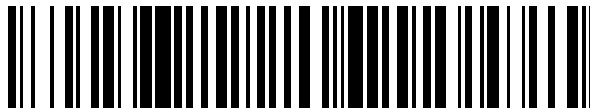


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 700 596**

51 Int. Cl.:

C12N 15/85 (2006.01)

C12N 9/22 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.10.2014 PCT/US2014/060788**

87 Fecha y número de publicación internacional: **18.06.2015 WO15088643**

96 Fecha de presentación y número de la solicitud europea: **15.10.2014 E 14790457 (7)**

97 Fecha y número de publicación de la concesión europea: **26.09.2018 EP 3080279**

54 Título: **Métodos y composiciones para la modificación dirigida de un genoma**

30 Prioridad:

11.12.2013 US 201361914768 P
26.06.2014 US 201462017416 P
25.07.2014 US 201462029261 P
19.09.2014 US 201462052906 P
03.10.2014 US 201462059527 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
18.02.2019

73 Titular/es:

REGENERON PHARMACEUTICALS, INC.
(100.0%)
777 Old Saw Mill River Road
Tarrytown NY 10591, US

72 Inventor/es:

FRENDEWEY, DAVID;
AUERBACH, WOJTEK;
LAI, KA-MAN VENUS;
VALENZUELA, DAVID M. y
YANCOPOULOS, GEORGE D.

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 700 596 T3

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

DESCRIPCIÓN

Métodos y composiciones para la modificación dirigida de un genoma

5 Referencia a un listado de secuencias presentado como un archivo de texto mediante efs web

La copia oficial del listado de secuencias se presenta por vía electrónica mediante EFS-Web como un listado de secuencias en formato ASCII con un archivo llamado 453460SEQLIST.TXT, creado el 15 de octubre de 2014 y que tiene un tamaño de 27,5 kilobytes, y se presenta junto con la descripción.

10

Antecedentes de la invención

15 Si bien las ratas se han considerado como un sistema de modelo animal importante que pueden recapitular la patología de varias enfermedades humanas, incluidas, pero sin limitarse a, enfermedades cardiovasculares (*por ejemplo*, hipertensión), metabólicas (*por ejemplo*, obesidad, diabetes), neurológicas (*por ejemplo*, patologías dolorosas), y una variedad de cánceres, el uso de ratas en el modelado de enfermedades humanas se ha limitado en comparación con los ratones, debido en parte a la no disponibilidad de células de rata pluripotentes transmisibles en la línea germinal, las que pueden mantener su pluripotencia después de una serie de modificaciones genéticas *in vitro*, *por ejemplo*, una o más electroporaciones en serie, y debido en parte a la falta de tecnologías de transformación eficientes que permitan la introducción o delección de grandes secuencias de ADN genómico, o la sustitución de grandes secuencias de ADN genómico con secuencias de ácido nucleico exógenas en células de rata pluripotentes.

20

25 Existe una necesidad en la técnica de composiciones y métodos que permitan cambios dirigidos precisos en el genoma de un organismo, que puedan abrir o expandir áreas actuales de descubrimiento de objetivos y validar agentes terapéuticos de manera más rápida y fácil.

30

Mali y otros (Science vol. 339, núm.6121, páginas 823-826) y Cong y otros (Science vol 339, núm. 6121, páginas 819-823) describen un método para modificar un genoma en un locus genómico de interés en una célula de ratón o una célula humana, que comprende el uso de ARN de CRISPR, una proteína Cas, ARNtracr y un vector que comprende un ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3'. Los documentos WO2006044962 y WO2007117410 describen inserciones de grandes fragmentos genómicos en el genoma de un ratón o humano.

35

Resumen

40

La presente invención es como se define en las reivindicaciones.

La invención se refiere a un método *in vitro* para modificar un genoma en un locus genómico de interés en una célula embrionaria (ES) de ratón, que comprende:

45

poner en contacto las células ES de ratón con una proteína Cas9, un ARN de CRISPR que se hibrida a una secuencia objetivo en el locus genómico de interés, un ARNtracr, y un vector de direccionamiento grande (LTVEC) que tiene un tamaño de al menos 10 kb y comprende un inserto de ácido nucleico flanqueado por:

50

(i) un brazo de homología 5' que es homólogo a una secuencia objetivo 5' en el locus genómico de interés; y
 (ii) un brazo de homología 3' que es homólogo a una secuencia objetivo 3' en el locus genómico de interés,
 45 en donde después de poner en contacto las células ES de ratón con la proteína Cas9, el ARN de CRISPR, y el ARNtracr en la presencia del LTVEC, el genoma de la célula ES de ratón se modifica para que comprenda una modificación genética dirigida que comprende la delección de una región del locus genómico de interés en donde la delección es de al menos 30 kb y/o la inserción del inserto de ácido nucleico en el locus genómico de interés en donde la inserción es de al menos 30 kb.

55

En una modalidad, la modificación genética dirigida es una modificación genética bialélica.

En una modalidad, el LTVEC es de al menos 15 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, o al menos 90 kb. En otra modalidad, el LTVEC es de al menos 100 kb, al menos 150 kb, o al menos 200 kb.

60

Esta descripción describe que la célula eucariota es una célula de mamífero. Esta descripción describe que la célula de mamífero es un fibroblasto.

65

Esta descripción describe que la célula eucariota es una célula pluripotente. Esta descripción describe que la célula pluripotente es una célula pluripotente humana. Esta descripción describe que la célula pluripotente humana es una célula madre embrionaria (ES) humana o una célula madre adulta humana. Esta descripción describe que la célula pluripotente humana es una célula progenitora humana con desarrollo limitado. Esta descripción describe que, la célula pluripotente humana es una célula madre pluripotente inducida (iPS) humana.

En una modalidad, la proteína Cas es Cas9.

En una modalidad, la secuencia objetivo está flanqueada por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM). En una modalidad, la secuencia objetivo está flanqueada inmediatamente en el extremo 3' por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).

5 En algunas modalidades, la suma total de los brazos de homología 5' y 3' es de aproximadamente 10 kb a aproximadamente 150 kb. En algunas modalidades, la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.

15 La descripción proporciona, además, que la modificación genética dirigida comprende: (a) una sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos homóloga u ortóloga; (b) una delección de una secuencia de ácidos nucleicos endógena; (c) una delección de una secuencia de ácidos nucleicos endógena, en donde la delección varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (d) inserción de una secuencia de ácidos nucleicos exógena; (e) inserción de una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb; (f) inserción de una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u ortóloga; (g) inserción de una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humana y una no humana; (h) inserción de un alelo condicional flanqueado por secuencias objetivo de recombinasa específica de sitio; (i) inserción de un gen marcador de selección o un gen reportero unido operativamente a un tercer promotor activo en la célula pluripotente; o (j) una combinación de estos.

35 En una modalidad, el locus genómico de interés comprende (i) una secuencia objetivo 5' que es homóloga al brazo de homología 5'; y (ii) una secuencia objetivo 3' que es homóloga al brazo de homología 3'.

40 En algunas modalidades, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas en al menos 5 kb pero menos de 3 Mb. En algunas modalidades, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas en al menos 5 kb pero menos de 10 kb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero menos de aproximadamente 1 Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 2 Mb pero menos de aproximadamente 2,5 Mb, o al menos aproximadamente 2,5 Mb pero menos de aproximadamente 3 Mb.

50 En una modalidad, el locus genómico de interés comprende el locus de la cadena gamma del receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2*, o ambos loci de *Rag1* y de *Rag2*.

En una modalidad, los constructos de expresión primero y segundo están en una única molécula de ácido nucleico.

55 Se proporciona, además, un método para modificar un genoma, que comprende exponer el genoma a una proteína Cas y a un ARN de CRISPR en la presencia de un vector de direccionamiento grande (LTVEC) que comprende una secuencia de ácidos nucleicos de al menos 10 kb, en donde después de la exposición a la proteína Cas, al ARN de CRISPR, y al LTVEC, el genoma se modifica para que contenga al menos una secuencia de ácidos nucleicos de 10 kb.

60 En algunos de tales métodos, el LTVEC comprende una secuencia de ácidos nucleicos de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, o al menos 90 kb. En algunos de tales métodos, el LTVEC comprende una secuencia de ácidos nucleicos de al menos 100 kb, al menos 150 kb, o al menos 200 kb.

65 Se proporciona, además, un método para modificar un genoma, que comprende poner en contacto el genoma con una proteína Cas, un ARN de CRISPR que se hibrida a una secuencia objetivo, y un ARNtracr en la presencia de un vector

- 5 de direccionamiento grande (LTVEC), en donde el LTVEC es de al menos 10 kb y comprende un primer ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde después de ponerse en contacto con la proteína Cas, el ARN de CRISPR, y el ARNtracr en la presencia del LTVEC, el genoma se modifica en un locus genómico de interés para que contenga el primer ácido nucleico. La secuencia objetivo puede estar en o cerca del locus genómico de interés.
- 10 En algunos de tales métodos, el genoma está en una célula eucariota, y la proteína Cas, el ARN de CRISPR, el ARNtracr, y el LTVEC se introducen en la célula eucariota. Algunos de tales métodos comprenden, además, identificar una célula eucariota modificada que comprende una modificación genética dirigida en el locus genómico de interés.
- 15 En algunos de tales métodos, el ARN de CRISPR y el ARNtracr se introducen juntos en la forma de un único ARN guía (ARNg). En otros métodos, el ARN de CRISPR y el ARNtracr se introducen por separado.
- 20 En algunos de tales métodos (a) la proteína Cas se introduce en la célula eucariota en la forma de una proteína, un ARN mensajero (ARNm) que codifica la proteína Cas, o un ADN que codifica la proteína Cas; (b) el ARN de CRISPR se introduce en la célula eucariota en la forma de un ARN o un ADN que codifica el ARN de CRISPR; y (c) el ARNtracr se introduce en la célula eucariota en la forma de un ARN o un ADN que codifica el ARNtracr.
- 25 En algunos métodos (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; (b) el ADN que codifica el ARN de CRISPR está en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica el ARN de CRISPR; y (c) el ADN que codifica el ARNtracr está en la forma de un tercer constructo de expresión que comprende un tercer promotor unido operativamente a un cuarto ácido nucleico que codifica el ARNtracr, en donde los promotores primero, segundo y tercero son activos en la célula eucariota. Opcionalmente, los constructos de expresión primero, segundo y/o tercero están en una única molécula de ácido nucleico.
- 30 En algunos métodos (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; y (b) el ADN que codifica el ARN de CRISPR y el ADN que codifica el ARNtracr están en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica un ARNg que comprende el ARN de CRISPR y el ARNtracr; en donde los promotores primero y segundo son activos en la célula eucariota. Opcionalmente, los constructos de expresión primero y segundo están en una única molécula de ácido nucleico.
- 35 En algunos métodos, la proteína Cas, el ARN de CRISPR, y el ARNtracr se introducen en la célula eucariota como un complejo de proteína-ARN.
- 40 En algunos métodos, la modificación genética dirigida comprende la delección simultánea de una secuencia de ácidos nucleicos endógena en el locus genómico de interés y la inserción del primer ácido nucleico en el locus genómico de interés. En algunos métodos, la secuencia de ácidos nucleicos endógena delecionada es de aproximadamente 30 kb a aproximadamente 110 kb, y el primer ácido nucleico insertado es de aproximadamente 40 kb a aproximadamente 140 kb. En algunos métodos, la secuencia de ácidos nucleicos endógena delecionada es de aproximadamente 38 kb a aproximadamente 110 kb, y el primer ácido nucleico insertado es de aproximadamente 43 kb a aproximadamente 134 kb.
- 45 En algunos métodos, la modificación genética dirigida es una modificación genética bialélica. Opcionalmente, la modificación genética bialélica comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico en el locus genómico de interés en dos cromosomas homólogos.
- 50 En algunos métodos, la célula eucariota modificada es heterocigota compuesta en el locus genómico de interés. En algunos métodos, la célula eucariota modificada es hemicigota en el locus genómico de interés. Opcionalmente, la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico. Opcionalmente, la modificación genética dirigida comprende: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en dos cromosomas homólogos; y (2) inserción del primer ácido nucleico en el locus genómico de interés en un primer cromosoma y alteración del locus genómico de interés en un segundo cromosoma. El primer cromosoma puede ser uno de los dos cromosomas homólogos, y el segundo cromosoma puede ser el otro cromosoma homólogo.
- 55 En algunos métodos, el LTVEC es de al menos 15 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, o al menos 90 kb. Opcionalmente, el LTVEC es de al menos 100 kb, al menos 150 kb, o al menos 200 kb.
- 60 En algunos métodos, el primer ácido nucleico es de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, o al menos 300 kb. En algunos métodos, el primer ácido nucleico es de aproximadamente 40 kb a aproximadamente 140 kb. En algunos métodos, el primer ácido nucleico es de aproximadamente 43 kb a aproximadamente 134 kb.
- 65

En algunos métodos de la descripción, la célula eucariota es una célula de mamífero, un fibroblasto, una célula pluripotente, una célula pluripotente no humana, una célula pluripotente de roedor, una célula madre embrionaria (ES) de ratón o rata, una célula pluripotente humana, una célula madre embrionaria humana (ES), una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, o una célula madre pluripotente inducida (iPS) humana.

En algunos métodos, la proteína Cas es Cas9. En algunos métodos, la secuencia objetivo está flanqueada inmediatamente por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).

En algunos métodos, la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 150 kb. Opcionalmente, la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.

En algunos métodos de la descripción, la modificación genética dirigida comprende: (a) una sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos homóloga u ortóloga; (b) una delección de una secuencia de ácidos nucleicos endógena; (c) una delección de una secuencia de ácidos nucleicos endógena, en donde la delección varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (d) inserción de una secuencia de ácidos nucleicos exógena; (e) inserción de una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb; (f) inserción de una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u ortóloga; (g) inserción de una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humana y una no humana; (h) inserción de un alelo condicional flanqueado por secuencias objetivo de recombinasa específica de sitio; (i) inserción de un gen marcador de selección o un gen reportero unido operativamente a un tercer promotor activo en la célula pluripotente; o (j) una combinación de estos.

En algunos métodos, el locus genómico de interés comprende (i) una secuencia objetivo 5' que es homóloga al brazo de homología 5'; y (ii) una secuencia objetivo 3' que es homóloga al brazo de homología 3'. Opcionalmente, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 3 Mb. Opcionalmente, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 10 kb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero menos de aproximadamente 1 Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 2 Mb pero menos de aproximadamente 2,5 Mb, o al menos aproximadamente 2,5 Mb pero menos de aproximadamente 3 Mb. Opcionalmente, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 110 kb, al menos 120 kb, al menos 130 kb, al menos 140 kb, al menos 150 kb, al menos 160 kb, al menos 170 kb, al menos 180 kb, al menos 190 kb, o al menos 200 kb. En algunos métodos, las secuencias objetivo 5' y 3' están separadas por aproximadamente 30 kb a aproximadamente 110 kb. En algunos métodos, las secuencias objetivo 5' y 3' están separadas por aproximadamente 38 kb a aproximadamente 110 kb.

En algunos métodos, el locus genómico de interés comprende el locus de la cadena gamma de receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2*, o ambos loci de *Rag1* y de *Rag2*. En otros métodos, el locus genómico de interés comprende el locus de *Adamts5*, el locus de *Trpa1*, el locus de *Folh1*, o el locus de *ErbB4*. Aún en otros métodos, el locus genómico de interés comprende el locus de *Lrp5*. Aún en otros métodos, el locus genómico de interés comprende el locus de *C5 (Hc)*, el locus de *Ror1*, o el locus de *Dpp4*.

Se proporciona, además, un método para producir un animal no humano de generación F0 que comprenda una modificación genética dirigida en un locus genómico de interés, el método que comprende: (a) poner en contacto el

5 genoma en una célula ES no humana con una proteína Cas, un ARN de CRISPR, y un ARNtracr en la presencia de un vector de direccionamiento grande (LTVEC) para formar una célula ES no humana modificada, en donde el LTVEC es de al menos 10 kb y comprende un primer ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3'; (b) identificar la célula ES no humana modificada que comprende la modificación genética dirigida en el locus genómico de interés; (c) introducir la célula ES no humana modificada en un embrión huésped no humano; y (d) gestar el embrión huésped no humano en una madre sustituta, en donde la madre sustituta produce el animal no humano de generación F0 que comprende la modificación genética dirigida en el locus genómico de interés.

10 En algunos de tales métodos, el ARN de CRISPR y el ARNtracr se introducen juntos en la forma de un único ARN guía (ARNg). En otros métodos tales, el ARN de CRISPR y el ARNtracr se introducen por separado.

15 En algunos de tales métodos, (a) la proteína Cas se introduce en la célula ES no humana en la forma de una proteína, un ARN mensajero (ARNm) que codifica la proteína Cas, o un ADN que codifica la proteína Cas; (b) el ARN de CRISPR se introduce en la célula ES no humana en la forma de un ARN o un ADN que codifica el ARN de CRISPR; y (c) el ARNtracr se introduce en la célula ES no humana en la forma de un ARN o un ADN que codifica el ARNtracr.

20 En algunos de tales métodos, (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; (b) el ADN que codifica el ARN de CRISPR está en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica el ARN de CRISPR; y (c) el ADN que codifica el ARNtracr está en la forma de un tercer constructo de expresión que comprende un tercer promotor unido operativamente a un cuarto ácido nucleico que codifica el ARNtracr, en donde los promotores primero, segundo y tercero son activos en la célula ES no humana. Opcionalmente, los constructos de expresión primero, segundo y tercero están en una única molécula de ácido nucleico.

25 En algunos de tales métodos, (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; y (b) el ADN que codifica el ARN de CRISPR y el ADN que codifica el ARNtracr están en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica un ARNg que comprende el ARN de CRISPR y el ARNtracr; en donde los promotores primero y segundo son activos en la célula ES no humana. Opcionalmente, los constructos de expresión primero y segundo están en una única molécula de ácido nucleico.

30 En algunos de tales métodos, la proteína Cas, el ARN de CRISPR, y el ARNtracr se introducen en la célula ES no humana como un complejo de proteína-ARN.

35 En algunos de tales métodos, la modificación genética dirigida comprende la delección simultánea de una secuencia de ácidos nucleicos endógena en el locus genómico de interés y la inserción del primer ácido nucleico en el locus genómico de interés.

40 En algunos de tales métodos, la modificación genética dirigida es una modificación genética bialélica. Opcionalmente, la modificación genética bialélica comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico en el locus genómico de interés en dos cromosomas homólogos.

45 En algunos de tales métodos, la célula ES no humana modificada es heterocigota compuesta en el locus genómico de interés. En algunos de tales métodos, la célula ES no humana modificada es hemocigota en el locus genómico de interés. Opcionalmente, la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico. Opcionalmente, la modificación genética dirigida comprende: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en dos cromosomas homólogos; y (2) inserción del primer ácido nucleico en el locus genómico de interés en un primer cromosoma y alteración del locus genómico de interés en un segundo cromosoma. El primer cromosoma puede ser uno de los dos cromosomas homólogos, y el segundo cromosoma puede ser el otro cromosoma homólogo.

50 En algunos de tales métodos, la proteína Cas es Cas9.

55 En esta descripción se proporcionan, además, los métodos para modificar un genoma en un locus genómico de interés en una célula eucariota, una célula de ratón, o una célula humana, que comprende poner en contacto el genoma con una proteína Cas, un ARN de CRISPR que se hibrida a una secuencia objetivo en el locus genómico de interés, y un ARNtracr en la presencia de un vector de direccionamiento grande (LTVEC), en donde el LTVEC es de al menos 10 kb y comprende un primer ácido nucleico flanqueado por un brazo de homología 5' que es homólogo a una secuencia objetivo 5' en el locus genómico de interés y un brazo de homología 3' que es homólogo a una secuencia objetivo 3' en el locus genómico de interés, en donde el primer ácido nucleico es de al menos 30 kb y/o la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 30 kb, en donde después de ponerse en contacto con la proteína Cas, el ARN de CRISPR, y el ARNtracr en la presencia del LTVEC, el genoma se modifica para que comprenda una modificación genética dirigida que comprende la inserción del primer ácido nucleico en el locus genómico de interés.

Cualquiera de los métodos anteriores de la descripción puede comprender, además, introducir la proteína Cas, el ARN de CRISPR, el ARNtracr, y el LTVEC en la célula eucariota, la célula de ratón, o la célula humana. Cualquiera de los métodos anteriores puede comprender, además, identificar la célula eucariota modificada, la célula de ratón modificada, o la célula humana modificada que comprende la modificación genética dirigida en el locus genómico de interés.

5 En algunos de los métodos anteriores, el ARN de CRISPR y el ARNtracr se introducen juntos en la forma de un transcrito único. En algunos de los métodos anteriores, el ARN de CRISPR y el ARNtracr se introducen por separado.

10 En algunos de los métodos anteriores de la descripción, (a) la proteína Cas se introduce en la célula eucariota, la célula de ratón, o la célula humana en la forma de una proteína, un ARN mensajero (ARNm) que codifica la proteína Cas, o un ADN que codifica la proteína Cas; (b) el ARN de CRISPR se introduce en la célula eucariota, la célula de ratón, o la célula humana en la forma de un ARN o un ADN que codifica el ARN de CRISPR; y (c) el ARNtracr se introduce en la célula eucariota, la célula de ratón, o la célula humana en la forma de un ARN o un ADN que codifica el ARNtracr. En algunos de los métodos anteriores, la proteína Cas, el ARN de CRISPR, y el ARNtracr se introducen en la célula eucariota, la célula de ratón, o la célula humana como un complejo de proteína-ARN.

15 En algunos de los métodos anteriores de la descripción, (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; (b) el ADN que codifica el ARN de CRISPR está en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica el ARN de CRISPR; y (c) el ADN que codifica el ARNtracr está en la forma de un tercer constructo de expresión que comprende un tercer promotor unido operativamente a un cuarto ácido nucleico que codifica el ARNtracr; en donde los promotores primero, segundo y tercero son activos en la célula eucariota, la célula de ratón, o la célula humana. En algunos de los métodos anteriores, los constructos de expresión primero, segundo y/o tercero están en una única molécula de ácido nucleico.

20 En algunos de los métodos anteriores de la descripción, (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; y (b) el ADN que codifica el ARN de CRISPR y el ADN que codifica el ARNtracr están en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica un ARNg que comprende el ARN de CRISPR y el ARNtracr en un transcrito único; en donde los promotores primero y segundo son activos en la célula eucariota, la célula de ratón, o la célula humana. En algunos de los métodos anteriores, los constructos de expresión primero y segundo están en una única molécula de ácido nucleico.

25 En algunos de los métodos anteriores, el LTVEC es de al menos 15 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, o al menos 90 kb. En algunos de los métodos anteriores, el LTVEC es de al menos 100 kb, al menos 150 kb, o al menos 200 kb.

30 En algunos de los métodos anteriores, el primer ácido nucleico es de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, o al menos 300 kb. En algunos de los métodos anteriores, el primer ácido nucleico es de aproximadamente 40 kb a aproximadamente 140 kb.

35 En algunos de los métodos anteriores, la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 150 kb. En algunos de los métodos anteriores, la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.

40 En algunos de los métodos anteriores, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 3 Mb. En algunos de los métodos anteriores, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 10 kb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero menos de aproximadamente 1 Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 2 Mb pero menos de aproximadamente 2,5 Mb, o al menos aproximadamente 2,5 Mb pero menos de aproximadamente 3 Mb. En algunos de los métodos anteriores, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 110 kb, al menos 120 kb, al menos 130 kb, al menos 140 kb, al menos 150 kb, al menos 160 kb, al menos 170 kb, al menos 180 kb, al menos 190 kb, o al menos 200 kb. En algunos de los métodos anteriores, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por de aproximadamente 30 kb a aproximadamente 110 kb.

5 En algunos de los métodos anteriores, la célula eucariota no es una célula de rata. En algunos de los métodos anteriores de la descripción, la célula eucariota es una célula pluripotente, una célula no pluripotente, una célula de mamífero, una célula humana, una célula de mamífero no humana, una célula de roedor, una célula de ratón, una célula de hámster, una célula pluripotente no humana, una célula pluripotente humana, una célula pluripotente de roedor, o un fibroblasto. En algunos de los métodos anteriores de la descripción, la célula eucariota es una célula primaria o una célula inmortalizada. En algunos de los métodos anteriores de la descripción, la célula pluripotente de roedor es una célula madre embrionaria (ES) de ratón o rata.

10 En algunos de los métodos anteriores de la descripción, la célula de ratón, o la célula humana es una célula primaria o una célula inmortalizada. En algunos de los métodos anteriores de la descripción, la célula de ratón, o la célula humana es una célula pluripotente. En algunos de los métodos anteriores, la célula pluripotente de ratón es una célula madre embrionaria (ES) de ratón. En algunos de los métodos anteriores de la descripción, la célula pluripotente humana es una célula madre embrionaria humana (ES), una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, o una célula madre pluripotente inducida (iPS) humana. En algunos de los métodos anteriores de la descripción, las células iPS humanas se mantienen en un medio que comprende un medio base y suplementos, en donde el medio comprende: (a) un polipéptido del factor inhibidor de la leucemia (LIF); (b) un inhibidor de la glucógeno sintasa quinasa (GSK3); y (c) un inhibidor de MEK; en donde el medio tiene una osmolalidad de aproximadamente 175 mOsm/kg a aproximadamente 280 mOsm/kg.

20 En algunos de los métodos anteriores, la proteína Cas es Cas9. En algunos de los métodos anteriores, la secuencia objetivo está flanqueada inmediatamente por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).

25 En algunos de los métodos anteriores, la modificación genética dirigida comprende la delección simultánea de una secuencia de ácidos nucleicos endógena en el locus genómico de interés y la inserción del primer ácido nucleico en el locus genómico de interés en una única etapa. En algunos de los métodos anteriores, la secuencia de ácidos nucleicos endógena delecionada es de aproximadamente 30 kb a aproximadamente 110 kb, y el primer ácido nucleico insertado es de aproximadamente 40 kb a aproximadamente 140 kb.

30 En algunos de los métodos anteriores, la modificación genética dirigida es una modificación genética bialélica. En algunos de los métodos anteriores, la modificación genética bialélica comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico en el locus genómico de interés en dos cromosomas homólogos. En algunos de los métodos anteriores, la célula eucariota modificada, la célula de ratón modificada, o la célula humana modificada es heterocigota compuesta en el locus genómico de interés. En algunos de los métodos anteriores, la célula eucariota modificada, la célula de ratón modificada, o la célula humana modificada es hemicigota en el locus genómico de interés. En algunos de los métodos anteriores, la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico. En algunos de los métodos anteriores, la modificación genética dirigida comprende: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en los cromosomas homólogos primero y segundo; y (2) inserción del primer ácido nucleico en el locus genómico de interés en el primer cromosoma homólogo y alteración del locus genómico de interés en el segundo cromosoma homólogo.

45 En algunos de los métodos anteriores descritos en la presente descripción, la modificación genética dirigida comprende: (a) una sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos homóloga u ortóloga; (b) una delección de una secuencia de ácidos nucleicos endógena; (c) una delección de una secuencia de ácidos nucleicos endógena, en donde la delección varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (d) inserción de una secuencia de ácidos nucleicos exógena; (e) inserción de una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb; (f) inserción de una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u ortóloga; (g) inserción de una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humana y no humana; (h) inserción de un alelo condicional flanqueado por secuencias objetivo de recombinasa específica de sitio; (i) inserción de un gen marcador de selección o un gen reportero unido operativamente a un promotor activo en la célula pluripotente; o (j) una combinación de estos.

65

En algunos de los métodos anteriores, el locus genómico de interés comprende el locus de la cadena gamma de receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2*, ambos loci de *Rag1* y de *Rag2*, el locus de *Adamts5*, el locus de *Trpa1*, el locus e *Folh1*, el locus de *ErbB4*, el locus de *Lrp5*, el locus de *C5 (Hc)*, el locus de *Ror1*, o el locus de *Dpp4*. En algunos de los métodos anteriores, el locus genómico de interés comprende ADN extracromosómico.

5

Se proporcionan, además, métodos para producir un animal no humano o ratón de generación F0 que comprenda una modificación genética dirigida en un locus genómico de interés, que comprende: (a) modificar una célula ES no humana o de ratón mediante el uso de cualquiera de los métodos anteriores; (b) identificar la célula ES no humana o de ratón modificada que comprende la modificación genética dirigida en el locus genómico de interés; (c) introducir la célula ES no humana o de ratón modificada en un embrión huésped no humano o de ratón; y (d) gestar el embrión huésped no humano o de ratón en una madre sustituta, en donde la madre sustituta produce el animal no humano o ratón de generación F0 que comprende la modificación genética dirigida en el locus genómico de interés.

10

Breve descripción de las figuras

15

La figura 1 representa las ESC de rata, que crecen como colonias esféricas compactas que se desprenden y flotan de manera rutinaria en el plato.

20

Las figuras 2A a la D representan varios marcadores de pluripotencia expresados por las ESC de rata: A representa Oct-4 (verde); B representa Sox-2 (rojo); C representa DAPI (azul); D representa una superposición de marcadores de pluripotencia expresados por las rESC.

La figura 3 representa que las ESC de rata expresan bajos niveles de fosfatasa alcalina (un marcador de pluripotencia).

25

La figura 4 representa el cariotipo de la línea DA.2B, que es 42X,Y. El cariotipado se hizo porque las ESC de rata frecuentemente se vuelven tetraploides; las líneas por lo tanto se pretamizaron mediante el recuento de propagaciones de cromosomas en metafase, y después las líneas con recuentos principalmente normales se cariotiparon normalmente.

30

Las figuras 5A-B proporcionan fotografías que muestran el análisis del número de cromosomas de la línea de células ES de rata ACI.G1.

Las figuras 6A-B proporcionan fotografías que muestran el análisis del número de cromosomas de la línea de células ES de rata DA.2B.

35

Las figuras 7A-B proporcionan fotografías que muestran el análisis del número de cromosomas de la línea de células ES de rata DA.2C.

La figura 8 representa una vista más cercana de una ESC de rata de la figura 1.

40

La figura 9 representa la producción de quimeras mediante la inyección de blastocistos y la transmisión del genoma de ESC de rata a través de la línea germinal. Las quimeras se produjeron mediante la inyección de blastocistos mediante el uso de las ESC de rata ACI.G1 progenitoras. Las quimeras de alto porcentaje usualmente tienen hocicos albinos.

45

La figura 10 representa crías de agutí F1 con compañeros de camada albinos, engendrados por la quimera ACI/SD marcada con un asterisco (*) en la figura 9.

50

La figura 11 proporciona un esquema del locus de *ApoE* de rata y se indica con barras grises el sitio de corte de las nucleasas de dedos de zinc (ZFN1 y ZFN2). Las regiones genómicas que corresponden a los brazos de homología 5' y 3' (5 kb y 5,4 kb, respectivamente) se indican por los recuadros gris oscuro. El exón 1 del gen de *ApoE* es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los tres intrones del gen de *ApoE* se indican como líneas. Los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco.

55

La figura 12 representa la transformación del locus de *Rosa26* de rata, que se encuentra entre los genes *Setd5* y *Thumpd3* en un ratón, con el mismo espaciado. El panel A muestra la estructura del locus de *Rosa26* de ratón. Los transcritos de *Rosa26* de ratón consisten en 2 o 3 exones. El panel B representa la estructura del locus de *Rosa26* de rata; el locus de rata contiene un segundo exón 1 (Ex1b) además del exón homólogo al exón 1 de ratón (Ex1a); no se ha identificado un tercer exón en rata. El panel C representa un alelo de *Rosa26* de rata transformado; los brazos de homología de 5 kb se clonaron cada uno mediante PCR mediante el uso de ADN genómico de rESC DA; el alelo transformado contiene un casete Aceptor de Corte y Empalme (SA)-lacZ-hUB-neo que sustituye una delección de 117 pb en el intrón de *Rosa26* de rata.

60

La figura 13A representa un cerebro control de una rata de tipo silvestre de 14 semanas de edad, el cual se tiñó con X-gal. El cerebro control mostró un bajo nivel de tinción de fondo para LacZ (vista dorsal).

65

ES 2 700 596 T3

La figura 13B representa la expresión de LacZ en el cerebro de una rata heterocigota para rRosa26 (14 semanas de edad). El reportero lacZ se expresó de manera ubicua en todo el cerebro del heterocigoto para rRosa26.

5 La figura 13C representa un corazón control y el timo (recuadro adicional) de una rata de tipo silvestre de 14 semanas de edad, que se trataron con X-gal. El corazón control y el timo mostraron un bajo nivel de tinción de fondo para LacZ.

10 La figura 13D representa la expresión de LacZ en el corazón y el timo (recuadro adicional) de una rata heterocigota para rRosa26 de 14 semanas de edad. El reportero lacZ se expresó de manera ubicua en todo el corazón y el timo del heterocigoto para rRosa26.

La figura 13E representa un pulmón control de una rata de tipo silvestre de 14 semanas de edad, que se trató con X-gal. El pulmón control mostró un bajo nivel de tinción de fondo para LacZ.

15 La figura 13F representa la expresión de LacZ en el pulmón de una rata heterocigota para rRosa26 de 14 semanas de edad. El reportero lacZ se expresó de manera ubicua en todo el pulmón del heterocigoto para rRosa26.

20 Las figuras 13G a H representan la expresión de LacZ en embriones de rata E12.5. A diferencia del embrión control de tipo silvestre (H), el cual muestra un bajo nivel de tinción de fondo para LacZ, el embrión heterocigoto para rRosa26 mostró expresión ubicua del reportero LacZ en todo el embrión.

Las figuras 13I a J representan la expresión de LacZ en embriones de rata E14.5. A diferencia del embrión control de tipo silvestre (J), el cual muestra un bajo nivel de tinción de fondo para LacZ, el embrión de rata heterocigoto para rRosa26 mostró expresión ubicua del reportero LacZ en todo el embrión.

25 La figura 14 ilustra un evento de recombinación homóloga o no homóloga que ocurre dentro de una célula ES de rata después de una electroporación de un vector de direccionamiento que comprende un casete de selección (casete lacZ-neo).

30 La figura 15 ilustra el mecanismo mediante el cual las endonucleasas de edición del genoma (por ejemplo, ZFN y TALEN) introducen una ruptura de la doble cadena (DSB) en una secuencia genómica objetivo y activan la unión de extremos no homólogos (NHEJ) en una célula ES.

35 La figura 16 ilustra una técnica de transformación de genes que utiliza las ZFN/TALEN para mejorar la eficiencia de la recombinación homóloga de un vector de direccionamiento. DSB representa la ruptura de la doble cadena.

La figura 17 muestra las quimeras ApoE-ZFN-AB5 producidas mediante la producción de quimeras y la transmisión a través de la línea germinal del locus de *ApoE* de rata modificado. La modificación dirigida fue asistida por las nucleasas de dedos de zinc.

40 La figura 18 proporciona un esquema del evento de transformación de IL2r-y en combinación con nucleasas de dedos de zinc que dirige ZFN U y ZFN D. Se muestra la región del locus de IL2r-y de rata transformado por ZFN U y ZFN D (sec. con núm. de ident.: 93). Los sitios de corte de ZFN se indican en la figura.

45 La figura 19 proporciona un esquema del evento de transformación de IL2r-y en combinación con nucleasas de dedos de zinc que es dirigido ZFN U y ZFN D o en combinación con los ARNg (ARNg1, ARNg2, ARNg3, ARNg4). Se muestran las regiones del locus de IL2r-y de rata a transformar por ZFN U y ZFN D o los ARNg 1-4, y se indican los sitios de corte de ZFN.

50 La figura 20 proporciona un esquema del locus de *ApoE* de rata y un plásmido de transformación. El esquema superior muestra la estructura genómica del locus de *ApoE* de rata y las regiones genómicas que corresponden a los brazos de homología 5' y 3' (5 kb y 5,4 kb respectivamente; recuadros gris oscuro). El exón 1 del gen de *ApoE* es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los tres intrones del gen de *ApoE* se indican como líneas. Los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco. El panel inferior muestra el plásmido de transformación. Los brazos de homología 5' y 3' (5 kb y 5,4 kb, respectivamente) se indican por los recuadros gris oscuro. El vector de direccionamiento comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por sitios loxP (flechas en blanco). El casete de autodelección comprende un promotor *Prrm1* de ratón unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

60 La figura 21A proporciona un esquema para la transformación del locus de *ApoE* en las células ES de rata mediante el uso de nucleasas de dedos de zinc y un vector de direccionamiento que comprende un gen reportero (*LacZ*) y un casete de autodelección que comprende un promotor *Prrm1* de ratón unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina. La figura 21B representa un locus de *ApoE* transformado homocigoto.

La figura 22 proporciona un esquema del locus de *ApoE* de rata y un vector de direccionamiento grande (LTVEC). El panel superior muestra la organización genómica del locus de *ApoE* de rata y las regiones genómicas que corresponden a los brazos de homología 5' y 3' (45 kb y 23 kb, respectivamente; los recuadros gris oscuro). El exón 1 de *ApoE* es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los tres intrones del gen de *ApoE* se indican como líneas y los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco. El panel inferior muestra el LTVEC para modificar el locus de *ApoE* de rata. Los brazos de homología 5' y 3' (45 kb y 23 kb, respectivamente) se indican por los recuadros gris oscuro. El LTVEC comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por los sitios loxP (flechas en blanco), que comprende un promotor *Prr1* de ratón unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

La figura 23 proporciona un esquema del locus de *ApoE* de rata y se indican con barras grises los sitios de corte de las nucleasas de dedos de zinc (ZFN1 y ZFN2) usadas junto con el vector de direccionamiento grande (LTVEC) para potenciar la recombinación homóloga entre el vector de direccionamiento y la región cromosómica afín objetivo.

La figura 24 representa el locus de IL2r-y de rata que se ha alterado por una delección de 3,2 kb y la inserción de un gen reportero (eGFP) y un casete de autodelección que comprende un casete de selección por fármacos (hUb-neo) y el gen de Crei unido operativamente a un promotor *Prr1* de ratón.

La figura 25 proporciona otra representación del locus de IL2r-y de rata que se ha alterado por una delección de 3,2 kb y la inserción de un gen reportero (eGFP) y un casete de autodelección que comprende el gen de Crei unido operativamente a un promotor *Prr1* de ratón y un casete de selección por fármacos (hUb-Neo).

La figura 26 proporciona un esquema del locus de *Rag2* de rata y un vector de direccionamiento grande (LTVEC) para modificar el locus de *Rag2* de rata. El panel superior muestra la organización genómica del locus de *Rag2* de rata y las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' (48 kb y 84 kb, respectivamente; recuadros gris oscuro). *Rag2* comprende un único exón indicado por el sombreado gris punteado. El panel inferior es el LTVEC. Los brazos de homología 5' y 3' (48 kb y 84 kb, respectivamente) se indican por los recuadros gris oscuro. El LTVEC comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por los sitios loxP (flechas en blanco) que contiene un promotor *Prr1* de rata unido operativamente al gen de Crei y un casete de selección por fármacos que contiene un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

La figura 27 proporciona la estructura genómica del locus de *Rag1/Rag2* de rata y las regiones genómicas delecionadas ya sea por transformación de *Rag2* (delección de *Rag2*) o por transformación doble de *Rag2/Rag1* (delección de *Rag2/Rag1*).

La figura 28 proporciona un esquema de los loci de *Rag2* y *Rag1* de rata y un vector de direccionamiento grande (LTVEC) usado para modificar los loci. El panel superior muestra la organización genómica de los loci de *Rag1* y *Rag2* y las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' (48 kb y 15 kb, respectivamente; recuadros gris oscuro). Cada uno de *Rag2* y *Rag1* comprende un único exón indicado por el sombreado gris punteado. El panel inferior es el LTVEC. Los brazos de homología 5' y 3' (48 kb y 15 kb, respectivamente) se indican por los recuadros gris oscuro. El LTVEC comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por los sitios loxP (flechas en blanco), que comprende un promotor *Prr1* de rata unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

La figura 29 muestra el análisis por citometría de flujo para la expresión de GFP y de marcador de células T CD3 (paneles A y D), marcador de células B B220 (paneles B y E), y marcador de células NK CD161a (paneles C y F) en células mononucleares de sangre periférica (PBMC) de una rata quimérica IL2rg-/y (paneles A-C) y una rata de tipo silvestre DA (paneles D-F). Las células doble positivas se muestran en el cuadrante R8. La figura 29 muestra que IL2rg-/y PBMC no expresa marcadores de linfocitos maduros.

La figura 30 muestra que los linfocitos positivos a GFP se detectaron en sangre periférica en 2 de las 3 quimeras IL2rg-/y.

La figura 31 proporciona un esquema del locus de *IL2rg* de rata y un plásmido de transformación para la completa humanización del locus de *IL2rg* de rata. El panel superior muestra la organización del locus de *IL2rg* de rata y las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' (4,3 kb y 4,0 kb, respectivamente; recuadros grises). El panel inferior es el plásmido de transformación. Los brazos de homología 5' y 3' (4,3 kb y 4,0 kb, respectivamente) se indican por los recuadros grises. El plásmido de transformación comprende la región genómica de IL-2rg humana y un casete de delección flanqueado por los sitios loxP (flechas en blanco) que contiene un casete de selección por fármacos que contiene un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

La figura 32 proporciona un esquema del locus de *IL2rg* de rata y un plásmido de transformación para la humanización del ectodominio del locus de *IL2rg* de rata. El panel superior muestra la organización genómica del locus de *IL2rg* de rata y las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' (4,3 kb y 4,0 kb, respectivamente; recuadros grises). El panel inferior es el plásmido de transformación. Los brazos de homología 5' y 3' (4,3 kb y 4,0 kb,

- respectivamente) se indican por los recuadros grises. El plásmido de transformación comprende el ectodominio humano de la región genómica de IL-2Rg y un casete de autodelección flanqueado por los sitios loxP (flechas en blanco) que contiene un promotor *Prrm1* de rata unido operativamente al gen de Crei y un casete de selección por fármacos que contiene un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.
- 5 La figura 33 proporciona una alineación de secuencias de la proteína IL-2rg humana (sec. con núm. de ident.: 20; NP_000197.1); la proteína IL-2rg de rata (sec. con núm. de ident.: 21; NP_543165.1); y la proteína IL-2rg quimérica (sec. con núm. de ident.: 22) que comprende el ectodominio humano de IL-2rg fusionado al resto de la proteína IL-2rg de rata. La unión entre la IL-2rg humana y de rata se indica por la línea vertical.
- 10 La figura 34 proporciona un esquema de humanización asistida por CRISPR/Cas9 del gen de *Lrp5* de ratón; el LTVEC se muestra en el panel superior y el locus de *Lrp5* de ratón se muestra en el panel inferior. La región humanizada es el ectodominio. Las flechas indican los sitios objetivo para cada ARNg (gA, gB, gB2, gC, gD, gE2, gE, gF) y ZFN (a-d).
- 15 La figura 35 representa el porcentaje de la eficiencia de transformación de los LTVEC que se dirigen a genes de tamaño creciente para la delección (figura 35A) y los LTVEC con inserciones de genes humanos de tamaño creciente (figura 35B). Los LTVEC se usaron solos (cuadrados o triángulos grises) o en combinación con las ZFN (cuadrados o triángulos negros).
- 20 La figura 36 proporciona un esquema de humanización asistida por CRISPR/Cas9 de la región codificante completa del gen de *Trpa1* de ratón; el LTVEC se muestra el panel superior y el locus de *Trpa1* de ratón se muestra en el panel inferior. Las flechas indican los sitios objetivos para cada ARNg (gA, gA2, gB, gC, gD, gE2, gE, gF).
- 25 La figura 37 proporciona un esquema de humanización asistida por CRISPR/Cas9 del ectodominio (exón 2 al codón de parada) del gen *Folh1* de ratón; el LTVEC se muestra en el panel superior y el locus de *Folh1* de ratón se muestra en el panel inferior. Las flechas indican los sitios objetivos para cada ARNg (gA, gA2, gB, gC, gD, gE, gE2, gF).
- 30 La figura 38 proporciona un esquema de humanización asistida por CRISPR/Cas9 de la región del exón 2 al codón de parada del gen *C5 (Hc)* de ratón; el LTVEC se muestra el panel superior y el locus de *C5 (Hc)* de ratón se muestra en el panel inferior. Las flechas indican los sitios objetivos para cada ARNg (gA, gB, gB2, gC, gD, gE2, gE, gF).
- 35 La figura 39 proporciona un esquema de humanización asistida por CRISPR/Cas9 de la región codificante completa del gen *Adamts5* de ratón; el LTVEC se muestra el panel superior y el locus de *Adamts5* de ratón se muestra en el panel inferior. Las flechas indican los sitios objetivos para cada ARNg (gA, gA2, gB, gC, gD, gE2, gE, gF).
- 40 La figura 40 proporciona un esquema de humanización asistida por CRISPR/Cas9 de los exones 4-15 del gen *ErbB4* de ratón; el LTVEC se muestra en el panel superior y el locus de *ErbB4* de ratón se muestra en el panel inferior. Las flechas indican los sitios objetivos para cada ARNg (gA, gB, gB2, gC, gD, gE2, gE, gF).
- 45 La figura 41 proporciona un esquema de humanización asistida por CRISPR/Cas9 de los exones 2-7 del gen *Ror1* de ratón; el LTVEC se muestra en el panel superior y el locus de *Ror1* de ratón se muestra en el panel inferior. Las flechas indican los sitios objetivos para cada ARNg (gA, gB, gC, gD, gE, gF).
- 50 La figura 42 proporciona un esquema de humanización asistida por CRISPR/Cas9 de la región del exón 2 al codón de parada del gen *Dpp4* de ratón; el LTVEC se muestra en el panel superior y el locus de *Dpp4* de ratón se muestra en el panel inferior. Las flechas indican los sitios objetivos para cada ARNg (gA, gB, gB2, gC, gD, gE2, gE, gF).
- 55 La figura 43 muestra cerebros de ratas hembras de 12 semanas de edad teñidos con X-gal. Las figuras 43A-C muestra un cerebro de una rata de tipo silvestre, y las figuras 43D-F muestra un cerebro de una rata *ApoE^{+/-}*. Las figuras 43A a D muestran las vistas dorsales, las figuras 43B a E muestran las vistas ventrales, y las figuras 43C a F muestran las vistas en primer plano.
- 60 La figura 44 muestra los corazones de ratas hembras de 12 semanas (A y C) y los primeros planos correspondientes de vasos sanguíneos (B y D) teñidos con X-gal. Las figuras 44A a B muestran un corazón y los vasos sanguíneos, respectivamente, de una rata de tipo silvestre, y las figuras 44C a D muestran un corazón y los vasos sanguíneos, respectivamente, de una rata *ApoE^{+/-}*. La tinción estuvo presente en las aurículas of del corazón y en algunos vasos (por ejemplo, la vena cava).
- 65 La figura 45 muestra los hígados de ratas hembras de 12 semanas de edad teñidos con X-gal. Las figuras 45A a B muestran un hígado de una rata de tipo silvestre, y las figuras 45C a D muestran un hígado de una rata *ApoE^{+/-}*. Las figuras 45B a D son los primeros planos de los hígados.
- La figura 46 muestra la detección de los niveles de colesterol, LDL, HDL, y triglicéridos (figuras 46A-D, respectivamente) en ratas transformadas en ApoE homocigotas, ratas transformadas en ApoE heterocigotas, y ratas de tipo silvestre en 6 semanas, 9 semanas, 12 semanas, y 15 semanas.

La figura 47 muestra un esquema del locus de *ApoE* de rata (panel superior) y un vector de direccionamiento grande (LTVEC) que se dirige al locus de *ApoE* de rata (panel inferior). El panel superior muestra la organización genómica del locus de *ApoE* de rata y las regiones genómicas que corresponden a los brazos de homología 5' y 3' (45 kb y 23 kb, respectivamente; los recuadros gris oscuro). El exón 1 de *ApoE* es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los tres intrones del gen de *ApoE* se indican como líneas y los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco. Se indican los sitios objetivo para el ARNg2 (sec. con núm. de ident.: 87) y el ARNg3 de *ApoE* (sec. con núm. de ident.: 88). El panel inferior muestra el LTVEC para modificar el locus de *ApoE* de rata. Los brazos de homología 5' y 3' (45 kb y 23 kb, respectivamente) se indican por los recuadros gris oscuro. El LTVEC comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por los sitios loxP (flechas en blanco), que comprende un promotor *Prr1* de ratón unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

La figura 48 muestra un esquema del locus de *Rag2* de rata (panel superior) y un vector de direccionamiento grande (LTVEC) que se dirige al locus de *Rag2* de rata (panel inferior). El panel superior muestra la organización genómica del locus de *Rag2* de rata y las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' (48 kb y 84 kb, respectivamente; recuadros gris oscuro). *Rag2* comprende un único exón indicado por el sombreado gris punteado. Se indican los sitios objetivo para el ARNg1 (sec. con núm. de ident.: 89) y el ARNg4 (sec. con núm. de ident.: 90) de *Rag2*. El panel inferior es el LTVEC. Los brazos de homología 5' y 3' (48 kb y 84 kb, respectivamente) se indican por los recuadros gris oscuro. El LTVEC comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por los sitios loxP (flechas en blanco) que contiene un promotor *Prr1* de rata unido operativamente al gen de Crei y un casete de selección por fármacos que contiene un promotor de ubiquitina humano unido operativamente a un gen de resistencia a higromicina.

La figura 49 muestra un esquema del locus de *Il2rg* de rata (panel superior) y un plásmido de transformación para la humanización del ectodominio del locus de *Il2rg* de rata (panel inferior). El panel superior muestra la organización genómica del locus de *Il2rg* de rata y las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' (4,3 kb y 4,0 kb, respectivamente; recuadros grises). Se indican los sitios objetivo para el ARNg2 (sec. con núm. de ident.: 91) y el ARNg4 (sec. con núm. de ident.: 92) de *Il2rg*. El panel inferior es el plásmido de transformación. Los brazos de homología 5' y 3' (4,3 kb y 4,0 kb, respectivamente) se indican por los recuadros grises. El plásmido de transformación comprende el ectodominio humano de la región genómica de IL-2Rg y un casete de autodelección flanqueado por los sitios loxP (flechas en blanco) que contiene un promotor *Prr1* de rata unido operativamente al gen de Crei y un casete de selección por fármacos que contiene un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

La figura 50 muestra un esquema de los loci de *Rag2* y *Rag1* de ratas y un vector de direccionamiento grande (LTVEC) usado para modificar los loci en células ES de ratas transformadas con *Il2rg* (clon *Il2rg*-CG12). El panel superior muestra la organización genómica de los loci de *Rag1* y *Rag2* y las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' (48 kb y 15 kb, respectivamente; recuadros grises). Cada *Rag2* y *Rag1* comprenden un único exón indicado por las flechas no sombreadas. El panel inferior es el LTVEC. Los brazos de homología 5' y 3' (48 kb y 15 kb, respectivamente) se indican por los recuadros grises. El LTVEC comprende un gen reportero (eGFP) y un gen de resistencia a puromicina separados por un sitio de entrada al ribosoma interno (IRES) y unidos operativamente a un promotor de actina. El LTVEC comprende, además, un casete de autodelección flanqueado por los sitios loxP (flechas en blanco), que comprende un promotor *Prr1* de rata unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina.

La figura 51 representa un esquema para la sustitución de una porción del locus de *ADAM6* humano con un ácido nucleico que comprende los loci de *Adam6a* de ratón y de *Adam6b* de ratón mediante el uso de un LTVEC y un ARN guía en células iPS humanas. El sitio objetivo del ARN guía se indica por la flecha.

La figura 52A representa la morfología mostrada por las células iPS humanas cultivadas durante 8 días en medio 2i. La figura 52B representa la morfología mostrada por las células iPS humanas cultivadas durante 12 días en medio 2i.

Las figuras 53A-53D representan la morfología de células iPS humanas cultivadas en medio mTeSR™-hLIF o medio VG2i de baja osmolalidad durante 6 días. Las figuras 53A y 53B representan la morfología de células iPS humanas cultivadas en medio mTeSR™-hLIF (figura 3A) o medio VG2i (figura 53B) durante 6 días. Las figuras 53C y 53D representan la morfología de células iPS humanas cultivadas en células alimentadoras de fibroblastos de prepucio humano de recién nacidos (NuFF) en medio mTeSR™-hLIF (figura 53C) o medio VG2i (figura 53D) durante 6 días.

La figura 54A representa células iPS humanas reprogramadas cultivadas en medio VG2i que se tiñeron para fosfatasa alcalina. Las figuras 54B y 54C representan células iPS humanas reprogramadas cultivadas en medio VG2i que se inmunotifieron para la expresión de NANOG.

Las figuras 55A-55C ilustran la disociación enzimática y el subcultivo de células iPS humanas reprogramadas cultivadas en medio VG2i. La figura 55A representa células iPS humanas reprogramadas cultivadas en medio VG2i antes de la

disociación enzimática con tripsina en la ausencia de un inhibidor de ROCK. La figura 55B representa células iPS humanas cultivadas en medio VG2i durante 1 día después del subcultivo. La figura 55C representa células iPS humanas cultivadas en medio VG2i durante 4 días después del subcultivo.

5 Descripción detallada de la invención

Se proporcionan composiciones y métodos para modificar un locus genómico de interés de rata, de eucariota, de eucariota no rata, de mamífero, de mamífero no humano, de humano, de roedor, de roedor no rata, de ratón, o de hámster a través de la recombinación homóloga bacteriana (BHR) en una célula procariota. Se proporcionan composiciones y métodos, además, para modificar genéticamente un locus genómico de interés, por ejemplo, un locus genómico de interés de rata, de eucariota, de eucariota no rata, de mamífero, de mamífero no humano, de humano, de roedor, de roedor no rata, o de ratón mediante el uso de un vector de direccionamiento grande (LTVEC) en combinación con endonucleasas. Se proporcionan composiciones y métodos, además, para producir un animal no humano genéticamente modificado, por ejemplo, una rata, ratón, roedor, o roedor no rata, que comprende una o más modificaciones genéticas dirigidas. Se proporcionan, además, células madre totipotentes o pluripotentes humanas y no humanas aisladas, en particular células madre embrionarias de ratas, que son capaces de mantener la pluripotencia después de una o más modificaciones genéticas en serie *in vitro*, y que son capaces de transmitir las modificaciones genéticas dirigidas a generaciones posteriores en toda la línea germinal.

20 Glosario

El término "célula madre embrionaria" o "célula ES" como se usa en la presente descripción incluye una célula totipotente o pluripotente derivada de embriones que es capaz de contribuir a cualquier tejido del embrión en desarrollo tras la introducción en un embrión. El término "célula pluripotente" como se usa en la presente descripción incluye una célula no diferenciada que posee la capacidad de desarrollarse en más de un tipo de células diferenciadas. El término "célula no pluripotente" incluye las células que no son células pluripotentes.

El término "ácido nucleico homólogo" como se usa en la presente descripción incluye una secuencia de ácidos nucleicos que es ya sea idéntica o sustancialmente similar a una secuencia referencia conocida. En una modalidad, el término "ácido nucleico homólogo" se usa para caracterizar una secuencia que tiene una secuencia de aminoácidos que es al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o incluso 100 % idéntica a una secuencia referencia conocida.

El término "ácido nucleico ortólogo" como se usa en la presente descripción incluye una secuencia de ácidos nucleicos de una especie que es funcionalmente equivalente a una secuencia referencia conocida en otra especie.

El término "vector de direccionamiento grande" o "LTVEC" como se usa en la presente descripción incluye vectores de direccionamiento grandes para célula eucariotas que se derivan de fragmentos de ADN genómico clonado más grandes que los usados típicamente por otros enfoques destinados a realizar la transformación de genes homólogos en células eucariotas. Los ejemplos de LTVEC, incluyen, pero sin limitarse a, cromosoma homólogo bacteriano (BAC) y cromosoma artificial de levadura (YAC).

El término "modificación de alelo" (MOA) como se usa en la presente descripción incluye la modificación de la secuencia de ADN exacta de un alelo de un(os) gene(s) o locus (loci) cromosómico(s) en un genoma. Los ejemplos de "modificación de alelo (MOA)" como los descritos en la presente descripción incluyen, pero sin limitarse a, deleciones, sustituciones, o inserciones de tan solo un único nucleótido o deleciones de muchos kilobases que abarcan un(os) gene(s) o locus (loci) cromosómico(s) de interés, así como también cualquiera de todas las modificaciones posibles entre estos dos extremos.

El término "sitio de recombinación" como se usa en la presente descripción incluye una secuencia de nucleótidos que es reconocida por una recombinasa específica de sitio y que puede servir como un sustrato para un evento de recombinación.

Las modificaciones genéticas "seriada" incluyen dos o más modificaciones conducidas independientemente a una célula (*por ejemplo*, una célula eucariota, una célula eucariota no de rata, una célula de mamífero, una célula humana, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto, o una célula de ovario de hámster chino (CHO)). La primera modificación puede lograrse mediante electroporación, o cualquier otro método conocido en la técnica. Después se hizo una segunda modificación al mismo genoma celular empleando un segundo constructo de ácido nucleico adecuado. La segunda modificación puede lograrse mediante una segunda electroporación, o cualquier otro método conocido en la técnica. En diversas modalidades, después de las modificaciones genéticas primera y segunda de la misma célula, una tercer, una cuarta, una quinta, una sexta, y así sucesivamente, modificaciones genéticas en serie (una tras otras) pueden lograrse mediante el uso, *por ejemplo*, electroporación en serie o cualquier otro método adecuado (seriadamente) conocido en la técnica.

El término "recombinasa específica de sitio" como se usa en la presente descripción incluye un grupo de enzimas que puede facilitar la recombinación entre los "sitios de recombinación" donde los dos sitios de recombinación están separados físicamente dentro de una única molécula de ácido nucleico o en moléculas de ácidos nucleicos separadas. Los ejemplos de "recombinasa específica de sitio" incluyen, pero sin limitarse a, recombinasas Cre, Flp, y Dre.

El término "línea germinal" en referencia a una secuencia de ácidos nucleicos incluye una secuencia de ácidos nucleicos que puede transferirse a la progenie.

La frase "cadena pesada," o "cadena pesada de inmunoglobulina" incluye una secuencia de cadena pesada de inmunoglobulina, que incluye una secuencia de región constante de cadena pesada de inmunoglobulina, de cualquier organismo. Los dominios variables de cadena pesada incluyen tres CDR de cadena pesada y cuatro regiones FR, a menos que se especifique de otra manera. Los fragmentos de cadenas pesadas incluyen CDR, CDR y FR, y combinaciones de estos. Una cadena pesada típica tiene, después del dominio variable (del N-terminal al C-terminal), un dominio C_{H1}, una región bisagra, un dominio C_{H2} y un dominio C_{H3}. Un fragmento funcional de una cadena pesada incluye un fragmento que es capaz de reconocer específicamente a epítipo (por ejemplo, reconocer el epítipo con una K_D en el intervalo micromolar, nanomolar o picomolar), que es capaz de expresarse y secretarse de una célula, y que comprende al menos una CDR. Un dominio variable de cadena pesada está codificado por una secuencia de nucleótidos de región variable, que generalmente comprende los segmentos V_H, D_H, y J_H derivados de un repertorio de segmentos V_H, D_H, y J_H presentes en la línea germinal. Las secuencias, localizaciones y nomenclatura de los segmentos de cadena pesada V, D, y J de varios microorganismos, pueden encontrarse en la base de datos IMGT, a la que se puede acceder a través del internet en la web mundial (www) en la URL "imgt.org."

La frase "cadena ligera" incluye una secuencia de cadena ligera de inmunoglobulina de cualquier organismo, y a menos que se especifique de cualquier otra manera incluye las cadenas ligeras humanas kappa (κ) y lambda (λ) y una VpreB, así como también cadenas ligeras sustitutas. Los dominios variables de cadena ligera incluyen típicamente tres CDR de cadena ligera y cuatro regiones de marco (FR), a menos que se especifique de cualquier otra manera. Generalmente, una cadena ligera en toda su longitud incluye, desde el amino terminal al carboxilo terminal, un dominio variable que incluye FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, y una secuencia de aminoácidos de región constante de cadena ligera. Los dominios variables de cadena ligera está codificado por la secuencia de nucleótidos de región variable de cadena ligera, que generalmente comprende los segmentos génicos V_L de cadena ligera y J_L de cadena ligera, derivados de un repertorio de los segmentos génicos de V y J de cadena ligera presentes en la línea germinal. Las secuencias, localizaciones y nomenclatura de los segmentos génicos de V y J de cadena ligera de varios organismos pueden encontrarse en la base de datos IMGT, a la que puede accederse a través del internet en la red mundial (www) en la URL "imgt.org." Las cadenas ligeras incluyen, *por ejemplo*, las que no se unen selectivamente ya sea a un primer o a un segundo epítipo unido selectivamente a la proteína de unión a epítipo en la que aparecen. Las cadenas ligeras incluyen, además, las que se unen y reconocen, o ayudan a la cadena pesada con la unión y el reconocimiento a, uno o más epítopos unidos selectivamente a la proteína de unión a epítipo en la que aparecen.

La frase "unido operativamente" comprende una relación en donde los componentes unidos operativamente funcionan de manera prevista. En un caso, una secuencia de ácidos nucleicos que codifica una proteína puede unirse operativamente a secuencias reguladoras (por ejemplo, secuencia promotora, potenciadora, silenciadora, etc.) para mantener la regulación transcripcional apropiada. En un caso, una secuencia de ácidos nucleicos de una región variable de inmunoglobulina (o segmentos V(D)J) puede unirse operativamente a una secuencia de ácidos nucleicos de una región constante de inmunoglobulina para permitir la recombinación apropiada entre las secuencias en una secuencia de cadena pesada o ligera de inmunoglobulina.

1. Locus objetivo que comprende un ácido nucleico

Se proporcionan varios métodos y composiciones, que permiten la integración de al menos un inserto de ácido nucleico en un locus objetivo. Como se usa en la presente descripción, un "locus genómico de interés" comprende cualquier segmento o región de ADN dentro del genoma al que se desea integrar un inserto de ácido nucleico. Los términos "locus genómico de interés" y "locus genómico de interés objetivo" pueden usarse indistintamente. El locus genómico de interés puede ser nativo para la célula, o alternativamente puede comprender un segmento de ADN heterólogo o exógeno que se integró en el genoma de la célula. Tales segmentos de ADN heterólogos o exógenos pueden incluir transgenes, casetes de expresión, polinucleótido que codifica marcadores de selección, o regiones heterólogas o exógenas de ADN genómico. El término "locus" se define en la presente descripción como un segmento de ADN dentro del ADN genómico. Las modificaciones genéticas como las descritas en la presente descripción pueden incluir una o más delecciones de un locus de interés, adiciones a un locus de interés, sustitución de un locus de interés, y/o cualquier combinación de estos. El locus de interés puede comprender regiones codificantes o regiones reguladoras no codificantes.

El locus genómico de interés puede comprender, además, cualquier componente de un sistema de integración dirigida que incluye, por ejemplo, un sitio de reconocimiento, un marcador de selección, un inserto de ácido nucleico integrado anteriormente, polinucleótidos que codifica agentes nucleasas, promotores, etc. Alternativamente, el locus genómico de interés puede localizarse dentro de un ADN extracromosómico dentro de la célula, tal como un cromosoma artificial de levadura (YAC), un cromosoma artificial bacteriano (BAC), un cromosoma artificial humano, o cualquier otra región genómica modificada genéticamente contenida en una célula huésped apropiada. La descripción describe que el locus

5 objetivo puede comprender la secuencia de ácidos nucleicos nativa, heteróloga o exógena de un procarionta, un eucariota, un eucariota no rata, levadura, bacteria, un mamífero no humano, una célula no humana, un roedor, un roedor no rata, un ser humano, una rata, un ratón, un hámster, un conejo, un cerdo, un bovino, un ciervo, una oveja, una cabra, una gallina, un gato, un perro, un hurón, un primate (*por ejemplo*, títí, mono rhesus), mamífero domesticado o un mamífero agrícola o cualquier otro organismo de interés o una combinación de estos. La descripción describe que el locus genómico de interés comprende una secuencia de ácidos nucleicos de un ser humano, un ratón, o una combinación de estos.

10 La descripción describe que el locus objetivo es de, por ejemplo, una célula eucariota, una célula eucariota no de rata, una célula de mamífero, célula humana, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto, o una célula CHO.

15 La descripción describe que el locus genómico de interés comprende un locus objetivo de un "ácido nucleico de rata." Tal región comprende un ácido nucleico de una rata que se integra dentro del genoma de una célula. Los ejemplos no limitantes del locus objetivo incluyen un locus genómico que codifica una proteína expresada en una célula B, un locus genómico que expresa un polipéptido en una célula B inmadura, un locus genómico que expresa un polipéptido en una célula B madura, unos loci de inmunoglobulina (Ig), o unos loci de receptor de células T, incluidos, por ejemplo, un locus alfa de receptor de células T. Los ejemplos adicionales de locus genómico objetivo incluyen un locus de *Fcgr1a*, un locus de *Tlr4*, un locus de *Prlr*, un locus de *Notch4*, un locus de *Accn2*, un locus de *Adamts5*, un locus de *Trpa1*, un locus de *Folh1*, un locus de *Lrp5*, un locus de receptor de IL2, incluido, por ejemplo, un locus de Receptor gamma de IL2 (*Il2rg*), un locus de *ApoE*, un locus de *Rag1*, un locus de *Rag2*, un locus de *Rag1/Rag2*, y un locus de *ErbB4*. Cualquiera de estos locus objetivo puede ser de una rata o pueden ser de una célula eucariota, una célula eucariota no de rata, una célula de mamífero, una célula humana, o una célula de mamífero no humana.

30 La descripción describe que, el locus objetivo codifica una secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina de mamífero. La descripción describe que el locus objetivo codifica una secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina de rata. La descripción describe que el locus objetivo comprende una secuencia de ADN genómico que comprende una secuencia de ácidos nucleicos de región variable de cadena pesada de inmunoglobulina de rata, ratón o humana no reordenada unida operativamente a una secuencia de ácidos nucleicos de región constante de cadena pesada de inmunoglobulina. La descripción describe que la secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina es una secuencia de ácidos nucleicos de región constante de cadena pesada de inmunoglobulina de rata, ratón o humana seleccionada de una CH1, una región bisagra, una CH2, una CH3, y una combinación de estos. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena pesada comprende una CH1-región bisagra-CH2-CH3. La descripción describe que el locus objetivo comprende una secuencia de ácidos nucleicos de región variable de cadena pesada de inmunoglobulina de rata, ratón o humana no reordenada unida operativamente a una secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina. La descripción describe que la secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina es una secuencia de ácidos nucleicos de región constante de cadena pesada de inmunoglobulina de rata, ratón o humana seleccionada de una CH1, una región bisagra, una CH2, una CH3, y una combinación de estos. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena pesada comprende una CH1-región bisagra-CH2-CH3.

45 En una modalidad, el locus objetivo comprende una secuencia de ADN genómico que codifica una secuencia de aminoácidos de región variable de cadena ligera de inmunoglobulina de mamífero. En una modalidad, la secuencia de ADN genómico comprende una secuencia de ácidos nucleicos de región variable de cadena ligera λ y/o κ de mamífero.

50 En una modalidad, la secuencia de ADN genómico comprende una secuencia de ácidos nucleicos de región variable de cadena ligera λ y/o κ de mamífero reordenada. En una modalidad, la secuencia de ácidos nucleicos de región variable de cadena ligera λ o κ no reordenada se une operativamente a una secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina de mamífero seleccionada de una secuencia de ácidos nucleicos de región constante de cadena ligera λ y una secuencia de ácidos nucleicos de región constante de cadena ligera κ . La descripción describe que la secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina de mamífero es una secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina de rata. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina de mamífero es una secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina de ratón. La descripción describe que la secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina de mamífero es una secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina humana.

60 Como se usa en la presente descripción, un locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2 (*Il2rg*), un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* comprenden las regiones respectivas del genoma (es decir, un genoma de mamífero, un genoma humano o un genoma no humano de mamífero) en las cuales se localizan cada uno de estos genes o combinaciones de genes. La modificación de cualquiera del locus de *ApoE*, locus de la cadena gamma de receptor de interleucina-2 (*Il2rg*), locus de *Rag2*, locus de *Rag1* y/o locus de *Rag2/Rag1* (es decir, un locus de mamífero, de un ser humano, o uno de mamífero no humano *ApoE*, el locus de la cadena gamma de receptor de

interleucina-2, el locus de *Rag2*, el locus de *Rag1* y/o el locus de *Rag2/Rag1* combinado) puede comprender cualquier alteración deseada al locus dado. Los ejemplos no limitantes de modificación al locus dado (es decir, un locus de mamífero, de un ser humano, o de un mamífero no humano locus) se analizan en mayor detalle en la presente descripción.

5 Por ejemplo, en la presente descripción, se modifica uno o más del locus de *ApoE*, locus de la cadena gamma de receptor de interleucina-2 (*IL2rg*), locus de *Rag2*, locus de *Rag1* y/o locus de *Rag2/Rag1* (es decir, un locus de mamífero, de un ser humano, o uno de mamífero no humano *ApoE*, un locus de la cadena gamma de receptor de interleucina-2 de mamífero, de un ser humano, o uno de mamífero no humano, un locus de *Rag2* de un de mamífero, de un ser humano, o uno de mamífero no humano, y/o el locus de *Rag2/Rag1*) de manera que disminuya la actividad y/o el nivel de la proteína ApoE codificada o la proteína la cadena gamma de receptor de interleucina-2 o la proteína Rag1 o la proteína Rag2 o una combinación de las proteínas Rag1 y Rag2. En otras modalidades, está ausente la actividad de la proteína ApoE, la proteína la cadena gamma de receptor de interleucina-2, la proteína Rag1, o la proteína Rag2, o una combinación de las proteínas Rag1 y Rag2.

15 Por "disminuido" se pretende cualquier disminución en el nivel o actividad del gen/proteína codificado en el locus de interés. Por ejemplo, una disminución en la actividad puede comprender ya sea (1) una disminución estadísticamente significativa en el nivel total de una proteína dada (es decir, ApoE, la cadena gamma de receptor de interleucina-2, Rag2, Rag2 o una combinación de Rag1 y Rag2) que incluye, por ejemplo, una disminución del nivel o actividad de 0,5 %, 1 %, 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 100 %, 120 % o más cuando se compara con un control apropiado. Los métodos para analizar una disminución en la concentración y/o la actividad de cualquiera de ApoE, la cadena gamma de receptor de interleucina-2, Rag1 y Rag2 se conocen en la técnica.

25 La descripción describe que, uno o más del locus de *ApoE* de un mamífero, de un ser humano, o uno de un mamífero no humano, el locus de la cadena gamma de receptor de interleucina-2 de un mamífero, de un ser humano, o uno de un mamífero no humano, locus de *Rag2* de un mamífero, de un ser humano, o uno de un mamífero no humano, locus de *Rag1* de un mamífero, de un ser humano, o uno de un mamífero no humano y/o locus de *Rag2/Rag1* de un mamífero, de un ser humano, o uno de un mamífero no humano comprenden una modificación de manera que aumenta la actividad y/o el nivel del polipéptido de ApoE codificado, el polipéptido de la cadena gamma de receptor de interleucina-2, el polipéptido de Rag2, el polipéptido de Rag1, o tanto el polipéptido de Rag1 y Rag2. Por "aumentado" se pretende cualquier aumento en el nivel o actividad del gen/polipéptido codificado en el locus de interés. Por ejemplo, un aumento en la actividad puede comprender ya sea (1) un aumento estadísticamente significativo en el nivel o actividad global de una proteína dada (es decir, ApoE, la cadena gamma de receptor de interleucina-2, Rag1, Rag2 o Rag1 y Rag2) que incluye, por ejemplo, un aumento en el nivel o actividad de 0,5 %, 1 %, 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 100 %, 120 % o más cuando se compara con un control apropiado. Los métodos para analizar un aumento en la concentración y/o la actividad de cualquiera de las proteínas ApoE, Rag1, Rag2 y la cadena gamma de receptor de interleucina-2 se conocen en la técnica.

35 La modificación genética al locus de *ApoE* de un mamífero, de un ser humano, o uno de un mamífero no humano, un locus de la cadena gamma de receptor de interleucina-2 de un mamífero, de un ser humano, o uno de un mamífero no humano, locus de *Rag2* de un mamífero, de un ser humano, o uno de un mamífero no humano, locus de *Rag1* de un mamífero, de un ser humano, o uno de un mamífero no humano y/o locus de *Rag2/Rag1* de un mamífero, de un ser humano, o uno de un mamífero no humano puede comprender una delección de una secuencia de ácidos nucleicos endógena en el locus genómico, una inserción de un ácido nucleico exógeno en el locus genómico, o una combinación de estos. La delección y/o inserción puede producirse en cualquier lugar dentro del locus dado como se analiza en otra parte en la presente descripción.

45 La descripción describe que la modificación de uno o más del locus de *ApoE* de mamífero, humano, o de mamífero no humano, locus de la cadena gamma de receptor de interleucina-2, locus de *Rag2*, locus de *Rag1* y/o locus de *Rag2/Rag1* mediante la sustitución de una porción del locus de *ApoE*, el locus de la cadena gamma de receptor de interleucina-2 (*IL2rg*), locus de *Rag2*, locus de *Rag1* y/o locus de *Rag2/Rag1* con la porción homóloga u ortóloga correspondiente de un locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2, un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* de otro organismo.

50 La descripción describe que la modificación of uno o más del locus de *ApoE* de mamífero, humano, o de mamífero no humano, el locus de la cadena gamma de receptor de interleucina-2, locus de *Rag2*, el locus de *Rag1*, y/o el locus de *Rag2/Rag1* se lleva a cabo mediante la sustitución de una porción del locus de *ApoE*, locus de la cadena gamma de receptor de interleucina-2 (*IL2rg*), locus de *Rag2*, locus de *Rag1* y/o locus de *Rag2/Rag1* con un inserto de nucleótido que comparte en toda su longitud al menos 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % a una porción de un locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2, un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* que reemplaza.

60 El inserto de polinucleótido dado y/o la región correspondiente del locus que se deleciona puede ser una región codificante, un intrón, un exón, una región no traducida, una región reguladora, un promotor, o un potenciador o cualquier combinación de estos o cualquier porción de estos. Además, el inserto de polinucleótido dado y/o la región del locus, por ejemplo, que se deleciona puede ser de cualquier longitud deseada, que incluye, por ejemplo, entre 10-100 nucleótidos de longitud, 100-500 nucleótidos de longitud, 500-1 kb de nucleótidos de longitud, 1 Kb a 1,5 kb de nucleótidos de longitud, 1,5 kb a 2

- kb nucleótidos de longitud, 2 kb a 2,5 kb nucleótidos de longitud, 2,5 kb a 3 kb nucleótidos de longitud, 3 kb a 5 kb nucleótidos de longitud, 5 kb a 8 kb nucleótidos de longitud, 8 kb a 10 kb nucleótidos de longitud o más. En otros casos, el tamaño de la inserción o la sustitución es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, de aproximadamente 350 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 800 kb, de aproximadamente 800 kb a 1 Mb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, de aproximadamente 2,5 Mb a aproximadamente 2,8 Mb, de aproximadamente 2,8 Mb a aproximadamente 3 Mb. En otras modalidades, el inserto de polinucleótido dado y/o la región del locus que se deleciona es de al menos 100, 200, 300, 400, 500, 600, 700, 800, o 900 nucleótidos o al menos 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 15 kb, 16 kb o más. En otras modalidades, el inserto de polinucleótido dado y/o la región del locus que se deleciona es de al menos 10 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, o al menos 300 kb o más.
- El inserto de polinucleótido dado puede ser de cualquier organismo, que incluye, por ejemplo, un roedor, un roedor no rata, una rata, un ratón, un hámster, un mamífero, un mamífero no humano, un eucariota, un eucariota no rata, un humano, un animal agrícola o un animal doméstico.
- Como se analiza en mayor detalle en la presente descripción, se proporcionan varios métodos para generar modificaciones objetivo de cualquier locus de interés, que incluye, por ejemplo, modificaciones objetivo en el locus de *ApoE*, locus de la cadena gamma de receptor de interleucina-2 (*Il2rg*), locus de *Rag2*, locus de *Rag1* y/o locus de *Rag2/Rag1*. En la presente descripción se proporcionan, además, animales no humanos genéticamente modificados, mamíferos no humanos genéticamente modificados, eucariotas no ratas genéticamente modificados, células no pluripotentes genéticamente modificadas, o células pluripotentes genéticamente modificadas (por ejemplo, una célula pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, o una célula iPS humana), que comprende una delección, una inserción, una sustitución y/o cualquier combinación de estos en el locus de la cadena gamma de receptor de interleucina-2, en el locus de *ApoE*, en el locus de *Rag2*, en el locus de *Rag1*, y/o en el locus de *Rag2/Rag1*. Tales modificaciones genéticas (incluidas aquellas que dan como resultado una ausencia, una disminución, un aumento o una modulación en la actividad del locus objetivo) y además son capaces de transmitirse a través de la línea germinal. En modalidades específicas, las modificaciones genéticas dan como resultado una inactivación del locus objetivo deseado. Tales animales no humanos, por ejemplo, encuentran uso en una variedad de sistemas experimentales como se analiza en cualquier otra parte de la presente descripción.
- Por ejemplo, las inactivaciones de *ApoE* (Apolipoproteína E) ofrecen un modelo animal para estudiar la función endotelial, que incluye, pero sin limitarse a, la formación de placa, cambios transcripcionales (Secuenciación por fragmentación aleatoria del transcriptoma completo (ARN-Seq), y la función *ex vivo*. *ApoE* es una molécula transportadora importante y puede transportar lípidos, tales como colesterol, a través del torrente sanguíneo. *ApoE* puede funcionar, además, en el sistema nervioso, por ejemplo, para eliminar β -amiloides del cerebro. Las modificaciones en *ApoE* se han implicado en varias afecciones, que incluyen, por ejemplo, aterosclerosis, hiperlipidemia, y enfermedad de Alzheimer. *Los animales con ApoE inactivado* muestran una eliminación deficiente de lipoproteínas de la sangre y desarrollan aterosclerosis. Por lo tanto, los animales con *ApoE* inactivado proporcionan un modelo para estudiar afecciones y/o procesos tales como, por ejemplo, función endotelial, la formación de placa, cambios transcripcionales (ARN-Seq), hiperlipidemia, aterosclerosis y enfermedad de Alzheimer. Los ensayos para medir la actividad de *ApoE* se conocen en la técnica. Por ejemplo, una disminución en la actividad de *ApoE* puede medirse mediante el análisis de una disminución en los niveles de *ApoE* en una muestra de sangre obtenida de un sujeto mediante inmunoensayos, tales como mediante ELISA o mediante técnicas de Inmunotransferencia. Además, el gran tamaño de las ratas facilita todos estos ensayos y mejora la calidad de los datos.
- RAG1 (Gen de Activación de la Recombinación 1) y RAG2 (Gen de Activación de la Recombinación 2) son enzimas que son parte de un complejo de múltiples subunidades y juegan un papel importante en el reordenamiento y la recombinación de genes de inmunoglobulinas y receptores de células T en linfocitos. RAG1 y RAG2 inducen una escisión del ADN bicatenario para facilitar la recombinación y unión de segmentos de los genes del receptor de células T y del receptor de células B (es decir, inmunoglobulina). La inactivación de RAG1 y/o RAG2 provoca una pérdida de células B y de células T en el animal que da como resultado una inmunodeficiencia grave. Los animales con RAG1 y/o RAG2 inactivados encuentran uso, por ejemplo, en estudios de xenoinjertos (es decir, xenoinjertos de células humanas en ratas), cáncer, desarrollo de vacunas, enfermedad autoinmunitaria, enfermedad infecciosa y enfermedad de huésped contra injerto (GVHD). Se conocen en la técnica diversos ensayos para medir la actividad de RAG1 y/o RAG2 e incluyen, por ejemplo, medir la eficiencia de recombinación o analizar la presencia o ausencia de células B y/o células T en un sujeto.
- El receptor de IL-2 (IL-2R) se expresa en la superficie de determinadas células inmunitarias y se une a la citocina interleucina-2 (IL-2). La IL-2R es una proteína integral de membrana que comprende al menos tres cadenas de

subunidades separadas, que incluye, una cadena alfa (IL-2Ra, CD25), una cadena beta (IL-2Rb, CD122) y una cadena gamma (IL2-Rg, CD132). La cadena de receptor de IL-2 gamma (también denominado como IL2r-y o IL2Rg) es una cadena gamma común que es compartida por varios receptores de citocina, que incluyen, por ejemplo, los receptores de IL-2, IL-4, IL-7, IL-9, IL-15 e IL-21. IL-2Rg comprende un ectodominio en la superficie extracelular de la célula, que contribuye a la unión del ligando, un dominio transmembrana, y un dominio intracelular, que puede interactuar con varias moléculas para inducir las vías intracelulares de transducción de señales. El gen *IL2rg* se encuentra en el cromosoma X en mamíferos y determinadas mutaciones en el gen de la cadena gamma en humanos puede provocar inmunodeficiencia combinada grave ligada al (XSCID) X humana caracterizada por un profundo defecto en las células T. Además, el ectodominio de la cadena gamma puede desprenderse del receptor de transmembrana y liberarse como un receptor de cadena gamma soluble. El receptor de cadena gamma soluble puede detectarse en la sangre de un sujeto y puede funcionar para regular la señalización de citocina.

En algunas modalidades, la cadena de IL-2Rg no humana se sustituye con la cadena de IL2-Rg humana de manera que el animal genéticamente modificado exprese una cadena de IL-2Rg completamente humana. En otros casos, puede ser útil sustituir sólo el ectodominio de una cadena de IL-2Rg no humana con el ectodominio de la cadena de IL-2Rg humana. En tales casos, la cadena de IL-2Rg humanizada expresada en un no humano comprende un ectodominio humano, con el resto de la molécula que es del organismo nativo.

La humanización en toda su longitud de IL-2Rg es útil porque los mamíferos no humanos que tienen este locus modificado producirán el IL-2Rg humano. Esto permitirá la detección de IL-2Rg humano en mamíferos no humanos con anticuerpos específicos a IL-2Rg humano. La humanización de ecto (es decir, sustitución del ecto-dominio de IL-2Rg de un mamífero no humano con el ectodominio humano de IL-2Rg) dará como resultado un polipéptido de IL-2Rg que se unirá a los ligandos humanos de IL2-Rg, pero porque el dominio citoplasmático todavía es del mamífero no humano, la forma humanizada de ecto de IL-2Rg interactuará, además, con la maquinaria de señalización del mamífero no humano.

2. Modificación de un locus objetivo

A. Vectores de transformación e insertos de ácido nucleico

i. Inserto de ácido nucleico

Como se usa en la presente descripción, el "inserto de ácido nucleico" comprende un segmento de ADN que se desea integrar en el locus objetivo. En una modalidad, el inserto de ácido nucleico comprende uno o más polinucleótidos de interés. En otras modalidades, el inserto de ácido nucleico puede comprender uno o más casetes de expresión. Un casete de expresión dado puede comprender un polinucleótido de interés, un polinucleótido que codifica un marcador de selección y/o un gen reportero junto con varios componentes reguladores que influyen en la expresión. Los ejemplos no limitantes de polinucleótidos de interés, marcadores de selección, y genes reporteros que pueden incluirse dentro del inserto de ácido nucleico se analizan en detalle en cualquier otra parte de la presente descripción.

En modalidades específicas, el inserto de ácido nucleico puede comprender un ácido nucleico de rata, que puede incluir un segmento de ADN genómico, un ADNc, una región reguladora, o cualquier porción o combinación de estos. En otras modalidades, el inserto de ácido nucleico puede comprender un ácido nucleico de un eucariota, un eucariota no rata, un mamífero, un ser humano, un mamífero no humano, un roedor, un roedor no rata, un ser humano, una rata, un ratón, un hámster, un conejo, un cerdo, un bovino, un ciervo, una oveja, una cabra, una gallina, un gato, un perro, un hurón, un primate (*por ejemplo*, tití, mono rhesus), un mamífero domesticado, o un mamífero agrícola o cualquier otro organismo de interés. Como se expone en mayor detalle en la presente descripción, el inserto de ácido nucleico empleado en los diversos métodos y composiciones puede dar como resultado la "humanización" de un locus objetivo de interés.

En una modalidad, el inserto de ácido nucleico comprende un alelo de inserción de secuencias génicas (knock-in) de al menos un exón de un gen endógeno. En una modalidad, el inserto de ácido nucleico comprende un alelo de inserción de secuencias génicas de todo el gen endógeno (es decir, "inserción de secuencias génicas por intercambio de genes").

En una modalidad, el inserto de ácido nucleico comprende un elemento regulador, que incluye, por ejemplo, un promotor, un potenciador, o un elemento de unión a represor transcripcional.

En modalidades adicionales, el inserto de ácido nucleico comprende un alelo condicional. En una modalidad, el alelo condicional es un alelo multifuncional, como se describe en el documento US 2011/0104799. En modalidades específicas, el alelo condicional comprende: (a) una secuencia de accionamiento en orientación con sentido con respecto a la transcripción de un gen objetivo, y un casete de selección por fármacos en orientación con sentido o antisentido; (b) en orientación antisentido una secuencia de nucleótidos de interés (NSI) y un módulo condicional por inversión (COIN, que utiliza un intrón de división de exón y un módulo tipo genetráp invertible; ver, por ejemplo, el documento US 2011/0104799; y (c) las unidades recombinables que se recombinan tras la exposición a una primera recombinasa para formar un alelo condicional que (i) carece de la secuencia de accionamiento y el DSC, y (ii) contiene la NSI en orientación con sentido y el COIN en orientación antisentido.

5 La descripción describe que el inserto de ácido nucleico varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.

10 Esta descripción describe que el inserto de ácido nucleico comprende una delección de, por ejemplo, una secuencia de ADN genómico de una célula eucariota, una célula eucariota no de rata, una célula de mamífero, una célula humana o una célula de mamífero no humana en el intervalo de aproximadamente 1 kb a aproximadamente 200 kb, de aproximadamente 2 kb a aproximadamente 20 kb, o de aproximadamente 0,5 kb a aproximadamente 3 Mb. La amplitud de la delección de la secuencia de ADN genómico es mayor que una longitud total del brazo de homología 5' y el brazo de homología 3'. En una modalidad, la amplitud de la delección de la secuencia de ADN genómico varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 70 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, de aproximadamente 190 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, de aproximadamente 350 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 800 kb, de aproximadamente 800 kb a 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb, a aproximadamente 2,5 Mb, de aproximadamente 2,5 Mb a aproximadamente 2,8 Mb, de aproximadamente 2,8 Mb a aproximadamente 3 Mb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb.

40 En una modalidad, el inserto de ácido nucleico comprende una inserción o una sustitución de una secuencia de ácidos nucleicos de un eucariota, un eucariota no rata, un mamífero, un ser humano o un mamífero no humano con una secuencia de ácidos nucleicos humana homóloga u ortóloga. En una modalidad, el inserto de ácido nucleico comprende una inserción o sustitución de una secuencia de ADN con una secuencia de ácidos nucleicos humana homóloga u ortóloga en un locus endógeno que comprende la secuencia de ADN correspondiente.

45 En una modalidad, la modificación genética es una adición de una secuencia de ácidos nucleicos. Esta descripción describe que la secuencia de nucleótidos añadida varía de 5 kb a 200 kb.

50 En una modalidad, el inserto de ácido nucleico comprende una modificación genética en una secuencia codificante. En una modalidad, la modificación genética comprende una mutación por delección de una secuencia codificante. En una modalidad, la modificación genética comprende una fusión de dos secuencias codificantes endógenas.

