

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 627 552**

51 Int. Cl.:

A01K 67/027 (2006.01)

C12N 15/11 (2006.01)

C12N 9/22 (2006.01)

C12N 15/85 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.12.2009 PCT/US2009/006365**

87 Fecha y número de publicación internacional: **10.06.2010 WO10065123**

96 Fecha de presentación y número de la solicitud europea: **03.12.2009 E 09830727 (5)**

97 Fecha y número de publicación de la concesión europea: **26.04.2017 EP 2352369**

54 Título: **Edición de genoma en ratas usando nucleasas con dedos de cinc**

30 Prioridad:

04.12.2008 US 200985 P

26.01.2009 US 205970 P

24.11.2009 US 263904 P

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

28.07.2017

73 Titular/es:

SIGMA ALDRICH COMPANY (50.0%)

3050 Spruce Street

St. Louis, MO 63103, US y

SANGAMO BIOSCIENCES, INC. (50.0%)

72 Inventor/es:

CUI, XIAOXIA;

GEURTS, ARON M. y

URNOV, FYODOR

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 627 552 T3

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

DESCRIPCIÓN

Edición de genoma en ratas usando nucleasas con dedos de cinc

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

5 La presente solicitud reivindica el beneficio de las solicitudes provisionales de EE.UU. N.º 61/200.985, presentada el 4 de diciembre de 2008; 61/205.970, presentada el 26 de enero de 2009, y 61/263.904, presentada el 24 de noviembre de 2009, cuyas divulgaciones se incorporan por este documento por referencia en su totalidad.

ESTABLECIMIENTO DE DERECHOS A INVENCIONES HECHAS BAJO INVESTIGACIÓN FEDERALMENTE PATROCINADA

No aplicable.

10 CAMPO TÉCNICO

La presente divulgación está en los campos de la ingeniería del genoma de ratas, que incluye rupturas génicas somáticas y heredables, alteraciones genómicas, generación de alelos que llevan mutaciones aleatorias en posiciones específicas de genes de rata e inducción de reparación dirigida por homología.

ANTECEDENTES

15 Las ratas (*Rattus norvegicus*) son un modelo animal ampliamente usado en los campos de la hipertensión, fisiología cardiovascular, diabetes, trastornos metabólicos, estudios conductuales y pruebas de toxicidad. Michalkiewicz et al. (2007) J. Amer. Phys. Society 293:H881-H894. La disponibilidad de estos sistemas modelo, avances en la genómica de la rata y la secuencia de los genomas de la rata, humano y de ratón han acelerado enormemente el uso de modelos de rata endogámicos para el descubrimiento de la base genética de enfermedades complejas y proporcionado modelos animales para el descubrimiento de fármacos terapéuticos.

20 Sin embargo, los avances en la información sobre el genoma de la rata no han sido acompañados por el progreso paralelo en la tecnología de modificación del genoma. A diferencia de los ratones, no se producen fácilmente clones de células madre embrionarias de rata para el direccionamiento génico. La inyección pronuclear también ha demostrado ser difícil y tiene poca tasa de éxito en la generación de ratas transgénicas. Michalkiewicz et al. (2007) J. Amer. Phys. Society 293:H881-H894 informan de la generación de ratas transgénicas usando una construcción lentiviral que expresa un gen indicador de proteína verde fluorescente potenciada (eGFP), donde se encontró que el transgén eGFP estaba presente en 1-4 copias integradas en sitios aleatorios dentro del genoma.

25 El documento WO 2003/080809 describe el uso de endonucleasas con dedos de cinc para potenciar la recombinación homóloga en una secuencia de ADN diana cromosómica endógena en una célula primaria.

30 Doyon et al. 2008, Nat. Biotechnol. 26(6): 702-708, describen el uso de nucleasas con dedos de cinc para rotura somática y de la línea germinal de genes en pez cebra. Xiangdong Meng et al. 2008, Nat. Biotechnol. 26(6): 695-701 enseñan la inactivación de genes dirigida en pez cebra usando nucleasas con dedos de cinc manipuladas.

35 El documento US 2009/0205083 se refiere a proteínas de dedos de cinc manipuladas que dirigen genes EPSPS en plantas y métodos de uso de tales proteínas de dedos de cinc en la modulación de la expresión génica, inactivación génica y modificación de gen dirigida.

El documento EP 1 348 335 describe un método de producción de un mamífero no humano que lleva mutaciones de la línea germinal en tanto los genes *p53* como *Fos*, en el que el método implica la ruptura de estos genes por recombinación homóloga.

40 Haas et al. 1998, Molecular Carcinogenesis, 21(2):128-134, caracteriza mutantes *p53* de ratas transgénicas para el antígeno T del virus 40 simio por albúmina tratadas con carcinógeno.

El documento WO 2008/151081 desvela un método de producción de animales transgénicos que carecen de Ig endógena y capaces de producir anticuerpos transgénicos usando nucleasas con dedos de cinc manipuladas.

45 Sigue existiendo una necesidad de métodos de modificación de genomas de la rata de un modo dirigido. La escisión específica de sitio dirigida con precisión de loci genómicos ofrece un complemento eficiente y/o alternativa a la recombinación homóloga convencional. La creación de una rotura bicatenaria (DSB) aumenta la frecuencia de recombinación homóloga en el locus dirigido superior a 1000 veces. Más simplemente, la reparación imprecisa de una DSB específica de sitio por unión de extremos no homólogos (NHEJ) también puede producir ruptura génica. La creación de dos DSB tales produce la delección de regiones arbitrariamente grandes. Las preferencias de reconocimiento de ADN modular de proteína de dedos de cinc permiten el diseño racional de proteínas de unión a ADN multi-dedos específicas de sitio. La fusión del dominio de nucleasa de la enzima de restricción de tipo II *Fok I* con proteínas de dedos de cinc específicas de sitio permite la creación de nucleasas específicas de sitio. Véanse, por ejemplo, las publicaciones de patente de Estados Unidos 20030232410; 20050208489; 20050026157;

20050064474; 20060188987; 20060063231; 20070134796; 2008015164; 20080131962; 2008015996 y las publicaciones internacionales WO 07/014275 y 2008/133938, que todas describen el uso de nucleasas con dedos de cinc.

SUMARIO

5 La presente invención se define por las reivindicaciones.

La presente invención se refiere a un método de modificación de un gen IgM endógeno en una célula de rata, comprendiendo el método: introducir, en la célula de rata, uno o más ARNm que codifican primera y segunda nucleasas con dedos de cinc (ZFN) que se unen a sitios diana en el gen IgM endógeno en condiciones tales que las ZFN escindan el uno o más genes celulares endógenos tal que el gen IgM endógeno se modifique, en el que la primera ZFN comprende una ZFP que tiene las regiones de hélice de reconocimiento que se muestran en las ZFP designadas 17747, 17759, 17756, 17767 y 17764 como se expone en una única fila de la Tabla 6.

En el presente documento se describen composiciones para la edición de genoma en rata, que incluyen, pero no se limitan a: escisión de uno o más genes en rata produciendo alteración dirigida (mutaciones de inserción, delección y/o sustitución) en uno o más genes de rata, que incluye la incorporación de estas alteraciones dirigidas en la línea germinal; introducción dirigida de secuencias de ácidos nucleicos no endógenas, la inactivación parcial o completa de uno o más genes en rata; métodos de inducción de reparación dirigida por homología y/o generación de mutaciones aleatorias que codifican novedosas formas alélicas de genes de rata.

En el presente documento se describe una proteína de dedos de cinc (ZFP) que se une a sitio diana en una región de interés en un genoma de rata, en la que la ZFP comprende uno o más dominios de unión de dedos de cinc manipulados. La ZFP puede ser una nucleasa con dedos de cinc (ZFN) que escinde una región genómica diana de interés en rata, en la que la ZFN comprende uno o más dominios de unión de dedos de cinc manipulados y un dominio de escisión de nucleasa o medio-dominio de escisión. Los dominios de escisión y medio-dominios de escisión pueden obtenerse, por ejemplo, de diversas endonucleasas de restricción y/o endonucleasas de migración. En una realización, los medio-dominios de escisión se derivan de una endonucleasa de restricción de tipo IIS (por ejemplo, *Fok I*). La ZFN puede escindir específicamente una secuencia de genes de rata particular. Alternativamente, la ZFN puede escindir dos o más secuencias de genes de rata homólogos.

La ZFN puede unirse a y/o escindir un gen de rata dentro de la región codificante del gen o en una secuencia no codificante dentro de o adyacente al gen, tal como, por ejemplo, una secuencia conductora, secuencia remolque o intrón, o dentro de una región no transcrita, tanto en la dirección 5' como en la dirección 3' de la región codificante. La ZFN puede unirse a y/o escindir una secuencia codificante o una secuencia reguladora del gen de rata diana.

En el presente documento se describen composiciones que comprenden una o más de las nucleasas con dedos de cinc descritas en el presente documento. En ciertas realizaciones, la composición puede comprender una o más nucleasas con dedos de cinc en combinación con un excipiente farmacéuticamente aceptable.

En el presente documento se describe un polinucleótido que codifica una o más ZFN descritas en el presente documento. Según la presente invención, el polinucleótido es ARNm.

En el presente documento se describe un vector de expresión de ZFN que comprende un polinucleótido, que codifica una o más ZFN descritas en el presente documento, operativamente unido a un promotor.

En el presente documento se describe una célula huésped de rata que comprende uno o más vectores de expresión de ZFN. La célula huésped de rata puede ser establemente transformada o transitoriamente transfectada o una combinación de los mismos con uno o más vectores de expresión de ZFP. La célula huésped de rata puede ser una célula madre embrionaria. El uno o más vectores de expresión de ZFP pueden expresar una o más ZFN en la célula huésped de rata. La célula huésped de rata puede comprender además una secuencia donante de polinucleótido exógeno. La célula huésped de rata puede comprender una célula de embrión, por ejemplo un embrión uni- o pluricelular.

En el presente documento se describe un método de escisión de uno o más genes en una célula de rata, comprendiendo el método: (a) introducir, en la célula de rata, uno o más polinucleótidos que codifican una o más ZFN que se unen a un sitio diana en el uno o más genes en condiciones tales que la(s) ZFN se exprese(n) y escindan el uno o más genes.

En el presente documento se describe un método de introducción de una secuencia exógena en el genoma de una célula de rata, comprendiendo el método las etapas de: (a) introducir, en la célula de rata, uno o más polinucleótidos que codifican una o más ZFN que se unen a un sitio diana en el uno o más genes en condiciones tales que la(s) ZFN se exprese(n) y el uno o más genes se escindan; y (b) poner en contacto la célula con un polinucleótido exógeno; de forma que la escisión del (de los) gen(es) estimule la integración del polinucleótido exógeno en el genoma por recombinación homóloga. En ciertas realizaciones, el polinucleótido exógeno se integra físicamente en el genoma.

En otras realizaciones, el polinucleótido exógeno se integra en el genoma copiando la secuencia exógena en el genoma de la célula huésped mediante procesos de replicación de ácido nucleico (por ejemplo, reparación dirigida

por homología de la rotura bicatenaria). En todavía otras realizaciones, la integración en el genoma se produce mediante integración dirigida no dependiente de homología (por ejemplo, "captura de extremos"). La una o más nucleasas pueden ser fusiones entre el dominio de escisión de una endonucleasa de restricción de tipo IIS y un dominio de unión de dedo de cinc manipulado.

5 En el presente documento se describe además un método de modificación de una o más secuencias de genes en el genoma de una célula de rata, comprendiendo el método (a) proporcionar una célula de rata que comprende una o más secuencias de genes diana; y (b) expresar la primera y segunda nucleasas con dedos de cinc (ZFN) en la célula, en el que la primera ZFN escinde en un primer sitio de escisión y la segunda ZFN escinde en un segundo sitio de escisión, en el que la secuencia de gen está localizada entre el primer sitio de escisión y el segundo sitio de escisión, en el que la escisión del primer y segundo sitios de escisión produce la modificación de la secuencia de gen por unión de extremos no homólogos. En ciertas realizaciones, la unión de extremos no homólogos produce una delección entre el primer y segundo sitios de escisión. El tamaño de la delección en la secuencia de gen se determina por la distancia entre el primer y segundo sitios de escisión. Por consiguiente, pueden obtenerse delecciones de cualquier tamaño, en cualquier región genómica de interés. Pueden obtenerse delecciones de 25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1.000 pares de nucleótidos, o cualquier valor entero de pares de nucleótidos dentro de este intervalo. Además, pueden obtenerse delecciones de una secuencia de cualquier valor entero de pares de nucleótidos superior a 1.000 pares de nucleótidos usando los métodos y composiciones descritos en el presente documento. La unión de extremos no homólogos puede también producir una inserción entre el primer y segundo sitios de escisión. Pueden usarse métodos de modificación del genoma de una rata como se describen en el presente documento para crear modelos de enfermedad animal (por ejemplo, humano), por ejemplo inactivando (parcialmente o completamente) un gen o creando mutaciones aleatorias en posiciones definidas de genes que permiten la identificación o selección de ratas transgénicas que llevan formas alélicas novedosas de aquellos genes, por inserción de genes de rata humanizados (para estudiar, a modo de un ejemplo no limitante, el metabolismo de fármacos) o por inserción de alelos mutantes de interés para examinar, por ejemplo, el efecto fenotípico de un alelo mutante tal.

En otro aspecto, en el presente documento se describe un método para la ruptura de la línea germinal de uno o más genes diana en rata, comprendiendo el método modificar una o más secuencias de genes en el genoma de una o más células de un embrión de rata por cualquiera de los métodos descritos en el presente documento y permitir que se desarrolle el embrión de rata, en el que las secuencias de genes modificadas están presentes en al menos una porción de gametos de la rata sexualmente madura.

En cualquiera de los métodos descritos en el presente documento, el polinucleótido que codifica la(s) nucleasa(s) con dedos de cinc puede comprender ADN, ARN o combinaciones de los mismos. El polinucleótido puede comprender un plásmido. Según la presente invención, el polinucleótido que codifica la nucleasa comprende ARNm.

Un aspecto adicional descrito en el presente documento es un método para la integración específica de sitio de una secuencia de ácidos nucleicos en un cromosoma. El método puede comprender: (a) inyectar un embrión con (i) al menos un vector de ADN, en el que el vector de ADN comprende una secuencia en la dirección 5' y una secuencia en la dirección 3' que flanquea la secuencia de ácidos nucleicos que va a integrarse, y (ii) al menos una molécula de ARN que codifica una nucleasa con dedos de cinc que reconoce el sitio de integración cromosómica, y (b) cultivar el embrión para permitir la expresión de la nucleasa con dedos de cinc, en el que se repara una rotura bicatenaria introducida en el sitio de integración por la nucleasa con dedos de cinc, mediante recombinación homóloga con el vector de ADN, de manera que se integre la secuencia de ácidos nucleicos en el cromosoma. Especies adecuadas de mamífero, ave, reptil, anfibio y peces. En términos generales, un embrión adecuado es un embrión que puede recogerse, inyectarse y cultivarse para permitir la expresión de una nucleasa con dedos de cinc. Embriones adecuados pueden incluir embriones de roedores, animales de compañía, ganado y primates. Ejemplos no limitantes de roedores pueden incluir ratones, ratas, hámsteres, jerbos y cobayas. Ejemplos no limitantes de animales de compañía pueden incluir gatos, perros, conejos, erizos y hurones. Ejemplos no limitantes de ganado pueden incluir caballos, cabras, ovejas, cerdos, llamas, alpacas y ganado vacuno. Ejemplos no limitantes de primates pueden incluir monos capuchinos, chimpancés, lémures, macacos, titis, tamarinos, monos araña, monos ardilla y monos verdes. Embriones adecuados pueden incluir embriones de peces, reptiles, anfibios o aves. Alternativamente, embriones adecuados pueden ser embriones de insecto, por ejemplo, un embrión de *Drosophila* o un embrión de mosquito.

También se describe un embrión que comprende al menos un vector de ADN, en el que el vector de ADN comprende una secuencia en la dirección 5' y una secuencia en la dirección 3' que flanquea la secuencia de ácidos nucleicos que va a integrarse, y al menos una molécula de ARN que codifica una nucleasa con dedos de cinc que reconoce el sitio de integración cromosómica. También se describen organismos derivados de cualquiera de los embriones que se describen en el presente documento.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **Figura 1** muestra resultados de ensayos de nucleasas ("CEL-I") Surveyor™ de pares de ZFN específicas de p53 en células C6 de rata. El par de ZFN usado en cada carril se muestra encima del carril y el porcentaje de actividad de NHEJ como se detecta por el ensayo de desapareamiento Surveyor se muestra debajo.

La **Figura 2** muestra resultados de ensayos de nucleasas ("CEL-I") Surveyor™ de pares de ZFN dirigidas a eGFP 16834/16833, 16856/16855 y 16859/16860 en células C6 de rata que llevan el gen eGFP. El par de ZFN usado se muestra encima de cada carril y el % de unión de extremos no homólogos (% de NHEJ) se indica debajo de cada carril, según convenga.

5 La **Figura 3** es un esquema que representa la modificación dirigida de un transgén de GFP usando ZFN en ratas GFP transgénicas.

10 La **Figura 4**, paneles A y B, muestra la rotura dirigida de GFP por ZFN en crías de rata nacidas de inyección pronuclear de ZFN dirigidas a GFP en embriones obtenidos de ratas GFP transgénicas. La Figura 4A muestra las 5 crías bajo luz ultravioleta, revelando 3 animales positivos para GFP y 2 animales que no expresan GFP (GFP-negativos). La Figura 4B muestra los resultados de análisis por PCR de biopsias de cola de crías GFP+ y GFP-.

15 La **Figura 5** representa la escisión mediada por ZFN en el exón 1 de *IgM* endógeno en células C6 usando pares de ZFN conducidos por el promotor del CMV (CMV) o el promotor de CAG. "Unido" se refiere a pares de ZFN en el mismo plásmido unidos por el péptido 2A, mientras que "no unido" se refiere a pares de ZFN no unidos por el péptido 2A.

La **Figura 6** representa análisis, por nucleasa Surveyor™, de ADN genómico preparado a partir de las colas de 43 animales resultantes de partos vivos de embriones unicelulares inyectados con ZFN de *IgM*. Como se indica, las ratas N.º 6, 7, 8, 19 y 46 dieron positivo para la modificación en el sitio *IgM*. Barras con números blancos indican crías nacidas de madres individuales (numeradas).

20 La **Figura 7** representa los resultados de análisis por PCR de ratas modificadas en *IgM* (N.º 6, 7, 8, 19 y 46 como se identifica en la Figura 6) para la inserción de los plásmidos de ZFN en el genoma.

La **Figura 8**, paneles A y B, muestra el análisis de CEL-I y de secuenciación de la rata N.º 19 modificada en *IgM*. El alineamiento de las secuencias de WT, alelo no mutante de rata 19 y alelo de delección de rata 19 (Panel 8B) demuestra las secuencias que han sido delecionadas del alelo de delección de rata 19.

25 La **Figura 9**, paneles A a C, representa el análisis de ratas modificadas en *IgM* (N.º 6, 7, 8, 19 y 46 como se identifica en la Figura 6) para la actividad en 8 sitios inespecíficos diferentes. Los sitios inespecíficos (Sitio 1, Sitio 2, etc.) son como se definen en la Tabla 9.

La **Figura 10** representa el análisis de secuencias de la modificación mediada por ZFN de *Rab38*. En esta figura se muestra un alineamiento del alelo no mutante con dos alelos de delección ($\Delta 6$ y $\Delta 42$).

30 La **Figura 11**, paneles A a C, representa el análisis de las crías obtenidas del cruce ZFN-ratas modificadas en *IgM* y una rata no mutante. La Figura 11A muestra análisis de PCR y CEL-I de las 5 crías (numeradas 224 a 228) de la rata de cruce N.º 19 (Ejemplo 3) con una rata no mutante. La Figura 11B muestra la confirmación del análisis de secuenciación de que 3 crías modificadas en *IgM* (N.º 225, 227 y 228 como se identifica en la figura) incluyen el mismo alelo de delección de 64 pares de bases en el locus *IgM* que la rata parental N.º 19. La Figura 11C muestra análisis de PCR y CEL-I de crías adicionales de rata N.º 19, además de crías de los cruces de ratas modificadas en *IgM* N.º 46 y N.º 8. La rata modificada en *IgM* parental se indica en la parte superior "F0" y los números de las crías se indican encima de cada carril.

35 La **Figura 12** es un esquema que representa los resultados de reparación después de una rotura bicatenaria inducida por ZFN dirigida. Las barras sombreadas representan el fragmento donante, mientras que las barras blancas representan el sitio diana para la rotura bicatenaria de ZFN.

La **Figura 13** es un esquema que representa la construcción de plásmidos donantes de RFLP. Se muestran el plásmido, y fragmentos amplificados por PCR izquierdo y derecho homólogos al sitio diana de integración. Se indican las enzimas de restricción usadas para la clonación. El fragmento izquierdo usó KpnI y NotI o PmeI. El fragmento derecho usó NotI o PmeI y SacII.

40 La **Figura 14** es un esquema que representa la construcción de plásmidos donantes que expresan GFP. El casete de GFP se amplificó por PCR a partir de un plásmido existente y se cerró en el donante de RFLP de NotI usando un sitio NotI.

45 La **Figura 15**, paneles A y B, representa métodos de detección de la integración de RFLP. La Figura 15A es un esquema que representa métodos de detección la integración de RFLP y digestión con enzimas de restricción. La Figura 15B es un esquema que representa la integración del casete de expresión de GFP usando amplificación por PCR.

50 La **Figura 16** es una imagen fotográfica de fragmentos de PCR fluorescentemente teñidos resueltos en un gel de agarosa. El carril más a la izquierda contiene una escalera de ADN. Los carriles 1 a 6 contienen fragmentos de PCR amplificados usando cebadores específicos de *Mdr1a* de ratón de una totalidad o una fracción de un

blastocisto de ratón. Los carriles 1 y 2 se amplificaron de 5/6 y 1/6 de un blastocisto, respectivamente. El carril 3 era de blastocisto completo. Los carriles 4 a 6 eran de 1/2, 1/3, y 1/6 del mismo blastocisto, respectivo. El carril 7 contiene un fragmento de PCR de control positivo amplificado usando los mismos cebadores de ADN de dedo de la pata de ratón extraído.

5 La **Figura 17**, paneles A y B, representa imágenes fotográficas de fragmentos de ADN fluorescentemente teñidos resueltos en un gel de agarosa. Los carriles más a la izquierda contienen una escalera de ADN. Los carriles 1 a 39 de la Figura 17A contienen fragmentos de PCR amplificados usando cebadores específicos de mMdr1a de 37 embriones de ratón cultivados *in vitro* después de ser microinyectados con ARN de ZFN contra Mdr1a de ratón y donante de RFLP con sitio NotI, junto con un control positivo y negativo para la amplificación por PCR. Los carriles 1 a 39 de la Figura 17B contienen los fragmentos de PCR de la Figura 17A después de realizar el ensayo de detección de mutaciones Surveyor™.

10 La **Figura 18**, paneles A y B, son imágenes fotográficas de fragmentos de ADN fluorescentemente teñidos resueltos en un gel de agarosa. Los carriles más a la izquierda y más a la derecha contienen una escalera de ADN. Los carriles contienen fragmentos de PCR amplificados usando cebadores específicos de mMdr1a de embriones de ratón mostrados en la Figura 17, y se digirieron con NotI sin purificar el producto de PCR. La Figura 18B es una serie más larga del mismo gel en la Figura 18A. Los productos de PCR no cortados son de aproximadamente 1,8 kb, y los productos digeridos son dos bandas de aproximadamente 900 pb.

15 La **Figura 19** es una imagen fotográfica de fragmentos de ADN fluorescentemente teñidos resueltos en un gel de agarosa. El carril más a la izquierda contiene una escalera de ADN. Los carriles 1 a 6 contienen algunos de los fragmentos de PCR como se muestra en la Figura 18 digeridos con NotI después de que los productos de PCR se purificaran en columna de manera que NotI pudiera funcionar en su tampón óptimo. Las líneas 7 y 8 son dos de las muestras digeridas con NotI (como en la Figura 18). Este gel muestra que la digestión de NotI en reacciones de PCR era completa.

20 La **Figura 20** es una imagen fotográfica de fragmentos de ADN fluorescentemente teñidos resueltos en un gel de agarosa. El carril más a la izquierda contiene una escalera de ADN. Los carriles 1 a 5 contienen fragmentos de PCR amplificados usando cebadores específicos de PXR de 1, 1/2, 1/6, 1/10, 1/30 de un blastocisto de rata. El carril 6 es un control positivo amplificado usando los mismos cebadores de ADN genómico de Sprague Dawley purificado.

25 La **Figura 21**, paneles A y B, son imágenes fotográficas de fragmentos de ADN fluorescentemente teñidos resueltos en un gel de agarosa. Los carriles más a la izquierda y más a la derecha contienen una escalera de ADN. La Figura 21A muestra fragmentos de PCR amplificados de embriones de rata cultivados *in vitro* después de la microinyección de ARNm de ZFN de PXR y el donante de RFLP de NotI, usando cebadores específicos de PXR y digeridos con NotI. La Figura 21B muestra los mismos fragmentos de PCR que en la Figura 21A después de realizar el ensayo de detección de mutaciones Surveyor™.

