

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 797 050**

51 Int. Cl.:

C12N 15/06 (2006.01)

C12N 5/00 (2006.01)

C12N 9/10 (2006.01)

C12N 5/071 (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.03.2015 PCT/US2015/018474**

87 Fecha y número de publicación internacional: **11.09.2015 WO15134488**

96 Fecha de presentación y número de la solicitud europea: **03.03.2015 E 15757751 (1)**

97 Fecha y número de publicación de la concesión europea: **22.04.2020 EP 3114222**

54 Título: **Células virales resistentes y sus usos**

30 Prioridad:

04.03.2014 US 201461947860 P

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

01.12.2020

73 Titular/es:

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

72 Inventor/es:

**LIN, NAN;
MASCARENHAS, JOAQUINA;
CHANG, AUDREY;
ONIONS, DAVID;
GEORGE, HENRY y
KAYSER, KEVIN**

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 797 050 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

Células virales resistentes y sus usos

Campo de la invención

5 La presente invención se refiere a líneas celulares de mamíferos manipuladas para tener resistencia viral y al uso de dichas líneas celulares para reducción o prevención de la contaminación viral de sistemas de producción biológica.

Antecedentes

10 El uso de proteínas terapéuticas producidas de manera recombinante para el tratamiento de numerosas enfermedades o afecciones tal como el cáncer y enfermedades autoinmunes continúa en aumento. Sin embargo, la producción a gran escala de tales productos terapéuticos proteicos aún continúa siendo un desafío. Por ejemplo, un proceso de fabricación comercial debe administrar un rendimiento confiablemente alto con procesos corriente abajo que producen un producto extremadamente puro permitiendo cantidades muy pequeñas de contaminantes, preferentemente ninguna cantidad.

15 El uso de medios libres de componentes animales ha reducido significativamente la incidencia de contaminación viral adventicia. Además, la implementación de procedimientos tal como ultrafiltración, procesamiento de intervalo breve a alta temperatura, y/o radiación UVC de materiales a granel ha reducido adicionalmente la incidencia de contaminación. No obstante, continúa presente el riesgo de contaminación viral. Un incidente de contaminación sería catastrófico para el fabricante en términos de pérdida de producto, retiro temporario del mercado, y costos extensivos de descontaminación. De este modo, existe una necesidad de lograr líneas celulares de mamíferos que tengan mayor resistencia a la infección viral.

20 Taber *et al.*, 1976 (Cell, Vol. 8, 529-533) y Jan *et al.*, 1999 (Journal of Virology, Vol. 73, No. 6, pp 4919-4924) describen líneas celulares de ovario de hámster chino (CHO) que han sido seleccionadas por su resistencia a la infección viral.

25 Los virus pueden ingresar a las células a través de unión específica a receptores complementarios sobre las superficies de las células. Para muchos virus, estos receptores de la superficie celular comprenden estructuras de glicano ligadas a las proteínas (o a los lípidos). Los glicanos que terminan en ácido siálico o sus derivados sirven como receptores para muchos virus (Matrosovich *et al.*, 2013, Top Curr Chem, DOI: 10.1007/128_2013_466). Nam *et al.*, 2006 (Journal of Biological Chemistry, Vol. 281, No. 35, pp. 25670-25677) describen diversos glicanos que muestran especificidad para virus diminutos del ratón (MVM). Yang *et al.* 2012 (Blood, Vol. 120, No. 5, pp.1015-1026) enseñan que la expresión conjunta de *St3Gal4* y *St3Gal6* contribuye a la sialilación que es responsable de la síntesis y la función de los ligandos de selectina.

Breve descripción de las figuras

35 La **FIG. 1** presenta una secuencia temporal de infección viral por MVM según lo ensayado por análisis de Southern blot. Trazada con la cantidad relativa de ADN viral MVM detectada a las 24, 48, 72 y 96 horas de células de tipo salvaje (CHOZN GS-/-), *Slc35A1* KO, clon F07 COSMC KO, COSMC KO, clon G03 COSMC KO, y clon H05 COSMC KO. El clon H05 tiene una eliminación dentro del marco de 12 pb (es decir, una mutación sin desviación de marco).

La **FIG. 2** muestra una secuencia temporal de infección viral por MVM según lo detectado por ensayos en placa. Es trazada la pfu/ml detectada a las 24, 48, 72 y 96 horas para células de tipo salvaje (CHOZN), *Slc35A1* KO, clon F07 COSMC KO, COSMC KO, clon G03COSMC KO, y clon H05 COSMC KO.

40 La **FIG. 3** ilustra el efecto de infección por MMV sobre el crecimiento celular. Es trazado el número de células viables (células/ml) durante el transcurso de 120 horas para células de tipo salvaje (2E3) (**A**) y clon F07 COSMC KO (**B**) en ausencia (líneas continuas) y presencia (líneas discontinuas) de virus MVM a MOI de 1 u 8.

45 La **FIG. 4** presenta una secuencia temporal de virus asociado con las células después de infección por MMV. Es trazada la copia del genoma del virus MVM (vgc) por célula (en base a la suposición de que cada célula infectada puede producir 2×10^4 vgc) según lo detectado a través de qPCR durante el transcurso de 120 horas para células de tipo salvaje (2E3) (líneas continuas) y clon F07 clon F07 COSMC KO, (líneas discontinuas) en presencia de virus MVM a MOI de 1 (**A**) u 8 (**B**).

La **FIG. 5** muestra la secuencia temporal de replicación de MMV en las líneas celulares indicadas. Es trazada la vgc/muestra a las 0 y 21 horas después de la infección con virus MVM a MOI de 0,3 (**A**) o 0,03 (**B**).

50 La **FIG. 6** presenta la secuencia temporal de replicación de Reovirus-3 en las líneas celulares indicadas. Es trazada la vgc/muestra relativa a células de tipo salvaje (2E3) a las 0 y 24 horas después de la infección con virus Reo-3.

La **FIG. 7** presenta ensayos de crecimiento para células crecidas en ausencia (UN) o presencia (IN) de virus

MVM. El panel **A** presenta la densidad de células viables (VCD) para tipo salvaje (GS), *Slc35A1* KO, clon F07 clon F07 COSMC KO, y clon G03 COSMC KO durante 10 días. El panel **B** presenta la VCD para tipo salvaje (GS), clones que producen IgG 71H1 y 71C3 derivados de clon F07 COSMC KO, durante 8 días.

5 La **FIG. 8** presenta ensayos de crecimiento para células crecidas en ausencia (UN) o presencia (IN) de virus MVM. Es trazada la VCD para tipo salvaje (GS), clon 7D10*St3Gal4* KO, clon 1B08 *St3Gal4* KO, y clon 1B10 *St3Gal4* KO, durante 9 días.

Sumario

10 Entre los diversos aspectos de la presente invención, y como es definido en las reivindicaciones, se encuentra la provisión de líneas celulares de mamíferos que están manipuladas para impedir o prevenir la entrada viral y/o la propagación viral en las células. En algunas realizaciones, las líneas celulares de mamíferos son modificadas genéticamente de modo tal que la entrada y/o propagación de virus diminutos de ratón (MVM) es reducida o eliminada en comparación con las líneas celulares progenitoras. En diversas realizaciones, las líneas celulares de mamíferos desveladas en la presente memoria comprenden al menos una secuencia cromosómica modificada que comprende una secuencia cromosómica inactivada que codifica la alfa-2,3-sialiltransferasa 4 de beta-galactósido St3 (*St3Gal4*). En realizaciones adicionales, la secuencia cromosómica modificada es inactivada de modo tal que la línea celular no produce, o produce niveles reducidos, del producto proteico codificado, *St3Gal4*.

20 En una realización, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la chaperona enzimática Núcleo 1 (COSMC). En otra realización, todas las copias de secuencias cromosómicas que codifican COSMC son inactivadas y la línea celular no produce COSMC. En una realización alternativa, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la familia de portador soluto 35 (transportador de ácido siálico-CMP), miembro A1 (*Slc35A1*). En otra realización, todas las copias de secuencias cromosómicas que codifican *Slc35A1* son inactivadas y la línea celular no produce *Slc35A1*. Incluso en otra realización, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la enzima de estiramiento núcleo 1 (*C1GalT1*). En una realización adicional, todas las copias de secuencias cromosómicas que codifican *C1GalT1* son inactivadas y la línea celular no produce *C1GalT1*. En una realización adicional, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la alfa-2,3-sialiltransferasa 1 de beta-galactósido St3 (*St3Gal1*). En otra realización, todas las copias de secuencias cromosómicas que codifican *St3Gal1* son inactivadas y la línea celular no produce *St3Gal1*. En una realización alternativa, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la alfa-2,3-sialiltransferasa 2 de beta-galactósido St3 (*St3Gal2*). En otra realización, todas las copias de secuencias cromosómicas que codifican *St3Gal2* son inactivadas y la línea celular no produce *St3Gal2*. Aún en otra realización, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la alfa-2,3-sialiltransferasa 3 de beta-galactósido St3 (*St3Gal3*). En una realización adicional, todas las copias de secuencias cromosómicas que codifican *St3Gal3* son inactivadas y la línea celular no produce *St3Gal3*. Incluso en otra realización alternativa, y de acuerdo con la presente invención, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la alfa-2,3-sialiltransferasa 4 de beta-galactósido St3 (*St3Gal4*). En otra realización, todas las copias de secuencias cromosómicas que codifican *St3Gal4* son inactivadas y la línea celular no produce *St3Gal4*. En una realización alternativa, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la alfa-2,3-sialiltransferasa 5 de beta-galactósido St3 (*St3Gal5*). En una realización adicional, todas las copias de secuencias cromosómicas que codifican *St3Gal5* son inactivadas y la línea celular no produce *St3Gal5*. Aún en otra realización, la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la alfa-2,3-sialiltransferasa 6 de beta-galactósido St3 (*St3Gal6*). En una realización adicional, todas las copias de secuencias cromosómicas que codifican *St3Gal6* son inactivadas y la línea celular no produce *St3Gal6*. En una realización alternativa, la línea celular comprende secuencias cromosómicas inactivadas que codifican cualquiera de dos de *St3Gal1*, *St3Gal2*, *St3Gal3*, *St3Gal4*, *St3Gal5* y *St3Gal6*. Aún en otra realización alternativa, la línea celular comprende secuencias cromosómicas inactivadas que codifican cualquiera de tres de *St3Gal1*, *St3Gal2*, *St3Gal3*, *St3Gal4*, *St3Gal5* y *St3Gal6*. En realizaciones adicionales, las líneas celulares que comprenden secuencias cromosómicas inactivadas que codifican COSMC, *Slc35A1*, *C1GalT1*, *St3Gal1*, *St3Gal2*, *St3Gal3*, *St3Gal4*, *St3Gal5*, y/o *St3Gal6* comprenden, adicionalmente, secuencias cromosómicas inactivadas que codifican (alfa-1,3-manosil)-glicoproteína beta-1,2-N-acetilglucosamiltransferasa 1 (*Mgat1*), manosil (alfa-1,6-)-glicoproteína beta-1,2-N-acetilglucosamiltransferasa 2 (*Mgat2*), manosil (alfa-1,4-glicoproteína beta-1,4-N-acetilglucosamiltransferasa 3 (*Mgat3*), (alfa-1,3-manosil)-glicoproteína beta-1,4-N-acetilglucosamiltransferasa 4 (*Mgat4*), y/o manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetilglucosamiltransferasa 5 (*Mgat5*). Las líneas celulares de mamíferos desveladas en la presente memoria que comprenden secuencia cromosómica modificada tienen mayor resistencia a la entrada viral y/o la propagación viral en comparación con las líneas celulares progenitoras no modificadas. Incluso en otras realizaciones, las líneas celulares que comprenden las secuencias cromosómicas inactivadas que codifican *St3Gal1*, *St3Gal2*, *St3Gal3*, *St3Gal4*, *St3Gal5*, y/o *St3Gal6* están manipuladas adicionalmente para sobreexpresar al menos una sialiltransferasa responsable de producir ligaduras 2,6 (por ej., *St6* beta-galactosamida alfa-2,6-sialiltransferasa 1 (*St6Gal1*), *St6* beta-galactosamida alfa-2,6-sialiltransferasa 2 (*St6Gal2*), *St6* (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 1 (*St6GalNac1*), *St6* (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 2 (*St6GalNac2*), *St6* (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 3 (*St6GalNac3*), *St6* (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 4 (*St6GalNac4*), *St6* (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 5 (*St6GalNac5*), y/o *St6* (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 6 (*St6GalNac6*).

5 En algunas realizaciones, las líneas celulares de mamíferos son líneas celulares no humanas. En otras realizaciones de acuerdo con la invención, las líneas celulares de mamíferos son líneas celulares de ovario de hámster chino (CHO). En realizaciones específicas, la línea celular es una línea celular de CHO que comprende secuencias cromosómicas inactivadas que codifican COSMC, Slc35A1, C1GalT1, St3Gal1, St3Gal2, St3Gal3, St3Gal4, St3Gal5, St3Gal6, o combinaciones de las anteriores.

10 En ciertas realizaciones, las líneas celulares de mamíferos desveladas en la presente memoria exhiben resistencia a la entrada y/o propagación de virus seleccionados de parvovirus, reovirus, rotavirus, virus de la gripe, virus adenoasociados, calcivirus, virus de parainfluenza, virus de rubeola, coronavirus, norovirus, virus de encefalomiocarditis, poliomavirus, o sus combinaciones. En algunas realizaciones, el parvovirus es virus diminuto de ratón (MVM), parvovirus tipo 1 de ratón, parvovirus tipo 2 de ratón, parvovirus tipo 3 de ratón, parvovirus 1 porcino, parvovirus 1 bovino, parvovirus B19 humano, parvovirus 4 humano, parvovirus 5 humano, o sus combinaciones. En realizaciones adicionales, el reovirus es reovirus 3 de mamífero, ortoreovirus de mamífero, ortoreovirus aviar, o sus combinaciones.

15 En diversas realizaciones, las líneas celulares de mamíferos desveladas en la presente memoria son preparadas por modificación de al menos una secuencia cromosómica usando una técnica de modificación del genoma mediada por endonucleasa diana. La endonucleasa diana puede ser nucleasas de dedos de cinc, endonucleasas CRISPR/Cas, nucleasa efectora similar al activador de transcripción (TALE), meganucleasa, una endonucleasa específica del sitio, o un agente inductor de ruptura de doble hebra de ADN diana artificial. En realizaciones específicas, la endonucleasa diana es un par de nucleasas de dedos de cinc.

20 En algunas realizaciones, las líneas celulares de mamíferos desveladas en la presente memoria además comprenden al menos un ácido nucleico que codifica una proteína recombinante seleccionada de un anticuerpo, un fragmento de anticuerpo, una vacuna, un factor de crecimiento, una citoquina, una hormona, un factor de coagulación u otra proteína terapéutica.

25 Otro aspecto de la presente divulgación abarca un procedimiento de reducción o prevención de contaminación viral de un producto de proteína recombinante, el procedimiento comprende obtener una línea celular de mamífero con resistencia viral como es desvelado en la presente memoria y expresar el producto de proteína recombinante en la línea celular.

30 Un aspecto adicional de la presente divulgación proporciona un procedimiento de reducción de riesgo de contaminación viral de un sistema de producción biológica, en el que el procedimiento comprende proporcionar una línea celular de mamífero con resistencia viral como es desvelado en la presente memoria para usar en el sistema de producción biológica.

Incluso otro aspecto de la presente divulgación abarca una composición que comprende una línea celular de mamífero con resistencia viral como es desvelado en la presente memoria y al menos un virus, en la que la línea celular exhibe resistencia a la infección producida por el al menos un virus.

35 Otro aspecto e iteraciones de la divulgación son descritos en mayor detalle a continuación.

Descripción detallada

40 La presente divulgación proporciona líneas celulares de mamíferos manipuladas para impedir o prevenir la entrada viral y/o la propagación viral en las líneas celulares. Las líneas celulares manipuladas que tienen resistencia viral pueden ser modificadas genéticamente para contener secuencias cromosómicas modificadas (por ej., inactivadas). También son proporcionados procedimientos de uso de las líneas celulares desveladas en la presente memoria para la producción de proteínas recombinantes, en los que los productos de proteína recombinante están esencialmente desprovistos de contaminación viral. El uso de las líneas celulares que son resistentes a la infección viral, por lo tanto, reduce o elimina el riesgo de contaminación viral de los sistemas de producción biológica y los productos de proteína resultantes.

(I) Líneas celulares con resistencia viral

50 Un aspecto de la presente invención abarca líneas celulares de mamíferos, como es definido por las reivindicaciones, que están manipuladas para tener resistencia viral. Indicado de otro modo, las líneas celulares desveladas en la presente memoria tienen mayor resistencia a la infección por al menos un virus en comparación con las líneas celulares progenitoras, no modificadas. Más específicamente, la entrada del virus MVM y/o la propagación del virus MVM es reducida o eliminada en las líneas celulares manipuladas descritas en esta memoria en comparación con las líneas celulares progenitoras no modificadas. En algunas realizaciones y de acuerdo con la invención, las líneas celulares de mamíferos están modificadas genéticamente y contienen al menos una secuencia cromosómica modificada. En general, la secuencia modificada comprende una mutación. En algunas realizaciones, la secuencia cromosómica modificada es inactivada (o bloqueada) de modo tal que el producto de proteína codificada no es producido por la línea celular.

55 En general, la resistencia (o susceptibilidad) a infección viral puede ser determinada por comparación de la respuesta

de las líneas celulares de mamífero manipuladas con la exposición a uno o varios virus con la respuesta de las células progenitoras (no manipuladas) no modificadas a la misma provocación viral. La infección viral de la línea celular y/o la propagación viral en la línea celular puede ser analizada por una variedad de técnicas. Los ejemplos no limitativos de técnicas adecuadas incluyen procedimientos de detección de ácido nucleico (por ej., ensayo Southern blot con ácido nucleico para detectar la presencia de ácidos nucleicos virales específicos, PCR o RT-PCR para detectar ácidos nucleicos virales, procedimientos de secuencia y similares), técnicas a base de anticuerpos (por ej., técnicas de inmunotransferencia Western usando anticuerpos de proteínas antivirales, procedimientos ELISA, y demás), bioensayos, (por ej., ensayos de placas, ensayos de efecto citopático y similares), y técnicas microscópicas (por ej., microscopio electrónico para detectar partículas virales, etc.). En algunas realizaciones, la infección y/o propagación del virus dentro de las líneas celulares de mamífero manipuladas pueden ser reducidas en al menos aproximadamente 10%, al menos aproximadamente 20%, al menos aproximadamente 40%, al menos aproximadamente 60%, al menos aproximadamente 80%, al menos aproximadamente 90%, al menos aproximadamente 99%, o más de aproximadamente 99% con relación a las células progenitoras no modificadas. En realizaciones específicas, las líneas celulares manipuladas de mamíferos son resistentes a la infección viral, es decir, el virus es incapaz de entrar y/o de propagarse en las líneas celulares manipuladas de mamíferos.

Las líneas celulares de mamíferos desveladas en la presente memoria pueden ser manipuladas en una variedad de modos diferentes para impartir resistencia viral. En algunas realizaciones, las líneas celulares pueden ser modificadas de modo tal que la entrada viral a través de los receptores de superficie sea reducida o eliminada. En otras realizaciones, las líneas celulares pueden ser manipuladas para expresar moléculas que inhiban o bloqueen las proteínas virales específicas involucradas en la replicación y/o infectividad. Incluso en otras realizaciones, las líneas celulares pueden ser manipuladas para sobreexpresar proteínas antivirales celulares específicas. En algunas realizaciones, la manipulación puede ser genética, en la que las secuencias genómicas o cromosómicas son modificadas. Dicho de otro modo, las líneas celulares son modificadas genéticamente. En otras realizaciones la manipulación puede ser epigenética o extracromosómica.

(a) Mecanismo de resistencia viral

(i) Perturbación de receptores de superficie celular

En ciertas realizaciones, las líneas celulares de mamíferos son manipuladas para contener receptores de la superficie celular alterados. Los virus pueden ingresar a las células a través de la unión específica a receptores complementarios sobre las superficies de las células. Para numerosos virus, estos receptores de la superficie celular comprenden estructuras de glicano ligadas a las proteínas (o a los lípidos). Los glicanos que terminan en ácido siálico o sus derivados sirven como receptores para numerosos virus (Matrosovich *et al.*, 2013, Top Curr Chem, DOI: 10.1007/128_2013_466). De este modo, las glicoproteínas que comprenden glicanos O-ligados o N-ligados con residuos terminales de ácido siálico pueden funcionar como los receptores de la superficie celular para numerosos virus. Ácido siálico se refiere a derivados del ácido neuramínico, por ejemplo, ácido N-acetilneuramínico (Neu5Ac o NANA) y ácido N-glicolilneuramínico (Neu5Gc o NGNA).

En algunas realizaciones, las líneas celulares son manipuladas para comprender glicoproteínas que carecen de residuos de ácido siálico terminal. Los residuos de ácido siálico terminal pueden ser eliminados por eliminación (es decir, noqueo) o alteración de enzimas y/o proteínas involucradas en la síntesis de las cadenas de glicano. Los blancos adecuados incluyen enzimas o proteínas involucradas en la síntesis de glicanos O-ligados, enzimas o proteínas involucradas en la síntesis de glicanos N-ligados, y/o enzimas o proteínas involucradas en la síntesis o el transporte del ácido siálico.