55 Esta descripción describe que el inserto de ácido nucleico comprende una inserción o una sustitución de una secuencia de ácidos nucleicos de un eucariota, un eucariota no rata, un mamífero, un ser humano, o un mamífero no humano con una secuencia de ácidos nucleicos humana homóloga u ortóloga. Esta descripción describe que el inserto de ácido nucleico comprende una inserción o la sustitución de una secuencia de ADN de rata con una secuencia de ácidos nucleicos humana homóloga u ortóloga en un locus de rata endógeno que comprende la secuencia de ADN de rata correspondiente.

60 En una modalidad, la modificación genética comprende una delección de una secuencia no codificante de proteínas, pero no comprende una delección de una secuencia codificante de proteínas. En una modalidad, la delección de la secuencia no codificante de proteínas comprende una delección de un elemento regulador. En una modalidad, la modificación genética comprende una delección de un promotor. En una modalidad, la modificación genética comprende una adición de un promotor o un elemento regulador. En una modalidad, la modificación genética comprende una sustitución de un promotor o un elemento regulador.

65 Esta descripción describe que la secuencia de ácidos nucleicos del vector de direccionamiento puede comprender un polinucleótido que cuando se integra en el genoma producirá una modificación genética de una región del locus de *ApoE*

de mamífero, ser humano, o un mamífero no humano, en donde la modificación genética en el locus de *ApoE* da como resultado una disminución de la actividad de ApoE, un aumento de la actividad de ApoE, o una modulación de la actividad de ApoE. En una modalidad, se genera una inactivación de *ApoE* ("alelo nulo").

5 Esta descripción describe que la secuencia de ácidos nucleicos del vector de direccionamiento puede comprender un polinucleótido que cuando se integra en el genoma producirá una modificación genética de una región del locus de receptor de interleucina 2 de mamífero, célula humana, o mamífero no humano, en donde la modificación genética en el locus del receptor de interleucina 2 da como resultado una disminución de la actividad del receptor de interleucina 2. En una modalidad, se genera una inactivación del receptor de interleucina 2 ("alelo nulo").

10 La descripción describe que el inserto de ácido nucleico da como resultado la sustitución de una porción del locus de *ApoE* de mamífero, célula humana, o mamífero no humano, el locus de la cadena gamma de receptor de interleucina-2 y/o locus de *Rag2*, y/o locus de *Rag1* y/o locus de *Rag2/Rag1* con la porción homóloga u ortóloga correspondiente de un locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2, un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* de otro organismo.

15 Aún en otras modalidades, el inserto de ácido nucleico comprende un polinucleótido que comparte en toda su longitud al menos el 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % con una porción de un locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2, un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* que se sustituye.

20 Esta descripción describe que el inserto de polinucleótido dado y la región correspondiente del locus de mamífero, de célula humana, de mamífero no humano que se sustituye puede ser una región codificante, un intrón, un exón, una región no traducida, una región reguladora, un promotor, o un potenciador o cualquier combinación de estos. Además, el inserto de polinucleótido dado y/o la región del locus de mamífero, de célula humana, o de mamífero no humano que se deleciona puede ser de cualquier longitud deseada, que incluye por ejemplo, entre 10-100 nucleótidos de longitud, 100-500 nucleótidos de longitud, 500-1 kb nucleótidos de longitud, 1 Kb a 1,5 kb nucleótidos de longitud, 1,5 kb a 2 kb nucleótidos de longitud, 2 kb a 2,5 kb nucleótidos de longitud, 2,5 kb a 3 kb nucleótidos de longitud, 3 kb a 5 kb nucleótidos de longitud, 5 kb a 8 kb nucleótidos de longitud, 8 kb a 10 kb nucleótidos de longitud o más. En otros casos, el tamaño de la inserción o la sustitución es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, de aproximadamente 350 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 800 kb, de aproximadamente 800 kb a 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb, a aproximadamente 2,5 Mb, de aproximadamente 2,5 Mb a aproximadamente 2,8 Mb, de aproximadamente 2,8 Mb a aproximadamente 3 Mb. La descripción describe que el inserto de polinucleótido dado y/o la región del locus de mamífero, de célula humana, o de mamífero no humano que se deleciona es de al menos 100, 200, 300, 400, 500, 600, 700, 800, o 900 nucleótidos o al menos 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 15 kb, 16 kb o más.

45 En una modalidad, el promotor es un promotor constitutivamente activo.

En una modalidad, el promotor es un promotor inducible. En una modalidad, el promotor inducible es un promotor químicamente regulado. En una modalidad, el promotor químicamente regulado es un promotor regulado por alcohol. En una modalidad, el promotor regulado por alcohol es un promotor del gen de la alcohol deshidrogenasa (*alcA*). En una modalidad, el promotor químicamente regulado es un promotor regulado por tetraciclina. En una modalidad, el promotor regulado por tetraciclina es un promotor sensible a tetraciclina. En una modalidad, el promotor regulado por tetraciclina es un promotor tet-Activado. En una modalidad, el promotor regulado por tetraciclina es un promotor tet-Desactivado. En una modalidad, el promotor químicamente regulado es un promotor regulado por esteroides. En una modalidad, el promotor regulado por esteroides es un promotor de un receptor de glucocorticoides de rata. En una modalidad, el promotor regulado por esteroides es un promotor de un receptor de estrógenos. En una modalidad, el promotor regulado por esteroides es un promotor de un receptor de ecdisona. En una modalidad, el promotor químicamente regulado es un promotor regulado por metales. En una modalidad, el promotor regulado por metales es un promotor de metaloproteína. En una modalidad, el promotor inducible es un promotor físicamente regulado. En una modalidad, el promotor físicamente regulado es un promotor regulado por temperatura. En una modalidad, el promotor regulado por temperatura es un promotor de choque térmico. En una modalidad, el promotor físicamente regulado es un promotor regulado por la luz. En una modalidad, el promotor regulado por la luz es un promotor inducible por la luz. En una modalidad, el promotor regulado por la luz es un promotor reprimible por la luz.

65 En una modalidad, el promotor es un promotor específico de tejidos. En una modalidad, el promotor es un promotor específico de neuronas. En una modalidad, el promotor es un promotor específico de la glía. En una modalidad, el promotor es un promotor específico de células musculares. En una modalidad, el promotor es un promotor específico de

5 células cardíacas. En una modalidad, el promotor es un promotor específico de células renales. En una modalidad, el promotor es un promotor específico de células óseas. En una modalidad, el promotor es un promotor específico de células endoteliales. En una modalidad, el promotor es un promotor específico de células del sistema inmunitario. En una modalidad, el promotor de células del sistema inmunitario es un promotor de células B. En una modalidad, el promotor de células del sistema inmunitario es un promotor de células T.

10 En una modalidad, el promotor es un promotor regulado por el desarrollo. En una modalidad, el promotor regulado por el desarrollo es activo solo durante una etapa embrionaria de desarrollo. En una modalidad, el promotor regulado por el desarrollo es activo solo en una célula adulta.

15 La descripción describe que el promotor puede seleccionarse basado en el tipo celular. Por lo tanto los diversos promotores encuentran uso en una célula eucariota, una célula eucariota no de rata, una célula de mamífero, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto o una célula CHO.

20 En algunas modalidades, el inserto de ácido nucleico comprende un ácido nucleico flanqueado por secuencias objetivo de recombinación específica de sitio. Se reconoce que mientras que el inserto de ácido nucleico completo puede ser flanqueado por tales secuencias objetivo de recombinación específica de sitio, cualquier región o polinucleótido de interés individual dentro del inserto de ácido nucleico también puede estar flanqueado por tales sitios. La recombinasa específica de sitio puede introducirse en la célula mediante cualquier medio, incluido mediante la introducción del polipéptido de recombinasa en la célula o mediante la introducción de un polinucleótido que codifica la recombinasa específica de sitio en la célula huésped. El polinucleótido que codifica la recombinasa específica de sitio puede localizarse dentro del inserto de ácido nucleico o dentro de un polinucleótido separado. La recombinasa específica de sitio puede unirse operativamente a un promotor activo en la célula que incluye, por ejemplo, un promotor inducible, un promotor que es endógeno a la célula, un promotor que es heterólogo a la célula, un promotor específico de células, un promotor específico de tejidos, o un promotor específico de la etapa de desarrollo. Las secuencias objetivo de recombinación específica de sitio, que pueden flanquear el inserto de ácido nucleico o cualquier polinucleótido de interés en el inserto de ácido nucleico puede incluir, pero sin limitarse a, loxP, lox511, lox2272, lox66, lox71, loxM2, lox5171, FRT, FRT11, FRT71, attP, att, FRT, rox, y una combinación de estos.

35 En algunas modalidades, los sitios de recombinación específica de sitio flanquean un polinucleótido que codifica un marcador de selección y/o un gen reportero contenido dentro del inserto de ácido nucleico. En tales casos después de la integración del inserto de ácido nucleico en el locus objetivo pueden eliminarse las secuencias entre los sitios de recombinación específica de sitio.

40 En una modalidad, el inserto de ácido nucleico comprende un polinucleótido que codifica un marcador de selección. El marcador de selección puede estar contenido en un casete de selección. Tales marcadores de selección incluyen, pero sin limitarse a, neomicina fosfotransferasa (neo^r), higromicina B fosfotransferasa (hyg^r), puromicina-N-acetiltransferasa (puro^r), blasticidin S deaminasa (bsr^r), xantina/guanina fosforribosil transferasa (gpt), o timidina cinasa del virus del herpes simple (HSV-k), o una combinación de estos. En una modalidad, el polinucleótido que codifica el marcador de selección se une operativamente a un promotor activo en una célula, una célula de rata, una célula pluripotente de rata, una célula ES de rata, una célula eucariota, una célula eucariota no de rata, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula de mamífero, una célula de mamífero no humano, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de ratón, una célula de hámster, un fibroblasto, o una célula CHO. Cuando se traslapan en serie los polinucleótidos de interés en un locus objetivo, el marcador de selección puede comprender un sitio de reconocimiento para un agente de nucleasa, como se indicó anteriormente. En una modalidad, el polinucleótido que codifica el marcador de selección es flanqueado por unas secuencias objetivo de recombinación específica de sitio.

50 El inserto de ácido nucleico puede comprender, además, un gen reportero unido operativamente a un promotor, en donde el gen reportero codifica una proteína reportera seleccionada del grupo que consiste en o que comprende LacZ, mPlum, mCherry, tdTomato, mStrawberry, J-Red, DsRed, mOrange, mKO, mCitrine, Venus, YPet, proteína fluorescente amarilla potenciada (eYFP), Emerald, proteína fluorescente verde potenciada (EGFP), CyPet, proteína fluorescente ciano (CFP), Cerulean, T-Sapphire, luciferasa, fosfatasa alcalina, y/o una combinación de estos. Tales genes reporteros pueden unirse operativamente a un promotor activo en la célula. Tales promotores pueden ser un promotor inducible, un promotor que es endógeno al gen reportero o a la célula, un promotor que es heterólogo al gen reportero o a la célula, un promotor específico de células, un promotor específico de tejidos, o un promotor específico de la etapa de desarrollo.

65 En una modalidad, el inserto de ácido nucleico que puede comprender un ácido nucleico de mamífero comprende un locus genómico que codifica una proteína expresada en el sistema nervioso, el sistema esquelético, el sistema digestivo, el sistema circulatorio, el sistema muscular, el sistema respiratorio, el sistema cardiovascular, el sistema linfático, el sistema endocrino, el sistema urinario, el sistema reproductor, o una combinación de estos. En una modalidad, el ácido nucleico de mamífero comprende un locus genómico que codifica una proteína expresada en una médula ósea o una

célula derivada de médula ósea. En una modalidad, el ácido nucleico comprende un locus genómico que codifica una proteína expresada en una célula de bazo.

5 En una modalidad, el ácido nucleico de mamífero comprende un locus genómico que codifica una proteína expresada en el sistema nervioso, el sistema esquelético, el sistema digestivo, el sistema circulatorio, el sistema muscular, el sistema respiratorio, el sistema cardiovascular, el sistema linfático, el sistema endocrino, el sistema urinario, el sistema reproductor, o una combinación de estos. En una modalidad, el ácido nucleico de mamífero comprende un locus genómico que codifica una proteína expresada en una médula ósea o una célula derivada de médula ósea. En una modalidad, el ácido nucleico comprende un locus genómico que codifica una proteína expresada en una célula de bazo. La descripción describe que el locus genómico comprende una secuencia de ADN genómico de ratón, una secuencia de ADN genómico de rata, una secuencia de ADN genómico de eucariota, una secuencia de ADN genómico de eucariota no rata, una secuencia de ADN genómico de mamífero, una secuencia de ADN genómico humano, o una secuencia de ADN de mamífero no humano, o una combinación de estos. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de rata y humano. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de ratón y humano. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de ratón y rata. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de rata, ratón y humano.

20 La descripción describe que el locus genómico comprende una secuencia de ADN genómico de ratón, una secuencia de ADN genómico de rata, una secuencia de ADN genómico de hámster, una secuencia de ADN genómico humano, una secuencia de ADN genómico de eucariota, una secuencia de ADN genómico de eucariota no rata, una secuencia de ADN genómico de mamífero, o una combinación de ADN de mamífero no humano, o una combinación de estos. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de rata y humano. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de ratón y humano. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de ratón y rata. La descripción describe que el locus genómico comprende, en cualquier orden, secuencias de ADN genómico de rata, ratón y humano.

30 La descripción describe que la modificación genética comprende al menos un alelo de enfermedad humana de un gen humano. La descripción describe que la enfermedad humana es una enfermedad neurológica. La descripción describe que la enfermedad humana es una enfermedad cardiovascular. La descripción describe que la enfermedad humana es una enfermedad renal. La descripción describe que la enfermedad humana es una enfermedad muscular. La descripción describe que la enfermedad humana es una enfermedad de la sangre. La descripción describe que la enfermedad humana es un cáncer. La descripción describe que la enfermedad humana es una enfermedad del sistema inmunitario.

35 La descripción describe que el alelo de enfermedad humana es un alelo dominante. La descripción describe que el alelo de enfermedad humana es un alelo recesivo. La descripción describe que el alelo de enfermedad humana comprende un alelo de polimorfismo de nucleótido único (SNP).

40 En una modalidad, la modificación genética produce una forma mutante de una proteína con una característica de unión alterada, localización alterada, expresión alterada, y/o patrón de expresión alterado.

45 En una modalidad, el inserto de ácido nucleico comprende un casete de selección. En una modalidad, el casete de selección comprende una secuencia de ácidos nucleicos que codifica un marcador de selección, en donde la secuencia de ácidos nucleicos se une operativamente a un promotor activo en células ES de rata. En una modalidad, el marcador de selección se selecciona de o comprende un gen de resistencia a higromicina o un gen de resistencia a neomicina.

50 En una modalidad, el ácido nucleico comprende un locus genómico que codifica una proteína expresada en una célula B. En una modalidad, el ácido nucleico comprende un locus genómico que codifica una proteína expresada en una célula B inmadura. En una modalidad, el ácido nucleico comprende un locus genómico que codifica una proteína expresada en una célula B madura.

55 En una modalidad, el inserto de ácido nucleico comprende un elemento regulador. En una modalidad, el elemento regulador es un promotor. En una modalidad, el elemento regulador es un potenciador. En una modalidad, el elemento regulador es un elemento de unión a represor transcripcional.

60 En una modalidad, la modificación genética comprende una delección de una secuencia no codificante de proteínas, pero no comprende una delección de una secuencia codificante de proteínas. En una modalidad, la delección de la secuencia no codificante de proteínas comprende una delección de un elemento regulador. En una modalidad, la modificación genética comprende una delección de un elemento regulador. En una modalidad, la modificación genética comprende una adición de un promotor o un elemento regulador. En una modalidad, la modificación genética comprende una sustitución de un promotor o un elemento regulador.

65 ii. Casetes de expresión

En la presente descripción se proporcionan polinucleótidos o moléculas de ácido nucleico que comprenden los diversos componentes empleados en un sistema de integración genómica objetivo proporcionado en la presente descripción (es decir, cualquiera de o cualquier combinación de agentes nucleasas, sitios de reconocimiento, inserto de ácidos nucleicos, polinucleótidos de interés, vectores de direccionamiento, marcadores de selección, y otros componentes).

Los términos "polinucleótido," "secuencia de polinucleótidos," "secuencia de ácidos nucleicos," y "fragmento de ácido nucleico" se usan indistintamente en la presente descripción. Estos términos abarcan secuencias de nucleótidos y similares. Un polinucleótido puede ser un polímero de ARN o ADN que es mono o bicatenario, que opcionalmente contiene bases de nucleótidos sintéticas, no naturales o alteradas. Un polinucleótido en la forma de un polímero de ADN puede estar formado por uno o más segmentos de ADNc, ADN genómico, ADN sintético, o mezclas de estos. Los polinucleótidos pueden comprender desoxirribonucleótidos y ribonucleótidos incluidas tanto moléculas de origen natural como análogos sintéticos, y cualquiera de sus combinaciones. Los polinucleótidos proporcionados en la presente descripción abarcan, además, todas las formas de secuencias que incluyen, pero sin limitarse a, formas monocatenarias, formas bicatenarias, horquillas, estructuras de tallo y bucle, y similares.

Se proporcionan, además, polinucleótidos recombinantes que comprenden los diversos componentes del sistema de integración genómica objetivo. Los términos "polinucleótido recombinante" y "constructo de ADN recombinante" se usan indistintamente en la presente descripción. Un constructo recombinante comprende una combinación artificial o heteróloga de secuencias de ácidos nucleicos, *por ejemplo*, secuencias reguladoras y recombinantes que no se encuentran juntas en la naturaleza. En otras modalidades, un constructo recombinante puede comprender secuencias reguladoras y secuencias codificantes que se derivan de diferentes fuentes, o secuencias reguladoras y secuencias codificantes derivadas de la misma fuente, pero reordenadas de manera diferente a la encontrada en la naturaleza. Un constructo tal puede usarse solo o puede usarse junto con un vector. Si se usa un vector, entonces la elección del vector depende del método que se usa para transformar las células huésped como es bien conocido por los expertos en la técnica. Por ejemplo, puede usarse un vector plasmídico. Se proporcionan, además, los elementos genéticos necesarios para transformar, seleccionar, y propagar exitosamente las células huésped que comprenden cualquiera de los fragmentos de ácido nucleico aislados proporcionados en la presente descripción. El tamizaje puede lograrse mediante el análisis de transferencia de Southern de ADN, análisis de transferencia de Northern de la expresión de ARNm, análisis de inmunotransferencia de la expresión de proteínas, o análisis fenotípicos, entre otros.

En modalidades específicas, uno o más de los componentes del sistema de integración genómica objetivo descritos en la presente descripción pueden proporcionarse en un casete de expresión para la expresión en una célula procarionota, una célula eucariota, una célula eucariota no de rata, una célula bacteriana, una célula de levadura, o una célula de mamífero u otro organismo o tipo celular de interés. El casete puede incluir secuencias reguladoras 5' y 3' unidas operativamente a un polinucleótido proporcionado en la presente descripción. "Unido operativamente" comprende una relación en donde los componentes unidos operativamente funcionan en su forma prevista. Por ejemplo, una unión operativa entre un polinucleótido de interés y una secuencia reguladora (es decir, un promotor) es un enlace funcional que permite la expresión del polinucleótido de interés. Los elementos unidos operativamente pueden ser contiguos o no contiguos. Cuando se usa para referirse a la unión de dos regiones codificantes de proteínas, unido operativamente significa que las regiones codificantes están en el mismo marco de lectura. En otro caso, una secuencia de ácidos nucleicos que codifica una proteína puede unirse operativamente a secuencias reguladoras (por ejemplo, secuencia promotora, potenciadora, silenciadora, etc.) para mantener la regulación transcripcional apropiada. En un caso, una secuencia de ácidos nucleicos de una región variable de inmunoglobulina (o segmentos V(D)J) puede unirse operativamente a una secuencia de ácidos nucleicos de una región constante de inmunoglobulina para permitir la recombinación apropiada entre las secuencias en una secuencia de cadena pesada o ligera de inmunoglobulina.

El casete puede contener, adicionalmente, al menos un polinucleótido adicional de interés a cointroducir en el organismo. Alternativamente, el polinucleótido adicional de interés puede proporcionarse en múltiples casetes de expresión. Un casete de expresión tal se proporciona con una pluralidad de sitios de restricción y/o sitios de recombinación para la inserción de un polinucleótido recombinante que estará bajo la regulación transcripcional de las regiones reguladoras. El casete de expresión puede contener, adicionalmente, genes marcadores de selección.

El casete de expresión puede incluir en la dirección 5'-3' de transcripción, una región de iniciación transcripcional y traduccional (es decir, un promotor), un polinucleótido recombinante proporcionado en la presente descripción, y una región de terminación transcripcional y traduccional (es decir, región de terminación) funcional en la célula de mamífero o una célula huésped de interés. Las regiones reguladoras (es decir, promotores, regiones reguladoras transcripcionales, y regiones de terminación traduccionales) y/o un polinucleótido proporcionado en la presente descripción pueden ser nativos/análogos a la célula huésped o entre sí. Alternativamente, las regiones reguladoras y/o un polinucleótido proporcionados en la presente descripción pueden ser heterólogos a la célula huésped o entre sí. Por ejemplo, un promotor unido operativamente a un polinucleótido heterólogo es de una especie diferente a las especies a partir de las cuales se obtuvo el polinucleótido, o, si es de la misma/análoga especie, uno o ambos se modifican sustancialmente a partir de su forma y/o locus genómico original, o el promotor no es el promotor nativo para el polinucleótido unido operativamente. Alternativamente, las regiones reguladoras y/o un polinucleótido recombinante proporcionados en la presente descripción pueden ser totalmente sintéticos.

La región de terminación puede ser nativa con la región de iniciación transcripcional, puede ser nativa con el polinucleótido recombinante unido operativamente, puede ser nativa con la célula huésped, o puede derivarse de otra fuente (es decir, extraña o heteróloga) al promotor, el polinucleótido recombinante, la célula huésped, o cualquier combinación de estos.

5 Durante la preparación del casete de expresión, los diversos fragmentos de ADN pueden manipularse, para proporcionar las secuencias de ADN en la orientación adecuada. Hacia este fin, los adaptadores o enlazadores pueden emplearse para unir los fragmentos de ADN o pueden estar implicadas otras manipulaciones para proporcionar sitios de restricción convenientes, eliminación de ADN superfluo, eliminación de sitios de restricción, o similares. Para este propósito, puede estar implicada la mutagénesis in vitro, reparación con cebadores, restricción, apareamiento, resustituciones, por ejemplo, transiciones y transversiones.

10 Pueden usarse una serie de promotores en los casetes de expresión proporcionados en la presente descripción. Los promotores pueden seleccionarse basado en el resultado deseado. Se reconoce que pueden mejorarse diferentes aplicaciones mediante el uso de diferentes promotores en los casetes de expresión para modular el momento, la localización y/o el nivel de expresión del polinucleótido de interés. Tales constructos de expresión pueden contener, además, si es conveniente, una región reguladora promotora (por ejemplo, una que confiere expresión inducible, constitutiva, regulada en desde el punto de vista ambiental o de desarrollo, específica/selectiva de células o tejidos), un sitio de inicio de la iniciación de la transcripción, un sitio de unión a ribosomas, una señal de procesamiento de ARN, un sitio de terminación de la transcripción, y/o una señal de poliadenilación.

20 El casete de expresión que contiene los polinucleótidos proporcionados en la presente descripción puede comprender, además, un gen marcador de selección para la selección de células transformadas. Los genes marcadores de selección se utilizan para la selección de células o tejidos transformados.

25 Cuando sea apropiado, las secuencias empleadas en los métodos y composiciones (es decir, el polinucleótido de interés, el agente nucleasa, etc.) pueden optimizarse para aumentar la expresión en la célula. Esto es, los genes pueden sintetizarse mediante el uso de codones preferidos en una célula de interés dada que incluyen, por ejemplo, codones preferidos de mamífero, codones preferidos de ser humano, codones preferidos de roedor, codones preferidos de roedor no rata, codones preferidos de ratón, codones preferidos de rata, codones preferidos de hámster, etc. para una expresión mejorada.

30 Los diversos métodos y composiciones proporcionados en la presente descripción pueden emplear marcadores de selección. Pueden usarse diversos marcadores de selección en los métodos y composiciones descritas en la presente descripción. Tales marcadores de selección pueden, por ejemplo, impartir resistencia a un antibiótico tal como G418, higromicina, blastidina, neomicina, o puromicina. Tales marcadores de selección incluyen neomicina fosfotransferasa (neo^r), higromicina B fosfotransferasa (hyg^r), puromicina-N-acetiltransferasa (puro^r), y blastidina S deaminasa (bsr^r). Aún en otras modalidades, el marcador de selección se une operativamente a un promotor inducible y la expresión del marcador de selección es tóxica para la célula. Los ejemplos no limitantes de tales marcadores de selección incluyen xantina/guanina fosforribosil transferasa (gpt), hipoxantina-guanina fosforribosiltransferasa (HGPRT) o timidina cinasa del virus del herpes simple (HSV-TK). El polinucleótido que codifica los marcadores de selección se unen operativamente a un promotor activo en la célula.

iii. Vectores de direccionamiento

45 Los vectores de direccionamiento se emplean para introducir el inserto de ácido nucleico en el locus objetivo del ácido nucleico de rata, de eucariota, de eucariota no rata, de mamífero, de mamífero no humano, de ser humano, de roedor, de roedor no rata, de ratón o de hámster. El vector de direccionamiento comprende el inserto de ácido nucleico y comprende, además, un brazo de homología 5' y uno 3', los cuales flanquean el inserto de ácido nucleico. Los brazos de homología, los cuales flanquean el inserto de ácido nucleico, corresponden a regiones dentro del locus objetivo del ácido nucleico de rata, de eucariota, de eucariota no rata, de mamífero, de mamífero no humano, de ser humano, de roedor, de roedor no rata, de ratón o de hámster. Para facilitar la referencia, las regiones genómicas afines correspondientes dentro del locus genómico objetivo se refieren en la presente descripción a "sitios objetivo". Por ejemplo, un vector de direccionamiento puede comprender un primer inserto de ácido nucleico flanqueado por un primer y un segundo brazo de homología complementario a un primer y un segundo sitio objetivo. Como tal, el vector de direccionamiento de este modo ayuda a la integración del inserto de ácido nucleico en el locus objetivo del ácido nucleico de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster a través del evento de recombinación homóloga que se produce entre los brazos de homología y los sitios objetivo complementarios dentro del genoma de la célula.

60 La descripción describe que el locus objetivo del ácido nucleico de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster ácido nucleico comprende una primera secuencia de ácidos nucleicos que es complementaria al brazo de homología 5' y una segunda secuencia de ácidos nucleicos que es complementaria al brazo de homología 3'. En una modalidad, las secuencias de ácidos nucleicos primera y segunda están separadas por al menos 5 kb. En otra modalidad, las secuencias de ácidos nucleicos primera y segunda están separadas por al menos 5 kb pero menos de 200 kb. En una modalidad, las secuencias de ácidos nucleicos primera y segunda están separadas por al menos 10 kb. En una modalidad, las secuencias de ácidos nucleicos primera y segunda están separadas por al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al

menos 90 kb, al menos 100 kb, al menos 110 kb, al menos 120 kb, al menos 130 kb, al menos 140 kb, al menos 150 kb, al menos 160 kb, al menos 170 kb, al menos 180 kb, al menos 190 kb, o al menos 200 kb. Aún en otras modalidades, las secuencias de ácidos nucleicos primera y segunda están separadas por al menos 5 kb pero menos de 10 kb, al menos 5 kb pero menos de 3 Mb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero menos de aproximadamente 1 Mb, al menos aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos aproximadamente 2 Mb pero menos de 2,5 Mb, al menos aproximadamente 2,5 Mb pero menos de 3 Mb, o al menos aproximadamente 2 Mb pero menos de aproximadamente 3 Mb.

Un brazo de homología del vector de direccionamiento puede ser de cualquier longitud que sea suficiente para promover el evento de recombinación homóloga con un sitio objetivo correspondiente, que incluye por ejemplo, al menos 5-10 kb, 5-15 kb, 10-20 kb, 20-30 kb, 30-40 kb, 40-50 kb, 50-60 kb, 60-70 kb, 70-80 kb, 80-90 kb, 90-100 kb, 100-110 kb, 110-120 kb, 120-130 kb, 130-140 kb, 140-150 kb, 150-160 kb, 160-170 kb, 170-180 kb, 180-190 kb, 190-200 kb de longitud o más. Como se indica con más detalle más adelante, los vectores de direccionamiento grandes pueden emplear brazos de transformación de mayor longitud. En una modalidad específica, la suma total del brazo de homología 5' y el brazo de homología 3' es al menos 10 kb o la suma total del brazo de homología 5' y el brazo de homología 3' es al menos aproximadamente 16 kb a aproximadamente 100 kb o aproximadamente 30 kb a aproximadamente 100 kb. En otras modalidades, el tamaño de la suma total del total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 150 kb, aproximadamente 10 kb a aproximadamente 100 kb, aproximadamente 10 kb a aproximadamente 75 kb, aproximadamente 20 kb a aproximadamente 150 kb, aproximadamente 20 kb a aproximadamente 100 kb, aproximadamente 20 kb a aproximadamente 75 kb, aproximadamente 30 kb a aproximadamente 150 kb, aproximadamente 30 kb a aproximadamente 100 kb, aproximadamente 30 kb a aproximadamente 75 kb, aproximadamente 40 kb a aproximadamente 150 kb, aproximadamente 40 kb a aproximadamente 100 kb, aproximadamente 40 kb a aproximadamente 75 kb, aproximadamente 50 kb a aproximadamente 150 kb, aproximadamente 50 kb a aproximadamente 100 kb, o aproximadamente 50 kb a aproximadamente 75 kb, aproximadamente 10 kb a aproximadamente 30 kb, aproximadamente 20 kb a aproximadamente 40 kb, aproximadamente 40 kb a aproximadamente 60 kb, aproximadamente 60 kb a aproximadamente 80 kb, aproximadamente 80 kb a aproximadamente 100 kb, aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a aproximadamente 150 kb. En una modalidad, el tamaño de la delección es igual o similar al tamaño de la suma total de los brazos de homología 5' y 3' del LTVEC.

En una modalidad, el locus genómico de interés comprende (i) una secuencia objetivo 5' que es homóloga al brazo de homología 5'; y (ii) una secuencia objetivo 3' que es homóloga al brazo de homología 3'. En una modalidad, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 3 Mb. Aún en otras modalidades, la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 10 kb, al menos 5 kb pero menos de 3 Mb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero menos de aproximadamente 1 Mb, al menos aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos aproximadamente 2 Mb pero menos de 2,5 Mb, al menos aproximadamente 2,5 Mb pero menos de aproximadamente 3 Mb, o al menos aproximadamente 2 Mb pero menos de aproximadamente 3 Mb.

Cuando se emplean agentes nucleasas, las regiones genómicas afines que corresponden a los brazos de homología 5' y 3' de un vector de direccionamiento se "localizan en una cercanía suficiente" a los sitios objetivo de nucleasas para promover la ocurrencia de un evento de recombinación homóloga entre las regiones genómicas afines y los brazos de homología tras una mella o ruptura de la doble cadena en el sitio de reconocimiento. Por ejemplo, los sitios objetivo de las nucleasas pueden localizarse en cualquier parte entre las regiones genómicas afines que corresponden a los brazos de homología 5' y 3'. En modalidades específicas, el sitio de reconocimiento está inmediatamente adyacente a al menos una o ambas regiones genómicas afines.

Como se usa en la presente descripción, un brazo de homología y un sitio objetivo (es decir, región genómica afín) "se complementan" o son "complementarios" a otro cuando las dos regiones comparten un nivel suficiente de identidad de secuencia con respecto a otra para funcionar como sustratos para una reacción de recombinación homóloga. Por "homología" se entiende las secuencias de ADN que son idénticas o comparten identidad de secuencia con respecto a una secuencia correspondiente o "complementaria". La identidad de secuencia entre un sitio objetivo dado y el brazo de homología correspondiente encontrada en el vector de direccionamiento puede ser cualquier grado de identidad de secuencia que permita que se produzca la recombinación homóloga. Por ejemplo, la cantidad de identidad de secuencia compartida por el brazo de homología del vector de direccionamiento (o un fragmento de este) y el sitio objetivo (o un fragmento de este) puede ser al menos 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86

%, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia, de manera que las secuencias experimentan recombinación homóloga. Además, una región de homología complementaria entre el brazo de homología y el sitio objetivo complementario puede ser de cualquier longitud que sea suficiente para promover la recombinación homóloga en el sitio de reconocimiento escindido. Por ejemplo, un brazo de homología y/o sitio objetivo complementario dado puede comprender regiones de homología complementarias que son de al menos 5-10 kb, 5-15 kb, 10-20 kb, 20-30 kb, 30-40 kb, 40-50 kb, 50-60 kb, 60-70 kb, 70-80 kb, 80-90 kb, 90-100 kb, 100-110 kb, 110-120 kb, 120-130 kb, 130-140 kb, 140-150 kb, 150-160 kb, 160-170 kb, 170-180 kb, 180-190 kb, 190-200 kb, 200 kb a 300 kb de longitud o más (tal como se describe en los vectores LTVEC descritos en cualquier otra parte de la presente descripción) de manera que el brazo de homología tiene suficiente homología para experimentar recombinación homóloga con los sitios objetivo correspondientes dentro del genoma de la célula. Para facilitar la referencia los brazos de homología se refieren en la presente descripción como un brazo de homología 5' y uno 3'. Esta terminología se refiere a la posición relativa de los brazos de homología con respecto al inserto de ácido nucleico dentro del vector de direccionamiento.

Los brazos de homología del vector de direccionamiento se diseñan, por lo tanto, para que sean complementarios a un sitio objetivo con el locus objetivo. Por lo tanto, los brazos de homología pueden ser complementarios a un locus que es nativo a la célula, o alternativamente pueden ser complementarios a una región de un segmento de ADN heterólogo o exógeno que se integró en el genoma de la célula, que incluye, pero sin limitarse a, transgenes, casetes de expresión, o regiones heterólogas o exógenas de ADN genómico. Alternativamente, los brazos de homología del vector de direccionamiento pueden ser complementarios a una región de un cromosoma artificial humano o cualquier otra región genómica modificada genéticamente contenida en una célula huésped apropiada. Aún más, los brazos de homología del vector de direccionamiento pueden ser complementarios a una región de una biblioteca de BAC, una biblioteca de cósmidos, o una biblioteca de fago P1. Por lo tanto, esta descripción describe que los brazos de homología del vector de direccionamiento son complementarios a un locus genómico de una rata, un eucariota, un eucariota no rata, un mamífero, un mamífero no humano, un ser humano, un roedor, un roedor no rata, un ratón o un hámster que es nativo, heterólogo o exógeno a una célula dada. En modalidades adicionales, los brazos de homología son complementarios a un locus genómico de una rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster que no puede transformarse mediante el uso de un método convencional o puede transformarse solo de manera incorrecta o solo con una eficiencia significativamente baja, en la ausencia de una mella o ruptura de la doble cadena inducida por un agente nucleasa. En una modalidad, los brazos de homología se derivan de un ADN sintético.

Aún en otros ejemplos, los brazos de homología 5' y 3' son complementarios al mismo genoma que el genoma objetivo. En un ejemplo, los brazos de homología son de un genoma relacionado, por ejemplo, el genoma objetivo es un genoma de rata de una primera cepa, y los brazos de transformación son de un genoma de rata de una segunda cepa, en donde la primera cepa y la segunda cepa son diferentes. En otro ejemplo, los brazos de homología son del genoma del mismo animal o son del genoma de la misma cepa, por ejemplo, el genoma objetivo es un genoma de rata de una primera cepa, y los brazos de transformación son de un genoma de rata de la misma rata o de la misma cepa.

El vector de direccionamiento (tal como un vector de direccionamiento grande) puede comprender, además, un casete de selección o un gen reportero como se analiza en cualquier otra parte de la presente descripción. El casete de selección puede comprender una secuencia de ácidos nucleicos que codifica un marcador de selección, en donde la secuencia de ácidos nucleicos se une operativamente a un promotor. El promotor puede ser activo en una célula procarionta de interés y/o activo en una célula eucariota de interés. Tales promotores pueden ser un promotor inducible, un promotor que es endógeno con respecto al gen reportero o a la célula, un promotor que es heterólogo al gen reportero o a la célula, un promotor específico de células, un promotor específico de tejidos o un promotor específico de la etapa de desarrollo. En una modalidad, el marcador de selección se selecciona de o comprende neomicina fosfotransferasa (neo^r), higromicina B fosfotransferasa (hyg^r), puromicina-N-acetiltransferasa (puro^r), blasticidina S deaminasa (bsr^r), xantina/guanina fosforribosil transferasa (gpt), y timidina cinasa del virus del herpes simple (HSV-k), y/o una combinación de estos. El marcador de selección del vector de direccionamiento puede estar flanqueado por los brazos de homología 5' y 3' o puede encontrarse ya sea 5' o 3' con respecto a los brazos de homología.

En una modalidad, el vector de direccionamiento (tal como un vector de direccionamiento grande) comprende un gen reportero unido operativamente a un promotor, en donde el gen reportero codifica una proteína reportera seleccionada del grupo que consiste en o comprende LacZ, mPlum, mCherry, tdTomato, mStrawberry, J-Red, DsRed, mOrange, mKO, mCitrine, Venus, YPet, proteína fluorescente amarilla potenciada (EYFP), Emerald, proteína fluorescente verde potenciada (EGFP), CyPet, proteína fluorescente ciano (CFP), Cerulean, T-Sapphire, luciferasa, fosfatasa alcalina, y/o una combinación de estos. Tales genes reporteros pueden unirse operativamente a un promotor activo en la célula. Tales promotores pueden ser un promotor inducible, un promotor que es endógeno al gen reportero o a la célula, un promotor que es heterólogo al gen reportero o a la célula, un promotor específico de células, un promotor específico de tejidos o un promotor específico de la etapa de desarrollo.

En una modalidad, el uso combinado del vector de direccionamiento (que incluye, por ejemplo, un vector de direccionamiento grande) con el agente nucleasa da como resultado un aumento de la eficiencia de transformación en comparación con el uso del vector de direccionamiento solo. En una modalidad, cuando el vector de direccionamiento se usa junto con el agente nucleasa, la eficiencia de transformación del vector de direccionamiento aumenta al menos en

dos veces, al menos tres veces, o al menos 4 veces cuando se compara con el vector de direccionamiento cuando se usa solo.

5 Cuando se emplea un vector de direccionamiento, el diseño del vector de esta descripción puede ser tal como para permitir la inserción de una secuencia dada que es de aproximadamente 5 kb a aproximadamente 200 kb como se describe en la presente descripción. En un ejemplo, la inserción es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 70 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, o de aproximadamente 190 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.

25 Cuando se emplea un vector de direccionamiento, el diseño del vector puede ser tal como para permitir la sustitución de una secuencia dada que es de aproximadamente 5 kb a aproximadamente 200 kb o de aproximadamente 5 kb a aproximadamente 3,0 Mb como se describe en la presente descripción. En un ejemplo, la sustitución es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 70 kb a aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, de aproximadamente 190 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb.

45 En una modalidad, el vector de direccionamiento comprende un gen de recombinasa específica de sitio. En una modalidad, el gen de recombinasa específica de sitio codifica una Cre recombinasa. En una modalidad, el gen de Cre recombinasa es Crei, en donde dos exones que codifican la Cre recombinasa están separados por un intrón para evitar su expresión en una célula procarionota.

50 En una modalidad, el gen de Cre recombinasa comprende, además, una señal de localización nuclear para facilitar la localización de Cre (o cualquier recombinasa o agente nucleasa) al núcleo (por ejemplo, el gen es un gen de NL-Cre). En una modalidad específica, el gen de Cre recombinasa comprende, además, una señal de localización nuclear y un intrón (por ejemplo, NL-Crei).

55 En diversas modalidades, un promotor adecuado para la expresión del agente nucleasa (que incluye la recombinasa Cre o Crei analizada anteriormente) se selecciona de o comprende un *Prrm1*, *Blimp1*, *Gata6*, *Gata4*, *Igf2*, *Lhx2*, *Lhx5*, y/o *Pax3*. En una modalidad específica, el promotor es el promotor *Gata6* o *Gata4*. Los diversos promotores pueden ser de cualquier organismo, incluido, por ejemplo, un roedor tal como un ratón o una rata, un roedor no rata, un eucariota, un eucariota no rata, un mamífero no humano, un mamífero, un ser humano o un hámster. En otra modalidad específica, el promotor es un promotor *Prrm1*. En otra modalidad específica, el promotor es un promotor *Prrm1* de rata. En otra modalidad específica, el promotor es un promotor *Prrm1* de ratón. En otra modalidad específica, el promotor es un promotor *Blimp1* o un fragmento de este, por ejemplo, un fragmento de 1 kb o 2 kb de un promotor *Blimp1*. Ver, por ejemplo, la patente de los Estados Unidos 8,697,851 y la publicación de solicitud de patente de los Estados Unidos 2013-0312129.

65 iv. Vectores de direccionamiento grandes

El término "vector de direccionamiento grande" o "LTVEC" como se usa en la presente descripción comprende vectores de direccionamiento grandes que comprenden los brazos de homología que corresponden a y se derivan de secuencias de ácidos nucleicos más grandes que las usadas típicamente por otros enfoques destinados a realizar la transformación homóloga en células y/o que comprenden insertos de ácidos nucleicos que comprenden secuencias de ácidos nucleicos más grandes que las usadas típicamente por otros enfoques destinados a realizar la transformación mediante recombinación homóloga en células. Por ejemplo, el LTVEC hace posible la modificación de grandes loci que no pueden ser acomodados por los vectores de direccionamiento basados en plásmido tradicionales debido a sus limitaciones de tamaño. En modalidades específicas, los brazos de homología y/o el inserto de ácido nucleico del LTVEC comprenden la secuencia genómica de una célula eucariota o una célula eucariota no de rata. El tamaño del LTVEC es muy grande para permitir el tamizaje de eventos de transformación mediante ensayos convencionales, *por ejemplo*, transferencia de southern y PCR de gran alcance (*por ejemplo*, 1 kb-5 kb). Los ejemplos del LTVEC, incluyen, pero sin limitarse a, vectores derivados de un cromosoma artificial bacteriano (BAC), un cromosoma artificial humano o un cromosoma artificial de levadura (YAC). Los ejemplos no limitantes de LTVEC y los métodos para producirlos se describen, *por ejemplo*, en la patente de los Estados Unidos núm. 6,586,251, 6,596,541, 7,105,348, y los documentos WO 2002/036789 (PCT/USO1/45375), y US 2013/0137101.

La descripción describe que el LTVEC puede ser de cualquier longitud, que incluye, pero sin limitarse a, de aproximadamente 20 kb a aproximadamente 400 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 75 kb, de aproximadamente 75 kb a aproximadamente 100 kb, de aproximadamente 100 kb a 125 kb, de aproximadamente 125 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 175 kb, de aproximadamente 175 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 225 kb, de aproximadamente 225 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 275 kb o de aproximadamente 275 kb a aproximadamente 300 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, de aproximadamente 350 kb a aproximadamente 400 kb, de aproximadamente 350 kb a aproximadamente 550 kb. En una modalidad, el LTVEC es de aproximadamente 100 kb.

En algunas modalidades, el LTVEC es de al menos 10 kb, al menos 15 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb o al menos 200 kb.

En algunas modalidades, el LTVEC comprende una secuencia de ácidos nucleicos de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb o al menos 200 kb.

La descripción describe que el LTVEC comprende un inserto de ácido nucleico en el intervalo de aproximadamente 5 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 0,5 kb a aproximadamente 30 kb, de aproximadamente 0,5 kb a aproximadamente 40 kb, de aproximadamente 30 kb a aproximadamente 150 kb, de aproximadamente 0,5 kb a aproximadamente 150 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, o de aproximadamente 190 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.

En una modalidad, el LTVEC comprende una secuencia de ácidos nucleicos de al menos 100 kb, al menos 150 kb, o al menos 200 kb.

Cuando se emplea un LTVEC, el diseño del vector puede ser tal como para permitir la sustitución de una secuencia dada que es de aproximadamente 5 kb a aproximadamente 200 kb o de aproximadamente 5 kb a aproximadamente 3 Mb como se describe en la presente descripción. En una modalidad, la sustitución es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de

aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, de aproximadamente 190 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb.

En una modalidad, los brazos de homología del LTVEC se derivan de una biblioteca de BAC, una biblioteca de cósmidos, o una biblioteca de fago P1. En otras modalidades, los brazos de homología se derivan del locus genómico a transformar de la célula y en algunos casos el locus genómico objetivo, al que se dirige el LTVEC diseñado no puede transformarse mediante el uso de un método convencional. Aún en otras modalidades, los brazos de homología se derivan de un ADN sintético.

En una modalidad, una suma total del brazo de homología 5' y el brazo de homología 3' en el LTVEC es de al menos 10 kb. En otras modalidades, la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 30 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de 100 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 200 kb. En una modalidad la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 30 kb a aproximadamente 100 kb. En otras modalidades, el tamaño de la suma total del total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 150 kb, aproximadamente 10 kb a aproximadamente 100 kb, aproximadamente 10 kb a aproximadamente 75 kb, aproximadamente 20 kb a aproximadamente 150 kb, aproximadamente 20 kb a aproximadamente 100 kb, aproximadamente 20 kb a aproximadamente 75 kb, aproximadamente 30 kb a aproximadamente 150 kb, aproximadamente 30 kb a aproximadamente 100 kb, aproximadamente 30 kb a aproximadamente 75 kb, aproximadamente 40 kb a aproximadamente 150 kb, aproximadamente 40 kb a aproximadamente 100 kb, aproximadamente 40 kb a aproximadamente 75 kb, aproximadamente 50 kb a aproximadamente 150 kb, aproximadamente 50 kb a aproximadamente 100 kb, o aproximadamente 50 kb a aproximadamente 75 kb, aproximadamente 10 kb a aproximadamente 30 kb, aproximadamente 20 kb a aproximadamente 40 kb, aproximadamente 40 kb a aproximadamente 60 kb, aproximadamente 60 kb, aproximadamente 60 kb a aproximadamente 80 kb, aproximadamente 80 kb a aproximadamente 100 kb, aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a aproximadamente 150 kb. En una modalidad, el tamaño de la delección es igual o similar al tamaño de la suma total de los brazos de homología 5' y 3' del LTVEC.

En otras modalidades, el brazo de homología 5' varía de aproximadamente 5 kb a aproximadamente 100 kb. En una modalidad, el brazo de homología 3' varía de aproximadamente 5 kb a aproximadamente 100 kb. En otras modalidades, la suma total de los brazos de homología 5' y 3' son de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 70 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, de aproximadamente 190 kb a aproximadamente 200 kb, o de aproximadamente 30 kb a aproximadamente 100 kb, aproximadamente 10 kb a aproximadamente 30 kb, aproximadamente 20 kb a aproximadamente 40 kb, aproximadamente 40 kb a aproximadamente 60 kb, aproximadamente 60 kb a aproximadamente 80 kb, aproximadamente 80 kb a aproximadamente 100 kb, aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a aproximadamente 150 kb.

En una modalidad, el LTVEC comprende un inserto de ácido nucleico que es homólogo u ortólogo a una secuencia de ácidos nucleicos de rata flanqueado por los brazos de homología del LTVEC. En una modalidad, la secuencia de ácidos nucleicos del inserto es de una especie diferente de una rata. En una modalidad, la secuencia de ácidos nucleicos del inserto es de un eucariota. En una modalidad, el inserto de ácido nucleico que es homólogo u ortólogo a la secuencia de ácidos nucleicos de rata es un ácido nucleico de mamífero. En una modalidad, el inserto de ácido nucleico que es homólogo u ortólogo a la secuencia de ácidos nucleicos de rata es un ácido nucleico no humano de mamífero. En una modalidad, el ácido nucleico de mamífero es un ácido nucleico de ratón. En una modalidad, el ácido nucleico de mamífero es un ácido nucleico humano. En una modalidad, el ácido nucleico de mamífero es un ácido nucleico de hámster. En una

modalidad, el inserto de ácido nucleico es un ADN genómico. En una modalidad, el inserto es de 5 kb a 200 kb como se describió anteriormente.

En una modalidad, el LTVEC comprende un casete de selección o un gen reportero. Diversas formas del casete de selección y del gen reportero que pueden emplearse se analizan en cualquier otra parte de la presente descripción.

Como se describe en cualquier otra parte de la presente descripción, el LTVEC puede usarse, además, en los métodos proporcionados en la presente descripción en combinación con un agente nucleasa que promueve la recombinación homóloga entre el vector de direccionamiento y el locus objetivo de un ácido nucleico de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster en una célula pluripotente o no pluripotente de rata, eucariota, eucariota no de rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster.

En una modalidad, el vector de direccionamiento grande (LTVEC) comprende un gen de recombinasa específica de sitio. En una modalidad, el gen de recombinasa específica de sitio codifica una Cre recombinasa. En una modalidad, el gen de Cre recombinasa es Crei, en donde dos exones que codifican la Cre recombinasa están separados por un intrón para evitar su expresión en una célula procarionota. En una modalidad, el gen de Cre recombinasa comprende, además, una señal de localización nuclear para facilitar la localización de Cre (o cualquier recombinasa o agente nucleasa) al núcleo (por ejemplo, el gen es un gen de NL-Cre). En una modalidad específica, el gen de Cre recombinasa comprende, además, una señal de localización nuclear y un intrón (por ejemplo, NL-Crei)

En diversas modalidades, un promotor adecuado para la expresión del agente nucleasa (que incluye la recombinasa Cre o Crei analizada anteriormente) se selecciona de o comprende un *Prr1*, *Blimp1*, *Gata6*, *Gata4*, *Igf2*, *Lhx2*, *Lhx5*, y/o *Pax3*. En una modalidad específica, el promotor es el promotor *Gata6* o *Gata4*. Los diversos promotores pueden ser de cualquier organismo, incluido, por ejemplo, un roedor tal como un ratón o una rata, un roedor no rata, un eucariota, un eucariota no rata, un mamífero no humano, un mamífero, un ser humano o un hámster. En otra modalidad específica, el promotor es un promotor *Prr1*. En otra modalidad específica, el promotor es un promotor *Prr1* de rata. En otra modalidad específica, el promotor es un promotor *Prr1* de ratón. En otra modalidad específica, el promotor es un promotor *Blimp1* o un fragmento de este, por ejemplo, un fragmento de 1 kb o 2 kb de un promotor *Blimp1*. Ver, por ejemplo, la patente de los Estados Unidos 8,697,851 y la publicación de solicitud de patente de los Estados Unidos 2013-0312129.

En una modalidad, el LTVEC comprende un inserto de ácido nucleico que puede producir una delección, adición, sustitución o una combinación de estos de una región del locus de *ApoE* de rata, un eucariota, un eucariota no rata, un mamífero, un mamífero no humano, un ser humano, un roedor, un roedor no rata, un ratón o un hámster, el locus de *Il2rg*, el locus de *Rag2*, el locus de *Rag1* y/o el locus de *Rag2/Rag1* como se analiza en detalle en cualquier otra parte de la presente descripción. En modalidades específicas, la modificación genética en el locus de *ApoE* da como resultado una disminución, un aumento o una modulación de la actividad de ApoE, actividad de IL-2Rg, actividad de *Rag2*, actividad de *Rag1* y/o *Rag2* y actividad de *Rag1*. En una modalidad, se genera una inactivación de *ApoE*, e inactivación de *Il2rg*, una inactivación de *Rag2*, una inactivación de *Rag1*, una inactivación de *Rag2/Rag1*. Como se analiza más adelante, pueden emplearse agentes nucleasas con cualquiera de los sistemas de transformación de LTVEC para transformar cualquier locus genómico de interés.

En otra modalidad, el genoma se expone a una proteína Cas y un ARN de CRISPR en la presencia de un vector de direccionamiento grande (LTVEC) que comprende una secuencia de ácidos nucleicos de al menos 10 kb. En tales casos, después de la exposición a la proteína Cas, al ARN de CRISPR, y al LTVEC, el genoma se modifica para que contenga al menos 10 kb de secuencia de ácidos nucleicos. En modalidades específicas, el LTVEC comprende una secuencia de ácidos nucleicos de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb o al menos 200 kb.

v. Agentes nucleasas y sitios de reconocimiento para agentes nucleasas

Como se expone en esta descripción, pueden utilizarse agentes nucleasas en los métodos y composiciones descritas en la presente descripción para ayudar en la modificación del locus objetivo tanto en una célula procarionota o dentro de una célula pluripotente o no pluripotente de rata, eucariota, eucariota no de rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster. Un agente nucleasa tal puede promover la recombinación homóloga entre el vector de direccionamiento y el locus objetivo. En una modalidad, el agente nucleasa comprende un agente endonucleasa.

Como se usa en la presente descripción, el término "sitio de reconocimiento de un agente nucleasa" comprende una secuencia de ADN en la cual se induce una mella o ruptura de la doble cadena por un agente nucleasa. El sitio de reconocimiento de un agente nucleasa puede ser endógeno (o nativo) a la célula o el sitio de reconocimiento puede ser exógeno a la célula. En modalidades específicas, el sitio de reconocimiento es exógeno a la célula y por lo tanto no es de origen natural en el genoma de la célula. Aún en otras modalidades, el sitio de reconocimiento es exógeno a la célula y a los polinucleótidos de interés que se desean ubicar en el locus genómico objetivo. En modalidades adicionales, el sitio de reconocimiento exógeno o endógeno está presente solo una vez en el genoma de la célula huésped. En modalidades específicas, se identifica un sitio endógeno o nativo que se produce solo una vez dentro del genoma. Un sitio tal puede usarse después para diseñar agentes nucleasas que producirán una mella o ruptura de la doble cadena en el sitio de reconocimiento endógeno.

La longitud del sitio de reconocimiento puede variar, e incluye, por ejemplo, sitios de reconocimiento que son de al menos 4, 6, 8, 10, 12, 14, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70 o más nucleótidos de longitud. En una modalidad, cada monómero del agente nucleasa reconoce un sitio de reconocimiento de al menos 9 nucleótidos. En otras modalidades, el sitio de reconocimiento es de aproximadamente 9 a aproximadamente 12 nucleótidos de longitud, de aproximadamente 12 a aproximadamente 15 nucleótidos de longitud, de aproximadamente 15 a aproximadamente 18 nucleótidos de longitud, o de aproximadamente 18 a aproximadamente 21 nucleótidos de longitud, y cualquier combinación de tales subintervalos (por ejemplo, 9-18 nucleótidos). El sitio de reconocimiento podría ser palindrómico, lo que significa que, la secuencia en una cadena se lee igual en la dirección opuesta en la cadena complementaria. Se reconoce que un agente nucleasa dado puede unirse al sitio de reconocimiento y escindir el sitio de unión o alternativamente, el agente nucleasa puede unirse a una secuencia que es diferente del sitio de reconocimiento. Además, el término sitio de reconocimiento comprende tanto el sitio de unión del agente nucleasa y el sitio de mella/escisión independientemente de si el sitio de mella/escisión está dentro de o fuera del sitio de unión del agente nucleasa. En otra variación, la escisión por el agente nucleasa puede producirse en las posiciones nucleotídicas inmediatamente opuestas entre sí para producir un corte de extremo romo o, en otros casos, las incisiones pueden escalonarse para producir protuberancias monocatenarias, también llamadas "extremos cohesivos", que pueden ser protuberancias 5', o protuberancias 3'.

Cualquier agente nucleasa que induce una mella o ruptura de la doble cadena en un sitio de reconocimiento deseado puede usarse en los métodos y composiciones descritas en la presente descripción. Un agente nucleasa de origen natural o nativo puede emplearse siempre que el agente nucleasa induzca una mella o ruptura de la doble cadena en un sitio de reconocimiento deseado. Alternativamente, puede emplearse un agente nucleasa modificado o transformado mediante ingeniería genética. Un "agente nucleasa transformado mediante ingeniería genética" comprende una nucleasa que se modifica mediante ingeniería genética (modificada o derivada) a partir de su forma nativa para reconocer específicamente e inducir una mella o ruptura de la doble cadena en el sitio de reconocimiento deseado. Por lo tanto, un agente nucleasa transformado mediante ingeniería genética puede derivarse de un agente nucleasa nativo, de origen natural o puede generarse o sintetizarse artificialmente. La modificación del agente nucleasa puede ser tan pequeña como un aminoácido en un agente de escisión de proteínas o un nucleótido en un agente de escisión de ácidos nucleicos. En algunas modalidades, la nucleasa transformada mediante ingeniería genética induce una mella o ruptura de la doble cadena en un sitio de reconocimiento, en donde el sitio de reconocimiento no era una secuencia que hubiera sido reconocida por un agente nucleasa nativo (no transformado mediante ingeniería genética o no modificado). La producción de una mella o ruptura de la doble cadena en un sitio de reconocimiento u otro ADN puede denominarse en la presente descripción como "corte" o "escisión" del sitio de reconocimiento u otro ADN.

Se proporcionan, además, las variantes y fragmentos activos de los sitios de reconocimiento ejemplificados. Tales variantes activas pueden comprender al menos 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o más de identidad de secuencia con respecto al sitio de reconocimiento dado, en donde las variantes activas mantienen la actividad biológica y por lo tanto son capaces de ser reconocidos y escindidos por un agente nucleasa de una manera específica de secuencia. Los ensayos para medir la ruptura de la doble cadena de un sitio de reconocimiento por un agente nucleasa se conocen en la técnica y en general miden la capacidad de una nucleasa de cortar el sitio de reconocimiento.

El sitio de reconocimiento del agente nucleasa puede ubicarse en cualquier parte en o cerca del locus objetivo. El sitio de reconocimiento puede localizarse dentro de una región codificante de un gen, o dentro de las regiones reguladoras, que influyen en la expresión del gen. Por lo tanto, un sitio de reconocimiento del agente nucleasa puede localizarse en un intrón, un exón, un promotor, un potenciador, una región reguladora, o cualquier región no codificante de proteínas.

La descripción describe que el agente nucleasa es una Nucleasa Efectora de tipo Activador de la Transcripción (TALEN). Las nucleasas efectoras TAL son una clase de nucleasas específicas de secuencia que pueden usarse para producir rupturas de la doble cadena en secuencias objetivo específicas en el genoma de un organismo procarionta o eucarionta. Las nucleasas efectoras TAL se generan mediante la fusión de un efector de tipo activador de la transcripción (TAL) nativo o transformado mediante ingeniería genética, o parte funcional de este, al dominio catalítico de una endonucleasa, tal como, por ejemplo, *FokI*. El dominio de unión al ADN del efector TAL único, modular permite el diseño de proteínas potencialmente con cualquier especificidad de reconocimiento del ADN dada. Por lo tanto, los dominios de unión al ADN de las nucleasas efectoras TAL pueden transformarse mediante ingeniería genética para reconocer los sitios objetivo de ADN específicos y por lo tanto, usarse para producir rupturas de la doble cadena en las secuencias objetivo deseadas. Ver, el documento WO 2010/079430; Morbitzer y otros. (2010) PNAS 10.1073/pnas.1013133107; Scholze & Boch (2010) Virulence 1:428-432; Christian y otros. Genetics (2010) 186:757-761; Li y otros. (2010) Nuc. Acids Res. (2010) doi:10.1093/nar/gkq704; y Miller y otros. (2011) Nature Biotechnology 29:143-148.

Los ejemplos de nucleasas TAL adecuadas, y los métodos para preparar las nucleasas TAL adecuadas, se describen, por ejemplo, en la solicitud de patente de los Estados Unidos núm. 2011/0239315 A1, 2011/0269234 A1, 2011/0145940 A1, 2003/0232410 A1, 2005/0208489 A1, 2005/0026157 A1, 2005/0064474 A1, 2006/0188987 A1, y 2006/0063231 A1. En diversas modalidades, las nucleasas efectoras TAL se transforman mediante ingeniería genética que cortan en o cerca de una secuencia de ácidos nucleicos objetivo en, *por ejemplo*, un locus genómico de interés, en donde la secuencia de ácidos nucleicos objetivo está en o cerca de una secuencia a modificar por un vector de direccionamiento. Las nucleasas

TAL adecuadas para el uso con los varios métodos y composiciones proporcionados en la presente descripción incluyen las que se diseñan específicamente para unirse en o cerca de las secuencias de ácidos nucleicos objetivo a ser modificadas por los vectores de direccionamiento como se describe en la presente descripción.

5 La descripción describe que cada monómero de la TALEN comprende 12-25 repeticiones TAL, en donde cada repetición TAL se une a un subsitio de 1 bp. En una modalidad, el agente nucleasa es una proteína quimérica que comprende un dominio de unión al ADN basado en repeticiones TAL unido operativamente a una nucleasa independiente. En una modalidad, la nucleasa independiente es una endonucleasa FokI. En una modalidad, el agente nucleasa comprende un primer dominio de unión al ADN basado en repeticiones TAL y un segundo dominio de unión al ADN basado en repeticiones TAL, en donde cada uno de los dominios de unión al ADN basado en repeticiones TAL primero y segundo se unen operativamente a una nucleasa FokI, en donde los dominios de unión al ADN basado en repeticiones TAL primero y segundo reconocen dos secuencias de ADN objetivo contiguas en cada cadena de la secuencia de ADN objetivo separadas por aproximadamente un sitio de escisión de 6 pb a aproximadamente 40 bp, y en donde las nucleasas FokI dimerizan y generan una ruptura de la doble cadena en una secuencia objetivo.

15 La descripción describe que el agente nucleasa comprende un primer dominio de unión al ADN basado en repeticiones TAL y un segundo dominio de unión al ADN basado en repeticiones TAL, en donde cada uno de los dominios de unión al ADN basado en repeticiones TAL primero y segundo se unen operativamente a una nucleasa FokI, en donde los dominios de unión al ADN basado en repeticiones TAL primero y segundo reconocen dos secuencias de ADN objetivo contiguas en cada cadena de la secuencia de ADN objetivo separadas por un sitio de escisión de 5 pb o 6 pb, y en donde las nucleasas FokI dimerizan y generan una ruptura de la doble cadena.

25 El agente nucleasa empleado en los diversos métodos y composiciones descritas en la presente descripción pueden comprender, además, una nucleasa de dedos de zinc (ZFN). La descripción describe que cada monómero de la ZFN comprende 3 o más dominios de unión al ADN basados en dedos de zinc, en donde cada dominio de unión al ADN basado en dedos de zinc se une a un subsitio de 3 pb. En otras modalidades, la ZFN es una proteína quimérica que comprende un dominio de unión al ADN basado en dedos de zinc unido operativamente a una nucleasa independiente. La descripción describe que la endonucleasa independiente es una endonucleasa FokI. La descripción describe que el agente nucleasa comprende una primera ZFN y una segunda ZFN, en donde cada una de la primera ZFN y la segunda ZFN se une operativamente a una nucleasa FokI, en donde la primera y la segunda ZFN reconocen dos secuencias de ADN objetivo contiguas en cada cadena de la secuencia de ADN objetivo separadas por aproximadamente un sitio de escisión de 6 pb a aproximadamente 40 pb o sitio de escisión de aproximadamente a 5 pb a aproximadamente 6 pb, y en donde las nucleasas FokI dimerizan y generan una ruptura de la doble cadena. Ver, por ejemplo, el documento US20060246567; el documento US20080182332; el documento US20020081614; el documento US20030021776; el documento WO/2002/057308A2; el documento US20130123484; el documento US20100291048; y, el documento WO/2011/017293A2.

35 La descripción describe que, de los métodos proporcionados en la presente descripción, el agente nucleasa comprende (a) una proteína quimérica que comprende un dominio de unión al ADN basado en dedos de zinc fusionado a una endonucleasa FokI; o (b) una proteína quimérica que comprende una Nucleasa Efectora de tipo Activador de la Transcripción (TALEN) fusionada a una endonucleasa FokI.

45 La descripción describe que el agente nucleasa es una meganucleasa. Las meganucleasas se han clasificado en cuatro familias basado en motivos de secuencia conservados, las familias son las familias de LAGLIDADG (sec. con núm. de ident.: 16), GIY-YIG, H-N-H, y caja de His-Cys. Estos motivos participan en la coordinación de iones metálicos y en la hidrólisis de enlaces fosfodiéster. Las HEasas son notables por sus grandes sitios de reconocimiento, y por su tolerancia a algunos polimorfismos de secuencia en sus sustratos de ADN. Se conocen los dominios, estructura y función de las meganucleasas, ver por ejemplo, Guhan y Muniyappa (2003) Crit Rev Biochem Mol Biol 38:199-248; Lucas y otros, (2001) Nucleic Acids Res 29:960-9; Jurica y Stoddard, (1999) Cell Mol Life Sci 55:1304-26; Stoddard, (2006) Q Rev Biophys 38:49-95; y Moure y otros, (2002) Nat Struct Biol 9:764. En algunos ejemplos se usa una variante de origen natural, y/o derivado de meganucleasa transformado mediante ingeniería genética. Se conocen los métodos para modificar la cinética, interacciones con el cofactor, expresión, condiciones óptimas, y/o especificidad por el sitio de reconocimiento, y el tamizaje para la actividad, ver por ejemplo, Epinat y otros, (2003) Nucleic Acids Res 31:2952-62; Chevalier y otros, (2002) Mol Cell 10:895-905; Gimble y otros, (2003) Mol Biol 334:993-1008; Seligman y otros, (2002) Nucleic Acids Res 30:3870-9; Sussman y otros, (2004) J Mol Biol 342:31-41; Rosen y otros, (2006) Nucleic Acids Res 34:4791-800; Chames y otros, (2005) Nucleic Acids Res 33:e178; Smith y otros, (2006) Nucleic Acids Res 34:e149; Gruen y otros, (2002) Nucleic Acids Res 30:e29; Chen y Zhao, (2005) Nucleic Acids Res 33:e154; documento WO2005105989; documento WO2003078619; documento WO2006097854; documento WO2006097853; documento WO2006097784; y documento WO2004031346.

60 La descripción describe que puede usarse cualquier meganucleasa en la presente descripción, que incluye, pero sin limitarse a, I-SceI, I-SceII, I-SceIII, I-SceIV, I-SceV, I-SceVI, I-SceVII, I-CeuI, I-CeuAIIIP, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIIVP, I-TilI, I-PpoI, PI-PspI, F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-Amal, I-Anil, I-Chul, I-Cmoel, I-Cpal, I-Cpall, I-Csml, I-Cvul, I-CvuAIP, I-Ddil, I-Ddill, I-DirI, I-Dmol, I-Hmul, I-Hmull, I-HsNIP, I-L1al, I-Msol, I-Naal, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-Njal, I-Nsp236IP, I-PakI, I-PbolIP, I-PculIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PbpIP, I-SpBetaIP, I-ScaI, I-SexIP, I-SexIP, I-SnelIP, I-SpomI, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquilP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiSTe3bP, I-TdelIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPAIP, I-UarHGPA13P, I-VinIP, I-ZbilP,

PI-Mtul, PI-MtuHIP PI-MtuHIIP, PI-Pful, PI-Pfull, PI-Pkol, PI-Pkoll, PI-Rma43812IP, PI-SpBetaIP, PI-Scel, PI-Tful, PI-Tfull, PI-Thyl, PI-Tlil, PI-Tlill, o cualquiera de las variantes o fragmentos de estas.

5 La descripción describe que la meganucleasa reconoce secuencias de ADN bicatenario de 12 a 40 pares de bases. En una modalidad, la meganucleasa reconoce una secuencia objetivo perfectamente coincidente en el genoma. En una modalidad, la meganucleasa es una nucleasa de asentamiento. En una modalidad, la nucleasa de asentamiento es una familia de nucleasas de asentamiento de LAGLIDADG (sec. con núm. de ident.: 16). En una modalidad, la familia de nucleasas de asentamiento de LAGLIDADG (sec. con núm. de ident.: 16) se selecciona de I-Scel, I-Crel, y I-Dmol.

10 Los agentes nucleasas pueden comprender, además, endonucleasas de restricción, que incluyen endonucleasas Tipo I, Tipo II, Tipo III, y Tipo IV. Las endonucleasas de restricción de Tipo I y Tipo III reconocen sitios de reconocimiento específicos, pero escinden típicamente en una posición variable del sitio de unión a la nucleasa, que puede estar a cientos de pares de bases del sitio de escisión (sitio de reconocimiento). En los sistemas de Tipo II la actividad de restricción es independiente de cualquier actividad metilasa, y la escisión se produce típicamente en sitios específicos dentro de o cerca
15 del sitio de unión. La mayoría de las enzimas de Tipo II cortan secuencias palindrómicas, sin embargo, las enzimas de Tipo IIa reconocen sitios de reconocimiento no palindrómicos y escinden fuera del sitio de reconocimiento, las enzimas de Tipo IIb cortan secuencias dos veces con ambos sitios fuera del sitio de reconocimiento, y las enzimas de Tipo IIc reconocen un sitio de reconocimiento asimétrico y escinden en un lado y a una distancia definida de aproximadamente 1-
20 20 nucleótidos del sitio de reconocimiento. Las enzimas de restricción de Tipo IV se dirigen al ADN metilado. Las enzimas de restricción se describen, además, y se clasifican, por ejemplo, en la base de datos REBASE (página web en rebase.neb.com; Roberts y otros, (2003) *Nucleic Acids Res* 31:418-20), Roberts y otros, (2003) *Nucleic Acids Res* 31:1805-12, y Belfort y otros, (2002) en *Mobile ADN II*, páginas 761-783, Eds. Craigie y otros, (ASM Press, Washington, DC).

25 El agente nucleasa empleado en los varios métodos y composiciones puede comprender, además, un sistema CRISPR/Cas. Tales sistemas pueden emplear, por ejemplo, una nucleasa Cas9, que, en algunos casos, se optimiza en codones para el tipo celular deseado en el cual va a expresarse. Tales sistemas pueden emplear, además, un ARN guía (ARNg) que comprende dos moléculas separadas. Un ARNg de dos moléculas ilustrativo comprende una molécula de tipo ARNcr ("ARN de CRISPR" o "ARN transformador" o "ARNcr" o "repetición de ARNcr") y una molécula de tipo ARNtracr correspondiente ("ARN de transactivación de CRISPR" o "ARN activador" o "ARNtracr" o "andamio"). Un ARNcr comprende tanto el segmento de transformación de ADN (monocatenario) del ARNg como un tramo de nucleótidos que forma una mitad de un dúplex de ARN bicatenario (ARNbc) del segmento de unión a proteínas del ARNg. Un ARNtracr correspondiente (ARN activador) comprende un tramo de nucleótidos que forma la otra mitad del dúplex de ARNbc del segmento de unión a proteínas del ARNg. Por lo tanto, un tramo de nucleótidos de un ARNcr es complementario a y se
30 hibrida con un tramo de nucleótidos de un ARNtracr para formar el dúplex de ARNbc del dominio de unión a proteínas del ARNg. Como tal, puede decirse que cada ARNcr tiene un ARNtracr correspondiente. El ARNcr proporciona adicionalmente el segmento de transformación de ADN monocatenario. En consecuencia, un ARNg comprende una secuencia que se hibrida a una secuencia objetivo, y un ARNtracr. Por lo tanto, un ARNcr y un ARNtracr (como un par correspondiente) se hibridan para formar un ARNg. Si se usa para la modificación dentro de una célula, la secuencia y/o longitud exacta de una molécula de ARNcr o ARNtracr dada puede diseñarse para que sea específica a las especies en las que se usarán las moléculas de ARN.

45 Los genes de origen natural que codifican los tres elementos (Cas9, ARNtracr y ARNcr) se organizan, típicamente, en operón(es). Los ARN de origen natural de CRISPR difieren en dependencia del sistema Cas9 y el organismo, pero contienen, frecuentemente, un segmento de transformación de entre 21 a 72 nucleótidos de longitud, flanqueado por dos repeticiones directas (DR) de una longitud de entre 21 a 46 nucleótidos (ver, por ejemplo, el documento WO2014/131833). En el caso de *S. pyogenes*, las DR son de 36 nucleótidos de largo y el segmento de transformación es de 30 nucleótidos de largo. La DR ubicada en dirección 3' es complementaria a y se hibrida con el ARNtracr correspondiente, el que a su vez se une a la proteína Cas9.

50 Alternativamente, el sistema emplea además un constructo de ARNcr-ARNtracr fusionado (es decir, un transcrito único) que funciona con la Cas9 optimizada en codones. Este ARN único se denomina frecuentemente como ARN guía o ARNg. Dentro de un ARNg, la porción de ARNcr se identifica como la 'secuencia objetivo' para el sitio de reconocimiento dado y el ARNtracr se denomina frecuentemente como el 'andamio.' En resumen, un fragmento de ADN corto que contiene la secuencia objetivo se inserta en un plásmido de expresión de ARN guía. El plásmido de expresión de ARNg comprende la secuencia objetivo (en algunas modalidades alrededor de 20 nucleótidos), una forma de la secuencia de ARNtracr (el andamio) así como también un promotor adecuado que es activo en la célula y los elementos necesarios para el procesamiento apropiado en célula eucariotas. Muchos de los sistemas se basan en oligos personalizados, complementarios que se aparean para formar un ADN bicatenario y después se clonan en el plásmido de expresión de ARNg. El casete de expresión de ARNg y el casete de expresión de Cas9 se introducen después en la célula. Ver, por ejemplo, Mali P y otros. (2013) *Science* 2013 Feb 15;339(6121):823-6; Jinek M y otros. *Science* 2012 Ago 17;337(6096):816-21; Hwang WY y otros. *Nat Biotechnol* 2013 Mar;31(3):227-9; Jiang W y otros. *Nat Biotechnol* 2013 Mar;31(3):233-9; y Cong L y otros. *Science* 2013 Feb 15;339(6121):819-23. Ver también, por ejemplo, los documentos WO/2013/17672A1, WO/2014/065596A1, WO/2014/089290A1, WO/2014/093622A2, WO/2014/099750A2, y
65 WO/2013142578A1.

- 5 En algunas modalidades, la nucleasa Cas9 puede proporcionarse en la forma de una proteína. En algunas modalidades, la proteína Cas9 puede proporcionarse en la forma de un complejo con el ARNg. En otras modalidades, la nucleasa Cas9 puede proporcionarse en la forma de un ácido nucleico que codifica la proteína. El ácido nucleico que codifica la nucleasa Cas9 puede ser ARN (por ejemplo, ARN mensajero (ARNm)) o ADN.
- 10 En algunas modalidades, el ARNg puede proporcionarse en la forma de ARN. En otras modalidades, el ARNg puede proporcionarse en la forma de ADN que codifica el ARN. En algunas modalidades, el ARNg puede proporcionarse en la forma de moléculas de ARNcr y de ARNtracr separadas, o moléculas de ADN separadas que codifican el ARNcr y el ARNtracr, respectivamente.
- 15 En una modalidad, el método para modificar un locus genómico de interés en una célula comprende, además, introducir en la célula: (a) un primer constructo de expresión que comprende un primer promotor unido operativamente a una primera secuencia de ácidos nucleicos que codifica una proteína (Cas) asociada a Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR); (b) un segundo constructo de expresión que comprende un segundo promotor unido operativamente a una secuencia objetivo genómica unida a un ARN guía (ARNg), en donde la secuencia objetivo genómica está flanqueada por un Motivo Adyacente a los Protoespaciadores. Opcionalmente, la secuencia objetivo genómica está flanqueada en el extremo 3' por una secuencia de un Motivo Adyacente a los Protoespaciadores (PAM). La descripción describe que la célula comprende una célula eucariota, una célula eucariota no de rata, una célula de mamífero, una célula humana, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto, o una célula CHO.
- 20 En una modalidad, la secuencia objetivo genómica comprende la secuencia nucleotídica de GNNNNNNNNNNNNNNNNNNNGG (GN₁₋₂₀GG; sec. con núm. de ident.: 1). En una modalidad, la secuencia objetivo genómica comprende la sec. con núm. de ident.: 23, en donde N es de entre 1 y 20 nucleótidos de longitud. En otra modalidad, la secuencia objetivo genómica comprende entre 14 y 20 nucleótidos de longitud de la sec. con núm. de ident.: 1.
- 25 En una modalidad, el ARNg comprende una tercera secuencia de ácidos nucleicos que codifica un ARN de Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (ARNcr) y un ARN de CRISPR de transactivación (ARNtracr). En modalidades específicas, la proteína Cas es Cas9.
- 30 En algunas modalidades, el ARNg comprende (a) el ARN quimérico de la secuencia de ácidos nucleicos 5'-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAU AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUU UU-3' (sec. con núm. de ident.: 2); o (b) el ARN quimérico de la secuencia de ácidos nucleicos 5'-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCG-3' (sec. con núm. de ident.: 3).
- 35 En otra modalidad, el ARNcr comprende 5'-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAU-3' (sec. con núm. de ident.: 4); 5'-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAG (sec. con núm. de ident.: 5); o 5'-GAGUCCGAGCAGAAGAAGAAGUUUA-3' (sec. con núm. de ident.: 6).
- 40 Aún en otras modalidades, el ARNtracr comprende, 5'-AAGGCUAGUCCG-3' (sec. con núm. de ident.: 7) o 5'-AAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU-3' (sec. con núm. de ident.: 8).
- 45 En una modalidad, la proteína Cas es una proteína Cas de Tipo I. En una modalidad, la proteína Cas es una proteína Cas de Tipo II. En una modalidad, la proteína Cas de Tipo II es Cas9. En una modalidad, la primera secuencia de ácidos nucleicos codifica una proteína Cas humana optimizada en codones.
- 50 En determinadas modalidades, la proteína Cas es una "enzima generadora de mellas" que puede generar rupturas de una única cadena (es decir, "mellas") en el sitio objetivo sin cortar ambas cadenas del ADN bicatenario (ADNbc). Cas9, por ejemplo, comprende dos dominios nucleasa-un dominio nucleasa de tipo RuvC y un dominio nucleasa de tipo HNH-que son responsables de la escisión de cadenas de ADN opuestas. La mutación en cualquiera de estos dominios puede generar una enzima generadora de mellas. Los ejemplos de mutaciones que generan enzimas generadoras de mellas pueden encontrarse, por ejemplo, en los documentos WO/2013/176772A1 y WO/2013/142578A1.
- 55 En determinadas modalidades, dos proteínas Cas separadas (por ejemplo, enzimas generadoras de mellas) específicas para un sitio objetivo en cada cadena del ADNbc pueden generar secuencias protuberantes complementarias a secuencias protuberantes en otro ácido nucleico, o una región separada en el mismo ácido nucleico. Los extremos protuberantes generados cuando se pone en contacto un ácido nucleico con dos enzimas generadoras de mellas específicas de los sitios objetivo en ambas cadenas de ADNbc pueden ser extremos protuberantes ya sean 5' o 3'. Por ejemplo, una primera enzima generadora de mellas puede generar una ruptura de cadena única en la primera cadena de ADNbc, mientras que una segunda enzima generadora de mellas puede generar una ruptura de cadena única en la segunda cadena de ADNbc de manera que se generan secuencias protuberantes. Los sitios objetivo de cada enzima
- 60
- 65

5 generadora de mellas que generan la ruptura de cadena única pueden seleccionarse de manera que las secuencias de extremos protuberantes generadas sean complementarias a secuencias de extremos protuberantes en una molécula de ácido nucleico diferente. Los extremos protuberantes complementarios de las dos moléculas de ácido nucleico diferentes pueden aparearse mediante los métodos descritos en la presente descripción. En algunas modalidades, el sitio objetivo de la enzima generadora de mellas en la primera cadena es diferente del sitio objetivo de la enzima generadora de mellas en la segunda cadena.

10 En una modalidad, el primer ácido nucleico comprende una mutación que altera al menos un residuo de aminoácido de los sitios activos de nucleasa en la proteína Cas, en donde la proteína Cas mutante genera una ruptura en solo una cadena de la región de ADN objetivo, y en donde la mutación disminuye la recombinación no homóloga en la región de ADN objetivo.

15 En una modalidad, el primer ácido nucleico que codifica la proteína Cas comprende, además, una señal de localización nuclear (NLS). En una modalidad, la señal de localización nuclear es una señal de localización nuclear de SV40.

20 En una modalidad, el segundo promotor que conduce la expresión de la secuencia objetivo genómica y el ARN guía (ARNg) es un promotor de ARN polimerasa III. En una modalidad, el promotor de ARN polimerasa III es un promotor U6 humano. En una modalidad, el promotor de ARN polimerasa III es un promotor de U6 polimerasa III de rata. En una modalidad, el promotor de ARN polimerasa III es un promotor de U6 polimerasa III de ratón.

25 En una modalidad, las secuencias de ácidos nucleicos que codifican el ARNcr y el ARNtracr se unen a través de un bucle sintético, en donde, tras la expresión, el ARNcr y el ARNtracr forman un dúplex de ARNcr:ARNtracr.

30 Esta descripción describe que el sistema CRISPR/Cas como se describió anteriormente puede usarse en combinación con vectores de direccionamiento grandes con cualquiera de los siguientes tipos celulares: una célula eucariota, una célula eucariota no de rata, una célula de mamífero, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto o una célula CHO.

35 En una modalidad, el primer constructo de expresión y el segundo constructo de expresión se expresan a partir de un mismo plásmido.

40 En una modalidad, los constructos de expresión primero y segundo se introducen junto con el LTVEC. En una modalidad, los constructos de expresión primero y segundo se introducen por separado del LTVEC en un periodo de tiempo.

45 En una modalidad, el método comprende introducir una pluralidad del segundo constructo y una pluralidad del LTVEC para la edición múltiple de los loci objetivo diferentes como se describe en la presente descripción.

50 Se proporcionan, además, variantes y fragmentos activos de los agentes nucleasas (es decir, un agente nucleasa transformado mediante ingeniería genética). Tales variantes activas pueden comprender al menos 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o más de identidad de secuencia con respecto a los agentes nucleasas nativos, en donde las variantes activas mantienen la capacidad de cortar en un sitio de reconocimiento deseado y por lo tanto mantienen la actividad de inducir mella o ruptura de la doble cadena. Por ejemplo, cualquiera de los agentes nucleasas descritos en la presente descripción pueden modificarse a partir de una secuencia endonucleasa nativa y diseñarse para reconocer e inducir una mella o ruptura de la doble cadena en un sitio de reconocimiento que no es reconocido por los agentes nucleasas nativos. Por lo tanto, en algunas modalidades, la nucleasa transformada mediante ingeniería genética tiene una especificidad para inducir una mella o ruptura de la doble cadena en un sitio de reconocimiento que es diferente del sitio de reconocimiento del agente nucleasa nativo correspondiente. Los ensayos para la actividad de inducción de una mella o ruptura de la doble cadena se conocen y en general miden la actividad global y la especificidad de la endonucleasa sobre los sustratos de ADN que contienen el sitio de reconocimiento.

55 El agente nucleasa puede introducirse en la célula mediante los medios conocidos en la técnica. El polipéptido que codifica el agente nucleasa puede introducirse directamente en la célula. Alternativamente, un polinucleótido que codifica el agente nucleasa puede introducirse en la célula. Cuando un polinucleótido que codifica el agente nucleasa se introduce en la célula, el agente nucleasa puede expresarse transitoriamente, condicionalmente o constitutivamente dentro de la célula. Por lo tanto, el polinucleótido que codifica el agente nucleasa puede estar contenido en un casete de expresión y puede unirse operativamente a un promotor condicional, un promotor inducible, un promotor constitutivo, o un promotor específico de tejidos. Tales promotores de interés se analizan con mayor detalle en cualquier otra parte de la presente descripción. Alternativamente, el agente nucleasa se introduce en la célula como un ARNm que codifica o que comprende un agente nucleasa.

60 En una modalidad, el ARNcr y el ARNtracr se expresan como transcritos de ARN separados.

65

En modalidades específicas, el polinucleótido que codifica el agente nucleasa se integra de manera estable en el genoma de la célula y se une operativamente a un promotor activo en la célula. En otras modalidades, el polinucleótido que codifica el agente nucleasa está en el mismo vector de direccionamiento que comprende el inserto de ácido nucleico, mientras que en otros casos el polinucleótido que codifica el agente nucleasa está en un vector o un plásmido que está separado del vector de direccionamiento que comprende el inserto de ácido nucleico.

Cuando el agente nucleasa se proporciona a la célula mediante la introducción de un polinucleótido que codifica el agente nucleasa, un polinucleótido tal que codifica un agente nucleasa puede modificarse para sustituir codones que tienen una mayor frecuencia de uso en la célula de interés, en comparación con la secuencia de polinucleótidos de origen natural que codifica el agente nucleasa. Por ejemplo el polinucleótido que codifica el agente nucleasa puede modificarse para sustituir codones que tienen una mayor frecuencia de uso en una célula procarionota o eucariota de interés, incluida una célula bacteriana, una célula de levadura, una célula humana, una célula no humana, una célula eucariota no de rata, una célula de mamífero, una célula de roedor, una célula de roedor no rata, una célula de ratón, una célula de rata, una célula de hámster o cualquier otra célula huésped de interés, en comparación con la secuencia de polinucleótidos de origen natural.

En una modalidad, el agente endonucleasa se introduce junto con el LTVEC. En una modalidad, el agente endonucleasa se introduce por separado del LTVEC en un periodo de tiempo. En una modalidad, el agente endonucleasa se introduce antes de la introducción del LTVEC. En una modalidad, el agente endonucleasa se introduce en la célula ES de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster después de la introducción del LTVEC.

En una modalidad, el agente endonucleasa es un constructo de expresión que comprende una secuencia de ácidos nucleicos que codifica una endonucleasa, en donde la secuencia de ácidos nucleicos se une operativamente a un promotor. En una modalidad, el promotor es un promotor constitutivamente activo. En una modalidad, el promotor es un promotor inducible. En una modalidad, el promotor es activo en la célula pluripotente o no pluripotente de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster. En una modalidad, el agente endonucleasa es un ARNm que codifica una endonucleasa.

B. Métodos para integrar un polinucleótido de interés en un locus objetivo

Se proporcionan los métodos para modificar un locus objetivo de interés. Esta descripción describe que un locus objetivo en una célula pluripotente o no pluripotente de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster es objetivo de modificación genética. Un método tal comprende: (a) introducir en la célula pluripotente o no pluripotente de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster y un brazo de homología 3' de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster; y (b) identificar una célula pluripotente o no pluripotente genéticamente modificada de rata, eucariota, eucariota no de rata, de mamífero, de mamífero no ser humano, ser humano, de roedor, de roedor no rata, de ratón o de hámster que comprende la modificación genética dirigida en el locus objetivo, en donde la modificación genética dirigida es capaz de transmitirse a través de la línea germinal. En modalidades específicas, la suma total del brazo de homología 5' y el brazo de homología 3' es de al menos 10 kb y/o se emplea un vector de direccionamiento grande.

En otras modalidades, el tamaño de la suma total del total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 150 kb, aproximadamente 10 kb a aproximadamente 100 kb, aproximadamente 10 kb a aproximadamente 75 kb, aproximadamente 20 kb a aproximadamente 150 kb, aproximadamente 20 kb a aproximadamente 100 kb, aproximadamente 20 kb a aproximadamente 75 kb, aproximadamente 30 kb a aproximadamente 150 kb, aproximadamente 30 kb a aproximadamente 100 kb, aproximadamente 30 kb a aproximadamente 75 kb, aproximadamente 40 kb a aproximadamente 150 kb, aproximadamente 40 kb a aproximadamente 100 kb, aproximadamente 40 kb a aproximadamente 75 kb, aproximadamente 50 kb a aproximadamente 150 kb, aproximadamente 50 kb a aproximadamente 100 kb, o aproximadamente 50 kb a aproximadamente 75 kb, aproximadamente 10 kb a aproximadamente 30 kb, aproximadamente 20 kb a aproximadamente 40 kb, aproximadamente 40 kb a aproximadamente 60 kb, aproximadamente 60 kb a aproximadamente 80 kb, aproximadamente 80 kb a aproximadamente 100 kb, aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a aproximadamente 150 kb. En una modalidad, el tamaño de la delección es igual o similar al tamaño de la suma total de los brazos de homología 5' y 3' del LTVEC.