30 La **Figura 22** es una imagen fotográfica de fragmentos de ADN fluorescentemente teñidos resueltos en un gel de agarosa. Los primeros 4 carriles se amplifican por PCR de 4 fetos bien desarrollados 12,5 días después de la concepción de embriones inyectados con ARNm de ZFN de mMdr1a con el donante de RFLP de NotI. La PCR se digirió con NotI. El carril 4 es positivo. Los carriles 5-8 son 4 implantaciones abortadas de deciduas. Los cuatro fueron negativos.

35 La **Figura 23**, paneles A a E, son imágenes esquemáticas y fotográficas de fragmentos de ADN fluorescentemente teñidos resueltos en un gel de agarosa. La Figura 23A es un esquema que muestra la localización de los cebadores usados. Las Figuras 23B y 23C muestran resultados de los cebadores PF y GR. Las Figuras 23D y 23E muestran resultados de los cebadores PR + GF. El tamaño de fragmento esperado es 2,4 kb. Dos de los cuarenta fetos fueron positivos para GFP.

40 La **Figura 24** es una imagen fotográfica de fragmentos de ADN resueltos en un gel de agarosa. El carril 8 representa un feto de 13 dpc positivo para el sitio NotI.

DESCRIPCIÓN DETALLADA

En el presente documento se describen composiciones y métodos para la edición genómica en rata (por ejemplo, escisión de genes; alteración de genes, por ejemplo por escisión seguida de inserción (inserción física o inserción por replicación mediante reparación dirigida por homología) de una secuencia exógena y/o escisión seguida de la unión de extremos no homólogos (NHEJ); inactivación parcial o completa de uno o más genes; generación de alelos con mutaciones aleatorias para crear expresión alterada de genes endógenos; etc.) y alteraciones del genoma de rata que son llevadas en la línea germinal. También se describen métodos de preparación y uso de estas composiciones (reactivos), por ejemplo para editar (alterar) uno o más genes en una célula de rata diana. Así, los métodos y composiciones descritos en el presente documento proporcionan métodos altamente eficientes para la alteración génica dirigida (por ejemplo, activación) y/o inactivación (parcial o completa) de uno o más genes de rata y/o para la mutación aleatorizada de la secuencia de cualquier alelo diana, y, por tanto, permiten la generación de modelos animales de enfermedades humanas.

Las composiciones y métodos descritos en el presente documento proporcionan rotura dirigida rápida, completa y permanente de loci endógenos en ratas sin la necesidad de selección y/o cribado laborioso y con mínimos efectos inespecíficos. Las inactivaciones de genes animales completos también pueden ser fácilmente generadas en una única etapa inyectando ARNm de ZFN o casetes de expresión de ZFN.

5 Parte general

La práctica de los métodos, además de la preparación y uso de las composiciones desveladas en el presente documento, emplean, a menos que se indique lo contrario, técnicas convencionales en biología molecular, bioquímica, estructura y análisis de cromatina, química computacional, cultivo celular, ADN recombinante y campos relacionados que están dentro de la experiencia de la materia. Estas técnicas se explican completamente en la bibliografía. Véanse, por ejemplo, Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, Segunda Edición, Cold Spring Harbor Laboratory Press, 1989 y Tercera Edición, 2001; Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 y actualizaciones periódicas; la serie METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Tercera Edición, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman and A. P. Wolffe, eds.), Academic Press, San Diego, 1999; y METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Definiciones

Los términos "ácido nucleico", "polinucleótido" y "oligonucleótido" se usan indistintamente y se refieren a un polímero de desoxirribonucleótidos o ribonucleótidos, en conformación lineal o circular, y en tanto forma mono- como bicatenaria. Para los fines de la presente divulgación, estos términos no deben interpretarse como limitantes con respecto a la longitud de un polímero. Los términos pueden englobar análogos conocidos de nucleótidos naturales, además de nucleótidos que están modificados en los restos de bases, azúcares y/o fosfato (por ejemplo, esqueletos de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad por apareamiento de bases; es decir, un análogo de A se apareará con base con T.

Los términos "polipéptido", "péptido" y "proteína" se usan indistintamente para referirse a un polímero de restos de aminoácidos. El término también se aplica a polímeros de aminoácidos en los que uno o más aminoácidos son análogos químicos o derivados modificados de aminoácidos que existen de forma natural correspondientes.

"Unión" se refiere a una interacción no covalente específica de secuencia, entre macromoléculas (por ejemplo, entre una proteína y un ácido nucleico). No todos los componentes de una interacción de unión necesitan ser específicos de secuencia (por ejemplo, contactos con restos de fosfato en un esqueleto de ADN), en tanto que la interacción en conjunto sea específica de secuencia. Tales interacciones se caracterizan generalmente por una constante de disociación (K_d) de $10^{-6} M^{-1}$ o más baja. "Afinidad" se refiere a la intensidad de la unión: estando la elevada afinidad de unión correlacionada con una K_d más baja.

Una "proteína de unión" es una proteína que es capaz de unirse no covalentemente a otra molécula. Una proteína de unión puede unirse a, por ejemplo, una molécula de ADN (una proteína de unión a ADN), una molécula de ARN (una proteína de unión a ARN) y/o una molécula de proteína (una proteína de unión a proteína). En el caso de una proteína de unión a proteína, puede unirse a ella misma (para formar homodímeros, homotrímeros, etc.) y/o puede unirse a una o más moléculas de una proteína diferente o proteínas. Una proteína de unión puede tener más de un tipo de actividad de unión. Por ejemplo, las proteínas de dedos de cinc tienen actividad de unión a ADN, unión a ARN y unión a proteína.

Una "proteína de unión a ADN de dedos de cinc" (o dominio de unión) es una proteína, o un dominio dentro de una proteína más grande, que se une a ADN de una manera específica de secuencia mediante uno o más dedos de cinc, que son regiones de secuencia de aminoácidos dentro del dominio de unión cuya estructura se estabiliza mediante coordinación de un ión cinc. El término proteína de unión a ADN de dedos de cinc se abrevia frecuentemente proteína de dedos de cinc o ZFP.

Los dominios de unión de dedos de cinc pueden ser "manipulados" para unirse a una secuencia de nucleótidos predeterminada. Ejemplos no limitantes de métodos de manipulación de proteínas de dedos de cinc son diseño y selección. Una proteína de dedos de cinc diseñada es una proteína que no se encuentra en la naturaleza cuyo diseño/composición resulta principalmente de criterios racionales. Criterios racionales para el diseño incluyen la aplicación de reglas de sustitución y algoritmos computerizados para procesar información en una base de datos que guarda información de diseños de ZFP existentes y datos de unión. Véanse, por ejemplo, las patentes de EE.UU. 6.140.081; 6.453.242; y 6.534.261; véanse también los documentos WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 y WO 03/016496.

Una proteína de dedos de cinc "seleccionada" es una proteína no encontrada en la naturaleza cuya producción resulta principalmente de un proceso empírico tal como presentación en fagos, trampa de interacción o selección híbrida. Véanse, por ejemplo, los documentos US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197 y WO 02/099084.

El término "secuencia" se refiere a una secuencia de nucleótidos de cualquier longitud, que puede ser ADN o ARN; puede ser lineal, circular o ramificada y puede ser tanto monocatenaria como bicatenaria. El término "secuencia donante" se refiere a una secuencia de nucleótidos que se inserta en un genoma. Una secuencia donante puede ser de cualquier longitud, por ejemplo, entre 2 y 10.000 nucleótidos de longitud (o cualquier valor de número entero entremedias o por encima), preferentemente entre aproximadamente 100 y 1.000 nucleótidos de longitud (o cualquier número entero entremedias), más preferentemente entre aproximadamente 200 y 500 nucleótidos de longitud.

Una "secuencia no idéntica homóloga" se refiere a una primera secuencia que comparte un grado de identidad de secuencias con una segunda secuencia, pero cuya secuencia no es idéntica a la de la segunda secuencia. Por ejemplo, un polinucleótido que comprende la secuencia no mutante de un gen mutante es homólogo y no idéntico a la secuencia del gen mutante. En ciertas realizaciones, el grado de homología entre las dos secuencias es suficiente para permitir la recombinación homóloga entremedias, utilizando mecanismos celulares normales. Dos secuencias no idénticas homólogas pueden ser de cualquier longitud y su grado de no homología puede ser tan pequeño como un único nucleótido (por ejemplo, para la corrección de una mutación puntual genómica por recombinación homóloga dirigida) o de hasta 10 o más kilobases (por ejemplo, para la inserción de un gen en un sitio ectópico predeterminado en un cromosoma). Dos polinucleótidos que comprenden las secuencias no idénticas homólogas no necesitan ser de la misma longitud. Por ejemplo, puede usarse un polinucleótido exógeno (es decir, polinucleótido donante) de entre 20 y 10.000 nucleótidos o pares de nucleótidos.

En la técnica también se conocen técnicas para determinar la identidad de secuencias de ácidos nucleicos y de aminoácidos. Normalmente, tales técnicas incluyen determinar la secuencia de nucleótidos del ARNm para un gen y/o determinar la secuencia de aminoácidos así codificada, y comparar estas secuencias con un segundo nucleótido o secuencia de aminoácidos. También pueden determinarse y compararse secuencias genómicas de esta forma. En general, identidad se refiere a una correspondencia exacta de nucleótido con nucleótido o aminoácido con aminoácido de dos polinucleótidos o secuencias de polipéptidos, respectivamente. Pueden compararse dos o más secuencias (polinucleótido o aminoácido) determinando su identidad en porcentaje. La identidad en porcentaje de dos secuencias, tanto de secuencias de ácidos nucleicos como de aminoácidos, es el número de apareamientos exactos entre dos secuencias alineadas dividido entre la longitud de las secuencias más cortas y multiplicado por 100. Un alineamiento aproximado para secuencias de ácidos nucleicos se proporciona por el algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). Este algoritmo puede aplicarse a secuencias de aminoácidos usando la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*, M.O. Dayhoff ed., 5 supl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., EE.UU., y normalizado por Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). Una implementación a modo de ejemplo de este algoritmo para determinar la identidad en porcentaje de una secuencia se proporciona por Genetics Computer Group (Madison, WI) en la aplicación informática "BestFit". Los parámetros por defecto para este método se describen en *Wisconsin Analysis Package Program Manual*, versión 8 (1995) (disponible de Genetics Computer Group, Madison, WI). Un método preferido de establecimiento de la identidad en porcentaje en el contexto de la presente divulgación es usar el paquete MPSRCH de programas con los derechos registrados por la Universidad de Edimburgo, desarrollado por John F. Collins y Shane S. Sturrok, y distribuido por IntelliGenetics, Inc. (Mountain View, CA). De este juego de paquetes puede emplearse el algoritmo de Smith-Waterman cuando se usan parámetros por defecto para la tabla de puntuación (por ejemplo, penalización por hueco abierto de 12, penalización por extensión de hueco de uno, y un hueco de seis). De los datos generados, el valor de "Apareamiento" refleja la identidad de secuencias. Otros programas adecuados para calcular la identidad en porcentaje o similitud entre secuencias son generalmente conocidos en la técnica, por ejemplo, otro programa de alineamiento es BLAST, usado con parámetros por defecto. Por ejemplo, pueden usarse BLASTN y BLASTP usando los siguientes parámetros por defecto: código genético = estándar; filtro = ninguno; cadena = ambas; corte = 60; esperado = 10; Matriz = BLOSUM62; Descripción = 50 secuencias; clasificado por = PUNTUACIÓN ALTA; Bases de datos = no redundantes, traducciones de GenBank + EMBL + DDBJ + PDB + GenBank CDS + proteína Swiss + Spupdate + PIR. Detalles de estos programas pueden encontrarse en la siguiente dirección de internet: <http://www.ncbi.nlm.gov/cgi-bin/BLAST>. Con respecto a las secuencias descritas en el presente documento, el intervalo de grados deseados de identidad de secuencias es aproximadamente del 80 % al 100 % y cualquier valor de número entero entremedias. Normalmente, las identidades en porcentaje entre secuencias son de al menos el 70-75 %, preferentemente el 80-82 %, más preferentemente el 85-90 %, incluso más preferentemente el 92 %, todavía más preferentemente el 95 %, y lo más preferentemente el 98 % de identidad de secuencias.

Alternativamente, el grado de similitud de secuencias entre polinucleótidos puede determinarse por hibridación de polinucleótidos en condiciones que permiten la formación de dúplex estables entre regiones homólogas, seguido de digestión con nucleasa(s) específicas monocatenarias, y determinación del tamaño de los fragmentos digeridos. Dos secuencias de ácidos nucleicos, o dos de polipéptidos, son sustancialmente homólogas entre sí cuando las secuencias presentan al menos aproximadamente el 70 %-75 %, preferentemente el 80 %-82 %, más preferentemente el 85 %-90 %, incluso más preferentemente el 92 %, todavía más preferentemente el 95 %, y lo más preferentemente el 98 % de identidad de secuencias sobre una longitud definida de las moléculas, como se determina usando los métodos anteriores. Como se usa en el presente documento, sustancialmente homólogo también se refiere a secuencias que muestran identidad completa con una secuencia de ADN o de polipéptidos especificada. Las secuencias de ADN que son sustancialmente homólogas pueden identificarse en un experimento

de hibridación Southern bajo, por ejemplo, condiciones rigurosas, como se definen para ese sistema particular. La definición de condiciones de hibridación apropiadas está dentro de la experiencia de la materia. Véase, por ejemplo, Sambrook et al., arriba; *Nucleic Acid Hybridization: A Practical Approach*, editors B.D. Hames and S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

5 La hibridación selectiva de dos fragmentos de ácido nucleico puede determinarse del siguiente modo. El grado de identidad de secuencias entre dos moléculas de ácidos nucleicos afecta la eficiencia e intensidad de los eventos de hibridación entre tales moléculas. Una secuencia de ácidos nucleicos parcialmente idéntica inhibirá al menos parcialmente la hibridación de una secuencia completamente idéntica con una molécula diana. La inhibición de la hibridación de la secuencia completamente idéntica puede evaluarse usando ensayos de hibridación que son muy conocidos en la técnica (por ejemplo, transferencia Southern (ADN), transferencia Northern (ARN), hibridación en disolución, o similares, véase Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, segunda edición, (1989) Cold Spring Harbor, N.Y.). Tales ensayos pueden realizarse usando grados variables de selectividad, por ejemplo, usando condiciones variables de baja a alta rigurosidad. Si se emplean condiciones de baja rigurosidad, la ausencia de unión no específica puede evaluarse usando una sonda secundaria que carece de incluso un grado parcial de identidad de secuencias (por ejemplo, una sonda que tiene menos de aproximadamente el 30 % de identidad de secuencias con la molécula diana), de forma que, en ausencia de eventos de unión no específicos, la sonda secundaria no se hibridará con la diana.

Si se utiliza un sistema de detección basado en hibridación, se elige una sonda de ácido nucleico que es complementaria a una secuencia de ácidos nucleicos de referencia, y luego por selección de condiciones apropiadas la sonda y la secuencia de referencia se hibridan selectivamente, o se unen, entre sí para formar una molécula de dúplex. Una molécula de ácido nucleico que puede hibridarse selectivamente con una secuencia de referencia bajo condiciones de hibridación de rigurosidad moderada normalmente se hibrida bajo condiciones que permitan la detección de una secuencia de ácidos nucleicos diana de al menos aproximadamente 10-14 nucleótidos de longitud que al menos aproximadamente el 70% de identidad de secuencias con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación rigurosas permiten normalmente la detección de secuencias de ácidos nucleicos diana de al menos aproximadamente 10-14 nucleótidos de longitud que tienen una identidad de secuencias superior a aproximadamente el 90-95 % con la secuencia de la sonda de ácido nucleico seleccionada. Las condiciones de hibridación útiles para la hibridación de secuencias de la sonda/de referencia, en las que la secuencia de la sonda y de referencia tienen un grado específico de identidad de secuencias, pueden determinarse como se conoce en la técnica (véase, por ejemplo, *Nucleic Acid Hybridization: A Practical Approach*, editors B.D. Hames and S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

Las condiciones para la hibridación son muy conocidas para aquellos expertos en la materia. Rigurosidad de la hibridación se refiere al grado al que las condiciones de hibridación desfavorecen la formación de híbridos que contienen nucleótidos desapareados, correlacionándose la mayor rigurosidad con una menor tolerancia por híbridos desapareados. Los factores que afectan la rigurosidad de hibridación son muy conocidos para aquellos expertos en la materia e incluyen, pero no se limitan a, temperatura, pH, fuerza iónica y concentración de disolventes orgánicos tales como, por ejemplo, formamida y sulfóxido de dimetilo. Como es conocido por aquellos expertos en la materia, la rigurosidad de hibridación aumenta por mayores temperaturas, menor fuerza iónica y menores concentraciones de disolvente.

40 Con respecto a las condiciones de rigurosidad para la hibridación, es muy conocido en la técnica que puedan emplearse numerosas condiciones equivalentes para establecer una rigurosidad particular variando, por ejemplo, los siguientes factores: la longitud y la naturaleza de las secuencias, composición básica de las diversas secuencias, concentraciones de sales y otros componentes de la disolución de hibridación, la presencia o ausencia de agentes de bloqueo en las disoluciones de hibridación (por ejemplo, sulfato de dextrano y polietilenglicol), parámetros de temperatura de reacción y de tiempo de hibridación, además de condiciones de lavado variables. La selección de un conjunto particular de condiciones de hibridación se selecciona siguiendo métodos estándar en la técnica (véanse, por ejemplo, Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, segunda edición, (1989) Cold Spring Harbor, N.Y.).

"Recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. Para los fines de la presente divulgación, "recombinación homóloga (HR)" se refiere a la forma especializada de tal intercambio que tiene lugar, por ejemplo, durante la reparación de roturas bicatenarias en células mediante mecanismos de reparación dirigidos por homología. Este proceso requiere homología de secuencias de nucleótidos, usa una molécula "donante" para la reparación del molde de una molécula "diana" (es decir, la que experimentó la rotura bicatenaria), y se conoce indistintamente como "conversión génica no de entrecruzamiento" o "conversión génica de extensión corta", debido a que conduce a la transferencia de información genética del donante a la diana. Sin desear ceñirse a teoría alguna particular, tal transferencia puede implicar la corrección de desapareamientos del ADN de heterodúplex que se forma entre la diana rota y el donante, y/o la "hibridación de hebras dependiente de la síntesis", en la que el donante se usa para resintetizar información genética que será parte de la diana, y/o procesos relacionados. Tal HR especializada frecuentemente produce una alteración de la secuencia de la molécula diana de forma que parte o toda la secuencia del polinucleótido donante se incorpora en el polinucleótido diana.

En los métodos descritos en el presente documento, una o más nucleasas dirigidas como se describe en el presente documento crean una rotura bicatenaria en la secuencia diana (por ejemplo, cromatina celular) en un sitio predeterminado, y un polinucleótido "donante", que tiene homología con la secuencia de nucleótidos en la región de la rotura, puede introducirse en la célula. Se ha mostrado que la presencia de la rotura bicatenaria facilita la integración de la secuencia donante. La secuencia donante puede ser físicamente integrada o, alternativamente, el polinucleótido donante se usa como molde para la reparación de la rotura mediante recombinación homóloga, produciendo la introducción de toda o parte de la secuencia de nucleótidos como en el donante en la cromatina celular. Así, puede alterarse una primera secuencia en la cromatina celular y, en ciertas realizaciones, puede convertirse en una secuencia presente en un polinucleótido donante. Así, el uso de los términos "sustituir" o "sustitución" puede entenderse que representa la sustitución de una secuencia de nucleótidos por otra (es decir, sustitución de una secuencia en el sentido informativo), y no requiere necesariamente la sustitución física o química de un polinucleótido por otro.

En cualquiera de los métodos descritos en el presente documento, pueden usarse pares adicionales de proteínas de dedos de cinc para la escisión bicatenaria adicional de sitios diana adicionales dentro de la célula.

En ciertas realizaciones de métodos para la recombinación dirigida y/o sustitución y/o alteración de una secuencia en una región de interés en cromatina celular, se altera una secuencia cromosómica por recombinación homóloga con una secuencia de nucleótidos "donante" exógena. Tal recombinación homóloga se estimula por la presencia de una rotura bicatenaria en la cromatina celular, si están presentes secuencias homólogas a la región de la rotura.

En cualquiera de los métodos descritos en el presente documento, la primera secuencia de nucleótidos (la "secuencia donante") puede contener secuencias que son homólogas, pero no idénticas, a secuencias genómicas en la región de interés, estimulando así la recombinación homóloga para insertar una secuencia no idéntica en la región de interés. Porciones de la secuencia donante que son homólogas a secuencias en la región de interés pueden presentar entre aproximadamente el 80 y el 99 % (o cualquier número entero intermedio) de identidad de secuencia con la secuencia genómica que se sustituye. La homología entre la secuencia donante y genómica puede ser superior al 99 %, por ejemplo si solo se diferencia 1 nucleótido como entre secuencias donantes y genómicas de más de 100 pares de bases contiguos. En ciertos casos, una porción no homóloga de la secuencia donante puede contener secuencias no presentes en la región de interés, de forma que se introducen nuevas secuencias en la región de interés. En estos casos, la secuencia no homóloga está generalmente flanqueada por secuencias de 50-1.000 pares de bases (o cualquier valor entero intermedio) o cualquier número de pares de bases superior a 1.000, que son homólogos o idénticos a secuencias en la región de interés. En otras realizaciones, la secuencia donante es no homóloga a la primera secuencia, y se inserta en el genoma por mecanismos de recombinación no homóloga.

Cualquiera de los métodos descritos en el presente documento puede usarse para la inactivación parcial o completa de una o más secuencias diana en una célula por integración dirigida de secuencia donante que rompe la expresión del (de los) gen(es) de interés. También se describen líneas celulares con genes parcialmente o completamente inactivados.

Además, también pueden usarse los métodos de integración dirigida como se describen en el presente documento para integrar una o más secuencias exógenas. La secuencia de ácidos nucleicos exógena puede comprender, por ejemplo, uno o más genes o moléculas de ADNc, o cualquier tipo de secuencia codificante o no codificante, además de uno o más elementos de control (por ejemplo, promotores). Además, la secuencia de ácidos nucleicos exógena puede producir una o más moléculas de ARN (por ejemplo, ARN de horquilla pequeña (ARNhp), ARN inhibidores (ARNi), microARN (miARN), etc.).

"Escisión" se refiere a la rotura del esqueleto covalente de una molécula de ADN. La escisión puede iniciarse mediante una variedad de métodos que incluyen, pero no se limitan a, hidrólisis enzimática o química de un enlace fosfodiéster. Tanto la escisión monocatenaria como la escisión bicatenaria son posibles, y la escisión bicatenaria puede producirse como resultado de dos eventos de escisión monocatenaria distintos. La escisión de ADN puede producir la producción de tanto extremos romos como extremos en bisel. En ciertas realizaciones, los polipéptidos de fusión se usan para la escisión de ADN bicatenario dirigida.

Un "medio-dominio de escisión" es una secuencia de polipéptidos que, conjuntamente con un segundo polipéptido (tanto idéntico como diferente), forma un complejo que tiene actividad de escisión (preferentemente actividad de escisión bicatenaria). Los términos "primer y segundo medio-dominios de escisión", "medio-dominios de escisión + y -" y "medio-dominios de escisión derecho e izquierdo" se usan indistintamente para referirse a pares de medio-dominios de escisión que se dimerizan.

Un "medio-dominio de escisión manipulado" es un medio-dominio de escisión que ha sido modificado para formar heterodímeros obligados con otro medio-dominio de escisión (por ejemplo, otro medio-dominio de escisión manipulado). Véanse, por tanto, las publicaciones de patente de EE.UU. N.º 2005/0064474; 2007/0218528 y 2008/0131962.

"Cromatina" es la estructura de nucleoproteína que comprende el genoma celular. La cromatina celular comprende ácido nucleico, principalmente ADN, y proteína, que incluye histonas y proteínas cromosómicas no histonas. La

mayoría de la cromatina celular eucariota existe en forma de nucleosomas, en la que un núcleo de nucleosoma comprende aproximadamente 150 pares de bases de ADN asociados a un octámero que comprende dos de cada una de las histonas H2A, H2B, H3 y H4; y el ADN de conector (de longitud variable dependiendo del organismo) se extiende entre los núcleos de nucleosomas. Una molécula de histona H1 está generalmente asociada al ADN de conector. Para los fines de la presente divulgación, se indica que el término "cromatina" engloba todos los tipos de nucleoproteína celular, tanto procariota como eucariota. La cromatina celular incluye tanto cromatina cromosómica como episómica.

Un "cromosoma" es un complejo de cromatina que comprende toda o una porción del genoma de una célula. El genoma de una célula se caracteriza frecuentemente por su cariotipo, que es el conjunto de todos los cromosomas que comprenden el genoma de la célula. El genoma de una célula puede comprender uno o más cromosomas.

