



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 347 684**

51 Int. Cl.:

C12N 9/22 (2006.01)

C12N 15/55 (2006.01)

C12N 15/10 (2006.01)

C12N 15/64 (2006.01)

C12N 15/79 (2006.01)

C12N 15/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06744673 .2**

96 Fecha de presentación : **15.03.2006**

97 Número de publicación de la solicitud: **1863909**

97 Fecha de publicación de la solicitud: **12.12.2007**

54

Título: **Variantes de la meganucleasa I-Crel con especificidad modificada, método de preparación y usos de las mismas.**

30

Prioridad: **15.03.2005 WO PCT/IB05/000981**
19.09.2005 WO PCT/IB05/003083

45

Fecha de publicación de la mención BOPI:
03.11.2010

45

Fecha de la publicación del folleto de la patente:
03.11.2010

73

Titular/es: **CELLECTIS**
102, route de Noisy
93235 Romainville, Cédex, FR

72

Inventor/es: **Paques, Frederic y**
Duchateau, Philippe

74

Agente: **Martín Santos, Victoria Sofía**

ES 2 347 684 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).

VARIANTES DE LA MEGANUCLEASA I-CREI CON ESPECIFICIDAD MODIFICADA,
MÉTODO DE PREPARACION Y USOS DE LAS MISMAS.

DESCRIPCIÓN

5 La presente invención se refiere a un método para preparar variantes de meganucleasa I-*CreI* que tienen una especificidad de escisión modificada. La invención se refiere también a las variantes de meganucleasa I-*CreI* que pueden obtenerse mediante dicho método y a sus aplicaciones tanto para la escisión de una nueva diana de ADN como para ingeniería genética e ingeniería genómica para propósitos no terapéuticos.

10 La invención también se refiere a ácidos nucleicos que codifican dichas variantes, a los casetes de expresión que comprenden dichos ácidos nucleicos, a los vectores que comprenden dichos casetes de expresión, a células u organismos, plantas o animales excepto seres humanos, transformados mediante dichos vectores.

Las meganucleasas son endonucleasas específicas de secuencia que reconocen sitios de escisión de ADN grandes (>12pb; habitualmente 14-40 pb) (Thierry y Dujon, 1992). En la naturaleza, las meganucleasas están representadas esencialmente por endonucleasas de movimiento dirigido, generalmente codificadas por elementos genéticos móviles tales como inteínas e intrones de clase I (Belfort y Roberts, 1997; Chevalier y Stoddard, 2001). El movimiento dirigido se refiere a la movilización de estos elementos, que se basa en la reparación de una ruptura de doble cadena de ADN (DSB), iniciada por la actividad endonucleasa de la meganucleasa. Estudios tempranos de las proteínas HO (Haber, 1998; Klar *et al.*, 1984; Kostriken *et al.*, 1983), I-*SceI* (Colleaux *et al.*, 1988; Jacquier y Dujon, 1985; Perrin *et al.*, 1993; Plessis *et al.*, 1992) e I-*TevI* (Bell-Pedersen *et al.*, 1990; Bell-Pedersen *et al.*, 1989; Bell-Pedersen *et al.*, 1991; Mueller *et al.*, 1996) han ilustrado la biología del proceso de movimiento dirigido. Por otro lado, estos estudios también han proporcionado un paradigma para el estudio de la reparación de DSB en células vivas.

La asimetría general de las secuencias diana de endonucleasa de movimiento dirigido contrasta con la simetría de díada característica de la mayoría de los sitios de reconocimiento de enzimas de restricción. Se ha demostrado que varias endonucleasas de movimiento dirigido codificadas por ORF de intrones o inteínas promueven el movimiento dirigido de sus elementos genéticos respectivos a sitios alélicos sin intrones o sin inteínas. Al realizar una rotura de doble cadena específica de sitio en los alelos sin intrón y sin inteína, estas nucleasas crean extremos recombinantes, que entablan un proceso de conversión génica que duplica la secuencia codificante y conduce a la inserción de un intrón o una secuencia intercalada a nivel de ADN.

Las endonucleasas de movimiento dirigido se incluyen en 4 familias separadas en base a motivos aminoacídicos muy bien conservados [para revisión, véase Chevalier y Stoddard (Nucleic Acids Research, 2001, 29, 3757-3774)]. Una de ellas es la familia de los dodecapéptidos (dodecámero, DOD, D1-D2, LAGLIDADG, P1-P2). Esta es la familia más grande de proteínas agrupadas por su motivo de secuencia conservada más general: una o dos copias (amplia mayoría) de una secuencia de doce restos: el

dodecapéptido. Las endonucleasas de movimiento dirigido con un dodecapéptido (D) tienen alrededor de 20 kDa de masa molecular y funcionan como homodímeros. Las que tienen dos copias (DD) varían desde 25 kDa (230 aminoácidos) a 50 kDa (HO, 545 aminoácidos) con de 70 a 150 restos entre cada motivo y actúan como un monómero. La escisión se da dentro del sitio de reconocimiento, dejando un corte escalonado de 4 nucleótidos con salientes 3'OH. Las enzimas que contienen una copia única del motivo LAGLIDADG, tales como I-CeuI e I-CreI actúan como homodímeros y reconocen un sitio de movimiento dirigido casi palindrómico.

La secuencia y la estructura de la endonucleasa de movimiento dirigido I-CreI (código de acceso pdb 1g9y) se ha determinado (Rochaix JD *et al.*, NAR, 1985, 13, 975-984; Heath PJ *et al.*, Nat. Struct. Biol., 1997, 4, 468-476; Wang *et al.*, NAR, 1997, 25, 3767-3776; Jurica *et al.* Mol. Cell, 1998, 2, 469-476) y se han generado modelos estructurales usando cristalografía de rayos X (Heath *et al.*, 1997).

I-CreI comprende 163 aminoácidos (código de acceso pdb 1g9y); dicha endonucleasa se corta como un dímero. El motivo LAGLIDADG corresponde a los restos 13 a 21; a cada lado de las α -hélices LAGLIDADG, lámina β cuádruple (posiciones 21-29; 37-48; 66-70 y 73-78) proporciona una interfaz de unión a ADN que dirige la interacción de la proteína con el semisitio de la secuencia de ADN diana. La interfaz de dimerización implica a las dos hélices LAGLIDADG así como otros restos.

El sitio de movimiento dirigido reconocido y escindido por I-CreI es de 22-24 pb de longitud y es un palíndromo degradado (véase figura 2 de Jurica MS *et al.*, 1998 y SEC ID N°:65). De manera más precisa, el sitio de movimiento dirigido de dicha I-CreI es una secuencia semi-palindrómica de 22 pb, con 7 de 11 pb idénticos en cada semisitio (Seligman LM *et al.*, NAR, 2002, 30,3870-3879).

La interfaz de endonucleasa-ADN también se ha descrito (véase figura 4 de Jurica MS *et al.*, 1998) y ha conducido a varias predicciones acerca de contactos específicos proteína-ADN (Seligman LM *et al.*, Genetics, 1997, 147, 1653-1664; Jurica MS *et al.*, 1998; Chevalier B. *et al.*, Biochemistry, 2004, 43, 14015-14026).

Surge de dichos documentos que:

- los restos G19, D20, Q47, R51, K98 y D137 son parte del sitio endonucleolítico de I-CreI;
- las secuencias de sitio de movimiento dirigido deben tener al menos 20 pb para conseguir una afinidad de enlace máxima de 0,2 nM;
- los contactos específicos de secuencia se distribuyen a lo largo de la longitud completa del sitio de movimiento dirigido;
- las sustituciones de pares de bases pueden tolerarse en muchas posiciones diferentes del sitio de movimiento dirigido, sin perturbar seriamente la unión o escisión del sitio de movimiento dirigido;
- R51 y K98 se localizan en el sitio de activo de la enzima y son candidatos para actuar como ácido de Lewis o para activar un donante de protones en la región de escisión; las mutaciones en cada uno de estos restos se ha observado que reducen en gran medida la actividad endonucleolítica de I-CreI (R51G, K98Q);

- cinco restos adicionales, que cuando mutan suprimen la actividad endonucleasa de I-CreI se localizan en o cerca del sitio activo de enzima (R70A, L39R, L91R, D75G, Q47H).

Estos estudios han preparado el terreno para un uso general de meganucleasa para ingeniería genómica. La dirección génica homóloga es la forma más precisa para modificar de forma estable un locus cromosómico en células vivas, pero su baja eficacia sigue siendo un inconveniente principal. Puesto que la DSB inducida por meganucleasa estimula la recombinación homóloga hasta 10.000 veces, las meganucleasas son actualmente la mejor manera para mejorar la eficacia de la inserción génica dirigida en células mamíferas (Choulika *et al.*, 1995; Cohen-Tannoudji *et al.*, 1998; Donoho *et al.*, 1998; Elliott *et al.*, 1998; Rouet *et al.*, 1994), y para llevarlas a eficacias viables en organismos tales como plantas (Puchta *et al.*, 1993; Puchta *et al.*, 1996) e insectos (Rong y Golic, 2000; Rong y Golic, 2001; Rong *et al.*, 2002).

Las meganucleasas se han usado para inducir diversos tipos de sucesos de recombinación homóloga, tales como recombinación de repetición directa en células de mamífero (Liang *et al.*, 1998), plantas (Siebert y Puchta, 2002), insectos (Rong *et al.*, 2002), y bacterias (Posfai *et al.*, 1999), o recombinación intercromosómica (Moynahan y Jasin, 1997; Puchta, 1999; Richardson *et al.*, 1998).

Sin embargo, esta tecnología aun está limitada por el bajo número de sitios diana naturales potenciales para meganucleasas: aunque se han identificado varios cientos de endonucleasas de movimiento dirigido naturales (Belfort y Roberts, 1997; Chevalier y Stoddard, 2001), la probabilidad de tener una meganucleasa natural que escinda un gen de interés es extremadamente baja. La realización de meganucleasas artificiales con especificidades dedicadas evitaría esta limitación.

Se han hecho endonucleasas artificiales con nuevas especificidades, basándose en la fusión de dominios de endonucleasas a dominios de unión a ADN de dedo de cinc (Bibikova *et al.*, 2003; Bibikova *et al.*, 2001; Bibikova *et al.*, 2002; Porteus y Baltimore, 2003).

Las endonucleasas de movimiento dirigido también se han usado como armazones para realizar endonucleasas nuevas, bien mediante fusión de diferentes dominios de proteínas (Chevalier *et al.*, 2002; Epinat *et al.*, 2003) o bien mediante mutación de restos de aminoácidos específicos únicos (Seligman *et al.*, 1997, 2002; Sussman *et al.*, 2004; Solicitud Internacional PCT WO 2004/067736).

La Solicitud Internacional PCT WO 2004/067736 describe un método general para producir una meganucleasa a medida derivada de una meganucleasa inicial, siendo capaz dicha variante de meganucleasa de escindir una secuencia diana de ADN que es diferente del sitio de reconocimiento y de escisión de la meganucleasa inicial. Este método general comprende los pasos de preparar una biblioteca de variantes de meganucleasa que tienen mutaciones en posiciones que contactan con la secuencia diana de ADN o que interactúan directa o indirectamente con dicho ADN diana y seleccionan las variantes capaces de escindir la secuencia diana de ADN. Cuando la meganucleasa inicial es la proteína I-CreI N75 una biblioteca, en la que los restos 44, 68 y 70 se han mutado, se construyó y se exploró contra una serie de seis dianas cercanas al sitio diana natural de I-CreI; los mutantes explorados tienen perfiles de unión alterados en comparación con la proteína de armazón I-CreI N75; sin embargo, escinden el sitio diana natural de I-CreI.

Seligman *et al.*, 2002, describe mutaciones que alteran la especificidad de escisión de I-*CreI*. Más específicamente, han estudiado el papel de nueve aminoácidos de I-*CreI* que se predice que contactan directamente con la diana de ADN (Q26, K28, N30, S32, Y33, Q38, Q44, R68 y R70). Entre estos nueve aminoácidos, se cree que siete interactúan con nucleótidos en posiciones simétricas (S32, Y33, N30, Q38, R68, Q44 y R70). Se construyeron mutantes que tienen cada uno de dichos nueve aminoácidos y un décimo (T140), que se predice que participan en interacción mediada por agua, convertidos a alanina, y se ensayaron en un ensayo basado en *E. coli*.

Los mutantes I-*CreI* resultantes se incluyen en cuatro clases fenotípicas definidas en relación con el sitio de movimiento dirigido de tipo silvestre:

- 10 - los contactos de S32A y T140A parecen ser los menos importantes para el reconocimiento de sitio de movimiento dirigido,
- N30A, Q38A y Q44A presentaron niveles intermedios de actividad en cada ensayo,
- Q26A, R68A y Y33A están inactivos,
- K28A y R70A están inactivos y no son tóxicos.

15 Surge de los resultados que los mutantes de I-*CreI* en las posiciones 30, 38, 44, 26, 68, 33, 28 y 70 tienen un comportamiento modificado en relación con el sitio de movimiento dirigido de I-*CreI* tipo silvestre.

Con respecto a las mutaciones que alteran las siete posiciones simétricas en el sitio de movimiento dirigido de I-*CreI*, surge de los resultados obtenidos que cinco de las siete posiciones simétricas de cada semisitio parece ser esencial para un reconocimiento de sitio eficaz *in vivo* mediante I-*CreI* de tipo silvestre: 2/21, 3/20, 7/16, 8/15 y 9/14 (correspondientes a las posiciones -10/+10, -9/+9, -5/+5, -4/+4 y -3/+3 en SEC ID N°:65). Todos los mutantes alterados en estas posiciones eran resistentes a escisión por I-*CreI* de tipo silvestre *in vivo*; sin embargo, el ensayo *in vitro* usando *E. coli* pareció ser más sensible que el ensayo *in vivo* y permitió la detección de sitios de movimiento dirigido de I-*CreI* de tipo silvestre de forma más eficaz que el ensayo *in vivo*; así pues el ensayo *in vitro* muestra que la diana de ADN de I-*CreI* de tipo silvestre puede ser la siguiente: el triplete gtc (sitio de movimiento dirigido reconocido en todos los documentos citados), gcc o gtt en las posiciones -5 a -3, en referencia a la SEC ID N°:65.

Seligman *et al.* también han estudiado la interacción entre la posición 33 de I-*CreI* y las bases 2 y 21 del sitio de movimiento dirigido (± 10) o entre la posición 32 de I-*CreI* y las bases 1 y 22 del sitio de movimiento dirigido (± 11); se descubrió que los mutantes Y33C, Y33H, Y33R, Y33L, Y33S e Y33T escindían un sitio de movimiento dirigido modificado en las posiciones ± 10 que no se escinde mediante I-*CreI* (Tabla 3). Por otro lado, se descubrió que S32K y S32R escindían un sitio de movimiento dirigido modificado en las posiciones ± 11 que se escinde por I-*CreI* (Tabla 3).

35 Sussman *et al.*, 2004, presenta estudios en los que la endonucleasa de movimiento dirigido de LAGLIDADG homodimérica I-*CreI* está alterada en las posiciones 26, y eventualmente 66, o en la posición 33, en contacto con las bases del sitio de movimiento dirigido de las posiciones ± 6 y ± 10 , respectivamente. Las construcciones enzimáticas resultantes (Q26A, Q26C, Y66R, Q26C/Y66R, Y33C,

Y33H) conducen una eliminación específica de ADN diana seleccionadas *in vivo* y muestran especificidades cambiadas de unión y escisión a ADN *in vitro*.

El resultado global de la selección y caracterización de mutantes puntuales enzimáticos contra variantes del sitio diana individual es tanto un cambio como una ampliación de la especificidad de unión y de la cinética de la escisión del sustrato.

Cada mutante presenta una constante de disociación más alta (menor afinidad) contra el sitio diana de tipo silvestre original que la enzima de tipo silvestre y cada mutante presenta una constante de disociación menor (mayor afinidad) contra su nueva diana que la enzima de tipo silvestre.

Los mutantes enzimáticos presentan una cinética similar de escisión de sustrato, con cambios y ampliación en las preferencias de sustratos similares a las descritas para afinidades de enlace.

Para alcanzar un mayor número de secuencias diana de ADN, sería extremadamente valioso generar nuevas variantes de I-CreI con nueva especificidad, es decir capaz de escindir dianas de ADN a las que no escinde I-CreI o las pocas variantes que se han aislado hasta el momento.

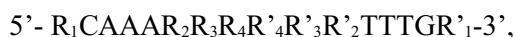
Dichas variantes serían de interés particular para ingeniería genética y genómica.

Aquí los inventores han descubierto mutaciones en las posiciones 44, 68 y 70 de I-CreI que dan como resultado variantes capaces de escindir al menos un sitio de movimiento dirigido modificado en las posiciones ± 3 a 5.

Por lo tanto, el tema de la presente invención es un método para preparar una variante de meganucleasa I-CreI que tenga una especificidad de escisión modificada, comprendiendo dicho método:

(a) reemplazar los aminoácidos Q44, R68 y/o R70 en referencia al código de acceso pdb de I-CreI 1g9y, con un aminoácido seleccionado del grupo que consiste en A, D, E, G, H, K, N, P, Q, R, S, T e Y;

(b) seleccionar las variantes de meganucleasa I-CreI obtenidas en el paso (a) que tienen al menos uno de los siguientes perfiles de escisión de tripletes R_3 en referencia a las posiciones -5 a -3 en una diana de ADN bicatenario, correspondiendo dichas posiciones -5 a -3 a R_3 de la siguiente fórmula I:



en la que:

R_1 está ausente o presente; y cuando está presente representa un fragmento de ácido nucleico que comprende de 1 a 9 nucleótidos que corresponden a una secuencia de ácidos nucleicos aleatoria o a un fragmento de un sitio de movimiento dirigido de meganucleasa I-CreI situado desde la posición -20 a la -12 (de 5' a 3'), correspondiendo R_1 al menos a la posición -12 de dicho sitio de movimiento dirigido,

R_2 representa el doblete de ácido nucleico ac o ct y corresponde a las posiciones -7 a -6 de dicho sitio de movimiento dirigido,

R_3 representa un triplete de ácido nucleico que corresponde a dichas posiciones -5 a -3, seleccionadas de entre g, t, c y a, excepto los siguientes tripletes: gtc, gcc, gtg, gtt y gct; por lo tanto dicho triplete de ácido nucleico se selecciona preferiblemente de entre los siguientes tripletes: ggg, gga, ggt, ggc, gag, gaa, gat, gac, gta, gcg, gca, tgg, tga, tgt, tgc, tag, taa, tat, tac, ttg, tta, ttt, ttc, tcg, tca, tct, tcc, agg, aga, agt, agc, aag,

aaa, aat, aac, atg, ata, att, atc, acg, aca, act, acc, cgg, cga, cgt, cgc, cag, caa, cat, cac, ctg, cta, ctt, ctc, ccg, cca, cct y ccc y más preferiblemente entre los siguientes tripletes: ggg, ggt, ggc, gag, gat, gac, gta, gcg, gca, tag, taa, tat, tac, ttg, ttt, ttc, tcg, tct, tcc, agg, aag, aat, aac, att, atc, act, acc, cag, cat, cac, ctt, ctc, ccg, cct y ccc,

5 R₄ representa el doblete de ácido nucleico gt o tc y corresponde a las posiciones -2 a -1 de dicho sitio de movimiento dirigido,

R'₁ está ausente o presente; y cuando está presente representa un fragmento de ácido nucleico que comprende de 1 a 9 nucleótidos que corresponden a una secuencia de ácidos nucleicos aleatoria o a un fragmento de un sitio de movimiento dirigido de meganucleasa I-*CreI* situado desde la posición +12 a +20 (de 5' a 3'), correspondiendo R'₁ al menos a la posición +12 de dicho sitio de movimiento dirigido,

10 R'₂ representa el doblete de ácido nucleico ag o gt y corresponde a las posiciones +7 a +6 de dicho sitio de movimiento dirigido,

R'₃ representa un triplete de ácido nucleico que corresponde a dichas posiciones +3 a +5, seleccionadas de entre g, t, c y a; siendo R'₃ diferente de gac, ggc, cac, aac y agc, cuando R₃ y R'₃ no son palindrómicos,

15 R'₄ representa el doblete de ácido nucleico ga o ac y corresponde a las posiciones +1 a +2 de dicho sitio de movimiento dirigido.

Definiciones

- Los restos de aminoácidos en una secuencia polipeptídica se designan en el presente documento de acuerdo con el código de una letra, en el que, por ejemplo, Q se refiere a Gln o resto de Glutamina, R se refiere a Arg o resto de Arginina y D se refiere a Asp o resto de ácido aspártico.

20 - En la presente invención, a no ser que se mencione lo contrario, los números de resto se refieren a la numeración del aminoácido de la secuencia de I-*CreI* SWISSPROT P05725 o el código de acceso pdb 1g9y. De acuerdo con esta definición, una variante denominada “ADR” es la meganucleasa I-*CreI* en la que los restos de aminoácido Q44 y R68 se han reemplazado por alanina y ácido aspártico, respectivamente mientras que R70 no se ha reemplazado. Otras mutaciones que no alteran actividad de escisión de la variante no se indican y la nomenclatura adoptada aquí no limita las mutaciones a sólo las tres posiciones 44, 68 y 70.

30 - Los nucleótidos se designan como sigue: el código de una letra se usa para designar la base de un nucleósido: a es adenina, t es timina, c es citosina y g es guanina. Para los nucleótidos degradados, r representa g o a (nucleótidos de purina), k representa g o t, s representa g o c, w representa a o t, m representa a o c, y representa t o c (nucleótidos de pirimidina), d representa g, a o t, v representa g, a o c, b representa g, t o c, h representa a, t o c y n representa g, a, t o c.

35 - En la presente solicitud, cuando se da una secuencia para ilustrar un sitio de reconocimiento o de movimiento dirigido, debe entenderse que representa, desde 5' a 3', solamente una cadena del polinucleótido de doble cadena.

- La expresión “secuencia parcialmente palindrómica”, “secuencia parcialmente simétrica”, “palíndromo degradado”, “secuencia pseudo-palindrómica” se usan indiscriminadamente para designar una secuencia palindrómica que tiene una simetría rota. Por ejemplo, la secuencia de 22 pb: c-₁₁a-₁₀a-₉a-₈a-₇c-₆g-₅t-

$4C-3g-2t-1g+1a+2g+3a+4C+5a+6g+7t+8t+9t+10g+11$ (SEC ID N°: 71) es una secuencia parcialmente palindrómica en la que la simetría está rota en los pares de bases +/- 1, 2, 6 y 7. De acuerdo con otra formulación, las secuencias de nucleótidos de las posiciones +/- 8 a 11 y +/- 3 a 5 son secuencias palindrómicas. El eje de simetría se sitúa entre los pares de bases de las posiciones - 1 y +1. Usando otra numeración, desde el extremo 5' al extremo 3', las secuencias palindrómicas están en las posiciones 1 a 4 y 19 a 22, y 7 a 9 y 14 a 16, se rompe la simetría en los pares de bases 5, 6, 10, 11, 12, 13, 17 y 18, y el eje de simetría se sitúa entre los pares de bases de las posiciones 11 y 12.

- Como se usa en el presente documento, la expresión "I-CreI de tipo silvestre" designa una meganucleasa I-CreI que tiene la secuencia SWISSPROT P05725 o código de acceso pdb 1g9y.

10 - Las expresiones "sitio de reconocimiento", "secuencia de reconocimiento", "diana", "secuencia diana", "diana de ADN", "sitio de reconocimiento de movimiento dirigido", "sitio de movimiento dirigido", "sitio de escisión" se usan indiscriminadamente para designar una secuencia de 14 a 40 pb de doble cadena, palindrómica, no palindrómica o parcialmente palindrómica que se reconoce y se escinde mediante una meganucleasa. Estas expresiones se refieren a una localización de ADN específica, preferiblemente una localización cromosómica, en la que se va a inducir una rotura de doble cadena (escisión) mediante la meganucleasa.

Por ejemplo, el sitio de reconocimiento de movimiento dirigido de I-CreI de tipo silvestre se representa mediante la secuencia de 22 pb 5'- caaaacgtcgtgagacagtttg-3' (SEC ID N°: 71) o la secuencia de 24 pb 5'- tcaaaacgtcgtgagacagtttgg-3' presentada en la figura 2A (denominada aquí C1234, SEC ID N°: 65; gtc en las posiciones -5 a -3 y gac en las posiciones +3 a +5). Este sitio particular se denomina también en lo sucesivo "sitio diana natural de I-CreI". De la diana natural pueden derivarse dos secuencias palindrómicas mediante mutación de los nucleótidos en las posiciones +1,+2, +6 y +7 o -1,-2,-6 y -7: C1221 (SEC ID N°: 12) y C4334 (SEC ID N°:66), presentados en la figura 2A. Ambas tienen gtc en las posiciones -5 a -3 y gac en las posiciones +3 a +5, y se cortan mediante I-CreI, *in vitro* y en levadura.

25 - La expresión "especificidad modificada" se refiere a una variante de meganucleasa capaz de escindir un sitio de movimiento dirigido que no está escindido, en las mismas condiciones mediante la meganucleasa inicial (proteína de armazón) de la que deriva; dicha proteína inicial o de armazón puede ser la meganucleasa de tipo silvestre o un mutante de la misma.

De hecho, cuando se usa un ensayo *in vivo* en una cepa de levadura, los inventores han descubierto que la I-CreI de tipo silvestre no sólo escinde los sitios de movimiento dirigido en los que la secuencia palindrómica de las posiciones -5 a -3 es gtc (como en C1234, C1221 o C4334), sino también gcc, gac, ggc, ate, etc y ttc (figura 9a).

El mutante D75N de I-CreI (I-CreI N75) que también puede usarse como una proteína de armazón para realizar variantes con nueva especificidad, escinde no sólo los sitios de movimiento dirigido en los que la secuencia palindrómica de las posiciones -5 a -3 es gtc, sino también gcc, gtt, gtg, o gct (figura 8 y 9a).

