo (LRA) y un brazo de recombinación derecho (RRA) que flanquean una región de ADN transgénico que se va a incorporar en un genoma huésped por medio de recombinación homóloga.

65 Como se utiliza en el presente documento, el término “raro” en relación con sitios de endonucleasa se refiere a un sitio de endonucleasa que se presenta relativamente infrecuentemente en un genoma. Como se utiliza en el presente documento, la expresión “brazo de recombinación” se refiere a secuencias de nucleótidos que facilitan la recombinación homóloga entre el ADN transgénico y el ADN genómico. La recombinación satisfactoria necesita la presencia de un brazo de recombinación izquierdo (LRA) y un brazo de recombinación derecho (RRA) que flanquean una región de ADN transgénico que se va a incorporar en un genoma huésped por medio de recombinación homóloga.

Como se utiliza en el presente documento, el término "recombinación" se refiere al proceso para utilizar enzimas recombinasas aleatorias o selectivas del sitio en conjunto con secuencias de ADN sobre las que pueden actuar enzimas recombinasas para trasladar una parte del material génico de una molécula de ADN a una molécula de ADN diferente.

5 Como se utiliza en el presente documento, la expresión "gen indicador" se refiere a secuencias de nucleótido que codifican una proteína útil para controlar la actividad de un promotor particular de interés. Como se utiliza en el presente documento, la expresión "vector lanzadera" se refiere a un vector plasmídico de clonación especializado que se utiliza en la invención para producir una molécula intermedia que modificará los extremos de un fragmento de ADN.

Como se utiliza en el presente documento, la expresión "secuencia marcadora" (TAG) se refiere a secuencias de nucleótidos que codifican una región proteica única que permite que sea detectada, o en algunos casos, distinguible de cualquier equivalente endógena.

15 Como se utiliza en el presente documento, la expresión "región no traducida" (UTR) se refiere a secuencias de nucleótidos que engloban la región no codificante de proteínas de una molécula de ARNm. Estas regiones no traducidas pueden residir en el extremo 5' (5' UTR) o el extremo 3' (3' UTR) de una molécula de ARNm.

20 La presente descripción proporciona un método para coger un transgén recién producido que contiene los módulos y retirar selectivamente uno o más de los módulos y volver a colocarlos con una inserción diferente. Este proceso se llama en el presente documento "segundo paso" y "enhebrado múltiple". La invención proporciona además un método para crear una matriz de diferentes transgenes, cada uno con una inserción de Promotor, de Expresión y Reguladora, incorporando múltiples inserciones de Promotor, de Expresión y Reguladoras en un vector plasmídico de clonación en una única etapa. La presente descripción también proporciona un método que comprende las etapas de proporcionar vectores plasmídicos de clonación que tienen inserciones de Promotor, Expresión y Reguladoras recién introducidas combinadas juntas, retirando la combinación completa como un armazón de vector, e insertando un número múltiple de armazones de vectores en un único vector plasmídico de clonación.

30 La presente descripción que comprende la presente invención también proporciona un método para crear un vector plasmídico de clonación modular para la síntesis de un transgén u otra construcción de ADN complicada proporcionando un armazón que tiene puntos de acoplamiento en el mismo. Cada punto de acoplamiento representa un área en la que hay preferentemente al menos un sitio fijo de endonucleasa raro no variable, y más preferentemente agrupamientos de dos sitios fijos de endonucleasa raros no variables, y más preferentemente agrupamientos de tres sitios fijos de endonucleasa raros no variables. Un sitio de restricción particular de cada punto de acoplamiento se escinde por su enzima endonucleasa equivalente. Esto creará un extremo 5' o 3' deseado que es compatible con el extremo 5' o 3' complementario de una de las inserciones pre-construidas que contienen una secuencia de nucleótidos de elección, tal como una secuencia de nucleótidos de Promotor, de Expresión o Reguladora. Al menos dos inserciones, que tienen cada una extremos 5' y 3' que son compatibles con el punto de acoplamiento escindido de interés, se pueden añadir junto con el vector plasmídico de clonación escindido a una mezcla de reacción apropiada, y, asumiendo el medio termodinámico apropiado, las inserciones pueden integrarse, es decir, en una única etapa, en el vector plasmídico de clonación. Durante esta adición singular y reacción de unión, los puntos de acoplamiento se vuelven a formar y el vector plasmídico de clonación se convierte en modular, en el que los puntos de acoplamiento y la conexión entre los dos módulos se pueden volver a escindir con las enzimas de restricción adecuadas. Entonces el módulo se puede retirar más tarde, y se puede colocar un nuevo módulo en su lugar.

50 Una realización de la presente descripción que comprende la presente invención se refiere a un método para construir un transgén, que comprende las etapas de proporcionar un vector plasmídico de clonación con un armazón capaz de aceptar una disposición secuencial de inserciones, proporcionando al menos una primera inserción y una segunda inserción que se van a incluir en el transgén, y transfiriendo tanto la primera inserción como la segunda inserción al armazón en una reacción única. Más preferentemente, las inserciones consisten en tres inserciones, específicamente al menos un módulo de Promotor, de Expresión y Regulador.

55 Otra realización de la presente descripción que comprende la invención es un método para producir un transgén que comprende las etapas de proporcionar un vector plasmídico de clonación que comprende un primer y un segundo punto de acoplamiento, introducir secuencias de nucleótido de Promotor que se van a incluir en el transgén en un vector lanzadera de promotor, introducir secuencias de nucleótido de Expresión que se van a incluir en el transgén en un vector lanzadera de Expresión, introducir secuencias de nucleótido Reguladoras que se van a incluir en el transgén en un vector lanzadera Regulador, transferir simultáneamente las secuencias de nucleótido de Promotor, de Expresión y Reguladoras de los vectores lanzadera de Promotor, de Expresión y Reguladores al vector plasmídico de clonación, entre el primer y el segundo punto de acoplamiento.

65 Se prefiere que los extremos 5' y 3' de cada uno de los puntos de acoplamiento y cada una de las inserciones sean compatibles con un extremo correspondiente de otro punto de acoplamiento o inserción. Por ejemplo, si un primer punto de acoplamiento contiene un sitio de restricción para una enzima de restricción rara no variable tal como SgrAI y ese punto de acoplamiento se escinde a continuación, entonces una primera inserción que se pretenda insertar en

el extremo 3' del primer punto de acoplamiento escindido contendrá un extremo 5' compatible para crear un sitio de restricción para SgrAI cuando la inserción se combine con el primer punto de acoplamiento. Un segundo punto de acoplamiento en el plásmido puede, por ejemplo, tener un sitio de restricción para una enzima de restricción no variable tal como Swal. Cualquier segunda inserción que tenga en su extremo 3' una secuencia de nucleótido compatible para combinarse con el extremo 5' escindido del segundo punto de acoplamiento escindido para crear un sitio de restricción par Swal. Además, el extremo 3' de la primera inserción y el extremo 5' de la segunda inserción, con el fin de insertarse simultáneamente en el vector plasmídico de clonación modular y también para retirarse después en el mismo punto, tienen que contener extremos compatibles para crear un tercer sitio de restricción para una enzima de restricción rara no variable tal como PacI o Sall.

Los elementos secuenciales que codifican la estructura modular de la presente invención puede comprender específicamente: tres sitios de restricción comunes únicos y no variables, un sitio de un oligonucleótido cebador 5', un único sitio de HE en una orientación directa, un par de sitios de restricción comunes, únicos y no variables que flanquean secuencias de nucleótido aleatorias, un agrupamiento fijo de sitios de restricción raros no variables que definen una parte 5' de un módulo promotor, secuencias de nucleótido aleatorias, un agrupamiento fijo de sitios de restricción raros no variables que definen una unión compartida entre una posición 3' relativa a un módulo de Promotor/intrón y una posición 5' relativa a un módulo de Expresión, secuencias de nucleótido aleatorias, un agrupamiento fijo de sitios de restricción raros no variables de define una posición 3' relacionada con un módulo Regulador 3', secuencias aleatorias de nucleótidos, un agrupamiento fijo de sitios de restricción raros no variables que definen una posición 3' relativa a un módulo Regulador 3', un par de sitios de restricción comunes, únicos y no variables que flanquean secuencias de nucleótidos aleatorias, un único sitio HE en orientación inversa que es el mismo sitio HE que el que se sitúa 3' del sitio de oligonucleótido cebador 5', un sitio de oligonucleótido cebador 3' en orientación inversa, y cuatro sitios de restricción comunes, únicos y no variables que definen un sitio de inserción 3'.