Un "episoma" es un ácido nucleico replicante, complejo de nucleoproteína u otra estructura que comprende un ácido nucleico que no es parte del cariotipo cromosómico de una célula. Ejemplos de episomas incluyen plásmidos y ciertos genomas virales.

Un "sitio diana" o "secuencia diana" es una secuencia de ácidos nucleicos que define una porción de un ácido nucleico con el que se unirá una molécula de unión, proporcionando condiciones suficientes para que exista la unión. Por ejemplo, la secuencia 5'-GAATTC-3' es un sitio diana para la endonucleasa de restricción Eco RI.

Una molécula "exógena" es una molécula que no está normalmente presente en una célula, pero puede introducirse en una célula por uno o más métodos genéticos, bioquímicos u otros. La "presencia normal en la célula" se determina con respecto a la etapa de desarrollo particular y las condiciones medioambientales de la célula. Así, por ejemplo, una molécula que está presente solo durante el desarrollo embrionario de músculo es una molécula exógena con respecto a una célula de músculo adulta. Similarmente, una molécula inducida por choque térmico es una molécula exógena con respecto a una célula no de choque térmico. Una molécula exógena puede comprender, por ejemplo, una versión con funcionalidad de una molécula endógena disfuncional o una versión disfuncional de una molécula endógena que funciona normalmente. Una molécula exógena puede comprender, por ejemplo, una versión que funciona de una molécula endógena que funciona mal o una versión que funciona mal de una molécula endógena que funciona normalmente. Una molécula exógena también puede ser una molécula normalmente encontrada en otras especies, por ejemplo, una secuencia humana introducida en un genoma de rata.

Una molécula exógena puede ser, entre otras cosas, una molécula pequeña, tal como se genera por un proceso de química combinatoria, o una macromolécula tal como una proteína, ácido nucleico, hidrato de carbono, lípido, glicoproteína, lipoproteína, polisacárido, cualquier derivado modificado de las moléculas anteriores, o cualquier complejo que comprende una o más de las moléculas anteriores. Los ácidos nucleicos incluyen ADN y ARN, pueden ser mono- o bicatenarios; pueden ser lineales, ramificados o circulares; y pueden ser de cualquier longitud. Los ácidos nucleicos incluyen aquellos que pueden formar dúplex, además de ácidos nucleicos que forman tríplex. Véanse, por ejemplo, las patentes de EE.UU. N.º 5.176.996 y 5.422.251. Las proteínas incluyen, pero no se limitan a, proteínas de unión a ADN, factores de transcripción, factores de remodelación de cromatina, proteínas de unión a ADN metilado, polimerasas, metilasas, desmetilasas, acetilasas, desacetilasas, cinasas, fosfatasa, integrasas, recombinasas, ligasas, topoisomerasas, girasas y helicasas.

Una molécula exógena puede ser del mismo tipo de molécula que una molécula endógena, por ejemplo, una proteína exógena o ácido nucleico. Por ejemplo, un ácido nucleico exógeno puede comprender un genoma viral infectante, un plásmido o episoma introducido en una célula, o un cromosoma que normalmente no está presente en la célula. Métodos para la introducción de moléculas exógenas en células son conocidos para aquellos expertos en la materia e incluyen, pero no se limitan a, transferencia mediada por lípidos (es decir, liposomas, que incluyen lípidos neutros y catiónicos), electroporación, inyección directa, fusión de células, bombardeo de partículas, co-precipitación con fosfato de calcio, transferencia mediada por DEAE-dextrano y transferencia mediada por vector viral.

Por el contrario, una molécula "endógena" es una que normalmente está presente en una célula particular en una etapa de desarrollo particular bajo condiciones medioambientales particulares. Por ejemplo, un ácido nucleico endógeno puede comprender un cromosoma, el genoma de una mitocondria, cloroplasto u otro orgánulo, o un ácido nucleico episómico que se produce naturalmente. Moléculas endógenas adicionales pueden incluir proteínas, por ejemplo, factores de transcripción y enzimas.

Una molécula de "fusión" es una molécula en la que dos o más moléculas de subunidad están asociadas, preferentemente covalentemente. Las moléculas de subunidad pueden ser el mismo tipo químico de molécula, o pueden ser tipos químicos diferentes de moléculas. Ejemplos del primer tipo de molécula de fusión incluyen, pero no se limitan a, proteínas de fusión (por ejemplo, una fusión entre un dominio de unión de ADN de ZFP y un dominio de escisión) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica la proteína de fusión descrita arriba). Ejemplos del segundo tipo de molécula de fusión incluyen, pero no se limitan a, una fusión entre un ácido nucleico formador de tríplex y un polipéptido, y una fusión entre un ligando de unión al surco menor y un ácido nucleico.

La expresión de una proteína de fusión en una célula puede resultar de la administración de la proteína de fusión a la célula o por administración de un polinucleótido que codifica la proteína de fusión a una célula, en la que el polinucleótido se transcribe, y el transcrito se traduce, para generar la proteína de fusión. También pueden participar corte y empalme en trans, escisión de polipéptidos y ligación de polipéptidos en la expresión de una proteína en una célula. Métodos para la administración de polinucleótidos y polipéptidos a células se presentan en cualquier parte en la presente divulgación.

Un "gen", para los fines de la presente divulgación, incluye una región de ADN que codifica un producto génico (véase más adelante), además de todas las ADN regiones que regulan la producción del producto génico, tanto si tales secuencias reguladoras son adyacentes o no a secuencias de codificación y/o transcritas. Por consiguiente, un gen incluye, pero no se limita necesariamente a, secuencias promotoras, terminadores, secuencias reguladoras de la traducción tales como sitios de unión al ribosoma y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aislantes, elementos de frontera, orígenes de replicación, sitios de unión a la matriz y regiones de control de locus.

"Expresión génica" se refiere a la conversión de la información, contenida en un gen, en un producto génico. Un producto génico puede ser el producto de la transcripción directa de un gen (por ejemplo, ARNm, ARNt, ARNr, ARN antisentido, ribozima, ARN estructural o cualquier otro tipo de ARN) o una proteína producida por la traducción de un ARNm. Productos génicos también incluyen ARN que se modifican, por procesos tales como polarización, poliadenilación, metilación y edición, y proteínas modificadas por, por ejemplo, metilación, acetilación, fosforilación, ubiquitinación, ribosilación de ADP, miristilación y glucosilación.

"Modulación" de la expresión génica se refiere a un cambio en la actividad de un gen. La modulación de la expresión puede incluir, pero no se limita a, activación génica y represión génica. Puede usarse edición de genoma (por ejemplo, escisión, alteración, inactivación, mutación aleatoria) para modular la expresión. Inactivación génica se refiere a cualquier reducción en la expresión génica en comparación con una célula que no incluye una ZFP como se describe en el presente documento. Así, la inactivación génica puede ser parcial o completa.

Una "región de interés" es cualquier región de cromatina celular, tal como, por ejemplo, un gen o una secuencia no codificante dentro de o adyacente a un gen, en la que se desea unir una molécula exógena. La unión puede ser para los fines de escisión de ADN dirigida y/o recombinación dirigida. Una región de interés puede estar presente en, por ejemplo, un cromosoma, un episoma, un genoma organular (por ejemplo, mitocondria, cloroplasto), o un genoma viral infectante. Una región de interés puede estar dentro de la región codificante de un gen, dentro de regiones no codificantes transcritas tales como, por ejemplo, secuencias conductoras, secuencias de transporte o intrones, o dentro de regiones no transcritas, tanto en la dirección 5' como en la dirección 3' de la región codificante. Una región de interés puede ser tan pequeña como un único par de nucleótidos o hasta 2.000 pares de nucleótidos de longitud, o cualquier valor entero de pares de nucleótidos.

Los términos "enlace operativo" y "operativamente unido" (u "operablemente unido") se usan indistintamente con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en los que los componentes están dispuestos de forma que ambos componentes funcionan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar en una función que es ejercida por al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está operativamente unida a una secuencia codificante si la secuencia reguladora transcripcional controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o ausencia de uno o más factores reguladores de la transcripción. Una secuencia reguladora de la transcripción está generalmente operativamente unida en *cis* con una secuencia codificante, pero no necesita ser directamente adyacente a ella. Por ejemplo, un potenciador es una secuencia reguladora transcripcional que está operativamente unida a una secuencia codificante, aún cuando no sean contiguas.

Con respecto a polipéptidos de fusión, el término "operativamente ligado" puede referirse al hecho de que cada uno de los componentes realice la misma función en el enlace con el otro componente que haría si no estuviera así ligado. Por ejemplo, con respecto a un polipéptido de fusión en el que un dominio de unión de ADN de ZFP está fusionado con un dominio de escisión, el dominio de unión de ADN de ZFP y el dominio de escisión están en enlace operativo si, en el polipéptido de fusión, la porción del dominio de unión de ADN de ZFP puede unirse a su sitio diana y/o su sitio de unión, mientras que el dominio de escisión puede escindir ADN en la proximidad del sitio diana.

Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, pero retiene la misma función que la proteína, polipéptido o ácido nucleico de longitud completa. Un fragmento funcional puede poseer más, menos, o el mismo número de residuos que la molécula nativa correspondiente, y/o puede contener uno o más sustituciones de aminoácidos o de nucleótidos. Métodos para determinar la función de un ácido nucleico (por ejemplo, función codificante, capacidad para hibridarse con otro ácido nucleico) son muy conocidos en la técnica. Similarmente, métodos para determinar la función de proteínas son muy conocidos. Por ejemplo, la función de unión del ADN de un polipéptido puede determinarse, por ejemplo, por unión a filtro, desplazamiento de la movilidad electroforética o ensayos de inmunoprecipitación. La escisión de ADN puede ensayarse por electroforesis en gel. Véase Ausubel et al., arriba. La capacidad de una proteína para interactuar con otra proteína puede determinarse,

por ejemplo, por co-inmunoprecipitación, ensayos de dos híbridos o complementación, tanto genética como bioquímica. Véase, por ejemplo, Fields et al., (1989) *Nature* 340:245-246; patente de EE.UU. N.º 5.585.245 y documento PCT WO 98/44350.

Nucleasas con dedos de cinc

- 5 En el presente documento se describen nucleasas con dedos de cinc (ZFN) que pueden usarse para la edición genómica (por ejemplo, escisión, alteración, inactivación y/o mutación aleatoria) de uno o más genes de rata. Las ZFN comprenden una proteína de dedos de cinc (ZFP) y un dominio (de escisión) de nucleasa (por ejemplo, medio-dominio de escisión).

A. Proteínas de dedos de cinc

- 10 Los dominios de unión de dedos de cinc pueden manipularse para unirse a una secuencia de elección. Véanse, por ejemplo, Beerli et al., (2002) *Nature Biotechnol.* 20:135-141; Pabo et al., (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan et al., (2001) *Nature Biotechnol.* 19:656-660; Segal et al., (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo et al., (2000) *Curr. Opin. Struct. Biol.* 10:411-416. Un dominio de unión de dedos de cinc manipulado puede tener una especificidad de unión novedosa, en comparación con una proteína de dedos de cinc que se produce naturalmente.
- 15 Los métodos de manipulación incluyen, pero no se limitan a, diseño racional y diversos tipos de selección. El diseño racional incluye, por ejemplo, usar bases de datos que comprenden secuencias de triplete (o cuadruplete) de nucleótidos y secuencias de aminoácidos de dedos de cinc individuales, en las que cada secuencia de nucleótidos de triplete o cuadruplete está asociada con una o más secuencias de aminoácidos de dedos de cinc que se unen a la secuencia de triplete o cuadruplete particular. Véanse, por ejemplo, las patentes de EE.UU. del mismo solicitante 6.453.242 y 6.534.261, incorporadas por referencia en el presente documento en sus totalidades.

- Métodos de selección a modo de ejemplo, que incluyen presentación en fagos y sistemas de dos híbridos, se desvelan en las patentes de EE.UU. 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; además de los documentos WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237. Además, se ha descrito el potenciamiento de la especificidad de unión por dominios de unión de dedos de cinc, por ejemplo, en el documento WO 02/077227 del mismo solicitante.

- 25 Selección de sitios diana; ZFP y métodos de diseño y construcción de proteínas de fusión (y polinucleótidos que codifican las mismas) son conocidos para aquellos expertos en la materia y se describen en detalle en las publicaciones de solicitud de patente de EE.UU. N.º 20050064474 y 20060188987, incorporadas por referencia en sus totalidad en el presente documento.

- 30 Además, como se ha desvelado en estas y otras referencias, los dominios de dedos de cinc y/o las proteínas de dedos de cinc multi-dedos pueden unirse juntas usando cualquier secuencia conectora adecuada, que incluye, por ejemplo, conectores de 5 o más aminoácidos de longitud (por ejemplo, TGEKP (SEQ ID NO:1), TGGQRP (SEQ ID NO:2), TGQKP (SEQ ID NO:3) y/o TGSQKP (SEQ ID NO:4)). Véanse, por tanto, las patentes de EE.UU. N.º 6.479.626; 6.903.185; y 7.153.949 para secuencias conectoras a modo de ejemplo de 6 o más aminoácidos de longitud. Las proteínas descritas en el presente documento pueden incluir cualquier combinación de conectores adecuados entre los dedos de cinc individuales de la proteína.

- Como se describe más adelante, un dominio de unión de cuatro, cinco o seis dedos está fusionado con un medio-dominio de escisión, tal como, por ejemplo, el dominio de escisión de una endonucleasa de restricción de tipo IIS tal como *Fok I*. Se usan uno o más pares de tales fusiones de medio-dominios de dedo de cinc/nucleasa para la escisión dirigida, como se desvela, por ejemplo, en la publicación de patente de EE.UU. N.º 20050064474.

- 40 Para la escisión dirigida, los bordes próximos de los sitios de unión pueden separarse 5 o más pares de nucleótidos, y cada una de las proteínas de fusión puede unirse a una hebra opuesta del ADN diana. Todas las combinaciones en parejas 1 pueden usarse para la escisión dirigida de un gen de rata. Siguiendo la presente divulgación, las ZFN pueden dirigirse a cualquier secuencia en el genoma de rata.

- 45 En algunas realizaciones, el dominio de unión de ADN es un dominio manipulado de un efector TAL derivado del patógeno de planta *Xanthomonas* (véanse Boch et al, (2009) *Science* 29 Oct 2009 (10.1126/science, 117881) y Moscou y Bogdanove, (2009) *Science* 29 Oct 2009 (10.1126/science, 1178817).

B. Dominios de escisión

- 50 Las ZFN también comprenden una nucleasa (dominio de escisión, medio-dominio de escisión). La porción del dominio de escisión de las proteínas de fusión desveladas en el presente documento puede obtenerse de cualquier endonucleasa o exonucleasa. Endonucleasas a modo de ejemplo de las que puede derivarse un dominio de escisión incluyen, pero no se limitan a, endonucleasas de restricción y endonucleasas de asentamiento. Véase, por ejemplo, 2002-2003 Catalogue, New England Biolabs, Beverly, MA; y Belfort et al., (1997) *Nucleic Acids Res.* 25:3379-3388. Se conocen enzimas adicionales que escinden ADN (por ejemplo, nucleasa S1; nucleasa de judías mungo; DNasa I pancreática; nucleasa microcócica; endonucleasa HO de levadura; véase también Linn et al., (eds.) *Nucleases*, Cold

Spring Harbor Laboratory Press,1993). Una o más de estas enzimas (o fragmentos funcionales de las mismas) pueden usarse como fuente de dominios de escisión y medio-dominios de escisión.

5 Similarmente, un medio-dominio de escisión puede derivarse de cualquier nucleasa o porción de la misma, como se expone anteriormente, que requiere dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden medio-dominios de escisión. Alternativamente, puede usarse una única proteína que comprende dos medio-dominios de escisión. Los dos medio-dominios de escisión pueden derivarse de la misma endonucleasa (o fragmentos funcionales de la misma), o cada medio-dominio de escisión puede derivarse de una endonucleasa diferente (o fragmentos funcionales de la misma). Además, los sitios diana para las dos proteínas de fusión están preferentemente dispuestos, el uno con respecto al otro, tal que la unión de las dos proteínas de fusión a sus sitios diana respectivos ponga los medio-dominios de escisión en una orientación espacial entre sí que permita que los medio-dominios de escisión formen un dominio de escisión funcional, por ejemplo, por dimerización. Así, en ciertas realizaciones, los bordes próximos de los sitios diana están separados por 5-8 nucleótidos o por 15-18 nucleótidos. Sin embargo, puede interponerse cualquier número entero de nucleótidos o pares de nucleótidos entre dos sitios diana (por ejemplo, de 2 a 50 pares de nucleótidos o más). En general, el sitio de escisión se encuentra entre los sitios diana.

10 Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unión específica de secuencia a ADN (en un sitio de reconocimiento), y escindir ADN en o cerca del sitio de unión. Ciertas enzimas de restricción (por ejemplo, tipo IIS) escinden ADN en sitios eliminados del sitio de reconocimiento y tienen dominios de unión y de escisión separables. Por ejemplo, la enzima de tipo IIS Fok I cataliza la escisión de ADN bicatenario, en 9 nucleótidos desde su sitio de reconocimiento sobre una cadena y 13 nucleótidos desde su sitio de reconocimiento sobre la otra. Véanse, por ejemplo, las patentes de EE.UU. 5.356.802; 5.436.150 y 5.487.994; además de Li et al., (1992) Proc. Natl. Acad. Sci. USA 89:4275-4279; Li et al., (1993) Proc. Natl. Acad. Sci. USA 90:2764-2768; Kim et al., (1994a) Proc. Natl. Acad. Sci. USA 91:883-887; Kim et al., (1994b) J. Biol. Chem. 269:31,978-31,982. Así, en una realización, las proteínas de fusión comprenden el dominio de escisión (o medio-dominio de escisión) desde al menos una enzima de restricción de tipo IIS y uno o más dominios de unión de dedos de cinc, que pueden o pueden no estar manipulados.

20 Una enzima de restricción de tipo IIS a modo de ejemplo, cuyo dominio de escisión es separable del dominio de unión, es *Fok I*. Esta enzima particular es activa como dímero. Bitinaite et al., (1998) Proc. Natl. Acad. Sci. USA 95:10.570-10.575. Por consiguiente, para los fines de la presente divulgación, la porción de la enzima *Fok I* usada en las proteínas de fusión desveladas se considera un medio-dominio de escisión. Así, para la escisión bicatenaria dirigida y/o sustitución dirigida de secuencias celulares usando fusiones de dedo de cinc-*Fok I*, pueden usarse dos proteínas de fusión, cada una que comprende un medio-dominio de escisión de *FokI*, para reconstituir un dominio de escisión catalíticamente activo. Alternativamente, también pueden usarse una única molécula de polipéptido que contiene un dominio de unión de dedos de cinc y dos medio-dominios de escisión de *Fok I*. Parámetros para la escisión dirigida y alteración de secuencias dirigida usando fusiones de dedo de cinc-*Fok I* se proporcionan en cualquier parte en la presente divulgación.

30 Un dominio de escisión o medio-dominio de escisión puede ser cualquier porción de una proteína que retiene actividad de escisión, o que retiene la capacidad para multimerizarse (por ejemplo, dimerizarse) para formar un dominio de escisión funcional.

40 Enzimas de restricción de tipo IIS a modo de ejemplo se describen en la publicación internacional WO 07/014275. Enzimas de restricción adicionales también contienen dominios de unión y de escisión separables, y éstos se contemplan por la presente divulgación. Véase, por ejemplo, Roberts et al., (2003) Nucleic Acids Res. 31:418-420.

45 El dominio de escisión puede comprender uno o más medio-dominios de escisión manipulados (también denominados mutantes de dominio de dimerización) que minimizan o previenen la homodimerización, como se describe, por ejemplo, en las publicaciones de patente de EE.UU. N.º 20050064474; 20060188987 y 20080131962. Restos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 de *Fok I* son todos dianas para influir en la dimerización de los medio-dominios de escisión de *Fok I*.

50 Medio-dominios de escisión a modo de ejemplo manipulados de *Fok I* que forman heterodímeros obligados incluyen un par en el que un primer medio-dominio de escisión incluye mutaciones en restos de aminoácidos en las posiciones 490 y 538 de *Fok I* y un segundo medio-dominio de escisión incluye mutaciones en restos de aminoácidos 486 y 499.

55 Así, en una realización, una mutación en 490 sustituye Glu (E) con Lys (K); la mutación en 538 sustituye Iso (I) con Lys (K); la mutación en 486 sustituye Gln (Q) con Glu (E); y la mutación en la posición 499 sustituye Iso (I) con Lys (K). Específicamente, los medio-dominios de escisión manipulados descritos en el presente documento se prepararon mutando las posiciones 490 (E→K) y 538 (I→K) en un medio-dominio de escisión para producir un medio-dominio de escisión manipulado designado "E490K:I538K" y mutando las posiciones 486 (Q→E) y 499 (I→L) en otro medio-dominio de escisión para producir un medio-dominio de escisión manipulado designado "Q486E:I499L". Los medio-dominios de escisión manipulados descritos en el presente documento son mutantes de

heterodímero obligados en los que la escisión anormal se minimiza o suprime. Véase, por ejemplo, el Ejemplo 1 de la publicación de patente de EE.UU. N.º 2008/0131962.

5 Los medio-dominios de escisión manipulados descritos en el presente documento pueden prepararse usando cualquier método adecuado, por ejemplo, por mutagénesis dirigida al sitio de medio-dominios de escisión no mutantes (*Fok I*) como se describe en la publicación de patente de EE.UU. N.º 20050064474 (N.º de serie 10/912.932, Ejemplo 5).

C. Métodos adicionales para la escisión dirigida en rata

10 Puede usarse cualquier nucleasa que tenga un sitio diana en cualquier gen de rata en los métodos descritos en el presente documento. Por ejemplo, las endonucleasas de migración y las meganucleasas tienen secuencias de reconocimiento muy largas, algunas de las cuales es probable que estén presentes, en una base estadística, una vez en un genoma de tamaño humano. Cualquier nucleasa tal que tenga un sitio diana en un gen de rata puede usarse en lugar de, o además de, una nucleasa con dedos de cinc, para la escisión dirigida en un gen de rata.

15 Endonucleasas de migración a modo de ejemplo incluyen *I-SceI*, *I-CeuI*, *PI-PspI*, *PI-Sce*, *I-SceIV*, *I-CsmI*, *I-PanI*, *I-Scell*, *I-PpoI*, *I-ScellI*, *I-CreI*, *I-TevI*, *I-TevII* y *I-TevIII*. Sus secuencias de reconocimiento son conocidas. Véase también la patente de EE.UU. N.º 5.420.032; patente de EE.UU. N.º 6.833.252; Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J. Mol. Biol.* 263:163-180; Argast et al. (1998) *J. Mol. Biol.* 280:345-353 y el catálogo de New England Biolabs.

20 Aunque la especificidad por escisión de la mayoría de las endonucleasas de migración no es absoluta con respecto a sus sitios de reconocimiento, los sitios son de longitud suficiente como para que pueda obtenerse un único evento de escisión por el genoma de tamaño de mamífero expresando una endonucleasa de migración en una célula que contiene una única copia de su sitio de reconocimiento. También se ha informado que la especificidad de endonucleasas de migración y meganucleasas puede manipularse para unirse a sitios diana no naturales. Véanse, por ejemplo, Chevalier et al. (2002) *Molec. Cell* 10:895-905; Epinat et al. (2003) *Nucleic Acids Res.* 31:2952-2962; Ashworth et al. (2006) *Nature* 441:656-659; Paques et al. (2007) *Current Gene Therapy* 7:49-66.

Administración

Las ZFN descritas en el presente documento pueden administrarse a una célula de rata diana por cualquier medio adecuado, que incluye, por ejemplo, mediante inyección de ARNm de ZFN. Véase Hammerschmidt et al. (1999) *Methods Cell Biol.* 59:87-115

30 Métodos de administración de proteínas que comprenden dedos de cinc se describen, por ejemplo, en las patentes de EE.UU. N.º 6.453.242; 6.503.717; 6.534.261; 6.599.692; 6.607.882; 6.689.558; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824.

35 Las ZFN, como se describen en el presente documento, también pueden administrarse usando vectores que contienen secuencias que codifican una o más de las ZFN. Puede usarse cualquier sistema de vector que incluye, pero no se limita a, vectores plasmídicos, vectores retrovirales, vectores lentivirales, vectores de adenovirus, vectores de poxvirus; vectores de virus del herpes y vectores de virus adeno-asociado, etc. Véanse, por tanto, las patentes de EE.UU. N.º 6.534.261; 6.607.882; 6.824.978; 6.933.113; 6.979.539; 7.013.219; y 7.163.824. Además, será evidente que cualquiera de estos vectores puede comprender una o más secuencias codificantes de ZFN. Así, cuando uno o más pares de ZFN se introducen en la célula, las ZFN pueden ser llevadas en el mismo vector o en vectores diferentes. Cuando se usan múltiples vectores, cada vector puede comprender una secuencia que codifica una ZFN o múltiples.