- La forma heterodimérica puede obtenerse por ejemplo mediante el avance hacia la fusión de los dos monómeros. La meganucleasa heterodimérica resultante es capaz de escindir al menos un sitio diana que no se escinde mediante la forma homodimérica. Por lo tanto una variante de meganucleasa es aun parte

de la invención cuando se usa en forma heteromérica. El otro monómero seleccionado para la formación de la meganucleasa heterodimérica puede ser otro monómero variante, pero también puede ser un monómero de tipo silvestre, por ejemplo un monómero de I-*CreI* o un monómero de I-*DmoI*.

Así pues, los inventores construyeron una biblioteca de variantes de I-*CreI* a partir de una proteína de armazón de I-*CreI* (I-*CreI* D75N), presentando cada una de ellas al menos una mutación en los restos aminoácidos de las posiciones 44, 68 y/o 70 (código pdb 1g9y) y siendo capaz cada una de ellas de escindir al menos un sitio diana no escindido por la proteína de armazón de I-*CreI*.

En este enfoque particular, la mutación consiste en el reemplazo de al menos un resto de aminoácido en la posición 44, 68 y/o 70 por otro resto seleccionado del grupo que comprende A, D, E, G, H, K, N, P, Q, R, S, T e Y. Cada resto de aminoácido mutado se cambia independientemente de los otros restos y los restos de aminoácidos seleccionados pueden ser los mismos o pueden ser diferentes de los otros restos de aminoácidos en las posiciones 44, 68 y/o 70. En este enfoque, el sitio de movimiento dirigido, escindido por la variante de meganucleasa I-*CreI* de acuerdo con la invención pero que no se escinde por la proteína de armazón de I-*CreI*, es la misma que se ha descrito anteriormente y que se ilustra en la figura 2, excepto que la secuencia triplete de las posiciones -5 a -3 (correspondiente a R₃ en la fórmula I) y/o secuencia de triplete de las posiciones +3 a +5 (correspondientes a R'₃ de la fórmula I) difiere de la secuencia de triplete en las mismas posiciones en los sitios de movimiento dirigido escindidos por la proteína de armazón de I-*CreI*.

Inesperadamente, las variantes de meganucleasa I-*CreI*, obtenibles por el método descrito anteriormente, es decir con una “especificidad modificada” son capaces de escindir al menos una diana que difiere de la diana de proteína de armazón I-*CreI* en las posiciones -5 a -3 y/o en las posiciones +3 a +5. Debe observarse que dicha diana de ADN no es necesariamente palindrómica en las posiciones +/- 3 a 5. I-*CreI* está activa en su forma homodimérica, pero puede estar activa en una forma heterodimérica. Por lo tanto las variantes de I-*CreI* de acuerdo con la presente invención pueden estar activas no sólo en forma homodimérica, sino también en una forma heterodimérica y en ambos casos, podrían reconocer una diana tanto con secuencia palindrómica como no palindrómica en la posición +/-3 a 5, con tal que cuando la proteína I-*CreI* N75 se usa como armazón, el triplete en la posición -5 a -3 y/o +3 a +5 difiere de gtc, gcc, gtg, gtt y gct y de gac, ggc, cac, aac y agc, respectivamente. Puesto que cada monómero de la variante de I-*CreI* se une a una mitad del sitio de movimiento dirigido, una variante capaz de escindir una pluralidad de dianas también pueden escindir una diana cuya secuencia en la posición +/- 3 a 5 no es palindrómica. Adicionalmente, una variante puede actuar tanto en forma homodimérica como en una forma heterodimérica. La variante de I-*CreI* podría formar una meganucleasa heterodimérica, en la que la otra variante puede ser un monómero de I-*CreI* de tipo silvestre, otro monómero de meganucleasa de tipo silvestre, tal como I-*DmoI*, otro monómero variante de I-*CreI* o un monómero de una variante de otra meganucleasa que no sea I-*CreI*.

De acuerdo con una realización ventajosa de dicho método, la variante de meganucleasa I-*CreI* obtenida en el paso (b) se selecciona del grupo que consiste en: A44/A68/A70, A44/A68/G70, A44/A68/H70, A44/A68/K70, A44/A68/N70, A44/A68/Q70, A44/A68/R70, A44/A68/S70, A44/A68/T70,

A44/D68/H70, A44/D68/K70, A44/D68/R70, A44/G68/H70, A44/G68/K70, A44/G68/N70, 44/G68/P70,
 A44/G68/R70, A44/H68/A70, A44/H68/G70, A44/H68/H70, A44/H68/K70, A44/H68/N70,
 A44/H68/Q70, A44/H68/R70, A44/H68/S70, A44/H68/T70, A44/K68/A70, A44/K68/G70,
 44/K68/H70, A44/K68/K70, A44/K68/N70, A44/K68/Q70, A44/K68/R70, A44/K68/S70, A44/K68/T70,
 5 A44/N68/A70, A44/N68/E70, A44/N68/G70, A44/N68/H70, A44/N68/K70, A44/N68/N70, A44/N68/Q70,
 A44/N68/R70, A44/N68/S70, A44/N68/T70, A44/Q68/A70, A44/Q68/D70, A44/Q68/G70, 44/Q68/H70,
 A44/Q68/N70, A44/Q68/R70, A44/Q68/S70, A44/R68/A70, A44/R68/D70, A44/R68/E70, A44/R68/G70,
 A44/R68/H70, A44/R68/K70, A44/R68/L70, A44/R68/N70, A44/R68/R70, A44/R68/S70, A44/R68/T70,
 A44/S68/A70, A44/S68/G70, A44/S68/K70, A44/S68/N70, A44/S68/Q70, A44/S68/R70, A44/S68/S70,
 10 A44/S68/T70, A44/T68/A70, A44/T68/G70, A44/T68/H70, A44/T68/K70, A44/T68/N70, A44/T68/Q70,
 A44/T68/R70, A44/T68/S70, A44/T68/T70, D44/D68/H70, D44/N68/S70, D44/R68/A70, D44/R68/K70,
 D44/R68/N70, D44/R68/Q70, D44/R68/R70, D44/R68/S70, D44/R68/T70, E44/H68/H70, E44/R68/A70,
 E44/R68/H70, E44/R68/N70, E44/R68/S70, E44/R68/T70, E44/S68/T70, G44/H68/K70, G44/Q68/H70,
 G44/R68/Q70, G44/R68/R70, G44/T68/D70, G44/T68/P70, G44/T68/R70, H44/A68/S70, H44/A68/T70,
 15 H44/R68/A70, H44/R68/D70, H44/R68/E70, H44/R68/G70, H44/R68/N70, H44/R68/R70, H44/R68/S70,
 H44/R68/T70, H44/S68/G70, H44/S68/S70, H44/S68/T70, H44/T68/S70, H44/T68/T70, K44/A68/A70,
 K44/A68/D70, K44/A68/E70, K44/A68/G70, K44/A68/H70, K44/A68/N70, K44/A68/Q70, K44/A68/S70,
 K44/A68/T70, K44/D68/A70, K44/D68/T70, K44/E68/G70, K44/E68/N70, K44/E68/S70, K44/G68/A70,
 K44/G68/G70, K44/G68/N70, K44/G68/S70, K44/G68/T70, K44/H68/D70, K44/H68/E70, K44/H68/G70,
 20 K44/H68/N70, K44/H68/S70, K44/H68/T70, K44/K68/A70, K44/K68/D70, K44/K68/H70, K44/K68/T70,
 K44/N68/A70, K44/N68/D70, K44/N68/E70, K44/N68/G70, K44/N68/H70, K44/N68/N70,
 K44/N68/Q70, K44/N68/S70, K44/N68/T70, K44/P68/H70, K44/Q68/A70, K44/Q68/D70, K44/Q68/E70,
 K44/Q68/S70, K44/Q68/T70, K44/R68/A70, K44/R68/D70, K44/R68/E70, K44/R68/G70, K44/R68/H70,
 K44/R68/N70, K44/R68/Q70, K44/R68/S70, K44/R68/T70, K44/S68/A70, K44/S68/D70,
 25 K44/S68/H70, K44/S68/N70, K44/S68/S70, K44/S68/T70, K44/T68/A70, K44/T68/D70, K44/T68/E70,
 K44/T68/G70, K44/T68/H70, K44/T68/N70, K44/T68/Q70, K44/T68/S70, K44/T68/T70, N44/A68/H70,
 N44/A68/R70, N44/H68/N70, N44/H68/R70, N44/K68/G70, N44/K68/H70, N44/K68/R70, N44/K68/S70,
 N44/N68/R70, N44/P68/D70, N44/Q68/H70, N44/Q68/R70, N44/R68/A70, N44/R68/D70, N44/R68/E70,
 N44/R68/G70, N44/R68/H70, N44/R68/K70, N44/R68/N70,
 30 N44/R68/R70, N44/R68/S70, N44/R68/T70, N44/S68/G70, N44/S68/H70, N44/S68/K70, N44/S68/R70,
 N44/T68/H70, N44/T68/K70, N44/T68/Q70, N44/T68/R70, N44/T68/S70, P44/N68/D70, P44/T68/T70,
 Q44/A68/A70, Q44/A68/H70, Q44/A68/R70, Q44/G68/K70, Q44/G68/R70, Q44/K68/G70,
 Q44/N68/A70, Q44/N68/H70, Q44/N68/S70, Q44/P68/P70, Q44/Q68/G70, Q44/R68/A70, Q44/R68/D70,
 Q44/R68/E70, Q44/R68/G70, Q44/R68/H70, Q44/R68/N70, Q44/R68/Q70, Q44/R68/S70, Q44/S68/H70,
 35 Q44/S68/R70, Q44/S68/S70, Q44/T68/A70, Q44/T68/G70, Q44/T68/H70,
 Q44/T68/R70, R44/A68/G70, R44/A68/T70, R44/G68/G70, R44/H68/D70, R44/H68/T70, R44/N68/T70,
 R44/R68/A70, R44/R68/D70, R44/R68/E70, R44/R68/G70, R44/R68/N70, R44/R68/Q70, R44/R68/S70,
 R44/R68/T70, R44/S68/G70, R44/S68/N70, R44/S68/S70, R44/S68/T70, S44/D68/K70, S44/H68/R70,

S44/R68/G70, S44/R68/N70, S44/R68/R70, S44/R68/S70, T44/A68/K70, T44/A68/R70, T44/H68/R70, T44/K68/R70, T44/N68/P70, T44/N68/R70, T44/Q68/K70, T44/Q68/R70, T44/R68/A70, T44/R68/D70, T44/R68/E70, T44/R68/G70, T44/R68/H70, T44/R68/K70, T44/R68/N70, T44/R68/Q70, T44/R68/R70, T44/R68/S70, T44/R68/T70, T44/S68/K70, T44/S68/R70, T44/T68/K70 y T44/T68/R70.

5 De acuerdo con otra realización ventajosa de dicho método, el paso (b) de seleccionar dicha variante de meganucleasa I-*CreI* se realiza *in vivo* en células de levadura.

La materia objeto de la presente invención es también el uso de una variante de meganucleasa I-*CreI* como se ha definido anteriormente en el presente documento, es decir obtenible por el método que se ha descrito anteriormente, *in vitro* o *in vivo* para propósitos no terapéuticos, para la
 10 escisión de una diana de ácido nucleico bicatenario que comprende al menos una secuencia parcialmente palindrómica de 20-24 pb, en la que al menos la secuencia de las posiciones +/- 8 a 11 es palindrómica, y el triplete de nucleótidos de las posiciones -5 a -3 y/o el triplete de nucleótidos de las posiciones +3 a +5 difiere de gtc, gcc, gtg, gtt y gct, y de gac, ggc, cac, aac y agc, respectivamente. La fórmula I describe una diana tal de ADN.

15 De acuerdo con una realización ventajosa de dicho uso, dicha variante de meganucleasa I-*CreI* se selecciona del grupo que consiste en: A44/A68/A70, A44/A68/G70, A44/A68/H70, A44/A68/K70, A44/A68/N70, A44/A68/Q70, A44/A68/R70, A44/A68/S70, A44/A68/T70, A44/D68/H70, A44/D68/K70, A44/D68/R70, A44/G68/H70, A44/G68/K70, A44/G68/N70, A44/G68/P70, A44/G68/R70, A44/H68/A70, A44/H68/G70, A44/H68/H70, A44/H68/K70, A44/H68/N70, A44/H68/Q70, A44/H68/R70,
 20 A44/H68/S70, A44/H68/T70, A44/K68/A70, A44/K68/G70, A44/K68/H70, A44/K68/K70, A44/K68/N70, A44/K68/Q70, A44/K68/R70, A44/K68/S70, A44/K68/T70, A44/N68/A70, A44/N68/E70, A44/N68/G70, A44/N68/H70, A44/N68/K70, A44/N68/N70, A44/N68/Q70, A44/N68/R70, A44/N68/S70, A44/N68/T70, A44/Q68/A70, A44/Q68/D70, A44/Q68/G70, A44/Q68/H70, A44/Q68/N70, A44/Q68/R70, A44/Q68/S70, A44/R68/A70, A44/R68/D70, A44/R68/E70, A44/R68/G70, A44/R68/H70, A44/R68/K70,
 25 A44/R68/L70, A44/R68/N70, A44/R68/R70, A44/R68/S70, A44/R68/T70, A44/S68/A70, A44/S68/G70, A44/S68/K70, A44/S68/N70, A44/S68/Q70, A44/S68/R70, A44/S68/S70, A44/S68/T70, A44/T68/A70, A44/T68/G70, A44/T68/H70, A44/T68/K70, A44/T68/N70, A44/T68/Q70, A44/T68/R70, A44/T68/S70, A44/T68/T70, D44/D68/H70, D44/N68/S70, D44/R68/A70, D44/R68/K70, D44/R68/N70, D44/R68/Q70, D44/R68/R70, D44/R68/S70, D44/R68/T70, E44/H68/H70, E44/R68/A70,
 30 E44/R68/H70, E44/R68/N70, E44/R68/S70, E44/R68/T70, E44/S68/T70, G44/H68/K70, G44/Q68/H70, G44/R68/Q70, G44/R68/R70, G44/T68/D70, G44/T68/P70, G44/T68/R70, H44/A68/S70, H44/A68/T70, H44/R68/A70, H44/R68/D70, H44/R68/E70, H44/R68/G70, H44/R68/N70, H44/R68/R70, H44/R68/S70, H44/R68/T70, H44/S68/G70, H44/S68/S70, H44/S68/T70, H44/T68/S70, H44/T68/T70, K44/A68/A70, K44/A68/D70, K44/A68/E70, K44/A68/G70, K44/A68/H70, K44/A68/N70,
 35 K44/A68/Q70, K44/A68/S70, K44/A68/T70, K44/D68/A70, K44/D68/T70, K44/E68/G70, K44/E68/N70, K44/E68/S70, K44/G68/A70, K44/G68/G70, K44/G68/N70, K44/G68/S70, K44/G68/T70, K44/H68/D70, K44/H68/E70, K44/H68/G70, K44/H68/N70, K44/H68/S70,

- K44/H68/T70, K44/K68/A70, K44/K68/D70, K44/K68/H70, K44/K68/T70, K44/N68/A70, K44/N68/D70, K44/N68/E70, K44/N68/G70, K44/N68/H70, K44/N68/N70, K44/N68/Q70, K44/N68/S70, K44/N68/T70, K44/P68/H70, K44/Q68/A70, K44/Q68/D70, K44/Q68/E70, K44/Q68/S70, K44/Q68/T70, K44/R68/A70, K44/R68/D70, K44/R68/E70, K44/R68/G70,
- 5 K44/R68/H70, K44/R68/N70, K44/R68/Q70, K44/R68/S70, K44/R68/T70, K44/S68/A70, K44/S68/D70, K44/S68/H70, K44/S68/N70, K44/S68/S70, K44/S68/T70, K44/T68/A70, K44PT68/D70, K44/T68/E70, K44/T68/G70, K44/T68/H70, K44/T68/N70, K44/T68/Q70, K44/T68/S70, K44/T68/T70, N44/A68/H70, N44/A68/R70, N44/H68/N70, N44/H68/R70, N44/K68/G70, N44/K68/H70, N44/K68/R70, N44/K68/S70, N44/N68/R70, N44/P68/D70, N44/Q68/H70, N44/Q68/R70,
- 10 N44/R68/A70, N44/R68/D70, N44/R68/E70, N44/R68/G70, N44/R68/H70, N44/R68/K70, N44/R68/N70, N44/R68/R70, N44/R68/S70, N44/R68/T70, N44/S68/G70, N44/S68/H70, N44/S68/K70, N44/S68/R70, N44/T68/H70, N44/T68/K70, N44/T68/Q70, N44/T68/R70, N44/T68/S70, P44/N68/D70, P44/T68/T70, Q44/A68/A70, Q44/A68/H70, Q44/A68/R70, Q44/G68/K70, Q44/G68/R70, Q44/K68/G70, Q44/N68/A70, Q44/N68/H70, Q44/N68/S70, Q44/P68/P70, Q44/Q68/G70, Q44/R68/A70, Q44/R68/D70,
- 15 Q44/R68/E70, Q44/R68/G70, Q44/R68/H70, Q44/R68/N70, Q44/R68/Q70, Q44/R68/S70, Q44/S68/H70, Q44/S68/R70, Q44/S68/S70, Q44/T68/A70, Q44/T68/G70, Q44/T68/H70, Q44/T68/R70, R44/A68/G70, R44/A68/T70, R44/G68/T70, R44/H68/D70, R44/H68/T70, R44/N68/T70, R44/R68/A70, R44/R68/D70, R44/R68/E70, R44/R68/G70, R44/R68/N70, R44/R68/Q70, R44/R68/S70, R44/R68/T70, R44/S68/G70, R44/S68/N70, R44/S68/S70, R44/S68/T70, S44/D68/K70,
- 20 S44/H68/R70, S44/R68/G70, S44/R68/N70, S44/R68/R70, S44/R68/S70, T44/A68/K70, T44/A68/R70, T44/H68/R70, T44/K68/R70, T44/N68/P70, T44/N68/R70, T44/Q68/K70, T44/Q68/R70, T44/R68/A70, T44/R68/D70, T44/R68/E70, T44/R68/G70, T44/R68/H70, T44/R68/K70, T44/R68/N70, T44/R68/Q70, T44/R68/R70, T44/R68/S70, T44/R68/T70, T44/S68/K70, T44/S68/R70, T44/T68/K70, y T44/T68/R70.

De acuerdo con otra realización ventajosa de dicho uso, la variante de meganucleasa I-*CreI* es un homodímero.

De acuerdo con otra realización ventajosa de dicho uso, dicha variante de meganucleasa I-*CreI* es un heterodímero.

Dicho heterodímero puede ser una molécula quimérica monocatenaria que consiste en la fusión de dos variantes de I-*CreI* diferentes como se ha definido en la presente invención o de una proteína de armazón I-*CreI* con una variante de I-*CreI* como se ha definido en la presente invención. Como alternativa, dicho heterodímero puede consistir en dos monómeros separados seleccionados de dos variantes de I-*CreI* diferentes como se ha definido en la presente invención o una proteína de armazón I-*CreI* y una variante de I-*CreI* como se ha definido en la presente invención.

De acuerdo con dicho uso:

- 35 - o bien la variante de meganucleasa I-*CreI* es capaz de escindir una diana de ADN cuya secuencia en las posiciones +/- 3 a 5 es palindrómica,
- o bien dicha variante de meganucleasa I-*CreI* es capaz de escindir una diana de ADN cuya secuencia en las posiciones +/- 3 a 5 no es palindrómica

De acuerdo con otra realización ventajosa de dicho uso la diana de ácido nucleico escindida es una diana de ADN en la que las secuencias palindrómicas en las posiciones -11 a -8 y +8 a +11 son caaa y ttgt, respectivamente.

De acuerdo con otra realización ventajosa de dicho uso, dicha variante de meganucleasa I-*CreI* comprende adicionalmente una mutación en la posición 75, preferiblemente una mutación en un aminoácido no cargado, más preferiblemente una asparagina o una valina (D75N o D75V).

De acuerdo con otra realización ventajosa más de dicho uso, dicha variante de meganucleasa I-*CreI* tienen una alanina (A) o una asparagina (N) en la posición 44, para escindir una diana de ADN que comprende un nucleótido a en la posición -4 y/o t en la posición +4.

De acuerdo con otra realización ventajosa más de dicho uso, dicha variante meganucleasa I-*CreI* tiene una glutamina (Q) en la posición 44, para escindir una diana de ADN que comprende un nucleótido t en la posición -4 o a en la posición +4.

De acuerdo con otra realización ventajosa más de dicho uso, dicha variante de meganucleasa I-*CreI* tiene una lisina (K) en la posición 44, para escindir una diana que comprende un nucleótido c en la posición -4 y/o g en la posición +4.

La materia objeto de la presente invención también son las variantes de meganucleasa I-*CreI*:

- obtenibles por el método de preparación como se ha definido anteriormente;
- que tienen una mutación de al menos uno de los restos de aminoácidos en las posiciones 44, 68 y 70 de I-*CreI*; dichas mutaciones pueden ser las únicas dentro de los aminoácidos que contactan directamente con la diana de ADN; y
- Que tienen una especificidad de escisión modificada en las posiciones ± 3 a 5.

Dichas meganucleasas I-*CreI* nuevas pueden usarse como endonucleasas muy específicas en digestión *in vitro*, para su uso en restricción o mapeo, tanto *in vivo* como *ex vivo* como herramientas para ingeniería genómica. Además, cada una puede usarse como un nuevo armazón para un segundo ciclo de mutagénesis y selección/exploración, con el propósito de realizar nuevas endonucleasas de movimiento dirigido de segunda generación.

Las variantes de meganucleasa I-*CreI* de acuerdo con la invención se mutan solamente en las posiciones 44, 68 y/o 70 del dominio de unión a ADN. Sin embargo, la presente invención también incluye diferentes proteínas capaces de formar heterodímeros: la heterodimerización de dos diferentes proteínas de la lista anterior da como resultado también una escisión de secuencias no palindrómicas, hechas de dos mitades de los sitios escindidos mediante las proteínas parentales solamente. Esto puede obtenerse *in vitro* mediante la adición de dos variantes de I-*CreI* diferentes en el tampón de reacción e *in vivo* o *ex vivo* mediante co-expresión. Otra posibilidad es construir una molécula monocatenaria, como se describe en Epinat *et al.* (Epinat *et al.*, 2003). Esta molécula monocatenaria sería la fusión de dos variantes de I-*CreI* diferentes y también debería dar como resultado la escisión de secuencias quiméricas no palindrómicas.

De acuerdo con una realización ventajosa de dicha variante de meganucleasa I-CreI, el resto de aminoácido seleccionado para el reemplazo del aminoácido en las posiciones 44, 68 y/o 70 se selecciona del grupo que comprende A, D, E, G, H, K, M, P, Q, R, S, T e Y.

De acuerdo con otra realización ventajosa, dicha variante de meganucleasa I-CreI se
 5 selecciona del grupo que consiste en: A44/A68/A70, A44/A68/G70, A44/A68/H70, A44/A68/K70, A44/A68/N70, A44/A68/Q70, A44/A68/S70, A44/A68/T70, A44/D68/H70, A44/D68/K70, A44/D68/R70, A44/G68/H70, A44/G68/K70, A44/G68/N70, A44/G68/P70, A44/H68/A70, A44/H68/G70, A44/H68/H70, A44/H68/K70, A44/H68/N70, A44/H68/Q70, A44/H68/S70, A44/H68/T70, A44/K68/A70, A44/K68/G70, A44/K68/H70, A44/K68/N70, A44/K68/Q70, A44/K68/R70, A44/K68/S70, A44/K68/T70,
 10 A44/N68/A70, A44/N68/E70, A44/N68/G70, A44/N68/H70, A44/N68/K70, A44/N68/N70, A44/N68/Q70, A44/N68/R70, A44/N68/S70, A44/N68/T70, A44/Q68/A70, A44/Q68/D70, A44/Q68/G70, A44/Q68/H70, A44/Q68/N70, A44/Q68/S70, A44/R68/E70, A44/R68/K70, A44/R68/L70, A44/S68/A70, A44/S68/G70, A44/S68/N70, A44/S68/Q70, A44/S68/R70, A44/S68/S70, A44/S68/T70, A44/T68/A70, A44/T68/G70, A44/T68/H70, A44/T68/N70, A44/T68/Q70, A44/T68/S70, A44/T68/T70, D44/D68/H70,
 15 D44/N68/S70, D44/R68/A70, D44/R68/N70, D44/R68/Q70, D44/R68/R70, D44/R68/S70, D44/R68/T70, E44/H68/H70, E44/R68/A70, E44/R68/H70, E44/R68/N70, E44/R68/S70, E44/R68/T70, E44/S68/T70, G44/H68/K70, G44/Q68/H70, G44/R68/Q70, G44/T68/D70, G44/T68/P70, G44/T68/R70, H44/A68/S70, H44/A68/T70, H44/R68/D70, H44/R68/E70, H44/R68/G70, H44/R68/N70, H44/R68/R70, H44/R68/S70, H44/S68/G70, H44/S68/S70, H44/S68/T70, H44/T68/S70, H44/T68/T70, K44/A68/A70, K44/A68/D70,
 20 K44/A68/E70, K44/A68/G70, K44/A68/H70, K44/A68/N70, K44/A68/Q70, K44/D68/A70, K44/D68/T70, K44/E68/G70, K44/E68/S70, K44/G68/A70, K44/G68/G70, K44/G68/N70, K44/G68/S70, K44/G68/T70, K44/H68/D70, K44/H68/E70, K44/H68/G70, K44/H68/N70, K44/H68/S70, K44/H68/T70, K44/K68/A70, K44/K68/D70, K44/K68/H70, K44/K68/T70, K44/N68/A70, K44/N68/D70, K44/N68/E70, K44/N68/G70, K44/N68/H70, K44/N68/N70, K44/N68/Q70, K44/N68/S70, K44/N68/T70,
 25 K44/P68/H70, K44/Q68/A70, K44/Q68/D70, K44/Q68/E70, K44/Q68/S70, K44/Q68/T70, K44/R68/A70, K44/R68/D70, K44/R68/E70, K44/R68/G70, K44/R68/H70, K44/R68/N70, K44/R68/S70, K44/S68/A70, K44/S68/D70, K44/S68/H70, K44/S68/N70, K44/S68/S70, K44/S68/T70, K44/T68/A70, K44/T68/D70, K44/T68/E70, K44/T68/G70, K44/T68/H70, K44/T68/N70, K44/T68/Q70, K44/T68/S70, K44/T68/T70, N44/A68/H70, N44/H68/N70, N44/H68/R70, N44/K68/G70,
 30 N44/K68/H70, N44/K68/R70, N44/K68/S70, N44/P68/D70, N44/Q68/H70, N44/R68/A70, N44/R68/D70, N44/R68/E70, N44/R68/K70, N44/S68/G70, N44/S68/H70, N44/S68/K70, N44/S68/R70, N44/T68/H70, N44/T68/K70, N44/T68/Q70, N44/T68/S70, P44/N68/D70, P44/T68/T70, Q44/G68/K70, Q44/G68/R70, Q44/K68/G70, Q44/N68/A70, Q44/N68/H70, Q44/N68/S70, Q44/P68/P70, Q44/Q68/G70, Q44/R68/D70, Q44/R68/E70, Q44/R68/G70, Q44/R68/Q70, Q44/S68/S70, Q44/T68/A70, Q44/T68/G70, Q44/T68/H70,
 35 R44/A68/G70, R44/A68/T70, R44/G68/T70, R44/H68/D70, R44/H68/T70, R44/N68/T70, R44/R68/A70, R44/R68/D70, R44/R68/E70, R44/R68/G70, R44/R68/Q70, R44/R68/S70, R44/R68/T70, R44/S68/G70, R44/S68/N70, R44/S68/S70, R44/S68/T70, S44/D68/K70, S44/R68/R70, S44/R68/S70, T44/A68/K70, T44/N68/P70, T44/N68/R70, T44/R68/E70, T44/R68/Q70 y T44/S68/K70; dicha variante de

meganucleasa I-CreI es capaz de escindir al menos una diana, como se ha definido anteriormente, que no se escinde mediante la proteína de armazón de I-CreI N75.