Otros elementos secuenciales que codifican la estructura modular de la presente invención pueden comprender específicamente: dos sitios de restricción comunes únicos y no variables que definen un sitio de inserción 5', un sitio de oligonucleótido cebador, un par de sitios HE únicos en orientación opuesta flanqueando secuencias de nucleótidos aleatorias, un sitio de restricción único, común y no variable que permite la clonación de un módulo de vector lanzadera secuencia abajo del par de sitios HE únicos, un agrupamiento fijo de sitios de restricción raros no variables, secuencias de nucleótidos aleatorias, un agrupamiento fijo de sitios de restricción raros no variables, un sitio HE único en una orientación directa, un par de sitios de restricción comunes, únicos y no variables flanqueando secuencias de nucleótidos aleatorias, un sitio de oligonucleótido cebador, un par de sitios BstXI únicos en orientación opuesta (en el que la región de nucleótidos variable en el sitio de reconocimiento de BstXI se define por nucleótidos idénticos a las colas no complementarias generados por la ordenación de dos sitios de reconocimiento de HE idénticos dispuestos en orientación complementaria inversa), un par de sitios HE únicos en orientaciones opuestas que flanquean secuencias de nucleótidos aleatorias; un sitio de oligonucleótido cebador en orientación inversa, un par de sitios de restricción comunes, únicos y no variables flanqueando secuencias de nucleótidos aleatorias, un sitio HE único en orientación inversa, siendo el sitio HE el mismo sitio HE de orientación directa, un agrupamiento fijo de sitios de restricción raros no variables, secuencias de nucleótidos aleatorias, un agrupamiento fijo de sitios de restricción raros no variables, un sitio de restricción común, único y no variable, un par de sitios HE únicos en orientación opuesta flanqueando secuencias de nucleótidos aleatorias, un oligonucleótido cebador en orientación inversa, y tres sitios de restricción comunes únicos y no variables.

La presente descripción que comprende la presente invención proporciona un grupo de métodos para el ensamblaje de varios fragmentos de ADN en una construcción de ADN *de novo* o transgén utilizando vectores de clonación optimizados para reducir la cantidad de manipulación que se necesita frecuentemente.

El vector primario, al que se hace referencia en el presente documento como Plásmido de Acoplamiento, contiene un sitio de clonación múltiple (MCS) con preferentemente 3 grupos de sitios de endonucleasa raros dispuestos con un patrón lineal. Esta disposición define una arquitectura modular que permite al usuario ensamblar múltiples inserciones en una única construcción transgénica sin alterar la integridad de los elementos de ADN incorporados en las etapas de clonación previas en el plásmido de acoplamiento.

Dos sitios de reconocimiento para al menos tres HE se sitúan en orientación opuesta para flanquear tres regiones modulares con el fin de crear un sitio receptor de casete génico que no se pueda auto-hibridar. Debido a que los sitios HE son asimétricos y no palindrómicos, es posible generar colas cohesivas que protruyen no complementarias 3' para situar dos sitios de reconocimiento de HE en orientación opuesta. Por lo tanto la HE I-SceI corta su sitio de reconocimiento equivalente como se indica por una "f":

5'...TAGGGATAA/CAGGGTAAT...3',
3'...ATCCC/TATTGTCCCATTA...5'.

El emplazamiento inverso de un segundo sitio en un MCS generaría dos colas cohesivas que protruyen no complementarias:

5'...TAGGGATAACCCTA...3'
3'...ATCCCAATAGGGAT...5'.

5 Esto es particularmente útil cuando es necesario subclonar grandes transgenes en un vector. Debido al tamaño de la inserción, es más favorable termodinámicamente que un vector se auto-hibride a que acepte una inserción grande. La presencia de colas no complementarias generadas por este emplazamiento de sitios de restricción proporciona fuerzas químicas que actúan contra la inclinación termodinámica por la auto-unión.

10 La naturaleza asimétrica de la mayoría de las colas que protruyen por HE también crean una poderosa herramienta de clonación cuando se utiliza en combinación con el sitio de enzima endonucleasa BstX I (5' CCANNNNN/NTGG 3', donde 'N' puede ser cualquier nucleótido). El dominio de secuencia neutra de BstX I se puede utilizar para generar extremos cohesivos compatibles para dos colas que protruyen por HE orientadas inversamente, a la vez que se evita la auto-hibridación.

15 BstX I (I-Sce I Dir.) I-Sce Directo I-Sce Inverso BstX I (I-Sce I Inv.)

5'-CCAGATAA
CAGGGTAAT//ATTACCCTGTTAT
GTGG-3'
20 3'-GGTC
TATTGTCCCATTA//TAATGGGAC A
ATACACC-5'

25 Los sitios de endonucleasa que se utilizan en la invención se escogieron de acuerdo con la jerarquía de presentación. Con el fin de determinar la frecuencia de presentación de sitios de endonucleasa, se analizó la información de secuencia de ADN correspondiente con diecinueve genes diferentes utilizando el software Vector NTI. Esta búsqueda cubría un total de 110.530 nucleótidos de la secuencia de ADN. Los resultados de estos análisis se calcularon de acuerdo con el número de veces que se presentaba un sitio de endonucleasa en los 110.530 nucleótidos analizados. Se les asignó entonces a los sitios de endonucleasa una designación jerárquica de acuerdo con cuatro clasificación, en las que los sitios "comunes" se presentan más de 25 veces en 110.530 nucleótidos, los "sitios de 6 pb de baja frecuencia" se presentan entre 6-24 veces por 110.530 nucleótidos, y los sitios "raros" se presentan entre 0-5 veces por 110.530 nucleótidos. Se enumera una lista parcial de enzimas "adecuadas" de acuerdo con su clasificación de presentación:

35 Enzimas endonucleasas comunes:

Ase I, BamH I, Bgl II, Bln I, BstX I, EcoR I, Hinc II, Hind III, Nco I, Pst I, Sac I, Sac II, Sph I, Stu I, Xba I

40 Enzimas endonucleasas que tienen sitios de reconocimiento de 6 pb, pero tienen una baja frecuencia de presentación:

Aar I, Aat II, Afl II, Age I, ApaL I, Avr II, BseA I, BspD I, BspE I, BstB I, Cla I, Eag I, Eco0109 I, EcoR V, Hpa I, Kpn I, Mfe I, Nar I, Nde I, NgoM IV, Nhe I, Nsi I, Pml I, SexA I, Sma I, Spe I, Xho I

45 Enzimas endonucleasas raras:

Acl I, Asc I, AsiS I, BsiW I, Fse I, Mlu I, Not I, Nru I, Pac I, Pme I, Pvu I, Rsr II, Sal I, Sbf I, Sfi I, SgrA I, SnaB I, Swa I, PI-Sce I, I-Sce I, I-Ceu I, PI-Psp I, I-Ppo I, I-Tli I. También se pueden utilizar otras endonucleasas no incluidas en estos listados, que mantienen la misma funcionalidad y el espíritu e intención de la invención.

50 Los vectores secundarios de la presente descripción que comprende la invención, conocidos en el presente documento como vectores lanzadera, contienen múltiples sitios de clonación con sitios de endonucleasa comunes flanqueados por sitios de endonucleasa raros. Los vectores lanzadera se diseñan para la clonación de fragmentos de ADN en los sitios de endonucleasa comunes entre los sitios raros. Los fragmentos clonados se pueden liberar posteriormente por escisión en el sitio o sitios de endonucleasa raros, e incorporarse en el plásmido de acoplamiento utilizando el mismo sitio o sitios de endonucleasa raros que se encuentran en los vectores lanzadera.

