

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 733 248**

51 Int. Cl.:

C12P 21/00 (2006.01)

C12N 15/87 (2006.01)

C12N 9/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.01.2013 PCT/US2013/020947**

87 Fecha y número de publicación internacional: **18.07.2013 WO13106515**

96 Fecha de presentación y número de la solicitud europea: **10.01.2013 E 13735783 (6)**

97 Fecha y número de publicación de la concesión europea: **03.04.2019 EP 2802654**

54 Título: **Producción de proteínas recombinantes con glicoformas simples**

30 Prioridad:

11.01.2012 US 201261585508 P
22.04.2012 US 201261636680 P
30.10.2012 US 201261720268 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.11.2019

73 Titular/es:

SIGMA-ALDRICH CO. LLC (100.0%)
3050 Spruce Street
St. Louis, MO 63103, US

72 Inventor/es:

LIN, NAN;
SEALOVER, NATALIE;
GEORGE, HENRY y
KAYSER, KEVIN

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 733 248 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

Producción de proteínas recombinantes con glicoformas simples

Campo de la invención

La presente invención se refiere a proteínas recombinantes con glicoformas simples. En particular, la presente invención se refiere a composiciones y procedimientos para producir glicoproteínas recombinantes que tiene uno o más residuos de manosa terminal.

Antecedentes de la invención

La glicosilación N-ligada es una modificación proteica común que se encuentra en muchas glicoproteínas secretadas y unidas a la membrana. La síntesis de glicoformas complejas requiere una serie de reacciones enzimáticas llevadas a cabo de manera gradual en el retículo endoplásmico y el aparato de Golgi de células eucarióticas. Si bien la mayoría de las proteínas terapéuticas recombinantes requieren la presencia de glicanos complejos para evitar la inmunogenicidad y aumentar la vida media de las proteínas, también existen aplicaciones para las glicoproteínas con estructuras simples de N-glicano. Por ejemplo, las proteínas recombinantes con glicoformas simples son ventajosas para los estudios de cristalografía de rayos X. Adicionalmente, las glicoformas simples con solo residuos de manosa terminales pueden llevar a una mayor eficacia de algunas terapias al facilitar la captación mediada por el receptor de manosa para estas proteínas.

Sin embargo, la mayoría de las líneas celulares utilizadas para la producción de glicoproteínas terapéuticas producen proteínas con glicoformas unidas a N heterogéneas o complejas. Por lo tanto, hay una necesidad de líneas celulares que estén diseñadas para producir glicoproteínas recombinantes con glicoformas simples. Mientras que Zhong et al. (Biotechnol. Bioeng., 2012, 109 (7): 1723-1734) describió tal línea celular, la tasa de crecimiento no era adecuada para uso comercial. Es deseable que estas líneas celulares manipuladas genéticamente produzcan proteínas con residuos de manosa terminales, mientras que aún conservan una alta productividad de proteínas y un crecimiento robusto en medios de cultivo definidos químicamente.

Sumario

Entre los diversos aspectos de la presente descripción es la provisión de una línea celular de mamífero que es deficiente en manosa(alfa-1,3-)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa I (Mgat1) que comprende una supresión de 2 a 300bp en la secuencia cromosómica que codifica Mgat1. En una realización, la línea celular deficiente en Mgat1 no produce proteína de Mgat1. En otra realización, la línea celular deficiente en Mgat1 produce una cantidad reducida de proteína de Mgat1 con respecto a la línea celular progenitora que no es deficiente en Mgat1. En una realización, la línea celular deficiente en Mgat1 comprende una secuencia cromosómica inactivada que codifica Mgat1. La secuencia cromosómica que codifica Mgat1 se puede inactivar con una endonucleasa de direccionamiento. En ciertas realizaciones, la endonucleasa de direccionamiento se selecciona de una nucleasa de dedos de cinc, una nucleasa efectora tipo activador de transcripción (TALEN), una meganucleasa, y una endonucleasa específica de sitio. En una realización, la endonucleasa de direccionamiento es una nucleasa de dedos de cinc. En ciertas realizaciones, la secuencia cromosómica que codifica Mgat1 está inactivada debido a una supresión 2-55 pares de bases. En una realización la línea celular deficiente en Mgat1 también es deficiente en glutamina sintetasa (GS), dihidrofolato reductasa (DHFR), hipoxantina-guanina fosforribosiltransferasa (HPRT), o una de sus combinaciones. En otra realización, la línea celular deficiente en Mgat1 es una línea celular CHO. En una realización, la línea celular deficiente en Mgat1 es una línea celular CHO en que la copia única de la secuencia cromosómica que codifica Mgat1 está inactivada por una supresión y la línea celular no produce o Mgat1. En otra realización, la línea celular deficiente en Mgat1 es una línea celular CHO GS -/- en que copia única de la secuencia cromosómica que codifica Mgat1 está inactivada por una supresión y la línea celular no produce Mgat1. En una realización, la línea celular deficiente en Mgat1 expresa al menos una glicoproteína que tiene una glicoforma simple, en la que la glicoforma simple comprende uno o más residuos de manosa terminal. En una realización, la línea celular deficiente en Mgat1 tiene una tasa de crecimiento comparable con una línea celular que no tiene deficiencia en Mgat1. En una realización, la línea celular deficiente en Mgat1 produce un nivel de proteína comparable con una línea celular que no tiene deficiencia en Mgat1.

Otro aspecto de la descripción abarca un procedimiento para producir una célula deficiente en Mgat1, como se describió anteriormente. El procedimiento comprende transfectar una célula huésped con una endonucleasa de direccionamiento que se dirige a una secuencia específica en una secuencia cromosómica que codifica Mgat1. La endonucleasa de direccionamiento puede ser una nucleasa de dedos de cinc, una nucleasa efectora tipo activador de transcripción (TALEN), una meganucleasa, o una endonucleasa específica de sitio. En una realización la endonucleasa de direccionamiento es una nucleasa de dedos de cinc. En una realización, la célula huésped es una célula mamífera. En otra realización la célula huésped también es deficiente en glutamina sintetasa (GS), dihidrofolato reductasa (DHFR), hipoxanthineguanine fosforribosiltransferasa (HPRT), o una de sus combinaciones. En una realización, la célula huésped es una célula CHO. En otra realización, la célula huésped es una célula CHO GS -/-. En una realización, la célula huésped transfectada se incuba con aglutinina-I de *Ricinus communis*.

Aún otro aspecto de la presente descripción es un procedimiento para producir una proteína recombinante que tiene

uno o más residuos de manosa terminal. El procedimiento comprende introducir un ácido nucleico que codifica la proteína recombinante en una línea celular de la invención, y expresar la proteína recombinante.

Otros aspectos e iteraciones de la descripción se presentan con más detalle a continuación.

Breve descripción de los dibujos

- 5 La FIG. 1 es un diagrama esquemático que muestra el papel de Mgat1 en la síntesis de N-glicanos. Mgat1 cataliza la transferencia de un residuo de GlcNAc a una glicofoma Man5GlcNAc2 (Man5).
- La FIG. 2 ilustra el diseño y la validación de nucleasas de dedos de zinc (ZFN) dirigidas a Mgat1. (A) presenta una ilustración esquemática de la posición en la que se diseña las ZFN para unirse y cortar en la región codificadora de Mgat1. (B) muestra la validación de la actividad de ZFN mediante un ensayo de nucleasa Cel1.
- 10 Se muestran los digestos de ADN genómico en el día 3 y el día 10 después de la transfección de células transfectadas de forma simulada, células transfectadas con ADN y células transfectadas con ARN. Dos fragmentos de ADN, de 220 y 197 pb, están presentes tanto en los digestos del día 3 como en el día 10 post-transfección.
- La FIG. 3 presenta el análisis Cel1 antes y después del enriquecimiento con RCA-I. Se muestran los digestos de ADN genómico de células transfectadas con ADN o ARN antes y después del enriquecimiento con RCA-I, así como células transfectadas de manera simulada. El porcentaje modificado, determinado por densitometría, se indica debajo de cada calle.
- 15 La FIG. 4 muestra los glicoperfiles de IgG de células transfectadas con ZFN Mgat1 con o sin enriquecimiento con RCA-I. (A) presenta el porcentaje de cada glicofoma (Man5, G0, G0F, G1F y G2F) en las IgG producidas por células transfectadas de forma simulada, células transfectadas con ADN o ARN de ZFN sin selección de RCA-I y células transfectadas con ADN o ARN de ZFN con selección de RCA-I. (B) diagrama de la estructura de cada glicofoma analizada.
- 20 La FIG. 5 presenta los genotipos de los clones productores de IgG alterados por Mgat1. Se muestra la secuencia de alineamiento de las siguientes líneas celulares: tipo salvaje no transfectado (wt); dos clones de tipo salvaje (#15, #46) que se sometieron a la transfección ZFN y al aislamiento del clon; dos líneas celulares alteradas por Mgat1 seleccionadas sin RCA-I (#73, #92) y cuatro líneas celulares alteradas por Mgat1 seleccionadas en presencia de RCA-I (#21, #25, #27, #37). El área sombreada marca el sitio de corte de la ZFN.
- 25 La FIG. 6 ilustra la insensibilidad de los clones knockout de Mgat1 a RCA-I. Se grafica la densidad celular viable máxima en cultivos por lotes de clones de tipo silvestre (wt, #15, #46) y knockout Mgat1 (#73, #92, #21, #25, #27, #37) cuando se cultivan en ausencia o presencia de RCA-I.
- 30 La FIG. 7 muestra los glicoperfiles de IgG de tipo salvaje y los clones productores de IgG knockout Mgat1. (A) muestra los porcentajes de las diferentes glicofomas en muestras de IgG recolectadas el día 5 de los clones no transfectados (wt) y dos de tipo salvaje (#15, #46). (B) presenta los porcentajes de las diferentes glicofomas en muestras de IgG recolectadas en el día 5 de los seis clones knockout Mgat1 (#73, #92, #21, #25, #27, #37). (C) muestra los porcentajes de las diferentes glicofomas en muestras de IgG recolectadas el día 8 de los clones de tipo salvaje. (D) presenta los porcentajes de las diferentes glicofomas en muestras de IgG recolectadas el día 8 de los seis clones knockout Mgat1. ND = no detectado.
- 35 La FIG. 8 ilustra el crecimiento de clones knockout de tipo salvaje y Mgat1 a lo largo del tiempo. (A) presenta el porcentaje de células viables en los clones de tipo salvaje (wt, #15, #46) y Mgat1 (#73, #92, #21, #25, #27, #37) durante 13 días en cultivo por lotes. (B) muestra la densidad de células viables en cultivos por lotes de tipo salvaje y seis clones knockout Mgat1 durante 13 días.
- 40 La FIG. 9 muestra la productividad IgG de los clones knockout de tipo salvaje y Mgat1. Se grafica la productividad volumétrica máxima en los clones de tipo salvaje (wt, #15, #46) y clones knockout Mgat1 (#73, #92, #21, #25, #27, #37).
- 45

Descripción detallada de la invención

- La presente divulgación proporciona composiciones y procedimientos para la producción de proteínas recombinantes que tienen glicofomas simples, en la que una glicofoma simple comprende uno o más residuo de manosa terminal.
- 50 En particular, la presente descripción proporciona una línea celular que es deficiente en manosil(alfa-1,3-)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa I (Mgat1) que comprende una supresión de 2 a 300bp en la secuencia cromosómica que codifica Mgat1. En algunas realizaciones, la línea celular deficiente en Mgat1 comprende una secuencia cromosómica inactivada o knocked-out que codifica Mgat1 de modo que la línea celular no produce Mgat1. Como consecuencia, la línea celular produce glicoproteínas que tienen uno o más residuo de manosa terminal.

Las líneas celulares deficientes en Mgat1 descritas en la presente son útiles para la producción de vacunas debido a que las glicoproteínas con residuos de manosa terminales facilitan la captación mediada por el receptor de manosa. Además, las glicoproteínas con glicoformas simples son útiles para los estudios de cristalografía de rayos X. Además, las células deficientes en Mgat1 descritas en la presente evitan posibles preocupaciones regulatorias asociadas con la mutagénesis química aleatoria y mantienen características de crecimiento y productividad que son comparables a las de las líneas celulares progenitoras.

La presente divulgación también proporciona procedimientos para preparar las células deficientes en Mgat1. Por ejemplo, las células deficientes en Mgat1 se pueden preparar utilizando endonucleasas de direccionamiento para inactivar la secuencia cromosómica que codifica Mgat1. También se proporcionan procedimientos para producir glicoproteínas recombinantes con glicoformas simples usando las células deficientes en Mgat1 descritas en la presente memoria.

(I) Líneas celulares deficientes en Mgat1

(a) Propiedades de la línea celular

La presente descripción abarca una línea celular que comprenden una supresión de 2 a 300 bp en la secuencia cromosómica que codifica Mgat1. Mgat1 también se conoce como N-acetilglucosaminiltransferasa I (GlcAc-TI o GnTI) o N-glicosil-oligosacárido-glicoproteína N-acetilglucosaminiltransferasa I. Mgat1 cataliza la transferencia de N-acetilglucosamina (GlcNAc) en el núcleo de oligomanosa (es decir, resto Man5GlcNAc2 (Man5) de un N-glicano en crecimiento (ver FIG. 1).

En algunas realizaciones, la línea celular deficiente en Mgat1 tiene niveles reducidos de Mgat1 con respecto a la línea celular progenitora. Por ejemplo, los niveles de Mgat1 en la línea celular deficiente en Mgat1 se pueden reducir de aproximadamente 5% a aproximadamente 10%, de aproximadamente 10% a aproximadamente 20%, de aproximadamente 20% a aproximadamente 30%, de aproximadamente 30% a aproximadamente 40%, de aproximadamente 40% a aproximadamente 50%, de aproximadamente 50% a aproximadamente 60%, de aproximadamente 60% a aproximadamente 70%, de aproximadamente 70% a aproximadamente 80%, de aproximadamente 80% a aproximadamente 90%, o de aproximadamente 90% a aproximadamente 99.9% con respecto a la línea celular progenitora que no es deficiente en Mgat1. En otras realizaciones, la línea celular deficiente en Mgat1 no produce esencialmente Mgat1. Como se usa en la presente, el término "esencialmente no Mgat1" significa que no se puede detectar ARNm y/o proteína de Mgat1 en las líneas celulares deficientes de Mgat1 o lisados derivados de las mismas usando procedimientos bien conocidos en la técnica. Los ejemplos no limitantes de procedimientos adecuados para determinar el nivel de ARNm o proteína incluyen ensayos de PCR, qPCR, transferencia Western y ELISA.

Las líneas celulares deficientes en Mgat1 producen glicoproteínas con glicoformas simples. Una glicoforma simple se refiere a una glicoforma que comprende uno o más residuos de manosa terminal. En general, una glicoforma simple está desprovista de residuos de galactosa y/o ácido siálico. Un ejemplo de glicoforma simple es Man5GlcNAc2 o Man5. En algunas realizaciones, las líneas celulares deficientes en Mgat1 pueden producir glicoproteínas en las que los niveles de residuo de manosa terminales aumentan de aproximadamente 1,1 veces a aproximadamente 3 veces, de aproximadamente 3 veces a aproximadamente 10 veces, de aproximadamente 10 veces a aproximadamente 30 veces, de aproximadamente 30 veces a aproximadamente 100 veces, o más de aproximadamente 100 veces. En otras realizaciones, las líneas celulares deficientes en Mgat1 pueden producir glicoproteínas en las que los niveles de residuos de galactosa y/o ácido siálico terminales disminuyen de aproximadamente 1,1 veces a aproximadamente 3 veces, de aproximadamente 3 veces a aproximadamente 10 veces, de aproximadamente 10 veces a aproximadamente 30 veces, de aproximadamente 30 veces a aproximadamente 100 veces, o más de aproximadamente 100 veces.