De acuerdo con otra realización ventajosa más la variante de meganucleasa I-CreI tiene una alanina (A) o una asparagina (N), en la posición 44 y escinde una diana que comprende el nucleótido a en una posición -4 y/o t en una posición +4, con la exclusión de las variantes presentadas en la Tabla 4 y Tabla 5 de la Solicitud Internacional PCT WO 2004/067736, preferiblemente dicha variante tiene una alanina o una asparagina.

De acuerdo con otra realización ventajosa más, la variante de meganucleasa I-CreI tiene glutamina (Q) y escinde una diana que comprende el nucleótido t en una posición -4 y/o a en una posición +4 en la posición 44, con la exclusión de las variantes presentadas en la Tabla 3, Tabla 4 y Tabla 5 de la Solicitud Internacional PCT WO 2004/067736.

De acuerdo con otra realización ventajosa más, la variante de meganucleasa I-CreI de la invención tiene una lisina (K) en la posición 44 y escinde una diana que comprende c en una posición -4 y/o g en una posición +4, con la exclusión de la variante presentada en la Tabla 5 de la Solicitud Internacional PCT WO 2004/067736.

Como se ha especificado anteriormente en el presente documento, en el marco de la definición de la variante de meganucleasa I-CreI en la solicitud en uso, dicha variante de meganucleasa I-CreI puede ser un homodímero o un heterodímero. Puede ser capaz de escindir una diana de ADN palindrómica o no palindrómica. Puede comprender adicionalmente una mutación en la posición 75, como se ha especificado anteriormente en el presente documento.

La materia objeto de la presente invención también es un polinucleótido, **caracterizado por que** codifica una variante de meganucleasa I-CreI de acuerdo con la invención.

Adicionalmente, la materia objeto de la presente invención es un casete de expresión que comprende dicho polinucleótido y secuencias de regulación tal como un promotor y un vector de expresión que comprende dicho casete de expresión. Cuando dicha variante es un heterodímero que consiste en dos diferentes monómeros, cada monómero puede expresarse de un vector único (vector de expresión dual) o de dos vectores diferentes.

La materia objeto de la presente invención también es un vector de expresión, como se ha descrito anteriormente, que comprende adicionalmente una construcción de ADN diana.

El término “vector” se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha unido. Un tipo de vector preferido es un episoma, es decir, un ácido nucleico capaz de replicación extra-cromosómica. Son vectores preferidos los capaces de replicación autónoma y/o expresión de ácidos nucleicos a los que están unidos. Los vectores capaces de dirigir la expresión de genes a los que están unidos operativamente se conocen en el presente documento como “vectores de expresión”. Un vector de acuerdo con la presente invención comprende, pero no se limita a, un YAC (cromosoma artificial de levadura), un BAC (artificial de bacteria), un vector de báculo virus, un fago, un fagémido, un cósmido, un vector viral, un plásmido, un vector de ARN o una molécula de ADN o ARN lineal o circular que puede consistir en ADN cromosómico, no cromosómico, semi-sintético o sintético. En general, los

vectores de expresión de utilidad en técnicas de ADN recombinante están con frecuencia en forma de “plásmidos” que generalmente se refiere a bucles de ADN bicatenario circular que, en su forma de vector no están unidos al cromosoma. Los expertos en la materia conocen un gran número de vectores adecuados y disponibles en el mercado, tal como los siguientes vectores bacterianos: pQE7O, pQE6O, pQE-9 (Qiagen), pbs, pDIO, phagescript, psiX174. Pbluescript SK, pbsks, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); pWLNEO, pSV2CAT, pOG44, pXTI, pSG (Stratagene); pSVK3, pBPV, pMSG, pSVL (Pharmacia); pQE-30 (QIAexpress), pET (Novagen).

Los vectores virales incluyen retrovirus, adenovirus, parvovirus (por ejemplo virus adenosociados), coronavirus, virus de ARN de cadena negativa tales como ortomixovirus (por ejemplo, virus de la gripe), rabdovirus (por ejemplo, virus de la rabia y estomatitis vesicular), para-mixovirus (por ejemplo, sarampión y Sendai), virus de ARN de cadena positiva tales como picornavirus y alfavirus y virus de ADN bicatenario incluyendo adenovirus, herpesvirus (por ejemplo, virus Herpes Simplex tipos 1 y 2, virus de Epstein-Barr, citomegalovirus) y poxvirus (por ejemplo, vaccinia, viruela aviar y viruela del canario). Otros virus incluyen el virus Norwalk, togavirus, flavivirus, reovirus, papovavirus, hepadnavirus y virus de la hepatitis, por ejemplo.

Los vectores pueden comprender marcadores seleccionables, por ejemplo: neomicin fosfotransferasa, histidinol deshidrogenasa, dihidrofolato reductasa, higromicin fosfotransferasa, timidina quinasa de virus de herpes simplex, adenosina desaminasa, glutamina sintetasa e hipoxantina-guanina fosforibosil transferasa para cultivo de células eucariotas; TRP1 para *S. cerevisiae*; resistencia a tetraciclina, rifampicina o ampicilina en *E. coli*.

Preferiblemente dichos vectores son vectores de expresión, en los que las secuencias que codifican los polipéptidos de la invención se sitúan bajo el control de elementos de control transcripcional y traduccional apropiados para permitir la producción de síntesis de dichos polipéptidos. Por lo tanto, dichos polinucleótidos están comprendidos en un casete o casetes de expresión. Mas particularmente, el vector comprende un origen de replicación, un promotor unido operativamente a dicho polinucleótido codificante, un sitio de unión a ribosoma, un sitio de corte y empalme de ARN (cuando se usa ADN genómico), un sitio de poliadenilación y un sitio de terminación de la transcripción. También puede comprender un potenciador. La selección del promotor dependerá de la célula en la que se expresa el polipéptido.

De acuerdo con una realización ventajosa de dicho vector de expresión, dicha construcción de ADN diana comprende una secuencia que comparte homologías con la región circundante del sitio de escisión de la variante de meganucleasa I-CreI de la invención.

De acuerdo con una realización ventajosa de dicho vector de expresión, dicha construcción de ADN diana comprende:

- a) secuencias que comparten homologías con la región circundante del sitio de escisión de la variante de meganucleasa I-CreI de acuerdo con la reivindicación, y
- b) secuencias a introducir flanqueadas por la secuencia de a).

La materia objeto de la presente invención también es una célula, **caracterizada por que** está modificada por un polinucleótido como se ha definido anteriormente o por un vector como se ha definido anteriormente.

5 La materia objeto de la presente invención también es una planta transgénica, **caracterizada por que** comprende un polinucleótido como se ha definido anteriormente o un vector como se ha definido anteriormente.

La materia objeto de la presente invención también es un mamífero transgénico no humano, **caracterizado por que** comprende un polinucleótido como se ha definido anteriormente o un vector como se ha definido anteriormente.

10 Las secuencias polinucleotídicas que codifican los polipéptidos como se han definido en la presente invención pueden prepararse por cualquier método conocido por el experto en la materia. Por ejemplo, se amplifican a partir de un molde de ADNc, mediante la reacción en cadena de la polimerasa con cebadores específicos. Preferiblemente los codones de dicho ADNc se dirigen para favorecer la expresión de dicha proteína en el sistema de expresión deseado.

15 El vector recombinante que comprende dichos polinucleótidos puede obtenerse e introducirse en una célula hospedadora mediante las técnicas bien conocidas de ingeniería genética y ADN recombinante.

20 La meganucleasa heterodimérica de la invención se produce mediante la expresión de dos polipéptidos como se ha definido anteriormente; preferiblemente dichos polipéptidos están co-expresados en una célula hospedadora modificada por dos vectores de expresión, comprendiendo cada uno un fragmento polinucleotídico que codifica un polipéptido diferente como se ha definido anteriormente o mediante un vector de expresión dual que comprende ambos fragmentos polinucleotídicos como se ha definido anteriormente, en condiciones adecuadas para la co-expresión de los polipéptidos y la meganucleasa heterodimérica se recupera del cultivo celular hospedador.

25 La materia objeto de la presente invención es el uso adicional de una variante de meganucleasa I-CreI, uno o dos polinucleótido o polinucleótidos, preferiblemente ambos incluidos en un vector de expresión (vector de expresión dual) o incluido cada uno en un vector de expresión diferente, una célula, una planta transgénica, un mamífero transgénico no humano, como se ha definido anteriormente, para biología molecular, para ingeniería genética *in vivo* o *in vitro* y para ingeniería genómica *in vivo* o *in vitro*, para propósitos no terapéuticos.

30 Los propósitos no terapéuticos incluyen por ejemplo (i) fijación de objetivos génicos de locus específicos en líneas de empaquetamiento celular para producción proteínica, (ii) fijación de objetivos génicos de locus específicos en plantas de cultivo, para mejoras en cepas e ingeniería metabólica, (iii) recombinación dirigida para la retirada de marcadores en plantas de cultivo modificadas genéticamente, (iv) recombinación dirigida para la retirada de marcadores en cepas de microorganismos
35 modificados genéticamente (para producción de antibióticos por ejemplo).

De acuerdo con una realización ventajosa de dicho uso, éste es para inducir una rotura de doble cadena en un sitio de interés que comprende una secuencia diana de ADN, induciendo de este modo un suceso de recombinación de ADN, una pérdida de ADN o muerte celular.

De acuerdo con la invención, dicha rotura de doble cadena es para: reparar una secuencia
5 específica, modificar una secuencia específica, restaurar un gen funcional en lugar de uno mutado, atenuar o activar un gen endógeno de interés, introducir una mutación en un sitio de interés, introducir un gen exógeno o una parte del mismo, inactivar o suprimir un gen endógeno o una parte del mismo, traslocar un brazo cromosómico o dejar el ADN sin reparar y degradado.

De acuerdo con otra realización ventajosa de dicho uso, dicha variante de meganucleasa I-
10 *CreI*, polinucleótido, vector, célula, planta transgénica o mamífero transgénico no humano se asocia con una construcción de ADN diana como se ha definido anteriormente.

La materia objeto de la presente invención también es un método de ingeniería genética, **caracterizado por que** comprende un paso de rotura de ácido nucleico de doble cadena en un sitio de interés localizado en un vector, que comprende una diana de ADN de una variante de meganucleasa I-*CreI*
15 como se ha definido anteriormente en el presente documento, mediante el contacto de dicho vector con una variante de meganucleasa I-*CreI* como se ha definido anteriormente, induciendo de este modo una recombinación homóloga con otro vector que representa homología con la secuencia que rodea el sitio de escisión de dicha variante de meganucleasa I-*CreI*.

La materia objeto de la presente invención también es un método de ingeniería genómica, **caracterizado por que** comprende los siguientes pasos: 1) rotura de doble cadena de un locus genómico
20 que comprende al menos un sitio de reconocimiento y escisión de una variante de meganucleasa I-*CreI* como se ha definido anteriormente, mediante el contacto de dicho sitio de escisión con dicha variante de meganucleasa I-*CreI*; 2) mantener dicho locus genómico roto en condiciones apropiadas para la recombinación homóloga con una construcción de ADN diana que comprende la secuencia a introducir en dicho locus, flanqueada por secuencias que comparten homología con el locus diana.
25

La materia objeto de la presente invención también es un método de ingeniería genómica, **caracterizado por que** comprende los siguientes pasos: 1) rotura de doble cadena de un locus genómico que comprende al menos un sitio de reconocimiento y escisión de una variante de meganucleasa I-*CreI* como se ha definido anteriormente, mediante el contacto de dicho sitio de escisión con dicha variante de
30 meganucleasa I-*CreI*; 2) mantener dicho locus genómico roto en condiciones apropiadas para la recombinación homóloga con ADN cromosómico que comparte homología con regiones circundantes del sitio de escisión.

La materia objeto de la presente invención también es una composición **caracterizada por que** comprende al menos una variante de meganucleasa I-*CreI*, un polinucleótido o un vector como se ha
35 definido anteriormente.

En una realización preferida de dicha composición, ésta comprende una construcción de ADN diana que comprende la secuencia que repara el sitio de interés flanqueado por secuencias que comparten homología con el locus diana.

La materia objeto de la presente invención también es el uso de al menos una variante de meganucleasa I-CreI, un polinucleótido o un vector, como se ha definido para la preparación de un medicamento para la prevención, mejora o cura de una enfermedad genética en un individuo que lo necesite, administrándose dicho medicamento por cualquier medio a dicho individuo.

5 La materia objeto de la presente invención también es el uso de al menos una variante de meganucleasa I-CreI, un polinucleótido o un vector como se ha definido para la preparación de un medicamento para la prevención, mejora o cura de una enfermedad causada por un agente infeccioso que presenta un intermedio de ADN, en un individuo que lo necesite, administrándose dicho medicamento por cualquier medio a dicho individuo.

10 La materia objeto de la presente invención también es el uso de al menos una variante de meganucleasa I-CreI, un polinucleótido o un vector, como se ha definido anteriormente, *in vitro*, para inhibir la propagación, inactivar o suprimir un agente infeccioso que presenta un intermedio de ADN, en productos derivados biológicos o productos que se pretenden para usos biológicos o para desinfectar un objeto.

15 En una realización particular, dicho agente infeccioso es un virus.

Además de las características precedentes, la invención comprende además otras características que surgirán de la descripción a continuación, que se refiere a ejemplos que ilustran las variantes de meganucleasa I-CreI y sus usos de acuerdo con la invención, así como los dibujos anexos en los que:

20 - La figura 1 ilustra el fundamento de los experimentos. (a) Estructura de I-CreI unida a su diana de ADN. (b) Ampliación de la estructura que muestra los restos 44, 68, 70 elegidos para selección aleatoria, D75 y pares de bases que interactúan. (c) diseño de la biblioteca y las dianas. Las interacciones de los restos de I-CreI Q44, R68 y R70 con dianas de ADN están indicadas (parte superior). Otros restos de aminoácidos que interactúan directa o indirectamente con la diana de ADN no se muestran. El resto de arginina (R) en la posición 44 de un monómero I-CreI directamente interactúa con la guanina de la posición -5 de la secuencia diana, mientras que el resto de glutamina (Q) de la posición 44 y el resto de arginina (R) de la posición 70 interactúa directamente con adenina en la posición +4 y guanina en la posición +3 de la cadena complementaria, respectivamente. La diana descrita aquí (C1221, SEC ID N°: 12) es un palíndromo derivado de la diana natural de I-CreI (C1234, SEC ID N°:65) y escindido mediante I-CreI (Chevalier *et al.*, 2003, citado anteriormente). Las posiciones de escisión se indican mediante flechas. En la biblioteca, los restos 44, 68 y 70 se reemplazan con ADEGHKPNRST. Puesto que I-CreI es un homodímero, la biblioteca se exploró con dianas palindrómicas. Se generaron 64 dianas palindrómicas resultantes de sustituciones en las posiciones ± 3 , ± 4 y ± 5 . Se muestran varios ejemplos de dichas dianas (parte inferior; SEC ID N°: 1 a 7).

35 - La figura 2 ilustra la diana usada en el estudio. A. Dos dianas palindrómicas derivadas de la diana natural de I-CreI (denomina aquí C1234, SEC ID N°: 65). La diana natural de I-CreI contiene dos palíndromos, marcados en cajas grises: los nucleótidos -8 a -12 y +8 a +12 por un lado, y los nucleótidos -5 a -3 y +3 a +5 por otro. La línea punteada vertical, de la que están numeradas las bases nucleotídicas,

representa el eje de simetría para las secuencias palindrómicas. Desde la diana natural pueden derivarse dos secuencias palindrómicas, C1221 (SEC ID N°: 12) y C4334 (SEC ID N°:66). Ambas se cortan mediante I-CreI, *in vitro* y en levaduras. Solamente una cadena de cada sitio diana se muestra. B. Las 64 dianas. Las 64 dianas (SEC ID N°: 1 a 64) se derivan de C1221 (SEC ID N°: 12) un palíndromo derivado de la diana natural de I-CreI (C1234, SEC ID N°:65) y escindido por I-CreI (Chevalier *et al.*, 2003, citado anteriormente). Corresponden a todos los palíndromos de 24 pares de bases resultantes de las sustituciones en las posiciones -5, -4, -3, +3, +4 y +5.

- La figura 3 ilustra la exploración de las variantes. (a) Se transforman levaduras con el vector de expresión de meganucleasa, marcado con el gen *LEU2* y se cruzan individualmente con levadura transformada con el plásmido informador, marcado mediante el gen *TRP1*. En el plásmido informador, un gen informador LacZ está interrumpido con un inserto que contiene el sitio de interés, flanqueado por dos repeticiones directas. En diploides (*LEU2 TRP1*), la escisión del sitio diana mediante la meganucleasa (óvalo blanco) induce una recombinación homóloga entre las dos repeticiones LacZ, dando como resultado un gen de beta galactosidasa funcional (óvalo gris), que puede controlarse mediante tinción con X-Gal. (b) Esquema del experimento. Una biblioteca de variantes de I-CreI se construye usando PCR, se clona en un vector de expresión replicativo de levadura y se transforma en la cepa FYC2-6A de *S. cerevisiae* (*MAT α* , *trp1 Δ 63*, *leu2 Δ 1*, *his3 Δ 200*). Las 64 dianas palindrómicas se clonan en un vector informador de levadura basado en LacZ y los clones resultantes se transforman en la cepa FYBL2-7B (*MAT α* , *ura3 Δ 851*, *trp1 Δ 63*, *leu2 Δ 1*, *lys2 Δ 202*). El entramado asistido por robot en membrana de filtro se usa para realizar cruzamiento entre clones individuales que expresan variantes de la meganucleasa y clones individuales que albergan un plásmido informador. Después de la exploración primaria de alto rendimiento, el ORF de los clones positivos se amplifica mediante PCR y se secuencian. Se identificaron 410 variables diferentes entre los 2100 positivos y se ensayaron a densidad baja para establecer patrones concretos y se validaron 350 clones. También se reclonaron 294 mutantes en vectores de levadura y se ensayaron en una exploración secundaria y los resultados confirmaron los obtenidos sin reclonación. Los clones elegidos se ensayaron después con respecto a su actividad de escisión en un ensayo similar basado en CHO y eventualmente *in vitro*.

- La figura 4 representa la secuencia de ADNc que codifica la proteína de armazón I-CreI N75 y los cebadores degenerados usados para la construcción de la biblioteca Ulib2. A. La secuencia codificante (CDS) de la proteína de armazón (SEC ID N° 69) es desde el par de bases 1 al par de bases 501 y el codón de terminación TGA (no mostrado) sigue al par de bases 501. Además de la mutación D75N, la proteína contiene mutaciones adicionales que no alteran su actividad; en la secuencia de proteína (SEC ID N° 70), los dos primeros restos N terminales son metionina y alanina (MA) y los tres restos C terminales alanina, alanina y ácido aspártico (AAD). B. Cebadores degenerados (SEC ID N° 67, 68).

- La figura 5 representa el mapa del vector de expresión de meganucleasa pCLS0542. El vector de expresión de la meganucleasa está marcado con *LEU2*. Los ADNc que codifican las variantes de meganucleasa I-CreI se clonan en este vector digerido con *NotI* y *EagI*, para tener la expresión de la variante conducida por el promotor Gal10 inducible.

- La figura 6 representa el mapa del vector informador pCLS0042. El vector informador está marcado con *TRP1* y *URA3*. Las repeticiones en tándem LacZ comparten 800 pb de homología y están separadas por 1,3 kb de ADN. Están rodeadas por secuencias promotoras *ADH* y terminadoras. Los sitios diana se clonan en el sitio *SmaI*.
- 5 - La figura 7 ilustra el perfil de escisión de 292 variantes de meganucleasa I-*CreI* con una especificidad modificada. Las variantes derivan de la proteína de armazón I-*CreI* N75. Las proteínas se definen mediante el aminoácido presente en las posiciones 44, 68 y 70 (tres primeras columnas). La numeración de los aminoácidos es de acuerdo con el código de acceso de pdb 1g9y. Las dianas se definen mediante los nucleótidos en las posiciones -5 a -3. Para cada proteína, la escisión observada (1) o la
- 10 escisión no observada (0) se muestra para cada una de las 64 dianas.
- La figura 8 ilustra ocho ejemplos de patrón de escisión de variantes de I-*CreI*. Las meganucleasas se ensayan 4 veces contra las 64 dianas descritas en la figura 2B. La posición de las diferentes dianas se indica en el panel superior izquierdo. Las variantes que derivan de la proteína de armazón I-*CreI* N75 se identifican mediante los aminoácidos en las posiciones 44, 68 y 70 (ej: KSS es K44, S68, S70 y N75 o
- 15 K44/S68/S70). La numeración de los aminoácidos es de acuerdo con el código pdb 1g9y. QRR corresponde a I-*CreI* N75. Las dianas escindidas se indican junto a los paneles.
- La figura 9 ilustra los patrones de escisión de las variantes. Los mutantes se identifican mediante tres letras, correspondientes a los restos en las posiciones 44, 68 y 70. Cada mutante se ensaya contra las 64 dianas derivadas de las dianas naturales I-*CreI* y una serie de dianas de control. El mapa de la diana se
- 20 indica en el panel superior derecho. (a) Los patrones de escisión en levadura (izquierda) y células de mamífero (derecha) para las proteínas I-*CreI* de tipo silvestre (I-*CreI*) e I-*CreI* N75 (QRR) y 7 derivados de la proteína I-*CreI* N75. Para la levadura se muestran los datos sin procesar iniciales (filtro). Para las células CHO se muestran los datos sin procesar cuantitativos (medición ONPG), los valores superiores a 0,25 están marcados con una caja, los valores superiores a 0,5 se destacan en gris medio, los valores superiores a 1 en gris oscuro. LacZ: control positivo. 0: sin diana. U1, U2 y U3: tres controles diferentes no
- 25 escindidos. (b) Escisión *in vitro*. I-*CreI* y cuatro mutantes se ensayan contra un conjunto de 2 ó 4 dianas, incluyendo la diana que da como resultado la señal más fuerte en levadura y CHO. Las digestiones se realizan a 37°C durante 1 hora, con sustrato linealizado 2 nM, como se describe en Métodos. Se muestran los datos sin procesar para I-*CreI* con dos dianas diferentes. Tanto con ggg como con cct, la escisión no se
- 30 detecta con I-*CreI*.
- La figura 10 representa el análisis estadístico. (a) Dianas escindidas: las dianas escindidas con variantes de I-*CreI* se colorean en gris. El número de proteínas que escinden cada diana se muestra debajo y el nivel de coloración gris es proporcional a la intensidad de señal media obtenida con estos cortadores en levadura. (b) Análisis de tres de los siete grupos. Para cada grupo mutante (grupos 1, 3 y 7), se
- 35 calcularon las intensidades acumuladas para cada diana y un diagrama de barras (columna izquierda) muestra en orden decreciente las intensidades normalizadas. Para cada grupo, el número de aminoácidos de cada tipo en cada posición (44, 68 y 70) se muestra como un histograma codificado en la columna derecha. La leyenda del código de colores de aminoácidos está en la parte de abajo de la figura. (b) Agrupación

jerárquica de datos de mutante y diana en levadura. Tanto los mutantes como las dianas se agruparon usando agrupamiento jerárquico con distancia euclídea y método de Ward (Ward, J.H., American statist. Assoc., 1963, 58, 236-244). La agrupación se realizó con hclusp del paquete R. Los dendogramas de los mutantes y dianas se reordenaron para optimizar las posiciones de los grupos y el dendograma de mutantes se cortó a una altura de 8 con los grupos deducidos. El mutante de QRR y la diana de GTC se indican con una flecha. Los niveles de gris reflejan la intensidad de la señal.

- La figura 11 ilustra un ejemplo de sitio híbrido o quimérico: gtt (SEC ID N°: 79) y cct (SEC ID N°: 77) son dos sitios palindrómicos derivados del sitio I-CreI. El sitio híbrido gtt/cct (SEC ID N°: 80) presenta la secuencia gtt en la cadena superior en -5, -4, -3 y la secuencia cct en la cadena inferior en 5, 4, 3.

