

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 671 733**

51 Int. Cl.:

C12N 5/10 (2006.01)

C12P 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.06.2012** **PCT/US2012/045027**

87 Fecha y número de publicación internacional: **03.01.2013** **WO13003767**

96 Fecha de presentación y número de la solicitud europea: **29.06.2012** **E 12803827 (0)**

97 Fecha y número de publicación de la concesión europea: **04.04.2018** **EP 2726604**

54 Título: **Células deficientes en ácido CMP-N-acetilneuramínico hidroxilasa y/o glucoproteína alfa-1,3-galactosil transferasa**

30 Prioridad:

30.06.2011 US 201161503436 P

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

08.06.2018

73 Titular/es:

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

72 Inventor/es:

LIN, NAN;
GEORGE, HENRY J.;
MASCARENHAS, JOAQUINA;
COLLINGWOOD, TREVOR N.;
KAYSER, KEVIN J. y
ACHTIEN, KATHERINE

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 671 733 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).

DESCRIPCION

Células deficientes en ácido CMP-N-acetilneuramínico hidroxilasa y/o glucoproteína alfa-1,3-galactosil transferasa

Campo de la invención

La presente divulgación se refiere de manera general a células útiles para la producción de proteínas y, en particular, proteínas terapéuticas. Más específicamente, la presente divulgación se refiere a células deficientes en ciertas enzimas para el procesamiento de N-glucanos, procedimientos de producción de las células, y procedimientos para uso de las células para generar proteínas que tienen ciertos patrones de glucosilación.

Antecedentes de la invención

Aproximadamente el 70% de las proteínas terapéuticas tales como anticuerpos, factores de crecimiento, citocinas, hormonas, y factores coagulantes son glucoproteínas, las cuales son proteínas post-traducionalmente modificadas mediante la unión de glucanos. Las glucoproteínas terapéuticas las más recombinantes se producen en sistemas de expresión de mamíferos dado que la localización, número, y estructura de los N-glucanos se ha mostrado que afecta a la bioactividad, solubilidad, estabilidad, farmacocinéticas, inmunogenicidad, e índice de aclaramiento de las glucoproteínas terapéuticas. Dos diferencias en la maquinaria de glucosilación de proteínas de humanos y otros mamíferos explican las diferencias en los patrones de glucosilación de glucoproteínas producidas por células humanas y glucoproteínas producidas por otras células de mamíferos tales como células de roedores.

En primer lugar, los humanos no pueden sintetizar un resto galactosa-alfa-1,3-galactosa terminal (también conocido como alfa-Gal o α -Gal) sobre N-glucanos. La enzima glucoproteína alfa-1,3-galactosiltransferasa (Ggta1) forma el resto α -Gal mediante ligamiento a un resto galactosa a través de un enlace α -1,3 glucosídico a un terminal galactosa del N-glucano. Aparentemente los humanos tienen un gen GGTA1, pero es un pseudogen expresado que codifica una proteína truncada no funcional que contiene los cuatro primeros exones traducidos pero que ha perdido los dos exones catalíticos. Incluso aún cuando los humanos no expresan una enzima Ggta1 funcional y, por ello, no sintetizan restos α -Gal, muchos humanos producen anticuerpos contra esta estructura.

En segundo lugar, los humanos no pueden sintetizar el ácido siálico, ácido N-glucolilneuramínico (Neu5Gc). El Neu5Gc se produce por la hidroxilación del ácido CMP-N-acetilneuramínico (Neu5Ac) al CMP-Neu5Gc mediante la enzima CMP-Neu5Ac hidroxilasa (Cmah). Aunque el gen CMAH humano está irreversiblemente mutado, lo que previene la expresión de Neu5Ac, han sido detectadas trazas de Neu5Gc en suero humano. Parece que el Neu5Gc no humano puede ser incorporado metabólicamente dentro de tejidos humanos procedentes de ciertos alimentos de origen mamífero, de manera tal que esencialmente todos los humanos tienen anticuerpos específicos de Neu5Gc, a veces a altos niveles.

Las células de ovarios de hamster chino (CHO) han sido ampliamente usadas para la fabricación de proteínas terapéuticas, en parte debido a que se ha asumido que producen proteínas con patrones de glucosilación de tipo humano. Por ejemplo, de manera general se ha aceptado que las células CHO carecen de la maquinaria de biosíntesis para producir glucoproteínas con restos α -Gal. Además, incluso aunque las células CHO expresan Cmah y producen proteínas que tienen unidades Neu5Gc, la relación de unidades Neu5Gc a Neu5Ac puede reducirse modificando las condiciones de cultivo de células CHO. A pesar de la aceptación general que las células CHO fueran incapaces de sintetizar restos α -Gal, recientemente ha sido identificado el ortólogo CHO de Ggta1 (Bosques y otros, Nature Biotechnol., vol. 28 (nº. 11), págs. 1153-1156 (2010)). Yuko Naito y otros (Mol.Cell. Biol., vol. 27 (nº. 8), págs. 3008-3022, (2007)) han divulgado una línea de células madre en las cuales se ha interrumpido el gen Cmah, en tanto que la Patente de EE.UU. 2007/154982 divulga células CHO deficientes en Cmah. Debido al valor potencial para las reacciones de hipersensibilidad a las glucoproteínas terapéuticas recombinantes, existe una necesidad por líneas celulares CHO y otras líneas celulares de mamífero no humanas que produzcan glucoproteínas exentas de restos α -Gal y Neu5Gc.

Sumario de la invención

En resumen, de acuerdo con ello, un aspecto de la presente divulgación proporciona una línea celular de mamífero no humanas que comprende una secuencia cromosómica inactivada que codifica ácido citidino monofosfato-N-acetilneuramínico hidroxilasa (Cmah) y una secuencia cromosómica inactivada que codifica glucoproteína alfa-1,3-galactosiltransferasa (Ggla1), opcionalmente en la que la línea celular es una línea celular de ovario de hamster chino (CH). En ciertas realizaciones, la secuencia cromosómica inactivada que codifica Cmah comprende una delección de al menos un nucleótido, una inserción de al menos un nucleótido, una sustitución de al menos un nucleótido, o combinaciones de los mismos. En una realización, la secuencia cromosómica inactivada que codifica Cmah comprende secuencia introducida no exógenamente. En otra realización, la secuencia cromosómica inactivada que codifica Cmah es monoalélica y la línea celular produce una cantidad reducida de Cmah. En otra realización aún, la secuencia cromosómica inactivada que codifica Cmah es bialélica y la línea celular no produce Cmah. En una realización, la secuencia cromosómica está inactivada con una endonucleasa dirigida, por ejemplo, una meganucleasa, una TALEN, una endonucleasa específica del sitio, o una nucleasa de dedos de cinc. En cualquiera de estas realizaciones, la línea celular puede producir proteínas que carecen de restos de ácido N-glucolilneuramínico (Neu5Gc).

En ciertas realizaciones, la secuencia cromosómica inactivada que codifica Ggta1 comprende una delección de al menos un nucleótido, una inserción de al menos un nucleótido, una sustitución de al menos un nucleótido, o combinaciones de los mismos. En una realización, la secuencia cromosómica inactivada que codifica Ggta1 comprende una secuencia introducida no exógenamente. En una realización, la secuencia cromosómica inactivada que codifica Ggta1 es monoalélica y la línea celular produce una cantidad reducida de Ggta1. En otra realización, la secuencia cromosómica inactivada que codifica Ggta1 es bialélica y la línea celular no produce Ggta1. En otra realización, la secuencia cromosómica está inactivada con una endonucleasa dirigida, por ejemplo, una meganucleasa, una TALEN, una endonucleasa específica del sitio, o una nucleasa de dedos de cinc. En cualquiera de estas realizaciones, la línea celular de mamífero no humano puede producir proteínas que adicionalmente carecen de restos de galactosa-alfa-1,3-galactosa (alfa-Gal).

En una realización, la línea celular comprende una inactivación monoalélica de la secuencia cromosómica que codifica Cmah y una inactivación monoalélica de la secuencia cromosómica que codifica Ggta1, y la línea celular produce una cantidad reducida de Cmah y una cantidad reducida de Ggta1. En otra realización, la línea celular comprende una inactivación bialélica de la secuencia cromosómica que codifica Cmah y una inactivación bialélica de la secuencia cromosómica que codifica Ggta1, y la línea celular no produce Cmah o Ggta1. En una realización particular, la línea celular de mamífero no humano produce proteínas que carecen de restos de ácido N-glicolilneuramínico (Neu5Gc) y de restos de galactosa-alfa-1,3-galactosa (alfa-Gal).