La célula pluripotente, por ejemplo, una célula de rata, puede ser una célula madre embrionaria, por ejemplo, una célula madre embrionaria de rata. Esta descripción describe que (a) la célula ES de rata se deriva de una cepa DA o una cepa AC1; o (b) la célula ES de rata se caracteriza por la expresión de un marcador de pluripotencia que comprende Oct-4, Sox-2, fosfatasa alcalina, o una combinación de estos. En otros casos, la célula madre embrionaria de rata empleada comprende una célula ES de rata como se describe en la solicitud de patente de los Estados Unidos núm. 14/185,103, presentada el 20 de febrero de 2014.

Esta descripción describe que cualquier célula pluripotente o no pluripotente puede usarse en los métodos proporcionados en la presente descripción. Por ejemplo, la célula pluripotente o no pluripotente puede ser de un eucariota, un eucariota no rata, un mamífero no ser humano, un mamífero, un roedor, un roedor no rata, una rata, un ratón, un ser humano o un hámster.

5 Como se describe en cualquier otra parte de la presente descripción, el inserto de ácido nucleico puede ser cualquier secuencia de ácidos nucleicos. Esta descripción describe que (a) el inserto de ácido nucleico comprende una sustitución de una secuencia de ácidos nucleicos endógena de rata, eucariota, eucariota no rata, mamífero, ser humano, roedor, roedor no rata, ratón o hámster con una secuencia de ácidos nucleicos homóloga u ortóloga de mamífero; (b) el inserto de ácido nucleico comprende una delección de una secuencia de ácidos nucleicos endógena de rata, eucariota, eucariota no rata, mamífero, ser humano, roedor, roedor no rata, ratón o hámster; (c) el inserto de ácido nucleico comprende una delección de una secuencia de ácidos nucleicos endógena de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster, en donde la delección varía de 5 kb a 200 kb o de 5 kb a 3 Mb (como se analiza en detalle en cualquier otra parte de la presente descripción); (d) el inserto de ácido nucleico comprende una adición de una secuencia de ácidos nucleicos exógena (que incluye por ejemplo una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb); (e) el inserto de ácido nucleico comprende una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u ortóloga; (f) la secuencia de ácidos nucleicos homóloga u ortóloga de (a) en donde la secuencia de ácidos nucleicos es una secuencia de ácidos nucleicos humana; (g) el inserto de ácido nucleico comprende la secuencia de ácidos nucleicos homóloga u ortóloga de (a) en donde la secuencia de ácidos nucleicos es una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos de humana y una de rata; (h) el inserto de ácido nucleico comprende la secuencia de ácidos nucleicos exógena (e), en donde el inserto de ácido nucleico varía de aproximadamente 5 kb a aproximadamente 200 kb; (i) el inserto de ácido nucleico comprende un alelo condicional flanqueado por secuencias objetivo de recombinasa específica de sitio; (j) el inserto de ácido nucleico comprende un gen reportero unido operativamente a un promotor; (k) el inserto de ácido nucleico comprende uno o más segmentos del gen de V_H de cadena pesada de inmunoglobulina humana no reordenados, uno o más segmentos del gen de D de cadena pesada de inmunoglobulina humana no reordenados, y uno o más segmentos del gen de J_H de cadena pesada de inmunoglobulina humana no reordenados, que se unen operativamente a una secuencia de ácidos nucleicos de región constante de cadena pesada de roedor; (l) el inserto de ácido nucleico comprende una secuencia de ácidos nucleicos de la región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácidos nucleicos de región constante de cadena pesada de roedor; (m) el inserto de ácido nucleico comprende uno o más segmentos del gen de V_K o V_λ de inmunoglobulina humana y uno o más segmentos del gen de J_K o J_λ de inmunoglobulina humana no reordenados, que se unen operativamente a una secuencia de ácidos nucleicos de región constante de cadena ligera λ o κ de inmunoglobulina de mamífero; (n) el inserto de ácido nucleico comprende una secuencia de ácidos nucleicos de región variable de cadena ligera λ o κ de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácidos nucleicos de región constante de cadena ligera λ o κ de inmunoglobulina de mamífero; (o) la secuencia de ácidos nucleicos de región constante de cadena pesada de mamífero de (k) y/o (l) comprende una secuencia de ácidos nucleicos de región constante de rata, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos; o (p) el ácido nucleico de región constante de cadena ligera λ o κ de inmunoglobulina de mamífero de (m) y/o (n) comprende una secuencia de ácidos nucleicos de región constante de rata, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos.

50 En una modalidad, el inserto de ácido nucleico comprende uno o más segmentos del gen de V_H humano funcional que comprenden V_{H1-2} , V_{H1-3} , V_{H1-8} , V_{H1-18} , V_{H1-24} , V_{H1-45} , V_{H1-46} , V_{H1-58} , V_{H1-69} , V_{H2-5} , V_{H2-26} , V_{H2-70} , V_{H3-7} , V_{H3-9} , V_{H3-11} , V_{H3-13} , V_{H3-15} , V_{H3-16} , V_{H3-20} , V_{H3-21} , V_{H3-23} , V_{H3-30} , $V_{H3-30-3}$, $V_{H3-30-5}$, V_{H3-33} , V_{H3-35} , V_{H3-38} , V_{H3-43} , V_{H3-48} , V_{H3-49} , V_{H3-53} , V_{H3-64} , V_{H3-66} , V_{H3-72} , V_{H3-73} , V_{H3-74} , V_{H4-4} , V_{H4-28} , $V_{H4-30-1}$, $V_{H4-30-2}$, $V_{H4-30-4}$, V_{H4-31} , V_{H4-34} , V_{H4-39} , V_{H4-59} , V_{H4-61} , V_{H5-51} , V_{H6-1} , V_{H7-4-1} , V_{H7-81} , o una combinación de estos.

55 En una modalidad, el inserto de ácido nucleico comprende uno o más segmentos del gen de D humano funcional que comprenden $D1-1$, $D1-7$, $D1-14$, $D1-20$, $D1-26$, $D2-2$, $D2-8$, $D2-15$, $D2-21$, $D3-3$, $D3-9$, $D3-10$, $D3-16$, $D3-22$, $D4-4$, $D4-11$, $D4-17$, $D4-23$, $D5-12$, $D5-5$, $D5-18$, $D5-24$, $D6-6$, $D6-13$, $D6-19$, $D6-25$, $D7-27$, o una combinación de estos.

60 En una modalidad, el inserto de ácido nucleico comprende uno o más segmentos del gen de J_H funcional que comprenden J_{H1} , J_{H2} , J_{H3} , J_{H4} , J_{H5} , J_{H6} , o una combinación de estos. En una modalidad, el inserto de ácido nucleico comprende uno o más segmentos del gen de V_K humano que comprenden V_{K4-1} , V_{K5-2} , V_{K7-3} , V_{K2-4} , V_{K1-5} , V_{K1-6} , V_{K3-7} , V_{K1-8} , V_{K1-9} , V_{K2-10} , V_{K3-11} , V_{K1-12} , V_{K1-13} , V_{K2-14} , V_{K3-15} , V_{K1-16} , V_{K1-17} , V_{K2-18} , V_{K2-19} , V_{K3-20} , V_{K6-21} , V_{K1-22} , V_{K1-23} , V_{K2-24} , V_{K3-25} , V_{K2-26} , V_{K1-27} , V_{K2-28} , V_{K2-29} , V_{K2-30} , V_{K3-31} , V_{K1-32} , V_{K1-33} , V_{K3-34} , V_{K1-35} , V_{K2-36} , V_{K1-37} , V_{K2-38} , V_{K1-39} , V_{K2-40} , o una combinación de estos.

65

En una modalidad, el inserto de ácido nucleico comprende uno o más segmentos del gen de Vλ humano que comprenden Vλ3-1, Yλ4-3, Yλ2-8, Yλ3-9, Vλ3-10, Vλ2-11, Vλ3-12, Vλ2-14, Vλ3-16, Vλ2-18, Vλ3-19, Vλ3-21, Vλ3-22, Vλ2-23, Vλ3-25, Vλ3-27, o una combinación de estos.

5 En una modalidad, el inserto de ácido nucleico comprende uno o más segmentos del gen de JK humano que comprenden Jk1, Jk2, Jk3, Jk4, Jk5, o una combinación de estos.

La descripción describe que tras la modificación del locus objetivo en una célula pluripotente o no pluripotente de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster, la modificación genética se transmite a través de la línea germinal.

10 La descripción describe que la secuencia de ácidos nucleicos del inserto comprende un polinucleótido que cuando se integra en el genoma producirá una modificación genética de una región del locus de *ApoE* de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster, en donde la modificación genética en el locus de *ApoE* da como resultado una disminución de la actividad de ApoE, una disminución de la actividad de ApoE o una modulación de la actividad de ApoE. En una modalidad, se genera una inactivación de *ApoE*.

15 La descripción describe que la secuencia de ácidos nucleicos del inserto comprende un polinucleótido que cuando se integra en el genoma producirá una modificación genética de una región del locus de la cadena gamma de receptor de interleucina-2 de rata, eucariota, eucariota no rata, mamífero, humano, mamífero no humano, roedor, roedor no rata, ratón o hámster, en donde la modificación genética en el locus de la cadena gamma de receptor de interleucina-2 da como resultado una disminución de la actividad del receptor de interleucina 2, una disminución de la actividad de la cadena gamma de receptor de interleucina-2, o una modulación de la actividad del receptor de interleucina 2. En una modalidad, se genera una inactivación del receptor de interleucina 2.

20 La descripción describe que la secuencia de ácidos nucleicos del inserto comprende un polinucleótido que cuando se integra en el genoma producirá una modificación genética de una región del locus de *Rag1* de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster, del locus de *Rag2* de rata, eucariota, eucariota no rata, mamífero no humano, mamífero, ser humano, roedor, roedor no rata, ratón o hámster y/o del locus de *Rag2/Rag1* de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster, en donde la modificación genética en el locus de *Rag1*, *Rag2* y/o *Rag2/Rag1* de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster da como resultado una disminución de la actividad de las proteínas *Rag1*, *Rag2* o *Rag1* y *Rag2*, una disminución de la actividad de las proteínas *Rag1*, *Rag2* o *Rag1* y *Rag2*, o una modulación de la actividad de las proteínas *Rag1*, *Rag2* o *Rag1* y *Rag2*. En una modalidad, se genera una inactivación de *Rag1*, *Rag2* o *Rag2/Rag1*.

25 La descripción describe que el inserto de ácido nucleico da como resultado la sustitución de una porción del locus de *ApoE* de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster, del locus de la cadena gamma de receptor de interleucina-2 y/o locus de *Rag2*, y/o locus de *Rag1* y/o locus de *Rag2/Rag1* con la porción ortóloga correspondiente de un locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2, un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* de otro organismo.

30 Aún en otras modalidades, el inserto de ácido nucleico comprende un polinucleótido que comparte en toda su longitud al menos 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % con respecto a una porción de un locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2, un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* que reemplaza.

35 La descripción describe que el inserto de polinucleótido dado y la región correspondiente del locus de rata, eucariota, eucariota no rata, mamífero, mamífero no humano, ser humano, roedor, roedor no rata, ratón o hámster que se reemplaza puede ser una región codificante, un intrón, un exón, una región no traducida, una región reguladora, un promotor, o un potenciador o cualquier combinación de estos. Además, el inserto de polinucleótido dado y/o la región del locus de rata, eucariota, eucariota no rata, mamífero, ser humano, mamífero no humano, roedor, roedor no rata, ratón o hámster que se reemplaza puede ser de cualquier longitud deseada, que incluye por ejemplo, entre 10-100 nucleótidos de longitud, 100-500 nucleótidos de longitud, 500-1 kb de nucleótidos de longitud, 1 kb a 1,5 kb de nucleótidos de longitud, 1,5 kb a 2 kb nucleótidos de longitud, 2 kb a 2,5 kb nucleótidos de longitud, 2,5 kb a 3 kb nucleótidos de longitud, 3 kb a 5 kb nucleótidos de longitud, 5 kb a 8 kb nucleótidos de longitud, 8 kb a 10 kb nucleótidos de longitud o más. En otros casos, el tamaño de la inserción o la sustitución es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, de aproximadamente 350 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 800 kb, de aproximadamente 800 kb a 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb, de aproximadamente 2,5 Mb, de aproximadamente 2,5 Mb a aproximadamente 2,8 Mb, de aproximadamente 2,8 Mb a aproximadamente 3 Mb. En otras modalidades, el inserto de polinucleótido dado y/o la región del locus de rata, eucariota, eucariota no rata, mamífero no humano, mamífero, ser humano, roedor, roedor no rata, ratón

o hámster que se reemplaza es de al menos 100, 200, 300, 400, 500, 600, 700, 800, o 900 nucleótidos o al menos 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 15 kb, 16 kb o más.

i. Métodos para modificar un locus objetivo de un ácido nucleico a través de recombinación homóloga bacteriana (BHR)

Esta descripción proporciona métodos y composiciones para modificar un locus objetivo de un ácido nucleico de eucariota, eucariota no rata, mamífero, ser humano o mamífero no humano, a través de la recombinación homóloga bacteriana (BHR) en una célula procariota. Tales métodos encuentran uso en la utilización de la recombinación homóloga bacteriana en una célula procariota para modificar genéticamente un locus objetivo de un ácido nucleico de eucariota, eucariota no rata, mamífero, ser humano o mamífero no humano para generar un vector de direccionamiento. Un vector de direccionamiento tal que comprende el locus objetivo modificado genéticamente puede introducirse en una célula eucariota, por ejemplo, una célula eucariota, célula eucariota no de rata, una célula de mamífero, una célula humana, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto, o una célula CHO. La "recombinación homóloga" incluye el intercambio de fragmentos de ADN entre dos moléculas de ADN de sitios de entrecruzamiento dentro de regiones de homología. Por lo tanto, "recombinación homóloga bacteriana" o "BHR" incluye la recombinación homóloga que se produce en bacterias.

Se proporcionan en la presente descripción los métodos para modificar un locus objetivo de un ácido nucleico de una célula eucariota, célula eucariota no de rata, una célula de mamífero, una célula humana, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto, o una célula CHO a través de la recombinación homóloga bacteriana (BHR). Los métodos comprenden introducir en una célula procariota un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde la célula procariota comprende un locus objetivo de un ácido nucleico y es capaz de expresar una recombinasa que media la BHR en el locus objetivo. Tales vectores de direccionamiento pueden incluir cualquiera de los vectores de direccionamiento grandes descritos en la presente descripción.

En una modalidad, el método comprende introducir en una célula procariota: (i) un primer constructo que comprende un ácido nucleico que tiene una secuencia de ADN de interés; (ii) un segundo constructo de transformación que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', y (iii) un tercer constructo que codifica una recombinasa que media la recombinación homóloga bacteriana. En una modalidad, el primer, el segundo, y el tercer constructo se introducen en la célula procariota por separado en un periodo de tiempo. En una modalidad, la célula procariota comprende un ácido nucleico que codifica la recombinasa, y el método no requiere la introducción del tercer constructo. En una modalidad, la recombinasa se expresa bajo el control de un promotor inducible.

En una modalidad el primer constructo que comprende el ácido nucleico, se deriva de un cromosoma artificial bacteriano (BAC) o un cromosoma artificial de levadura (YAC). Puede seleccionarse una célula procariota que comprende el inserto de ácido nucleico en el locus genómico objetivo. Este método puede repetirse en serie como se describe en la presente descripción para permitir la introducción de múltiples insertos de ácido nucleicos en el locus a transformar en la célula procariota. Una vez que el locus de ácido nucleico objetivo se "construye" dentro de la célula procariota, puede aislarse un vector de direccionamiento que comprende el locus objetivo modificado de la célula procariota e introducirse en un locus genómico objetivo dentro de una célula eucariota, célula eucariota no de rata, una célula de mamífero, una célula humana, una célula de mamífero no humana, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula iPS humana, una célula humana, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster, un fibroblasto, o una célula CHO de esta descripción.

Las células de ratas preferidas para recibir los vectores de direccionamiento de esta descripción se describen en la solicitud de los Estados Unidos 14/185,703, presentada el 20 de febrero de 2014, cuyos contenidos se resumen en la presente descripción. Estas células de ratas son células pluripotentes de ratas capaces de mantener su pluripotencia después de una o más modificaciones genéticas dirigidas *in vitro*, y son capaces de transmitir las modificaciones genéticas dirigidas a través de la línea germinal.

Las células pluripotentes electroporadas, por ejemplo, se cultivan en placas a una alta densidad para la selección de células resistentes a fármacos que comprenden el vector de direccionamiento. El proceso de selección de fármacos elimina la mayoría de las células sembradas (~99 %), dejando atrás colonias individuales, cada una de las cuales es un clon derivado de una célula única. De las células restantes, la mayoría de las células (~ 80-100 %) contienen el vector de direccionamiento (que comprende un casete de selección por fármacos) integrado a una localización aleatoria en el genoma. Por lo tanto, las colonias se seleccionan individualmente y se genotipan para identificar células ES que portan

el vector de direccionamiento en la localización genómica correcta (por ejemplo, mediante el uso del ensayo de modificación de alelo descrito a continuación).

5 Un ensayo cuantitativo de alto rendimiento, a saber, ensayo de modificación de alelo (MOA), puede usarse para el genotipado. Un ensayo tal permite un tamizaje a gran escala de un(os) alelo(s) modificado(s) en un cromosoma progenitor después de una modificación genética. El ensayo de MOA puede llevarse a cabo a través de diversas técnicas analíticas, que incluyen, pero sin limitarse a, una PCR cuantitativa, por ejemplo, una PCR en tiempo real (qPCR). Por ejemplo, la PCR en tiempo real comprende un primer conjunto de cebadores que reconocen el locus objetivo y un segundo conjunto de cebadores que reconocen un locus de referencia que no se direcciona. Además, el conjunto de cebadores comprende una sonda fluorescente que reconoce la secuencia amplificada. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de Invader Probes®. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de MMP assays®. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de Baliza Molecular TaqMan®. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de la tecnología de sondas Eclipse™. (Ver, por ejemplo, el documento US2005/0144655).

15 La célula pluripotente seleccionada (es decir, una célula pluripotente no humana, una célula ES no humana) que comprende la modificación genética dirigida puede introducirse después en un embrión huésped, por ejemplo, un embrión en etapa premórula o etapa de blastocisto e implantarse en el útero de una madre sustituta para generar un animal no humano fundador (animal F0). Posteriormente, el animal fundador, por ejemplo, puede reproducirse con un animal de tipo silvestre para generar la progenie F1 heterocigota para la modificación genética. El apareamiento del animal F1 heterocigoto puede producir la progenie homocigota para la modificación genética. El apareamiento del animal F1 heterocigoto puede producir la progenie homocigota para la modificación genética. En algunas modalidades, varias de las modificaciones genéticas de los loci objetivos descritos en la presente descripción pueden llevarse a cabo mediante el uso de un vector de direccionamiento grande (LTVEC) como se describe en detalle en cualquier otra parte de la presente descripción. Por ejemplo, un LTVEC puede derivarse del ADN del cromosoma artificial bacteriano (BAC) mediante el uso de la tecnología de modificación genética de VELOCIGENE® (ver, por ejemplo, la patente de los Estados Unidos núm. 6,586,251 y Valenzuela, D. M. y otros. (2003), High-throughput engineering of the mice genome coupled with high-resolution expression analysis, Nature Biotechnology 21(6): 652-659).

30 El uso de la recombinación homóloga bacteriana (BHR) para generar un vector de direccionamiento grande (LTVEC) evita las limitaciones de los plásmidos de acomodar un gran fragmento de ADN genómico y la consecuente baja eficiencia de introducir una modificación dirigida en un locus endógeno en células pluripotentes o no pluripotentes. Una o más modificaciones genéticas dirigidas pueden realizarse en la generación de un LTVEC. Un LTVEC ilustrativo producido en la célula procariota puede comprender un inserto de ácido nucleico que porta una secuencia genómica con una o más modificaciones genéticas o un ácido nucleico exógeno (por ejemplo, uno homólogo u ortólogo de un ácido nucleico de rata), que está flanqueada por brazos homólogos, complementarios a regiones genómicas específicas.

35 Se proporcionan, además, células procariotas huésped que comprenden los diversos vectores de direccionamiento descritos en la presente descripción. Tales células procariotas incluyen, pero sin limitarse a, bacterias tales como *E. coli*. En una modalidad, una célula procariota huésped comprende un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde el inserto de ácido nucleico varía de aproximadamente 5 kb a aproximadamente 200 kb.

45 La célula procariota huésped puede comprender, además, un ácido nucleico que codifica un polipéptido de recombinasa o el ácido nucleico que codifica el polipéptido de recombinasa se une operativamente a un promotor inducible.

50 En la presente se proporcionan, además, diversos métodos y composiciones, que emplean el LTVEC como se describe en la presente descripción en combinación con una célula procariota para producir las modificaciones genéticas dirigidas. Tales composiciones y métodos se analizan en cualquier otra parte de la presente descripción.

55 Se proporcionan métodos para modificar un locus objetivo de un ácido nucleico a través de la recombinación homóloga bacteriana (BHR) que comprenden introducir en una célula procariota un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde la célula procariota comprende ácidos nucleicos que corresponden a los brazos de homología 5' y 3' y la célula procariota es capaz de expresar una recombinasa que media la BHR en el locus objetivo. Tales vectores de direccionamiento pueden incluir cualquiera de los vectores de direccionamiento grandes descritos en la presente descripción. Tales métodos pueden emplear un LTVEC como se analiza en detalle en la presente descripción y además emplear el sistema CRISPR/Cas como se analiza en cualquier otra parte de la presente descripción.

60 En una modalidad, el sistema CRISPR/Cas puede ser controlado por un promotor activo en una célula procariota, tal como, por ejemplo, *E. coli*.

ii. Métodos para modificar un locus objetivo de interés en una célula pluripotente o célula no pluripotente.

65 La descripción proporciona, además, un método para modificar un locus objetivo de interés en una célula pluripotente o célula no pluripotente a través de la modificación genética dirigida, que comprende (a) introducir en la célula pluripotente

o célula no pluripotente un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde la suma total del brazo de homología 5' y el brazo de homología 3' es de al menos 10 kb; y (b) identificar una célula pluripotente o no pluripotente modificada genéticamente que comprende la modificación genética dirigida en el locus objetivo de interés. En una modalidad, la suma total del brazo de homología 5' y el brazo de homología 3' es de al menos aproximadamente 16 kb a aproximadamente 30 kb. En modalidades específicas, la modificación genética dirigida es capaz de transmitirse a través de la línea germinal. Tales vectores de direccionamiento pueden incluir cualquiera de los vectores de direccionamiento grandes descritos en la presente descripción.

La descripción describe que pueden usarse diversas células, además, en los métodos para modificar un locus objetivo de interés proporcionadas en la presente descripción, por ejemplo, la célula es una célula eucariota, célula eucariota no de rata, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula pluripotente inducida (iPS) humana, una célula de mamífero, una célula humana, un fibroblasto, una célula de roedor, una célula de roedor no rata, una célula de ratón, una célula de hámster o una célula CHO.

En un aspecto, se proporciona un método para modificar un locus genómico de interés en una célula pluripotente a través de la modificación genética dirigida, que comprende: (a) proporcionar una célula pluripotente que es capaz de mantener su pluripotencia después de al menos una modificación genética dirigida de su genoma y es capaz de transmitir la modificación dirigida a una línea germinal de una generación F1; (b) introducir un vector de direccionamiento grande (LTVEC) en la célula pluripotente, en donde el LTVEC comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde el brazo de homología 5' y el brazo de homología 3' comprenden un fragmento de ADN genómico; y (c) identificar una célula pluripotente genéticamente modificada que comprende la modificación genética dirigida.

Pueden usarse varios métodos para identificar las células que tienen el inserto de ácido nucleico integrado en el locus objetivo de interés. La inserción del inserto de ácido nucleico en el locus objetivo de interés da como resultado una "modificación de alelo". El término "modificación de alelo" y los métodos para la detección del alelo modificado se analizan en mayor detalle en cualquier otra parte de la presente descripción.

En un aspecto, se proporciona un método para modificar un locus genómico de interés en una célula no pluripotente o una célula pluripotente a través de la transformación de genes mediada por endonucleasas, el método que comprende: (a) proporcionar una célula no pluripotente aislada o una célula pluripotente aislada que es capaz de transmitir el genoma genéticamente modificado a una línea germinal de una generación F1; (b) introducir en la célula no pluripotente o la célula pluripotente un agente endonucleasa; en donde el agente endonucleasa genera una mella o una ruptura de la doble cadena en una secuencia de ADN objetivo localizada en el locus genómico de interés, y en donde la mella o la ruptura de la doble cadena en la secuencia de ADN objetivo en la célula no pluripotente o la célula pluripotente induce: (i) reparación del ADN mediada por la unión de extremos no homólogos (NHEJ) de la mella o la ruptura de la doble cadena, en donde la reparación del ADN mediada por NHEJ genera un alelo mutante que comprende una inserción o una deleción de una secuencia de ácidos nucleicos en la secuencia de ADN objetivo; o (ii) la reparación del ADN mediada por recombinación homóloga que da como resultado la restauración de una secuencia de ácidos nucleicos de tipo silvestre; y (c) identificar el locus genómico de interés modificado.

En un aspecto, se proporciona un método para modificar un locus genómico de interés en una célula madre embrionaria (ES) aislada a través de un agente nucleasa, que comprende: (a) proporcionar una célula ES aislada que es capaz de transmitir la modificación genética dirigida a una línea germinal de una generación F1; (b) introducir en la célula ES: (i) un vector de direccionamiento grande (LTVEC) que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde el inserto es una secuencia de ácidos nucleicos que es de al menos 5 kb; y (ii) un agente endonucleasa, en donde el agente endonucleasa produce una mella o una ruptura de la doble cadena en una secuencia de ADN objetivo ubicada en el locus genómico de interés, y en donde la secuencia objetivo no está presente en el inserto de ácido nucleico; y (c) identificar la modificación genética dirigida en la célula madre embrionaria (ES).

La descripción describe que, se proporciona un método para modificar un locus genómico de interés en una célula no pluripotente o una célula pluripotente a través de la modificación genética del genoma guiada por ARN, el método que comprende: (a) proporcionar una célula no pluripotente o una célula pluripotente que sea capaz de transmitir el genoma genéticamente modificado a una línea germinal de una generación F1; (b) introducir en la célula no pluripotente o la célula pluripotente: (i) un primer constructo de expresión que comprende un primer promotor unido operativamente a una primera secuencia de ácidos nucleicos que codifica una proteína asociada a Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (Cas), (ii) un segundo constructo de expresión que comprende un segundo promotor unido operativamente a una secuencia objetivo genómica unida a un ARN guía (ARNg), en donde la secuencia objetivo genómica está flanqueada por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM). Opcionalmente, la secuencia objetivo genómica está flanqueada en el extremo 3' por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM). En una modalidad, la proteína Cas y el ARN de CRISPR y/o ARNtracr no se producen juntos de manera natural (por ejemplo, la proteína Cas y el ARN de CRISPR no se producen juntos de manera natural). En una modalidad, la secuencia objetivo genómica comprende la secuencia nucleotídica de

GNNNNNNNNNNNNNNNNNNNNNGG (GN₁₋₂₀GG; sec. con núm. de ident.: 1). En una modalidad, la secuencia objetivo genómica comprende la sec. con núm. de ident.: 1, en donde N es de entre 14 y 20 nucleótidos de longitud. En una modalidad, el ARNg comprende una tercera secuencia de ácidos nucleicos que codifica un ARN de Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (ARNcr) y una cuarta secuencia de ácidos nucleicos que codifica un ARN de CRISPR de transactivación (ARNtracr). En una modalidad, tras la expresión, la proteína Cas forma un complejo CRISPR-Cas que comprende el ARNcr y el ARNtracr, y el complejo CRISPR-Cas produce una mella o una ruptura de la doble cadena en una secuencia de ADN objetivo ubicada en el locus genómico de interés, y en donde la mella o la ruptura de la doble cadena en la secuencia de ADN objetivo en la célula no pluripotente o la célula pluripotente induce: (i) reparación del ADN mediada por la unión de extremos no homólogos (NHEJ) de la mella o la ruptura de la doble cadena generada por el complejo CRISPR-Cas, en donde la NHEJ genera un alelo mutante que comprende una inserción o una delección de una secuencia de ácidos nucleicos en la secuencia de ADN objetivo; o (ii) la reparación del ADN mediada por recombinación homóloga que da como resultado la restauración de una secuencia de ácidos nucleicos de tipo silvestre; y (c) identificar el locus genómico de interés modificado.

En un aspecto, se proporciona un método para modificar un locus genómico de interés en una célula no pluripotente o una célula pluripotente a través de la modificación genética del genoma guiada por ARN, el método que comprende introducir en la célula no pluripotente o la célula pluripotente que es capaz de transmitir el genoma modificado a través de la línea germinal: (i) una proteína asociada a Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (Cas) o un ácido nucleico que codifica la proteína Cas; y (ii) un ARN guía o un ADN que codifica el ARNg, en donde el ARNg comprende una secuencia de nucleótidos que se hibrida a la secuencia objetivo genómica y a un ARN de CRISPR de transactivación (ARNtracr); en donde la secuencia objetivo genómica está flanqueada por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).

La descripción describe que la proteína Cas puede introducirse en la célula no pluripotente o la célula pluripotente como una proteína aislada. En algunas modalidades, la proteína Cas puede comprender, además, un dominio de penetración a célula que facilita la entrada celular de la proteína. En otras modalidades, la proteína Cas puede introducirse en la célula como una molécula de ARN mensajero (ARNm) que codifica la proteína Cas. En otras modalidades, la proteína Cas puede introducirse en la célula como una molécula de ADN que codifica la proteína Cas. Por ejemplo, la molécula de ADN que codifica la proteína Cas puede proporcionarse en un constructo y unirse operativamente a un promotor capaz de expresar en la célula no pluripotente o la célula pluripotente. En determinadas modalidades, el ácido nucleico que codifica la proteína Cas se optimiza en codones para la expresión en la célula no pluripotente o la célula pluripotente.

La descripción describe que el ARNg puede introducirse en la célula no pluripotente o la célula pluripotente como una molécula de ARN. Por ejemplo, la molécula de ARNg puede transcribirse *in vitro*. La descripción describe que, el ARNg puede introducirse en la célula no pluripotente o la célula pluripotente como una molécula de ADN que codifica el ARNg. Por ejemplo, la molécula de ADN que codifica el ARNg puede estar en un constructo y unirse operativamente a un promotor capaz de expresar el ARNg en la célula no pluripotente o la célula pluripotente. En otras modalidades, el ARNg puede sintetizarse químicamente.

La descripción describe que el ARNg puede introducirse en la célula no pluripotente o la célula pluripotente como una molécula de ARNcr-ARNtracr fusionados (es decir, un transcrito único). La descripción describe que el ARNg puede introducirse en la célula no pluripotente o la célula pluripotente como moléculas de ARNcr y ARNtracr separadas (es decir, transcritos separados). La descripción describe que el ARNg puede introducirse en la célula no pluripotente o la célula pluripotente como moléculas de ADN separadas que codifican el ARNcr y el ARNtracr, respectivamente. Por ejemplo, las moléculas de ADN separadas que codifican el ARNcr y el ARNtracr pueden estar en constructos separados y unirse operativamente a promotores capaces de expresar en la célula no pluripotente o la célula pluripotente. En cualquiera de las modalidades anteriores, cualquier combinación de los constructos puede estar en moléculas de ácidos nucleicos separadas o juntos en una única molécula de ácido nucleico.

La descripción describe que la proteína Cas y el ARNg puede introducirse en la célula no pluripotente o la célula pluripotente simultáneamente o secuencialmente. De manera similar, el ARNcr y el ARNtracr del ARNg puede introducirse en la célula no pluripotente o la célula pluripotente simultáneamente o secuencialmente. La relación de la proteína Cas (o ácido nucleico codificante) con respecto al ARNg (o ADN codificante) y/o la relación del ARNcr con respecto al ARNtracr puede ser aproximadamente estequiométrica de manera que puedan formar un complejo de ARN-proteína.

La descripción describe que la proteína Cas puede introducirse en la célula no pluripotente o la célula pluripotente en la forma de un complejo con el ARNg.

La descripción describe que la célula pluripotente es una célula madre pluripotente inducida (iPS). En una modalidad, la célula pluripotente es una célula progenitora con desarrollo limitado.

La presencia de una mella o una ruptura de la doble cadena en el sitio de reconocimiento dentro del marcador de selección, en varias modalidades, aumenta la eficiencia y/o frecuencia de recombinación entre un vector de direccionamiento (tal como a LTVEC) y el locus de interés a transformar. En una modalidad, la recombinación es la recombinación homóloga. En otra modalidad, la recombinación es una inserción por unión de extremos no homólogos. En diversas modalidades, en la presencia de la mella o ruptura de la doble cadena, la eficiencia de transformación de un vector de direccionamiento

(tal como un LTVEC) en el locus genómico objetivo es de al menos aproximadamente 2 veces mayor, al menos aproximadamente 3 veces mayor, al menos aproximadamente 4 veces mayor que en la ausencia de la mella o ruptura de la doble cadena (mediante el uso, por ejemplo, del mismo vector de direccionamiento y los mismos brazos de homología y los sitios objetivo correspondientes en el locus genómico de interés pero en la ausencia de un agente nucleasa añadido que genera la mella o la ruptura de la doble cadena).

En una modalidad, la modificación genética dirigida en el locus objetivo es bialélica. Por "bialélica" se entiende que ambos alelos de un gen comprenden la modificación genética dirigida. La modificación genética dirigida puede ser igual o diferente en cada alelo. Por ejemplo, una modificación bialélica puede ser el resultado de la misma modificación que se realiza a los alelos correspondientes en los cromosomas homólogos correspondientes, o de diferentes modificaciones que se hacen a alelos correspondientes en los cromosomas homólogos correspondientes. Por lo tanto, una modificación bialélica puede dar como resultado, por ejemplo, la homocigosidad para una modificación específica en un locus genómico de interés (es decir, la modificación específica en ambos alelos), heterocigosidad compuesta en un locus genómico de interés (por ejemplo, la modificación específica en un alelo y la inactivación o alteración del otro alelo), o hemicigosidad en un locus genómico de interés (por ejemplo, la modificación específica en un alelo y pérdida del otro alelo). En determinadas modalidades, el uso combinado de un vector de direccionamiento (que incluye, por ejemplo, un LTVEC) con un agente nucleasa da como resultado la modificación bialélica genética dirigida del locus genómico de interés en una célula en comparación con el uso del vector de direccionamiento solo. Cuando el vector de direccionamiento se usa junto con un agente nucleasa, la eficiencia de transformación bialélica aumenta al menos en dos veces, al menos tres veces, al menos 4 veces o más en comparación con el vector de direccionamiento cuando se usa solo. En modalidades adicionales, la eficiencia de transformación bialélica es al menos 0,2 %, 0,3 %, 0,4 %, 0,5 %, 0,6 %, 0,7 %, 0,8 %, 0,9 %, 1 %, 2 %, 3 %, 4 % o 5 % o mayor.

La modificación bialélica genética dirigida en el locus objetivo puede dar como resultado una célula genéticamente modificada homocigota. Por "homocigota" se entiende que ambos alelos del locus objetivo (es decir, los alelos en ambos cromosomas homólogos) se han modificado de la misma manera. En determinadas modalidades, el uso combinado de un vector de direccionamiento (que incluye, por ejemplo, un LTVEC) con un agente nucleasa da como resultado la modificación genética dirigida homocigota bialélica del locus genómico de interés en una célula. En una modalidad, la modificación genética bialélica comprende la delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en dos cromosomas homólogos (es decir, un par de los cromosomas homólogos primero y segundo) y la inserción del inserto de ácido nucleico en el locus genómico de interés en dos cromosomas homólogos (es decir, el par de los cromosomas homólogos primero y segundo). En algunas modalidades, el inserto de ácido nucleico sustituye la secuencia de ácidos nucleicos endógena en el locus genómico de interés en ambos cromosomas homólogos. En una modalidad, el inserto de ácido nucleico es homólogo u ortólogo a la secuencia de ácidos nucleicos endógena delecionada.

En una modalidad, la modificación genética dirigida en el locus objetivo da como resultado una célula genéticamente modificada hemicigota. Por "hemicigota" se entiende que solo un alelo (es decir, el alelo en uno de dos cromosomas homólogos) del locus objetivo está presente o solo un alelo es capaz de expresarse y ser funcional. En otras modalidades, la modificación genética dirigida da como resultado más generalmente la heterocigosidad compuesta. La heterocigosidad compuesta incluye situaciones en las que ambos alelos del locus objetivo (es decir, los alelos en ambos cromosomas homólogos) se han modificado, pero se han modificado de diferentes maneras (por ejemplo, una inserción en un alelo e inactivación o alteración del otro alelo). En determinadas modalidades, el uso combinado de un vector de direccionamiento (que incluye, por ejemplo, un LTVEC) con un agente nucleasa da como resultado una modificación genética dirigida hemicigota del locus genómico de interés en una célula. En determinadas modalidades, el uso combinado de un vector de direccionamiento (que incluye, por ejemplo, un LTVEC) con un agente nucleasa da como resultado modificaciones genéticas dirigidas que generan heterocigosidad compuesta en un locus genómico de interés en una célula. En una modalidad, la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógenos y la inserción del inserto de ácido nucleico. En otras modalidades, la modificación genética dirigida comprende: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en dos cromosomas homólogos; y (2) inserción del inserto de ácido nucleico en el locus genómico de interés en un primer cromosoma y alteración del locus genómico de interés en un segundo cromosoma. El primer cromosoma puede ser el primero de los dos cromosomas homólogos, y el segundo cromosoma puede ser el segundo de los dos cromosomas homólogos. En otras modalidades, la modificación dirigida comprende: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés e inserción del inserto de ácido nucleico en el locus genómico de interés en el primer cromosoma homólogo; y (2) alteración del locus genómico de interés en el segundo cromosoma homólogo. La alteración de la secuencia de ácidos nucleicos endógeno puede ser el resultado de, por ejemplo, cuando una ruptura de la doble cadena en el locus genómico de interés generada por el agente nucleasa se prepara mediante la reparación del ADN mediada por la unión de extremos no homólogos (NHEJ), lo que genera un alelo mutante que comprende una inserción o una delección de una secuencia de ácidos nucleicos en el locus genómico de interés y de esta manera provoca la alteración del locus genómico de interés. Los ejemplos de alteración incluyen la alteración de un elemento regulador (por ejemplo, promotor o potenciador) en el locus genómico de interés, una mutación con pérdida de sentido, una mutación de truncación, una mutación nula, o una inserción o delección de un pequeño número de nucleótidos (por ejemplo, que provoca una mutación con desplazamiento del marco). Otro ejemplo de alteración es una mutación sin sentido. La alteración puede dar como resultado la inactivación (es decir, pérdida de función) o la pérdida del alelo.

Las modificaciones genéticas dirigidas homocigotas y hemocigotas son ventajosas porque cuando se usan las células genéticamente modificadas que contienen estas mutaciones para generar animales genéticamente modificados como se analiza más adelante, el proceso para la generación de animales genéticamente modificados que no son heterocigotos (es decir, homocigotos o hemocigotos) para la modificación genética dirigida que se pretende es más eficiente y menos consumidor de tiempo porque se necesitan menos etapas de reproducción. Las modificaciones genéticas dirigidas que dan como resultado la heterocigosidad compuesta o hemocigosidad (por ejemplo, una inserción en un alelo e inactivación, alteración, o pérdida del otro alelo) pueden ser ventajosas por la misma razón.

La descripción describe que pueden usarse diversos tipos celulares, además, en cualquiera de los diversos métodos descritos anteriormente en la presente descripción para modificar un locus genómico a través de un agente nucleasa. Esta descripción describe que la célula es una célula eucariota, célula eucariota no de rata, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula pluripotente inducida (iPS) humana, una célula de mamífero, una célula humana, un fibroblasto, una célula de roedor, una célula de roedor no rata, una célula de ratón, una célula de hámster o una célula CHO.

Se proporcionan composiciones que comprenden un animal no humano genéticamente modificado, que tiene una modificación genética dirigida en el locus de la cadena gamma de receptor de interleucina-2 o en el locus de *ApoE*. Los diversos métodos y composiciones proporcionados en la presente descripción permiten que estos loci modificados se transmitan a través de la línea germinal.

En ejemplos específicos, un animal no humano genéticamente modificado, o una célula pluripotente o no pluripotente modificada genéticamente comprende un locus genómico que tiene una modificación genética dirigida en el locus de la cadena gamma de receptor de interleucina-2 o que tiene una modificación genética dirigida en el locus de *ApoE*, en donde el locus de la cadena gamma de receptor de interleucina-2 genómico o el locus de *ApoE* comprende: (i) una delección de al menos una porción del locus de la cadena gamma de receptor de interleucina-2 o al menos una porción del locus de *ApoE*; (ii) una inserción de una secuencia de ácidos nucleicos heteróloga en el locus de *ApoE* o en el locus de la cadena gamma de receptor de interleucina-2; o (iii) una combinación de estos, en donde el locus genómico genéticamente modificado es capaz de transmitirse a través de la línea germinal.

Se proporcionan, además, los métodos que permiten que se generen tales animales no humanos genéticamente modificados y tales células pluripotentes genéticamente modificadas. Tales métodos incluyen un método para modificar un locus genómico de *ApoE* o un locus de la cadena gamma de receptor de interleucina-2 en una célula pluripotente a través de la modificación genética dirigida. El método comprende (a) introducir en la célula pluripotente un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5', al locus de *ApoE* y un brazo de homología 3', al locus de *ApoE*, (b) identificar una célula pluripotente genéticamente modificada que comprende la modificación genética dirigida en el locus genómico de interés de *ApoE*, en donde la modificación genética dirigida es capaz de transmitirse a través de la línea germinal.

Los métodos adicionales incluyen (a) introducir en la célula pluripotente un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' al locus de la cadena gamma de receptor de interleucina-2 y un brazo de homología 3' al locus de la cadena gamma de receptor de interleucina-2, (b) identificar una célula pluripotente genéticamente modificada que comprende la modificación genética dirigida en el locus de la cadena gamma de receptor de interleucina-2, en donde la modificación genética dirigida es capaz de transmitirse a través de la línea germinal.

iii. Métodos para integrar múltiples polinucleótidos de interés en el locus objetivo

Los diversos métodos y composiciones proporcionados en la presente descripción permiten la integración dirigida de múltiples polinucleótidos de interés con un locus objetivo dado. Los diversos métodos expuestos anteriormente pueden repetirse secuencialmente para permitir la integración dirigida de cualquier número de insertos de ácido nucleicos en un locus objetivo dado. Por lo tanto, los diversos métodos proporcionan la inserción de al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más insertos de ácido nucleicos en el locus objetivo. En modalidades particulares, tales métodos de traslape secuencial permiten la reconstrucción de grandes regiones genómicas de una célula eucariota, por ejemplo, célula eucariota no de rata, una célula de mamífero (es decir, un ser humano, un no ser humano, un roedor, un roedor no rata, un ratón, un mono, una rata, un hámster, un mamífero domesticado o un animal agrícola) en un locus a transformar. En tales casos, la transferencia y reconstrucción de regiones genómicas que incluyen ambas regiones codificantes y no codificantes permiten conservar la complejidad de una región dada al retener, al menos en parte, las regiones codificantes, las regiones no codificantes y las variaciones del número de copias encontradas dentro de la región genómica nativa. Por lo tanto, los diversos métodos de la descripción proporcionan, por ejemplo, los métodos para generar regiones genómicas "heterólogas" o "exógenas" dentro de cualquier célula eucariota, cualquier célula eucariota no de rata, cualquier célula de mamífero o animal de interés, particularmente dentro de una célula huésped procarionta o dentro de una célula no pluripotente, una célula pluripotente o una célula ES. En un ejemplo no limitante, se genera una región genómica "humanizada" dentro de un animal no humano (es decir, dentro de una rata). Los métodos para generar regiones genómicas dentro de cualquier célula se proporcionan en la presente descripción. La descripción describe que la célula es una célula eucariota, una célula eucariota no de rata, una célula pluripotente, una célula no pluripotente, una célula

pluripotente no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula pluripotente inducida (iPS) humana, una célula de mamífero, una célula humana, un fibroblasto, una célula de roedor, una célula de roedor no rata, una célula de ratón, una célula de hámster o una célula CHO.

5

3. Un locus genómico humanizado

En la presente descripción se proporcionan diversos métodos y composiciones que comprenden un locus genómico humanizado. Como se usa en la presente descripción, por locus genómico "humanizado" se entiende una región de un genoma no humano que comprende al menos una secuencia de ácidos nucleicos humana. El locus genómico humanizado puede comprender una región de ADN de cualquier organismo que tiene insertada una secuencia de ADN humano. La descripción describe que el organismo es un eucariota, un eucariota no rata, un mamífero no humano, un mamífero, un ser humano, un roedor, un roedor no rata, una rata, un ratón o un hámster. Por ejemplo, un "locus de rata humanizado" comprende una región de ADN de rata que tiene insertada una secuencia de ADN humano.

15

La secuencia de ADN humano puede ser una secuencia de ADN humano de origen natural o puede modificarse a partir de su forma nativa. En modalidades específicas, el ADN humano comparte al menos 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, o 99 % de identidad de secuencia con respecto a una secuencia humana nativa. Si una secuencia humana no es una secuencia humana nativa al menos tiene mayor identidad de secuencia con respecto a una secuencia humana nativa que con respecto a una secuencia no humana ortóloga. Además, la secuencia de ADN humano puede comprender un ADNc, una región de ADN genómico humano, una región reguladora no codificante, o cualquier porción de una región codificante, genómica, o reguladora del ADN humano. La secuencia de ADN humano insertada en el locus no humano puede comprender cualquiera de los insertos de polinucleótidos como se describe en cualquier otra parte de la presente descripción. En modalidades específicas, la secuencia de ADN humano es ortóloga al locus no humano objetivo, mientras que en otros casos, la secuencia de ADN humano es homóloga al locus objetivo no humano.

20

En una modalidad, la modificación genética dirigida es una inserción o una sustitución de una secuencia de ácidos nucleicos endógena, con una secuencia de ácidos nucleicos humana homóloga u ortóloga. En una modalidad, la modificación genética dirigida comprende una inserción o la sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos humana homóloga u ortóloga en un locus endógeno que comprende la secuencia de ácidos nucleicos no humanos correspondientes.

30

Los métodos para generar un locus humanizado comprenden introducir en el locus objetivo que comprende un ácido nucleico una secuencia de ácidos nucleicos humana. En una modalidad, se proporciona un método para generar un animal no humano humanizado. Un método tal comprende (a) modificar un genoma de una célula pluripotente no humana o célula no pluripotente con un vector de direccionamiento que comprende un inserto de ácido nucleico que comprende una secuencia de ácidos nucleicos humana para formar una célula donante; (b) introducir la célula donante en un embrión huésped; y (c) gestar el embrión huésped en una madre sustituta; en donde la madre sustituta produce una progenie que comprende la secuencia de ácidos nucleicos humana. En modalidades específicas, el locus humanizado es capaz de transmitirse a través de la línea germinal. En una modalidad adicional, el vector de direccionamiento comprende un vector de direccionamiento grande (LTVEC) y el inserto de ácido nucleico que comprende una secuencia de ácidos nucleicos humana es de al menos 5 kb.

35

40

En otros métodos, el locus genómico humanizado se genera mediante la modificación de un locus objetivo de un ácido nucleico a través de la recombinación homóloga bacteriana (BHR). El método comprende introducir en una célula procariota un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde el inserto de ácido nucleico comprende una secuencia de ácidos nucleicos humana, y en donde la célula procariota comprende un ácido nucleico y es capaz de expresar una recombinasa que media la BHR en el locus objetivo.

45

50

El locus genómico humanizado puede comprender (a) una inserción de una secuencia de ácidos nucleicos humana homóloga u ortóloga; (b) una sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos humana homóloga u ortóloga; o (c) una combinación de estos. En modalidades específicas, el locus genómico humanizado es capaz de transmitirse a través de la línea germinal. Aún en otras modalidades, la secuencia ortóloga humana sustituye la secuencia correspondiente encontrada en el locus no humano.

55

Cualquier secuencia de ácidos nucleicos humana puede usarse en los métodos y composiciones proporcionados en la presente descripción. Los ejemplos no limitantes de secuencia de ácidos nucleicos humana que pueden usarse en los métodos y composiciones se analizan en detalle en cualquier otra parte de la presente descripción.

60

La secuencia de ácidos nucleicos humana para la inserción en un locus de interés puede ser de cualquier tamaño. La descripción describe que la secuencia de ácidos nucleicos humana puede ser de aproximadamente 500 nucleótidos a aproximadamente 200 kb, de aproximadamente 500 nucleótidos a aproximadamente 5 kb, de aproximadamente 5 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 60 kb a

65

- aproximadamente 70 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, o de aproximadamente 190 kb a aproximadamente 200 kb. La descripción describe que la secuencia de ácidos nucleicos humana es de al menos 5 kb.
- 5
- En una modalidad, se proporciona un locus genómico en donde la secuencia de ácidos nucleicos humana homóloga u ortóloga comprende (a) uno o más segmentos del gen de V_H de cadena pesada de inmunoglobulina humana no reordenados, uno o más segmentos del gen de D de cadena pesada de inmunoglobulina humana no reordenados, y uno o más segmentos del gen de J_H de cadena pesada de inmunoglobulina humana no reordenados, que se unen operativamente a una secuencia de ácidos nucleicos de la región constante de cadena pesada de mamífero; (b) una secuencia de ácidos nucleicos de la región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácidos nucleicos de la región constante de cadena pesada de inmunoglobulina de mamífero; (c) uno o más segmentos del gen de V_K o V_k de inmunoglobulina humana no reordenados y uno o más segmentos del gen de J_K o J_k de inmunoglobulina humana no reordenados, que se unen operativamente a una secuencia de ácidos nucleicos de la región constante de cadena ligera de cadena ligera λ o κ de inmunoglobulina de mamífero; o (d) una secuencia de ácidos nucleicos de la región variable de cadena ligera λ o κ de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácidos nucleicos de la región constante de cadena ligera de cadena ligera de cadena λ o κ de inmunoglobulina de mamífero.
- 10
- En otra modalidad, se proporciona un locus genómico en donde (a) la secuencia de ácidos nucleicos de región constante de cadena pesada de inmunoglobulina de mamífero es una secuencia de ácidos nucleicos de región constante, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos; o (b) secuencia de ácidos nucleicos de la región constante de cadena ligera de cadena ligera λ o κ de inmunoglobulina de mamífero es una secuencia de ácidos nucleicos de región constante de rata, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos.
- 15
- En una modalidad específica, se proporciona un locus genómico en donde la secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina se selecciona de o comprende una CH1, una región bisagra, una CH2, una CH3, y/o una combinación de estos.
- 20
- En una modalidad, el locus genómico comprende uno o más segmentos del gen de V_H humano funcional que comprenden V_{H1-2} , V_{H1-3} , V_{H1-8} , V_{H1-18} , V_{H1-24} , V_{H1-45} , V_{H1-46} , V_{H1-58} , V_{H1-69} , V_{H2-5} , V_{H2-26} , V_{H2-70} , V_{H3-7} , V_{H3-9} , V_{H3-11} , V_{H3-13} , V_{H3-15} , V_{H3-16} , V_{H3-20} , V_{H3-21} , V_{H3-23} , V_{H3-30} , $V_{H3-30-3}$, V_H 3-30-5, V_{H3-33} , V_{H3-35} , V_{H3-38} , V_{H3-43} , V_{H3-48} , V_{H3-49} , V_{H3-53} , V_{H3-64} , V_{H3-66} , V_{H3-72} , V_{H3-73} , V_{H3-74} , V_{H4-4} , V_{H4-28} , $V_{H4-30-1}$, $V_{H4-30-2}$, $V_{H4-30-4}$, V_{H4-31} , V_{H4-34} , V_{H4-39} , V_{H4-59} , V_{H4-61} , V_{H5-51} , V_{H6-1} , V_{H7-4-1} , V_{H7-81} , o una combinación de estos.
- 25
- En una modalidad, el locus genómico comprende uno o más segmentos del gen de D humano funcional que comprenden $D1-1$, $D1-7$, $D1-14$, $D1-20$, $D1-26$, $D2-2$, $D2-8$, $D2-15$, $D2-21$, $D3-3$, $D3-9$, $D3-10$, $D3-16$, $D3-22$, $D4-4$, $D4-11$, $D4-17$, $D4-23$, $D5-12$, $D5-5$, $D5-18$, $D5-24$, $D6-6$, $D6-13$, $D6-19$, $D6-25$, $D7-27$, o una combinación de estos.
- 30
- En una modalidad, el locus genómico comprende uno o más segmentos del gen de J_H funcional que comprenden J_{H1} , J_{H2} , J_{H3} , J_{H4} , J_{H5} , J_{H6} , y/o una combinación de estos. En una modalidad, el inserto de ácido nucleico comprende uno o más segmentos del gen de V_K humano que comprenden V_{K4-1} , V_{K5-2} , V_K 7-3, V_K 2-4, V_{K1-5} , V_{K1-6} , V_{K3-7} , V_{K1-8} , V_{K1-9} , V_{K2-10} , V_{K3-11} , V_{K1-12} , V_{K1-13} , V_{K2-14} , V_{K3-15} , V_{K1-16} , V_{K1-17} , V_{K2-18} , V_{K2-19} , V_{K3-20} , V_{K6-21} , V_{K1-22} , V_{K1-23} , V_{K2-24} , V_{K3-25} , V_{K2-26} , V_{K1-27} , V_{K2-28} , V_{K2-29} , V_{K2-30} , V_{K3-31} , V_{K1-32} , V_{K1-33} , V_{K3-34} , V_{K1-35} , V_{K2-36} , V_{K1-37} , V_{K2-38} , V_{K1-39} , V_{K2-40} , o una combinación de estos.
- 35
- En una modalidad, el locus genómico comprende uno o más segmentos del gen de V_λ humano que comprenden $V_{\lambda 3-1}$, $V_{\lambda 4-3}$, $V_{\lambda 2-8}$, $V_{\lambda 3-9}$, $V_{\lambda 3-10}$, $V_{\lambda 2-11}$, $V_{\lambda 3-12}$, $V_{\lambda 2-14}$, $V_{\lambda 3-16}$, $V_{\lambda 2-18}$, $V_{\lambda 3-19}$, $V_{\lambda 3-21}$, $V_{\lambda 3-22}$, $V_{\lambda 2-23}$, $V_{\lambda 3-25}$, $V_{\lambda 3-27}$, o una combinación de estos.
- 40
- En una modalidad, el locus genómico comprende uno o más segmentos del gen de J_λ humano que comprenden J_{k1} , J_{k2} , J_{k3} , J_{k4} , J_{k5} , o una combinación de estos.
- 45
- Aún en otra modalidad, el locus genómico, comprende un locus genómico humanizado que comprende una secuencia de ácidos nucleicos de receptor de interleucina 2 (IL2R) humano o se proporciona una variante o un fragmento de este. En modalidades específicas, la secuencia de ácidos nucleicos de IL2R comprende una cadena alfa de receptor de interleucina 2, una cadena beta de receptor de interleucina 2, o una secuencia de ácidos nucleicos de la cadena gamma de receptor de interleucina-2 o variantes o fragmentos de esta.
- 50
- En modalidades adicionales, un locus genómico, comprende un locus genómico humanizado que comprende una porción del locus de *ApoE* humano, el locus de la cadena gamma de receptor de interleucina-2 humano, el locus de *Rag2* humano, el locus de *Rag1* humano y/o el locus de *Rag2/Rag1* humano que sustituye la porción homóloga u ortóloga
- 55
- 60
- 65

correspondiente del locus de *ApoE* no humano, locus de la cadena gamma de receptor de interleucina-2, locus de *Rag2*, locus de *Rag1* y/o locus de *Rag2/Rag1*. En una modalidad, el ectodominio no humano de IL-2Rg se sustituye con el ectodominio de IL-2Rg humano, con el resto de la molécula que es del no humano.

5 En otra modalidad, se proporciona un animal no humano genéticamente modificado, que comprende un locus genómico humanizado. Tales animales no humanos genéticamente modificados comprenden (a) una inserción de una secuencia de ácidos nucleicos humana homóloga u ortóloga; (b) una sustitución de secuencia de ácidos nucleicos con una secuencia de ácidos nucleicos humana homóloga u ortóloga en un locus endógeno genómico; o (c) una combinación de estos, en donde el locus genómico humanizado es capaz de transmitirse a través de la línea germinal.

10 Se proporcionan, además, animales genéticamente modificados, que incluyen animales no humanos) que comprenden cualquiera de los diversos loci genómicos humanizados proporcionados en la presente descripción y descritos anteriormente.

15 4. Polinucleótidos de interés

Cualquier polinucleótido de interés puede estar contenido en los diversos insertos de ácidos nucleicos y de este modo integrarse en el locus objetivo. Los métodos descritos en la presente descripción, proporcionan al menos 1, 2, 3, 4, 5, 6 o más polinucleótidos de interés a integrar en el locus genómico a transformar.

20 Cuando el polinucleótido de interés dentro del inserto de ácido nucleico se integra en el locus genómico objetivo puede introducir una o más modificaciones genéticas en la célula. La modificación genética puede comprender una delección de una secuencia de ácidos nucleicos endógena y/o la adición de un polinucleótido exógeno o heterólogo u ortólogo en el locus genómico objetivo. En una modalidad, la modificación genética comprende una sustitución de una secuencia de ácidos nucleicos endógena con un polinucleótido de interés exógeno en el locus genómico objetivo. Por lo tanto, los métodos proporcionados en la presente descripción permiten la generación de una modificación genética que comprende una inactivación, una delección, una inserción, una sustitución ("inserción de secuencias génicas"), una mutación puntual, un intercambio de dominios, un intercambio de exones, un intercambio de intrones, una secuencia reguladora de intercambio, un intercambio de genes, o una combinación de estos. Tales modificaciones pueden producirse tras la integración del primero, segundo, tercero, cuarto, quinto, sexto, séptimo, o cualquiera de los insertos de ácidos nucleicos posteriores en el locus genómico objetivo.

35 El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo puede comprender una secuencia que es nativa a la célula en la que se introduce; el polinucleótido de interés puede ser heterólogo a la célula en la que se introduce; el polinucleótido de interés puede ser exógeno a la célula en la que se introduce; el polinucleótido de interés puede ser ortólogo a la célula en la que se introduce; o el polinucleótido de interés puede ser de una especie diferente a la de la célula en la que se introduce. Como se usa en la presente descripción "nativa" en referencia a una secuencia insertada en el locus objetivo es una secuencia que es nativa a la célula que tiene el locus objetivo o nativa a la célula a partir de la cual se derivó el locus objetivo (es decir, de una rata). Como se usa en la presente descripción, "heteróloga" en referencia a una secuencia incluye una secuencia que se origina de una especie extraña, o, si es de la misma especie, es sustancialmente diferente o está modificada de su forma nativa en la composición y/o locus genómico mediante intervención humana deliberada. Como se usa en la presente descripción, "exógena" en referencia a una secuencia es una secuencia que se origina de una especie foránea. El polinucleótido de interés puede ser de cualquier organismo de interés que incluye, pero sin limitarse a, un no ser humano, un roedor, un roedor no rata, un hámster, un ratón, una rata, un ser humano, un mono, un mamífero agrícola o un no mamífero agrícola. El polinucleótido de interés puede comprender, además, una región codificante, una región no codificante, una región reguladora, o un ADN genómico. Por lo tanto, el 1er, 2do, 3ro, 4to, 5to, 6to, 7mo, y/o cualquiera de los insertos de ácido nucleicos posteriores pueden comprender tales secuencias.

50 En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo es nativo a una secuencia de ácidos nucleicos de ratón, un ácido nucleico humano, un ácido nucleico no humano, un ácido nucleico de eucariota, un ácido nucleico de eucariota no rata, un ácido nucleico de mamífero no humano, un ácido nucleico de mamífero, un ácido nucleico de roedor, un ácido nucleico de roedor no rata, un ácido nucleico de rata, un ácido nucleico de hámster, un ácido nucleico de mono, un ácido nucleico de mamífero agrícola, o un ácido nucleico de mamífero no agrícola. Aún en otras modalidades, el polinucleótido de interés integrado en el locus objetivo es un fragmento de un ácido nucleico genómico. En una modalidad, el ácido nucleico genómico es un ácido nucleico genómico de ratón, un ácido nucleico genómico humano, un ácido nucleico no humano, un ácido nucleico de eucariota, un ácido nucleico de eucariota no rata, un ácido nucleico de mamífero no humano, un ácido nucleico de mamífero, un ácido nucleico de roedor, un ácido nucleico de roedor no rata, un ácido nucleico de rata, un ácido nucleico de hámster, un ácido nucleico de mono, un ácido nucleico de mamífero agrícola o un ácido nucleico de mamífero no agrícola o una combinación de estos.

65 En una modalidad, el polinucleótido de interés puede estar en el intervalo de aproximadamente 500 nucleótidos a aproximadamente 200 kb como se describió anteriormente. El polinucleótido de interés puede ser de aproximadamente 500 nucleótidos a aproximadamente 5 kb, de aproximadamente 5 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a

- aproximadamente 50 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, o de aproximadamente 190 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.
- 15 El polinucleótido de interés dentro del inserto de ácido nucleico y/o insertado en el locus genómico objetivo puede codificar un polipéptido, puede codificar un ARNmi, o puede comprender cualquiera de las regiones reguladoras o regiones no codificantes de interés que incluye, por ejemplo, una secuencia reguladora, una secuencia promotora, una secuencia potenciadora, una secuencia de unión a represor transcripcional, o una delección de una secuencia no codificante de proteínas, pero no comprende una delección de una secuencia codificante de proteínas. Además, el polinucleótido de interés dentro del inserto de ácido nucleico y/o insertado en el locus genómico objetivo puede codificar una proteína expresada en el sistema nervioso, el sistema esquelético, el sistema digestivo, el sistema circulatorio, el sistema muscular, el sistema respiratorio, el sistema cardiovascular, el sistema linfático, el sistema endocrino, el sistema urinario, el sistema reproductor, o una combinación de estos. En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o insertado en el locus genómico objetivo codifica una proteína expresada en una médula ósea o una célula derivada de médula ósea. En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo codifica una proteína expresada en una célula de bazo. Aún en otras modalidades, el polinucleótido de interés dentro del inserto de ácido nucleico y/o insertado en el locus objetivo codifica una proteína expresada en una célula B, codifica una proteína expresada en una célula B inmadura o codifica una proteína expresada en una célula B madura.
- 30 El polinucleótido de interés dentro del inserto de polinucleótido puede comprender una porción de un locus de *ApoE*, un locus de *Il2rg*, un locus de *Rag1*, un locus de *Rag2* y/o un locus de *Rag2/Rag1*. Tales porciones de estos loci dados se analizan en cualquier otra parte de la presente descripción, al igual que varias regiones homólogas y ortólogas de cualquier organismo de interés que pueda emplearse.
- 35 En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o insertado en el locus objetivo comprende una secuencia de ácidos nucleicos genómica que codifica una secuencia de aminoácidos de la región variable de cadena pesada de inmunoglobulina. La frase "cadena pesada," o "cadena pesada de inmunoglobulina" se describe en cualquier otra parte de la presente descripción.
- 40 En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo comprende una secuencia de ácidos nucleicos genómica que codifica una secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina humana.
- 45 En una modalidad, la secuencia de ácidos nucleicos genómica comprende uno o más segmentos del gen de V_H de cadena pesada de inmunoglobulina humana no reordenados, uno o más segmentos del gen de D de cadena pesada de inmunoglobulina humana no reordenados, y uno o más segmentos del gen de J_H de cadena pesada de inmunoglobulina humana no reordenados, que se unen operativamente a una secuencia de ácidos nucleicos de región constante de cadena pesada de mamífero. En una modalidad, la secuencia de ácidos nucleicos genómica comprende una secuencia de ácidos nucleicos de la región variable de cadena pesada de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácidos nucleicos de región constante de cadena pesada de mamífero. En una modalidad, la secuencia de ácidos nucleicos genómica comprende uno o más segmentos del gen de V_K o V_λ de inmunoglobulina humana no reordenados y uno o más segmentos del gen de J_K o J_λ de inmunoglobulina humana no reordenados, que se unen operativamente a una secuencia de ácidos nucleicos de región constante de cadena ligera de cadena ligera λ o κ de inmunoglobulina de mamífero. En una modalidad, la secuencia de ácidos nucleicos genómica comprende una secuencia de ácidos nucleicos de región variable de cadena ligera λ o κ de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácidos nucleicos de región constante de cadena ligera λ o κ de inmunoglobulina de mamífero. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena pesada comprende una secuencia de ácidos nucleicos de región constante de rata, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena ligera λ o κ de inmunoglobulina comprende una secuencia de ácidos nucleicos de región constante de rata, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos.
- 60 En una modalidad, la secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina se selecciona de o comprende un CH1, una región bisagra, un CH2, un CH3, y/o una combinación de estos. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena pesada comprende una CH1-región bisagra-CH2-CH3.
- 65

En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo comprende una secuencia de ácidos nucleicos genómica que codifica una secuencia de aminoácidos de región variable de cadena ligera de inmunoglobulina. La frase "cadena ligera" incluye una secuencia de cadena ligera de inmunoglobulina de cualquier organismo, y se describe en cualquier otra parte de la presente descripción.

5

En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus genómico objetivo comprende una secuencia de ácidos nucleicos genómica que codifica una secuencia de aminoácidos de región variable de cadena ligera de inmunoglobulina humana.

10

En una modalidad, la secuencia de ácidos nucleicos genómica comprende uno o más segmentos del gen de V_{κ} o V_{λ} de inmunoglobulina humana no reordenados y uno o más segmentos del gen de J_{κ} o J_{λ} de inmunoglobulina humana no reordenados, que se unen operativamente a una secuencia de ácidos nucleicos de región constante de cadena ligera de cadena ligera λ o κ de inmunoglobulina de roedor. En una modalidad, la secuencia de ácidos nucleicos genómica comprende una secuencia de ácidos nucleicos de región variable de cadena ligera λ o κ de inmunoglobulina humana reordenada unida operativamente a una secuencia de ácidos nucleicos de región constante de cadena ligera de cadena ligera λ o κ de inmunoglobulina de roedor. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena ligera comprende una secuencia de ácidos nucleicos de región constante de rata, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos. En una modalidad, la secuencia de ácidos nucleicos de región constante de cadena ligera λ o κ de inmunoglobulina comprende una secuencia de ácidos nucleicos de región constante de rata, una secuencia de ácidos nucleicos de región constante humana, o una combinación de estos.

15

20

El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo puede codificar una proteína extracelular o un ligando para un receptor. En modalidades específicas, el ligando codificado es una citocina. Las citocinas de interés incluyen una quimiocina seleccionada de o que comprende CCL, CXCL, CX3CL, y/o XCL. La citocina puede comprender, además, un factor de necrosis tumoral (TNF). Aún en otras modalidades, la citocina es una interleucina (IL). En una modalidad, la interleucina se selecciona de o comprende IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IL-34, IL-35, y/o IL-36. En una modalidad, la interleucina es IL-2. En modalidades específicas, tales polinucleótidos de interés dentro del inserto de ácido nucleico y/o integrado en el locus genómico objetivo son de un ser humano y, en modalidades más específicas, puede comprender la secuencia genómica humana.

25

30

El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus genómico objetivo puede codificar la Apolipoproteína E (ApoE).

35

El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo puede codificar una proteína citoplasmática o una proteína de membrana. En una modalidad, la proteína de membrana es un receptor, tal como, un receptor de citocina, un receptor de interleucina, una cadena alfa de receptor de interleucina-2, una cadena beta de receptor de interleucina-2, una cadena gamma de receptor de interleucina-2 o receptor de tirosina cinasa. En otros casos, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo puede comprender una región ortóloga u homóloga del locus objetivo.

40

El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo puede comprender un polinucleótido que codifica al menos una región de un receptor de células T, que incluye el receptor alfa de células T. En métodos específicos cada uno de los insertos de ácidos nucleicos comprenden una región genómica del locus de receptor de células T (es decir, el locus del receptor alfa de células T) de manera que, tras finalizar la integración seriada, una porción o la totalidad del locus genómico del receptor de células T se ha integrado en el locus objetivo. Tales insertos de ácidos nucleicos pueden comprender al menos uno o más de un segmento variable o un segmento de unión de un locus del receptor de células T (es decir, del locus del receptor alfa de células T). Aún en otras modalidades, el polinucleótido de interés que codifica la región del receptor de células T puede ser de, por ejemplo, un polinucleótido de un eucariota, un eucariota no rata, un mamífero, un mamífero no humano, roedor, roedor no rata, ratón, rata, un humano, un mono, un hámster, un mamífero agrícola o un mamífero doméstico que codifica una proteína mutante.

45

50

En otras modalidades, el polinucleótido de interés integrado en el locus objetivo codifica una proteína nuclear. En una modalidad, la proteína nuclear es un receptor nuclear. En modalidades específicas, tales polinucleótidos de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo son de un ser humano y, en modalidades más específicas, pueden comprender la secuencia genómica humana.

55

El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus genómico objetivo puede comprender una modificación genética en una secuencia codificante. Tales modificaciones genéticas incluyen, pero sin limitarse a, una mutación de delección de una secuencia codificante o la fusión de dos secuencias codificantes.

60

El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo puede comprender un polinucleótido que codifica una proteína mutante, que incluye, por ejemplo, una proteína mutante humana. En una modalidad, la proteína mutante se caracteriza por una característica de unión alterada, localización alterada, expresión alterada, y/o patrón de expresión alterado. En una modalidad, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo comprende al menos un alelo de enfermedad, que incluye, por ejemplo, un alelo

65

- de una enfermedad neurológica, un alelo de una enfermedad cardiovascular, un alelo de una enfermedad renal, un alelo de una enfermedad muscular, un alelo de una enfermedad de la sangre, un alelo de un gen que provoca cáncer, o un alelo de una enfermedad del sistema inmunitario. En tales casos, el alelo de enfermedad puede ser un alelo dominante o el alelo de enfermedad es un alelo recesivo. Además, el alelo de enfermedad puede comprender un alelo de polimorfismo de nucleótido único (SNP). El polinucleótido de interés que codifica la proteína mutante puede ser de cualquier organismo, que incluye, pero sin limitarse a, un polinucleótido de un eucariota, un eucariota no rata, un mamífero, un mamífero no humano, roedor, roedor no rata, ratón, rata, un humano, un hámster, un mono, un mamífero agrícola o un mamífero doméstico que codifica una proteína mutante.
- 5
- 10 En una modalidad, la modificación genética produce una forma mutante de una proteína con una característica de unión alterada, localización alterada, expresión alterada, y/o patrón de expresión alterado.
- En una modalidad, la modificación genética produce una delección, adición, sustitución o una combinación de estos de una región del locus de *ApoE*, por ejemplo, el locus de *ApoE* de rata, en donde la modificación genética en el locus de *ApoE* da como resultado una disminución de la actividad de *ApoE*. En una modalidad, se genera una inactivación de *ApoE*.
- 15
- En una modalidad, la modificación genética produce una delección, adición, sustitución o una combinación de estos de una región del locus de *Rag1*, por ejemplo, el locus de *Rag1* de rata, en donde la modificación genética en el locus de *Rag1* da como resultado una disminución de la actividad de *Rag1*. En una modalidad, se genera una inactivación de *Rag1*. En una modalidad, la modificación genética produce una delección, adición, sustitución o una combinación de estos de una región del locus de *Rag2*, por ejemplo, el *Rag2* de rata, en donde la modificación genética en el locus de *Rag2* da como resultado una disminución en la actividad de *Rag2*. En una modalidad, se genera una inactivación de *Rag2*. En una modalidad, la modificación genética produce una delección, adición, sustitución o una combinación de estos de una región del locus de *Rag1/Rag2*, por ejemplo, el locus de *Rag1/Rag2* de rata, en donde la modificación genética en el locus de *Rag1/Rag2* da como resultado una disminución de la actividad de *Rag1* y una disminución de la actividad de *Rag2*. En una modalidad, se genera una inactivación de *Rag1/Rag2*.
- 20
- 25
- En una modalidad, la modificación genética produce una delección, adición, sustitución o una combinación de estos de una región del locus de la cadena gamma de receptor de interleucina-2, por ejemplo, el locus de la cadena gamma de receptor de interleucina-2 de rata, en donde la modificación genética en el locus de la cadena gamma de receptor de interleucina-2 da como resultado una disminución de la cadena gamma de receptor de interleucina-2. En una modalidad, se genera una inactivación de la cadena gamma de receptor de interleucina-2.
- 30
- Como se analiza en cualquier otra parte de la presente descripción, las modalidades adicionales proporcionadas en la presente descripción comprende uno o más del locus de *ApoE*, el locus de la cadena gamma de receptor de interleucina-2, el locus de *Rag2*, el locus de *Rag1* y/o el locus de *Rag2/Rag1*, por ejemplo, el locus de *ApoE* de rata, el locus de la cadena gamma de receptor de interleucina-2 de rata, el locus de *Rag2*, el locus de *Rag1* y/o el locus de *Rag2/Rag1*, se modifica mediante la sustitución de una porción del locus de *ApoE* de rata, el locus de la cadena gamma de receptor de interleucina-2, el locus de *Rag2*, el locus de *Rag1* y/o el locus de *Rag2/Rag1* con la porción ortóloga correspondiente de locus de *ApoE*, un locus de la cadena gamma de receptor de interleucina-2, un locus de *Rag2*, un locus de *Rag1* y/o un locus de *Rag2/Rag1* de otro organismo.
- 35
- 40
- En una modalidad, se generan múltiples modificaciones genéticas. En una modalidad, una modificación genética produce una delección, adición, sustitución o una combinación de estos de una región del locus de la cadena gamma de receptor de interleucina-2, por ejemplo, el locus de la cadena gamma de receptor de interleucina-2 de rata, en donde la modificación genética en el locus de la cadena gamma de receptor de interleucina-2 da como resultado una disminución de la cadena gamma de receptor de interleucina-2 y una segunda modificación genética produce una delección, adición, sustitución o una combinación de estos de una región del locus de *Rag2* de rata, en donde la modificación genética en el locus de *Rag2* da como resultado una disminución de la actividad de *Rag2*. En una modalidad, se genera una inactivación de la cadena gamma de receptor de interleucina-2/*Rag2*. Una rata tal tiene un fenotipo de SCID.
- 45
- 50
- En una modalidad, el ácido nucleico de mamífero comprende un locus genómico que codifica una proteína expresada en el sistema nervioso, el sistema esquelético, el sistema digestivo, el sistema circulatorio, el sistema muscular, el sistema respiratorio, el sistema cardiovascular, el sistema linfático, el sistema endocrino, el sistema urinario, el sistema reproductor, o una combinación de estos. En una modalidad, el ácido nucleico de mamífero comprende un locus genómico que codifica una proteína expresada en una médula ósea o una célula derivada de médula ósea. En una modalidad, el ácido nucleico comprende un locus genómico que codifica una proteína expresada en una célula de bazo. En una modalidad, el locus genómico comprende una secuencia de ADN genómico de ratón, una secuencia de ADN genómico de rata, una secuencia de ADN genómico humano, o una combinación de estos. En una modalidad, el locus genómico comprende, en cualquier orden, las secuencias de ADN genómico de rata o ser humano. En una modalidad, el locus genómico comprende, en cualquier orden, las secuencias de ADN genómico de ratón o ser humano. En una modalidad, el locus genómico comprende, en cualquier orden, las secuencias de ADN genómico de ratón o rata. En una modalidad, el locus genómico comprende, en cualquier orden, las secuencias de ADN genómico de rata, ratón y ser humano.
- 55
- 60

En una modalidad, el inserto de ácido nucleico comprende una modificación genética en una secuencia codificante de un gen. En una modalidad, la modificación genética comprende una mutación de delección en la secuencia codificante. En una modalidad, la modificación genética comprende una fusión de dos secuencias codificantes endógenas.

5 En una modalidad, la modificación genética comprende una delección de una secuencia no codificante de proteínas, pero no comprende una delección de una secuencia codificante de proteínas. En una modalidad, la delección de la secuencia no codificante de proteínas comprende una delección de un elemento regulador. En una modalidad, la modificación genética comprende una adición de un promotor. En una modalidad, la modificación genética comprende una sustitución de un promotor o elemento regulador. En una modalidad, el elemento regulador es un potenciador. En una modalidad, el elemento regulador es un elemento de unión a represor transcripcional.

10 En una modalidad, la modificación genética comprende la ubicación de una secuencia de ácidos nucleicos humana que codifica una proteína humana mutante. En una modalidad, la modificación genética comprende al menos un alelo de enfermedad humana de un gen humano. En una modalidad, la enfermedad humana es una enfermedad neurológica. En una modalidad, la enfermedad humana es una enfermedad cardiovascular. En una modalidad, la enfermedad humana es una enfermedad renal. En una modalidad, la enfermedad humana es una enfermedad muscular. En una modalidad, la enfermedad humana es una enfermedad de la sangre. En una modalidad, la enfermedad humana es un cáncer. En una modalidad, la enfermedad humana es una enfermedad del sistema inmunitario. En una modalidad, el alelo de enfermedad humana es un alelo dominante. En una modalidad, el alelo de enfermedad humana es un alelo recesivo. En una modalidad, el alelo de enfermedad humana comprende un alelo de polimorfismo de nucleótido único (SNP).

15 El polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus objetivo puede comprender, además, una secuencia reguladora, que incluye, por ejemplo, una secuencia promotora, una secuencia potenciadora, o una secuencia de unión a represor transcripcional. En modalidades específicas, el polinucleótido de interés dentro del inserto de ácido nucleico y/o integrado en el locus genómico objetivo comprende un polinucleótido que tiene una delección de una secuencia no codificante de proteínas, pero no comprende una delección de una secuencia codificante de proteínas. En una modalidad, la delección de la secuencia no codificante de proteínas comprende una delección de una secuencia reguladora. En otra modalidad, la delección del elemento regulador comprende una delección de una secuencia promotora. En una modalidad, la delección del elemento regulador comprende una delección de una secuencia potenciadora. Un polinucleótido de interés tal puede ser de cualquier organismo, que incluye, pero sin limitarse a, un polinucleótido de un eucariota, un eucariota no rata, un mamífero, un mamífero no humano, un roedor, un roedor no rata, un ratón, una rata, un ser humano, un mono, un mamífero agrícola o un mamífero doméstico que codifica una proteína mutante.

35 5. Métodos para introducir secuencias y generación de animales transgénicos

Como se señaló anteriormente, se proporcionan métodos y composiciones en la presente descripción para permitir la integración dirigida de uno o más polinucleótidos de interés en un locus objetivo. Tales sistemas emplean una variedad de componentes y para fácil referencia, en la presente descripción el término "sistema de integración dirigida" genéricamente comprende todos los componentes necesarios para un evento de integración (es decir, en ejemplos no limitantes, los diversos agentes nucleasas, sitios de reconocimiento, inserto de polinucleótidos de ADN, vectores de direccionamiento, locus genómico objetivo, y/o polinucleótidos de interés).

45 Los métodos proporcionados en la presente descripción comprenden introducir en una célula uno o más polinucleótidos o constructos polipeptídicos que comprenden los diversos componentes del sistema de integración genómica objetivo. "Introducir" significa presentar a la célula la secuencia (polipéptido o polinucleótido) de una manera tal que la secuencia gana acceso al interior de la célula. Los métodos proporcionados en la presente descripción no dependen de un método particular para introducir cualquier componente del sistema de integración genómica objetivo en la célula, solo que el polinucleótido gana acceso al interior de al menos una célula. Los métodos para introducir polinucleótidos en diversos tipos celulares se conocen en la técnica e incluyen, pero sin limitarse a, métodos de transfección estable, métodos de transfección transitoria, y métodos mediados por virus.

50 En esta descripción, puede usarse cualquiera de las células de cualquier organismo en los métodos proporcionados en la presente descripción. Esta descripción describe que las células son de un eucariota, un eucariota no rata, un mamífero, un mamífero no humano, un ser humano, un roedor, un roedor no rata, una rata, un ratón o un hámster. En modalidades específicas, las células son una célula eucariota, una célula eucariota no de rata, una célula pluripotente, una célula no pluripotente, una célula pluripotente no humana, una célula de mamífero no humana, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula pluripotente inducida (iPS) humana, una célula de mamífero, una célula humana, un fibroblasto, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster o una célula CHO.

60 En algunas modalidades, las células empleadas en los métodos y composiciones tienen un constructo de ADN establemente incorporado en su genoma. "Establemente incorporado" o "establemente introducido" significa la introducción de un polinucleótido en la célula de manera que la secuencia nucleotídica se integra en el genoma de la célula y es capaz de ser heredada por la progenie de esta. Puede usarse cualquier protocolo para la incorporación estable de los constructos de ADN o los diversos componentes del sistema de integración genómica objetivo.

- Los protocolos de transfección, así como los protocolos para introducir polipéptidos o secuencias de polinucleótidos en células pueden variar. Los métodos de transfección no limitantes incluyen los métodos de transfección químicos que incluyen el uso de liposomas; nanopartículas; fosfato de calcio (Graham y otros. (1973). *Virology* 52 (2): 456–67, Bacchetti y otros. (1977) *Proc Natl Acad Sci USA* 74 (4): 1590–4 y, Kriegler, M (1991). *Transfer and Expression: A Laboratory Manual*. Nueva York: W. H. Freeman and Company. páginas 96–97); dendrímeros; o polímeros catiónicos tales como DEAE-dextrano o polietilimina. Los métodos no químicos incluyen electroporación; sonoporación; y transfección óptica. La transfección basada en partículas incluye el uso de una transfección con pistola de genes, asistida por imanes (Bertram, J. (2006) *Current Pharmaceutical Biotechnology* 7, 277–28). Los métodos virales pueden usarse, además, mediante transfección.
- En una modalidad, la introducción de uno o más de los polinucleótidos en una célula es mediante electroporación, mediante inyección intracitoplasmática, mediante una infección viral, mediante un adenovirus, mediante lentivirus, mediante retrovirus, mediante transfección, mediante transfección mediada por lípidos o mediante Nucleofection™.
- En una modalidad, la introducción de uno o más de los polinucleótidos en una célula comprende, además: introducir un constructo de expresión que comprende una secuencia de ácidos nucleicos de interés unido operativamente a un promotor. En una modalidad, el promotor es un promotor constitutivamente activo. En una modalidad, el promotor es un promotor inducible. En una modalidad, el promotor es activo en una célula madre, por ejemplo, una célula madre embrionaria.
- En una modalidad, el constructo de expresión se introduce junto con el LTVEC. En una modalidad, el constructo de expresión se introduce por separado del LTVEC en un periodo de tiempo.
- En una modalidad, la introducción de los uno o más polinucleótidos en la célula puede realizarse múltiples veces en un periodo de tiempo. En una modalidad, la introducción de los uno o más polinucleótidos en la célula se realizan al menos dos veces en un periodo de tiempo, al menos tres veces en un periodo de tiempo, al menos cuatro veces en un periodo de tiempo, al menos cinco veces en un periodo de tiempo, al menos seis veces en un periodo de tiempo, al menos siete veces en un periodo de tiempo, al menos ocho veces en un periodo de tiempo, al menos nueve veces en un periodo de tiempo, al menos diez veces en un periodo de tiempo, al menos once veces, al menos doce veces en un periodo de tiempo, al menos trece veces en un periodo de tiempo, al menos catorce veces en un periodo de tiempo, al menos quince veces en un periodo de tiempo, al menos dieciséis veces en un periodo de tiempo, al menos diecisiete veces en un periodo de tiempo, al menos dieciocho veces en un periodo de tiempo, al menos diecinueve veces en un periodo de tiempo, o al menos veinte veces en un periodo de tiempo.
- En una modalidad, el agente nucleasa se introduce en la célula simultáneamente con el vector de direccionamiento o el vector de direccionamiento grande (LTVEC). Alternativamente, el agente nucleasa se introduce por separado del vector de direccionamiento o el LTVEC en un periodo de tiempo. En una modalidad, el agente nucleasa se introduce antes de la introducción del vector de direccionamiento o el LTVEC, mientras que, en otras modalidades, el agente nucleasa se introduce después de la introducción del vector de direccionamiento o el LTVEC.
- En una modalidad, la etapa de tamizaje comprende un ensayo cuantitativo para evaluar la modificación de alelo (MOA) de un cromosoma progenitor. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de una PCR cuantitativa. En una modalidad, la PCR cuantitativa es una PCR en tiempo real (qPCR). En una modalidad, la PCR en tiempo real comprende un primer conjunto de cebadores que reconocen el locus objetivo y un segundo conjunto de cebadores que reconocen un locus de referencia que no se direcciona. En una modalidad, el conjunto de cebadores comprende una sonda fluorescente que reconoce la secuencia amplificada. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de la hibridación in situ mediada por fluorescencia (FISH). En una modalidad, el ensayo cuantitativo se lleva a cabo a través de la hibridación genómica comparativa. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de la amplificación de ADN isotérmica. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de la amplificación de ADN isotérmica. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de hibridación cuantitativa a una(s) sonda(s) inmovilizada(s). En una modalidad, el ensayo cuantitativo se lleva a cabo a través de Invader Probes®. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de MMP assays®. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de Baliza Molecular TaqMan®. En una modalidad, el ensayo cuantitativo se lleva a cabo a través de la tecnología de sondas Eclipse™. (Ver, por ejemplo, el documento US2005/0144655).
- Se proporciona, además, un método para generar un animal no humano humanizado, que comprende: (a) modificar un genoma de una célula pluripotente con un vector de direccionamiento que comprende un inserto de ácido nucleico que comprende una secuencia de ácidos nucleicos humana para formar una célula donante; (b) introducir la célula donante en un embrión huésped; y (c) gestar el embrión huésped en una madre sustituta; en donde la madre sustituta produce una progenie que comprende la secuencia de ácidos nucleicos humana. En una modalidad, la célula donante se introduce en un embrión huésped que se encuentra en la etapa de blastocisto o en una etapa de premórula (es decir, una etapa de 4 células o una etapa de 8 células). Además, la etapa (a) puede realizarse, además, con un vector de direccionamiento grande (LTVEC) y/o una secuencia de ácidos nucleicos humana de al menos 5 kb de longitud. Aún en otras modalidades, la modificación genética es capaz de transmitirse a través de la línea germinal.

dieciocho veces en un periodo de tiempo mediante el uso de una pluralidad de endonucleasas que reconocen diferentes secuencias objetivos, al menos diecinueve veces en un periodo de tiempo mediante el uso de una pluralidad de endonucleasas que reconocen diferentes secuencias objetivos, o al menos veinte veces en un periodo de tiempo mediante el uso de una pluralidad de endonucleasas que reconocen diferentes secuencias objetivos.

5 En una modalidad, la etapa de introducción está mediada por electroporación, por inyección intracitoplasmática, por un adenovirus, por lentivirus, por retrovirus, por transfección, por transfección mediada por lípidos o está mediada a través de Nucleofection™.

10 En una modalidad, el método comprende, además, introducir un ácido nucleico exógeno en la célula pluripotente genéticamente modificada. En una modalidad, el ácido nucleico exógeno es un transgén. En una modalidad, el ácido nucleico exógeno se introduce en un locus endógeno. En una modalidad, el ácido nucleico exógeno se introduce por vía ectópica (por ejemplo, en un locus diferente de su locus endógeno).

15 En un aspecto, se proporciona un método para generar un animal no humano genéticamente modificado, que comprende modificar un locus genómico de interés en una célula pluripotente empleando la modificación genética del genoma guiada por ARN para introducir una modificación en un locus genómico de interés para formar una célula pluripotente modificada, mantener la célula pluripotente modificada bajo condiciones suficientes para mantener la pluripotencia, emplear la célula pluripotente modificada como una célula donante en un embrión huésped y gestar el embrión huésped que comprende la
20 célula pluripotente modificada en una madre sustituta, en donde el embrión huésped es gestado por la madre sustituta y nace una progenie genéticamente modificada.