Las líneas celulares deficientes en Mgat1 tienen niveles de producción o productividad de proteína que son comparables con los de las líneas celulares que no son deficientes en Mgat1. Como se usa en la presente, "producción de proteína comparable" significa que las líneas celulares deficientes en Mgat1 producen aproximadamente el mismo nivel o más proteína que las líneas celulares que no están deficientes en Mgat1. La productividad de la proteína se puede cuantificar mediante la determinación de la cantidad de proteína producida por volumen unitario (es decir, productividad volumétrica pico). En algunas realizaciones, la cantidad de la proteína producida por las líneas celulares deficientes en Mgat1 puede variar aproximadamente $\pm 10\%$ con respecto a las producidas por las líneas celulares que no son deficientes en Mgat1. En otras realizaciones, la cantidad de la proteína producida por las líneas celulares deficientes en Mgat1 se puede aumentar de aproximadamente 10% a aproximadamente 20%, de aproximadamente 20% a aproximadamente 30%, de aproximadamente 30% a aproximadamente 40%, o de aproximadamente 40% a aproximadamente 50% con respecto a las producidas por las líneas celulares que no son deficientes en Mgat1.

Las líneas celulares deficientes en Mgat1 tienen tasas de crecimiento que son comparables a las de las líneas celulares que no son deficientes en Mgat1. Como se usa en la presente, "tasas de crecimiento comparables" se refieren a tasas de crecimiento que aumentan con respecto a las líneas celulares que no son deficientes en Mgat1 o que disminuyen en menos de aproximadamente el 50% con respecto a las líneas celulares que no son deficientes en Mgat1. Las tasas de crecimiento se pueden cuantificar mediante la determinación del porcentaje de células viables o el número de

células viables por volumen unitario (es decir, la densidad de células viables) después de un cierto período de tiempo en el cultivo. En algunas realizaciones, las líneas celulares deficientes en Mgat1 tienen tasas de crecimiento aumentadas con respecto a las líneas celulares que no son deficientes en Mgat1. Por ejemplo, el porcentaje de células viables o densidad celular viable en las líneas celulares deficientes en Mgat1 se puede aumentar de aproximadamente 1% a aproximadamente 10%, de aproximadamente 10% a aproximadamente 30%, de aproximadamente 30% a aproximadamente 50%, de aproximadamente 50% a aproximadamente 100%, o más de 100% con respecto a las líneas celulares que no son deficientes en Mgat1. En otras realizaciones, el porcentaje de células viables o densidad celular viable en las líneas celulares deficientes en Mgat1 se puede disminuir de aproximadamente 1% a aproximadamente 10%, de aproximadamente 10% a aproximadamente 30%, de aproximadamente 30% a aproximadamente 50% con respecto a líneas celulares que no son deficientes en Mgat1.

Las líneas celulares deficientes en Mgat1 descritas en la presente tienen fenotipos estables. Por ejemplo, las líneas celulares deficientes en Mgat1 producen glicoproteínas con glicofomas simples durante períodos prolongados de tiempo y/o por innumerables generaciones de células. Dicho de otra manera, las líneas celulares descritas en la presente no vuelven a producir glicoproteínas con glicofomas complejas. Además, las líneas celulares deficientes en Mgat1 mantienen altos niveles de productividad de la proteína y excelentes tasas de crecimiento durante períodos prolongados de tiempo y/o por innumerables generaciones de células.

(b) Secuencia cromosómica inactivada que codifica Mgat1

En algunas realizaciones, la línea celular deficiente en Mgat1 comprende una secuencia cromosómica inactivada que codifica Mgat1. Por ejemplo, a edición del genoma de la línea celular inactiva la secuencia cromosómica que codifica Mgat1. Como se usa en la presente, el término "secuencia cromosómica inactivada" se refiere a una secuencia cromosómica que es incapaz de generar un producto de proteína. En las realizaciones en las que la línea celular comprende células euploides, la secuencia cromosómica inactivada que codifica Mgat1 puede ser monoalélica de modo que la línea celular produce niveles reducidos de Mgat1. En otras realizaciones en las que la línea celular comprende células euploides, la secuencia cromosómica inactivada que codifica Mgat1 puede ser bialélica de modo que la línea celular no produce Mgat1. En otras realizaciones en que la línea celular es aneuploide, una o más copias de la secuencia cromosómica que codifica Mgat1 se puede inactivar de modo que la célula produce una cantidad reducida de Mgat1. En aún otras realizaciones en que la línea celular es aneuploide, todas las copias de la secuencia cromosómica que codifica Mgat1 se pueden inactivar de modo que la línea celular no produce Mgat1. En un ejemplo de realización, la línea celular es haploide para la secuencia cromosómica que codifica Mgat1, y la inactivación de la secuencia cromosómica que codifica Mgat1 produce una pérdida completa de la expresión de Mgat1.

Como consecuencia de la supresión, la secuencia codificadora de Mgat1 experimenta un cambio en el marco de lectura, de este modo impide la producción de un producto de proteína. La secuencia cromosómica que codifica Mgat1 se puede inactivar utilizando la tecnología de edición genómica mediada por endonucleasas de direccionamiento como se detalla a continuación en la sección (II). En diversas realizaciones, la secuencia cromosómica que codifica Mgat1 se puede inactivar mediante la supresión de la totalidad o parte de la región codificadora exónica, supresión de la totalidad o parte de una región de control y/o supresión de un sitio de empalme de manera de anular la expresión de Mgat1. En otras realizaciones, la secuencia cromosómica que codifica Mgat1 se puede inactivar mediante supresiones, inserciones y/o sustituciones de nucleótidos para introducir un codón de detención prematuro, nuevos sitios de empalme y/o SNP en la secuencia cromosómica de tal manera que la línea celular no pueda producir Mgat1.

En una realización, la línea celular deficiente en Mgat1 comprende una supresión que varía de aproximadamente 2 bp a aproximadamente 55 bp, de aproximadamente 55 bp a aproximadamente 100 bp, o de aproximadamente 100 bp a aproximadamente 300 bp de la región codificadora de Mgat1. En ejemplos de realización, la línea celular deficiente en Mgat1 comprende una supresión de 2 bp, 4 bp, 7 bp, 10, bp, 18 bp, 24 bp, 28 bp, o 55 bp de la secuencia cromosómica que codifica Mgat1. En otro ejemplo de realización, la línea celular deficiente en Mgat1 comprende una supresión de 157 bp que se reemplaza mediante una inserción de 104 bp de la secuencia cromosómica que codifica Mgat1.

(c) Deficiencias adicionales opcionales

En algunas realizaciones, la línea celular deficiente en Mgat1 también comprende una deficiencia en glutamina sintetasa (GS), dihidrofolato reductasa (DHFR), hipoxantina-guanina fosforibosiltransferasa (HPRT), o sus combinaciones. La deficiencia en GS, DHFR, y/o HPRT puede ser natural. Alternativamente, la deficiencia en GS, DHFR, y/o HPRT se puede manipular genéticamente. Por ejemplo, la célula puede ser deficiente en GS, DHFR, o HPRT porque la célula comprende una secuencia cromosómica inactivada que codifica GS, DHFR, o HPRT, respectivamente. En algunas realizaciones, la secuencia cromosómica que codifica GS, DHFR, o HPRT se puede inactivar mediante edición genómica mediada por endonucleasas de direccionamiento. En ejemplos de realización, todas las copias de la secuencia cromosómica que codifica GS, DHFR, o HPRT se inactivan de modo que las líneas celulares no producen GS, DHFR, o HPRT, respectivamente. En un ejemplo de realización, la línea celular deficiente en Mgat1 es GS ^{-/-}.

(d) Tipos de línea celular

El tipo de línea celular que es deficiente en Mgat1 puede variar. La línea celular puede ser una línea celular humana,

una línea celular de mamífero no humana, una línea celular de vertebrado no mamífero, una línea celular de invertebrado, o una línea celular de levadura.

Los ejemplos no limitantes de líneas celulares de mamífero adecuadas incluyen células de ovario de hámster chino (CHO), células de riñón de hámster bebé (BHK); células NS0 de mieloma de ratón, células 3T3 de fibroblasto embrionario de ratón (NIH3T3), células A20 de linfoma B de ratón; células B16 de melanoma de ratón; células C2C12 de mioblasto de ratón; células SP2 de mieloma de ratón/0; células mesenquimales embrionarias de ratón C3H-1 0T1/2; células CT26 de carcinoma de ratón, células DuCuP de próstata de ratón; células EMT6 de mama de ratón; células Hepa1c1c7 de hepatoma de ratón; células J5582 de mieloma de ratón; células MTD-1A epiteliales de ratón; células de MyEnd de miocardio de ratón; células RenCa renales de ratón; células RIN-5F pancreáticas de ratón; células X64 de melanoma de ratón; células YAC-1 de linfoma de ratón; células 9L de glioblastoma de rata; células RBL de linfoma B de rata; células B35 de neuroblastoma de rata; células de hepatoma de rata (HTC); células de BRL 3A de hígado de rata búfalo; células renales caninas (MDCK); células mamarias caninas (CMT); células D17 de osteosarcoma de rata; células DH82 de monocitos/macrófagos de rata; células de fibroblastos (COS7) transformados con SV-40 de riñón de mono; células CVI-76 de riñón de mono; células de riñón de mono verde africano (VERO-76); células renales embrionarias humanas (HEK293, HEK293T); células de carcinoma cervical humano (HELA); células pulmonares humanas (W138); células hepáticas humanas (Hep G2); células de osteosarcoma humano U2-OS, células A549 humanas, células A-431 humanas y células K562 humanas. Se puede encontrar una lista extensa de líneas celulares de mamíferos en el catálogo del American Type Culture Collection catalog (ATCC, Mamassas, VA)

Los ejemplos de líneas celulares no mamíferas adecuadas incluyen, pero sin limitación, líneas celulares de *Xenopus* (tales como S3, XTC, BB7, ff-2, 15/0 y 15/40); líneas celulares de pez cebra (tales como, ZF4, PAC2, etc.); líneas celulares de insecto (tales como SF9, S2 y similares); y líneas celulares de levadura tales como líneas celulares de *Pichia* y líneas celulares de *Saccharomyces*.

En ejemplos de realización, la línea celular es un tipo que se usa ampliamente para la producción de proteínas recombinantes, tales como anticuerpos, glicoproteínas y similares. En ejemplos de realización, la línea celular es una línea celular CHO. Numerosas líneas celulares de CHO están disponibles en ATCC. Las líneas celulares CHO adecuadas incluyen, pero no se limitan a, células CHO-K1 y derivados de las mismas.

(e) Ácido nucleico opcional

En algunas realizaciones, la línea celular deficiente en Mgat1 puede comprender además al menos una secuencia de ácido nucleico que codifica una proteína recombinante. En general, la proteína recombinante es heteróloga, lo que significa que la proteína no es nativa de la célula. La proteína recombinante puede ser, sin límite, un anticuerpo, un fragmento de un anticuerpo, un anticuerpo monoclonal, un anticuerpo humanizado, un anticuerpo monoclonal humanizado, un anticuerpo quimérico, una molécula de IgG, una cadena pesada de IgG, una cadena liviana de IgG, una molécula de IgA, una molécula de IgD, una molécula de IgE, una molécula de IgM, una vacuna, un factor de crecimiento, una citoquina, un interferón, una interleuquina, una hormona, un factor coagulante (o coagulación), un componente sanguíneo, una enzima, una proteína terapéutica, una proteína nutracéutica, un fragmento funcional o variante funcional de cualquiera de los anteriores, o una proteína de fusión que comprende cualquiera de las proteínas anteriores y/o fragmentos funcionales o variantes de los mismos.

En algunas realizaciones, la secuencia de ácido nucleico que codifica la proteína recombinante se puede unir a una secuencia de ácido nucleico que codifica hipoxantina-guanina fosforribosiltransferasa (HPRT), dihidrofolato reductasa (DHFR) y/o glutamina sintetasa (GS), tal como HPRT, DHFR y/o GS se puede utilizar como un marcador seleccionable amplificable.

En algunas realizaciones, la secuencia de ácido nucleico que codifica la proteína recombinante puede ser extracromosómica. Es decir, el ácido nucleico que codifica la proteína recombinante se puede expresar de forma transitoria a partir de un plásmido, un cósmido, un cromosoma artificial, un minicromosoma u otro constructo extracromosómico. El constructo de expresión puede comprender secuencias de control de expresión adicionales (por ejemplo, secuencias potenciadoras, secuencias de Kozak, secuencias de poliadenilación, secuencias de terminación de la transcripción, etc.), secuencias de marcadores seleccionables (por ejemplo, genes de resistencia a antibióticos), orígenes de replicación y similares. Información adicional se puede hallar en "Current Protocols in Molecular Biology" Ausubel et al., John Wiley & Sons, New York, 2003 o "Molecular Cloning: A Laboratory Manual" Sambrook & Russell, Cold Spring Harbor Press, Cold Spring Harbor, NY, 3rd edition, 2001.

En otras realizaciones, la secuencia de ácido nucleico que codifica la proteína recombinante se puede integrar cromosómicamente en el genoma de la célula. Por consiguiente, la proteína recombinante se puede expresar de forma estable. En algunas iteraciones de esta realización, la secuencia de ácido nucleico que codifica la proteína recombinante se puede unir operativamente a una secuencia de control de expresión heteróloga apropiada (es decir, promotor). En otras iteraciones, la secuencia de ácido nucleico que codifica la proteína recombinante se puede colocar bajo el control de una secuencia de control de expresión endógena. La secuencia de ácido nucleico que codifica la proteína recombinante se puede integrar en el genoma de la línea celular mediante la recombinación homóloga, edición genómica mediada por endonucleasas de direccionamiento, vectores virales, transposones, plásmidos y otros medios bien conocidos. Se puede encontrar orientación adicional en Ausubel et al. 2003, supra y Sambrook & Russell,

2001, supra.

(f) Ejemplos de realizaciones

En un ejemplo de realización, la línea celular es una línea celular CHO en la que la copia única de la secuencia cromosómica que codifica Mgat1 está inactivada (es decir, el genotipo de la célula es Mgat1 ⁻/0). En otro ejemplo de realización, la línea celular es una línea celular CHO GS ⁻/⁻ en la que la copia única de la secuencia cromosómica que codifica Mgat1 está inactivada (es decir, el genotipo de la célula es GS ⁻/⁻, Mgat1 ⁻/0). En general, la secuencia cromosómica que codifica Mgat1 está inactivada por una supresión de parte de la secuencia codificadora.

(II) Procedimientos para preparar líneas celulares deficientes en Mgat1

Las líneas celulares deficientes en Mgat1 se puede preparar mediante una variedad de procedimientos. En ciertas realizaciones, la línea celular deficiente en Mgat1 se puede preparar mediante un procedimiento de edición genómica mediada por endonucleasas de direccionamiento. En otras realizaciones, la línea celular deficiente en Mgat1 se puede preparar mediante procedimientos de ARNi, mutagénesis aleatoria, sistemas de recombinación específicos de sitio, u otros procedimientos conocidos en la técnica.

Las líneas celulares deficientes en Mgat1 producen glicoproteínas que tienen uno o más residuos de manosa terminales. En realizaciones adicionales, las líneas celulares deficientes de Mgat1 se pueden enriquecer adicionalmente mediante incubación con aglutinina-I de *Ricinus communis* (RCA-I). RCA-I es una lectina citotóxica que no se une a los residuos de manosa terminales y, por lo tanto, permite la selección de células desprovistas de actividad de Mgat1 porque tales células producen glicoproteínas con residuos de manosa terminales.

(a) Edición genómica mediada por endonucleasas de direccionamiento

La endonucleasa de direccionamientos se puede usar para editar (es decir, inactivar o modificar) una secuencia cromosómica específica. Una secuencia cromosómica específica se puede inactivar mediante la introducción en célula de una endonucleasa de direccionamiento o un ácido nucleico que codifica la endonucleasa de direccionamiento, que se manipula genéticamente para dirigirse a una secuencia cromosómica específica. En una realización, la endonucleasa de direccionamiento reconoce y se une a la secuencia cromosómica específica e introduce una ruptura de doble cadena que se repara mediante un procedimiento de reparación de unión de extremos no homólogos (NHEJ). Debido a que la NHEJ es propensa a errores, se puede producir una supresión, inserción y/o sustitución de al menos un nucleótido, lo que altera el marco de lectura de la secuencia cromosómica de modo que no se produce ningún producto de proteína. En otra realización, las endonucleasas de direccionamiento también se pueden usar para editar una secuencia cromosómica a través de una reacción de recombinación homóloga mediante la cointroducción de un polinucleótido que tiene una identidad de secuencia sustancial con una parte de la secuencia cromosómica dirigida. La ruptura de la cadena doble introducida por la endonucleasa de direccionamiento se repara mediante un procedimiento de reparación dirigido por homología, de manera que la secuencia cromosómica se intercambia con el polinucleótido de una manera que produce la edición de la secuencia cromosómica.