40 Pueden usarse métodos de transferencia génica basados en virus y no virales convencionales para introducir ácidos nucleicos que codifican ZFP manipuladas en células de rata. Tales métodos también pueden usarse para administrar ácidos nucleicos que codifican ZFP a células de rata *in vitro*. En ciertas realizaciones, los ácidos nucleicos que codifican ZFP se administran para usos *in vivo* o *ex vivo*.

45 Sistemas de administración de vectores no virales incluyen electroporación, lipofección, microinyección, biolística, virosomas, liposomas, inmunoliposomas, conjugados de polimerización o de lípido:ácido nucleico, ADN desnudo, viriones artificiales y captación de ADN potenciada por agente. También puede usarse sonoporación usando, por ejemplo, el sistema Sonitron 2000 (Rich-Mar) para la administración de ácidos nucleicos. Sistemas de administración de vectores virales incluyen virus de ADN y ARN, que tienen tanto genomas episómicos como integrados después de la administración a la célula. Sistemas de administración de ácidos nucleicos a modo de ejemplo adicionales incluyen aquellos proporcionados por Amaxa Biosystems (Colonia, Alemania), Maxcyte, Inc. (Rockville, Maryland), sistemas de administración molecular BTX (Holliston, MA) y Copernicus Therapeutics Inc. (véase, por ejemplo, el documento US6008336). La lipofección se describe en, por ejemplo, los documentos US 5.049.386, US 4.946.787; y US 4.897.355) y los reactivos de lipofección se venden comercialmente (por ejemplo, Transfectam™ y Lipofectin™). Lípidos catiónicos y neutros que son adecuados para la eficiente lipofección por reconocimiento de receptores de polinucleótidos incluyen aquellos de Felgner, documentos WO 91/17424, WO 91/16024. La administración puede ser

a células (administración *ex vivo*) o tejidos diana (administración *in vivo*). La preparación de complejos de lípido:ácido nucleico, que incluyen liposomas dirigidos tales como complejos de inmunolípido, es muy conocida para un experto en la materia (véase, por ejemplo, Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291-297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); patentes de EE.UU. N.º 4.186.183, 4.217.344, 4.235.871, 4.261.975, 4.485.054, 4.501.728, 4.774.085, 4.837.028 y 4.946.787).

Como se observa anteriormente, los métodos y composiciones descritos pueden usarse en cualquier tipo de célula de rata. También puede usarse progenie, variantes y derivados de células de rata.

Aplicaciones

Los métodos y composiciones descritos pueden usarse para la edición genómica de cualquier gen o genes de rata. En ciertas aplicaciones, los métodos y composiciones pueden usarse para la inactivación de secuencias genómicas de rata. En otras aplicaciones, los métodos y composiciones permiten la generación de mutaciones aleatorias, que incluyen generación de novedosas formas alélicas de genes con expresión diferente en comparación con genes no editados o integración de genes de rata humanizados, que a su vez permite la generación de modelos animales. En otras aplicaciones, los métodos y composiciones pueden usarse para crear mutaciones aleatorias en posiciones definidas de genes que permiten la identificación o selección de animales que llevan novedosas formas alélicas de aquellos genes. En otras aplicaciones, los métodos y composiciones permiten la integración dirigida de una secuencia exógena (donante) en cualquier área seleccionada del genoma de rata. Podrían integrarse secuencias reguladoras (por ejemplo, promotores) en un modo dirigido en un sitio de interés. Por "integración" se indica tanto inserción física (por ejemplo, en el genoma de una célula huésped) como, además, integración copiando la secuencia donante en el genoma de célula huésped mediante los procesos de replicación de ácido nucleico. Las secuencias donantes también pueden comprender ácidos nucleicos tales como ARNhp, miARN, etc. Estos donantes de ácidos nucleicos pequeños pueden usarse para estudiar sus efectos sobre genes de interés dentro del genoma de rata. La edición genómica (por ejemplo, inactivación, integración y/o mutación dirigida o aleatoria) de un gen de rata puede lograrse, por ejemplo, por un único evento de escisión, por escisión seguida de unión de extremos no homólogos, por escisión seguida de reparación dirigida por mecanismos de homología, por escisión seguida de integración física de una secuencia donante, por escisión en dos sitios seguida de unión de manera que se deleccione la secuencia entre los dos sitios de escisión, por recombinación dirigida de un codón de sentido alterado o finalizador en la región codificante, por recombinación dirigida de una secuencia irrelevante (es decir, una secuencia de "relleno") en el gen o su región reguladora, de manera que se rompa el gen o la región reguladora, o dirigiendo la recombinación de una secuencia aceptora de corte y empalme en un intrón para producir el corte y empalme erróneo del transcrito. Véanse las publicaciones de patente de EE.UU. N.º 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; y la publicación internacional WO 07/014275,

Hay una variedad de aplicaciones para la edición genómica mediada por ZFN de rata. Los métodos y composiciones descritos en el presente documento permiten la generación de modelos de rata de enfermedades humanas. Por ejemplo, la edición del gen p53 permite la generación de una "rata con cáncer" que proporciona un modelo animal para estudiar el cáncer y probar terapias para el cáncer.

EJEMPLOS

Ejemplo 1: Las ZFN inducen rotura dirigida en células C6 de rata (ejemplo comparativo)

Se diseñaron ZFN dirigidas a p53 de rata y se incorporaron en plásmidos esencialmente como se describe en Urnov et al. (2005) Nature 435(7042):646-651. Las hélices de reconocimiento para los diseños de p53 de rata representativos se muestran a continuación en la Tabla 1. Los sitios diana para estas ZFN se muestran en la Tabla 2.

Tabla 1: Diseños de ZFN específicas de p53 de rata

Nombre de ZFN	F1	F2	F3	F4
10356	RSDDLTR (SEQ ID NO:16)	RSDHLSR (SEQ ID NO:44)	DNPNLNR (SEQ ID NO:55)	RSDDLTR (SEQ ID NO:100)
10358	DNPNLNR (SEQ ID NO:55)	RSDDLTR (SEQ ID NO:100)	NSQHLTE (SEQ ID NO:101)	QSSHLSR (SEQ ID NO:102)
10359	QSGNLAR (SEQ ID NO:21)	RSDDLTR (SEQ ID NO:16)	NSQHLTE (SEQ ID NO:101)	QSSHLSR (SEQ ID NO:102)
10357	RSDDLTR (SEQ ID NO:16)	RSDHLSR (SEQ ID NO:44)	QSGNLAR (SEQ ID NO:21)	RSDDLTR (SEQ ID NO:16)

Nombre de ZFN	F1	F2	F3	F4
10360	RSDNLAR (SEQ ID NO:103)	RSDHLTT (SEQ ID NO:104)	RSDNLSQ (SEQ ID NO:105)	ASNDRKK (SEQ ID NO:106)
10362	RSDHLSE (SEQ ID NO:87)	RSAALAR (SEQ ID NO:107)	RSDHLSE (SEQ ID NO:87)	RNQHRIT (SEQ ID NO:108)
10361	RSDNLAR (SEQ ID NO:103)	RSDHLTT (SEQ ID NO:104)	RSDNLSE (SEQ ID NO:43)	DSRSRIN (SEQ ID NO:109)
10363	DRSHLSR (SEQ ID NO:110)	RSDDLTR (SEQ ID NO:16)	RSDHLSR (SEQ ID NO:44)	DRSHLAR (SEQ ID NO:12)

Tabla 2: Dianas de ZFN específicas de p53 de rata

Nombre de ZFN	Sitio diana (5' a 3')
10356	aaGCGGAAGGGGCGggccatagcccggg (SEQ ID NO:111)
10358	caGGACGTGCGGAAtgcgftaagggaaat (SEQ ID NO:112)
10359	caGGACGTGCGGAAtgcgftaagggaaat (SEQ ID NO:112)
10357	aaGCGGAAGGGGCGggccatagcccggg (SEQ ID NO:111)
10360	ctTCCCAGTGGGAGgtgacagaaccctg (SEQ ID NO:113)
10362	acCGGCGGGTGGGgcgactgcactta (SEQ ID NO:114)
10361	ctTCCCAGTGGGAGgtgacagaaccctg (SEQ ID NO:113)
10363	ccGGCGGGtGCGGGCggactgcacttag (SEQ ID NO:115)

5 Se transfectaron plásmidos que codificaban ZFN en células C6 de rata. Para determinar la actividad de ZFN en el locus p53, se realizaron ensayos de desapareamiento de CEL-I esencialmente según las instrucciones del fabricante (Trangenomic SURVEYOR™). Se recogieron las células y se preparó ADN cromosómico usando un kit Quickextract™ según indicaciones del fabricante (Epicentre®). Se amplificó por PCR la región apropiada del locus p53 usando ADN polimerasa Accuprime™ High-fidelity (Invitrogen). Se calentaron reacciones de PCR a 94 °C, y se enfriaron gradualmente hasta temperatura ambiente. Se mezclaron aproximadamente 200 ng del ADN hibridado con 10 0,33 µl de enzima CEL-I y se incubaron durante 20 minutos a 42 °C. Los productos de reacción se analizaron por electroforesis en gel de poliacrilamida en 1X tampón Tris-borato-EDTA.

Los resultados se muestran en la Figura 1 donde se probaron en combinación diversos pares de ZFN específicas de p53 descritas en las Tablas 1 y 2. El porcentaje de desapareamientos, una medida de la actividad de NHEJ, se muestra debajo de cada carril. Los resultados indican que estas ZFN son activas contra este locus de rata.

15 Se diseñaron ZFN dirigidas a GFP y se incorporaron en plásmidos esencialmente como se describe en Urnov et al. (2005) *Nature* 435(7042):646-651. Se cribaron pares de ZFN para la actividad en un sistema cromosómico basado en levadura como se describe en U.S. N.º de serie 12/284.887, titulada "Rapid in vivo Identification of Biologically Active Nucleases". Brevemente, se transformaron ZFN inducibles por galactosa en una cepa de levadura que contiene un indicador de hibridación monocatenaria (ySSA) integrado, que consistió en la secuencia de eGFP completa insertada entre dos segmentos de solapamiento del gen *MEL1* conducido por el promotor PGK. La expresión de las ZFN se indujo durante 6 horas, luego se reprimió durante 18 horas, tiempo después del cual se usó un ensayo colorimétrico estándar para cuantificar la cantidad de proteína MEL1 en el sobrenadante.

20 Las hélices de reconocimiento para diseños de dedos de cinc de GFP representativos se muestran a continuación en la Tabla 3.

25

Tabla 3: Diseños de dedos de cinc de GFP

Nombre de ZFN	F1	F2	F3	F4	F5	F6
16833 "33"	RSAHLR (SEQ ID NO:5)	TSANLSR (SEQ ID NO:6)	RSDNLSV (SEQ ID NO:7)	DRSNLTR (SEQ ID NO:8)		
16834 "34"	RSDTLSQ (SEQ ID NO:9)	QRDBRIK (SEQ ID NO:10)	DRSNLSR (SEQ ID NO:11)	DRSHLAR (SEQ ID NO:12)	DRSNLTR (SEQ ID NO:8)	
16855 "55"	RSDHLSA (SEQ ID NO:13)	DSSTRKT (SEQ ID NO:14)	TSGLSR (SEQ ID NO:15)	RDDLTR (SEQ ID NO:16)	TSANLSR (SEQ ID NO:6)	
16856 "56"	RSDNLST (SEQ ID NO:17)	DSSRIK (SEQ ID NO:18)	RAVLSE (SEQ ID NO:19)	TNSRIT (SEQ ID NO:20)	RSAHLR (SEQ ID NO:5)	QSGNLR (SEQ ID NO:21)
16859 "59"	TSGLSR (SEQ ID NO:15)	QSGLTR (SEQ ID NO:22)	TSGLSR (SEQ ID NO:15)	QSSDLRR (SEQ ID NO:23)	RSDALSR (SEQ ID NO:24)	TSGLTR (SEQ ID NO:25)
16860 "60"	RSANLSV (SEQ ID NO:30)	DRANLSR (SEQ ID NO:29)	DRSDLR (SEQ ID NO:28)	RSDLSV (SEQ ID NO:27)	DSSARKK (SEQ ID NO:26)	

Se muestran sitios diana de los diseños de dedos de cinc de GFP a continuación en la Tabla 4. Los nucleótidos en el sitio diana que se ponen en contacto por las hélices de reconocimiento de ZFN se indican en letras mayúsculas; nucleótidos que no se ponen en contacto se indican en minúscula.

Tabla 4: Sitios diana de dedos de cinc de GFP

Nombre de ZFN	Sitio diana (5' a 3')
16833	GACCAGGATGGG (SEQ ID NO:31)
16834	GACGGCGACgTAAACG (SEQ ID NO:32)
16855	GATGCGGTTcACCAGG (SEQ ID NO:33)
16856	GAAGGGCATCGAcTTCAAG (SEQ ID NO:34)
16859	GTTGTGGCTGTTGTAGTT (SEQ ID NO:35)
16860	ATCATGGCCGACAAG (SEQ ID NO:36)

Se transfectaron construcciones de expresión activas de ZFN dirigidas a GFP en células C6 de rata que contenían una construcción de expresión de GFP.

10 Como se muestra en la FIG. 2, todos los pares de ZFN probados escindieron el gen GFP en las células diana.

Ejemplo 2: Las ZFN inducen la rotura dirigida en ratas transgénicas (ejemplo comparativo)

ZFN específicas de GFP como se describe en el Ejemplo 1 también se introdujeron por inyección pronuclear (PNI) o inyección citoplásmica (de ARNm de ZFN) a concentraciones variables en embriones unicelulares obtenidos de ratas transgénicas que expresan GFP descritos en Michalkiewicz et al. (2007) J. Amer. Phys. Society 293:H881-H894. Véase la Fig. 3.

Los embriones inyectados se cultivaron durante 2-3 días hasta que alcanzaron el estadio de 2-4 células. Algunos de los embriones de 2-4 células se transfirieron entonces a hembras pseudo-preñadas. Se extrajo ADN de tanto embriones cultivados como embriones transferidos y se evaluó la escisión del gen GFP.

Los resultados del diferente modo de inyección y la concentración de ZFN inyectada en los embriones inyectados usando el par de ZFN 16859/16860 se muestran en la Tabla 5 a continuación.

TABLA 5

Método de inyección	Conc. de ZFN (ng/μl)	Embriones inyectados	Sobrevivieron	Dividieron en 2 células	%	Transferidos	Nacidos
PNI	1	40	23	19	83		
	2	39	29	19	66		
	1,5	36	25			25	5
citoplásmica	5	40	26	16	62		
	10	39	32	16	50		
	20	38	31	24	77		
	10	256	138			138	

5 La obtención de imágenes de GFP de inyecciones citoplásmicas de ARNm de ZFN mostró que muchos más embriones que contenían ZFN dejaron de expresar GFP que los embriones no inyectados, que indica que no estaba presente mosaicismo en las células en las que las ZFN eran activas.

Nacieron cinco crías de inyección pronuclear (PNI) de ZFN en embriones que se transfirieron en hembras pseudo-preñadas. Véase la Tabla 5. Como se muestra en la FIG. 4A, 3 de las cinco crías expresaron GFP mientras que 2 crías no.

10 Se preparó ADN genómico a partir de las colas de los dos animales GFP-negativos y se cribaron para la modificación mediante PCR. Cuando se comparó con el locus de eGFP no mutante, las regiones fronterizas con el sitio dirigido por el par de ZFN 59/60 se redujeron significativamente, sugiriendo deleciones de aproximadamente 150 pb para ambos animales GFP-negativos. Otra vez, no fue evidente mosaicismo en la biopsia de cola como se indica por la ausencia de una banda de eGFP no mutante. Estas deleciones se analizaron entonces directamente por secuenciación, que reveló deleciones de 162 nt y 156 nt, produciendo las bandas más pequeñas evidentes en la FIG. 4B. Además, como se muestra en la FIG. 4B, no fue evidente mosaicismo en crías GFP-negativas, ya que no se detectó banda de eGFP no mutante.

Así, las ZFN modificaron satisfactoriamente el transgén de GFP cromosómico.

Ejemplo 3: Las ZFN escinden loci de rata endógenos

Se diseñaron ZFN para escindir loci endógenos como se describe a continuación.

20 A. *IgM*

En un experimento, se diseñaron ZFN que escindían el gen *IgM* de rata endógeno, como se ha descrito anteriormente, y se probaron para la actividad de escisión en células C6 de rata. ZFN dirigidas a *IgM* de rata a modo de ejemplo se muestran a continuación en la Tabla 6 a continuación.

Tabla 6: Diseños de dedos de cinc de *IgM*

Nombre de ZFN	F1	F2	F3	F4	F5	F6
17747	DRSHLTR (SEQ ID NO:41)	RSDALTQ (SEQ ID NO:40)	DRSDLSR (SEQ ID NO:28)	RSDALAR (SEQ ID NO:39)	RSDLSLA (SEQ ID NO:38)	TSSNRKT (SEQ ID NO:37)
17749	NKVGLIE (SEQ ID NO:46)	TSSDLSR (SEQ ID NO:45)	RSDHLSR (SEQ ID NO:44)	RSDNLSE (SEQ ID NO:43)	QNAHRKT (SEQ ID NO:42)	
17759	DRSALSR (SEQ ID NO:51)	TSGHLSR (SEQ ID NO:52)	RSDNLST (SEQ ID NO:53)	HNATRIN (SEQ ID NO:54)	DRSALSR (SEQ ID NO:51)	QSGNLAR (SEQ ID NO:21)

Nombre de ZFN	F1	F2	F3	F4	F5	F6
17756	RSANLAR (SEQ ID NO:56)	RSDNLRE (SEQ ID NO:57)	TSGLSLR (SEQ ID NO:58)	QSGSLTR (SEQ ID NO:59)	RSDVLSE (SEQ ID NO:60)	TSGSLTR (SEQ ID NO:25)
17767	QSSDLR (SEQ ID NO:61)	RSDALAR (SEQ ID NO:39)	TSGHLR (SEQ ID NO:52)	RSDALR (SEQ ID NO:39)	DRSDLR (SEQ ID NO:28)	
17764	RSDALAR (SEQ ID NO:39)	RSDHLST (SEQ ID NO:62)	HSNARKN (SEQ ID NO:63)	DRSDLR (SEQ ID NO:28)	TSGHLR (SEQ ID NO:52)	
17782	RSANLSV (SEQ ID NO:30)	DRANLSR (SEQ ID NO:29)	RSDALAR (SEQ ID NO:39)	DRSDLR (SEQ ID NO:28)	RSDDLTR (SEQ ID NO:16)	
17778	RSAHLR (SEQ ID NO:5)	QSGDLTR (SEQ ID NO:64)	RSDALAR (SEQ ID NO:39)	RSDTLR (SEQ ID NO:65)	DNSTRIK (SEQ ID NO:66)	

Se muestran sitios diana de los diseños de dedos de cinc dirigidos a *IgM* de rata a continuación en la Tabla 7. Los nucleótidos en el sitio diana que se ponen en contacto por las hélices de reconocimiento de ZFP se indican en letras mayúsculas; nucleótidos que no se ponen en contacto se indican en minúscula.

5

Tabla 7: Sitios diana de dedos de cinc de *IgM*

Nombre de ZFN	Sitio diana (5' a 3')
17747	AATTTGGTGGCCATGGGC (SEQ ID NO:47)
17749	AGACAGGGGGCTCTC (SEQ ID NO:48)
17759	ctGAAGTCATGCAGGGTGTcagaacctt (SEQ ID NO:67)
17756	ttGTTCTGGTAGTTcCAGGAGaaggaaa (SEQ ID NO:68)
17767	gtGCTGTGGGTGTGGCTagtgtttgat (SEQ ID NO:69)
17764	aaGGTGCCATTGGGGTGactttccatga (SEQ ID NO:70)
17782	gaGAGGACcGTGGACAAGtccactgga (SEQ ID NO:71)
17778	tcACCATGtGTGGCAGGGcctcgtggcc (SEQ ID NO:72)

10

Todas las ZFN dirigidas a *IgM* contuvieron las mutaciones EL/KK *Fok I* como se describe en la publicación de patente de EE.UU. N.º 2008/0131962. La expresión de ZFN fue conducida por tanto CAG como el promotor del CMV. Se transfectoron ZFN (1 µg de cada una) en 200.000 células C6 mediante nucleofección de Amaxa usando la disolución SF y el nucleofector de 96 pocillos Amaxa Shuttle. Se amplificó por PCR el locus de *IgM* usando GJC153F (5'-ggaggcaagaagatggattc-3') y GJC154R (5'-gaatcgccatgcagatct-3') y se ensayó la escisión de ZFN con la nucleasa Surveyor™ como se describe, por ejemplo, en las publicaciones de patente de EE.UU. N.º 20080015164; 20080131962 y 20080159996.

15

En células C6, el par de ZFN 17747/17749 escindió el 3 % de los cromosomas cuando se usó el promotor del CMV y aproximadamente el 1 % cuando se usó el promotor CAG (Figura 4). Este par de ZFN escindió el gen *IgM* de rata en la región codificante del exón 1. Se inyectaron ovocitos de rata con 10 ng/ul de un plásmido que codificaba el par de ZFN 17747/17749 bajo el control del promotor CAG usando técnicas convencionales. Se fecundaron los ovocitos y se implantaron en hembras pseudo-preñadas. De los 430 ovocitos inyectados e implantados, resultaron 43 partos vivos. Se preparó ADN genómico a partir de las colas de estos 43 animales y se cribaron para modificación usando la nucleasa Surveyor™.

20

Como se muestra en la Figura 7, cinco de los 43 animales (ratas N.º 6, 7, 8, 19 y 46) dieron positivo para modificación en el locus *IgM*. Los patrones de digestión con nucleasa Surveyor™ fueron idénticos tanto con como

sin la adición de ADN genómico de rata no mutante, sugiriendo que ninguna de las ratas tiene una mutación homocigótica.

Se clonaron productos de PCR GJC153F/GJC154R de las ratas positivas y se secuenciaron. Una descripción de los alelos mutados está en la Tabla 8.

5

Tabla 8

Rata	Alelo	Cifra	% de NHEJ aprox.	Notas
6	No mutante	8	49	
6	$\Delta 9$	2		delección en marco de DEN
7	No mutante	5	31	
7	$\Delta 5$	1		fuera de marco
7	$\Delta 13$	1		fuera de marco
7	$\Delta 15$	3		delección en marco de SDENL
7	$\Delta 18$	1		delección en marco de DENLA
7	$\Delta 39$	1		del. en marco de SCESPLSDENLVA
8	No mutante	7	25	
8	$\Delta 3$, mut. de 7b pb	3		delección en marco de D, E->P
8	$\Delta 23$	2		fuera de marco
19	No mutante	7	70	
19	$\Delta 64$	17		delección más grande, fuera de marco
46	No mutante	9	47	
46	$\Delta 5$	2		fuera de marco
Cifra se refiere al número de veces que se aisló una secuencia particular.				
% de NHEJ es el porcentaje aproximado de cromosomas modificados en el ADN de la cola				

La secuenciación del locus de *IgM* en estas ratas confirmó los resultados del ensayo de nucleasas Surveyor™. Todas las deleciones solapan los sitios de unión de ZFN. El espectro de deleciones pequeñas observado aquí es típico de mutación mediada por NHEJ. Las ratas 7 y 8 tienen más de un alelo mutado y son, por tanto, mosaicos para la mutación de *IgM*. Aunque la secuenciación de ratas 6, 19 y 46 solo dio un alelo mutado, puede ser mosaico para la modificación de *IgM* en otros tejidos.

10

Así, las ZFN modificaron satisfactoriamente el locus de *IgM* de rata endógeno.

15

Para determinar si el propio plásmido de ZFN se integró en el genoma de rata, se desarrolló un ensayo basado en PCR para probar la integración de plásmidos de ZFN. Brevemente, se mezclaron ADN genómico de rata y el plásmido de ZFN para imitar una frecuencia de inserción de plásmido de una vez por genoma. El realizar 35 ciclos de amplificación por PCR de esta mezcla con un oligonucleótido en el promotor CAG (5'-GCT AAC CAT GTT CAT GCC TTC-3') (SEQ ID NO:49) y otro oligonucleótido en la región 2A del plásmido (5'-CAT CCT AGG GCC GGG ATT CTC-3') (SEQ ID NO:50) dio una banda de 1338 pb (Figura 7, carril 3). Cuando se analizó ADN genómico de ratas no mutantes y las cinco modificadas con ZFN, no fue detectable producto de PCR; que indica que la inserción del plásmido en el genoma de rata no es un evento de alta frecuencia.

20

Además, la rata N.º 19 modificada en *IgM* se analizó adicionalmente por ensayo de CEL-I y secuenciación. Como se muestra en las Figuras 7A y 7B, las ZFN de *IgM* produjeron una delección de 64 pares de bases en esta rata en el locus *IgM*.

25

Finalmente, también se evaluó la escisión de ZFN en sitios inespecíficos. Se usó un algoritmo informático para predecir la localización de los sitios inespecíficos más probables (Doyon et al (2008) Nature Biotechnology 26(6):702-708). Se ensayaron todos los sitios inespecíficos probables para la modificación de ZFN usando el ensayo

de nucleasas Surveyor™ como se ha descrito anteriormente. Los resultados de este análisis se muestran en la Tabla 9 y las Figs. 9A-C.