En una modalidad, el método tiene una tasa de transformación en el intervalo de aproximadamente 2 % a aproximadamente 80 %.

25 En una modalidad, el método comprende cointroducir una pluralidad del segundo constructo de expresión que comprende diferentes secuencias objetivo genómicas para la edición múltiple de diferentes loci genómicos. En una modalidad, el método comprende introducir una pluralidad del segundo constructo de expresión que comprende diferentes secuencias objetivo genómicas para la edición múltiple de diferentes loci genómicos en un periodo de tiempo.

30 En una modalidad, la etapa de introducción se realiza múltiples veces en un periodo de tiempo. En una modalidad, la etapa de introducción (b) se realiza al menos dos veces en un periodo de tiempo, al menos tres veces en un periodo de tiempo, al menos cuatro veces en un periodo de tiempo, al menos cinco veces en un periodo de tiempo, al menos seis veces en un periodo de tiempo, al menos siete veces en un periodo de tiempo, al menos ocho veces en un periodo de tiempo, al menos nueve veces en un periodo de tiempo, al menos diez veces en un periodo de tiempo, al menos once veces en un periodo de tiempo, al menos doce veces en un periodo de tiempo, al menos trece veces en un periodo de tiempo, al menos catorce veces en un periodo de tiempo, al menos quince veces en un periodo de tiempo, al menos dieciséis veces en un periodo de tiempo, al menos diecisiete veces en un periodo de tiempo, al menos dieciocho veces en un periodo de tiempo, al menos diecinueve veces en un periodo de tiempo, al menos veinte veces en un periodo de tiempo.

En una modalidad, el primer constructo de expresión y el segundo constructo de expresión se expresan a partir de un mismo plásmido.

45 En una modalidad, la etapa de introducción está mediada por electroporación, por inyección intracitoplasmática, por un adenovirus, por lentivirus, por retrovirus, por transfección, por transfección mediada por lípidos o está mediada a través de Nucleofection™.

50 En una modalidad, el método comprende, además, introducir un ácido nucleico exógeno en la célula pluripotente que comprende el alelo mutante.

En una modalidad, el ácido nucleico exógeno es un transgén. En una modalidad, el ácido nucleico exógeno se introduce en un locus endógeno. En una modalidad, el ácido nucleico exógeno se ubica por vía ectópica (por ejemplo, en un locus diferente de su locus endógeno).

55 En una modalidad, el método comprende, además, introducir un ácido nucleico exógeno en la célula pluripotente genéticamente modificada. En una modalidad, el ácido nucleico exógeno es un transgén. En una modalidad, el ácido nucleico exógeno se introduce en un locus endógeno. En una modalidad, el ácido nucleico exógeno se introduce por vía ectópica (por ejemplo, en un locus diferente de su locus endógeno).

60 En un aspecto, se proporciona un método para generar un animal no humano humanizado, que comprende modificar un genoma de una célula pluripotente con un LTVEC que comprende un inserto que comprende una secuencia humana de al menos 5 kb, y emplear la célula pluripotente como una célula donante, introducir la célula donante en un embrión huésped, y gestar el embrión huésped en una madre sustituta, en donde la madre sustituta da a luz una progenie que comprende la humanización.

65

Se proporcionan otros métodos para generar un animal no humano genéticamente modificado que comprende en su línea germinal una o más modificaciones genéticas como se describe en la presente descripción, que comprenden: (a) modificar un locus objetivo contenido en una célula procariota empleando los diversos métodos descritos en la presente descripción; (b) seleccionar una célula procariota modificada que comprende la modificación genética en el locus transformado; (c) aislar el vector de direccionamiento genéticamente modificado a partir del genoma de la célula procariota modificada; (d) introducir el vector de direccionamiento genéticamente modificado en una célula pluripotente para generar una célula pluripotente genéticamente modificada que comprende el inserto de ácido nucleico en el locus genómico transformado; (e) seleccionar la célula pluripotente genéticamente modificada; (f) introducir la célula pluripotente genéticamente modificada en un embrión huésped en una etapa de premórula; y (g) implantar el embrión huésped que comprende la célula pluripotente genéticamente modificada en una madre sustituta para generar una generación F0 derivada de la célula pluripotente genéticamente modificada. En tales métodos el vector de direccionamiento puede comprender un vector de direccionamiento grande. La célula pluripotente puede ser una célula ES. En modalidades adicionales, la etapa de aislamiento (c) comprende, además, (c1) linealizar el vector de direccionamiento genéticamente modificado (es decir, el LTVEC genéticamente modificado). Aún en otras modalidades, la etapa de introducción (d) comprende, además, (d1) introducir un agente nucleasa como se describe en la presente descripción en la célula pluripotente. En una modalidad, las etapas de selección (b) y/o (e) se llevan a cabo mediante la aplicación de un agente de selección como se describe en la presente descripción a la célula procariota o la célula pluripotente. En una modalidad, las etapas de selección (b) y/o (e) se llevan a cabo a través de un ensayo de modificación de alelo (MOA) como se describe en la presente descripción.

Se proporcionan los métodos adicionales para modificar un locus genómico objetivo de una célula de mamífero a través de la recombinación homóloga bacteriana (BHR) en una célula procariota y comprenden: (a) proporcionar una célula procariota que comprende un locus objetivo que comprende un ácido nucleico, (b) introducir en la célula procariota un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde el inserto de ácido nucleico comprende una región de mamífero (que incluye, por ejemplo, un inserto de ADN de un humano), y (c) seleccionar una célula procariota transformada que comprende el inserto de ácido nucleico en el locus objetivo, en donde la célula procariota es capaz de expresar una recombinasa que media la BHR. La etapa (a1) puede comprender proporcionar una célula procariota que comprende un locus objetivo que comprende un ácido nucleico que comprende un primer polinucleótido que comprende un primer sitio de reconocimiento para un primer agente nucleasa, y la etapa (b1) puede comprender, además, expresar en la célula procariota un agente nucleasa que genera una mella o ruptura de la doble cadena en o cerca del primer sitio de reconocimiento. Las etapas (a)-(c) pueden repetirse en serie como se describe en la presente descripción para permitir la introducción de múltiples insertos de ácidos nucleicos en el locus objetivo en la célula procariota. Una vez que el locus genómico transformado se "construye" con la célula procariota, puede aislarse un vector de direccionamiento que comprende el locus objetivo modificado a partir de la célula procariota e introducirse en un locus genómico objetivo dentro de una célula pluripotente. Las células pluripotentes (es decir, células ES) que comprenden el locus genómico modificado pueden generarse después en animales no humanos genéticamente modificados.

En algunas modalidades, diversas modificaciones genéticas de los loci genómicos objetivo descritos en la presente descripción pueden llevarse a cabo mediante una serie de reacciones de recombinación homóloga (BHR) en célula bacterianas mediante el uso de un LTVEC derivado del ADN de cromosoma artificial bacteriano (BAC) mediante el uso de la tecnología de modificación genética de VELOCIGENE®, la patente de los Estados Unidos núm. 6,586,251 y Valenzuela, D. M. y otros (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, *Nature Biotechnology* 21(6): 652-659).

En algunas modalidades, las células ES transformadas que comprenden varias modificaciones genéticas como se describe en la presente descripción se usan como insertos de células ES y se introducen en un embrión en etapa de premórula de un organismo correspondiente, por ejemplo, un embrión de ratón en etapa de 8 células, a través del método de VELOCIMOUSE® (ver, por ejemplo, los documentos US 7,576,259, US 7,659,442, US 7,294,754, y US 2008-0078000 A1). El embrión que comprende las células ES genéticamente modificadas se incuba hasta la etapa de blastocisto y después se implanta en una madre sustituta para producir una F0. Los animales que portan el locus genómico genéticamente modificado pueden identificarse a través del ensayo de modificación de alelo (MOA) como se describe en la presente descripción. El animal no humano de generación F0 resultante derivado de las células ES genéticamente modificadas se cruza con un animal no humano de tipo silvestre para obtener la descendencia de generación F1. Después del genotipado con cebadores y/o sondas específicas, los animales no humanos F1 que son heterocigotos para el locus genómico genéticamente modificado se cruzan entre sí para producir animales que son homocigotos para el locus genómico genéticamente modificado. Alternativamente, un animal no humano hembra F0 y un animal no humano macho F0 que tienen cada uno la modificación genética pueden cruzarse para obtener un animal no humano F1 homocigoto para la modificación genética.

En un aspecto, se proporciona, por ejemplo, un genoma de rata genéticamente modificado, que comprende una modificación dirigida de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos homóloga u ortóloga de otro organismo.

En una modalidad, la secuencia de ácidos nucleicos homólogas u ortólogas es de una longitud de aproximadamente 5 kb a aproximadamente 200 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 5 kb a aproximadamente 10 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata

homóloga u ortóloga varía de aproximadamente 10 kb a aproximadamente 20 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 20 kb a aproximadamente 30 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 30 kb a aproximadamente 40 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 40 kb a aproximadamente 50 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 50 kb a aproximadamente 60 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 60 kb a aproximadamente 70 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 70 kb a aproximadamente 80 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 80 kb a aproximadamente 90 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 90 kb a aproximadamente 100 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 100 kb a aproximadamente 110 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 110 kb a aproximadamente 120 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 120 kb a aproximadamente 130 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 140 kb a aproximadamente 150 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 150 kb a aproximadamente 160 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 160 kb a aproximadamente 170 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 170 kb a aproximadamente 180 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 180 kb a aproximadamente 190 kb. En una modalidad, la secuencia de ácidos nucleicos no de rata homóloga u ortóloga varía de aproximadamente 190 kb a aproximadamente 200 kb. Diversos polinucleótidos de interés que pueden emplearse en el inserto de ácido nucleico se describen en cualquier otra parte de la presente descripción.

Se proporciona los métodos adicionales para la modificación del genoma objetivo de un animal no humano. Tales métodos pueden comprender (a) modificar un locus genómico de interés en una célula pluripotente no humana de acuerdo con cualquiera de los diversos métodos proporcionados en la presente descripción para modificar un locus genómico de interés, produciendo de este modo una célula pluripotente genéticamente modificada no humana que comprende una modificación del genoma objetivo; (b) introducir la célula pluripotente modificada no humana de la etapa (a) en un embrión huésped no humano; y (c) gestar el embrión huésped no humano que comprende la célula pluripotente modificada en una madre sustituta, en donde la madre sustituta produce la progenie F0 que comprende la modificación del genoma objetivo, y en donde la modificación del genoma objetivo es capaz de transmitirse a través de la línea germinal.

En algunas modalidades, la modificación del genoma objetivo comprende la delección simultánea de una secuencia de ácidos nucleicos endógena en el locus genómico de interés y la inserción de un ácido nucleico exógeno en el locus genómico de interés (es decir, la delección y la inserción en una única etapa). En algunas modalidades, la modificación del genoma objetivo comprende una modificación genética bialélica. La modificación genética bialélica puede comprender la delección de una secuencia de ácidos nucleicos endógena y la inserción de un ácido nucleico exógeno en el locus genómico de interés en dos cromosomas homólogos (es decir, un par de los cromosomas homólogos primero y segundo).

En otras modalidades, la modificación del genoma objetivo genera una célula pluripotente modificada que es heterocigota compuesta en el locus genómico de interés. En otras modalidades, la modificación del genoma objetivo genera una célula pluripotente modificada que es hemicigota en el locus genómico de interés. En algunas modalidades, la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción de un ácido nucleico exógeno. Por ejemplo, la modificación genética dirigida puede comprender: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en dos cromosomas homólogos; y (2) inserción de un ácido nucleico exógeno en el locus genómico de interés en un primer cromosoma y alteración del locus genómico de interés en un segundo cromosoma. El primer cromosoma puede ser el primero de los dos cromosomas homólogos, y el segundo cromosoma puede ser el segundo de los dos cromosomas homólogos.

6. Células

Los diversos métodos y composiciones descritos en la presente descripción emplean un sistema de transformación de locus genómico en una célula. En una modalidad, la célula es una célula pluripotente. La descripción describe que la célula es una célula no pluripotente. En una modalidad, la célula pluripotente es una célula pluripotente no humana. En una modalidad, la célula pluripotente no humana es una célula pluripotente de mamífero. La descripción describe que la célula pluripotente es una célula madre pluripotente inducida (iPS) humana.

La descripción describe que la célula es una célula eucariota, una célula eucariota no de rata, una célula pluripotente humana, una célula ES humana, una célula madre adulta humana, una célula progenitora humana con desarrollo limitado, una célula de mamífero no humana, una célula de mamífero, una célula humana, un fibroblasto, una célula de roedor, una célula de roedor no rata, una célula de rata, una célula de ratón, una célula de hámster o una célula CHO.

La descripción describe que una célula eucariota es una célula primaria. Las células primarias incluyen células o cultivos de células que se han aislado directamente de un organismo, órgano, o tejido. Las células primarias incluyen células que no son transformadas ni inmortalizadas. Estas incluyen cualquier célula obtenida de un organismo, órgano, o tejido que no se ha sometido a pases en cultivo de tejidos o se han sometido a pases en cultivo de tejidos, pero es incapaz de ser sometida a pases indefinidamente en cultivo de tejidos. Tales células pueden aislarse mediante técnicas convencionales e incluyen, por ejemplo, células hematopoyéticas, células endoteliales, células epiteliales, fibroblastos, células mesenquimales, queratinocitos, melanocitos, monocitos, células mononucleares, adipocitos, preadipocitos, neuronas, células gliales, hepatocitos, mioblastos esqueléticos, y células del músculo liso. En algunas modalidades, las células primarias se derivan de tejidos conectivos, tejidos musculares, tejidos del sistema nervioso, o tejidos epiteliales.

La descripción describe que una célula eucariota es una célula inmortalizada. Las células inmortalizadas incluyen células de un organismo multicelular que podrían no proliferar normalmente pero, debido a una mutación o alteración, han evadido la senescencia celular normal y en su lugar pueden seguir experimentando división. Tales mutaciones o alteraciones pueden producirse de manera natural o inducirse intencionalmente. Los ejemplos de células inmortalizadas incluyen células de ovario de hámster chino (CHO), células renales embrionarias humanas (por ejemplo, células HEK 293), y células de fibroblastos embrionarios de ratón (por ejemplo, células 3T3). Se conocen en la técnica numerosos tipos de células inmortalizadas.

La descripción describe que las células inmortalizadas se derivan de células de cáncer. En otra modalidad, una célula primaria o inmortalizada es una que se usa típicamente para cultivar o para expresar genes o proteínas recombinantes.

En otras modalidades, la célula pluripotente es capaz de mantener su pluripotencia después de al menos una modificación genética dirigida de su genoma y es capaz de transmitir la modificación dirigida a una línea germinal de una generación F1.

En una modalidad, la célula pluripotente es un huevo no humano fertilizado en la etapa de célula única. En una modalidad, el huevo no humano fertilizado es un huevo fertilizado de mamífero. En una modalidad, el huevo fertilizado de mamífero es un huevo fertilizado de roedor en la etapa de célula única. La descripción describe que el huevo fertilizado de mamífero es un huevo fertilizado de rata o ratón en la etapa de célula única.

La descripción describe que las diversas células empleadas en el método y las composiciones descritas en la presente descripción pueden comprender, además, células procariotas, tales como una célula bacteriana, que incluye *E. coli*. En ejemplos específicos, la célula procariota es una cepa de *E. coli* competente para recombinación. En un ejemplo, la célula procariota comprende un ácido nucleico que codifica la recombinasa, mientras que, en otros casos, la célula procariota no comprende el ácido nucleico que codifica la recombinasa, y el ácido nucleico que codifica la recombinasa se introduce en la célula procariota. En una modalidad, el ácido nucleico que codifica la recombinasa comprende un ADN o un ARNm. En algunas modalidades, el ácido nucleico que codifica la recombinasa es pABG. En una modalidad, la recombinasa se expresa bajo el control de un promotor inducible. En una modalidad, la expresión de la recombinasa es controlada por arabinosa.

A. Medio de baja osmolalidad para generar y mantener células madre pluripotentes inducidas humanas

Se proporciona un medio de cultivo celular para el uso en los métodos y composiciones de la invención. La descripción describe que el medio es adecuado para generar una población de células iPS humanas. La descripción describe que el medio es adecuado para mantener células iPS humanas en cultivo. La descripción describe que las células iPS humanas son vírgenes o parecen vírgenes.

El medio proporcionado en la presente descripción comprende al menos un medio base, suplementos, un polipéptido del factor inhibidor de la leucemia (LIF), un inhibidor de glucógeno sintasa cinasa 3 (GSK3), y un inhibidor de MEK.

El medio presente es un medio de baja osmolalidad. En un ejemplo, la osmolalidad es de entre aproximadamente 175-280 mOsm/kg. En ejemplos adicionales, la osmolalidad del medio es de aproximadamente 180-270 mOsm/kg, aproximadamente 200-250 mOsm/kg, aproximadamente 220-240 mOsm/kg, o aproximadamente 225-235 mOsm. En una modalidad particular, la osmolalidad del medio es de aproximadamente 233 mOsm/kg.

El medio base proporcionado por la invención es un medio base de baja osmolalidad al que se le añaden suplementos. El medio base presente difiere de los medios se usados típicamente para mantener células iPS humanas en cultivo, que incluyen Medio Eagle Modificado de Dulbecco (DMEM), en diversas formas (por ejemplo, Invitrogen DMEM, Cat. núm. 1 1971 -025), y un DMEM bajo en sales disponible comercialmente como KO-DMEM™ (Invitrogen Cat. núm. 10829-018).

El medio base proporcionado en la presente descripción es un medio de baja osmolalidad, pero muestra características que no se limitan a baja osmolalidad. Por ejemplo, la formulación de DMEM mostrada en la Tabla A puede hacerse adecuada para los fines de la invención mediante la alteración de las concentraciones de cloruro de sodio y/o bicarbonato de sodio como se proporciona en la presente descripción, lo que dará como resultado una osmolalidad diferente en comparación con el medio base DMEM estándar o el medio base DMEM bajo en sales (KO-DMEM) mostrado en la Tabla A.

ES 2 700 596 T3

Tabla A: Formulación del medio base DMEM.

	Componente	Mg/L	mM
5	Glicina	30	0,4
	L-Arginina•HCl	84	0,398
	L-Cisteína•2HCl	63	0,201
10	L-glutamina	584	4
	L-Histidina•HCl•H ₂ O	42	0,2
	L-Isoleucina	105	0,802
	L-Leucina	105	0,802
15	L-Lisina•HCl	146	0,798
	L-Metionina	30	0,201
	L-Fenilalanina	66	0,4
20	L-Serina	42	0,4
	L-Treonina	95	0,798
	L-Triptófano	16	0,0784
25	Sal de sodio dihidratada de L-Tirosina	104	0,398
	L-Valina	94	0,803
	Cloruro de colina	4	0,0286
30	Pantotenato de D-Calcio	4	8,39 x 10 ⁻³
	Ácido fólico	4	9,07 x 10 ⁻³
	Niacinamida	4	0,0328
35	Piridoxina•HCl	4	0,0196
	Riboflavina	0,4	1,06 x 10 ⁻³
	Tiamina•HCl	4	0,0119
40	i-Inositol	7.2	0,04
	Cloruro de calcio (CaCl ₂) (anhidro)	200	1.8
	Nitrato férrico (Fe(NO ₃) ₃ •9H ₂ O)	0,1	2,48 x 10 ⁻⁴
45	Sulfato de Magnesio (MgSO ₄) (anhid.)	97,67	0,814
	Cloruro de potasio (KCl)	400	5,33
	D-Glucosa (Dextrosa)	4500	25
	Rojo fenol	15	0,0399
50	Contenido de NaCl/NaHCO ₃ de DMEM		
	Bicarbonato de sodio (NaHCO ₃)	3700	44,05
	Cloruro de sodio (NaCl)	6400	110,34
55	Contenido de NaCl/NaHCO ₃ de DMEM bajo en sales (KO-DMEM)		
	Bicarbonato de sodio (NaHCO ₃)	2200	26
60	Cloruro de sodio (NaCl)	5100	87.7
	Contenido de NaCl/NaHCO ₃ de DMEM de baja osmolalidad		

ES 2 700 596 T3

Bicarbonato de sodio (NaHCO ₃)	2200	26
Cloruro de sodio (NaCl)	3000	50

5

El medio base presente puede incluir una sal de un metal alcalino y un haluro, tal como cloruro de sodio (NaCl). Las concentraciones ilustrativas de NaCl en el medio base incluyen 50 ± 5 mM o aproximadamente 3 mg/mL.

10

En otra modalidad, el medio base muestra una concentración de una sal de ácido carbónico. La sal de ácido carbónico puede ser una sal de sodio. En un ejemplo tal, la sal de sodio puede ser bicarbonato de sodio. En una modalidad particular, el bicarbonato de sodio está presente en el medio base a una concentración de aproximadamente 26 ± 5 mM o aproximadamente 2,2 mg/mL.

15

Aún en otra modalidad, el medio base es un medio base de baja osmolalidad. La osmolalidad del medio base puede encontrarse dentro de un intervalo de aproximadamente 175-280 mOsm/kg, aproximadamente 180-250 mOsm/kg, aproximadamente 190-225 mOsm/kg, o aproximadamente 195-205 mOsm/kg. Una osmolalidad ilustrativa del medio base puede ser 200, 214, 216, o 218 mOsm/kg. En un ejemplo particular, la osmolalidad del medio base es de 200 mOsm/kg. La osmolalidad puede determinarse cuando las células se cultivan en diferentes concentraciones de CO₂. En algunos ejemplos, las células se cultivan en CO₂ al 3 % o CO₂ al 5 %.

20

En una modalidad preferida, el medio base comprende NaCl a una concentración de 3,0 mg/mL, bicarbonato de sodio a una concentración de aproximadamente 2,2 mg/mL, y tiene una osmolalidad de 200 mOsm/kg.

25

Los suplementos formulados con el medio base de la invención son adecuados para generar, mantener, o enriquecer poblaciones de células iPS humanas descritas en la presente descripción. Tales suplementos se indican como "suplementos" o "+ suplementos" en esta descripción. El término "suplementos" o la frase "+ suplementos," incluye uno o más elementos adicionales añadidos a los componentes del medio base descritos en la Tabla A. Por ejemplo, los suplementos pueden incluir, sin limitación, medio F-12® (Gibco), suplemento N2® (Gibco; solución a 100X), medio NEUROBASAL® (Gibco), suplemento B-27® (Gibco; solución a 50X), L-glutamina, glucosa, 2-mercaptoetanol, un polipéptido de Factor Inhibidor de la Leucemia (LIF), un inhibidor de glucógeno sintasa cinasa 3, un inhibidor de MEK, o cualquier combinación de estos.

30

35

En una modalidad particular, el polipéptido de LIF es un polipéptido de LIF humano (hLIF). En algunos ejemplos, un polipéptido de hLIF se usa a una concentración de aproximadamente 1-1 000 unidades/mL, aproximadamente 20-800 unidades/mL, aproximadamente 50-500 unidades/mL, aproximadamente 75-250 unidades/mL, o aproximadamente 100 unidades/mL.

40

En otra modalidad particular, el inhibidor de GSK3 comprende CHIR99021. En algunos ejemplos, CHIR99021 se usa a una concentración de aproximadamente 0,1 a 10 µM, aproximadamente 1-5 µM, aproximadamente 2-4 µM, o aproximadamente 3 µM.

45

En otra modalidad particular, el inhibidor de MEK comprende PD0325901. En algunos ejemplos, PD0325901 se usa a una concentración de aproximadamente 0,1-5 µM, aproximadamente 0,2-1 µM, aproximadamente 0,3-0,7 µM, o aproximadamente 0,5 µM.

50

Un medio ilustrativo comprende un medio base de baja osmolalidad descrito en la presente descripción a aproximadamente 24,75 % (v/v), medio F-12 a aproximadamente 24,75 % (v/v), suplemento N2 a aproximadamente 0,5 % (v/v), medio NEUROBASAL a aproximadamente 49 % (v/v), suplemento B-27 a aproximadamente 1 % (v/v), L-glutamina a aproximadamente 2 mM, 2-mercaptoetanol a aproximadamente 0,1 mM, hLIF a aproximadamente 100 unidades/mL, CHIR99021 a aproximadamente 3 µM, y PD0325901 a aproximadamente 0,5 µM.

En otra modalidad particular, el medio puede o no puede comprender factor de crecimiento de fibroblastos básico (bFGF, también conocido como FGF2 o FGF-β). Preferentemente el medio presente no comprende bFGF.

55

B. Células madre pluripotentes inducidas humanas

Esta descripción describe métodos y composiciones para generar una población de células iPS humanas. Los métodos y composiciones se proporcionan, además, en esta descripción para mantener células iPS humanas en cultivo. Las células iPS humanas que se producen o mantienen en cultivo se proporcionan, además, en esta descripción.

60

El término "célula pluripotente" o "célula madre pluripotente" incluye una célula no diferenciada que posee la capacidad de desarrollarse en más de un tipo celular diferenciado. Tales células pluripotentes pueden ser, por ejemplo, una célula madre embrionaria (célula ES) de mamífero o una célula madre pluripotente inducida (célula iPS) de mamífero. Los ejemplos de célula pluripotentes incluyen células iPS humanas.

65

El término "célula madre embrionaria" o "célula ES" significa una célula madre totipotente o pluripotente derivada de embrión, derivada de la masa celular interna de un blastocisto, que puede mantenerse en un cultivo *in vitro* bajo condiciones adecuadas. Las células ES son capaces de diferenciarse en células de cualquiera de las tres capas germinales de vertebrados, por ejemplo, el endodermo, el ectodermo, o el mesodermo. Las células ES se caracterizan, además, por su capacidad de propagarse indefinidamente bajo condiciones de cultivo *in vitro* adecuadas. Ver, por ejemplo, Thomson y otros. (Science (1998) Vol. 282(5391), páginas 1145-1147).

El término "célula madre pluripotente inducida" o "célula iPS" incluye una célula madre pluripotente que puede derivarse directamente de una célula adulta diferenciada. Las células iPS humanas pueden generarse mediante la introducción de conjuntos específicos de factores de reprogramación en una célula no pluripotente que puede incluir, por ejemplo, Oct3/4, factores de transcripción de la familia Sox (por ejemplo, Sox1, Sox2, Sox3, Sox15), factores de transcripción de la familia Myc (por ejemplo, c-Myc, l-Myc, n-Myc), factores de transcripción de la familia de tipo Krüppel (KLF) (por ejemplo, KLF1, KLF2, KLF4, KLF5), y/o factores de transcripción relacionados, tales como NANOG, LIN28, y/o Glis1. Las células iPS humanas también pueden generarse, por ejemplo, mediante el uso de ARNm, moléculas pequeñas que simulan las funciones de factores de transcripción, o especificadores de linaje. Las células iPS humanas se caracterizan por su capacidad de diferenciarse en cualquier célula de las tres capas germinales de vertebrados, por ejemplo, el endodermo, el ectodermo, o el mesodermo. Las células iPS humanas se caracterizan, además, por su capacidad de propagarse indefinidamente bajo condiciones de cultivo *in vitro* adecuadas. Ver, por ejemplo, Takahashi y Yamanaka (Cell (2006) Vol. 126(4), páginas 663-676).

Los términos "virgen" e "imprimado" identifican diferentes estados de pluripotencia de las células iPS humanas. El término "de apariencia virgen" identifica una célula que expresa un estado pluripotente que muestra una o más características de una célula pluripotente virgen. Las células iPS humanas de apariencia virgen pueden denominarse, además, células iPS humanas "de tipo virgen". La descripción describe que las células iPS humanas de apariencia virgen muestran una o más características morfológicas de células iPS humanas vírgenes, tal como una morfología caracterizada por colonias en forma de domos compactos. La descripción describe que las células iPS humanas de apariencia virgen expresan uno o más de los marcadores de pluripotencia descritos en la presente descripción. La descripción describe que las células iPS humanas vírgenes o de apariencia virgen son células iPS humanas vírgenes. La descripción describe que las células iPS humanas vírgenes o de apariencia virgen son células iPS humanas de apariencia virgen.

Las características de las células iPS vírgenes e imprimadas se describen en la técnica. Ver, por ejemplo, Nichols y Smith (Cell Stem Cell (2009) Vol. 4(6), páginas 487-492). Las células iPS humanas vírgenes muestran un estado de pluripotencia similar al de las células ES de la masa celular interna de un embrión antes de la implantación. Tales células vírgenes no son imprimadas para la especificación y el compromiso de linaje. Las células iPS vírgenes hembras se caracterizan por dos cromosomas X activos. En cultivo, la autorrenovación de células iPS humanas vírgenes es dependiente de factor inhibidor de leucemia (LIF) y otros inhibidores. Las células iPS humanas vírgenes cultivadas muestran una morfología clonal caracterizada por colonias en forma de domo redondas y una pérdida de polaridad ápico basal. Las células vírgenes cultivadas pueden mostrar, además, uno o más marcadores de pluripotencia como se describe en cualquier otra parte de la presente descripción. Bajo las condiciones apropiadas, el tiempo de duplicación de las células iPS humanas vírgenes en cultivo puede ser de entre 16 y 24 horas.

Las células iPS humanas imprimadas expresan un estado de pluripotencia similar al de las células de epiblasto después de la implantación. Tales células son imprimadas para la especificación y el compromiso del linaje. Las células iPS imprimadas hembras se caracterizan por un cromosoma X activo y un cromosoma X inactivo. En cultivo, la autorrenovación de las células iPS humanas imprimadas es dependiente del factor de crecimiento de fibroblastos (FGF) y de activina. Las células iPS humanas imprimadas cultivadas muestran una morfología clonal caracterizada por una monocapa epitelial y muestran polaridad ápico basal. Bajo las condiciones apropiadas, el tiempo de duplicación de las células iPS humanas imprimadas en cultivo puede ser de 24 horas o más.

La descripción describe que las células iPS humanas pueden derivarse de células no pluripotentes transformadas para expresar un estado pluripotente. Tales células transformadas incluyen, por ejemplo, las células que se han transformado para expresar genes de reprogramación que inducen pluripotencia. Un estado pluripotente puede incluir, por ejemplo, la expresión de uno o más de los marcadores de pluripotencia descritos en la presente descripción. Tales células (tales como fibroblastos de prepucio humano) pueden transformarse para expresar genes de reprogramación, o cualquiera de los genes de interés adicionales, mediante los medios conocidos en la técnica. Ver, por ejemplo, Takahashi y Yamanaka (Cell (2006) Vol. 126(4), páginas 663-676). Por ejemplo, estos pueden introducirse en las células mediante el uso de uno o más los plásmidos, vectores lentivirales, o vectores retrovirales. En algunos casos, los vectores se integran en el genoma y pueden eliminarse después de que finaliza la reprogramación. La descripción describe que, las células no pluripotentes se transforman con genes de reprogramación que comprenden Oct4, Sox2, Klf4, Myc, o cualquier combinación de estos. La descripción describe que las células transformadas comprenden células iPS humanas imprimadas.

La descripción describe que las células iPS humanas cultivadas en el medio de baja osmolalidad descrito en la presente descripción expresan uno o más fenotipos, perfiles de expresión de genes, o marcadores característicos de un estado virgen. En un ejemplo, las células iPS humanas expresan uno o más marcadores de pluripotencia cuya expresión es indicativa de un estado virgen. Tales marcadores de pluripotencia pueden incluir fosfatasa alcalina, NANOG, 5T4, ABCG2, Activin RIB/ALK-4, Activina RIIIB, E-Caderina, Cbx2, CD9, CD30/TNFRSF8, CD117/c-kit, CDX2, CHD1, Cripto, DNMT3B,

DPPA2, DPPA4, DPPA5/ESG1, EpCAM/TROP1, ERR beta/NR3B2, ESGP, proteína de F-box 15/FBXO15, FGF-4, FGF-5, FoxD3, GBX2, GCNF/NR6A1, GDF-3, Gi24/VISTA/B7-H5, integrina alfa 6/CD49f, integrina alfa 6 beta 1, integrina alfa 6 beta 4, integrina beta 1/CD29, KLF4, KLF5, L1TD1, Lefty, Lefty-1, Lefty-A, LIN-28A, LIN-28B, LIN-41, cMaf, cMyc, Oct-3/4, Oct-4A, Podocalixina, Rex-1/ZFP42, Smad2, Smad2/3, SOX2, SSEA-1, SSEA-3, SSEA-4, STAT3, Stella/Dppa3, SUZ12, TBX2, TBX3, TBX5, TERT, TEX19, TEX19.1, THAP11, TRA-1-60(R), TROP-2, UTF1, y/o ZIC3. En un ejemplo específico, el marcador de pluripotencia expresado es fosfatasa alcalina, NANOG, o ambos.

La descripción describe que las células iPS humanas cultivadas en el medio de baja osmolalidad descrito en la presente descripción muestran características morfológicas indicativas de un estado virgen. Una morfología ilustrativa se caracteriza por células que tienen colonias en forma de domos compactos en cultivo.

La descripción describe que las células iPS humanas cultivadas en el medio de baja osmolalidad descrito en la presente descripción pueden disociarse mecánicamente o enzimáticamente en una suspensión de célula única, someterse a pases, y/o subcultivarse. En un ejemplo, la disociación enzimática puede realizarse mediante el uso de tripsina. Cuando se cultivan en el medio de baja osmolalidad presente, las células iPS humanas pueden proporcionar mayor eficiencia de transformación debido a la disociación potenciada en una suspensión de célula única. Con otros tipos de medio (por ejemplo, medio mTeSR™ o medio 2i) típicamente usados para mantener células iPS humanas en cultivo, la disociación de células iPS humanas deberían realizarse mecánicamente o con enzimas tales como las colagenasas que son menos agresivas que la tripsina. En consecuencia, las células no se disocian de manera tan eficaz o completa. Por el contrario, con el medio de baja osmolalidad presente, puede usarse la tripsina para disociar las células, y la disociación potenciada da como resultado el aumento de la eficiencia de transformación. Además, a diferencia de otros tipos de medios típicamente usados para mantener células iPS humanas en cultivo (por ejemplo, medio mTeSR™ o medio 2i), la disociación enzimática de células iPS humanas cultivadas con el medio de baja osmolalidad presente (preferentemente un medio de baja osmolalidad que no comprende bFGF) puede realizarse en la ausencia de uno o más inhibidores que son generalmente necesarios para someter a pases tales células. Un inhibidor ilustrativo que puede omitirse es un inhibidor de proteína cinasa asociada a Rho (ROCK). Un inhibidor de ROCK es generalmente necesario cuando las células iPS humanas se someten a pases para inhibir la activación de rutas proapoptóticas.

La descripción describe que células iPS humanas subcultivadas cultivadas en el medio de baja osmolalidad descrito en la presente descripción pueden mantener un estado virgen o de apariencia virgen después de la disociación enzimática y el subcultivo. En algunos ejemplos, las células iPS humanas subcultivadas pueden continuar mostrando una morfología caracterizada por colonias en forma de domos compactos. Las células iPS humanas subcultivadas pueden continuar, además, expresando uno o más marcadores de pluripotencia como se describen en la presente descripción.

C. Métodos para generar y mantener una población de células madre pluripotentes inducidas humanas

La descripción describe métodos y composiciones para generar células iPS humanas en un cultivo *in vitro*. La descripción describe métodos y composiciones para mantener células iPS humanas en un cultivo *in vitro*.

El término "generar" incluye cultivar las células no pluripotentes transformadas para expresar uno o más factores de reprogramación como se describe en la presente descripción, bajo condiciones adecuadas para inducir un cambio en el fenotipo celular, expresión de genes, o ambos, de manera que las células muestran un estado virgen o de apariencia virgen, es decir, expresan una o más características de células iPS humanas vírgenes. Un estado virgen o de apariencia virgen puede expresarse en respuesta a condiciones de cultivos particulares, por ejemplo, el cultivo en un medio de baja osmolalidad como se describe en la presente descripción. En algunos ejemplos, la proporción de células que expresan un estado virgen o de apariencia virgen es de al menos aproximadamente 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, y hasta 100 % de las células en cultivo.

La descripción describe un método que enriquece un cultivo *in vitro* de una población de células iPS humanas vírgenes o de apariencia virgen. La descripción describe que las células iPS humanas vírgenes o de apariencia virgen pueden propagarse en cultivo preferentemente sobre células que no expresan un estado virgen o de apariencia virgen. La descripción describe que las células iPS humanas vírgenes o de apariencia virgen pueden seleccionarse de un cultivo, disociarse enzimáticamente, y subcultivarse para producir una población enriquecida de células iPS humanas vírgenes o de apariencia virgen.

La descripción describe que, las células no pluripotentes a transformar para que expresen un estado pluripotente, se cultivan *in vitro* en un medio proporcionado en la presente descripción que es adecuado para inducir la expresión de un estado virgen o de apariencia virgen durante un periodo de al menos 1, 2, 5, 7, 10, 14, 21, o 28 días, o cualquier periodo de tiempo suficiente para inducir la expresión de un estado virgen o de apariencia virgen en cultivo. Las células a transformar pueden cultivarse en el medio presente durante al menos 1, 2, 3, o 4 semanas. Algunas veces las células a transformar se cultivan durante 1-4 semanas. La expresión de un estado virgen o de apariencia virgen puede determinarse mediante la observación de características morfológicas o la expresión de marcadores de pluripotencia, característica de un estado virgen o de apariencia virgen, que se describe en cualquier otra parte de la presente descripción.

La descripción describe que las células no pluripotentes transformadas para expresar un estado pluripotente, se cultivan en el medio de baja osmolalidad presente hasta que estas expresan características de un estado virgen o de apariencia

virgen. Después las células pueden cultivarse en el medio presente para mantener un estado virgen o de apariencia virgen. En otra modalidad, las células no pluripotentes transformadas para expresar un estado pluripotente, primero se cultivan en un medio de alta osmolalidad antes del cultivo en el medio de baja osmolalidad presente. Tal medio de alta osmolalidad muestra una osmolalidad mayor que el medio de baja osmolalidad presente y puede comprender bFGF. Algún medio de alta osmolalidad comprende uno o más de albúmina de suero bovino, bFGF, factor de crecimiento transformante β (TGF β), cloruro de litio, ácido piperídico, y ácido gamma aminobutírico (GABA). Los ejemplos de un medio de alta osmolalidad incluyen el medio mTeSR™ (Stemcell Technologies).

La descripción describe que las células no pluripotentes transformadas para expresar un estado pluripotente, primero pueden cultivarse en medio de alta osmolalidad que comprende bFGF hasta que comienzan a expresar características de un estado virgen o de apariencia virgen, momento en el cual las células se cultivan en el medio de baja osmolalidad presente. En un ejemplo, las células pueden cultivarse en medio de alta osmolalidad que comprende bFGF durante un periodo de al menos 1, 2, 5, 10, 30, 60, o 90 días, un periodo de 1, 2, 4, 8, o 12 semanas, o un periodo entre 1 día a 3 meses. Un periodo de tiempo ilustrativo para el cultivo en un medio de alta osmolalidad que comprende bFGF es de 2 meses.

La descripción describe que las células no pluripotentes transformadas para expresar un estado pluripotente, primero pueden cultivarse en medio de alta osmolalidad que comprende bFGF hasta que comienzan a mostrar una morfología caracterizada por agrupaciones celulares tridimensionales, momento en el cual las células se cultivan en el medio de baja osmolalidad presente. La descripción describe que las células que muestran agrupaciones tridimensionales pueden seleccionarse, disociarse (por ejemplo, con tripsina), y transferirse a un nuevo cultivo en el medio de baja osmolalidad descrito en la presente descripción.

Los términos "mantiene," "mantener," y "mantenimiento" incluyen la conservación de al menos una o más de las características o fenotipos de las células iPS humanas descritas en la presente descripción. Tales características pueden incluir mantener la pluripotencia, morfología celular, perfiles de expresión de genes, y/u otras características funcionales de células vírgenes. Los términos "mantiene," "mantener," y "mantenimiento" pueden abarcar, además, la propagación de células y/o una disminución del número de células vírgenes que se cultivan. Los términos incluyen condiciones de cultivo que evitan que las células se conviertan a un estado imprimado o no pluripotente. Los términos incluyen, además, condiciones de cultivo que permiten que las células permanezcan pluripotentes y/o vírgenes, mientras que las células pueden o no pueden continuar dividiéndose o aumentar en número.

La descripción describe que las células iPS humanas se cultivan *in vitro* en un medio proporcionado en la presente descripción que es adecuado para mantener tales células en un estado virgen o de apariencia virgen. En un ejemplo particular, las células iPS humanas pueden cultivarse en un medio adecuado durante un periodo de 1, 2, 5, 7, 10, 14, 21, o 28 días, o durante un periodo de aproximadamente 2 semanas, aproximadamente 3 semanas, aproximadamente 4 semanas, o más, siempre que las células cultivadas se mantienen en un estado virgen o de apariencia virgen. Las células pueden cultivarse durante al menos 1, 2, 3 o 4 semanas. A veces las células se cultivan durante 1-4 semanas. Las células iPS humanas pueden mantenerse, por ejemplo, durante cualquier periodo de tiempo suficiente para la propagación de las células en cultivo, la modificación genética de las células, y/o el subcultivo de las células.

La descripción describe que las células iPS humanas o las células no pluripotentes transformadas para expresar un estado pluripotente, pueden cultivarse en un sustrato o capa adecuada de células alimentadoras para el cultivo *in vitro*. En un ejemplo particular, las células se cultivan en MATRIGEL™ (BD Biosciences). En otro ejemplo, las células se cultivan en células alimentadoras de fibroblastos de prepucio humano de recién nacido (NuFF). En otro ejemplo, las células se cultivan en GELTREX™ (Life Technologies).

La descripción describe que el tiempo de duplicación de células iPS humanas cultivadas en el medio de baja osmolalidad presente se reduce en comparación con las células iPS humanas imprimadas o las células no pluripotentes transformadas para expresar un estado pluripotente. En un ejemplo particular, el tiempo de duplicación de las células iPS humanas presentes es de entre aproximadamente 16-24 horas.

7. Identidad de secuencias

Los métodos y composiciones proporcionados en la presente descripción emplean una variedad de diferentes componentes del sistema de integración genómica objetivo (es decir, agentes nucleasas, sitios de reconocimiento, insertos de ácido nucleicos, polinucleótidos de interés, vectores de direccionamiento, marcadores de selección y otros componentes). Se reconoce en toda la descripción que algunos componentes del sistema de integración genómica objetivo pueden ser variantes y fragmentos activos. Tales componentes incluyen, por ejemplo, agentes nucleasas (es decir, agentes nucleasas modificados mediante ingeniería genética), sitios de reconocimiento para agentes nucleasas, polinucleótidos de interés, sitios objetivo y brazos de homología correspondientes del vector de direccionamiento. La actividad biológica de cada uno de estos componentes se describe en cualquier otra parte de la presente descripción.

Como se usa en la presente descripción, "identidad de secuencia" o "identidad" en el contexto de dos secuencias de polinucleótidos o polipéptidos hace referencia a los residuos en las dos secuencias que son iguales cuando se alinean para la máxima correspondencia en una ventana de comparación específica. Cuando el porcentaje de identidad de

- 5 secuencia se usa en referencia a proteínas se reconoce que las posiciones de los residuos que no son idénticos frecuentemente difieren en sustituciones aminoacídicas conservadoras, donde los residuos de aminoácidos se sustituyen por otros residuos de aminoácidos con propiedades químicas similares (por ejemplo, carga o hidrofobicidad) y por lo tanto no cambian las propiedades funcionales de la molécula. Cuando las secuencias difieren en las sustituciones conservadoras, el por ciento de identidad de secuencia puede ajustarse hacia arriba para corregir la naturaleza conservadora de la sustitución. Las secuencias que difieren en tales sustituciones conservadoras se dice que tienen "similitud de secuencias" o "similitud". Los medios para generar este ajuste son bien conocidos por un experto en la técnica. Típicamente esto implica calificar una sustitución conservadora como una no coincidencia parcial en lugar de total, lo que aumenta de este modo el porcentaje de identidad de secuencia. Por lo tanto, por ejemplo, cuando a un aminoácido idéntico se le asigna una calificación de 1 y a una sustitución no conservadora se le asigna una calificación de cero, a una sustitución conservadora se le asigna una calificación entre cero y 1. La calificación de sustituciones conservadoras se calcula, por ejemplo, como se implementó en el programa PC/GENE (Intelligenetics, Mountain View, California).
- 10
- 15 Como se usa en la presente descripción, "porcentaje de identidad de secuencia" significa el valor determinado mediante la comparación de dos secuencias alineadas de manera óptima en una ventana de comparación, en donde la porción de la secuencia de polinucleótidos en la ventana de comparación puede comprender adiciones o deleciones (es decir, interrupciones) en comparación con la secuencia de referencia (que no comprenden adiciones o deleciones) para la alineación óptima de las dos secuencias. El porcentaje se calcula mediante la determinación del número de posiciones en las que se produce la base de ácido nucleico o el residuo de aminoácido idéntico en ambas secuencias para obtener el número de posiciones coincidentes, la división del número de posiciones coincidentes por el número total de posiciones en la ventana de comparación, y la multiplicación del resultado por 100 para obtener el porcentaje de identidad de secuencia.
- 20
- 25 A menos que se indique de cualquier otra manera, los valores de identidad de secuencia/similitud proporcionados en la presente descripción se refieren al valor obtenido mediante el uso de GAP Versión 10 mediante el uso de los siguientes parámetros: % de identidad y % de similitud de una secuencia de nucleótidos mediante el uso de un Peso de interrupción de 50 y un Peso de longitud de 3, y la matriz de calificación nwsgapdna.cmp; % de identidad y % de similitud de una secuencia de aminoácidos mediante el uso de un Peso de interrupción de 8 y un Peso de longitud de 2, y la matriz de calificación BLOSUM62; o cualquier programa equivalente de estos. "Programa equivalente" significa cualquier programa de comparación de secuencias que, para cualquiera de dos secuencias en cuestión, genera una alineación que tiene coincidencias de nucleótidos o residuos de aminoácidos idénticas y un por ciento de identidad de secuencias idéntico cuando se compara con la alineación correspondiente generada por GAP Versión 10.
- 30
- 35 A menos que se defina de cualquier otra manera, todos los términos técnicos y científicos usados en la presente descripción tienen el mismo significado que entiende comúnmente un experto en la técnica a la que pertenece esta invención. Aunque cualquiera de los métodos y materiales similares o equivalentes a los descritos en la presente descripción también pueden usarse en la práctica o prueba de la invención descrita, a continuación, se describen métodos y materiales preferidos.
- 40
- 45 Debe señalarse que como se usa en la presente descripción y las reivindicaciones adjuntas, las formas en singular "un", "una", y "el/la" incluyen las referencias plurales a menos que el contexto lo indique claramente de cualquier otra manera. Todos los términos técnicos y científicos usados en la presente descripción tienen el mismo significado.
- 50
- 55 Las publicaciones analizadas en la presente descripción se proporcionan solamente para su descripción antes de la fecha de presentación de la presente solicitud. En la presente descripción nada debe interpretarse como una admisión de que la invención descrita no tiene derecho a anteceder a tal publicación en virtud de la invención anterior. Además, las fechas de publicación proporcionadas pueden ser diferentes de las fechas de publicación reales, que pueden necesitar confirmación independiente.
- 60
- 65 La invención descrita puede realizarse en otras formas específicas sin apartarse del espíritu o atributos esenciales de esta y, en consecuencia, debe hacerse referencia a las reivindicaciones adjuntas, en lugar de la descripción anterior, como indica el alcance de la invención.
- La descripción describe:
1. Un método para la modificación dirigida de un locus genómico de interés en una célula pluripotente de rata, que comprende (a) introducir en la célula pluripotente de rata un vector de direccionamiento grande (LTVEC) que comprende un inserto de ácido nucleico flanqueado por un brazo de homología de rata 5' y un brazo de homología de rata 3', en donde la suma total de los brazos de homología 5' y 3' es de al menos 10 kb pero menos de 150 kb; y (b) identificar una célula pluripotente genéticamente modificada de rata que comprende la modificación genética dirigida en el locus genómico de interés, en donde la modificación genética dirigida es capaz de transmitirse a través de la línea germinal.
 2. El método del ejemplo 1, en donde la modificación genética dirigida es bialélica.
 3. El método del ejemplo 1 o 2, en donde la célula pluripotente de rata es una célula madre embrionaria (ES) de rata.

4. El método del ejemplo 1, 2 o 3, en donde la célula pluripotente de rata se deriva de una cepa DA o una cepa ACI.
5. El método de cualquiera de los ejemplos 1-4, en donde la célula pluripotente de rata se caracteriza por la expresión de al menos un marcador de pluripotencia que comprende Dnmt3L, Eras, Err-beta, Fbxo15, Fgf4, Gdf3, Klf4, Lef1, receptor de LIF, Lin28, Nanog, Oct4, Sox15, Sox2, Utf1, o una combinación de estos.
6. El método de cualquiera de los ejemplos 1-4 en donde la célula pluripotente de rata se caracteriza por una o más de las siguientes características:
- (a) pérdida de expresión de uno o más marcadores de pluripotencia que comprenden c-Myc, Ecat1, y/o Rexo1; (b) pérdida de expresión de marcadores mesodérmicos que comprenden Brachyury y/o Bmpr2; (c) pérdida de expresión de uno o más marcadores endodérmicos que comprenden Gata6, Sox17 y/o Sox7; o (d) pérdida de expresión de uno o más marcadores neuronales que comprenden Nestin y/o Pax6.
7. El método de cualquiera de los ejemplos 1-6, en donde la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 30 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.
8. El método de cualquiera de los ejemplos 1-6, en donde la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 16 kb a aproximadamente 150 kb.
9. El método de cualquiera de los ejemplos 1-8, en donde la modificación genética dirigida comprende: (a) una sustitución de una secuencia de ácidos nucleicos de rata endógena con una secuencia de ácidos nucleicos homóloga u ortóloga; (b) una delección de una secuencia de ácidos nucleicos de rata endógena; (c) una delección de una secuencia de ácidos nucleicos de rata endógena, en donde la delección varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (d) una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb; (e) una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u ortóloga; (f) una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humana y de rata; (g) un alelo condicional flanqueado por secuencias objetivo de recombinasa específica de sitio; o (h) un gen reportero unido operativamente a un promotor activo en una célula de rata.
10. El método de cualquiera de los ejemplos 1-9, en donde el locus genómico de interés comprende (i) una primera secuencia de ácidos nucleicos que es complementaria al brazo de homología de rata 5'; y (ii) una segunda secuencia de ácidos nucleicos que es complementaria al brazo de homología de rata 3'.
11. El método del ejemplo 10, en donde las secuencias de ácidos nucleicos primera y segunda están separadas por al menos 5 kb pero menos de 3 Mb.
12. El método del ejemplo 10, en donde las secuencias de ácidos nucleicos primera y segunda están separadas por al menos 5 kb pero menos de 10 kb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero menos de aproximadamente 1 Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 2 Mb pero menos de aproximadamente 2,5 Mb, o al menos aproximadamente 2,5 Mb pero menos de aproximadamente 3 Mb.
13. El método de cualquiera de los ejemplos 1-12, en donde la etapa de introducción (a) comprende, además, introducir un segundo ácido nucleico que codifica un agente nucleasa que promueve la recombinación homóloga entre el constructo de transformación y el locus genómico de interés en la célula pluripotente de rata.

14. El método del ejemplo 13, en donde el agente nucleasa comprende (a) una proteína quimérica que comprende un dominio de unión al ADN basado en dedos de zinc fusionado a una endonucleasa FokI; o (b) una proteína quimérica que comprende una Nucleasa Efectora de tipo Activador de la Transcripción (TALEN) fusionada a una endonucleasa FokI.
- 5 15. El método de cualquiera de los ejemplos 1-12, en donde la etapa de introducción (a) comprende, además, introducir en la célula pluripotente de rata: (i) un primer constructo de expresión que comprende un primer promotor unido operativamente a una primera secuencia de ácidos nucleicos que codifica una proteína asociada a Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (Cas), (ii) un segundo constructo de expresión que comprende un segundo promotor unido operativamente a la secuencia objetivo genómica unida a un ARN guía (ARNg), en donde la secuencia objetivo genómica está flanqueada inmediatamente en el extremo 3' por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).
- 10 16. El método del ejemplo 15, en donde el locus genómico de interés comprende la secuencia nucleotídica de sec. con núm. de ident.: 1.
- 15 17. El método del ejemplo 15 o 16, en donde el ARNg comprende una tercera secuencia de ácidos nucleicos que codifica un ARN de Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (ARNcr) y un ARN de CRISPR de transactivación (ARNtracr).
- 20 18. El método del ejemplo 15, 16 o 17, en donde la proteína Cas es Cas9.
19. El método del ejemplo 15, 16, 17, o 18, en donde el ARNg comprende: (a) el ARN quimérico de la secuencia de ácidos nucleicos de sec. con núm. de ident.: 2; o (b) el ARN quimérico de la secuencia de ácidos nucleicos de sec. con núm. de ident.: 3.
- 25 20. El método del ejemplo 17, en donde el ARNcr comprende la sec. con núm. de ident.: 4; sec. con núm. de ident.: 5; o sec. con núm. de ident.: 6.
- 30 21. El método del ejemplo 17, en donde el ARNtracr comprende la sec. con núm. de ident.: 7 o la sec. con núm. de ident.: 8.
22. Un locus genómico de rata modificado que comprende: (i) una inserción de una secuencia de ácidos nucleicos humana homóloga u ortóloga; (ii) una sustitución de una secuencia de ácidos nucleicos de rata endógena con la secuencia de ácidos nucleicos humana homóloga u ortóloga; o (iii) una combinación de estos, en donde el locus genómico de rata modificado es capaz de transmitirse a través de la línea germinal.
- 35 23. El locus genómico de rata modificado del ejemplo 22, en donde el tamaño de la inserción o la sustitución es de aproximadamente 5 kb a aproximadamente 400 kb.
- 40 24. El locus genómico de rata del ejemplo 22, en donde el tamaño de la inserción o la sustitución es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.
- 45 25. Un método para generar una rata humanizada, que comprende: (a) transformar un locus genómico de interés en una célula pluripotente de rata con un constructo de transformación que comprende un ácido nucleico humano para formar una célula pluripotente genéticamente modificada de rata; (b) introducir la célula pluripotente genéticamente modificada de rata en un embrión de rata huésped; y (c) gestar un embrión de rata huésped en una madre sustituta; en donde la madre sustituta produce la progenie de ratas que comprende un locus genómico modificado que comprende: (i) una inserción de una secuencia de ácidos nucleicos humana; (ii) una sustitución de la secuencia de ácidos nucleicos de rata en el locus genómico de interés con una secuencia de ácidos nucleicos humana homóloga u ortóloga; (iii) una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humana y de rata; o (iv) una combinación de estos, en donde el locus genómico modificado es capaz de transmitirse a través de la línea germinal.
- 50 26. El método del ejemplo 25, en donde el constructo de transformación es un vector de direccionamiento grande (LTVEC), y la suma total de los brazos de homología 5' y 3' del LTVEC es de al menos 10 kb pero menos de 150 kb.
- 55 27. El método del ejemplo 26, en donde la suma total de los brazos de homología 5' y 3' del constructo de transformación es de aproximadamente 10 kb a aproximadamente 30 kb, de aproximadamente 20 kb a 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, o de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.
- 60 65

28. El método del ejemplo 25, 26 o 27, en donde la secuencia de ácidos nucleicos humana es de al menos 5 kb pero menos de 400 kb.
- 5 29. El método del ejemplo 25, 26, o 27, en donde la secuencia de ácidos nucleicos humana es de al menos 5 kb pero menos de 10 kb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, al menos 150 kb pero menos de 200 kb, al menos 200 kb pero menos de 250 kb, al menos 250 kb pero menos de 300 kb, al menos 300 kb pero menos de 350 kb, o al menos 350 kb pero menos de 400 kb.
- 10 30. El método de cualquiera de los ejemplos 25-29, en donde la célula pluripotente de rata es una célula madre embrionaria (ES) de rata.
31. El método de cualquiera de los ejemplos 25-30, en donde la célula pluripotente de rata se deriva de una cepa DA o una cepa ACI.
- 15 32. El método de cualquiera de los ejemplos 25-31, en donde la célula pluripotente de rata se caracteriza por la expresión de al menos un marcador de pluripotencia que comprende Dnmt3L, Eras, Err-beta, Fbxo15, Fgf4, Gdf3, Klf4, Lef1, receptor de LIF, Lin28, Nanog, Oct4, Sox15, Sox2, Utf1, o una combinación de estos.
- 20 33. El método de cualquiera de los ejemplos 25-31, en donde la célula pluripotente de rata se caracteriza por una o más de las siguientes características: (a) pérdida de expresión de uno o más marcadores de pluripotencia que comprenden c-Myc, Ecat1, y/o Rexo1; (b) pérdida de expresión de uno o más marcadores mesodérmicos que comprenden Brachyury y/o Bmpr2; (c) pérdida de expresión de uno o más marcadores endodérmicos que comprenden Gata6, Sox17, y/o Sox7; o (d) pérdida de expresión de uno o más marcadores neuronales que comprenden Nestin y/o Pax6.
- 25 34. Una rata modificada que comprende un locus genómico humanizado, en donde el locus genómico humanizado comprende: (i) una inserción de una secuencia de ácidos nucleicos humana homóloga u ortóloga; (ii) una sustitución de una secuencia de ácidos nucleicos de rata en un locus endógeno genómico con una secuencia de ácidos nucleicos humana homóloga u ortóloga; (iii) una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humano y de rata o (iv) una combinación de estos, en donde el locus genómico humanizado es capaz de transmitirse a través de la línea germinal.
- 30 35. Una rata o célula de rata que comprende una modificación genética dirigida en su locus genómico, en donde el locus genómico es un locus de la cadena gamma del receptor de interleucina-2, un locus de *ApoE*, un locus de *Rag1*, un locus de *Rag2*, o un locus de *Rag2/Rag1*, en donde la modificación genética dirigida comprende: (a) una delección de una secuencia de ácidos nucleicos de rata endógena en el locus genómico; (b) una inserción de un ácido nucleico homólogo, un ácido nucleico ortólogo, o un ácido nucleico quimérico que comprende una secuencia de ácidos nucleicos humana y de rata, o (c) una combinación de estos, en donde la modificación genética dirigida puede transmitirse a través de la línea germinal de la rata o propagarse a partir de la célula de rata.
- 35 40 36. La rata o célula de rata del ejemplo 35, en donde (a) la delección del ácido nucleico de rata endógeno en el locus genómico es de al menos aproximadamente 10 kb; o (b) la delección del ácido nucleico de rata endógeno en el locus genómico es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (c) la inserción de la secuencia de ácidos nucleicos exógena en el locus genómico es de al menos aproximadamente 5 kb; o (d) la inserción de la secuencia de ácidos nucleicos exógena en el locus genómico es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.
- 45 50 55 60 65 37. La rata o célula de rata del ejemplo 35 o 36, en donde (a) la modificación genética dirigida en el locus de la cadena gamma de receptor de interleucina-2 da como resultado una disminución o ausencia de actividad de la proteína gamma del receptor de interleucina-2; (b) la modificación genética dirigida en el locus de *ApoE* da como resultado una disminución o ausencia de actividad de la proteína ApoE; (c) la modificación genética dirigida en el locus de *Rag1* da como resultado una disminución o ausencia de actividad de la proteína Rag1; (d) la modificación genética dirigida en el locus de *Rag2* da como resultado una disminución o ausencia de actividad de la proteína Rag2; o (e) la modificación genética dirigida en el

locus de *Rag2/Rag1* da como resultado una disminución o ausencia de actividad de la proteína Rag2 y la actividad de Rag1.

- 5 38. La rata o célula de rata del ejemplo 35, 36, o 37, en donde la modificación genética dirigida del locus de la cadena gamma de receptor de interleucina-2 comprende: (a) una delección de la región codificante entera de la cadena gamma del receptor de interleucina-2 de rata o una porción de esta; (b) una sustitución de la región codificante entera de la cadena gamma del receptor de interleucina-2 de rata o una porción de esta con una región codificante de la cadena gamma del receptor de interleucina-2 humano o una porción de esta; (c) una sustitución de un ectodominio de la región codificante de la cadena gamma del receptor de interleucina-2 de rata con el ectodominio de una cadena gamma del receptor de interleucina-2 humano; o (d) al menos una delección de 3 kb del locus de la cadena gamma de receptor de interleucina-2.
- 10
- 15 39. La rata o célula de rata de cualquiera de los ejemplos 35-37, en donde la modificación genética dirigida del locus de *ApoE* comprende: (a) una delección de la región codificante entera de *ApoE* o una porción de esta; o (b) al menos una delección de 1,8 kb del locus de *ApoE* que comprende la región codificante de *ApoE*.
- 20 40. La rata o célula de rata de cualquiera de los ejemplos 35-37, en donde la modificación genética dirigida del locus de *Rag2* comprende: (a) una delección de la región codificante entera de *Rag2* o una porción de esta; (b) al menos una delección de 5,7 kb del locus de *Rag2* que comprende la región codificante de *Rag2*.
- 25 41. La rata o célula de rata de cualquiera de los ejemplos 35-37, en donde la modificación genética dirigida del locus de *Rag2/Rag1* comprende: (a) una delección de la región codificante entera de *Rag2* o una porción de esta y una delección de la región codificante entera de *Rag1* o porción de esta; o (b) una delección de al menos 16 kb del locus de *Rag2/Rag1* que comprende la región codificante de *Rag2*.
- 30 42. La rata o célula de rata de cualquiera de los ejemplos 35-41, en donde la modificación genética dirigida comprende una inserción de un casete de expresión que comprende un marcador de selección en el locus de la cadena gamma de receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2*, o el locus de *Rag2/Rag1*.
- 35 43. La rata o célula de rata de cualquiera de los ejemplos 42, en donde el casete de expresión comprende un gen lacZ unido operativamente al promotor endógeno en el locus genómico y un promotor de ubiquitina humano unido operativamente a un marcador de selección.
- 40 44. La rata o célula de rata de cualquiera de los ejemplos 35-43, en donde la modificación genética dirigida en el locus de la cadena gamma de receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2* o el locus de *Rag2/Rag1* comprende la inserción de un casete de selección de autodelección.
- 45 45. La rata o célula de rata del ejemplo 44, en donde el casete de selección de autodelección comprende un gen de marcador de selección unido operativamente a un promotor activo en la célula de rata y un gen de recombinasa unido operativamente a un promotor específico de células germinales masculinas, en donde el casete de autodelección está flanqueada por sitios de reconocimiento de recombinación reconocidos por la recombinasa.
- 50 46. La rata o célula de rata del ejemplo 45, en donde (a) el promotor específico de células germinales masculinas es un promotor de Protamina-1; o (b) el gen de recombinasa codifica Cre, y los sitios de reconocimiento de recombinación son sitios loxP.
- 55 47. La rata o célula de rata de cualquiera de los ejemplos 35-46, en donde la inserción de la secuencia de ácidos nucleicos exógena en el locus genómico comprende un ácido nucleico reportero unido operativamente a un promotor de la cadena gamma del receptor de interleucina-2 endógeno, un promotor de *ApoE* endógeno, un promotor de *Rag1* endógeno, o un promotor de *Rag2* endógeno.
- 60 48. La rata o célula de rata del ejemplo 47, en donde el ácido nucleico reportero codifica un reportero que comprende β -galactosidasa, mPlum, mCherry, tdTomato, mStrawberry, J-Red, DsRed, mOrange, mKO, mCitrine, Venus, YPet, proteína fluorescente amarilla potenciada (EYFP), Emerald, proteína fluorescente verde potenciada (EGFP), CyPet, proteína fluorescente ciano (CFP), Cerulean, T-Sapphire, luciferasa, fosfatasa alcalina, o una combinación de estos.
- 65 49. La célula de rata de cualquiera de los ejemplos 35-48, en donde la célula de rata es una célula pluripotente de rata o una célula madre embrionaria (ES) de rata.
50. La célula de rata del ejemplo 49, en donde la célula pluripotente de rata o la célula madre embrionaria (ES) de rata (a) se deriva de una cepa DA o una cepa ACI; (b) se caracteriza por la expresión de al menos un marcador de pluripotencia que comprenden Dnmt3L, Eras, Err-beta, Fbxo15, Fgf4, Gdf3, Klf4, Lef1, receptor de LIF, Lin28, Nanog, Oct4, Sox15, Sox2, Utf1, o una combinación de estos; o (c) se caracteriza por una o más de las siguientes características: (i) pérdida de expresión de uno o más marcadores de pluripotencia que comprenden c-Myc, Ecat1, y/o Rexo1; (ii) pérdida de la expresión de marcadores mesodérmicos que comprenden Brachyury y/o Bmpr2; (iii) pérdida de la expresión de uno o más marcadores endodérmicos que comprenden Gata6, Sox17 y/o Sox7; o (iv) pérdida de la expresión de uno o más marcadores neuronales que comprenden Nestin y/o Pax6.

51. Un método para modificar un locus genómico objetivo en un locus de la cadena gamma de receptor de interleucina-2, un locus de *ApoE*, un locus de *Rag1*, un locus de *Rag2* o un locus de *Rag2/Rag1* en una célula pluripotente de rata, el método que comprende: (a) introducir en la célula pluripotente de rata un vector de direccionamiento que comprende un inserto de ácido nucleico flanqueado por los brazos de homología de rata 5' y 3' homólogos al locus genómico objetivo, (b) identificar una célula pluripotente genéticamente modificada de rata que comprende una modificación genética dirigida en el locus genómico objetivo, en donde la modificación genética dirigida es capaz de transmitirse a través de la línea germinal de una rata propagada a partir de la célula pluripotente de rata.
52. El método del ejemplo 51, en donde el vector de direccionamiento es un vector de direccionamiento grande (LTVEC) en donde la suma total de los brazos de homología de rata 5' y 3' es de al menos aproximadamente 10 kb pero menos de aproximadamente 150 kb.
53. El método del ejemplo 51 o 52, en donde introducir el vector de direccionamiento en la célula pluripotente de rata conduce a: (i) una deleción de una secuencia de ácidos nucleicos de rata endógena en el locus genómico objetivo; (ii) una inserción de una secuencia de ácidos nucleicos exógena en el locus genómico objetivo; o (iii) una combinación de estos.
54. El método del ejemplo 53, en donde (a) la deleción del ácido nucleico de rata endógeno en el locus genómico es de al menos aproximadamente 10 kb; o (b) la deleción del ácido nucleico de rata endógeno en el locus genómico es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (c) la inserción de la secuencia de ácidos nucleicos exógena en el locus genómico es de al menos aproximadamente 5 kb; o (d) la inserción de la secuencia de ácidos nucleicos exógena en el locus genómico es de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.
55. El método de cualquiera de los ejemplos 51-54, en donde (a) la modificación genética dirigida en el locus de la cadena gamma de receptor de interleucina-2 da como resultado una disminución o ausencia de actividad de la proteína gamma del receptor de interleucina-2; (b) la modificación genética dirigida en el locus de *ApoE* da como resultado una disminución o ausencia de actividad de la proteína ApoE; (c) la modificación genética dirigida en el locus de *Rag1* da como resultado una disminución o ausencia de actividad de la proteína Rag1; (d) la modificación genética dirigida en el locus de *Rag2* da como resultado una disminución o ausencia de actividad de la proteína Rag2; o (e) la modificación genética dirigida en el locus de *Rag2/Rag1* da como resultado una disminución o ausencia de actividad de la proteína Rag2 y de la actividad de la proteína Rag1.
56. El método de cualquiera de los ejemplos 51-54, en donde la modificación genética dirigida del locus de la cadena gamma de receptor de interleucina-2 comprende (a) una deleción de la región codificante entera de la cadena gamma del receptor de interleucina-2 de rata o una porción de esta; (b) una sustitución de la región codificante entera de la cadena gamma del receptor de interleucina-2 de rata o una porción de esta con una región codificante de la cadena gamma del receptor de interleucina-2 humano o una porción de esta; (c) una sustitución de un ectodominio de la región codificante de la cadena gamma del receptor de interleucina-2 de rata con el ectodominio de una cadena gamma del receptor de interleucina-2 humano; o (d) al menos una deleción de 3 kb del locus de la cadena gamma de receptor de interleucina-2 que comprende la región codificante de la cadena gamma del receptor de interleucina-2.
57. El método de cualquiera de los ejemplos 51-55, en donde la modificación genética dirigida del locus de *ApoE* comprende: (a) una deleción de la región codificante entera de *ApoE* o una porción de esta; o (b) al menos una deleción de 1,8 kb del locus de *ApoE* que comprende la región codificante de *ApoE*.
58. El método de cualquiera de los ejemplos 51-55, en donde la modificación genética dirigida del locus de *Rag2* comprende: (a) una deleción de la región codificante entera de *Rag2* o una porción de esta; o (b) al menos una deleción de 5,7 kb del locus de *Rag2* que comprende la región codificante de *Rag2*.
59. El método de cualquiera de los ejemplos 51-55, en donde la modificación genética dirigida del locus de *Rag1/Rag2* comprende: (a) una deleción de la región codificante entera de *Rag2* o una porción de esta y una deleción de la región codificante entera de *Rag1* o porción de esta; o (b) una deleción de al menos 16 kb del locus de *Rag2/Rag1* que comprende las regiones codificantes de *Rag2* y *Rag1*.

60. El método de cualquiera de los ejemplos 51-59, en donde el inserto de ácido nucleico comprende un casete de expresión que comprende un polinucleótido que codifica un marcador de selección.
- 5 61. El método del ejemplo 60, en donde el casete de expresión comprende un gen *lacZ* unido operativamente a un promotor endógeno en el locus genómico y un promotor de ubiquitina humano unido operativamente a un gen de marcador de selección.
62. El método de cualquiera de los ejemplos 51-60, en donde el inserto de ácido nucleico comprende un casete de selección de autodelección.
- 10 63. El método del ejemplo 62, en donde el casete de selección de autodelección comprende un marcador de selección unido operativamente a un promotor activo en la célula pluripotente de rata y un polinucleótido que codifica una recombinasa unido operativamente a un promotor específico de células germinales masculinas, en donde el casete de autodelección está flanqueado por sitios de reconocimiento de recombinación reconocidos por la recombinasa.
- 15 64. El método del ejemplo 63, en donde (a) el promotor específico de células germinales masculinas es un promotor de Protamina-1; o (b) el gen de recombinasa codifica Cre y los sitios de reconocimiento de recombinación son sitios loxP.
- 20 65. El método del ejemplo 53, en donde la inserción de la secuencia de ácidos nucleicos exógena en el locus genómico comprende una secuencia de ácidos nucleicos reportera unida operativamente a un promotor de la cadena gamma del receptor de interleucina-2 endógeno, un promotor de *ApoE* endógeno, un promotor de *Rag1* endógeno, o un promotor de *Rag2* endógeno.
- 25 66. El método del ejemplo 65, en donde la secuencia de ácidos nucleicos reportera codifica un reportero que comprende β -galactosidasa, mPlum, mCherry, tdTomato, mStrawberry, J-Red, DsRed, mOrange, mKO, mCitrine, Venus, YPet, proteína fluorescente amarilla potenciada (EYFP), Emerald, proteína fluorescente verde potenciada (EGFP), CyPet, proteína fluorescente ciano (CFP), Cerulean, T-Sapphire, luciferasa, fosfatasa alcalina, o una combinación de estos.
- 30 67. El método de cualquiera de los ejemplos 51-66, en donde la célula pluripotente de rata es una célula madre embrionaria (ES) de rata.
68. El método de cualquiera de los ejemplos 51-67, en donde la célula pluripotente de rata (a) se deriva de una cepa DA o una cepa ACI; o (b) se caracteriza por la expresión de un marcador de pluripotencia que comprende Oct-4, Sox-2, fosfatasa alcalina, o una combinación de estos; o (c) se caracteriza por una o más de las siguientes características: (i) pérdida de expresión de uno o más marcadores de pluripotencia que comprenden c-Myc, Ecat1, y/o Rexo1; (ii) pérdida de la expresión de marcadores mesodérmicos que comprenden Brachyury y/o Bmpr2; (iii) pérdida de la expresión de uno o más marcadores endodérmicos que comprenden Gata6, Sox17 y/o Sox7; o (iv) pérdida de la expresión de uno o más marcadores neuronales que comprenden Nestin y/o Pax6.
- 35 69. El método de cualquiera de los ejemplos 51-68, que comprende, además, identificar la modificación genética dirigida en el locus genómico objetivo, en donde la etapa de identificación emplea un ensayo cuantitativo para evaluar una modificación de alelo (MOA) en el locus genómico objetivo.
- 40 70. El método de cualquiera de los ejemplos 51-69, en donde la etapa de introducción (a) comprende, además, introducir un segundo ácido nucleico que codifica un agente nucleasa que promueve la recombinación homóloga entre el vector de direccionamiento y el locus genómico objetivo en la célula pluripotente de rata.
- 45 71. El método del ejemplo 70, en donde el agente nucleasa comprende una proteína quimérica que comprende un dominio de unión al ADN basado en dedos de zinc fusionado a una endonucleasa FokI.
- 50 72. El método del ejemplo 71, en donde el método da como resultado la modificación bialélica del locus genómico objetivo.
73. El método de cualquiera de los ejemplos 51-70, en donde la etapa de introducción (a) comprende, además, introducir en la célula pluripotente de rata: (i) un primer constructo de expresión que comprende un primer promotor unido operativamente a una primera secuencia de ácidos nucleicos que codifica una proteína asociada a Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (Cas), (ii) un segundo constructo de expresión que comprende un segundo promotor unido operativamente a la secuencia objetivo genómica unida a un ARN guía (ARNg), en donde la secuencia objetivo genómica está flanqueada inmediatamente en el extremo 3' por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).
- 55 60 74. El método del ejemplo 73, en donde el locus genómico de interés comprende la secuencia nucleotídica de la sec. con núm. de ident.: 1.
- 65 75. El método del ejemplo 73 o 74, en donde el ARNg comprende una tercera secuencia de ácidos nucleicos que codifica un ARN de Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR) (ARNcr) y un ARN de CRISPR de transactivación (ARNtracr).

76. El método del ejemplo 73, en donde la proteína Cas es Cas9.
77. El método del ejemplo 73, 74, o 75, en donde el ARNg comprende: (a) el ARN quimérico de la secuencia de ácidos nucleicos de sec. con núm. de ident.: 2; o (b) el ARN quimérico de la secuencia de ácidos nucleicos de sec. con núm. de ident.: 3.
78. El método del ejemplo 75, en donde el ARNcr comprende la sec. con núm. de ident.: 4; sec. con núm. de ident.: 5; o sec. con núm. de ident.: 6.
79. El método del ejemplo 75, en donde el ARNtracr comprende la sec. con núm. de ident.: 7 o la sec. con núm. de ident.: 8.
80. La rata o célula de rata de cualquiera de los ejemplos 35-50, en donde la rata o célula de rata comprende las modificaciones genéticas dirigidas en el locus de la cadena gamma de receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2*, y/o el locus de *Rag2/Rag1*.
81. La rata o célula de rata del ejemplo 80, en donde la rata o célula de rata comprende las modificaciones genéticas dirigidas en el locus de la cadena gamma de receptor de interleucina-2 y el locus de *Rag2/Rag1*.
- La descripción describe:
1. Un método para modificar un locus genómico de interés en una célula eucariota, que comprende: (a) introducir en la célula eucariota: (i) un vector de direccionamiento grande (LTVEC) que comprende un primer ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3', en donde el LTVEC es de al menos 10 kb; (ii) un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica una proteína Cas, (iii) un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica un ARN guía (ARNg) que comprende una secuencia de nucleótidos que se hibrida a una secuencia objetivo y un ARN de CRISPR de transactivación (ARNtracr), en donde los promotores primero y segundo son activos en la célula eucariota; y (b) identificar una célula eucariota modificada que comprende una modificación genética dirigida en el locus genómico de interés.
 2. El método del ejemplo 1, en donde la modificación genética dirigida es una modificación genética bialélica.
 3. El método del ejemplo 1, en donde el LTVEC es de al menos 15 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, o al menos 90 kb.
 4. El método del ejemplo 1, en donde el LTVEC es de al menos 100 kb, al menos 150 kb, o al menos 200 kb.
 5. El método del ejemplo 1, en donde la célula eucariota es una célula de mamífero.
 6. El método del ejemplo 5, en donde la célula de mamífero es un fibroblasto.
 7. El método del ejemplo 1, en donde la célula eucariota es una célula pluripotente.
 8. El método del ejemplo 7, en donde la célula pluripotente es una célula pluripotente humana.
 9. El método del ejemplo 8, en donde la célula pluripotente humana es una célula madre embrionaria humana (ES) o una célula madre adulta humana.
 10. El método del ejemplo 8, en donde la célula pluripotente humana es una célula progenitora humana con desarrollo limitado.
 11. El método del ejemplo 8, en donde la célula pluripotente humana es una célula madre pluripotente inducida (iPS) humana.
 12. El método de la modalidad 1, en donde la proteína Cas es Cas9.
 13. El método del ejemplo 1, en donde la secuencia objetivo está flanqueada inmediatamente en el extremo 3' por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).
 14. El método del ejemplo 1, en donde la suma total de los brazos de homología 5' y 3' es de aproximadamente 10 kb a aproximadamente 150 kb.
 15. El método del ejemplo 1, en donde la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a

aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.

5 16. El método del ejemplo 1, en donde la modificación genética dirigida comprende:(a) una sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos homóloga u ortóloga; (b) una delección de una
 10 secuencia de ácidos nucleicos endógena; (c) una delección de una secuencia de ácidos nucleicos endógena, en donde la delección varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de
 15 aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de
 20 aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (d)
 25 inserción de una secuencia de ácidos nucleicos exógena; (e) inserción de una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de
 aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de
 aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de
 aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb; (f) inserción de una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u
 ortóloga; (g) inserción de una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humana y una no humana; (h) inserción de un alelo condicional flanqueado por secuencias objetivo de recombinasa
 específica de sitio; (i) inserción de un gen marcador de selección o un gen reportero unido operativamente a un tercer
 promotor activo en la célula pluripotente; o (j) una combinación de estos.

30 17. El método del ejemplo 1, en donde el locus genómico de interés comprende (i) una secuencia objetivo 5' que es homóloga al brazo de homología 5'; y (ii) una secuencia objetivo 3' que es homóloga al brazo de homología 3'.

18. El método del ejemplo 17, en donde la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 3 Mb.

35 19. El método del ejemplo 17, en donde la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 10 kb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al
 40 menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al
 menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero
 menos de aproximadamente 1 Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos
 aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 2 Mb pero menos de
 aproximadamente 2,5 Mb, o al menos aproximadamente 2,5 Mb pero menos de aproximadamente 3 Mb.

45 20. El método del ejemplo 1, en donde el locus genómico de interés comprende el locus de la cadena gamma de receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2*, o ambos loci de *Rag1* y de *Rag2*.

21. El método del ejemplo 1, en donde los constructos de expresión primero y segundo están en una única molécula de ácido nucleico.

50 22. Un método para modificar un genoma, que comprende exponer el genoma a una proteína Cas y un ARN de CRISPR en la presencia de un vector de direccionamiento grande (LTVEC) que comprende una secuencia de ácidos nucleicos de al menos 10 kb, en donde después de la exposición a la proteína Cas, el ARN de CRISPR, y el LTVEC, el genoma se modifica para que contenga al menos una secuencia de ácidos nucleicos de 10 kb.

55 23. El método del ejemplo 22, en donde el LTVEC comprende una secuencia de ácidos nucleicos de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, o al menos 90 kb.

24. El método de la modalidad 22, en donde el LTVEC comprende una secuencia de ácidos nucleicos de al menos 100 kb, al menos 150 kb, o al menos 200 kb.

60 25. Un método para modificar un genoma, que comprende poner en contacto el genoma con una proteína Cas, un ARN de CRISPR que se hibrida a una secuencia objetivo, y un ARNtracr en la presencia de un vector de direccionamiento grande (LTVEC), en donde el LTVEC es de al menos 10 kb y comprende un primer ácido nucleico flanqueado por un
 65 brazo de homología 5' y un brazo de homología 3', en donde después de ponerse en contacto con la proteína Cas, el ARN de CRISPR, y el ARNtracr en la presencia del LTVEC, el genoma se modifica en un locus genómico de interés para que contenga el primer ácido nucleico.

26. El método del ejemplo 25, en donde el genoma está en una célula eucariota, y la proteína Cas, el ARN de CRISPR, el ARNtracr, y el LTVEC se introducen en la célula eucariota
- 5 27. El método del ejemplo 26, que comprende, además, identificar una célula eucariota modificada que comprende una modificación genética dirigida en el locus genómico de interés.
28. El método del ejemplo 26 o 27, en donde el ARN de CRISPR y el ARNtracr se introducen juntos en la forma de un único ARN guía (ARNg).
- 10 29. El método del ejemplo 26 o 27, en donde el ARN de CRISPR y el ARNtracr se introducen por separado.
30. El método de cualquiera de los ejemplos 26-29, en donde: (a) la proteína Cas se introduce en la célula eucariota en la forma de una proteína, un ARN mensajero (ARNm) que codifica la proteína Cas, o un ADN que codifica la proteína Cas; (b) el ARN de CRISPR se introduce en la célula eucariota en la forma de un ARN o un ADN que codifica el ARN de CRISPR; y (c) el ARNtracr se introduce en la célula eucariota en la forma de un ARN o un ADN que codifica el ARNtracr.
- 15 31. El método del ejemplo 30, en donde la proteína Cas, el ARN de CRISPR, y el ARNtracr se introducen en la célula eucariota como un complejo de proteína-ARN.
- 20 32. El método del ejemplo 30, en donde: (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; (b) el ADN que codifica el ARN de CRISPR está en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica el ARN de CRISPR; y (c) el ADN que codifica el ARNtracr está en la forma de un tercer constructo de expresión que comprende un tercer promotor unido operativamente a un cuarto ácido nucleico que codifica el ARNtracr, en donde los promotores primero, segundo y
- 25 33. El método del ejemplo 32, en donde los constructos de expresión primero, segundo y/o tercero están en una única molécula de ácido nucleico.
- 30 34. El método del ejemplo 30, en donde: (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; y (b) el ADN que codifica el ARN de CRISPR y el ADN que codifica el ARNtracr están en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica un ARN guía que comprende el ARN de CRISPR y el ARNtracr; en donde los promotores primero y segundo son activos en la célula eucariota.
- 35 35. El método del ejemplo 34, en donde los constructos de expresión primero y segundo están en una única molécula de ácido nucleico.
- 40 36. El método de cualquiera de las modalidades 27-35, en donde la modificación genética dirigida comprende la delección simultánea de una secuencia de ácidos nucleicos endógena en el locus genómico de interés y la inserción del primer ácido nucleico en el locus genómico de interés.
- 45 37. El método de cualquiera de los ejemplos 27-36, en donde la modificación genética dirigida es una modificación genética bialélica.
38. El método del ejemplo 37, en donde la modificación genética bialélica comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico en el locus genómico de interés en dos cromosomas homólogos.
- 50 39. El método de cualquiera de los ejemplos 27-36, en donde la célula eucariota modificada es hemicigota en el locus genómico de interés.
- 55 40. El método del ejemplo 39, en donde la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico.
41. El método del ejemplo 39, en donde la modificación genética dirigida comprende: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en dos cromosomas homólogos; y (2) inserción del primer ácido nucleico en el locus genómico de interés en un primer cromosoma y alteración del locus genómico de interés en un segundo cromosoma.
- 60 42. El método de cualquiera de los ejemplos 25-41, en donde el LTVEC es de al menos 15 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, o al menos 90 kb.
- 65

ES 2 700 596 T3

43. El método de cualquiera de los ejemplos 25-42, en donde el LTVEC es de al menos 100 kb, al menos 150 kb, o al menos 200 kb.
- 5 44. El método de cualquiera de los ejemplos 25-43, en donde el primer ácido nucleico es de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, o al menos 300 kb.
45. El método de cualquiera de los ejemplos 26-44, en donde la célula eucariota es una célula de mamífero.
- 10 46. El método del ejemplo 45, en donde la célula de mamífero es un fibroblasto.
47. El método de cualquiera de los ejemplos 26-43, en donde la célula eucariota es una célula pluripotente.
48. El método del ejemplo 47, en donde la célula pluripotente es una célula pluripotente no humana.
- 15 49. El método del ejemplo 48, en donde la célula pluripotente humana no humana es una célula pluripotente de roedor.
50. El método del ejemplo 49, en donde la célula pluripotente de roedor es una célula madre embrionaria (ES) de ratón o humana.
- 20 51. El método del ejemplo 47, en donde la célula pluripotente es una célula pluripotente humana.
52. El método del ejemplo 51, en donde la célula pluripotente humana es una célula madre embrionaria humana (ES) o una célula madre adulta humana.
- 25 53. El método del ejemplo 51, en donde la célula pluripotente humana es una célula progenitora humana con desarrollo limitado.
54. El método del ejemplo 51, en donde la célula pluripotente humana es una célula madre pluripotente inducida (iPS) humana.
- 30 55. El método de cualquiera de los ejemplos 25-54, en donde la proteína Cas es Cas9.
56. El método de cualquiera de los ejemplos 25-55, en donde la secuencia objetivo está flanqueada inmediatamente por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).
- 35 57. El método de cualquiera de los ejemplos 25-56, en donde la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 150 kb.
- 40 58. El método de cualquiera de los ejemplos 25-57, en donde la suma total de los brazos de homología 5' y 3' del LTVEC es de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.
- 45 59. El método de cualquiera de los ejemplos 27-58, en donde la modificación genética dirigida comprende: (a) una sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos homóloga u ortóloga; (b) una delección de una secuencia de ácidos nucleicos endógena; (c) una delección de una secuencia de ácidos nucleicos endógena, en donde la delección varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb; (d) inserción de una secuencia de ácidos nucleicos exógena; (e) inserción de una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb; (f) inserción de una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u ortóloga; (g) inserción de una secuencia de ácidos nucleicos química que comprende una secuencia de ácidos nucleicos humana y una no humana; (h) inserción de un alelo condicional flanqueado por secuencias
- 50 55 60 65

objetivo de recombinasa específica de sitio; (i) inserción de un gen marcador de selección o un gen reportero unido operativamente a un tercer promotor activo en la célula pluripotente; o (j) una combinación de estos.

60. El método de cualquiera de los ejemplos 25-59, en donde el locus genómico de interés comprende (i) una secuencia objetivo 5' que es homóloga al brazo de homología 5'; y (ii) una secuencia objetivo 3' que es homóloga al brazo de homología 3'.

61. El método del ejemplo 60, en donde la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 3 Mb.

62. El método del ejemplo 60, en donde la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 5 kb pero menos de 10 kb, al menos 10 kb pero menos de 20 kb, al menos 20 kb pero menos de 40 kb, al menos 40 kb pero menos de 60 kb, al menos 60 kb pero menos de 80 kb, al menos aproximadamente 80 kb pero menos de 100 kb, al menos 100 kb pero menos de 150 kb, o al menos 150 kb pero menos de 200 kb, al menos aproximadamente 200 kb pero menos de aproximadamente 300 kb, al menos aproximadamente 300 kb pero menos de aproximadamente 400 kb, al menos aproximadamente 400 kb pero menos de aproximadamente 500 kb, al menos aproximadamente 500 kb pero menos de aproximadamente 1Mb, al menos aproximadamente 1 Mb pero menos de aproximadamente 1,5 Mb, al menos aproximadamente 1,5 Mb pero menos de aproximadamente 2 Mb, al menos aproximadamente 2 Mb pero menos de aproximadamente 2,5 Mb, o al menos aproximadamente 2,5 Mb pero menos de aproximadamente 3 Mb.

63. El método del ejemplo 60, en donde la secuencia objetivo 5' y la secuencia objetivo 3' están separadas por al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 110 kb, al menos 120 kb, al menos 130 kb, al menos 140 kb, al menos 150 kb, al menos 160 kb, al menos 170 kb, al menos 180 kb, al menos 190 kb, o al menos 200 kb.