(i) endonucleasas de direccionamiento

Una variedad de endonucleasas de direccionamiento se puede usar para editar la secuencia cromosómica. La endonucleasa de direccionamiento puede ser una proteína natural o una proteína manipulada genéticamente. En una realización, la endonucleasa de direccionamiento puede ser una meganucleasa. Las meganucleasas son endonucleasas caracterizadas por secuencias de reconocimiento largas, es decir, la secuencia de reconocimiento generalmente varía de aproximadamente 12 pares de bases a aproximadamente 40 pares de bases. Como consecuencia de este requerimiento, la secuencia de reconocimiento generalmente ocurre solo una vez en cualquier genoma dado. Entre las meganucleasas, la familia de endonucleasas de asentamiento llamada LAGLIDADG se ha convertido en una herramienta valiosa para el estudio de los genomas y la ingeniería genómica. Una meganucleasa se puede dirigir a una secuencia cromosómica específica mediante la modificación de su secuencia de reconocimiento utilizando técnicas bien conocidas por los expertos en la técnica.

En otra realización, la endonucleasa de direccionamiento puede ser una nucleasa efectora tipo activador de la transcripción (TALE). Los TALE son factores de transcripción del patógeno de plantas *Xanthomonas* que se pueden manipular genéticamente en forma sencilla para unirse a nuevas dianas de ADN. Las TALEs o versiones truncadas de las mismas se pueden unir al dominio catalítico de las endonucleasas tales como FokI para crear endonucleasas de direccionamiento llamadas nucleasas TALE o TALENs (Sanjana et al., 2012, Nat Protoc, 7 (1): 171-192)).

En otra realización más, la endonucleasa de direccionamiento puede ser una endonucleasa específica del sitio. En particular, la endonucleasa específica del sitio puede ser una endonucleasa de "corte raro" cuya secuencia de reconocimiento ocurre rara vez en un genoma. En general, la secuencia de reconocimiento de la endonucleasa específica del sitio ocurre solo una vez en un genoma. En una realización adicional alternativa, la endonucleasa de direccionamiento puede ser un agente inductor de ruptura de cadena doble de ADN dirigido artificial.

En ejemplos de realización, la endonucleasa de direccionamiento puede ser una nucleasa de dedos de cinc (ZFN). Típicamente, una ZFN comprende un dominio de unión a ADN (es decir, dedos de zinc) y un dominio de escisión (es

decir, nucleasa), los cuales se describen a continuación.

Dominio de unión de dedo de zinc. Los dominios de unión a los dedos de zinc se pueden manipular genéticamente para reconocer y unirse a cualquier secuencia de ácido nucleico de elección. Ver, por ejemplo, Beerli et al. (2002) Nat. Biotechnol. 20:135-141; Pabo et al. (2001) Ann. Rev. Biochem. 70:313-340; Isalan et al. (2001) Nat. Biotechnol. 19:656-660; Segal et al. (2001) Curr. Opin. Biotechnol. 12:632-637; Choo et al. (2000) Curr. Opin. Struct. Biol. 10:411-416; Zhang et al. (2000) J. Biol. Chem. 275(43):33850-33860; Doyon et al. (2008) Nat. Biotechnol. 26:702-708; and Santiago et al. (2008) Proc. Natl. Acad. Sci. USA 105:5809-5814. Un dominio de unión de dedos de zinc manipulado genéticamente puede tener una nueva especificidad de unión en comparación con una proteína de dedos de cinc natural. Los procedimientos de manipulación genética incluyen, pero sin limitación, diseño racional y varios tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos de doblete, triplete y/o cuadruplete y secuencias de aminoácidos de dedos de cinc individuales, en las que cada secuencia de nucleótidos de doblete, triplete o cuadruplete se asocia con una o más secuencias de aminoácidos de dedos de cinc que unen la secuencia particular de triplete o cuadruplete. Ver, por ejemplo, las patentes U.S.Nros. 6.453.242 y 6.534.261. Como ejemplo, el algoritmo descrito en la patente U.S.6.453.242 se puede usar para diseñar un dominio de unión de dedos de zinc para dirigirse a una secuencia preseleccionada. Procedimientos alternativos, tales como el diseño racional que utiliza una tabla de códigos de reconocimiento no degenerados, también se pueden usar para diseñar un dominio de unión de dedos de zinc para dirigirse a una secuencia específica (Sera et al. (2002) Biochemistry 41: 7074-7081). En la técnica se conocen herramientas basadas en la web disponibles públicamente para identificar sitios diana potenciales en secuencias de ADN, así como para diseñar dominios de unión de dedos de zinc. Por ejemplo, las herramientas para identificar potenciales dianas en secuencias de ADN se pueden encontrar en <http://www.zincfingertools.org>. Las herramientas para diseñar dominios de enlace de dedos de zinc se pueden encontrar en <http://zifit.partners.org/ZiFiT>. (Ver también, Mandell et al. (2006) Nuc. Acid Res. 34:W516-W523; Sander et al. (2007) Nuc. Acid Res. 35:W599-W605.)

Un dominio de unión de dedos de cinc se puede diseñar para reconocer y unir una secuencia de ADN que varía de aproximadamente 3 nucleótidos a aproximadamente 21 nucleótidos de longitud. En una realización, el dominio de unión de dedos de cinc se puede diseñar para reconocer y unir una secuencia de ADN que varía de aproximadamente 9 a aproximadamente 18 nucleótidos de longitud. En general, los dominios de unión de dedos de cinc de las nucleasas de dedos de cinc usadas en el presente comprenden al menos tres regiones de reconocimiento de dedos de zinc o dedos de zinc, en los que cada dedo de zinc une 3 nucleótidos. En una realización, el dominio de unión de dedos de cinc comprende cuatro regiones de reconocimiento de dedos de zinc. En otra realización, el dominio de unión de dedos de cinc comprende cinco regiones de reconocimiento de dedos de zinc. En aún otra realización, el dominio de unión de dedos de cinc comprende seis regiones de reconocimiento de dedos de zinc. A dominio de unión de dedos de cinc se puede diseñar para unir cualquier secuencia de ADN diana adecuada. Ver, por ejemplo, la patente U.S. Nº 6.607.882; 6.534.261 y 6.453.242.

Los ejemplos de procedimientos para seleccionar una región de reconocimiento de dedos de cinc incluyen el despliegue en fagos y los sistemas de dos híbridos, que se describen en las patentes U.S. Nros. 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; y 6.242.568; así como el documento WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237. Además, la mejora de la especificidad de unión para los dominios de unión de dedos de zinc se ha descrito, por ejemplo, en el documento WO 02/077227.

Los expertos en la técnica conocen los dominios de unión con el dedo de cinc y los procedimientos para el diseño y la construcción de proteínas de fusión (y polinucleótidos que codifican los mismos) y se describen en detalle, por ejemplo, en la patente U.S. Núm. 7.888.121. Las regiones de reconocimiento de los dedos de zinc y/o las proteínas de dedos de zinc de múltiples dedos se pueden unir entre sí utilizando secuencias ligadoras adecuadas, que incluyen, por ejemplo, ligadores de cinco o más aminoácidos de longitud. Ver, Patente U.S.Núm. 6.479.626; 6.903.185; y 7.153.949, para ejemplos no limitantes de secuencias ligadoras de seis o más aminoácidos de longitud. El dominio de unión de dedos de cinc descrito en la presente puede incluir una combinación de ligadores adecuados entre los dedos de zinc individuales de la proteína.

En algunas realizaciones, la nucleasa de dedos de cinc comprende además una señal o secuencia de localización nuclear (NLS). Una NLS es una secuencia de aminoácidos que facilita el direccionamiento de la proteína de la nucleasa de dedos de cinc en el núcleo para introducir una ruptura de doble cadena en la secuencia diana del cromosoma. Las señales de localización nuclear son conocidas en la técnica. Ver, por ejemplo, Makkerh et al. (1996) Current Biology 6:1025-1027.

Dominio de escisión. Una nucleasa de dedos de cinc también incluye un dominio de escisión. La porción del dominio de escisión de la nucleasa de dedos de cinc se puede obtener a partir de cualquier endonucleasa o exonucleasa. Los ejemplos no limitantes de endonucleasas a partir de los cuales se puede derivar un dominio de escisión incluyen, pero sin limitación, endonucleasas de restricción y endonucleasas de asentamiento. Ver, por ejemplo, New England Biolabs Catalog o Belfort et al. (1997) Nucleic Acids Res. 25: 3379-3388. Se conocen enzimas adicionales que escinden el ADN (por ejemplo, nucleasa S1, nucleasa de frijol mungo, ADNasa I pancreática, nucleasa microcócica, endonucleasa HO de levadura). Ver también Linn et al. (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993. Una o más de estas enzimas (o sus fragmentos funcionales) se pueden usar como fuente de dominios de escisión.

Un dominio de escisión también se puede derivar de una enzima o porción de la misma, como se describe anteriormente, que requiere la dimerización para la actividad de escisión. Se pueden requerir dos nucleasas de dedos de zinc para la escisión, ya que cada nucleasa comprende un monómero del dímero de enzima activa. Alternativamente, una nucleasa de dedos de zinc única puede comprender ambos monómeros para crear un dímero de enzima activa. Como se usa en la presente, un "dímero de enzima activa" es un dímero de enzima capaz de escindir una molécula de ácido nucleico. Los dos monómeros de escisión se pueden derivar de la misma endonucleasa (o fragmentos funcionales de la misma), o cada monómero puede derivar de una endonucleasa diferente (o fragmentos funcionales de la misma).

Cuando se usan dos monómeros de escisión para formar un dímero de enzima activa, los sitios de reconocimiento para las dos nucleasas de dedos de cinc se disponen preferiblemente de modo que la unión de las dos nucleasas de dedos de cinc a sus respectivos sitios de reconocimiento coloca los monómeros de escisión en una orientación espacial entre sí eso permite que los monómeros de escisión formen un dímero de enzima activa, por ejemplo, por dimerización. Como resultado, los bordes cercanos de los sitios de reconocimiento pueden estar separados por aproximadamente 5 a aproximadamente 18 nucleótidos. Por ejemplo, los bordes cercanos pueden estar separados por aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 o 18 nucleótidos. Sin embargo, se entenderá que cualquier número entero de nucleótidos o pares de nucleótidos puede intervenir entre dos sitios de reconocimiento (por ejemplo, de aproximadamente 2 a aproximadamente 50 pares de nucleótidos o más). Los bordes cercanos de los sitios de reconocimiento de las nucleasas de dedos de zinc, tal como por ejemplo los descritos en detalle en la presente, pueden estar separados por 6 nucleótidos. En general, el sitio de escisión se encuentra entre los sitios de reconocimiento.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unirse a la secuencia específica del ADN (en un sitio de reconocimiento), y de escindir el ADN en o cerca del sitio de unión. Ciertas enzimas de restricción (por ejemplo, Tipo IIS) escinden el ADN en sitios eliminados del sitio de reconocimiento y tienen dominios de unión y escisión separables. Por ejemplo, la enzima tipo IIS FokI cataliza la escisión de la cadena doble de ADN, a 9 nucleótidos de su sitio de reconocimiento en una cadena y 13 nucleótidos de su sitio de reconocimiento en la otra. Ver, por ejemplo, la patente U.S. Núm. 5.356.802; 5.436.150 y 5.487.994; así como Li et al. (1992) Proc. Natl Acad Sci. USA 89: 4275-4279; Li et al. (1993) Proc. Natl Acad Sci. USA 90: 2764-2768; Kim et al. (1994a) Proc. Natl Acad Sci. USA 91: 883-887; Kim et al. (1994b) J. Biol. Chem. 269: 31978-31982. Por lo tanto, una nucleasa de dedos de cinc puede comprender el dominio de escisión de al menos una enzima de restricción IS del tipo I y uno o más dominios de unión de dedo de cinc, que pueden estar o no manipulados genéticamente. Los ejemplos de enzimas de restricción tipo IIS se describen, por ejemplo, en la publicación internacional WO 07/014,275. Las enzimas de restricción adicionales también contienen dominios de unión y escisión separables, y estos también se contemplan en la presente descripción. Ver, por ejemplo, Roberts et al. (2003) Nucleic Acids Res. 31: 418-420.

Un ejemplo de enzima de restricción tipo IIS, cuyo dominio de escisión es separable del dominio de unión, es FokI. Esta enzima particular está activa como un dímero (Bitinaite et al. (1998) Proc. Natl. Acad. Sci. USA 95: 10, 570-10, 575). Por consiguiente, para los propósitos de la presente descripción, la porción de la enzima FokI usada en una nucleasa de dedos de zinc se considera un monómero de escisión. Por lo tanto, para la escisión de cadena doble dirigida utilizando un dominio de escisión de FokI, se pueden usar dos nucleasas de dedos de cinc, cada una de las cuales comprende un monómero de escisión de FokI, para reconstituir un dímero de enzima activa. Alternativamente, también se puede usar una molécula polipeptídica única que contiene un dominio de unión a los dedos de zinc y dos monómeros de escisión FokI.

En ciertas realizaciones, el dominio de escisión comprende uno o más monómeros de escisión manipulados genéticamente para minimizar o evitar la homodimerización. A modo de ejemplo no limitante, los residuos de aminoácidos en las posiciones 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 y 538 de FokI son todos dianas para influir en la dimerización de los semi-dominios de escisión de FokI. Los ejemplos de monómeros de escisión de FokI manipulados genéticamente que forman heterodímeros obligados incluyen un par en el que un primer monómero de escisión incluye mutaciones en las posiciones de residuos de aminoácidos 490 y 538 de FokI y un segundo monómero de escisión que incluye mutaciones en las posiciones de residuos de aminoácidos 486 y 499.

Por lo tanto, en una realización de los monómeros de escisión manipulados genéticamente, una mutación en la posición del aminoácido 490 reemplaza a Glu (E) con Lys (K); una mutación en el residuo de aminoácido 538 reemplaza a Iso (I) con Lys (K); una mutación en el residuo de aminoácido 486 reemplaza Gin (Q) con Glu (E); y una mutación en la posición 499 reemplaza Iso (I) con Lys (K). Específicamente, los monómeros de escisión manipulados genéticamente se pueden preparar mediante la mutación de las posiciones 490 de E a K y 538 de I a K en un monómero de escisión para producir un monómero de escisión denominado como "E490K: I538K" y mutación de las posiciones 486 de Q a E y 499 de I a K en otro monómero de escisión para producir un monómero de escisión denominado como "Q486E: I499K". Los monómeros de escisión manipulados genéticamente descritos anteriormente son mutantes heterodímeros obligados en los que se minimiza o anula la escisión aberrante. Los monómeros de escisión manipulados genéticamente se pueden preparar usando un procedimiento adecuado, por ejemplo, mediante mutagénesis dirigida al sitio de monómeros de escisión de tipo salvaje (FokI) como se describe en la Patente U.S. Núm. 7.888.121.

(ii) polinucleótido opcional

En algunas realizaciones, el procedimiento para la edición del genoma dirigido comprende además la introducción en la célula de al menos un polinucleótido que comprende una secuencia que tiene una identidad de secuencia sustancial con una secuencia en al menos un lado del sitio de escisión diana. Por ejemplo, el polinucleótido puede comprender una primera secuencia que tiene una identidad de secuencia sustancial con la secuencia en un lado del sitio de escisión diana y una segunda secuencia que tiene una identidad de secuencia sustancial con la secuencia en el otro lado del sitio de escisión diana. Alternativamente, el polinucleótido puede comprender una primera secuencia que tiene una identidad de secuencia sustancial con la secuencia en un lado del sitio de escisión diana y una segunda secuencia que tiene una identidad de secuencia sustancial con una secuencia ubicada lejos del sitio de escisión diana. La secuencia ubicada lejos del sitio de escisión diana puede ser de decenas, centenas o miles de nucleótidos corriente arriba o corriente abajo del sitio de escisión diana.