60 Por lo tanto, a diferencia de los vectores de clonación convencionales, el diseño de MCS permite "casetes" o módulos de fragmentos de ADN que se van a insertar en las regiones modulares del plásmido de acoplamiento. Al igual, cada uno se puede retirar fácilmente utilizando las mismas enzimas endonucleasas raras, y reemplazados con cualquier otro fragmento de ADN de interés. Esta característica permite al usuario cambiar la dirección de un proyecto experimental rápida y fácilmente sin tener que reconstruir la construcción completa de ADN. Por lo tanto, los vectores plasmídicos de clonación de la presente invención permiten al usuario clonar un fragmento de ADN en un vector intermediario utilizando sitios de endonucleasa comunes, creando un módulo receptor de un casete, y luego transferir es fragmento al sitio modular deseado en la construcción final por medio de sitios de endonucleasa raros. Además, esto permite futuras alteraciones en la molécula reemplazando módulos individuales en el plásmido

de acoplamiento con otros módulos de casete.

Los componentes individuales de un transgén (la región el amplificador de Promotor P, de la proteína expresada E, y/o la región reguladora 3' 3) se pueden ensamblar como módulos transferidos desde los vectores lanzadera en una Estación de acoplamiento PE3. Si se necesitan complejidades de mayor orden, los transgenes ensamblados, o se pueden transferir entonces otras secuencias de nucleótidos en un plásmido de acoplamiento primario. Cada uno de los cinco tipos de vectores plasmídicos de clonación se explicará con más detalle para proporcionar el entendimiento de los componentes incorporados en cada uno, comenzando por la estación de acoplamiento PE3 más complejo en el plásmido y el plásmido de acoplamiento primario.

El plásmido de acoplamiento PE3 comprende un armazón pUC19 con las siguientes modificaciones, en las que las secuencias se numeran de acuerdo con el archivo de secuencia de Genbank del pUC19, nº de Registro LO9137:

1. Solo se utilizaron las secuencias de 806 a 2617 (Afl3-Aat2) en el Plásmido de Acoplamiento,
2. El sitio BspH1 en 1729 de pUC19 está mutado de TCATGA a GCATGA,
3. El sitio Acl1 en 1493 en pUC19 está mutado de AACGTT a AACGCT,
4. El sitio Acl1 en 1120 en pUC19 está mutado de AACGTT a CACGCT,
5. El sitio Ahd1 en pUC19 está mutado de GAC-NNNNNGTC a CACNNNNNGTC,
6. Las secuencias que codifican BspH1/I-Ppo 1/BspH1 están insertadas en el único sitio restante BspH1 de pUC19 seguido por la etapa de mutación 2 de la lista anterior.

El sitio de clonación múltiple (MCS) del Plásmido de Acoplamiento PE3 comprende los siguientes elementos secuenciales en el orden listado:

1. Tres sitios de endonucleasa comunes, únicos y no variables que definen un sitio de inserción 5' para el vector pUC19 mutado descrito anteriormente (mostrado, pero no limitado a, Aat II, Bln I, y Eco0109 I);
2. Un sitio de cebador T7;
3. Un único sitio HE (por ejemplo, I-SceI (orientación directa));
4. Un par de sitios de endonucleasa comunes, únicos y no variables que flanquean secuencias de nucleótidos aleatorias que pueden servir como un módulo receptor de un dominio de modificación de cromatina (RNAS-CMD-1) (por ejemplo, Kpn y Avr II);
5. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 5' del módulo promotor (por ejemplo, AsiS I, Pac I, y Sbf I);
6. Secuencias de nucleótidos aleatorias que pueden servir como un módulo receptor de Promotor/intrón (RNAS-P);
7. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la unión compartida entre la parte 3' del módulo Promotor/intrón y la parte 5' del módulo de Expresión (por ejemplo, SgrA I, Ascl, y MluI);
8. Secuencias de nucleótidos aleatorios que pueden servir como un módulo receptor de expresión (RNAS-E);
9. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la unión de la parte 3' del módulo de Expresión y la parte 5' del módulo Regulador 3' (por ejemplo, SnaBI, Not I, y Sal I);
10. Secuencia de nucleótidos aleatoria que puede servir como una molécula receptora del dominio regulador 3' (RNAS-3);
11. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 3' del módulo regulador 3' (por ejemplo, Swa I, Rsr II, y BsiW I);
12. Un par de sitios de endonucleasa comunes, únicos y no variables que flanquean una secuencia de nucleótidos de ADN que sirven como un módulo receptor del dominio de modificación de cromatina (RNAS-CMD-2) (por ejemplo, Xho I y Nhe I); Un único sitio HE en orientación inversa que es idéntico al del punto 3 anterior;
13. Un sitio de cebador en orientación inversa; y
14. Cuatro sitios de endonucleasa comunes únicos y no variables que definen un sitio de inserción 3' para el vector pUC19 mutado descrito anteriormente (por ejemplo, BspEI, Pme I, Sap I, y PspH I).

El plásmido de acoplamiento primario se puede utilizar para ensamblar dos transgenes completos que se construyen primero en Plásmidos de Acoplamiento PE3, o dos brazos de homología necesarios para construir un transgén que se dirige a un gen, o para introducir dos tipos de elementos de selección negativa o positiva. El sitio de clonación múltiple (MCS) en el Plásmido de Acoplamiento Primario comprende los siguientes elementos secuenciales en el orden enumerado:

1. Dos sitios de endonucleasa comunes únicos y no variables que definen un sitio de inserción 5' para el vector pUC19 descrito anteriormente (por ejemplo, Aat II y Bln I);
2. Un sitio de cebador M13 Inv.;
3. Un par de sitios de endonucleasa únicos que flanquean una secuencia de nucleótidos de ADN que pueden servir como un módulo receptor de un selector génico huésped de expresión genómico (RNAS-GEH-S1);
4. Un sitio de endonucleasa común, único y no variable que permite la clonación de un módulo del vector lanzadera secuencia abajo del par de HE (por ejemplo, Eco0109I);
5. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen una parte 5' de un módulo de Brazo de Recombinación Izquierdo (por ejemplo, AsiS I, Pac I y Sbf I);

6. Secuencias de nucleótidos aleatorios que funcionan como un módulo receptor de Brazo de Recombinación Izquierdo;
7. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 3' del módulo receptor del Brazo de Recombinación Izquierdo (por ejemplo, SgrA I, MluI, y AscI);
8. Un sitio HE único (por ejemplo, I-Ceu I (orientación directa));
9. Un par de sitios de endonucleasa comunes, únicos y no variables que flanquean una secuencia de nucleótido de ADN que puede funcionar como un módulo receptor del dominio de modificación de cromatina (RNAS-CMD-1) (por ejemplo, Kpn I y Avr II);
10. Un sitio de cebador T7;
11. Un par de sitios únicos BstX I en orientación opuesta (en el que la región variable de nucleótidos del sitio BstX I se define por nucleótidos idénticos a las colas no complementarias generadas por la ordenación de dos sitios de reconocimiento HE idénticos dispuestos en orientación complementaria inversa; por ejemplo, PI-Scel (orientación directa) y PI-Scel (orientación inversa) que flanquean una secuencia de nucleótidos aleatorias de ADN que pueden funcionar como un módulo receptor transgénico complejo (RNAS-PE3-1);
12. Un par de sitios de endonucleasa únicos que flanquean una secuencia de nucleótidos aleatorios de ADN que puede funcionar como un módulo receptor transgénico complejo (RNAS-PE3-2);
13. Un sitio cebador T3 en orientación inversa;
14. Un par de sitios de endonucleasa comunes, únicos y no variables que flanquean una secuencia de nucleótidos aleatorios de ADN que pueden funcionar como un módulo receptor de un dominio de modificación de cromatina (RNAS-CMD-2) (por ejemplo, xho I y Nhe I);
15. Un sitio de HE único en orientación inversa que es idéntico al del punto 8 anterior;
16. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 5' de un módulo de Brazo de Recombinación Derecho (por ejemplo, SnaB I, Sal I, y Not I);
17. Secuencias de nucleótidos aleatorios que pueden funcionar como un módulo receptor del Brazo de Recombinación Derecho (RNAS-RRA);
18. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 3' del módulo receptor del Brazo de Recombinación Derecho (por ejemplo, Rsr II, Swa I, y BsiW I);
19. Un sitio de endonucleasa común, único y no variable que permite la clonación de un módulo del vector lanzadera (por ejemplo, BspE I);
20. Un par de sitios de endonucleasa únicos que flanquean una secuencia de nucleótidos aleatoria de ADN que funciona como un módulo receptor del selector génico huésped de expresión genómica (RNAS-GEH-S2);
21. Un sitio de cebador directo M13 situado en orientación inversa; y
22. Tres sitios de endonucleasa comunes, únicos y no variables que definen un sitio de inserción 3' para el vector pUC19 mutado que se describe anteriormente (por ejemplo, Pme I, Sap I, y BspH I).