64. El método de cualquiera de los ejemplos 25-63, en donde el locus genómico de interés comprende el locus de la cadena gamma de receptor de interleucina-2, el locus de *ApoE*, el locus de *Rag1*, el locus de *Rag2*, o ambos loci de *Rag1* y de *Rag2*.

65. El método de cualquiera de los ejemplos 25-63, en donde el locus genómico de interés comprende el locus de *Adamts5*, el locus de *Trpa1*, el locus de *Folh1*, o el locus de *ErbB4*.

66. El método de cualquiera de los ejemplos 25-63, en donde el locus genómico de interés comprende el locus de *Lrp5*.

67. Un método para producir un animal no humano de generación F0 que comprende una modificación genética dirigida en un locus genómico de interés, el método que comprende: (a) poner en contacto el genoma en una célula ES no humana con una proteína Cas, un ARN de CRISPR, y un ARNtracr en la presencia de un vector de direccionamiento grande (LTVEC) para formar una célula ES no humana modificada, en donde el LTVEC es de al menos 10 kb y comprende un primer ácido nucleico flanqueado por un brazo de homología 5' y un brazo de homología 3'; (b) identificar la célula ES no humana modificada que comprende la modificación genética dirigida en el locus genómico de interés; (c) introducir la célula ES no humana modificada en un embrión huésped no humano; y (d) gestar el embrión huésped no humano en una madre sustituta, en donde la madre sustituta produce el animal no humano de generación F0 que comprende la modificación genética dirigida en el locus genómico de interés.

68. El método de la modalidad 67, en donde el ARN de CRISPR y el ARNtracr se introducen juntos en la forma de un único ARN guía (ARNg).

69. El método de la modalidad 67, en donde el ARN de CRISPR y el ARNtracr se introducen por separado.

70. El método de cualquiera de las modalidades 67-69, en donde: (a) la proteína Cas se introduce en la célula ES no humana en la forma de una proteína, un ARN mensajero (ARNm) que codifica la proteína Cas, o un ADN que codifica la proteína Cas; (b) el ARN de CRISPR se introduce en la célula ES no humana en la forma de un ARN o un ADN que codifica el ARN de CRISPR; y (c) el ARNtracr se introduce en la célula ES no humana en la forma de un ARN o un ADN que codifica el ARNtracr.

71. El método de la modalidad 70, en donde la proteína Cas, el ARN de CRISPR, y el ARNtracr se introducen en la célula ES no humana como un complejo de proteína-ARN.

72. El método de la modalidad 70, en donde: (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; (b) el ADN que codifica el ARN de CRISPR está en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica el ARN de CRISPR; y (c) el ADN que codifica el ARNtracr está en la forma de un tercer constructo de expresión que comprende un tercer promotor unido operativamente a un cuarto ácido nucleico que codifica el ARNtracr, en donde los promotores primero, segundo y tercero son activos en la célula ES no humana.

73. El método de la modalidad 72, en donde los constructos de expresión primero, segundo y tercero están en una única molécula de ácido nucleico.

74. El método de la modalidad 70, en donde: (a) el ADN que codifica la proteína Cas está en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un segundo ácido nucleico que codifica la proteína Cas; y (b) el ADN que codifica el ARN de CRISPR y el ADN que codifica el ARNtracr están en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un tercer ácido nucleico que codifica un ARNg que comprende el ARN de CRISPR y el ARNtracr; en donde los promotores primero y segundo son activos en la célula ES no humana.

75. El método de la modalidad 74, en donde los constructos de expresión primero y segundo están en una única molécula de ácido nucleico.

76. El método de cualquiera de las modalidades 67-75, en donde la modificación genética dirigida comprende la delección simultánea de una secuencia de ácidos nucleicos endógena en el locus genómico de interés y la inserción del primer ácido nucleico en el locus genómico de interés.

77. El método de cualquiera de las modalidades 67-76, en donde la modificación genética dirigida es una modificación genética bialélica.

78. El método de la modalidad 77, en donde la modificación genética bialélica comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico en el locus genómico de interés en dos cromosomas homólogos.

79. El método de cualquiera de las modalidades 67-76, en donde la célula ES no humana modificada es hemicigota en el locus genómico de interés.

80. El método de la modalidad 79, en donde la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del primer ácido nucleico.

81. El método de la modalidad 79, en donde la modificación genética dirigida comprende: (1) delección de una secuencia de ácidos nucleicos endógena en el locus genómico de interés en dos cromosomas homólogos; y (2) inserción del primer ácido nucleico en el locus genómico de interés en un primer cromosoma y alteración del locus genómico de interés en un segundo cromosoma.

82. El método de cualquiera de las modalidades 67-81, en donde la proteína Cas es Cas9.

EJEMPLOS

Los siguientes ejemplos se proponen para proporcionar a los expertos en la técnica una descripción completa de cómo realizar y usar la presente invención, y no se destinan a limitar el alcance de lo que los inventores consideran como su invención ni se destinan para representar que los experimentos a continuación son todos o los únicos experimentos realizados. Se han realizado esfuerzos para asegurar la precisión con respecto a los números usados (*por ejemplo*, cantidades, temperatura, etcétera) pero deben tenerse en cuenta algunos errores experimentales y desviaciones. A menos que se indique de cualquier otra manera, las partes son partes en peso, el peso molecular es peso molecular promedio, la temperatura es en grados Centígrados, y la presión es la atmosférica o cercana a esta.

Ejemplo 1. Derivación y caracterización de células ES de rata

1.1. Caracterización de células ES de rata

Como se muestra en la figura 1, las ESC de rata crecieron como colonias esféricas compactas, las que rutinariamente se desprenden y flotan en el plato (primer plano, figura 8). Las ESC de rata expresan marcadores de pluripotencia que incluyen Oct-4 (figura 2A) y Sox2 (figura 2B), y expresan altos niveles de fosfatasa alcalina (figura 3). El cariotipo de la línea DA.2B es 42X,Y (figura 4). Las ESC de rata frecuentemente se vuelven tetraploides; por lo tanto, las líneas se pretamizan mediante el recuento de las propagaciones de cromosomas en metafase; las líneas con recuentos en su mayoría normales después se cariotiparon formalmente.

Los blastocistos ACI se recolectaron a partir de hembras superovuladas obtenidas comercialmente. Los blastocistos DA se cultivaron a partir de los embriones de 8 células congeladas obtenidos comercialmente. Las zonas pelúcidas se eliminaron con Tirodes Ácida; y los blastocistos se sembraron en placas sobre MEF mitóticamente inactivadas. Los crecimientos se seleccionaron y se expandieron mediante el uso de métodos estándar. Todos los blastocistos se sembraron en placas, se cultivaron y se expandieron mediante el uso de medios 2i (Li y otros. (2008) Germline competent embryonic stem cells derived from rat blastocysts, Cell 135:1299-1310.

ES 2 700 596 T3

		Tabla 1. Derivación de células ES de rata	
		ACI	DA
5	Fuente de embriones	Blastocistos (Superovulación)	Embriones de 8 células congeladas cultivadas a blastocisto
	Blastocistos sembrados en placas:	107	22
10	Crecimientos:	32 (30 % de blastos)	10 (45 % de blastos)
	Líneas:	16 (50 % de excrecencias)	9 (90 % de excrecencias)
15	Cariotipado:	3; todos 42X,Y	6: 3 42X,X 3 42X,Y
	GLT validado:	1 (ACI.G1)	1 42X,X (DA.2C) 1 42X,Y (DA.2B)

1.2.: Producción de ratas

Las ratas quiméricas se produjeron mediante la inyección de blastocistos y la transmisión del genoma de ESC de rata. Las quimeras producidas mediante la microinyección de blastocistos mediante el uso de las ESC de ratas ACI.G1 progenitoras se muestran en la figura 9. Las crías de agutí F1 con compañeros de camada albinos, engendradas por la quimera ACI/SD marcada con asterisco (*) en la figura 9 se muestran en la figura 10.

Transmisión de la línea germinal de ESC de ratas progenitoras.

Tres líneas ESC de ratas euploides se evaluaron para la pluripotencia mediante microinyección en blastocistos SD albinos. Las quimeras se identificaron por el color de pelaje agutí, lo que indica la contribución de las ESC a las ratas (ver la figura 10). Para cada línea, una mayoría de quimeras transmitieron el genoma de rESC a la descendencia F1 (Tabla 2).

Tabla 2. Transmisión de la línea germinal de rESC progenitora					
Línea (Cariotipo)	Cultivo de quimeras	Trasmisores de la línea germinal	Crías totales de quimeras GLT	crías derivadas de rESC	Eficiencia de GLT (%)
ACI.G1 (XY)	5	3(60 %)	103	11	11
DA.2B (XY)	5	4(80 %)	129	11	9
DA.2C (XX)	3	2(66 %)	45	7	16

1.3.: Derivación de células madre embrionarias de rata.

Protocolo de superovulación, ratas

Día 0: inyectar con suero de yegua embarazada: IP, 20 U (0,4 ml).

Día 1: sin acción

Día 2: (46 horas después): inyectar con hCG, IP, 50 U (1 ml).

establecer parejas de hembras únicas.

Día 3: verificar los tapones vaginales. A las hembras se les introdujeron tapones vaginales. Este es el día 0,5.

Día 6 (e3.5): Sacrificar las hembras y descargar los embriones.

Protocolo de obtención de células ES (Superovulación)

Día 0:

ES 2 700 596 T3

1) Sacrificar ratas hembras con CO₂.

2) Frotar el abdomen ventral con etanol al 70 %; mediante el uso de tijeras, abrir la pared del cuerpo ventral para exponer las vísceras.

3) Diseccionar los oviductos y los cuernos uterinos y colocarlos en un plato de cultivo de tejidos que contiene medio N2B27 tibio. Lavar la mayor cantidad de sangre posible y transferirla a un nuevo plato con N2B27.

4) Mediante el uso de una jeringa de 1 ml y una aguja roma de 27 g, descargar el medio a través de los cuernos uterinos y los oviductos para extraer los blastocistos en el medio.

5) Recolectar los blastocistos con una pipeta bucal y transferirlos a un plato del cultivo de embriones que contiene KSOM + 2i (PD0325901 1 μM, CHIR99021 3 μM). KSOM es un medio de cultivo producido por Millipore. El número de catálogo es MR-106-D.

6) cultivar durante toda la noche a 37 °; CO₂ al 7,5 %.

Protocolo de obtención de células ES (embriones congelados)

Día 0:

1) Descongelar los embriones de 8 células congelados (obtenidos comercialmente) en medio M2. Cultivar 10 minutos a temperatura ambiente.

2) Transferir a KSOM + 2i y cultivar durante toda la noche.

Protocolo de obtención de células ES (igual para ambos)

Día 1:

1) Transferir los embriones cavitados a medio 2i y cultivarlos durante toda la noche.

2) Continuar el cultivo de los embriones no cavitados en KSOM +2i

Día 2:

1) Transferir todos los embriones restantes al medio 2i (ya sea que hayan cavitado o no).

2) Cultivar durante toda la noche; continuar el cultivo de embriones tempranos en medio 2i.

Día 3:

1) Transferir los embriones durante 30 - 60 segundos con Tirodes Ácida para eliminar la zona pelúcida.

2) Lavar los embriones 3x en medio 2i para eliminar la Tirodes Ácida.

3) Depositar cada embrión en un pocillo separado de una placa alimentadora de 96 pocillos (el pocillo contiene una monocapa de fibroblastos embrionarios de ratón (MEF) mitóticamente inactivados.

4) Cultivar durante toda la noche en medio 2i.

Día 4 - 5:

1) Monitorear los embriones sembrados en placas para la presencia de una excrecencia (una masa no diferenciada amorfa de células). Las excrecencias están listas para su transferencia cuando tengan aproximadamente dos veces el tamaño del embrión sembrado en placas.

2) Cada día: eliminar el medio gastado con una micropipeta y reemplazarlo con medio 2i fresco.

3) Transferir las excrecencias a nuevos pocillos alimentadores:

a. Eliminar el medio gastado y lavar suavemente los pocillos con PBS.

b. Eliminar el PBS y añadir 30 μl de tripsina al 0,05 %; incubar durante 10 minutos.

c. Parar la reacción con tripsina mediante la adición de 30 μl de 2i + FBS al 10 %.

ES 2 700 596 T3

d. Disociar suavemente las células con una micropipeta y transferir todos los contenidos del pocillo a un nuevo pocillo en una placa alimentadora de 24 pocillos. Este fue el Pase 1 (P1).

e. Cultivar durante toda la noche en medio 2i.

5 Día 5 - 8: (el momento depende de qué tan rápido se expande cada línea)

1) Cambiar el medio cada día (medio 2i) y monitorear para la presencia de colonias con una morfología de ESC.

10

2) Cuando aparecen las colonias, continuar el cultivo hasta que las colonias se expanden a una confluencia de ~ 50 %.

3) Someter a tripsina las colonias y a pases como antes; sembrar en placas sobre alimentadores, 1 pocillo por línea, en una placa de 6 pocillos. Este fue el Pase 2 (P2).

En proceso:

15

1) Continuar la alimentación y monitorear cada línea hasta aproximadamente una confluencia del 50 %.

2) Someter a tripsina las células como de costumbre.

20

3) Parar la tripsina con 2i + FBS al 10 %; sedimentar las células mediante centrifugación (5', 1 200 rpm en una centrífuga de mesa de Beckman-Coulter).

4) Aspirar el sobrenadante y resuspender las células suavemente en 400 µl de Medio de congelación (2i al 70 %, FBS al 20 %, DMSO al 10 %).

25

5) Distribuir las células en 2 viales y congelar a -80 °. Este fue el Pase 3 (P3).

6) Para el almacenamiento a largo plazo, transferir los viales a almacenamiento de N₂ líquido.

30

El medio 2i se preparó como sigue en la Tabla 3.

Reactivo	Vendedor	Concentración
Medio basal DMEM/F12	Invitrogen/Life Technologies	1x
Medio neurobasal	Invitrogen/Life Technologies	1x
Penicilina/ estreptomycin	Invitrogen/Life Technologies	1 %
L-glutamina	Invitrogen/Life Technologies	4 mM
2-Mercaptoetanol	Invitrogen/Life Technologies	0,1 mM
Suplemento de N2	Invitrogen/Life Technologies	1x
Suplemento B27	Invitrogen/Life Technologies	1x
LIF	Millipore	100 U/ml
PD0325901 (inhibidor de MEK).	Stemgent	1 uM
CHIR99021 (inhibidor de GSK).	Stemgent	3 uM

50

Materiales: Gonadotropina de suero de yegua preñada (PMSG)

Gonadotropina coriónica humana en orina de embarazadas (HCG)

Ratas hembras (5-12 semanas de edad)

55

Ratas machos (12 semanas a 8 meses de edad), uno por jaula

Jeringas/agujas

Sala de animales con luces encendidas 6:00-18:00

Procedimiento:

60

Día 1: 8:00-10:00 AM

Inyectar las hembras con 20 IU de PMSG (0,4 ml), IP

65

Desechar el PMSG sin usar.

ES 2 700 596 T3

Día 3: 8:00-10:00 AM (48 horas después de la inyección de PMSG)

Inyectar las hembras con 50 IU de HCG (1 ml), IP
Colocar una hembra por macho en una jaula de apareamiento.

5

Desechar la HCG no usada.

Día 4: 8:00-10:00 AM (24 horas después de la inyección de HCG)

10

Verificar los tapones vaginales de las hembras.

Suministradores de hormonas

15

PMSG: Sigma #G-4877 (1 000 IU). Resuspender en PBS a una [] final de 50 IU/ml. Almacenar a -20° en alícuotas de 1 ml.

HCG: Sigma #CG-5 (5 000 IU). Resuspender en PBS a una [] final de 50 IU/ml. Almacenar a -20° en alícuotas de 1 ml.

20

1.4.: Cariotipado de líneas de células madre embrionarias de rata

Las líneas de células ES de rata generadas en la presente descripción se cariotiparon, y los resultados se resumen en las Tablas 4-7.

25

Tabla 4

Resultados del cariotipado de ACI.G1	Número de células
Número de células cariotipadas	7
Número de células analizadas	20
Número de células 42, XY	18
Número de células anormales	2
40, XY, -5, -9	1
41, XY, -14	1
42, XY	18
Otras notas:	
A las dos células analizadas les faltaban diferentes autosomas, lo que puede ser de ocurrencia esporádica debido a un artefacto técnico. El 90 % de las células analizadas tenían un cariotipo 42, XY de macho normal.	
La figura 5 proporciona una fotografía que muestra el análisis del número de cromosomas de la línea de células ES de rata ACI.G1.	

40

45

50

Tabla 5

Resultados del cariotipado de DA.2B	Número de células
Número de células cariotipadas	6
Número de células analizadas	20
Número de células 42, XY	20
Número de células anormales	0
42, XY	20
Otras notas:	
Todas las células analizadas tenían un cariotipo 42, XY diploide normal.	
La figura 6 proporciona una fotografía que muestra el análisis del número de cromosomas de la línea de células ES de rata DA.2B.	

55

60

65

Tabla 6

Resultados del cariotipado de DA.2C		Número de células
5	Número de células cariotipadas	5
	Número de células analizadas	20
	Número de células 42, XX	20
10	Número de células anormales	0
	42, XX	
	Otras notas:	
15	El 100 % de las células analizadas tenían un cariotipo de rata XX femenino.	
	La figura 7 proporciona una fotografía que muestra el análisis del número de cromosomas de la línea de células ES de rata DA.2C.	

Tabla 7.

cepa	Blastocistos sembrados en placas	Líneas establecidas	Líneas Cariotipadas	Cariotipos	
25	BN x SD F1	41	8(20 %)	5	todas las líneas fueron poliploides complejas en un alto %
30	ACI	27	16 (60 %)	3	G1: 90 % 42 XY; otras fueron 70-85 % euploides
35	DA	20	9 (45 %)	6	2B: 100 % 42 XY; 2C: 100 % 42 XX; otras fueron 95-100 % euploides
	F344	4	1 (25 %)	0	
	Totales	92	34 (37 %)		

1.5.: Electroporación del vector en célula madre embrionaria de rata

1. Someter las células ES de rata 24-48 horas antes de la electroporación.
2. Cambiar el medio a RVG2i + ROCKi (Y-27632 10 µM) 24 horas antes de la electroporación
3. Cambiar el medio 30' antes de la tripsinización.
4. Alicuotear el ADN para ser electroporado.
5. Dejar calentar el ADN a TA durante >10 min.
6. Calentar el ADN durante 5' @ 62 °C. Colocar el ADN en hielo.
7. Someter las células a tripsina:
 - a. Recolectar las colonias flotantes. Lavar la placa para recolectar tantas colonias flotantes como sea posible.
 - b. Sedimentar las colonias: 3' @ 750 rpm.
 - c. Lavar el sedimento 1x con 5-10 ml de PBS y volver a centrifugar/sedimentar
 - d. Aspirar el sobrenadante; añadir tripsina de 500λ, 0,05 % + suero de gallina al 1 %.
 - i. No agrupar más de 1 placa de 10 cm de colonias por tubo. Si hay demasiadas colonias empacadas en el fondo del tubo durante la tripsinización se agruparán y la mayoría de las células se perderán.

ES 2 700 596 T3

- e. 4' @ 37 °. Pipetear varias veces las colonias para minimizar el agrupamiento.
- f. Repetir las etapas 1-2 X: 4' @ 37 °.
- 5 g. Parar la tripsina con RVG2i 500λ + FBS al 10 %.
8. Sedimentar las células: 5' @ 1 200 rpm.
- 10 9. Resuspender las células en 10 ml de PBS. Hacer el recuento de dos alícuotas de 20λ para determinar el número total de células.
10. Sedimentar las células (5'/1 200rpm); calcular el número total de células y el volumen total de resuspensión para lograr la concentración celular correcta (objetivo #/75 μl de tampón EP).
- 15 11. Resuspender en un volumen mínimo de tampón EP; medir el volumen total y ajustar el volumen objetivo con tampón EP. El tampón de electroporación es vendido por Millipore. El número de catálogo es ES-003-D. Ver, Valenzuela y otros. (2003) Nature Biotechnology 21:652-659.
- 20 12. Añadir las células de 75λ a ADN de 50λ; transferir la solución de células/ADN a 125λ a un pocillo de una cubeta de 48 pocillos BTX.
- a. Llenar los pocillos vacíos en la misma columna con tampón EP de 125λ.
- 25 13. Someter a pulsos la cubeta una vez en el electroporador BTX:
- a. Ajustes: 400 V; Ω; 100 μF (los ajustes pueden variar)
14. Colocar la cubeta en hielo durante 15' para recuperarla.
- 30 15. Remover las células en 5 ml de RVG2i + ROCKi 10 μM.
16. Añadir a la placa de 15 cm, 20 ml de RVG2i + ROCKi 10 μM. La placa tiene MEF neoR 2x (u otras MEF en dependencia del proyecto). El marcador de selección neoR es el gen de neomicina fosfotransferasa (neo) de Beck y otros. (1982) Gene, 19:327-36 o en la patente de los Estados Unidos núm., 7,205,148 o 6,596,541.
- 35 17. Incubar @ 37 °. Comenzar la selección 48 horas después.
- El inhibidor de ROCK usado fue Y-27632.
- 40 1.6: Selección de una modificación genética dirigida en una célula madre embrionaria de rata.
1. Someter a pases las células durante 24-48 horas antes de electroporación.
- 45 2. Cambiar el medio a RVG2i + ROCKi (Y-27632 10 μM) 24 horas antes de la electroporación
3. Cambiar el medio 30' antes de la tripsinización.
4. Alicuotear el ADN para ser electroporado.
- 50 5. Dejar calentar el ADN a TA durante >10 min.
6. Calentar el ADN durante 5' @ 62 °C. Colocar el ADN en hielo.
7. Someter las células a tripsina:
- 55 a. Recolectar las colonias flotantes. Lavar la placa para recolectar tantas colonias flotantes como sea posible.
- b. Sedimentar las colonias: 3' @ 750 rpm.
- 60 c. Lavar el sedimento 1x con 5-10 ml de PBS y volver a centrifugar/sedimentar
- d. Aspirar el sobrenadante; añadir tripsina de 500λ, 0,05 % + suero de gallina al 1 %.
- 65 i. No agrupar más de 1 placa de 10 cm de colonias por tubo. Si hay demasiadas colonias empacadas en el fondo del tubo durante la tripsinización se agruparán y la mayoría de las células se perderán.

ES 2 700 596 T3

- e. 4' @ 37 °. Pipetear las colonias varias veces para minimizar el agrupamiento
- f. Repetir 1-2 X: 4' @ 37 °.
- 5 g. Parar la tripsina con RVG2i 500λ + FBS al 10 %.
8. Sedimentar las células: 5' @ 1 200 rpm.
- 10 9. Resuspender las células en 10 ml de PBS. Hacer el recuento de dos alícuotas de 20λ para determinar el número total de células.
10. Sedimentar las células (5'/1 200rpm); calcular el número total de células y el volumen total de resuspensión para lograr la concentración celular correcta (objetivo #/75 μl de tampón EP).
- 15 11. Resuspender en un volumen mínimo de tampón EP; medir el volumen total y ajustar hasta el volumen objetivo con tampón EP.
12. Añadir células de 75λ a ADN de 50λ; transferir la solución de células/ADN de 125λ a un pocillo de una cubeta de 48 pocillos BTX.
- 20 a. Llenar los pocillos vacíos en la misma columna con tampón EP de 125λ.
13. Someter a pulsos la cubeta una vez en el electroporador BTX:
- 25 a. Ajustes: 400 V; 100 μF (los ajustes pueden variar)
14. Colocar la cubeta en hielo durante 15' para recuperarla.
15. Eliminar las células en 5 ml de RVG2i + ROCKi 10 μM.
- 30 16. Añadir a la placa de 15 cm, 20 ml de RVG2i + ROCKi 10 μM. La placa tenía MEF neoR 2x (u otras MEF en dependencia del proyecto).
17. Incubar @ 37 °. Comenzar la selección 48 horas después.
- 35 18. El protocolo de selección de G418 fue como sigue:
- a. Día 2 (2^{do} día después de EP): incubar las células en medio 2i + G418, 75 μg/ml.
- 40 b. Día 3: incubar las células en medio 2i sin G418
- c. Día 4: incubar las células en medio 2i + G418, 75 μg/ml.
- 45 d. Día 5: incubar las células en medio 2i sin G418
- e. Día 6: incubar las células en medio 2i + G418, 75 μg/ml.
- f. Día 7: incubar las células en medio 2i sin G418
- 50 g. Día 8: incubar las células en medio 2i + G418, 75 μg/ml.
- h. Día 9: incubar las células en medio 2i sin G418
- i. Día 10: incubar las células en medio 2i + G418, 75 μg/ml.
- 55 j. Día 11: incubar las células en medio 2i sin G418
- k. Día 12: seleccionar las colonias para expandirlas para tamizaje. Cada colonia se disoció en tripsina 0,05 % + suero de pollo al 1 % durante 10 minutos y después se colocó en 1 pocillo de una placa alimentadora de 96 pocillos.
- 60 19. Expandir las colonias durante 3 días en medio 2i.
20. Someter a pases los clones 1:1 a nuevas placas alimentadoras de 96 pocillos.
- 65 21. Expandir los clones durante 3 días en medio 2i.

22. Para cada clon, disociar las colonias en tripsina. Congelar 2/3 de cada clon y almacenar a -80 °; colocar el 1/3 restante en placas de laminina (placas de 96 pocillos recubiertas con 10 µg/ml de laminina).

23. Cuando las placas de laminina fueron confluentes, pasaron al laboratorio de tamizaje para el genotipado de los clones.

1.7. Firma molecular de las células madre embrionarias de rata

Los genes enumerados en la Tabla 8 se expresaron 20 veces menos en las células ES de rata que los genes correspondientes en células ES de ratón. Los genes enumerados en la Tabla 9 se expresaron a niveles 20 veces más altos en células ES de rata que los genes correspondientes en células ES de ratón.

Los datos de microarreglos en las Tablas 8 y 9 de generaron como sigue. Las células ES de rata (AC1.G2 y DA.2B) y las células ES de ratón (F1H4) se cultivaron en medio 2i durante 3 pases hasta que confluyeron. Las células F1H4 se cultivaron en placas recubiertas con gelatina en la ausencia de alimentadores. Las células ES de ratón F1H4 se obtuvieron de embriones heterocigotos 129S6/SvEvTac y C57BL/6NTac (ver, por ejemplo, la patente de los Estados Unidos núm. 7,294,754 y Poueymirou, W.T., Auerbach, W., Friendewey, D., Hickey, J.F., Escaravage, J.M., Esau, L., Dore, A.T., Stevens, S., Adams, N.C., Dominguez, M.G., Gale, N.W., Yancopoulos, G.D., DeChiara, T.M., Valenzuela, D.M. (2007).

El siguiente protocolo se usó para la preparación de muestras: Los tubos Eppendorf de 1,5 mL se marcaron con la ID de la muestra. Las células cultivadas en una placa se enjuagaron en Solución salina tamponada con fosfato (PBS) a 37 °C. El PBS se eliminó y se añadieron 300 µl de Trizol®. Se usó un raspador para romper las células en Trizol® (Life Technology). Las células lisadas se recolectaron en Trizol® en un tubo Eppendorf de 1,5 mL. Para las células cultivadas en suspensión, las células se enjuagaron en PBS a 37 °C y se recolectaron en un tubo de 1,5 mL. Las células se centrifugaron rápidamente; se eliminó el PBS; y se añadieron 300 µl de Trizol® a las células. Las membranas celulares se rompieron mediante pipeteo. Las muestras se clasificaron por FACS con 10 a 10⁵ células, el volumen se concentró a menos de 100 µL. Se añadieron 4 volúmenes de tampón de lisis de ARN y se mezclaron mediante pipeteo. Para la muestra, se añadieron 320 µL de tampón de lisis de ARN a 80 µL de muestra. Las muestras se almacenaron a -20 °C.

Se usó ARN-Seq para medir el nivel de expresión de genes de ratón y rata. Las lecturas de secuenciación se mapearon con respecto al genoma de referencia de ratón y rata por Tophat, y los RPKM (fragmentos por kilobase de exón por millones de fragmentos mapeados) se calcularon para genes de ratón y rata. Se seleccionaron los genes de homología basado en el símbolo de los genes, y después se usó la prueba t para comparar el nivel de expresión de cada gen entre ratón y rata. miR-32 estuvo en el top 10 más alto expresado en las ESC de rata pero no se expresó en las células ES de ratón. Aunque no existen datos comparativos de miR-632, basado en el nivel de su expresión en comparación con otros genes expresados en las ESC de rata y su función conocida en el desarrollo embrionario, miR-632 se seleccionó como un marcador para células ES de rata.

Tabla 8. Los genes enumerados se expresaron a niveles 20 veces más bajos en células ES de rata que los genes correspondientes en células ES de ratón.

ID	Símbolo	Nombre en Entrez Gene	Localización	Tipo(s)
Abcb1b	Abcb1b	Casete de unión a ATP, subfamilia B (MDR/TAP), miembro 1B	Membrana plasmática	transportador
Acta2	ACTA2	actina, alfa 2, músculo liso, aorta	Citoplasma	otro
Actg2	ACTG2	actina, gamma 2, músculo liso, entérico	Citoplasma	otro
Aebp1	AEBP1	proteína de unión a AE 1	Núcleo	peptidasa
Angptl2	ANGPTL2	tipo angiopoyetina 2	Espacio extracelular	otro
Ankrd1	ANKRD1	dominio de repeticiones de anquirina 1 (músculo cardíaco)	Citoplasma	regulador de la transcripción
Anxa1	ANXA1	anexina A1	Membrana plasmática	otro
Anxa6	ANXA6	anexina A6	Membrana plasmática	otro
Anxa8	ANXA8L2	tipo anexina A8 2	Membrana plasmática	otro
Arhgef25	ARHGEF25	Factor de intercambio de nucleótidos de guanina Rho (GEF) 25	Citoplasma	otro

ES 2 700 596 T3

Axl	AXL	receptor tirosina cinasa AXL	Membrana plasmática	cinasa
Basp1	BASP1	proteína de señales unida a la membrana 1, abundante en cerebro	Núcleo	regulador de la transcripción
Bgn	BGN	biglicano	Espacio extracelular	otro
Bst2	BST2	antígeno de células estromales de médula ósea 2	Membrana plasmática	otro
Btf3	BTF3	factor de la transcripción básico 3	Núcleo	regulador de la transcripción
Btg2	BTG2	familia de BTG, miembro 2	Núcleo	regulador de la transcripción
Capsl	CAPSL	tipo calcifosina	Otro	otro
Cav1	CAV1	caveolina 1, proteína caveolae, 22 kDa	Membrana plasmática	receptor de transmembrana
Ccdc80	CCDC80	que contiene dominio de hélice enrollada	Núcleo	otro
		que contiene 80		
Ccnd2	CCND2	ciclina D2	Núcleo	otro
Cd248	CD248	molécula CD248, endosialina	Membrana plasmática	otro
Cd44	CD44	molécula de CD44 (grupo sanguíneo de la India)	Membrana plasmática	enzima
Cd97	CD97	molécula CD97	Membrana plasmática	receptor acoplado a proteína G
Cdc42ep5	CDC42EP5	proteína efectora de CDC42 (de unión a GTPasa Rho) 5	Citoplasma	otro
Cdh11	CDH11	caderina 11, Tipo 2, OB-caderina (osteoblasto)	Membrana plasmática	otro
Cdkn2a	CDKN2A	inhibidor cinasa dependiente de ciclina 2A	Núcleo	regulador de la transcripción
Cdo1	CDO1	cisteína dioxigenasa de tipo 1	Citoplasma	enzima
Clip3	CLIP3	proteína enlazadora que contiene el dominio CAP-GLY 3	Citoplasma	otro
Cln5	CLN5	ceroide-lipofuscinosis, neuronal 5	Citoplasma	otro
Cnn1	CNN1	calponina 1, básico, músculo liso	Citoplasma	otro
Col1a1	COL1A1	colágeno, tipo I, alfa 1	Espacio extracelular	otro
Col1a2	COL1A2	colágeno, tipo I, alfa 2	Espacio extracelular	otro
Col3a1	COL3A1	colágeno, tipo III, alfa 1	Espacio extracelular	otro
Col5a2	COL5A2	colágeno, tipo V, alfa 2	Espacio extracelular	otro
Col6a2	COL6A2	colágeno, tipo VI, alfa 2	Espacio extracelular	otro
Cryab	CRYAB	cristalina, alfa B	Núcleo	otro
Csf1	CSF1	factor estimulante de colonias 1 (macrófago)	Espacio extracelular	citocina

ES 2 700 596 T3

Cth	CTH	cistationasa (cistationina gamma-liasa)	Citoplasma	enzima
Cthrc1	CTHRC1	que contiene repeticiones de triple hélice de colágeno 1	Espacio extracelular	otro
Ctsc	CTSC	catepsina C	Citoplasma	peptidasa
Cyr61	CYR61	rico en cisteína, inductor angiogénico, 61	Espacio extracelular	otro
Ddx58	DDX58	polipéptido 58 de caja DEAD (Asp-Glu-Ala-Asp)	Citoplasma	enzima
Dkk3	DKK3	inhibidor 3 de la ruta de señalización WNT de dickkopf	Espacio extracelular	citocina
Dmc1	DMC1	recombinasa meiótica de ADN 1	Núcleo	enzima
Dpysl3	DPYSL3	tipo dihidropirimidinas 3	Citoplasma	enzima
Dse	DSE	dermatán sulfato epimerasa	Citoplasma	enzima
Dusp1	DUSP1	fosfatasa de doble especificidad 1	Núcleo	fosfatasa
Dusp27	DUSP27	fosfatasa de doble especificidad 27 (putativa)	Otro	fosfatasa
Dusp9	DUSP9	fosfatasa de doble especificidad 9	Núcleo	fosfatasa
Ece2	ECE2	enzima de conversión de endotelina 2	Membrana plasmática	peptidasa
Ecm1	ECM1	proteína de matriz extracelular 1	Espacio extracelular	transportador
Egr1	EGR1	respuesta al crecimiento temprano 1	Núcleo	regulador de la transcripción
Emp1	EMP1	proteína de membrana epitelial 1	Membrana plasmática	otro
Emp3	EMP3	proteína de membrana epitelial 3	Membrana plasmática	otro
Ephx2	EPHX2	epóxido hidrolasa 2, citoplasmática	Citoplasma	enzima
F3	F3	factor de coagulación III (tromboplastina, factor tisular)	Membrana plasmática	receptor de transmembrana
Fau	FAU	virus del sarcoma murino Finkel-Biskis-Reilly (FBR-MuSV) expresado de manera ubicua	Citoplasma	otro
Fbn1	FBN1	fibrilina 1	Espacio extracelular	otro
Fbxo15	FBXO15	proteína de F-box 15	Otro	regulador de la transcripción
Fhl2	FHL2	cuatro y la mitad de dominios LIM 2	Núcleo	regulador de la transcripción
Finc	FLNC	filamina C, gamma	Citoplasma	otro
Fos	FOS	oncogén homólogo viral de osteosarcoma murino FBJ	Núcleo	regulador de la transcripción
Fundc2	FUNDC2	que contiene dominio FUN 14 2	Citoplasma	otro
Gjb3	GJB3	proteína de unión de interrupciones, beta 3, 31 kDa	Membrana plasmática	transportador
Gpa33	GPA33	glicoproteína A33 (transmembrana)	Membrana plasmática	otro

ES 2 700 596 T3

Gpbp111	GPBP1L1	tipo proteína 1 de unión a promotores ricos en GC 1	Otro	otro
Gpc3	GPC3	glicoproteína 3	Membrana plasmática	otro
Grb10	GRB10	proteína unida a receptor de factor de crecimiento 10	Citoplasma	otro
Gstm1	GSTM5	glutatión S-transferasa mu 5	Citoplasma	enzima
Hap1	HAP1	proteína asociada a huntingtina 1	Citoplasma	otro
Hist1h2bc	HIST2H2B E (incluye otros)	grupo de histonas 2, H2be	Núcleo	otro
Hmga2	HMGA2	grupo de alta movilidad AT-gancho 2	Núcleo	enzima
Hmgn3	Hmgn3	grupo de alta movilidad de dominio de unión a nucleosoma 3	Núcleo	otro
Hormad1	HORMAD 1	que contiene dominio HORMA 1	Núcleo	otro
Hsd17b14	HSD17B14	hidroxiesteroide (17-beta) deshidrogenasa 14	Citoplasma	enzima
Hspb1	HSPB1	proteína de 27 kDa de choque térmico 1	Citoplasma	otro
Hspb8	HSPB8	proteína de 22 kDa de choque térmico 8	Citoplasma	cinasa
Htra1	HTRA1	HtrA serina peptidasa 1	Espacio extracelular	peptidasa
Ifi204	Ifi204 (incluye otros)	gen activado de interferón 204	Núcleo	regulador de la transcripción
Ifi44	IFI44	proteína inducida por interferón 44	Citoplasma	otro
Ifit1	IFIT1B	proteína inducida por interferón con repeticiones de tetratricopéptidos 1B	Citoplasma	otro
Ifitm3	IFITM2	proteína de transmembrana 2 inducida por interferón	Citoplasma	otro
Igf2	IGF2	factor de crecimiento de tipo insulina 2 (somatomedina A)	Espacio extracelular	factor de crecimiento
Igfbp7	IGFBP7	proteína de unión a factor de crecimiento de tipo insulina 7	Espacio extracelular	transportador
Il1r1	IL1RL1	de tipo receptor de interleucina-1 1	Membrana plasmática	receptor de transmembrana
Inhba	INHBA	inhibina, beta A	Espacio extracelular	factor de crecimiento
Inhbb	INHBB	inhibina, beta B	Espacio extracelular	factor de crecimiento
Irf7	IRF7	factor regulador de interferón 7	Núcleo	regulador de la transcripción
Isg15	ISG15	modificador de tipo ubiquitina ISG15	Espacio extracelular	otro
Itga5	ITGA5	integrina, alfa 5 (receptor de fibronectina, polipéptido alfa)	Membrana plasmática	receptor de transmembrana
Jun	JUN	jun protooncogén	Núcleo	regulador de la transcripción

ES 2 700 596 T3

Junb	JUNB	jun B protooncogén	Núcleo	regulador de la transcripción
Lgals3bp	LGALS3B P	lectina, de unión a galactósido, soluble, proteína de unión 3	Membrana plasmática	receptor de transmembrana
Lgals9	LGALS9	lectina, de unión a galactósido, soluble, 9	Espacio extracelular	otro
Lmna	LMNA	lamina A/C	Núcleo	otro
Lox	LOX	lisil oxidasa	Espacio extracelular	enzima
Loxl2	LOXL2	tipo lisil oxidasa 2	Espacio extracelular	enzima
Loxl3	LOXL3	tipo lisil oxidasa 3	Espacio extracelular	enzima
Lrp1	LRP1	proteína relacionada con receptores de lipoproteína de baja densidad 1	Membrana plasmática	receptor de transmembrana
Mageb16	MAGEB16	familia B de antígeno de melanoma, 16	Otro	otro
Mcam	MCAM	molécula de adhesión de células de melanoma	Membrana plasmática	otro
Mgp	MGP	proteína Gla de matriz	Espacio extracelular	otro
Mmp2	MMP2	metalopeptidasa 2 de matriz (gelatinasa A, gelatinasa de 72kDa, colagenasa tipo IV de 72 kDa)	Espacio extracelular	peptidasa
Mxra8	MXRA8	asociada a remodelación de matriz 8	Otro	otro
MyI9	MYL9	miosina, cadena ligera 9, reguladora	Citoplasma	otro
MyIpf	MYLPF	cadena ligera de miosina, fosforilable, músculo esquelético rápido	Citoplasma	otro
Nab2	NAB2	proteína de unión a NGFI-A 2 (proteína de unión a EGR1 2)	Núcleo	regulador de la transcripción
Ndufb4	NDUFB4	subcomplejo beta 1 de NADH deshidrogenasa (ubiquinona), 4, 15 kDa	Citoplasma	transportador
Npm1	NPM1	nucleofosmina (fosfoproteína nucleolar B23, numatrina)	Núcleo	regulador de la transcripción
Nr0b1	NR0B1	subfamilia del receptor nuclear 0, grupo B, miembro 1	Núcleo	receptor nuclear dependiente de ligando
Nr4a1	NR4A1	subfamilia de receptor nuclear 4, grupo A, miembro 1	Núcleo	receptor nuclear dependiente de ligando
Nrp2	NRP2	neuropilina 2	Membrana plasmática	cinasa
Oas1a	OAS1	2'-5'-oligoadenilato sintetasa 1, 40/46 kDa	Citoplasma	enzima
Oas12	Oas12	tipo 2'-5'-oligoadenilato sintetasa 2	Otro	enzima
P4ha2	P4HA2	proliil 4-hidroxilasa, alfa polipéptido II	Citoplasma	enzima
Parp3	PARP3	familia de la poli (ADP-ribosa) polimerasa, miembro 3	Núcleo	enzima
Pcolce	PCOLCE	potenciador de procolágeno C-endopeptidasa	Espacio extracelular	otro

ES 2 700 596 T3

Pcyl1b	PCYT1B	fosfato citidililtransferasa 1, colina, beta	Citoplasma	enzima
Pdgfc	PDGFC	factor de crecimiento C derivado de plaquetas	Espacio extracelular	factor de crecimiento
Phlda1	PHLDA1	dominio de tipo homología a pleckstrin, familia A, miembro 1	Citoplasma	otro
Phlda2	PHLDA2	dominio de tipo homología a pleckstrin, familia A, miembro 2	Citoplasma	otro
Pla2g1b	PLA2G1B	fosfolipasa A2, grupo IB (páncreas)	Espacio extracelular	enzima
Pla2g4a	PLA2G4A	fosfolipasa A2, grupo IVA (citósólica, dependiente de calcio)	Citoplasma	enzima
Porcn	PORCN	homólogo de puercoespín (<i>Drosophila</i>)	Citoplasma	otro
Postn	POSTN	perioquina, factor específico de osteoblastos	Espacio extracelular	otro
Prrx1	PRRX1	homeocaja 1 relacionada en pares	Núcleo	regulador de la transcripción
Prss23	PRSS23	proteasa, serina, 23	Espacio extracelular	peptidasa
Psmb8	PSMB8	subunidad proteasoma (prosome, macropain), beta tipo, 8	Citoplasma	peptidasa
Ptgs2	PTGS2	prostaglandina-endoperoxidase sintasa 2 (prostaglandina G/H sintasa y ciclooxigenasa)	Citoplasma	enzima
Ptn	PTN	pleiotrofina	Espacio extracelular	factor de crecimiento
Ptfr	PTRF	polimerasa I y factor de liberación de transcrito	Núcleo	regulador de la transcripción
Rarg	RARG	receptor de ácido retinoico, gamma	Núcleo	receptor nuclear dependiente de ligando
Rgs16	RGS16	regulador de la señalización de proteína G 16	Citoplasma	otro
Rn45s	Rn45s	ARN pre-ribosomal 45S	Otro	otro
Rpl10a	RPL10A	proteína ribosomal L10a	Otro	otro
Rpl31	RPL31	proteína ribosomal L31	Otro	otro
Rpl37a	RPL37A	proteína ribosomal L37a	Citoplasma	otro
Rps10	RPS10-NUDT3	transcriptura de RPS10-NUDT3	Citoplasma	otro
Rps14	RPS14	proteína ribosomal S14	Citoplasma	regulador de la traducción
Rps20	RPS20	proteína ribosomal S20	Citoplasma	otro
Rps26	RPS26	proteína ribosomal S26	Citoplasma	otro
Rps9	RPS9	proteína ribosomal S9	Citoplasma	regulador de la traducción
S100a4	S100A4	proteína de unión a calcio S100 A4	Citoplasma	otro
S100a6	S100A6	proteína de unión a calcio S100 A6	Citoplasma	transportador
Schip1	SCHIP1	proteína de interacción con schwannomin 1	Citoplasma	otro

ES 2 700 596 T3

Sdc2	SDC2	sindecano 2	Membrana plasmática	otro
Serpine1	SERPINE1	inhibidor de serpina peptidasa, clade E (nexina, inhibidor del activador de plasminógeno tipo 1), miembro 1	Espacio extracelular	otro
Serpine2	SERPINE2	inhibidor de serpina peptidasa, clade E (nexina, inhibidor del activador de plasminógeno tipo 1), miembro 2	Espacio extracelular	otro
Serpinf1	SERPINF1	inhibidor de serpina peptidasa, clade F (antiplasmina alfa-2, factor derivado de epitelio pigmentario), miembro 1	Espacio extracelular	otro
Sh3gl2	SH3GL2	tipo GRB2 de dominio SH3 2	Membrana plasmática	enzima
Slc19a2	SLC19A2	familia 19 de portadores de soluto (transportador de tiamina), miembro 2	Membrana plasmática	transportador
Slc25a5	SLC25A5	familia 25 de portadores de soluto (portador mitocondrial; translocador de nucleótidos de adenina), miembro 5	Citoplasma	transportador
Slc29a1	SLC29A1	familia 29 de portadores de soluto (transportador de nucleósidos en equilibrio), miembro 1	Membrana plasmática	transportador
Slc35f2	SLC35F2	familia 35 de portadores de soluto, miembro F2	Otro	otro
Snrpn	SNRPN	polipéptido N de ribonucleoproteína nuclear pequeña	Núcleo	otro
Snx22	SNX22	sorting nexina 22	Otro	transportador
Sparc	SPARC	proteína secretada, ácida, rica en cisteínas (osteonectina)	Espacio extracelular	otro
Spp1	SPP1	fosfoproteína secretada 1	Espacio extracelular	citocina
Sult4a1	SULT4A1	familia 4A de la sulfotransferasa, miembro 1	Citoplasma	enzima
Tagln	TAGLN	transgelina	Citoplasma	otro
Tcea3	TCEA3	factor A de elongación de la transcripción (SII), 3	Núcleo	regulador de la transcripción
Tgfb3	TGFB3	factor de crecimiento transformante, beta 3	Espacio extracelular	factor de crecimiento
Thbs 1	THBS1	trombospondina 1	Espacio extracelular	otro
Thbs2	THBS2	trombospondina 2	Espacio extracelular	otro
Tm4sf1	TM4SF1	miembro 1 de seis familias de transmembrana 4 L	Membrana plasmática	otro
Tmbim1	TMBIM1	que contiene un motivo de inhibidor BAX de transmembrana 1	Citoplasma	otro
Tmem176b	TMEM176	proteína de transmembrana	Otro	otro
	B	176B		
Tnc	TNC	tenascina C	Espacio extracelular	otro
Tpd5211	TPD52L1	tipo D52 de proteína de tumor 1	Citoplasma	otro

ES 2 700 596 T3

Tpm2	TPM2	tropomiosina 2 (beta)	Citoplasma	otro
Usp18	USP18	peptidasa 18 especifica de ubiquitina	Citoplasma	peptidasa
Vim	VIM	vimentina	Citoplasma	otro
Wfdc2	WFDC2	dominio núcleo de cuatro disulfuros WAP 2	Espacio extracelular	otro
Wisp2	WISP2	proteína 2 de la ruta de señalización inducible de WNT1	Espacio extracelular	factor de crecimiento
Ybx1	YBX1	proteína 1 de unión a caja Y	Núcleo	regulador de la transcripción

Tabla 9. Los genes enumerados se expresaron a niveles 20 veces mayores en células ES de rata que los genes correspondientes en células ES de ratón.

ID	Símbolo	Nombre en Entrez Gene	Localización	Tipo(s)
Ajap1	Ajap1	proteína 1 asociada a uniones adherentes	Otro	otro
Amd1	AMD1	adenosilmetionina descarboxilasa 1	Citoplasma	enzima
Ankrd2	ANKRD 2	dominio de repeticiones de anquirina 2 (músculo sensible al estiramiento)	Núcleo	regulador de la transcripción
Arhgef9	ARHGE F9	factor 9 de intercambio de nucleótidos de guanina Cdc42 (GEF)	Citoplasma	otro
Atp5h	Atp5h	ATP sintasa, transportador de H+, complejo F0 mitocondrial, subunidad d	Citoplasma	enzima
Btg3	BTG3	familia de BTG, miembro 3	Núcleo	otro
Car6	CA6	anhidrasa carbónica VI	Espacio extracelular	enzima
Camk4	CAMK4	proteína cinasa IV dependiente de calcio/calmodulina	Núcleo	cinasa
Capn12	CAPN12	calpaina 12	Otro	peptidasa
Cct6b	CCT6B	chaperonina que contiene TCP1, subunidad 6B (zeta 2)	Citoplasma	transportador
Cdx2	CDX2	homeocaja 2 tipo caudal	Núcleo	regulador de la transcripción
Cldn5	CLDN5	claudina 5	Membrana plasmática	otro
Clec3a	CLEC3A	familia 3 del dominio de lectina tipo C, miembro A	Otro	otro
Clc6	CLIC6	canal intracelular de cloruro 6	Membrana plasmática	canal iónico
Dhrsx	DHRXS	deshidrogenasa/reductasa (familia SDR) ligada a X	Otro	enzima
Dpysl2	DPYSL2	tipo dihidropirimidinasa 2	Citoplasma	enzima
Dusp26	DUSP26	fosfatasa de doble especificidad 26 (putativa)	Citoplasma	enzima
Eci3	Eci3	enoil-Coenzima A delta isomerasa 3	Otro	enzima
Eef2k	EEF2K	factor-2 de alargamiento eucariota cinasa	Citoplasma	cinasa
Efna1	EFNA1	efrina-A1	Membrana plasmática	otro
Epha4	EPHA4	receptor A4 de EPH	Membrana plasmática	cinasa

ES 2 700 596 T3

Fank1	FANK1	fibronectina tipo III y dominio de repeticiones de anquirinas 1	Núcleo	regulador de la transcripción
Fhit	FHIT	tríada de histidina frágil	Citoplasma	enzima
Filip1	FILIP1	proteína 1 de interacción con filamina A	Citoplasma	otro
Fmod	FMOD	fibromodulina	Espacio extracelular	otro
Foxe1	FOXE1	forkhead caja E1 (factor 2 de transcripción de tiroides)	Núcleo	regulador de la transcripción
Fry	FRY	homólogo a furry (Drosophila)	Espacio extracelular	otro
Gjb5	GJB5	proteína de unión a interrupciones, beta 5, 31,1 kDa	Membrana plasmática	transportador
Gpx2	GPX2	glutación peroxidasa 2 (gastrointestinal)	Citoplasma	enzima
Grxcr2	GRXCR 2	glutaredoxina, rica en cisteínas 2	Otro	otro
Hecw2	HECW2	proteína E3 ubiquitina ligasa 2 que contiene dominios HECT, C2 y WW	Espacio extracelular	enzima
Hey2	HEY2	genes de HES relacionados con el motivo 2 de YRPW	Núcleo	regulador de la transcripción
Icos	Icos	coestimulador de células T inducibles	Membrana plasmática	otro
Ifitm1	IFITM1	proteína de transmembrana 1 inducida por interferón	Membrana plasmática	receptor de transmembrana
Il1f8	IL1F8 (IL36B)	miembro de la familia de interleucina-1 (Interleucina 36 beta)	Espacio extracelular	citocina
Il28ra	IL-28RA	receptor de interleucina 28, alfa	Membrana plasmática	Receptor de citocina
Igfbp1	IGFBPL 1	tipo proteína de unión a factor de crecimiento de tipo insulina 1	Otro	otro
Ipcef1	IPCEF1	proteína de interacción para factores de intercambio de citohesina 1	Citoplasma	enzima
Lctl	Lctl	tipo lactasa	Citoplasma	otro
Ldhd	LDHD	lactato deshidrogenasa D	Citoplasma	enzima
Lef1	LEF1	factor de unión a potenciador linfoide 1	Núcleo	regulador de la transcripción
Lefty1	LEFTY1	factor de determinación de izquierda derecha 1	Espacio extracelular	factor de crecimiento
Lifr	LIFR	receptor alfa del factor inhibidor de leucemia	Membrana plasmática	receptor de transmembrana
Lpar2	LPAR2	receptor de ácido lisofosfatídico 2	Membrana plasmática	receptor acoplado a proteína G
Mog	MOG	glicoproteína de mielina de oligodendrocitos	Espacio extracelular	otro
Morn5	MORN5	que contiene repeticiones de MORN 5	Otro	otro
Pigz	NCBP2	subunidad 2 de la proteína de unión a caperuza nuclear, 20 kDa	Núcleo	otro
Nptxr	NPTXR	receptor de pentraxina neuronal	Membrana plasmática	receptor de transmembrana
Ntm	NTM	neurotrimina	Membrana plasmática	otro

ES 2 700 596 T3

Nutf2	NUTF2	factor del transporte nuclear 2	Núcleo	transportador
Ocln	OCLN	occludina	Membrana plasmática	enzima
Olr1	OLR1	receptor 1 de lipoproteínas de baja densidad oxidada (tipo lectina)	Membrana plasmática	receptor de transmembrana
Pabpc4	PABPC4	proteína de unión a poli(A), citoplasmática 4 (forma inducible)	Citoplasma	regulador de la traducción
Pde11a	PDE11A	fosfodiesterasa 11A	Citoplasma	enzima
Pdyn	PDYN	prodinorfina	Espacio extracelular	transportador
Per3	PER3	reloj del periodo circadiano 3	Núcleo	otro
Pllp	PLLP	plasmolipina	Plasma	transportador
			Membrana	
Ppp1r14c	PPP1R14 C	proteína fosfatasa 1, subunidad 14C reguladora (inhibidora)	Citoplasma	otro
Pramel6	Pramel6	antígeno preferentemente expresado en melanoma tipo 6	Otro	otro
Ptpn18	PTPN18	proteína tirosina fosfatasa, noreceptor tipo 18 (derivada de cerebro)	Núcleo	fosfatasa
Pycr1	PYCR1	pirrolina-5 -carboxilato reductasa 1	Citoplasma	enzima
Rab26	RAB26	RAB26, miembro de la familia de oncogenes RAS	Membrana plasmática	enzima
Ramp2	RAMP2	proteína 2 modificadora de la actividad del receptor (acoplado a proteína G)	Membrana plasmática	transportador
Rbm24	RBM24	proteína 24 de motivo de unión a ARN	Otro	otro
Rhag	RHAG	glicoproteína asociada a Rh	Membrana plasmática	peptidasa
Rpl3	RPL3	proteína ribosomal L3	Citoplasma	otro
Sall3	SALL3	tipo sal 3 (Drosophila)	Núcleo	otro
Satb1	SATB1	homeocaja 1 de SATB	Núcleo	regulador de la transcripción
Scg2	SCG2	secretogranina II	Espacio extracelular	citocina
Slc15a1	SLC15A 1	familia de portadores de soluto 15 (transportador de oligopéptido), miembro 1	Membrana plasmática	transportador
Slc1a1	SLC1A1	familia de portadores de soluto 1 (transportador de glutamato de alta afinidad neuronal/epitelial, sistema Xag), miembro 1	Membrana plasmática	transportador
Slc24a5	Slc24a5	familia de portadores de soluto 24 (intercambiador de sodio/potasio/calcio), miembro 5	Otro	otro
Slc37a2	SLC37A 2	familia de portadores de soluto 37 (transportador de glucosa-6-fosfato), miembro 2	Otro	transportador
40424	SNTB1	sintrofina, beta 1 (proteína asociada a distrofina A1, 59 kDa, componente básico 1)	Membrana plasmática	otro
St6galnac3	ST6GAL NAC3	ST6 (alfa-N-acetil-neuraminil-2, 3-beta-galactosil-1,3)-N-acetilgalactosaminida alfa-2,6-sialiltransferasa 3	Citoplasma	enzima

ES 2 700 596 T3

Tex12	TEX12	expresado en testículo 12	Núcleo	otro
Tex15	TEX15	expresado en testículo 15	Espacio extracelular	otro
Tfap2a	TFAP2A	factor AP-2 alfa de transcripción (proteína 2 alfa de unión a potenciador de activación)	Núcleo	regulador de la transcripción
Tmc1	TMC1	tipo canal de transmembrana 1	Membrana plasmática	otro
Tmem130	TMEM130	proteína de transmembrana 130	Otro	otro
Tmem30b	TMEM30B	proteína de transmembrana 30B	Otro	otro
Tomm20	TOMM20	translocasa de homólogo de membrana mitocondrial externa 20 (levadura)	Citoplasma	transportador
Tox3	TOX3	miembro 3 de la familia de cajas de grupo de alta movilidad TOX	Otro	otro
Ttc25	TTC25	dominio de repeticiones de tetratricopéptido 25	Citoplasma	otro
Tymp	TYMP	timidina fosforilasa	Espacio extracelular	factor de crecimiento
Ubb	Ubb	ubiquitina B	Citoplasma	otro
Vamp7	VAMP7	proteína de membrana asociada a vesícula 7	Citoplasma	transportador
Wfdc12	Wfdc12	dominio núcleo de cuatro disulfuros WAP 12	Espacio extracelular	otro
Wfdc15a	Wfdc15a	dominio núcleo de cuatro disulfuros WAP 15A	Otro	otro
Wfdc6a	Wfdc6a	dominio núcleo de cuatro disulfuros WAP 6A	Otro	otro

Tabla 10. Un subconjunto de genes de Tabla 9, que se expresan a niveles 20 veces mayores en células ES de rata que los genes correspondientes en células ES de ratón.

ID	Nombre en Entrez Gene
Ajap1	Proteína asociadas a uniones adherentes
Cldn5	Claudina 5
Arhgef9	factor 9 de intercambio de guanina de Cdc42
Camk4	proteína cinasa IV dependiente de calcio/calmodulina
Efna1	efrina-A1
Epha4	receptor A4 de EPH
Gjb5	proteína de unión a interrupciones beta 5
Igfbp1	tipo proteína de unión a factor de crecimiento de tipo insulina 1
Il1f8	Interleucina 36 beta
Il28ra	receptor de interleucina 28, alfa
Lefty1	factor de determinación de izquierda derecha 1
Lifr	receptor alfa del factor inhibidor de leucemia
Lpar2	receptor de ácido lisofosfatídico 2
Ntm	receptor de pentraxina neuronal
Ptpn18	tipo no receptor de proteína tirosina cinasa 18
Cdx2	homeocaja 2 tipo caudal

ES 2 700 596 T3

Fank1	Fibronectina tipo III y dominio de repeticiones de anquirinas 1
Foxe1	Forkhead caja E1 (factor 2 de transcripción de tiroides)
Hey2	genes de HES relacionados con el motivo 2 de YRPW
Lef1	factor de unión a potenciador linfoide 1
Sall3	tipo Sal 3 (Drosophila)
Satb1	homeocaja 1 de SATB

Se ha desarrollado, además, una firma molecular adicional empleando los marcadores/genes de pluripotencia para las células ES de rata. La Tabla 11 proporciona una lista de genes y sus rangos de expresión a partir de los datos de perfiles de ARN. El ARNm se aisló a partir de las células ES de rata y los niveles de expresión de diversos marcadores se compararon entre sí. El término "rango" significa los niveles de expresión comparativos de genes individuales: a mayor rango (1 es el más alto), mayor la expresión. Por ejemplo, el rango de Oct4's de 13 significa que, de todos los genes analizados, tuvo mayor expresión que todos excepto para 12 genes. El antecedente en este experimento fue cualquier valor de expresión por debajo de 30; 6107 genes tenían valores de expresión de 30 o mayores.

Tabla 11. Firma molecular de células ES de rata empleando diversos marcadores/genes de pluripotencia, mesodérmicos, endodérmicos, neuronales y trofectodérmicos

Pluripotencia	Rango pluripotencia	Rango de Mesodérmico	Rango mesodérmico	Endodérmico	Rango Endodérmico	Neuronal	Rango neuronal	Trofectodérmico	Rango trofectodérmico
c-Myc	8248	Brachyury	7542	Gata6	11195	Nestin	7761	Cdx2	739
Dnmi3L	127	Fkl1	No evaluado	Sox17	11418	Pax6	13570		
Dppa2	No evaluado	Nodal	3050	Hhex1	4571	Sox2	681		
Dppa5	No evaluado	Bmp4	3072	Nodal	3050				
Ecat1	9714	Bmpr2	6382	Ext1	6091				
Eras	2541			Sox7	10284				
Err-beta	1368								
Fbxo15	1369								
Fgf4	3440								
Fthl17	No evaluado								
Gdf3	2771								
Klf4	836								
Lef1	1313								
receptor de LIF	724								
Lin28	828								
Nanog	774								
Oct4	13								
Rexo1	6119								
Sox15	4524								
Sox2	681								
SSEA1	No evaluado								
SSEA4	No evaluado								
Stella	No evaluado								
Tcl1	No evaluado								
Utf1	1501								

Ejemplo 2: Inactivación de los loci genómicos en ratas

2.1: Inactivación de los loci genómicos endógenos mediante el uso de un agente endonucleasa

5 Para introducir un alelo mutante en un locus de rata endógeno genómico, las células ES de rata descritas en la presente descripción se electroporan con vectores de expresión (o ARNm) que expresan las ZFN 1 y 2 (o las TALEN 1 y 2). Estas proteínas se unen a sus secuencias objetivo en cadenas opuestas, separadas por aproximadamente 6 pb a aproximadamente 40 bp. Se forma una ruptura de la doble cadena dentro del locus objetivo, que la célula intenta reparar mediante la Unión de Extremos no Homólogos (NHEJ). En muchos casos, la NHEJ da como resultado la creación de una
 10 deleción, que frecuentemente altera la función del gen (con mayor frecuencia mediante la producción de una mutación de desplazamiento de marco). Para identificar un clon positivo que comprende un alelo mutante, las células electroporadas se cultivan en placas a baja densidad, porque no se hace la selección por fármacos. Las colonias se seleccionan y se analizan en el sitio objetivo para ver si se produjo una mutación (*por ejemplo*, mediante el uso de un ensayo de modificación de alelo (MOA) descrito anteriormente). Las células ES seleccionadas que comprenden el alelo mutante se introducen después en un embrión de rata huésped, por ejemplo, un embrión de rata en etapa de premórula o etapa de blastocisto, y se implantan en el útero de una madre sustituta para generar una rata fundadora (rata F0). Posteriormente, la rata fundadora se reproduce con una rata tipo silvestre para generar la progenie F1 heterocigota para el alelo mutante. El apareamiento de la rata F1 heterocigota puede producir la progenie homocigota para el alelo mutante.

20 2.2.: Transformación de ESC de rata para la inactivación del gen de Apolipoproteína E (ApoE) de rata mediante el uso de nucleasas de dedos de zinc

Las nucleasas de dedos de zinc usan dominios de unión a ADN modulares específicos de secuencia para dirigir la actividad endonucleasa a una única secuencia objetivo en el genoma. Las ZFN se transforman mediante ingeniería genética como un par de monómeros. Cada monómero contiene un dominio de escisión no específica de endonucleasa *FokI* fusionado a 3 o más dominios de unión a ADN de dedos de zinc. Cada dedo de zinc se une a un subsitio de 3 pb y la especificidad se logra por los sitios objetivo combinados de ambos monómeros. Las ZFN producen rupturas de la doble cadena (DSB) en el ADN, y las mutaciones (inserciones o deleciones) frecuentemente se producen durante la unión de extremos no homólogos (NHEJ). La figura 15 ilustra el mecanismo mediante el cual las endonucleasas de edición del genoma tales como las ZFN y las TALEN introducen rupturas de la doble cadena en una secuencia objetivo genómica y activan la NHEJ en una célula. Las DSB también estimulan la reparación dirigida por homología (HDR) mediante la recombinación homóloga si se proporciona una secuencia donante con ZFN.

Tales ZFN se emplearon en combinación con los varios métodos y composiciones descritos en la presente descripción para mejorar la eficiencia de transformación. El locus de Apolipoproteína E (*ApoE*) de rata se transformó como se describe en el ejemplo 3.2(a)(i), excepto que los vectores de expresión que expresan las ZFN 1 y 2 también se introdujeron en las células ES de rata. Ver la figura 11, que proporciona un esquema del evento de transformación de *ApoE* en combinación con rTZFNIP y rTZFN2P. La eficiencia de transformación se determinó como se analiza más adelante en el Ejemplo 5 y los resultados se muestran en la Tabla 12. Para tamizar para la transformación heterocigota, la transformación homocigota, y las dobles "mixtas" (por ejemplo, transformación heterocigota compuesta), se usaron cebadores y sondas específicas para determinar el genotipo. Sorprendentemente, la eficiencia de transformación aumentó 8-10 veces.

Tabla 12. ZFN para *ApoE* de rata: Eficiencia de transformación mejorada.

ADN	Colonias	Tamizadas	Transformadas heterocigotos	Transformadas homocigotos	Dobles "mixtas"	Corte/ no transformadas
vector	330	184	15 (8,2 %)	0	0	N/A
vector + ZFN 1	560	192	132 (68,8 %)	6 (3,1 %)	18 (9,4 %)	17 (8,9 %)
vector + ZFN 2	410	192	136 (70,8 %)	2 (1,0 %)	6 (3,1 %)	18 (9,4 %)

55 Se construyó un vector de direccionamiento plasmídico con un casete de selección por fármacos de autodeleción y un gen lacZ como un gen reportero (ver la figura 14 para una ilustración de los eventos de recombinación homóloga y no homóloga que pueden producirse tras la electroporación de un vector de direccionamiento que comprende un casete de selección). Se logró una buena eficiencia de transformación y se produjeron quimeras de alto %. Las nucleasas de dedos de zinc (ZFN) se analizaron, además, en combinación con vectores de direccionamiento para examinar su efecto en la mejora de la eficiencia de transformación (ver la figura 16 para una ilustración de la técnica de transformación de genes que utiliza las ZFN o las TALEN para mejorar la eficiencia de la recombinación homóloga de un vector de direccionamiento). El vector de direccionamiento se coexpresó con los vectores de expresión para 2 pares de ZFN que cortan el locus de *ApoE*. Los clones de ESC de rata electroporados con el vector de direccionamiento y un conjunto de las ZFN mostraron una eficiencia de transformación de 8-10 veces mayor que los clones de ESC de rata electroporados con un vector de direccionamiento solo. Además, se detectó la transformación homocigota bialélica en aproximadamente el 2 % de los clones. Se obtuvieron quimeras de alto % de dos de estos clones transformados.

Los clones de ESC de rata transformados en *ApoE* (con la ayuda de la ZFN) se microinyectaron en blastocistos SD, que se transfirieron después a hembras receptoras SD pseudopreñadas, mediante el uso de técnicas estándar. Las quimeras se identificaron por el color del pelaje (ver la figura 17, que muestra las quimeras ApoE-ZFN-AB5 (es decir, quimeras ApoE^{-/-}); las quimeras F0 machos se reprodujeron con hembras SD. Las crías F1 de la línea germinal se genotiparon para la presencia del alelo de *ApoE* transformado (Tabla 13). Se obtuvieron quimeras de alto % a partir de dos de estos clones transformados.

Tabla 13. Resultados de la microinyección.

		(% de quimera)
ApoE-ZFN1-AB5 (transformado homocigoto)	12	4 (90, 90, 80, 80)
ApoE-ZFN1-AE5 (transformado homocigoto)	6	3 (90, 80, 70)

Una rata inactivada en *ApoE* proporciona un medio para estudiar diversos tipos de trastornos y enfermedades. En seres humanos, la Apolipoproteína se encuentra en quilomión, HDL, LDL y VLDL. ApoE es esencial para el catabolismo normal de constituyentes de lipoproteínas ricas en triglicéridos. Los defectos en APOE da como resultado numerosos estados de la enfermedad que incluyen, por ejemplo, hipercolesterolemia familiar, hiperlipidemia, betalipoproteinemia, disbetalipoproteinemia familiar, hiperlipoproteinemia tipo III (HLP III), riesgo de enfermedad arterial coronaria. Una isoforma (ApoE4) se asocia con la enfermedad de Alzheimer familiar de aparición tardía y esporádica, posiblemente con MS también.

En ratones, ApoE se encuentra principalmente en HDL; transporta colesterol, como en seres humanos. *Los ratones deficientes de ApoE* (2 KO independientes) tienen 5 veces más colesterol en plasma normal; desarrollaron depósitos ricos en células espumosas en sus aortas proximales a la edad de 3 meses (comparable al síndrome humano).

Las inactivaciones de ApoE en ratas ofrecen un modelo animal para estudiar la función endotelial, que incluye, pero sin limitarse a, la formación de placas, los cambios transcripcionales (ARN-Seq), función *ex vivo*. Además, un tamaño mayor de ratas podría facilitar todos estos ensayos y potencialmente mejorar la calidad de los datos de ARN-Seq.

2.3. Inactivación del locus de la cadena gamma de receptor de interleucina-2 (IL2r-γ) de rata mediante el uso de nucleasas de dedos de zinc

El locus de la cadena gamma de receptor de interleucina-2 (IL2r-γ o *Il2rg*) se transformó como se describe en el Ejemplo 3.3(a), excepto que los vectores de expresión que expresan ZFN U (ZFN corriente arriba) y ZFN D (ZFN corriente abajo) también se introdujeron en las células ES de rata. La figura 18 proporciona un esquema del evento de transformación de IL2r-γ en combinación con ZFN U y ZFN D. La secuencia del locus de IL2r-γ a la cual se unen los dedos de zinc se indica en la figura 18 dentro de la sec. con núm. de ident.: 93. La eficiencia de transformación se determinó como se analiza más adelante en el Ejemplo 3.3(a) y los resultados se muestran en la Tabla 14. En resumen, los clones transformados de manera homocigota se confirmaron mediante PCR. Para el par de ZFN1: 173 clones mutantes de 192 tamizados (90 %) y para el par de ZFN2: 162 clones de 192 (84 %) tamizados.

Tabla 14. Transformación del locus de IL2r-γ de rata.

Placa	Colonias tamizadas	Transformadas	Eficiencia
7/18: Vector solo	96	4	4,2 %
7/18: Vector + ZFN	96	3	3,1 %

Los clones de ESC de rata transformados en IL2r-γ (con ayuda de ZFN) se microinyectaron en blastocistos SD, los cuales se transfirieron después a hembras receptoras SD pseudopreñadas, mediante el uso de técnicas estándar. Las quimeras se identificaron por el color del pelaje; las quimeras F0 machos se reprodujeron con hembras SD. Las crías F1 de línea germinal se genotiparon para la presencia del alelo de IL2r-γ transformado.

2.4.: Inactivación de la cadena gamma de receptor de interleucina-2 (IL2r-γ) mediante el uso de CRISPR/Cas9

El locus de IL2r-γ de rata se transformó como se describe en el Ejemplo 3.3(a), excepto que el sistema CRISPR/Cas9 se introdujo también en las células ES de rata para apoyar la eficiencia de transformación. SBI: Se emplearon vectores todo en uno de System Biosciences Cas9 "SmartNuclease" y se condujo la expresión de Cas9 por el promotor CAG, EF1a, PGK, o CMV. El ARNg personalizado se ligó en un vector y fue expresado por el promotor HI. Se diseñaron 4 ARNg contra *Il2rg*. Las regiones del locus de IL2r-γ de rata que son objetivo de los ARNg1-4 se muestran en la figura 19. Para tamizar la transformación (por ejemplo, transformación heterocigota, transformación homocigota, y transformación heterocigota compuesta), se usaron cebadores y sondas específicas para determinar el genotipo. Los resultados de la transformación

cuando se emplean diversos ARN guías se muestran en la Tabla 15. "Fuerte" y "débil" se refieren a la fortaleza de la evidencia basada en el tamizaje de que la colonia tiene una modificación dirigida.

Tabla 15. Transformación del locus de *Il2rg* de rata con los ARN guía.

Construccto(s)	ADN (ug)	Colonias	Candidatos (potencialmente transformados)
vector plasmídico de <i>Il2rg</i>	6 ug	30	3 débiles
plásmido + ARNg1 de SBI	6 ug/40 ug	22	1 fuerte, 1 débil
plásmido + ARNg2 de SBI	6 ug/40 ug	45	2 fuertes, 1 débil
plásmido + ARNg3 de SBI	6 ug/40 ug	66	1 fuerte, 2 débil
plásmido + ARNg4 de SBI	6 ug/40 ug	59	0

2.5.: Inactivación del gen de la hipoxantina guanina fosforribosil transferasa (*Hprt*) de ratón mediante el uso de CRISPR/Cas9

El locus de *Hprt* de ratón se transformó en células ES de ratón mediante el uso de los LTVEC solos o en combinación con CRISPR/Cas9. La secuencia codificante de *Hprt* completa de 32,9 kb se transformó para delección y sustitución con el casete de selección de resistencia a puromicina pCAGG-Puro, el cual también expresó eGFP. Los extremos de la delección fueron los codones de inicio y de parada. La secuencia de ARN guía usada fue 5'-GACCCGCAGUCCAGCGUCG-3' (sec. con núm. de ident.: 84), que se dirigía al exón 1 del gen de *Hprt* de ratón. La posición de la escisión predicha del sitio objetivo fue 22 pares de bases a partir del extremo 5' de la delección. La eficiencia de escisión en el objetivo de Cas9/ARNg observada en las células ES fue $\geq 93\%$. Un resumen se muestra en la Tabla 16. El uso de CRISPR/Cas9 para apoyar la transformación del locus de *Hprt* de 32,9 kb completo dio como resultado una potenciación de cinco veces de la transformación sobre el uso de LTVEC solo.

Tabla 16. Resumen de la delección asistida por CRISPR del gen *Hprt*

Resumen de la delección asistida por CRISPR del gen <i>Hprt</i>						
Gen objetivo	Delección (kb)	Brazo de homología 5' (kb)	Brazo de homología 3' (kb)	Eficiencia de transformación (%)		Veces de mejoramiento
				LTVEC solo	LTVEC + CRISPR/Cas9	
<i>Hprt</i>	32,9	88	66	5,0	25,4	5,1

Ejemplo 3: Modificación dirigida de los loci genómicos de rata

3.1: Transformación de ESC de rata: El locus de *Rosa26* de rata.

El locus de *Rosa26* de rata se encuentra entre los genes *Setd5* y *Thumpd3* como en ratón, con el mismo espaciado. El locus de *Rosa26* de rata (Figura 12, Panel B) difiere del locus de *Rosa26* de ratón (Figura 12, Panel A). Los transcritos de *Rosa26* de ratón consisten en 2 o 3 exones. El locus de rata contiene un 2do exón 1 (Ex1b) además del exón homólogo al exón 1 de ratón (Ex1a). No se identificó un 3er exón en rata. La transformación de un alelo de *Rosa26* de rata se describe en la figura 12C, donde los brazos de homología de 5 kb se clonaron cada uno mediante PCR mediante el uso de ADN genómico de la ESC de rata DA. El alelo transformado contiene un casete SA (aceptor de corte y empalme)-lacZ-hUbe-neo que reemplaza una delección de 117 pb en el intrón de *Rosa26* de rata.

Se determinó la eficiencia de transformación en el locus de *Rosa26* de rata (Tabla 17). El vector linealizado se electroporó en las ESC de rata DA o ACI, y las colonias transfectadas se cultivaron en medio 2i + G418, mediante el uso de técnicas estándar. Las colonias individuales se seleccionaron y se tamizaron mediante el uso de un ensayo de Pérdida de Alelo (LOA) (Valenzuela, D. y otros. (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, Nature Biotech. 21:652-660.

Tabla 17. Eficiencia de transformación de Rosa26 de rata

Línea celular	Colonias seleccionadas	Positivos vueltos a confirmar	Eficiencia de transformación (%)
DA.2B	192	4	2,1
ACI.G1	96	4	4,2

5

10 Producción de quimeras y transmisión a través de la línea germinal mediante el uso de clones de ESC de rata transformados en Rosa26. Los clones de ESC de rata transformados en Rosa26 se microinyectaron en blastocistos SD, los que se transfirieron después a hembras receptoras SD pseudopreñadas, mediante el uso de técnicas estándar. Las quimeras se identificaron por el color del pelaje; las quimeras F0 machos se reprodujeron con hembras SD. Las crías F1 de la línea germinal (agutí) se genotiparon para la presencia del alelo de Rosa26 transformado; nueve de 22 crías agutí se genotiparon como heterocigotas en el locus de Rosa26 (Tabla 18).

15

Tabla 18. Transmisión a través de la línea germinal mediante el uso de rESC con Rosa26 transformado

Línea celular	clones de R26 inyectados	Clones que producen quimeras	Clones que se transmiten a través de la línea germinal	Crías totales	Crías derivadas de rESC	Crías derivadas de ESC (%)
DA.2B	4	3	2	AH7: 64	AH7:41	AH7: 63
				AE3: 112	AE3:6	AE3:3
ACI.G1	4	4	1	DE9: 39	DE9: 4	10

20

25

30

35 Para confirmar que el alelo genéticamente modificado en el locus de Rosa26 se transmitió a través de la línea germinal, la expresión de lacZ se confirmó mediante la tinción con X-gal en ratas transformadas en Rosa26. La tinción con X-gal del cerebro, corazón y el timo, y un pulmón de rata transformada en Rosa26- heterocigota de 14 semanas de edad demostró la expresión de lacZ (Figura 13B, D, y F, respectivamente), mientras que los controles de tipo silvestre de la misma edad demostraron un bajo nivel de tinción con X-gal de fondo (Figura 13A, C, y E, respectivamente). La tinción con X-gal en embriones de rata transformada en Rosa26 heterocigotas E12.5 y E 14.5 mostraron expresión ubicua de lacZ (Figura 13G e I, respectivamente), mientras que los embriones de rata control mostraron bajos niveles de tinción con X-gal de fondo (Figura 13H y J, respectivamente).

40

3.2.(a)(i) : Transformación del locus de Apolipoproteína E (ApoE) de rata.

45 El locus de Apolipoproteína E (*ApoE*) de rata se transformó para alterar la función de ApoE. La transformación del locus de *ApoE* se hizo mediante el uso de un vector de direccionamiento que comprende un casete lacZ-hUb-neo flanqueado por brazos de homología 5' y 3' homólogos al locus de *ApoE*. La figura 20 representa un locus de *ApoE* de rata genéticamente modificado que se ha alterado por una delección de 1,8 kb y la inserción de un casete lacZ-hUb-neo, que incluye, además, un casete Cre de autodelección que comprende un gen de Crei conducido por un promotor de protamina. Las condiciones de electroporación fueron las siguientes: 6 ug de ADN; 2,05 x 10⁶ de células; 400 V; 200 uF: 342 V, 593 usec; placa en MEF neoR densas 2x de 15 cm en 2i + ROCKi 10 uM.

50

55 La eficiencia de transformación en el locus de *ApoE* se determinó y se muestra en la Tabla 19. El vector linealizado se electroporó en las ESC de rata DA.2B derivadas de la cepa DA, y las colonias transfectadas se cultivaron mediante el uso de técnicas estándar. Las colonias individuales se seleccionaron y se tamizaron mediante el uso de un ensayo de pérdida de alelo (LOA).

55

Tabla 19. Eficiencia de transformación de ApoE de rata

Línea celular	Vector	Colonias seleccionadas	Transformadas	Eficiencia de transformación (%)
DA.2B	ApoE-mSDC	192	7	3,7
DA.2B	ApoE-mSDC	192	15	7,8

60

65

Se realizó la producción de quimeras y la transmisión a través de la línea germinal mediante el uso de clones de ESC de rata transformados en *ApoE*. Los clones de ESC de rata transformados en *ApoE* se microinyectaron en blastocistos SD, los que se transfirieron después a hembras receptoras SD pseudopreñadas, mediante el uso de técnicas estándar. Las quimeras se identificaron por el color del pelaje; las quimeras F0 machos se reprodujeron con hembras SD. Se logró la transmisión a través de la línea germinal. Las crías F1 se genotiparon para la presencia del alelo de *ApoE* transformado (Tabla 20).

Tabla 20 Resultados de las microinyecciones

Exp	Clon	Crías	Quimeras (% de quimera)
1	ApoE-AF5	4	3 (90, 90, 90)
2	ApoE-BC4	5	0

La expresión de LacZ conducida por el promotor de *ApoE* endógeno se confirmó mediante la tinción con X-gal en ratas hembras *ApoE*^{+/-} de 12 semanas de edad en el cerebro, vasos sanguíneos, e hígado (Figuras 43-45, respectivamente). Las figuras 43-45 muestran un patrón de expresión para lacZ que refleja el patrón de expresión de *ApoE* endógeno. Los controles de tipo silvestre de la misma edad mostraron un bajo nivel de tinción con X-gal de fondo.

Se estudiaron adicionalmente los fenotipos de ratas con *ApoE* deleciónado. Se realizaron estudios químicos longitudinales en suero para medir los niveles de colesterol, LDL, HDL, y triglicéridos a intervalos de tres semanas. La figura 46A-D muestra los niveles de colesterol, LDL, HDL, y triglicéridos en suero en ratas transformadas homocigotas, transformadas heterocigotas, y de tipo silvestres a las 6 semanas, 9 semanas, 12 semanas, y 15 semanas de edad. Las hemorragias oculares se realizaron en una cohorte de la misma edad que consistía en 2 ratas de tipo silvestre, 7 heterocigotas y 8 homocigotas. No se observaron diferencias significativas entre machos y hembras. Las ratas con *ApoE* deleciónado homocigotas mostraron niveles de colesterol y de LDL elevados y niveles de HDL disminuidos. A diferencia de los ratones *ApoE*^{-/-}, no se observó aumento significativo en los triglicéridos en ratas con *ApoE* deleciónado.

El análisis fenotípico adicional que se realiza incluye la histología/imagenología *ex vivo* para la formación de placas del arco aórtico, imagenología *in vivo* para la formación de placas del arco aórtico, y cambios transcripcionales (secuenciación por fragmentación aleatoria del transcriptoma completo (ARN-Seq)) del endotelio del arco aórtico. El tiempo de estos ensayos depende de la línea de tiempo de la formación de placa. Las placas son detectables en ratones *ApoE*^{-/-} a las 24 semanas.

Los datos de transformación adicionales para *ApoE* se proporcionan, además, en la Tabla 22.

3.2.(a)(ii). Transformación de *ApoE* en ratas con un vector de direccionamiento

La figura 20 proporciona un esquema del locus de *ApoE* de rata y un plásmido de transformación. El esquema superior de la figura 20 muestra la estructura genómica del locus de *ApoE* de rata y las regiones genómicas que corresponden a los brazos de homología 5' y 3' (5 kb y 5,4 kb, respectivamente; recuadros gris oscuro). El exón 1 de *ApoE* es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los 3 intrones de *ApoE* se indican como líneas y los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco.

El esquema inferior en la figura 20 es el vector de direccionamiento. Los brazos de homología 5' y 3' (5 kb y 5,4 kb respectivamente) se indican por los recuadros gris oscuro. El vector de direccionamiento comprende un gen reportero (*lacZ*) y un casete de autodeleción flanqueado por sitios loxP (flechas en blanco). El casete de autodeleción comprende el gen de Crei unido operativamente a un promotor *Prr1* de ratón y un casete de selección que comprende un gen de resistencia a neomicina unido operativamente a un promotor de ubiquitina humano.

El gen de Crei comprende dos exones que codifican una Cre recombinasa, que están separados por un intrón (Crei) para evitar su expresión en una célula procariota. Ver, por ejemplo, la patente de los Estados Unidos 8,697,851 y la publicación de solicitud de patente de los Estados Unidos 2013-0312129, que describen el casete de autodeleción en detalle. Mediante el empleo del promotor *Prr1*, el casete de autodeleción puede deleccionarse específicamente en células germinales machos de ratas F0. El vector de direccionamiento se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas en MEF resistentes a neomicina densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron, se seleccionaron, y se mantuvieron como se describe en el Ejemplo 1.

Como se muestra en la Tabla 44, se tamizaron 384 colonias y se obtuvieron 23 clones transformados. La eficiencia de transformación fue de 5,99 %. Se inyectaron 3 clones en blastocistos como se describe en la presente descripción en el Ejemplo 1. Se obtuvieron 3 clones que producen quimeras y 1 de los clones transmitió la modificación dirigida a través de la línea germinal.

3.2.(a)(iii). Transformación de ApoE en ratas con un vector de direccionamiento en combinación con nucleasas de dedos de zinc

El vector de direccionamiento empleado en el Ejemplo 3.2(a)(ii) se usó en combinación con nucleasas de dedos de zinc para dirigirse al locus de ApoE de rata. La Tabla 21 proporciona un resumen de la organización genómica del locus de ApoE de rata. Las posiciones mostradas en la Tabla 21 se tomaron de la construcción 5.0 de la secuencia de referencia del genoma de rata (ENSEMBL). ApoE está en el cromosoma 1 en la cadena (-).

Tabla 21. Resumen del locus de ApoE de rata y las posiciones de los sitios de unión y los sitios de corte de las nucleasas de dedos de zinc.

Característica	Inicio	Final	longitud	Notas
Exón 1	81881110	81881182	73	5' no codificante
Exón 2	81880269	81880332	64	contiene ATG
ATG	81880309	81880311	3	codón de inicio
Exón 3	81879607	81879775	169	
sitio de unión a ZFN1a	81879707	81879693	15	CAGGCCCTGAACCGC (sec. con núm. de ident.: 10)
sitio de corte de ZFN1	81879692	81879687	6	TTCTGG (sec. con núm. de ident.: 11)
sitio de unión a ZFN1b	81879686	81879671	16	GATTACCTGCGCTGGG (sec. con núm. de ident.: 12)
Intrón 3-4	81879776	81879207	400	
sitio de unión a ZF21a	81879591	81879577	15	TTCACCCTCCGCACC (sec. con núm. de ident.: 13)
sitio de corte de ZFN2	81879576	81879570	7	TGCTGAG (sec. con núm. de ident.: 14)
sitio de unión a ZF21b	81879569	81879552	18	TATCCAGATCCAGGGGTT (sec. con núm. de ident.: 15)
Exón 4	81878371	81879208	838	contiene TGA
TGA	81878482	81878484	3	
deleción de ApoE	81878482	81880311	1830	

La figura 11 proporciona un esquema del locus de ApoE de rata y se indica con barras grises el sitio de corte de ZFN1 y ZFN2. El sitio de corte de ZFN1 está en el exón 3 y el sitio de corte de ZFN2 está en el intrón 3. La posición exacta de los dos sitios de ZFN se expone en la Tabla 21. Las regiones genómicas que corresponden a los brazos de homología 5' y 3' (5 kb y 5,4 kb, respectivamente) se indican por los recuadros gris oscuro. El exón 1 de ApoE es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los tres intrones del gen de ApoE se indican como líneas y los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco.

El vector de direccionamiento empleado fue el mismo que el del Ejemplo 3.2(a)(ii) y se muestra en la figura 20, y la figura 21A proporciona un esquema de la transformación del locus de ApoE en células ES de rata mediante el uso de nucleasas de dedos de zinc y el vector de direccionamiento descrito en la figura 20. Las ZFN se introdujeron como dos plásmidos de expresión, uno para cada mitad del par de ZFN. Se usaron 20 ug del plásmido para ZFN1 y 20 ug del plásmido para ZFN2. Las ZFN se compraron de Sigma. La expresión de cada ZFN fue conducida por el promotor de CMV.

El vector de direccionamiento se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas en MEF neoR densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron, se seleccionaron y se mantuvieron como se describe en el Ejemplo 1.

Como se muestra en la Tabla 22 y la Tabla 44, se tamizaron 384 colonias y se obtuvieron 290 clones transformados. La eficiencia de transformación fue de 75,52 %. Se inyectaron 2 clones en blastocistos como se describe en la presente descripción en el Ejemplo 1. Se obtuvieron dos clones que producen quimeras y uno de los clones transmitió la modificación dirigida a través de la línea germinal.

Además, el empleo de ZFN1 y ZFN2 produjo 8 clones transformados bialélicos con una eficiencia del 2,08 %.

Tabla 22. Transformación del locus de *ApoE*.