Tabla 9

Sitio	Puntuación	Secuencia	Mm	Gen	PCR	Frag. A	Frag. B	Acierto
1	8,18E-17	AGtcAGChCCTGTCTAGAAGA GAAcTgGGTGICATGGGCC (SEQ ID NO:73)	8		320	221	99	No
2	2,90E-18	CaAatGCCaCCTGTCTGAATG GttTaTgcTGGCaATGGGCT (SEQ ID NO:74)	9		325	222	103	No
3	1,67E-18	GGIGAGaCCCCTGTCTTAACA AAAgatGGgGGggtTGGGaA (SEQ ID NO:75)	9		379	239	40	No
4	7,75E-19	GatCCAaGGCCACCAAcTgGA GTTTAAGACAaaGGGCTCTgC (SEQ ID NO:76)	8		322	218	104	No
5	6,44E-19	TGICCATGGCCICClccTcTTTG CTAGgcGGiGGCTCTCA (SEQ ID NO:77)	9	Pde4d	396	200	196	No
6	1,49E-19	GGANGCCCCCTGTcATGCAC AGcATaTGGTGGCCATaGatG (SEQ ID NO:78)	8	LOC499 913	342	200	142	No
7	1,14E-19	GGAGAagCCCaTGTgTACTCT TtAgTTGGTGGCtcTGGGaG (SEQ ID NO:79)	9		567	317	250	No
8	1,07E-19	GcCCataGGCCAaCAAcTcTCA GGCTAGACAacGGGCTCTCA (SEQ ID NO:80)	9	Actn1	354	255	99	No

Mm: desapareamientos con respecto al sitio diana previsto

Frag. A, B: tamaños esperados de los productos de escisión por nucleasa Surveyor™

Acierto: Ratas que muestran productos de escisión por nucleasa Surveyor™ correctos

5 Como se muestra, ningún sitio inespecífico probado mostró evidencia de modificación. Como se muestra, ningún sitio inespecífico probado mostró evidencia de modificación. El análisis de secuenciación de la rata N.º 19 positiva para CEL-I mostrado en la Figura 9A y cinco de sus crías mostradas en la Figura 11 en el Sitio 1 revelaron que la señal positiva para CEL-I era debida a un SNP cerca del posible sitio inespecífico. El desapareamiento se produce debido a que las ratas son heterocigóticas para este SNP que también se encontró en ratas no tratadas (datos no mostrados). Aunque está presente en el 50 % de los cromosomas en animales positivos para CEL-I, el SNP es malamente reconocido por la enzima CEL-I, produciendo productos de escisión de intensidad inesperadamente más baja.

B. *Rab38* (ejemplo comparativo)

15 También se diseñaron ZFN para dirigir el locus *Rab38* endógeno en ratas, particularmente el exón 1 del gen *Rab38* de rata. Diseños de dedos de cinc de *Rab38* a modo de ejemplo se muestran en la Tabla 8 a continuación.

Tabla 10: Diseños de dedos de cinc de *Rab38*

Nombre de ZFN	F1	F2	F3	F4	F5	F6
18160	DRSNLSS (SEQ ID NO:81)	RSHSLLR (SEQ ID NO:82)	RSDLSA (SEQ ID NO:38)	TSGSLTR (SEQ ID NO:25)	QSGNLAR (SEQ ID NO:21)	QSGHLR (SEQ ID NO:83)
18181	TSGHLR (SEQ ID NO:52)	HKWQRNK (SEQ ID NO:84)	DRSVLRR (SEQ ID NO:85)	DSSTRKK (SEQ ID NO:86)	RSDHLSE (SEQ ID NO:87)	DKSNRKK (SEQ ID NO:88)

Nombre de ZFN	F1	F2	F3	F4	F5	F6
16897	RSDLTSE (SEQ ID NO:89)	QKRNRTK (SEQ ID NO:90)	RSDLSA (SEQ ID NO:38)	TSGSLTR (SEQ ID NO:25)	QSGNLAR (SEQ ID NO:21)	QSGHLSR (SEQ ID NO:83)
16898	RSDHLSK (SEQ ID NO:91)	HNDSRTN (SEQ ID NO:92)	DRSDLSR (SEQ ID NO:28)	RSDHLSK (SEQ ID NO:87)	DKSNRKK (SEQ ID NO:88)	N/A
18173	RSDYLPR (SEQ ID NO:93)	QSNLNS (SEQ ID NO:94)	DRSDLSR (SEQ ID NO:28)	RSDHLSK (SEQ ID NO:87)	DKSNRKK (SEQ ID NO:88)	N/A
18174	RSDYLPR (SEQ ID NO:93)	QRVTRDA (SEQ ID NO:95)	DRSDLSR (SEQ ID NO:28)	RSDHLSK (SEQ ID NO:87)	DKSNRKK (SEQ ID NO:88)	N/A
18175	HSNARKT (SEQ ID NO:96)	ASKTRTN (SEQ ID NO:97)	DRSDLSR (SEQ ID NO:28)	RSDHLSK (SEQ ID NO:87)	DKSNRKK (SEQ ID NO:88)	N/A
18161	RSHLLR (SEQ ID NO:82)	RSDLSA (SEQ ID NO:38)	TSGSLTR (SEQ ID NO:25)	QSGNLAR (SEQ ID NO:21)	QSGHLSR (SEQ ID NO:83)	N/A
18183	RSHLLR (SEQ ID NO:82)	RSDYLPR (SEQ ID NO:93)	DRSVLRR (SEQ ID NO:85)	DSSTRKK (SEQ ID NO:86)	RSDHLSK (SEQ ID NO:87)	DKSNRKK (SEQ ID NO:88)

Se muestran sitios diana de los diseños de dedos de cinc dirigidos a *Rab38* de rata a continuación en la Tabla 11. Los nucleótidos en el sitio diana que se ponen en contacto por las hélices de reconocimiento de ZFP se indican en letras mayúsculas; los nucleótidos que no se ponen en contacto se indican en minúscula.

5

Tabla 11: Sitios diana de dedos de cinc de Rab38

Nombre de ZFN	Sitio diana (5' a 3')
18161	gaGGAGAAGTTTTGGTGCACgtagcgct (SEQ ID NO:98)
18181	acTACCGGGCCACCATTGGTgtggactt (SEQ ID NO:99)
16897	gaGGAGAAGTTTTGgTGCACGtagcgct (SEQ ID NO:98)
16898	acTACCGGGCCacCATTGgtgtggactt (SEQ ID NO:99)
18173	acTACCGGGCCaCCATTGgtgtggactt (SEQ ID NO:99)
18174	acTACCGGGCCaCCATTGgtgtggactt (SEQ ID NO:99)
18175	acTACCGGGCCACCATTggtgtggactt (SEQ ID NO:99)
18160	gaGGAGAAGTTTTGGTgcacgtagcgct (SEQ ID NO:98)
18183	acTACCGGGCCACCATTGGTgtggactt (SEQ ID NO:99)

Todas las ZFN dirigidas a *Rab38* contuvieron las mutaciones EL/KK *Fok I* como se describe en la publicación de patente de EE.UU. N.º 2008/0131962. La expresión de ZFN se condujo por tanto el promotor CAG como del CMV. Se transfectaron ZFN (1 µg de cada una) en 200.000 células C6 mediante nucleofección de Amaxa usando la disolución SF y el nucleofactor de 96 pocillos Amaxa Shuttle. La escisión se ensayó con la nucleasa Surveyor™ de CEL-I como se describe, por ejemplo, en las publicaciones de patente de EE.UU. N.º 20080015164; 20080131962 y 20080159996.

10

5 Se linealizaron plásmidos de expresión que codifican ZFN de *Rab38* con *XbaI*, se extrajeron con fenol-cloroformo y se precipitaron. Se transcribió *in vitro* ARN mensajero usando el kit de transcripción MessageMax™ T7 ARCA-Capped Message (Epicentre Biotechnologies). La síntesis resultante se purificó usando MegaClear Kit™ (Ambion) antes de la resuspensión en 0,1X TE (Tris-Cl 1 mM a pH 8,0, EDTA 0,1 mM) libre de RNasa, se cuantificó usando un NanoDrop-1000 (Thermo Scientific) y se guardó a -80 °C hasta uso. Los ARN mensajeros que codifican ZFN de *Rab38* se mezclaron a una concentración total final de 5 ng/μl en 0,1X TE. Se inyectaron embriones con ZFN de *Rab38* bajo tiempo y presión constantes (Pi= 65, Pc= 20, ti=1,5s) en el citoplasma y se incubaron a 37,5 °C y 5 % de CO2 en KSOM (Millipore) durante la noche como se describe previamente en Filipiak et al. (2006) Transgenic Res 15:673-686 para el análisis molecular.

10 Se usó una estrategia de enriquecimiento de mutaciones como se describe en Lloyd et al. ((2005) Proc Natl Acad Sci U S A 102:2232-2237) para detectar alteraciones del exón diana *Rab38* en cromosomas de ADN extraído de embriones cultivados durante 48 horas después de la inyección.

15 Como se muestra en la Figura 10, y como se demuestra para tanto los loci GFP como *IgM* anteriormente, podrían detectarse múltiples alelos de *Rab38* mutantes en los genomas de tan solo 16 embriones de dos células y la secuenciación reveló deleciones en el sitio diana.

Así, estos datos confirman que múltiples loci genómicos son dianas adecuadas para la edición de genoma mediada por ZFN.

Ejemplo 4: Modificaciones de la línea germinal mediadas por ZFN

20 Se aparearon las ratas N.º 19, N.º 46 y N.º 8 modificadas en *IgM* como se describe en Ejemplo 3 con una rata no mutante y se tomaron biopsias de la cola, se aisló ADN genómico y entonces se realizaron ensayos de CEL-I y PCR en el ácido nucleico purificado de las crías.

25 Como se muestra en las Figuras 11A y 11C, las crías (numeradas 225, 227, 228, 229, 230, 231, 234 y 235) resultantes de un cruce entre la rata N.º 19 y una rata no mutante llevaron la deleción de 64 pares de bases de la rata N.º 19 modificada en *IgM*, como se ha determinado por ensayo de PCR y CEL-I. Además, el análisis de secuenciación confirma que 3 crías de rata N.º 19 (crías N.º 225, 227 y 228) se modificaron en el locus *IgM*. Véase la Figura 11B. Además, como se muestra en la Figura 11C, una cría resultante de aparear la rata N.º 46 con una rata no mutante llevó la misma modificación de *IgM* que la rata parental N.º 46 (véase, cría número 236 de la Figura 11 C).

Estos datos demuestran que la ruptura mediada por ZFN de un locus de rata se transmite en la línea germinal.

30 **Ejemplo 5: Construcción de ácido nucleico donante de polimorfismo en la longitud de los fragmentos de restricción (RFLP) para la integración dirigida en la región de ácido nucleico PXR del genoma de rata (ejemplo comparativo)**

35 Hay dos posibles resultados de reparación de ADN después de una rotura bicatenaria inducida por ZFN dirigida (**FIG. 12**). La rotura puede ser reparada por la unión de extremos no homólogos (NHEJ), que conduce a mutaciones que contienen deleciones o adiciones de bases o, en presencia de un ADN donante, el ADN donante puede usarse como molde para reparar la rotura bicatenaria por recombinación homóloga (HR). Si el ADN donante codifica cambios de secuencia específicos, estas mutaciones deliberadas se incorporarán en el genoma del organismo en el sitio diana.

40 Para probar la integración dirigida en el genoma de rata usando inyección pronuclear, se diseñaron construcciones y se prepararon para la integración dirigida en la región de gen PXR del genoma de rata. Se ensamblaron construcciones para introducir tanto un sitio de polimorfismo en la longitud de los fragmentos de restricción NotI como PmeI (RFLP) en la región génica PXR (**FIG. 13**). Las construcciones se diseñaron con tanto 200, 800 como 2000 pares de bases de homología de secuencias para el sitio diana del gen PXR flanqueando los sitios RFLP a introducir. Se usaron los tres tamaños de regiones de homología para determinar el tamaño de homología requerido para el eficiente direccionamiento y recombinación homóloga.

45 Se ensamblaron los clones usando amplificación por PCR para introducir sitios de restricción convenientes para la clonación, y el sitio RFLP en las extremidades de las regiones de homología PXR (**FIG. 12**). Los cebadores de PCR usados para amplificar la región de homología PXR se describen en la Tabla 12. Se usó la ADN polimerasa Accuprime HF para la reacción de amplificación por PCR. Se usó una extensión de 30 s para los fragmentos de 200 pb, se usó una extensión de 1,5 min para los fragmentos de 800 pb y se usó una extensión de 4 min para los fragmentos de 2 Kbp. Entonces, los fragmentos de PCR se digirieron con las enzimas de restricción apropiadas y se clonaron en pBluescript usando ligación de tres vías para producir seis plásmidos enumerados en la Tabla 13.

Tabla 12: Secuencias de cebadores	
Nombre	Secuencia
PXR 200 pb F KpnI	5'- aaaaggtacctctgtgttttccgttctagtccag (SEQ ID NO: 116)
PXR 200 pb R SacII	5'- aaaaccgcggtgaagtatactggctctcttga (SEQ ID NO:117)
PXR diana F NotI	5'- gtgtagcggccgacagaaggccaatggctatcac (SEQ ID NO:118)
PXR diana F PmeI	5'- gtgtagttaaacgacaaggccaatggctatcac (SEQ ID NO:119)
PXR diana R NotI	5'- ttgtcgcggccgctacacggcagafflgaagacctc (SEQ ID NO:120)
PXR diana R PmeI	5'- ttgtcgtttaaactacacggcagatttgaagacctc (SEQ ID NO:121)
PXR 800 pb F KpnI	5'- aaaaggtacctcagactggtccagatttagamaagggg (SEQ ID NO:122)
PXR 800 pb R SacII	5'- aaaaccgcgataaatctactggttcgccaagctag (SEQ ID NO:123)
PXR 2Kb F KpnI	5'- aaaaggtaccgagtagtaggaaatgcactc (SEQ ID NO:124)
PXR 2Kb R SacII	5'- aaaaccgcggaagagaattattgctgacagtc (SEQ ID NO:125)
PXR 50 pb F	5'- gagcctatcaacgtagatgagg (SEQ ID NO:126)
PXR 50 pb R	5'- cttacatccttcacaggatcatgac (SEQ ID NO:127)

Tabla 13: Plásmidos construidos	
RFLP introducido	Longitud de la región de homología
NotI	200 pb
NotI	800
NotI	2 Kpb
PmeI	200 pb
PmeI	800 pb
PmeI	2K pb

5 **Ejemplo 6: Construcción de ácido nucleico donante de polimorfismo en la longitud de los fragmentos de restricción (RFLP) para la integración dirigida en la región de ácido nucleico rRosa26 del genoma de rata.** (ejemplo comparativo)

También se construyeron plásmidos para dirigir la integración de sitios de RFLP NotI y PmeI en la región de ácido nucleico rRosa26 del genoma de rata. El diseño y la construcción de los plásmidos fueron como se describe en el **Ejemplo 5** anterior. Los pares de cebadores de PCR usados para amplificar la región rRosa26 de homología se describen en la **Tabla 14**.

Tabla 14: Secuencias de cebadores	
Nombre	Secuencia
rRosa26 200 pb F KpnI	aaaaggtaccgggagtgatgaaggagttg (SEQ ID NO:128)
rRosa26 200 pb R SacII	aaaaccgcgcgatcacaagcaataat (SEQ ID NO:129)
rRose26 diana F NotI	cttcgcccgcgatctgcaactggagtcttc (SEQ ID NO:130)

rRosa26 diana F Pmel	cttcgtttaaacgatctgcaactggagctcttc (SEQ ID NO:131)
rRosa26 diana F NotI	gatcgcggccgcgaagaagggggaaggaatc (SEQ ID NO:132)
rRosa26 diana R Pmel	gatcgtttaaacgaagaagggggaaggaatc (SEQ ID NO:133)
rRosa26 800 pb F KpnI	aaaaggtaccgcgtgtgaaaacacaaatgg (SEQ ID NO:134)
rRosa26 800 pb R SacI	aaaaccgcggaaggaagaggcattcatgg (SEQ ID NO:135)
rRosa26 2Kb F KpnI	aaaaggtaccattatggaggggaggactgg (SEQ ID NO:136)
rRosa26 2Kb R SacI	aaaaccgcggacatgtggcaaacaggaga (SEQ ID NO:137)
rRosa26 50 pb F	tgtctctgaggaccgcc (SEQ ID NO:138)
rRosa26 50 pb R	ctgccagaagactccgc (SEQ ID NO:139)

Ejemplo 7: Construcción de ácido nucleico donante de polimorfismo en la longitud de los fragmentos de restricción (RFLP) para la integración dirigida en la región de ácido nucleico Mdr1a del genoma de ratón o de rata (ejemplo comparativo)

- 5 Se construyeron plásmidos para dirigir la integración de sitios de RFLP NotI y Pmel en la región de ácidos nucleicos mMdr1a del genoma ratón o la región de ácidos nucleicos de rMdr1a del genoma de rata. El diseño y la construcción de los plásmidos fueron como se describe en el **Ejemplo 5** anterior. Los pares de cebadores de PCR usados para amplificar la región Mdr1a de homología se describen en las Tablas **15 y 16**. "m" representa ratón y "r" representa rata.

Nombre	Secuencia
mMdr1a 200 pb F KpnI	aaaaggraccaacaacactaggctcaggag (SEQ ID NO:140)
mMdr1a 200 pb R SacI	aaaaccgcggcacatggctaagcacagcatg (SEQ ID NO:141)
mMdr1a diana F NotI	cctgcggccgcggactgtcagctgtattg (SEQ ID NO:142)
mMdr1a diana F Pmel	cctgtttaaacggactgtcagctgtattg (SEQ ID NO:143)
mMdr1a diana R NotI	gtccgcggccgcagggtgatggccaaaatc (SEQ ID NO:144)
mMdr1a diana R Pmel	gtccgtttaaacagggtgatggccaaaatc (SEQ ID NO:145)
mMdr1a 800 pb F KpnI	aaaaggtaccatgctgtgaagcagatacc (SEQ ID NO:146)
mMdr1a 800 pb R SacI	aaaaccgcggctgaaaactgaatgagacattgc (SEQ ID NO:147)
mMdr1a 2KB F KpnI	aaaaggtaccgtaattgtccaattgcatctcc (SEQ ID NO:148)
mMdr1a 2KB R SacI	aaaaccgcggctctcagttctctgctgtg (SEQ ID NO:149)
mMdr1a 50 pb F	gattaccctgtgctggaag (SEQ ID NO:150)
mMdr1a 50 pb R	ctggactcatggactcacc (SEQ ID NO:151)

10

Nombre	Secuencia
rMdr1a 200 pb F KpnI	aaaaggtacctggctcaggagaaaattgtg (SEQ ID NO:152)
rMdr1a 200 pb R SacI	aaaaccgcggcagcgttaaagacagcatga (SEQ ID NO:153)

Tabla 15	
rMdr1a diana F NotI	ccctgcggccgcggactgtcagctggtatttg (SEQ ID NO:154)
rMdr1a diana F Pmel	ccctgtttaaacggactgtcagctggtatttg (SEQ ID NO:155)
rMdr1a diana R NotI	gtccgcggccgcagggtgatggccaaaatc (SEQ ID NO:156)
rMdr1a diana R Pmel	gtccgtttaaacagggtgatggccaaaatc (SEQ ID NO:157)
rMdr1a 800 pb F KpnI	aaaaggtagcggagataggctggttgacg (SEQ ID NO:158)
rMdr1a 700 pb R SacII	aaaaccgcggatggtgtagttcggatgg (SEQ ID NO: 159)
rMdr1a 2Kb F KpnI	aaaaaggtagcggatggttcttgagatgtgc (SEQ ID NO:160)
rMdr1a 2Kb T SacII	aaaaccgcggctccttggtggtgagttt (SEQ ID NO:161)
rMdr1a 50 pb F	gatttactcggcgtggaag (SEQ ID NO:162)
rMdr1a 50 pb R	ctggactcaggcctcac (SEQ ID NO:163)

Ejemplo 8: Construcción de casete de integración de expresión de GFP (ejemplo comparativo)

Para probar la integración dirigida de fragmentos de ácido nucleico mayores que RFLP, se diseñaron construcciones y se prepararon para la integración dirigida de un casete de expresión de GFP en las regiones genómicas de ácido nucleico PXR y rRosa26 de la rata y las regiones genómicas de ácido nucleico mMdr1a del ratón. Brevemente, se amplificó un casete de expresión de GFP que contenía el promotor PGK humano, el marco de lectura abierto de GFP y una señal de poliadenilación usando PCR para introducir sitios de restricción NotI en las extremidades (**FIG. 14**) usando los siguientes cebadores: PGKGFP-F NotI (5'-aaagcggccgctggggtgcgcctttcc) (SEQ ID NO:164) y PGKGFP-R NotI (5'-aaaagcggccgcatagaccaccgcac) (SEQ ID NO:165). El fragmento de PCR se clonó entonces en los plásmidos que contienen NotI construidos en los Ejemplos 5-7.

Ejemplo 9: Preparación de ARNm de dedos de cinc para la integración dirigida

Se diseñaron un par de nucleasas con dedos de cinc para cada sitio de integración dirigida y se clonaron como se describe en el sitio web de Sigma. Para más información, véase Science (2009) 325:433, incorporado en el presente documento por referencia. Entonces se produjeron *in vitro* ARNm que expresaban ZFN digiriendo primero 20 µg de cada ADN de plásmido de expresión de ZFN maxi-preparado en reacciones de 100 µl que contenían 10 µl de tampón 2 (NEB, N.º B7002S), 10 µl 10x BSA (diluido de 100x BSA, NEB, N.º B9001S), 8 µl de XbaI (NEB, N.º R0145S), a 37 °C durante 2 h. Las reacciones se extrajeron con 100 µl de fenol/cloroformo (Sigma, P2069), se centrifugaron a más de 20.000 x g durante 10 min. El sobrenadante acuoso se precipitó con 10 µl de NaOAc 3 M (Sigma, S7899) y 250 µl de etanol al 100 % y se centrifugó a velocidad máxima durante 25 min a temperatura ambiente. El sedimento resultante se lavó añadiendo 300 µl de etanol al 70 % filtrado a través de un filtro de 0,02 µM. El sedimento se secó al aire y se resuspendió en 20 µl de 0,1xTE filtrado en 0,02 µM.

Entonces se usó el ADN digerido purificado para producir transcritos de ZFN usando transcripción *in vitro* con el kit de transcripción MessageMax T7 Capped Message (N.º MMA60710) de Epicentre Biotechnologies como se describe. En resumen, se precalientan los componentes del kit hasta temperatura ambiente, y se combinaron los componentes de reacción para una reacción de 20 µl a temperatura ambiente en el siguiente orden: 5 µl de agua libre de RNasa filtrada en 0,02 µm, 1 µl de molde preparado, 2 µl de lox tampón de transcripción, 8 µl de premezcla 2-Way Cap/NTP, 2 µl de DTT 100 mM y 2 µl de disolución de enzima MessageMax T7. Las reacciones se incubaron entonces en una estufa de incubación a 37 °C durante 30 min.

Se añadieron colas al ARN terminado con poliA usando el kit de adición de colas A-Plus Poly (A) Polymerase (Epicentre, N.º PAP5 104H) como se describe. Los componentes de reacción se combinaron a temperatura ambiente en el siguiente orden dado: 55,5 µl de agua libre de RNasa filtrada en 0,02 µm, 10 µl de 10x tampón de reacción A-Plus, 10 µl de ATP 10 mM, 2,5 µl de inhibidor de RNasa ScriptGuard (40 unidad/µl), 20 µl de reacción de terminación de la transcripción *in vitro*, 2 µl de poli A polimerasa A-Plus. La reacción se incubó entonces a 37 °C durante 30 min. El ARNm con colas de poli A terminado resultante se purificó mediante precipitación con un volumen igual de NH₄Oac 5 M dos veces. Entonces se secó al aire el sedimento de ARNm y se resuspendió en 30 µl de tampón de inyección filtrado (Tris 1 mM, pH 7,4, EDTA 0,25 mM), y se midió la concentración de ARN usando un espectrofotómetro Nanodrop.

Ejemplo 10: Integración dirigida en embriones

Para integrar los ácidos nucleicos en el genoma de rata o ratón, se mezclaron ARNm de nucleasa de dedos de cinc con el ADN diana maxi-preparado filtrado con filtros de 0,02 μm . La mezcla de ácido nucleico consistió en una parte ARNm de ZFN con respecto a una parte de ADN donante. Entonces, la mezcla de ácido nucleico se microinyectó en el pronúcleo de un embrión unicelular usando métodos conocidos. Los embriones inyectados tanto se incubaron *in vitro* como se transfirieron a pseudo-madres. Se recogieron los embriones/feto resultantes, o el dedo del pie/cola cortada de animales nacidos vivos para la extracción de ADN y análisis.