En ciertas realizaciones, la línea celular puede ser deficiente en al menos una de las enzimas o proteínas específicas mencionadas con anterioridad (o cualquier combinación de las dianas mencionadas con anterioridad). Como se usa en la presente memoria, "deficiente" se refiere a niveles reducidos o no detectables de enzimas o proteínas específicas, o actividad reducida o no detectable de las enzimas o proteínas específicas. La cantidad o la actividad de la enzima o la proteína diana puede ser reducida o eliminada por modificación de al menos una secuencia cromosómica que codifica la proteína o enzima diana. Por ejemplo, la secuencia cromosómica puede ser modificada para contener una eliminación de al menos un nucleótido, una inserción de al menos un nucleótido, una sustitución de al menos un nucleótido, o una de sus combinaciones. De esta manera, las eliminaciones, inserciones y/o sustituciones pueden cambiar el marco de lectura de la secuencia cromosómica de modo tal que no sea producido producto de proteína (es decir, la secuencia cromosómica es inactivada). Alternativamente, las eliminaciones, inserciones y/o sustituciones en la secuencia cromosómica modificada pueden conducir a la producción de un producto de proteína alterada. La modificación de secuencias cromosómicas de interés puede ser lograda usando técnicas de edición genómica específica mediada por endonucleasa, que se detallan a continuación en la sección (III)(a). En casos en los que una secuencia cromosómica que codifica la enzima o proteína diana es inactivada, la línea celular manipulada produce niveles reducidos de enzima o proteína diana. En otros casos en los que todas las copias de las secuencias que codifican la enzima o proteína diana son inactivadas, la línea celular manipulada no produce enzima o proteína diana (es decir, la línea celular es un knock out o KO). Incluso en otras realizaciones, el nivel de enzima o proteína diana puede ser reducido o eliminado usando mecanismos mediados por interferencia de ARN, que se describen a continuación en la sección (III)(b).

En algunas realizaciones, el nivel de enzima o proteína diana puede ser reducido en al menos aproximadamente 5%, de aproximadamente 5 a 10%, de aproximadamente 10 a 20%, de aproximadamente 20 a 30%, de aproximadamente 30 a 40%, de aproximadamente 40 a 50%, de aproximadamente 50 a 60%, de aproximadamente 60 a 70%, de aproximadamente 70 a 80%, de aproximadamente 80 a 90%, o de aproximadamente 90 a aproximadamente 100%.

5 En otras realizaciones, el nivel de enzima o proteína diana puede ser reducido hasta niveles no detectables usando técnicas estándares en el campo (por ej., ensayos de inmunotransferencia Western, ensayos enzimáticos ELISA, y similares).

En algunas realizaciones, la línea celular puede ser deficiente en enzimas o proteínas involucradas en la glicosilación O-ligada. Por ejemplo, la línea celular puede ser deficiente en enzima de estiramiento núcleo 1 (también llamada núcleo 1 sintasa, glicoproteína-N-acetilgalactosamina 3-beta-galactosiltransferasa, 1, o C1GalT1), núcleo 1 chaperona enzimática (también llamada chaperona específica de C1GalT1 o COSMC), o ambas. COSMC facilita el plegado, la estabilidad y la actividad de C1GalT1, que cataliza la transferencia de un residuo de glucosa en el residuo de N-acetilgalactosamina (GalNAc) que está O-ligado a serina o treonina de una proteína. En realizaciones específicas, la línea celular es deficiente en C1GalT1 o COSMC, o ambas. La deficiencia puede deberse a las secuencias cromosómicas inactivadas que codifican C1GalT1 y/o COSMC de modo tal que son logrados niveles reducidos o nada de proteína C1GalT1 y/o COSMC. En algunos casos, al menos una secuencia cromosómica que codifica C1GalT1 y/o COSMC es inactivada. En otros casos, todas las copias de las secuencias cromosómicas que codifican C1GalT1 y/o COSMC son inactivadas, de modo tal que la línea celular es desprovista de proteína C1GalT1 y/o proteína COSMC.

En otras realizaciones, la línea celular puede ser deficiente en al menos una sialiltransferasa (ST). La sialiltransferasa puede ser una sialiltransferasa que agrega ácido siálico a la galactosa en una conformación de ligación alfa-2,3, una sialiltransferasa que agrega ácido siálico a galactosa o N-acetilgalactosamina en la conformación de ligación alfa-2,6, o una sialiltransferasa que agrega ácido siálico a otras unidades de ácido siálico en una conformación de ligación alfa-2,8. Los ejemplos no limitativos de sialiltransferasas adecuadas incluyen con St3 beta-galactósido alfa-2,3-sialiltransferasa 1 (St3Gal1), St3 beta-galactósido alfa-2,3-sialiltransferasa 2 (St3Gal2), St3 beta-galactósido alfa-2,3-sialiltransferasa 3 (St3Gal3), St3 beta-galactósido alfa-2,3-sialiltransferasa 4 (St3Gal4), St3 beta-galactósido alfa-2,3-sialiltransferasa 5 (St3Gal5), St3 beta-galactósido alfa-2,3-sialiltransferasa 6 (St3Gal6), St6 beta-galactosamida alfa-2,6-sialiltransferasa 1 (St6Gal1), St6 beta-galactosamida alfa-2,6-sialiltransferasa 2 (St6Gal2), St6 (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 1 (St6GalNac1), St6 (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 2 (St6GalNac2), St6 (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 3 (St6GalNac3), St6 (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 4 (St6GalNac4), St6 (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 5 (St6GalNac5), St6 (alfa-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida 6 (St6GalNac6), St8 alfa-N-acetil-neuraminida alfa-2,8-sialiltransferasa 1 (St8Sia1), St8 alfa-N-acetil-neuraminida alfa-2,8-sialiltransferasa 2 (St8Sia2), St8 alfa-N-acetil-neuraminida alfa-2,8-sialiltransferasa 3 (St8Sia3), St8 alfa-N-acetil-neuraminida alfa-2,8-sialiltransferasa 4 (St8Sia4), St8 alfa-N-acetil-neuraminida alfa-2,8-sialiltransferasa 5 (St8Sia5) o St8 alfa-N-acetil-neuraminida alfa-2,8-sialiltransferasa 6 (St8Sia6).

En realizaciones específicas, la línea celular puede ser deficiente en al menos una alfa-2,3-sialiltransferasa (por ej., St3Gal1, St3Gal2, St3Gal3, St3Gal4, St3Gal5, y/o St3Gal6). La deficiencia puede ser debida a secuencias cromosómicas inactivadas que codifican la al menos una alfa-2,3-sialiltransferasa de modo tal que son producidos niveles reducidos o nada de la al menos una alfa-2,3-sialiltransferasa. En algunos casos, al menos una secuencia cromosómica que codifica al menos una alfa-2,3-sialiltransferasa puede ser inactivada. En otros casos, todas las copias de las secuencias cromosómicas que codifican la al menos una alfa-2,3-sialiltransferasa puede ser inactivada, de modo tal que la línea celular esté desprovista de al menos una alfa-2,3-sialiltransferasa. Dicho de otro modo, las secuencias cromosómicas que codifican St3Gal 1, 2, 3, 4, 5 y/o 6 pueden ser bloqueadas.

En algunas instancias en las que la línea celular comprende al menos una secuencia cromosómica inactivada que codifica la al menos una alfa-2,3-sialiltransferasa (por ej., St3Gal1, St3Gal2, St3Gal3, St3Gal4, St3Gal5 y/o St3Gal6), la línea celular puede ser manipulada adicionalmente para sobreexpresar al menos una sialiltransferasa responsable de producir las ligaduras 2,6 (por ej., St6Gal 1, St6Gal 2, St6GalNac1, St6GalNac2, St6GalNac3, St6GalNac4, St6GalNac5 y/o St6GalNac6). De esta manera, dichas líneas celulares pueden comprender glicoproteínas que tienen números reducidos de (o ningún) residuos de ácido siálico 2,3-ligado terminal y mayores números de residuos de ácido siálico 2,6-ligado terminal.

En realizaciones adicionales, la línea celular puede ser deficiente en al menos una enzima o proteína involucrada en la síntesis o el transporte de ácido siálico. Los ejemplos de enzimas o proteínas involucradas en la síntesis o en el transporte de ácido siálico incluyen, sin límite, glucosamina (UDP-N-acetil)-2-epimerasa/N-acetilmanosamina quinasa (GNE), sintasa del ácido N-acetilneuramínico (NANS), fosfatasa del ácido N-acetilneuramínico (NANP), citidina monofosfato de sintetasa del ácido N-acetilneuramínico (CMAS), e citidina monofosfato de hidroxilasa del ácido N-acetilneuramínico (CMAH), familia 35 de portador soluto (transportador de ácido siálico-CMP), miembro A1 (Slc35A1). En ciertas realizaciones, la línea celular puede ser deficiente en Slc35A1, que transporta CMP-ácido siálico en Golgi. En algunos casos, al menos una secuencia cromosómica que codifica Slc35A1 puede ser inactivada. En otros casos, todas las copias de las secuencias cromosómicas que codifican Slc35A1 pueden ser inactivadas, de modo tal que la célula es desprovista de proteína Slc35A1.

En realizaciones adicionales, la línea celular puede ser deficiente en al menos una enzima o proteína involucrada en la N-glicosilación. En algunas instancias, la enzima o la proteína involucrada en la N-glicosilación puede ser una N-acetilglucosilaminil-transferasa, que añade un residuo de GlcNAc a un residuo de manosa beta-ligado de un glicano N-ligado. Ejemplos incluyen (alfa-1,3-manosil)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa 1 (Mgat-1), manosil (alfa-1,6-)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa 2 (Mgat-2), manosil (alfa-1,4-)-glicoproteína beta-1,4-N-acetilglucosaminiltransferasa 3 (Mgat-3), (alfa-1,3-manosil)-glicoproteína beta-1,4-N-acetilglucosaminiltransferasa 4 (Mgat-4), y manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetilglucosaminiltransferasa 5 (Mgat-5). En otras instancias, la enzima o la proteína involucrada en la N-glicosilación puede ser galactosiltransferasa, que añade un residuo de galactosa en un residuo 1,4-ligadura beta a un residuo GlcNAc de un glicano N-ligado. La galactosiltransferasa, puede ser UDP-Gal:BetaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 1 (B4GalT1), UDP-Gal:BetaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 2 (B4GalT2), UDP-Gal:BetaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 3 (B4GalT3), UDP-Gal:BetaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 4 (B4GalT4), UDP-Gal:BetaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 5 (B4GalT5), UDP-Gal:BetaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 6 (B4GalT6), o UDP-Gal:BetaGlcNAc beta 1,4-galactosiltransferasa, polipéptido 7 (B4GalT7). En ciertos casos, la línea celular puede ser deficiente en al menos una N-acetilglucoilaminiltransferasa y/o galactosiltransferasa. En algunas iteraciones, al menos una secuencia cromosómica que codifica al menos una N-acetilglucoilaminiltransferasa y/o galactosiltransferasa puede ser inactivada. En otros casos, todas las copias de las secuencias cromosómicas que codifican la al menos una N-acetilglucoilaminiltransferasa y/o galactosiltransferasa puede ser inactivada, de modo tal que la línea celular esté desprovista de al menos una N-acetilglucoilaminiltransferasa y/o galactosiltransferasa.

(ii) Interferencia con proteínas virales

En otras realizaciones, las líneas celulares de mamíferos pueden ser manipuladas para expresar moléculas que inhiban o bloqueen la replicación y/o infectividad viral. Por ejemplo, las líneas celulares pueden manipularse para expresar de manera estable al menos un agente de interferencia de ARN (iARN) contra las proteínas virales específicas que están involucradas en la replicación y/o infectividad. Ejemplos no limitativos de proteínas virales adecuadas incluyen proteínas no estructurales tales como NS1 o NS2 y proteínas de la cápside tal como VP1 o VP2. Los agentes de iARN son unidos a las transcritos diana y previenen la expresión de las proteínas por mediación de la escisión de la escisión por transcripción o la traducción por interrupción de la transcripción.

En algunas realizaciones, el agente de iARN puede ser un ARN de interferencia breve (ARNsi). En general, el ARNsi comprende una molécula de ARN de doble hebra que oscila de aproximadamente 15 a aproximadamente 29 nucleótidos de longitud, o más generalmente de aproximadamente 19 a aproximadamente 23 nucleótidos de longitud. En realizaciones específicas, el ARNsi puede ser de aproximadamente 21 nucleótidos de longitud. El ARNsi puede comprender opcionalmente, además, una o dos salientes de hebra simple, por ej., una saliente 3' en uno o ambos extremos. El ARNsi puede ser formado a partir de dos moléculas de ARN hibridadas en conjunto o, alternativamente, puede ser generado a partir de un ARN de horquilla corta (ARNsh) (véase a continuación). En algunas realizaciones, las dos hebras de ARNsi pueden ser completamente complementarias, de modo tal que no existan faltas de coincidencia o protuberancias en el dúplex formado entre las dos secuencias. En otras realizaciones, las dos hebras de ARNsi pueden ser sustancialmente complementarias, de modo tal que existan una o más faltas de coincidencia y/o protuberancias en el dúplex formado entre las dos secuencias. En ciertas realizaciones, uno o ambos de los extremos 5' del ARNsi pueden tener un grupo fosfato, mientras que en otras realizaciones uno o ambos de los extremos 5' pueden carecer de un grupo fosfato.

Una hebra del ARNsi, a la que se hace referencia como "hebra antisentido" o "hebra guía", incluye una porción hibridada con la transcrito diana. En algunas realizaciones, la hebra antisentido del ARNsi puede ser completamente complementar a una región del transcrito diana, es decir, se hibrida a la transcrito diana, sin una falta de coincidencia individual o protuberancia a través de la longitud del ARNsi. En otras realizaciones, la hebra antisentido puede ser sustancialmente complementaria a la región específica, es decir, una o más faltas de equivalencia y/o protuberancias pueden existir en el dúplex formado por la hebra antisentido y la transcrito diana. Normalmente, los ARNsi están dirigidos a secuencias exónicas del transcrito diana. Las personas con experiencia en la técnica están familiarizadas con programas, algoritmos y/o servicios comerciales que diseñan ARNsi para transcritos diana.

En otras realizaciones, el agente de iARN puede ser un ARN de horquilla breve (ARNsh). En general, un ARNsh y una molécula de ARN que comprende al menos dos porciones complementarias hibridadas o son capaces de hibridación para formar una estructura de hebra doble lo suficientemente larga para mediar la interferencia de ARN (como se describe con anterioridad), y al menos una porción de hebra simple que forma un bucle que conecta las regiones de ARNsh que forma el dúplex. La estructura también puede ser llamada estructura de vástago-bucle, siendo el vástago la porción del dúplex. En algunas realizaciones, la porción de dúplex de la estructura puede ser completamente complementaria, de modo tal que no existan faltas de coincidencia o protuberancias en la región del dúplex del ARNsh. En otras realizaciones, la porción del dúplex de la estructura puede ser sustancialmente complementaria, de modo tal que existan una o más faltas de coincidencia y/o protuberancias en la porción del dúplex del ARNsh. El bucle de la estructura puede ser de aproximadamente 1 a aproximadamente 20 nucleótidos de longitud, específicamente de aproximadamente 6 a aproximadamente 9 nucleótidos de longitud. El bucle puede estar ubicado indistintamente en el extremo 5' o 3' de la región que es complementaria a la transcrito diana (es decir, la porción antisentido del ARNsh).

- El ARNsh puede además comprender una saliente en el extremo 5' o 3'. La saliente opcional puede ser de aproximadamente 1 a aproximadamente 20 nucleótidos de longitud, o más específicamente de aproximadamente 2 a aproximadamente 15 nucleótidos de longitud. En algunas realizaciones, la saliente puede comprender uno o más residuos U, por ej., entre aproximadamente 1 y aproximadamente 5 residuos U. En algunas realizaciones, el extremo 5' del ARNsh puede tener un grupo fosfato. En general, los ARNsh son procesados en ARNsi por la maquinaria de ARNi celular conservada. De este modo, los ARNsh son precursores de los ARNsi que son similarmente capaces de inhibir la expresión de una transcripción blanco que es complementaria de la porción de ARNsh (es decir, la porción antisentido del ARNsh). Las personas expertas en la técnica están familiarizadas con los recursos disponibles para el diseño y la síntesis de ARNsh. Un ejemplo es ARNsh MISSION® (Sigma-Aldrich).
- El ARNsi o ARNsh puede ser expresado *in vivo* desde un constructo de expresión de ARNi. Los constructos adecuados incluyen vectores plásmidos, fagémidos, cósmidos, mini-cromosomas/cromosomas artificiales, transposones y vectores virales (por ej., vectores lentivirales, vectores virales adenoasociados, etc.). En una realización, la construcción de expresión de ARNi puede ser un vector de plásmido (por ej., pUC, pBR322, pET, pBluescript, y las variantes de éstos). La construcción de expresión de ARNi puede comprender dos secuencias de control del promotor, en las que cada una es una secuencia de codificación adecuada operativamente ligada de modo tal que pueden transcribirse hebras de ARNsi complementarias. Las dos secuencias de control del promotor pueden estar en la misma orientación o en orientaciones opuestas. En otra realización, el vector de expresión de ARNi puede contener una secuencia de control del promotor que conduce la transcripción de una molécula de ARN simple que comprende dos regiones complementarias, de modo tal que la transcripción forma un ARNsh. En general, la(s) secuencia(s) de control del promotor serán promotores de polimerasa III (Pol III) de ARN tales como los promotores U6 o H1. En otras realizaciones, pueden ser usadas las secuencias de control del promotor de polimerasa II (Pol II) del ARN (algunos ejemplos se presentan a continuación). Las construcción de expresión de ARNi pueden contener elementos de secuencia adicionales, tales como las secuencias de terminación de transcripción, secuencias de marcador seleccionable, etc., La construcción de expresión del ARNi puede ser introducida en la línea celular de interés usando procedimientos estándar. La construcción de expresión de ARNi puede estar cromosómicamente integrada en la línea celular para la expresión estable. Alternativamente, la construcción de expresión de ARNi puede ser extracromosómica (por ej., episomal) en la línea celular para la expresión estable.
- Incluso en otras realizaciones, las líneas celulares pueden ser manipuladas para expresar de manera estable formas negativas dominantes de proteínas virales involucradas en la replicación y/o infectividad. Una forma negativa dominante de una proteína es alterada o mutada tal que compite con, o inhibe, la proteína de tipo salvaje. Los ejemplos no limitativos de proteínas adecuadas incluyen proteínas virales no estructurales tal como NS1 o NS2 y proteínas virales de la cápside tales como VP1 o VP2. En realizaciones específicas, la línea celular puede ser manipulada para expresar una forma negativa dominante de una o más proteínas NS1.
- Una proteína negativa dominante puede tener una eliminación, una inserción, y/o una sustitución relativa a la proteína de tipo salvaje (Lagna *et al.*, 1998, Curr. Topics Dev. Biol, 36:75-98). La eliminación, inserción y/o sustitución puede estar en la ubicación de extremo terminal N-, extremo terminal C o interna de la proteína. Los medios para generar proteínas mutantes (a través de mutagénesis dirigida al sitio, mutagénesis basada en PCR, mutagénesis aleatoria, etc.) son conocidos en la técnica, como lo son los medios para identificar aquellas que tienen efectos negativos dominantes. Las líneas celulares pueden ser transfectadas con constructos de expresión (véase más arriba) que comprenden la secuencia que codifica las proteínas negativas dominantes, en las que la secuencia de codificación está operativamente ligada a una secuencia de control del promotor Pol II para expresión. La secuencia de control del promotor puede ser constitutiva, regulada o específica del tejido.
- Las secuencias de control constitutivas del promotor adecuadas incluyen, pero sin limitación, promotor temprano inmediato de citomegalovirus (CMV), promotor del virus de simio (SV40), promotor tardío principal de adenovirus, promotor del virus de sarcoma de Rous (RSV), promotor del virus de tumor mamario de ratón (MMTV), promotor de quinasa de fosfoglicerato (PGK), promotor de estiramiento del factor (ED1)-alfa, promotores de ubiquitina, promotores de actina, promotores de tubulina, promotores de inmunoglobulina, fragmentos de los anteriores, o combinaciones de cualquiera de los anteriores. Los ejemplos de secuencias de control del promotor regulado adecuado incluyen sin límite aquellos regulados por shock térmico, metales, esteroides, antibióticos o alcohol. Ejemplos no limitativos de promotores específicos del tejido incluyen promotor B29, promotor CD14, promotor CD43, promotor CD45, promotor CD68, promotor de desmina, promotor de elastasa-1, promotor de endoglina, promotor de fibronectina, promotor Flt-1, promotor GFAP, promotor GPIIb, promotor ICAM-2, promotor INF- β , promotor Mb, promotor Nphsl, promotor OG-2, promotor SP-B, promotor SYN1, y promotor WASP. La secuencia promotora puede ser tipo salvaje o puede ser modificada para una expresión más eficiente o eficaz.
- La construcción de expresión puede comprender secuencias adicionales de control de expresión (por ej., secuencias potenciadoras, secuencias de Kozak, secuencias de poliadenilación, secuencias de terminación de transcripción, etc.), secuencias de marcador seleccionable (por ej., genes de resistencia a los antibióticos), orígenes de replicación, y similares. Información adicional puede ser hallada 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.

(iii) Sobreexpresión de proteínas celulares involucradas en respuestas antivirales

En realizaciones alternativas, las líneas celulares de mamíferos pueden ser manipuladas para sobreexpresar proteínas celulares involucradas en las respuestas antivirales de la célula huésped. Los ejemplos no limitativos de proteínas involucradas en repuestas antivirales incluyen quinasa R de la proteína con ARN activado de hebra doble (PKR, también denominada quinasa 2 del factor eucariótico iniciador de la traducción 2-alfa o Eif2ak2), quinasa 2 de la proteína que interactúa con el receptor (RIPK2), interferones (por ej., Tipo I y Tipo II), interleucinas (por ej., IL-1 e IL-6), factor alfa de necrosis tumoral, factor 1 regulador del interferón, STATs, p53, factor 3 de activación de transcripción, NF-κB, factor 2 de iniciación eucariótica (eIF2), inhibidores de proteínas de apoptosis (IAPs), y proteína antiviral de dedos de cinc (ZAP). En realizaciones específicas, las líneas celulares pueden ser manipuladas para PKR.