Las longitudes de la primera y la segunda secuencia en el polinucleótido que tienen una identidad de secuencia sustancial con las secuencias en la secuencia cromosómica pueden variar y variarán. En general, cada una de las primera y segunda secuencias del polinucleótido tiene una longitud de al menos aproximadamente 10 nucleótidos de longitud. En diversas realizaciones, las secuencias de polinucleótidos que tienen una identidad de secuencia sustancial con las secuencias cromosómicas pueden variar de aproximadamente 10 a 30 nucleótidos, de aproximadamente 30 a aproximadamente 100 nucleótidos, de aproximadamente 100 a aproximadamente 300 nucleótidos, de aproximadamente 300 a aproximadamente 1000 nucleótidos, de aproximadamente 1000 a aproximadamente 3000 nucleótidos, o más de 3000 de longitud.

La frase "identidad de secuencia sustancial" significa que las secuencias en el polinucleótido tienen al menos aproximadamente el 75% de identidad de secuencia con las secuencias cromosómicas de interés. Por ejemplo, al menos una de las secuencias en el polinucleótido puede ser idéntica a la secuencia cromosómica diana, excepto que contiene una supresión, inserción y/o sustitución de al menos un nucleótido de tal manera que, después de la recombinación homóloga, se introduce una secuencia alterada en la secuencia cromosómica. En diversas realizaciones, las secuencias en el polinucleótido pueden tener aproximadamente 75%, 76%, 77%, 78%, 79%, 80%, 81 %, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% de identidad de secuencia con la secuencia cromosómica de interés.

En realizaciones alternativas, el polinucleótido puede comprender una secuencia adicional que está flanqueada por las secuencias que tienen una identidad de secuencia sustancial con la secuencia cromosómica. Después de la recombinación homóloga, la secuencia adicional se puede integrar en la secuencia cromosómica, de este modo se inactiva la secuencia cromosómica y/o modifica la secuencia cromosómica.

La longitud del polinucleótido puede y podrá variar. Por ejemplo, el polinucleótido puede variar de aproximadamente 20 nucleótidos de longitud up a aproximadamente 200.000 nucleótidos de longitud. En varias realizaciones, el polinucleótido varía de aproximadamente 20 nucleótidos a aproximadamente 100 nucleótidos de longitud, de aproximadamente 100 nucleótidos a aproximadamente 1000 nucleótidos de longitud, de aproximadamente 1000 nucleótidos a aproximadamente 10.000 nucleótidos de longitud, de aproximadamente 10.000 nucleótidos a aproximadamente 100.000 nucleótidos de longitud, o de aproximadamente 100.000 nucleótidos a aproximadamente 200.000 nucleótidos de longitud.

Típicamente, el polinucleótido será ADN. El ADN puede ser de cadena simple o cadena doble. El polinucleótido donante puede ser un plásmido de ADN, un cromosoma artificial bacteriano (BAC), un cromosoma artificial de levadura (YAC), un vector viral, un fragmento lineal de ADN, un fragmento de PCR, un ácido nucleico desnudo o un ácido nucleico complejo con un vehículo de administración tal como un liposoma o poloxámero.

En algunas realizaciones, el polinucleótido puede comprender además un marcador. Los ejemplos no limitantes de marcadores adecuados incluyen sitios de restricción, proteínas fluorescentes o marcadores seleccionables. Tales marcadores permiten la detección de integraciones dirigidas.

(iii) introducción en la célula

La endonucleasa de direccionamiento se puede introducir en la célula como una proteína o como un ácido nucleico que codifica la endonucleasa de direccionamiento. El ácido nucleico que codifica la endonucleasa de direccionamiento puede ser ADN o ARN (es decir, ARNm). En realizaciones en las que el ácido nucleico codificador es ARNm, el ARNm puede estar terminado en 5'y/o 3' poliadenilado. En realizaciones en las que el ácido nucleico codificador es ADN, el ADN puede ser lineal o circular. El ADN puede formar parte de un vector, en el que el ADN codificador está opcionalmente unido operativamente a un promotor adecuado. Se puede encontrar información adicional sobre vectores, promotores, otros elementos de control y medios para introducir el vector en la célula apropiados, por ejemplo, en Ausubel et al, 2003, supra y/o Sambrook & Russell, 2001, supra.

La endonucleasa de direccionamiento o ácido nucleico que codifica la endonucleasa de direccionamiento y el polinucleótido opcional descrito anteriormente se pueden introducir en la célula por una variedad de medios. Los medios de administración adecuados incluyen microinyección, electroporación, sonoporación, biolística, transfección

mediada por fosfato de calcio, transfección catiónica, transfección de liposomas, transfección de dendrímeros, transfección por choque térmico, transfección de nucleofección, magnetofección, lipofección, impalefacción, transfección óptica, adaptación mejorada por agente patentado de ácidos nucleicos y administración a través de liposomas, inmunoliposomas, virosomas o viriones artificiales. En ciertas realizaciones, la molécula de endonucleasa de direccionamiento y los polinucleótidos opcionales se introducen en una célula por nucleofección o electroporación.

En realizaciones en las que se introducen más de una molécula de endonucleasa de direccionamiento y más de un polinucleótido en una célula, las moléculas se pueden introducir de forma simultánea o secuencial. Por ejemplo, las moléculas de endonucleasas de direccionamiento, cada una específica para un sitio de escisión diana (y polinucleótidos opcionales) se pueden introducir al mismo tiempo. Alternativamente, cada molécula de endonucleasa de direccionamiento, así como los polinucleótidos opcionales, se pueden introducir secuencialmente.

La relación de la molécula de endonucleasa de direccionamiento (o ácido nucleico codificador) al polinucleótido opcional puede variar y variará. En general, la relación de la molécula de endonucleasa de direccionamiento al polinucleótido puede variar de aproximadamente 1:10 a aproximadamente 10: 1. En varias realizaciones, la relación de la molécula de endonucleasa de direccionamiento al polinucleótido es aproximadamente 1:10, 1:9, 1:8, 1:7, 1:6, 1:5, 1:4, 1:3, 1:2, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, o 10:1.. En una realización, la relación es de aproximadamente 1:1.

(iv) ejemplos de realizaciones

En algunas realizaciones, la línea celular deficiente en Mgat1 se prepara mediante la manipulación genética de una nucleasa de dedos de zinc (ZFN) para escindir una secuencia específica en el gen Mgat1. Después de la introducción de la ZFN o ácido nucleico que codifica la ZFN en la línea celular progenitora, la ZFN se une y escinde la secuencia específica en el gen Mgat1. El procedimiento de NHEJ propenso a errores repara la ruptura de la cadena doble en el gen Mgat1, de este modo introduce una supresión, inserción y/o sustitución de al menos una pb, de manera que el gen Mgat1 está inactivado. En general, una región de la secuencia codificadora de Mgat1 se suprime de tal manera que el marco de lectura se altera y no se produce ninguna proteína Mgat1 por la línea celular deficiente en Mgat1. En algunas realizaciones, la línea celular deficiente en Mgat1 generada a través de la edición del genoma mediada por ZFN se enriquece aún más mediante la selección en presencia de RCA-I. En algunas realizaciones, la línea celular deficiente en Mgat1 se prepara mediante el contacto de una línea celular CHO con una ZFN dirigida a Mgat1. En algunos casos, la línea celular CHO es GS -/-.

(b) Interferencia de ARN

En otra realización, la línea celular deficiente en Mgat1 se puede preparar usando un agente de interferencia de ARN (ARNi) que inhibe la expresión de un ARNm o transcrito diana. El agente de ARNi puede llevar a la escisión del ARNm o transcrito diana. Alternativamente, el agente de ARNi puede evitar o alterar el ARNm diana en proteína.

En algunas realizaciones, el agente de ARNi puede ser un ARN de interferencia corto (ARNsi). En general, un ARNsi comprende una molécula de ARN de cadena doble que varía de aproximadamente 15 a aproximadamente 29 nucleótidos de longitud. El ARNsi puede tener aproximadamente 16-18, 17-19, 21-23, 24-27 o 27-29 nucleótidos de longitud. En una realización específica, el ARNsi tiene aproximadamente 21 nucleótidos de longitud. El ARNsi opcionalmente puede comprender además uno o dos proyecciones de cadena simple, por ejemplo, una proyección 3' en uno o ambos extremos. El ARNsi se puede formar a partir de dos moléculas de ARN que se hibridan entre sí o, alternativamente, se puede generar a partir de un ARN de horquilla corto (ARNsh) (ver más abajo). En algunas realizaciones, las dos cadenas del ARNsi son completamente complementarias, de modo que no existen errores de apareamiento o protuberancias en el dúplex formado entre las dos secuencias. En otras realizaciones, las dos cadenas del ARNsi son sustancialmente complementarias, de modo que pueden existir uno o más errores de apareamiento y/o protuberancias en el dúplex formado entre las dos secuencias. En ciertas realizaciones, uno o ambos extremos 5' del ARNsi tienen un grupo fosfato, mientras que en otras realizaciones uno o ambos extremos 5' carecen de un grupo fosfato. En otras realizaciones, uno o ambos extremos 3' del ARNsi tienen un grupo hidroxilo, mientras que en otras realizaciones uno o ambos extremos 5' carecen de un grupo hidroxilo.

Una cadena del ARNsi, que se denomina como "cadena antisentido" o "cadena guía", incluye una porción que se hibrida con el transcrito diana. En ciertas realizaciones, la cadena antisentido del ARNsi es completamente complementaria con una región del transcrito diana, es decir, se hibrida con el transcrito diana sin un apareamiento erróneo o protuberancia únicos sobre una región diana entre aproximadamente 15 y aproximadamente 29 nucleótidos de longitud, preferiblemente al menos 16 nucleótidos de longitud, y más preferiblemente aproximadamente de 18-20 nucleótidos de longitud. En otras realizaciones, la cadena antisentido es sustancialmente complementaria de la región diana, es decir, pueden existir uno o más errores de apareamiento y/o protuberancias en el dúplex formado por la cadena antisentido y el transcrito diana. Típicamente, los ARNsi están dirigidos a secuencias exónicas del transcrito diana. Los expertos en la técnica están familiarizados con los programas, algoritmos y/o servicios comerciales que diseñan los ARNsi para transcritos diana. Un ejemplo es el algoritmo de diseño de ARNsi de Rosetta (Rosetta Inpharmatics, North Seattle, WA) y el ARNsi de MISSION® (Sigma-Aldrich, St. Louis, MO). El ARNsi se puede sintetizar enzimáticamente in vitro usando procedimientos bien conocidos por los expertos en la técnica. Alternativamente, el ARNsi se puede sintetizar químicamente usando técnicas de síntesis de oligonucleótidos que son

bien conocidas en la técnica.

En otras realizaciones, el agente de ARNi puede ser un ARN de horquilla corto (ARNsh). En general, un ARNsh es una molécula de ARN que comprende al menos dos porciones complementarias que están hibridadas o que se pueden hibridar para formar una estructura de doble cadena suficientemente larga para mediar la interferencia del ARN (como se describe anteriormente), y al menos una porción de la cadena simple que forma un bucle que conecta las regiones del ARNsh que forman el dúplex. La estructura también se denomina estructura de tallo-bucle, y el tallo es la porción del dúplex. En algunas realizaciones, la porción del dúplex de la estructura es completamente complementaria, de modo que no existen errores de apareamiento o protuberancias en la región dúplex del ARNsh. En otras realizaciones, la porción del dúplex de la estructura es sustancialmente complementaria, de modo que existen uno o más errores de apareamiento y/o protuberancias en la porción del dúplex del ARNsh. El bucle de la estructura puede tener de aproximadamente 1 a aproximadamente 20 nucleótidos, preferiblemente de aproximadamente 4 a aproximadamente 10 aproximadamente nucleótidos de longitud, y más preferiblemente de aproximadamente 6 a aproximadamente 9 nucleótidos de longitud. El bucle se puede ubicar en el extremo 5' o 3' de la región que es complementaria al transcripto diana (es decir, la porción antisentido del ARNsh).

El ARNsh también puede comprender una proyección en el extremo 5' o 3'. La proyección opcional puede ser de aproximadamente 1 a aproximadamente 20 nucleótidos de longitud, y más preferiblemente de aproximadamente 2 a aproximadamente 15 nucleótidos de longitud. En algunas realizaciones, la proyección comprende uno o más residuos U, por ejemplo, entre aproximadamente 1 y aproximadamente 5 residuos U. En algunas realizaciones, el extremo 5' del ARNsh tiene un grupo fosfato, mientras que en otras realizaciones no lo tiene. En otras realizaciones, el extremo 3' del ARNsh tiene un grupo hidroxilo, mientras que en otras realizaciones no lo tiene. En general, los ARNsh se procesan en ARNsi mediante la maquinaria de ARNi celular conservada. Por lo tanto, los ARNsh son precursores de los ARNsi y son igualmente capaces de inhibir la expresión de un transcripto diana que es complementario con una porción del ARNsh (es decir, la porción antisentido del ARNsh). Los expertos en la técnica están familiarizados con los recursos disponibles (como se detalla anteriormente) para el diseño y la síntesis de los ARNsh. Un ejemplo de ello son los ARNsh de MISSION® (Sigma-Aldrich).

En otras realizaciones más, el agente de ARNi puede ser un vector de expresión de ARNi. Típicamente, un vector de expresión de ARNi se usa para la síntesis intracelular (in vivo) de agentes de ARNi, tales como ARNsi o ARNsh. En una realización, dos cadenas de ARNsi complementarias, separadas se transcriben utilizando un vector único que contiene dos promotores, cada uno de los cuales dirige la transcripción de una cadena simple de ARNsi (es decir, cada promotor está unido de manera operativa a un molde para el ARNsi, de modo que se puede producir la transcripción). Los dos promotores pueden estar en la misma orientación, en cuyo caso cada uno está unido operativamente a un molde para una de las cadenas de ARNsi complementarias. Alternativamente, los dos promotores pueden estar en orientaciones opuestas, flanqueando un molde único de modo que la transcripción de los promotores produzca la síntesis de dos cadenas de ARNsi complementarias. En otra realización, el vector de expresión de ARNi puede contener un promotor que dirige la transcripción de una molécula de ARN única que comprende dos regiones complementarias, de manera que la transcripción forma un ARNsh.

Los expertos en la técnica apreciarán que es preferible que los agentes de ARNsi y ARNsh se produzcan in vivo a través de la transcripción de más de una unidad de transcripción. En términos generales, los promotores utilizados para dirigir la expresión in vivo de una o más unidades de transcripción de ARNsi o ARNsh pueden ser promotores de la ARN polimerasa III (Pol III). Ciertos promotores Pol III, tales como los promotores U6 o H1, no requieren elementos reguladores de acción en cis dentro de la región transcrita, y por lo tanto, se prefieren en ciertas realizaciones. En otras realizaciones, los promotores para Pol II se pueden usar para dirigir la expresión de una o más unidades de transcripción de ARNsi o ARNsh. En algunas realizaciones, se pueden usar promotores de Pol II específicos de tejido, específicos de células o inducibles.

Un constructo que proporciona un molde para la síntesis de ARNsi o ARNsh se puede producir usando procedimientos estándar de ADN recombinante e insertar en cualquiera de una amplia variedad de diferentes vectores adecuados para la expresión en células eucarióticas. Las técnicas de ADN recombinante se describen en Ausubel et al, 2003, supra y Sambrook & Russell, 2001, supra. Los expertos en la técnica también aprecian que los vectores pueden comprender secuencias reguladoras adicionales (por ejemplo, secuencia de terminación, secuencia de control de traducción, etc.), así como secuencias marcadoras seleccionables. Los plásmidos de ADN son conocidos en la técnica, incluidos los basados en pBR322, PUC, y similares. Debido a que muchos vectores de expresión ya contienen un promotor o promotores adecuados, puede que solo sea necesario insertar la secuencia de ácido nucleico que codifica el agente de ARNi de interés en una ubicación apropiada con respecto al promotor o promotores. Los vectores virales también se pueden usar para proporcionar la expresión intracelular de agentes de ARNi. Los vectores víricos adecuados incluyen vectores retrovirales, vectores lentivirales, vectores adenovirales, vectores de virus adenoasociados, vectores de virus del herpes y similares. En una realización específica, el vector de expresión de ARNi es un vector basado en ARNsh lentiviral o partícula lentiviral, tal como el proporcionado en productos de ARNsh MISSION®TRC (Sigma Aldrich).