Se conocen tres vectores plasmídicos de la invención como Vectores Lanzadera. Los Vectores Lanzadera tales como los Plásmidos PE3 y de Acoplamiento Primarios, también se construyen a partir de un armazón pUC19. Al igual que los Plásmidos PE3 y de Acoplamiento Primarios, cada Vector Lanzadera tiene las mismas modificaciones del armazón de pUC19 que se enumeran en los puntos 1 a 6 anteriores. Los vectores lanzadera individuales (SV) se identifican como vector lanzadera de Promotor/intrón (P), vector lanzadera de expresión (E), y vector lanzadera Regulador 3' (3); que por lo tanto se exponen como SVP, SVE, y SV3, respectivamente. Cada uno se describe completamente posteriormente.

Vector lanzadera P (SVP):

El SVP es un vector plasmídico de clonación que se puede utilizar para preparar secuencias de promotor e intrón para su ensamblaje en una construcción transgénica. Un ejemplo de un plásmido SVP puede comprender los siguientes elementos secuenciales en el MCS, en el orden enumerado:

1. Dos sitios de endonucleasa comunes, únicos y no variables que definen un sitio de inserción 5' para el vector pUC19 mutado descrito anteriormente (por ejemplo, AatII y BlnI);
2. Un sitio de cebador T7;
3. Un sitio de endonucleasa común único y no variable que permite una clonación eficaz de un vector lanzadera secuencia abajo del sitio de cebador T7 (por ejemplo, Eco0109I);
4. Un agrupamiento fijo de sitios de endonucleasas raros no variables que definen la parte 5' del módulo de promotor (por ejemplo, AsiSI, PacI, y Sbf I). Estos sitios de endonucleasa raros no variables proporcionan el punto de acoplamiento representado por la estrella en el extremo 5' del vector de promotor de la Figura 2;
5. Un MCS variable que comprende cualquier agrupamiento de sitios de endonucleasa raros o comunes que son únicas para el vector lanzadera;
6. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 3' del módulo promotor (por ejemplo, SgrA I, Asc I, y MluI). Estos sitios de endonucleasa raros no variables proporcionan el punto de acoplamiento representado por el círculo en el extremo 3' del vector de Promotor de la Figura 2;
7. Un sitio de endonucleasa común, único y no variable que permite la clonación eficaz de un módulo de vector lanzadera secuencia arriba del sitio de cebador T3 (por ejemplo, BspEI);
8. Un sitio de cebador T3 de orientación inversa; y
9. Dos sitios de endonucleasa comunes, únicos y no variables que definen un sitio de inserción 3' para el vector

pUC19 mutado descrito anteriormente (por ejemplo, Pmel y SapI).

Vector lanzadera E (SVE):

- 5 Este es un vector plasmídico de clonación que se puede utilizar para preparar secuencias que se van a expresar por un transgén por ensamblaje en una construcción transgénica. Un ejemplo de plásmido SVE comprende los siguientes elementos secuenciales en el MCS, en el orden que se enumera:
- 10 1. Dos sitios de endonucleasa comunes, únicos y no variables que definen un sitio de inserción 5' para el vector pUC19 descrito anteriormente (por ejemplo, AatII y BlnI);
 2. Un sitio de cebador T7;
 3. Un sitio de endonucleasa común, único y no variable que permite la clonación eficaz de un módulo del vector lanzadera secuencia abajo del sitio de cebador T7 (por ejemplo, Eco0109I);
 - 15 4. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 5' del módulo de expresión (por ejemplo, SgrA I, Ascl, y MluI). Estos sitios de endonucleasa raros no variables proporcionan el punto de acoplamiento representado por el círculo en el extremo 5' del vector de Expresión de la Figura 2;
 5. Un MCS variable que consiste en cualquier agrupamiento de sitios de endonucleasa raros o comunes que son únicos para el vector lanzadera;
 - 20 6. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 3' del módulo de expresión (por ejemplo, SnaBI, NotI, y Sall). Estos sitios de endonucleasa raros no variables proporcionan el punto de acoplamiento representado por el triángulo en el extremo 3' del vector de expresión de la Figura 2;
 7. Un sitio de endonucleasa común, único y no variable que permite la clonación eficaz de un módulo de vector lanzadera secuencia arriba del sitio de cebador T3 (por ejemplo, BspEI);
 8. Un sitio de cebador T3 con orientación inversa; y
 - 25 9. Dos sitios de restricción comunes, únicos y no variables que definen el sitio de inserción para el vector pUC19 mutado descrito anteriormente (por ejemplo, Pmel y SapI).

Vector lanzadera 3 (SV3):

- 30 Este es un vector plasmídico de clonación que se puede utilizar para preparar secuencias reguladoras 3' para su ensamblaje en una construcción transgénica. Un ejemplo de un plásmido SV3 puede comprender los siguientes elementos en el MCS, en el orden que se enumera:
- 35 1. Dos sitios de endonucleasa comunes, únicos y no variables que definen un sitio de inserción 5' para el vector pUC19 mutado descrito anteriormente (por ejemplo, AatII y BlnI);
 2. Un sitio de cebador T7;
 3. Un sitio de endonucleasa común, único y no variable que permite la clonación eficaz de un módulo de vector lanzadera secuencia abajo del cebador T7 (por ejemplo, Eco0109I);
 - 40 4. Un agrupamiento fijo de sitios de endonucleasa raros no variables que define la parte 5' del módulo regulador 3' (por ejemplo, SnaBI, NotI, y Sall). Estos sitios de endonucleasa raros no variables proporcionan el punto de acoplamiento representado por el triángulo en el extremo 5' del vector Regulador de la Figura 2;
 5. Un MCS variable que consiste en cualquier agrupamiento de sitios de endonucleasa raros o comunes que son únicos para el vector lanzadera;
 - 45 6. Un agrupamiento fijo de sitios de endonucleasa raros no variables que definen la parte 3' del módulo regulador 3' (por ejemplo, SwaI, RsrII, y BsiWI). Estos sitios de endonucleasa raros no variables proporcionan el punto de acoplamiento representado por el cuadrado del extremo 3' del vector Regulador de la Figura 2;
 7. Un sitio de endonucleasa no raro, único y no variable que permite la clonación eficaz de un módulo de vector lanzadera secuencia arriba del sitio de cebador T3 (por ejemplo, BspEI);
 8. Un sitio de cebador T3 con orientación inversa; y
 - 50 9. Dos sitios de endonucleasa no raros, únicos y no variables que definen el sitio de inserción 3' para el vector pUC19 mutado descrito anteriormente (por ejemplo, Pmel y SapI).

Aunque la presente descripción que comprende la presente invención desvela métodos para construir transgenes en vectores plasmídicos de clonación, se pueden utilizar métodos similares para construir transgenes en moléculas de ADN extracromosómico más grandes tales como cósmidos o cromosomas artificiales incluyendo cromosomas bacterianos artificiales (BAC). Para su uso en plantas, se puede utilizar también un vector T1. La amplia variedad de elementos génicos que se pueden incorporar en los vectores plasmídicos de clonación también permiten transferir los productos del transgén final a una amplia variedad de organismos huésped con poca o ninguna manipulación.