La sobreexpresión puede ser lograda mediante la introducción de una o más copias exógenas de una secuencia de ácido nucleico que codifica la proteína de interés. La secuencia que codifica la proteína celular de interés generalmente está ligada operativamente a una secuencia de control del promotor Pol II (véase con anterioridad). Múltiples copias de la secuencia de codificación pueden estar relacionadas en tándem y colocadas bajo control de una sola secuencia de control del promotor. Las secuencias que codifican la proteína de interés pueden ser introducidas en las líneas celulares como parte de un constructo de expresión (véase con anterioridad). De esta manera, la construcción de expresión puede ser insertada en una ubicación cromosómica o, alternativamente, la construcción de expresión puede ser extracromosómica (por ej., episomal) para la expresión estable.

La sobreexpresión también puede ser lograda por modificación de la secuencia de control del promotor de la secuencia cromosómica endógena que codifica la proteína de interés. Por ejemplo, la secuencia de control del promotor endógeno puede ser modificada por inserción de al menos una secuencia de control del promotor "fuerte" exógeno (es decir, que tiene alta afinidad por la polimerasa de ARN y/o factores asociados) (cuyos ejemplos son presentados anteriormente). Alternativamente, la secuencia de control del promotor endógeno puede ser modificada para imitar las secuencias de control del promotor "fuerte". Las secuencias cromosómicas endógenas pueden ser modificadas usando técnicas de modificación específica de genoma mediadas por endonucleasa, detalladas a continuación en la sección (III).

(b) Tipo de células

Las líneas celulares resistentes a los virus desveladas en la presente memoria son líneas celulares de mamíferos. En algunas realizaciones, las líneas celulares que tienen resistencia a la infección viral pueden ser derivadas de células de ovario de hámster chino (CHO); células de mieloma NSO de ratón; células de riñón de hámster bebé (BHK); células de fibroblasto 3T3 de embrión de ratón (NIH3T3); células de linfoma A20 de ratón B; células de melanoma B16 de ratón; células de mioblasto C2C12 de ratón; células de mieloma SP2/0 de ratón; células mesenquimales C3H- 10T1/2 de embrión de ratón; células de carcinoma CT26 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 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 B linfoma de rata; células B35 de neuroblastoma de rata; células de hepatoma de rata (HTC); células BRL 3A de hígado de rata búfalo; células renales de canino (MDCK); células mamarias de canino (CMT); células D17 de osteosarcoma de rata; células DH82 de monocito/macrófago de rata; células de fibroblasto transformadas SV-40 (COS7) de riñón de mono; células CVI-76 de riñón de mono; células de riñón de mono verde africano (VERO-76); células de riñón de embrión humano (HEK293, HEK293T); células de carcinoma cervical humano (HELA); células de pulmón humano (W138); células de hígado humano (Hep G2); células de osteosarcoma U2-OS humano, células A549 humanas, células A-431 humanas, o células K562 humanas. Una lista extensiva de líneas celulares de mamíferos pueden ser halladas en el catálogo de la American Type Culture Collection (ATCC, Manassas, VA). En otras realizaciones, las líneas celulares con resistencia viral son líneas celulares de mamíferos no humanos. En realizaciones adicionales, las líneas celulares con resistencia viral son líneas celulares de mamíferos, no ratón, no humanos. En ciertas realizaciones, las líneas celulares con resistencia viral son líneas celulares de CHO. Numerosas líneas celulares de CHO están disponibles de ATCC. Las líneas celulares de CHO incluyen, pero sin limitación, células CHO-K1 y sus derivados.

En diversas realizaciones, las líneas celulares pueden ser deficientes en glutamina sintasa (GS), dihidrofolato reductasa (DHFR), hipoxantina-guanina fosforibosiltransferasa (HPRT), o una de sus combinaciones. Por ejemplo, las secuencias cromosómicas que codifican GS, DHFR, y/o HPRT pueden ser inactivadas. En realizaciones específicas, todas las secuencias cromosómicas que codifican GS son inactivadas en las líneas celulares.

(c) Virus

Las líneas celulares manipuladas que tienen resistencia viral pueden ser resistentes a una variedad de virus de mamíferos. El virus puede ser un virus de ADN o un virus de ARN, y el virus puede ser con envoltura o sin envoltura ("desnudo"). Ejemplos no limitativos de virus que pueden unirse e ingresar a las células de mamíferos a través de los receptores de ácido siálico incluyen parvovirus, reovirus, rotavirus, virus de la gripe, virus adenoasociados, calcivirus, virus de parainfluenza, virus de rubeola, coronavirus, norovirus, virus de encefalomiocarditis y poliomavirus. En algunas realizaciones, las líneas celulares de mamíferos manipuladas son resistentes a la infección por al menos un parvovirus. Ejemplos no limitativos de parvovirus adecuados incluyen el virus diminuto del ratón (MVM) (que también se conoce como virus diminuto del ratón (MMV) o protoparvovirus 1 del roedor), tipo-1 de parvovirus del ratón (MPV-

1), tipo-2 de parvovirus del ratón (MPV-2), tipo-3 de parvovirus del ratón (MPV-3), parvovirus 1 porcino, parvovirus 1 bovino, y parvovirus humano (por ej., parvovirus B19 humano, parvovirus 4 humano, parvovirus 5 humano, etc.). En realizaciones particulares y de acuerdo con la presente invención, el parvovirus puede ser MVM. En otras realizaciones, el virus puede ser un reovirus, tal como reovirus-3 de mamífero, ortoreovirus de mamífero, ortoreovirus aviar, y similares).

En algunas realizaciones, las líneas celulares de mamíferos manipuladas que tienen resistencia a la infección viral también pueden tener resistencia a la infección por organismos del orden de *Mollicutes*. En particular, las líneas celulares divulgadas en esta memoria pueden ser resistentes a la infección por el género *mycoplasma* o *spiroplasma*.

(d) Ácido nucleico opcional que codifica la proteína recombinante

En algunas realizaciones, las líneas celulares de mamíferos que tienen resistencia a la infección viral pueden comprender, además, al menos un ácido nucleico que codifica una proteína recombinante. En general, la proteína recombinante es heteróloga, lo que significa que la proteína no es nativa para la célula. La proteína recombinante puede ser, sin límite, una proteína terapéutica seleccionada de 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 interleucina, una hormona, un factor de coagulación, un componente de la sangre, 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 que anteceden y/o sus fragmentos o variantes funcionales.

En algunas realizaciones, el ácido nucleico que codifica la proteína recombinante puede ser relacionado con la secuencia que codifica la hipoxantina-guanina fosforibosiltransferasa (HPRT), dihidrofolato reductasa (DHFR), y/o glutamina sintasa (GS), de modo tal que HPRT, DHFR, y/o GS pueden ser usados como marcador seleccionable susceptible de amplificarse. El ácido nucleico que codifica la proteína recombinante también puede ser relacionado con la secuencia que codifica al menos un gen de resistencia a los antibióticos y/o la secuencia que codifica las proteínas marcadoras tales como las proteínas fluorescentes. En algunas realizaciones, el ácido nucleico que codifica la proteína recombinante puede ser parte de un constructo de expresión. Como es detallado en otra parte, los constructos o los vectores de expresión pueden comprender secuencias adicionales de control de expresión (por ej., secuencias potenciadoras, secuencias de Kozak, secuencias de poliadenilación, secuencias de terminación de transcripción, etc.), secuencias de marcador seleccionable, orígenes de replicación, y similares. Información adicional puede ser hallada en Ausubel *et al.* 2003, *supra*, y Sambrook & Russell, 2001, *supra*.

En algunas realizaciones, el ácido nucleico que codifica la proteína recombinante puede estar ubicado extracromosómicamente.

Es decir, el ácido nucleico que codifica la proteína recombinante puede ser expresado transitoriamente a partir de un plásmido, un cósmido, un cromosoma artificial, un minicromosoma u otra construcción extracromosómica. En otras realizaciones, el ácido nucleico que codifica la proteína recombinante puede ser integrado cromosómicamente en el genoma de la célula. La integración puede ser aleatoria o específica. De esta manera, la proteína recombinante puede ser expresada de manera estable. En algunas iteraciones de esta realización, la secuencia de ácido nucleico que codifica la proteína recombinante puede estar ligada operativamente a una secuencia de control de expresión heteróloga adecuada (es decir, un promotor). En otras iteraciones, la secuencia de ácido nucleico que codifica la proteína recombinante puede ser colocada bajo control de una secuencia de control de expresión endógena. La secuencia de ácido nucleico que codifica la proteína recombinante puede ser integrada en el genoma de la línea celular usando recombinación homóloga, edición del genoma mediada por endonucleasa específica, vectores virales, transposones, plásmidos, y otros medios conocidos. Directrices adicionales pueden ser halladas en Ausubel *et al.* 2003, *supra* y Sambrook & Russell, 2001, *supra*.

(e) Composiciones

En ciertas realizaciones, las líneas celulares de mamíferos, como son definidas en las reivindicaciones, y manipuladas para exhibir resistencia viral, pueden ser parte de una composición, que también comprende al menos un virus. La composición, por lo tanto, comprende la línea celular específica (que comprende opcionalmente un ácido nucleico que codifica una proteína recombinante) desvelada en la presente memoria y un virus, en la que la entrada y/o la propagación de al menos un virus es reducida o eliminada en la línea celular manipulada de mamífero. De este modo, las células en la composición son capaces de propagarse, pero el virus en la composición es incapaz de propagarse porque su entrada y/o replicación dentro de las células es reducida o eliminada. La composición puede comprender, además, al menos un medio de crecimiento celular para soportar el crecimiento de la línea celular manipulada de mamífero. En algunas instancias, el medio de crecimiento celular es un medio libre de componentes animales.

(f) Realizaciones de ejemplo

En realizaciones específicas y de acuerdo con la invención, las líneas celulares de mamíferos que tienen resistencia viral son líneas celulares de CHO. Las líneas celulares de CHO con resistencia viral pueden ser resistentes a la

infección por virus diminuto del ratón (MVM) (que también es denominado virus del ratón diminuto (MMV) o protoparvovirus 1 del roedor) y/o reovirus 3 del mamífero. Específicamente, las líneas celulares de CHO modificadas genéticamente tienen mayor resistencia a la infección por MVM o reovirus-3 en comparación con las líneas celulares de CHO progenitoras no modificadas. En algunas realizaciones, la línea celular progenitora no modificada es una línea celular de CHO (GS -/-).

Las líneas celulares de CHO con resistencia viral comprenden al menos una secuencia cromosómica inactivada que codifica St3Gal4. En algunas realizaciones, todas las copias de secuencias cromosómicas que codifican St3Gal4 son inactivadas o bloqueadas en las líneas celulares de CHO. En otras realizaciones, las células CHO que comprenden secuencias cromosómicas inactivadas que codifican St3Gal4 comprenden adicionalmente secuencias cromosómicas inactivadas que codifican Mgat1. En realizaciones adicionales, las células CHO que comprenden secuencias cromosómicas inactivadas que codifican St3Gal4 están manipuladas además para sobreexpresar St6Gal 1, St6Gal 2, St6GalNac1, St6GalNac2, St6GalNac3, ST6GalNac4, St6GalNac5, y/o St6GalNac6.

(II) Procedimientos para reducción o prevención de la contaminación viral

Otro aspecto de la presente divulgación abarca procedimientos para reducción o prevención de la contaminación viral de un producto de proteína recombinante, o reducción del riesgo de contaminación viral de un sistema de producción biológica. En general, los procedimientos comprenden proporcionar líneas celulares de mamíferos manipuladas en las que la entrada y/o la propagación de al menos un virus es reducida o eliminada, y usar dichas líneas celulares para la producción de la proteína recombinante. Las líneas celulares de mamíferos manipuladas son detalladas anteriormente en las secciones (I) y en las reivindicaciones. Las líneas celulares de mamíferos manipuladas exhiben mayor resistencia a la infección viral en comparación con las líneas progenitoras modificadas. Las líneas celulares de mamíferos manipuladas se exhiben resistentes a los virus descritos en la sección (I)(c). Las proteínas recombinantes adecuadas son descritas en la sección (I)(d). Los medios para producir o fabricar proteínas recombinantes son conocidos en el campo (véase por ej., "Biopharmaceutical Production Technology", Subramanian (ed), 2012, Wiley-VCH; ISBN: 978-3-527-33029-4). En las realizaciones específicas, las líneas celulares de mamíferos manipuladas son genéticamente modificadas para comprender al menos una secuencia cromosómica modificada de modo tal que la línea celular sea resistente a la infección viral.

En general, el uso de líneas celulares de mamíferos manipuladas desveladas en la presente memoria reduce la capacidad de que los virus se repliquen en un fermentador u otro recipiente de bioproducción de modo tal que el nivel de virus replicable esté a un nivel ínfimo o, idealmente, a un nivel que no sea detectable por las mejores prácticas estándar de la industria. Los procedimientos adecuados incluyen procedimientos de detección de ácido nucleico (por ej., Southern blot para detectar la presencia de ácidos nucleicos virales específicos, PCR o RT-PCR para detectar ácidos nucleicos virales, procedimientos de secuencia y similares), técnicas a base de anticuerpos (por ej., inmunotransferencia Western usando anticuerpos de proteínas antivirales, procedimientos ELISA, y demás), y técnicas microscópicas (por ej., ensayos de efecto citopático, microscopio electrónico para detectar partículas virales, etc.).

(III) Procedimientos para preparar líneas celulares con resistencia viral

Incluso otro aspecto de la presente divulgación proporciona procedimientos de preparación de células con resistencia viral que tienen receptores de superficie celular alterados, como es detallado con anterioridad en la sección (I)(a)(i). Las secuencias cromosómicas que codifican enzimas o proteínas involucradas en la síntesis de cadenas de glicano pueden ser reducidas o eliminadas usando una variedad de técnicas para generar las líneas celulares con resistencia viral. En algunas realizaciones, las líneas celulares con resistencia viral pueden ser preparadas por un proceso de modificación del genoma mediado por endonucleasa específica. En otras realizaciones, las líneas celulares con resistencia viral pueden ser preparadas por mecanismos mediados por interferencia de ARN. Incluso en otras realizaciones, las líneas celulares con resistencia viral pueden ser preparadas por sistemas de recombinación específica del sitio, mutagénesis aleatoria u otros procedimientos conocidos en la técnica.

(a) Edición del genoma mediada por endonucleasa diana

Las endonucleasas diana pueden ser usadas para modificar (es decir, inactivar o alterar) las secuencias cromosómicas específicas de interés. Una secuencia cromosómica específica puede ser inactivada por introducción en una célula de una endonucleasa específica o un nucleico que codifica la endonucleasa específica, que se dirige a una secuencia cromosómica específica. En una realización, la endonucleasa específica reconoce y se une a la secuencia cromosómica específica e introduce una ruptura de hebra doble que es reparada por un proceso de reparación de unión final no homóloga (NHEJ). Dado que la NHEJ es propensa a error, puede ocurrir una eliminación, inserción y/o sustitución de al menos un nucleótido, interrumpiendo por consiguiente el marco de lectura de la secuencia cromosómica de modo tal que no se produzca producto de proteína. En otra realización, las endonucleasas diana también pueden ser usadas para alterar una secuencia cromosómica a través de una reacción de recombinación homóloga por la introducción conjunta de un polinucleótido que tiene identidad sustancial de secuencia con una porción de la secuencia cromosómica específica. La ruptura de hebra doble introducida por la endonucleasa es reparada por un proceso de reparación dirigida por homología de modo tal que la secuencia cromosómica es intercambiada con el polinucleótido de un modo que genera que la secuencia cromosómica sea cambiada o alterada.

Una variedad de endonucleasas diana puede ser usada para modificar las secuencias cromosómicas de interés. La endonucleasa específica puede ser una proteína natural o una proteína manipulada. Las endonucleasas diana adecuadas incluyen, sin límite, nucleasas de dedos de cinc (ZFN), endonucleasas CRISPR/Cas, nucleasas efectoras similares al activador de transcripción (TALE), meganucleasas, endonucleasas diana del sitio, y agentes inductores de ruptura de doble hebra de ADN diana artificial.

(i) Nucleasas de dedos de cinc

En realizaciones específicas, la endonucleasa específica puede ser una nucleasa de dedos de cinc (ZFN). Las ZFN son unidas a una secuencia diana específica buscada e introducen una ruptura de hebra doble en la secuencia específica. Normalmente, una ZFN comprende un dominio de unión al ADN (es decir, dedos de cinc) y un dominio de escisión (es decir, nucleasa), cada uno de los cuales es descrito a continuación.

Dominio de unión al ADN. Un dominio de unión al ADN o los dedos de cinc pueden ser manipulados para reconocer y ser unidos a cualquier secuencia de ácido nucleico de elección. Véase, 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; y Santiago *et al.* (2008) Proc. Natl. Acad. Sci. USA 105:5809-5814. Un dominio de unión de los dedos de cinc manipulado puede tener una especificidad de unión novedosa en comparación con una proteína de dedos de cinc natural. Los procedimientos de manipulación incluyen, pero sin limitación, el diseño racional y diversos tipos de selección. El diseño racional incluye, por ejemplo, el uso de bases de datos que comprenden secuencias de nucleótidos en duplete, triplete y/o cuadruplete y secuencias de aminoácidos de dedos de cinc individuales, en las que cada secuencia de nucleótidos en duplete, triplete o cuadruplete es asociada con una o más secuencias de aminoácidos de los dedos de cinc que son unidos a la secuencia en triplete o cuadruplete particular. Véase, por ejemplo, la Patente de los Estados Unidos Núm. 6.453.242 y 6.534.261 como ejemplo, el algoritmo descrito en la Patente de los Estados Unidos 6.453.242 puede ser usado para diseñar un dominio de unión a los dedos de cinc para buscar una secuencia preseleccionada. Los procedimientos alternativos, tales como el diseño racional usando una tabla del código de reconocimiento no degenerado también pueden ser usados para diseñar un dominio de unión a los dedos de cinc para buscar una secuencia específica (Sera *et al.* (2002) Biochemistry 41:7074-7081). Las herramientas de base web disponibles al público para identificar los sitios blanco potenciales en las secuencias de ADN así como diseñar los dominios de unión a los dedos de cinc son conocidas en la técnica. Por ejemplo, las herramientas para identificar sitios blancos potenciales en las secuencias de ADN pueden ser encontrados en <http://www.zincfingertools.org>. Las herramientas para diseñar dominios de unión de los dedos de cinc pueden ser halladas en <http://zifit.partners.org/ZiFiT>. (Véase 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 los dedos de cinc puede ser diseñado para reconocer y unirse a una secuencia de ADN que oscila de aproximadamente 3 nucleótidos a aproximadamente 21 nucleótidos de longitud. En una realización, el dominio de unión de los dedos de cinc puede ser diseñado para reconocer y unirse a una secuencia de ADN que oscila de aproximadamente 9 a aproximadamente 18 nucleótidos de longitud. En general, los dominios de unión de los dedos de cinc de las nucleasas con dedos de cinc usadas en esta memoria comprenden al menos tres regiones de reconocimiento con dedos de cinc o dedos de cinc, en las que cada dedo de cinc es unido a 3 nucleótidos. En una realización, el dominio de unión de los dedos de cinc comprende cuatro regiones de reconocimiento con dedos de cinc. En otra realización, el dominio de unión de los dedos de cinc comprende cinco regiones de reconocimiento con dedos de cinc. Incluso en otra realización, el dominio de unión de los dedos de cinc comprende seis regiones de reconocimiento con dedos de cinc. Un dominio de unión de los dedos de cinc puede ser diseñado para unirse a cualquier secuencia de ADN diana adecuada. Véase, por ejemplo, las Patentes de los Estados Unidos Núm. 6.607.882; 6.534.261 y 6.453.242.

Los procedimientos de ejemplo de selección de una región de reconocimiento con dedos de cinc incluyen sistemas de visualización de fago y de dos híbridos, que son descritos en la Patente de los Estados Unidos Núm. 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 por los dominios de unión de los dedos de cinc ha sido descrita, por ejemplo, en el documento WO 02/077227.

Los dominios de unión de los dedos de cinc y procedimientos para diseño y construcción de las proteínas de fusión (y los polinucleótidos que codifican las mismas) son conocidos para los expertos en la técnica y son descritos en detalle, por ejemplo, en la Patente de los Estados Unidos Núm. 7.888.121, las regiones de reconocimiento de los dedos de cinc y/o las proteínas dedos de cinc con múltiples dedos de cinc pueden ser ligadas entre sí usando secuencias ligadoras adecuadas, incluyendo, por ejemplo, ligadores de cinco o más aminoácidos de longitud. Véase la Patente de los Estados Unidos Núm. 6.479.626; 6.903.185 y 7.153.949, para conocer ejemplos no limitativos o secuencias ligadoras de seis o más aminoácidos de longitud. El dominio de unión de los dedos de cinc descritos en esta memoria puede incluir una combinación de ligaduras adecuadas entre los dedos de cinc individuales de la proteína.

En algunas realizaciones, la nucleasa con dedos de cinc comprende de manera adicional una señal o secuencia de localización nuclear (NLS). Una NLS es una secuencia de aminoácidos que facilita la especificación de la proteína de nucleasa con dedos de cinc en el núcleo para introducir una ruptura de hebra doble en la secuencia específica en el

cromosoma. Las señales de localización nuclear son conocidas en la técnica. Véase, por ejemplo, Makkerh *et al.* (1996) *Current Biology* 6:1025-1027.

Dominio de escisión. La nucleasa con dedos de cinc también incluye un dominio de escisión. La porción del dominio de escisión de la nucleasa con dedos de cinc puede ser obtenida de cualquier endonucleasa o exonucleasa. Los ejemplos no limitativos de endonucleasas de los cuales puede ser derivado un dominio de escisión incluyen pero sin limitación, endonucleasas de restricción y endonucleasas de dirección. Véase, por ejemplo, New England Biolabs Catalog o Belfort *et al.* (1997) *Nucleic Acids Res.* 25:3379-3388. Las enzimas adicionales que escinden el ADN son conocidas (por ej., Nucleasa S1; nucleas de porotos mung; DNase I pancreática; nucleasa de micrococo; endonucleasa HO de levadura). Véase también Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993. Una o más de estas enzimas (o sus fragmentos funcionales) pueden ser usadas como fuente de dominios de escisión.