En una realización particular de la invención, la línea celular es una línea celular de ovario de hámster chino (CHO). En una realización, la línea celular CHO comprende una inactivación monoalélica de la secuencia cromosómica que codifica Cmah, y produce una cantidad reducida de Cmah. En otra realización, la línea celular CHO comprende una inactivación bialélica de la secuencia cromosómica que codifica Cmah, y no produce Cmah. En otra realización, la línea celular CHO comprende una inactivación monoalélica de la secuencia cromosómica que codifica Ggta1, y produce una cantidad reducida de Cmah y una cantidad reducida de Ggta1. En otra realización aún, la línea celular CHO comprende una inactivación bialélica de la secuencia cromosómica que codifica Cmah, y una inactivación bialélica de la secuencia cromosómica que codifica Ggta1, y no produce Cmah o Ggta1. En una realización, la línea celular CHO produce proteínas que carecen de restos de ácido N-glicolilneuramínico (Neu5Gc) y restos de galactosa-alfa-1,3-galactosa (alfa-Gal).

En otro aspecto, la divulgación abarca procedimientos de producción de una línea celular deficiente en Cmah y Ggta1. En una realización, el procedimiento comprende la introducción dentro de la línea celular una endonucleasa dirigida o un ácido nucleico que codifica una endonucleasa dirigida orientada a una secuencia cromosómica que codifica Cmah. En otra realización, el procedimiento comprende la introducción dentro de una línea celular que es deficiente en Cmah, una endonucleasa dirigida o un ácido nucleico que codifica una endonucleasa dirigida orientada a una secuencia cromosómica que codifica Ggta1. En otra realización, el procedimiento comprende la introducción dentro de una línea celular una endonucleasa dirigida o un ácido nucleico que codifica una endonucleasa dirigida a una secuencia cromosómica que codifica Cmah y una endonucleasa dirigida o un ácido nucleico que codifica una endonucleasa dirigida orientada a una secuencia cromosómica que codifica Ggta1.

Otro aspecto de la divulgación abarca un procedimiento de producción de una proteína recombinante con un patrón de glucosilación de tipo humano. El procedimiento comprende la expresión de la proteína en una línea celular de mamífero no humano deficiente en Cmah y Ggta1. En una realización específica, la línea celular es una línea celular de ovario de hámster chino (CHO). En una realización, la línea celular comprende una secuencia cromosómica inactivada que codifica Cmah y una secuencia cromosómica inactivada que comprende Ggta1. En una realización, la secuencia cromosómica inactivada que codifica Cmah y Ggta1 es monoalélica y la línea celular produce una cantidad reducida de Cmah y Ggta1. En otra realización, la secuencia cromosómica inactivada que codifica Cmah y/o Ggta1 es bialélica, y la línea celular no produce Cmah y/o Ggta1. En otra realización, la proteína recombinante carece de restos de ácido N-glicolilneuramínico (Neu5Gc) y/o restos de galactosa-alfa-1,3-galactosa (alfa-Gal). En una realización, la proteína recombinante tiene al menos una propiedad que está mejorada con relación a una proteína recombinante similar producida mediante una línea celular comparable no deficiente en Cmah y Ggta1, por ejemplo, inmunogenicidad reducida, biodisponibilidad incrementada, eficacia incrementada, estabilidad incrementada, solubilidad incrementada, vida media mejorada, aclaramiento mejorado, farmacocinéticas mejoradas, y combinaciones de las mismas. La proteína recombinante puede ser cualquier proteína, incluyendo una proteína terapéutica. Los ejemplos de proteínas incluyen las seleccionadas procedentes de un anticuerpo, un fragmento de anticuerpo, un factor de crecimiento, una citocina, una hormona, un factor de coagulación, y fragmento funcional o variantes de los mismos.

Otros aspectos e iteraciones de la divulgación se describen con más detalle a continuación.

Descripción de las figuras

La FIG. 1 ilustra la escisión mediada por ZFN del locus Ggta1 en células CHO. Muestra los resultados de un ensayo de nucleasa topógrafa Cel-1. Las flechas indican productos de escisión de 215 bp y 100 bp en células CHO transfectadas con ARNm de ZFN (1) o ADN de ZFN (2). No se detectaron productos de escisión en células transfectadas falsas (3).

La FIG. 2 documenta la escisión mediada por ZFN del locus Cmah en células CHO, detectada mediante un ensayo de nucleasa topógrafa Cel-1. La flecha indica un producto de escisión.

La FIG. 3 ilustra la escisión mediada por ZFN del locus Ggta1 en células Cmah (-/-), detectada mediante un ensayo de nucleasa topógrafa Cel-1. Las células transfectadas con ZFN (marcadas con "#1" y "#2"), pero no las células transfectadas falsas, tenían los fragmentos de escisión de 215 bp y 100 bp.

Descripción detallada de la invención

La presente divulgación proporciona líneas celulares de mamífero no humano deficientes en Cmah y Ggta1. En una realización, las líneas celulares comprenden secuencias cromosómicas inactivadas que codifican Cmah y Ggta1 de manera tal que las líneas celulares producen cantidades reducidas de Cmah y Ggta1. En otra realización, las líneas celulares comprenden secuencias cromosómicas inactivadas que codifican Cmah y Ggta1 de manera tal que las líneas celulares no producen Cmah y/o Ggta1. Igualmente, se proporcionan en la presente invención procedimientos de obtención de las líneas celulares divulgadas en la presente invención y procedimientos para el uso de las líneas celulares divulgadas en la presente invención a fin de producir proteínas recombinantes con patrones de glucosilación de tipo humano. Dado que las líneas celulares son deficientes en Cmah y Ggta1, las líneas celulares producen glucoproteínas recombinantes con contenido reducido en Neu5Gc y α -Gal o glucoproteínas que carecen de Neu5Gc y/o α -Gal.

(I) Líneas de células deficientes en Cmah y Ggta1

Un aspecto de la presente divulgación proporciona una línea celular de mamífero no humano deficiente en ácido citidino monofosfato-N-acetilneuramínico hidroxilasa (Cmah) y glucoproteína alfa-1,3-galactosiltransferasa (Ggta1).

(a) Cmah y Ggta1

Las Cmah y Ggta1 son enzimas implicadas en la generación de N-glucanos sobre glucoproteínas. La Cmah cataliza la conversión del ácido siálico Neu5Ac en su derivado hidroxilado Neu5Gc. La Ggta1 liga un resto galactosa a través de un enlace α -1,3 glucosídico a una galactosa sobre el N-glucano para formar un resto Gal- α -1,3-Gal (es decir, α -Gal) terminal. En una realización, la línea celular es deficiente en Cmah. En otra realización, la línea celular es deficiente en Ggta1. En otra realización aún, la línea de celular es deficiente tanto en Cmah como en Ggta1.

En algunos casos, la línea celular deficiente en Cmah y Ggta1 puede tener niveles reducidos de Cmah y Ggta1 con relación a la línea celular madre. Por ejemplo, los niveles de Cmah Ggta1 pueden estar reducidos desde aproximadamente 5% hasta aproximadamente 10%, desde aproximadamente 10% hasta aproximadamente 20%, desde aproximadamente 20% hasta aproximadamente 30%, desde aproximadamente 30% hasta aproximadamente 40%, desde aproximadamente 40% hasta aproximadamente 50%, desde aproximadamente 50% hasta aproximadamente 60%, desde aproximadamente 60% hasta aproximadamente 70%, desde aproximadamente 70% hasta aproximadamente 80%, desde aproximadamente 80% hasta aproximadamente 90%, o desde aproximadamente 90% hasta aproximadamente 99,9% con relación a la línea celular madre que no es deficiente en Cmah y Ggta1. La línea celular que tiene niveles reducidos de Cmah y Ggta1 producirá, generalmente, proteínas con contenido reducido en Neu5Gc y α -Gal con relación a proteínas producidas mediante células comparables que no son deficientes en Cmah y Ggta1.