Para extraer ADN, se lisó tejido en 100 μl de QuickExtract de Epicentre a 50 °C durante 30 min, seguido de incubación a 65 °C durante 10 min, y 98 °C durante 3 min. Para determinar si se produjo la integración dirigida, se usó PCR para amplificar la región diana usando cebadores apropiados. Para experimentos donde se integró RFLP en el genoma del animal, los productos de PCR se digirieron con la enzima de RFLP introducida para detectar la integración (**FIG. 15A**). Además, se usó un ensayo de endonucleasa Cel-I usando fragmentos de PCR no mutantes y fragmentos de PCR derivados de embriones inyectados para demostrar que el ARNm de ZFN era funcional en los embriones detectando NHEJ, que es independiente de la integración dirigida. Para experimentos donde GFP se integró en el genoma del animal, un desplazamiento en el tamaño del fragmento de PCR es indicativo de la integración (**FIG. 15B**). Alternativamente, se usó la amplificación de la unión de integración, donde aterriza un cebador solo en el casete de GFP, para evaluar la integración del ácido nucleico donante.

Ejemplo 11: Prueba de la extracción de ADN y amplificación por PCR del sitio diana de mMdr1a en el genoma de ratón (ejemplo comparativo)

Se probaron condiciones de PCR para amplificar ácido nucleico diana extraído de tejido usando embriones con 1-64 células extraídas como se describe en el **Ejemplo 10**. Se amplificó un fragmento de 900 pb que contenía la región diana de mMdr1a de ratón usando 36 ciclos de amplificación con 4 min de extensión a 60 °C en reacciones que contenían hasta 5 μl de la disolución QuickExtract de Epicentre en reacciones de 50 μl (**FIG. 16**). Estos resultados muestran que QuickExtract no interfiere con la amplificación por PCR, y que el ADN puede amplificarse a partir de la muestra extraída de solo 1-10 células. Para potenciar la sensibilidad, puede aumentarse el número de ciclos de PCR, o pueden realizarse reacciones de PCR anidada.

Ejemplo 12: Integración de RFLP donante de NotI en la región genómica PXR de rata (ejemplo comparativo)

Se inyectó un plásmido donante (con un brazo de 800 pb) para integrar un sitio de RFLP NotI en la región PXR del genoma de rata en embriones de rata con ARNm de ZFN como se ha descrito anteriormente. Se realizaron PCR, seguida de análisis de enzimas de restricción NotI y análisis de endonucleasas Cel-I, usando ADN extraído de varios embriones. La amplificación por PCR fue satisfactoria con varios embriones (**FIG. 17A**), y el análisis de endonucleasas Cel-I reveló que la mayoría de los fragmentos tenía cambios de secuencia de ácidos nucleicos en la diana deseada (**FIG. 17B**).

Ejemplo 13: Integración de RFLP donante de NotI en la región genómica mMdr1a de ratón (ejemplo comparativo)

Se repitió la integración dirigida de RFLP de NotI en la región mMdr1a de ratón como se describe en el Ejemplo 8. Se amplificó la región mMdr1a usando PCR y se digirió con NotI. La amplificación por PCR fue satisfactoria con varios embriones (**FIG. 18**), y la digestión con NotI reveló que varios embriones comprendieron el sitio de RFLP integrado (véanse, por ejemplo, los carriles 13, 17, 19, 20 y 23). En total, integración dirigida en 7 de los 32 embriones para los que se generaron datos.

Estos resultados se confirmaron repitiendo la reacción de digestión de NotI después de limpiar adicionalmente el producto de reacción de PCR (**FIG. 19**).

Ejemplo 14: Prueba de extracción de ADN y amplificación por PCR del sitio diana PXR en el genoma de rata (ejemplo comparativo)

Se probó la amplificación por PCR de la región de PXR de blastocistos para determinar el nivel de sensibilidad. La reacción de PCR contuvo 5 μl de molde, 5 μl de tampón de PCR, 5 μl de cada cebador, 0,5 μl de enzima Taq polimerasa y 33,5 μl de agua para una reacción de 50 μl . El molde consistió en ADN no diluido extraído de blastocistos de rata o ADN diluido a una relación de 1:2, 1:6, 1:10 y 1:30 (**FIG. 20**).

Ejemplo 15: Integración de RFLP donante de NotI en la región genómica PXR de rata (ejemplo comparativo)

Se inyectó un plásmido donante (con brazos de homología de 800 pb) para integrar un sitio de RFLP NotI en la región PXR del genoma de rata en embriones de rata con ARNm de ZFN como se ha descrito anteriormente. Se inyectaron un total de 123 embriones, y sobrevivieron 106. Se inyectaron concentraciones de ácidos nucleicos decrecientes para probar la toxicidad. De los 51 embriones inyectados con 5 ng de ácidos nucleicos, 17 sobrevivieron y se dividieron en embriones de dos células en el día dos. De los 23 embriones inyectados con 2 ng de ácidos nucleicos, 14 sobrevivieron y se dividieron en embriones de dos células en el día dos. De los 29 embriones

inyectados con 10 ng de ácidos nucleicos, 12 sobrevivieron y se dividieron en embriones de dos células en el día dos. De los diez embriones de control no inyectados, todos sobrevivieron y se dividieron en embriones de dos células en el día dos.

- 5 Se realizaron amplificación por PCR de la región PXR, seguido de análisis de endonucleasas NotI y Cel-I, usando ADN extraído de varios embriones. La amplificación por PCR fue satisfactoria con varios embriones, y el análisis de endonucleasas NotI y Cel-I reveló que 18 de los 47 embriones tenían cambios de secuencia de ácidos nucleicos en la diana deseada (**FIG. 21**).

Ejemplo 16: Integración dirigida de RFLP en la región diana mMdr1a del genoma de ratón en feto (ejemplo comparativo)

- 10 Se inyectó un plásmido donante (con brazos de homología de 800 pb) para introducir NotI en la región mMdr1a del genoma de ratón en embriones de ratón con ARNm de ZFN como se ha descrito anteriormente. Uno de los cuatro fetos bien desarrollados 12,5 dpc fue positivo para el sitio NotI. Las cuatro deciduas fueron negativas. (**FIG. 22**).

Ejemplo 17: Integración dirigida de GFP en el locus mMdr1a de un feto (ejemplo comparativo)

- 15 Se inyectó un plásmido donante (con brazos de homología de 800 pb) para introducir el casete de GFP en la región mMdr1a del genoma de ratón en embriones de ratón con ARNm de ZFN como se ha descrito anteriormente. Dos de los cuarenta fetos 12,5 dpc fueron positivos para el casete de GFP (**FIG. 23**).

Ejemplo 18. Integración dirigida de RFLP de la región diana PXR del genoma de rata en un feto (ejemplo comparativo)

- 20 Se inyectó un plásmido donante (con brazos de homología de 800 pb) para introducir NotI en la región PXR del genoma de rata en embriones de ratón con ARNm de ZFN como se ha descrito anteriormente. Uno de los ocho fetos 13 dpc fue positivo para el sitio NotI (**FIG. 24**).

REIVINDICACIONES

1. Un método *in vitro* de modificación de un gen IgM endógeno en una célula de rata, comprendiendo el método:
 - 5 introducir, en la célula de rata, uno o más ARNm que codifican primera y segunda nucleasas con dedos de cinc (ZFN) que se unen a sitios diana en el gen IgM endógeno en condiciones tales que las ZFN escindan el uno o más genes celulares endógenos tal que el gen IgM endógeno se modifique, en el que la primera ZFN comprende una ZFP que tiene las regiones de hélice de reconocimiento como se muestra en las ZFP designadas 17747, 17759, 17756, 17767 y 17764 como se expone en una única fila de la Tabla 6.
- 10 2. El método de la reivindicación 1, en el que la modificación comprende introducir una secuencia exógena en el genoma de una célula de rata por recombinación homóloga estimulada por escisión del uno o más genes celulares endógenos.
3. El método de la reivindicación 2, en el que la secuencia exógena está integrada físicamente en el genoma.
4. El método de la reivindicación 2, en el que la secuencia exógena se integra en el genoma copiando la secuencia exógena en el genoma de la célula huésped mediante procesos de replicación de ácido nucleico.
- 15 5. El método de la reivindicación 4, en el que el proceso de replicación de ácido nucleico comprende reparación dirigida por homología de la rotura bicatenaria.
6. El método de la reivindicación 4, en el que el proceso de replicación de ácido nucleico comprende integración dirigida no dependiente de homología.
7. El método de la reivindicación 1, en el que la modificación resulta de unión de extremos no homólogos tras la escisión.
- 20 8. El método de la reivindicación 7, en el que la primera y segunda nucleasas con dedos de cinc escinden el genoma en dos sitios y además en el que la unión de extremos no homólogos produce una delección entre el primer y segundo sitios de escisión.
9. El método de cualquiera de las reivindicaciones 1 a 8, en el que las nucleasas con dedos de cinc comprenden un dominio de escisión para medio-dominio de escisión de una endonucleasa de restricción de tipo IIS.
- 25 10. Un método para la rotura de la línea germinal de uno o más genes diana en rata, comprendiendo el método modificar una o más secuencias de genes en el genoma de una o más células de un embrión de rata, comprendiendo el método:
 - 30 modificar uno o más de los genes diana en una o más células de un embrión de rata según el método de cualquiera de las reivindicaciones 1 a 9; y
 - permitir que el embrión de rata se desarrolle, en el que secuencias de genes modificadas están presentes en al menos una porción de gametos de la rata sexualmente madura.

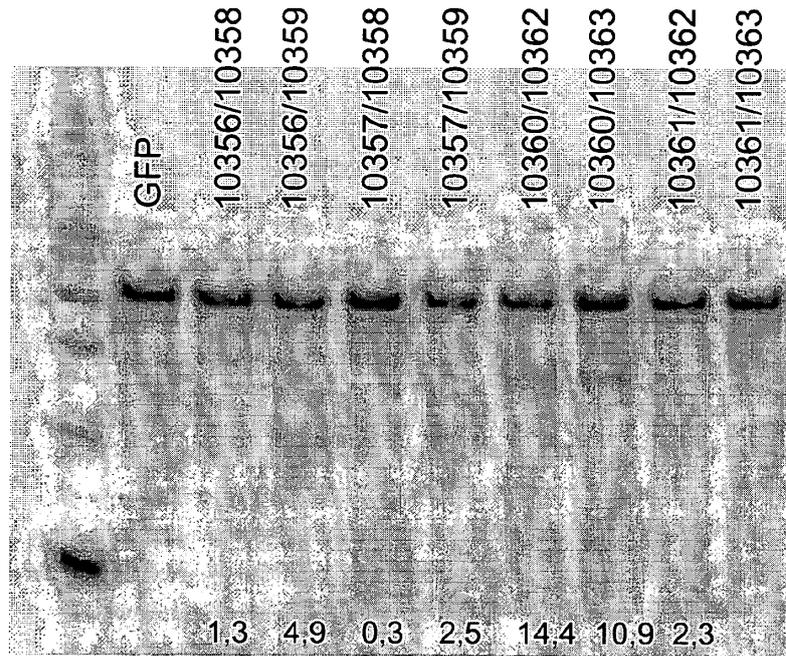


FIG. 1

Análisis de ZFN dirigidas a eGFP
en células C6 de rata - día 3, ³²P Cel-1

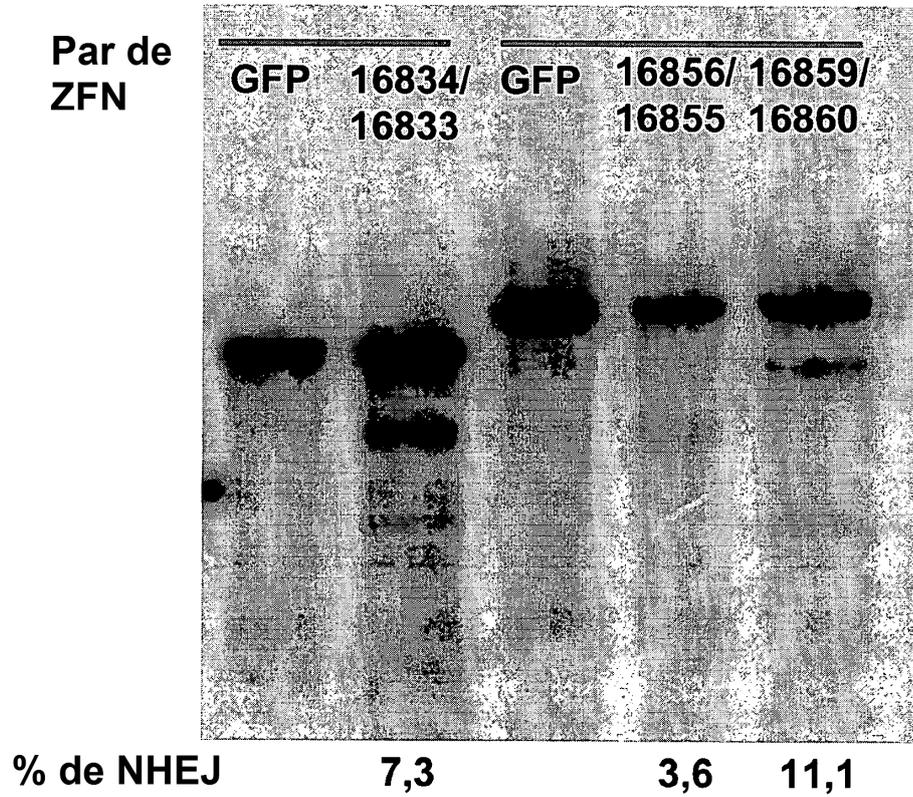


FIG. 2

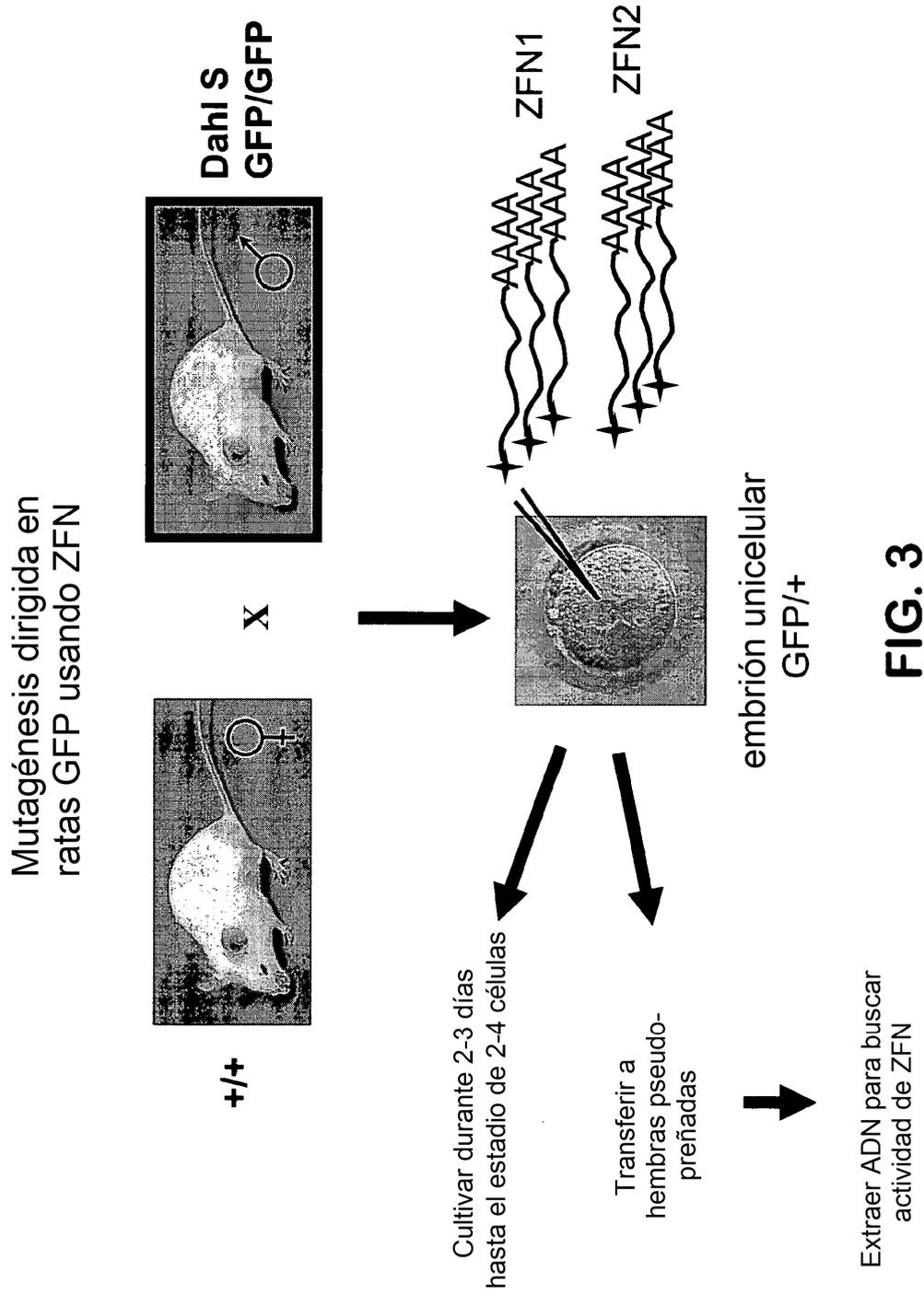


FIG. 3

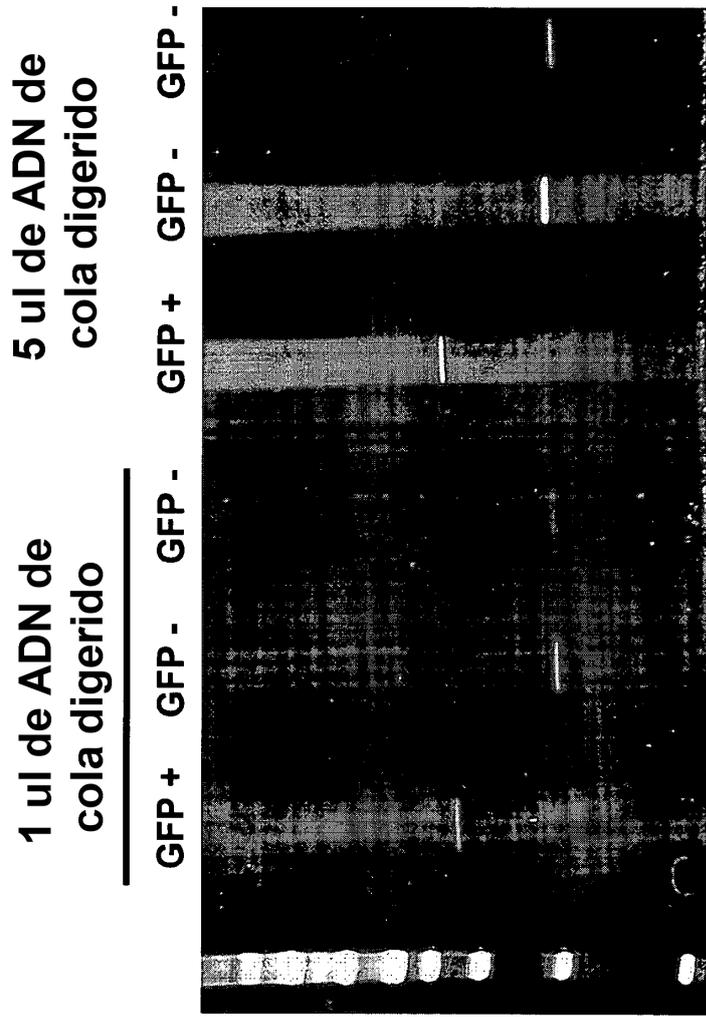
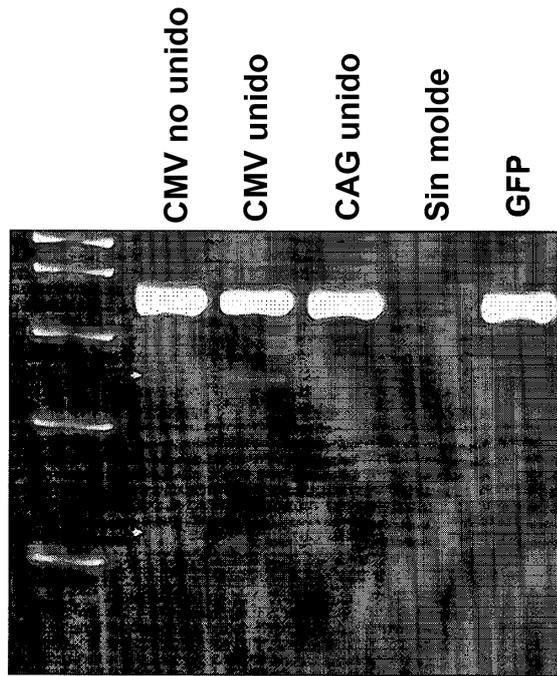


FIG. 4B



FIG. 4A



17747 / 17749
Exón 1

FIG. 5

PCR +/- la adición de ADN no mutante (líneas emparejadas)

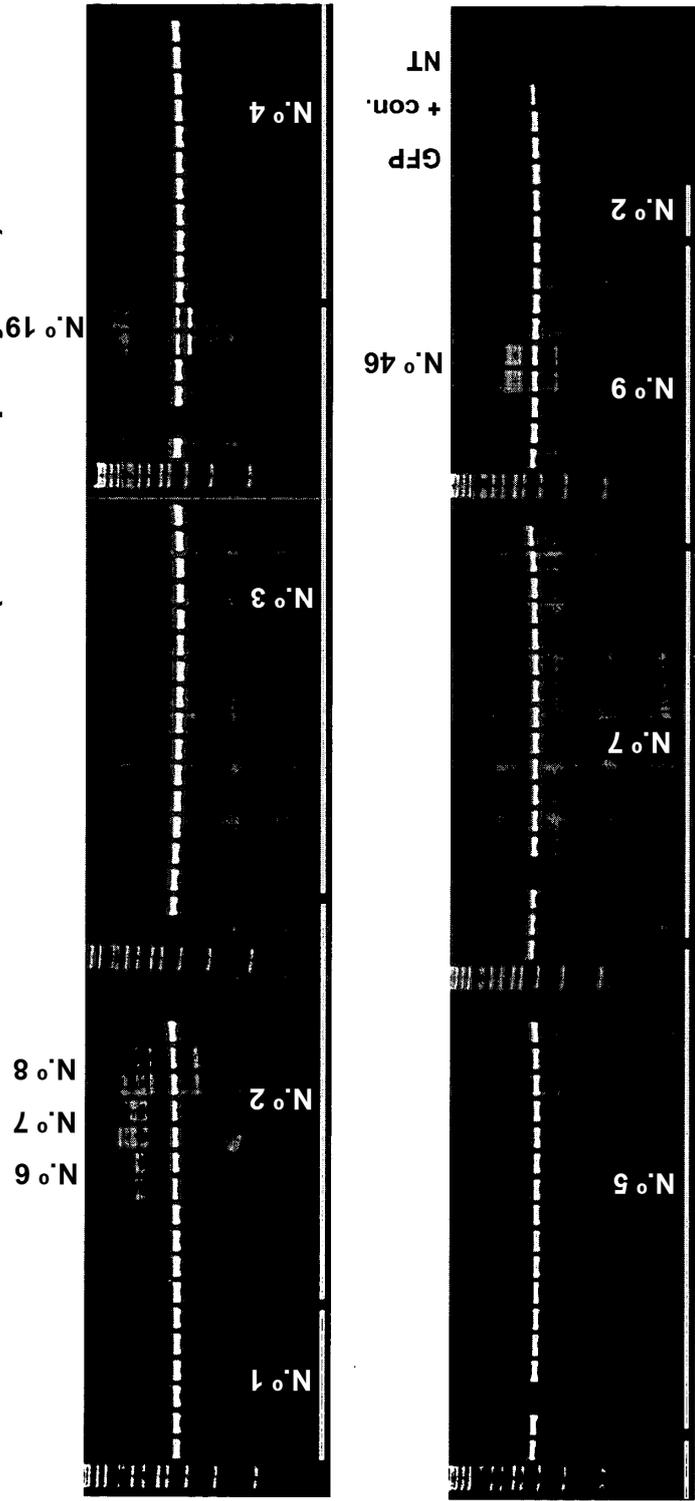
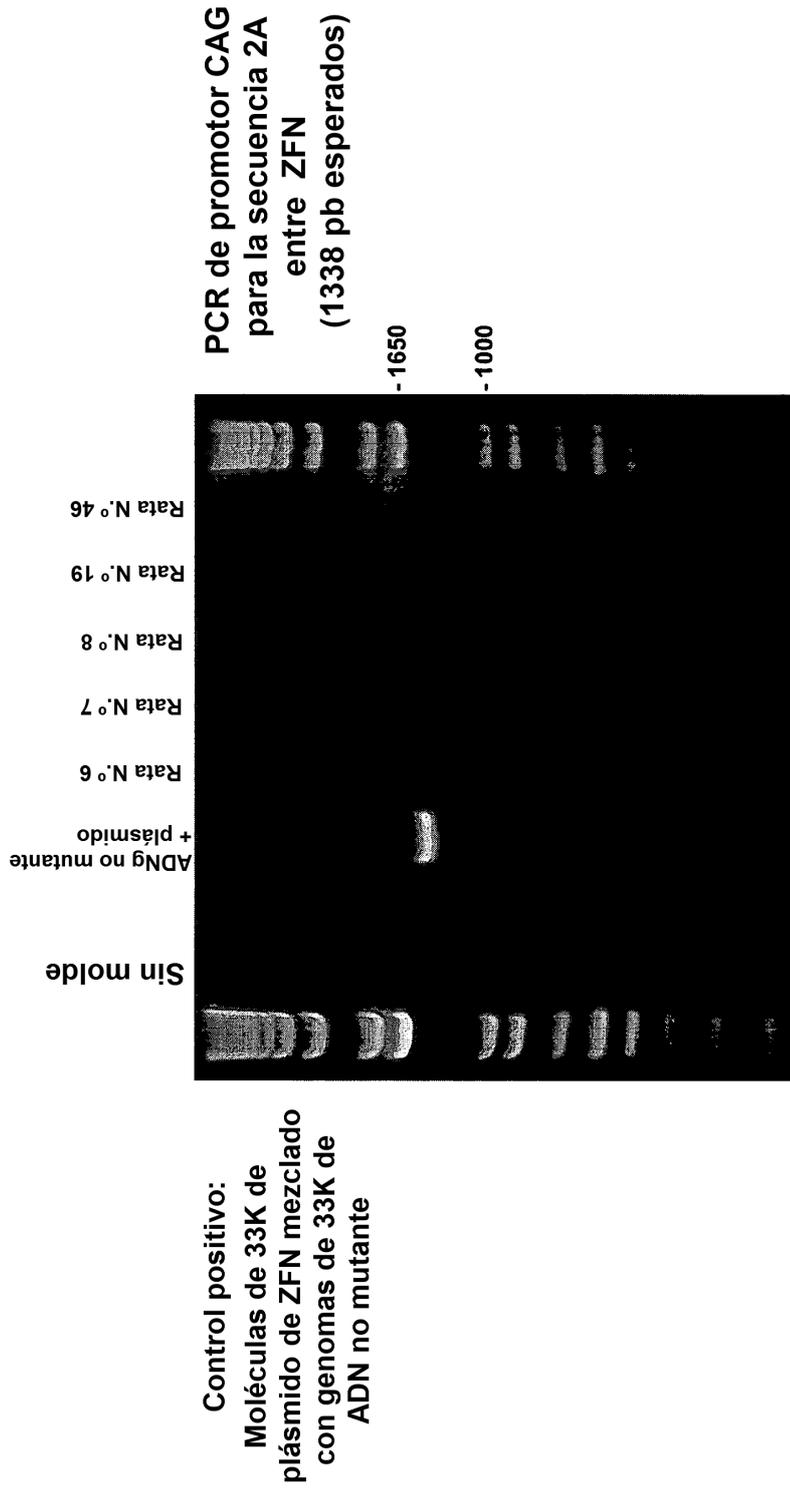