Un dominio de escisión también puede ser derivado de una enzima o porción de la misma, como es descrito anteriormente, que requiere dimerización para la actividad de escisión. Dos nucleasas con dedos de cinc pueden ser requeridas para la escisión, ya que cada nucleasa comprende un monómero de dímero enzimático activo. Alternativamente, una nucleasa con dedos de cinc puede comprender ambos monómeros para crear un dímero enzimático activo. Como se usa en la presente memoria, un "dímero enzimático activo" es un dímero enzimático capaz de escindir una molécula de ácido nucleico. Los dos monómeros de escisión pueden derivarse de la misma endonucleasa (o fragmentos funcionales de esta), o cada monómero puede derivar de una endonucleasa diferente (o fragmentos funcionales de esta).

Cuando dos monómeros de escisión son usados para formar un dímero enzimático activo, los sitios de reconocimiento para las dos nucleasas con dedos de cinc son preferentemente descartados de modo tal que la unión de las nucleasas con dedos de cinc a sus respectivos sitios de reconocimiento coloca los monómeros de escisión en una orientación espacial entre sí que permite que los monómeros de escisión formen un dímero enzimático activo, por ej., por dimerización. Como resultado, los bordes cercanos de los sitios de reconocimiento pueden ser separados por aproximadamente 5 a aproximadamente 18 nucleótidos. Por ejemplo, los bordes cercanos pueden ser 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 integral de nucleótidos o pares de nucleótido pueden intervenir entre dos sitios de reconocimiento (por ej., de aproximadamente 2 a aproximadamente 50 pares de nucleótidos o más). Los bordes cercanos de los sitios de reconocimiento de las nucleasas con dedos de cinc, tales como por ejemplo aquellas desveladas en la presente memoria en detalle, pueden ser separadas por 6 nucleótidos. En general, el sitio de escisión yace entre los sitios de reconocimiento.

Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unión específica de la secuencia al ADN (en el sitio de reconocimiento), y escindir el ADN en el sitio de unión o cerca del mismo. Ciertas enzimas de restricción (por ej., Tipo IIS) escinden el ADN en los sitios retirados del sitio de reconocimiento y tienen unión separable y dominios de escisión. Por ejemplo, la enzima FokI tipo IIS cataliza la escisión de hebra doble del ADN, en 9 nucleótidos de su sitio de reconocimiento en una hebra y 13 nucleótidos de su sitio de reconocimiento en la otra. Véase, por ejemplo, la Patente de los Estados Unidos 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. De esta manera, una nucleasa con dedos de cinc puede comprender el dominio de escisión de al menos una enzima de restricción tipo IIS y uno o más dominios de unión de los dedos de cinc, que pueden o no ser manipulados. Las enzimas de restricción Tipo IIS ejemplificativas son descritas 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 están contemplados por la presente descripción. Véase, por ejemplo, Roberts *et al.* (2003) *Nucleic Acids Res.* 31:418-420.

Una enzima de restricción Tipo IIS ejemplificativa, cuyo dominio de escisión es separable del dominio de unión, es FokI. Esta enzima particular es activa como dímero (Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95: 10, 570-10, 575). De esta manera, para los fines de la presente divulgación, la porción de la enzima FokI usada en una nucleasa con dedos de cinc es considerada un monómero de escisión. De este modo, para la escisión de hebra doble específica con el uso de un dominio de escisión FokI, dos nucleasas con dedos de cinc, comprendiendo cada una un monómero de escisión FokI, pueden ser usadas para reconstituir un dímero enzimático.

Alternativamente, una molécula de polipéptido individual que contiene un dominio de unión de los dedos de cinc y dos monómeros de escisión FokI también pueden ser usados.

En ciertas realizaciones, el dominio de escisión comprende uno o más monómeros de escisión manipulados que minimizan o previenen la homodimerización. A modo de ejemplo no limitativo, los residuos 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 medios dominios de escisión FokI. Los monómeros de escisión manipulados ejemplificativos de FokI que forman heterodímeros obligados incluyen un par en el cual un primer monómero de escisión incluye mutaciones en las posiciones del residuo de aminoácidos 490 y 538 de FokI y un segundo monómero de escisión que incluye mutaciones en las posiciones del residuo de aminoácidos 486 y 499.

De este modo, en una realización de los monómeros de escisión manipulados, la 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 a Gln (Q) con Glu (E); y una mutación en la posición 499 reemplaza a Iso (I) con Lys (K). Específicamente, los monómeros de escisión manipulados pueden ser preparados por 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 manipulado designado "E490K:I538K" y por 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 manipulado "Q486E:I499K." Los monómeros de escisión manipulados descritos más arriba son mutantes de heterodímero obligados en los cuales la escisión aberrante es minimizada o abolida. Los monómeros de escisión manipulados pueden ser preparados usando un procedimiento adecuado, por ejemplo, por mutagénesis dirigida al sitio de monómeros de escisión tipo salvaje (FokI) como se describe en la Patente de los Estados Unidos Núm. 7.888.121, que se incorpora en la presente memoria en su totalidad.

Dominios adicionales. En algunas realizaciones, la nucleasa con dedos de cinc además comprende al menos una secuencia de localización nuclear (NLS). Una NLS es una secuencia de aminoácidos que facilita la especificación de la proteína de nucleasa con dedos de cinc en el núcleo para introducir una ruptura de hebra doble en la secuencia específica en el cromosoma. Las señales de localización nuclear son conocidas en la técnica (véase por ej., Lange *et al.*, J. Biol. Chem., 2007, 282:5101-5105). Por ejemplo, en una realización, la NLS puede ser una secuencia monopartita, tal como PKKKRKV (SEC ID NÚM.: 1) o PKKKRRV (SEC ID NÚM.: 2). En otra realización, la NLS puede ser una secuencia bipartita. Incluso en otra realización, la NLS puede ser KR-PAATKKAGQAKKKK (SEC ID NÚM.: 3). La NLS puede estar ubicada en el extremo terminal N, el C-término o en una ubicación interna de la proteína.

En realizaciones adicionales, la nucleasa con dedos de cinc también puede comprender al menos un dominio que penetra en la célula. En una realización, el dominio que penetra en la célula puede ser una secuencia que penetra en la célula derivada de la proteína HIV-1 TAT. Como ejemplo, la secuencia que penetra en la célula TAT puede ser GRKKRRQRRRPPQPKKKRKV (SEC ID NÚM.:4). En otra realización, el dominio de penetración en la célula puede ser TLM (PLSSIFSRIGDPPKKKRKV; SEC ID NÚM.: 5), una secuencia de péptido que penetra en la célula derivada del virus de la hepatitis B. Incluso en otra realización, el dominio de penetración en la célula puede ser MPG (GALFLGWLGAAGSTMGAPKKKKRKV; SEC ID NÚM.: 6 o GALFLGFLGAAGSTMGAWSQP- KKKRKV; SEC ID NÚM.: 7). En una realización adicional, el dominio de penetración en la célula puede ser Pep-1 (KETWWETWW-TEWSQPKKKKRKV; SEC ID NÚM.: 8), VP22, un péptido que penetra en la célula del virus del Herpes simple, o una secuencia de péptido de poliarginina. El dominio que penetra en la célula puede estar ubicado en el extremo terminal N, el extremo terminal C o en una ubicación interna de la nucleasa con dedos de cinc.

Incluso en otras realizaciones, la nucleasa con dedos de cinc puede comprender, además, al menos un dominio marcador. Los ejemplos no limitativos de dominios marcadores incluyen proteínas fluorescentes, marcas de purificación y marcas de epítopes. En una realización, el dominio marcador puede ser una proteína fluorescente. Ejemplos no limitativos de proteínas fluorescentes adecuadas incluyen proteínas fluorescentes verdes (por ej., GFP, GFP-2, tagGFP, turboGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarillas (por ej. YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), proteínas fluorescentes azules (por ej. EBFP, EBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-sapphire), proteínas fluorescentes cian (por ej. ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes rojas (mKate, mKate2, mPlum, monómero DsRed, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monómero, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), y proteínas fluorescentes naranja (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) o cualquier otra proteína fluorescente adecuada. En otra realización, el dominio marcador puede ser una marca de purificación y/o una marca de epítopo. Los marcadores adecuados incluyen, pero sin limitación, glutiona-S-transferasa (GST), proteína que se une a quitina (CBP), proteína que se une a la maltosa, tioredoxina (TRX), poli(NANP), marca de purificación por afinidad en tándem (TAP), myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu- Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, proteína portadora de carboxilo de biotina (BCCP), y calmodulina. El dominio marcador puede estar ubicado en el extremo terminal N, el extremo terminal C o en una ubicación interna de la nucleasa con dedos de cinc.

El dominio marcador puede estar ligado a la nucleasa con dedos de cinc por un péptido 2A (Szymczak *et al.*, 2004, Nat. Biotechnol., 589(5):589-94). Un péptido 2A fue caracterizado originalmente en virus de ARN de hebra positiva, que producen una poliproteína que es "escindida" durante la traducción en proteínas individuales maduras. Más específicamente, la región del péptido 2A (~20 aminoácidos) media la "escisión" en su propio extremo terminal C para liberarse de la región corriente abajo de la poliproteína. En general, una secuencia de péptido 2A termina con una glicina y un residuo de prolina. Durante la traducción de un péptido 2A, el ribosoma pausa después del residuo de glicina, lo que resulta en la liberación de la cadena de polipéptidos naciente. La traducción se reanuda, con el residuo de prolina de la secuencia 2A que se convierte en el primer aminoácido de la proteína corriente abajo.

(ii) Endonucleasas CRISPR/Cas

En otras realizaciones, la endonucleasa diana puede ser una endonucleasa CRISPR/Cas. Las endonucleasas CRISPR/Cas son endonucleasas guiadas por ARN derivadas de sistemas CRISPR/Cas. Las bacterias y arqueas han evolucionado un sistema inmunitario adaptativo a base de ADN que usa CRISPR (del inglés, repetición palindrómica

cortas agrupadas y regularmente espaciada) y proteínas Cas (asociadas con CRISPR) para detectar y destruir los virus o plásmidos invasores. Las endonucleasas CRISPR/Cas pueden ser programadas para introducir rupturas específicas del sitio de doble hebra al proporcionar ARN guía sintéticos específicos del sitio (Jinek *et al.*, 2012, Science, 337:816-821).

5 **Endonucleasa.** La endonucleasa CRISPR/Cas puede ser derivada de un sistema CRISPR/Cas tipo I, tipo II, o tipo III. Ejemplos no limitativos de proteínas CRISPR/Cas adecuadas incluyen Cas3, Cas4, Cas5, Cas5e (o CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9, Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (o CasA), Cse2 (o CasB), Cse3 (o CasE), Cse4 (o CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csz1, Csx15, Csf1, Csf2, Csf3, Csf4 y Cu1966.

10 En una realización, la endonucleasa CRISPR/Cas es derivada de un sistema CRISPR/Cas tipo II. En realizaciones ejemplificativas, la endonucleasa CRISPR/Cas es derivada de una proteína Cas9. La proteína Cas9 puede ser de *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus sp.*, *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*,
15 *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycoloides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas sp.*, *Crocospaera watsonii*, *Cyanothece sp.*, *Microcystis aeruginosa*, *Synechococcus sp.*, *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccsii*, *Candidatus Desulfurudis*, *Clostridium botulinum*, *Clostridium difficile*, *Fingoldia magna*,
20 *Natranaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium vinosum*, *Marinobacter sp.*, *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc sp.*, *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira sp.*, *Lyngbya sp.*, *Microcoleus chthonoplastes*, *Oscillatoria sp.*, *Petrotoga mobilis*, *Thermosipho africanus* o *Acaryochloris marina*. En una realización específica, la proteína Cas9 es de *Streptococcus pyogenes*.

25 En general, las proteínas CRISPR/Cas comprenden al menos un dominio de reconocimiento del ARN y/o de unión al ARN. Los dominios de reconocimiento del ARN y/o de unión al ARN interactúan con el ARN guía de modo tal que la proteína CRISPR/Cas es dirigida a una secuencia cromosómica específica o una secuencia cromosómica (es decir, sitio específico). Las proteínas CRISPR/Cas también pueden comprender dominios de nucleasa (es decir, dominios de DNasa o RNasa), dominios que son unidos al ADN, dominios de helicasa, dominios de interacción proteína-proteína, dominios de dimerización, así como otros dominios.

30 La endonucleasa CRISPR/Cas puede ser derivada de una proteína CRISPR/Cas de tipo salvaje, una proteína CRISPR/Cas modificada o un fragmento de una proteína de tipo salvaje o proteína CRISPR/Cas modificada. La proteína CRISPR/Cas puede ser modificada para aumentar la afinidad de unión y/o la especificidad del ácido nucleico, alterar una actividad enzimática, y/o cambiar otra propiedad de la proteína. Por ejemplo, los dominios de nucleasa (es decir DNasa, RNasa) de la proteína CRISPR/Cas pueden ser modificados, eliminados o inactivados. La proteína CRISPR/Cas puede ser truncada para retirar los dominios que no son esenciales para la función de la proteína. La proteína CRISPR/Cas también puede ser truncada o modificada para optimizar la actividad de la proteína o un dominio efector con la proteína CRISPR/Cas.

35 En algunas realizaciones, la endonucleasa CRISPR/Cas puede ser derivada de una proteína Cas9 de tipo salvaje o sus fragmentos. En otras realizaciones, la endonucleasa CRISPR/Cas puede ser derivada de una proteína Cas9 modificada. Por ejemplo, la secuencia de aminoácidos de la proteína Cas9 puede ser modificada para alterar una o más propiedades (por ej., actividad de nucleasa, afinidad, estabilidad, etc.) de la proteína. Alternativamente, los dominios de la proteína Cas9 no involucrados en la escisión guiada por el ARN pueden ser eliminados de la proteína de modo tal que la proteína Cas9 modificada sea más pequeña que la proteína Cas9 tipo salvaje.

40 En general, una proteína Cas comprende al menos dos dominios de nucleasa (es decir DNasa). Por ejemplo, una proteína Cas9 puede comprender un dominio de nucleasa similar a RuvC y un dominio de nucleasa similar a HNH. Los dominios RuvC y HNH trabajan juntos para cortar las hebras simples para hacer una ruptura de doble hebra en el ADN (Jinek *et al.*, Science, 337: 816-821). En una realización, la endonucleasa basada en CRISPR es derivada de una proteína Cas9 y comprende dominios de nucleasa de dos funciones.

45 Los sitios específicos reconocidos por los sistemas CRISPR/Cas naturales que tienen longitudes de aproximadamente 14-15 pb (Cong *et al.*, Science, 339:819-823). El sitio específico no tiene limitación de secuencia excepto que la secuencia complementaria al extremo 5' del ARN guía (es decir, llamada secuencia protoseparadora) es inmediatamente seguida por (3' o más abajo) una secuencia de consenso. Esta secuencia de consenso también es conocida como motivo protoseparador adyacente (o PAM). Los ejemplos de PAM incluyen, pero sin limitación, NGG, NGGNG, y NNAGAAW (en los que N es definido como cualquier nucleótido y W es definido indistintamente como A o T). A la longitud típica, solo aproximadamente 5-7% de los sitios específicos serían únicos dentro del genoma específico, indicando que los efectos de compensación del objetivo podrían ser significativos. La longitud del sitio específico puede ser expandida por requisito de dos eventos de unión. Por ejemplo, las endonucleasas basadas en
55 CRISPR pueden ser modificadas de modo tal que puedan solo escindir una hebra de una secuencia de hebra doble
60

(es decir, convertida en nicasas). De este modo, el uso de una nicasa basada en CRISPR en combinación con dos ARN guía diferentes sería esencialmente el doble de la longitud del sitio específico, mientras que aún efectúa una ruptura de hebra doble.

5 En algunas realizaciones, por lo tanto, la endonucleasa derivada de Cas9 puede ser modificada para contener solo un dominio de nucleasa funcional (indistintamente un dominio de nucleasa similar a RuvC o similar a HNH). Por ejemplo, la proteína derivada de Cas9 puede ser modificada de modo tal que uno de los dominios de nucleasa es eliminado o mutado de modo tal que ya no es funcional (es decir, el dominio carece de actividad de nucleasa). En algunas realizaciones, en las cuales uno de los dominios de nucleasa es inactivo, la proteína derivada de Cas9 es capaz de introducir un deterioro en un ácido nucleico de hebra doble (dicha proteína es denominada "nicasa"), pero no disocia el ADN de hebra doble. Por ejemplo, una conversión de aspartato a alanina (D10D) en un dominio similar a RuvC convierte la proteína derivada de Cas9 en una nicasa "HNH". Del mismo modo, una conversión de histidina a alanina (H840A) en algunas instancias, la histidina está ubicada en la posición 839) en un dominio de HNH convierte la proteína derivada de Cas9 en una nicasa "RuvC". De este modo, por ejemplo, en una realización, la nicasa derivada de Cas9 tiene una conversión de aspartato a alanina (D10A) en un dominio similar a RuvC. En otra realización, la nicasa derivada de Cas9 tiene una conversión de histidina a alanina (H840A o H839A) en un dominio de HNH. Los dominios de nucleasa similares a RuvC o similares a HNH de la nicasa derivada de Cas9 pueden ser modificados usando procedimientos conocidos, tales como mutagénesis dirigida al sitio, mutagénesis mediada por PCR y síntesis genética total, así como otros procedimientos conocidos en la técnica.

20 **Dominios adicionales.** La endonucleasa o nicasa CRISPR/Cas generalmente comprende al menos una señal de localización nuclear (NLS). Por ejemplo, en una realización, la NLS puede ser una secuencia monopartita, tal como PKKKRKY (SEC ID NÚM.: 1) o PKKKRRV (SEC ID NÚM.: 2). En otra realización, la NLS puede ser una secuencia bipartita. Incluso en otra realización, la NLS puede ser KRPAATKKAGQAKKKK (SEC ID NÚM.: 3). La NLS puede estar ubicada en el extremo terminal N, el extremo terminal C o en una ubicación interna de la proteína.

25 En algunas realizaciones, la endonucleasa CRISPR/Cas o nicasa puede comprender además al menos un dominio que penetra en la célula. El dominio que penetra en la célula puede ser una secuencia que penetra en la célula derivada de la proteína HIV-1 TAT. Como ejemplo, la secuencia que penetra en la célula TAT puede ser GRKKRRQRRRPPQPKKKRKY (SEC ID NÚM.: 4). En otra realización, el dominio de penetración en la célula puede ser TLM (PLSSIFSRIGDPPKKKKRKY; SEC ID NÚM.: 5), una secuencia de péptido que penetra en la célula derivada del virus de la hepatitis B. Incluso en otra realización, el dominio de penetración en la célula puede ser MPG (GALFLGWLGAAGSTMGAPKKKKRKY; SEC ID NÚM.: 6 o GALFLGFLGAAGSTMGAWSQP- KKKRKY; SEC ID NÚM.: 7). En una realización adicional, el dominio de penetración en la célula puede ser Pep-1 (KETWWETWW-TEWSQPKKKRKY; SEC ID NÚM.: 8), VP22, un péptido que penetra en la célula del virus del Herpes simple, o una secuencia de péptido de poliarginina. El dominio que penetra en la célula puede estar ubicado en el extremo terminal N, el extremo terminal C o en una ubicación interna de la proteína.

35 Incluso en otras realizaciones, la endonucleasa CRISPR/Cas o nicasa puede comprender además al menos un dominio marcador. Ejemplos no limitativos de dominios marcadores incluyen proteínas fluorescentes, marcas de purificación y marcas de epítopes. En una realización, el dominio marcador puede ser una proteína fluorescente. Ejemplos no limitativos de proteínas fluorescentes adecuadas incluyen proteínas fluorescentes verdes (por ej., GFP, GFP-2, tagGFP, turboGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarillas (por ej. YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), proteínas fluorescentes azules (por ej. EBFP, EBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-sapphire), proteínas fluorescentes cian (por ej. ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes rojas (mKate, mKate2, mPlum, monómero DsRed, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monómero, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), y proteínas fluorescentes naranja (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) o cualquier otra proteína fluorescente adecuada. En otra realización, el dominio marcador puede ser una marca de purificación y/o una marca de epítipo. Los marcadores adecuados incluyen, pero sin limitación, glutationa-S-transferasa (GST), proteína que se une a quitina (CBP), proteína que se une a la maltosa, tioredoxina (TRX), poli(NANP), marca de purificación por afinidad en tándem (TAP), myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu- Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, proteína portadora de carboxilo de biotina (BCCP), y calmodulina. El dominio marcador puede estar ubicado en el extremo terminal N, el extremo terminal C o en una ubicación interna de la proteína. El dominio marcador puede estar ligado a la endonucleasa CRISPR/Cas o nicasa por un péptido 2A (Szymczak *et al.*, 2004, Nat. Biotechnol., 589(5):589-94).

55 **ARN guía.** La endonucleasa CRISPR/Cas es guiada al sitio especificado por un ARN guía. Un ARN guía interactúa con ambos, la endonucleasa CRISPR/Cas y el sitio específico en el cromosoma, en el que el sitio de la endonucleasa o nicasa CRISPR/Cas escinde al menos una hebra de la secuencia de hebra doble. El ARN guía puede ser introducido en la célula junto con la endonucleasa CRISPR/Cas o el ácido nucleico que codifica la endonucleasa CRISPR/Cas. Alternativamente, el ADN que codifica ambos, la endonucleasa CRISPR/Cas y el ARN guía puede ser introducido en la célula.

60 Un ARN guía comprende tres regiones: una primera región en el extremo 5' que es complementaria a la secuencia en el sitio específico, una región interna que forma una estructura en bucle de vástago, y una tercera región 3' que

permanece esencialmente con hebra simple. La primera región de cada ARN guía es diferente de modo tal que cada ARN guía guíe una endonucleasa o nicasa CRISPR/Cas a un sitio blanco específico. La segunda y tercera regiones (también llamada región andamio) de cada ARN guía pueden ser iguales en todos los ARN guía.