En otros casos, la línea celular deficiente en Cmah y Ggta1 puede esencialmente no producir Cmah y/o Ggta1. Tal como se usa en la presente invención, el término "esencialmente no producir Cmah y/o Ggta1" significa que no puede detectarse ARNm o proteína de Cmah y/o Ggta1 en las células deficientes o los lisatos obtenidos a partir de las mismas usando procedimientos bien conocidos en la técnica. Los ejemplos no limitativos de procedimientos adecuados para la determinación del nivel de ARNm o proteína incluyen ensayos PCR, qPCR, transferencia Western y ELISA. De acuerdo con ello, el nivel ARNm y/o proteína de Cmah y/o Ggta1 ARNm detectado en las células deficientes o lisatos es esencialmente el mismo que los niveles de la base genética. La línea celular exenta de Cmah y/o Ggta1 producirá, generalmente, proteínas que carecen de restos Neu5Gc y/o α -Gal.

En algunas realizaciones, el genoma de la línea celular deficiente en Cmah y Ggta1 puede editarse de manera tal que la secuencia cromosómica que codifica Cmah y la secuencia cromosómica que codifica Ggta1 está inactivada. Tal como se usa en la presente invención, el término "secuencia cromosómica inactivada" se refiere a una secuencia cromosómica que es incapaz de generar un producto de gen funcional. En una realización en la cual la línea celular comprende células euploides, la secuencia cromosómica inactivada puede ser monoalélica, de manera tal que la célula produce niveles reducidos de Cmah y Ggta1. En otra realización en la cual la línea celular es euploide, la secuencia cromosómica inactivada puede ser bialélica, de manera tal que la célula esencialmente no produce Cmah y/o Ggta1 y la célula puede denominarse una célula "inactivada". Como alternativa, en otras realizaciones en las cuales la línea celular es aneuploide, una o más copias de la secuencia(s) cromosómica(s) que codifica Cmah y Ggta1 están inactivadas, dando como resultado una cantidad reducida de Cmah y Ggta1. En otra realización en la cual la línea celular es aneuploide, todas las copias de la secuencia(s) cromosómica(s) que codifica Cmah y/o Ggta1 están inactivadas, dando como resultado una completa pérdida de la expresión del gen Cmah y Ggta1.

La secuencia cromosómica inactivada que codifica Cmah y Ggta1 puede comprender una delección de al menos un nucleótido, una inserción de al menos un nucleótido, o una sustitución de al menos un nucleótido. La secuencia cromosómica que codifica Cmah y Ggta1 puede inactivarse usando tecnología de edición de genoma mediada por endonucleasa dirigida tal como se detalla más adelante en la sección (II). En varias realizaciones, la secuencia cromosómica que codifica Cmah y Ggta1 puede inactivarse mediante la delección de la totalidad o parte de la región que codifica exones, delección de la totalidad o parte de una región de control, y/o delección de un sitio de corte y empalme, de manera tal que la línea celular es incapaz de producir Cmah y/o Ggta1. En otras realizaciones, la secuencia cromosómica que codifica Cmah y Ggta1 puede inactivarse a través de delecciones, inserciones, y/o sustituciones de nucleótidos para introducir un codón de parada prematuro, nuevo sitio de corte y empalme, y/o SNPs dentro de la secuencia cromosómica, de manera tal que la línea celular es incapaz de producir Cmah y Ggta1.

En una realización, la línea celular puede comprender una secuencia cromosómica inactivada que codifica Cmah debido a una delección, inserción, y/o sustitución de al menos un nucleótido dentro de la secuencia cromosómica que codifica Cmah. Por ejemplo, la secuencia cromosómica que codifica Cmah puede estar inactivada debido a una delección, inserción, y/o sustitución de al menos un nucleótido dentro del exón 5 de la secuencia cromosómica que codifica Cmah. En otra realización, la línea celular puede comprender una secuencia cromosómica inactivada que codifica Ggta1 debido a una delección, inserción, y/o sustitución de al menos un nucleótido dentro de la secuencia cromosómica que codifica Ggta1. Por ejemplo, la secuencia cromosómica que codifica Ggta1 puede inactivarse debido a una delección, inserción, y/o sustitución de al menos un nucleótido dentro del exón 9 de la secuencia cromosómica que codifica Ggta1.

En algunas realizaciones, la línea celular deficiente en Cmah y Ggta1 puede ser igualmente deficiente en glutamina sintasa (GS), dihidrofolato reductasa (DHFR), hipoxantina-guanina fosforibosiltransferasa (HPRT), o combinaciones de las mismas. Además, la línea celular que comprende deficiencias en GS, DHFR, y/o HPRT puede ser deficiente en GS, DHFR, y/o HPRT debido a secuencias cromosómicas inactivadas que codifican GS, DHFR, y/o HPRT.

(b) tipos de células

El tipo de línea celular que es deficiente en Cmah y Ggta1 puede ser cualquiera de un cierto número de tipos de células adecuadas. En general, la línea celular es una línea celular de mamífero no humano. Las líneas celulares de mamífero no humano adecuadas incluyen, pero sin limitarse a ellas, células de ovario de hámster chino (CHO), células de riñón de hámster bebé (BHK); células NSO de mieloma de ratón, células 3T3 de fibroblasto embrionario de ratón, células A20 de linfoma B de ratón; células B16 de melanoma de ratón; células C2C12 de mioblasto de ratón; células SP2/0 de mieloma de ratón; células C3H-10T1/2 mesenquimales embrionarias de ratón; células CT26 de carcinoma de ratón; células DuCuP de próstata de ratón; células EMT6 de mama de ratón; células Hepa1c1c7 de hepatoma de ratón; células J5582 de mieloma de ratón; células MTD-1A epiteliales de ratón; células MyEnd de miocardio de ratón; células RenCa renal de ratón; células RIN-5F pancreáticas de ratón; células X64 de melanoma de ratón; células YAC-1 de linfoma de ratón; células 9L de glioblastoma de rata; células RBL de linfoma B de rata; células B35 de neuroblastoma de rata; células (HTC) de hepatoma de rata; células BRL3A de hígado rata búfalo; células (MDCK) de riñón canino; células (CMT) de mamífero canino; células D17 de osteosarcoma de rata; células DH82 de monocito/macrófago de rata; células (COS7) de fibroblasto SV-40 transformado de riñón de mono; células CVI-76 de riñón de mono; y células (VERO-76) de riñón de mono verde africano. Una lista extensa de líneas celulares de mamífero no humano puede encontrarse en el catálogo de la American Type Culture Collection (ATCC, Manassas, VA). En una realización, la línea celular que es deficiente en Cmah y Ggta1 es distinta de una línea celular de ratón. En otra realización aún, la línea celular que es deficiente en Cmah y Ggta1 es distinta de una línea celular de porcino.

En algunas realizaciones, la línea celular es de un tipo que es ampliamente usado para la producción de glucol proteínas recombinantes. En una realización a modo de ejemplo, la línea celular es una línea celular CHO. Numerosas líneas celulares CHO se encuentran disponibles de la ATCC y de vendedores comerciales. Las líneas celulares CHO adecuadas incluyen, pero sin limitarse a ellas, células CHO-K1 y derivadas de las mismas, células CHO-K1SV, células CHO DG44, células CHO-S, células CHO P12, células CHO pro3, células CHO/DHFR, CHO/GS, y células CHO DXB11.

(c) ácido nucleico opcional

En algunas realizaciones, la línea celular de mamífero no humano divulgada en la presente invención puede comprender además al menos una secuencia de ácido nucleico que codifica una proteína recombinante. La proteína recombinante puede ser, sin limitación, un anticuerpo, un fragmento de anticuerpo, un anticuerpo monoclonal, un anticuerpo humanizado, un anticuerpo monoclonal humanizado, un anticuerpo quimérico, una molécula IgG, una cadena pesada IgG, una cadena ligera IgG, una molécula IgA, una molécula IgD, una molécula IgE, una molécula IgM, una glucoproteína, un factor de crecimiento, una citocina, un interferón, una interleucina, una hormona, un factor de coagulación (o coagulante), un componente de la sangre, una enzima, una proteína terapéutica, una proteína nutricional, una vacuna, un fragmento funcional o variante funcional de cualquiera de los anteriores, o una proteína de fusión que comprende cualquiera de las proteínas anteriores y/o fragmentos funcionales o variantes de los mismos.