Los agentes de ARNi o los vectores de expresión de ARNi se pueden introducir en la célula usando procedimientos bien conocidos por los expertos en la técnica. Estas técnicas se describen, por ejemplo, en Ausubel et al., 2003, supra o Sambrook & Russell, 2001, supra. En ciertas realizaciones, el vector de expresión de ARNi, por ejemplo, un vector

viral, se integra de forma estable en el genoma de la célula, de modo que la expresión de Mgat1 se altera a lo largo de las generaciones celulares posteriores.

(c) recombinación específica del sitio

En realizaciones alternativas, la línea celular deficiente en Mgat1 se puede preparar usando técnicas de recombinación específicas del sitio. Por ejemplo, las técnicas de recombinación específicas del sitio se pueden usar para suprimir la totalidad o parte de una secuencia cromosómica de interés, o introducir polimorfismos de nucleótido único (SNP) en la secuencia cromosómica de interés. En una realización, la secuencia cromosómica de interés se dirige utilizando un sistema de recombinación específico del sitio Cre-loxP, un sistema de recombinación específico del sitio FIp-FRT, o variantes del mismo. Tales sistemas de recombinación están disponibles comercialmente, y se encuentran enseñanzas adicionales para estas técnicas en por ejemplo, Ausubel et al., 2003, supra.

(III) Procedimientos para producir proteínas recombinantes con glicoformas simples

También se proporcionan procedimientos para usar las células deficientes en Mgat1 para producir proteínas recombinantes con glicoformas simples. Una glicoforma simple se refiere a una glicoforma que comprende uno o más residuos de manosa terminales. En general, una glicoforma simple carece de residuos de galactosa y/o ácido siálico. Un ejemplo de glicoforma simple es Man5GlcNac2 o Man5. Como se mencionó anteriormente, las glicoproteínas con glicoformas simples son útiles para los estudios de cristalografía de rayos X. Las glicoproteínas con glicoformas simples tienen aplicaciones adicionales. Por ejemplo, las células deficientes en Mgat1 se pueden usar para la producción de vacunas. Las células deficientes en Mgat1 producen glicoproteínas que contienen residuos de manosa terminales. Los residuos de manosa terminales en el antígeno permiten su captación por los receptores de manosa en las células presentadoras de antígeno.

Las células presentadoras de antígeno (es decir, macrófagos o células dendríticas) posteriormente presentan el antígeno para el reconocimiento por parte de las células T, de este modo se monta una respuesta inmune. El uso de antígenos que comprenden residuos de manosa terminales puede mejorar la eficiencia de la presentación de antígenos a las células T. Además, las glicoproteínas con solo residuos de manosa terminales pueden ser útiles para otros fines terapéuticos al facilitar la captación mediada por el receptor de manosa de estas proteínas.

Se pueden preparar glicoproteínas con glicoformas simples. Un aspecto adicional de la presente descripción abarca un procedimiento para producir una proteína de recombinación con una glicoforma simple. El procedimiento comprende la introducción de un ácido nucleico que codifica la proteína recombinante en una línea celular deficiente en Mgat1 y que expresa la proteína de recombinación, en donde la proteína recombinante expresada comprende uno o más residuos de manosa terminales. Las líneas celulares deficientes de Mgat1 se describen anteriormente en la sección (I).

La glicoproteína recombinante producida en la línea celular deficiente en Mgat1 puede ser cualquier glicoproteína adecuada, que incluye las proteínas terapéuticas y proteínas biológicas. Por ejemplo, la proteína recombinante puede ser, sin límite, un anticuerpo, un fragmento de un anticuerpo, un anticuerpo monoclonal, un anticuerpo humanizado, un anticuerpo monoclonal humanizado, un anticuerpo quimérico, una molécula de IgG, una cadena pesada de IgG, una cadena liviana de IgG, una molécula de IgA, una molécula de IgD, una molécula de IgE, una molécula de IgM, una vacuna, un factor de crecimiento, una citoquina, un interferón, una interleuquina, una hormona, un factor coagulante (o de coagulación), un componente sanguíneo, un enzima, una proteína nutracéutica, un fragmento funcional o variante funcional de cualquiera de los anteriores, o una proteína de fusión que comprende cualquiera de las proteínas anteriores y/o fragmentos funcionales o variantes de los mismos. En ejemplos de realizaciones, la glicoproteína recombinante es una proteína humana.

Los procedimientos para producir proteína recombinante son bien conocidos en la técnica, y proporcionan enseñanzas adicionales por Ausubel et al., 2003, supra. En general, la proteína recombinante se expresa a partir de un ácido nucleico introducido de forma exógena. Como se detalla anteriormente en la sección (I) (e), el ácido nucleico que codifica la proteína recombinante puede ser extracromosómico o el ácido nucleico que codifica la proteína recombinante se puede integrar en el genoma.

Los procedimientos para cultivar la línea celular de manera que la proteína recombinante se exprese son bien conocidos en la técnica. Los medios y sistemas de cultivo apropiados son conocidos en la técnica y están disponibles comercialmente. En una realización, la proteína recombinante se produce por las líneas celulares descritas en la presente a través de un cultivo en suspensión libre de suero.

Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en la presente tienen el significado comúnmente entendido por un experto en la técnica a la que pertenece esta invención. Las siguientes referencias proporcionan a los expertos una definición general de muchos de los términos utilizados en esta invención: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991). Como se usa en la presente, los siguientes

términos tienen los significados que se les asigna a menos que se especifique lo contrario.

Quando se introducen elementos de la presente descripción las realizaciones preferidas de la misma, los artículos "un", "una", y "dicho" se considera que significan que hay uno o más de los elementos. Los términos "que comprende", "que incluye" y "que tiene" se consideran inclusivos y significan que puede haber elementos adicionales distintos de los elementos enumerados.

Como se usa en la presente, el término "secuencia endógena" se refiere a una secuencia cromosómica que es nativa de la célula.

El término "secuencia exógena" se refiere a una secuencia cromosómica que no es nativa de la célula, o una secuencia cromosómica cuya ubicación cromosómica nativa se encuentra en una ubicación diferente en un cromosoma.

Los términos "edición", "edición genómica" o "edición cromosómica" se refieren a un procedimiento mediante el cual se cambia una secuencia cromosómica específica. La secuencia cromosómica editada puede comprender una inserción de al menos un nucleótido, una supresión de al menos un nucleótido, y/o una sustitución de al menos un nucleótido.

Un "gen", como se usa en la presente, se refiere a una región de ADN (que incluye exones e intrones) que codifica un producto génico, así como a todas las regiones de ADN que regulan la producción del producto génico, ya sea que dichas secuencias reguladoras sean o no adyacente a las secuencias codificadoras y/o transcritas. Por consiguiente, un gen incluye, pero no está necesariamente limitado a, secuencias promotoras, terminadores, secuencias reguladoras de la traducción tales como sitios de unión a ribosomas y sitios de entrada de ribosoma internos, potenciadores, silenciadores, aisladores, elementos de límite, orígenes de replicación, sitios de unión de matriz, y regiones de control de locus.

El término "heterólogo" se refiere a una entidad que no es nativa de la célula o especie de interés.

Los términos "ácido nucleico" y "polinucleótido" se refieren a un polímero de desoxirribonucleótido o ribonucleótido, en conformación lineal o circular, y en forma de cadena simple o doble. Para los fines de la presente descripción, estos términos no se deben interpretar como limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en la base, azúcar y/o restos fosfato (por ejemplo, esqueletos de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de apareamiento de bases; es decir, un análogo de A se apareará con T.

El término "nucleótido" se refiere a desoxirribonucleótidos o ribonucleótidos. Los nucleótidos pueden ser nucleótidos estándares (es decir, adenosina, guanosina, citidina, timidina y uridina) o análogos de nucleótidos. Un análogo de nucleótido se refiere a un nucleótido que tiene una base de purina o pirimidina modificada o un resto de ribosa modificada. Un análogo de nucleótido puede ser un nucleótido natural (por ejemplo, inosina) o un nucleótido no natural. Los ejemplos no limitantes de modificaciones en los restos de azúcar o base de un nucleótido incluyen la adición (o eliminación) de grupos acetilo, grupos amino, grupos carboxilo, grupos carboximetilo, grupos hidroxilo, grupos metilo, grupos fosforilo y grupos tiol, así también como la sustitución de los átomos de carbono y nitrógeno de las bases con otros átomos (por ejemplo, 7-deaza purinas). Los análogos de nucleótidos también incluyen nucleótidos didesoxi, nucleótidos 2'-O-metilo, ácidos nucleicos bloqueados (LNA), ácidos nucleicos peptídicos (PNA) y morfolinós.

Los términos "polipéptido" y "proteína" se usan indistintamente para referirse a un polímero de residuos de aminoácidos.

El término "recombinación" se refiere a un procedimiento de intercambio de información genética entre dos polinucleótidos. Para los fines de esta descripción, "recombinación homóloga" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de rupturas de cadena doble en las células. Este procedimiento requiere una secuencia de similitud entre los dos polinucleótidos, utiliza una molécula "donante" o "de intercambio" para reparar el molde de una molécula "diana" (es decir, la que experimentó la ruptura de la cadena doble), y se conoce de diversas maneras como "conversión del gen no cruzada" o "conversión del gen del tracto corto", porque conduce a la transferencia de información genética del donante al diana. Sin estar limitado por ninguna teoría en particular, dicha transferencia puede involucrar la corrección de errores de apareamiento de ADN heterodúplex que se forma entre el diana roto y el donante, y/o el "apareamiento de cadenas dependiente de la síntesis", en el que el donante se utiliza para resintetizar la información genética que se convertirá en parte del diana, y/o procedimientos relacionados.

Dicha recombinación homóloga especializada a menudo produce una alteración de la secuencia de la molécula diana, de manera que parte o toda la secuencia del polinucleótido donante se incorpora en el polinucleótido diana.

Como se usa en la presente, los términos "sitio diana" o "secuencia diana" se refiere a una secuencia de ácidos nucleicos que define una porción de una secuencia cromosómica para editar y en la que una endonucleasa de direccionamiento se manipula genéticamente para reconocer, unir y escindir.

Los términos "corriente arriba" y "corriente abajo" se refieren a ubicaciones en una secuencia de ácido nucleico en

relación con una posición fija. Corriente arriba se refiere a la región que es 5' (es decir, cerca del extremo 5' de la cadena) a la posición y corriente abajo se refiere a la región que es 3' (es decir, cerca del extremo 3' de la cadena) a la posición.

Las técnicas para determinar la identidad de la secuencia de aminoácidos y aminoácidos son conocidas en la materia.

5 Típicamente, tales técnicas incluyen determinar la secuencia de nucleótidos del ARNm para un gen y/o determinar la secuencia de aminoácidos codificada de ese modo, y comparar estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también se pueden determinar y comparar de esta manera. En general, la identidad se refiere a una correspondencia exacta de nucleótido a nucleótido o de aminoácido a aminoácido de dos secuencias depolinucleótidos o polipeptidos, respectivamente. Se pueden comparar dos o más secuencias (polinucleótido o aminoácido) mediante la determinación de su porcentaje de identidad. El porcentaje de identidad de dos secuencias, ya sean secuencias de ácido nucleico o aminoácidos, es el número de coincidencias exactas entre dos secuencias alineadas divididas por la longitud de las secuencias más cortas y multiplicadas por 100. Unaproximado para las secuencias de ácido nucleico se proporciona mediante el algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981). Este algoritmo se puede aplicar a las secuencias de aminoácidos utilizando la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA y normalizada por Gribskov, *Nucl. Acidos Res.* 14 (6): 6745-6763 (1986). Un ejemplo de implementación de este algoritmo para determinar el porcentaje de identidad de una secuencia se proporciona mediante Genetics Computer Group (Madison, Wis) en la aplicación de utilidad "BestFit". Otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias son generalmente conocidos en la técnica, por ejemplo, otro programa de alineamiento es BLAST, utilizado con parámetros predeterminados. Por ejemplo, BLASTN y BLASTP se pueden usar usando los siguientes parámetros predeterminados: código genético = estándar; filtro = ninguno; cadena = ambos; corte = 60; espera = 10; Matriz = BLOSUM62; Descripciones = 50 secuencias; ordenar por = ALTA PUNTUACIÓN; Bases de datos = no redundante, GenBank+EMBL+DDBJ+PDB+GenBank CDS translations+Swiss protein+Spupdate+PIR. Los detalles de estos programas se pueden encontrar en el sitio web de GenBank. Con respecto a las secuencias descritas en la presente, el intervalo de grados deseados de identidad de secuencia es aproximadamente de 80% a 100% y cualquier valor entero entre ellos. Típicamente, el porcentaje de identidades entre secuencias es al menos 70-75%, preferiblemente 80-82%, más preferiblemente 85-90%, incluso más preferiblemente 92%, aún más preferiblemente 95%, y de máxima preferencia 98% de identidad de secuencia mediante la determinación de su porcentaje de identidad. El porcentaje de identidad de dos secuencias, ya sean secuencias de ácido nucleico o aminoácidos, es el número de coincidencias exactas entre dos secuencias alineadas divididas por la longitud de las secuencias más cortas y multiplicadas por 100. Un alineamiento aproximado para las secuencias de ácido nucleico se proporciona mediante el algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981). Este algoritmo se puede aplicar a secuencias de aminoácidos utilizando la matriz de puntuación desarrollada por Dayhoff, *Protein Sequences and Structure*, M. O. Dayhoff ed., 5 suppl. 3: 353-358, National Biomedical Research Foundation, Washington, D.C., USA, y normalizada por Gribskov, *Nucl. Acidos Res.* 14 (6): 6745-6763 (1986). Un ejemplo de implementación de este algoritmo para determinar el porcentaje de identidad de una secuencia se proporciona mediante Genetics Computer Group (Madison, Wis) en la aplicación de utilidad "BestFit". Otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias son generalmente conocidos en la materia, por ejemplo, otro programa de alineación es BLAST, utilizado con parámetros predeterminados. Por ejemplo, BLASTN y BLASTP se pueden usar usando los siguientes parámetros predeterminados: código genético = estándar; filtro = ninguno; cadena = ambos; corte = 60; espera = 10; Matriz = BLOSUM62; Descripciones = 50 secuencias; ordenar por = ALTA PUNTUACIÓN; Bases de datos = no redundante, GenBank+EMBL+DDBJ+PDB+GenBank CDS translations+Swiss protein+Spupdate+PIR. Los detalles de estos programas se pueden encontrar en el sitio web de GenBank. Con respecto a las secuencias descritas en la presente, el intervalo de grados deseados de identidad de secuencia es aproximadamente del 80% a 100% y cualquier valor entero entre ellos. Típicamente, el porcentaje de identidades entre secuencias es al menos 70-75%, preferiblemente 80-82%, más preferiblemente 85-90%, incluso más preferiblemente 92%, aún más preferiblemente 95%, y con máxima preferencia 98% de identidad de secuencia.

50 Como se podrían realizar varios cambios en las células y los procedimientos descritos anteriormente sin apartarse del alcance de la invención, se pretende que toda la materia contenida en la descripción anterior y en los ejemplos que se dan a continuación se interprete como ilustrativa y no como un ejemplo en sentido limitante.

Ejemplos

Los siguientes ejemplos ilustran ciertos aspectos de la invención.

55 Ejemplo 1: Alteración de Mgat1 en células CHO usando ZFNs

A. Preparación de vectores de expresión de ZFN y ARNm de ZFN dirigidos a Mgat1

Se emplearon técnicas knockout mediada por ZFN para knockout o inactivar el gen Mgat1 en células CHO, de este modo se generan células que producen proteínas con un perfil de N-glicano simple homogéneo que contiene restos Man5 terminales. Los cóntigos genómicos de Mgat1 que contienen el locus Mgat1 (Xu et al., 2011, *Nat Biotechnol*, 29: 735-741) se anotaron usando homología de ratón. Las ZFN dirigidas a un sitio específico dentro de la región

codificadora de Mgat1 CHO se diseñaron utilizando un algoritmo patentado. La FIG. **2A** diagrama la ubicación del sitio diana en Mgat1. La secuencia diana en Mgat1 fue 5'- AACAAAGTTCAGTTCccagcaGCTGTGGTAGTGGAGGAC-3' (SEQ ID NO: 1; sitios de unión a ZFN en mayúscula y el sitio de escisión en minúscula). El constructo de expresión Mgat1 ZFN se preparó utilizando procedimientos estándares y se produjo ARNm de Mgat1 ZFN a partir de ADN plasmídico ZFN como se describe en la información del producto COMPOZR® nucleasa de dedos de cinc knockout (ZFN) (Sigma-Aldrich) usando `procedimientos de transcripción in vitro, poli-adenilación de ARNm y terminación. En resumen, el ADN del plásmido ZFN se linealizó y se purificó utilizando extracción de ADN con fenol/cloroformo. Se usó el kit de transcripción de mensajes de MessageMax™ T7 ARCAcapped(Cell Script Inc.) para terminar el ADN linealizado. Se utilizó un kit de Poly(A) Polymerase Tailing (EpiCentre) para agregar una cola de poli (A). El ARNm de ZFN se purificó utilizando el kit MEGAClear™ (Ambion).