FIG. 6



Inserción no detectable de ADN de plásmido

FIG. 7

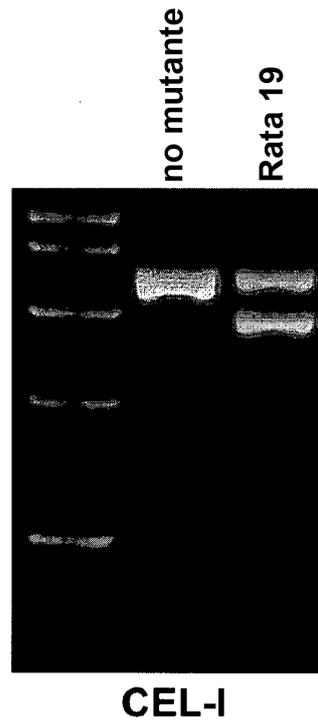


FIG. 8A

WT
 Alelo no mutante de rata 19
 Alelo de delección de rata 19

CCTGCGAGAGCCCCCTGTCTGATGAGAAATTGGTGGCCATGGGCTGCCCGGGGACTTCTGCCCCAGCTCCATTTCTTCTCCCTGGAACTTACCAGAAACAACACTGAAG
 CCTGCGAGAGCCCCCTGTCTGATGAGAAATTGGTGGCCATGGGCTGCCCGGGGACTTCTGCCCCAGCTCCATTTCTTCTCCCTGGAACTTACCAGAAACAACACTGAAG
 CCTGCGAGAGCCCCCTGTCTGATGAGAAATTGGTGGCCATGGGCTGCCCGGGGACTTCTGCCCCAGCTCCATTTCTTCTCCCTGGAACTTACCAGAAACAACACTGAAG

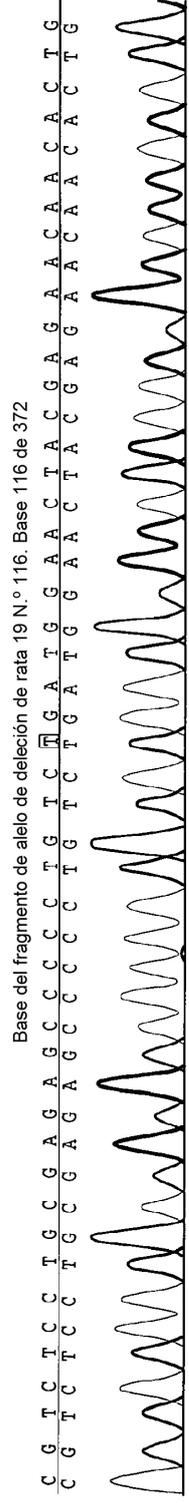
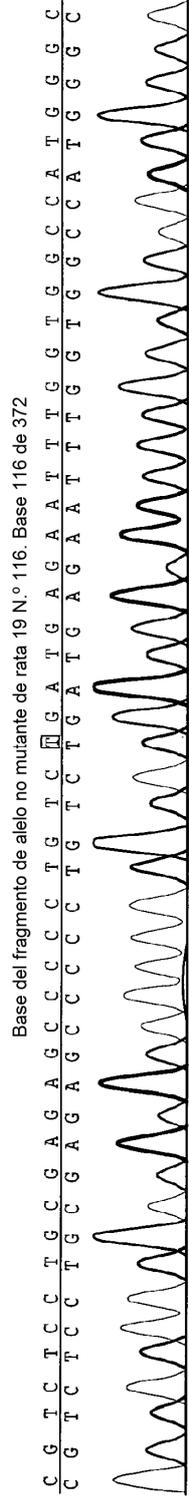
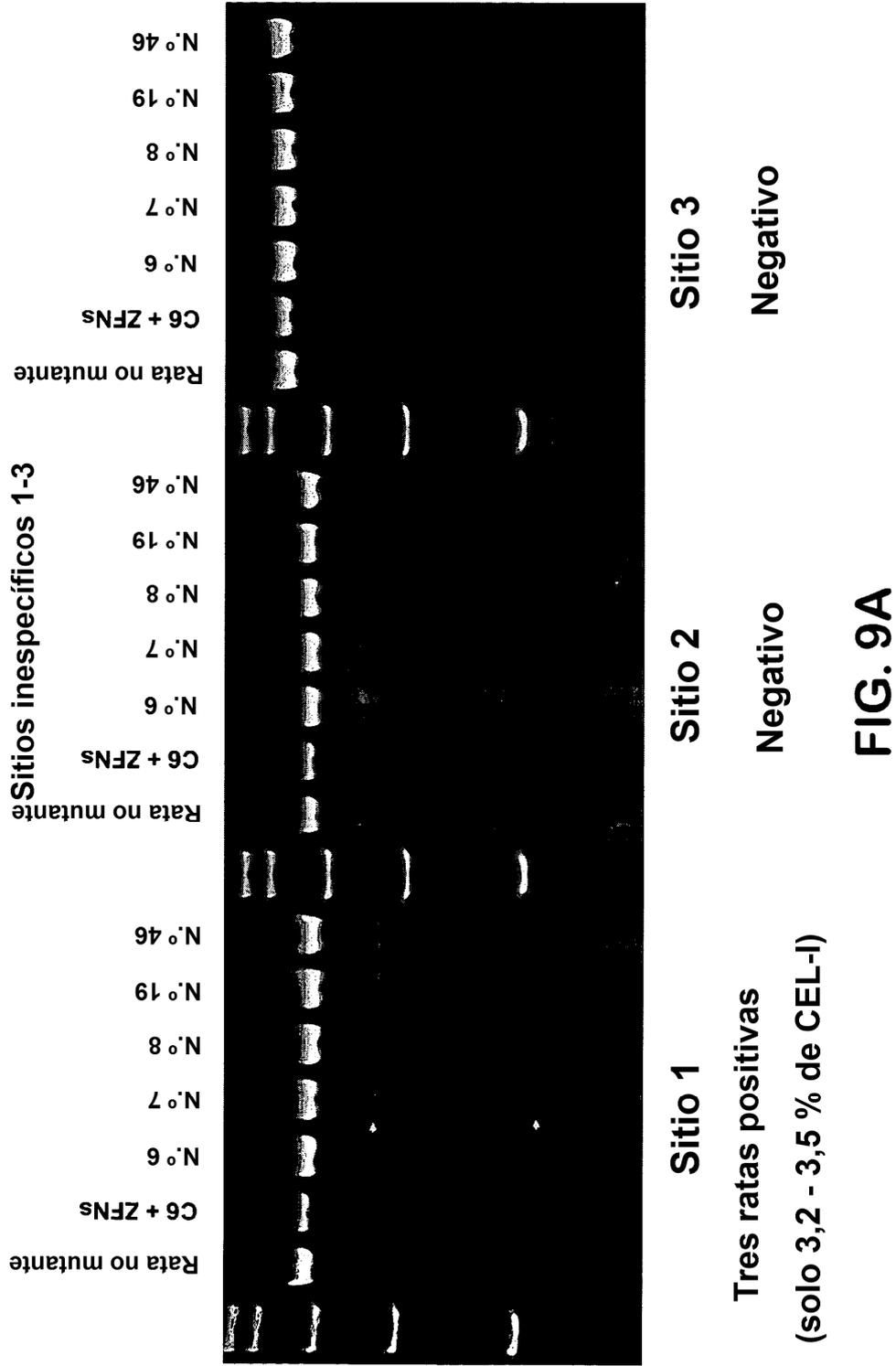


FIG. 8B



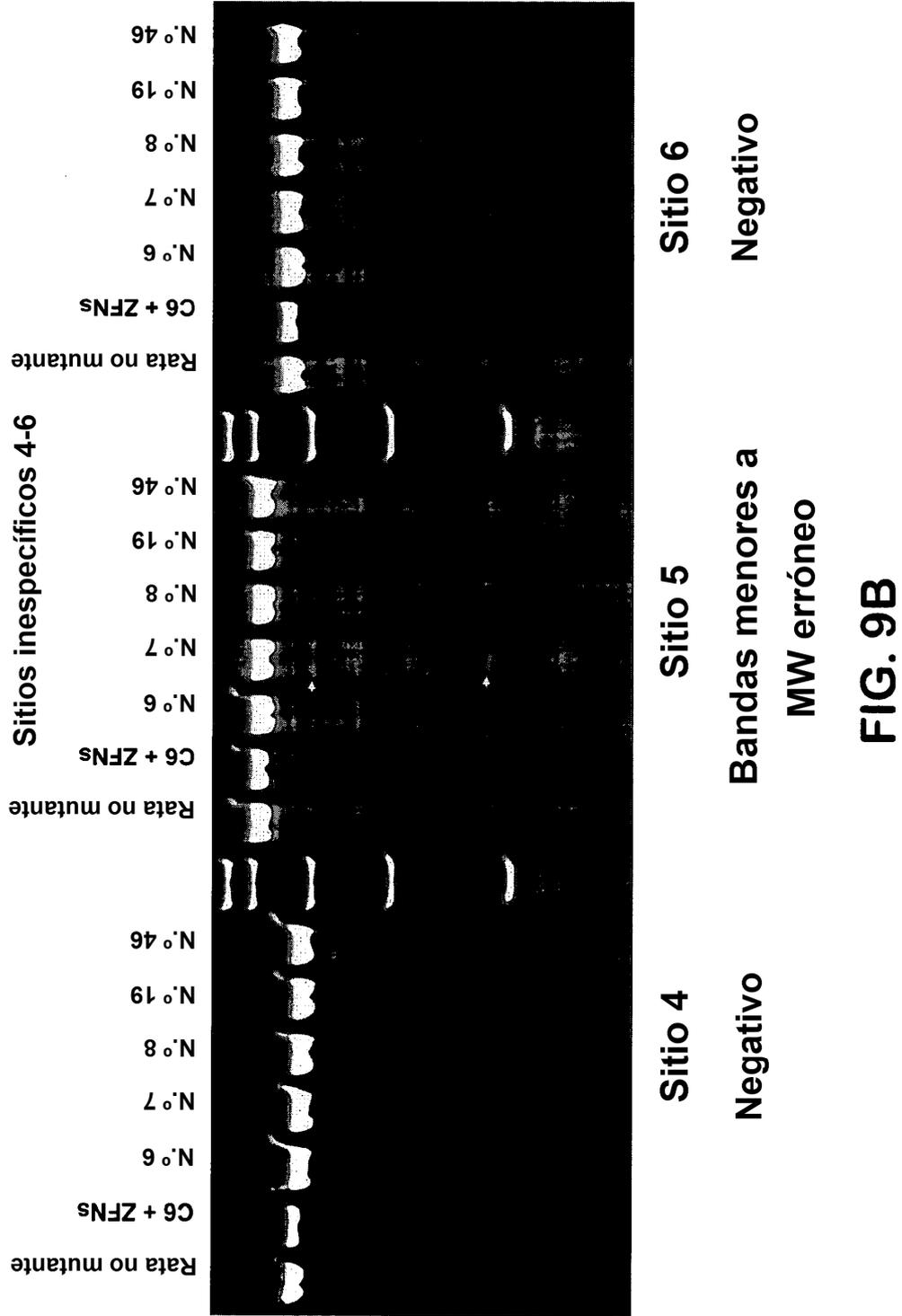


FIG. 9B

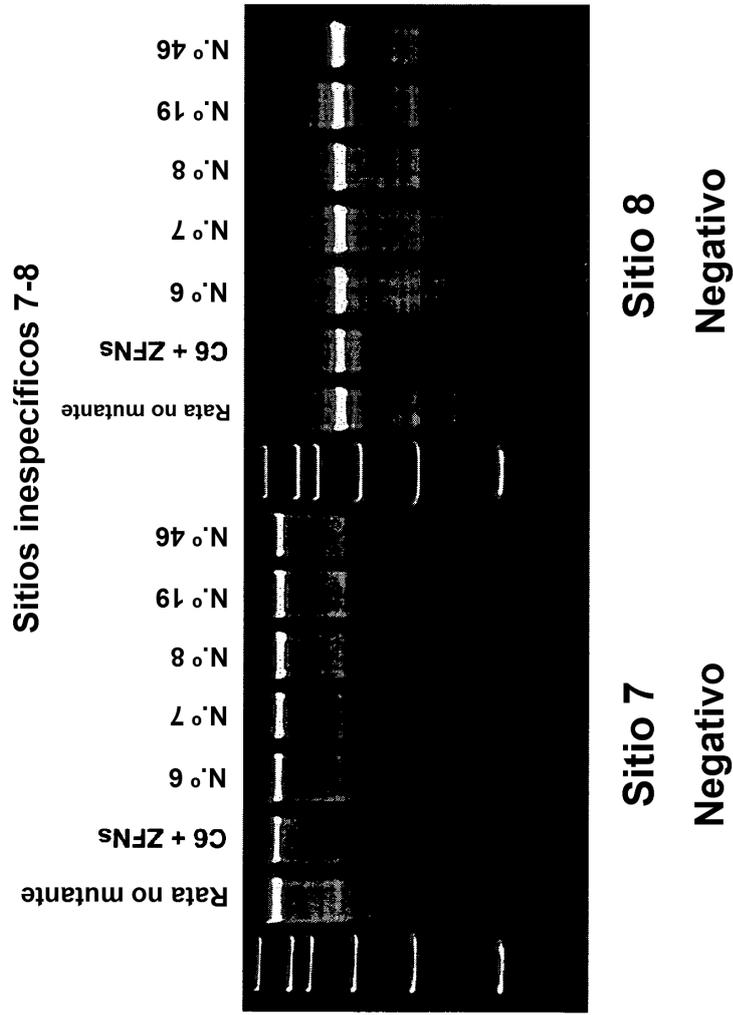


FIG. 9C

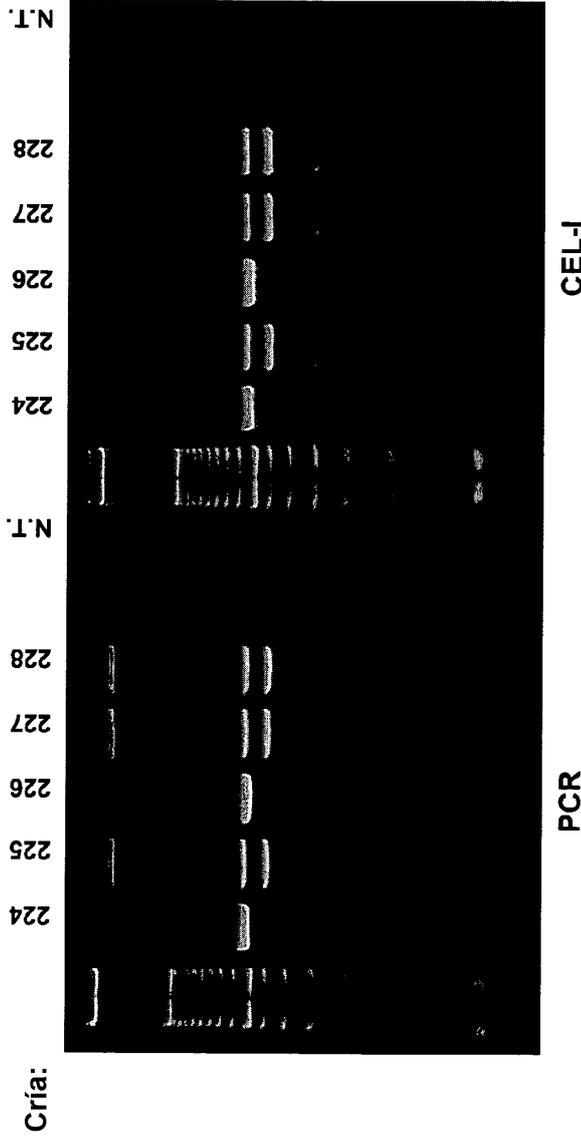


FIG. 11A

Alelo mutado de rata 19 parental
 No mutante 153/154
 Alelo no mutante de rata 225
 Alelo mutante de rata 225
 Alelo no mutante de rata 227
 Alelo mutante de rata 227
 Alelo no mutante de rata 228
 Alelo mutante de rata 228

```

CTGGGAGAGCCCCCTGTCT|
CTGGGAGAGCCCCCTGTCT|GATGAGAAATTGGTGGCCATGGGCTGGCCCGGGACTTCCTGGCCAGCTCCATTTCTTCTCTGGAACTACCAAGAAACAACACT
CTGGGAGAGCCCCCTGTCT|GATGAGAAATTGGTGGCCATGGGCTGGCCCGGGACTTCCTGGCCAGCTCCATTTCTTCTCTGGAACTACCAAGAAACAACACT
CTGGGAGAGCCCCCTGTCT|GATGAGAAATTGGTGGCCATGGGCTGGCCCGGGACTTCCTGGCCAGCTCCATTTCTTCTCTGGAACTACCAAGAAACAACACT
CTGGGAGAGCCCCCTGTCT|GATGAGAAATTGGTGGCCATGGGCTGGCCCGGGACTTCCTGGCCAGCTCCATTTCTTCTCTGGAACTACCAAGAAACAACACT
CTGGGAGAGCCCCCTGTCT|GATGAGAAATTGGTGGCCATGGGCTGGCCCGGGACTTCCTGGCCAGCTCCATTTCTTCTCTGGAACTACCAAGAAACAACACT
CTGGGAGAGCCCCCTGTCT|GATGAGAAATTGGTGGCCATGGGCTGGCCCGGGACTTCCTGGCCAGCTCCATTTCTTCTCTGGAACTACCAAGAAACAACACT
CTGGGAGAGCCCCCTGTCT|
    
```