En algunas realizaciones, la secuencia de ácido nucleico que codifica la proteína recombinante puede estar ligada a una secuencia de ácido nucleico que codifica hipoxantina-guanina fosforibosiltransferasa (HPRT), dihidrofolato reductasa (DHFR), y/o glutamina sintasa (GS), de manera tal que HPRT, DHFR, y/o GS pueden usarse como un marcador seleccionable amplificable.

- 5 En algunas realizaciones, la secuencia de ácido nucleico que codifica la proteína recombinante puede ser extracromosómica. Es decir, el ácido nucleico que codifica la proteína recombinante puede ser expresada transitoriamente a partir de un plásmido, un cósmido, un cromosoma artificial, un minicromosoma, y similares. Los expertos en la técnica están familiarizados con constructos de expresión adecuados, secuencias de control de expresión apropiadas, y procedimientos de introducción de dichos constructos dentro de células.
- 10 En otras realizaciones, la secuencia de ácido nucleico que codifica la proteína recombinante puede estar cromosómicamente integrada dentro del genoma de la célula, de manera tal que la proteína recombinante está expresada de manera estable. En algunas iteraciones de la presente realización, la secuencia de ácido nucleico que codifica la proteína recombinante puede estar ligada de manera operable a una secuencia de control de expresión heteróloga apropiada (es decir, promotor). En otras iteraciones, la secuencia de ácido nucleico que codifica la proteína recombinante puede disponerse bajo el control de una secuencia de control de expresión endógena. La secuencia de
- 15 ácido nucleico que codifica la proteína recombinante puede estar integrada dentro del genoma de la línea celular usando técnicas bien conocidas.

Los procedimientos, vectores, y técnicas de clonación para la preparación e introducción de secuencias de ácido nucleico exógenas (por ejemplo, las que codifican una proteína recombinante) son bien conocidas en la técnica (véase, por ejemplo, "Current Protocols in Molecular Biology", Ausubel y otros, 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)).

(d) ejemplos de realizaciones

- 25 En una realización, la línea celular es una línea celular CHO que comprende una inactivación monoalélica o bialélica de la secuencia cromosómica que codifica Cmah y una inactivación monoalélica o bialélica de la secuencia cromosómica que codifica Ggta1.

(II) Procedimientos de preparación de líneas celulares deficientes en Cmah y Ggta1

- La línea celular deficiente en Cmah y Ggta1 puede prepararse mediante una diversidad de procedimientos. En ciertas realizaciones, la línea celular deficiente en Cmah y Ggta1 puede prepararse mediante un procedimiento de edición de genoma mediado por endonucleasa dirigida. En otras realizaciones, la línea celular deficiente en Cmah y
- 30 Ggta1 puede prepararse mediante mutagénesis aleatoria, sistemas de recombinación específicos del sitio, u otros procedimientos conocidos en la técnica.

(a) edición de genoma mediado por endonucleasa dirigida

- Las endonucleasas dirigidas puede usarse para editar (es decir, inactivar o modificar) una secuencia cromosómica específica. Una secuencia cromosómica específica puede inactivarse mediante la introducción dentro de una endonucleasa dirigida o un ácido nucleico que codifica la endonucleasa dirigida, la cual está manipulada genéticamente para identificar una secuencia cromosómica específica. En una realización, la endonucleasa dirigida reconoce y une la secuencia cromosómica específica e introduce una doble hebra rota que se repara mediante un procedimiento de reparación de unión de extremos no homólogos (NHEJ). Dado que el NHEJ es propenso al error, puede producirse una delección, inserción, o sustitución de al menos un nucleótido, interrumpiendo, de esta forma, el marco de lectura de la secuencia cromosómica, de manera tal que no se produce producto de proteína. En otra realización, las endonucleasas dirigidas pueden igualmente usarse para editar una secuencia cromosómica a través de una reacción de recombinación de homólogos mediante la co-introducción de un polinucleótido que tenga una identidad de secuencia substancial con una porción de la secuencia cromosómica dirigida. La doble hebra rota introducida mediante la endonucleasa dirigida se repara mediante un procedimiento de reparación dirigido a la homología, de manera tal que la secuencia cromosómica se intercambia con el polinucleótido de manera tal que da como resultado la edición de la secuencia cromosómica.
- 45

(i) endonucleasas dirigidas

- Para editar la secuencia cromosómica pueden usarse una diversidad de endonucleasas dirigidas. La endonucleasa dirigida puede ser una proteína de origen natural o una proteína genéticamente manipulada. En una realización, la endonucleasa dirigida puede ser una meganucleasa. Las meganucleasas son endodeoxiribonucleasas caracterizadas por secuencias de reconocimiento largas, es decir, la secuencia de reconocimiento generalmente está comprendida desde aproximadamente 12 pares de bases hasta aproximadamente 40 pares de bases. Como una consecuencia de esta exigencia, la secuencia de reconocimiento generalmente se produce solamente una vez en cualquier genoma dado. Entre las meganucleasas, la familia de endonucleasas de inicio denominadas LAGLIDADG se han transformado en una herramienta valiosa para el estudio de genomas y manipulación genética de genomas.
- 50
- 55

Una meganucleasa puede ser dirigida a una secuencia cromosómica específica mediante la modificación de su secuencia de reconocimiento usando técnicas bien conocidas para los expertos en la técnica.

En otra realización, la endonucleasa dirigida puede ser una nucleasa efectora de tipo activador de la transcripción (TALE). Los TALEs son factores de transcripción procedentes de la planta patógena *Xantomonas*, la cual puede fácilmente manipularse genéticamente para unir nuevas ADN dianas. Los TALEs, o versiones truncadas de los mismos, pueden estar ligados al dominio catalítico de endonucleasas tales como FokI para crear la endonucleasa dirigida denominada TALE nucleasas o TALENs.

En otra realización aún, la endonucleasa dirigida puede ser una endonucleasa específica del sitio. En particular, la endonucleasa específica del sitio puede ser una endonucleasa de "corte raro" cuya secuencia de reconocimiento se produce raramente en un genoma. Preferiblemente, la secuencia de reconocimiento de la endonucleasa específica del sitio se produce solamente una vez en un genoma. En una realización alternativa además, la endonucleasa dirigida puede ser un agente de inducción de rotura de doble hebra de ADN dirigido artificial.

En otras realizaciones, la endonucleasa dirigida puede ser una nucleasa de dedo de cinc (ZFN). Típicamente, una nucleasa de dedo de cinc comprende un dominio de unión de ADN (es decir, dedo de cinc) y un dominio de escisión (es decir, nucleasa), ambos de los cuales se describen más adelante.

Dominio de unión de dedo de cinc

Los dominios de unión de dedos de cinc pueden manipularse genéticamente para reconocer y unir a cualquier secuencia de ácido nucleico a elegir. Véase, por ejemplo, Beerli y otros, *Nat. Biotechnol.*, vol. 20, págs. 135-141, (2002); Pabo y otros, *Ann.Rev. Biochem.*, vol. 70, págs. 313-340, (2001); Isalan y otros, *Nat. Biotechnol.*, vol. 19, págs. 656-660, (2001); Segal y otros, *Curr. Opin. Biotechnol.*, vol. 12, págs. 632-637, (2001); Choo y otros, *Curr. Opin. Struct. Biol.*, vol. 10, pág. 411-416, (2000); Zhang y otros, *J. Biol. Chem.*, vol. 275, (nº. 43), págs. 33850-33860, (2000); Doyon y otros, *Nat. Biotechnol.*, vol. 26, págs. 702-708, (2008); y Santiago y otros, *Proc. Natl. Acad. Sci. USA.*, vol. 105, págs. 5809-5814, (2008). Un dominio de unión de dedo de cinc manipulado genéticamente puede tener una nueva especificidad de unión comparada a una proteína de dedo de cinc producida de manera natural. Los procedimientos de manipulación genética incluyen, pero sin limitarse a ellos, 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 dobles, tripletes y/o cuadrupletes y secuencias de aminoácidos de dedos de cinc individuales, en las cuales cada secuencia de nucleótido doblete, triplete o cuadruplete está asociada con una o más secuencias de aminoácidos de dedos de cinc que unen la secuencia triplete o cuadruplete particular. Véase, por ejemplo, Las Patentes de EE.UU. Nos. 6.453.242 y 6.534.261. Como un ejemplo, el algoritmo descrito en la Patente de EE.UU. No. 6.453.242 puede usarse para diseñar un dominio de unión de dedo de cinc para identificar una secuencia preseleccionada. Igualmente, pueden usarse procedimientos alternativos, tales como de diseño racional usando una tabla de códigos de reconocimiento no degenerada, para diseñar un dominio de unión de dedo de cinc para identificar una secuencia específica (Sera y otros, *Biochemistry*, vol. 41, págs. 7074-7081, (2002). Las herramientas de bases web públicamente disponibles para la identificación de sitios diana potenciales en secuencias de ADN, así como de diseño de dominios de unión de dedos de cinc, son conocidos en la técnica. Por ejemplo, las herramientas para la identificación de sitios diana potenciales en secuencias de ADN pueden encontrarse en <http://www.zincfingertools.org>. Pueden encontrarse herramientas para el diseño de dominios de unión de dedos de cinc en <http://bindr.gdcb.iastate.edu/ZiFiT/>. (Véase, igualmente, Mandell y otros, *Nuc. Acid. Res.*, vol. 34, págs. W516-W523, (2006)); Sander y otros, *Nuc. Acid. Res.*, vol. 35, págs. W599-W605, (2007)).