5 La primera región del ARN guía es complementaria a la secuencia (es decir, secuencia protoseparadora) en el sitio específico de tal modo que la primera región del ARN guía puede ser par base con la secuencia en el sitio específico. En general, no hay malas equivalencias entre la secuencia de la primera región del ARN guía y la secuencia en el sitio específico (es decir, la complementariedad es total). En diversas realizaciones, la primera región del ARN guía puede comprender de aproximadamente 10 nucleótidos a más de aproximadamente 25 nucleótidos. Por ejemplo, la región del apareamiento de bases entre la primera región del ARN guía y el sitio específico en la secuencia cromosómica
10 puede ser de aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, o más de 25 nucleótidos de longitud. En realizaciones ejemplificativas, la primera región del ARN guía es de aproximadamente 19 o 20 nucleótidos de longitud.

15 El ARN guía también comprende una segunda región que forma una estructura secundaria. En algunas realizaciones, la estructura secundaria comprende un vástago (u horquilla) y un bucle. La longitud del bucle y la horquilla pueden variar. Por ejemplo, el bucle puede oscilar de aproximadamente 3 a aproximadamente 10 nucleótidos de longitud, y el vástago puede oscilar de aproximadamente 6 a aproximadamente 20 pares de base de longitud. El vástago puede comprender una o más salientes de 1 a aproximadamente 10 nucleótidos. De este modo, la longitud general de la segunda región puede oscilar de aproximadamente 16 a aproximadamente 60 nucleótidos de longitud. En una realización ejemplificativa, el bucle es de aproximadamente 4 nucleótidos de longitud y el vástago comprende
20 aproximadamente 12 pares de bases.

25 El ARN guía también comprende una tercera región en el extremo 3' que permanece esencialmente de hebra simple. De este modo, la tercera región no tiene complementariedad con ninguna secuencia cromosómica en la célula de interés y no tiene complementariedad con el resto del ARN guía. La longitud de la tercera región puede variar. En general, la tercera región es de más de aproximadamente 4 nucleótidos de longitud. Por ejemplo, la longitud general de la tercera región puede oscilar de aproximadamente 5 a aproximadamente 60 nucleótidos de longitud.

La longitud combinada de la segunda y tercera regiones (o andamio) del ARN guía pueden oscilar de aproximadamente 30 a aproximadamente 120 nucleótidos de longitud. En un aspecto, la longitud combinada de la segunda y tercera regiones del ARN guía oscila de aproximadamente 70 a aproximadamente 100 nucleótidos de longitud.

30 En algunas realizaciones, el ARN guía comprende una molécula que comprende las tres regiones. En otras realizaciones, el ARN guía puede comprender dos moléculas separadas. La primera molécula de ARN puede comprender la primera región del ARN guía y una mitad del "vástago" de la segunda región del ARN guía. La segunda molécula de ARN puede comprender la otra mitad del "vástago" de la segunda región del ARN guía y la tercera región del ARN guía. De este modo, en esta realización, la primera y segunda moléculas de ARN contienen, cada una, una
35 secuencia de nucleótidos que son complementarias entre sí. Por ejemplo, en una realización, la primera y segunda moléculas de ARN comprenden, cada una, una secuencia (de aproximadamente 6 a aproximadamente 20 nucleótidos) que forman pares de base con la otra secuencia para formar un ARN guía funcional.

(iii) Otras endonucleasas diana

40 En realizaciones adicionales, la endonucleasa diana puede ser una meganucleasa. Las meganucleasas son endodesoxiribunucleasas caracterizadas por secuencias extensas de reconocimiento, es decir, la secuencia de reconocimiento generalmente oscila de aproximadamente 12 pares de bases a aproximadamente 40 pares de bases. Como consecuencia de este requisito, la secuencia de reconocimiento generalmente ocurre solo una vez en cualquier genoma dado. Entre las meganucleasas, la familia de endonucleasas de conducción denominada LAGLIDADG se ha convertido en una herramienta valiosa para el estudio de genomas y manipulación de genomas (véase, por ej., Arnould
45 *et al.*, 2011, *Protein Eng Des Sel*, 24(1-2):27-31). Una meganucleasa puede dirigirse a una secuencia cromosómica diana por modificación de su secuencia de reconocimiento usando técnicas muy conocidos por las personas con experiencia en la técnica.

50 En realizaciones adicionales, la endonucleasa diana puede ser una nucleasa efectora similar al activador de transcripción (TALE). Los TALE son factores de transcripción de los Xanthomas patogénicos de las plantas que pueden ser manipuladas fácilmente para unirse a nuevos blancos de ADN. Los TALE o sus versiones truncadas pueden ser ligados al dominio catalítico de las endonucleasas tales como FokI para crear la endonucleasa diana llamadas nucleasas TALE o TALEN (Sanjana *et al.*, 2012, *Nat Protoc*, 7(1): 171-192).

55 Incluso en otras realizaciones, la endonucleasa diana puede ser una endonucleasa diana del sitio. En particular, la endonucleasa diana del sitio puede ser una nucleasa "cortadora infrecuente" cuya secuencia de reconocimiento ocurre raramente en un genoma. Alternativamente, la endonucleasa diana del sitio puede manipularse para escindir un sitio de interés (Friedhoff *et al.*, 2007, *Methods Mol Biol* 352:1110123). Generalmente, la secuencia de reconocimiento de la endonucleasa diana del sitio ocurre solo una vez en un genoma. En realizaciones alternativas, la endonucleasa diana puede ser un agente inductor de ruptura de la hebra doble del ADN específico artificial.

(iv) Polinucleótido opcional

El procedimiento para modificación del genoma especificado puede comprender, adicionalmente, introducir en la célula al menos un polinucleótido que comprende una secuencia que tiene identidad sustancial de secuencia con una secuencia en al menos un lado del sitio de escisión especificado de modo tal que la ruptura de hebra doble introducida por la endonucleasa diana puede ser reparada por un proceso de reparación dirigido por homología y la secuencia del polinucleótido es intercambiada con la secuencia cromosómica endógena, modificando de esta manera la secuencia cromosómica endógena. Por ejemplo, el polinucleótido comprende una primera secuencia que tiene identidad de secuencia sustancial con la secuencia en un lado del sitio de escisión especificado y una segunda secuencia que tiene identidad de secuencia sustancial con la secuencia en el otro lado del sitio de escisión especificado. Alternativamente, el polinucleótido comprende una primera secuencia que tiene identidad de secuencia sustancial con la secuencia en un lado del sitio de escisión diana y una segunda secuencia que tiene identidad de secuencia sustancial con una secuencia ubicada lejos del sitio de escisión diana. La secuencia ubicada lejos del sitio de escisión especificado puede estar decenas, cientos o miles de nucleótidos corriente arriba o corriente abajo del sitio de escisión diana.

Las longitudes de las primeras y segundas secuencias en el polinucleótido que tienen identidad de secuencia sustancial con las secuencias en la secuencia cromosómica especificada pueden variar y varían. En general, cada una de las primeras y segundas secuencias en el polinucleótido es de al menos 10 nucleótidos de longitud. En diversas realizaciones, las secuencias de polinucleótidos que tienen identidad de secuencia sustancial con las secuencias cromosómicas pueden ser de aproximadamente 15 nucleótidos, aproximadamente 20 nucleótidos, aproximadamente 25 nucleótidos, aproximadamente 30 nucleótidos, aproximadamente 40 nucleótidos, aproximadamente 50 nucleótidos, aproximadamente 100 nucleótidos, o más de 100 nucleótidos de longitud.

La frase "identidad de secuencia sustancial" significa que las secuencias en el polinucleótido tienen al menos aproximadamente 75% de identidad de secuencia con las secuencias cromosómicas de interés. En algunas realizaciones, las secuencias en el polinucleótido tiene 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 las secuencias cromosómicas de interés.

La longitud del polinucleótido puede variar y varía. Por ejemplo, el polinucleótido puede oscilar de aproximadamente 20 nucleótidos de longitud a aproximadamente 200.000 nucleótidos de longitud. En diversas realizaciones, el polinucleótido oscila 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.

Normalmente, el polinucleótido es ADN. El ADN puede ser de hebra simple o de doble hebra. El polinucleótido puede ser un plásmido de ADN, un cromosoma bacteriano artificial (BAC), un cromosoma de levadura (YAC), un vector viral, una parte lineal de ADN, un fragmento de PCR, un ácido nucleico desnudo, o un ácido nucleico en complejo con un vehículo de administración tal como un liposoma o poloxámero. En ciertas realizaciones, el polinucleótido es de hebra simple. En realizaciones ejemplificativas, el polinucleótido es un oligonucleótido de hebra simple que comprende menos de aproximadamente 200 nucleótidos.

En algunas realizaciones, el polinucleótido comprende, además, un marcador. Dicho marcador puede permitir seleccionar integraciones especificadas. En algunas formas de realización, el marcador es un sitio de endonucleasa de restricción. En otras realizaciones, el dominio marcador es una proteína fluorescente, una marca de purificación o una marca de epítipo. Ejemplos no limitativos de proteínas fluorescentes adecuadas incluyen proteínas fluorescentes verdes (por ej., GFP, GFP-2, tagGFP, turboGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarillas (por ej. YFP, EYFP, Citrine, Venus, YFP, PhiYFP, ZsYellow1), proteínas fluorescentes azules (por ej. EBFP, EBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-sapphire), proteínas fluorescentes cian (por ej. ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes rojas (mKate, mKate2, mPlum, monómero DsRed, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monómero, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), y proteínas fluorescentes naranja (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) o cualquier otra proteína fluorescente adecuada. En otras realizaciones, el marcador puede ser una marca de purificación y/o una marca de epítipo. Los marcadores ejemplificativos incluyen, pero sin limitación, glutatona-S-transferasa (GST), proteína que se une a quitina (CBP), proteína que se une a la maltosa, tioredoxina (TRX), poli(NANP), marca de purificación por afinidad en tándem (TAP), myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu- Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, proteína portadora de carboxilo de biotina (BCCP), y calmodulina.

(v) Administración a la célula

El procedimiento comprende introducir la endonucleasa diana en la célula de interés. La endonucleasa puede ser introducida en la célula como proteína aislada purificada o como ácido nucleico que codifica la endonucleasa diana. El ácido nucleico puede ser ARN o ADN. En realizaciones en las que el ácido nucleico que codifica es ARN, el ARNm puede ser 3' tapado y/o 3' poliadenilado. En realizaciones en las que el ácido nucleico que codifica es ADN, el ADN

puede ser lineal o circular. El ADN puede ser parte de un vector, en el que el ADN que codifica puede estar ligado operativamente a un promotor adecuado. Las personas con experiencia en la técnica están familiarizadas con vectores adecuados, promotores, otros elementos de control, y medios para introducir el vector en la célula de interés.

5 Las moléculas de endonucleasa diana y los polinucleótidos opcionales descritos anteriormente pueden ser introducidos 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 shock térmico, transfección por nucleofección, magnetofección, lipofección, impalefacción, transfección óptica, captación de ácidos nucleicos potenciada por el agente exclusivo, y distribución a través de liposomas, inmunoliposomas, virosomas, o viriones artificiales. En una realización específica, las moléculas de endonucleasa y los polinucleótidos diana son introducidos en la célula por nucleofección.

10 En realizaciones en las que más de una molécula de endonucleasa y más de un polinucleótido son introducidos en una célula, las moléculas pueden ser introducidas simultáneamente o consecutivamente. Por ejemplo, las moléculas de endonucleasa diana, cada una específica para un sitio de escisión específica (y polinucleótidos opcionales) pueden ser introducidas al mismo tiempo. Alternativamente, cada molécula de endonucleasa diana, además de los polinucleótidos opcionales puede ser introducida secuencialmente.

15 La relación de las moléculas de endonucleasa diana a los polinucleótidos opcionales puede variar y varía. En general, la relación de las moléculas de endonucleasa diana a los polinucleótidos oscila de aproximadamente 1:10 a aproximadamente 10:1. En diversas realizaciones, la relación de las moléculas de especificación a polinucleótidos puede ser de 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.

(vi) Cultivo de la célula

25 El procedimiento además comprende mantener la célula en las condiciones adecuadas de modo tal que la ruptura de hebra doble introducida por la endonucleasa puede ser reparada por (i) un proceso de reparación de unión final no homóloga de modo tal que la secuencia cromosómica sea modificada por una eliminación, inserción y/o sustitución de al menos un nucleótido u, opcionalmente, (ii) un proceso de reparación dirigido por homología de modo tal que la secuencia cromosómica sea intercambiada con la secuencia del polinucleótido de modo tal que la secuencia cromosómica sea modificada. En realizaciones en las que los ácidos nucleicos que codifican las endonucleasas son introducidos en la célula, el procedimiento comprende mantener la célula en las condiciones adecuadas de modo tal que la célula expresa las endonucleasas diana.

30 En general, la célula es mantenida en las condiciones adecuadas para el crecimiento y/o mantenimiento celular. Las condiciones adecuadas de cultivo celular son muy conocidas en la técnica y son descritas, por ejemplo, en Santiago *et al.* (2008) PNAS 105:5809-5814; Moehle *et al.* (2007) PNAS 104:3055-3060; Urnov *et al.* (2005) Nature 435:646-651; y Lombardo *et al.* (2007) Nat. Biotechnology 25:1298-1306. Las personas con experiencia en la técnica apreciarán que los procedimientos para cultivar células son conocidos en la técnica y pueden variar, y varían, dependiendo del tipo de célula. Puede ser usada optimización de rutina, en todos los casos, para determinar las mejores técnicas para un tipo de célula particular.

35 Durante esta etapa de proceso, las endonucleasas diana reconocen, unen y crean una ruptura de hebra doble en los sitios de escisión diana en la secuencia cromosómica, y durante la reparación de las rupturas de hebra doble una eliminación, inserción, y/o sustitución de al menos un nucleótido es introducida en la secuencia cromosómica diana. En realizaciones específicas, la secuencia cromosómica diana es inactivada.

40 Tras la confirmación de que la secuencia cromosómica de interés ha sido modificada, los clones celulares individuales pueden ser aislados y su genotipo determinado (a través de secuenciación de ADN y/o análisis de proteínas). Las células que comprenden una secuencia cromosómica modificada pueden someterse a uno o más ciclos adicionales de modificación especificada del genoma para modificar las secuencias cromosómicas adicionales (por ej., véase el Ejemplo 1).

(b) Interferencia de ARN

45 En otra realización, una línea celular con resistencia viral puede ser preparada usando un agente de interferencia de ARN (iARN) que inhibe la expresión de un ARNm o transcrito diana. El agente de iARN puede conducir a la escisión del ARNm o transcrito diana. Alternativamente, el agente de iARN puede prevenir o interrumpir la traducción del ARNm específico en la proteína.

50 En algunas realizaciones, el agente de iARN puede ser un ARN de interferencia breve (ARNsi). En general, el ARNsi comprende una molécula de ARN de doble hebra que oscila de aproximadamente 15 a aproximadamente 29 nucleótidos de longitud. El ARNsi puede ser de aproximadamente 16-18, 17-19, 21-23, 24-27 o 27-29 nucleótidos de longitud. En una realización específica, el ARNsi es de aproximadamente 21 nucleótidos de longitud. El ARNsi puede comprender opcionalmente, además, una o dos salientes de hebra simple, por ej., una saliente 3' en uno o ambos extremos. El ARNsi puede ser formado a partir de dos moléculas de ARN hibridadas juntas o, alternativamente, puede

ser generado a partir de un ARN de horquilla corta (ARNsh) (véase a continuación). En algunas realizaciones, las dos hebras de ARNsi son completamente complementarias, de modo tal que no existan faltas de coincidencia o protuberancias en el dúplex formado entre las dos secuencias. En otras realizaciones, las dos hebras de ARNsi son sustancialmente complementarias, de modo tal que pueda existir una o más faltas de coincidencia y/o protuberancias en el dúplex formado entre las dos secuencias. En ciertas realizaciones, uno o ambos de los extremos 5' del ARNsi tienen un grupo fosfato, mientras que en otras realizaciones uno o ambos de los extremos 5' carecen de un grupo fosfato. En otras realizaciones, uno o ambos de los extremos 3' del ARNsi tienen un grupo hidroxilo, mientras que en otras realizaciones uno o ambos de los extremos 5' carecen de un grupo hidroxilo.

Una hebra del ARNsi, a la que se hace referencia como "hebra antisentido" o "hebra guía", incluye una porción hibridada con la transcrito diana. En ciertas realizaciones, la hebra antisentido del ARNsi es completamente complementaria con una región del transcrito diana, es decir, hibridada a la transcrito diana sin una falta de equivalencia o saliente individual sobre una región específica entre aproximadamente 15 y aproximadamente 29 nucleótidos de longitud, preferentemente al menos 16 nucleótidos de longitud, y más preferentemente aproximadamente 18-20 nucleótidos de longitud. En otras realizaciones, la hebra antisentido es sustancialmente complementaria a la región específica, es decir, una o más faltas de equivalencia y/o protuberancias pueden existir en el dúplex formado por la hebra antisentido y la transcrito diana. Normalmente, los ARNsi están dirigidos a secuencias exónicas del transcrito diana. Las personas con experiencia en la técnica están familiarizadas con programas, algoritmos y/o servicios comerciales que diseñan ARNsi para transcritos diana. Un ejemplo es Rosetta siRNA Design Algorithm (Rosetta Inpharmatics, North Seattle, WA) y MISSION® siRNA (Sigma-Aldrich, St. Louis, MO). El ARNsi puede ser sintetizado enzimáticamente *in vitro* usando procedimientos muy conocidos para las personas con experiencia en la técnica. Alternativamente, el ARNsi puede ser sintetizado químicamente usando técnicas de síntesis de oligonucleótidos que son muy conocidas en la técnica.

En otras realizaciones, el agente de iARN puede ser un ARN de horquilla breve (ARNsh). En general, un ARNsh y una molécula de ARN que comprende al menos dos porciones complementarias hibridadas o son capaces de hibridación para formar una estructura de hebra doble lo suficientemente larga para mediar la interferencia de ARN (como se describe con anterioridad), y al menos una porción de hebra simple que forma un bucle que conecta las regiones de ARNsh que forma el dúplex. La estructura también es llamada estructura de vástago-bucle, siendo el vástago la porción del dúplex. En algunas realizaciones, la porción de dúplex de la estructura es completamente complementaria, de modo tal que no existan faltas de coincidencia o protuberancias en la región del dúplex del ARNsh. En otras realizaciones, la porción del dúplex de la estructura es sustancialmente complementaria, de modo tal que existan una o más faltas de coincidencia y/o protuberancias en la porción del dúplex del ARNsh. El bucle de la estructura puede ser de aproximadamente 1 a aproximadamente 20 nucleótidos de longitud, preferentemente de aproximadamente 4 a aproximadamente 10 nucleótidos de longitud, y más preferentemente de aproximadamente 6 a aproximadamente 9 nucleótidos de longitud. El bucle puede estar ubicado indistintamente en el extremo 5' o 3' de la región que es complementaria a la transcrito diana (es decir, la porción antisentido del ARNsh).

El ARNsh puede además comprender una saliente en el extremo 5' o 3'. La saliente opcional puede ser de aproximadamente 1 a aproximadamente 20 nucleótidos de longitud, y más preferentemente de aproximadamente 2 a aproximadamente 15 nucleótidos de longitud. En algunas realizaciones, la saliente comprende uno o más residuos U, por ej., 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. En otras realizaciones, el extremo 3' del ARNsh tiene un grupo hidroxilo, mientras que en otras realizaciones no. En general, los ARNsh son procesados en ARNsi por la maquinaria de ARNi celular conservada. De este modo, los ARNsh son precursores de los ARNsi que son similarmente capaces de inhibir la expresión de una transcripción blanco que es complementaria a la porción de ARNsh (es decir, la porción antisentido del ARNsh). Las personas expertas en la técnica están familiarizadas con los recursos disponibles (como es detallado con anterioridad) para el diseño y la síntesis de ARNsh.

Incluso en otras realizaciones, el agente de iARN puede ser un vector de expresión de iARN. Normalmente, el vector de expresión de iARN es usado para la síntesis extracelular (*in vivo*) de agentes de iARN, tal como ARNsi o ARNsh. En una realización, dos hebras de ARNsi complementarias, separadas, son transcritas usando un vector individual que contiene dos promotores, cada uno de los cuales dirige la transcripción de una hebra de ARNsi individual (es decir, cada promotor está ligado operativamente a una plantilla para el ARNsi de modo que pueda ocurrir la transcripción). Los dos promotores pueden estar en la misma orientación, en cuyo caso cada uno está ligado operativamente a una plantilla para una de las hebras de ARNsi complementarias. Alternativamente, los dos promotores pueden estar en orientaciones opuestas, flanqueando un solo patrón de modo tal que la transcripción para los promotores resulta en la síntesis de dos hebras de ARNsi complementarias. En otra realización, el vector de expresión de ARNi puede contener un promotor que conduce la transcripción de una molécula de ARN simple que comprende dos regiones complementarias, de modo tal que el transcrito forma un ARNsh.

Las personas con experiencia en la técnica apreciarán que es preferente que los agentes de ARNsi o ARNsh sean producidos *in vivo* a través de la transcripción de más de una unidad de transcripción. En términos generales, los promotores usados para dirigir la expresión *in vivo* de una o más unidades de transcripción ARNsi o ARNsh pueden ser promotores de polimerasa III del ARN (Pol III). Ciertos promotores Pol III, tal como U6 o H1, no requieren elementos regulatorios que actúen sobre Cis dentro de la región transcrita, y de este modo se prefieren en ciertas realizaciones. En otras realizaciones, los promotores para Pol II pueden ser usados para conducir la expresión de las una o más

unidades de transcripción de ARNsi o ARNsh. En algunas realizaciones, pueden ser usados promotores Pol II específicos del tejido, específicos de la célula o inducibles.