60 Las FIG. 2 y 3 son una ilustración general de la modularidad de la invención. Como se muestra en la FIG. 2, hay uno de cada vector lanzadera de Promotor, de Expresión, y Regulador 3'. Flanqueando cada inserción en los vectores lanzadera hay sitios de restricción de endonucleasa que son específicos para crear un punto de acoplamiento. Más específicamente, en la FIG. 2, la inserción de Promotor (P) está flanqueada por un primer grupo de uno o más sitios de endonucleasa de restricción representados por una estrella en el extremo 5' y un segundo grupo de uno o más sitios de restricción de endonucleasa representadas por un círculo en el extremo 3'; el módulo de Expresión está flanqueado por el segundo grupo de sitios de endonucleasa de restricción representados por un círculo en el

extremo 3' y un tercer grupo de uno o más sitios de endonucleasa de restricción representados por un triángulo en el extremo 3' y el módulo Regulador 3' (3) está flanqueado por el tercer grupo de sitios de endonucleasa de restricción representados por un triángulo en el extremo 5' y un cuarto grupo de uno o más sitios de restricción de endonucleasa representados por un cuadrado en el extremo 3'. Escindiendo cada sitio de reconocimiento de endonucleasa por la endonucleasa específica de cada sitio se crean extremos pegajosos en el extremo 5' y 3' de cada módulo como se indica en la parte inferior de la FIG. 2 por las inserciones en el final de las flechas con líneas discontinuas. Los módulos se pueden ahora combinar con un vector plasmídico de clonación que también se ha escindido en sus dos puntos de acoplamiento fijos por endonucleasas específicas para el primer grupo de sitios de endonucleasa de restricción (representados por la estrella) y el cuarto grupo de sitios de endonucleasa de restricción (representado por el cuadrado). Cuando los vectores modulares se colocan con el vector plasmídico de clonación en una mezcla de reacción apropiada, los extremos pegajosos escindidos (representados por estrellas huecas, círculos, triángulos, y cuadrados) de cada vector modular se auto-orientan en el plásmido y se unen secuencialmente, con los extremos estrella escindidos que combinan, los extremos de círculo escindidos que combinan, los extremos de triángulo escindidos que combinan, y los extremos de cuadrado que combinan. Esto da como resultado una armazón de vector ensamblado que se muestra en la FIG. 3. Además cada uno de los grupos de sitios de endonucleasa representados por la estrella, círculo, triángulo y cuadrado pueden escindirse una vez de nuevo por su endonucleasa específica correspondiente, de manera que más tarde una inserción particular se puede retirar y remplazarse por otra inserción de interés.

Se pueden insertar múltiples armazones de vector (ejemplo, PE3-1 y PE3-2) en un único plásmido de acoplamiento. La naturaleza asimétrica de las colas que protruyen por una endonucleasa tales como I-Sce I, como con otras HE, crea una herramienta potente cuando se utilizan en combinación con el sitio de la enzima endonucleasa BstX I (5' CCANNNNN/NTGG 3', donde 'N' puede ser cualquier nucleótido). El dominio de secuencia neutra de BstX I se puede utilizar para generar extremos cohesivos compatibles para dos colas que protruyen por I-Sce I orientadas inversamente, a la vez que se evita la auto-hibridación. Con este método, se puede colocar una primera inserción PE3-1, que tiene un sitio I-Sce I en sus extremos en un vector plasmídico de clonación escindiendo el plásmido en los sitios de endonucleasa BstXI/Scel. Se puede entonces cortar de nuevo con I-Sce I e insertar una PE3-2 que tenga un sitio I-Sce I en sus extremos. Este armazón completo puede entonces escindirse de su plásmido de acoplamiento por PI-Sce I e insertarse en otro plásmido de acoplamiento que contiene sitios de endonucleasa BstX I/PI-Sce I. Este segundo plásmido de acoplamiento tiene también sitios de endonucleasa para PI-Sce I, en el cual se puede insertar otro módulo de un plásmido de acoplamiento, que contenga posiblemente un PE3-3 y PE3-4, (no mostrado). De esta manera, un investigador puede obtener más información de una célula, es decir, se pueden insertar múltiples genes en el contexto de un vector único, lo cual no se había conseguido previamente por los expertos en la técnica. Dicho proceso puede ahorrar tanto dinero como tiempo para los investigadores que trabajan en este campo.

Ejemplos

EJEMPLO 1 (no según la presente invención): Plásmido de Acoplamiento PE3. Como ejemplo del método, puede construirse un transgén que contenga estos elementos:

1. Secuencias de nucleótidos del promotor humano de la proteína C tensioactiva (SP-C);
2. Secuencias que codifican el producto proteico del gen de ratón del receptor beta c del factor estimulante de colonias de macrófagos-granulocitos (GMR β c).
3. Secuencias de intrón de betaglobina de conejo; y
4. Señal de poli-A SV40.

Las secuencias SP-C contienen sitios BamHI internos, y se pueden liberar de su plásmido parental solo con NotI y EcoRI. El GMR β c tiene un sitio NotI interno, y se puede cortar de su plásmido parental con BamHI y XhoI. Las secuencias de intrón de betaglobina de conejo se pueden cortar de su plásmido parental con EcoRI. La cola poli-A SV40 se puede cortar de su plásmido parental con XhoI y SacI. Debido a la redundancia de varios sitios de endonucleasa, ninguno de los plásmidos parentales se puede utilizar para ensamblar todos los fragmentos necesarios.

Las etapas utilizadas para construir el transgén deseado en el Plásmido de Acoplamiento PE3 de la invención son las siguientes.

1. Como NotI y PspOMI generan extremos cohesivos compatibles, las secuencias de promotor de SP-C humano se escinden con NotI y EcoRI y se clonan en los sitios PspOMI y EcoRI del vector lanzadera P. El producto de esta reacción se llama pSVP-SPC.
2. A continuación de la propagación y las etapas de recuperación bien conocidas por los expertos en la técnica, se clonan las secuencias de intrón de betaglobina de conejo en el sitio EcoRI de pSVP-SPC. La orientación del intrón en la construcción intermedia resultante se verifica por secuenciación del producto, llamado pSVP-SPC-r β G.
3. El promotor y el intrón se escinden y se aíslan como un fragmento contiguo a partir del pSVP-SPC-r β G utilizando AsiSI y Ascl. Al mismo tiempo, el Plásmido de Acoplamiento PE3 se corta con AsiSI y Ascl para prepararlo para la unión con el segmento promotor/intrón. El fragmento promotor/intrón se une al Plásmido de

Acoplamiento, se propaga y se recupera.

4. El sitio XhoI del fragmento GMR β c se llena para crear un extremo truncado, utilizando técnicas bien conocidas por los expertos en la técnica. Se clona después en el sitio BamHI y el sitio Pvu2 con extremos truncados de pSVP-SPC-r β G. El plásmido resultante (pDP-SPC-GMR- β c-r β c) se propaga y recupera.

5. La etapa final de clonación es la adición de la cola Poli-A SV40. El fragmento poli-A SV40 se corta con XhoI y SacI, como están en el vector receptor pDS1-SPC-GMR β c-r β G. Ambas porciones de ADN se purifican en gel y se recuperan. Se prepara una mezcla de unión con una relación molar de 10:1 de poliA SV40 respecto a pDS1-SPC-GMR β c-r β G. Los productos de la unión se propagan y recolectan. El nuevo plásmido pDS1-SPC-GMR- β c-r β G-pA contiene todos los elementos necesarios para el transgén, incluyendo un único sitio de endonucleasa en el extremo 3' con el que se puede alinear el plásmido pDS1-SPC-GMR- β c-r β G-pA completo para la transfección en células eucariotas o microinyección en el pronúcleo de un óvulo fertilizado.

EJEMPLO 2 (no según la presente invención): Ensamblaje Dinámico de Vector

El ensamblaje dinámico del vector se ilustra en el siguiente ejemplo:

1. Se insertaron secuencias promotoras de citomegalovirus humano (CMV) en un vector lanzadera P (SVP), que tienen endonucleasas AsiSI y AscI en las partes 5' y 3', respectivamente. Se amplificaron los plásmidos, y el módulo promotor se escindió del vector por digestión con endonucleasas AsiSI y AscI y se aisló.

2. Se insertaron secuencias que codifican una proteína luciferasa en un Vector Lanzadera de Expresión (SVE), que tiene endonucleasas AscI y NotI en las partes 5' y 3', respectivamente. Los plásmidos se amplificaron, y el módulo de Expresión se escindió el vector por digestión con endonucleasas AscI y NotI y se aisló.

3. Se insertaron secuencias que codifican un intrón de mamífero y un sitio de poliadenilación SV40 en un Vector Lanzadera Regulador 3' (SV3), que tiene endonucleasas NotI y BsiWI en las partes 5' y 3', respectivamente. Los plásmidos se amplificaron, y se escindió el módulo Regulador del vector por digestión con endonucleasas NotI y BsiWI y se aisló.