Puede diseñarse un dominio de unión de dedo de cinc para reconocer y unir una secuencia de ADN comprendida desde aproximadamente 3 nucleótidos hasta aproximadamente 21 nucleótidos de longitud, o preferiblemente desde aproximadamente 9 hasta aproximadamente 18 nucleótidos de longitud. En general, los dominios de unión de dedos de cinc de las nucleasas de dedos de cinc divulgadas en la presente invención comprenden al menos tres regiones de reconocimiento de dedos de cinc (es decir, dedos de cinc). En una realización, el dominio de unión de dedos de cinc comprende cuatro regiones de reconocimiento de dedos de cinc. En otra realización, el dominio de unión de dedos de cinc comprende cinco regiones de reconocimiento de dedos de cinc. En otra realización aún, el dominio de unión de dedos de cinc comprende seis regiones de reconocimiento de dedos de cinc. Puede diseñarse un dominio de unión de dedos de cinc para unir a cualquier secuencia de ADN diana adecuada. Véanse, por ejemplo, las Patentes de EE.UU. Nos. 6.607.882; 6.534.261 y 6.453.242.

Los ejemplos de procedimientos de selección de una región de reconocimiento de dedos de cinc pueden incluir la representación de fagos y sistemas bi-híbridos, los cuales se encuentran descritos en las Patentes de EE.UU. Nos. 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 en las WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 y GB 2.338.237. Además, la potenciación de la especificidad de unión para los dominios de unión de dedos de cinc ha sido descrita, por ejemplo, en la WO 02/077227.

Los dominios de unión de dedos de cinc y procedimientos para el diseño y construcción de proteínas de fusión (y polinucleótidos que las codifican) son conocidos para los expertos en la técnica y se encuentran descritos en detalle, por ejemplo, en la Patente de EE.UU. No. 7.888.121. Las regiones de reconocimiento de dedos de cinc y/o proteínas de dedos de cinc multi-dedos, pueden ligarse conjuntamente usando secuencias ligadoras adecuadas, incluyendo,

por ejemplo, ligadores de cinco o más aminoácidos de longitud. Véanse, Patentes de EE.UU. Nos. 6.479.626; 6.903.185; y 7.153.949, para ejemplos no limitativos de secuencias ligadoras de seis o más aminoácidos de longitud. El dominio de unión de dedos de cinc descrito en la presente invención puede incluir una combinación de ligadores adecuados entre los dedos de cinc individuales de la proteína.

- 5 En algunas realizaciones, la nucleasa de dedos de cinc comprende además una señal o secuencia de localización nuclear (NLS). Una NLS es una secuencia de aminoácido que facilita la identificación de la proteína nucleasa de dedos de cinc dentro del núcleo, para introducir una doble hebra rota junto a la secuencia diana en el cromosoma. Las señales de localización nuclear son conocidas en la técnica. Véase, por ejemplo, Makkerh y otros, Current Bio-

logy, vol. 6, págs. 1025-1027, (1996).

10 Dominio de escisión

Una nucleasa de dedos de cinc incluye, igualmente, un dominio de escisión. La porción de dominio de escisión de la nucleasas de dedos de cinc puede obtenerse a partir de cualquier endonucleasa o exonucleasa. Los ejemplos no limitativos de endonucleasas a partir de las cuales puede obtenerse un dominio de escisión incluyen, por sin limitarse a ellos, endonucleasas de restricción y endonucleasas de inicio. Véase, por ejemplo, New England Biolabs Catalog o Belfort y otros, Nucleic Acids Res., vol. 25, pág. 3379-3388, (1997). Son conocidas enzimas adicionales que escinden ADN (por ejemplo, S1 Nuclease; nucleasa de judía de Mung; DNasa I pancreática; nucleasa microcócica; endonucleasa HO de levadura). Véase igualmente Linn y otros (eds.) Nucleases, Cold Spring Harbor Laboratory Press, (1993). Una o más de estas enzimas (o fragmentos funcionales de las mismas) pueden usarse como una fuente de dominios de escisión.

- 20 Igualmente, puede obtenerse un dominio de escisión a partir de una enzima o porción de la misma, tal como se ha descrito anteriormente, que requiere la dimerización para la actividad de escisión. Pueden requerirse dos nucleasas de dedos de cinc para la escisión, puesto que cada nucleasa comprende un monómero del dímero de enzima activo. Como alternativa, una única nucleasa de dedos de cinc puede comprender ambos monómeros para crear un dímero de enzima activo. Tal como se usa en la presente invención, un "dímero de enzima activo" es un dímero de enzima capaz de escisión de una molécula de ácido nucléico. Los dos monómeros de escisión pueden obtenerse a partir de la misma endonucleasa (o fragmentos funcionales de la misma), o cada monómero puede obtenerse a partir de una endonucleasa diferente (o fragmentos funcionales de la misma).

- 30 Cuando se usan dos monómeros de escisión para formar un dímero de enzima activo, los sitios de reconocimiento para las dos nucleasas de dedos de cinc están dispuestas, preferiblemente, de forma tal que la unión de las dos nucleasas de dedos de cinc a sus sitios de reconocimiento respectivos dispone los monómeros de escisión en una orientación espacial entre sí de manera tal que permite a los monómeros de escisión formar un dímero de enzima activo, por ejemplo, mediante dimerización. Como un resultado de ello, los bordes próximos de los sitios de reconocimiento pueden estar separados mediante aproximadamente 5 hasta aproximadamente 18 nucleótidos. Por ejemplo, los bordes próximos pueden estar separados mediante aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 35 17 ó 18 nucleótidos. No obstante, se sobreentiende que cualquier número entero de nucleótidos o de pares de nucleótidos puede interponerse entre dos sitios de reconocimiento (por ejemplo, desde aproximadamente 2 hasta aproximadamente 50 pares de nucleótidos o más). Los bordes próximos de los sitios de reconocimiento de las nucleasas de dedos de cinc, tales como, por ejemplo, los descritos detalladamente en la presente invención, pueden estar separados por 6 nucleótidos. En general, el sitio de escisión está situado entre los sitios de reconocimiento.

- 40 Las endonucleasas de restricción (enzimas de restricción) están presentes en muchas especies y son capaces de unión específica de la secuencia a un ADN (a un sitio de reconocimiento), y escisión del ADN en o cerca del sitio de unión. Ciertas enzimas de restricción (por ejemplo Tipo IIS) escinden el ADN en los sitios separados del sitio de reconocimiento y tienen dominios de unión y escisión separables. Por ejemplo, la enzima Tipo IIS FokI cataliza la escisión de doble hebra de ADN, a 9 nucleótidos a partir de su sitio de reconocimiento sobre una hebra y 13 nucleótidos a partir de su sitio de reconocimiento sobre la otra. Véanse, por ejemplo, las Patentes de EE.UU. Nos. 5.356.802; 5.436.150 y 5.487.994; así como Li y otros, Proc. Natl. Acad. Sci. USA, vol. 89, pág. 4275-4279, (1992); Li y otros, Proc. Natl. Acad. Sci. USA, vol. 90, pág. 2764-2768, (1993); Kim y otros, Proc. Natl. Acad. Sci. USA, vol. 91, pág. 883-887, (1994a); Kim y otros, J. Biol. Chem., vol. 269, pág. 31978-31982, (1994b). De acuerdo con ello, una nucleasa de dedos de cinc puede comprender el dominio de escisión a partir de al menos una enzima de restricción Tipo IIS y uno o más dominios de unión de dedos de cinc, los cuales pueden o no estar manipulados genéticamente. Los ejemplos de enzimas de restricción Tipo IIS se describen, por ejemplo, en la International Publication WO 07/014.275. Las enzimas de restricción adicionales contienen, igualmente, dominios de unión y escisión separables, estando igualmente contemplados estos por la presente divulgación. Véase, por ejemplo, Roberts y otros, Nucleic Acids Res., vol. 31, págs. 418-420, (2003).