FIG. 11B

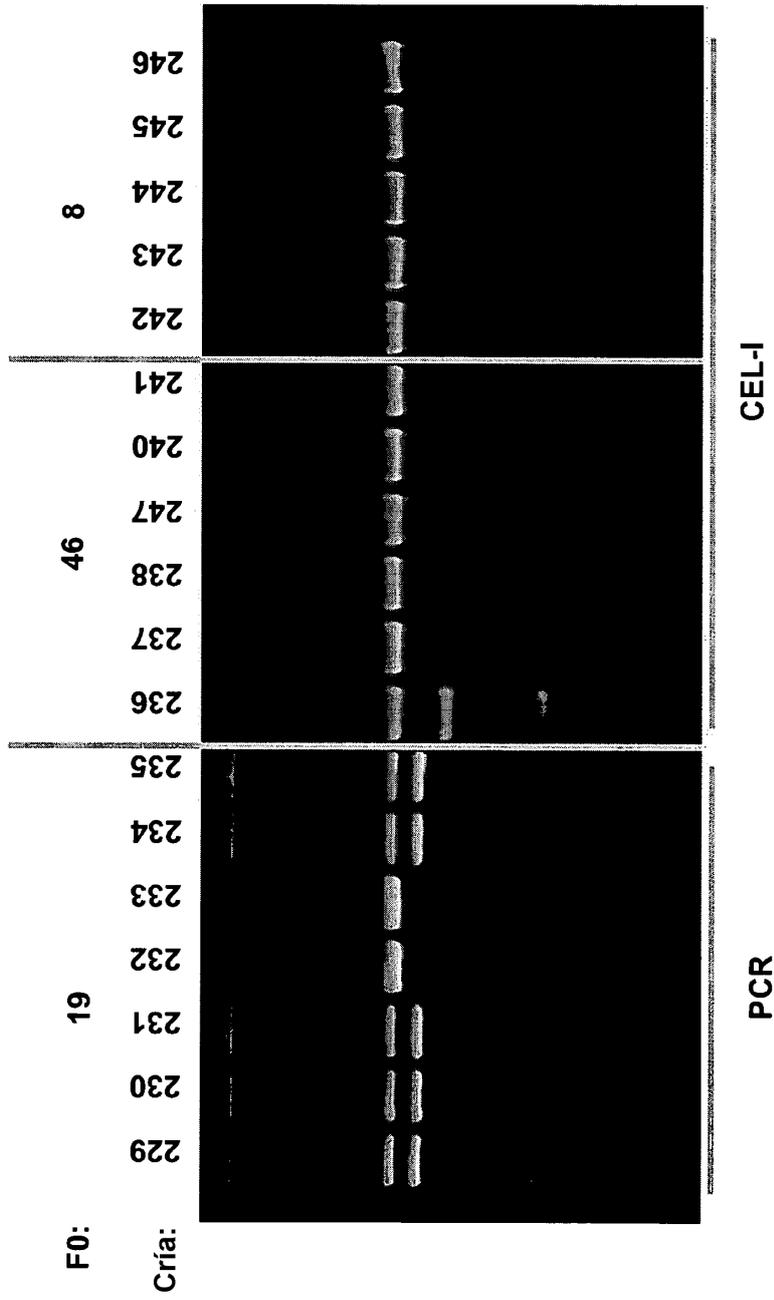


FIG. 11C

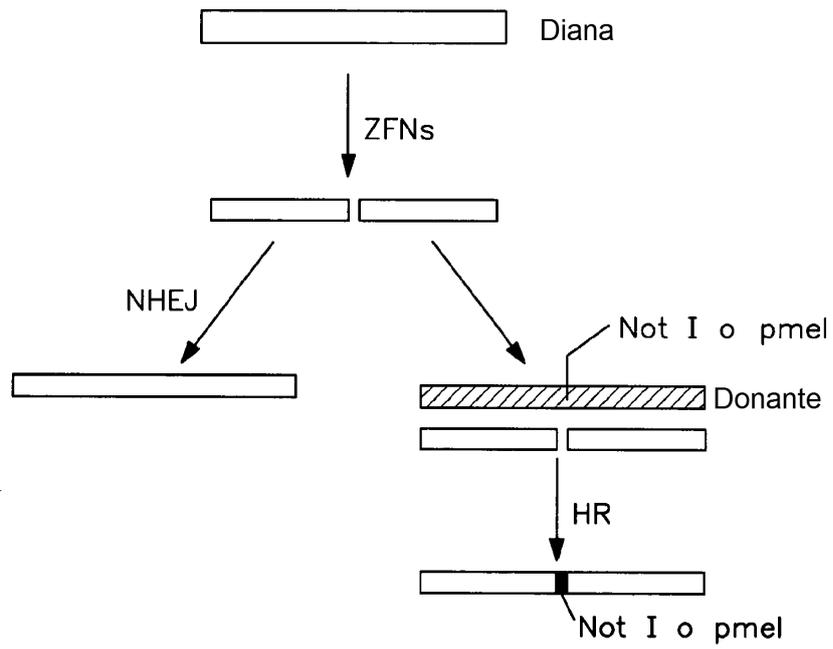


FIG. 12

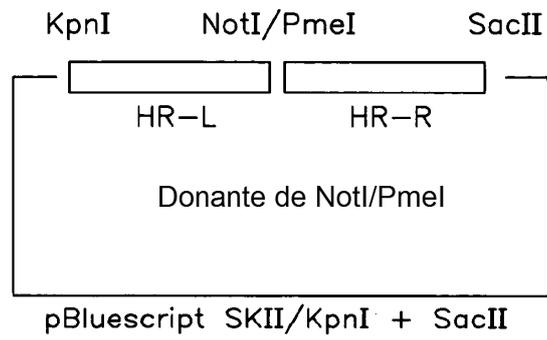


FIG. 13

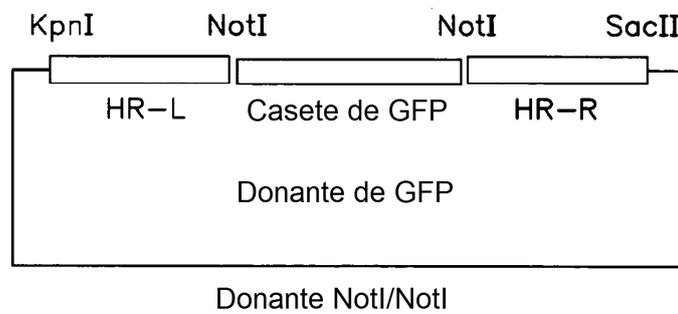


FIG. 14

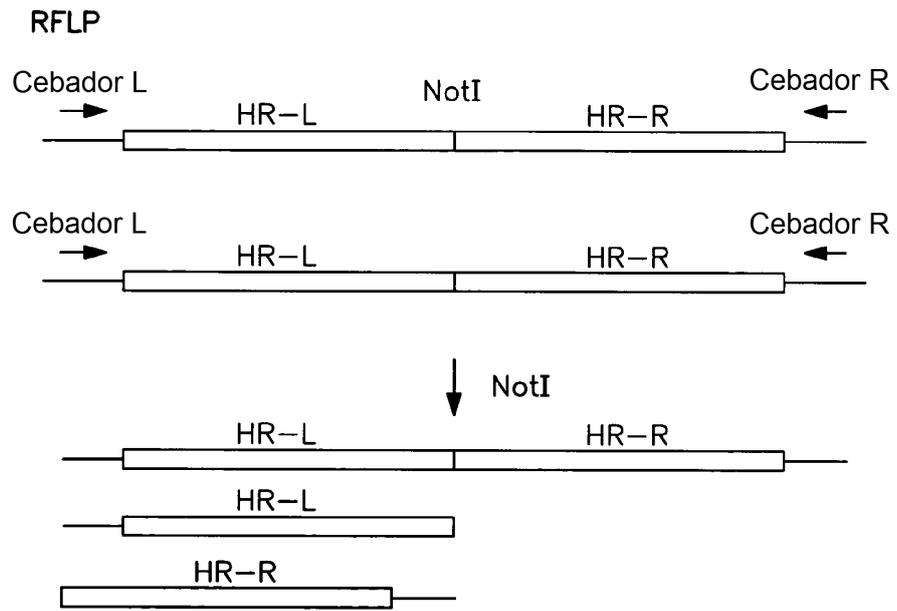


FIG. 15A

Integraciones mayores (por ejemplo, casete de GFP)

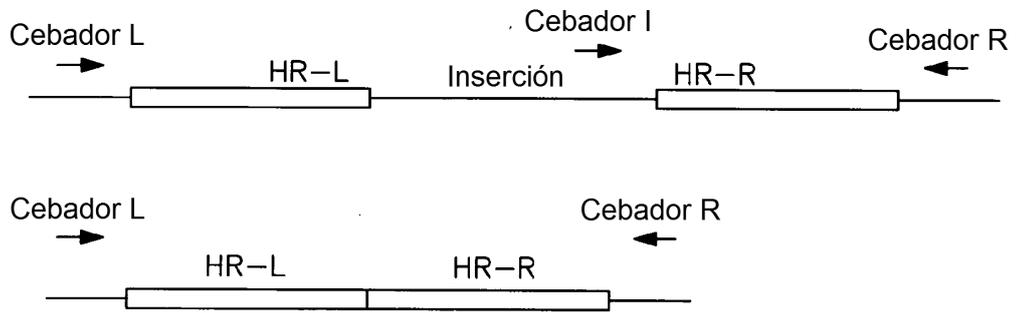


FIG. 15B

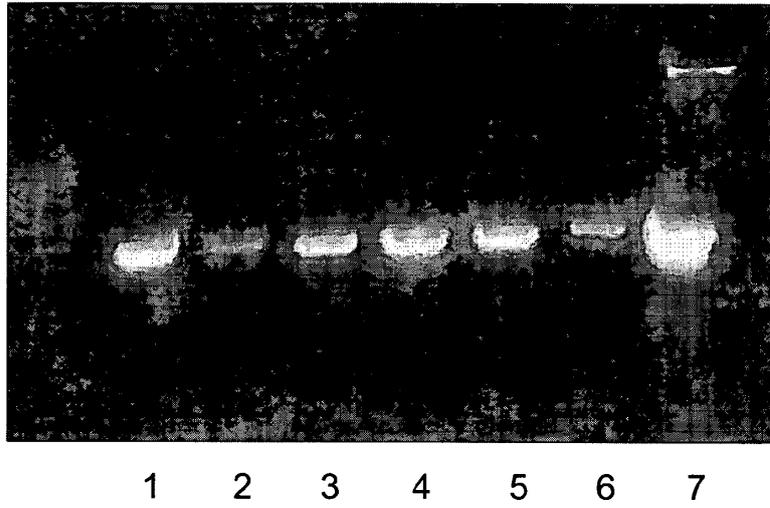


FIG. 16

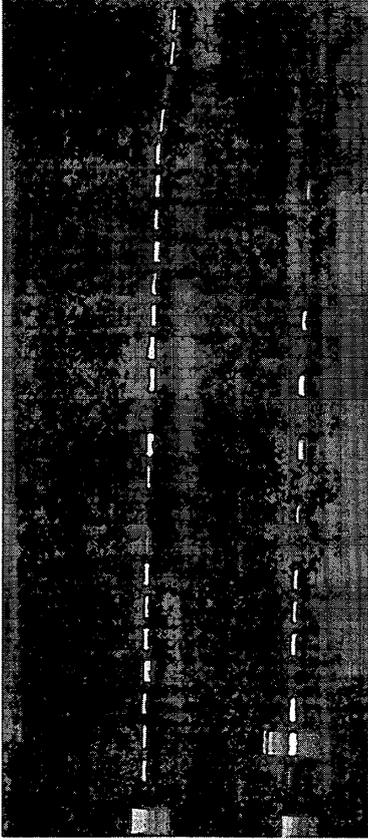


FIG. 17A

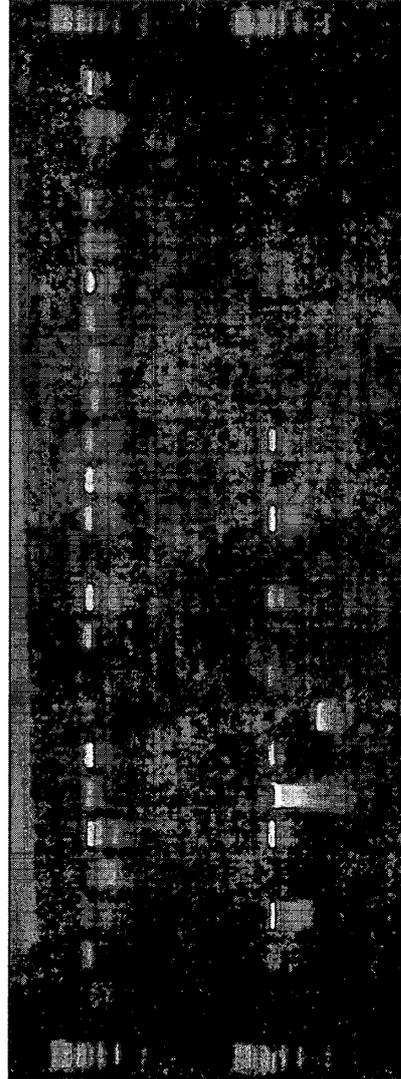


FIG. 17B

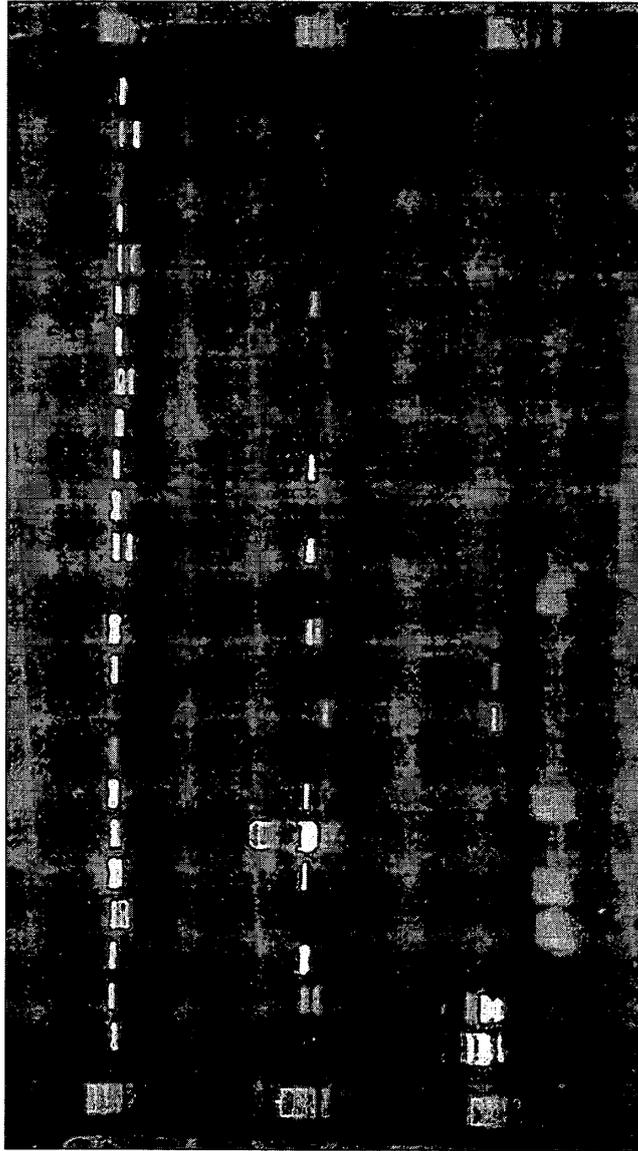


FIG. 18A

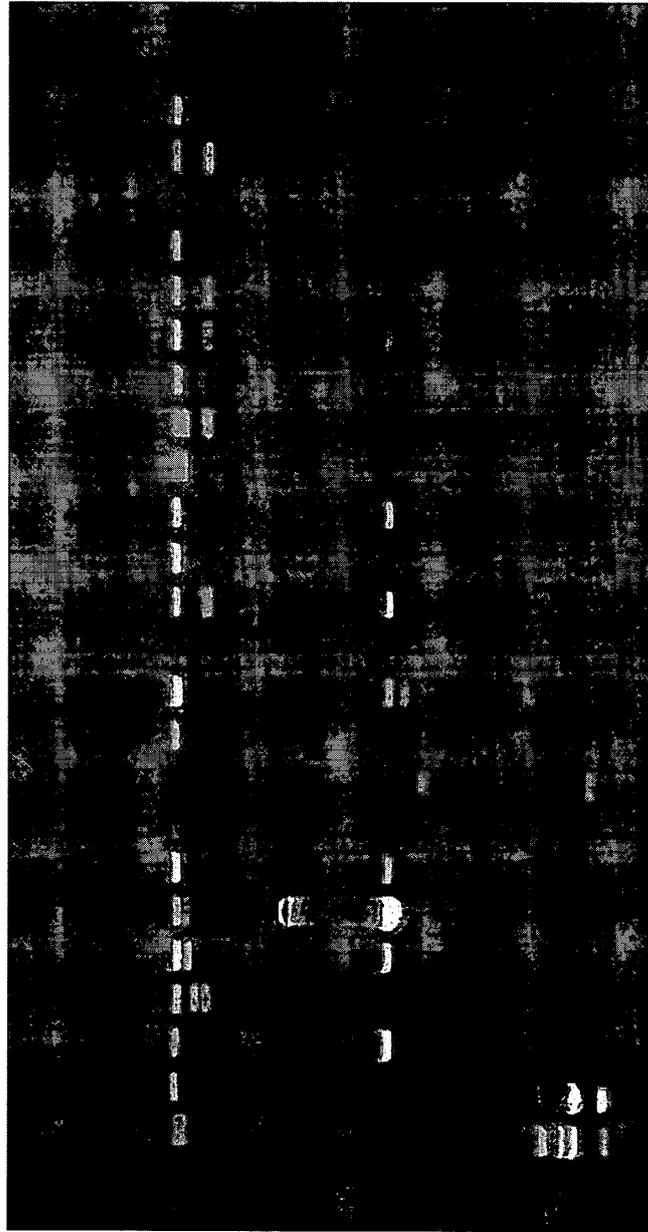


FIG. 18B

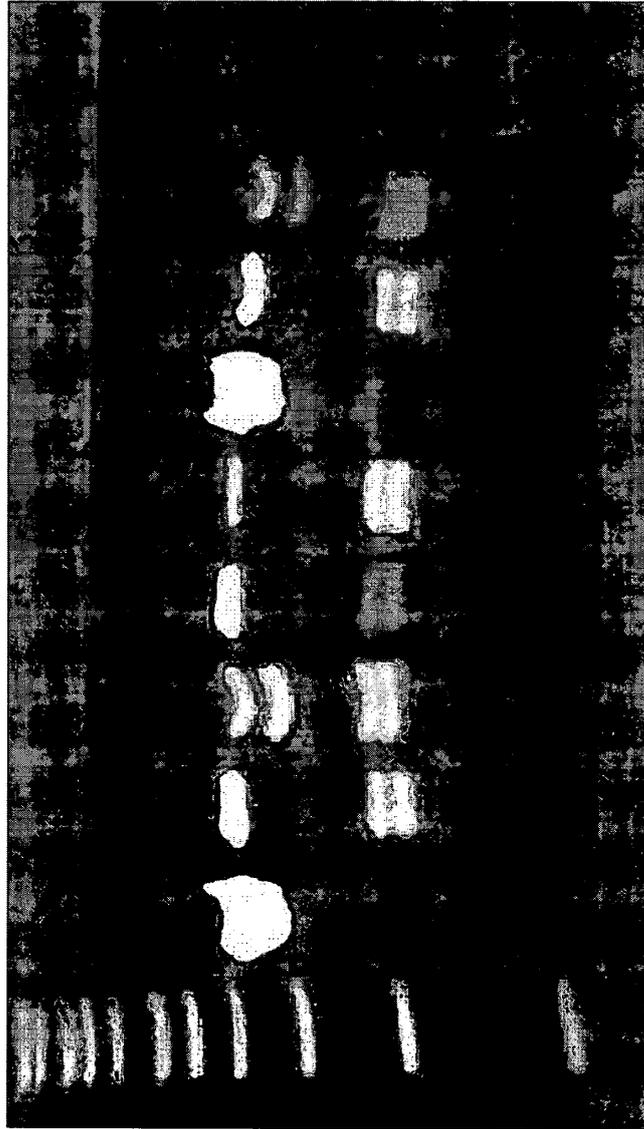


FIG. 19

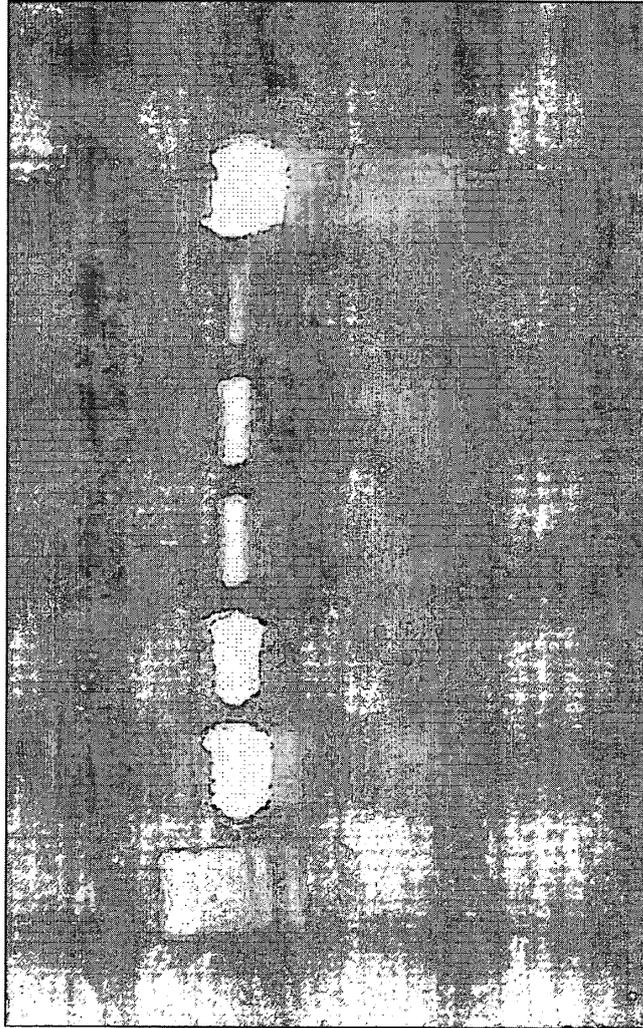


FIG. 20

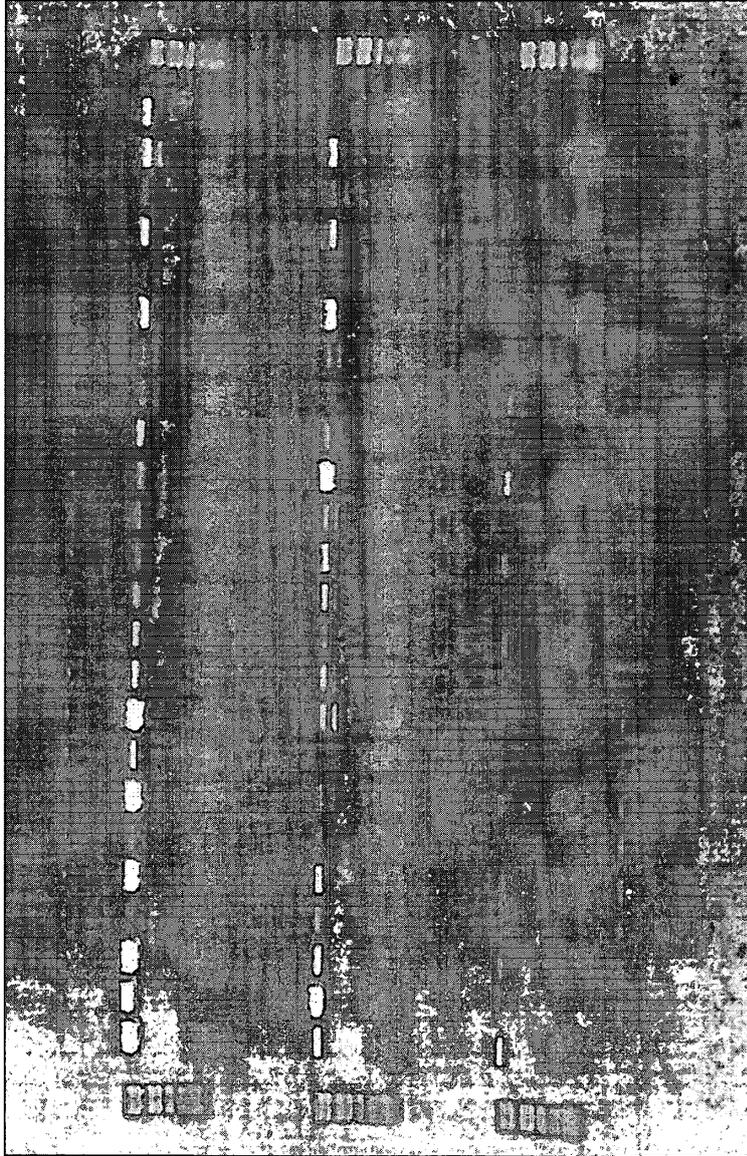


FIG. 21A

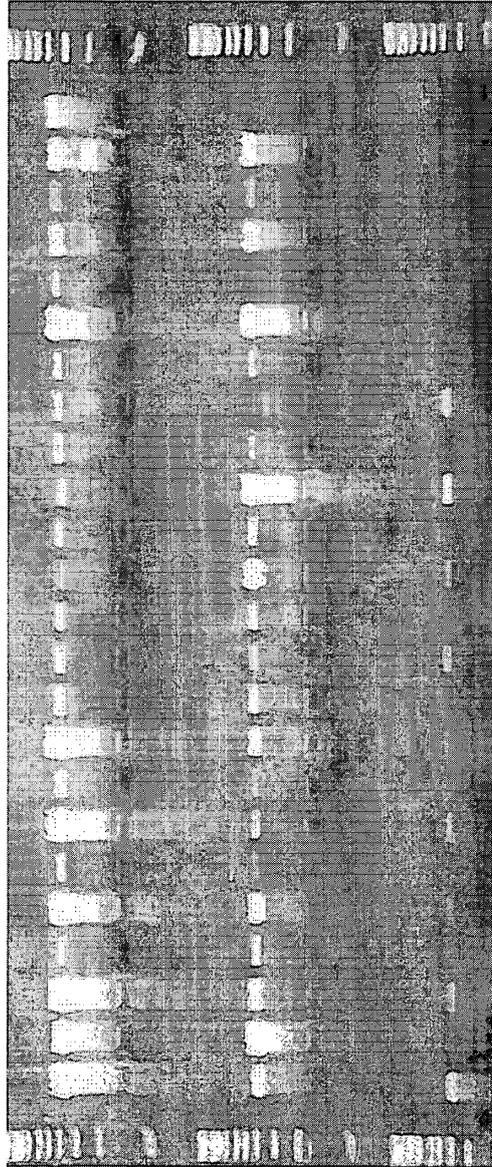


FIG. 21B



1 2 3 4 5 6 7 8

FIG. 22



1 2 3 4 5 6 7 8

FIG. 22

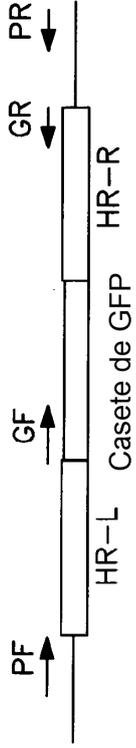


FIG. 23A



FIG. 23B



FIG. 23C



FIG. 23D



FIG. 23E

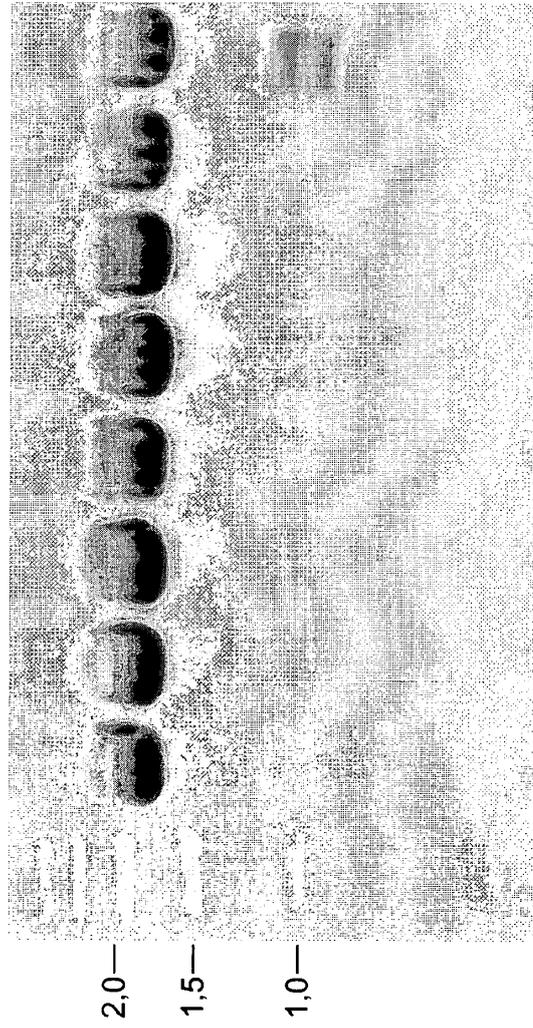


FIG. 24