4. Los sitios de reconocimiento de endonucleasa en un Vector plasmídico de Acoplamiento que tienen endonucleasas AsiSI y BsiWI en la parte 5' del módulo promotor y la parte 3' del módulo regulador, respectivamente, se escindieron con endonucleasas AsiSI y BsiWI y se aislaron.

5. Los módulos Promotor, de Expresión, y Regulador se combinaron en el Vector plasmídico de Acoplamiento en una mezcla de unión. Tras una incubación de dos horas, la mezcla de unión se utilizó para transformar *E. coli*, la cual se diseminó en una placa LB agar con ampicilina. La placa se incubó a 37 °C durante una noche. Las colonias se aislaron y propagaron en cultivos de caldo LB líquidos. El plásmido ADN se aisló de cada cultivo con caldo LB. Se analizó el ADN por mapeo de endonucleasa para determinar si los plásmidos de cada colonia contenían las tres inserciones modulares (de Promotor, de Expresión y Reguladora). Se identificó un plásmido que contenía las tres inserciones modulares como el transgén pCMV-luc-SV40 pA. Se puede alinear utilizando endonucleasa I-SceI y se inyectó en pronúcleos de ratón para generar ratones CMV-luciferasa. El promotor CMV en este ejemplo dirige la expresión del gen de luciferasa en todos los tejidos de un organismo huésped, tal como un ratón CMV-luciferasa.

EJEMPLO 3 (no según la presente invención): Rediseño de un Ensamblaje dinámico de vector

Si el investigador deseara ahora refinar el patrón de expresión de manera que la luciferasa solo se expresase en un tejido o tipo celular en particular, podría rápida y fácilmente reemplazar el promotor CMV con lo que proporciona un patrón de expresión restringido. El siguiente ejemplo ilustra el uso para facilitar el rápido rediseño del pCMV-luc-pA:

1. Un promotor específico neuronal, Enolasa específica de neurona (NSE), se inserta en un vector lanzadera P (SVP) y se prepara como el módulo promotor en el ejemplo anterior.

2. Se escinde el pCMV-luc-pA con AsiSI y AscI para retirar el módulo promotor CMV. El resto del vector plasmídico que contiene los módulos de expresión y regulador se aíslan.

3. El módulo promotor NSE se coloca en una mezcla de unión con el resto de vector plasmídico de acoplamiento que contiene los módulos de expresión y regulador intactos. Después de la incubación durante 3 horas, se utiliza la nueva mezcla de unión para transformar *E. coli*. La mezcla con *E. coli* se disemina en una placa LB agar con ampicilina como en el ejemplo anterior. Las colonias se aislaron al día siguiente, se propagaron, y se aisló el plásmido ADN de cada uno. Se utilizó el mapeo de endonucleasa para identificar los plásmidos que contenían el módulo promotor NSE deseado.

EJEMPLO 4: Matriz de transgenes

El siguiente ejemplo es una ilustración del uso de la invención para ensamblar rápidamente una matriz de transgenes, que contiene cada una de las diferentes combinaciones de módulos Promotor, de Expresión y Regulador. Se utilizaría una serie de seis vectores lanzadera y un vector con una estación de acoplamiento PE3 para generar ocho productos de vectores distintos utilizando un ensamblaje combinatorio. La serie de seis vectores lanzadera consiste en dos lanzaderas P (SVP), dos lanzaderas E (SVE), y dos lanzaderas 3 (SV3). Los dos lanzaderas P (SVP) distintos contienen cada uno un promotor de citomegalovirus humano (CMV) o un promotor SPC específico de pulmón de ratón, y cada uno tiene sitios para endonucleasas AsiSI y AscI en las partes 5' y 3',

respectivamente. Los dos lanzaderas E distintos contienen cada uno un ADNc de luciferasa o un ADNc EGFP, y cada uno tiene sitios de endonucleasa Asc I y Not I en las partes 5' y 3', respectivamente. Los dos vectores lanzadera 3 distintos contienen cada uno una señal poliA SV40 o la región reguladora 3' de la hormona de crecimiento humana (hGH), y cada uno tiene sitios de endonucleasa Not I y BsiWI en las partes 5' y 3', respectivamente.

Los módulos de promotor se liberan de sus respectivos vectores lanzadera SVP por digestión individual apropiada del vector lanzadera con endonucleasas AsiSI y Ascl. Los productos de restricción resultantes se someten individualmente a electroforesis en gel y la banda de ADN correspondiente al módulo de promotor apropiado se somete a purificación en gel. Este procedimiento rendirá tanto el módulo promotor CMV como un módulo promotor SPC unidos en el lado 5' por una protuberancia AsiSI y por una protuberancia Asc I en el extremo 3'.

Los módulos de expresión se liberan de sus vectores lanzadera SVE respectivos por digestión individual apropiada del vector lanzadera con las endonucleasas de restricción Ascl y NotI. Los productos de restricción resultantes se someten individualmente a electroforesis en gel y la banda de ADN correspondiente al módulo de expresión apropiado se somete a purificación en gel. Este procedimiento rendirá un módulo de expresión de luciferasa o un módulo de expresión EGFP unido en el lado 5' por una protuberancia Ascl y por una protuberancia Not I en el extremo 3'.

Los módulos reguladores 3' se liberan de sus respectivos vectores lanzadera SV3 por digestión individual apropiada del vector lanzadera con las endonucleasas de restricción NotI y BsiWI. Los productos de restricción resultantes se someten individualmente a electroforesis en gel y la banda de ADN correspondiente al módulo regulador 3' apropiado se somete a purificación en gel. Este procedimiento producirá un módulo regulador 3' SV40 o un módulo regulador 3' hGH unido en el lado 5' por una protuberancia NotI y por una protuberancia BsiWI en el extremo 3'.

El vector con la estación de acoplamiento PE3 se prepara por digestión con las endonucleasas de restricción AsiSI y BsiWI. Para ayudar a evitar que el vector se vuelva a unir, se expone la digestión de restricción del vector a fosfatasa intestinal bovina (CIP) durante una hora a 37 °C. El producto de restricción del vector tratado con CIP resultante se somete entonces a electroforesis en gel y la banda de ADN correspondiente al armazón del vector PE3 alineado se somete a purificación en gel.

Las muestras de los siete fragmentos de ADN purificados en gel se analizaron en cuanto a la identidad, integridad, pureza, y cantidad ejecutando un gel electroforético diagnóstico. Los datos cuantitativos con respecto a la abundancia relativa del vector con estación de acoplamiento PE3 purificado y los módulos de ADN respectivos se utilizaron para definir la cantidad de cada componente necesario para una reacción de unión combinatoria.

Cuando se lleva a cabo una reacción de unión, hay dos estrategias que dan lugar frecuentemente a resultados satisfactorios. La primera estrategia es llevar a cabo las mezclas de reacción de unión en las que la relación entre la inserción respecto al vector es aproximadamente de 3:1. La segunda estrategia, que se utiliza cuando se va a introducir más de una inserción en un único vector simultáneamente, es introducir un equivalente molar de cada módulo génico que se insertará en el vector. Esto se puede conseguir añadiendo un volumen variable de los módulos en el contenedor de reacción con el fin de obtener una equivalencia molar en el contexto de la mezcla de reacción de unión, o añadiendo una solución tampón neutra a cada uno de los módulos purificados de manera que sus concentraciones sean equivalentes con una base de relación molar. En este ejemplo, el vector purificado en gel y los fragmentos de inserciones se ajustaron en equivalencia molar utilizando el tampón 10 mM Tris, pH 8,0. El volumen de reacción de unión total se fijó en 150 microlitros. La mezcla de reacción de unión consiste en los siguientes constituyentes: 39 microlitros de agua ultrapura, 15 microlitros de 10x tampón de ligasa, 5 microlitros del armazón de vector PE3 purificado, 15 microlitros del módulo Promotor CMV purificado, 15 microlitros de módulo promotor SPC purificado, 15 microlitros del módulo de expresión de luciferasa purificado, 15 microlitros de módulo de expresión EGFP purificado, 15 microlitros de módulo regulador 3' SV40 purificado, 15 microlitros del módulo regulador 3' hGH purificado, y 1 microlitro de enzima ligasa. Los componentes de reacción resultantes se mezclaron concienzudamente y luego se incubaron durante una noche a 16 °C.