- 55 Un ejemplo de enzima de restricción Tipo IIS, cuyo dominio de escisión es separable del dominio de unión, es FokI. Esta enzima particular es activa en forma de un dímero (Bitinaite y otros, Proc. Natl. Acad. Sci. USA, vol. 95, pág. 10.570-10.575, (1998)). De acuerdo con ello, para los fines de la presente divulgación, la porción de la enzima FokI usada en una nucleasa de dedos de cinc se considera un monómero de escisión. Así, para la escisión de doble hebra dirigida usando un dominio de escisión FokI, puede usarse dos nucleasas de dedos de cinc, comprendiendo cada una un monómero de escisión FokI, para reconstituir un dímero de enzima activo. Como alternativa, puede

usarse igualmente una molécula polipéptida sencilla conteniendo un dominio de unión de dedos de cinc y dos monómeros de escisión Folk.

En ciertas realizaciones, el dominio de escisión comprende uno o más monómeros de escisión manipulados genéticamente que minimizan o previenen la homodimerización. A modo de ejemplo no limitativo, los restos aminoácidos en las posiciones 446, 447, 449, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, y 538 de Folk son todos ellos dianas para influenciar la dimerización de los semi-dominios de escisión Folk. Los ejemplos de monómeros de escisión manipulados genéticamente de Folk que forman heterodímeros obligados, incluyen un par en el cual un primer monómero de escisión incluye mutaciones en las posiciones de los restos de aminoácidos 490 y 538 de Folk y un segundo monómero de escisión que incluye mutaciones en las posiciones de los restos de aminoácidos 486 y 499.

De acuerdo con ello, en una realización, una mutación en la posición del aminoácido 490 reemplaza Glu (E) con Lys (K); una mutación en el resto del aminoácido 538 reemplaza Iso (I) con Lys (K); una mutación en el resto del aminoácido 486 reemplaza Gln (Q) con Glu (E); y una mutación en la posición 499 reemplaza Iso (I) con Lys (K). Específicamente, los monómeros de escisión manipulados genéticamente pueden prepararse mediante mutación en las posiciones 490 desde E hasta K y 538 desde I hasta K en un monómero de escisión para producir un monómero de escisión manipulado genéticamente designado "E490K:I538K", y mediante mutación en las posiciones 486 desde Q hasta E y 499 desde I hasta L en otro monómero de escisión, para producir un monómero de escisión manipulado genéticamente designado "Q486E:I499L". Los monómeros de escisión manipulados genéticamente anteriormente descritos son mutantes heterodímeros obligados en los cuales la escisión aberrante está minimizada o suprimida. Los monómeros de escisión manipulados genéticamente pueden prepararse usando un procedimiento adecuado, por ejemplo, mediante mutagénesis dirigida al sitio de monómeros de escisión de tipo salvaje (Folk), tal como se describe en la Patente de EE.UU. No. 7.888.121.

(ii) polinucleótido opcional

El procedimiento para edición de genoma dirigido puede comprender además la introducción dentro de la célula de al menos un polinucleótido que comprende una secuencia que tiene identidad de secuencia substancial a una secuencia sobre al menos un lado del sitio de escisión dirigido. Por ejemplo, el polinucleótido puede comprender una primera secuencia que tiene identidad de secuencia substancial a la secuencia sobre un lado del sitio de escisión dirigido y una segunda secuencia que tiene identidad de secuencia substancial a la secuencia sobre el otro lado del sitio de escisión dirigido. Como alternativa, el polinucleótido puede comprender una primera secuencia que tiene identidad de secuencia substancial a la secuencia sobre un lado del sitio de escisión dirigido y una segunda secuencia que tiene identidad de secuencia substancial a una secuencia localizada separada del sitio de escisión dirigido. La secuencia localizada separada del sitio de escisión dirigido puede ser de decenas, centenas, o miles de nucleótidos más arriba o más abajo del sitio de escisión dirigido.

Las longitudes de las secuencias primera y segunda en el polinucleótido que tienen identidad de secuencia substancial a la secuencias en la secuencia cromosómica puede variar y varía. En general, cada una de las secuencias primera y segunda en el polinucleótido son al menos de aproximadamente 10 nucleótidos de longitud. En varias realizaciones, las secuencias polinucleótidas que tienen identidad de secuencia substancial con las secuencias cromosómicas son aproximadamente de 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 substancial" 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 tienen 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%, ó 99% de identidad de secuencia con la secuencia cromosómica de interés.

La longitud de polinucleótido puede variar y varía. Por ejemplo, el polinucleótido puede estar comprendido desde aproximadamente 20 nucleótidos de longitud hasta aproximadamente 200.000 nucleótidos de longitud. En varias realizaciones, el polinucleótido está comprendido desde aproximadamente 20 nucleótidos hasta aproximadamente 100 nucleótidos de longitud, desde aproximadamente 100 nucleótidos hasta aproximadamente 1000 nucleótidos de longitud, desde aproximadamente 1000 nucleótidos hasta aproximadamente 10.000 nucleótidos de longitud, desde aproximadamente 10.000 nucleótidos hasta aproximadamente 100.000 nucleótidos de longitud, o desde aproximadamente 100.000 nucleótidos hasta aproximadamente 200.000 nucleótidos de longitud,

Típicamente, el polinucleótido será ADN. El ADN puede ser de una sola hebra o de doble hebra. El polinucleótido donante puede ser un ADN plásmido, un cromosoma artificial bacteriano (BAC), un cromosoma artificial de levadura (YAC), un vector vírico, una pieza lineal de ADN, un fragmento PCR, un ácido nucléico desprotegido, o un ácido nucléico acomplejado con un vehículo de suministro tal como un liposoma o poloxámero.

En algunas realizaciones, el polinucleótido puede comprender además un marcador. Los ejemplos no limitativos de marcadores adecuados incluyen sitios de restricción, proteínas fluorescentes, o marcadores seleccionables. Dichos marcadores hacen posible el cribado de integraciones dirigidas.

(iii) introducción dentro de la célula

- 5 La endonucleasa dirigida puede introducirse dentro de la célula como una proteína o como un ácido nucleico que codifica la endonucleasa dirigida. El ácido nucleico que codifica la endonucleasa dirigida puede ser ADN o ARN (es decir, ARNm). En realizaciones en las cuales el ácido nucleico que codifica es ARNm, el ARNm puede ser con terminación 5' y/o poliadenilada 3'. En realizaciones en las cuales 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 está opcionalmente unido
- 10 de manera operable a un promotor adecuado. Los expertos en la técnica están familiarizados con los vectores, promotores, u otros elementos de control apropiados, y medios de introducción del vector dentro de la célula de interés.

- La endonucleasa dirigida o ácido nucleico que codifica la endonucleasa dirigida y el polinucleótido opcional descrito anteriormente puede introducirse dentro de la célula mediante una diversidad de medios. Los medios de suministro adecuados incluyen microinyección, electroporación, sonoporación, biolísticos, transfección mediada por fosfato
- 15 cálcico, transfección catiónica, transfección de liposoma, transfección de dendrímero, transfección por choque térmico, transfección por nucleofección, magnetofección, lipofección, impalefacción, transfección óptica, fijación potenciada por agente patentado de ácidos nucleicos, y suministro mediante liposomas, inmunoliposomas, virosomas, o viriones artificiales. En ciertas realizaciones, la molécula de endonucleasa dirigida y polinucleótidos opcionales se introducen dentro de una célula mediante nucleofección o electroporación.