10 - La figura 12 ilustra la actividad de escisión de las variantes heterodiméricas. Las levaduras se co-transformaron con las variantes KTG y QAN. La organización diana se muestra en el panel superior: diana con un semisitio gtt, cct o ggc único están en negrita; las dianas con dos semisitios tales, que se espera que se escindan mediante homo y/o heterodímeros están en negrita y destacados en gris; 0: sin diana. Los resultados se muestran en los tres paneles inferiores. Se observan señales tenues e inesperadas solamente para gtc/cct y gtt/gtc, escindidas por KTG y QAN, respectivamente.

15 - La figura 13 representa el análisis cuantitativo de la actividad de escisión de las variantes heterodiméricas. (a) Co-transformación de mutantes seleccionados en levadura. Para una mayor claridad, sólo se muestran los resultados de dianas híbridas relevantes. La diana aac/acc se muestra siempre como un ejemplo de diana no relacionada. Para la pareja KTGxAGR, las dianas tac y tct palindrómicas, aunque no se muestran, se escinden mediante AGR y KTG, respectivamente. La escisión de la diana cat mediante el mutante RRN es muy baja y no pudo cuantificarse en levadura. (b) Co-transfección transitoria en células CHO. Para (a) y (b), barras negras: señal para el primer mutante solo; barras grises: señal para el segundo mutante solamente; barras a rayas: señal obtenida mediante co-expresión o co-transfección.

20 - La figura 14 ilustra la actividad del heterodímero ensamblado ARS-KRE en la diana de ADN seleccionada del cromosoma 17 de ratón. La línea celular CHO-K1 se co-transfectó con equimolar de plásmido diana LagoZ, plásmidos de expresión ARS y KRE y se midió la actividad beta galactosidasa. Las células co-transfectadas con el plásmido LagoZ y el plásmido recombinante I-SceI, I-CreI, ARS o KRE o un plásmido vacío se usaron como control.

25 Los siguientes ejemplos se presentan aquí solamente para ilustrar la invención y no para limitar el ámbito de la misma. Otras variantes, obtenidas de un ADNc, cuya secuencia difiere de la SEC ID N°: 69 y que usan cebadores apropiados, aun son parte de la invención.

Ejemplo 1: exploración para nuevas endonucleasas funcionales

El método para producir variantes de meganucleasa y los ensayos basados en la recombinación inducida por escisión en células de mamífero o levadura, que se usan para variantes de exploración con especificidad alterada, se describen en la Solicitud Internacional PCT WO 2004/067736.

5 Estos ensayos dan como resultado un gen informador LacZ que puede controlarse mediante métodos convencionales (figura 3a).

A) Material y métodos**a) Construcción de bibliotecas mutantes**

Se sintetizaron marcos abiertos de lectura de I-CreI wt e I-CreI D75N (o I-CreI N75)
10 (SEC ID N°:69, figura 4A), como se ha descrito anteriormente (Epinat *et al.*, N.A.R., 2003, 31, 2952-2962). La mutación D75N se introdujo mediante el reemplazamiento del codón 75 con aac. La diversidad de la biblioteca de meganucleasas se generó mediante PCR que usa cebadores degenerados de Sigma que albergan el codón VVK (18 codones, aminoácidos ADEGHKPNQRST) en las posiciones 44, 68 y 70 que interactúan directamente con las bases de las posiciones 3 a 5 y como molde de ADN, el gen I-CreI D75N.
15 Dichos cebadores permiten la mutación de los restos 44, 68 y 70 con una diversidad teórica de 12. Brevemente, el cebador directo (5'-gtttaaacaatcagctaagcttgacctttvkgctgacttcaaagaccag-3', SEC ID N°: 67) y el cebador reverso (5'-gatgtagttggaacggatccmbbtacgtaccaacgcc-3', SEC ID N°:68) se usaron para amplificar un fragmento de PCR en reacciones de PCR de 50 µl: los productos de PCR se agruparon, se precipitaron con EtOH y se resuspendieron en 50 µl de Tris 10 mM. Los productos de PCR se clonaron
20 en un vector de expresión pET que contenía el gen I-CreI D75N, digerido con enzimas de restricción apropiadas. La digestión del ADN de vector y de inserto se llevó a cabo en dos pasos (digestión de enzima sencilla) entre los que la muestra de ADN se extrajo (usando métodos clásicos basados en fenol; cloroformo: isoamilalcohol) y se precipitó con EtOH. Se usaron 10 µg de ADN del vector digerido para ligación, con un exceso 5:1 de ADN inserto. Las células de TG1 de *E. coli* se transformaron con el vector
25 resultante mediante electroporación. Para producir varios clones celulares por encima de la diversidad teórica de la biblioteca, se produjeron 6×10^4 clones. Los clones bacterianos se rasparon de las placas y los vectores plasmídicos correspondientes se extrajeron y se purificaron.

La biblioteca se reclonó en el vector de levadura pCLS0542 (figura 5), mediante la subclonación de un fragmento de ADN de *NcoI-EagI* que contenía el ORF de I-CreI D75N completo. En este
30 vector replicativo basado en 2 micrones marcado con el gen *LEU2*, las variantes de I-CreI están bajo el control de un promotor inducible por galactosa (Epinat *et al.*, citado anteriormente). Después de la electroporación en *E. coli*, se obtuvieron 7×10^4 clones, que representan 12 veces la diversidad teórica a nivel de ADN ($18^3 = 5832$). El ADN se extrajo y se transformó en la cepa FYC2-6A de *S. cerevisiae* (*MATa*, *trp1Δ63*, *leu2Δ1*, *his3Δ200*). Se recogieron 13824 colonias usando un recolector de colonias
35 (QpixII, GENETIX) y se cultivaron en placas de 144 microtítulos.

b) Construcción de clones diana

El palíndromo de veinticuatro pb C1221 (tcaaacgtcgtgacgacgtttga, SEC ID N°: 12) es una repetición del semisitio de la diana natural de I-CreI casi palindrómica (tcaaacgtcgtgagacagttgg, SEC ID

Nº: 65). C1221 se escinde tan eficazmente como la diana natural de I-CreI *in vitro* y *ex vivo* tanto en células de levadura como de mamíferos. Las 64 dianas palindrómicas se derivaron como sigue: 64 pares de oligonucleótidos (ggcatacaagtttcaaaacnngtacnngtttgacaatcgctgtca (SEC ID Nº: 72) y las secuencias complementarias reversas) que corresponden a las dos cadenas de las 64 dianas de ADN, con 12 pb de secuencia extra no palindrómica en cada lado, se encargaron de Sigma, se hibridaron y se clonaron en pGEM-T Easy (PROMEGA). A continuación se escindió un fragmento de 400 pb de PvuII de cada uno de los 64 vectores derivados de pGEM-T y se clonaron en un vector de levadura pFL39-ADH-LACURAZ, descrito previamente (Epinat *et al.*, citado anteriormente), también llamado pCLS0042 (Figura 6), dando como resultado 64 vectores informadores de levadura. Los pasos de escisión, digestión y ligación se realizaron usando métodos típicos conocidos por los expertos en la materia. La inserción de la secuencia diana se realiza en el sitio SmaI de pCLS00402. Las 64 dianas palindrómicas se describen en la figura 2B (posiciones -5 a -3 y +3 a +5, SEC ID Nº: 1 a SEC ID Nº: 64).

c) Cepas de levadura y transformación

La biblioteca de variantes de expresión de meganucleasa y la variante A44/R68/L70, se transformaron en la cepa FYC2-6A (*MAT α* , *trp1 Δ 63*, *leu2 Δ 1*, *his3 Δ 200*).

Los plásmidos diana se transformaron en la cepa de levadura FYBL2-7B: (*MAT α* , *ura3 Δ 851*, *trp1 Δ 63*, *leu2 Δ 1*, *lys2 Δ 202*).

Para la transformación, puede usarse un protocolo clásico de choque químico/térmico y rutinariamente da 10^6 transformantes independientes por μ g de ADN; los transformantes se seleccionaron en medio sintético de separación de leucina (Gietz y Woods, 2002).

d) Cruzamiento de clones de expresión de meganucleasa y exploración en levadura

Los clones de variantes de I-CreI así como las cepas informadoras de levadura se reservaron en reserva de glicerol (20%) y se replicaron en nuevas micro placas. Los mutantes se reticularon en filtros de nailon que cubrían placas YPD, usando una densidad de reticulación alta (aproximadamente 20 puntos/cm²). Un segundo proceso de reticulación se realizó en los mismos filtros para aplicar puntualmente una segunda capa que consiste en 64 ó 75 cepas diferentes de levadura que albergan informador para cada variante. Brevemente, cada cepa informadora se aplicó puntualmente 13.824 veces en una membrana de nailon y en cada uno de estos puntos se aplicó puntualmente uno de los 13.824 clones de levadura que expresan una variante de meganucleasa. Las membranas se colocaron en un medio de agar sólido rico en YPD y se incubaron a 30°C durante una noche, para permitir el cruzamiento. A continuación, se transfirieron los filtros a medio sintético, carente de leucina y triptófano, con galactosa (1%) como fuente de carbono (y con G418 para los experimentos de co-expresión) y se incubaron durante cinco días a 37°C, para seleccionar los diploides que portan los vectores de expresión y diana. Después de 5 días, se colocaron los filtros en un medio de agarosa sólida con X-Gal al 0,02% en tampón de fosfato de sodio 0,5 M, pH 7,0, SDS al 0,1%, dimetil formamida (DMF) al 6%, β -mercaptoetanol 7 mM, agarosa al 1% y se incubaron a 37°C, para controlar la actividad de β -galactosidasa. Los clones positivos se identificaron después de dos días de incubación, de acuerdo con la tinción. Los resultados se analizaron mediante exploración y se realizó la cuantificación usando un software patentado. Para la exploración secundaria, se

siguió el mismo procedimiento con los 292 positivos seleccionados, excepto que cada mutante se ensayó 4 veces en la misma membrana (véase figura 8 y 9a).

e) Secuencia y Re-Clonación de éxitos primarios.

El marco abierto de lectura (ORF) de clones positivos identificados durante la exploración primaria en levadura se amplificó mediante PCR y se secuenció. Después, los ORF se reclonaron usando el protocolo Gateway (Invitrogen). Los ORF se amplificaron mediante PCR en colonias de levadura (Akada *et al.*, *Biotechniques*, 28, 668-670, 672-674), usando los cebadores:

ggggacaagttgtacaaaaaagcaggcttcgaaggagatagaacctggccaataccaatacaaaagattcc (SEC ID N°: 73) y
ggggaccactttgtacaagaaagctgggttagtcggccgccggggaggattcttctctcgc (SEC ID N°: 74) de PROLIGO. Los productos de PCR se clonaron en: (i) vector de expresión Gateway de levadura que alberga un promotor inducible por galactosa, *LEU2* o KanR como marcador seleccionable y un origen de replicación de 2 micrones y (ii) un vector pET 24d(+) de NOVAGEN. Los clones resultantes se verificaron mediante secuenciación (MILLEGEN).

B) Resultados

I-CreI es una endonucleasa de movimiento dirigido dimérica que escinde una diana pseudo-palindrómica de 22 pb. El análisis de la estructura de *I-CreI* unida a su diana natural ha mostrado que en cada monómero ocho restos establecen interacciones directas con siete bases (Jurica *et al.*, 1998, citado anteriormente). Los restos Q44, R68, R70 entran en contacto con tres pares de bases consecutivos en las posiciones 3 a 5 (y -3 a -5, figura 1). Se llevó a cabo un enfoque exhaustivo de biblioteca de proteínas contra biblioteca de dianas para desarrollar por ingeniería genética localmente esta parte de la interfaz de unión a ADN. Primero, el armazón de *I-CreI* se mutó de D75 a N para disminuir los probables esfuerzos energéticos causados por el reemplazo de los restos básicos R68 y R70 en la biblioteca que satisface el potencial aceptor de hidrógeno del D75 enterrado en la estructura de *I-CreI*. Los homodímeros del mutante D75N (purificado de células *E. coli* en las que se sobre-expresaba usando un vector de expresión pET) se mostró que escindían el sitio de movimiento dirigido *I-CreI*. La mutación D75N no afectó a la estructura proteínica, pero disminuyó la toxicidad de *I-CreI* en los experimentos de sobre-expresión. A continuación, la posiciones 44, 68 y 70 se seleccionaron aleatoriamente y se generaron las 64 dianas palindrómicas resultantes de sustituciones en las posiciones ± 3 , ± 4 y ± 5 de una diana palindrómica escindida por *I-CreI* (Chevalier *et al.*, 2003, citado anteriormente), como se describe en la figura 1 y 2B. Eventualmente, mutantes de la biblioteca de proteínas correspondieron a combinaciones independientes de cualquiera de los 12 aminoácidos codificados por el codón vvk en posiciones de tres restos. En consecuencia, la diversidad máxima (teórica) de la biblioteca de proteínas era de 12^3 ó 1.728. Sin embargo, en términos de ácidos nucleicos, la diversidad es de 18^3 ó 5.832.

La biblioteca resultante se clonó en un vector de expresión replicativo de levadura que portaba un gen marcador auxotrófico *LEU2* y se transformó en una cepa de levadura haploide mutante *leu2* (FYC2-6A). Las 64 dianas se clonaron en el vector informador de levadura apropiado y se transformaron en una cepa haploide (FYBL2-7B), dando como resultados 64 cepas de ensayo.

Un protocolo de cruzamiento asistido por robot se usó para explorar un gran número de meganucleasas de nuestra biblioteca. La estrategia de exploración general se describe en la figura 3B. Se aplicaron puntualmente 138.247 clones de expresión de meganucleasa (aproximadamente 2,3 veces la diversidad teórica) a alta densidad (20 puntos/cm²) en filtros de nailon y se ensayaron individualmente
5 contra cada uno de las 64 cepas diana (884.608 puntos). Se aislaron 2.100 clones que mostraron una actividad contra al menos una diana (figura 3b) y el ORF que codifica la meganucleasa se amplificó por PCR y se secuenció. Se identificaron 410 secuencias diferentes y un número similar de clones correspondientes se seleccionaron para análisis adicional. La densidad de puntos se redujo a 4 puntos/cm² y cada clon se ensayó contra las 64 cepas informadoras por cuadruplicado, creando de este modo perfiles
10 completos (como en la figura 8 y 9a). Se pudieron confirmar 350 positivos. A continuación, para evitar la posibilidad de cepas que contuvieran más de un clon, se amplificaron los ORF mutantes mediante PCR y se reclonaron en el vector de levadura. Los plásmidos resultantes se transformaron individualmente de nuevo en levadura. Se obtuvieron 294 de tales clones y se ensayaron a densidad baja (4 puntos/cm²). Se observaron diferencias con la exploración primaria principalmente para señales débiles, apareciendo ahora
15 28 cortadores débiles como negativos. Sólo un clon positivo presentó un patrón diferente del que se observó en el perfil primario.

Ejemplo 2: variantes de meganucleasa I-CreI con diferentes perfiles de escisión

Los clones validados del ejemplo 1 mostraron **patrones** muy diversos. Algunos de estos nuevos perfiles compartían alguna similitud con el armazón inicial mientras que muchos otros eran
20 totalmente diferentes. Se muestran diversos ejemplos de perfiles, incluyendo I-CreI de tipo silvestre e I-CreI N75 en las figuras 8 y 9a. Los resultados globales (sólo para las 292 variantes con especificidad modificada) se resumen en la figura 7.

Las endonucleasas de movimiento dirigido pueden habitualmente alojar algo de degeneración en sus secuencias diana y uno de nuestros primeros hallazgos fue que la proteína I-CreI
25 original en sí misma escinde 7 dianas diferentes en levadura. Muchos de nuestros mutantes siguieron esta regla también, variando el número de secuencias escindidas de 1 a 21 con una media de 5,0 secuencias escindidas (desviación típica = 3,6). De manera interesante, en 50 mutantes (14%), la especificidad se alteró de tal modo que escindían exactamente una diana. 37 (11%) escindían 2 dianas, 61 (17 %) escindían 3 dianas y 58 (17 %) escindían 4 dianas. Para 5 dianas y más, los porcentajes eran menores del 10 %. En
30 total, los mutantes escindieron 38 dianas (Figura 10a). Es digno de atención que la escisión apenas se observaba en dianas con una A en posición ± 3 y nunca con dianas con tgn y cgn en posición ± 5 , ± 4 y ± 3 .

Estos resultados no limitan el ámbito de la invención, puesto que la figura 7 sólo muestra resultados obtenidos con 292 variantes (291 de las 1.728 (ó 12³) variantes de meganucleasa I-CreI obtenibles en una biblioteca completa).

Ejemplo 3: Nuevas meganucleasas pueden escindir nuevas dianas mientras que mantienen una alta actividad y una especificidad estrecha**A) Material y métodos****a) Construcción de clones diana**

5 Las 64 dianas palindrómicas se clonaron pGEM-T Easy (PROMEGA), como se ha descrito en el ejemplo 1. A continuación, un fragmento de 400 pb de *PvuII* se escindió y clonó en el vector de mamífero pcDNA3.1-LACURAZ-AURA, descrito anteriormente (Epinat *et al.*, citado anteriormente). Las 75 secuencias diana híbridas se clonaron como sigue: se designaron oligonucleótidos que contenían dos semisitios diferentes para cada palíndromo mutante (PROLIGO).

10 b) Reclonación de éxitos primarios

El marco abierto de lectura (ORF) de clones positivos identificados durante la exploración primaria en levadura se reclonaron en: (i) un vector de expresión gateway CHO pCDNA6.2, siguiendo las instrucciones del proveedor (INVITROGEN) y ii) un vector pET 24d(+) de NOVAGEN. Los clones resultantes se verificaron mediante secuenciación (MILLEGEN).

15 c) Ensayo de células de mamífero

La línea celular CHO-K1 de la Colección Americana de Cultivos Tipo (American Type Culture Collection) (ATCC) se cultivó en medio F12K de Ham suplementado con suero bovino fetal al 10%. Para ensayos de hibridación de cadena sencilla transitorios (SSA), se sembraron las células en placas de doce pocillos a 13.10^3 células por pocillo un día antes de la transfección. La co-transfección se llevó a cabo el día siguiente con 400 ng de ADN usando el kit de transfección EFFECTENE (QIAGEN). Se usaron cantidades equimolares del plásmido diana LagoZ y plásmido de expresión. El día siguiente, se reemplazó el medio y las células se incubaron durante otras 72 horas. Las monocapas celulares CHO-K1 se lavaron una vez con PBS. Las células se alisaron después con 150 μ l de tampón de lisis/revelado añadido para un ensayo líquido de β -galactosidasa (100 ml de tampón de lisis (Tris-HCl 10 mM pH 7,5, NaCl 150 mM, Triton X100 0,1 %, BSA 0,1 mg/ml, inhibidores de proteasa) y 900 ml de tampón de revelado (10 ml de tampón Mg 100X (MgCl₂ 100 mM, (β -mercaptoetanol 35 %), 110 ml de ONPG (8 mg/ml) y 780 ml de fosfato de sodio 0.1 M pH 7,5), 30 minutos en hielo. La actividad de Beta-galactosidasa se ensayó mediante la medición de la densidad óptica a 415 nm. El proceso completo se realizó en una plataforma automatizada Velocity11 BioCel. La actividad de beta-galactosidasa se calcula como unidades relativas normalizadas para concentración de proteína, tiempo de incubación y eficacia de transfección.

30 d) Expresión de purificación de proteínas

Las proteínas marcadas con His se sobre-expresaron en células BL21 (DE3) pLysS de *E. coli* usando vectores pET-24d(+) (NOVA-GEN). La inducción con IPTG (0,3 mM), se realizó a 25°C. Las células se sometieron a ultrasonidos en una solución de fosfato de sodio 50 mM (pH 8), cloruro de sodio 300 mM que contenía inhibidores de proteasa (Comprimidos completos sin EDTA, Roche) y glicerol al 5 % (v/v). Los lisados celulares se centrifugaron a 100.000 g durante 60 min. Las proteínas marcadas con His se purificaron después por afinidad, usando columnas HP quelantes Hi-Trap de 5 ml (Amersham Biosciences) cargadas con cobalto. Se recogieron varias fracciones durante la elución con un gradiente

lineal de imidazol (hasta imidazol 0,25 M, seguido de meseta de imidazol 0,5 M, NaCl 0,3 M y fosfato de sodio 50 mM pH 8). Se aplicaron fracciones ricas en proteínas (determinadas por SDS-PAE) a la segunda columna. Las muestras purificadas en bruto se tomaron a pH 6 y se aplicaron a una columna de HP de Heparina HiTrap de 5ml (Amersham Biosciences) equilibrada con fosfato de sodio 20 mM pH 6,0. Las proteínas unidas se eluyen con un gradiente continuo de cloruro de sodio con fosfato de sodio 20 mM y cloruro de sodio 1M. Las fracciones purificadas se sometieron a SDS-PAGE y se concentraron (Sistema Amicon Ultra centriprep con punto de corte de 10 kDa), congelado en nitrógeno líquido y almacenado a -80°C. Se desalaron las proteínas purificadas usando columnas PD10 (Sephadex G-25M, Amersham Biosciences) en PBS o tampón Tris-HCl 10 mM (pH 8).

10 e) Ensayos de escisión *in vitro*

Los plásmidos pGEM con sitios de corte diana de ADN de meganucleasas sencillos se linealizaron primero con XmnI. Los ensayos de escisión se realizaron a 37°C en Tris-HCl 10 mM (pH 8), NaCl 50 mM, MgCl₂ 10mM, DTT 1mM y BSA 50 µg/ml. Se usó 2 nM como concentración de sustrato diana. Se usó un intervalo de dilución entre 0 y 85 nM para cada proteína en volumen de reacción final de 15 25 ml. Las reacciones se detuvieron después de 1 hora mediante la adición de 5µl de glicerol al 45%, EDTA 95mM (pH 8), SDS 1,5 % (p/v), proteinasa K 1,5 mg/ml y azul de bromofenol al 0,048 % (p/v) (6X tampón de parada) y se incubaron a 37°C durante 30 minutos. Los productos de digestión se analizaron en gel de electroforesis de agarosa y se cuantificaron los fragmentos después de tinción con bromuro de etidio, para calcular el porcentaje de escisión.

20 **B) Resultados**

Se seleccionaron ocho mutantes representativos (que pertenecen a 6 grupos diferentes, véase abajo) para una caracterización adicional (Figura 9). Primero, los datos en la levadura se confirmaron en las células de mamífero, mediante el uso de un ensayo basado en la co-transfección transitoria de un vector que expresa meganucleasa y un vector diana, como se ha descrito en un informe anterior. Los 8 ORF mutantes y las 64 dianas se clonaron en vectores apropiados y se usó un protocolo basado en microtítulos asistido por robot para co-transfectar en células CHO cada variante seleccionada con cada uno de los 64 plásmidos informadores diferentes. La recombinación inducida por meganucleasa se midió mediante un ensayo de ONPG cuantitativo convencional que controla la restauración de un gen funcional de β-galactosidasa. Se descubrió que los perfiles eran reproducibles cualitativa y cuantitativamente en 25 cinco experimentos independientes. Como se muestra en la figura 9a casi siempre se observaron señales fuertes y medias tanto con células de levadura como CHO (con la excepción de ADK), validándose de este modo la relevancia del proceso HTS de levadura. Sin embargo, no se detectaron con frecuencia señales observadas débiles en levadura en las células CHO, probablemente debido a una diferencia en el nivel de detección (véase QRR y dianas gtg, gct, y ttc). También se produjeron cuatro mutantes en *E. Coli* y se 35 purificaron mediante cromatografía de afinidad metálica. Se determinó que sus eficacias de escisión relativas *in vitro* contra el sitio de tipo silvestre y sus sitios afines. El alcance de la escisión en condiciones convencionales se evaluó a través de un amplio intervalo de concentraciones para los mutantes (Figura 9b). De forma similar, se analizó la actividad de I-CreI wt en estas dianas. En muchos casos, no se pudo

alcanzar un 100% de escisión del sustrato, reflejando probablemente el hecho de que estas proteínas pueden tener una renovación baja o nula (Perrin *et al.*, EMBO J., 1993, 12, 2939-2947; Wang *et al.*, Nucleic Acids Res., 1997, 25, 3767-3776). En general, el ensayo *in vitro* confirmó los datos obtenidos en levadura y células CHO, pero sorprendentemente, la diana gtt se escindió eficazmente por I-CreI.

5 Los cambios de especificidad fueron obvios a partir de los perfiles obtenidos en levadura y CHO: la diana gtc favorita de I-CreI no se escindió o se escindió escasamente, mientras que se observaron señales con nuevas dianas. Este cambio de especificidad se confirmó para QAN, DRK, RAT y KTG mediante análisis *in vitro*, como se muestra en la figura 9b. Además, estos cuatro mutantes, que presentan diversos niveles de actividad en levadura y CHO (Figura 9a) mostraron que escindían el 17-60% de su
10 diana favorita *in vitro* (Figura 9), con cinética similar a I-CreI (la mitad de la escisión máxima mediante 13-25 nM). Así pues, la actividad se conservó mayoritariamente mediante ingeniería genética. Tercero, el número de dianas escindidas varió entre los mutantes: cortadores fuertes tales como QRR, QAN, ARL y KTG tienen un espectro de escisión en el intervalo de lo que se observa con I-CreI (5-8 señales detectables en levadura, 3-6 en CHO). La especificidad es más difícil de comparar con mutantes que escinden
15 débilmente. Por ejemplo, una señal única débil se observa con DRK pero podría representar la única señal detectable que resulta de la atenuación de un patrón más complejo. Sin embargo, el comportamiento de las variantes que escinden fuertemente muestra que el desarrollo por ingeniería genética conserva una especificidad muy estrecha.

Ejemplo 4: La agrupación jerárquica define siete familias de variantes de I-CreI.

20 A) Material y métodos.

La agrupación se realizó usando hclust del paquete R. Usamos datos cuantitativos de la exploración primaria de baja densidad. Tanto las variantes como las dianas se agruparon usando agrupación jerárquica convencional con distancia Euclídea y método de Ward (Ward, J.H., American Stat. Assoc., 1963, 58, 236-244). Los dendrogramas de mutantes y dianas se reordenaron para optimizar las
25 posiciones de los grupos y el dendrograma mutante se cortó a la altura 8 para definir el grupo.