Un constructo que proporciona una plantilla para la síntesis de ARNsi o ARNsh puede ser producido usando procedimientos estándar de ADN recombinante y ser insertado en cualquiera de una amplia variedad de diferentes vectores adecuados para expresión en células eucarióticas. Las técnicas de ADN recombinantes se describen en Ausubel *et al.*, 2003, *supra* y Sambrook & Russell, 2001, *supra*. Las personas con experiencia en la técnica apreciarán que los vectores pueden comprender secuencias regulatorias adicionales (por ej., secuencia de terminación, secuencia de control de traducción, etc.), además de secuencias del marcador seleccionable. Los plásmidos de ADN son conocidos en la técnica, incluyendo aquellos basados en pBR322, PUC, y demás. Dado que muchos vectores de expresión pueden contener ya un promotor o promotores adecuados, puede ser solo necesario insertar la secuencia de ácido nucleico que codifica el agente de ARNi de interés en la ubicación adecuada con respecto a los promotores. Los vectores virales también pueden ser usados para proporcionar expresión intracelular de los agentes de iARN. Los vectores virales adecuados incluyen vectores retrovirales, vectores lentivirales, vectores adenovirales, vectores del virus adenoasociados, vectores del herpes virus y demás. En una realización específica, el vector de expresión de iRNA es un vector shARN de base lentiviral o partícula lentiviral, tal como se proporciona en los productos de ARNsh MISSION®TRC (Sigma- Aldrich).

Los agentes de iARN o los vectores de expresión de iARN pueden ser introducidos en la célula usando procedimientos muy conocidos por los expertos en la técnica. Dichas técnicas son descritas en Ausubel *et al.*, 2003, *supra* o Sambrook & Russell, 2001, *supra*, por ejemplo. En ciertas realizaciones, el vector de expresión de iARN, por ej., un vector viral, es integrado de manera estable en el genoma de la célula, de modo tal que la expresión de Mgat1 es interrumpida durante generaciones celulares subsiguientes.

(c) Recombinación específica del sitio

En realizaciones alternativas, las líneas celulares con resistencia viral pueden ser preparadas usando técnicas de recombinación específicas del sitio. Por ejemplo, las técnicas de recombinación específicas del sitio pueden ser usadas para eliminar la totalidad o parte de la secuencia cromosómica de interés, o introducir polimorfismos de nucleótidos individuales (SNP) en la secuencia cromosómica de interés. En una realización, la secuencia cromosómica de interés es localizada usando un sistema de recombinación específica del sitio Cre-loxP, un sistema de recombinación específica del sitio Flp-FRT, o variantes de ellos. Dichos sistemas de recombinación están disponibles en el mercado, y la enseñanza adicional para estas técnicas es hallada en Ausubel *et al.*, 2003, *supra*, por ejemplo.

Definiciones

A menos que se defina de otro modo, todos los términos científicos y técnicos usados en la presente memoria tienen el significado comprendido por las personas con experiencia en la técnica a la cual pertenece la presente invención. Las siguientes referencias proporcionan al experto una definición general de muchos de los términos usados en la presente 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); y Hale & Marham, The Harper Collins Dictionary of Biology (1991). Como se usa en la presente memoria, los siguientes términos tienen los significados adscritos a ellos a menos que se especifique de otro modo.

Al presentar elementos de la presente divulgación o sus realizaciones preferentes, los artículos "un/una", "el/la/los/las", "dicho/a" significan que hay un elemento o más. Los términos "que comprende", "que incluye" y "que tiene" pretenden ser inclusivos y significan que puede haber elementos adicionales además de los mencionados.

Como se usa en la presente memoria, "deficiente" se refiere a niveles reducidos o no detectables de enzimas o proteínas específicas, o actividad reducida o no detectable de las enzimas o proteínas específicas.

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

El término "secuencia exógena" se refiere a una secuencia cromosómica que no es nativa para la célula, o una secuencia cromosómica que se mueve a una ubicación cromosómica diferente.

Una célula "modificada genéticamente" se refiere a una célula en la cual el genoma ha sido modificado, es decir, la célula contiene al menos una secuencia cromosómica que ha sido manipulada para contener una inserción de al menos un nucleótido, una eliminación de al menos un nucleótido y/o una sustitución de al menos un nucleótido.

Los términos "modificación del genoma" y "edición del genoma" se refieren a procesos por los cuales una secuencia cromosómica específica es cambiada de modo tal que la secuencia cromosómica es modificada. La secuencia cromosómica puede ser modificada para comprender una inserción de al menos un nucleótido, una eliminación de al menos un nucleótido, y/o una sustitución de al menos un nucleótido. La secuencia cromosómica modificada es inactivada de manera tal que no es realizado ningún producto. Alternativamente, la secuencia cromosómica puede ser modificada de modo tal que es realizado un producto alterado.

Un "gen" como se usa en la presente memoria, se refiere a una región de ADN (incluyendo exones e intrones) que codifican un producto genético, así como todas las regiones de ADN que regulan la producción del producto genético, tanto si dichas secuencias regulatorias son adyacentes o no a las secuencias de codificación y/o transcritas. De esta manera, un gen incluye, pero no se limita necesariamente a, las secuencias promotoras, terminadores, secuencias reguladoras de transcripción tal como los sitios que son unidos al ribosoma y sitios internos de entrada al ribosoma, potenciadores, silenciadores, insultadores, elementos de límite, orígenes de replicación, sitios de unión a la matriz, y regiones de control del locus.

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

Los términos "ácido nucleico" y "polinucleótido" se refieren a un polímero desoxirribonucleótido o ribonucleótido, en conformación lineal o circular. Para los fines de la presente divulgación, esos términos no deben considerarse limitativos con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, además de nucleótidos que están modificados en las fracciones base, azúcar y/o fosfato. En general, un análogo de un nucleótido particular tiene la misma especificidad de formación de pares de bases, es decir, un análogo de A formará pares de bases con T. Los nucleótidos de un ácido nucleico o polinucleótido pueden estar ligados por fosfodiéster, fosfotioato, fosforamida, uniones fosfordiamidato, o sus combinaciones.

El término "nucleótido" se refiere a desoxirribonucleótidos o ribonucleótidos. Los nucleótidos pueden ser nucleótidos estándar (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 purina o pirimidina modificadas o una fracción de ribosa modificada. Un análogo de nucleótido puede ser un nucleótido natural (por ej., inosina) un nucleótido no natural. Los ejemplos no limitativos de modificaciones de las fracciones azúcar o base de un nucleótido incluyen la adición (o eliminación) de los grupos acetilo, grupos amino, grupos carboxilo, grupos carboximetilo, grupos hidroxilo, grupos metilo, grupos fosforilo y grupos tiol, así como la sustitución de los átomos de carbono y nitrógeno de las bases con otros átomos (por ej., purinas 7-deaza). Los análogos de nucleótidos también incluyen didesoxi nucleótidos, nucleótidos de 2'-O-metilo, ácidos nucleicos trabados (LNA), ácidos nucleicos peptídicos (PNA) y morfolinos.

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

Como se usan en la presente memoria, los términos "sitio diana" o "secuencia diana" se refieren a una secuencia de ácido nucleico que define una porción de una secuencia cromosómica que será modificada o editada y a la que la endonucleasa diana es manipulada para reconocer y unirse, a condición de que existan condiciones suficientes para la unión.

Los términos "corriente arriba" y "corriente abajo" se refieren a ubicaciones en la secuencia de ácido nucleico relativa a una posición fija. Corriente arriba se refiere a la región que está 5' (es decir, cerca del extremo 5' de la hebra) a la posición y corriente abajo se refiere a la región que está 3' (es decir, cerca del extremo 3' de la hebra) a la posición.

Como se usa en la presente memoria, "resistencia viral" se refiere a la capacidad de las células de resistir a la infección viral. Más específicamente, la entrada de un virus y/o la propagación de un virus es reducida o eliminada en las líneas celulares manipuladas desveladas en la presente memoria en comparación con las líneas celulares progenitoras no modificadas.

Las técnicas para determinar la identidad de la secuencia de ácido nucleico y de aminoácidos son conocidas en la técnica. Normalmente, dichas técnicas incluyen determinar la secuencia de nucleótidos del ARNm para un gen y/o determinar la secuencia de aminoácidos codificada por ésta, y comparar estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también pueden ser determinadas y comparadas de este modo. En general, identidad se refiere a una correspondencia exacta de nucleótido a nucleótido o de aminoácido a aminoácido de dos secuencias de polinucleótidos o polipéptidos, respectivamente. Dos o más secuencias (polinucleótido o aminoácido) pueden ser comparadas por determinación de su porcentaje de identidad. El porcentaje de identidad de dos secuencias, ya sea secuencias de ácido nucleico como de aminoácidos, es el número de coincidencias exactas entre dos secuencias alineadas dividido por la longitud de las secuencias más cortas y multiplicado por 100. Una alineación aproximada para las secuencias de ácido nucleico es proporcionada por el algoritmo de homología de Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). Este algoritmo puede ser aplicado a las secuencias de aminoácidos por el uso de 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. Acids Res.* 14(6):6745-6763 (1986). Una implementación de ejemplo de este algoritmo para determinar el porcentaje de identidad de una secuencia es proporcionada por 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 alineación es BLAST, usado con parámetros por defecto. Por ejemplo, BLASTN y BLASTP pueden ser usados usando los siguientes parámetros por defecto: código genético=estándar; filtro=ninguno; hebra=ambas; corte=60; expectativa=10; Matriz=BLOSUM62; Descripciones=50 secuencias; clasificar por=HIGH SNÚCLEO; Bases de datos=no redundantes, traducciones GenBank+EM- BL+DDBJ+PDB+GenBank CDS+proteína Swiss+Spupdate+PIR. Los detalles de estos programas pueden ser encontrados en la página en internet del GenBank.

Con respecto a las secuencias descritas en la presente memoria, el intervalo de grados deseados de identidad de secuencia es de aproximadamente 80% a 100% y cualquier valor entero entre estos. Normalmente, los porcentajes de identidad entre las secuencias son al menos del 70-75%, preferentemente 80-82%, más preferentemente 85-90%, incluso más preferentemente 92%, aún más preferentemente 95%, y mucho más preferentemente 98% de identidad de secuencia.

Dado que diversos cambios pueden ser realizados en las células descritas con anterioridad y en los procedimientos sin alejarse del ámbito de la invención, se pretende que el tópicamente contenido en la descripción que antecede y en los ejemplos que presentados a continuación, sea interpretado como ilustrativo y no en un sentido limitativo.

Ejemplos

Los siguientes ejemplos ilustran ciertos aspectos de la invención.

Ejemplo 1: Preparación de las líneas celulares de CHO modificadas genéticamente diana

Fueron empleadas técnicas de modificación genética mediada por ZFN para inactivar (es decir, bloquear) genes que codifican enzimas o proteínas involucradas en las reacciones de glicosilación N-ligadas u O-ligadas. En general, los pares de AFN localizados en sitios específicos dentro de la región de codificación de los genes de interés son diseñados usando un algoritmo exclusivo. Los constructos de expresión de ZFN fueron preparados usando procedimientos estándar y el ARNm de ZFN fue producido a partir de ADN del plásmido de ZFN como es descrito en la información del producto COMPOZr® Knockout Zinc Finger Nuclease (ZFN) (Sigma-Aldrich) usando transcripción *in vitro*, poliadrenilación de ARNm y procedimientos de tapado. Brevemente, el ADN del plásmido ZFN fue linealizado y purificado usando extracción de ADN con fenol/cloroformo. El MessageMax™ T7 ARCA-Capped Message Transcription Kit ARCA (Cell Script Inc.) fue usado para tapar el ADN linealizado. Un Poly(A) Polymerase Tailing Kit (EpiCentre) fue usado para añadir una cola de poli(A). El ARNm de ZFN fue purificado usando el kit MEGAclean™ (Ambion). Las células progenitoras fueron mantenidas como cultivos en suspensión en medio EX-CELL® CHO CD Fusion (Sigma-Aldrich). Las células fueron sembradas a $0,5 \times 10^6$ células/ml en tubos de biorreactor un día antes de la transfección. Normalmente, cada transfección contenía 1×10^6 células en 150 µg de medios de crecimiento y 5 µg de ADN o ARNm de ZFN. Las transfecciones fueron conducidas por electroporación a 140 V y 950 mF en cubetas de 0,2 cm. Las células electroporadas fueron colocadas en 2 ml de medio de crecimiento en un cultivo estático en placa de 6 pocillos.

Los días 3 y 10 después de la transfección, las células son retiradas del cultivo y el ADN genómico es aislado usando el GeneElute™ Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich). La escisión inducida por ZFN fue verificada usando un ensayo de nucleasa Cel-1, como es descrito en la información del producto CompoZr® Knockout ZFN. Este ensayo es conducido para determinar la eficiencia de la mutación genética mediada por ZFN descrita con anterioridad (Miller *et al.*, Nat. Biotechnol. 2007, 25:778-785). El ensayo detecta los alelos del locus específico que se desvían del tipo salvaje como resultado de la reparación imperfecta mediada por unión de extremo no homólogo (NHEJ) de las rupturas de hebra doble del ADN inducidas por ZFN. La amplificación por PCR de la región especificada de un conjunto de células tratadas con ZFN genera una mezcla de amplicones de tipo salvaje (WT) y mutantes. La fusión y el reemplazo de esta mezcla resultan en faltas de coincidencia que se forman entre los heterodúplex de los alelos WT y mutantes. Una "burbuja" de ADN formada en este sitio de falta de coincidencia es escindida por la nucleasa sobreviviente Cel-1, y los productos de escisión pueden ser resueltos por electroforesis con gel.

Tras la confirmación de la actividad de ZFN, las células transfectadas con ZFN son clonadas en célula individual usando dilución limitativa. Para esto, las células fueron colocadas en placas a una densidad de aproximadamente 0,5 células/pocillo usando una mezcla de medio de clonación libre de suero de CHO al 80%, 20% medio acondicionado y L-glutamina 4 mM. La capacidad de clonación y el crecimiento fueron verificados microscópicamente los días 7 y 14 después de la plantación, respectivamente. Los clones con crecimiento fueron expandidos y el genotipo fue determinado por PCR y secuenciación de ADN.

Línea celular KO de Mgat1. Mgat1 añade GlnNac a la estructura de glicano Man5GlcNAc2 N-ligado como parte de la síntesis compleja de N-glicano. Un par de ZFN fue diseñado para localizar 5'-TTCCTGGACCACTTCCCaccgggt-GGCCGGCAGGATGG-3' (SEC ID NÚM.: 9; los sitios de unión a ZFN en el caso superior y el sitio de escisión en el caso inferior) en el *gen Mgat1Mgat1* de CHO. La línea celular progenitora es CHOK1 (GS -/-) (Sigma Aldrich) en la que la glutamato sintetasa es bloqueada. Esta línea celular progenitora produce estructuras N- y O-glicano de tipo salvaje. Las líneas celulares de CHOK1 (GS -/-) son transfectadas con ADN *Mgat1Mgat1* de ZFN esencialmente como es detallado con anterioridad. Un ensayo de Cel-1 confirma la presencia de dos fragmentos de escisión, 220 y 197 pb, en ambos días 3 y 10 después de la transfección, indicando actividad de ZFN. Los clones de células individuales son identificados con eliminaciones que oscilan de 2 pb a 55 pb en un alelo (un segundo alelo no fue detectado). Esta línea celular produce estructuras N-ligadas truncadas que terminan con cinco residuos de manosa (es decir, glicofomas Man5NeuAc2) y estructuras de O glicano de tipo salvaje.

Línea celular KO COSMC. La segunda etapa biosintética más común en la O-glicosilación es la estiramiento núcleo-1. La estiramiento núcleo-1 es catalizado por una enzima individual C1GalT1 (también denominada T-sintasa) que requiere una chaperona exclusiva, COSMC, para funcionar. COSMC es una proteína de ER que aparenta unirse

específicamente a T-sintasa y asegura su actividad completa en el Golgi. Para separar el gen COSMC, las células CHOK1 (GS -/-) son transfectadas con ADN o ARN que codifican ZFN diseñadas para localizar la secuencia 5'-GCCTTCTCAGTGTCCGGAAaagtgTCCTGAACAAGGTGGGAT-3' (SEC ID NÚM.: 10) en el gen COSMC de CHO. Un ensayo de nucleasa A Cel-1 confirma la presencia de tres fragmentos de escisión (es decir, 318, 184, 134 pb) en las células transfectadas con ZFN. Fue aislado un clon de célula individual que tenía una eliminación de 4 pb en un alelo de COSMC (un segundo alelo no fue detectado). La línea celular produce estructuras de O-glicano truncadas (es decir, inmaduras) y N-glicanos de tipo salvaje.

Línea celular KO COSMC/Mqat3. Un par de ZFN fue diseñado para localizar 5'-TTCCTGGACCACTTCCCaccggt-GGCCGGCAGGATGGC-3' (SEC ID NÚM.: 11) en la región de codificación de *Mgat3*. La línea celular COSMC KO, detallada anteriormente, es transfectada con ADN de ZFN como es detallado con anterioridad. Después de la confirmación de la escisión del ZFN, son aislados clones de células individuales. La secuenciación revela que los clones mutantes tienen eliminaciones de 9, 10, 11, o 41 pb en el gen *Mgat3*. Esta línea celular produce N-glicanos de tipo salvaje y O-glicanos truncados (es decir, es similar a la línea celular progenitora).

Línea celular KO COSMC/Mqat3/Mqat5. La línea celular KO *COSMC/Mgat3*, detallada con anterioridad, es transfectada con ADN de plásmido que codifica ZFN diseñados para localizar 5'-TTCTGCACTTCACCATCCAgcagcgGACTCAGCCTGAGAG-CAGCT-3' (SEC ID NÚM.: 12) en la región de codificación de *Mgat5*. Es aislado un clon de célula individual que tiene una eliminación de 129 pb en el gen *Mgat5*. Esta línea celular produce N-glicanos con menos ramificación lateral y O-glicanos truncados.

Ejemplo 2: Infecciones virales y ensayos de prueba de la resistencia viral

Las líneas celulares de CHO descritas con anterioridad son cultivadas y probadas para determinar la capacidad de soportar o resistir la infección después de la provocación con el virus de MVM prototipo (cepa MVMp). Brevemente, las células son cultivadas en medios adecuados y el virus MVMp es detectado a una multiplicidad de infección (MOI) de cualquiera de 1 o 10, y las células infectadas incubadas durante un adicional de 24, 48 o 72 horas antes del ensayo. Como controles, las células son cultivadas e incubadas sin virus. Las células no infectadas son tratadas también con neucaminidasa enzimática para retirar ácidos siálicos terminales de los glicanos superficiales (tanto N- como O-ligados) y posteriormente infectados a las MOI indicadas.

Las células infectadas y no infectadas son examinadas visualmente para determinar la presencia del efecto citopático (CPE) y en los puntos de tiempo indicados, las células son cosechadas por centrifugado. La infectividad y la producción de ADN viral es seleccionada usando análisis por Southern blot sobre el ADN genómico total, y análisis de inmunotransferencia Western usando un anticuerpo de proteína antiviral. Para el análisis Southern, gránulos celulares de muestras duplicadas tomadas a las 24 horas fueron cosechadas y el ADN genómico total aislado. El ADN fue cuantificado por espectroscopía Nanodrop y las muestras fueron normalizadas para asegurar la carga equitativa sobre geles de agarosa para fraccionamiento de tamaño. El ADN con tamaño fraccionado fue transferido a membranas cargadas (Southern blot), y la membrana fue sondeada para síntesis de ADN viral usando una sonda de ADN viral con marca 32P. La cuantificación de las bandas de ADN de hebra doble viral con marca 32P fue adquirida por imágenes de fósforo y los valores relativos son informados en la Tabla 1.

Tabla 1. Niveles virales

Línea Celular	MOI	Tratamiento con neuroaminidasa	ADN viral (valor relativo)
Tipo salvaje	10		133,35
Tipo salvaje	10	Sí	25,59
<i>Mgat1</i> KO <i>Mgat1</i>	10		49,56
<i>COSMC</i> KO	10		24,01
Tipo salvaje	10		121,36
Tipo salvaje	10	Sí	31,38
<i>COSMC/Mgat3</i> KO	10		23,78
<i>COSMC/Mgat3/Mgat5</i> KO	10		28,87
Tipo salvaje	1		36,28
Tipo salvaje	1	Sí	4,16
<i>Mgat1</i> KO	1		17,263
<i>COSMC</i> KO	1		1,58
Tipo salvaje	1		85,61
Tipo salvaje	1	Sí	11,65
<i>COSMC/Mgat3</i> KO	1		0,45
<i>COSMC/Mgat3/Mgat5</i> KO	1		0,54

5 Para el análisis Western, los gránulos celulares cosechados a las 24 horas fueron lisados en tampones de SDS, y las proteínas fueron separadas usando SDS-PAGE. Después de la electroforesis, las proteínas fueron transfectadas sobre membranas de PVDF, fueron bloqueadas e inmunotransferidas (inmunoensayo Western) para determinar la presencia de la proteína NS1 viral usando un anticuerpo anti-NS1. El nivel de proteína viral detectada a través de Western blot es indicado en la Tabla 2.

Tabla 2. Resumen del análisis Western

Línea celular	Estructuras de glicano producidas	Producción de virus
Tipo salvaje	Tipo salvaje O- y N-ligado	+++++ en ambos MOI
Tipo salvaje + neuraminidasa	Retirado de ácido siálico terminal (O- y N-ligado)	+ en MOI=10
<i>Mgat1</i> KO	Tipo salvaje O-ligado; Truncado N-ligado	+++ en ambos MOI
<i>COSMC</i> KO	Tipo salvaje N-ligado; Truncado O-ligado	- en MOI=1 + en MOI=10
<i>COSMC/Mgat3</i> KO	Truncado O-ligado; Supuestamente tipo salvaje	- en MOI=1 + en MOI=10
<i>COSMC/Mgat3/Mgat5</i>	Truncado O-ligado; Modificado N-ligado (menos ramificación lateral)	- en MOI=1 + en MOI=10

Cuando está infectada con la cepa MVMp del virus, la línea celular progenitora CHOK1 (GS -/-) (es decir, con las estructuras de glicano tipo salvaje) muestra infectividad clásica y producción de ADN viral y síntesis de proteínas

según lo esperado. Los productos de ADN viral y proteína viral son evidentes 24 horas después de la infección cuando las células están infectadas a cualquier MOI. La producción fuerte de productos virales también es observada en los últimos puntos de tiempo. Las células de tipo salvaje no infectadas no muestran evidencia de infección, según lo esperado, independientemente de los puntos de tiempo.