- 20 En realizaciones en las cuales se introducen más de una molécula de endonucleasa dirigida y más de un polinucleótido dentro de una célula, las moléculas pueden introducirse simultáneamente o secuencialmente. Por ejemplo, las moléculas de endonucleasa dirigidas, cada una específica para un sitio de escisión dirigido (y polinucleótidos opcionales) pueden introducirse al mismo tiempo. Como alternativa, cada molécula de endonucleasa dirigida, así como el polinucleótido(s) opcional(es) puede introducirse secuencialmente.

- 25 La relación de la molécula de endonucleasa dirigida (o ácido nucleico que codifica) con respecto al polinucleótido opcional puede variar y varía. En general, la relación de la molécula de endonucleasa dirigida al polinucleótido puede variar desde aproximadamente 1:10 hasta aproximadamente 10:1. En varias realizaciones, la relación de la molécula de endonucleasa dirigida al polinucleótido es aproximadamente de 1:10, 1:9, 1:8, 1:7, 1:6, 1:5, 1:4, 1:3, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, ó 10:1. En una realización, la relación es aproximadamente de 1:1.

(b) mutagénesis aleatoria

- En otras realizaciones aún, la línea de células deficiente en Cmah y Ggta1 puede prepararse usando mutagénesis aleatoria. En una realización, se genera una mutación aleatoria mediante la exposición de la célula a un compuesto químico tal como N-etil-N-nitrosourea (ENU), N-etil-N-nitrosourea (NMU), metanosulfonato de etilo (EMS), ácido
- 35 nitroso (NA), u otro compuesto químico mutagénico. En otra realización, se genera una mutación aleatoria usando un sistema de tipo transposón para insertar aleatoriamente secuencias cortas en el genoma, interrumpiendo, de esta forma, la expresión de la secuencia cromosómica en la cual se ha insertado una secuencia. En otra realización, se genera una mutación aleatoria usando radiación ionizante.

(c) recombinación específica del sitio

- En realizaciones alternativas, la línea de células deficiente en Cmah y Ggta1 puede prepararse usando técnicas de combinación específicas del sitio. Por ejemplo, pueden usarse técnicas de recombinación específicas del sitio para borrar la totalidad o parte de una secuencia cromosómica de interés, o introducir polimorfismo nucleótidos sencillos (SNPs) dentro de la secuencia cromosómica de interés. En una realización, la secuencia cromosómica de interés está dirigida usando un sistema de recombinación específico del sitio Cre-*loxP*, un sistema de recombinación específico del sitio Flp-FRT, o variantes de los mismos. Dichos sistemas de recombinación se encuentran comercialmente disponibles, y, por ejemplo, en Ausubel y otros, véase anteriormente, se encuentran enseñanzas adicionales para estas técnicas.
- 45

(III) Procedimientos de producción de proteínas recombinantes

- Un aspecto adicional de la presente divulgación abarca un procedimiento de producción de una proteína recombinante con un patrón de glucosilación de tipo humano. En general, una glucoproteína que tiene un patrón de glucosilación de tipo humano carece de restos α -Gal y Neu5Gc. El procedimiento comprende la expresión de la proteína recombinante en una línea celular de mamífero no humano deficiente en Cmah y Ggta1. Las líneas celulares deficientes en Cmah y Ggta1 han sido descritas anteriormente en la sección (I).
- 50

- En una realización a modo de ejemplo, la línea celular puede comprender inactivaciones bialélicas de las secuencias cromosómicas que codifican Cmah y Ggta1, de manera tal que la línea celular no produce Cmah y Ggta1 y la proteína recombinante producida por la célula carece de restos α -Gal y Neu5Gc. En una realización a modo de ejemplo adicional, en la cual la línea celular es aneuploide, todas las copias de la secuencia cromosómica que codifica Cmah
- 55

y todas las copias de la secuencia cromosómica que codifica Ggta1 están inactivadas, de manera tal que la línea celular no produce Cmah o Ggta1 y la proteína recombinante producida por la célula carece de restos α -Gal y Neu5Gc.

En general, la proteína recombinante producida por la línea celular deficiente en Cmah y Ggta1 tienen al menos una propiedad que está mejorada con respecto a la misma proteína producida por una línea celular comparable que no es deficiente en Cmah y Ggta1. Los ejemplos no limitativos de propiedades mejoradas incluyen inmunogenicidad reducida, biodisponibilidad incrementada, eficacia incrementada, estabilidad incrementada, solubilidad incrementada, vida media mejorada, aclaramiento mejorado, farmacocinéticas mejoradas, y combinaciones de las mismas. Por ejemplo, dado que la proteína recombinante producida por el procedimiento divulgado en la presente invención carece de restos α -Gal y Neu5Gc, la proteína recombinante producida tiene inmunogenicidad reducida y potencial reducido para la inducción de reacciones de hipersensibilidad en sujetos humanos con respecto a una proteína comparable que contenga restos α -Gal y Neu5Gc.

La proteína recombinante producida en la línea celular deficiente en Cmah y Ggta1 puede ser cualquier proteína adecuada, incluyendo proteínas terapéuticas y productos biológicos proteínicos. Por ejemplo, la proteína recombinante puede ser, sin limitarse a ellas, un anticuerpo, un fragmento de un anticuerpo, un anticuerpo monoclonal, un anticuerpo humanizado, un anticuerpo monoclonal humanizado, un anticuerpo quimérico, una molécula IgG, una cadena pesada IgG, una cadena ligera IgG, una molécula IgA, una molécula IgD, una molécula IgE, una molécula IgM, una glucoproteína, un factor de crecimiento, una citocina, un interferón, una interleucina, una hormona, un factor coagulante (o de coagulación), un componente de la sangre, una enzima, una proteína nutricional, una vacuna, un fragmento funcional o variante funcional de cualquiera de los anteriores, o una proteína de fusión que comprenda cualquiera de las proteínas anteriores y/o fragmentos funcionales o variantes de los mismos.

Los procedimientos de producción de proteína recombinante son bien conocidos en la técnica, y, mediante Ausubel y otros, véase anteriormente, se proporcionan enseñanzas adicionales. En general, la proteína recombinante está expresada a partir de un ácido nucleico introducido exógenamente. Tal como se ha detallado anteriormente en la sección (I)(a), el ácido nucleico que codifica la proteína recombinante puede ser extracromosómico o el ácido nucleico que codifica la proteína recombinante puede estar integrado dentro del genoma.

Los procedimientos para el cultivo de la línea celular, de manera tal que la proteína recombinante esté expresada, son bien conocidos en la técnica. Los medios y sistemas de cultivo apropiados son conocidos en la técnica y se encuentran comercialmente disponibles. En una realización, la proteína recombinante se produce mediante líneas celulares divulgadas en la presente invención mediante cultivo en suspensión exento de suero.

Definiciones

Salvo que se defina de otra manera, todos los términos técnicos y científicos usados en la presente invención tienen el significado comúnmente entendido por una persona experta en la técnica a la cual pertenece la presente invención. Las referencias siguientes proporcionan a un experto una definición general de muchos de los términos usados en la presente invención: Singleton y otros, 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 y otros (eds.), Springer Verlag (1991); y Hale & Marham, The Harper Collins Dictionary of Biology (1991). Tal como se usa en la presente invención, los términos siguientes tienen los significados adscritos a ellos, salvo que se especifique lo contrario.

Cuando se introducen elementos de la presente divulgación o la(s) realización(es) preferida(s) de la misma, los artículos "un", "uno", "el" y "dicho" se entiende que significan que existe uno o más de los elementos. Los términos "comprende", "incluye" y "tiene" están destinados a ser inclusivos y significan que pueden ser elementos adicionales distintos de los elementos listados.

Tal como se usa en la presente invención, el término "secuencia endógena" se refiere a una secuencia cromosómica que es natural de la célula.

El término "secuencia exógena" se refiere a una secuencia cromosómica que no es natural de la célula, o una secuencia cromosómica cuya localización natural está en una localización cromosómica diferente.