B) Resultados

A continuación, la agrupación jerárquica se usó para determinar si las familias podían identificarse entre los numerosos y diversos patrones de escisión de las variantes. Puesto que la exploración primaria y secundaria dio resultados concluyentes, se usaron datos cuantitativos del primer
30 ciclo de exploración de baja densidad de levadura para el análisis, para permitir un mayor tamaño de muestra. Tanto las variantes como las dianas se agruparon usando la agrupación jerárquica convencional con distancia Euclídea y método de Ward (Ward, J.H., citado anteriormente) y se definieron siete grupos (Figura 10b). Se muestra el análisis detallado para los 3 (Figura 10c) y los resultados se resumen en la Tabla I.

Tabla I: Análisis de grupos

grupo	ejemplos (Fig. 3a)	Tres dianas preferidas		Nucleótido en posición 4 (%) ¹		aminoácidos preferidos ²		
		secuencia	% escisión	1	4	44	68	70
1	QAN	gtt	46,2	g	0,5	Q		
		gtc	18,3	a	2,0	80,5%		
		gtg	13,6	t	82,4	(62/77)		
			$\Sigma= 78,1$	c	15,1			
2	QRR	gtt	13,4	g	0	Q	R	
		gtc	11,8	a	4,9	100,0%	100,0%	
		tct	11,4	t	56,9	(8/8)	(8/8)	
			$\Sigma= 36,6$	c	38,2			
3	ARL	gat	27,9	g	2,4	A	R	
		tat	23,2	a	88,9	63,0%	33,8%	
		gag	15,7	t	5,7	(41/65)	(22/65)	
			$\Sigma= 66,8$	c	3,0			
4	AGR	gac	22,7	g	0,3	A&N	R	R
						51,6% &		
		tac	14,5	a	91,9	35,4%	48,4%	67,7%
		gat	13,4	t	6,6	(16&11/31)	15/31	21/31
		$E= 50,6$	c	1,2				
5	ADK DRK	gat	29,21	g	1,6			
		tat	15,4	a	73,8			
		gac	11,4	t	13,4			
			$E= 56.05.9$	c	11,2			
6	KTG RAT	cct	30,1	g	0	K		
		tct	19,6	a	4,0	62,7%		
		tcc	13,9	t	6,3	(32/51)		
			$\Sigma= 63,6$	c	89,7			
7		cct	20,8	g	0	K		
		tct	19,6	a	0,2	91,9%		
		tcc	15,3	t	14,4	(34/37)		
			$\Sigma= 55,7$	c	85,4			

¹ frecuencias de acuerdo con el índice de escisión, como se ha descrito en la figura 10c

² en cada posición, se indican los restos presentes en más de un tercio de los grupos.

Para cada grupo, se puede identificar un conjunto de dianas preferidas en base a la frecuencia e intensidad de la señal (Figura 10c). Las tres dianas preferidas para cada grupo se indican en la Tabla 1, con sus frecuencias de escisión. La suma de estas frecuencias es una medida de la especificidad del grupo. Por ejemplo, en el grupo 1, las tres dianas preferidas (gtt/c/g), representan el 78,1% de la escisión observada, con un 46,2% solo para gtt lo que revela una especificidad muy estrecha. De hecho, este grupo incluye varias proteínas que, como QAN, que escinde principalmente gtt (Figura 9a). Por el contrario, las tres dianas preferidas del grupo 2 representan solo el 36,6% de todas las señales observadas. De acuerdo con los patrones relativamente amplios y diversos observados en este grupo, QRR escinde 5 dianas (figura 9a), mientras que la actividad de otros miembros del grupo no está restringida a estas cinco dianas.

El análisis de los restos encontrados en cada grupo mostró fuertes sesgos para la posición 44: Q está abundantemente representada en los grupos 1 y 2, mientras que A y N son más frecuentes en los grupos 3 y 4 y K en los grupos 6 y 7. Mientras tanto, estos sesgos se correlacionan con fuertes preferencias de bases para las posiciones de ADN ± 4 , con una gran mayoría de pares de bases t:a en los grupos 1 y 2, a:t en los grupos 3, 4 y 5 y t:g en los grupos 6 y 7 (véase tabla 1). La estructura de I-CreI unida a su diana muestra que el resto Q44 interactúa con la cadena inferior en la posición -4 (y la cadena superior de la posición +4, véase figura 1b y 1c). Estos resultados sugieren que esta interacción está ampliamente conservada en nuestros mutantes y revela un "código", en el que Q44 establecería contacto con adenina, A44 (o menos frecuentemente N44) con timina y K44 con guanina. Dicha correlación no se observó para las posiciones 68 y 70.

Ejemplo 5: las variantes pueden ensamblarse en heterodímeros funcionales para escindir nuevas secuencias diana de ADN

A) Material y Métodos

Las 75 secuencias diana híbridas se clonaron como sigue: los oligonucleótidos se diseñaron para contener dos semisitios diferentes de cada palíndromo mutante (PROLIGO). Se clonó el ADN diana bicatenario, generado por amplificación de PCR de los oligonucleótidos monocatenarios, usando el protocolo Gateway (INVITROGEN) en vectores informadores de levadura y mamífero. Los vectores informadores de levadura se transformaron en la cepa FYBL2-7B de *S. cerevisiae* (*MAT α* , *ura3 Δ 851*, *trp1 Δ 63*, *leu2 Δ 1*, *lys2 Δ 202*).

B) Resultados

Las variantes son homodímeros capaces de escindir sitios palindrómicos. Para ensayar si la lista de dianas escindibles puede extenderse mediante la creación de heterodímeros que escindirían sitios de escisión híbridos (como se describe en la figura 11), se seleccionó un subconjunto de variantes de I-CreI con perfiles definidos y se clonó en dos vectores diferentes de levadura marcados con genes *LEU2* o *KAN*. Las combinaciones de mutantes que tienen mutaciones en las posiciones 44, 68 y/o 70 y N en la posición 75, se co-expresaron después en levadura con un conjunto de dianas de ADN quiméricas, palindrómicas y no palindrómicas. Se muestra un ejemplo en la figura 12: la co-expresión de los mutantes K44, T68, G70, N75 (KTG) y Q44, A68, N70, N75 (QAN) dio como resultado la escisión de dos dianas quiméricas gtt/gcc

y gtt/cct, que no se escindían por ninguno de los mutantes por sí mismos. Las dianas gtt, cct y gcc palindrómicas (y otras dianas de KTG y QAN) también se escindieron, probablemente como resultado de la formación de especies homodiméricas pero las dianas no relacionadas no se escindieron. Además, un semisitio gtt, cct o gcc no era suficiente para permitir la escisión, puesto que dichas dianas eran completamente resistentes (véase ggg/gcc, gat/gcc, gcc/tac y muchos otros, en la Figura 12). La escisión inesperada se observó solamente con gtc/cct y gtt/gtc, con homodímeros KTG y QAN, respectivamente, pero la señal permaneció muy débil. Así pues, la escisión eficaz requiere la unión cooperativa de dos monómeros mutantes. Estos resultados demuestran un buen nivel de especificidad para las especies heterodiméricas.

En total, un total de 112 combinaciones de 14 proteínas diferentes se ensayaron en levadura y el 37,5% de las combinaciones (42/112) reveló una señal positiva en su diana quimérica predicha. Los datos cuantitativos se muestran para seis ejemplos en la figura 13a y para las mismas seis combinaciones los resultados se confirmaron en células CHO en experimentos de co-transfección transitoria, con un subconjunto de dianas relevantes (Figura 13 b) Como regla general, los heterodímeros funcionales se obtuvieron siempre cuando una de las dos proteínas expresadas daba una señal fuerte como homodímero. Por ejemplo DRN y RRN, dos mutantes de actividad baja, dieron heterodímeros funcionales con cortadores fuertes tales como KTG o QRR (Figura 13a y 113b) mientras que no se pudo detectar escisión de dianas quiméricas mediante la co-expresión de los mismos mutantes débiles.

Ejemplo 6: Escisión de una diana de ADN natural mediante heterodímero ensamblado

A) Materiales y Métodos

a) Examen genómico

Una diana natural potencialmente escindida por una variante de I-CreI, se identificó mediante exploración de las bases de datos públicas para secuencias genómicas que coinciden con el patrón caaaacnnnnnnnnngtttg, en el que n es a, t, c, o g (SEC ID N°: 78). La secuencia de ADN diana natural caaaactatgtgagggtttg (SEC ID N°: 75) se identificó en el cromosoma 17 de ratón.

Esta secuencia de ADN se escinde potencialmente por una combinación de dos variantes de I-CreI que escinde las secuencias tcaaaactatgtgaatagtttga (SEC ID N°: 76) y tcaaaaccctgtgaagggtttga (SEC ID N°: 77), respectivamente.

b) Aislamiento de variantes de meganucleasa

Las variantes se seleccionaron del ensayo de recombinación inducido por escisión en levadura, como se ha descrito en el ejemplo 1, usando la secuencia tcaaaactatgtgaatagtttga (SEC ID N°: 76) o la secuencia tcaaaaccctgtgaagggtttg (SEC ID N°: 77) como dianas.

c) Construcción del plásmido diana

Se diseñaron Oligonucleótidos que contenían dos semisitios diferentes de cada palíndromo mutante (PROLIGO). El ADN diana bicatenario, generado por amplificación por PCR de los oligonucleótidos monocatenarios, se clonó usando el protocolo Gateway (INVITROGEN) en el vector informador de mamíferos pcDNA3.1-LACURAZ-AURA, descrito anteriormente (Epinat *et al.*, citado anteriormente), para generar el plásmido diana LagoZ.

d) Construcción del vector de expresión de meganucleasa

Los marcos abiertos de lectura (ORF) de los clones identificados durante la exploración en levadura se amplificaron mediante PCR en una colonia de levaduras y se clonaron individualmente en el vector de expresión de CHO pCDNA6.2 (INVITROGEN), como se ha descrito en el ejemplo 1. Las variantes de I-*CreI* se expresaron bajo el control de un promotor CMV.

e) Ensayo de células de mamífero

La línea celular CHO-K1 se co-transfectó transitoriamente con cantidades equimolares del plásmido diana LagoZ y los plásmidos de expresión y se midió la actividad beta galactosidasa como se ha descrito en los ejemplos 3 y 5.

B) Resultados

Una diana de ADN natural, escindida potencialmente por variantes de I-*CreI* se identificó mediante la realización de un examen genómico de secuencias que coinciden con el patrón caaacnntnnnnnnngtrigt (SEC ID N°: 78). Una secuencia de ADN seleccionada aleatoriamente (SEC ID N°: 78) identificada en el cromosoma 17 del ratón se clonó en un plásmido informador. Esta diana de ADN se escindía potencialmente por una combinación de variantes de I -*CreI* A44,R68,S70,N75 (ARS) y K44,R68,E70,N75 (KRE).

La co-expresión de estas dos variantes en células CHO conduce a la formación de proteína heterodimérica funcional como se muestra en la figura 14. Es más, cuando las variantes de I-*CreI* se expresaron individualmente, virtualmente no se pudo detectar ninguna actividad de escisión en la diana de ADN de ratón aunque la proteína KRE mostró una actividad residual. Por el contrario, cuando estas dos variantes se co-expresaron junto con el plásmido que porta la diana potencial, se pudo medir una actividad de la galactosidasa fuerte. Todos estos datos juntos revelaron que se produjo heterodimerización en las células CHO y que los heterodímeros eran funcionales.

Estos datos demuestran que las proteínas heterodiméricas creadas mediante el ensamblaje de variantes homodiméricas, extiende la lista de secuencias diana de ADN de origen natural a todas las dianas escindibles híbridas potenciales que resultan de todas las posibles combinaciones de las variantes.

Además, estos datos demostraron que es posible predecir las secuencias de ADN que pueden escindirse por una combinación de variantes conociendo su diana individual de ADN de homodímero. Además, los nucleótidos de las posiciones 1 y 2 (y -1 y -2) de la diana pueden ser diferentes de gtac, lo que indica que juegan un papel pequeño en la interacción ADN/proteína.

Referencias

- Belfort, M. y Roberts, R.J. (1997) Homing endonucleases: keeping the house in order. *Nucleic Acids Res*, 25, 3379-3388.
- Bell-Pedersen, D., Quirk, S., Clyman, J. y Belfort, M. (1990) Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. *Nucleic Acids Res*, 18, 3763-3770.

- Bell-Pedersen, D., Quirk, S.M., Aubrey, M. y Belfort, M. (1989) A site-specific endonuclease and co-conversion of flanking exons associated with the mobile td intron of phage T4. *Gene*, 82, 119-126.
- Bell-Pedersen, D., Quirk, S.M., Bryk, M. y Belfort, M. (1991) I-TevI, the endonuclease encoded by the mobile td intron, recognizes binding and cleavage domains on its DNA target. *Proc Natl Acad Sci U S A*, 88, 7719-7723.
- 5 Bibikova, M., Beumer, K., Trautman, J.K. y Carroll, D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science*, 300, 764.
- Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.G. y Chandrasegaran, S. (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol*, 21, 289-297.
- 10 Bibikova, M., Golic, M., Golic, K.G. y Carroll, D. (2002) Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc- finger nucleases. *Genetics*, 161, 1169-1175.
- Chevalier, B., Sussman, D., Otis, C., Noël A.J., Turmel, M., Lemieux, C., Stephens, K., Monnat, R.Jr., Stoddard, B.L., (2004) Metal-Dependant DNA cleavage mechanism of the I-CreI LAGLIDADG homing endonuclease. *Biochemistry*, 43, 14015-14026.
- Chevalier, B.S., Kortemme, T., Chadsey, M.S., Baker, D., Monnat, R.J. y Stoddard, B.L. (2002) Design, 15 activity, and structure of a highly specific artificial endonuclease. *Mol Cell*, 10, 895-905.
- Chevalier, B.S. y Stoddard, B.L. (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res*, 29, 3757-3774.
- Choulika, A., Perrin, A., Dujon, B. y Nicolas, J.F. (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol*, 15, 1968-1973.
- 20 Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjou, F., Babinet, C., Louvard, D. y Jaisser, F. (1998) I-SceI-induced gene replacement at a natural locus in embryonic stem cells. *Mol Cell Biol*, 18, 1444-1448.
- Colleaux, L., D'Auriol, L., Galibert, F. y Dujon, B. (1988) Recognition and cleavage site of the intron-encoded omega transposase. *Proc Natl Acad Sci U S A*, 85, 6022-6026.
- Donoho, G., Jasin, M. y Berg, P. (1998) Analysis of gene targeting and intrachromosomal homologous 25 recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol*, 18, 4070-4078.
- Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J.A. y Jasin, M. (1998) Gene conversion tracts from double-strand break repair in mammalian cells. *Mol Cell Biol*, 18, 93-101.
- Epinat, J.C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., Patin, A., Zanghellini, A., Paques, F. y Lacroix, E. (2003) A novel engineered meganuclease induces homologous recombination in 30 yeast and mammalian cells. *Nucleic Acids Res*, 31, 2952-2962.
- Gietz, R.D. y Woods, R.A. (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/poly-ethylene glycol method. *Methods Enzymol*, 350, 87-96.
- Haber, J.E. (1998) Mating-type gene switching in *Saccharomyces cerevisiae*. *Annu Rev Genet*, 32, 561-599.
- Jacquier, A. y Dujon, B. (1985) An intron-encoded protein is active in a gene conversion process that 35 spreads an intron into a mitochondrial gene. *Cell*, 41, 383-394.
- Jurica, M.S., Monnat, R.J., Jr. y Stoddard, B.L. (1998) DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-CreI. *Mol Cell*, 2, 469-476.

- Klar, A.J., Strathern, J.N. y Abraham, J.A. (1984) Involvement of double-strand chromosomal breaks for mating-type switching in *Saccharomyces cerevisiae*. *Cold Spring Harb Symp Quant Biol*, 49, 77-88.
- Kostriken, R., Strathern, J.N., Klar, A.J., Hicks, J.B. y Heffron, F. (1983) A site-specific endonuclease essential for mating-type switching in *Saccharomyces cerevisiae*. *Cell*, 35, 167-174.
- 5 Liang, F., Han, M., Romanienko, P.J. y Jasin, M. (1998) Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc Natl Acad Sci U S A*, 95, 5172-5177.
- Moynahan, M.E. y Jasin, M. (1997) Loss of heterozygosity induced by a chromosomal double-strand break. *Proc Natl Acad Sci U S A*, 94, 8988-8993.
- Mueller, J.E., Smith, D. y Belfort, M. (1996) Exon coconversion biases accompanying intron homing:
10 battle of the nucleases. *Genes Dev*, 10, 2158-2166.
- Perrin, A., Buckle, M. y Dujon, B. (1993) Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions. *Embo J*, 12, 2939-2947.
- Plessis, A., Perrin, A., Haber, J.E. y Dujon, B. (1992) Site-specific recombination determined by I-SceI, a mito-chondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics*, 130, 451-
15 460.
- Porteus, M.H. y Baltimore, D. (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science*, 300, 763.
- Posfai, G., Kolisnychenko, V., Berezki, Z. y Blattner, F.R. (1999) Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. *Nucleic Acids Res*, 27, 4409-4415.
- Puchta, H. (1999) Double-strand break-induced recombination between ectopic homologous
20 sequences in somatic plant cells. *Genetics*, 152, 1173-1181.
- Puchta, H., Dujon, B. y Hohn, B. (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res*, 21, 5034-5040.
- Puchta, H., Dujon, B. y Hohn, B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc Natl Acad Sci U S A*, 93,
25 5055-5060.
- Richardson, C., Moynahan, M.E. y Jasin, M. (1998) Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. *Genes Dev*, 12, 3831-3842.
- Rong, Y.S. y Golic, K.G. (2000) Gene targeting by homologous recombination in *Drosophila*. *Science*, 288, 2013-2018.
- Rong, Y.S. y Golic, K.G. (2001) A targeted gene knockout in *Drosophila*. *Genetics*, 157, 1307-1312.
- 30 Rong, Y.S., Titen, S.W., Xie, H.B., Golic, M.M., Bastiani, M., Bandyopadhyay, P., Olivera, B.M., Brodsky, M., Rubin, G.M. y Golic, K.G. (2002) Targeted mutagenesis by homologous recombination in *D. melanogaster*. *Genes Dev*, 16, 1568-1581.
- Rouet, P., Smith, F. y Jasin, M. (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol*, 14, 8096-8106.
- Seligman et al., 1997,
35 *Genetics*, 1997, 147, 1653-1664
- Seligman, L.M., Chisholm, K.M., Chevalier, B.S., Chadsey, M.S., Edwards, S.T., Savage, J.H. y Veillet, A.L. (2002) Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res*, 30, 3870-3879.
- Siebert, R. y Puchta, H. (2002) Efficient repair of genomic double-strand breaks by

homologous recombination between directly repeated sequences in the plant genome. *Plant Cell*, 14, 1121-1131.

Sussman, D., Chadsey, M., Fauce, S., Engel, A., Bruett, A., Monnat, R., Jr., Stoddard, B.L. y Seligman, L.M. (2004) Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J Mol Biol*, 342, 31-41.

Thierry, A. y Dujon, B. (1992) Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes. *Nucleic Acids Res*, 20, 5625-5631.

LISTA DE SECUENCIAS

10 <110> Celectis

Duchateau, Philippe

Pâques, Frédéric

<120> Variantes de meganucleasa I-CreI con especificidad modificada, método de preparación y usos de las mismas.

<130> F1546/5PCT

<160> 81

20

<170> PatentIn versión 3.1

<210> 1

<211> 24

25 <212> ADN

<213> Secuencia Artificial

<220>

<223> ggg

30

<400> 1

tcaaacggg gtaccocgtt ttga 24

<210> 2

35 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

<223> gga

<400> 2

5 taaaacgga gtactccggt ttga 24

<210> 3

<211> 24

<212> ADN

10 <213> Secuencia Artificial

<220>

<223> ggt

15 <400> 3

taaaacggt gtacaccggt ttga 24

<210> 4

<211> 24

20 <212> ADN

<213> Secuencia Artificial

<220>

<223> ggc

25

<400> 4

taaaacggc gtacgccggt ttga 24

<210> 5

30 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

35 <223> gag

<400> 5

taaaacgag gtacctcggt ttga 24

- <210> 6
<211> 24
<212> ADN
5 <213> Secuencia Artificial
- <220>
<223> gaa
- 10 <400> 6
tcaaaacgaa gtacttcggt ttga 24
- <210> 7
<211> 24
15 <212> ADN
<213> Secuencia Artificial
- <220>
<223> gat
- 20 <400> 7
tcaaaacgat gtacatcggt ttga 24
- <210> 8
25 <211> 24
<212> ADN
<213> Secuencia Artificial
- <220>
30 <223> gac
- <400> 8
tcaaaacgac gtacgctggt ttga 24
- 35 <210> 9
<211> 24
<212> ADN
<213> Secuencia Artificial

<220>

<223> gtg

5 <400> 9

tcaaaacgtg gtaccacggt ttga 24

<210> 10

<211> 24

10 <212> ADN

<213> Secuencia Artificial

<220>

<223> gta

15

<400> 10

tcaaaacgta gtactacggt ttga 24

<210> 11

20 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

25 <223> gtt

<400> 11

tcaaaacggt gtacaacggt ttga 24

30 <210> 12

<211> 24

<212> ADN

<213> Secuencia Artificial

35 <220>

<223> gtc (= C1221)

<400> 12

tcaaaacgtc gtacgacggt ttga 24

<210> 13

<211> 24

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> gcg

10

<400> 13

tcaaaacgcg gtaccgctt ttga 24

<210> 14

15 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

20 <223> gca

<400> 14

tcaaaacgca gtactgcggt ttga 24

25 <210> 15

<211> 24

<212> ADN

<213> Secuencia Artificial

30 <220>

<223> gct

<400> 15

tcaaaacgct gtacgacggt ttga 24

35

<210> 16

<211> 24

<212> ADN

<213> Secuencia Artificial

<220>

<223> gcc

5

<400> 16

tcaaacgcc gtacggcgtt ttga 24

<210> 17

10 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

15 <223> agg

<400> 17

tcaaacagg gtaccctgtt ttga 24

20 <210> 18

<211> 24

<212> ADN

<213> Secuencia Artificial

25 <220>

<223> aga

<400> 18

tcaaacaga gtactctgtt ttga 24

30

<210> 19

<211> 24

<212> ADN

<213> Secuencia Artificial

35

<220>

<223> agt

<400> 19

tcaaaacagt gtacactggt ttga 24

<210> 20

5 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

10 <223> agc

<400> 20

tcaaaacagc gtacgctggt ttga 24

15 <210> 21

<211> 24

<212> ADN

<213> Secuencia Artificial

20 <220>

<223> aag

<400> 21

tcaaaacaag gtacctggt ttga 24

25

<210> 22

<211> 24

<212> ADN

<213> Secuencia Artificial

30

<220>

<223> aaa

<400> 22

35 tcaaaacaaa gtacttgggt ttga 24

<210> 23

<211> 24

<212> ADN

<213> Secuencia Artificial

<220>

5 <223> aat

<400> 23

tcaaacaat gtacattgtt ttga 24

10 <210> 24

<211> 24

<212> ADN

<213> Secuencia Artificial

15 <220>

<223> aac

<400> 24

tcaaacaac gtacgttgtt ttga 24

20

<210> 25

<211> 24

<212> ADN

<213> Secuencia Artificial

25

<220>

<223> atg

<400> 25

30 taaaacatg gtacatgtt ttga 24

<210> 26

<211> 24

<212> ADN

35 <213> Secuencia Artificial

<220>

<223> ata

<400> 26

tcaaaacata gtactatggt ttga 24

5 <210> 27

<211> 24

<212> ADN

<213> Secuencia Artificial

10 <220>

<223> att

<400> 27

tcaaaacatt gtacaatggt ttga 24

15

<210> 28

<211> 24

<212> ADN

<213> Secuencia Artificial

20

<220>

<223> atc

<400> 28

25 tcaaaacatc gtacgatggt ttga 24

<210> 29

<211> 24

<212> ADN

30 <213> Secuencia Artificial

<220>

<223> acg

35 <400> 29

tcaaaacacg gtaccgtggt ttga 24

<210> 30

<211> 24

<212> ADN

<213> Secuencia Artificial

5 <220>

<223> aca

<400> 30

tcaaaacaca gtactgtgtt ttga 24

10

<210> 31

<211> 24

<212> ADN

<213> Secuencia Artificial

15

<220>

<223> act

<400> 31

20 tcaaaacact gtacagtgtt ttga 24

<210> 32

<211> 24

<212> ADN

25 <213> Secuencia Artificial

<220>

<223> acc

30 <400> 32

tcaaaacacc gtacggtgtt ttga 24

<210> 33

<211> 24

35 <212> ADN

<213> Secuencia Artificial

<220>

<223> tgg

<400> 33

tcaaaactgg gtaccagtt ttga 24

5

<210> 34

<211> 24

<212> ADN

<213> Secuencia Artificial

10

<220>

<223> tga

<400> 34

15 tcaaaactga gtactcagtt ttga 24

<210> 35

<211> 24

<212> ADN

20 <213> Secuencia Artificial

<220>

<223> tgt

25 <400> 35

tcaaaactgt gtacacagtt ttga 24

<210> 36

<211> 24

30 <212> ADN

<213> Secuencia Artificial

<220>

<223> tgc

35

<400> 36

tcaaaactgc gtacgcagtt ttga 24

<210> 37

<211> 24

<212> ADN

<213> Secuencia Artificial

5

<220>

<223> tag

<400> 37

10 tcaaaactag gtacctagtt ttga 24

<210> 38

<211> 24

<212> ADN

15 <213> Secuencia Artificial

<220>

<223> taa

20 <400> 38

tcaaaactaa gtacttagtt ttga 24

<210> 39

<211> 24

25 <212> ADN

<213> Secuencia Artificial

<220>

<223> tat

30

<400> 39

tcaaaactat gtacatagtt ttga 24

<210> 40

35 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

<223> tac

<400> 40

5 tcaaaactac gtacgtagt ttga 24

<210> 41

<211> 24

<212> ADN

10 <213> Secuencia Artificial

<220>

<223> ttg

15 <400> 41

tcaaaacttg gtaccaagt ttga 24

<210> 42

<211> 24

20 <212> ADN

<213> Secuencia Artificial

<220>

<223> tta

25

<400> 42

tcaaaactta gtactaagt ttga 24

<210> 43

30 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

35 <223> ttt

<400> 43

tcaaaacttt gtacaaagt ttga 24

<210> 44

<211> 24

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ttc

10 <400> 44

tcaaaacttc gtacgaagt ttga 24

<210> 45

<211> 24

15 <212> ADN

<213> Secuencia Artificial

<220>

<223> tcg

20

<400> 45

tcaaaactcg gtaccgagtt ttga 24

<210> 46

25 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

30 <223> tca

<400> 46

tcaaaactca gtactgagtt ttga 24

35 <210> 47

<211> 24

<212> ADN

<213> Secuencia Artificial

<220>

<223> tct

5 <400> 47

tcaaaactct gtacagagtt ttga 24

<210> 48

<211> 24

10 <212> ADN

<213> Secuencia Artificial

<220>

<223> tcc

15

<400> 48

tcaaaactcc gtaccgagtt ttga 24

<210> 49

20 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

25 <223> cgg

<400> 49

tcaaaaccgg gtaccgggtt ttga 24

30 <210> 50

<211> 24

<212> ADN

<213> Secuencia Artificial

35 <220>

<223> cga

<400> 50

tcaaaaccga gtactcgggt ttga 24

<210> 51

<211> 24

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> cgt

10

<400> 51

tcaaaaccgt gtacacgggt ttga 24

<210> 52

15 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

20 <223> cgc

<400> 52

tcaaaaccgc gtacgcgggt ttga 24

25 <210> 53

<211> 24

<212> ADN

<213> Secuencia Artificial

30 <220>

<223> cag

<400> 53

tcaaaaccag gtacctgggt ttga 24

35

<210> 54

<211> 24

<212> ADN

<213> Secuencia Artificial

<220>

<223> caa

5

<400> 54

tcaaaaccaa gtacttggt ttga 24

<210> 55

10 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

15 <223> cat

<400> 55

tcaaaaccat gtacatggt ttga 24

20 <210> 56

<211> 24

<212> ADN

<213> Secuencia Artificial

25 <220>

<223> cac

<400> 56

tcaaaaccac gtacgtggt ttga 24

30

<210> 57

<211> 24

<212> ADN

<213> Secuencia Artificial

35

<220>

<223> ctg

<400> 57

tcaaaacctg gtaccagggt ttga 24

<210> 58

5 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

10 <223> cta

<400> 58

tcaaaacctg gtactagggt ttga 24

15 <210> 59

<211> 24

<212> ADN

<213> Secuencia Artificial

20 <220>

<223> ctt

<400> 59

tcaaaacctt gtacaagggt ttga 24

25

<210> 60

<211> 24

<212> ADN

<213> Secuencia Artificial

30

<220>

<223> ctc

<400> 60

35 tcaaaacctc gtacgagggt ttga 24

<210> 61

<211> 24

<212> ADN

<213> Secuencia Artificial

<220>

5 <223> ccg

<400> 61

tcaaaaccg gtaccgggtt ttga 24

10 <210> 62

<211> 24

<212> ADN

<213> Secuencia Artificial

15 <220>

<223> cca

<400> 62

tcaaaacca gtactgggtt ttga 24

20

<210> 63

<211> 24

<212> ADN

<213> Secuencia Artificial

25

<220>

<223> cct

<400> 63

30 tcaaaaccct gtacagggtt ttga 24

<210> 64

<211> 24

<212> ADN

35 <213> Secuencia Artificial

<220>

<223> ccc

<400> 64
tcaaaacccc gtacggggtt ttga 24

5 <210> 65
<211> 24
<212> ADN
<213> Secuencia Artificial

10 <220>
<223> C1234 (=wt)