ADN	Transformados heterocigotos	Transformados homocigotos	Microinyectados	Quimeras (% de quimerismo)	Reproducción
vector solo	15/192 (8 %)	0			
vector + par 1 de ZFN	156/192 (81 %)	6/192 (3 %)	2 clones	7 (70-90 %)	
vector + par 2 de ZFN	134/192 (70 %)	2/192 (1 %)			

3.2.(b)(i): Modificación dirigida del locus de Apolipoproteína E (*ApoE*) de rata mediante el uso de un vector de direccionamiento grande (LTC).

La transformación del locus de *ApoE* se hace mediante el uso de un vector de direccionamiento grande (LTVEC) que comprende un casete *lacZ-Prm1* de ratón-Crei flanqueado por un brazo de homología 5' con respecto al locus de *ApoE* de aproximadamente 45 kb y un brazo de homología 3' con respecto al locus de *ApoE* de aproximadamente 23 Kb. La figura 22 representa el locus de *ApoE* de rata en el que el locus de *ApoE* se ha alterado con una delección de 1,83 kb y la inserción del gen *lacZ* y un casete de autodelección que comprende el casete *mPrm1-Crei* y un casete de selección hUneo. Los métodos empleados en el Ejemplo 3.2(a)(i) pueden usarse para introducir este vector en células ES de rata.

Ejemplo 3.2.(b)(ii). Transformación del locus de *ApoE* de rata con un vector de direccionamiento grande (LTVEC)

La figura 22 proporciona un esquema del locus de *ApoE* de rata y un vector de direccionamiento grande (LTVEC). El esquema superior de la figura 22 muestra la organización genómica del locus de *ApoE* de rata y las regiones genómicas que corresponden a los brazos de homología 5' y 3' (45 kb y 23 kb, respectivamente; recuadros gris oscuro). El exón 1 de *ApoE* es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los 3 intrones de *ApoE* se indican como líneas y los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco.

El esquema inferior en la figura 22 es el LTVEC. Los brazos de homología 5' y 3' (45 kb y 23 kb, respectivamente) se indican por los recuadros gris oscuro. El vector de direccionamiento comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por sitios *loxP* (flechas en blanco), que comprende el gen de Crei unido operativamente a un promotor *Prm1* de ratón y un casete de selección por fármacos que comprende un gen de resistencia a neomicina unido operativamente a un promotor de ubiquitina humano. El Crei comprende dos exones que codifican la Cre recombinasa los que están separados por un intrón (Crei) para evitar su expresión en una célula procariota. Ver, por ejemplo, la patente de los Estados Unidos 8,697,851 y la publicación de solicitud de patente de los Estados Unidos 2013-0312129, que describen el casete de autodelección con detalle. Mediante el empleo de un promotor *Prm1* de ratón, el casete de autodelección puede deleccionarse específicamente en células germinales masculinas de rata F0.

El LTVEC se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas en MEF neoR densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron, se seleccionaron, y se mantuvieron como se describe en el Ejemplo 1.

Como se muestra en la Tabla 44, se tamizaron 288 colonias y se obtuvieron 8 clones transformados. La eficiencia de transformación fue de 2,78 %. Se inyectaron 3 clones en un embrión huésped en una etapa de blastocisto como se describe en la presente descripción en el Ejemplo 2 para producir ratas quiméricas (F0). Además, se produjo un clon transformado bialélico proporcionando una eficiencia bialélica del 0,35 %.

3.2.(b)(iii). Transformación de *ApoE* en ratas con un vector de direccionamiento grande (LTVEC) en combinación con nucleasas de dedos de zinc

El LTVEC empleado en el Ejemplo 3.2.(b)(ii) se usó en combinación con nucleasas de dedos de zinc para dirigirse al locus de *ApoE* de rata. La Tabla 21 proporciona un resumen de la organización genómica del locus de *ApoE* de rata y las posiciones mostradas se tomaron de la construcción 5.0 de la secuencia de referencia del genoma de rata (ENSMBL).

La figura 23 proporciona un esquema del locus de *ApoE* de rata y se indica con barras grises el sitio de corte de ZFN1 y ZFN2. El sitio de corte para ZFN1 está en el exón 3 y el sitio de corte para ZFN2 está en el intrón 3. La posición exacta

de los dos sitios de ZFN se expone en la Tabla 21. Los brazos de homología 5' y 3' (45 kb y 23 kb, respectivamente) se indican por los recuadros gris oscuro. El exón 1 del gen de *ApoE* es no codificante y se muestra como un recuadro en blanco más cercano al brazo de homología 5'. Los tres intrones del gen de *ApoE* se indican como líneas. Los exones 2 y 3 comprenden regiones codificantes y se muestran como recuadros grises punteados. El exón 4 contiene tanto las secuencias codificantes como las no codificantes según se indica por el sombreado gris punteado y el recuadro en blanco.

El LTVEC empleado fue el mismo que el del Ejemplo 3.2(b)(ii) y mostrado en la figura 22. Las ZFN se introdujeron como dos plásmidos de expresión, uno para cada mitad del par de ZFN. Se usaron 20 ug del plásmido para ZFN 1 y 20 ug del plásmido para ZFN2. Las ZFN se compraron de Sigma. La expresión de cada ZFN fue conducida por el promotor de CMV.

El vector de direccionamiento se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas en MEF neoR densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron, se seleccionaron, y se mantuvieron como se describe en el Ejemplo 1.

Como se muestra en la Tabla 44, se tamizaron 288 colonias y se obtuvieron 16 clones transformados. La eficiencia de transformación fue de 5,56 %. Se inyectó un clon en blastocistos como se describe en la presente descripción en el Ejemplo 2.

Además, el empleo de ZFN1 y ZFN2 produjo un clon transformado bialélico, con una eficiencia de 0,35 %.

3.2.(b)(iv). Transformación de *ApoE* en ratas con un vector de direccionamiento grande (LTVEC) en combinación con CRISPR/Cas9

El LTVEC empleado en el Ejemplo 3.2.(b)(ii) se usó en combinación con CRISPR/Cas9 para dirigirse al locus de *ApoE* de rata. La Tabla 23 muestra una comparación de los resultados de los experimentos en los que se usó el LTVEC de *ApoE* solo para dirigirse al locus de *ApoE* de rata o se usó en combinación con un CRISPR/nucleasa Cas9 para dirigirse al locus de *ApoE* de rata. En cada experimento, las células electroporadas se sembraron en placas a una alta densidad y se sometieron a selección por fármacos para encontrar colonias que eran resistentes a fármacos. Las colonias resistentes a fármacos se seleccionaron y se tamizaron para la modificación dirigida mediante el uso del ensayo de modificación de alelo (MOA) como se describe en la presente descripción. Específicamente, 4×10^6 células se electroporaron con 2 ug de *ApoELTVEC* a una tensión de 400 V, una capacitancia de 100 uF, y una resistencia de 0. En el último experimento, 6 ug de plásmido de expresión de Cas9 y 3 ug de ARNg2 de *ApoE* o 3 ug de ARNg3 de *ApoE* también se electroporaron. La selección se hizo mediante el uso de 75 ug/mL de G418. El ARNg2 de *ApoE* tiene una secuencia de GCAGGCCCTGAACCGCTTCTTGG (sec. con núm. de ident.: 87) y se dirige a una región de 67 pb 3' del inicio del exón 3 de *ApoE* de rata. El ARNg3 de *ApoE* tiene una secuencia de CCTGCGCTGGGTGCAGACGCTTT (sec. con núm. de ident.: 88) y se dirige a una región de 97 pb 3' del inicio del exón 3 de *ApoE* de rata (ver la figura 47). Como se muestra en la Tabla 23, cuando Cas9 y cualquiera de los ARNg se introdujeron en las células junto con el LTVEC de *ApoE*, la eficiencia de transformación aumentó (del 43 % al 53 % o el 47 %). La transformación bialélica se observó en cinco colonias transformadas con el LTVEC de *ApoE* en combinación con ARNg2 o 3 de *ApoE*, pero no se observó transformación bialélica con el LTVEC de *ApoE* solo.

Tabla 23. Comparación de la transformación de LTVEC de *Rag2* con y sin CRISPR/Cas9

Vector	Cas9	ARNg	Colonias tamizadas	Clones transformados	Transformados bialélicos	Eficiencia de transformación
LTVEC de <i>ApoE</i>	No	No	75	32	0	43 %
LTVEC de <i>ApoE</i>	Si	ARNg2 de <i>ApoE</i>	80	42	1	53 %
LTVEC de <i>ApoE</i>	Si	ARNg3 de <i>ApoE</i>	89	42	4	47 %

3.3(a): Transformación del locus de la cadena gamma de receptor de interleucina-2 (IL2r-γ) de rata

Se transformó el locus de la cadena gamma de receptor de interleucina-2 (IL2r-γ o *Il2rg*) de rata para alterar la función de IL2r-γ. IL2r-γ juega un papel importante en la señalización de IL-2, IL-4, IL-7, IL-9, IL-15, IL-21 y las mutaciones en IL2r-γ se asocian con graves defectos en el desarrollo de células T, B y NK.

La transformación del locus de IL2r-γ se hizo mediante el uso de un vector de direccionamiento que comprende un casete eGFP-hUb-neo flanqueado por los brazos de homología 5' y 3' homólogos al locus de IL2r-γ, como se representa en la figura 24. La figura 25 representa la estructura genómica del locus de IL2r-γ de rata en la cual se ha alterado el locus de IL2r-γ locus por una delección de 3,2 kb. El locus de IL2r-γ objetivo también comprendía un gen de eGFP y un casete de autodelección que contiene Crei unido operativamente a un promotor de Protamina1 de ratón y un casete de selección por fármacos que comprende un promotor hUb unido operativamente a un gen de resistencia a neomicina.

La eficiencia de transformación en el locus de IL2r- γ se determinó y se muestra en la Tabla 24. El vector linealizado se electroporó en las ESC DA.2B de rata, y las colonias transfectadas se cultivaron mediante el uso de técnicas estándar. Las colonias individuales se seleccionaron y se tamizaron mediante el uso de un ensayo de pérdida de alelo (LOA).

Línea celular	Vector	Colonias seleccionadas	Transformadas	Eficiencia de transformación (%)	Quimeras (% de quimerismo)
DA.2B	Il2rg-floxed neo	136	1	0,7	5 (70-90 %)
DA.2B	Il2rg-mSDC	96	4	4,2	

Se realizó la producción de quimeras y la transmisión a través de la línea germinal mediante el uso de clones de ESC de rata transformados en IL2r- γ . Los clones de ESC de rata transformados en IL2r- γ se microinyectaron en blastocistos SD, los que se transfirieron después a hembras receptoras SD pseudopreñadas, mediante el uso de técnicas estándar. Las quimeras se identificaron por el color del pelaje; las quimeras F0 machos se reprodujeron con hembras SD. Las crías F1 de la línea germinal se genotiparon para la presencia del alelo de IL2r- γ transformado (Tabla 25). En otro experimento de microinyección con el clon Il2rg-CG12, la transmisión a través de la línea germinal se confirmó, además, por los colores del pelaje y el genotipaje.

Tabla 25. Resultados de la microinyección

Exp	Clon	crías	Quimeras (% de quimera)
1	Il2rg-AA1	5	2 (90, 70)
2	Il2rg-AA1	10	3 (90, 90, 80)
3	Il2rg-CG12	11	7 (95, 90, 90, 90, 80, 80, 80)

El fenotipo de la quimera Il2rg^{-Y} #3 se estudió más aún. Las células mononucleares de sangre periférica (PBMC) se tiñeron con anticuerpos que reconocen antígenos en varios linajes linfoides. Las PBMC positivas a GFP se detectaron de 2 de las quimeras, como se muestra en la figura 30. Además, las células GFP+ fueron negativas para el marcador de células T CD3 (figura 29A), y fueron principalmente negativas para el marcador de células B B220 y el marcador de células NK CD161a (figura 29B y C, respectivamente). Las PBMC de una rata de tipo silvestre se usaron como controles negativos para la expresión de GFP. Ver las figuras 29D-F. Las poblaciones doble positivas pequeñas son consistentes con el fenotipo de inactivación de Il2rg publicado en ratones. Estos datos se obtuvieron de una rata quimérica, que contiene células positivas a la cadena gamma de receptor de IL2, y esto puede complicar el análisis del fenotipo. El análisis de citometría de flujo puede realizarse, además, en poblaciones celulares de médula ósea y bazo para revelar las disminuciones correspondientes en el número de linfocitos. Ver Mashimo y otros. (2010) PLoS One 5(1):e8870.

3.3(b): Modificación dirigida del locus de la cadena gamma de receptor de interleucina-2 (IL2r- γ) de rata

El locus de la cadena gamma de Interleucina-2 (IL2r- γ) de rata se transformó para alterar la función de IL2r- γ en ratas. La figura 25 muestra la estructura genómica del locus de Il2rg de rata (panel superior de la figura 25) y el vector de direccionamiento introducido en el locus (panel inferior de la figura 25). Se escogió eGFP como un reportero para que el inmunofenotipo de las ratas genéticamente modificadas pudiera examinarse mediante el uso de FACS. El casete de autodelección (hUb-Neo; Prr1-Cre) se usó para delecionar el casete de selección por fármacos y el gen de Cre específicamente en células germinales masculinas de la rata F0. Además, el vector de direccionamiento se diseñó para delecionar la región codificante completa (aproximadamente 3,2 kb) del gen de Il2rg rata.

El tamaño de la delección en las ESC de rata se confirmó mediante PCR mediante el uso de cebadores específicos al locus de Il2rg de rata. Tras la microinyección de los clones transformados en embriones huéspedes en una etapa de blastocisto, se obtuvieron altos porcentajes de quimeras. Estas quimeras se han establecido para la reproducción. Para determinar si la transformación funcionó como se esperaba, la sangre periférica de las quimeras se recolectó antes de la reproducción, y el fenotipo de las células del sistema inmunitario en la sangre periférica se analizó a través de FACS. Como se muestra en la figura 30, las células positivas a GFP se detectaron en la sangre periférica en 2 de las 3 quimeras examinadas, y las ratas quiméricas contenían menos del 1 % de células T, menos del 1 % de células B, y menos del 1 % de células NK, que son positivas para GFP (es decir, células KO de Il2rg) (Figura 29A-C).

3.4(a)(i). Transformación del locus de Rag2 en ratas con un vector de direccionamiento grande (LTVEC)

La Tabla 26 proporciona un resumen de la organización genómica del locus de *Rag2* de rata y las posiciones mostradas se tomaron de la construcción 5.0 de la secuencia de referencia del genoma de la rata (ENSMBL). *Rag2* se encuentra en el cromosoma 3 en la cadena (+).

5 Tabla 26. Resumen de la organización genómica del locus de *Rag2* de rata.

Característica	Inicio	Final	Longitud	Notas
Exón 1	97,851,317	97,851,448	132	
Exón 2	97,854,635	97,854,693	59	
Exón 3	97,858,260	97,859,615	1,356	contiene la secuencia codificante completa
ATG	97,856,286	97,856,288	3	codón de inicio
TGA	97,857,867	97,857,869	3	codón de parada
delección de <i>Rag2</i>	97,856,289	97,859,784	3,496	

25 La figura 26 proporciona un esquema del locus de *Rag2* de rata y un vector de direccionamiento grande (LTVEC). El LTVEC es de 140 kb y se dirige a una porción de aproximadamente 5,7 kb del locus de *Rag2* de rata para la delección. El esquema superior de la figura 26 muestra la organización genómica del locus de *ApoE* de rata y las regiones genómicas que corresponden a los brazos de homología 5' y 3' (48 kb y 84 kb, respectivamente; recuadros gris oscuro). *Rag2* comprende un único exón indicado por el sombreado gris punteado.

30 El esquema inferior en la figura 26 es el LTVEC. Los brazos de homología 5' y 3' (48 kb y 84 kb, respectivamente) se indican por los recuadros gris oscuro. El LTVEC comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por sitios *loxP* (flechas en blanco). El casete de autodelección comprende un promotor *Prm1* de ratón unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina. Se generó otra versión del LTVEC en la cual el gen de resistencia a neomicina se sustituyó con un gen de resistencia a higromicina para permitir la retransformación de las células ES de rata transformadas en *I2rg*. El Crei comprende dos exones que codifican la Cre recombinasa que están separados por un intrón (Crei) para evitar su expresión en una célula procarionota. Ver, por ejemplo, la patente de los Estados Unidos 8,697,851 y la publicación de solicitud de patente de los Estados Unidos 2013-0312129, que describen el casete de autodelección. Mediante el empleo de un promotor *Prm1* de ratón, el casete de autodelección puede deleccionarse específicamente en células germinales masculinas de ratas F0.

45 El LTVEC se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas en MEF neoR densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron y se mantuvieron como se describe en el Ejemplo 1.

Las colonias se tamizaron como se describe en cualquier otra parte de la presente descripción y se obtuvieron los clones transformados. Los clones transformados se inyectan después en un embrión huésped como se describe en cualquier otra parte de la presente descripción para producir una rata F0.

50 3.4(a)(ii). Transformación del locus de *Rag2* en ratas con un vector de direccionamiento grande (LTVEC) y CRISPR/Cas9

La Tabla 27 muestra una comparación de los resultados de experimentos en los cuales se usó una versión del LTVEC de *Rag2* que tiene un gen de resistencia a higromicina (ver la figura 48) sola para dirigirse al locus de *Rag2* de rata o se usó en combinación con una CRISPR/nucleasa Cas9 para dirigirse al locus de *Rag2* de rata. En cada experimento, las células electroporadas se sembraron en placas a una alta densidad y se sometieron a selección por fármacos para encontrar colonias que eran resistentes a fármacos. Las colonias resistentes a fármacos se seleccionaron y se tamizaron para la modificación dirigida mediante el uso del ensayo de modificación de alelo (MOA) como se describe en la presente descripción. Específicamente, 4×10^6 células se electroporaron con 2 ug de LTVEC de *Rag2* a una tensión de 400 V, una capacitancia de 100 uF, y una resistencia de 0. En el último experimento, se electroporaron, además, 6 ug de plásmido de expresión de Cas9 y 3 ug de ARNg1 de *Rag2* o 3 ug de ARNg4 *Rag2*. La selección se hizo mediante el uso de 75 ug/mL de G418. El ARNg1 de *Rag2* tiene una secuencia de CCAGCTACTTGCTCGTACAA (sec. con núm. de ident.: 89) y se dirige a una región de 219 pb 3' del codón de inicio de *Rag2* de rata (ATG). El ARNg4 de *Rag2* tiene una secuencia de CCCCTCAGATTCACGTGCGT (sec. con núm. de ident.: 90) y se dirige a una región de 12 pb 3' del codón de parada de *Rag2* (TAG) (ver la figura 48). Como se muestra en la Tabla 27, cuando se introdujeron Cas9 y cualquiera de los ARNg en las células junto con el LTVEC de *Rag2*, la eficiencia de transformación aumentó (de 0 al 10 % o 38 %). La transformación bialélica se observó en una colonia.

Tabla 27. Comparación de la transformación de LTVEC de Rag2 con y sin CRISPR/Cas9

Vector	Cas9	ARNg	Colonias tamizadas	Clones transformados	Transformados bialélicos	Eficiencia de transformación
LTVEC de Rag2	No	No	36	0	0	0
LTVEC de Rag2	Si	ARNg1 de Rag2	23	5	1	22 %
LTVEC de Rag2	Si	ARNg4 de Rag2	16	1	0	6 %

3.4.(b)(i): Transformación del locus de Rag1 y el de Rag 2 en ratas

La figura 27 proporciona la estructura genómica del locus de *Rag1/Rag2* de rata. CDS indica la secuencia codificante y los recuadros grises representan exones. *Rag2* se encuentra en la cadena "más" con transcripción a la derecha. *Rag1* se encuentra en la cadena "menos" con transcripción a la izquierda. Mbp = millones de pares de bases.

La Tabla 28 proporciona un resumen de la organización genómica del locus de *Rag2* y *Rag1* de rata y las posiciones mostradas se toman de la construcción 5.0 de la secuencia de referencia del genoma de la rata (ENSEMBL). *Rag1* se encuentra en el cromosoma 3 en la cadena (-).

Tabla 28. Resumen de la organización genómica del locus de *Rag1* de rata.

Característica	Inicio	Final	Longitud	Notas
Exón 1	97,877,145	97,877,066	80	
Exón 2	97,872,503	97,866,047	6,457	contiene la secuencia codificante completa
ATG	97,872,489	97,872,487	3	codón de inicio
TAA	97,869,369	97,869,367	3	codón de parada
delección de Rag1-2	97,856,289	97,872,486	16,198	

La figura 28 proporciona un esquema del locus de *Rag2* y *Rag1* de rata y un vector de direccionamiento grande (LTVEC). El LTVEC es de aproximadamente 70 kb y se dirige a un locus genómico de rata de aproximadamente 16,6 kb que comprende los loci de *Rag1* y *Rag2* para delección. El esquema superior de la figura 28 muestra la organización genómica de los loci de *Rag1* y *Rag2* y las regiones genómicas que corresponden a los brazos de homología 5' y 3' (48 kb y 15 kb, respectivamente; recuadros gris oscuro). Cada uno de *Rag2* y *Rag1* comprende un único exón indicado por el sombreado gris punteado. El esquema inferior en la figura 28 es el LTVEC. Los brazos de homología 5' y 3' (48 kb y 15 kb, respectivamente) se indican por los recuadros gris oscuro. El LTVEC comprende un gen reportero (*lacZ*) y un casete de autodelección flanqueado por sitios loxP (flechas en blanco). El casete de autodelección comprende un promotor *P_{rm1}* de rata unido operativamente al gen de Crei y un casete de selección por fármacos que comprende un promotor de ubiquitina humano unido operativamente a un gen de resistencia a neomicina. Se generó otra versión del LTVEC en la cual el gen de resistencia a neomicina se sustituyó con un gen de resistencia a higromicina para permitir la retransformación de las células ES de rata transformadas en *Il2rg*. El Crei comprende dos exones que codifican la Cre recombinasa y que están separados por un intrón (Crei) para evitar su expresión en una célula procariota. Ver, por ejemplo, la patente de los Estados Unidos 8,697,851 y la publicación de solicitud de patente de los Estados Unidos 2013-0312129, que describen el casete de autodelección con detalle. Mediante el empleo de un promotor *P_{rm1}* de rata que conduce la expresión de Crei específicamente en células germinales masculinas, el casete de autodelección puede deleccionarse de las células germinales masculinas de ratas F0.

El LTVEC se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas en MEF neoR densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron y se mantuvieron como se describe en el Ejemplo 1.

Las colonias se tamizaron como se describe en cualquier otra parte de la presente descripción y se obtuvieron los clones transformados. Los clones transformados se inyectan después en un embrión huésped como se describe en cualquier otra parte de la presente descripción para producir una rata F0.

3.4.(b)(ii): Retransformación del locus de Rag1 y de Rag2 en las células ES de rata en las cuales el locus de Il2rg ya se ha transformado

Un LTVEC como en la figura 50 se preparó para que se dirija a los loci de *Rag1* y *Rag2* para delección. La longitud total del LTVEC fue de 72 kb. El LTVEC se electroporó en células ES de rata que ya se habían transformado para la delección del locus de *Il2rg* en el Ejemplo 3.3. Específicamente, las células ES de rata eran del clon *Il2rg*-CG12, para el cual la transmisión a través de la línea germinal se confirmó en el Ejemplo 3.3(a). Las células ES de rata transformadas se cultivaron y se mantuvieron como se describe en el Ejemplo 1. Los clones doble transformados se tamizaron como se describe en cualquier otra parte de la presente descripción, y se obtuvieron los clones transformados. *Las células Il2rg*-CG12 se retransformaron a una eficiencia del 85 %, y las mutaciones de *Il2rg* estuvieron aún presentes en los clones transformados. La electroporación se llevó a cabo como se describe en cualquier otra parte de la presente descripción, y la selección con antibióticos se llevó a cabo mediante el uso de 1,5 ug/ml de puomicina. Los clones transformados después se inyectarán en un embrión huésped como se describe en cualquier otra parte de la presente descripción para producir una rata F0. La retransformación es ventajosa porque es más rápida que la reproducción entre ratas transformadas en *Rag1/Rag2*- con ratas transformadas en *Il2rg*.

Ejemplo 4. Humanización

4.1. Humanización de los loci genómicos de rata

La humanización de los loci genómicos de rata se lleva a cabo empleando las células ES de rata descritas en la presente descripción, que son capaces de mantener su pluripotencia después de uno o más electroporaciones *in vitro*, y son capaces de transmitir las modificaciones genéticas dirigidas a generaciones posteriores. Además, para evitar las limitaciones de los plásmidos en acomodar un gran fragmento de ADN genómico, y para superar la baja eficiencia de introducir una modificación genética dirigida en un locus endógeno en células ES de rata, una o más modificaciones genéticas dirigidas se llevan a cabo en bacterias, *por ejemplo*, *E. coli*, mediante la utilización de la recombinación homóloga bacteriana (BHR) y el empleo de un vector de direccionamiento grande (LTVEC). El LTVEC descrito en la presente descripción, *por ejemplo*, incluye un gran fragmento de una secuencia genómica de rata endógena con una o más modificaciones o comprende un ácido nucleico exógeno (*por ejemplo*, un ácido nucleico humano homólogo u ortólogo) flanqueado por brazos de homología de rata complementarios a regiones genómicas específicas.

4.2. Humanización de los loci de inmunoglobulina de rata

La humanización de un locus de cadena pesada de inmunoglobulina de rata endógeno se lleva a cabo mediante la eliminación de una o más secuencias de ácidos nucleicos de cadena pesada de inmunoglobulina de rata endógenas (*por ejemplo*, uno o más segmentos del gen de V_H, uno o más segmentos del gen de D humano, y uno o más segmentos del gen de J_H humano); e introducir en el locus de inmunoglobulina modificado un vector de direccionamiento, *por ejemplo*, un vector de direccionamiento grande (LTVEC) que comprende: (i) una o más secuencias de ácidos nucleicos de región variable humana no reordenadas (*por ejemplo*, uno o más segmentos del gen de V_H humano, uno o más segmentos del gen de D humano, y uno o más segmentos del gen de J_H humano), o una o más secuencias de ácidos nucleicos de región variable humana no reordenada (*por ejemplo*, uno o más segmentos del gen de V-D-J reordenados humanos); (ii) un casete de selección (*por ejemplo*, gen de resistencia a neomicina flanqueado por sitios loxP); y (iii) brazos de homología de rata 5' y 3'.

En resumen, uno o más segmentos del gen de región variable de cadena pesada de inmunoglobulina de rata endógenos (*es decir*, uno o más segmentos del gen de V_H, uno o más segmentos del gen de D humano, y uno o más segmentos del gen de J_H humano) en un clon de BAC de rata se eliminan o inactivan mediante la transformación del locus de la cadena pesada de inmunoglobulina de rata endógeno con un casete de selección flanqueado por los brazos de homología de rata. Más específicamente, un vector de direccionamiento se construye para que contenga un casete de selección (*por ejemplo*, un gen de resistencia a neomicina flanqueado por sitios loxP) flanqueado por los brazos de homología de rata 5' y 3' que son complementarios a secuencias genómicas de rata objetivos (*por ejemplo*, secuencias de ADN genómicas corriente arriba y corriente abajo que abarcan uno o más segmentos del gen de V_H de rata, uno o más segmentos del gen de D humano, y uno o más segmentos del gen de J_H humano).

Después, las células bacterianas que contienen un gran fragmento de ADN genómico de rata que abarca un locus de cadena pesada de inmunoglobulina de rata se seleccionan y se introducen con un plásmido (*por ejemplo*, pABG) que codifica una recombinasa unida operativamente a un promotor inducible de manera transitoria. El vector de direccionamiento construido antes se introduce después en las células bacterianas competentes para recombinación. Después de la electroporación, las células bacterianas se tratan con un inductor (*por ejemplo*, arabinósido) para iniciar la recombinación homóloga entre el vector de direccionamiento y la secuencia genómica de rata objetivo en el clon de BAC. Las células transformadas se siembran en placas a una alta densidad y se someten a selección por fármacos para encontrar colonias que son resistentes a fármacos. Las colonias resistentes a fármacos se seleccionan y se tamizan para la modificación dirigida.

Para facilitar la identificación de la modificación genética dirigida, se emplea un ensayo cuantitativo de alto rendimiento, a saber, ensayo de modificación de alelo (MOA), que permite un tamizaje a gran escala de un(os) alelo(s) modificado(s) en un cromosoma progenitor después de una modificación genética. El ensayo de MOA puede llevarse a cabo a través de

- diversas técnicas analíticas, que incluyen, pero sin limitarse a, una PCR cuantitativa, por ejemplo, una PCR en tiempo real (qPCR). Por ejemplo, la PCR en tiempo real comprende un primer conjunto de cebadores que reconocen el locus objetivo y un segundo conjunto de cebadores que reconocen un locus de referencia que no se direcciona. Además, el conjunto de cebadores puede comprender una sonda fluorescente que reconoce la secuencia amplificada.
- 5 Alternativamente, el ensayo cuantitativo puede llevarse a cabo a través de una variedad de técnicas analíticas, que incluyen, pero sin limitarse a, la hibridación in situ mediada por fluorescencia (FISH), la hibridación genómica comparativa, la amplificación de ADN isotérmica, hibridación cuantitativa a una(s) sonda(s) inmovilizadas, Invader Probes®, MMP assays®, Baliza Molecular TaqMan®, y tecnología de sonda Eclipse™. (Ver, por ejemplo, el documento US2005/0144655.
- 10 Las células bacterianas que comprenden el clon de BAC de rata modificado, es decir, un clon de BAC que contiene una secuencia de ADN genómico de rata en donde uno o más segmentos del gen de región variable de cadena pesada endógeno (segmentos del gen de V_H , D, y/o J_H) se han delecionado o inactivado, se electroporan después con un vector de direccionamiento grande (LTVEC) que comprende: (i) una o más secuencias de ácidos nucleicos de región variable humana no reordenadas (por ejemplo, uno o más segmentos del gen de V_H humano no reordenados, uno o más
- 15 segmentos del gen de D humano, y uno o más segmentos del gen de J_H humano), o una o más secuencias de ácidos nucleicos de región variable humanos no reordenadas (por ejemplo, uno o más segmentos del gen de V-D-J humanos reordenados).
- La iniciación de la recombinación homóloga en las células bacterianas y la selección de clones positivos se realizan como se describió anteriormente. Las secuencias de ácidos nucleicos de región variable de cadena pesada de inmunoglobulina humana no reordenadas o reordenadas, cuando se dirigen al locus de cadena pesada de inmunoglobulina endógeno, se unen operativamente a una secuencia de aminoácidos de región variable de cadena pesada de inmunoglobulina de rata endógena. Alternativamente, el locus de la región constante de cadena pesada de rata endógeno puede inactivarse, por
- 20 ejemplo, mediante la delección de uno o más segmentos del gen de región constante de cadena pesada de rata (CH) del locus de la región constante de cadena pesada endógeno, y puede sustituirse con una secuencia de ácidos nucleicos de región constante de cadena pesada humana.
- De manera similar, la humanización de un locus de cadena ligera κ o λ de inmunoglobulina de rata endógeno se lleva a cabo mediante la eliminación de una o más secuencias de ácidos nucleicos de región variable de cadena ligera κ o λ de
- 30 inmunoglobulina de rata endógenas (*por ejemplo*, uno o más segmentos del gen de V_κ de rata endógenos y uno o más segmentos del gen de J_κ de rata endógenos); y la transformación del locus de cadena ligera de inmunoglobulina modificado con un vector de direccionamiento, *por ejemplo*, un vector de direccionamiento grande (LTVEC), que comprende: (i) una o más secuencias de ácidos nucleicos de región variable de cadena ligera de inmunoglobulina humana no reordenadas (*por ejemplo*, uno o más humano segmentos del gen de V_κ humano y uno o más segmentos del gen de J_κ humano), o una o más secuencias de ácidos nucleicos de región variable humanas no reordenadas (*por ejemplo*, uno o más segmentos del gen de V_κ - J_κ no reordenados humanos); (ii) un casete de selección (*por ejemplo*, gen de resistencia a neomicina flanqueado por sitios loxP); y (iii) brazos de homología de rata 5' y 3'.
- 40 Las secuencias de ácidos nucleicos de la región variable de cadena ligera de inmunoglobulina humanas no reordenadas o reordenadas, cuando se dirigen al locus de la cadena ligera de inmunoglobulina endógeno, se unen operativamente a la secuencia de ácidos nucleicos de región constante de cadena ligera de inmunoglobulina de rata endógena.
- El LTVEC producido así en las células bacterianas comprende, por ejemplo, un inserto de ácido nucleico que contiene un locus de cadena pesada o cadena ligera de inmunoglobulina de rata humanizado en el cual uno o más segmentos del gen de la región variable de cadena pesada o ligera humanos se han sustituido con uno o más segmentos del gen de
- 45 región variable de cadena pesada o ligera humanos; y brazos homólogos de rata (*por ejemplo*, en el intervalo de 5 kb a 150 kb) complementarios a las secuencias objetivo genómicas específicas. El LTVEC que comprende la modificación genética descrita anteriormente después se linealiza y se electropora en las células ES de rata. Las células ES de rata electroporadas se cultivan en placas a una alta densidad para seleccionar células ES resistentes a fármacos que comprenden el vector de direccionamiento. El proceso de selección por fármacos elimina la mayoría de las células sembradas (-99 %), dejando atrás colonias individuales, cada una de las cuales es un clon derivado de una célula única. De las células restantes, la mayoría de las células (~ 80-100 %) contienen el vector de direccionamiento integrado a una localización aleatoria en el genoma. Por lo tanto, las colonias se seleccionan y se genotipan de manera individual para
- 50 identificar células ES de rata que comprenden el vector de direccionamiento en la localización genómica correcta (por ejemplo, mediante el uso del ensayo de modificación de alelo (MOA) descrito anteriormente).
- Para aumentar la eficiencia de la modificación genética dirigida, las células ES de rata se electroporan con vectores de expresión (o ARNm) que expresan las ZFN 1 y 2 (o las TALEN 1 y 2) junto con el LTVEC. Los brazos de homología del vector de direccionamiento están fuera del sitio objetivo de ZFN, por lo tanto, el vector de direccionamiento no es escindido por las ZFN. La ruptura de la doble cadena producida por las ZFN estimula la reparación dirigida por homología (HDR), que de cualquier otra manera representa un porcentaje muy pequeño de reparaciones producidas normalmente en células de mamífero (en comparación con la unión de extremos no homólogos; NHEJ).
- 60 Alternativamente, los vectores de expresión que contienen una nucleasa asociada con CRISPR tipo II (por ejemplo, Cas9), un ARN guía (que incluye CRISPR-ARN (cr-ARN) y ARN de CRISPR de transactivación (ARNtracr)), como se describe en la presente descripción, pueden introducirse en las células bacterianas junto con el LTVEC para aumentar la eficiencia
- 65

de la recombinación homóloga en el locus genómico objetivo. Las células electroporadas se siembran en placas a una alta densidad y se someten a selección por fármacos para encontrar colonias que son resistentes a fármacos. Las colonias resistentes a fármacos se seleccionan y se tamizan para la modificación dirigida mediante el uso del ensayo de modificación de alelo (MOA) como se describe en la presente descripción. Después de estos procedimientos, puede lograrse una mejora de la eficiencia de transformación. Por ejemplo, la cantidad de mejora puede ser pequeña (por ejemplo, mejora del 10 % al 15 %) o grande (por ejemplo, mejora del 10 % al 80 %).

Las células ES de rata seleccionadas que comprenden la modificación genética dirigida se introducen después en un embrión de rata huésped, por ejemplo, un embrión de rata en etapa de premórula o de blastocisto, y se implantan en el útero de una madre sustituta para generar una rata fundadora (rata F0). Posteriormente, la rata fundadora se reproduce con una rata de tipo silvestre para generar la progenie F1 heterocigota para la modificación genética. El apareamiento de la rata F1 heterocigota puede producir la progenie homocigota para la modificación genética.

4.3(a). Sustitución de IL2rg de rata con la cadena gamma de receptor de IL2 humano

La Tabla 29 proporciona un resumen de la organización genómica del locus de la cadena gamma de receptor de Interleucina 2 de rata y las posiciones mostradas se toman de la construcción 5.0 de la secuencia de referencia del genoma de la rata (ENSEMBL). *Il2rg* se encuentra en el cromosoma X en la cadena (-).

Tabla 29. Resumen de la organización genómica del locus de *Il2rg* de rata

Característica	Inicio	Final	longitud	Notas
Exón 1	72,021,388	72,021,516	129	contiene ATG
ATG	72,017,500	72,017,502	3	codón de inicio
Exón 2	72,021,007	72,021,160	154	
sitio de unión a ZFN1	72,021,014	72,021,028	15	CAGGCCCTGAACCGC (sec. con núm. de ident.: 17)
sitio de corte de ZFN1	72,021,008	72,021,013	6	TTCTGG (sec. con núm. de ident.: 18)
sitio de unión a ZFN1b	72,020,993	72,021,007	15	GATTACCTGCGCTGGG (sec. con núm. de ident.: 20)
Exón 3	72,020,606	72,020,790	185	
Exón 4	72,020,274	72,020,413	140	
Exón 5	72,019,662	72,019,824	163	
Exón 6	72,019,101	72,019,197	97	
Exón 7	72,018,844	72,018,910	67	
Exón 8	72,017,856	72,018,506	651	contiene TGA
TGA	72,018,321	72,018,323	3	codón de parada
delección de <i>Il2rg</i>	72,018,323	72,021,502	3,180	

El esquema inferior en la figura 25 es el vector de direccionamiento para la delección de 3,2 kb de *Il2rg*. El vector de direccionamiento comprende un gen reportero (eGFP) unido operativamente al promotor endógeno y un casete de autodelección flanqueado por sitios loxP (flechas en blanco). El casete de autodelección comprende el gen de Crei unido operativamente a un promotor *Prrm1* de ratón y un casete de selección que comprende un gen de resistencia a neomicina unido operativamente a un promotor de ubiquitina humano.

El gen de Crei comprende dos exones que codifican una Cre recombinasa, que están separados por un intrón (Crei) para evitar su expresión en una célula procarionota. Ver, por ejemplo, la patente de los Estados Unidos 8,697,851 y la publicación de solicitud de patente de los Estados Unidos 2013-0312129, que describen el casete de autodelección. Mediante el empleo del promotor *Prrm1* de ratón el casete de expresión Cre y el casete de selección por fármacos puede deleccionarse específicamente en células germinales masculinas de ratas F0. El vector de direccionamiento se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas en MEF resistentes a neomicina densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron, se seleccionaron, y se mantuvieron como se describe en el Ejemplo 1.

Un vector plasmídico de transformación se construyó para sustituir la región codificante de la cadena gamma de receptor de interleucina 2 de rata en toda su longitud con la región codificante de la cadena gamma del receptor de interleucina 2

humana en toda su longitud como se muestra en la figura 31. El vector de direccionamiento se electroporó en las células ES de rata obtenidas en el Ejemplo 1, y las células se sembraron en placas en MEF resistentes a neomicina densas 2x de 15 cm en 2i + ROCKi 10 uM. Específicamente, 4×10^6 de células se electroporaron con 2 ug del vector de humanización en toda su longitud de *Il2rg* a una tensión de 400 V, una capacitancia de 100 uF, y una resistencia de 0. La selección se hizo mediante el uso de 75 ug/mL de G418. Las células ES de rata transformadas se cultivaron, se seleccionaron, y se mantuvieron como se describe en el Ejemplo 1.

Como se muestra en la Tabla 44, se tamizaron 168 colonias y se obtuvieron 6 clones transformados. La eficiencia de transformación fue de 3,57 %. Un clon se inyectó en blastocistos como se describe en el Ejemplo 1, y se obtuvo un clon que produce quimeras.

Los clones se inyectaron en blastocistos como se describe en la presente descripción en el Ejemplo 1. Se obtuvieron clones que producen ratas quiméricas F0. Los blastocistos se transfirieron a hembras receptoras pseudopreñadas mediante el uso de técnicas estándar, y se obtuvieron ratas F0 quiméricas. Se obtuvieron las ratas F0 que transmiten la modificación dirigida a través de la línea germinal.

4.3(b)(i). Sustitución del ectodominio de *Il2rg* de rata con el ectodominio de *Il2rg* humano

La humanización en toda su longitud de la cadena gamma del receptor de IL2 es útil porque las ratas que tienen este locus modificado producirán *Il2rg* humano; y esto podría permitir la detección de *Il2rg* humano en ratas con anticuerpos específicos a *Il2rg* humano.

La humanización de ecto (es decir, la sustitución del ectodominio de rata de *Il2rg* con el ectodominio humano de *Il2rg*) dará como resultado un polipéptido de *Il2rg* que se unirá a los ligandos humanos para *Il2rg*, pero porque el dominio citoplasmático aún es de rata, esta forma humanizada en ecto de *Il2rg* interactuará, además, con la maquinaria de señalización de rata. La figura 33 proporciona una alineación de secuencias de la proteína *IL-2rg* humana (sec. con núm. de ident.: 20; NP_000197.1); la proteína *IL-2rg* de rata (sec. con núm. de ident.: 21; NP_543165.1); y la proteína *IL-2rg* quimérica (sec. con núm. de ident.: 22) que comprende el ectodominio humano de *IL-2rg* fusionado al resto de la proteína *IL-2rg* de rata. La unión entre la *IL-2rg* humana y de rata se indica por la línea vertical.

La Tabla 30 proporciona un resumen de la organización genómica del locus de la cadena gamma de receptor de Interleucina 2 de rata y las posiciones mostradas se toman de la construcción 5.0 de la secuencia de referencia del genoma de la rata (ENSEMBL). *Il2rg* se encuentra en el cromosoma X en la cadena (-). Además, se indica la posición del ectodominio de *Il2rg*.

Tabla 30. Resumen de la organización genómica del locus de *Il2rg* de rata

Característica	Inicio	Final	Longitud	Notas
Exón 1	71,111,444	71,111,543	100	contiene ATG
ATG	71,111,537	71,111,539	3	codón de inicio
Exón 2	71,110,897	71,111,050	154	
Exón 3	71,110,504	71,110,688	185	
Exón 4	71,110,156	71,110,295	140	
Exón 5	71,109,228	71,109,390	163	
Exón 6	71,108,599	71,108,645	47	contiene el dominio transmembrana
Exón 7	71,108,277	71,108,346	70	
Exón 8	71,107,404	71,107,921	518	contiene TGA
TGA	71,108,736	71,108,738	3	codón de parada
humanización en toda la longitud:	71,107,404	71,111,539	4,136	(ATG a TGA más 3' poli-A)
humanización de ecto	71,108,679	71,111,539	2,861	(ATG tal inicio del dominio transmembrana)

Un vector plasmídico de transformación se construyó para sustituir el ectodominio de rata de la región codificante de la cadena gamma de receptor de interleucina 2 con el ectodominio humano como se muestra en la figura 32. El vector de direccionamiento se electroporó en las células ES de rata obtenidas en el Ejemplo 1 y las células se sembraron en placas

en MEF resistentes a neomicina densas 2x de 15 cm en 2i + ROCKi 10 uM. Las células ES de rata transformadas se cultivaron, se seleccionaron, y se mantuvieron como se describe en el Ejemplo 1.

5 Como se muestra en la Tabla 44, se tamizaron 192 colonias y se obtuvieron 13 clones transformados. La eficiencia de transformación fue de 6,77 %.

Se inyectaron dos clones en blastocistos como se describe en la presente descripción en el Ejemplo 1, y se obtuvieron dos clones que producen quimeras. Se obtuvieron clones que producen ratas F0. Se obtuvieron las ratas F0 que transmiten la modificación dirigida a través de la línea germinal.

10

4.3(b)(ii). Sustitución del ectodominio de IL2rg de rata con el ectodominio de IL2rg humano mediante el uso de plásmido en combinación con CRISPR/Cas9

15

La Tabla 31 muestra una comparación de los resultados de los experimentos en los cuales una versión del vector de humanización del ectodominio de *Il2rg* mostrada en la figura 32 se usó solo para dirigirse al locus de *Il2rg* de rata o se usó en combinación con un CRISPR/nucleasa Cas9 para dirigirse al locus de *Il2rg* de rata. En cada experimento, las células electroporadas se sembraron en placas a una alta densidad y se sometieron a selección por fármacos para encontrar colonias que eran resistentes a fármacos. Las colonias resistentes a fármacos se seleccionaron y se tamizaron para la modificación dirigida mediante el uso del ensayo de modificación de alelo (MOA) como se describe en la presente descripción. Específicamente, 4×10^6 células se electroporaron con 2 ug del vector de humanización de ectodominio de *Il2rg* a una tensión de 400 V, una capacitancia de 100 uF, y una resistencia de 0 En el último experimento, se electroporaron, además, 6 ug de plásmido de expresión de Cas9 y 3 ug de ARNg2 de *Il2rg* o 3 ug de ARNg4 de *Il2rg*. La selección se hizo mediante el uso de 75 ug/mL de G418. El ARNg2 de *Il2rg* tiene una secuencia de GAAGCTCTTTCTATAACAATCTGG (sec. con núm. de ident.: 91) y se dirige a una región de 190 pb 3' del exón 1 de *Il2rg* de rata. El ARNg4 de *Il2rg* tiene una secuencia de CCCCGAAAGGAGGAGCCCTAGG (sec. con núm. de ident.: 92) y se dirige a una región de 80 pb 5' del codón de parada de *Il2rg* (TGA) (ver la figura 49).

20

25

Tabla 31. Comparación de la transformación con vector de humanización del ectodominio de *Il2rg* con y sin CRISPR/Cas9

30

Vector	Cas9	ARNg	Colonias tamizadas	Clones transformados	Eficiencia de transformación
vector plasmídico de <i>Il2rg</i>	No	No	77	46	60 %
vector plasmídico de <i>Il2rg</i>	Si	ARNg2 de <i>Il2rg</i>	84	54	64 %
vector plasmídico de <i>Il2rg</i>	Si	ARNg4 de <i>Il2rg</i>	88	50	57 %

35

40

4.4(a). Transformación potenciada por el sistema CRISPR/endonucleasas Cas9 de grandes deleciones de genes de animales no humanos con sustituciones simultáneas de genes humanos

45

Los fármacos de recién desarrollo para condiciones de enfermedades humanas, tales como los anticuerpos completamente humanos, son frecuentemente muy específicos para sus objetivos en células y tejidos humanos y no reconocen los objetivos homólogos en roedores. Este alto nivel de selectividad hace imposible probar la eficacia y el mecanismo de acción de los fármacos en roedores antes de su primer uso en seres humanos.

50

Una solución muy eficaz a este problema es generar un ratón o rata genéticamente modificada en la cual el gen humano que codifica el objetivo del fármaco sustituya al homólogo de roedor. Una vía para generar un alelo humanizado tal en un roedor es deleccionar primero el gen de roedor en una célula madre embrionaria (ES) y después, en un segundo evento de modificación génica, insertar el gen humano precisamente en el locus deleccionado. Las células ES se inyectan después en un embrión de roedor y se implantan en el útero de un roedor madre sustituto, que posteriormente da a luz a crías genéticamente modificadas que portan el alelo humanizado.

55

Un método más eficiente para generar la modificación génica humanizada es usar un vector de direccionamiento grande (LTVEC) que dirige la deleción simultánea del gen de roedor y la sustitución con su contraparte humana. Mediante el empleo de los métodos de ingeniería genética de VELOCIGENE®, tales humanizaciones de una sola etapa pueden lograrse con una eficiencia relativamente alta cuando la deleción génica de roedor y la inserción génica humana son más pequeñas que aproximadamente 20 pares de kilobases (kb). Las humanizaciones de una sola etapa más grandes que implican deleciones y sustituciones de más de 100 kb son posibles con los LTVEC y los métodos de ingeniería genética tales como los métodos de ingeniería genética de VELOCIGENE®, pero debido a eficiencias de transformación reducidas que a veces se presentan con las modificaciones muy grandes, el éxito frecuentemente requiere el tamizaje de cientos o miles de clones de células ES para encontrar uno que porte la modificación génica deseada.

60

65

Para mejorar la eficiencia de grandes humanizaciones se han desarrollado métodos que combinan la transformación del gen de LTVEC con endonucleasas Cas9 guiadas por ARN de repeticiones palindrómicas cortas agrupadas regularmente interespaciadas (CRISPR/Cas9). CRISPR/nucleasas Cas9 son enzimas ribonucleoproteicas que consisten en una endonucleasa de ADN Cas9 unida a un ARN de CRISPR que guía a Cas9 a escindir en una secuencia específica de ADN mediante el apareamiento de las bases de Watson-Crick entre el ARN guía y una cadena del ADN objetivo. Debido a la simplicidad del mecanismo de transformación, es fácil diseñar el sistema CRISPR/endonucleasas Cas9 que dirija una ruptura de la doble cadena en cualquier locus genómico cercano. Las rupturas de la doble cadena inducen la reparación genómica celular mediante las rutas de la unión de extremos no homólogos (NHEJ), que son propensas a errores y frecuentemente dan como resultado deleciones o inserciones en el sitio de la ruptura de la doble cadena. Un mecanismo de reparación alternativo de la ruptura de la doble cadena es la reparación dirigida por homología (HDR) en la cual una pieza endógena o exógena de ADN que comparte identidad de secuencia o similitud con el sitio roto repara a la perfección los extremos rotos mediante la acción de la maquinaria de la recombinación homóloga celular. La HDR puede dar como resultado una reparación perfecta que restaura la secuencia original en el sitio roto, o esta puede usarse para dirigir una modificación diseñada, tal como una deleción, inserción, o sustitución de la secuencia en el sitio de la ruptura de la doble cadena. CRISPR/nucleasas Cas9 puede mejorar enormemente la tasa de eventos de HDR modificados mediante ingeniería genética dirigiendo escisiones precisas de la doble cadena en los sitios de las modificaciones génicas pretendidas.

Para efectuar una deleción precisa, de una sola etapa de todo o parte de un gen de roedor y la sustitución simultánea con todo o parte de su homólogo humano, se introdujeron mediante electroporación tres moléculas de ácidos nucleicos en células ES de roedor: (1) un LTVEC; (2) un plásmido o ARNm que codifica una endonucleasa Cas9; y (3) un plásmido que codifica un ARN guía único de CRISPR (ARNsg) o el ARNsg en sí mismo. El LTVEC comprendía todo o parte de un gen humano que codifica el producto génico (proteína o ARN) flanqueado por los brazos de homología de ADN de roedor diseñado para dirigirse a un evento de HR que deleciona el gen de roedor e inserta el gen humano. El LTVEC de humanización también porta un casete de selección por fármacos que dirige la expresión de una enzima (por ejemplo, neomicina fosfotransferasa) que imparte resistencia a un fármaco antibiótico (por ejemplo, G418). Las células ES que captan el LTVEC y lo incorporan en sus genomas fueron capaces de crecer y formar colonias en un plato de Petri en un medio de cultivo que contiene el fármaco antibiótico. Debido a que se introdujeron 500 a 1 000 veces más moléculas de ácidos nucleicos que codifican CRISPR/Cas9 que moléculas de LTVEC, la mayoría de las colonias resistentes a fármacos que contenían LTVEC también contenían, al menos de manera transitoria, los componentes de CRISPR/Cas9. Se seleccionaron colonias resistentes a fármaco y se tamizaron mediante el método de pérdida de alelo (Valenzuela, D. y otros. (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, *Nature Biotech.* 21:652-660; Friendewey, D. y otros. (2010) The loss-of-allele assay for ES cell screening y mouse genotyping, *Methods Enzymol.* 476:295-307 para identificar clones que tenían el alelo humanizado correctamente transformado.

En un experimento particular el LTVEC se diseñó para generar una deleción de 68 kb del gen de *Lrp5* (proteína 5 relacionada con el receptor de lipoproteína de baja densidad) de ratón y una sustitución simultánea con un fragmento de 91 kb del gen de *LRP5* humano homólogo (Figura 34). EL LTVEC comprendía el fragmento de 91-kb del gen de *LRP5* humano flanqueado por los brazos de homología que contienen 7 kb y 33 kb de ADN genómico derivado de partes del locus de *Lrp5* de ratón que flanquean la secuencia de 68 kb del gen de *Lrp5* de ratón pretendida para la deleción. En experimentos separados, se combinó el LTVEC de humanización de *Lrp5* con un plásmido que codifica Cas9 y un segundo plásmido que codifica uno de ocho ARNsg (gA, gB, gB2, gC, gD, gE2, gE, gF) diseñados para generar las rupturas de la doble cadena dentro de la región del gen de *Lrp5* de ratón que se transformó para deleción. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *LRP5* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *Lrp5* se muestran in Tabla 32. Cuando el LTVEC se introdujo solo en las células ES, se encontró que el 1,0 % de los clones resistentes a fármacos tamizados portaban un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con endonucleasas Cas9 guiadas por siete de los ocho ARNsg analizados (ARNsg-5'A, ARNsg-5'B, ARNsg-5'B2, ARNsg-C, ARNsg-D, ARNsg-3'E2, y ARNsg-3'F; secuencias proporcionadas en la Tabla 33) produjo mutaciones heterocigotas monoalélicas correctamente transformadas a eficiencias que se encuentran en el intervalo de 2,1 a 7,3 %, lo que representa una mejora de 2 a 9 veces de la transformación de genes humanizados en una etapa en comparación con el LTVEC sin ayuda. Para la escisión por Cas9 guiada por ARNsg-5'B2, además de la transformación monoalélica, se detectó la humanización homocigota bialélica a una frecuencia del 1 %. Las células ES humanizadas en *Lrp5* homocigoto pueden convertirse por el método de ingeniería genética VELOCIMOUSE® (Poueymirou, W. T. y otros. (2007) F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses, *NatBiotechnol.* 25:91-99, directamente en ratones derivados de células completamente ES listos para estudios fenotípicos y de eficacia de fármacos.

Tabla 32. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de *Lrp5*.

Experimento	Clones tamizados	Actividad de CRISPR (%)	Frecuencia de mutaciones heterocigotas monoalélicas (%)	Frecuencia de mutaciones heterocigotas compuestas bialélicas (%)	Frecuencia de mutaciones homocigotas bialélicas (%)
LTVEC solo	96	N/A	1,0 (1/96)	0	0
LTVEC + Cas9 + ARNsg-5'A	96	75,6	7,3 (7/96)	0	0
LTVEC + Cas9 + ARNsg-5'B	96	79,5	4,2 (4/96)	0	0
LTVEC + Cas9 + ARNsg-5'B2	96	60,5	6,2 (6/96)	0	1,0 (1/96)
LTVEC + Cas9 + ARNsg-C	96	no analizado	4,2 (4/96)	0	0
LTVEC + Cas9 + ARNsg-D	96	no analizado	7,3 (7/96)	0	0
LTVEC + Cas9 + ARNsg-3'E2	96	84,5	2,1 (2/96)	0	0
LTVEC + Cas9 + ARNsg-3'E	96	52,4	0	0	0
LTVEC + Cas9 + ARNsg-3'F	96	79,8	6,2 (6/96)	0	0

Tabla 33. Secuencias de las porciones guías de seis ARNsg que transforman el gen de *Lrp5* de ratón.

ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	Secuencia guía (5' a 3')
ARNsg-5'A	50	GGAACCCACAGCATACTCC (sec. con núm. de ident.: 24)
ARNsg-5'B	500	GAATCATGCACGGCTACCCC (sec. con núm. de ident.: 25)
ARNsg-5'B2	1000	TGCTCCTATGGGGAGGCGCG (sec. con núm. de ident.: 26)
ARNsg-C	29900/ 38430	ACTGAGATCAATGACCCCGA (sec. con núm. de ident.: 85)
ARNsg-D	29950/ 38380	GGGTCGCCCCGGAACCTCTAC (sec. con núm. de ident.: 86)
ARNsg-3'E2	1000	CTTGGATAACATTGATACCC (sec. con núm. de ident.: 27)
ARNsg-3'E	500	GGGGCAGAGCCCTTATATCA (sec. con núm. de ident.: 28)
ARNsg-3'F	50	TCGCTCACATTAATCCCTAG (sec. con núm. de ident.: 29)

La transformación mejorada de la gran humanización de *Lrp5* por el sistema CRISPR/endonucleasas Cas9 es notable cuando se compara con experimentos equivalentes realizados con las nucleasas de dedos de zinc (ZFN). Se obtuvieron cuatro ZFN diseñadas para producir rupturas de la doble cadena en los sitios dentro de la región del gen de *Lrp5* de ratón que es objetivo de delección (Figura 34). Una ZFN dirigida a una secuencia cerca del extremo 5' de la delección (a), una dirigida a una secuencia en el medio de la delección (b), y dos dirigidas a secuencias cerca del extremo 3' de la delección (c, d). En experimentos separados, se combinó el LTVEC de humanización de *Lrp5* con un plásmido que codifica una de las cuatro ZFN (a-d) diseñadas para generar rupturas de la doble cadena dentro de la región del gen de *Lrp5* de ratón que es objetivo de delección. Se determinó que todas las ZFN eran activas y capaces de inducir mutaciones por NHEJ en el gen de *Lrp5* (datos no mostrados), pero cuando se combinan con el LTVEC, ninguna mejoró la transformación de genes mediada por HDR en comparación con el LTVEC solo.

La eficiencia de transformación mejorada de la gran humanización de *Lrp5* por el sistema CRISPR/endonucleasas Cas9 es notable, además, cuando se compara con una serie de experimentos de humanización asistida por ZFN. En estos experimentos, se realizaron una serie de humanizaciones asistidas por ZFN en las que las deleciones de genes objetivo de ratón y las inserciones de genes humanos eran en general de tamaño creciente (Tabla 34; Figura 35). La figura 35A representa el porcentaje de la eficiencia de transformación de los LTVEC que se dirigen a genes de tamaño creciente para su deleción. Los LTVEC se usaron solos (cuadros grises) o en combinación con las ZFN (cuadros negros). La figura 35B representa el porcentaje de la eficiencia de transformación de los LTVEC con inserciones de gen humano de tamaño creciente. De nuevo, los LTVEC se usaron solos (triángulos grises) o en combinación con las ZFN (triángulos negros). Como se muestra en la Tabla 34 y la figura 35, la capacidad de escisión del ADN mediada por las ZFN para potenciar la eficiencia de transformación de LTVEC desapareció cuando el tamaño de la deleción de gen objetivo de ratón fue mayor que 24,7 kb y cuando el tamaño de la inserción de gen humano fue mayor que 22,2 kb (Tabla 34; Figura 35A). Por el contrario, CRISPR/Cas9 fue capaz de mejorar la eficiencia de transformación de LTVEC del gen de *Lrp5*, la cual implica una deleción de gen de ratón de 68,3 kb y una inserción de gen humano de 91,0 kb (Tabla 32; Figura 34). Esto indica que el sistema CRISPR/endonucleasas Cas9 es capaz de potenciar la eficiencia de transformación por LTVEC en situaciones donde otras nucleasas (por ejemplo, nucleasas de dedos de zinc) no pueden.

20

25

30

35

40

45

50

55

60

65

Tabla 34. Resumen de las humanizaciones asistidas por ZFN.

Gen objetivo	Delección de gen de ratón (kb)	Inserción de gen humano (kb)	Brazo de homología 5' (kb)	Brazo de homología 3' (kb)	Eficiencia de escisión de ZFN (%)	Eficiencia de transformación (%)		Mejora en veces
						LTVEC solo	LTVEC + ZFN	
<i>Fcgr1a</i>	4,1	4,2	10,9	76,8	22,9	5,20	32,81	6,3
<i>Tlr4</i>	7,1	5,7	67,6	85,5	12,5	5,20	22,39	4,3
<i>Ptfr</i>	8,7	18,0	49,6	112,9	30,7	1,56	24,48	15,7
<i>Notch4</i>	18,7	22,2	50,1	34,9	27,1	10,41	12,50	1,2
<i>Accn2</i>	24,7	18,9	57,8	60,1	20,8	4,17	8,33	2,0
<i>Adams5</i>	37,6	39,6	83,3	61,5	4,2	0,00	0,00	n.a.
<i>Trpa1</i>	45,3	53,9	41,3	57,8	8,8	0,52	0,00	0
<i>Folh1</i>	55,1	61,3	18,4	114,7	8,8	1,04	0,00	0
<i>Lrp5</i>	68,3	91,0	6,9	33,4	35,9	2,08	1,04	(0,5)
<i>ErbB4</i>	101,6	126,7	47,8	26,0	n.d.	0,00	0,00	n.a.

n.d. = no determinado

n.a. = no aplicable

() = eficiencia de transformación menor con ZFN que sin ZFN

Se realizaron experimentos comparables para la humanización de otros genes de ratón. En un experimento, el LTVEC se diseñó para generar una delección de 45 kb del gen de *Trpa1* (miembro 1 de la subfamilia A de canales catiónicos receptores de potencial transitorio) de ratón y una sustitución simultánea con un fragmento de 55 kb del gen de *TRPA1* humano homólogo (Figura 36). El LTVEC comprendía el fragmento de 55 kb del gen de *TRPA1* humano flanqueado por los brazos de homología que contienen 41 kb y 58 kb de ADN genómico derivado de partes del locus de *Trpa1* de ratón que flanquean la secuencia de 45 kb del gen de *Trpa1* de ratón destinados a la delección. En experimentos separados, se combinó el LTVEC de humanización de *Trpa1* con un plásmido que codifica Cas9 y un segundo plásmido que codifica uno de ocho ARNsg (gA, gA2, gB, gC, gD, gE, gE2, y gF) diseñados para generar rupturas de la doble cadena dentro de la región del gen de *Trpa1* de ratón que es objetivo de delección. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *TRPA1* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *Trpa1* se muestran en la Tabla 35. Cuando el LTVEC se introdujo solo en células ES, se encontró que el 1,0 % de los clones resistentes a fármacos tamizados portaban un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con endonucleasa Cas9 guiada por seis de ocho ARNsg analizados (A, A2, B, C, D, y F; secuencias proporcionadas en la Tabla 43) produjo mutaciones heterocigotas monoalélicas correctamente transformadas o mutaciones homocigotas o heterocigotas compuestas bialélicas a eficiencias que se encuentran en el intervalo del 1,0 al 3,1 %. Para la escisión por Cas9 guiada por ARNg A y ARNg F, se detectaron mutaciones heterocigotas compuestas a una frecuencia del 1,0 %.

Tabla 35. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de *Trpa1*

Posición de ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	ARNg	Actividad de CRISPR (%)	Clones tamizados	Transformados heterocigotos	Heterocigotos compuestos	Transformados homocigotos
5'	100	ARNg A	30,9	96	0	1	0
5'	500	ARNg A2	no analizado	96	2	0	0
5'	1000	ARNg B	42,8	96	3	0	0
medio	25600 / 19740	ARNg C	no analizado	96	1	0	0
medio	26970/18370	ARNg D	no analizado	96	2	0	0
3'	1000	ARNg E2	no analizado	96	0	0	0
3'	500	ARNg E	22,6	96	0	0	0
3'	100	ARNg F	28,6	96	1	1	0
N/A	N/A	ninguna	N/A	96	1	0	0

En otro experimento, el LTVEC se diseñó para generar una delección de 55 kb del gen de *Folh1* (glutamato carboxipeptidasa 2) de ratón y una sustitución simultánea con un fragmento de 61 kb del gen de *FOLH1* humano homólogo (Figura 37). El LTVEC comprendía el fragmento de 61 kb del gen de *FOLH1* humano flanqueado por los brazos de homología que contienen 22 kb y 46 kb de ADN genómico derivado de partes del locus de *Folh1* de ratón que flanquean la secuencia de 55 kb del gen de *Folh1* de ratón destinado a la delección. En experimentos separados, se combinó el LTVEC de humanización de *Folh1* con un plásmido que codifica Cas9 y un segundo plásmido que codifica uno de seis ARNsg (gA, gA2, gC, gD, gE, y gE2) diseñados para generar rupturas de la doble cadena dentro de la región del gen de *Folh1* de ratón que es objetivo de delección. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *FOLH1* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *Folh1* se muestran en la Tabla 36. Cuando el LTVEC se introdujo solo en células ES, se encontró que ninguno de los 96 clones resistentes a fármacos tamizados portaba un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con endonucleasa Cas9 guiada por tres de seis ARNsg analizados (A, D, y E2; secuencias proporcionadas en la Tabla 43) produjo mutaciones heterocigotas monoalélicas correctamente transformadas a eficiencias que se encuentran en el intervalo del 1,0 a 3,1 %

Tabla 36. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de *Folh1*.

Posición de ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	ARNg	Actividad de CRISPR (%)	Clones tamizados	Transformados heterocigotos	Heterocigotos compuestos	Transformados homocigotos
5'	100	ARNg A	45,2	96	2	0	0
5'	500	ARNg A2	61,9	96	0	0	0
medio	30300 / 24800	ARNg C	7,1	96	0	0	0
medio	31290 / 23810	ARNg D	39,2	96	1	0	0
3'	500	ARNg E2	no analizado	96	1	0	0
3'	100	ARNg E	1,2	96	0	0	0
N/A	N/A	ninguna	N/A	96	0	0	0

En otro experimento, el LTVEC se diseñó para generar una delección de 76 kb del gen de ratón para el componente 5 del complemento (*C5* o *Hc*) y una sustitución simultánea con un fragmento de 97 kb del gen de *C5* humano homólogo (Figura 38). El LTVEC comprendía el fragmento de 97 kb del gen de *C5* humano flanqueado por los brazos de homología que contienen 34,1 kb y 31,2 kb de ADN genómico derivado de partes del locus de *C5(Hc)* de ratón que flanquean la secuencia de 76 kb secuencia del gen de *C5 (Hc)* de ratón destinado a la delección. En experimentos separados, se combinó el LTVEC de humanización de *C5 (Hc)* con un plásmido que codifica Cas9 y un segundo plásmido que codifica una de seis ARNsg (gA, gB, gC, gD, gE, y gE2) diseñados para generar rupturas de la doble cadena dentro de la región del gen de *C5 (Hc)* de ratón que es objetivo de delección. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *C5* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *C5 (Hc)* se muestran en la Tabla 37. Cuando el LTVEC se introdujo solo en células ES, se encontró que el 1,0 % de los clones resistentes a fármacos tamizados portaban un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con endonucleasa Cas9 guiada por todos los seis ARNsg analizados (A, B, C, D, E, y E2; secuencias proporcionadas en la Tabla 43) produjo mutaciones heterocigotas monoalélicas correctamente transformadas o mutaciones homocigotas o heterocigotas compuestas bialélicas a eficiencias que se encuentran en el intervalo del 4,2 al 16,7 %. Para la escisión por Cas9 guiada por ARNg A y E, se detectaron mutaciones heterocigotas compuestas a frecuencias del 5,2 % y 4,2 %, respectivamente.

Tabla 37. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de C5 (Hc).

Posición de ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	ARNg	Actividad de CRISPR (%)	Clones tamizados	Transformados heterocigotos	Heterocigotos compuestos	Transformados homocigotos
5'	100	ARNg A	64,3	96	11	5	0
5'	500	ARNg B	72,6	96	14	0	0
medio	38200 / 37500	ARNg C	47,6	96	11	0	0
medio	43500 / 32200	ARNg D	47,6	96	7	0	0
3'	500	ARNg E	25,0	96	0	4	0
3'	100	ARNg E2	27,4	96	6	0	0
N/A	N/A	ninguna	N/A	96	1	0	0

En otro experimento, el LTVEC se diseñó para generar una delección de 38 kb del gen de *Adamts5* (una desintegrina y metaloproteínasa con motivos de trombospondina 5) de ratón y una sustitución simultánea con un fragmento de 43 kb del gen de *ADAMTS5* humano homólogo (Figura 39). El LTVEC comprendía el fragmento de 43 kb del gen de *ADAMTS5* humano flanqueado por los brazos de homología que contienen 22 kb y 46 kb de ADN genómico derivado de partes del locus de *Adamts5* de ratón que flanquean la secuencia de 38 kb del gen de *Adamts5* de ratón destinado a la delección. En experimentos separados, se combinó el LTVEC de humanización de *Adamts5* con un plásmido que codifica Cas9 y un segundo plásmido que codifica uno de ocho ARNsg (gA, gA2, gB, gC, gD, gE, gE2, y gF) diseñados para generar rupturas de la doble cadena dentro de la región del gen de *Adamts5* de ratón que es objetivo de delección. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *ADAMTS5* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *Adamts5* se muestran en la Tabla 38. Cuando el LTVEC se introdujo solo en células ES, se encontró que ninguno de los 96 clones resistentes a fármacos tamizados portaba un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con la endonucleasa Cas9 guiada por dos de ocho ARNsg analizados (B y F; secuencias proporcionadas en la Tabla 43) produjo mutaciones heterocigotas monoalélicas correctamente transformadas o mutaciones heterocigotas compuestas bialélicas a una eficiencia del 1,0 %. Para la escisión por Cas9 guiada por ARNg E2, se detectaron mutaciones heterocigotas compuestas a una frecuencia del 1,0 %.

Tabla 38. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de *Adamts5*

Posición de ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	ARNg	Actividad de CRISPR (%)	Clones tamizados	Transformados heterocigotos	Heterocigotos compuestos	Transformados homocigotos
5'	100	ARNg A	85,7	96	0	0	0
5'	500	ARNg A2	54,8	96	0	0	0
5'	1000	ARNg B	66,7	96	1	0	0
medio	18700/18950	ARNg C	9,5	96	0	0	0
medio	18800/18850	ARNg D	4,8	96	0	0	0
3'	1000	ARNg F	36,9	96	0	1	0
3'	500	ARNg E	54,8	96	0	0	0
3'	100	ARNg E2	54,8	96	0	0	0
N/A	N/A	ninguna	N/A	96	0	0	0

En otro experimento, el LTVEC se diseñó para generar una delección de 102 kb del gen de *ErbB4* (receptor tirosina-proteína cinasa erbB-4) de ratón y una sustitución simultánea con un fragmento de 127 kb del gen de *ERBB4* humano homólogo (Figura 40). El LTVEC comprendía el fragmento de 127 kb del gen de *ERBB4* humano flanqueado por los brazos de homología que contienen 48 kb y 26 kb de ADN genómico derivado de partes del locus de *ErbB4* de ratón que flanquean la secuencia de 102 kb del gen de *ErbB4* de ratón destinado a la delección. En experimentos separados, se combinó el LTVEC de humanización de *ErbB4* con un plásmido que codifica Cas9 y un segundo plásmido que codifica uno de ocho ARNsg (gA, gB, gB2, gC, gD, gE, gE2, y gF) diseñados para generar rupturas de la doble cadena dentro de la región del gen de *ErbB4* de ratón que es objetivo de delección. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *ERBB4* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *ErbB4* se muestran en la Tabla 39. Cuando el LTVEC se introdujo solo en células ES, se encontró que ninguno de los 96 clones resistentes a fármacos tamizados portaba un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con endonucleasa Cas9 guiada por uno de ocho ARNsg analizados (D; secuencia proporcionada en la Tabla 43) produjo mutaciones heterocigotas monoalélicas correctamente transformadas o mutaciones heterocigotas compuestas bialélicas a una eficiencia del 1,0 %. Para la escisión por Cas9 guiada por ARN guía D, se detectaron mutaciones heterocigotas compuestas a una frecuencia del 1 %

Tabla 39. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de *ErbB4*

Posición de ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	ARNg	Actividad de CRISPR (%)	Clones tamizados	Transformados heterocigotos	Heterocigotos compuestos	Transformados homocigotos
5'	100	ARNg A	25,0	96	0	0	0
5'	500	ARNg B	no analizado	96	0	0	0
5'	1000	ARNg B2	47,6	96	0	0	0
medio	50200/51350	ARNg C	20,2	96	0	0	0
medio	50230/51320	ARNg D	42,8	96	0	1	0
3'	1000	ARNg F	15,5	96	0	0	0
3'	500	ARNg E	89,2	96	0	0	0
3'	100	ARNg E2	14,3	96	0	0	0
N/A	N/A	ninguna	N/A	96	0	0	0

En otro experimento, el LTVEC se diseñó para generar una delección de 110 kb del gen de *Rorl* (receptor de transmembrana de tirosina-proteína cinasa ROR1) de ratón y una sustitución simultánea con un fragmento de 134 kb del gen de *ROR1* humano homólogo (Figura 41). El LTVEC comprendía el fragmento de 134 kb del gen de *ROR1* humano flanqueado por los brazos de homología que contienen 41,8 kb y 96,4 kb de ADN genómico derivado de partes del locus de *Ror1* de ratón que flanquean la secuencia de 110 kb del gen de *Rorl* de ratón destinado a la delección. En experimentos separados, se combinó el LTVEC de humanización de *Rorl* con un plásmido que codifica Cas9 y un segundo plásmido que codifica uno de seis ARNsg (gA, gB, gC, gD, gE, y gF) diseñados para generar rupturas de la doble cadena dentro de la región del gen de *Ror1* de ratón que es objetivo de delección. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *ROR1* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *Rorl* se muestran en la Tabla 40. Cuando el LTVEC se introdujo solo en células ES, se encontró que ninguno de los 96 clones resistentes a fármacos tamizados portaba un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con endonucleasa Cas9 guiada por dos de seis ARNsg analizados (D y F; secuencias proporcionadas en la Tabla 43) produjo mutaciones bialélicas o heterocigotas monoalélicas correctamente dirigidas a eficiencias del 1,0 %. Para la escisión por Cas9 guiada por ARNg F, se detectaron, además, mutaciones heterocigotas compuestas a una frecuencia del 1 %.

Tabla 40. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de *Ror1*

Posición de ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	ARNg	Actividad de CRISPR (%)	Clones tamizados	Transformados heterocigotos	Heterocigotos compuestos	Transformados homocigotos
5'	200	ARNg A	no analizado	96	0	0	0
5'	1000	ARNg B	no analizado	96	0	0	0
medio	54300 / 55500	ARNg D	no analizado	96	1	0	0
medio	54500 / 55300	ARNg C	no analizado	96	0	0	0
3'	1000	ARNg E	no analizado	96	0	0	0
3'	200	ARNg F	no analizado	96	0	1	0
N/A	N/A	ninguna	N/A	96	0	0	0

En otro experimento, el LTVEC se diseñó para generar una delección de 79 kb del gen de *Dpp4* (dipeptidil peptidasa 4) de ratón y una sustitución simultánea con un fragmento de 82 kb del gen de *DPP4* humano homólogo (Figura 42). El LTVEC comprendía el fragmento de 82 kb del gen de *DPP4* humano flanqueado por los brazos de homología 5' y 3', cada uno que contiene 46 kb de ADN genómico derivado de partes del locus de *Dpp4* de ratón que flanquean la secuencia de 79 kb del gen de *Dpp4* de ratón destinado a la delección. En experimentos separados, se combinó el LTVEC de humanización de *Dpp4* con un plásmido que codifica Cas9 y un segundo plásmido que codifica uno de ocho ARNsg (gA, gB, gB2, gC, gD, gE, gE2, y gF) diseñados para generar rupturas de la doble cadena dentro de la región del gen de *Dpp4* de ratón que es objetivo de delección. Los ARNsg se diseñaron para evitar el reconocimiento de cualquier secuencia en la porción insertada del gen de *DPP4* humano.

Los resultados de la humanización asistida por CRISPR/Cas9 del gen de *Dpp4* se muestran en la Tabla 41. Cuando el LTVEC se introdujo solo en células ES, se encontró que el 2,1 % de los clones resistentes a fármacos tamizados portaban un alelo humanizado heterocigoto monoalélico correctamente transformado. Por el contrario, la combinación del LTVEC con endonucleasa Cas9 guiada por cualquiera de uno de ocho ARNsg analizados (A, B, B2, C, D, E, E2, y F; secuencias proporcionadas en la Tabla 43) produjo mutaciones heterocigotas monoalélicas correctamente transformadas a eficiencias que se encuentran en el intervalo del 2,1 al 7,3 %.

Tabla 41. Resultados del tamizaje para la humanización asistida por CRISPR/Cas9 del gen de *Dpp4*

Posición de ARNsg	Distancia aproximada a partir del extremo de la delección (bp)	ARNg	Actividad de CRISPR (%)	Clones tamizados	Transformados heterocigotos	Heterocigotos compuestos	Transformados homocigotos
5'	50	ARNg A	no analizado	96	7	0	0
5'	400	ARNg B	no analizado	96	2	0	0
5'	900	ARNg B2	no analizado	96	5	0	0
medio	38800/40200	ARNg C	no analizado	96	3	0	0
medio	40800/38100	ARNg D	no analizado	96	3	0	0
3'	900	ARNg E2	no analizado	96	2	0	0
3'	500	ARNg E	no analizado	96	6	0	0
3'	200	ARNg F	no analizado	96	5	0	0
N/A	N/A	ninguna	N/A	96	2	0	0

Una tabla que resume los resultados para la humanización asistida por CRISPR/Cas9 de los diversos genes de ratón se proporciona en la Tabla 42. La primera fila indica el locus del gen que se direcciona. La segunda fila indica el tamaño de la delección (Del) del locus de ratón endógeno y el tamaño de la inserción (Ins) del locus humano correspondiente. Las filas restantes muestran el número de colonias (de un total de 96) para cada condición que tenían mutaciones heterocigotas monoalélicas correctamente transformadas, mutaciones heterocigotas compuestas bialélicas, o mutaciones homocigotas bialélicas. "No ARNg" representa el LTVEC solo, mientras que las otras columnas representan el LTVEC más los ARNg correspondientes (indicado por la posición relativa dentro del locus de delección).

Tabla 42. Resumen de la humanización asistida por CRISPR/Cas9 de genes de ratón.

	<i>Lrp5</i>	<i>Trpa1</i>	<i>Folh1</i>	<i>C5 (Hc)</i>	<i>Adamts5</i>	<i>ErbB4</i>	<i>Ror1</i>	<i>Dpp4</i>
Del/Ins (kb)	68/91	45/55	55/61	76/97	38/43	102/127	110/134	79/82
Más 5'	7	1	2	16	0	0	0	7
5'	4	2	0	14	0	0	0	2
5'	7	3	N/A	N/A	1	0	N/A	5
Medio	4	1	0	11	0	0	1	3
Medio	7	2	1	7	0	1	0	3
3'	2	0	N/A	N/A	1	0	N/A	2
3'	0	0	1	4	0	0	0	6
Más 3'	6	2	0	6	0	0	1	5
No ARNg	1	1	0	1	0	0	0	2

Tabla 43. Secuencias de ARN guía usadas para la humanización asistida por CRISPR/Cas9 de genes de ratón.

ARNg	Secuencia guía (5' a 3')	Sec. con núm. de ident.
<i>ARNg A para Trpa1</i>	GTACTGGGGAATCGGTGGTC	30
<i>ARNg A2 para Trpa1</i>	CACGCACTCCAAATTTATCC	31
<i>ARNg B para Trpa1</i>	CTAAGTGTGTATCAGTACAT	32
<i>ARNg C para Trpa1</i>	TGCCCTGCACAATAAGCGCA	33
<i>ARNg D para Trpa1</i>	ACTCATTGAAACGTTATGGC	34
<i>ARNg E2 para Trpa1</i>	AGTAAGGGTGGATTAATTC	35
<i>ARNg E para Trpa1</i>	GCCATCTAGATTCATGTAAC	36
<i>ARNg F para Trpa1</i>	GACTAGAAATGTTCTGCACC	37
<i>ARNg A para Folh1</i>	TGAACCAATTGTGTAGCCTT	38
<i>ARNg A2 para Folh1</i>	AATAGTGGTAAAGCACCATG	39
<i>ARNg B para Folh1</i>	GTGTGCTAAGGATCGAAGTC	40
<i>ARNg C para Folh1</i>	CACCGAGATGCTTGGGTATT	41
<i>ARNg D para Folh1</i>	TGTAACCGCCCTGAATGACC	42
<i>ARNg E para Folh1</i>	AAAAGGGCATCATAAATCCC	43
<i>ARNg E2 para Folh1</i>	TCAAAAATAGTCATACACCT	44
<i>ARNg F para Folh1</i>	GGTCTCTAGTACATTGTAGA	45
<i>ARNg A para C5 (Hc)</i>	ATCACAAACCAGTTAACCGG	46
<i>ARNg B para C5 (Hc)</i>	TTTCAGACGAGCCGACCCGG	47
<i>ARNg B2 para C5 (Hc)</i>	CTGTCAACAGTGCCGCGTTT	48
<i>ARNg C para C5 (Hc)</i>	TGTGTGTCATAGCGATGTCG	49
<i>ARNg D para C5 (Hc)</i>	AACAGGTACCCTATCCTCAC	50

ES 2 700 596 T3

	<i>ARNg E2 para C5 (Hc)</i>	TCGTGGTTGCATGCGCACTG	51
	<i>ARNg E para C5 (Hc)</i>	GGCCCGGACCTAGTCTCTCT	52
5	<i>ARNg F para C5 (Hc)</i>	AGTCTGTAAAGTTAGCAGTC	53
	<i>ARNg A para Adamts5</i>	GGTGGTGGTGCTGACGGACA	54
	<i>ARNg A2 para Adamts5</i>	TATGAGATCAACACTCGCTA	55
10	<i>ARNg B para Adamts5</i>	CCAAGGACTTCCCCACGTTA	56
	<i>ARNg C para Adamts5</i>	TGCTTCCCTTATGCAAGATT	57
	<i>ARNg D para Adamts5</i>	TTAGGTACCCTATTTGAATA	58
15	<i>ARNg E2 para Adamts5</i>	TGCAGTGGGTGACAGGTCCA	59
	<i>ARNg E para Adamts5</i>	AGGGTTATACTGACGTTGTG	60
	<i>ARNg F para Adamts5</i>	TGTCTTTCAAGGAGGGCTAC	<u>61</u>
20	<i>ARNg A para Erbb4</i>	TGATGTGCAGTCAGACAAAG	62
	<i>ARNg B para Erbb4</i>	TGCACTATGGTTGACTATGA	63
	<i>ARNg B2 para Erbb4</i>	GGAATATTCTAATAGGAAGT	64
25	<i>ARNg C para Erbb4</i>	AAGTGCTGTACCATTCTAGC	65
	<i>ARNg D para Erbb4</i>	TAATCAATAGACAACCTCGT	66
	<i>ARNg E2 para Erbb4</i>	TCATTCCTAATGGTATTATA	67
30	<i>ARNg E para Erbb4</i>	AGGGTACATAGATGGCATCG	68
	<i>ARNg F para Erbb4</i>	CTCTTTAACAATTACCACTT	<u>69</u>
	<i>ARNg A para Ror1</i>	TGTGGGCCTTTGCTGATCAC	70
35	<i>ARNg B para Ror1</i>	AATCTATGATCCTATGGCCT	71
	<i>ARNg D para Ror1</i>	TGCCAATAGCAGTGACTTGA	72
	<i>ARNg C para Ror1</i>	GGGAAGAATGGGCTATTGTC	73
40	<i>ARNg E para Ror1</i>	GGTTGTTTGTGCTGATGACG	74
	<i>ARNg F para Ror1</i>	CCGTCCTAGGCCTTCTACGT	75
	<i>ARNg A para Dpp4</i>	ACTAGTAGACCTGAGGGGTT	76
45	<i>ARNg B para Dpp4</i>	GCTCCAGTGTTTAGGCCTTG	77
	<i>ARNg B2 para Dpp4</i>	GGCAAGCTGAAAACGCATGC	78
	<i>ARNg C para Dpp4</i>	GTAGATCGCTTTCCACTACC	79
50	<i>ARNg D para Dpp4</i>	GAACTCCACTGCTCGTGAGC	80
	<i>ARNg E2 para Dpp4</i>	ATAGGTGGGCACTATTGAAG	81
	<i>ARNg E para Dpp4</i>	ATGGGAAGGTTTATACCAGC	82
55	<i>ARNg F para Dpp4</i>	CGGTGTAAAAACAACGGGAA	83

60 Ejemplo 5. Resumen de la modificación dirigida de los loci genómicos de rata

Tabla 44. Resumen de la transformación de ratas con varios tipos de vectores y agentes nucleasa como se analiza en los Ejemplos 3 y 4.

65

Tabla 44. Resumen de la transformación de ratas

Ejemplo #	Locus	Vector	Colonias tamizadas	Clones transformados	Eficiencia de transformación	Transformados bialélicos	Eficiencia bialélica	Clones inyectados	Clones que producen quimeras	Clones que transmiten a través de la línea germinal	Notas
3.2(a)(ii)	ApoE	plásmido	384	23	5,99 %	0	0	3	3	1	
3.2(a)(iii)	ApoE + ZFN	plásmido	384	290	75,52 %	8	2,08 %	2	2	1	Estos 2 clones son transformados bialélicos
3.3(a)	Il2rg	plásmido	232	5	2,16 %	N/A	N/A	6	5	1	
3.2(b)(ii)	LTVEC de ApoE	LTVEC	288	8	2,78 %	1	0,35 %	3	1	0	
3.2(b)(iii)	LTVEC de ApoE + ZFN	LTVEC	288	16	5,56 %	1	0,35 %	1	N/A	0	Este clon es transformado bialélico
3.2(b)(iv)	LTVEC de ApoE	LTVEC	75	32	42,67 %	0	0				
3.2(b)(iv)	LTVEC de ApoE + CRISPR/Cas9	LTVEC	169	84	50 %	5	3 %	0	0	0	
4.3(a)	Humanización 1 de Il2rg	plásmido	168	6	3,57 %	N/A	N/A	1	1	0	sustituye el Il2rg de rata completo con Il2rg humano
4.3(b)(i)	Humanización 2 de Il2rg	plásmido	192	13	6,77 %	N/A	N/A	2	2	0	sustituye el ectodominio de Il2rg de rata con el ectodominio de Il2rg humano
4.3(b)(ii)	Humanización 2 de Il2rg	plásmido	77	46	59,74 %						
4.3(b)(ii)	Humanización 2 de Il2rg + CRISPR/Cas9	plásmido	172	104	60,47 %	N/A	N/A	0	0	0	sustituye el ectodominio de Il2rg de rata con el ectodominio de Il2rg humano
3.4(a)(i)	Rag2	LTVEC	270	0	0	0	0	0	0	0	Delección de 5,7 KB predicha
3.4(a)(ii)	LTVEC de Rag2	LTVEC	36	0	0	0	0				
3.4(a)(ii)	LTVEC de Rag2 + CRISPR/Cas9	LTVEC	39	6	15,38 %	1	2,5 %	1	1	0	
3.4(b)(i)	Rag1-2	LTVEC	256	1	0,39 %	0	0	1	1	0	Delección de 16,2 kb predicha
3.4(b)(ii)	Rag1-2	LTVEC	94	80	85 %	0	0	0	0	0	El LTVEC Rag1-2 se electroporó en el clon mutante de Il2rg-CG12 (retransformación)

La Tabla 45 muestra un resumen de la transformación de células ES de rata ya sea con plásmidos o LTVEC en combinación con CRISPR/Cas9. Se analizaron dos ARN guías por separado para cada locus transformado: *Rag2*, *ApoE*, y *Il2rg*. La eficiencia de escisión de CRISPR/Cas9 fue > 20 % en los tres loci. Se observó un aumento de la eficiencia de transformación y aumento de la transformación bialélica cuando se usó CRISPR/Cas9 en combinación con los plásmidos de transformación y los LTVEC.

Tabla 45. Resumen de la transformación de células ES de rata con plásmidos o LTVEC en combinación con CRISPR/Cas9

Condición	Eficiencia de transformación	de	Transformación bialélica
Rag2 (LTVEC)	0		0
Rag2 (LTVEC+ CRISPR)	6 - 22 %		0 - 4 %
ApoE (LTVEC)	43 %		0
ApoE (LTVEC + CRISPR)	47 - 53 %		1 - 4 %
Humanización de <i>Il2rg</i> (vector plasmídico)	60 %		N/A (ligado a X)
Humanización de <i>Il2rg</i> (plásmido + CRISPR)	57 - 64 %		N/A (ligado a X)

La Tabla 46 muestra un resumen de los datos de transmisión a través de la línea germinal para la modificación dirigida de los loci genómicos de rata. La transmisión a través de la línea germinal se confirmó para ratas transformadas en *ApoE* y ratas transformadas en *Il2rg*. Las células ES de rata eran XY (macho) y eran transformadas heterocigotas. Por lo tanto, cuando las células ES transformadas contribuyen a la línea germinal, aproximadamente el 50 % del esperma derivado de células ES portarán el alelo mutante y producirán crías F1 heterocigotas.