Los productos de unión previstos incluyen los siguientes:

- pCMV-EGFP-SV40
- pCMV-EGFP-hGH
- pCMV-Luciferasa-SV40
- pCMV-Luciferasa-hGH
- pSPC-EGFP-SV40
- pSPC-EGFP-hGH
- pSPC-Luciferasa-SV40
- pSPC-Luciferasa-hGH

La mezcla de unión se utiliza entonces para transformar *E. coli*, que se diseminan entonces en una placa LB agar con ampicilina. La placa se incubó a 37 °C durante una noche. Las colonias se aíslan y propagan en cultivos

líquidos de caldo LB individuales. El plásmido ADN se aísla de cada cultivo en caldo LB. El ADN se analiza por mapeo de endonucleasa para determinar la identidad del vector resultante incorporado en cada colonia. En el ejemplo anterior, uno de los productos de vector previstos (pCMV-EGFP-SV40) no se produjo durante el primer proceso combinatorio. Un vector que se produjo satisfactoriamente (pC-MV-Luciferasa-SV40) pudo, sin embargo, servir como armazón de vector para producir el vector pCMV-EGFP-SV40 deseado. Puede hacerse referencia a esta técnica como "Ensamblaje en Segundo Paso".

EJEMPLO 5 (no según la presente invención): Ensamblaje en segundo paso

Con el fin de construir el vector pCMV-EGFP-SV40 deseado, el producto vector pC-MV-Luciferasa-SV40 del Ejemplo 4 se digirió con *Ascl* y *NotI*, se trató con CIP y posteriormente se purificó en gel. Este fragmento de vector alineado, en el que se había retirado el módulo de luciferasa, se incubaba en una mezcla de unión que contiene el módulo EGFP que se produjo en el ejemplo anterior de ensamblaje combinatorio. La mezcla de unión se utiliza para transformar *E. coli*, que entonces se disemina en una placa LB agar con ampicilina. La placa se incubaba a 37 °C durante una noche. Las colonias se aíslan y propagan en cultivos líquidos individuales de caldo LB. El plásmido ADN se aísla de cada cultivo en caldo LB. El ADN se analiza por mapeo de endonucleasa para determinar si los plásmidos de cada colonia contienen la inserción EGFP.

Entre las muchas ventajas de la presente invención, se puede apreciar fácilmente que se puede ensamblar rápidamente una matriz de transgenes que contiene cada una de las diferentes combinaciones de módulos de Promotor, de Expresión y Regulador, en un periodo de tiempo muy corto, así como variar o rediseñar rápida y fácilmente un ensamblaje nuevo de transgén. En el pasado, para variar un transgén ensamblado utilizando los métodos conocidos para crear una matriz de diferentes transgenes, que tuviera cada uno de los diferentes módulos de Promotor, de Expresión y Regulador habitualmente duraría un año o más de tiempo en el laboratorio. Utilizando los métodos de la presente invención, se puede hacer el mismo número de transgenes deseados en días o semanas, y luego hacer el ensayo deseado de cada uno, ahorrando de esta manera al investigador la gran cantidad de tiempo de anteriormente. Además, tanto el Ensamblaje Dinámico de Vector, en el que una inserción de Promotor, de Expresión o Reguladora se puede insertar en un único armazón al mismo tiempo, y el método de combinación descrito, en el que dos lanzaderas P, dos lanzaderas E y dos lanzaderas reguladoras se combinan para crear ocho tipos diferentes de transgenes, que se pueden utilizar para ahorrar un tiempo precioso y dinero a los investigadores. Las lanzaderas que se crearon originalmente por métodos de síntesis *de novo*, recombinación, y PCR de clonación de protuberancias terminadoras se pueden coger y usar con la tecnología de punto de acoplamiento de la presente invención para ensamblar estos elementos pre-producidos en una multitud de transgenes.

35

REIVINDICACIONES

1. Un método para sintetizar simultáneamente una matriz de transgenes, que comprende las etapas de:

- 5 a. proporcionar un vector plasmídico de clonación primario que comprende un armazón, comprendiendo el armazón
- 10 (i) al menos un primer punto de acoplamiento y un segundo punto de acoplamiento estando cada punto de acoplamiento fijado en el armazón y que comprenden al menos un sitio de restricción raro de más de 6 nucleótidos para una enzima de restricción rara no variable y
- (ii) un único sitio para HE en orientación directa localizado secuencia arriba desde el extremo 5' del primer punto de acoplamiento y un único sitio para HE en orientación inversa localizado secuencia abajo desde el extremo 3' del segundo punto de acoplamiento;
- 15 b. escindir el primer punto de acoplamiento con una primera enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción raros de más de 6 nucleótidos del primer punto de acoplamiento;
- c. escindir el segundo punto de acoplamiento con una segunda enzima de restricción rara no variable que se corresponde con uno de los sitios de restricción del segundo punto de acoplamiento;
- 20 d. proporcionar al menos una inserción de Promotor que comprende una secuencia de Promotor de interés, con un extremo 5' que es compatible con el extremo 3' del primer punto de acoplamiento;
- e. proporcionar al menos una inserción de Expresión que comprende una secuencia de Expresión de interés, con un extremo 5' que es compatible con el extremo 3' de la al menos una inserción de Promotor para formar un sitio de restricción raro de más de 6 nucleótidos para una tercera enzima de restricción rara no variable;
- 25 f. proporcionar al menos una inserción Reguladora que comprende una secuencia Reguladora de interés, con un extremo 5' que es compatible con el extremo 3' de la al menos una inserción de Expresión para formar un sitio de restricción raro de más de 6 nucleótidos para una cuarta enzima de restricción rara no variable, y un extremo 3' que es compatible con el extremo 5' del segundo punto de acoplamiento que se escindió en la etapa 'c', en donde la primera, la segunda, la tercera y la cuarta enzimas de restricción raras no variables son diferentes;
- 30 g. colocar al menos dos tipos diferentes de al menos una de las inserciones de Promotor, de Expresión y Reguladora, al menos una de cada una de las inserciones restantes y el vector plasmídico de clonación escindido en una mezcla de reacción apropiada para producir la unión simultánea, la auto-orientación y la colocación secuencial de cada una de las inserciones de Promotor, de Expresión y Reguladora entre el primer y el segundo puntos de acoplamiento del armazón, creando de esta manera una matriz de plásmidos que tienen diferentes combinaciones de inserciones de Promotor, de Expresión y Reguladoras en su armazón;
- 35 m. escindir el armazón en cada uno de los sitios para HE únicos con una enzima de restricción HE única;
- n. purificar la parte escindida que contiene las inserciones; y
- o. insertar la parte escindida en un genoma huésped de interés.

40 2. El método de la reivindicación 1, en el que la etapa 'g' comprende la colocación de al menos dos tipos diferentes de al menos dos de las inserciones de Promotor, de Expresión y Reguladora, al menos una de cada una de las inserciones restantes y el vector plasmídico de clonación escindido en una mezcla de reacción apropiada para producir la unión simultánea, la auto-orientación y la colocación secuencial de cada una de las inserciones de Promotor, de Expresión y Reguladora entre el primer y el segundo puntos de acoplamiento en el armazón, creando de esta manera una matriz de plásmidos que tiene diferentes combinaciones de inserciones de Promotor, de Expresión y Reguladora en su armazón.

50 3. El método de la reivindicación 1, en el que dicha etapa de colocación 'g' comprende la colocación de al menos dos tipos diferentes de cada una de las inserciones de Promotor, de Expresión y Reguladora y el vector plasmídico de clonación escindido en una mezcla de reacción apropiada para producir la unión simultánea, la auto-orientación y la colocación secuencial de cada una de las inserciones de Promotor, de Expresión y Reguladora entre el primer y el segundo puntos de acoplamiento en el armazón, creando de esta manera una matriz de plásmidos que tienen diferentes combinaciones de inserciones de Promotor, de Expresión y Reguladora en su armazón.