<400> 65
tcaaaacgtc gtgagacagt ttgg 24

15
<210> 66
<211> 24
<212> ADN
<213> Secuencia Artificial

20
<220>
<223> C4334

<400> 66
25 ccaaactgtc tcgagacagt ttgg 24

<210> 67
<211> 50
<212> ADN

30 <213> Secuencia Artificial

<220>
<223> Cebador vvk (directo)

35 <220>
<221> misc_feature
<223> k es g o t
v es a o g o c

<400> 67

gtttaacat cagctaagct tgaccttvv kgtgactca aaagaccag 50

5 <210> 68

<211> 45

<212> ADN

<213> Secuencia Artificial

<220>

10 <223> Cebador mbb (inverso)

<220>

<221> misc_feature

<223> b e s g o t o c

15 m e s a o c

<400> 68

gatgtagtg gaaacggatc cmbbatcmbb tacgtaacca acgcc 45

20 <210> 69

<211> 501

<212> ADN

<213> Secuencia Artificial

25 <220>

<223> CDS de proteína de armazón I-CreI N75

<220>

<221> CDS

30 <222> (1)..(501)

<223>

<400> 69

atg gcc aat acc aaa tat aac aaa gag ttc ctg ctg tac ctg gcc ggc	48
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly	
1 5 10 15	
ttt gtg gac ggt gac ggt agc atc atc gct cag att aaa cca aac cag	96
Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln	
20 25 30	
tct tat aag ttt aaa cat cag cta agc ttg acc ttt cag gtg act caa	144
Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln	
35 40 45	
aag acc cag cgc cgt tgg ttt ctg gac aaa cta gtg gat gaa att ggc	192
Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly	
50 55 60	
gtt ggt tac gta cgt gat cgc gga tcc gtt tcc aac tac atc tta agc	240
Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asn Tyr Ile Leu Ser	
65 70 75 80	
gaa atc aag ccg ctg cac aac ttc ctg act caa ctg cag ccg ttt ctg	288
Glu Ile Lys Pro Leu His Asn Phe Leu Thr 90 Gln Leu Gln Pro Phe Leu	
85 95	
aaa ctg aaa cag aaa cag gca aac ctg gtt ctg aaa att atc gaa cag	336
Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln	
100 105 110	
ctg ccg tct gca aaa gaa tcc ccg gac aaa ttc ctg gaa gtt tgt acc	384
Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr	
115 120 125	
tgg gtg gat cag att gca gct ctg aac gat tct aag acg cgt aaa acc	432
Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr	
130 135 140	
act tct gaa acc gtt cgt gct gtg ctg gac agc ctg agc gag aag aag	480
Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys	
145 150 155 160	
aaa tcc tcc ccg gcg gcc gac	501
Lys Ser Ser Pro Ala Ala Asp	
165	

<210> 70

<211> 167

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Proteína de armazón I-CreI N75

10

<400> 70

Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
 1 5 10 15
 Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
 20 25 30
 Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
 35 40 45
 Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
 50 55 60
 Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asn Tyr Ile Leu Ser
 65 70 75 80
 Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu
 85 90 95
 Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln
 100 105 110
 Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr
 115 120 125
 Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr
 130 135 140
 Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys
 145 150 155 160
 Lys Ser Ser Pro Ala Ala Asp
 165

<210> 71

<211> 22

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> Sitio de movimiento dirigido I-CreI

10

<400> 71

caaaacgtcg tgagacagtt tg 22

<210> 72

15 <211> 49

<212> ADN

<213> Secuencia Artificial

<220>

<223> oligonucleótido

5 <220>

<221> misc_feature

<222> (20)..(22)

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

10 <220>

<221> misc_feature

<222> (27)..(29)

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

15 <400> 72

ggcatacaag ttcaaaaacn nngtacnng tttgacaat cgtctgtca 49

<210> 73

<211> 77

20 <212> ADN

<213> Secuencia Artificial

<220>

<223> oligonucleótido

25

<400> 73

ggggacaagt ttgtacaaaa aagcaggctt cgaaggagat agaaccatgg ccaataccea 60

atataacaaa gagttcc 77

30 <210> 74

<211> 64

<212> ADN

<213> Secuencia Artificial

35 <220>

<223> oligonucleótido

<400> 74

ggggaccact ttgtacaaga aagctggggtt tagtcggccg ccggggagga tttcttcttc 60
 tcgc 64

<210> 75

<211> 22

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> oligonucleótido

10

<400> 75

caaaactatg tagagggtt tg 22

<210> 76

15 <211> 24

<212> ADN

<213> Secuencia Artificial

<220>

20 <223> oligonucleótido

<400> 76

tcaaaactat gtgaatagtt ttga 24

25 <210> 77

<211> 24

<212> ADN

<213> Secuencia Artificial

30 <220>

<223> oligonucleótido

<400> 77

tcaaaaccct gtgaagggtt ttga 24

35

<210> 78

<211> 22

<212> ADN

<213> Secuencia Artificial

<220>

<223> oligonucleótido

5

<220>

<221> misc_feature

<222> (7)..(16)

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

10

<400> 78

caaaacnnnn nnnnngttt tg 22

<210> 79

15

<211> 24

<212> ADN

<213> Secuencia Artificial

<220>

20

<223> oligonucleótido

<400> 79

taaaacggt gtacaacgt ttga 24

25

<210> 80

<211> 24

<212> ADN

<213> Secuencia Artificial

30

<220>

<223> oligonucleótido

<400> 80

taaaacggt gtacagggtt ttga 24

35

<210> 81

<211> 163

<212> PRT

<213> chlamydomonas reinhardtii

<400> 81

Met Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe
 1 5 10 15

Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser
 20 25 30

Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys
 35 40 45

Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val
 50 55 60

Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu
 65 70 75 80

Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys
 85 90 95

Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu
 100 105 110

Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp
 115 120 125

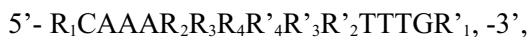
Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr
 130 135 140

Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys
 145 150 155 160

Ser Ser Pro

REIVINDICACIONES

1. Método de preparación de una variante de meganucleasa I-*CreI* que tiene una
 5 especificidad de escisión modificada, comprendiendo dicho método:
- (a) reemplazar los aminoácidos Q44, R68 y/o R70, en referencia al código de acceso pdb de I-*CreI* Ig9y, con un aminoácido seleccionado del grupo que consiste en A, D, E, G, H, K, N, P, Q, R, S, T e Y;
- (b) seleccionar las variantes de meganucleasa I-*CreI* obtenidas en el paso (a) que tienen al
 10 menos uno de los siguientes perfiles de escisión de tripletes R₃ en referencia a las posiciones -5 a -3 en una diana de ADN bicatenario, correspondiendo dichas posiciones -5 a -3 a R₃ de la siguiente fórmula I:



en la que:

- 15 R₁ está ausente o presente; y cuando está presente representa un fragmento de ácido nucleico que comprende de 1 a 9 nucleótidos que corresponden a una secuencia de ácido nucleico aleatoria o a un fragmento del sitio de movimiento dirigido de meganucleasas I-*CreI* situado desde la posición -20 a -12 (de 5' a 3'), correspondiendo R₁ al menos a la posición -12 de dicho sitio de movimiento dirigido,
- R₂ representa el doblete de ácido nucleico ac o ct y corresponde a las posiciones -7 a -6 de dicho sitio de
 20 movimiento dirigido,
- R₃ representa un triplete de ácido nucleico que corresponde a dichas posiciones -5 a -3, seleccionadas de entre g, t, c y a, excepto los siguientes tripletes: gtc, gcc, gtg, gtt y gct,
- R₄ representa el doblete de ácido nucleico gt o tc y corresponde a las posiciones -2 a -1 de dicho sitio de
 movimiento dirigido,
- 25 R'₁ está ausente o presente; y cuando está presente representa un fragmento de ácido nucleico que comprende de 1 a 9 nucleótidos que corresponden a una secuencia de ácido nucleico aleatoria o a un fragmento de un sitio de movimiento dirigido de meganucleasa I-*CreI* situado desde la posición +12 a +20 (de 5' a 3'), correspondiendo R'₁ al menos a la posición +12 de dicho sitio de movimiento dirigido,
- R'₂ representa el doblete de ácido nucleico ag o gt y corresponde a las posiciones +6 a +7 de dicho sitio de
 30 movimiento dirigido,
- R'₃ representa un triplete de ácido nucleico que corresponde a dichas posiciones +3 a +5, seleccionadas de entre g, t, c y a; siendo R'₃ diferente de gac, ggc, cac, aac y agc, cuando R₃ y R'₃ no son palindrómicos,
- R'₄ representa el doblete de ácido nucleico ga o ac y corresponde a las posiciones +1 a +2 de dicho sitio de
 movimiento dirigido.

35

2. Método de acuerdo con la reivindicación 1, **caracterizado por que** dicho triplete R₃ de ácido nucleico se selecciona preferiblemente de entre los siguientes tripletes: ggg., gga, ggt, ggc, gag, gaa, gat, gac, gta, gcg, gca, tgg, tga, tgt, tgc, tag, taa, tat, tac, ttg, tta, ttt, ttc, tcg, tca, tct, tcc, agg, aga, agt, agc,

aag, aaa, aat, aac, atg, ata, att, atc, acg, aca, act, acc, cgg, cga, cgt, cgc, cag, caa, cat, cac, ctg, cta, ctt, etc, ccg, cca, cct y ccc y más preferiblemente de entre los siguientes tripletes: ggg., ggt, ggc, gag, gat, gac, gta, gcg, gca, tag, taa, tat, tac, ttg, ttt, ttc, tcg, tct, tcc, agg, aag, aat, aac, att, atc, act, acc, cag, cat, cac, ctt, etc, ccg, cct y ccc.

5

3. Método de acuerdo con la reivindicación 1 o reivindicación 2, **caracterizado por que** la variante de meganucleasa I-*CreI* obtenida en el paso (b) se selecciona del grupo que consiste en:

A44/A68/A70,

10 A44/A68/G70,A44/A68/H70,A44/A68/K70,A44/A68/N70,A44/A68/Q70,A44/G68/H70,A44/H68/H70,A4
4/K68/A70,A44/K68/570,A44/N68/N70,A44/Q68/G70,A44/R68/E70,A44/R68/S70,A44/S68/R70,A44/T6
8/N70,D44/R68/A70,E44/H68/H70,G44/H68/K70,H44/A68/S70,H44/R68/R70,H44/T68/T70,K44/A68/Q
70,K44/E68/S70,K44/H68/E70,K44/K68/H70,K44/N68/N70,K44/Q68/E70,K44/R68/H70,K44/S68/H70,
K44/T68/G70,N44/A68/R70,N44/N68/R70,N44/R68/G70,N44/S68/G70,N44/T68/R70,Q44/G68/K70,A44
15 /A68/R70,A44/G68/K70,A44/H68/K70,A44/K68/G70,A44/K68/T70,A44/N68/Q70,A44/Q68/H70,A44/R
68/G70,A44/R68/T70,A44/S68/S70,A44/T68/Q70,D44/R68/K70,E44/R68/A70,G44/Q68/H70,H44/A68/T
70H44/R68/S70,K44/A68/A70,K44/A68/S70,K44/G68/A70,K44/H68/G70,K44/K68/T70,K44/N68/Q70;
K44/Q68/S70,K44/R68/N70,K44/S68/N70,K44/T68/H70,N44/H68/N70,N44/P68/D70,N44/R68/H70,N44
/S68/H70,N44/T68/S70,Q44/G68/R70,A44/A68/S70, A44/G68/N70, A44/H68/N70, A44/K68/H70,
A44/N68/A70, A44/N68/R70, A44/Q68/N70, A44/R68/H70, A44/S68/A70, A44/S68/T70, A44/T68/R70,
20 D44/R68/N70, E44/R68/H70, G44/R68/Q70, H44/R68/A70, H44/R68/T70, K44/A68/D70, K44/A68/T70,
K44/G68/G70, K44/H68/N70, K44/N68/A70, K44/N68/S70, K44/Q68/T70, K44/R68/Q70, K44/S68/S70,
K44/T68/N70, N44/H68/R70, N44/Q68/H70, N44/R68/K70, N44/S68/K70, P44/N68/D70,
Q44/K68/G70,A44/A68/T70, A44/G68/P70, A44/H68/Q70, A44/K68/K70, A44/N68/E70, A44/N68/S70,
A44/Q68/R70, A44/R68/K70, A44/S68/G70, A44/T68/A70, A44/T68/S70, D44/R68/Q70, E44/R68/N70,
25 G44/R68/R70, H44/R68/D70, H44/S68/G70, K44/A68/E70, K44/D68/A70, K44/G68/N70, K44/H68/S70,
K44/N68/D70, K44/N68/T70, K44/R68/A70, K44/R68/S70, K44/S68/T70, K44/T68/Q70, N44/K68/G70,
N44/Q68/R70, N44/R68/N70, N44/S68/R70, P44/T68/T70,A44/D68/H70, A44/G68/R70, A44/H68/R70,
A44/K68/N70, A44/N68/G70, A44/N68/T70, A44/Q68/S70, A44/R68/L70, A44/S68/K70, A44/T68/G70,
A44/T68/T70, D44/R68/R70, E44/R68/S70, G44/T68/D70, H44/R68/E70, H44/S68/S70, K44/A68/G70,
30 K44/D68/T70, K44/G68/S70, K44/H68/T70, K44/N68/E70, K44/P68/H70, K44/R68/D70, K44/R68/T70,
K44/T68/A70, K44/T68/S70, N44/K68/H70, N44/R68/A70, N44/R68/R70, N44/T68/H70,
Q44/A68/A70,Q44/N68/A70, Q44/N68/H70,A44/D68/K70, A44/H68/A70, A44/H68/S70, A44/K68/Q70,
A44/N68/H70, A44/Q68/A70, A44/R68/A70, A44/R68/N70, A44/S68/N70, A44/T68/H70, D44/D68/H70,
D44/R68/S70, E44/R68/T70, G44/T68/P70, H44/R68/G70, H44/S68/T70, K44/A68/H70, K44/E68/G70,
35 K44/G68/T70, K44/K68/A70, K44/N68/G70, K44/Q68/A70, K44/R68/E70, K44/S68/A70, K44/T68/D70,
K44/T68/T70, N44/K68/R70, N44/R68/D70, N44/R68/S70, N44/T68/K70, Q44/A68/H70,
Q44/N68/S70,A44/D68/R70, A44/H68/G70, A44/H68/T70, A44/K68/R70, A44/N68/K70, A44/Q68/D70,
A44/R68/D70, A44/R68/R70, A44/S68/Q70, A44/T68/K70, D44/N68/S70, D44/R68/T70, E44/S68/T70,

G44/T68/R70, H44/R68/N70, H44/T68/S70, K44/A68/N70, K44/E68/N70, K44/H68/D70, K44/K68/D70, K44/N68/H70, K44/Q68/D70, K44/R68/G70, K44/S68/D70, K44/T68/E70, N44/A68/H70, N44/K68/S70, N44/R68/E70, N44/R68/T70, N44/T68/Q70, Q44/A68/R70, Q44/P68/P70, Q44/Q68/G70, Q44/R68/Q70, Q44/T68/H70, R44/N68/T70, R44/R68/S70, S44/H68/R70, T44/H68/R70, T44/R68/D70, Q44/R68/A70, Q44/R68/S70, Q44/T68/R70, R44/R68/A70, R44/R68/T70, S44/R68/G70, T44/K68/R70, T44/R68/E70, Q44/R68/D70, Q44/S68/H70, R44/A68/G70, R44/R68/D70, R44/S68/G70, S44/R68/N70, T44/N68/P70, T44/R68/G70, Q44/R68/E70, Q44/S68/R70, R44/A68/T70, R44/R68/E70, R44/S68/N70, S44/R68/R70, T44/N68/R70, T44/R68/H70, Q44/R68/G70, Q44/S68/S70, R44/G68/T70, R44/R68/G70, R44/S68/S70, S44/R68/S70, T44/Q68/K70, T44/R68/K70, Q44/R68/H70, Q44/T68/A70, R44/H68/D70, R44/R68/N70, R44/S68/T70, T44/A68/K70, T44/Q68/R70, T44/R68/N70, Q44/R68/N70, Q44/T68/G70, R44/H68/T70, R44/R68/Q70, S44/D68/K70, T44/A68/R70, T44/R68/A70, T44/R68/Q70, T44/R68/R70, T44/R68/S70, T44/R68/T70, T44/S68/K70, T44/S68/R70, T44/T68/K70, y T44/T68/R70.

4. Método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, **caracterizado por** que la etapa (b) de selección de dicha variante de meganucleasa I-CreI se realiza *in vivo* en células de levadura.

5. Uso de una variante de meganucleasa I-CreI, obtenible por el método de acuerdo con una cualquiera de las reivindicaciones 1 a 4, *in vitro* o *in vivo* para propósitos no terapéuticos, para la escisión de una diana de ácido nucleico bicatenaria que comprende al menos una secuencia parcialmente palindrómica de 20-24 pb, en la que al menos la secuencia en las posiciones +/- 8 a 11 es palindrómica y el triplete de nucleótidos en las posiciones -5 a -3 y/o el triplete de nucleótidos de las posiciones +3 a +5 difiere de gtc, gcc, gtg, gtt y gct y de gac, ggc, cac, aac y agc, respectivamente.

6. El uso de acuerdo con la reivindicación 5, **caracterizado por que** dicha variante de meganucleasa I-CreI se selecciona del grupo que consiste en A44/A68/A70, A44/A68/G70, A44/A68/H70, A44/A68/K70, A44/A68/N70, A44/A68/Q70, A44/A68/R70, A44/A68/S70, A44/A68/T70, A44/D68/H70, A44/D68/K70, A44/D68/R70, A44/G68/H70, A44/G68/K70, A44/G68/N70, A44/G68/P70, A44/G68/R70, A44/H68/A70, A44/H68/G70, A44/H68/H70, A44/H68/K70, A44/H68/N70, A44/H68/Q70, A44/H68/R70, A44/H68/S70, A44/H6870, A44/K68/A70, A44/K68/G70, A44/K68/H70, A44/K68/K70, A44/K68/N70, A44/K68/Q70, A44/K68/R70, A44/K68/S70, A44/K68/T70, A44/N68/A70, A44/N68/K70, A44/Q68/D70, A44/R68/D70, A44/R68/R70, A44/S68/Q70, A44/T68/K70, D44/N68/S70, D44/R68/T70, A44/N68/E70, A44/N68/S70, A44/Q68/R70, A44/R68/K70, A44/S68/G70, A44/T68/A70, A44/T68/S70, D44/R68/Q70, E44/R68/N70, G44/R68/R70, H44/R8/D70, H44/S68/G70, K44/A68/E70, K44/D68/A70, K44/G68/N70, K44/H68/S70, K44/N68/D70, K44/N68/T70, K44/R68/A70, K44/R68/S70, K44/S68/T70, K44/T68/Q70, N44/K68/G70, N44/Q68/R70, N44/R68/N70, N44/S68/R70, P44/T68/T70, Q44/N68/A70, Q44/R68/E70, Q44/S68/R70, R44/A68/T70, R44/R68/E70, R44/S68/N70, A44/N68/G70, A44/N68/T70, A44/Q68/S70, A44/R68/L70, A44/S68/K70, A44/T68/G70, A44/T68/T70, D44

/R68/R70,E44/R68/S70,G44/T68/D70,H44/R68/E70,H44/S68/S70,K44/A68/G70,K44/D68/T70,K44/G68/
 S70,K44/H68/T70,K44/N68/E70,K44/P68/H70,K44/R68/D70,K44/R68/T70,K44/T68/A70,K44/T68/S70,
 N44/K68/H70,N44/R68/A70,N44/R68/R70,N44/T68/H70,A44/N68/H70, A44/Q68/A70, A44/R68/A70,
 A44/R68/N70, A44/S68/N70, A44/T68/H70, D44/D68/H70, D44/R68/S70, E44/R68/T70, G44/T68/P70,
 5 H44/R68/G70, H44/S68/T70, K44/A68/H70, K44/E68/G70, K44/G68/T70, K44/K68/A70, K44/N68/G70,
 K44/Q68/A70, K44/R68/E70, K44/S68/A70, K44/T68/D70, K44/T68/T70, N44/K68/R70, N44/R68/D70,
 N44/R68/S70, N44/T68/K70,A44/N68/Q70, A44/Q68/H70, A44/R68/G70, A44/R68/T70, A44/S68/S70,
 A44/T68/Q70, D44/R68/K70, E44/R68/A70,Q44/A68/A70, Q44/A68/H70,Q44/N68/H70,
 Q44/N68/S70,Q44/R68/G70,Q44/S68/S70,R44/G68/T70,R44/R68/G70,R44/S68/S70,A44/N68/N70,A44/
 10 Q68/G70,A44/R68/E70,A44/R68/S70,A44/S68/R70,A44/T68/N70,D44/R68/A70,E44/H68/H70,E44/S68/
 T70, G44/H68/K70, G44/Q68/H70,G44/T68/R70, H44/A68/S70, H44/A68/T70,H44/R68/N70,
 H44/R68/R70, H44/R68/S70,H44/T68/S70, H44/T68/T70, K44/A68/A70,K44/A68/N70, K44/A68/Q70,
 K44/A68/S70,K44/E68/N70, K44/E68/S70, K44/G68/A70,K44/H68/D70, K44/H68/E70,
 K44/H68/G70,K44/K68/D70, K44/K68/H70, K44/K68/T70,K44/N68/H70, K44/N68/N70,
 15 K44/N68/Q70,K44/Q68/D70, K44/Q68/E70, K44/Q68/S70,K44/R68/G70, K44/R68/H70,
 K44/R68/N70,K44/S68/D70, K44/S68/H70, K44/S68/N70,K44/T68/E70, K44/T68/G70,
 K44/T68/H70,N44/A68/H70, N44/A68/R70, N44/H68/N70,N44/K68/S70, N44/N68/R70,
 N44/P68/D70,N44/R68/E70, N44/R68/G70, N44/R68/H70,N44/R68/T70, N44/S68/G70,
 N44/S68/H70,N44/T68/Q70, N44/T68/R70, N44/T68/S70,Q44/A68/R70, Q44/G68/K70,
 20 Q44/G68/R70,Q44/P68/P70, Q44/Q68/G70, Q44/R68/A70,Q44/R68/H70, Q44/R68/N70, Q44/R68/Q70,
 Q44/R68/S70,Q44/T68/A70, Q44/T68/G70, Q44/T68/H70, Q44/T68/R70,R44/H68/D70, R44/H68/T70,
 R44/N68/T70, R44/R68/A70,R44/R68/N70, R44/R68/Q70, R44/R68/S70, R44/R68/T70,R44/S68/T70,
 S44/D68/K70, S44/H68/R70, S44/R68/G70,A44/N68/R70, A44/Q68/N70, A44/R68/H70, A44/S68/A70,
 A44/S68/T70, A44/T68/R70, D44/R68/N70, E44/R68/H70, G44/R68/Q70, H44/R68/A70, H44/R68/T70,
 25 K44/A68/D70, K44/A68/T70, K44/G68/G70, K44/H68/N70, K44/N68/A70, K44/N68/S70, K44/Q68/T70,
 K44/R68/Q70, K44/S68/S70, K44/T68/N70, N44/H68/R70, N44/Q68/H70, N44/R68/K70, N44/S68/K70,
 P44/N68/D70, Q44/K68/G70, Q44/R68/D70, Q44/S68/H70, R44/A68/G70, R44/R68/D70, R44/S68/G70,
 S44/R68/N70,S44/R68/R70, S44/R68/S70, T44/A68/K70, T44/A68/R70, T44/H68/R70, T44/K68/R70,
 T44/N68/P70,T44/N68/R70, T44/Q68/K70, T44/Q68/R70, T44/R68/A70, T44/R68/D70, T44/R68/E70,
 30 T44/R68/G70,T44/R68/H70, T44/R68/K70, T44/R68/N70, T44/R68/Q70, T44/R68/R70, T44/R68/S70,
 T44/R68/T70, T44/S68/K70, T44/S68/R70, T44/T68/K70.