Tabla 46. Datos de la transmisión a través de la línea germinal para la modificación dirigida de los loci genómicos de rata

Gen transformado	Clones microinyectados	Clones que producen quimeras	Clones que logran la transmisión a través de la línea germinal	Crías de la línea germinal/ Crías totales	Crías F1 heterocigotas*	Crías heterocigotas en las cuales se delecionó el casete neo
ApoE	3	3	1	7/79 (9 %)	4	4
<i>Il2rg</i>	5	5	1	11/257 (5 %)	5	5

Ejemplo 6. Generación, mantenimiento, y transformación células madre pluripotentes inducidas humanas

6.1. Generación de células iPS humanas

Este ejemplo describe la generación de células iPS humanas a partir de células humanas pluripotentes. Los vectores PiggyBac (System Biosciences) (PB-600A_CAGGS Bst XI (0,64 µg/µL) y PB-200 (0,99 µg/µL) que comprenden los genes que codifican cuatro factores de reprogramación (hOct4, hSox2, hKLF-4, hMYC) unido operativamente a un promotor CM7 se introdujeron en fibroblastos de prepucio humano neonatal mediante el uso de los reactivos de transfección RED y de BLUE GenIn™ (GlobalStem). Las células transfectadas se incubaron en células alimentadoras NuFF1 en medio E7 para permitir la incorporación de los vectores y la expresión de los factores de reprogramación. El medio E7 comprendía DMEM/F-12, NaHCO₃, ácido L-ascórbico, insulina, transferrina, selenio, y FGF-2.

La selección con puomicina comenzó 10 días después de la transfección mediante el uso de 2 µg/mL de puomicina en medio E7. En el día 21, se seleccionaron las colonias y se cultivaron en medio mTeSR™, que comprendía DMEM/F-12, NaHCO₃, ácido L-ascórbico, insulina, transferrina, selenio, FGF-2, TGF-β1, glutatión, L-glutamina, lípidos definidos, tiamina, elementos traza B y C, β-mercaptoetanol, albúmina de suero bovino, ácido pipercolico, cloruro de litio, y GABA. A los días 29 a 57, las células se propagaron y se sometieron a pases en medio mTeSR™ hasta que alcanzó ~50 % de confluencia en placas de 6 pocillos. A los días 65 a 73, la propagación y los pases continuaron mediante el uso de medio mTeSR™ y del reactivo de disociación celular suave (Stem Cell Technologies). El día 76, el medio se cambió a medio VG2i de baja osmolalidad para la propagación, pase, y mantenimiento adicional de las células que comprenden las hiPSC vírgenes o de apariencia virgen.

6.2. Transformación de LTVEC en células iPS humanas

Este ejemplo describe el uso de la transformación de LTVEC en células iPS humanas. Como se muestra en la figura 51, se introdujeron mediante electroporación en células iPS humanas propagadas en medio VG2i las siguientes moléculas de ácido nucleico: (1) un LTVEC (0,67 µg); (2) un plásmido que codifica una endonucleasa Cas9 (5 µg); y (3) un plásmido que codifica un ARN guía (ARNg) único de CRISPR (10 µg). En un conjunto de muestras, se excluyeron la Cas9 y el ARNg. Específicamente, 3×10^6 células se electroporaron a una tensión de 700 V, una capacitancia de 25 µF, y una resistencia de 400 ohm. El LTVEC comprendía un ácido nucleico de 16,7 kb que comprende genes de *Adam6a* y *Adam6b* de ratón flanqueados por los brazos de homología que contienen 34 kb y 105 kb de ADN genómico derivado de regiones genómicas que flanquean la secuencia de 4,1 kb del locus de ADAM6 humano destinado a la delección. El LTVEC también portaba un casete de selección por fármacos que dirige la expresión de una enzima que imparte resistencia a un fármaco antibiótico (higromicina). El ARNg de ADAM6 humano usado tenía la siguiente secuencia: GTATAGCCCTGTTACACATT (sec. con núm. de ident.: 94).

Las células que captan el LTVEC y lo incorporan en sus genomas fueron capaces de crecer y formar colonias en un plato de cultivo de tejidos recubierto con GELTRESTM en un medio de cultivo que contenían el fármaco antibiótico. Debido a que se introdujeron 500 a 1 000 veces más moléculas de ácidos nucleicos que codifican CRISPR/Cas9 que moléculas de LTVEC, la mayoría de las colonias resistentes a fármacos que contenían LTVEC también contenían, al menos de manera transitoria, los componentes de CRISPR/Cas9. Se escogieron colonias resistentes a fármacos y se tamizaron mediante el método de pérdida de alelos (Valenzuela y otros. (2003) Nat. Biotech. 21:652-660; Frenthewey y otros. (2010) Methods Enzymol. 476:295-307 para identificar clones que tenían el alelo correctamente transformado.

Los resultados de la transformación de LTVEC asistida por CRISPR/Cas9 del locus de ADAM6 se muestran en la Tabla 47.

Tabla 47. Transformación de LTVEC asistida por CRISPR/Cas9

	Eficiencia de transformación
LTVEC solo	3,1 %
LTVEC + CRISPR	7,3 %

Cuando el LTVEC se introdujo solo en las células iPS humanas, se observó una eficiencia de transformación del 3,1 %. Por el contrario, la combinación del LTVEC con Cas9 guiada por el ARNg de ADAM6 dio como resultado una eficiencia de transformación del 7,3 %.

6.3. Efecto del medio de baja osmolalidad en la morfología de células iPS humanas

Este ejemplo describe el efecto de la concentración salina, fuerza iónica, y/o osmolalidad en el estado de pluripotencia de las células iPS humanas en cultivo. Las células iPS humanas se cultivaron en un sustrato de MATRIGELTM o GELTRESTM en un medio descrito en la Tabla 48 o en el medio mTeSRTM-hLIF.

Tabla 48. Medio para el cultivo de células iPS.

Componente	Cantidad (v/v)
Medio base	24,75
medio F-12	24,75
Suplemento N2®	0,5
Medio Neurobasal®	49
Suplemento B-27®	1
Penicilina/Estreptomina	1
L-Glutamina (200 mM)	1
2-Mercaptoetanol (55 mM)	0,1836
hLIF (1 x 10 ⁴ unidades/mL)	0,001
CHIR99021 (10 mM)	0,03
PD0325901 (10 mM)	0,005

Cuando el medio base usado fue DMEM, este medio se denominó como medio 2i. Cuando el medio base usado fue VG-DMEM, este medio de baja osmolalidad se denominó como medio VG2i. La osmolalidad del medio VG2i (233 mOsm/kg) es menor que la osmolalidad del medio 2i tradicional (261 mOsm/kg).

Como se muestra en la figura 52, las células iPS humanas cultivadas en MATRIGEL™ en medio 2i durante un periodo de 8 días (Figura 52A) o 12 días (Figura 52B) mostraron una morfología característica de células iPS en un estado imprimado, particularmente el crecimiento en una monocapa epitelial y la aparición de la polaridad ápico basal.

5 El medio mTeSR-hLIF y el medio VG2i se evaluaron, además, para sus efectos en la morfología y el estado de pluripotencia de las células iPS humanas. En este estudio, las células iPS humanas se cultivaron en MATRIGEL™ o células alimentadoras NuFF en medio mTeSR™-hLIF (Figuras 53A y 53C) o en medio VG2i (Figuras 53B y 53D) durante un periodo de 6 días. Cuando se cultivan en medio mTeSR™-hLIF en MATRIGEL™ o células alimentadoras NuFF, las células iPS humanas mostraron una morfología característica de un estado de pluripotencia imprimado, particularmente el crecimiento en una monocapa epitelial y la aparición de la polaridad ápico basal. Alguna de las células cultivadas en medio mTeSR™-hLIF comenzaron a mostrar una morfología caracterizada por el agrupamiento tridimensional. Por el contrario, cuando se cultivan en medio VG2i en MATRIGEL™ o células alimentadoras NuFF, las células iPS humanas mostraron una morfología característica de un estado de pluripotencia virgen, particularmente el crecimiento en colonias redondas, en forma de domos y una pérdida de polaridad ápico basal.

15 6.4. Efecto del medio de baja osmolalidad en la expresión de marcadores de pluripotencia en células iPS humanas

Este ejemplo describe el efecto de la concentración salina, la fuerza iónica, y/o la osmolalidad en la expresión de marcadores de pluripotencia en células iPS humanas que se han reprogramado a partir de un estado imprimado a un estado virgen. Después de 24 días de cultivo en medio VG2i en un sustrato MATRIGEL™, las células iPS humanas vírgenes se tiñeron para la expresión de fosfatasa alcalina o NANOG. Se observó que las células reprogramadas expresaron fuertemente tanto la fosfatasa alcalina (Figura 54A) como NANOG (Figuras 54B y 54C), que son indicativos de un estado de pluripotencia virgen.

25 6.5. Efecto del medio de baja osmolalidad en la disociación enzimática y el subcultivo de células iPS humanas

En este ejemplo, las células iPS humanas que se reprogramaron a un estado virgen mediante el uso de medio VG2i de baja osmolalidad se disociaron enzimáticamente mediante el uso de tripsina para generar una suspensión celular única (Figura 55A). La suspensión celular se sometió a pases en nuevas placas recubiertas con GELTRES™ para el subcultivo en medio VG2i. Se observó después de 1 día (Figura 55B) y 4 días (Figura 55C) que las células subcultivadas continuaron mostrando una morfología característica de las células en un estado de pluripotencia virgen. En particular, las células crecieron como colonias en forma de domo redondas y no exhibieron una polaridad ápico basal. Fue notable que la disociación enzimática podría realizarse en la ausencia de un inhibidor de ROCK, el cual es típicamente necesario para evitar la activación de rutas proapoptóticas. Esto sugiere que las rutas proapoptóticas no se activan tan fuertemente durante la disociación enzimática y el subcultivo en células iPS humanas vírgenes bajo las condiciones identificadas en la presente descripción.

Listado de secuencias

40 <110> Frendewey, David
Auerbach, Wojtek
Lai, Ka-Man Venus
Mujica, Alexander O.
45 Lee, Jeffrey D. Droguett, Gustavo
Trzaska, Sean
Hunt, Charleen
Gagliardi, Anthony
Kuno, Junko
Valenzuela, David M.
50 Yancopoulos, George D.

<120> Métodos y composiciones para la modificación dirigida de un genoma

55 <130> 057766/453460

<150> US 61/914,768
<151> 2013-12-11

60 <150> US 62/017,416
<151> 2014-06-26

<150> US 62/029,261
<151> 2014-07-25

65 <150> US 62/052,906
<151> 2014-09-19

<150> US 62/059,527
 <151> 2014-10-03

 5 <150> US 62/064,384
 <151> 2014-10-15

 <160> 94

 10 <170> FastSEQ para Windows Versión 4.0

 <210> 1
 <211> 23
 <212> ADN
 <213> Secuencia Artificial

 15 <220>
 <223> una secuencia objetivo genómica que se une a un ARN guía (ARNg)

 20 <220>
 <221> característica_miscelánea
 <222> 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,
 19, 20, 21
 <223> n = A,T,C o G

 25 <400> 1
 gnnnnnnnnn nnnnnnnnnn ngg 23

 <210> 2
 <211> 80
 30 <212> ARN
 <213> Secuencia Artificial

 <220>
 <223> un ARN guía (ARNg)
 35 <400> 2

 guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu 60
 ggcaccgagu cggugcuuuu 80

 40 <210> 3
 <211> 42
 <212> ARN
 <213> Secuencia Artificial

 45 <220>
 <223> un ARN guía (ARNg)

 <400> 3
 50 guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cg 42

 <210> 4
 <211> 30
 <212> ARN
 55 <213> Secuencia Artificial

 <220>
 <223> un ARNcr

 60 <400> 4
 guuuuagagc uagaaauagc aaguuaaaau 30

 <210> 5
 <211> 33
 65 <212> ARN
 <213> Secuencia Artificial

ES 2 700 596 T3

<220>
<223> un ARNcr

5 <400> 5
guuuuagagc uagaaauagc aaguuaaaau aag 33

<210> 6
<211> 26
<212> RNA
10 <213> Secuencia Artificial

<220>
<223> un ARNcr

15 <400> 6
gaguccgagc agaagaagaa guuuua 26

<210> 7
<211> 12
20 <212> ARN
<213> Secuencia Artificial

<220>
<223> un ARNtracr

25 <400> 7
aaggcuaguc cg 12

<210> 8
<211> 50
30 <212> ARN
<213> Secuencia Artificial

<220>
<223> un ARNtracr

35 <400> 8
aaggcuaguc cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu 50

40 <210> 9
<211> 203
<212> PRT
<213> Mus Musculus

45 <400> 9

50

55

60

65

ES 2 700 596 T3

5 Met Lys Val Leu Ala Ala Gly Ile Val Pro Leu Leu Leu Leu Val Leu
 1 5 10 15
 His Trp Lys His Gly Ala Gly Ser Pro Leu Pro Ile Thr Pro Val Asn
 20 25 30
 Ala Thr Cys Ala Ile Arg His Pro Cys His Gly Asn Leu Met Asn Gln
 35 40 45
 Ile Lys Asn Gln Leu Ala Gln Leu Asn Gly Ser Ala Asn Ala Leu Phe
 50 55 60
 Ile Ser Tyr Tyr Thr Ala Gln Gly Glu Pro Phe Pro Asn Asn Val Glu
 65 70 75 80
 10 Lys Leu Cys Ala Pro Asn Met Thr Asp Phe Pro Ser Phe His Gly Asn
 85 90 95
 Gly Thr Glu Lys Thr Lys Leu Val Glu Leu Tyr Arg Met Val Ala Tyr
 100 105 110
 15 Leu Ser Ala Ser Leu Thr Asn Ile Thr Arg Asp Gln Lys Val Leu Asn
 115 120 125
 Pro Thr Ala Val Ser Leu Gln Val Lys Leu Asn Ala Thr Ile Asp Val
 130 135 140
 Met Arg Gly Leu Leu Ser Asn Val Leu Cys Arg Leu Cys Asn Lys Tyr
 145 150 155 160
 20 Arg Val Gly His Val Asp Val Pro Pro Val Pro Asp His Ser Asp Lys
 165 170 175
 Glu Ala Phe Gln Arg Lys Lys Leu Gly Cys Gln Leu Leu Gly Thr Tyr
 180 185 190
 Lys Gln Val Ile Ser Val Val Val Gln Ala Phe
 195 200
 25

<210> 10
 <211> 15
 <212> ADN
 30 <213> Secuencia Artificial

<220>
 <223> Sitio de unión a ZFN1a

35 <400> 10
 caggccctga accgc 15

<210> 11
 <211> 6
 40 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Sitio de corte de ZFN1

45 <400> 11
 ttctgg 6

<210> 12
 <211> 16
 50 <212> ADN
 <213> Secuencia Artificial

<220>
 55 <223> Sitio de unión a ZFN1b
 <400> 12
 gattacctgc gctggg 16

<210> 13
 <211> 15
 60 <212> ADN
 <213> Secuencia Artificial

<220>
 65 <223> Sitio de unión a ZF21a

<400> 13
 ttcaccctcc gcacc 15

5 <210> 14
 <211> 7
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> Sitio de corte de ZFN2

<400> 14
 tgctgag 7

15 <210> 15
 <211> 18
 <212> ADN
 <213> Secuencia Artificial

20 <220>
 <223> Sitio de unión a ZF21b

<400> 15
 tatccagatc caggggtt 18

25 <210> 16
 <211> 9
 <212> PRT
 <213> Secuencia Artificial

30 <220>
 <223> dominio conservado de una familia de nucleasas de asentamiento

<400> 16

35

Leu Ala Gly Leu Ile Asp Ala Asp Gly
 1 5

40 <210> 17
 <211> 15
 <212> ADN
 <213> Secuencia Artificial

45 <220>
 <223> Sitio de unión a ZFN1a

<400> 17
 cagggcctga accgc 15

50 <210> 18
 <211> 6
 <212> ADN
 <213> Secuencia Artificial

55 <220>
 <223> Sitio de corte de ZFN1

<400> 18
 ttctgg 6

60 <210> 19
 <211> 16
 <212> ADN
 <213> Secuencia Artificial

65

ES 2 700 596 T3

<220>

<223> Sitio de unión a ZFN1b

<400> 19

5 gattacctgc gctggg 16

<210> 20

<211> 369

<212> PRT

10 <213> Homo sapiens

<400> 20

15 Met Leu Lys Pro Ser Leu Pro Phe Thr Ser Leu Leu Phe Leu Gln Leu
 1 5 10 15
 Pro Leu Leu Gly Val Gly Leu Asn Thr Thr Ile Leu Thr Pro Asn Gly
 20 25 30
 Asn Glu Asp Thr Thr Ala Asp Phe Phe Leu Thr Thr Met Pro Thr Asp
 35 40 45
 Ser Leu Ser Ser Val Ser Thr Leu Pro Leu Pro Glu Val Gln Cys Phe Val
 50 55 60
 Phe Asn Val Glu Tyr Met Asn Cys Thr Trp Asn Ser Ser Ser Glu Pro
 65 70 75 80
 Gln Pro Thr Asn Leu Thr Leu His Tyr Trp Tyr Lys Asn Ser Asp Asn
 85 90 95
 25 Asp Lys Val Gln Lys Cys Ser His Tyr Leu Phe Ser Glu Glu Ile Thr
 100 105 110
 Ser Gly Cys Gln Leu Gln Lys Lys Glu Ile His Leu Tyr Gln Thr Phe
 115 120 125
 Val Val Gln Leu Gln Asp Pro Arg Glu Pro Arg Arg Gln Ala Thr Gln
 130 135 140
 30 Met Leu Lys Leu Gln Asn Leu Val Ile Pro Trp Ala Pro Glu Asn Leu
 145 150 155 160
 Thr Leu His Lys Leu Ser Glu Ser Gln Leu Glu Leu Asn Trp Asn Asn
 165 170 175
 Arg Phe Leu Asn His Cys Leu Glu His Leu Val Gln Tyr Arg Thr Asp
 180 185 190
 35 Trp Asp His Ser Trp Thr Glu Gln Ser Val Asp Tyr Arg His Lys Phe
 195 200 205
 Ser Leu Pro Ser Val Asp Gly Gln Lys Arg Tyr Thr Phe Arg Val Arg
 210 215 220
 40 Ser Arg Phe Asn Pro Leu Cys Gly Ser Ala Gln His Trp Ser Glu Trp
 225 230 235 240
 Ser His Pro Ile His Trp Gly Ser Asn Thr Ser Lys Glu Asn Pro Phe
 245 250 255
 45 Leu Phe Ala Leu Glu Ala Val Val Ile Ser Val Gly Ser Met Gly Leu
 260 265 270
 Ile Ile Ser Leu Leu Cys Val Tyr Phe Trp Leu Glu Arg Thr Met Pro
 275 280 285
 Arg Ile Pro Thr Leu Lys Asn Leu Glu Asp Leu Val Thr Glu Tyr His
 290 295 300
 50 Gly Asn Phe Ser Ala Trp Ser Gly Val Ser Lys Gly Leu Ala Glu Ser
 305 310 315 320
 Leu Gln Pro Asp Tyr Ser Glu Arg Leu Cys Leu Val Ser Glu Ile Pro
 325 330 335
 Pro Lys Gly Gly Ala Leu Gly Glu Gly Pro Gly Ala Ser Pro Cys Asn
 340 345 350
 55 Gln His Ser Pro Tyr Trp Ala Pro Pro Cys Tyr Thr Leu Lys Pro Glu
 355 360 365
 Thr

<210> 21

60 <211> 368

<212> PRT

<213> Rattus norvegicus

<400> 21

65

ES 2 700 596 T3

5 Met Leu Lys Pro Leu Leu Pro Ser Arg Ser Phe Leu Leu Leu Gln Leu
 1 5 10
 Leu Leu Leu Arg Val Gly Trp Ser Ser Lys Val Leu Met Ser Ser Gly
 20 25 30
 10 Asn Glu Asp Thr Lys Ser Asp Leu Leu Leu Thr Ser Met Asp Leu Lys
 35 40 45
 His Leu Ser Val Pro Thr Leu Pro Leu Pro Glu Val Gln Cys Phe Val
 50 55 60
 15 Phe Asn Val Glu Tyr Met Asn Cys Thr Trp Asn Ser Ser Ser Glu Pro
 65 70 75 80
 Gln Pro Thr Asn Leu Thr Met His Tyr Arg Tyr Lys Gly Ser Asp Asn
 85 90 95
 20 Asn Thr Phe Gln Glu Cys Ser His Tyr Leu Phe Ser Lys Glu Ile Thr
 100 105 110
 Ser Gly Cys Gln Ile Gln Lys Glu Asp Ile Gln Leu Tyr Gln Thr Phe
 115 120 125
 25 Val Val Gln Leu Gln Asp Pro Gln Lys Pro Gln Arg Arg Ala Glu Gln
 130 135 140
 Lys Leu Asn Leu Gln Asn Leu Val Ile Pro Trp Ala Pro Glu Asn Leu
 145 150 155 160
 30 Thr Leu Tyr Asn Leu Ser Glu Ser Gln Val Glu Leu Arg Trp Lys Ser
 165 170 175
 Arg Tyr Ile Glu Arg Cys Leu Gln Tyr Leu Val Gln Tyr Arg Ser Asn
 180 185 190
 35 Arg Asp Arg Ser Trp Thr Glu Gln Ile Val Asp His Glu Pro Arg Phe
 195 200 205
 Ser Leu Pro Ser Val Asp Glu Gln Lys Leu Tyr Thr Phe Arg Val Arg
 210 215 220
 40 Ser Arg Phe Asn Pro Ile Cys Gly Ser Thr Gln Gln Trp Ser Lys Trp
 225 230 235 240
 45 Ser Gln Pro Ile His Trp Gly Ser His Thr Ala Glu Glu Asn Pro Ser
 245 250 255
 Leu Phe Ala Leu Glu Ala Val Leu Ile Pro Val Gly Thr Met Gly Leu
 260 265 270
 50 Ile Ile Thr Leu Ile Phe Val Tyr Cys Trp Leu Glu Arg Met Pro Arg
 275 280 285
 Ile Pro Ala Ile Lys Asn Leu Glu Asp Leu Val Thr Glu Tyr His Gly
 290 295 300
 55 Asn Phe Ser Ala Trp Ser Gly Val Ser Lys Gly Leu Thr Glu Ser Leu
 305 310 315 320
 Gln Pro Asp Tyr Ser Glu Arg Phe Cys His Val Ser Glu Ile Pro Pro
 325 330 335
 60 Lys Gly Gly Ala Leu Gly Glu Gly Pro Gly Gly Ser Pro Cys Ser Leu
 340 345 350
 His Ser Pro Tyr Trp Pro Pro Pro Cys Tyr Ser Leu Lys Pro Glu Ala
 355 360 365

45 <210> 22
 <211> 368
 <212> PRT
 <213> Secuencia Artificial
 50 <220>
 <223> cadena gamma de receptor de IL-2 quimérica que comprende la proteína de la cadena gamma de receptor de IL-2 de rata que tiene el ectodominio de la cadena gamma de receptor de IL-2 de ser humano
 55 <400> 22
 60
 65

ES 2 700 596 T3

5 Met Leu Lys Pro Ser Leu Pro Phe Thr Ser Leu Leu Phe Leu Gln Leu
 1 5 10
 Pro Leu Leu Gly Val Gly Leu Asn Thr Thr Ile Leu Thr Pro Asn Gly
 20 25 30
 10 Asn Glu Asp Thr Thr Ala Asp Phe Leu Thr Thr Met Pro Thr Asp
 35 40 45
 Ser Leu Ser Val Ser Thr Leu Pro Leu Pro Glu Val Gln Cys Phe Val
 50 55 60
 Phe Asn Val Glu Tyr Met Asn Cys Thr Trp Asn Ser Ser Ser Glu Pro
 65 70 80
 15 Gln Pro Thr Asn Leu Thr Leu His Tyr Trp Tyr Lys Asn Ser Asp Asn
 85 90 95
 Asp Lys Val Gln Lys Cys Ser His Tyr Leu Phe Ser Glu Glu Ile Thr
 100 105 110
 Ser Gly Cys Gln Leu Gln Lys Lys Glu Ile His Leu Tyr Gln Thr Phe
 115 120 125
 Val Val Gln Leu Gln Asp Pro Arg Glu Pro Arg Arg Gln Ala Thr Gln
 130 135 140
 20 Met Leu Lys Leu Gln Asn Leu Val Ile Pro Trp Ala Pro Glu Asn Leu
 145 150 155 160
 Thr Leu His Lys Leu Ser Glu Ser Gln Leu Glu Leu Asn Trp Asn Asn
 165 170 175
 Arg Phe Leu Asn His Cys Leu Glu His Leu Val Gln Tyr Arg Thr Asp
 180 185 190
 Trp Asp His Ser Trp Thr Glu Gln Ser Val Asp Tyr Arg His Lys Phe
 195 200 205
 25 Ser Leu Pro Ser Val Asp Gly Gln Lys Arg Tyr Thr Phe Arg Val Arg
 210 215 220
 Ser Arg Phe Asn Pro Leu Cys Gly Ser Ala Gln His Trp Ser Glu Trp
 225 230 235 240
 Ser His Pro Ile His Trp Gly Ser Asn Thr Ser Lys Glu Asn Pro Phe
 245 250 255
 30 Leu Phe Ala Leu Glu Ala Val Leu Ile Pro Val Gly Thr Met Gly Leu
 260 265 270
 Ile Ile Thr Leu Ile Phe Val Tyr Cys Trp Leu Glu Arg Met Pro Arg
 275 280 285
 35 Ile Pro Ala Ile Lys Asn Leu Glu Asp Leu Val Thr Glu Tyr His Gly
 290 295 300
 Asn Phe Ser Ala Trp Ser Gly Val Ser Lys Gly Leu Thr Glu Ser Leu
 305 310 315 320
 Gln Pro Asp Tyr Ser Glu Arg Phe Cys His Val Ser Glu Ile Pro Pro
 325 330 335
 40 Lys Gly Gly Ala Leu Gly Glu Gly Pro Gly Gly Ser Pro Cys Ser Leu
 340 345 350
 His Ser Pro Tyr Trp Pro Pro Pro Cys Tyr Ser Leu Lys Pro Glu Ala
 355 360 365

45 <210> 23
 <211> 23
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia objetivo genómica que se une a un ARN guía (ARNg)
 55 <220>
 <221> característica_miscelánea
 <222> (2) ... (21)
 <223> n= A, T, C o G
 60 <220>
 <221> característica_miscelánea
 <222> (0)...(0)
 <223> n puede ser de 1-20 nucleótidos
 <400> 23
 65 gnnnnnnnnn nnnnnnnnnn ngg 23

<210> 24
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 24
 10 gggaacccac agcatactcc 20
 <210> 25
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 25
 gaatcatgca cggctacccc 20
 <210> 26
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 26
 tgctcctatg gggaggcgcg 20
 <210> 27
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 27
 cttggataac attgataccc 20
 45 <210> 28
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 28
 55 ggggcagagc ccttatatca 20
 <210> 29
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 29
 65 tcgctcacat taatccctag 20

<210> 30
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 30
 10 gtactgggga atcgggtggtc 20
 <210> 31
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 31
 cacgcactcc aaatttatcc 20
 <210> 32
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 32
 ctaagtgtgt atcagtacat 20
 <210> 33
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 33
 tgcctgcac aataagcgca 20
 45 <210> 34
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 34
 55 actcattgaa acggtatggc 20
 <210> 35
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 35
 65 agtaagggtg gattaaattc 20

<210> 36
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 36
 gccatctaga ttcattgtaac 20
 <210> 37
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 15
 <220>
 <223> una secuencia guía
 <400> 37
 gactagaaat gttctgcacc 20
 <210> 38
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 25
 <220>
 <223> una secuencia guía
 <400> 38
 tgaaccaatt ggtgtagcctt 20
 <210> 39
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 35
 <220>
 <223> una secuencia guía
 <400> 39
 aatagtggta aagcaccatg 20
 <210> 40
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 45
 <220>
 <223> una secuencia guía
 <400> 40
 gtgtgctaag gatcgaagtc 20
 <210> 41
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 55
 <220>
 <223> una secuencia guía
 <400> 41
 caccgagatg cttgggtatt 20

<210> 42
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 42
 10 tgtaaccgcc ctgaatgacc 20
 <210> 43
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 43
 aaaaggcat cataaatccc 20
 <210> 44
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 44
 tcaaaaatag tcatacacct 20
 <210> 45
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 45
 ggtctctagt acattgtaga 20
 45 <210> 46
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 46
 55 atcaciaaac agttaaccgg 20
 <210> 47
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 47
 65 ttcagacga gccgaccgg 20

<210> 48
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 48
 10 ctgtcaacag tgccgcgttt 20
 <210> 49
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 49
 tgtgtgtcat agcgatgtcg 20
 <210> 50
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 50
 aacaggtacc ctatcctcac 20
 <210> 51
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 51
 tcgtggtgc atgcgcactg 20
 45 <210> 52
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 52
 55 ggcccggacc tagtctctct 20
 <210> 53
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 53
 65 agtctgtaaa gtagcagtc 20

<210> 54
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 54
 10 ggtggtggtg ctgacggaca 20
 <210> 55
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 55
 tatgagatca acactcgcta 20
 <210> 56
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 56
 ccaaggactt ccccagtta 20
 <210> 57
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 57
 tgcttcctt atgcaagatt 20
 45 <210> 58
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 58
 55 ttaggtacc tattgaata 20
 <210> 59
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 59
 65 tgcagtgggt gacaggtcca 20

<210> 60
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 60
 10 aggggtatac tgacgttg 20
 <210> 61
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 61
 tgtcttcaa ggagggtac 20
 <210> 62
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 62
 tgatgtcag tcagacaaag 20
 <210> 63
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 63
 tgcactatgg ttgactatga 20
 45 <210> 64
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 64
 55 ggaatattct aataggaagt 20
 <210> 65
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 65
 65 aagtgtgta ccattctagc 20

<210> 66
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 66
 10 taatcaatag acaacctcgt 20
 <210> 67
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 67
 tcattcctaa tggattata 20
 <210> 68
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 68
 agggtagata gatggcatcg 20
 <210> 69
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 69
 ctcttaaca attaccact 20
 45 <210> 70
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 70
 55 tgtgggcctt tgctgatcac 20
 <210> 71
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 71
 65 aatctatgat cctatgcct 20

<210> 72
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 72
 10 tgccaatagc agtgactga 20
 <210> 73
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 73
 gggaagaatg ggctattgct 20
 <210> 74
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 74
 gggtgtttgt gctgatgacg 20
 <210> 75
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 75
 ccgtcctagg ccttctacgt 20
 45 <210> 76
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 76
 55 actagtagac ctgaggggtt 20
 <210> 77
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 77
 65 gctccagtgt ttaggccttg 20

5 <210> 78
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 10 <400> 78
 ggcaagctga aaacgcatgc 20
 <210> 79
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 79
 gtagatcgct ttccactacc 20
 <210> 80
 <211> 20
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 80
 gaactccact gctcgtgagc 20
 <210> 81
 <211> 20
 35 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> una secuencia guía
 <400> 81
 ataggtgggc actattgaag 20
 45 <210> 82
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 82
 atgggaaggt ttataccagc 20
 55 <210> 83
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 83
 65 cgggtgtaaaa acaacgggaa 20

5 <210> 84
 <211> 20
 <212> ARN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 10 <400> 84
 gacccgcagu cccagcgucg 20
 <210> 85
 <211> 20
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 85
 actgagatca atgaccccga 20
 <210> 86
 <211> 20
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 86
 gggctgccccg gaacctctac 20
 <210> 87
 <211> 23
 <212> ADN
 35 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 40 <400> 87
 gcaggccctg aaccgcttct tgg 23
 45 <210> 88
 <211> 23
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> una secuencia guía
 <400> 88
 55 cctgcgctgg gtcagacgc ttt 23
 <210> 89
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 60 <220>
 <223> una secuencia guía
 <400> 89
 65 ccagctactt gctcgtacaa 20

ES 2 700 596 T3

<210> 90
 <211> 20
 <212> ADN
 <213> Secuencia Artificial
 5
 <220>
 <223> una secuencia guía
 <400> 90
 10 cccctcagat tcacgtgcgt 20
 <210> 91
 <211> 23
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 20 <400> 91
 gaagctcttt ctatacaatc tgg 23
 <210> 92
 <211> 23
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 30 <400> 92
 cccccgaaag gaggagccct agg 23
 <210> 93
 <211> 56
 <212> ADN
 <213> Rattus norvegicus
 35 <400> 93
 40 ttctgagcct cagccgacca acctcactat gcactatagg tatgagaagg gggagg 56
 <210> 94
 <211> 20
 <212> ADN
 45 <213> Secuencia Artificial
 <220>
 <223> una secuencia guía
 50 <400> 94
 gtatagccct gttacacatt 20

Reivindicaciones

1. Un método *in vitro* para modificar un genoma en un locus genómico de interés en una célula madre embrionaria (ES) de ratón, que comprende:
 5 poner en contacto las células ES de ratón con una proteína Cas9, un ARN de CRISPR que se hibrida a una secuencia objetivo en el locus genómico de interés, un ARNtracr, y un vector de direccionamiento grande (LTVEC) que tiene un tamaño de al menos 10 kb y comprende un inserto de ácido nucleico flanqueado por:
 10 (i) un brazo de homología 5' que es homólogo a una secuencia objetivo 5' en el locus genómico de interés; y
 (ii) un brazo de homología 3' que es homólogo a una secuencia objetivo 3' en el locus genómico de interés,
 en donde después de poner en contacto las células ES de ratón con la proteína Cas9, el ARN de CRISPR, y el ARNtracr en la presencia del LTVEC, el genoma de la célula ES de ratón se modifica para que comprenda una modificación genética dirigida que comprende la delección de una región del locus genómico de interés en donde la delección es de al menos 30 kb y/o la inserción del inserto de ácido nucleico en el locus genómico de interés en donde la inserción es de al menos 30 kb.
 15
2. El método de conformidad con la reivindicación 1, en donde:
 (a) el ARN de CRISPR y el ARNtracr se introducen como un transcrito único que comprende el ARN de CRISPR y el ARNtracr, y opcionalmente en donde el transcrito único comprende el ARN de CRISPR y el ARNtracr fusionados juntos en la forma de un ARN guía único (ARNsg); o
 20 (b) el ARN de CRISPR y el ARNtracr se introducen por separado.
3. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde:
 (a) la proteína Cas9 se introduce en la forma de una proteína, un ARN mensajero (ARNm) que codifica la proteína Cas9, o un ADN que codifica la proteína Cas9;
 25 (b) el ARN de CRISPR se introduce en la forma de un ARN o un ADN que codifica el ARN de CRISPR; y
 (c) el ARNtracr se introduce en la forma de un ARN o un ADN que codifica el ARNtracr.
4. El método de conformidad con la reivindicación 3, en donde la proteína Cas9, el ARN de CRISPR, y el ARNtracr se introducen como un complejo de proteína-ARN.
 30
5. El método de conformidad con la reivindicación 3, en donde:
 (a) el ADN que codifica la proteína Cas9 se encuentra en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un primer ácido nucleico que codifica la proteína Cas9;
 35 (b) el ADN que codifica el ARN de CRISPR se encuentra en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un segundo ácido nucleico que codifica el ARN de CRISPR; y
 (c) el ADN que codifica el ARNtracr se encuentra en la forma de un tercer constructo de expresión que comprende un tercer promotor unido operativamente a un tercer ácido nucleico que codifica el ARNtracr;
 en donde los promotores primero, segundo y tercero son activos en las células ES de ratón, y
 40 en donde los constructos de expresión primero, segundo y tercero se encuentran en una única molécula de ácido nucleico o en múltiples moléculas de ácido nucleico.
6. El método de conformidad con la reivindicación 3, en donde:
 (a) el ADN que codifica la proteína Cas9 se encuentra en la forma de un primer constructo de expresión que comprende un primer promotor unido operativamente a un primer ácido nucleico que codifica la proteína Cas9; y
 45 (b) el ADN que codifica el ARN de CRISPR y el ADN que codifica el ARNtracr se encuentran en la forma de un segundo constructo de expresión que comprende un segundo promotor unido operativamente a un segundo ácido nucleico que codifica un ARN guía que comprende el ARN de CRISPR y el ARNtracr;
 en donde los promotores primero y segundo son activos en las células ES de ratón, y
 50 en donde los constructos de expresión primero y segundo se encuentran en una única molécula de ácido nucleico o en moléculas de ácidos nucleicos separadas.
7. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde
 (a) la modificación genética dirigida comprende la delección simultánea de una secuencia de ácidos nucleicos endógena en el locus genómico de interés y la inserción del inserto de ácido nucleico en el locus genómico de interés, opcionalmente en donde la secuencia de ácidos nucleicos endógena delecionada es de 30 kb a aproximadamente 110 kb, y el inserto de ácido nucleico es de aproximadamente 40 kb a aproximadamente 140 kb;
 55 (b) la modificación genética dirigida es una modificación genética bialélica, y opcionalmente en donde la modificación genética bialélica comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del inserto de ácido nucleico en el locus genómico de interés en dos cromosomas homólogos; o
 (c) en donde la célula ES de ratón modificada es heterocigota compuesta o hemicigota en el locus genómico de interés, y opcionalmente en donde la modificación genética dirigida en el locus genómico de interés en un cromosoma comprende la delección de una secuencia de ácidos nucleicos endógena y la inserción del inserto de ácido nucleico, y opcionalmente en donde la modificación genética dirigida comprende: (1) delección de una
 60 secuencia de ácidos nucleicos endógena en el locus genómico de interés en los cromosomas homólogos primero
 65

y segundo; y (2) inserción del inserto de ácido nucleico en el locus genómico de interés en el primer cromosoma homólogo y la alteración del locus genómico de interés en el segundo cromosoma homólogo.

- 5 8. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde el LTVEC es de al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, o al menos 200 kb.
- 10 9. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde modificación genética dirigida comprende la inserción del inserto de ácido nucleico, en donde el inserto de ácido nucleico es de al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, al menos 300 kb, o de aproximadamente 40 kb a aproximadamente 140 kb; o
 15 en donde la modificación genética dirigida comprende la inserción del inserto de ácido nucleico y la delección de una región del locus genómico de interés, en donde la delección es de al menos 30 kb, y el inserto de ácido nucleico es de al menos 10 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, o al menos 300 kb.
- 20 10. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la secuencia objetivo está flanqueada inmediatamente por una secuencia de Motivos Adyacentes a los Protoespaciadores (PAM).
- 25 11. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la suma total de los brazos de homología 5' y 3' del LTVEC es de 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 120 kb, o de aproximadamente 120 kb a 150 kb.
- 30 12. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la modificación genética dirigida comprende:
 (a) sustitución de una secuencia de ácidos nucleicos endógena con una secuencia de ácidos nucleicos homóloga u ortóloga;
 (b) delección de una secuencia de ácidos nucleicos endógena;
 (c) delección de una secuencia de ácidos nucleicos endógena,
 35 en donde la delección se encuentra en el intervalo de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb;
 (d) inserción de una secuencia de ácidos nucleicos exógena;
 (e) inserción de una secuencia de ácidos nucleicos exógena en el intervalo de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb;
 (f) inserción de una secuencia de ácidos nucleicos exógena que comprende una secuencia de ácidos nucleicos homóloga u ortóloga;
 (g) inserción de una secuencia de ácidos nucleicos quimérica que comprende una secuencia de ácidos nucleicos humana y una no humana;
 (h) inserción de un alelo condicional flanqueado por secuencias objetivo de recombinasa específica de sitio;
 (i) inserción de un gen marcador de selección o un gen reportero unido operativamente a un promotor activo en
 50 las células ES de ratón; o
 (j) una combinación de estos.
- 55 13. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la modificación genética dirigida comprende la delección de una región del locus genómico de interés,
 60 en donde la delección es de al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, al menos 300 kb; o
 en donde la modificación genética dirigida comprende, además, la inserción del inserto de ácido nucleico en el locus genómico de interés en donde la inserción es de al menos 30 kb, y la delección es de al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb,
 65 al menos 100 kb, al menos 150 kb, o al menos 200 kb.

14. El método de conformidad con cualquiera de las reivindicaciones anteriores, en donde la modificación genética dirigida comprende la delección de una región del locus genómico de interés en donde la delección es de al menos 30 kb y la inserción del inserto de ácido nucleico en el locus genómico de interés en donde la inserción es de al menos 30 kb.

5

15. El método de conformidad con cualquiera de las reivindicaciones anteriores, que comprende, además:
(b) identificar las células ES de ratón modificadas que comprenden la modificación genética dirigida en el locus genómico de interés;
(c) introducir las células ES de ratón modificadas en un embrión huésped de ratón; y
(d) gestar el embrión huésped de ratón en una madre sustituta, en donde la madre sustituta produce un ratón de generación F0 que comprende la modificación genética dirigida en el locus genómico de interés.

10

15

20

25

30

35

40

45

50

55

60

65

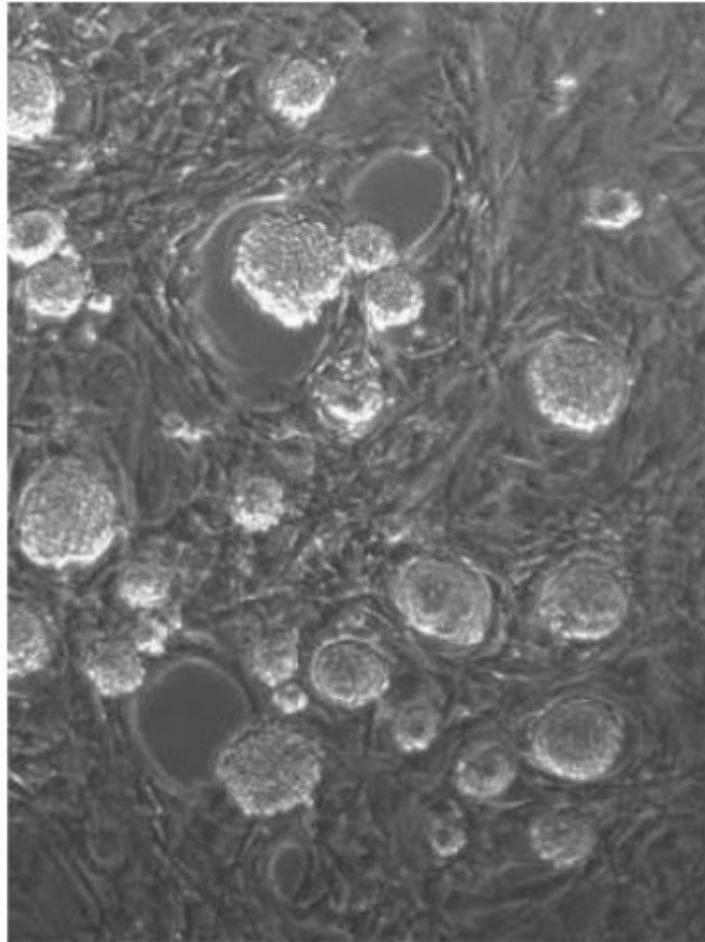
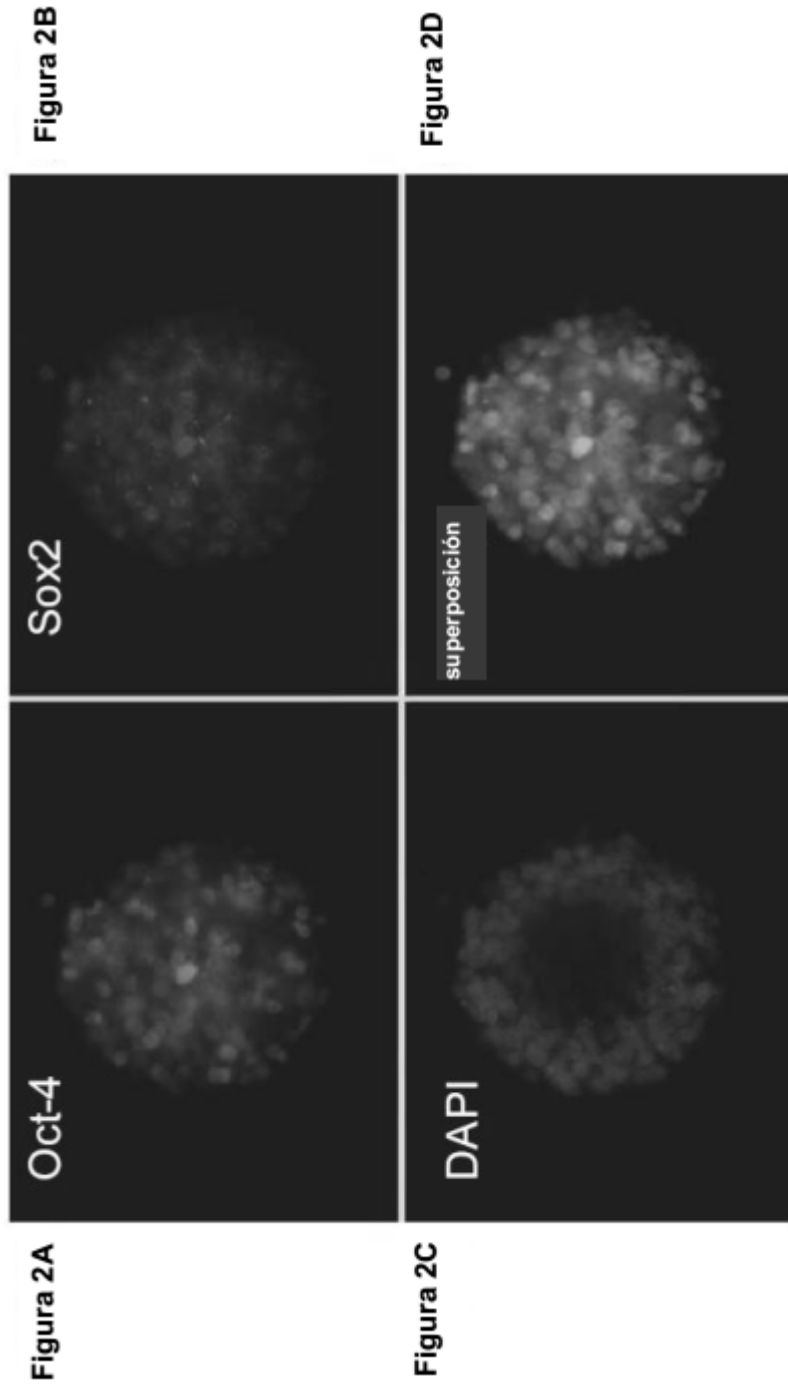


Figura 1



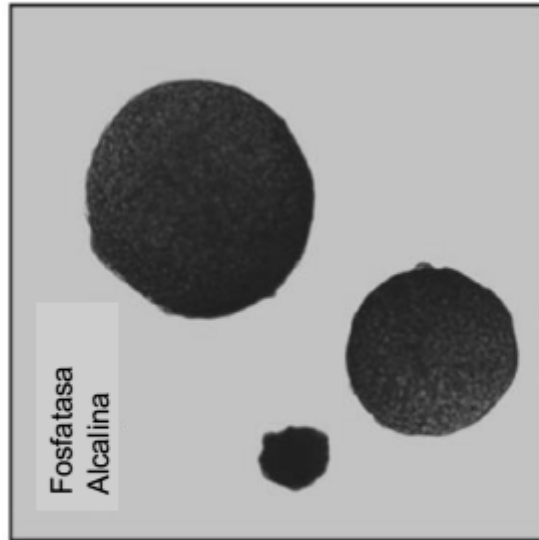
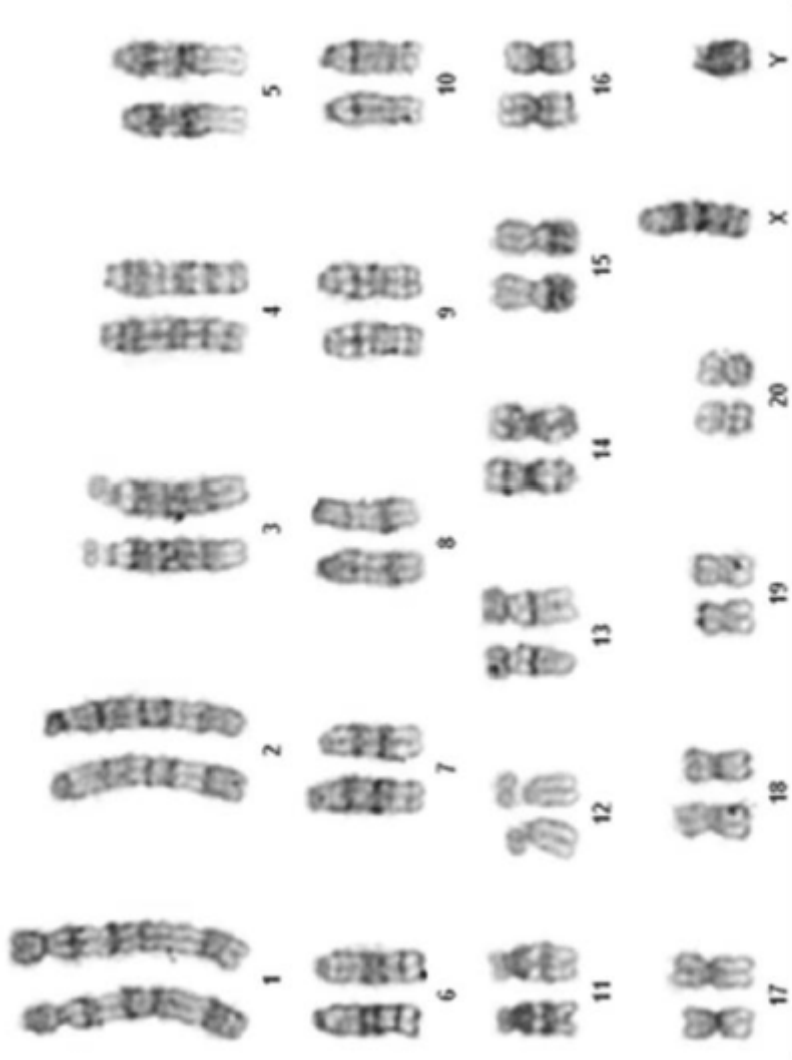


Figura 3



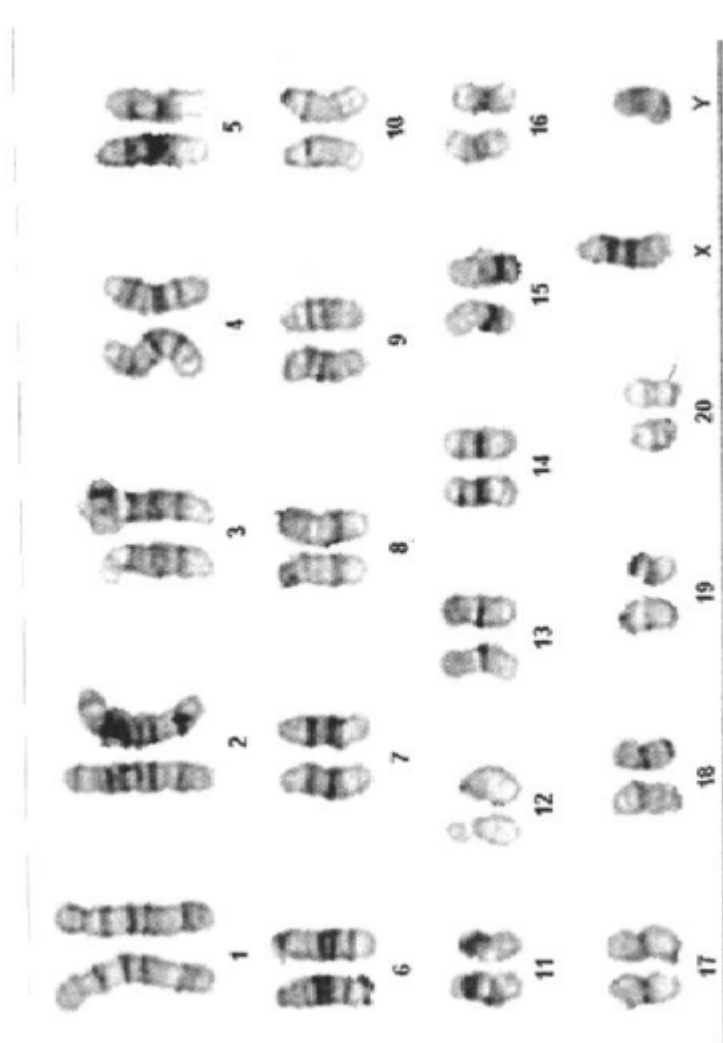
Resultados de las células: 42, XY

Figura 4



Resultados de las células: 42, XY

Figura 5A



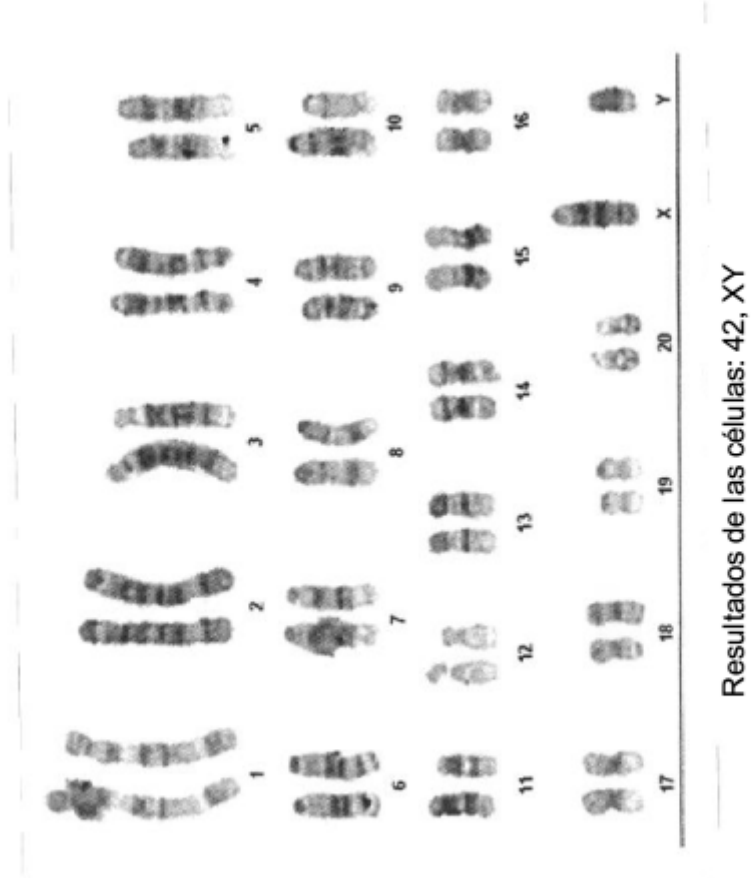
Resultados de las células: 42, XY

Figura 5B



Resultados de las células: 42, XY

Figura 6A



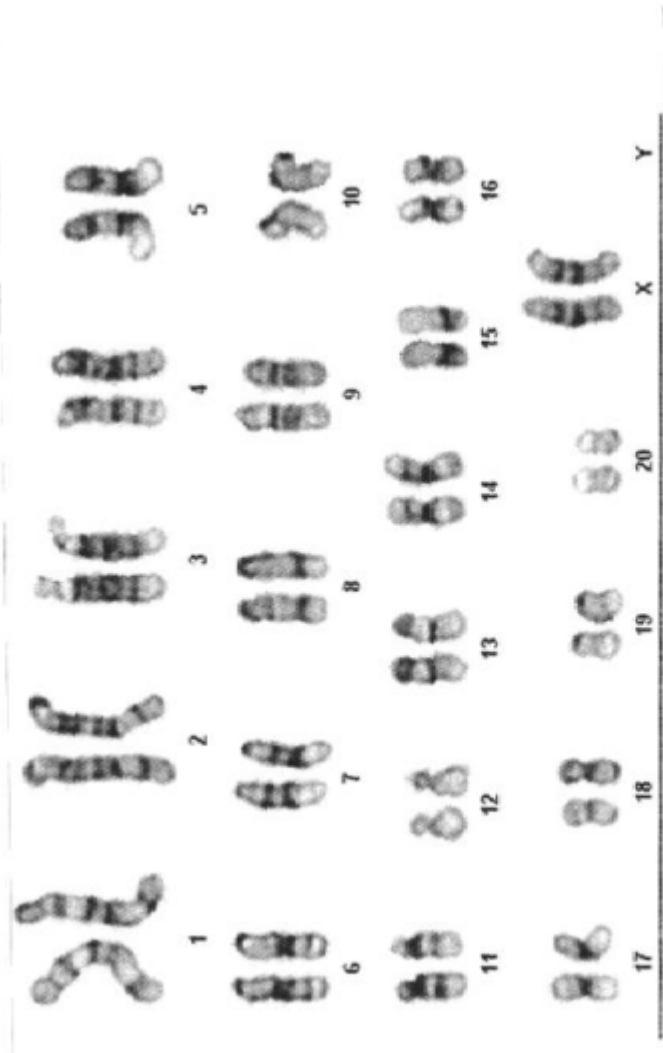
Resultados de las células: 42, XY

Figura 6B



Resultados de las células: 42, XX

Figura 7A



Resultados de las células: 42, XX

Figura 7B

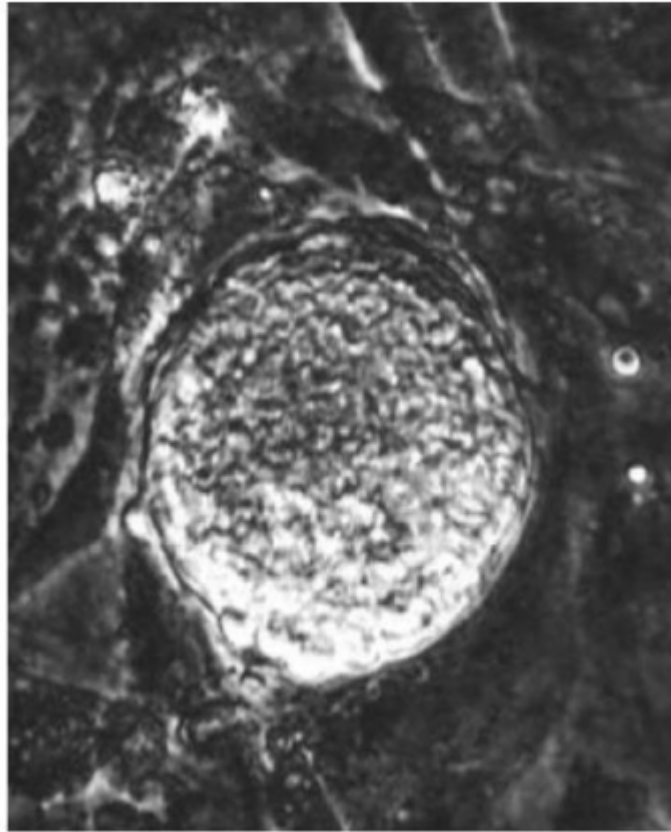


Figura 8



Figura 9



Figura 10

Locus de ApoE de rata y posición de los sitios de corte de los dedos de zinc

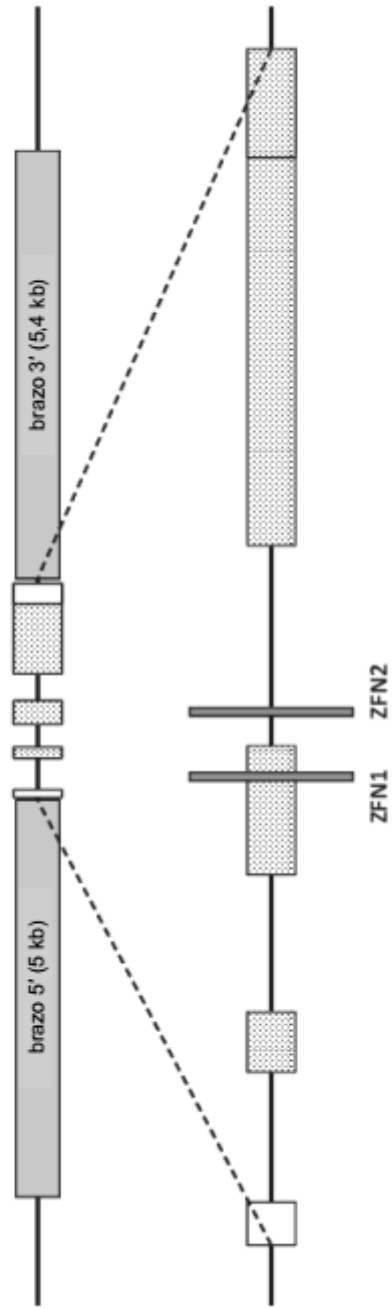
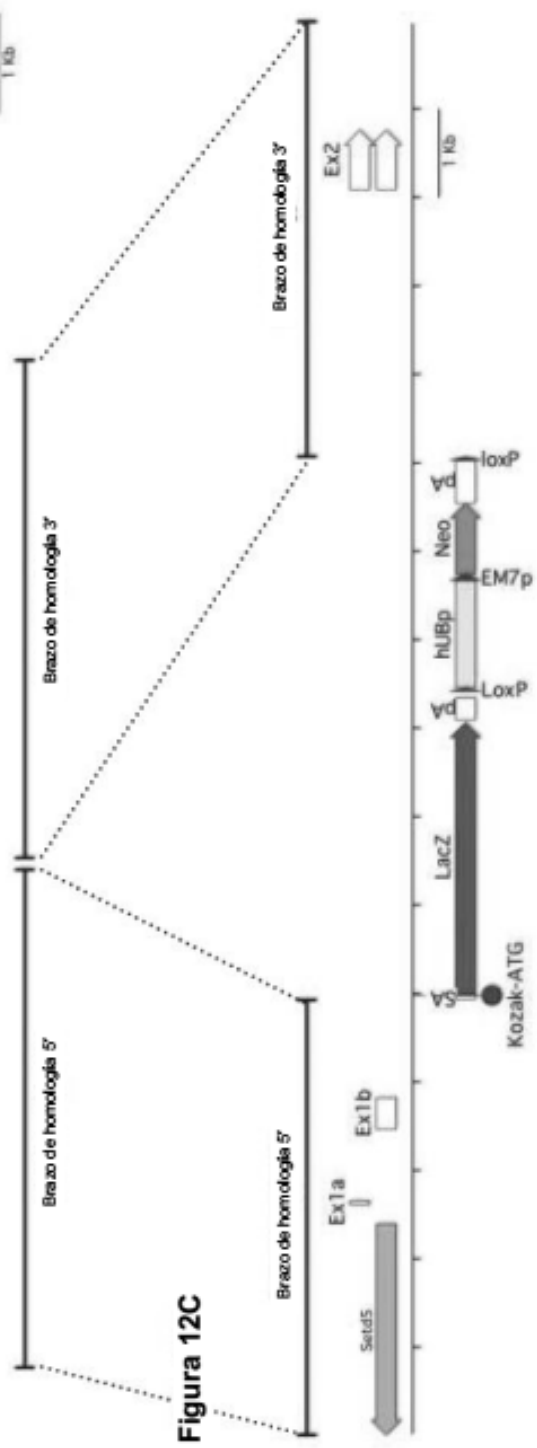
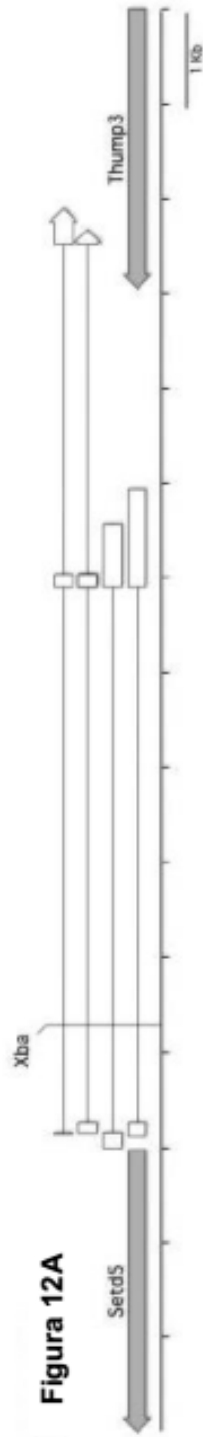


Figura 11



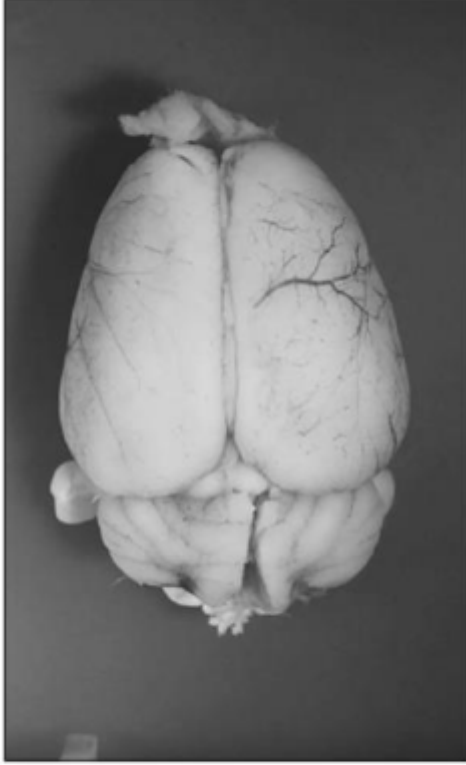


Figura 13A



Figura 13B

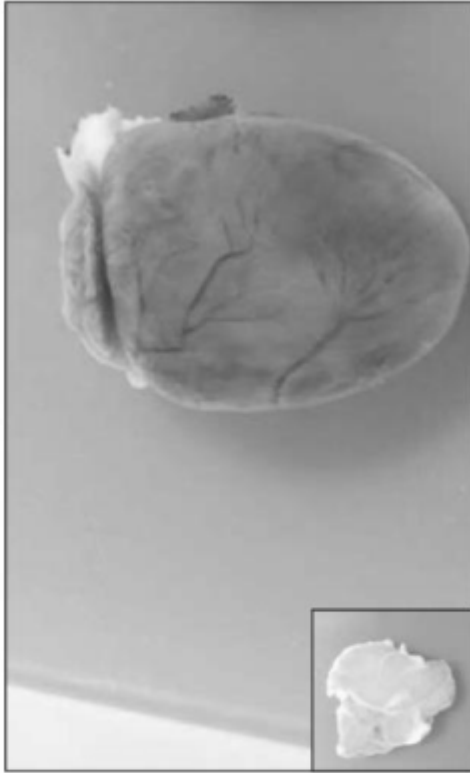


Figura 13C

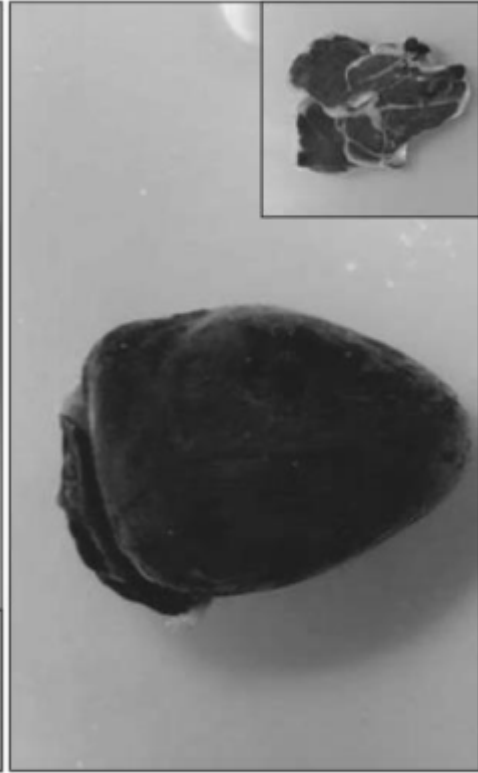


Figura 13D

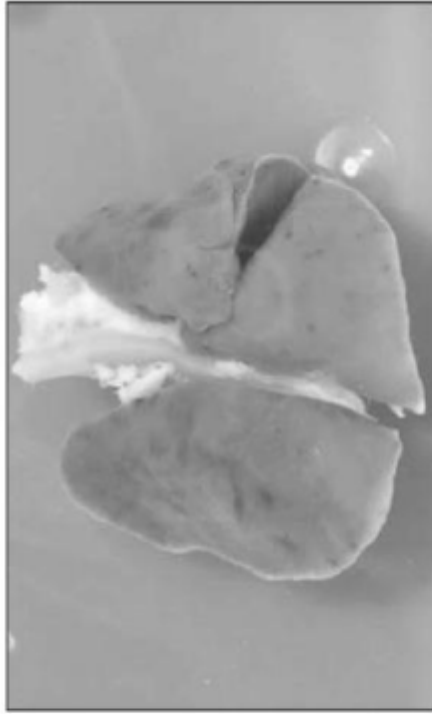


Figura 13E

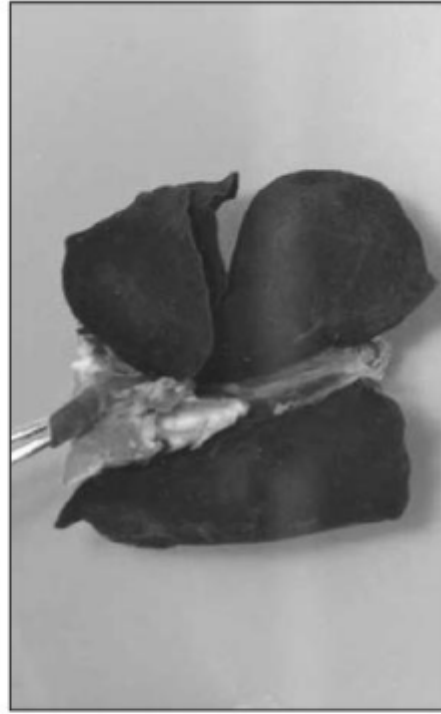


Figura 13F

Figura 13H

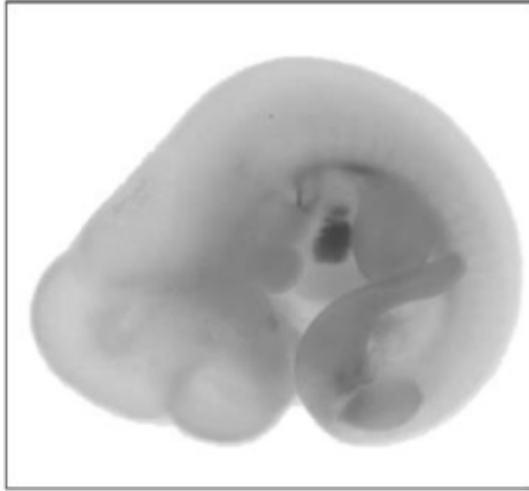


Figura 13G

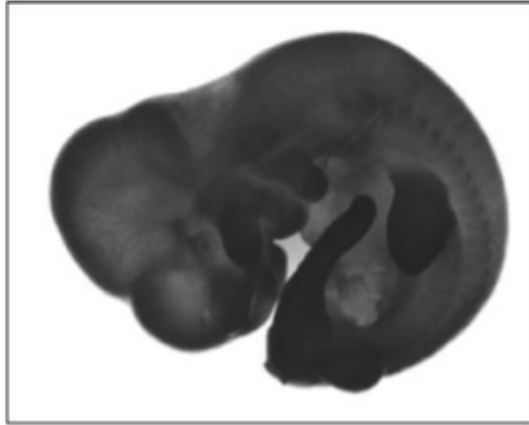


Figura 13J

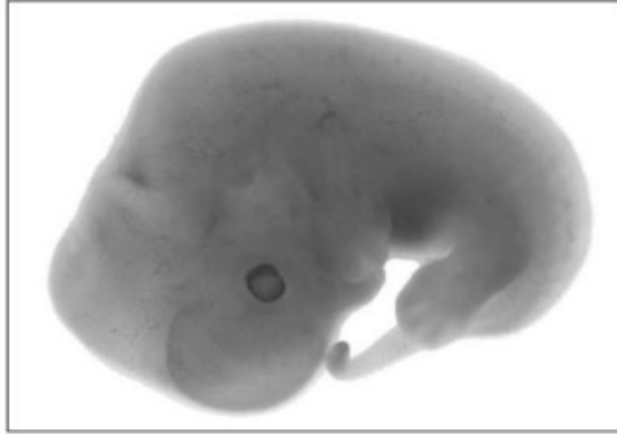


Figura 13I



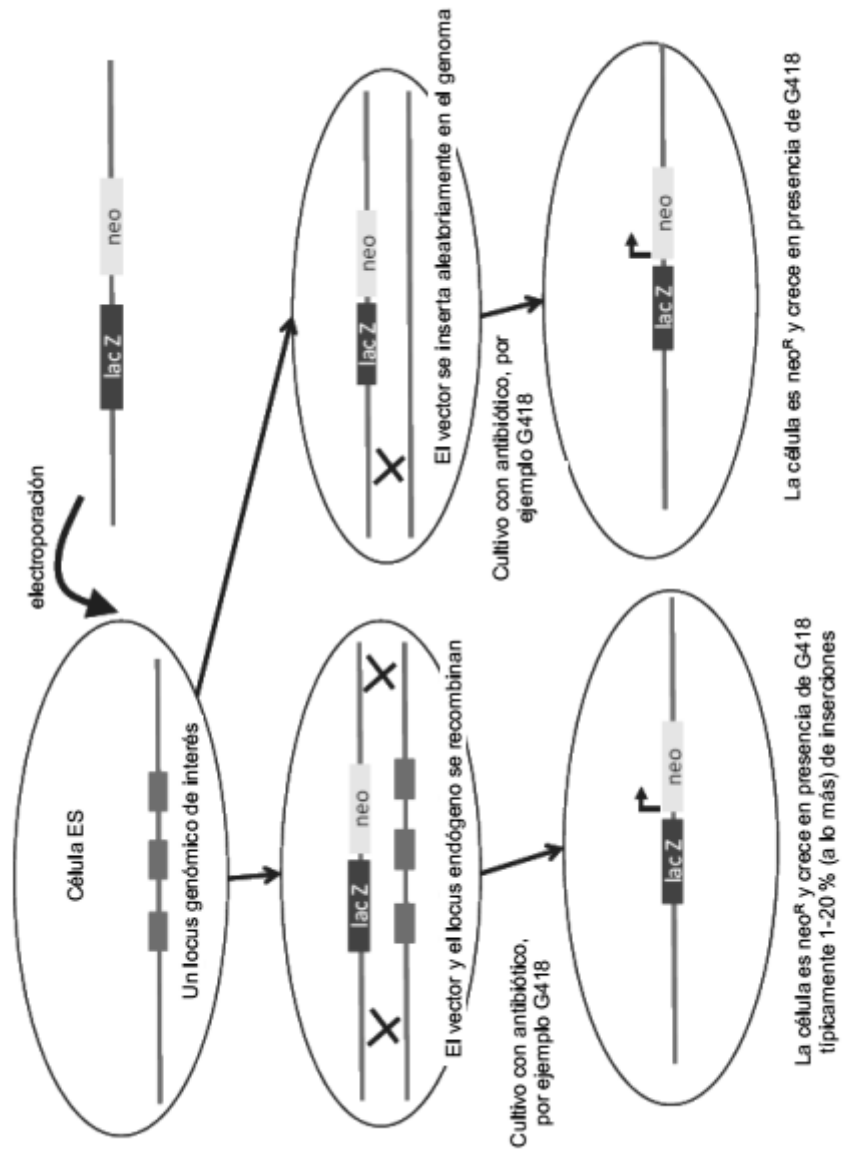


Figura 14

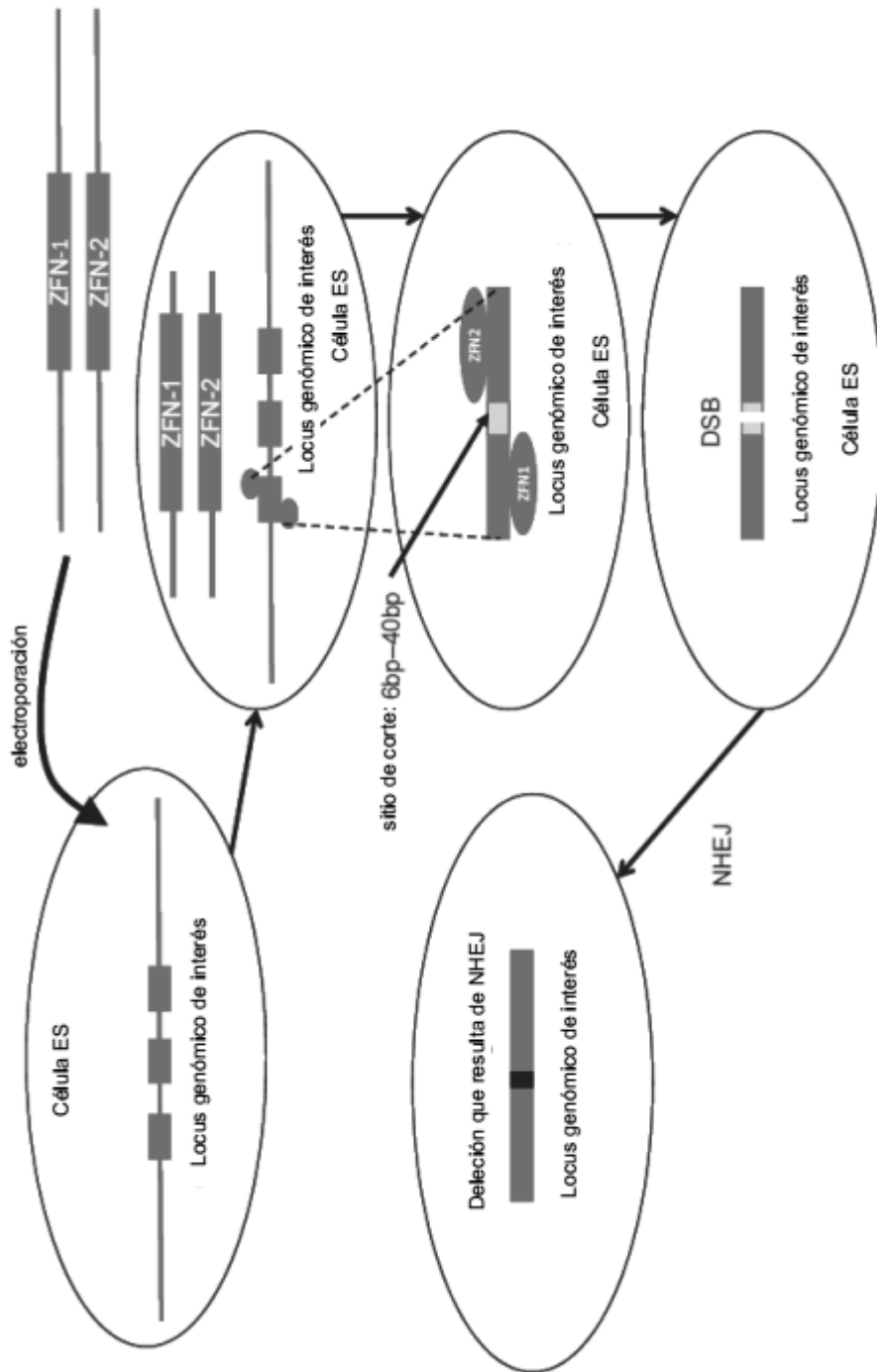


Figura 15

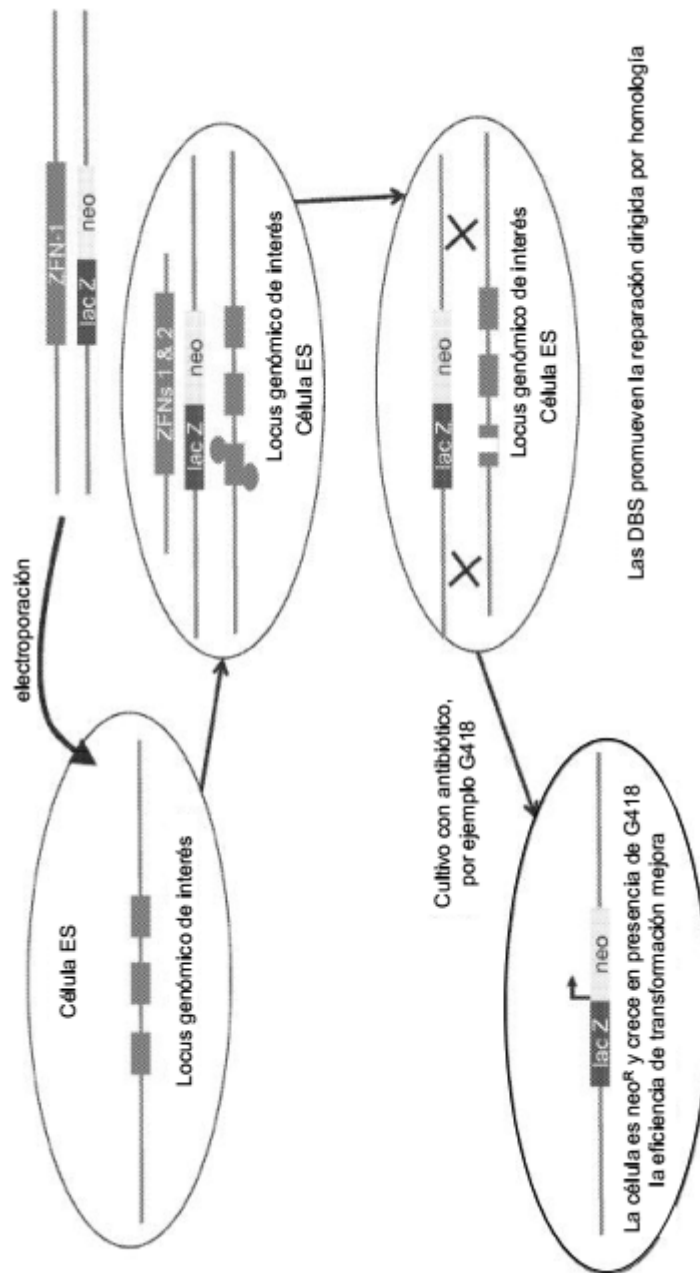


Figura 16



Figura 17

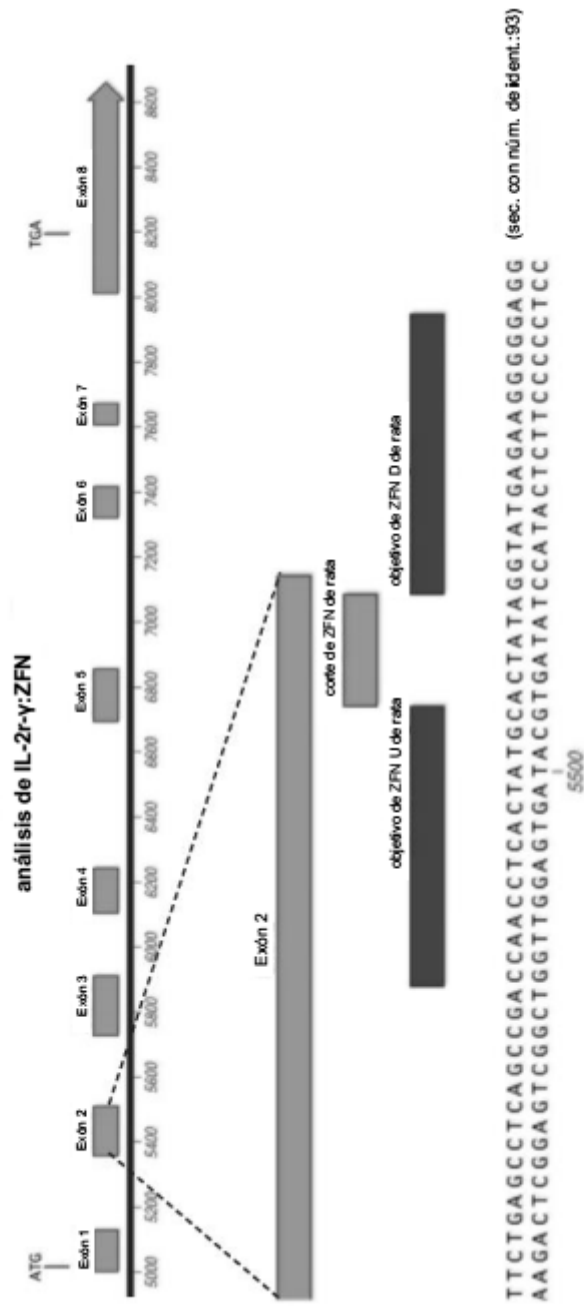
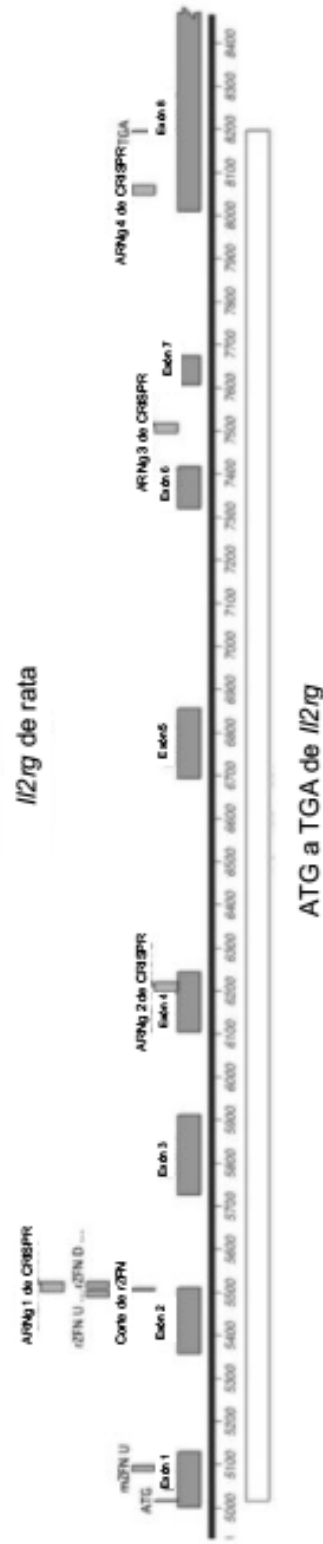