B. Transfección de las ZFN dirigidas a Mgat1 en una línea celular CHO productora de IgG

Las células CHOZN® (GS -/-) (Sigma Aldrich) que expresan IgG humana anti-rabia recombinante se mantuvieron como cultivos en suspensión en medios de fusión EX-CELL® CHO CD (Sigma-Aldrich) suplementados con metionina sulfoximina 25 µM (MSX). Las células se sembraron a razón de $0,5 \times 10^6$ células/ml en tubos biorreactores un día antes de la transfección. Se utilizaron 1×10^6 células en 150 µl de medio de crecimiento y 5 µg de ADN o ARNm de Mgat1 ZFN para cada transfección. Las transfecciones se realizaron por electroporación a 140 V y 950 µF en cubetas de 0,2 cm. Las células electroporadas se colocaron en 2 ml de medio de crecimiento en un cultivo estático en placa de 6 pocillos. En los días 3 y 10 pos-transfección, las células se extrajeron del cultivo y el ADN genómico se aisló utilizando el kit Miniprep de ADN genómico de mamífero GeneElute™ (Sigma-Aldrich). Se llevó a cabo un ensayo de nucleasa Cel-I, como se describe en la información del producto COMPOZR® Knockout ZFN para determinar la eficiencia de la escisión del gen mediada por ZFN.

La FIG. **2B** representa un gel que muestra los resultados de un ensayo de Cel-I en grupos de células transfectadas con ZFN (transfecciones de ADN o ARNm de plásmido de ZFN). Dos fragmentos de escisión, 220 y 197 pb, estuvieron presentes tanto en el Día 3 como en el Día 10 después de la transfección en los digestos de ADN genómico, lo que indica una actividad ZFN positiva en ambos puntos de tiempo. La actividad de ZFN en el día 3 y el día 10 fue de 8.8% y 9.6%, respectivamente. En ambos puntos de tiempo, las células transfectadas con ARN tenían bandas ligeramente más intensas. La actividad estable de ZFN a lo largo del tiempo no implica una citotoxicidad evidente o inhibición del crecimiento relacionada con la alteración de Mgat1. En otras palabras, no se observó ninguna ventaja de crecimiento de las células tipo salvaje respecto de las modificadas con ZFN en la población transfectada con ZFN.

Ejemplo 2: Selección de lectina y aislamiento de clones de célula única

Las células transfectadas con Mgat1 ZFN se trataron durante la noche con aglutinina-I de *Ricinus communis* (RCA-I), una lectina citotóxica que se une a galactosa pero no a manosa, para enriquecer las células con Mgat1 alterado. El día 14 después de la transfección, se sembró una alícuota de células para la clonación de células únicas usando FACS para células viables. Las células se sembraron en placas a razón de una célula por pocillo utilizando un clasificador de células FACS Aria™ III (Becton-Dickinson) en placas tratadas de cultivo de tejidos de 96 pocillos (Corning) en la mezcla de nutrientes F12 de Ham suplementada con 10% de suero fetal bovino (FBS) y 4 mM de L -glutamina. Las células restantes se expandieron hasta el día 18 y posteriormente se sembraron a razón de 1×10^6 células/ml en matraces T de cultivo en suspensión (Corning). Las células se incubaron con una concentración final de 10 µg/ml de RCA-I en medios de crecimiento durante la noche, posteriormente se lavaron con PBS y se devolvieron a los medios de crecimiento. Las células se expandieron a matraces T y se cultivaron durante 10 días y posteriormente se sembraron en placas para la clonación de células únicas como se describe anteriormente. Diez días después del tratamiento con RCA-I, las células se analizaron para determinar la actividad de ZFN.

La FIG. **3** representa un gel que muestra los resultados de un ensayo de Cel-I en grupos de células transfectadas con ZFN antes y después del enriquecimiento con RCA-1. Las células seleccionadas con RCA-I mostraron una actividad de ZFN de aproximadamente el 33-36% en comparación con el 3.8-5.5% de las células no seleccionadas, lo que indica un enriquecimiento exitoso para las células alteradas con Mgat1.

Ejemplo 3: Glicoperfiles de IgG humana anti-rabia recombinante

Se usó un flujo de trabajo de SEC-MS de alto rendimiento (Waters Acquity UPLC®/Q-ToF Premier TM) para el glicoanálisis. El flujo de trabajo utilizó la purificación con proteína A de IgG a partir de sobrenadantes de cultivos celulares, seguido de un análisis de masa exacto de los componentes de la cadena pesada de IgG intactos. La purificación de proteína A se realizó utilizando una placa de filtro de 96 pocillos con 50 ml de resina de proteína A y 750 ml de sobrenadante de clon por pocillo. La IgG se eluyó con 100 ml de citrato 25 mM, pH 3,0. El análisis de datos en masa se realizó por deconvolución y el procesamiento de los datos se realizó de forma automatizada utilizando el software BiopharmaLLynx™ (Waters). Las formas de glicopéptidos se confirmaron mediante digestión con tripsina de muestras seleccionadas y la verificación de la identificación y distribución relativa de glicanos (informada como un% del total de glicanos) a nivel de glicopéptidos por MALDI-TOF.

Los glicoperfiles de la IgG humana anti-rabia recombinante secretada se evaluaron para cada mezcla de células (transfectadas con ZFN y transfectadas de forma simulada, con o sin selección de RCA-I). Las mezclas enriquecidas

con RCA-I mostraron niveles aumentados de la glicoforma Man5 en relación con las células no expuestas a RCA-I (FIG. 4). Las células transfectadas simuladas demostraron un 0,3% de Man5 en comparación con el 13,7% de las células transfectadas con Mgat1 ZFN y el 92,8% de las células transfectadas con ZFN enriquecidas mediante la selección de RCA-I. Las glicoformas G0F y G1F fueron las más abundantes en las células transfectadas de forma simulada y en las células transfectadas con ZFN que no se sometieron a enriquecimiento de RCA-I.

Ejemplo 4: Caracterización del genotipo de clones de células únicas modificadas con Mgat1

El ADN genómico se aisló a partir de cultivos de los clones de células únicas (ya sea que producen IgG o las líneas celulares huésped GS (-/-)). Una porción del gen Mgat1 se amplificó por PCR y los productos de la PCR se clonaron utilizando el sistema TOPO (Life Technologies, Carlsbad, CA). Se prepararon muestras para secuenciación de ADN. Se secuenció un mínimo de 10 clones TOPO para la caracterización genotípica.

Para comparar las frecuencias de los clones modificados con Mgat1 en presencia o ausencia de enriquecimiento con RCA-I, las células transfectadas con ZFN cultivadas con o sin RCA-I se clonaron en célula única por FACS, como se describió anteriormente. Para las mezclas no enriquecidas, dos clones de 76 (2.63%) tenían supresiones cortas en el gen Mgat1 (ver Figura 5 y Tabla 1; clones #73 y #92). La secuenciación inicial de productos de PCR anidados a partir del lisado celular obtenido con la solución de extracción de ADN QuickExtract™ (Epicentre) indicó que todos los clones (52/52) de la población enriquecida con RCA-I contenían secuencias genéticas de Mgat1 alteradas. La secuenciación de estos clones reveló la supresión en Mgat1 de supresiones de 2 pb a supresiones de 55 pb (ver la Figura 5 y la Tabla 1, clones #21, #25, #27 y #37). Cada clon contenía un alelo alterado, sin que se detectara ninguna secuencia de tipo salvaje. Por lo tanto, el genotipo de estas células es MGAT-1 (-/0) y GS (-/-).

Tabla 1. Genotipos de clones de células únicas de la línea celular CHO productora de IgG knockout Mgat1			
Clon ID	Alelo 1	Alelo 2	Genotipo
#73	Supresión de 7 bp	No detectado	Knockout
#92	Supresión de 4 bp	No detectado	Knockout
#21	Supresión de 18 bp	No detectado	Knockout
#25	Supresión de 4 bp	No detectado	Knockout
#27	Supresión de 55 bp	No detectado	Knockout
#37	Supresión de 2 bp	No detectado	Knockout

Los siguientes clones knockout de Mgat1 que producen IgG se seleccionaron para una caracterización adicional: dos clones aislados sin selección de RCA-I (clones #73 y #92); dos clones que pasaron por el procedimiento de transfección con ZFN que retuvo la secuencia de tipo salvaje (#15 y #46); y cuatro de los clones seleccionados de RCA-I (#21, #25, #27 y #37), con un intervalo de longitudes de supresión. Ninguno de estos clones knockout mostró sensibilidad a RCA-I (FIG. 6), ni se observaron cambios significativos en la distribución de la productividad o crecimiento de IgG.

Ejemplo 5: Caracterización de clones de células productoras de IgG knockout Mgat1

Los glicoperfiles de la IgG humana anti-rabia recombinante secretada se evaluaron (esencialmente como se detalla anteriormente en el Ejemplo 3) para cada uno de los clones de células únicas mencionados anteriormente. La FIG. 7A-B presenta los glicoperfiles de clones de tipo salvaje y knockout, respectivamente, en cultivos de lotes del Día 5, y la FIG. 7C-D muestran los glicoperfiles de clones de tipo salvaje y knockout, respectivamente, en cultivos de lotes del Día 8. La especie de glicano más abundante es Man5 en los clones knockout Mgat1, mientras que los clones de tipo salvaje muestran los picos G0F más altos. El % de Man5 es significativamente mayor en los clones knockout Mgat1 ($p < 0.0001$, prueba t no apareada con corrección de Welch).

Para analizar el crecimiento y la productividad de los clones knockout Mgat1, se sembraron células a razón de $0,3 \times 10^6$ células/ml en 30 ml de medio de crecimiento en tubos de biorreactor TPP por duplicado. La densidad y viabilidad celular se determinaron utilizando un analizador de viabilidad celular ViCell (Beckmann Coulter). Como se muestra en la FIG. 8A-B, los clones de células knockout Mgat1 exhibieron el mismo o mejor crecimiento medido por la viabilidad celular y la densidad celular a lo largo del tiempo en comparación con las células que tienen Mgat1 de tipo salvaje.

La productividad de IgG volumétrica se determinó mediante interferometría utilizando un analizador ForteBio OCTET®

(Pall Life Sciences). Como se muestra en la FIG. 9, los clones de células knockout Mgat1 exhibieron la misma o mejor productividad volumétrica pico que las células de tipo salvaje.

Ejemplo 6: Aislamiento de clones knockoutMgat1 en una línea celular huésped CHO

Para evaluar adicionalmente este procedimiento para aplicaciones biofarmacéuticas, la transfección de ZFN y el aislamiento del clon se repitieron utilizando una línea celular huésped CHO K1 GS (-/-). Una línea celular huésped con GS (-/-) (CHOZN® Sigma-Aldrich) se transfeció con ARNm Mgat1 ZFNesencialmente como se describió anteriormente en el Ejemplo 1. Después de la confirmación de la actividad de ZFN utilizando el ensayo Cel-1, las células transfectadas clonaron en células únicas usando una dilución limitante a razón de 0,5 células/pocillo esencialmente como se describió anteriormente en el Ejemplo 2. La clonalidad y el crecimiento se verificaron microscópicamente en el Día 7 y el Día 14 después de la deposición, respectivamente. Estas células no se trataron con RCA-I para la selección adicional.

Cinco de los 90 (5,6%) clones generados a partir de la línea de células huésped CHOZN® GS (-/-) contenían alteraciones de Mgat1. Las supresiones identificadas en estos clones, generadas sin enriquecimiento de RCA-I, fueron de 7 a 28 pb (ver Tabla 2). Como se observó con los clones productores de IgG, estos clones progenitores tienen solo un alelo alterado y no se identificó ninguna secuencia de tipo salvaje. Se realizó un ensayo de citotoxicidad de lectina utilizando RCA-I, y de modo similar a los clones productores de IgG, no se observó citotoxicidad para ninguno de los clones knockout Mgat1.

Tabla 2. Genotipos de clones de células únicas de una línea celular huésped CHO Mgat1			
Clon ID	Alelo 1	Alelo 2	Genotipo
AB4	Supresión de 24 bp	No detectado	Knockout
BC9	Supresión de 7 bp	No detectado	Knockout
BD7	Supresión de 10 bp	No detectado	Knockout
BH11	Supresión de 157 bp reemplazada por supresión de 104 bp de inserción	No detectado	Knockout
CG10	Supresión de 28 bp	No detectado	Knockout

Ejemplo 7: Expresión de IgG estable y glicoperfilado de IgG de las líneas celulares huésped CHOknockoutMgat1

Se transfectaron cinco clones de la línea celular huésped knockout Mgat1 (enumerados en la Tabla 2) con un vector de expresión que contiene secuencias codificadoras de IgG anti-rabia humana y un casete de expresión de GS de hámster recombinante. Cuarenta y ocho horas antes de la transfección, las células se sembraron a razón de 5×10^5 células/ml. Las transfecciones se realizaron por electroporación como se describe en el Ejemplo 1 usando 30 µg de ADN plasmídico. Las células sometidas a electroporación se colocaron en matraces de cultivo celular de 25 cm² en suspensión en 5 ml de medio de crecimiento. Las células se transfirieron a medios deficientes de L-glutamina 48 horas después de la transfección y se cultivaron a 37 °C en CO₂ 5% hasta que se completó la selección (12-14 días) y se establecieron cultivos transfectados de manera estable. En una transfección separada, las células se sembraron a razón de 5000 células/pocillo en placas de 96 pocillos en 200 µl para la selección menos glutamina y se aumentaron de escala después de completar la selección a cultivos de "mini mezcla" a 24 pocillos (1 ml), matraces de 25 cm² (5 ml) y tubos de TPP (30 ml). A partir de estas líneas estables, los cultivos de TPP se establecieron de manera similar a como se describió anteriormente, y los cultivos se alimentaron con 2 g/l de glucosa junto con una alimentación definida químicamente el día 4 y el día 7. Los cultivos se recolectaron el día 8 y el sobrenadante se clarificó por centrifugación a 3400 rpm y posteriormente se dializó frente a 1xPBS usando un tubo de diálisis de membrana Spectra/Pore 4 (MWCO 12kDa, Spectrum Labs, Rancho Dominguez, CA) durante 24 horas a 4 °C y se purificó en un sistema ÄKTA FPLC (GE Healthcare, Pittsburgh, PA) utilizando una columna HiTrap Proteína A de flujo rápido de 1 ml (GE Healthcare). La IgG se eluyó en citrato 0,1 M, pH 3,0 y las fracciones se neutralizaron inmediatamente mediante la adición de 1/10 de volumen Tris 1,0 M HCl pH 9,0. El glicoperfilado se realizó esencialmente como se describe en el Ejemplo 3.

La IgG anti-rabia producida en las cinco líneas celulares huésped knockout Mgat1 demostró glicoperfiles muy similares, Man5 es la especie predominante sin glicanos complejos detectados. El análisis de MS demostró un desplazamiento de masa del pico de glicano predominante en las líneas celulares knockout Mgat1 (es decir, 50168 - 50172) en comparación con las células de tipo salvaje (es decir, 50400). La IgG producida en las "minimezclas" BC9 y BH11 demostró glicoperfiles comparables a sus mezclas estables seleccionadas "a granel".