5 Cuando las células de tipo salvaje son tratadas con la neuraminidasa (NA) enzimática para eliminar el ácido siálico de la superficie celular y las células posteriormente infectadas con MVMP, una producción reducida de virus (que indica la reducción de la infección) también es observada (oscilando de 5,32 veces a MOI=10 a 7,34 veces a MOI=1). Mientras que las células tratadas con NA aún muestran sensibilidad leve a la infección, es conocido que las células CHO regeneran rápidamente las estructuras de ácido siálico (SA) después de dicho tratamiento. Por consiguiente, el nivel bajo de infección observado puede ser de células con glicanos de SA terminal regenerados y por lo tanto proporcionan entrada del virus.

15 La infección viral de células KO *Mgat1* en las cuales el gen *Mgat1* de CHO es inactivado y no ha sido realizada enzima *Mgat1*, resultó solo en una resistencia leve (por ej., reducción de 2,1 a 2,7 veces, medida por imágenes de fósforo de transferencias) a la infección viral. Debido a que esta línea celular tiene N-glicanos truncados, estos resultados sugieren que los receptores de N-glicano de orden superior que tienen 2-3 ácidos siálicos ligados pueden cumplir solo una función menor en la unión a la cápsida viral inicial y la entrada viral a la célula.

20 Cuando las células KO COSMC son infectadas con virus MVMP, esta línea celular muestra resistencia significativa a la infección viral. La resistencia al plegado (en comparación con la línea celular de tipo salvaje) oscila de 5,5 veces (MOI=10) a 190 veces (MOI=1). El bloqueo genético adicional localizado en el gen *Mgat3* (supuesto pseudogen en células CHO) y el gen *Mgat5* (responsable de ramificación N-ligada de orden superior), dio resultados similares. El análisis por Western blot dio resultados similares, indicando que la resistencia viral ocurre cuando las estructuras de O-glicano fueron truncadas. Estos estudios revelaron que la separación de los receptores de ácido siálico redujo drásticamente la unión de MVM y/o la entrada a las células.

Ejemplo 3: Generación de líneas celulares COSMC KO y Slc35A1 KO adicionales

25 Nuevos clones de la línea celular COSMC KO son generados como es descrito anteriormente en el Ejemplo 1 usando células CHOK1 (GS *-/-*) y un par de ZFN diseñados para orientarse a la secuencia 5'-GCCTTCTCAGTGTCCGGAaaagtTCCTGAACAAGGTGGGAT-3' (SEC ID NÚM.: 10) en el exón 2 del gen COSMC de CHO. Un ensayo de nucleasa A Cel-1 confirma la presencia de tres fragmentos de escisión (es decir, 318, 184, 134 pb) en las células transfectadas con ZFN. Cinco clones de células individuales son aislados y la secuenciación revela eliminaciones de 1 a 12 pb, como es mostrado a continuación.

Clon	Genotipo
COSMC F07	Eliminación de 2 pb
COSMC G03	Inserción de 1 pb
COSMC H04	Inserción de 1 pb
COSMC A09	Eliminación de 9 pb
COSMC H05	Eliminación de 12 pb

El clon F07 fue modificado de manera adicional para expresar una IgG humana usando procedimientos estándar. De los numerosos clones identificados, dos clones que producen IgG son aislados (e identificados como 74H1, 71C3) para análisis adicional (véase a continuación).

35 El gen *Slc35A1*, que codifica un transportador de azúcar de nucleótido (transportador de CMP-ácido siálico), es eliminado en células CHOK1 (GS *-/-*) usando ZFN diseñados para localizar 5'-AGCTTATACCGTAGCTTTaagataCACAAAG-GACAACAGCTAAA-3' (SEC ID NÚM.: 13) o ZFN fue diseñado para dirigirse a 5'-TTCAAGCTATACTGCTTGGCAGTGAT-GACTCTGGTGGCT-3' (SEC ID NÚM.: 17) en el exón 1 del gen *Slc35A1* de CHO. Un ensayo de nucleasa A Cel-1 confirma la presencia de dos fragmentos de escisión en las células transfectadas con ZFN. Un clon de célula individual (B12) es aislado y la secuenciación revela una eliminación de 1 pb alrededor del sitio de unión a ZFN. La línea celular *Slc35A1* KO forma estructuras glicano N y O ligadas sin ácido siálico terminal.

45 La tinción con lectina II de *Maackia Amurensis* biotinilada (MALII; 20 µg/ml) y estreptavidina marcada con Alexa Fluor 647 (5 µg/ml) revela tinción reducida significativamente en las líneas celulares del clon F07 COSMC KO, clon G03 COSMC KO, y *Slc35A1* KO en comparación con la línea celular progenitora (indicando ausencia de residuos de ácido siálico terminales en las líneas celulares KO).

Ejemplo 4: Resistencia de las líneas celulares COSMC KO y Slc35A1 KO al virus MMV

Células de tipo salvaje (es decir, CHOZN GS-/-), COSMC KO (generadas con anterioridad en los Ejemplos 1 y 3), y Slc35A1 KO (generada arriba en el Ejemplo 3) fueron infectadas con virus MMVp a una MOI de 0,3 esencialmente como se describió anteriormente en el Ejemplo 2. La infección fue analizada a través del análisis Southern (esencialmente como se describió con anterioridad) y ensayos estándar en placa.

Como es mostrado en la FIG. 1, las células de tipo salvaje muestran niveles altos de ADN, mientras que las líneas celulares Slc35A1 KO y COSMC KO tienen niveles muy bajos de ADN viral; los clones COSMC F07 KO y COSMC G03 KO tienen niveles bajos de ADN viral; y el clon COSMC H05 FO, que tiene una eliminación de 12 pb, tiene niveles de ADN viral similares a la línea celular progenitora. Dado que la eliminación en H05 estaba en marco, no es sorprendente que las células exhibieran sensibilidad viral a niveles cercanos al tipo salvaje. Los resultados del ensayo de placas se muestran en la FIG. 2. Las células Slc35A1 KO, COSMC KO, COSMC F07 KO, y COSMC G03 KO tenían niveles virales extremadamente bajos.

El análisis adicional de la resistencia viral es realizado usando clones F07 de COSMC KO de tipo salvaje (es decir, CHOZN GS-/-, también denominados 2E3), y G03, y clon B12Slc35A1 KO. Son cultivadas células en medio adicionado con L-glutamina 6 mM e infectadas con MMVp a una MOI de 1 u 8 e incubadas en condiciones adecuadas. Las muestras de células son retiradas a las 0, 24, 48, 72, 96, y 120 horas; la viabilidad celular es evaluada con tinción azul tripano y la cuantificación de MMV fue determinada por PCRq.

La viabilidad celular es presentada en la FIG. 3. 120 horas después de la infección, la viabilidad celular es reducida hasta aproximadamente 50% en células de tipo salvaje (es decir, 2E3) (véase el panel A), mientras que en el clon F07 de COSMC al menos el 93% de las células eran viables a las 120 horas después de la infección (véase el panel B). La FIG. 4 muestra el curso temporal del virus asociado a las células (copia del genoma del virus (vgc) por célula) después de la infección con MMV en el clon F07 tipo salvaje (2E3) y COSMC a una MOI de 1 (véase el panel A) o MOI de 8 (véase el panel B) (el cálculo está basado en la suposición de que cada célula infectada puede producir 2×10^4 vgc). La fracción de células infectadas en cada condición es presentada a continuación.

Tabla 3. Porcentaje de células infectadas

MOI	Tipo de célula	Repetición	0 hs.	24 hs.	48 hs.	72 hs.	96 hs.	120 hs.
			% de células					
1	2E3	A	44	21	28	36	24	55
		B	41	27	26	30	32	14
	F07	A	5	2,1	1,7	1,6	0,5	1,9
		B	6	2,2	2,0	1,4	0,7	2,5
8	2E3	A	115	70	91	42	65	118
		B	105	85	113	55	60	90
	F07	A	32	14	12	6,9	1,1	5,5
		B	27	14	11	3,4	4,8	7,5

Células de los clones F07, G03 y H04 tipo salvaje (es decir, 2E3), COSMC KO, y el clon B12 de Slc35A1 KO son infectadas en el punto de tiempo de -3 horas con MMVp a una MOI de 0,3 o 0,03 indistintamente. En el punto de tiempo de 0, las células son lavadas tres veces con medio cultivado y luego son incubadas durante 21 horas (es decir, un solo ciclo de replicación). La replicación viral fue analizada por determinación del virus asociado a las células a las 0 y 21 horas. Como se muestra en la FIG. 5, la replicación viral fue reducida en las líneas celulares COSMC KO y eliminada en la línea celular Slc35A1 KO.

Ejemplo 5: Resistencia de las líneas celulares COSMC KO y Slc35A1 KO al Reovirus-3

Células de tipo salvaje (es decir, 2E3), clon F07 de COSMC KO, y Clon B12 de Slc35A1 KO son infectadas en el punto de tiempo de -3 horas con reovirus-3 en dos diluciones ($TCID_{50} = 5,6E+07$ y $5,6E+06$). En el punto de tiempo de 0, las células son lavadas tres veces y luego son incubadas durante 24 horas (es decir, un solo ciclo de replicación). La replicación viral es analizada por determinación del virus asociado a las células a las 0 y 24 horas. Los niveles de virus asociados a las células son reducidos drásticamente a las 24 horas en el clon de COSMC KO y casi eliminados en el clon de Slc35A1 KO (véase la FIG. 6).

Ejemplo 6: Ensayos de crecimiento de líneas celulares COSMC KO y Slc35A1 KO adicionales

Células de tipo salvaje (es decir, CHOZN GS^{-/-}), clones F07 y G03 de COSMC KO, y el clon B12 de Slc35A1 KO son cultivadas en ausencia o presencia de virus MVMP (a una MOI de 0,1) durante 8 a 10 días. El crecimiento celular es monitorizado por medición de la densidad de células viables (a través de tinción con azul tripano) los días 0, 1, 2, 3, 5, 6, 7, 8 y 10. Como es mostrado en el panel A de la FIG. 7, las varias células KO tuvieron perfiles similares de crecimiento en presencia o ausencia de virus, indicando resistencia a la inyección viral. En contraste, las células de tipo salvaje muestran infectividad clásica, crecimiento dañado y bajas viabilidades celulares en comparación con los cultivos no infectados.

El cultivo celular de tipo salvaje y los clones que producen IgG del clon F07 de COSMC KO (es decir, 71H1 y 71C3) es monitorizado en ausencia o presencia de virus esencialmente según lo detallado con anterioridad. Como es mostrado en el panel B de la FIG. 7, las líneas celulares KO que producen IgG también muestran resistencia significativa a la infección viral. Los cultivos infectados crecen con índices similares y hasta una densidad celular pico como los cultivos no infectados. Estos datos indican que secretar una IgG de la superficie celular no tiene impacto discernible sobre el grado de resistencia viral exhibida por el progenitor COSMC KO. Aunque ambas líneas celulares que producen IgG crecen con índices más lentos en comparación con las células de tipo salvaje, ambos productores de IgG muestran resistencia a la infección viral (es decir, sin diferencia en comparación con los cultivos no infectados).

Ejemplo 7: Generación de la línea celular St3Gal4 KO y ensayos de crecimiento celular

Las células St3Gal4 KO son generadas esencialmente como fue descrito con anterioridad en el ejemplo 1 usando ZFN diseñados para localizar 5'-GGCAGCCTCCAGTGTCTGTC gttgtgTTGTGGTGGGGAATGGGC (SEC ID NÚM.: 14) en el gen St3Gal4 de CHO. Un ensayo de nucleasa A Cel-1 confirma la presencia de tres fragmentos de escisión (es decir, 344 pb, 210 pb, y 135 pb) en las células transfectadas con ZFN. Cuatro clones celulares individuales son aislados y la secuenciación revela las siguientes mutaciones. Las células St3Gal4 KO tienen niveles reducidos de estructuras de ácido siálico.

Clon	Genotipo del alelo 1	Genotipo del alelo 2
<i>St3Gal4</i> 1B08	Inserción de 4 pb	Eliminación de 5 pb
<i>St3Gal4</i> 7D10	Eliminación de 2 pb	Eliminación de 8 pb
<i>St3Gal4</i> 1B10	Eliminación de 5 pb	Eliminación de 11 pb
<i>St3Gal4</i> 1H11	Inserción de 3 pb	Eliminación de 11 pb

Las células St3Gal4 KO (es decir, clones 7D10, 1B8 y 1B10) y de tipo salvaje son cultivadas en ausencia o presencia de virus MVMP a una MOI de 0,1 y el cultivo celular es monitorizado durante 9 días. Las células St3Gal4 KO también muestran resistencia significativa a la infección viral como es mostrado en la FIG. 8. Los cultivos KO infectados crecen con índices similares y hasta una densidad celular pico como los cultivos KO no infectados. Estos estudios revelan que la interrupción de la estructura específica de ácido siálico 2-3 ligada sobre la superficie celular también reduce drásticamente la entrada de MVM en dichas células.

Ejemplo 8: Generación de la línea celular doble KO St3Gal4 y St3Gal6

Las líneas celulares *St3Gal4* KO (es decir, clones 7D10, 1B8 y 1B10) son usadas como células iniciales para generar líneas celulares que también contienen un gen *St3Gal6* bloqueado. Los ZFN son diseñados para localizar 5'-CGGTACCTCTGATTTTGCTTgcccCTATGGGACAAGGCC-3' (SEC ID NÚM.: 15) y la edición genética es realizada esencialmente según lo descrito en el Ejemplo 1. Un ensayo de nucleasa A Cel-1 confirma la presencia de tres fragmentos de escisión (es decir, 308 pb, 171 pb, y 137 pb) en las células transfectadas con ZFN. Seis clones celulares individuales son aislados y la secuenciación revela las siguientes mutaciones.

Clon <i>St3Gal6</i>	Genotipo	Derivado del clon <i>St3Gal4</i>
1E8	Dos eliminaciones de 8 pb diferentes + sustitución de 1 pb	7D10
1H8	Eliminación de 8 pb + sustitución de 1 pb	7D10
2D12	Inserción de 4 pb + eliminación de 2 pb	1B08
4F7	Dos eliminaciones de 2 pb diferentes	1B08
2F11	Eliminación de 8 pb + sustitución de 1 pb y eliminación de 13 pb	7D10
3C1	Eliminación de 8 pb + sustitución de 1 pb y eliminación de 8 pb	7D10

Ejemplo 9: Generación de líneas celulares KO de *C1GalT1*

- 5 El gen *C1GalT1* es bloqueado en células CHOK1 (GS -/-) usando ZFN diseñados para localizar 5'-ACCCTCAT-GCTAGACattaGATGATAACGAACCCAGTC-3' (SEC ID NÚM.: 16) esencialmente como es descrito en el Ejemplo 1. El ensayo de nucleasa Cel-1 confirma la presencia de dos fragmentos de escisión (es decir, 223 pb y 148 pb) en las células transfectadas con ZFN. Siete clones celulares individuales son aislados y la secuenciación revela las siguientes mutaciones.

Clon <i>C1GalT1</i>	Genotipo del alelo 1	Genotipo del alelo 2
2A8	Eliminación de 8 pb	Eliminación de 8 pb
2A11	Eliminación de 5 pb	Inserción de 90 pb
2C12	Eliminación de 8 pb	Ninguna detectada
2G12	Eliminación de 5 pb	Ninguna detectada
3D12	Eliminación de 8 pb	Ninguna detectada
3E6	Eliminación de 1 pb	Eliminación de 2 pb e inserción de 97 pb
3F2	Eliminación de 5 pb	Eliminación de 5 pb e inserción de 31 pb

- 10 La tinción con MALII biotinilado y estreptavidina marcada con Alexa Fluor 647 revelan tinción significativamente reducida en el clon 2C12 de *C1GalT1* KO en comparación con la línea celular progenitora.

Ejemplo 10: Generación de líneas celulares que sobreexpresan *St6Gal1*

- 15 Dado que se llega a la hipótesis de que el virus MVM no se une a los ácidos siálicos ligados a alfa-2,6, son generadas líneas celulares de CHO que sobreexpresan *St6Gal1*. La secuencia de codificación de *St6Gal1* de hámster chino es obtenida del GenBank (AB492855) y de chogenome.org AQ2 (AFTD01061789 y AFTD01061790). El marco abierto de lectura es sintetizado comercialmente con una secuencia Kozak (5'-GCCGCCACCAatg-3'; SEC ID NÚM.: 18) añadida a la región 5'-no traducida (UTR). El fragmento sintetizado es clonado en el vector de expresión pJ602 (DNA2.0; Menlo Park, CA). La línea celular huésped CHO (GS-/-) y una línea celular de CHO que expresa IgG son transfectadas (por electroporación) con este constructo. Clones de células individuales son aislados usando un clasificador celular FACSAria™ III. Para esto, las células se tiñeron con lectina *Sambucus Nigra* conjugada con FITC (FITC-SNA), que se une al ácido siálico alfa-2,6 ligado, y las células con un máximo de 5% de fluorescencia son colocadas en placas a razón de 1 célula/pocillo y son cultivadas. Los clones de células individuales son sometidos a análisis FACS de dos colores en un Analizador MACSQuantVR (Miltenyi Biotec, San Diego, CA) tras la tinción con
- 20 FITC-SNA), que se une al ácido siálico alfa-2,6 ligado, y estreptavidina marcada con Alexa Fluor 647. Los clones de células individuales que sobreexpresan *St6Gal1* presentan mayores relaciones de FITC a AlexaFluor en comparación con las líneas celulares progenitoras no transfectadas.
- 25

La IgG es aislada de dos clones que sobreexpresan *St6Gal1* (es decir, el clon 31 y el clon 32) y el clon de la célula

- progenitora. Los extractos de proteína celular IgG o total son reducidos y carboxiamidometilados de acuerdo con procedimientos estándar antes de la tripsinización a 37°C durante la noche (12-16 h). La tripsina es desactivada por calentamiento a 100°C durante 5 min. La purificación de los fragmentos es llevada a cabo con un cartucho C18 SPE (Waters, envase de 300 mg). Después de un lavado con ácido acético al 5% (AcOH), los péptidos/glicopéptidos son eluidos consecutivamente en 20% de isopropanol/5% de AcOH, 40% de isopropanol/5% de AcOH, y 100% de isopropanol. El eluyente es secado, reconstituido en tampón de fosfato que contiene PNGasa F, e incubado a 37°C durante la noche. Los glicanos liberados son purificados usando un cartucho C18 permetilado y diluido en carbonato de litio 1 mM/50% MeOH e infundido directamente en un Espectrómetro de masa LTQ Orbitrap Discovery (Thermo Scientific) a un caudal de 0,5 ml/min para ionización por nanoaspersión. Un espectro de espectrometría de masa por transformada de Fourier (FTMS) completo es obtenido a una resolución de 30.000 para cada muestra para determinar qué glicanos contenían ácidos siálicos. La ligadura de ácido siálico de los glicanos sialilados es determinada por análisis de MSn. Cada muestra es sometida a selección de iones múltiples y etapas de fragmentación dentro de la trampa de iones para ruptura de un glicano del tipo complejo hasta obtener una galactosa individual y después el patrón de fragmentación es observado.
- Hubo diferencias sorprendentes entre los perfiles de N-glicano de la línea celular huésped GS (-/-) y las dos líneas celulares que sobreexpresan *St6Gal1* (es decir, los clones 31 y 32). En particular, no son identificados glicanos sialilados en la línea celular huésped, debido a su baja abundancia. En contraste, los N-glicanos mono-, bi-, tri-, y tetrasialilados son identificados en ambas líneas celulares que sobreexpresan *St6Gal1*. En particular, todas las poblaciones excepto una población sialilada en el clon 32 demuestran que contienen ambas ligaduras alfa-2,3 y alfa-2,6.