7. El uso de acuerdo con la reivindicación 5 o con la reivindicación 6 **caracterizado por que** dicha variante de meganucleasa I-CreI es un homodímero.

35

8. El uso de acuerdo con una cualquiera de las reivindicaciones 5 y 6, **caracterizado por que** dicha variante de meganucleasa I-CreI es un heterodímero.

9. El uso de acuerdo con una cualquiera de las reivindicaciones 5 a 8, **caracterizado por que** la secuencia de posiciones +/- 3 a 5 de dicha diana de ADN es palindrómica.
10. El uso, de acuerdo con una cualquiera de las reivindicaciones 5 a 6, **caracterizada por que** la secuencia en las posiciones +/- 3 a 5 de dicha diana de ADN es no palindrómica.
11. El uso de acuerdo con una cualquiera de las reivindicaciones 5 a 10, **caracterizado por que** dichas secuencias palindrómicas en las posiciones -11 a -8 y +8 a +11 son caaa y tttg, respectivamente.
- 10 12. El uso de acuerdo con una cualquiera de las reivindicaciones 5 a 11, **caracterizado por que** dicha variante de meganucleasa I-*CreI* comprende adicionalmente una mutación en la posición 75.
13. El uso de acuerdo con la reivindicación 12, **caracterizada por que** dicha mutación es D75N o D75V.
- 15 14. El uso de acuerdo con una cualquiera de las reivindicaciones 5 a 13, **caracterizado por que** dicha variante de meganucleasa I-*CreI* tiene una alanina (A) o una asparagina (N) en la posición 44 para escindir una diana de ADN que comprende un nucleótido a en una posición -4 y/o t en posición +4.
- 20 15. El uso de acuerdo con una cualquiera de las reivindicaciones 5 a 13, en el que dicha variante de meganucleasa tiene una lisina (K) en la posición 44 para escindir una diana que comprende un nucleótido c en la posición -4 y/o g en la posición +4.
- 25 16. El uso de acuerdo con una cualquiera de las reivindicaciones 6 a 13, en el que dicha variante de meganucleasa tiene una glutamina (Q) en la posición 44 para escindir una diana que comprende un nucleótido t en la posición -4 y/o a en la posición +4.
17. Una variante de meganucleasa I-*CreI*, obtenible por el método de preparación de acuerdo con las reivindicaciones 1 a 4; que se selecciona del grupo que consiste en: A44/A68/A70, A44/A68/G70, 30 A44/A68/H70, A44/A68/K70, A44/A68/N70, A44/A68/Q70, A44/A68/S70, A44/A68/T70, A44/G68/N70, A44/G68/P70, A44/H68/Q70, A44/H68/S70, A44/K68/Q70, A44/K68/R70, A44/N68/H70, A44/N68/K70, A44/Q68/A70, A44/Q68/D70, A44/R68/K70, A44/R68/L70, A44/S68/S70, A44/S68/T70, A44/T68/S70, A44/T68/T70, D44/R68/R70, D44/R68/S70, E44/R68/S70, E44/R68/T70, G44/T68/P70, G44/T68/R70, H44/R68/N70, H44/R68/R70, H44/T68/T70, K44/A68/A70, K44/A68/Q70, K44/D68/A70, K44/G68/N70, 35 K44/G68/S70, K44/H68/S70, K44/H68/T70, K44/N68/D70, K44/N68/E70, K44/N68/T70, K44/P68/H70, K44/R68/A70, K44/R68/D70, K44/S68/A70, K44/S68/D70, K44/T68/D70, K44/T68/E70, A44/D68/H70, A44/H68/A70, A44/H68/T70, A44/K68/S70, A44/N68/N70, A44/Q68/G70, A44/S68/A70, A44/T68/A70, D44/D68/H70, D44/R68/T70, E44/S68/T70, H44/A68/S70, H44/R68/S70, K44/A68/D70, K44/D68/T70,

K44/G68/T70, K44/K68/A70, K44/N68/G70, K44/Q68/A70, K44/R68/E70, K44/S68/H70,
 K44/T68/G70,A44/D68/K70, A44/H68/G70, A44/K68/A70, A44/K68/T70, A44/N68/Q70, A44/Q68/H70,
 A44/S68/G70, A44/T68/G70,A44/N68/S70, E44/H68/H70,G44/H68/K70, H44/A68/T70, H44/S68/G70,
 K44/A68/E70, K44/E68/G70, K44/H68/D70, K44/K68/D70, K44/N68/H70, K44/Q68/D70, K44/R68/G70,
 5 K44/S68/N70, K44/T68/H70,A44/D68/R70, A44/H68/H70, A44/K68/G70, A44/N68/A70, A44/N68/R70,
 A44/Q68/N70, A44/S68/N70, A44/T68/H70, D44/R68/A70, E44/R68/A70, G44/Q68/H70, H44/R68/D70,
 H44/S68/S70, K44/A68/G70, K44/E68/S70, K44/H68/E70, K44/K68/H70, K44/N68/N70, K44/Q68/E70,
 K44/R68/H70, K44/S68/S70, K44/T68/N70,A44/G68/H70, A44/H68/K70, A44/K68/H70, A44/N68/E70,
 A44/N68/S70, A44/Q68/S70, A44/S68/Q70, A44/T68/N70, D44/R68/N70, E44/R68/H70, G44/R68/Q70,
 10 H44/R68/E70, H44/S68/T70, K44/A68/H70, K44/G68/A70, K44/H68/G70, K44/K68/T70, K44/N68/Q70,
 K44/Q68/S70, K44/R68/N70, K44/S68/T70, K44/T68/Q70,A44/G68/K70, A44/H68/N70, A44/K68/N70,
 A44/N68/G70, A44/N68/T70, A44/R68/E70, A44/S68/R70, A44/T68/Q70, D44/R68/Q70, E44/R68/N70,
 G44/T68/D70, H44/R68/G70, H44/T68/S70, K44/A68/N70, K44/G68/G70, K44/H68/N70, K44/N68/A70,
 K44/N68/S70, K44/Q68/T70, K44/R68/S70, K44/T68/A70, K44/T68/S70,K44/T68/T70, N44/A68/H70,
 15 N44/K68/S70, N44/P68/D70, N44/S68/G70, N44/S68/H70, N44/T68/S70, P44/N68/D70, Q44/N68/H70,
 Q44/N68/S70, Q44/R68/Q70, Q44/S68/S70, R44/G68/T70, R44/H68/D70, R44/R68/G70, R44/R68/Q70,
 R44/S68/T70, S44/D68/K70, T44/R68/E70, T44/R68/Q70, y T44/S68/K70.

18. La variante de megaucleasa I-*CreI* de acuerdo con la reivindicación 17, que tiene una
 20 alanina (A) o una asparagina (N) en la posición 44 y escinde una diana que comprende un nucleótido a en
 la posición -4 y/o t en la posición +4.

19. La variante de megaucleasa I-*CreI* de acuerdo con la reivindicación 17, que tiene una
 lisina (K) en la posición 44 y escinde una diana que comprende un nucleótido c en la posición -4 y /o g en
 25 la posición +4.

20. La variante de megaucleasa I-*CreI* de acuerdo con la reivindicación 17, que tiene una
 glutamina (Q) en la posición 44 y escinde una diana que comprende un nucleótido t en la posición -4 y/o a
 en la posición +4.

30

21. La variante de meganucleasa I-*CreI* de acuerdo con una cualquiera de las reivindicaciones
 17 a 20, que es un homodímero.

22. La variante de meganucleasa I-*CreI* de acuerdo con una cualquiera de las reivindicaciones
 35 17 a 20, que es un heterodímero que consiste en dos monómeros, cada uno de una variante diferente como
 se define en una cualquiera de las reivindicaciones 17 a 20.

23. Un polinucleótido, **caracterizado por que** codifica una variante de meganucleasa I-CreI de acuerdo con una cualquiera de las reivindicaciones 17 a 22.
24. Un casete de expresión, que comprende un polinucleótido de acuerdo con la reivindicación 5 23 y las secuencias de regulación.
25. Un vector de expresión, **caracterizado por que** comprende un casete de expresión de acuerdo con la reivindicación 24.
- 10 26. Un vector de expresión, de acuerdo con la reivindicación 25, **caracterizado por que** comprende adicionalmente una construcción de ADN diana.
27. Un vector de expresión, de acuerdo con la reivindicación 26, **caracterizado por que** dicha construcción de ADN diana comprende una secuencia que comparte homologías con la región que rodea el 15 sitio de escisión de la variante de meganucleasa I-CreI de acuerdo con una cualquiera de las reivindicaciones 17 a 22.
28. El vector de expresión, de acuerdo con la reivindicación 27 **caracterizado por que** dicha construcción de ADN diana comprende:
- 20 a) secuencias que comparten homologías con la región que rodea el sitio de escisión de la variante de acuerdo con las reivindicaciones 17 a 22, y
b) secuencias a introducir flanqueadas por una secuencia como en a).
29. Una célula, **caracterizada por que** está modificada por un polinucleótido de acuerdo con 25 la reivindicación 23 o por un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28.
30. Una planta transgénica, **caracterizada por que** comprende un polinucleótido de acuerdo con la reivindicación 23 o un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28.
- 30 31. Un mamífero transgénico no humano, **caracterizado por que** comprende un polinucleótido de acuerdo con la reivindicación 23 o un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28.
- 35 32. Uso de una variante de meganucleasa I-CreI de acuerdo con una cualquiera de las reivindicaciones 17 a 22, un polinucleótido de acuerdo con la reivindicación 23, un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28, una célula de acuerdo con la reivindicación 29, una planta transgénica de acuerdo con la reivindicación 30, un mamífero transgénico no humano de acuerdo con la

reivindicación 31, para biología molecular, para ingeniería genética *in vivo* o *in vitro* y para ingeniería genómica *in vivo* o *in vitro*.

5 **33.** El uso de acuerdo con la reivindicación 32, para inducir una ruptura de ácido nucleico de doble cadena en un sitio de interés que comprende una secuencia diana de ADN, induciendo de este modo un suceso de recombinación de ADN, una pérdida de ADN o muerte celular.

10 **34.** El uso, de acuerdo con la reivindicación 32 o la reivindicación 33, **caracterizado por que** dicha rotura de ácido nucleico de doble cadena es para: reparar una secuencia específica, modificar una secuencia específica, restaurar un gen funcional en lugar de un gen mutado, atenuar o activar un gen endógeno de interés, introducir una mutación en un sitio de interés, introducir un gen exógeno o una parte del mismo, inactivar o suprimir un gen endógeno o una parte del mismo, traslocar un brazo cromosómico o dejar el ADN sin reparación y degradado.

15 **35.** El uso de acuerdo con una cualquiera de las reivindicaciones 32 a 34, **caracterizado por que** dicha variante de meganucleasa I-CreI, polinucleótido, vector, célula, planta transgénica o mamífero transgénico no humano se asocia con una construcción de ADN diana como se ha definido en la reivindicación 26.

20 **36.** El método de ingeniería genética, **caracterizado por que** comprende un paso de rotura de ácido nucleico de doble cadena en un sitio de interés localizado en un vector, que comprende una diana de ADN de una variante de meganucleasa de I-CreI de acuerdo con las reivindicaciones 17 a 22 mediante el contacto de dicho vector con una variante de meganucleasa I-CreI como se ha definido anteriormente, induciendo de este modo una recombinación homóloga con otro vector que presenta homología con la
25 secuencia que rodea el sitio de escisión de dicha variante de meganucleasa I-CreI.

37. Un método de ingeniería genómica, **caracterizado por que** comprende los siguientes pasos: 1) rotura de doble cadena de un locus genómico que comprende al menos un sitio de reconocimiento y escisión de una variante de meganucleasa I-CreI de acuerdo con las reivindicaciones 17 a 22, mediante el
30 contacto de dicho sitio de escisión con dicha variante de meganucleasa I-CreI; 2) mantener dicho locus genómico roto en condiciones apropiadas para recombinación homóloga con una construcción de ADN diana que comprende la secuencia a introducir en dicho locus, flanqueada por secuencias que comparten homologías con el locus diana.

35 **38.** Un método de ingeniería genómica, **caracterizado por que** comprende los siguientes pasos: 1) rotura de doble cadena de un locus genómico que comprende al menos un sitio de reconocimiento y escisión de una variante de meganucleasa I-CreI de acuerdo con las reivindicaciones 17 a 22, mediante el contacto de dicho sitio de escisión con dicha variante de meganucleasa I-CreI; 2) mantener dicho locus

genómico roto en condiciones apropiadas para recombinación homóloga con ADN cromosómico que comparte homologías con regiones que rodean el sitio de escisión.

- 5 **39.** Una composición **caracterizada por que** comprende al menos una variante de meganucleasa I-*CreI* de acuerdo con las reivindicaciones 17 a 22, un polinucleótido de acuerdo con la reivindicación 23 o un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28.
- 10 **40.** La composición de acuerdo con la reivindicación 39, **caracterizada por que** dicha composición comprende adicionalmente una construcción de ADN diana que comprende la secuencia que repara el sitio de interés flanqueado por secuencias que comparten homologías con el locus diana.
- 15 **41.** Uso de al menos una variante de meganucleasa I-*CreI* de acuerdo con las reivindicaciones 17 a 22, un polinucleótido de acuerdo con la reivindicación 23 o un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28 para la preparación de un medicamento para la prevención, mejora o cura de una enfermedad genética en un individuo que lo necesite, administrándose dicho medicamento por cualquier medio a dicho individuo.
- 20 **42.** Uso de al menos una variante de meganucleasa I-*CreI* de acuerdo con las reivindicaciones 17 a 22, un polinucleótido de acuerdo con la reivindicación 23 o un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28, para la preparación de un medicamento para prevenir, mejorar o curar una enfermedad causada por un agente infeccioso que presenta un intermedio de ADN en un individuo que lo necesite, administrándose dicho medicamento mediante cualquier medio a dicho individuo.
- 25 **43.** Uso de al menos una variante de meganucleasa I-*CreI* de acuerdo con las reivindicaciones 17 a 22, un polinucleótido de acuerdo con la reivindicación 23 o un vector de acuerdo con una cualquiera de las reivindicaciones 25 a 28, *in vitro*, para inhibir la propagación, inactivar o suprimir un agente infeccioso que presenta un intermedio de ADN, en productos derivados biológicos o productos dirigidos a usos biológicos o para desinfectar un objeto.
- 30 **44.** El uso de acuerdo con una cualquiera de las reivindicaciones 41 a 43, **caracterizado por que** dicho agente infeccioso es un virus.

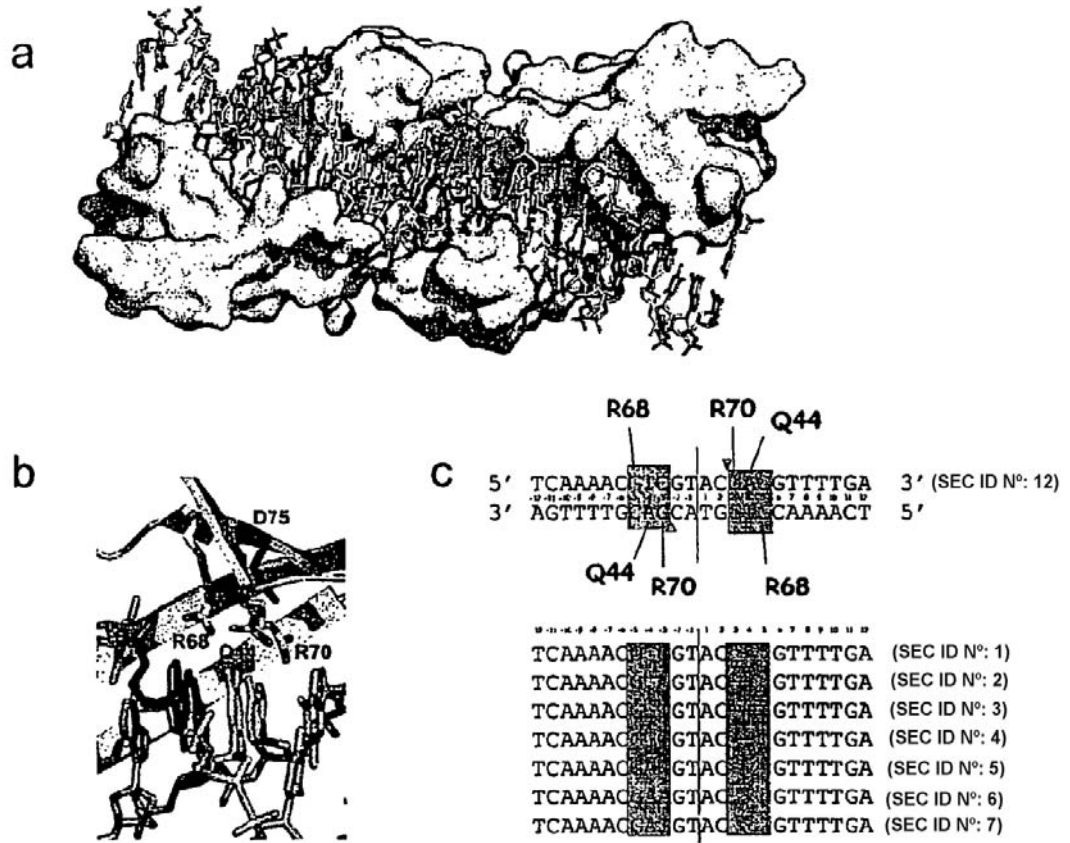


Figura 1

5

	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1		1	2	3	4	5	6	7	8	9	10	11	12
C1234 (SEC ID N°: 65)	T	C	A	A	A	A	C	G	T	C	G	T		G	A	G	A	C	A	G	T	T	T	G	G
C4334 (SEC ID N°: 66)	C	C	A	A	A	C	T	G	T	C	T	C		G	A	G	A	C	A	G	T	T	T	G	G
C1221 (SEC ID N°: 12)	T	C	A	A	A	A	C	G	T	C	G	T		A	C	G	A	C	G	T	T	T	T	G	A

Figura 2A

Figura 2B

	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12
T	C	A	A	A	A	C	G	G	G	G	T	A	C	C	C	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	G	A	G	T	A	C	T	C	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	G	T	G	T	A	C	A	C	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	G	C	G	T	A	C	G	C	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	A	G	G	T	A	C	C	T	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	A	A	G	T	A	C	T	T	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	A	T	G	T	A	A	A	T	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	A	C	G	T	A	C	G	T	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	T	G	G	T	A	C	C	A	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	T	A	G	T	A	C	T	A	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	T	G	G	T	A	C	A	A	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	T	C	G	T	A	C	G	A	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	C	G	G	T	A	C	C	G	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	C	A	G	T	A	C	T	G	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	C	T	G	T	A	C	A	G	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	G	C	C	G	T	A	C	G	G	C	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	G	G	G	T	A	C	C	C	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	G	A	G	T	A	C	T	C	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	G	T	G	T	A	C	G	T	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	A	G	G	T	A	C	C	T	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	A	A	G	T	A	C	T	T	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	A	T	G	T	A	C	A	T	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	A	C	G	T	A	C	G	T	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	T	G	G	T	A	C	C	A	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	T	A	G	T	A	C	T	A	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	T	T	G	T	A	C	A	A	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	T	G	G	T	A	C	C	A	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	C	G	G	T	A	C	C	G	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	C	A	G	T	A	C	T	G	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	C	T	G	T	A	C	A	G	T	G	T	T	T	T	G	A	
T	C	A	A	A	A	C	A	C	G	G	T	A	C	C	G	T	G	T	T	T	T	G	A	

Figura 2B (continuación)

-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12
T	C	A	A	A	A	C	T	G	G	G	T	A	C	C	C	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	G	A	G	T	A	C	T	C	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	G	T	G	T	A	C	A	C	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	G	C	G	T	A	C	G	C	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	A	G	G	T	A	C	C	T	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	A	A	G	T	A	C	T	T	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	A	C	G	T	A	C	G	T	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	T	G	G	T	A	C	C	A	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	T	A	G	T	A	C	T	A	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	T	C	G	T	A	C	G	A	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	C	G	G	T	A	C	C	G	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	C	A	G	T	A	C	T	G	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	T	C	T	G	T	A	C	A	G	A	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	G	G	G	T	A	C	C	C	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	G	A	G	T	A	C	T	C	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	G	T	G	T	A	C	G	T	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	G	C	G	T	A	C	A	T	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	A	G	G	T	A	C	C	T	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	A	A	G	T	A	C	T	T	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	A	T	G	T	A	C	A	T	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	A	C	G	T	A	C	G	T	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	T	A	G	T	A	C	T	A	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	T	T	G	T	A	C	A	A	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	G	T	G	G	T	A	C	C	A	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	C	G	G	T	A	C	C	G	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	C	A	G	T	A	C	T	G	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	C	T	G	T	A	C	A	G	G	G	T	T	T	T	G	A
T	C	A	A	A	A	C	C	C	G	G	T	A	C	C	G	G	G	T	T	T	T	G	A

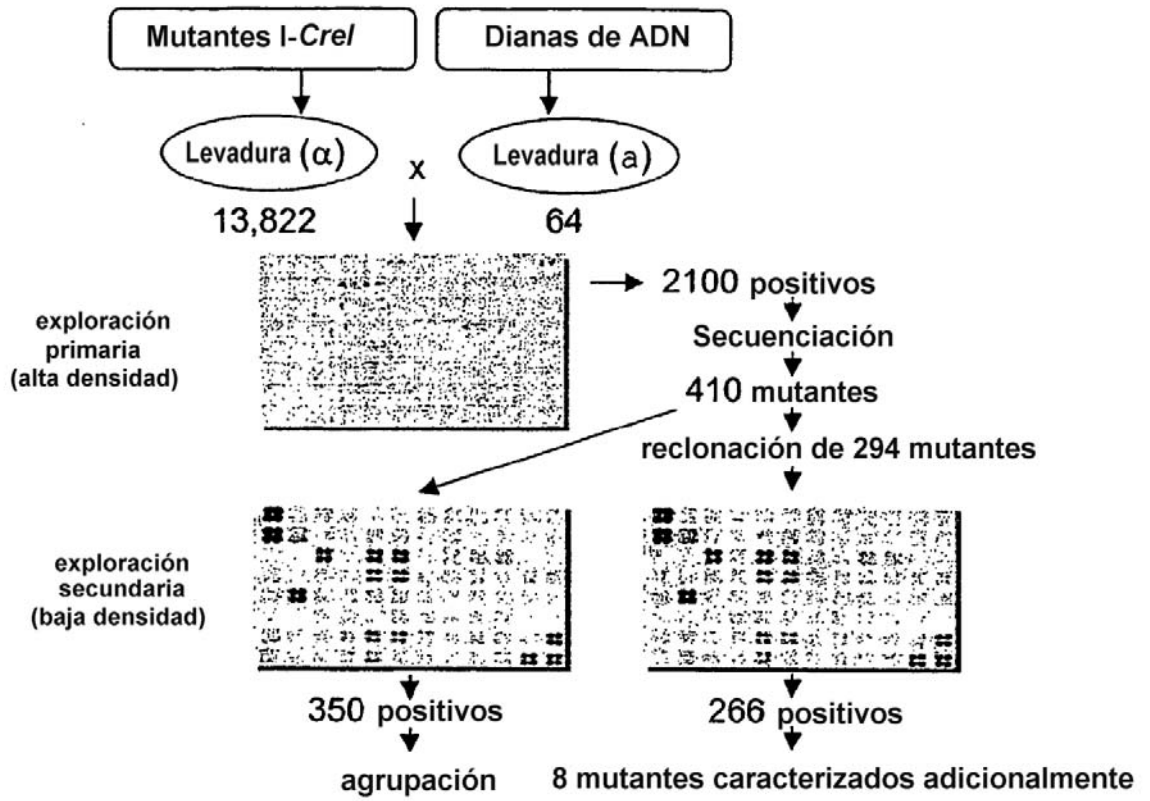


Figura 3b

SEC ID Nº: 69

```
1   ATGGCCAATA CCAAATATAA CAAAGAGTTC CTGCTGTACC TGGCCGGCTT
51  TGTGGACGGT GACGGTAGCA TCATCGCTCA GATTAAACCA AACCAGTCTT
101 ATAAGTTTAA ACATCAGCTA AGCTTGACCT TTCAGGTGAC TCAAAAGACC
151 CAGCGCCGTT GGTTTCTGGA CAAACTAGTG GATGAAATTG GCGTTGGTTA
201 CGTACGTGAT CGCGGATCCG TTTCCAATA CATCTTAAGC GAAATCAAGC
251 CGCTGCACAA CTCCTGACT CAACTGCAGC CGTTTCTGAA ACTGAAACAG
301 AAACAGGCAA ACCTGGTTCT GAAAATTATC GAACAGCTGC CGTCTGCAAA
351 AGAATCCCCG GACAAATTCC TGGAAGTTTG TACCTGGGTG GATCAGATTG
401 CAGCTCTGAA CGATTCTAAG ACGCGTAAAA CCACTTCTGA AACCGTTCGT
451 GCTGTGCTGG ACAGCCTGAG CGAGAAGAAG AAATCCTCCC CGGCGGCCGA
501 C
```