LISTADO DE SECUENCIAS

<110> SIGMA-ALDRICH CO. LLC
LIN, Nan
SEALOVER, Natalie
GEORGE, Henery
KAYSER, Kevin

<120> PRODUCCIÓN DE PROTEÍNAS RECOBINANTES COM GLICOFORMAS SIMPLES

<130> 047497-450042

<150> US 61/585,508

<151> 2012-01-11

<150> US 61/636,680

<151> 2012-04-22

<150> US 61/720,268

<151> 2012-10-30

<160> 8

<170> PatentIn version 3.5

<210> 1

<211> 39

<212> DNA

<213> Cricetulus griseus

<400> 1

aacaagttca agtccagc agctgtgga gtggaggac 39

ES 2 733 248 T3

<210> 2
 <211> 85
 <212> ADN
 <213> Cricetulus griseus

<400> 2

aagttcaagt tcccagcagc tgtggtagtg gaggacgac tggaggtggc accagaattc	60
tttgagtact tccaggccac ctacc	85

<210> 3
 <211> 78
 <212> ADN
 <213> Cricetulus griseus

<400> 3

aagttcaagt agctgtggtg gtggaggacg atctggaggt ggcaccagac ttctttgagt	60
acttccaggc cacctacc	78

<210> 4
 <211> 81
 <212> ADN
 <213> Cricetulus griseus

<400> 4

aagttcaagt tccagctgtg gtagtggagg acgatctgga ggtggcacca gacttctttg	60
agtacttcca ggccacctac c	81

<210> 5
 <211> 67
 <212> ADN
 <213> Cricetulus griseus

<400> 5

aagttcaagt tggaggacga tctggaggtg gcaccagact tctttgagta cttccaggcc	60
acctacc	67

<210> 6
 <211> 81
 <212> ADN
 <213> Cricetulus griseus

<400> 6

ES 2 733 248 T3

aagttcaagt tcaagctgtg gtagtggagg acgatctgga ggtggcacca gacttctttg 60

agtacttcca ggccacctac c 81

<210> 7

<211> 30

<212> ADN

<213> Cricetulus griseus

<400> 7

aagttcaagt tgaagatctg gccacctacc 30

<210> 8

<211> 83

<212> ADN

<213> Cricetulus griseus

<400> 8

aagttcaagt tccgcagctg tggtagtgga ggacgatctg gaggtggcac cagacttctt 60

tgagtacttc caggccacct acc 83

REIVINDICACIONES

- 5 1. Una línea celular de mamífero que es deficiente en mannosil(alfa-1,3-)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa 1 (Mgat1) que comprende una supresión de 2 a 300 bp en la secuencia cromosómica que codifica Mgat1.
2. La línea celular de la reivindicación 1, que comprende:
 - (i) una supresión que varía de 2bp a 55bp, o
 - (ii) una supresión de 157bp que se reemplaza por una inserción de 104bp, en la secuencia cromosómica que codifica Mgat1.
- 10 3. La línea celular de la reivindicación 1 o reivindicación 2, en la que la secuencia cromosómica que codifica Mgat1 está inactivada.
4. La línea celular de la reivindicación 3, en la que la secuencia cromosómica se inactiva con una endonucleasa de direccionamiento.
- 15 5. La línea celular de la reivindicación 4, en la que la endonucleasa de direccionamiento es una nucleasa de dedos de cinc.
6. La línea celular de una cualquiera de las reivindicaciones precedentes, en la que la línea celular no produce Mgat1.
7. La línea celular de una cualquiera de las reivindicaciones precedentes, en la que la línea celular además comprende una deficiencia en la glutamina sintetasa (GS), dihidrofolato reductasa (DHFR), hipoxantina-guanina fosforibosiltransferasa (HPRT), o una de sus combinaciones.
- 20 8. La línea celular de una cualquiera de las reivindicaciones precedentes, en la que la línea celular es una línea celular de ovario de hámster chino (CHO).
9. La línea celular de la reivindicación 8, en la que la línea celular CHO es una línea celular GS -/-.
- 25 10. La línea celular de una cualquiera de las reivindicaciones precedentes, en la que la línea celular expresa al menos una glicoproteína que comprende uno o más residuos de manosa terminal.
11. La línea celular de una cualquiera de las reivindicaciones precedentes, en la que la línea celular tiene una tasa de crecimiento comparable al de una línea celular que no tiene deficiencia de Mgat1.
12. La línea celular de una cualquiera de las reivindicaciones precedentes, en la que la línea celular produce un nivel de proteína comparable al de una línea celular que no tiene deficiencia de Mgat1.
- 30 13. Un procedimiento de producción de una célula de mamífero deficiente en mannosil(alfa-1,3-)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa 1 (Mgat1) que comprende una supresión de 2 a 300 bp en la secuencia cromosómica que codifica Mgat1 que comprende transfectar una célula huésped con una endonucleasa de direccionamiento que se dirige a una secuencia específica en una secuencia cromosómica que codifica Mgat1.
- 35 14. El procedimiento de la reivindicación 13, en el que la endonucleasa de direccionamiento es una nucleasa de dedos de cinc.
15. El procedimiento de cualquiera de las reivindicaciones 13 o 14 que además comprende incubar la célula transfectada con aglutinina-I (RCA-I) de *Ricinus communis*.
16. El procedimiento de una cualquiera de las reivindicaciones 13 a 15 que además comprende:
 - a) introducir un ácido nucleico que codifica una proteína recombinante en la célula; y
 - 40 b) expresar la proteína recombinante.
17. Un procedimiento de producción de una proteína recombinante que tiene uno o más residuos de manosa terminal, el procedimiento comprende:
 - a) introducir un ácido nucleico que codifica la proteína recombinante en la línea celular de cualquiera de las reivindicaciones 1-12; y
 - 45 b) expresar la proteína recombinante.