55 4. El método de las reivindicaciones 1, 2 o 3 que comprenden además las etapas 'h'-'l' de:

- h. ensayar los productos de la etapa 'g' utilizando cribado de alto rendimiento;
- i. aislar un producto específico de la etapa 'g' que es un vector plasmídico de clonación primario;
- 60 j. escindir el armazón del vector plasmídico primario aislado en la etapa 'i' en los extremos 5' y 3' de una de las inserciones de Promotor, de Expresión y Reguladora con el par de enzimas de restricción raras que se corresponden con ellos;
- k. proporcionar una sexta inserción, en la que el extremo 5' de la sexta inserción es compatible con el extremo 3' del armazón que se escindió en la etapa 'j', en donde el extremo 3' de la sexta inserción es compatible con el extremo 5' del armazón que se escindió en la etapa 'j'; y
- 65 l. colocar posteriormente la sexta inserción y el armazón que se escindió en la etapa 'j' en una mezcla de reacción apropiada para producir la unión simultánea y la auto-orientación de la sexta inserción en el armazón;

en donde las etapas 'h'-l' siguen a la etapa 'g' y preceden a la etapa 'm'.

5 El método de la reivindicación 4, que comprende además la etapa de repetir las etapas 'h', 'i', 'j', 'k' y 'l' para cualquier número de inserciones con el fin de reemplazar secuencialmente una inserción por otra inserción en el almacén.

10 6. El método de la reivindicación 1, en el que las inserciones se crean por un método que se selecciona de entre el grupo que consiste en síntesis *de novo*, recombinación, PCR, tecnología de clonación de protuberancia terminadora y mapeo de endonucleasas de restricción.

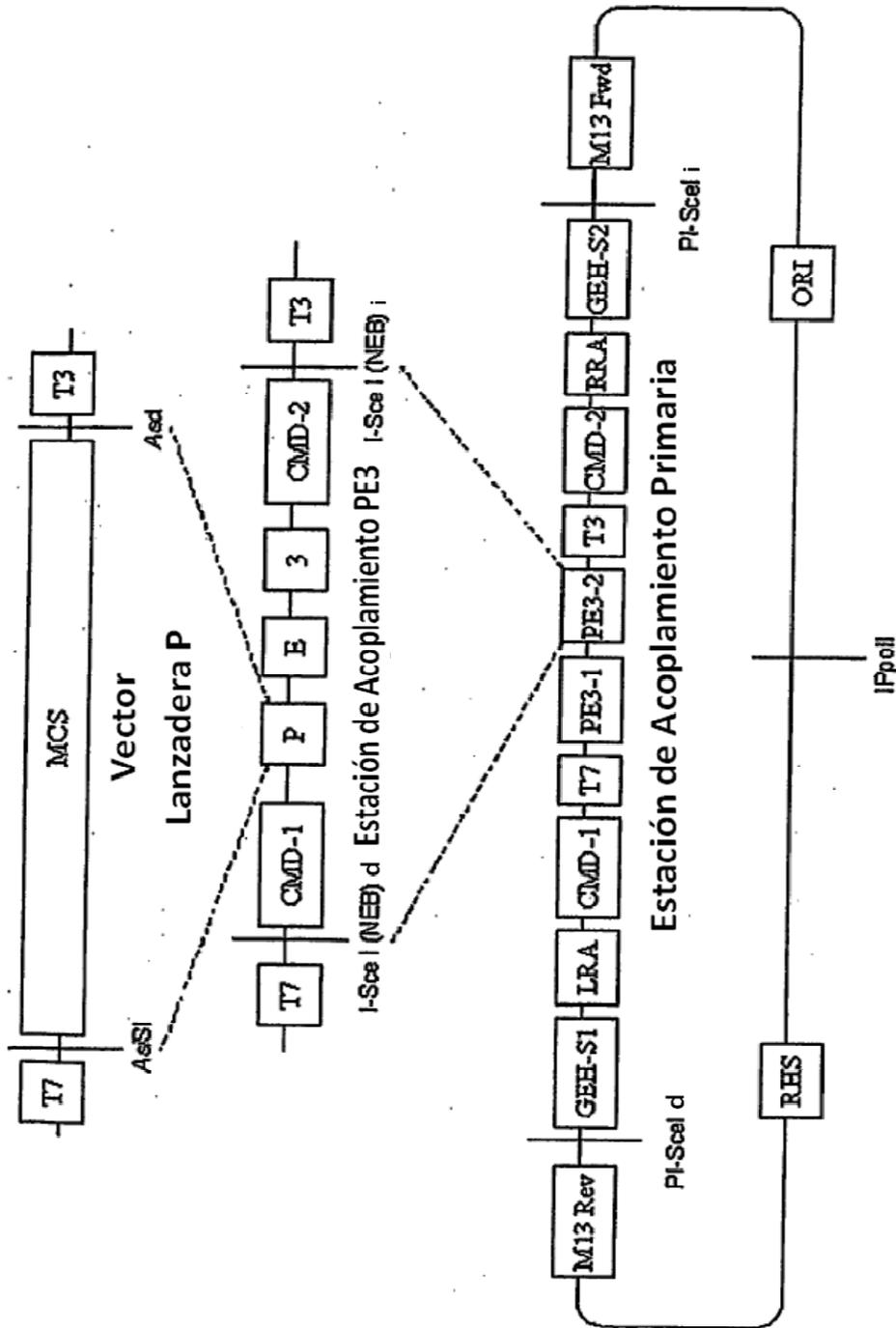


Fig. 1

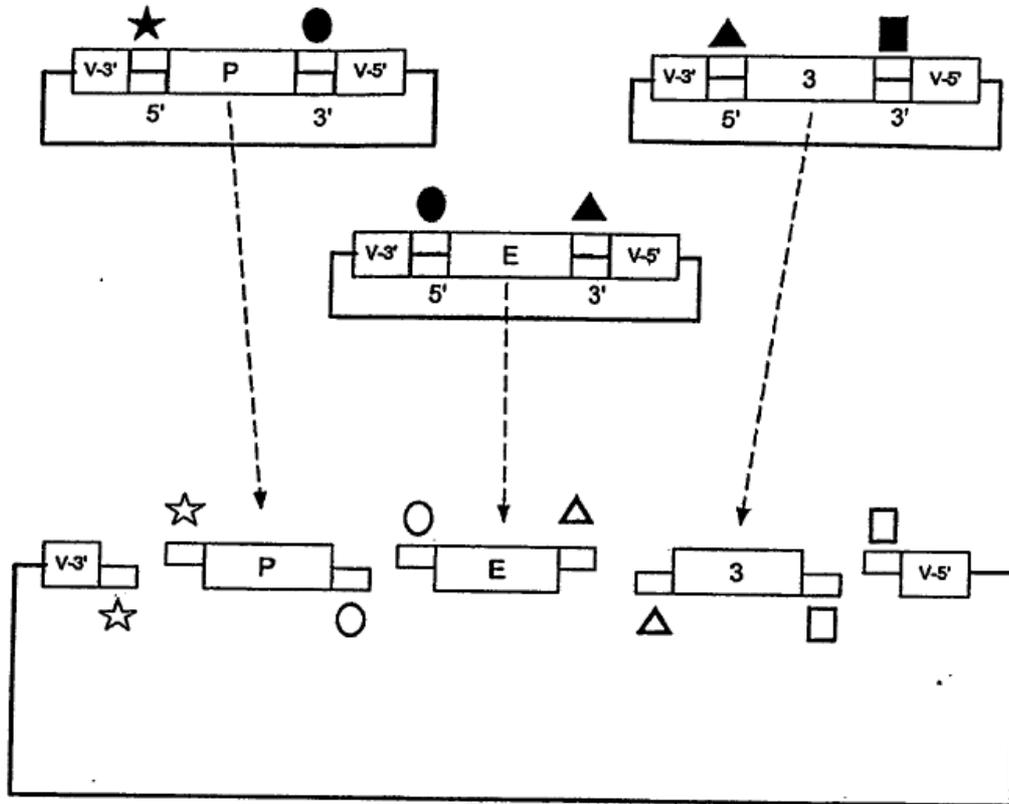


Fig. 2

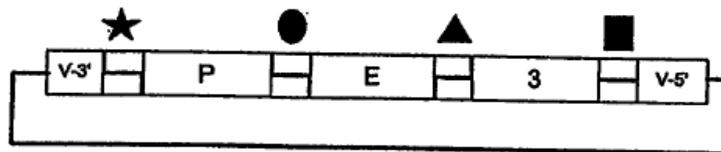


Fig. 3