Figura 4A

Ulibl1dir (SEC ID N°: 67)

GTTTAAACAT CAGCTAAGCT TGACCTTTVV KGTGACTCAA AAGACCCAG

Ulibl1rev (SEC ID N°: 68)

GATGTAGTTG GAAACGGATC CMBBATICMBB TACGTAACCA ACGCC

V = A o G o C

M = A o C

B = G o C o T

K = G o T

Figura 4B

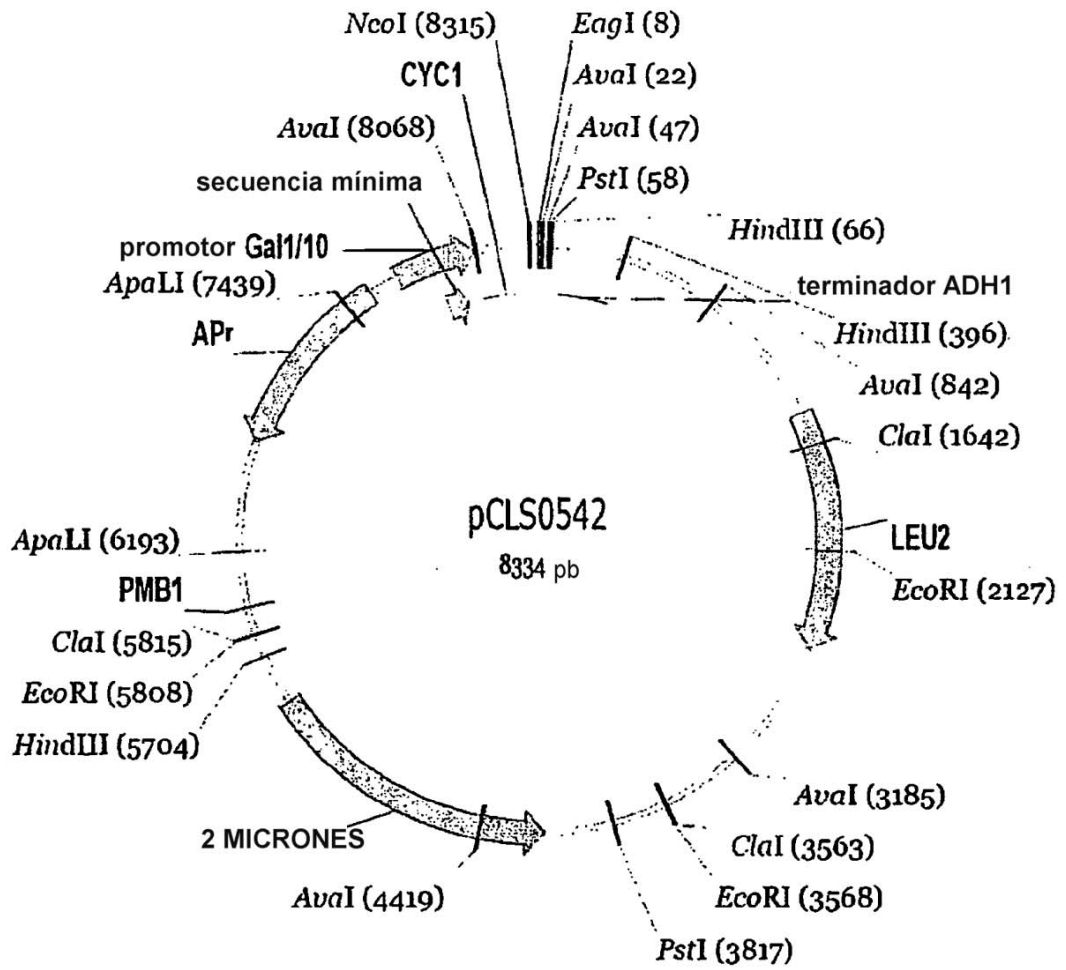


Figura 5

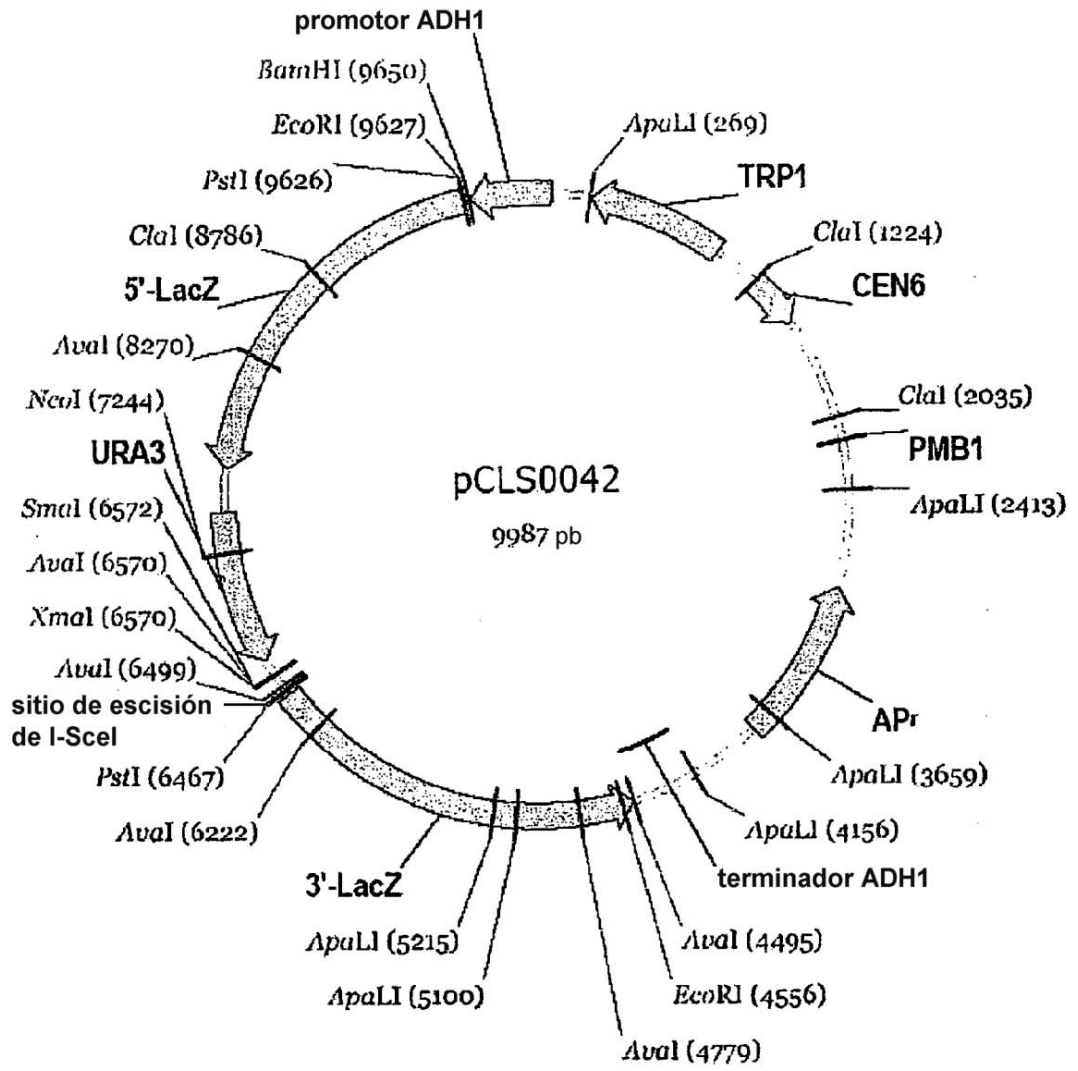


Figura 6

GGG	GGG	GCT	GCC	AGG	AGA	AAT	AAC
GAG	GAA	GAT	GAC	AAG	AAA	AAT	AAC
GUG	GUA	GTT	GTC	ATG	ATA	ATT	ATC
GCG	GCA	GCT	GCC	ACG	ACA	ACT	ACC
TGG	TGA	TGT	TGC	CGG	CGA	CGT	CGC
TAG	TAA	TAT	TAC	CAG	CAA	CAT	CAC
TUG	TUA	TTT	TTC	CTG	CTA	CTT	CTC
TGG	TGA	TGT	TGC	CCG	CCA	CCT	CCC

QRR



GTG
GTT
GTC
GCT
GCC

QAT



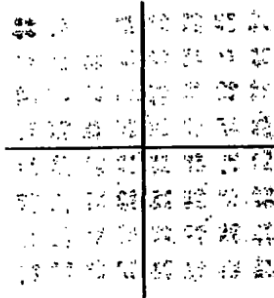
GTT

QAN



GTT

DRK



GGG

RAT



CCT

THR



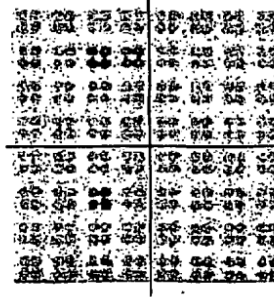
GAC

NAR



GAC
TAC

NRS



GAG
GAT
GAC
TAT

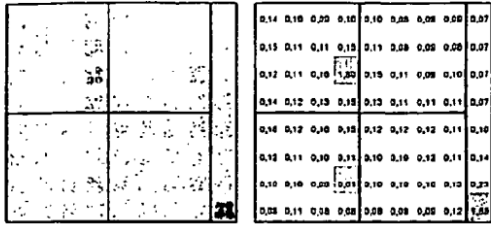
KSS



NCT
NCC
TTT
TTC
CTT
CTC

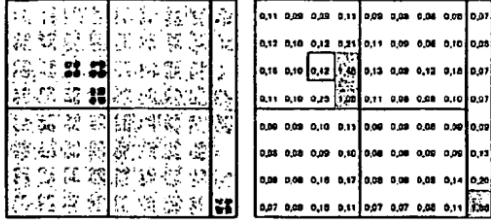
Figura 8

I-Crel

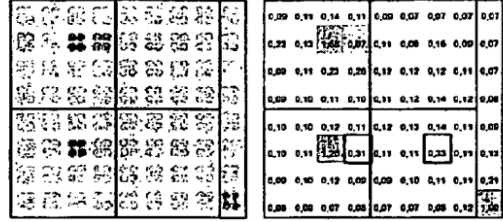


GCG	GGA	GCT	GCC	ACC	ACA	AGT	AGC	U1
GAG	GAA	GAT	GAC	AAG	AAA	AAT	AAC	U2
GUG	GUA	GUT	GUC	ATG	ATA	ATT	ATC	O
GCG	GGA	GCT	GCC	ACC	ACA	AGT	AGC	U1
TGG	TGA	TGT	TGC	CGG	CGA	CGT	CGC	U3
TAG	TAA	TAT	TAC	CAG	CAA	CAI	CAC	O
TTG	TTA	TTT	TTC	CTG	CTA	CTI	CTC	U1
TGG	TGA	TGT	TGC	CGG	CGA	CGT	CGC	U3

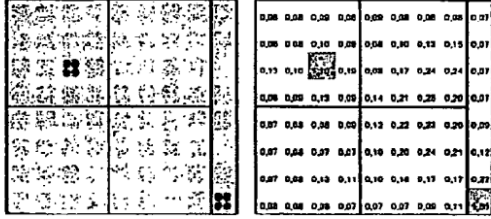
QRR



ARL



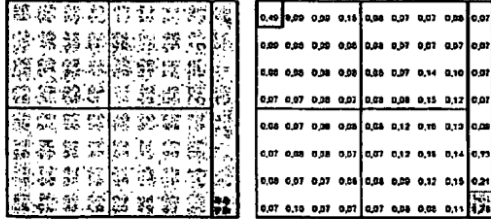
QAN



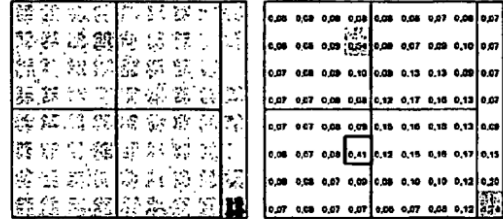
ADK



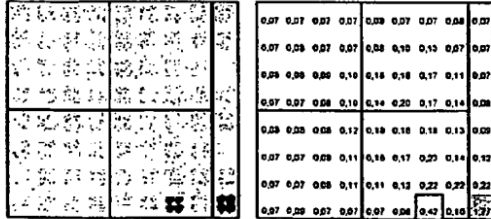
DRK



AGR



RAT



KTG

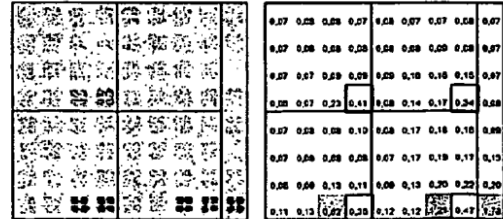


Figura 9a

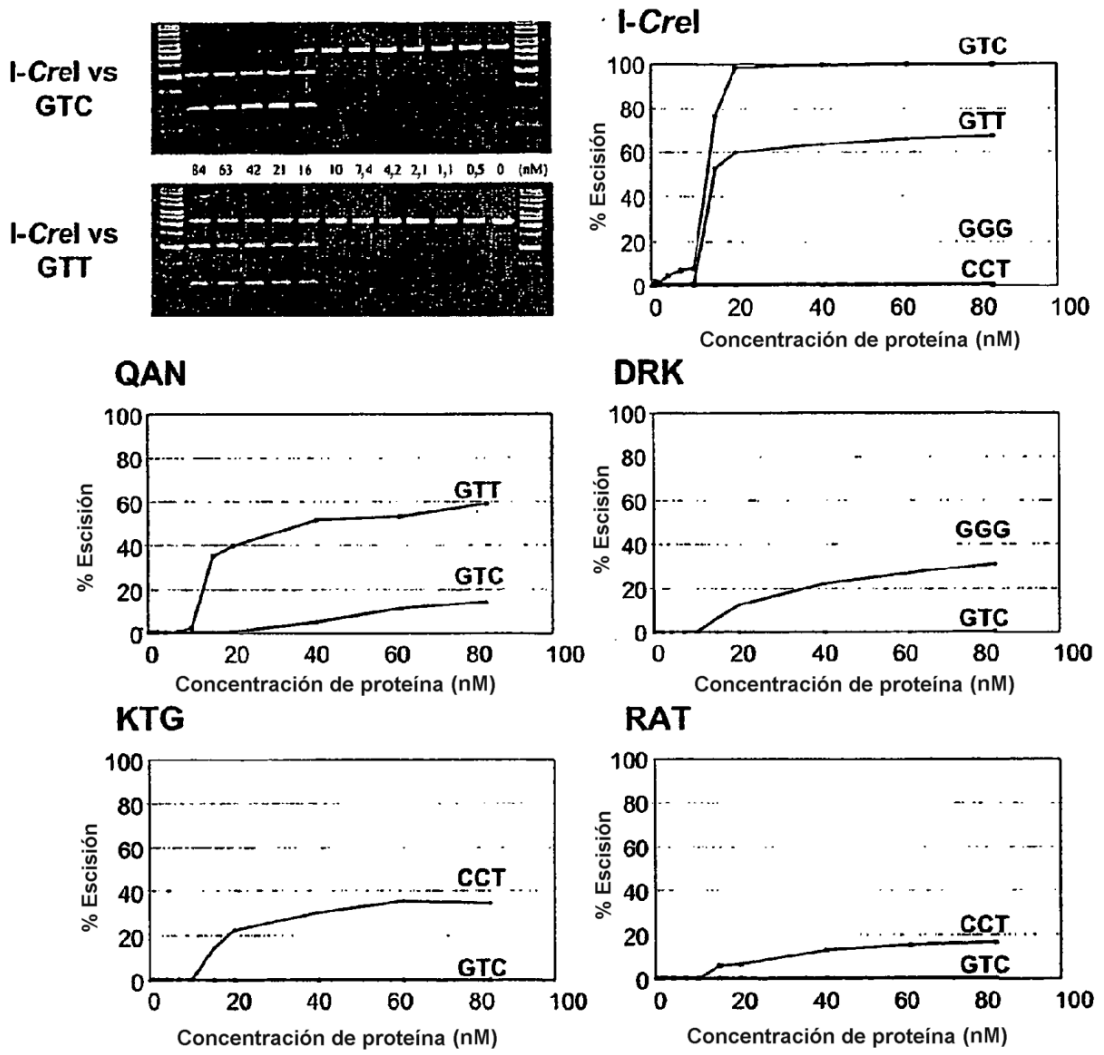


Figura 9b

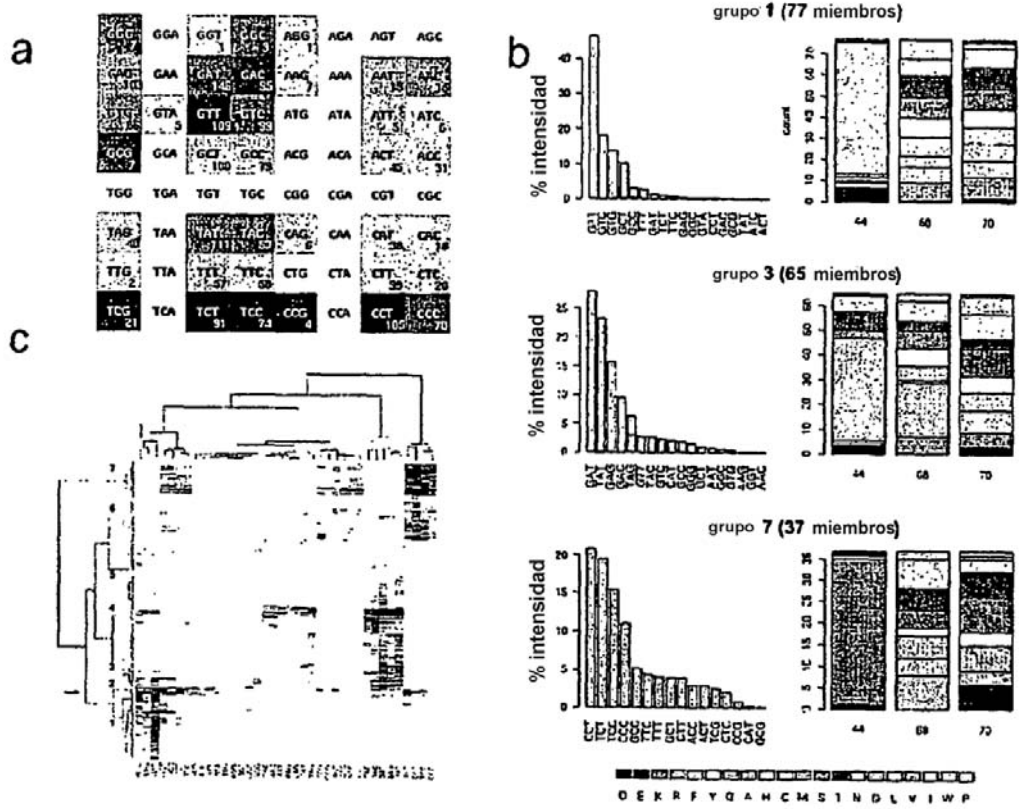


Figura 10



Figura 11

AAC/ACT	AAC/ACC	AAC/CCT	AAC/CCC	ACT/CAC	ACC/CAC	CAT/CCT	CAC/CCT	CAC/CCC	GGG/AAC	GCC
GGG/ACT	GGG/ACC	GGG/CAT	GGG/CAC	GGG/CCT	GGG/CCC	GGG/GAT	GGG/GAC	GGG/GTT	GGG/GTC	TAT
GGG/GCT	GGG/GCC	GGG/TAT	GGG/TAC	GGG/TCT	GGG/TCC	GAT/CAT	GAT/CCT	GAT/GAC	GAT/GTT	CAT
GAT/GTC	GAT/GCC	GAT/TAT	GAC/ACT	GAC/ACC	GAC/CAT	GAC/CCT	GAC/CCC	GAC/GTT	GAC/GTC	CCT
GAC/GCT	GAC/GCC	GAC/TAT	GAC/TCT	GAC/TCC	GTT/CAT	GTT/CCT	GTT/GTC	GTT/GCC	GTT/TAT	0
GTC/CAT	GTC/CCT	GTC/GCC	GTC/TAT	GCT/AAC	GCT/CAC	GCT/TAC	GCC/AAC	GCC/CAT	GCC/GAC	0
GCC/CCT	GCC/TAT	GCC/TAC	TAT/CAT	TAT/CCT	TAC/ACT	TAC/ACC	TAC/CCT	TAC/CCC	TAC/TCT	0
TAC/TCC	TCT/AAC	TCT/CAC	TCC/AAC	TCC/CAC	GGG	GAT	GAC	GTT	GTC	0

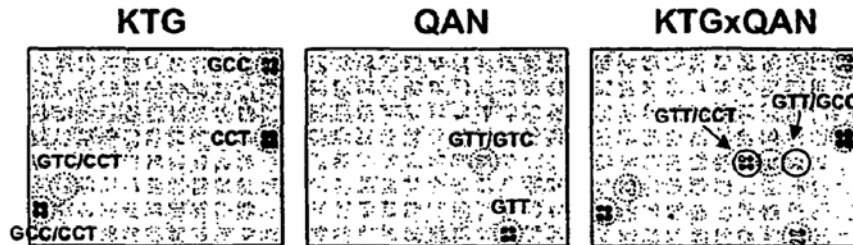


Figura 12

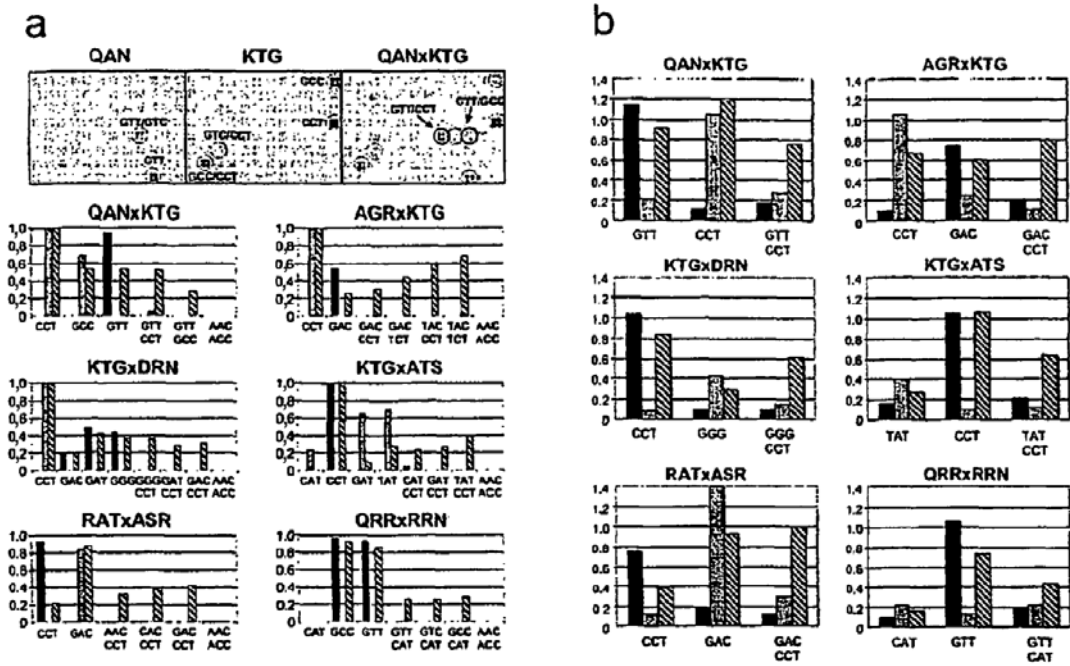
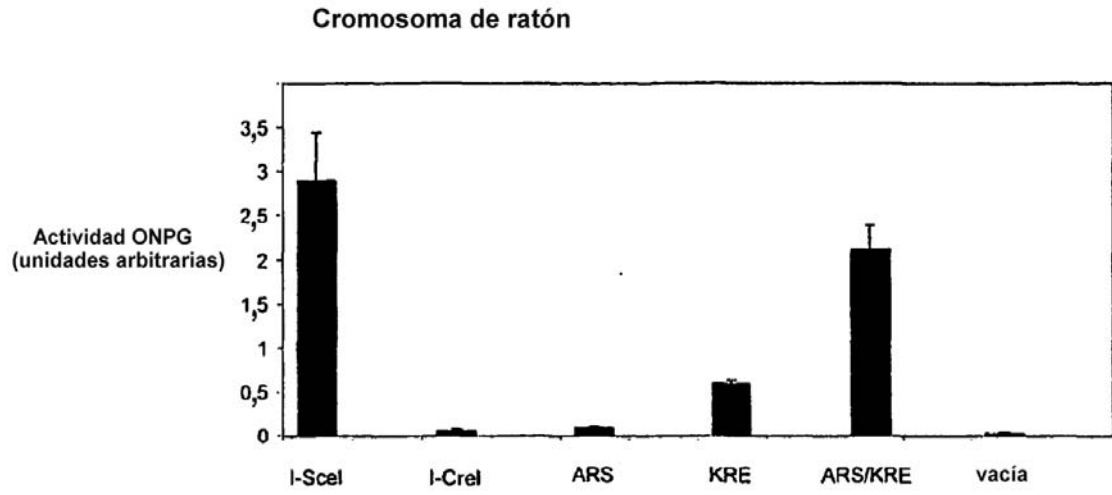


Figura 13



Meganucleasa

Figura 14