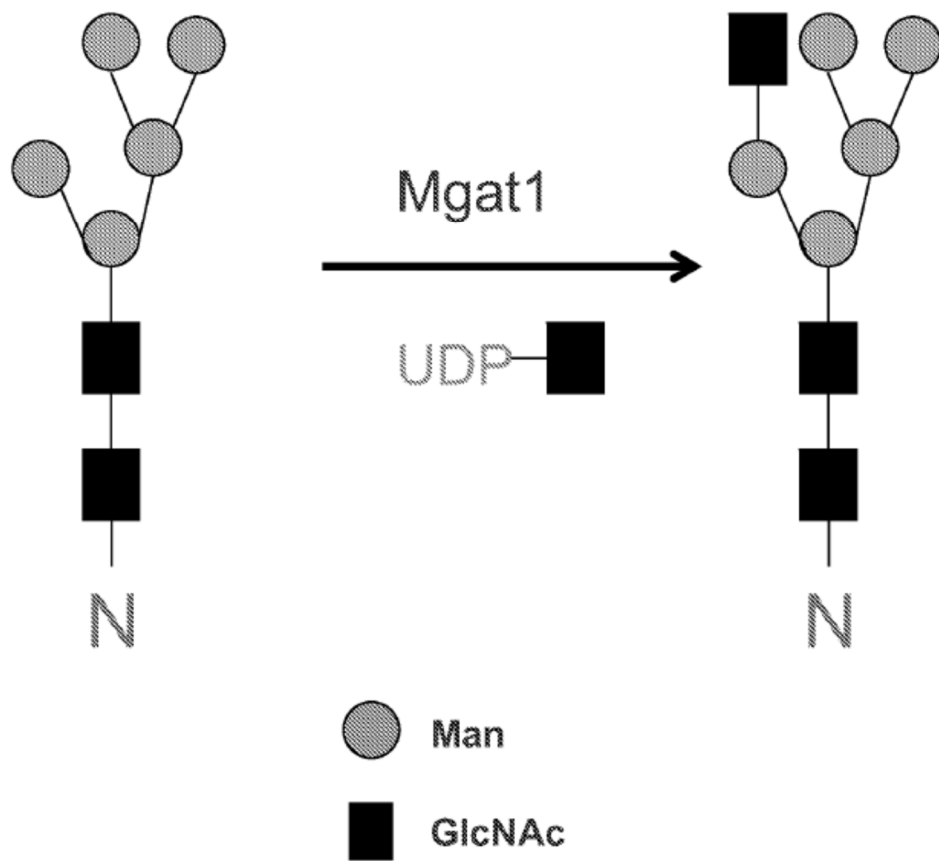


FIG. 1

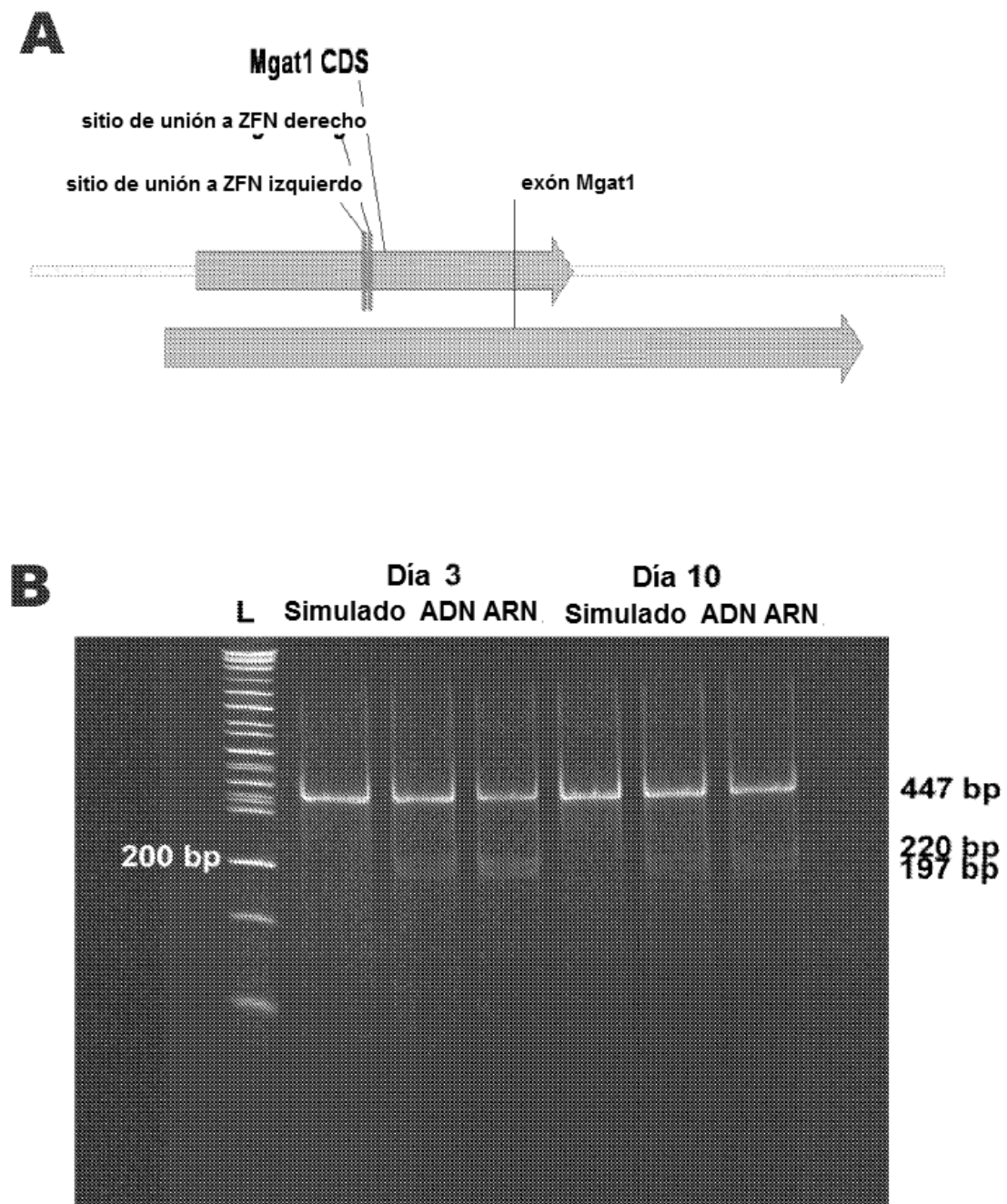


FIG. 2

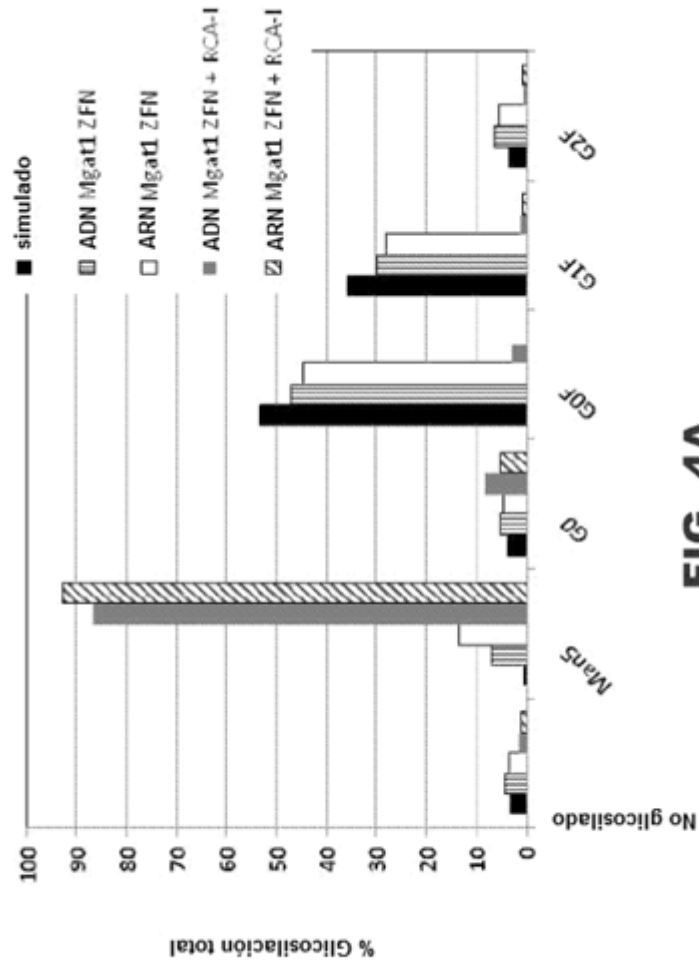


FIG. 4A

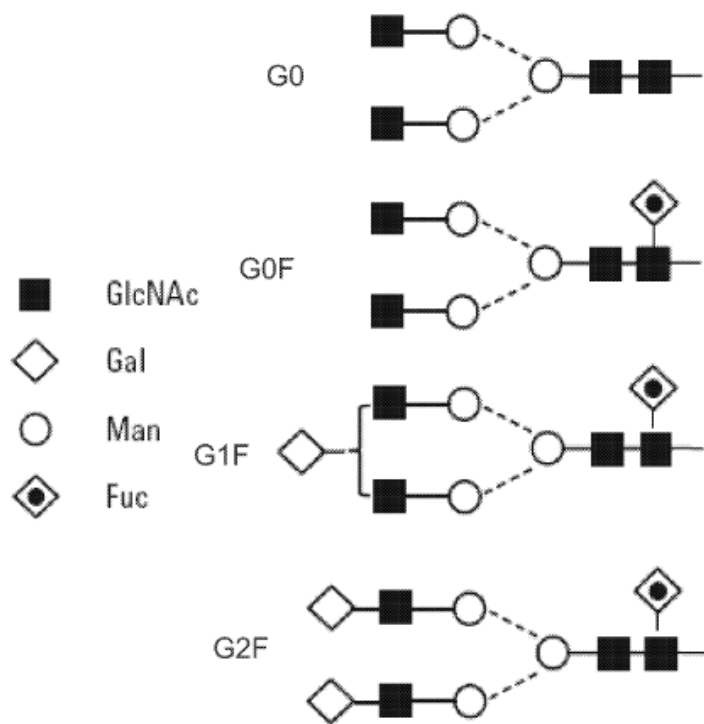


FIG. 4B

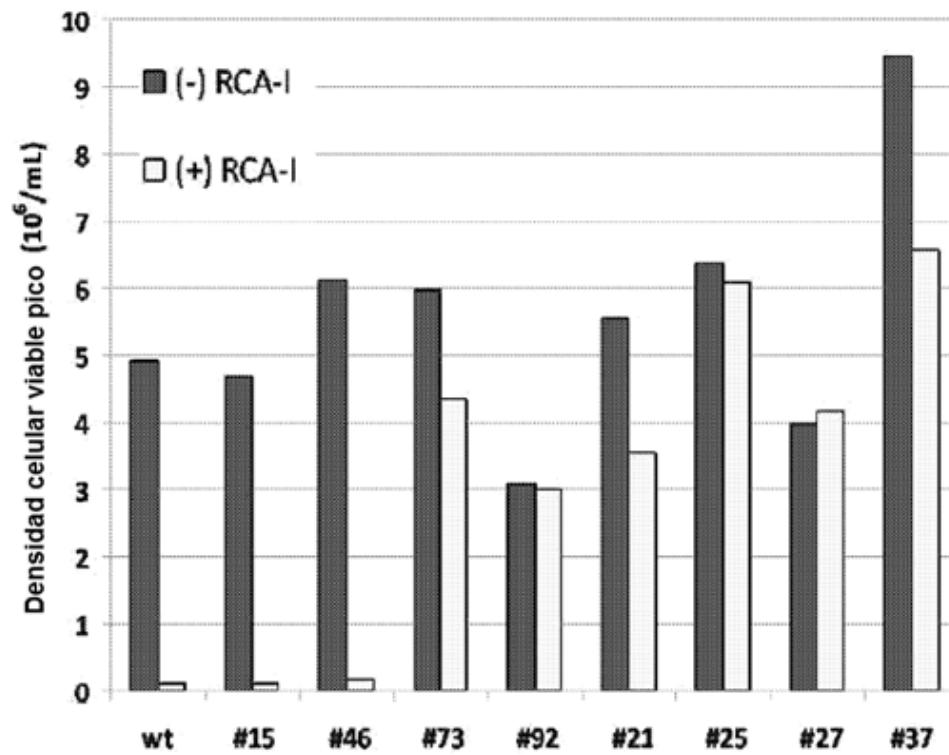


FIG. 6

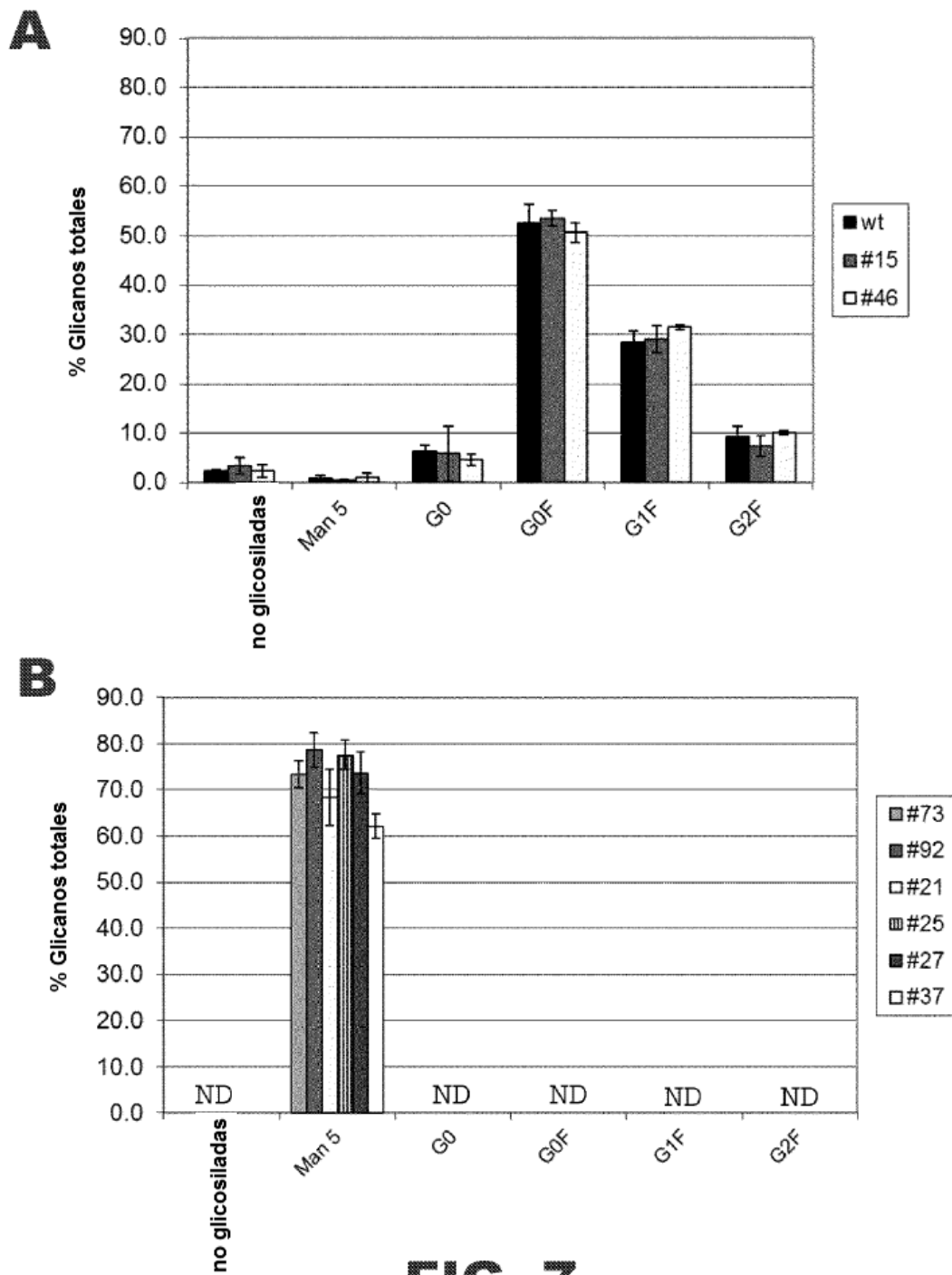


FIG. 7

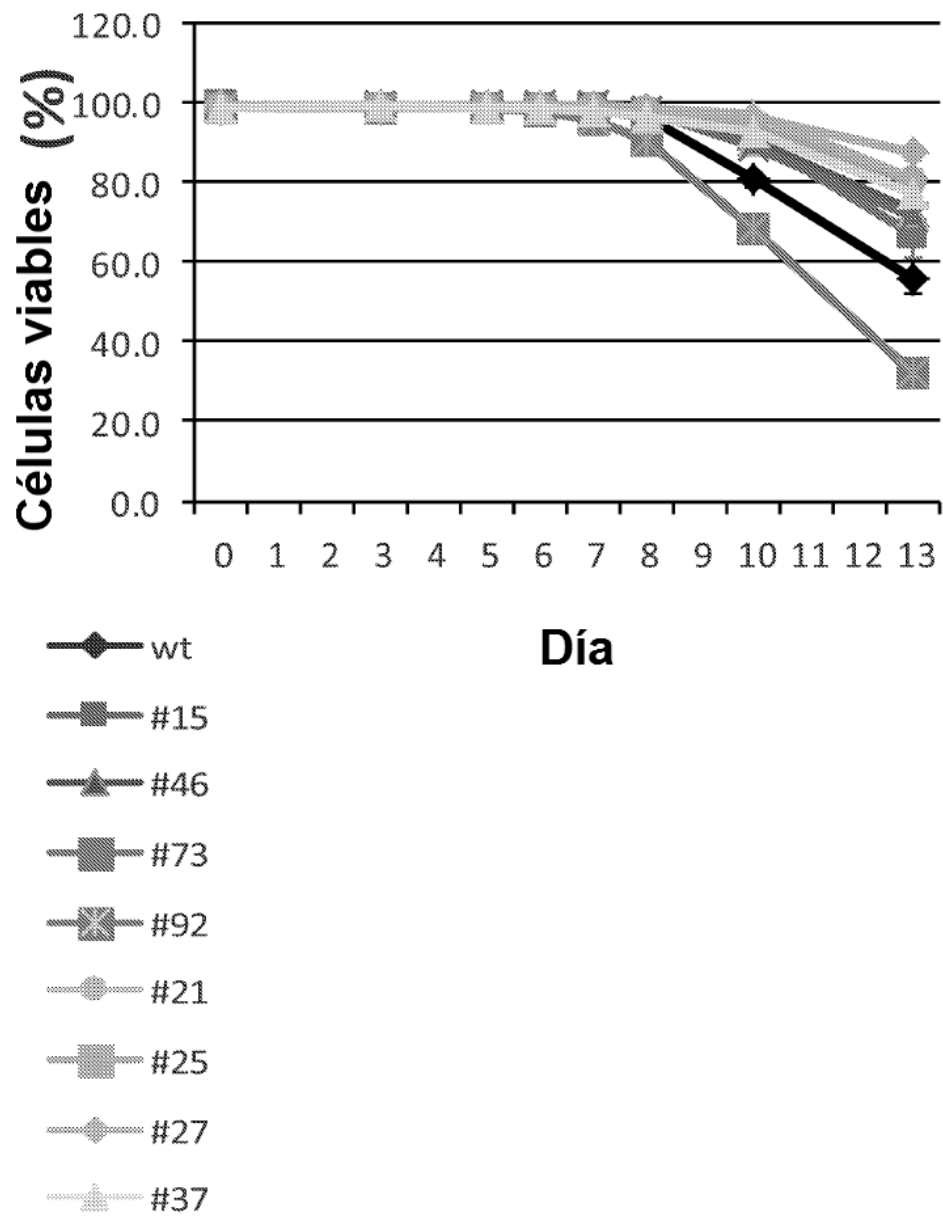
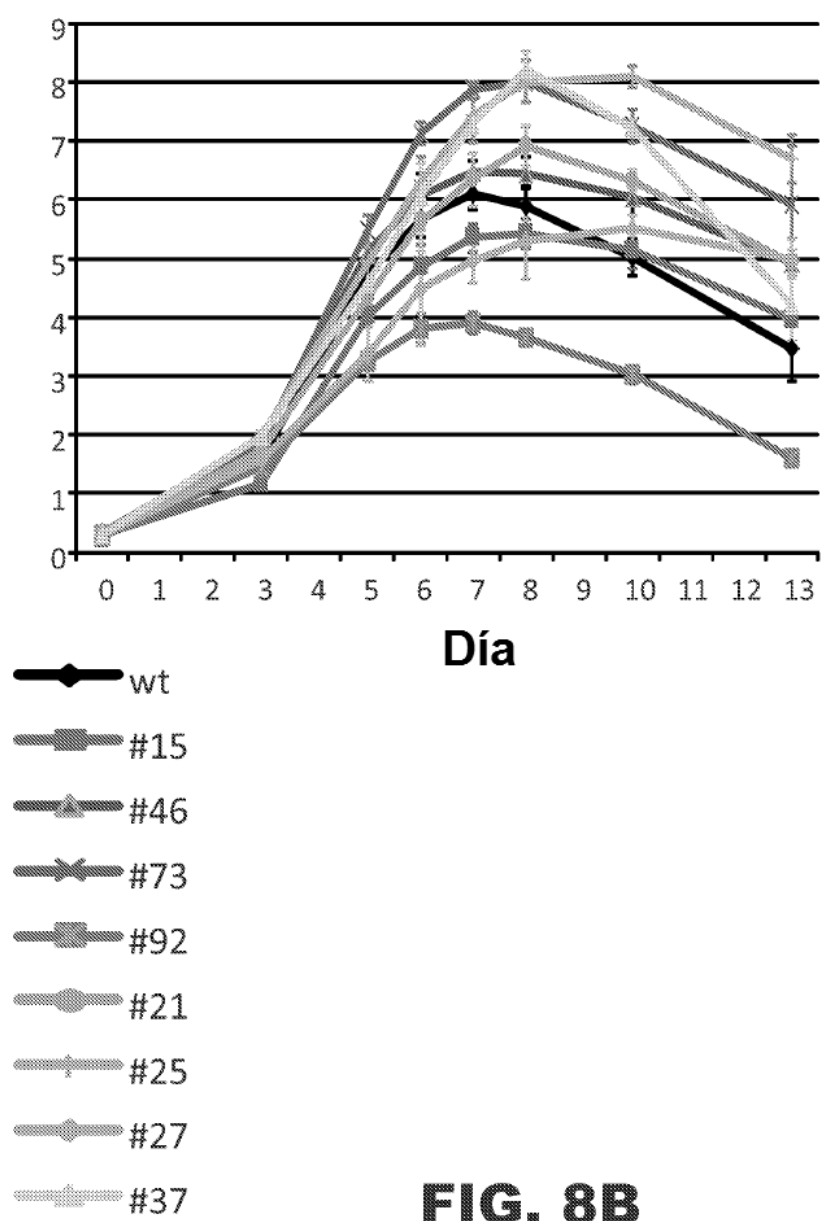


FIG. 8A



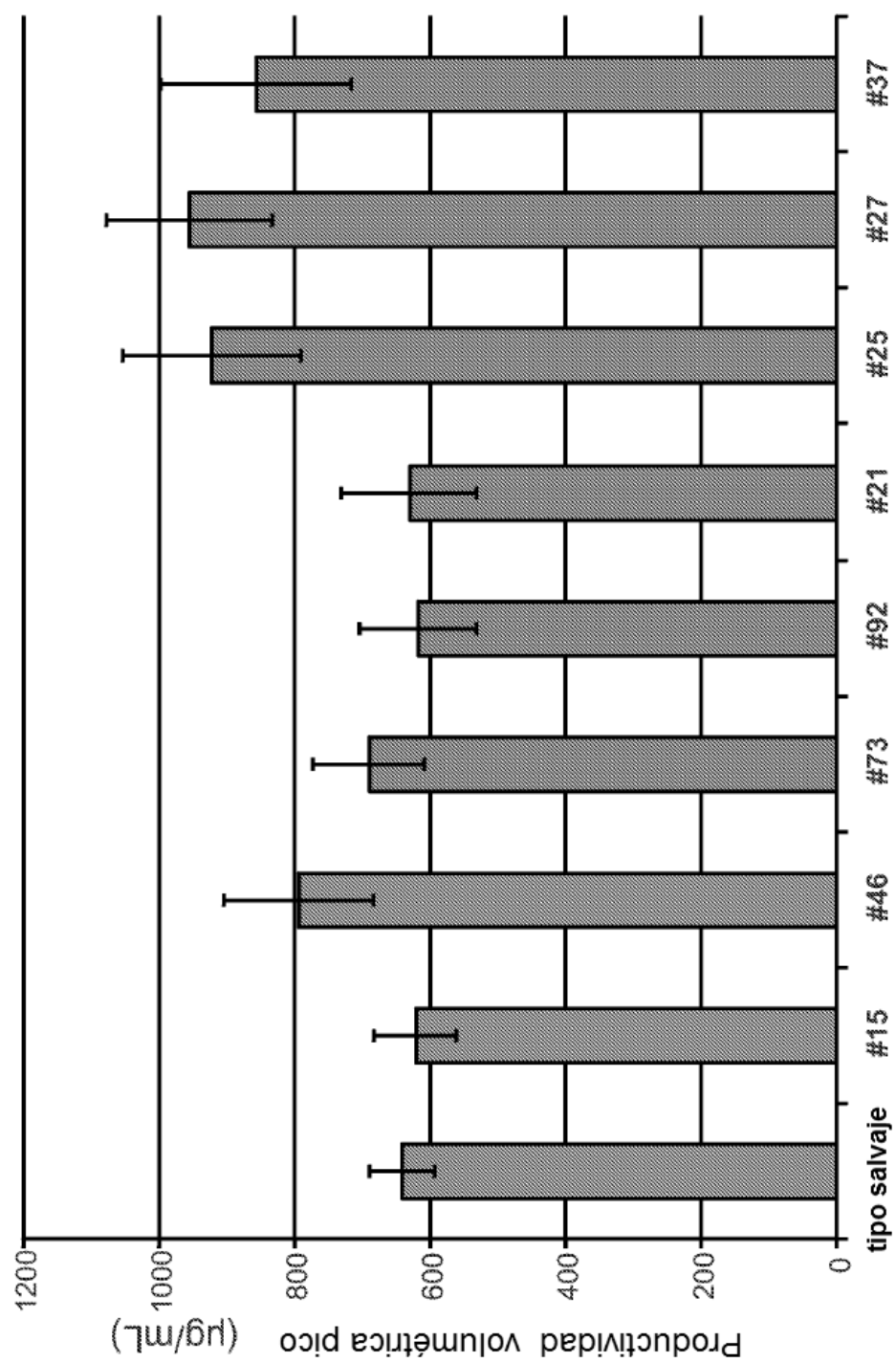


FIG. 9