Listado de secuencias

- <110> SIGMA-ALDRICH CO. LLC
 LIN, Nan
 MASCARENHAS, Joaquina
- 25 CHANG, Audrey
 ONIONS, David
 GEORGE, Henry
 KAYSER, Kevin
- <120> CÉLULAS VIRALES RESISTENTES Y SUS USOS
- 30 <130> 047497-469443
 <150> US 61/947.860
 <151> 04/03/2014
 <160> 18
 <170> PatentIn version 3.5
- 35 <210> 1
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
- 40 <223> SINTETIZADA
 <400> 1

Pro Lys Lys Lys Arg Lys Val
 1 5

- <210> 2
 <211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> SINTETIZADA

5 <400> 2

Pro Lys Lys Lys Arg Arg Val
1 5

<210> 3

<211> 16

<212> PRT

10 <213> Secuencia artificial

<220>

<223> SINTETIZADA

<400> 3

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

15 <210> 4

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

20 <223> SINTETIZADA

<400> 4

Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Pro Lys Lys
1 5 10 15

Lys Arg Lys Val
20

<210> 5

<211> 19

25 <212> PRT

<213> Secuencia artificial

<220>

<223> SINTETIZADA

<400> 5

ES 2 797 050 T3

Pro Leu Ser Ser Ile Phe Ser Arg Ile Gly Asp Pro Pro Lys Lys Lys
1 5 10 15

Arg Lys Val

<210> 6

<211> 24

<212> PRT

5 <213> Secuencia artificial

<220>

<223> SINTETIZADA

<400> 6

Gly Ala Leu Phe Leu Gly Trp Leu Gly Ala Ala Gly Ser Thr Met Gly
1 5 10 15

Ala Pro Lys Lys Lys Arg Lys Val
20

10 <210> 7

<211> 27

<212> PRT

<213> Secuencia artificial

<220>

15 <223> SINTETIZADA

<400> 7

Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly
1 5 10 15

Ala Trp Ser Gln Pro Lys Lys Lys Arg Lys Val
20 25

<210> 8

<211> 21

20 <212> PRT

<213> Secuencia artificial

<220>

<223> SINTETIZADA

<400> 8

Lys Glu Thr Trp Trp Glu Thr Trp Trp Thr Glu Trp Ser Gln Pro Lys
1 5 10 15

Lys Lys Arg Lys Val
20

25

<210> 9
 <211> 39
 <212> ADN
 <213> Cricetulus griseus
 5 <400> 9
 aacaagtca agttccagc agctgtgga gtggaggac 39
 <210> 10
 <211> 43
 <212> ADN
 10 <213> Cricetulus griseus
 <400> 10
 gcctctcag tgttcggaa aagtgtctg aacaaggtg gat 43
 <210> 11
 <211> 39
 15 <212> ADN
 <213> Cricetulus griseus
 <400> 11
 ttctggacc acttccaacc cggtgcoog caggatggc 39
 <210> 12
 20 <211> 45
 <212> ADN
 <213> Cricetulus griseus
 <400> 12
 ttctgactt caccatcag cagcggactc agcctgagag cagct 45
 <210> 13
 <211> 43
 <212> ADN
 <213> Cricetulus griseus
 <400> 13
 agctatacc gtagcttaa gatacacaag gacaacagct aaa 43
 30 <210> 14
 <211> 43
 <212> ADN
 <213> Cricetulus griseus
 35 <400> 14

ggcagcctcc agtgcgtcg ttgtgttg gggggaatg ggc 43

<210> 15

<211> 39

<212> ADN

5 <213> Cricetulus griseus

<400> 15

cggtacctt gatttgctt tgcctatgg gacaaggcc 39

<210> 16

<211> 39

10 <212> ADN

<213> Cricetulus griseus

<400> 16

acctcatgc tagacattta gatgataacg aaccagtc 39

<210> 17

15 <211> 39

<212> ADN

<213> Cricetulus griseus

<400> 17

ttcaagctat actgcttggc agtgatgact ctggtgct 39

20 <210> 18

<211> 13

<212> ADN

<213> Secuencia artificial

<220>

25 <223> SINTETIZADA

<400> 17

gocgocacca atg 13

30

REIVINDICACIONES

- 5 **1.** Una línea celular de ovario de hámster chino (CHO) modificada genéticamente en la que la entrada y/o propagación del virus diminuto del ratón (MVM) es reducida o eliminada en comparación con una línea celular progenitora no modificada, en la que la línea celular modificada genéticamente comprende una secuencia cromosómica modificada que comprende una secuencia cromosómica inactivada que codifica la St3 beta-galactósido alfa-2,3-sialiltransferasa 4 (St3Gal4).
- 10 **2.** La línea celular de mamífero modificada genéticamente de la reivindicación 1, que además comprende una secuencia cromosómica inactivada que codifica la (alfa-1,3-manosil)-glicoproteína beta-1,2-N-acetilglucosaminiltransferasa 1 (Mgat1).
- 15 **3.** La línea celular de mamífero modificada genéticamente de la reivindicación 1 o 2, en la que la secuencia cromosómica modificada es modificada usando una técnica de modificación del genoma mediada por endonucleasa diana.
- 20 **4.** La línea celular de mamífero modificada genéticamente de la reivindicación 3, en la que la endonucleasa diana es una nucleasa con dedos de cinc.
- 25 **5.** La línea celular de mamífero modificada genéticamente de una cualquiera de las reivindicaciones 1 a 4, que además comprende al menos un ácido nucleico que codifica una proteína recombinante seleccionada de un anticuerpo, un fragmento de anticuerpo, una vacuna, un factor de crecimiento, una citoquina, una hormona, o un factor de coagulación.
- 30 **6.** La línea celular de mamífero modificada genéticamente de una cualquiera de las reivindicaciones 1 a 5, que además comprende al menos una secuencia cromosómica inactivada que codifica la glutamina sintetasa.
- 35 **7.** La línea celular de mamífero modificada genéticamente de una cualquiera de las reivindicaciones 1 a 6, para uso en un sistema de producción biológica.
- 40 **8.** Una composición que comprende la línea celular de mamífero modificada genéticamente de una cualquiera de las reivindicaciones 1 a 7 y al menos un virus, en la que la línea celular exhibe resistencia a infección por MVM.
- 45 **9.** Un procedimiento de reducción del riesgo de contaminación viral de un sistema de producción biológica, el procedimiento comprende proporcionar en el sistema de producción biológica el uso de la línea celular de mamífero modificada genéticamente de una cualquiera de las reivindicaciones 1 a 6.
- 50 **10.** Un procedimiento de reducción o prevención de la contaminación viral de un producto de proteína recombinante, el procedimiento comprende:
 - a. obtener la línea celular de mamífero modificada genéticamente de una cualquiera de las reivindicaciones 1 a 6; y
 - b. expresar el producto de proteína recombinante en la línea celular de mamífero modificada genéticamente.

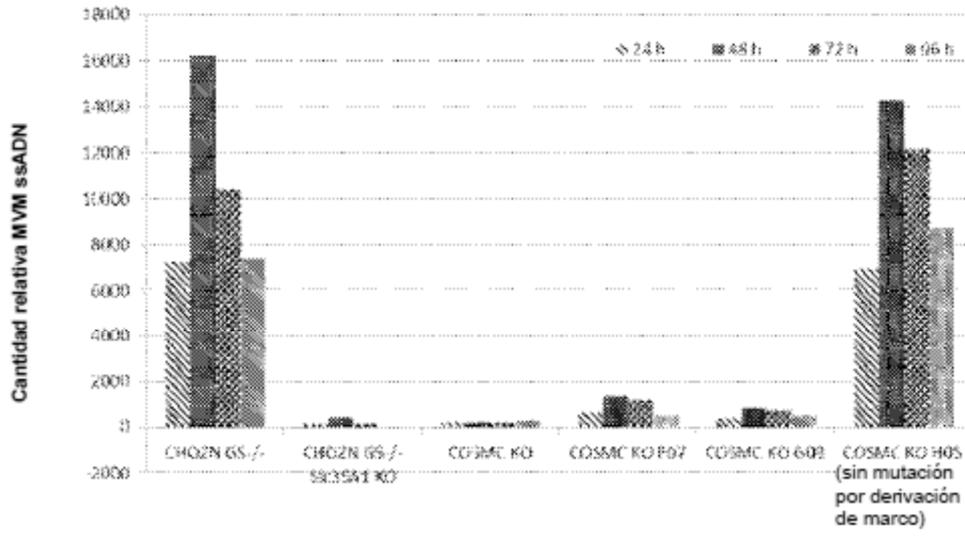


FIG. 1

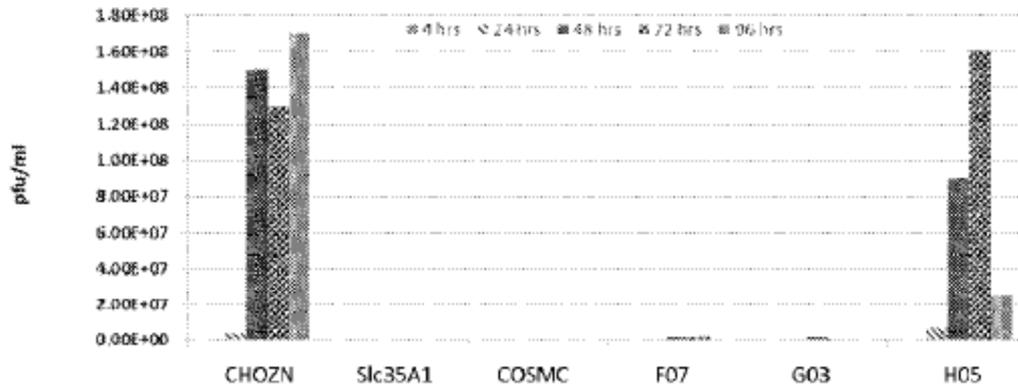
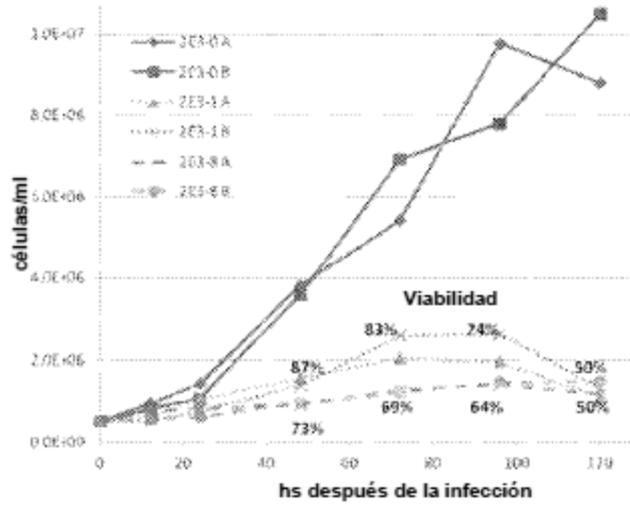


FIG. 2

A



B

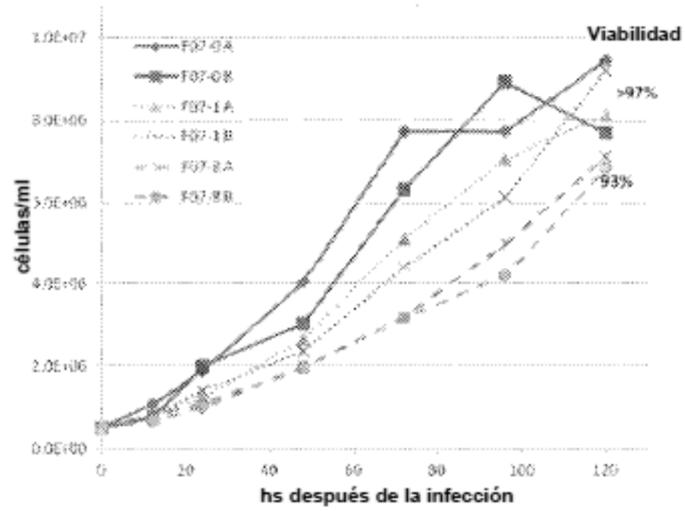
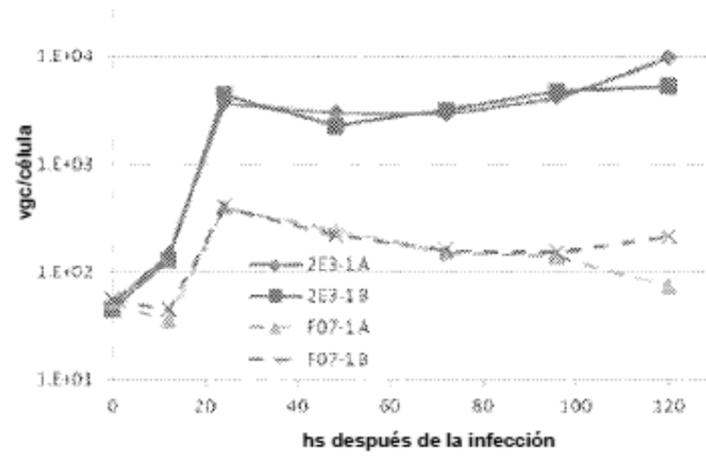


FIG. 3

A



B

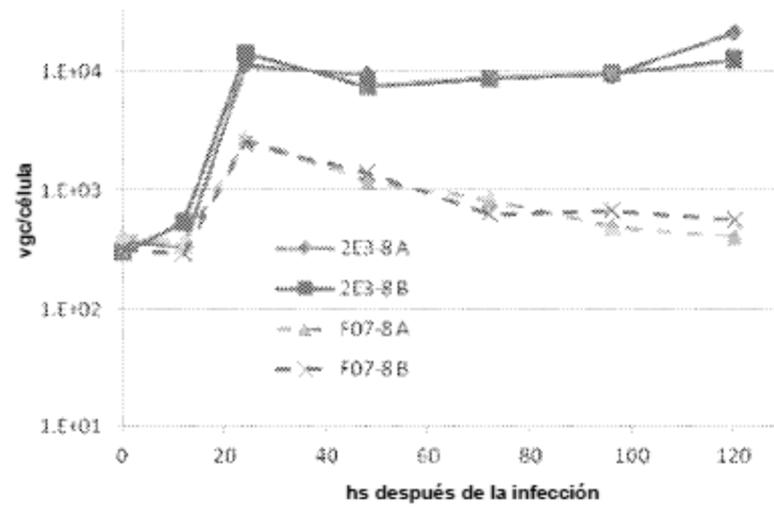
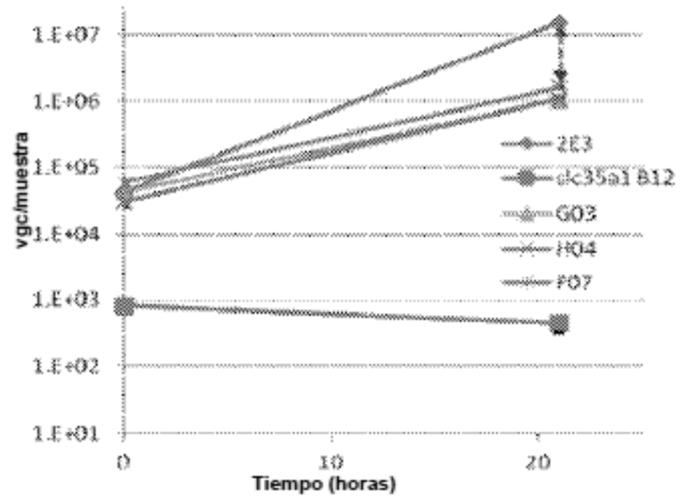


FIG. 4

A



B

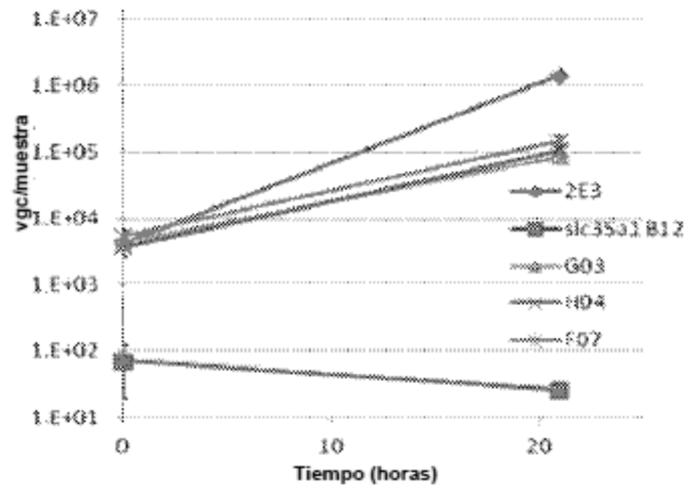


FIG. 5

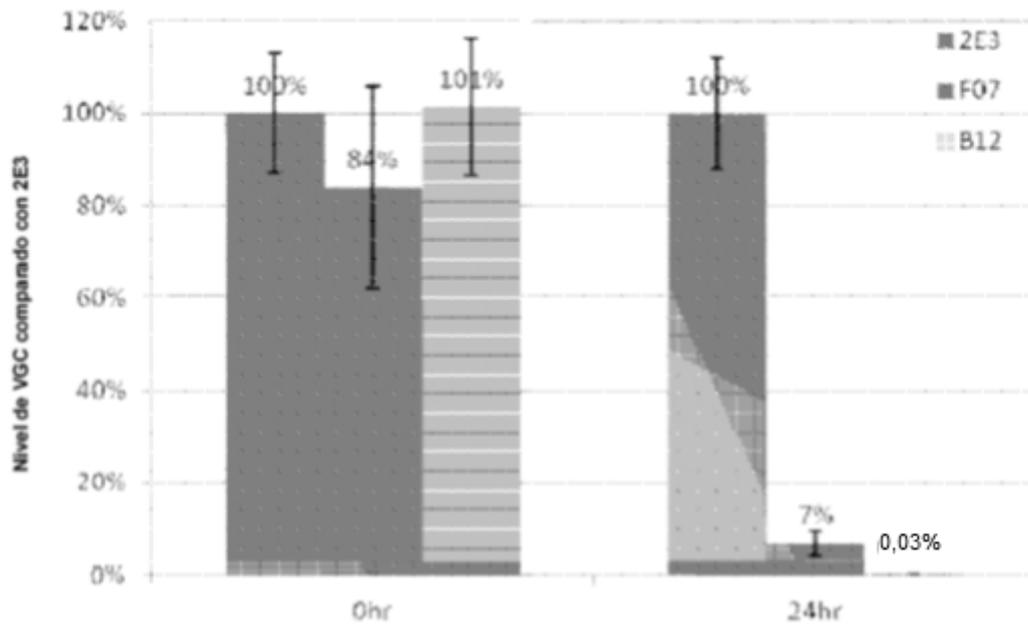
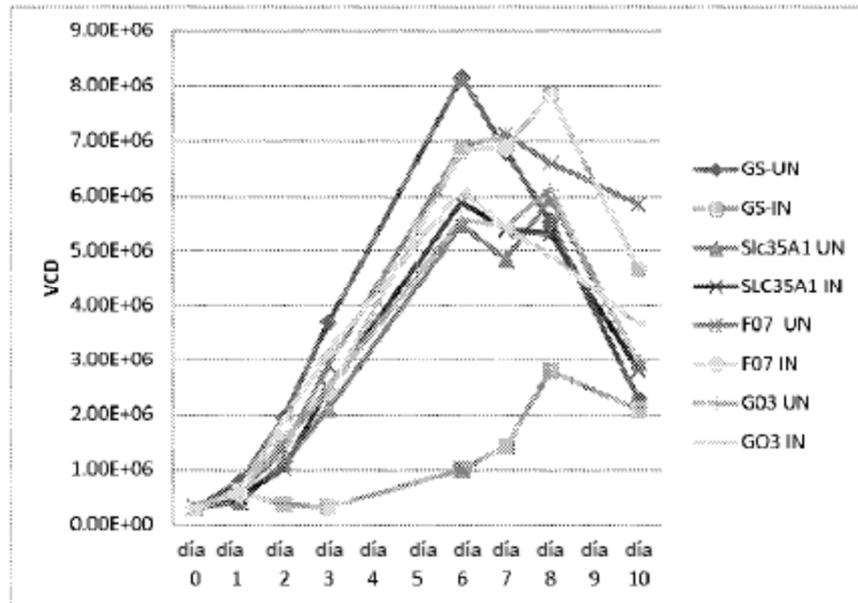


FIG. 6

A



B

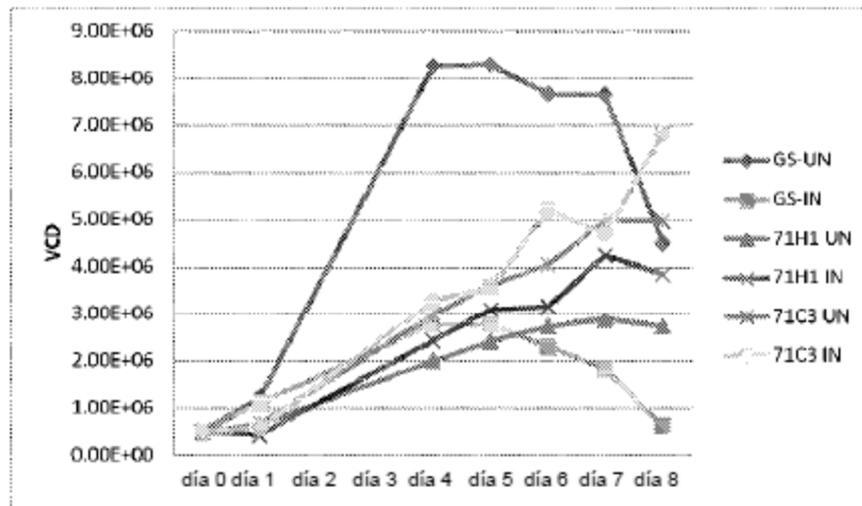


FIG. 7

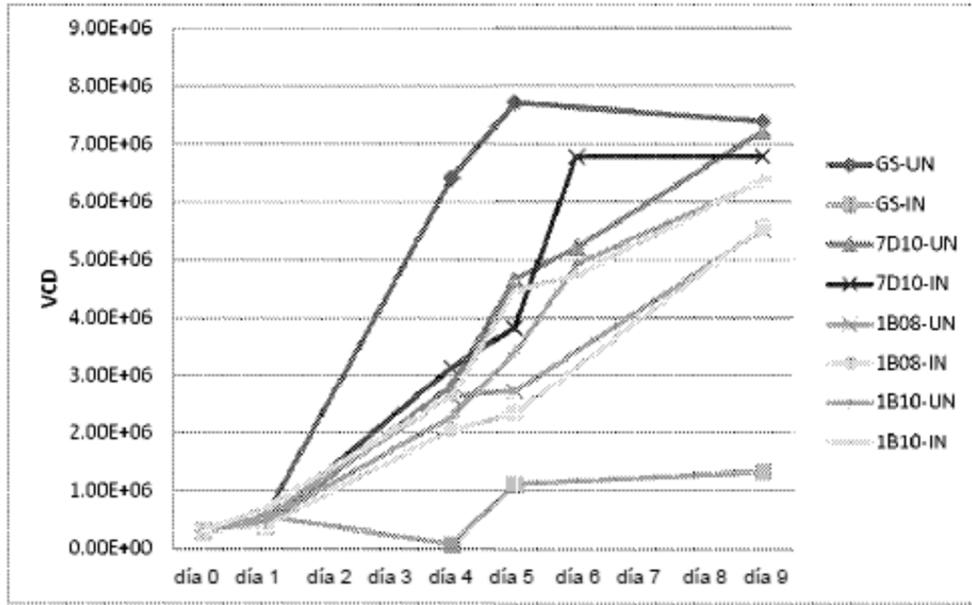


FIG. 8