Figura 18

Figura 19



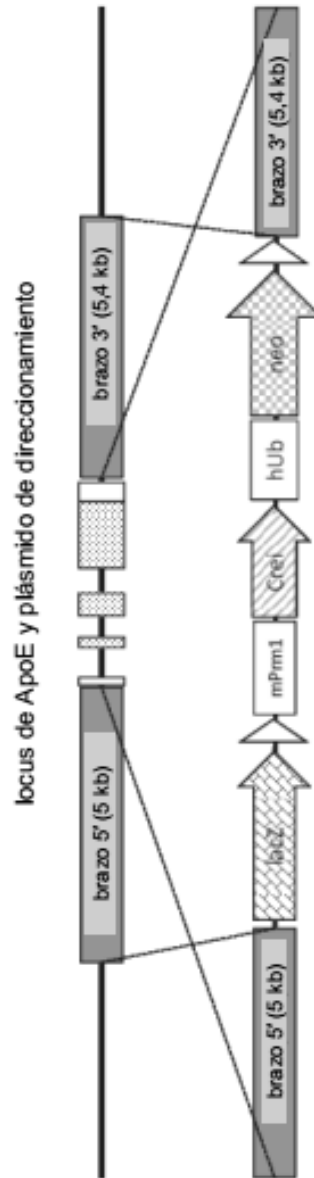
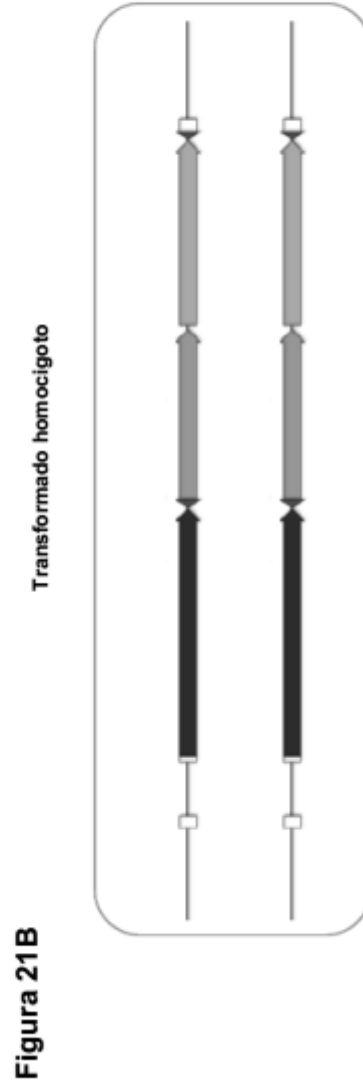
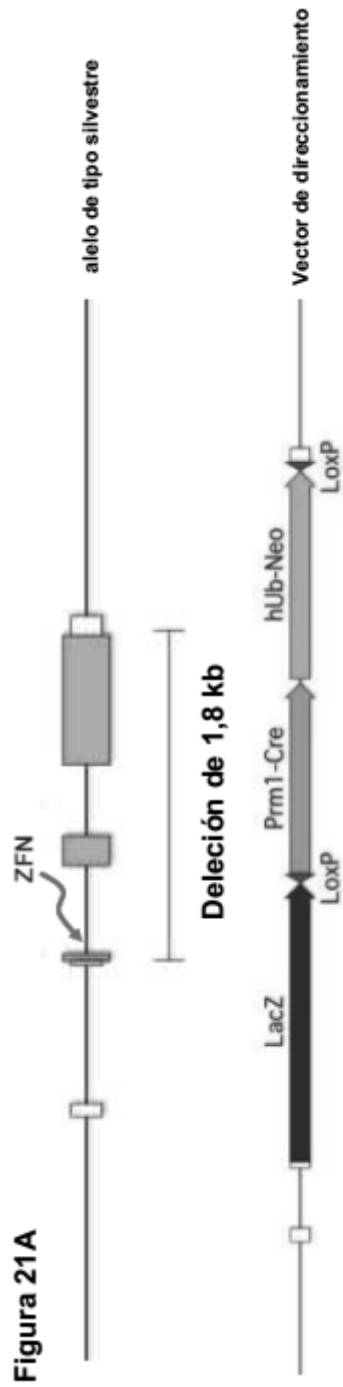


Figura 20



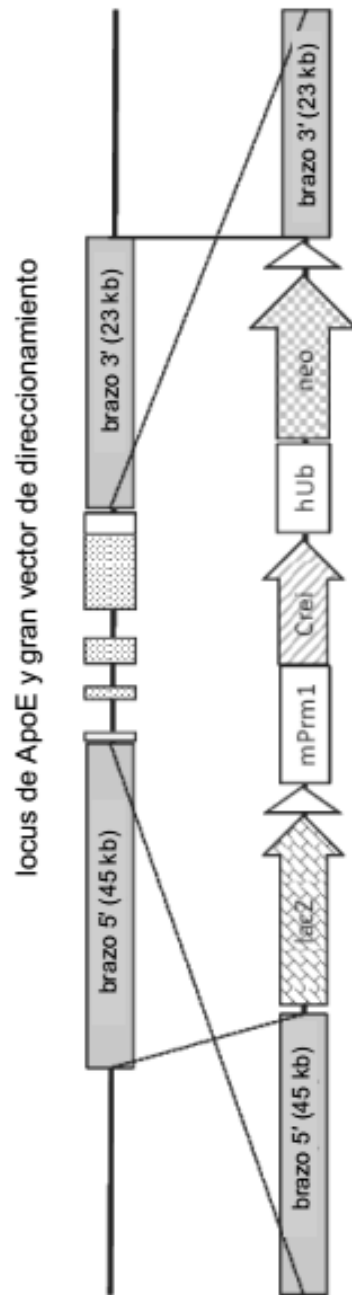


Figura 22

locus de ApoE de rata y posición de los sitios de corte de los dedos de zinc

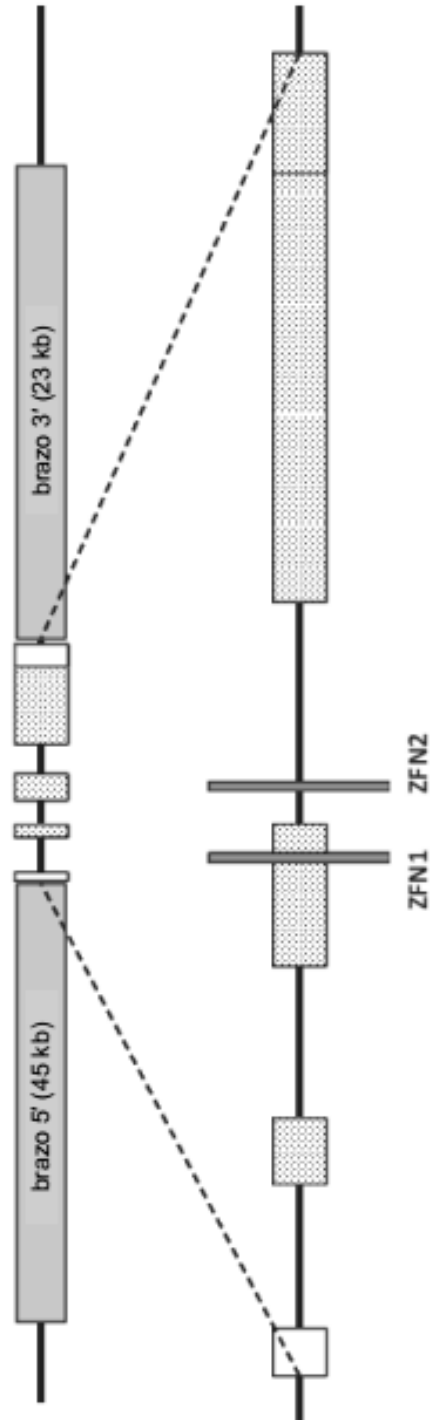


Figura 23

Delección de Il2rg

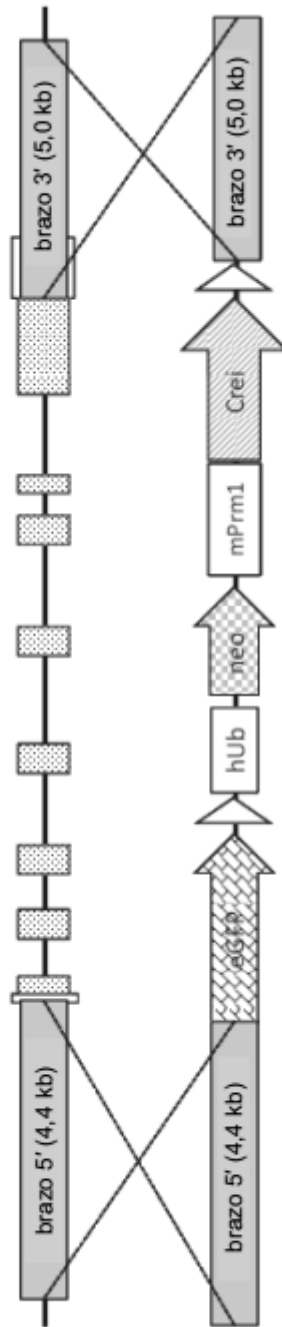


Figura 24

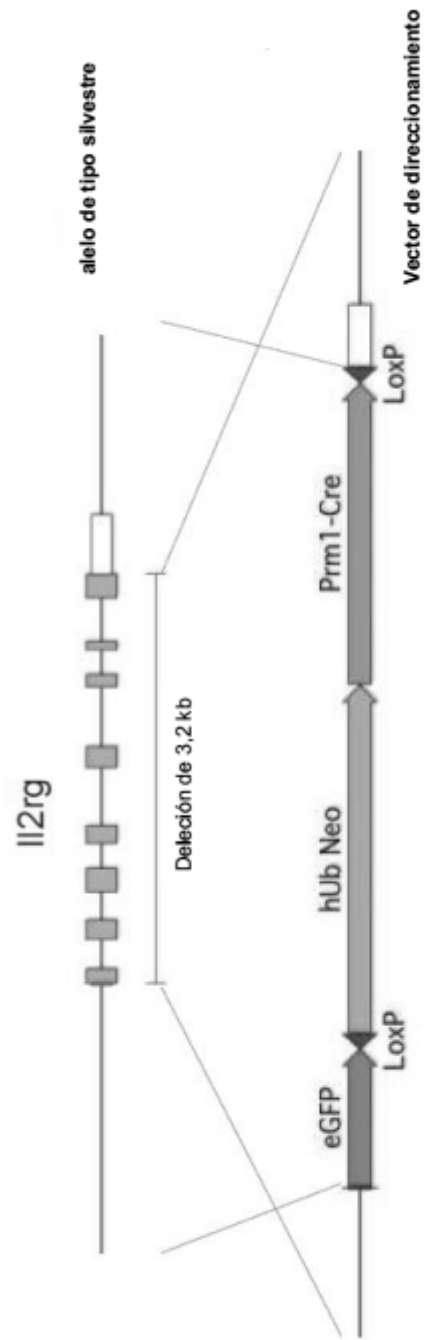


Figura 25

locus de Rag2 de rata y gran vector de direccionamiento

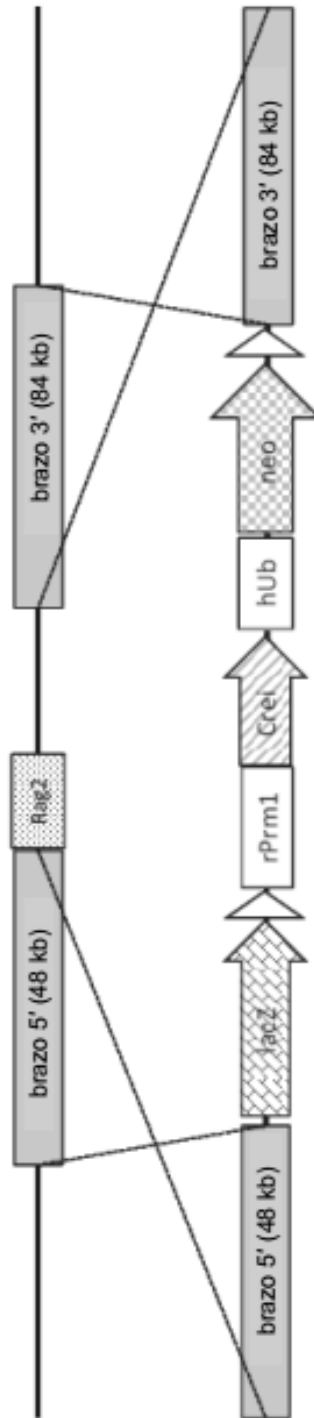


Figura 26

Transformación de Rag2/Rag1

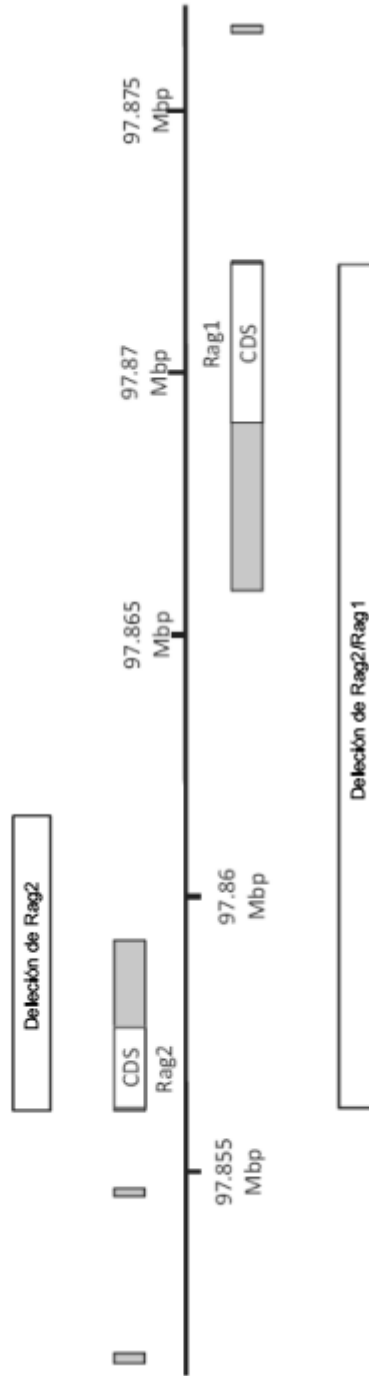


Figura 27

Transformación de Rag2/Rag1

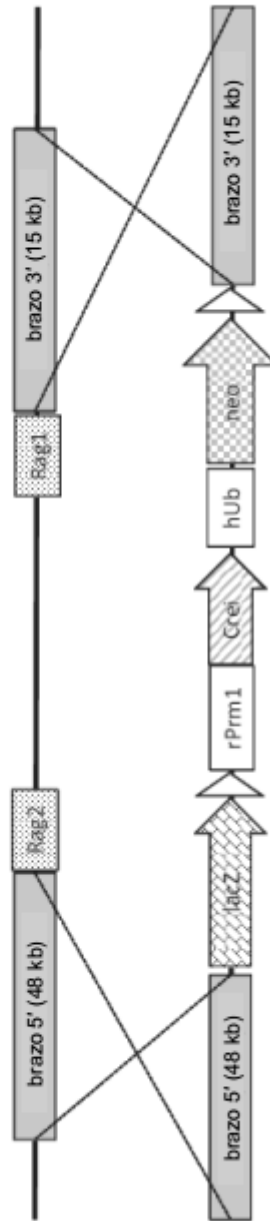
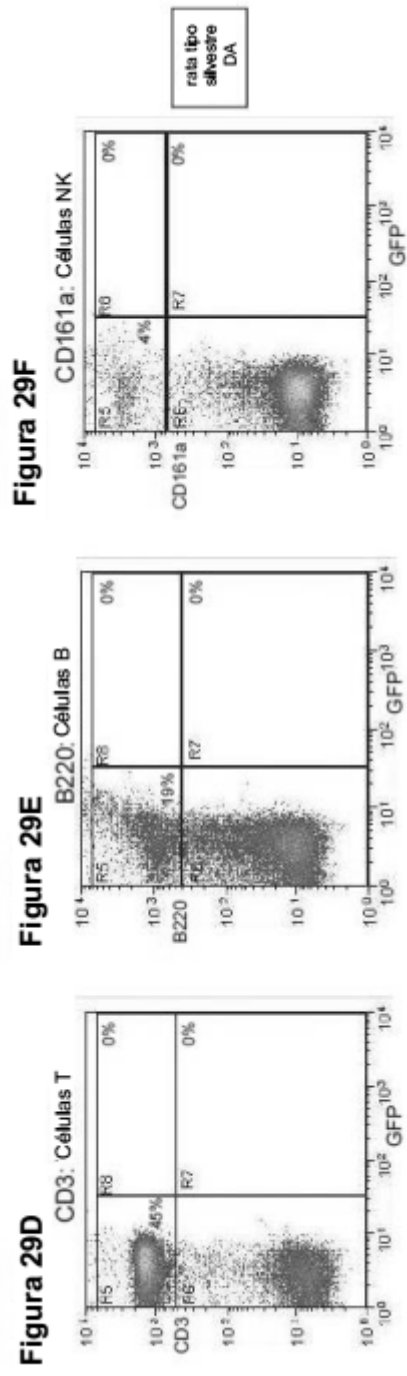
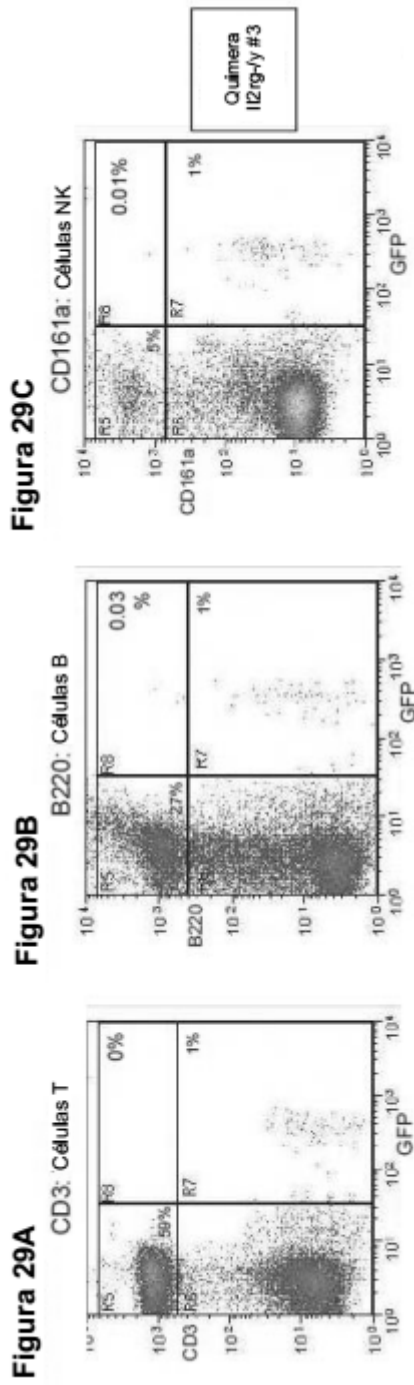


Figura 28



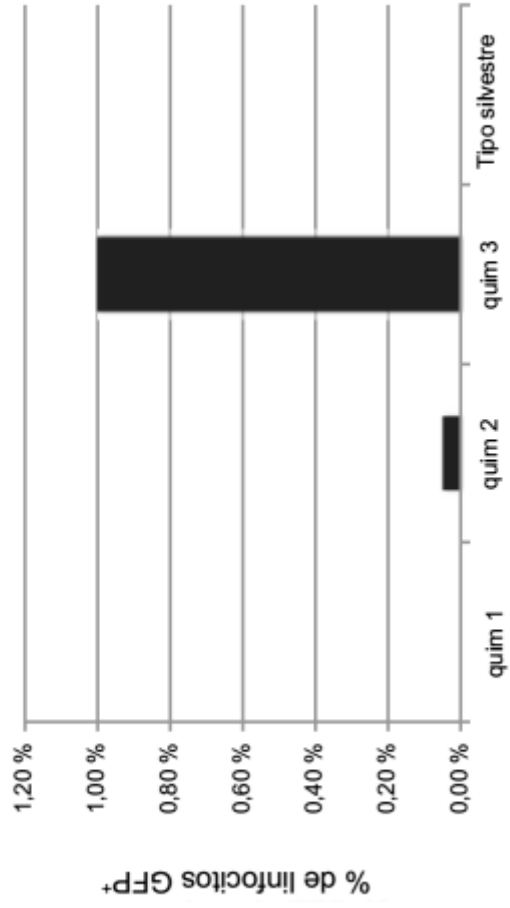


Figura 30

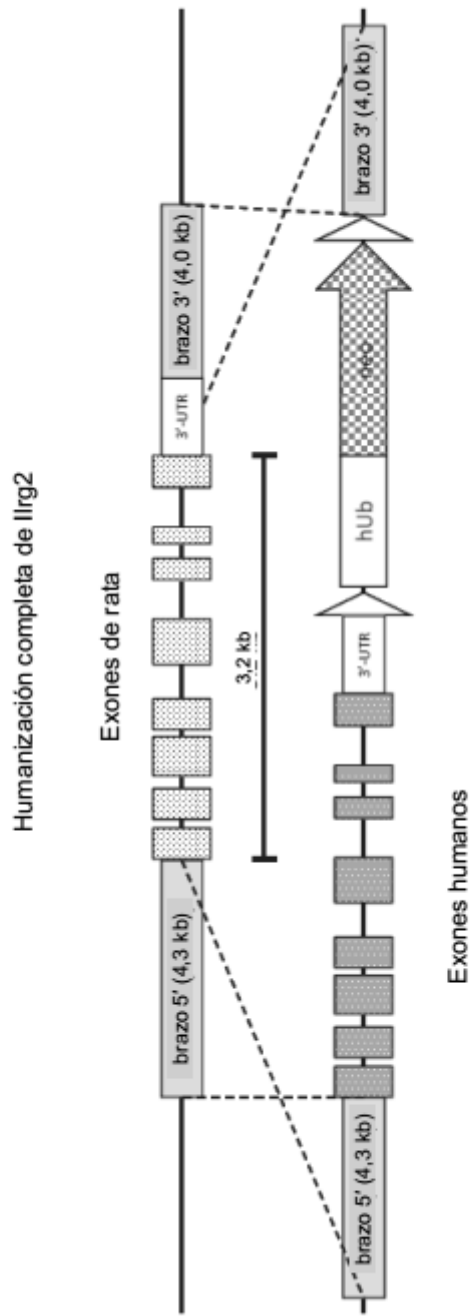


Figura 31

Humanización en ecto de Ilrg2

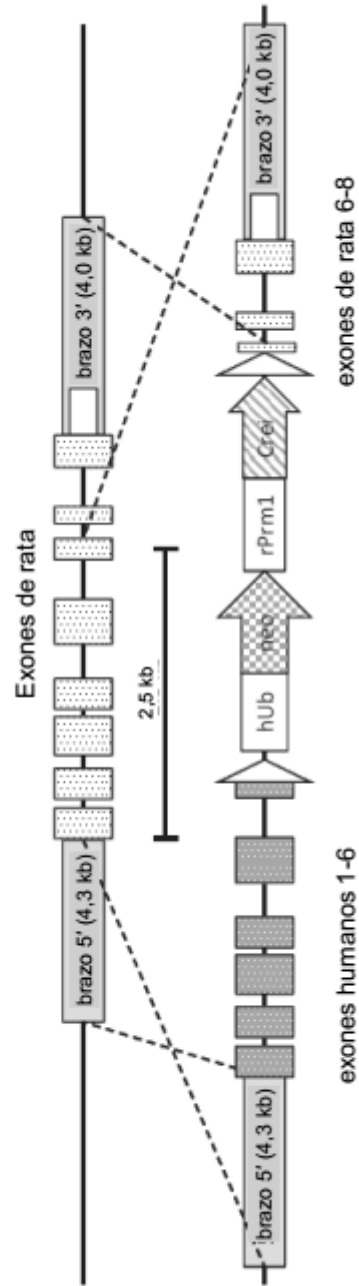
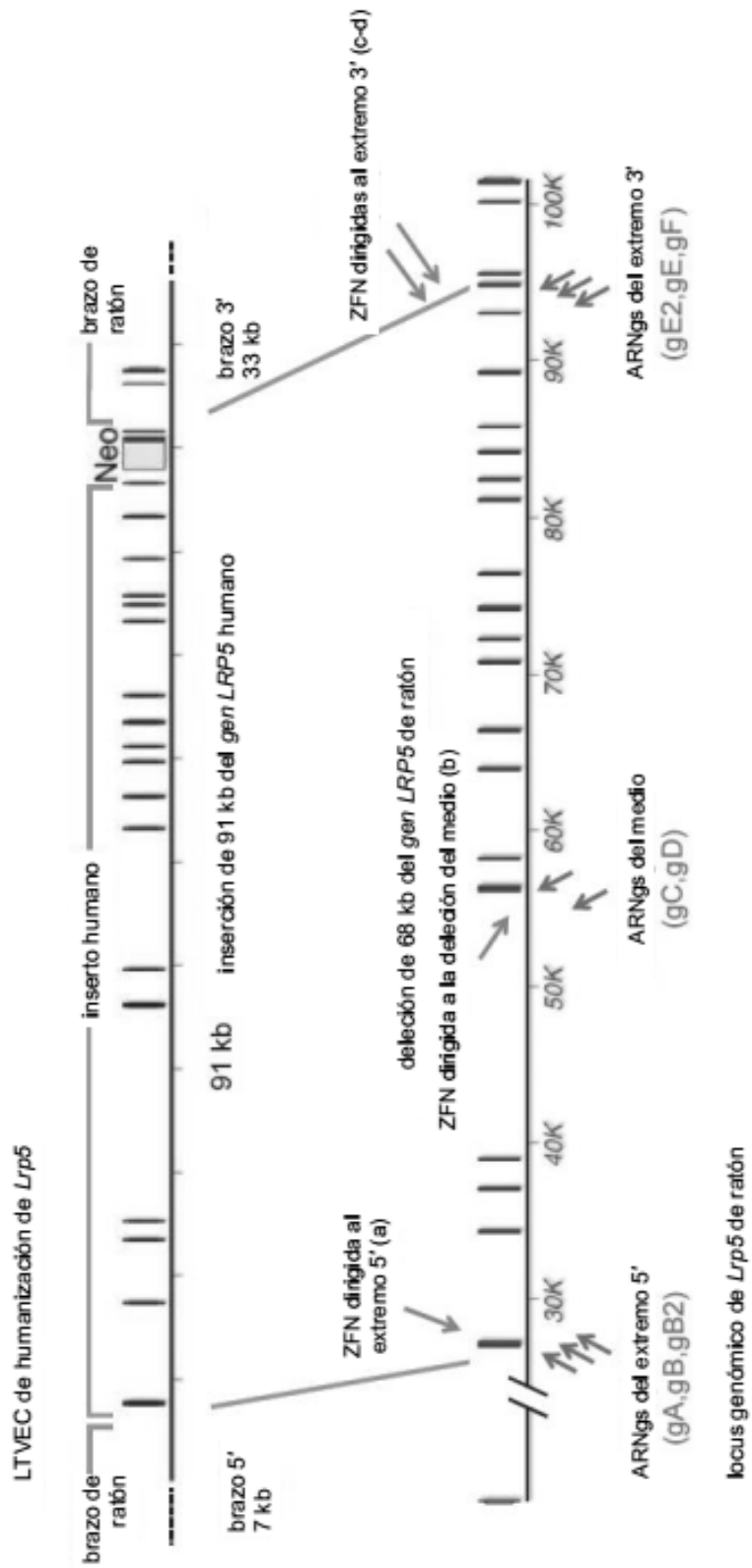


Figura 32

Figura 34



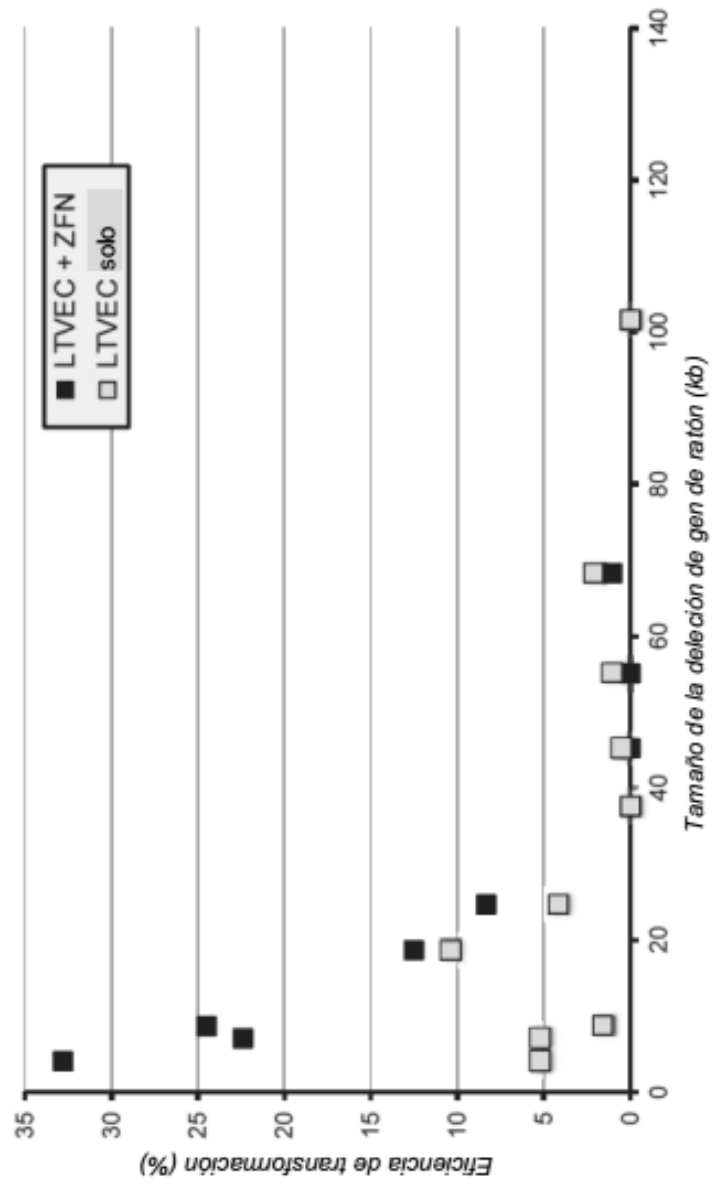


Figura 35A

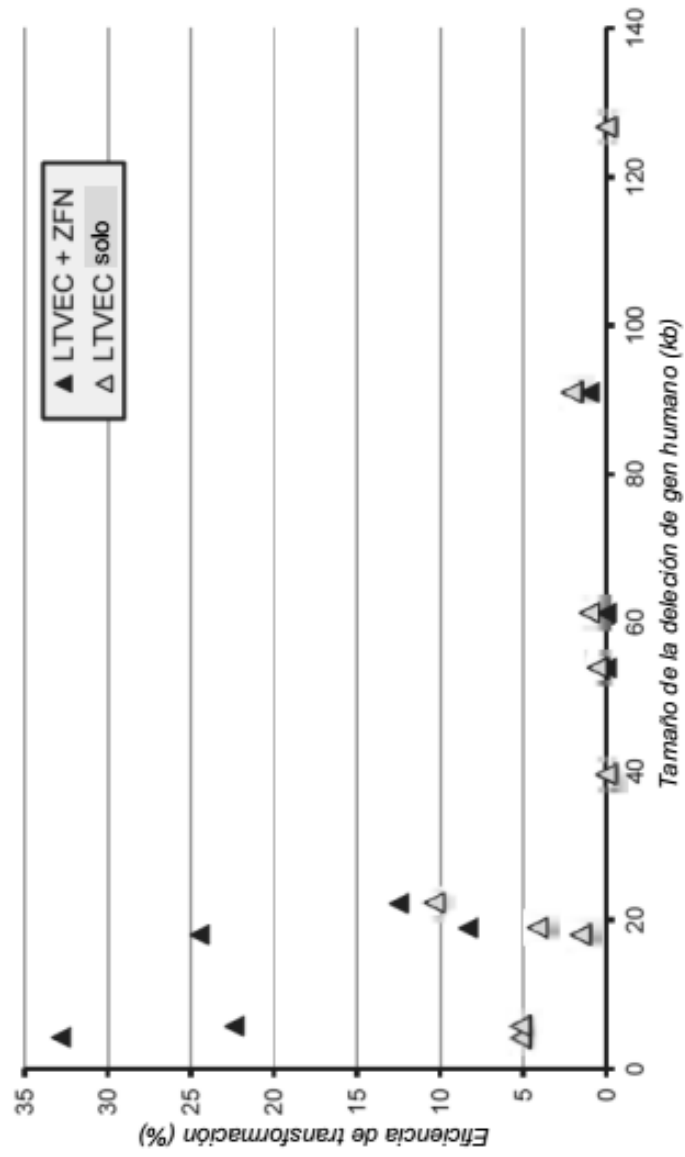


Figura 35B

LTVEC de humanización de *Trpa1*

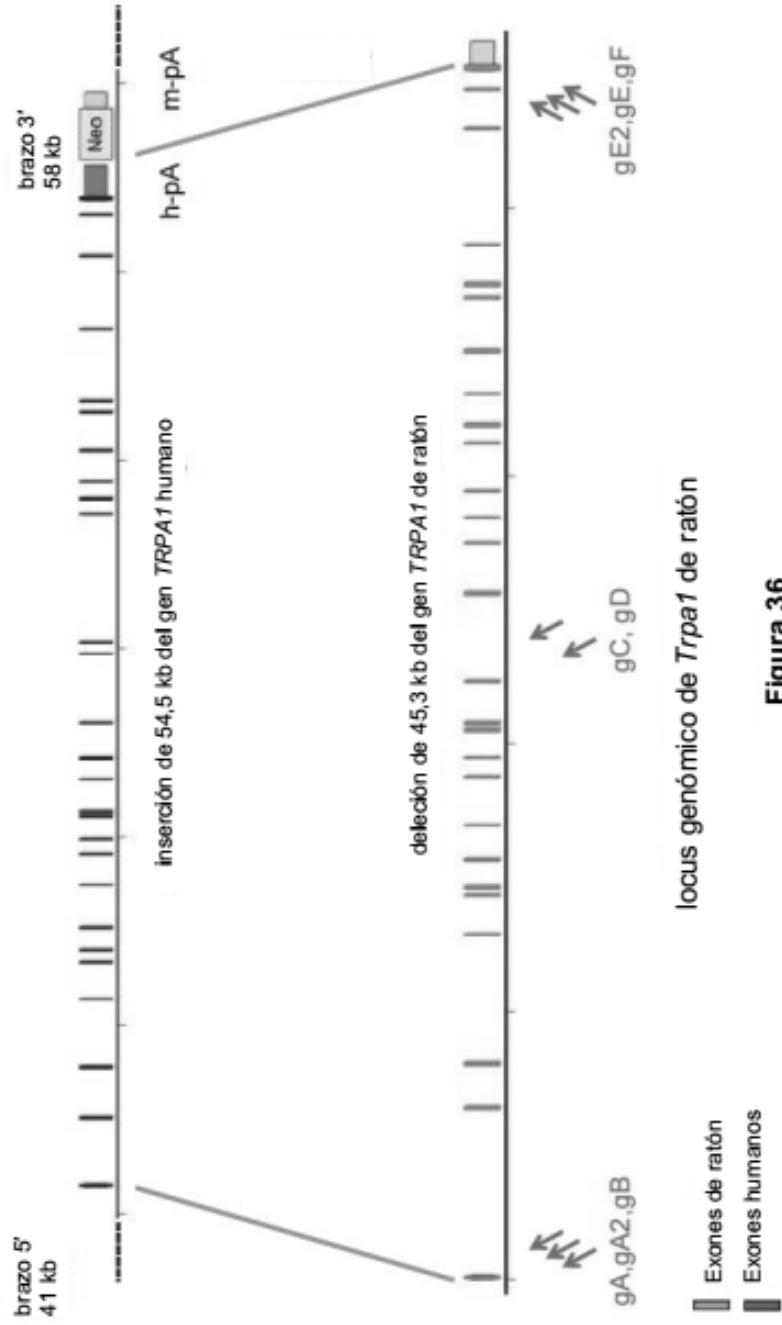


Figura 36

LTVEC de humanización de *Folh1*

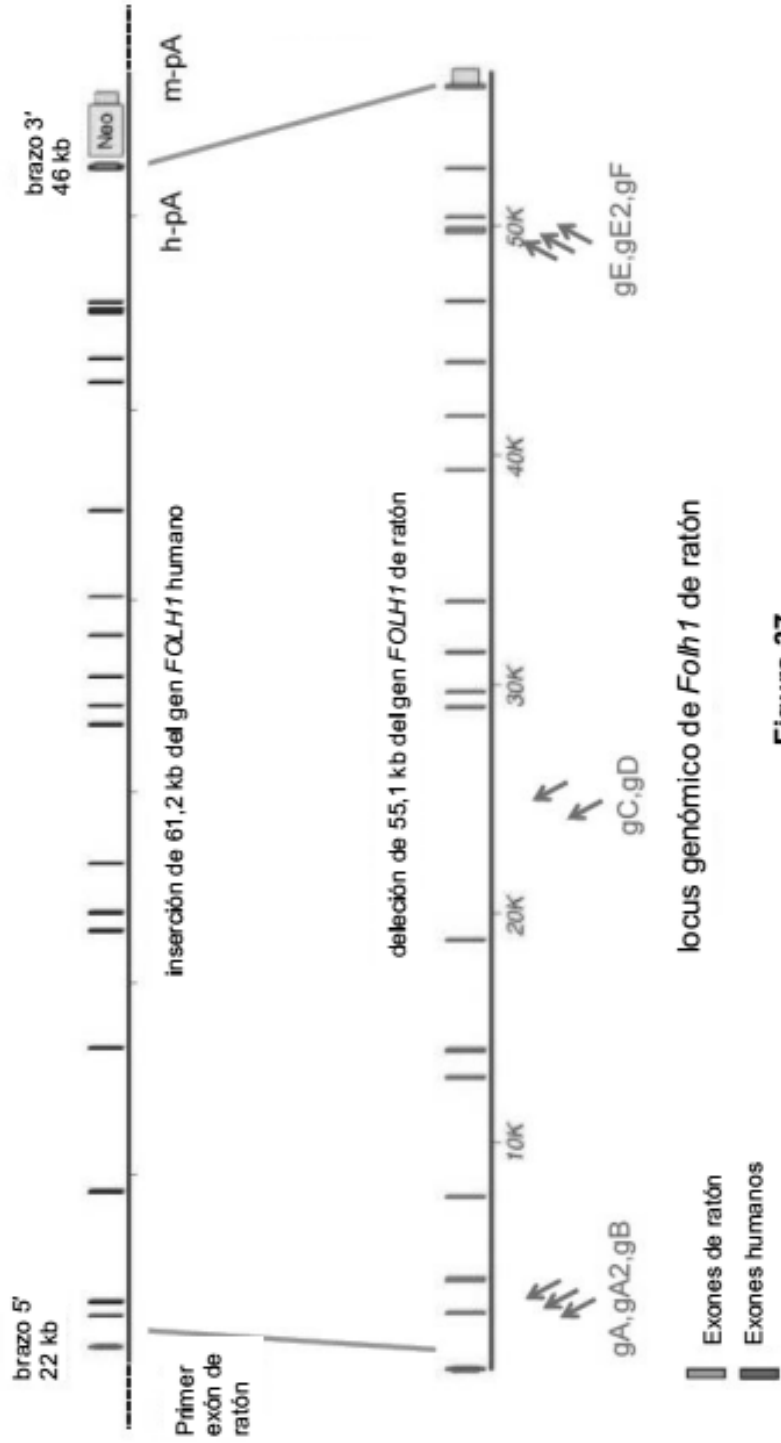


Figura 37

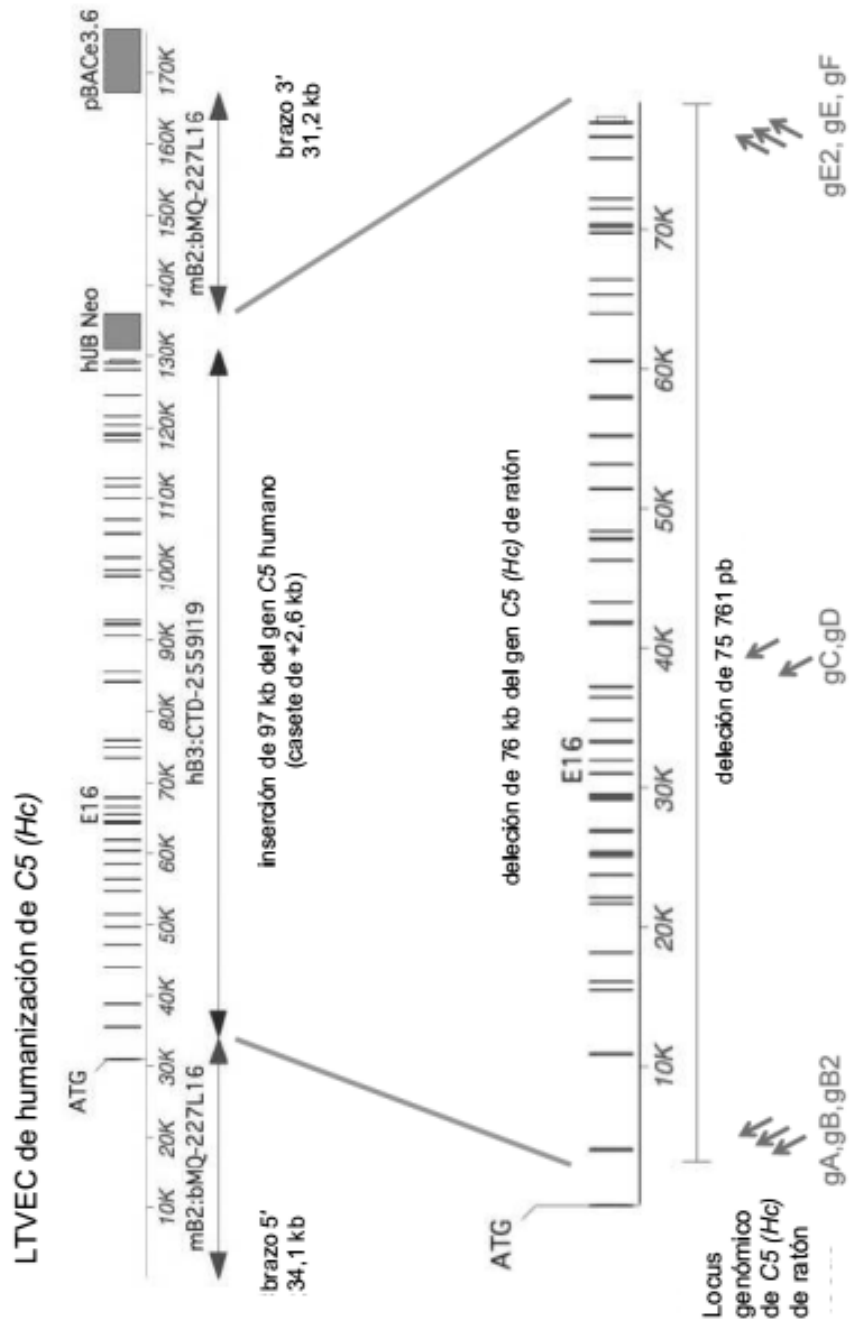


Figura 38

LTVEC de humanización de *Adamts5*

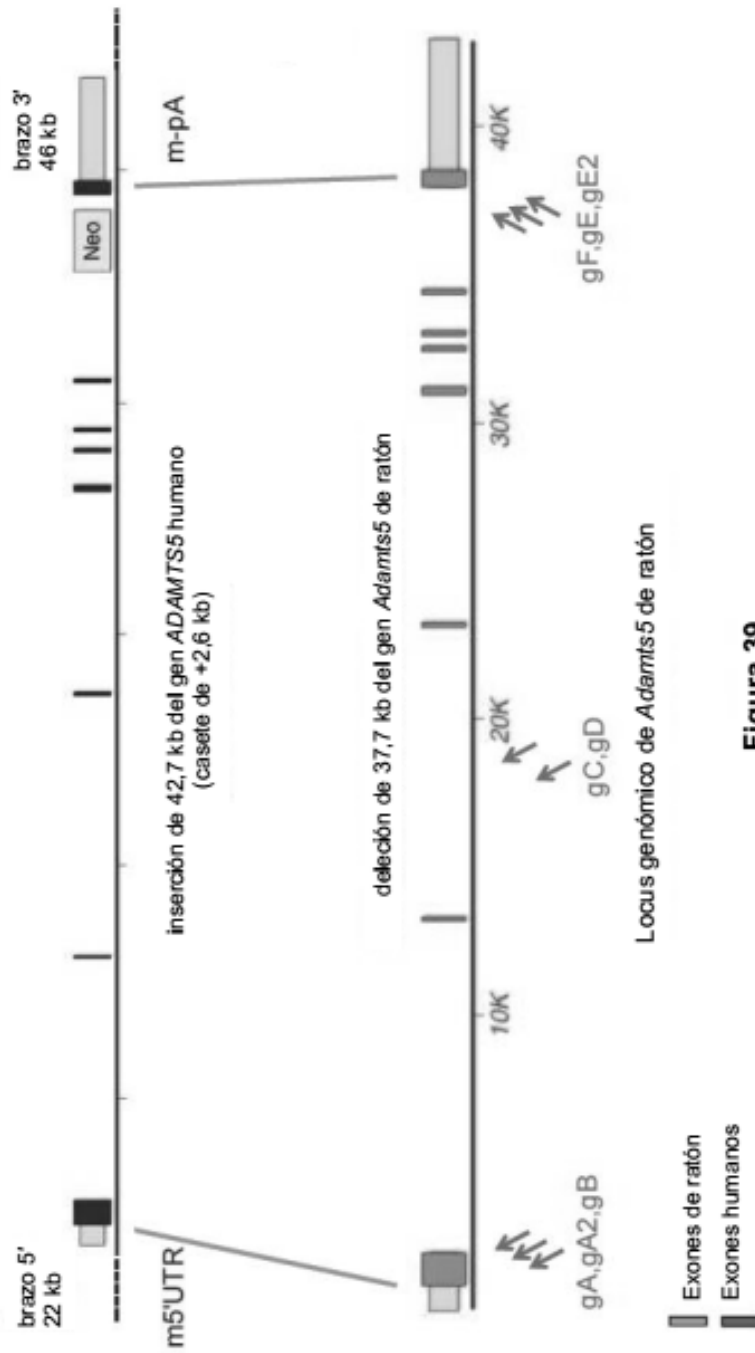


Figura 39

LTVEC de humanización de *ErbB4*

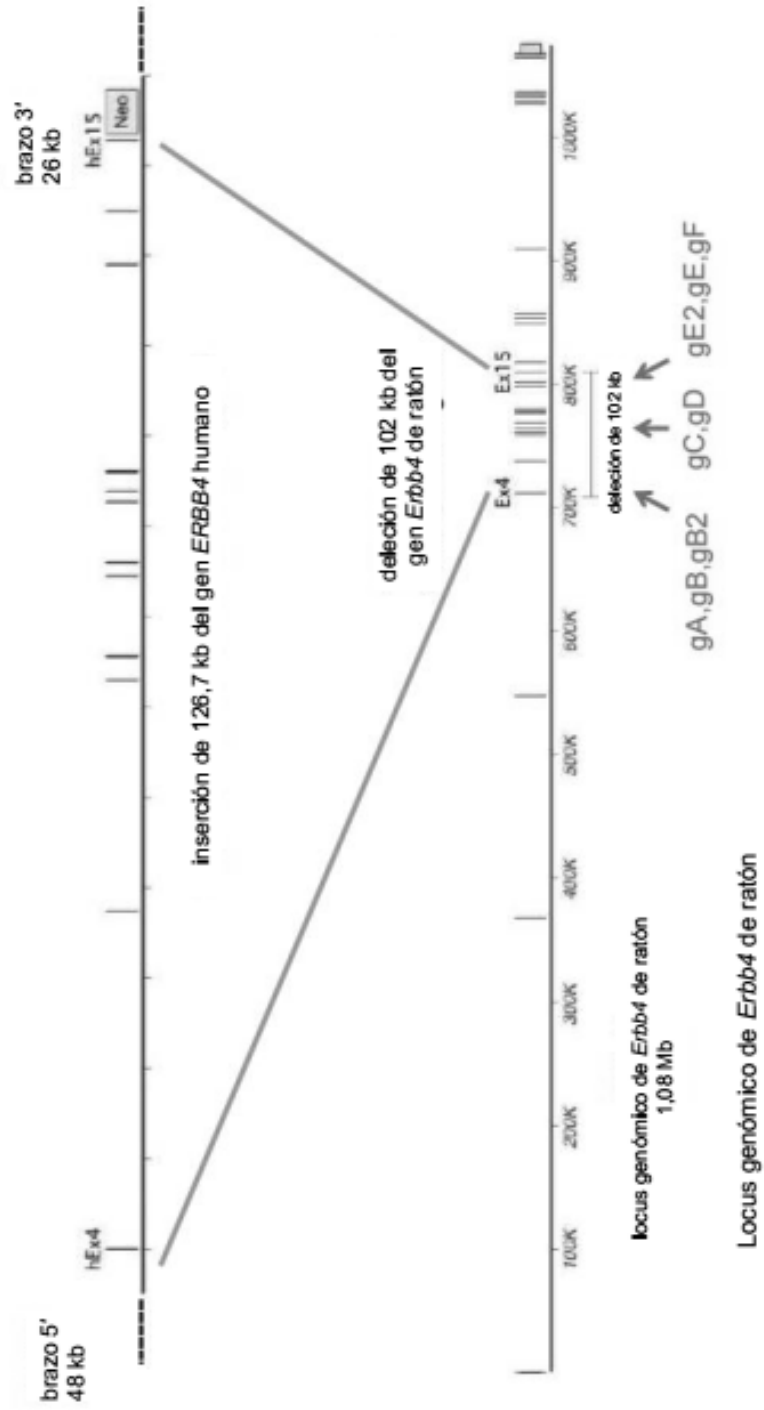


Figura 40

LTVEC de humanización de *Ror1*

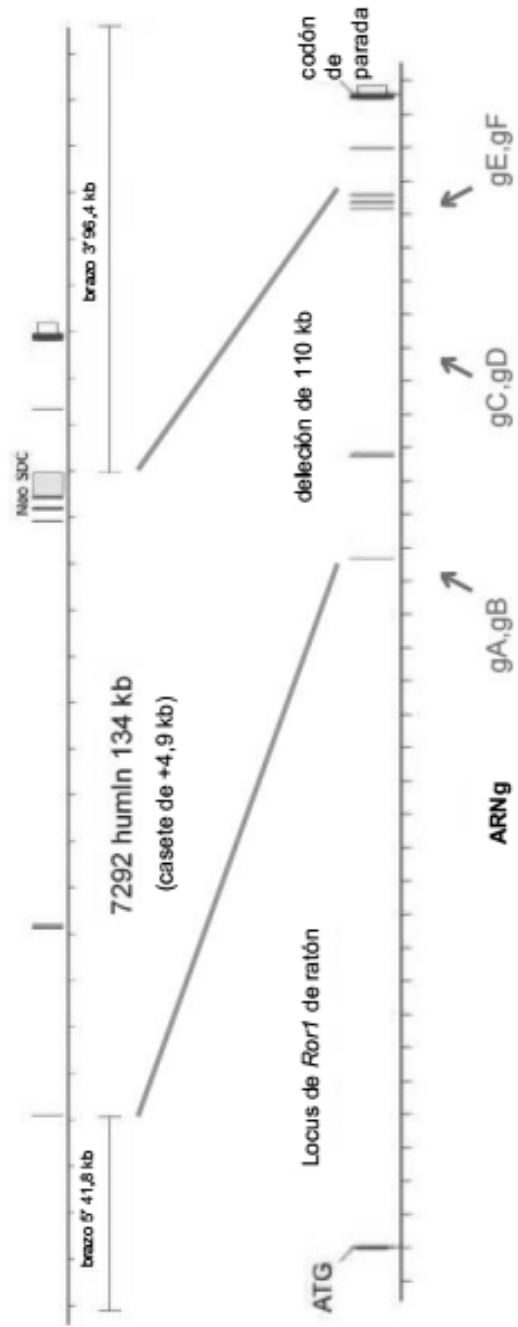


Figura 41

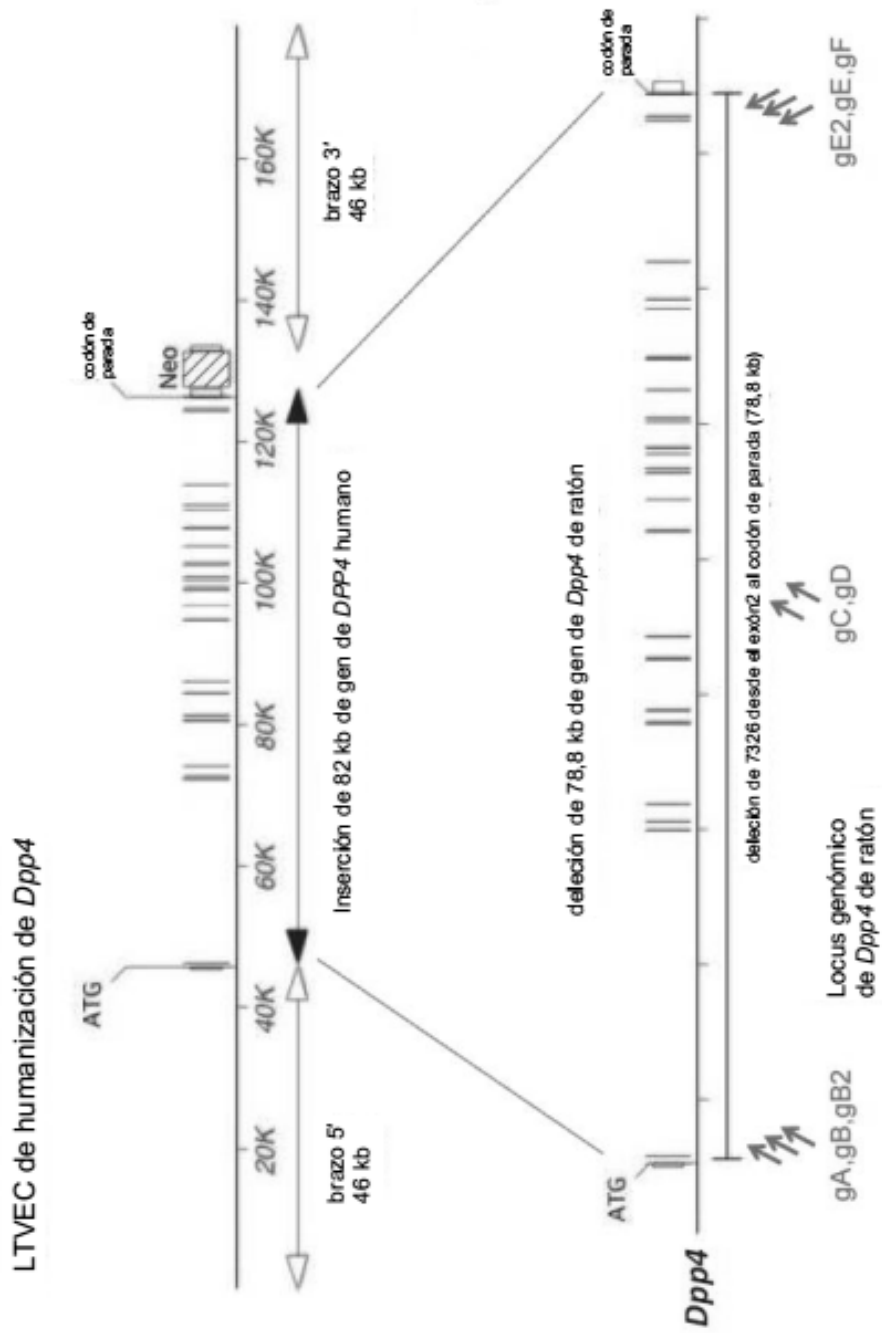


Figura 42

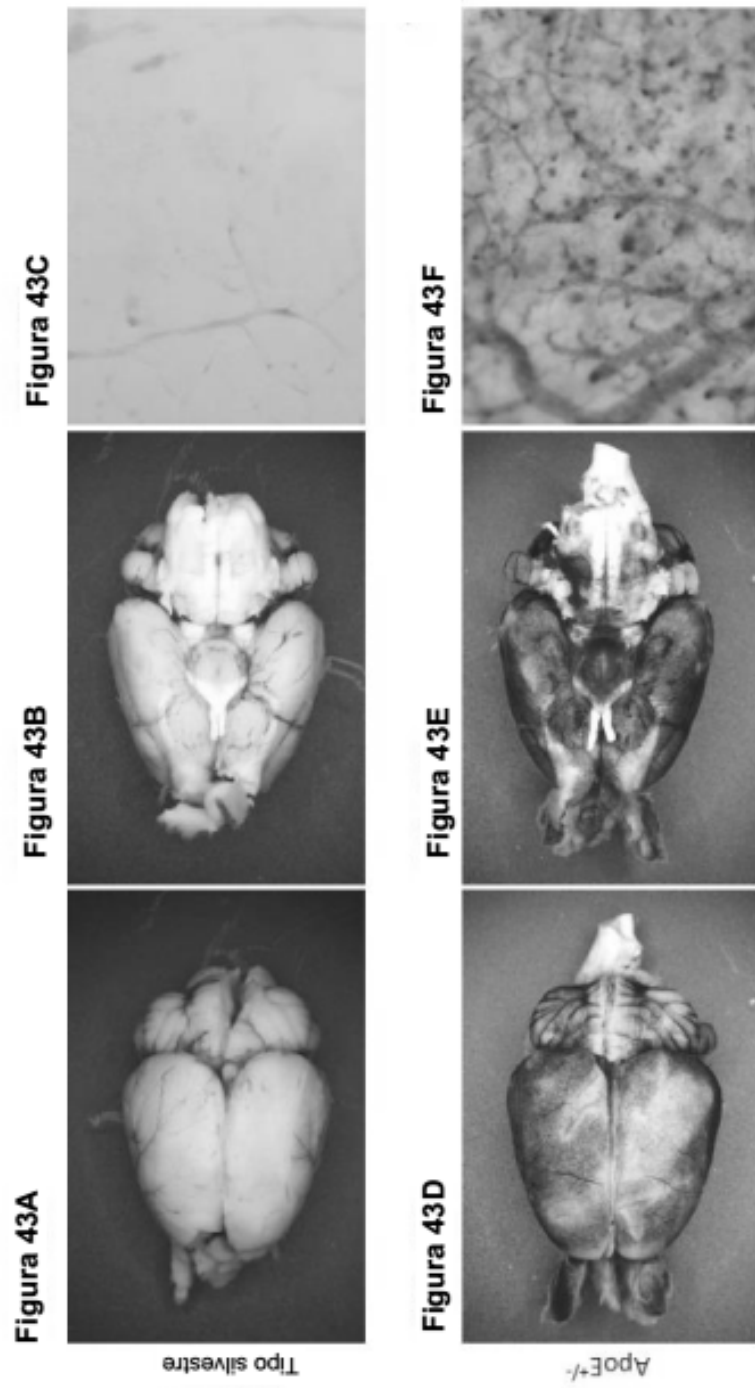


Figura 44B



Figura 44D

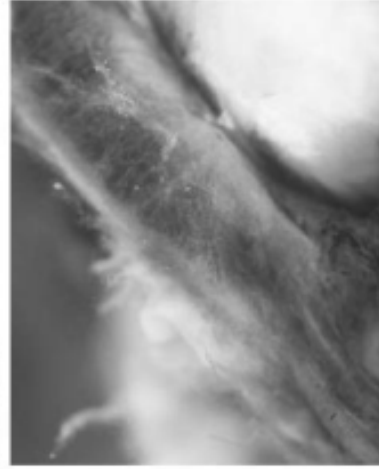
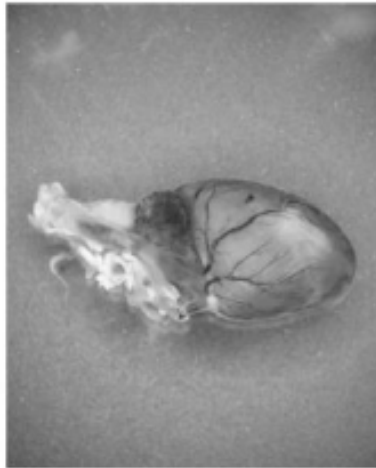
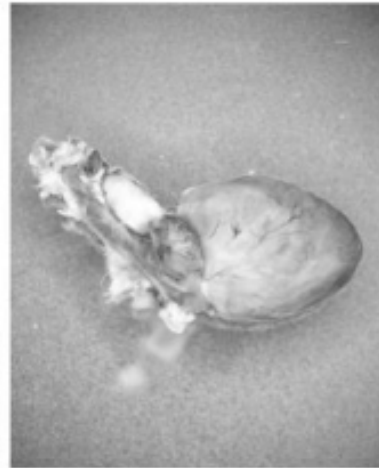


Figura 44A



Tipo silvestre

Figura 44C



ApoE^{+/-}

Figura 45B

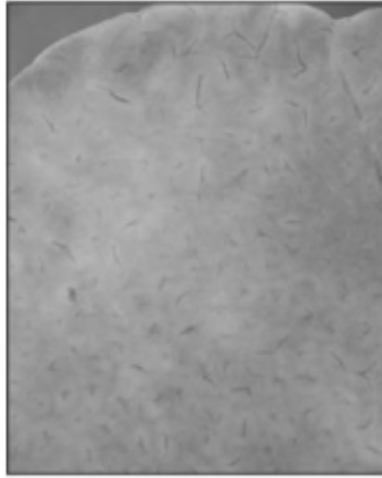
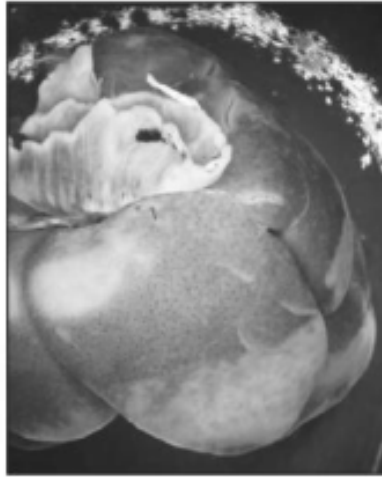


Figura 45D



Figura 45A



Tipo silvestre

Figura 45C



ApoE^{-/-}

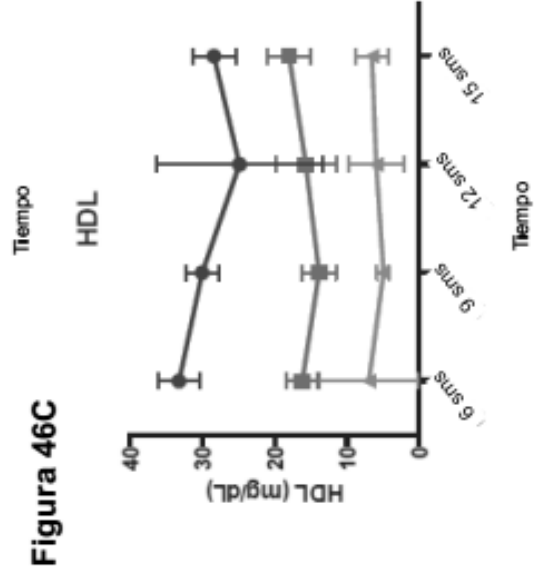
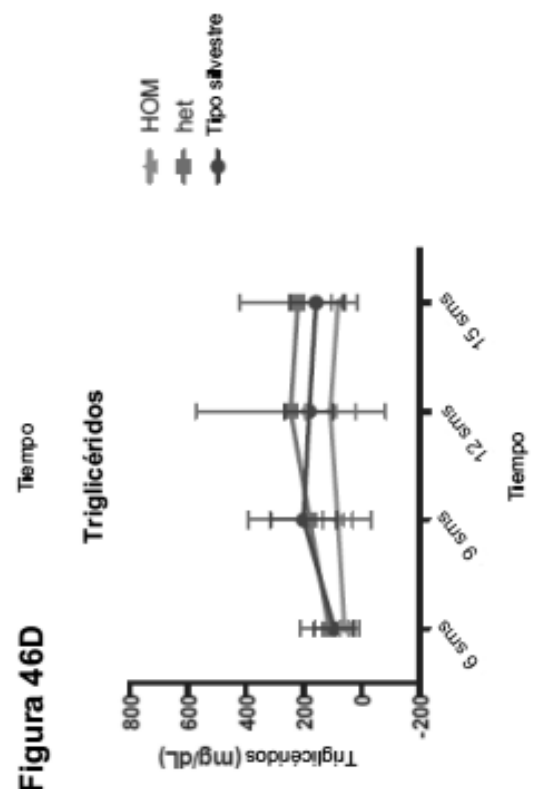
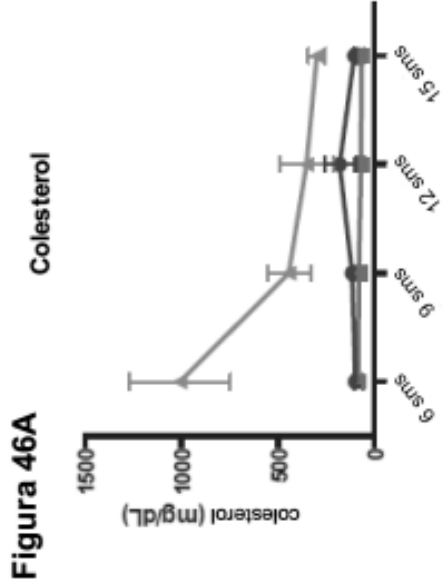
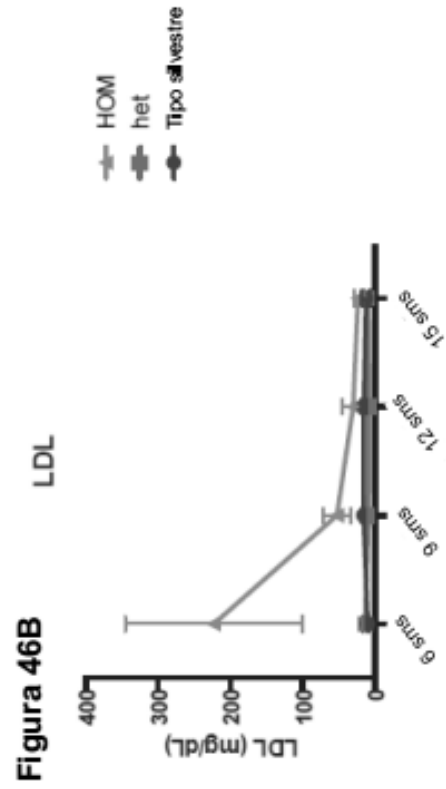


Figura 47

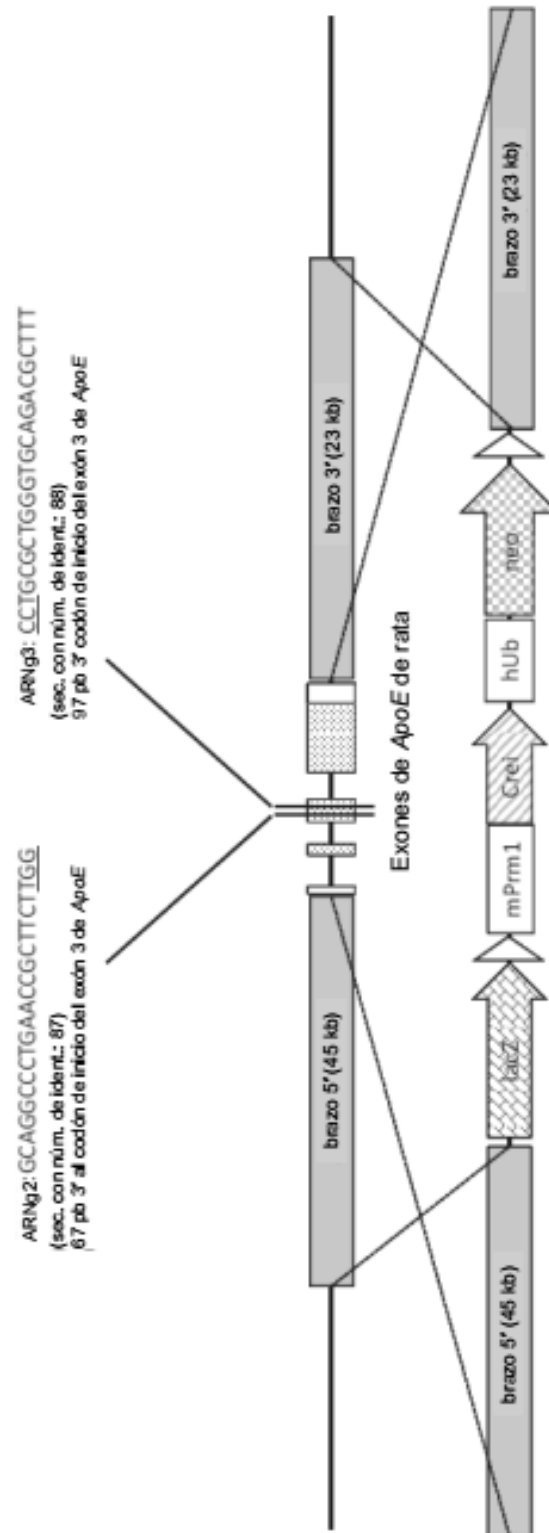


Figura 48

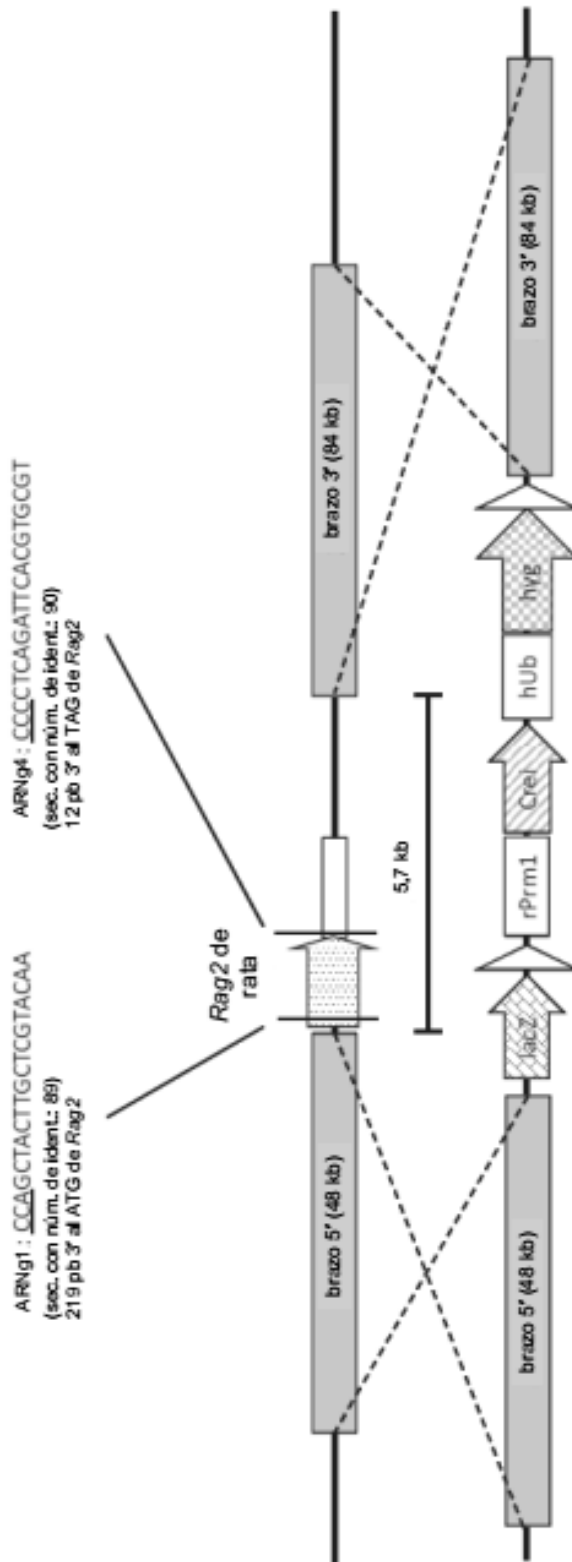
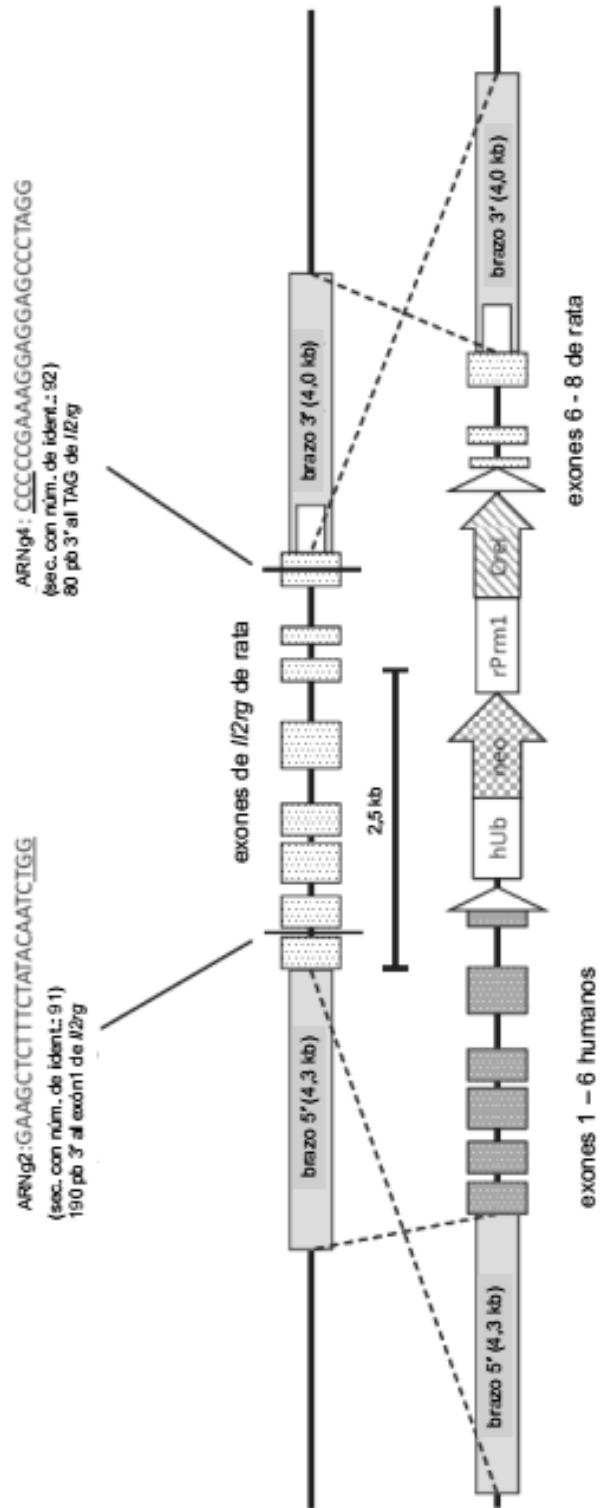


Figura 49



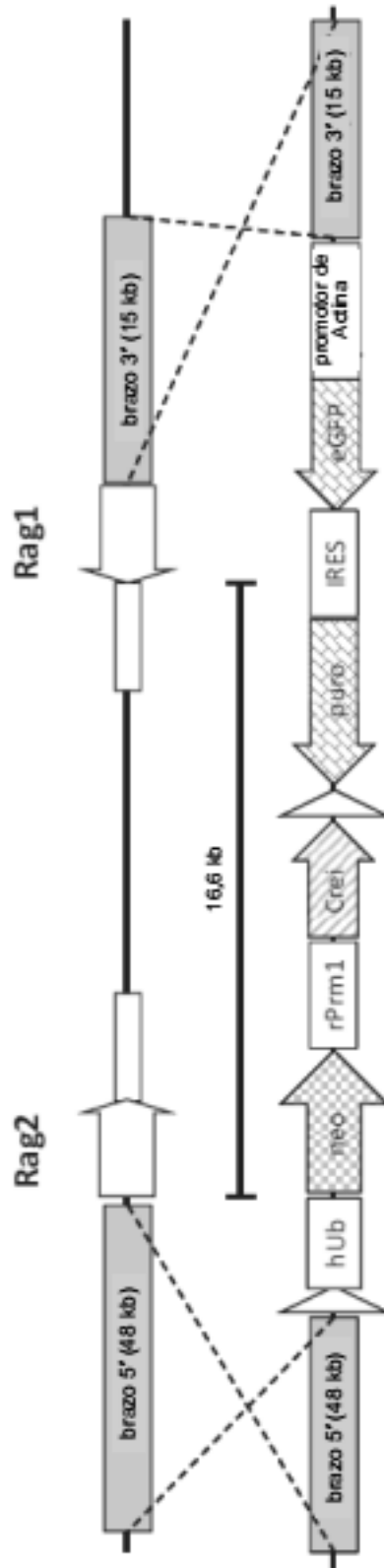


Figura 50

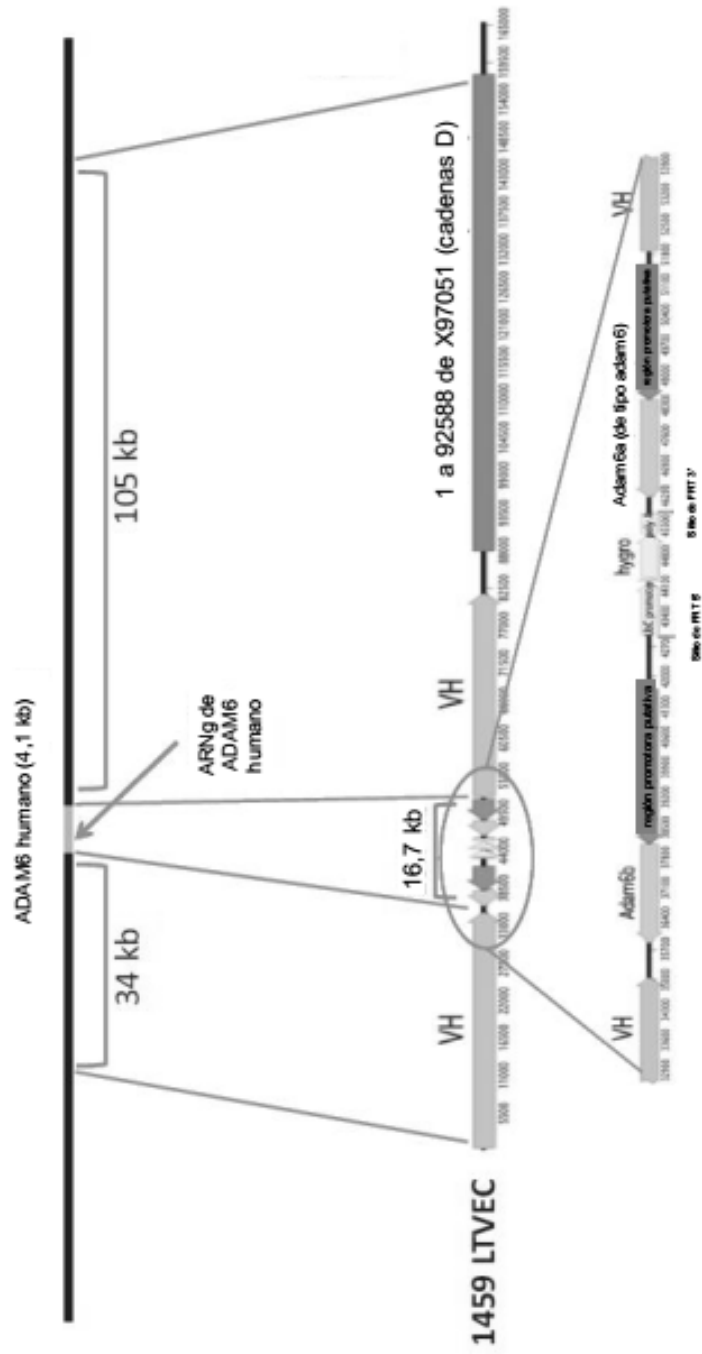
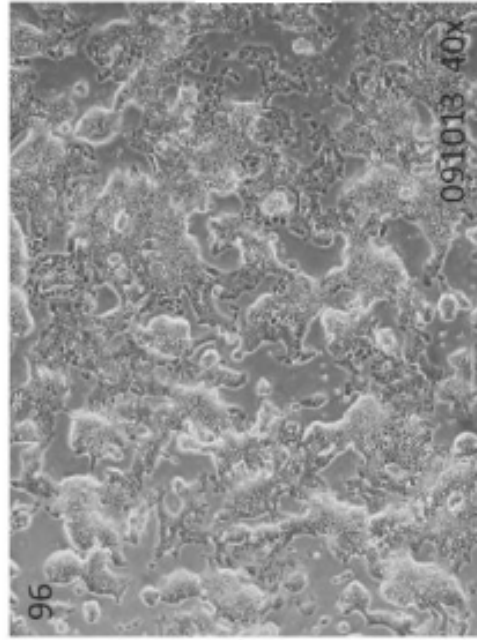


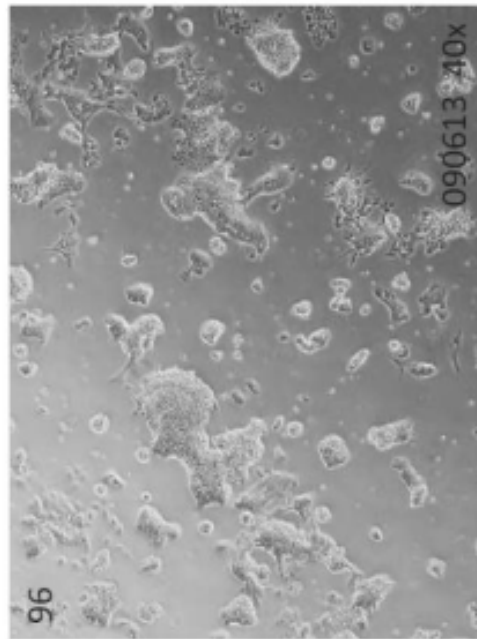
Figura 51

Figura 52B



12 días en 2i

Figura 52A



8 días en 2i

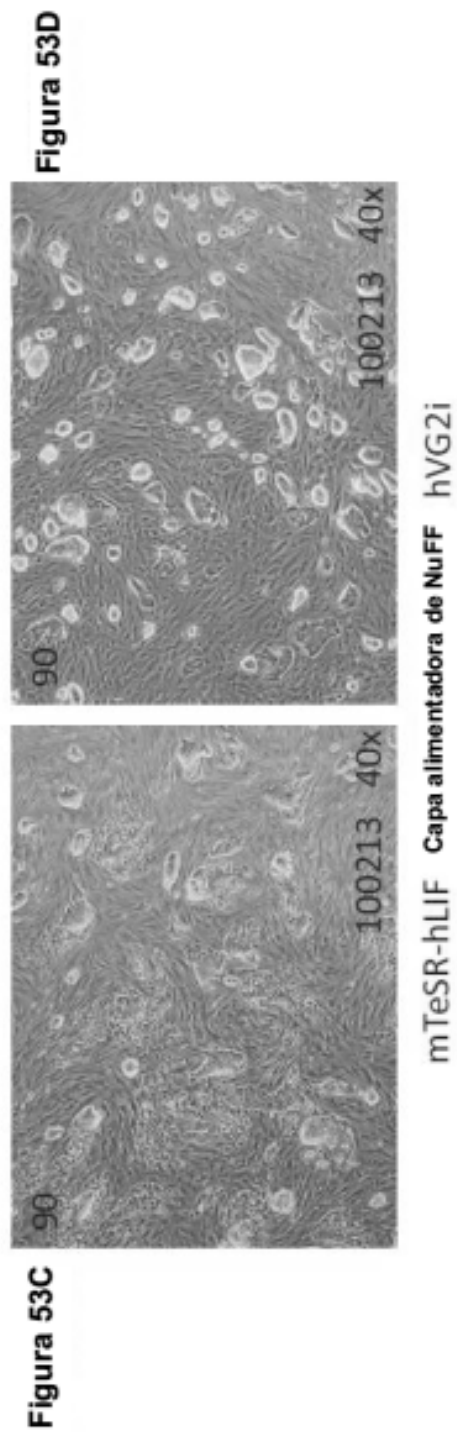
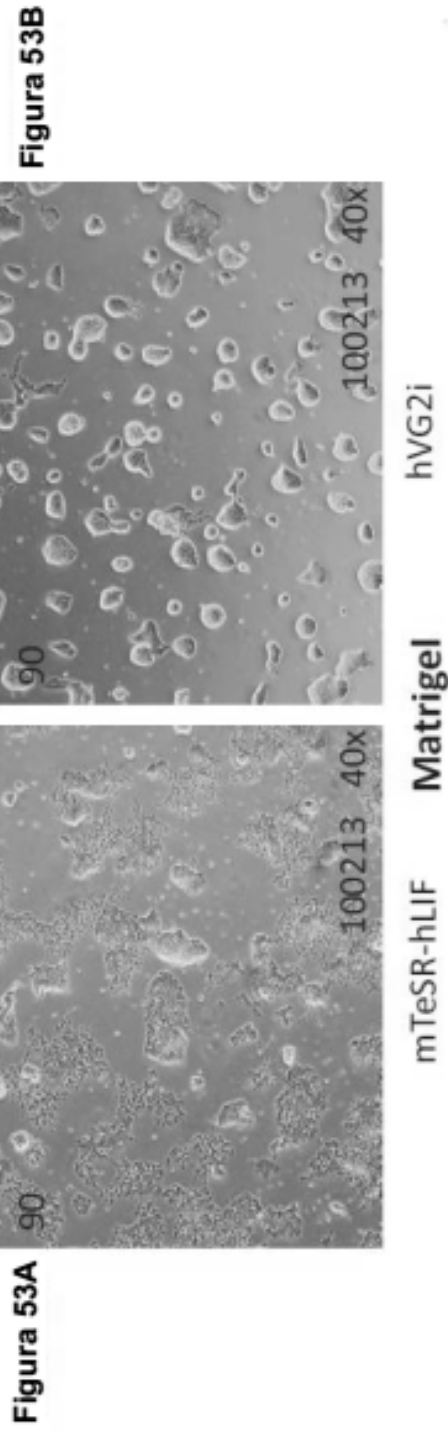


Figura 54C

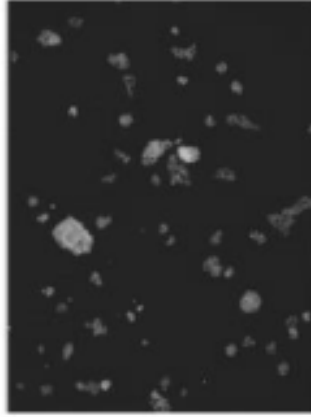


Figura 54B

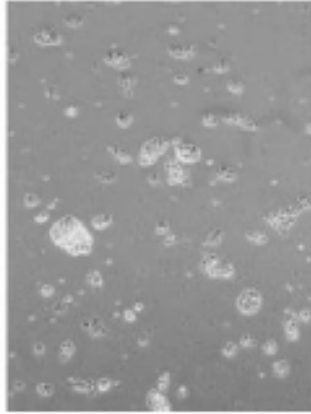


Figura 54A

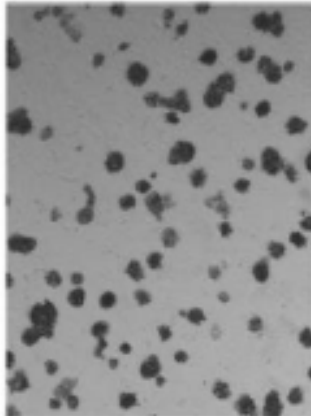
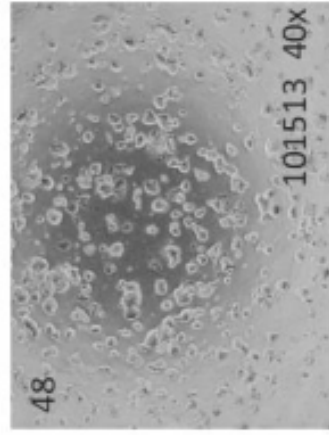
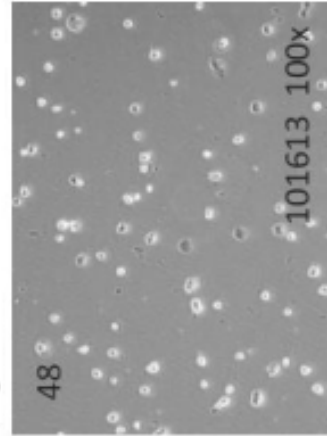


Figura 55A



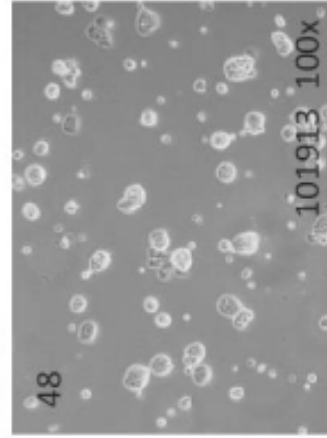
↙
pase celular único
con tripsina sin
inhibidor de ROCK
→

Figura 55B



1 días después del pase

Figura 55C



4 días después del pase