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

Un "gen", tal como se usa en la presente invención, se refiere a una región de ADN (incluyendo exones e intrones) que codifica un producto de gen, así como todas las regiones de ADN que regulan la producción del producto de gen, ya estén o no dichas secuencias reguladoras adyacentes a las secuencias de codificación y/o transcritas. De acuerdo con ello, un gen incluye, pero sin estar necesariamente limitado a ellas, secuencias promotoras, terminadores, secuencias reguladoras de traducción tales como sitios de unión de ribosomas y sitios de entrada de ribosomas

internos, potenciadores, silenciadores, aisladores, elementos de unión, orígenes de replicación, sitios de unión de matrices, y regiones de control de locus.

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

Los términos "ácido nucleico" y "polinucleótido" se refieren a un polímero desoxiribonucleótido o ribonucleótido, en conformación lineal o circular, y ya sea tanto en forma de hebra sencilla o doble. Para los fines de la presente divulgación estos términos no deben interpretarse como limitativos con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en la base, azúcar y/o restos fosfato (por ejemplo, cadenas principales fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de apareamiento de bases; es decir, un análogo de A comparte pares de bases con T.

El término "nucleótido" se refiere a desoxiribonucleótidos o ribonucleótidos. Los nucleótidos pueden ser nucleótidos convencionales (es decir, adenosina, guanosina, citidina, timidina, y uridina) o análogos nucleótidos. Un análogo nucleótido se refiere a un nucleótido que tiene una base purina o pirimidina modificada o un resto ribosa modificado. Un análogo nucleótido puede ser un nucleótido que aparece de manera natural (por ejemplo, inosina) o un nucleótido que aparece de manera no natural. Los ejemplos no limitativos de modificaciones sobre el azúcar o resto base de un nucleótido, incluyen la adición (o separación) de grupos acetilo, grupos amino, grupos carboxilo, grupos carboximetilo, grupos hidroxilo, grupos metilo, grupos fosforilo, y grupos tiol, así como la substitución de los átomos de carbono y nitrógeno de las bases con otros átomos (por ejemplo, 7-deaza purinas). Los análogos nucleótidos incluyen, igualmente, dideoxi nucleótidos, 2'-O-metil nucleótidos, ácidos nucleicos bloqueados (LNA), ácidos nucleicos péptidos (PNA), y morfolidos.

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

El término "recombinación" se refiere a un procedimiento de intercambio de información genética entre dos polinucleótidos. Para los fines de la presente divulgación, "recombinación de homólogos" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas de doble hebra en células. Este procedimiento requiere similitud de secuencia entre dos polinucleótidos, el uso de una molécula "donante" o de "intercambio" para reparar el molde de una molécula "diana" (es decir, la que ha experimentado la rotura de doble hebra), y que se conoce de varias formas tales como "conversión de gen de no cruzamiento" o "conversión de gen de segmento corto", dado que conduce a la transferencia de información genética desde el donante a la diana. Sin desea teorizar, dicha transferencia puede implicar la corrección de diferencias del ADN heteroduplex que se forma entre la diana rota y el donante, y/o "reasociación de hebra dependiente de la síntesis", en la cual el donante se usa para resintetizar la información genética que formará parte de la diana, y/o procedimientos relacionados. Dicha recombinación de homólogos especializados da lugar, frecuentemente, a una alteración de la secuencia de la molécula diana, de manera tal que parte o la totalidad de la secuencia del polinucleótido donante se incorpora dentro del polinucleótido diana.

Tal como se usa en la presente invención, los términos "sitio diana" o "secuencia diana" se refieren a una secuencia de ácido nucleico que define una porción de una secuencia cromosómica a editar y a la cual la endonucleasa dirigida está manipulada genéticamente para reconocer, unir, y escindir.

Los términos "más arriba" y "más abajo", se refiere a localizaciones en una secuencia de ácido nucleico con relación a una posición fijada. Más arriba se refiere a la región que está 5' (es decir, cerca de la terminación 5' de la hebra) con relación a la posición y más abajo se refiere a la región que está 3' (es decir, cerca de la terminación 3' de la hebra) con relación a la posición.

Las técnicas de determinación de identidad de secuencia de ácido nucleico y aminoácido son conocidas en la técnica. Típicamente, dichas técnicas incluyen la determinación de la secuencia nucleótida del ARNm para un gen y/o la determinación de la secuencia de aminoácido codificada de este modo, y comparación de estas secuencias con un segundo nucleótido o secuencia de aminoácido. Igualmente, las secuencias genómicas pueden determinarse y compararse de esta manera. En general, la identidad se refiere a una correspondencia de nucleótido-nucleótido o de aminoácido-aminoácido exacta de dos secuencias de polinucleótido o polipéptido, respectivamente. Pueden compararse dos o más secuencias (polinucleótido o aminoácido) mediante la determinación de su porcentaje de identidad. El porcentaje de identidad de dos secuencias, ya sean secuencias de ácido nucleico o aminoácido, es el número de coincidencias exactas entre dos secuencias alineadas divididas por la longitud de las secuencias más cortas y multiplicada por 100. Un alineamiento aproximado para secuencias de ácido nucleico es la proporcionada por el algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics*, vol. 2, pág. 482-489, (1981). Este algoritmo puede aplicarse a secuencias de aminoácidos usando la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*, M. O. Dayhoff, ed., 5 suppl. 3, págs. 353-358, National Biomedical Research Foundation, Washington, D.C., USA, y normalizado por Gribskov, *Nucl. Acids Res.*, vol. 14, (no. 6), págs. 6745-6763, (1986). Un ejemplo de implementación de este algoritmo para determinar el porcentaje de identidad de una secuencia está proporcionado por el Genetics Computer Group (Madison, Wis.) en la instalación de aplicación "BestFit". Otros programas adecuados para el cálculo del porcentaje de identidad o similitud entre secuencias son general-

mente conocidos en la técnica, por ejemplo, otro programa de alineamiento es BLAST, usado con parámetros de defectos. Por ejemplo, BLASTN y BLASTP pueden utilizarse usando los parámetros de defectos siguientes: código genético = estándar; filtro = ninguno; hebra = ambas; corte = 60; esperada = 10; Matriz = BLOSUM62; Descripciones = 50 secuencias; tipo = HIGH SCORE; Bases de datos = no redundante, GenBank+EMBL+DDBJ+PDB+GenBank translations+Swiss protein+Spudate+PIR. Los detalles de estos programas pueden encontrarse en la website GenBank. Con respecto a las secuencias descritas en la presente invención, el intervalo de grados deseados de identidad de secuencias es aproximadamente de 80% hasta 100% y cualquier valor entero entre ellos. Típicamente, el porcentaje de identidades entre secuencias son de al menos del 70-75%, preferiblemente 80-82%, más preferiblemente 85-90%, incluso más preferiblemente 92%, aún más preferiblemente 95%, y lo más preferiblemente 98% de identidad de secuencias.

Puesto que podrían realizarse cambios en las células y procedimientos anteriormente descritos sin apartarse del alcance de la invención, se pretende que todas las materias contenidas en la descripción anterior y en los ejemplos dados más adelante, deben de interpretarse como ilustrativas y no en un sentido limitativo.

Ejemplos

Los ejemplos siguientes ilustran ciertos aspectos de la invención.

Ejemplo 1: Células CHO que contienen el gen Ggta1

Para confirmar la presencia del gen Ggta1 en células CHO K1, se diseñaron cebadores para amplificar las regiones de exones 8 y 9 del gen Ggta1. Mediante PCR se amplificó el ADN procedente de células NSO y CHO K1 de mieloma mudo usando un par de cebadores. Se detectó una banda del tamaño esperado de aproximadamente 300 bp tanto en células NSO y CHO. Los fragmentos PCR aislados a partir de células CHO se secuenciaron y alinearon con el gen Ggta1 de ratón (UniProtKB/Swiss-Prot No. de registro: P23336). La identidad de secuencia fue aproximadamente del 85%.

Se usó PCR cuantitativo para evaluar la expresión ARNm en células CHO con relación a células de ratón. Después de normalización usando actina, una comparación de los valores del ciclo umbral (Ct) entre las células de ratón NSO y las células CHO sugiere que la expresión de Ggta1 fue significativamente menor en células CHO que en células NSO.

Ejemplo 2: Generación de células CHO que comprenden un locus Ggta1 inactivado

Se diseñó un par de ZFNs para identificar una región con exón 9 del gen Ggta1 en células CHO. Las secuencias de Ggta1 se obtuvieron a partir de una secuencia transcríptómica patentada y se verificaron usando RT-PCR. Los ZFNs que dirigen el gen se diseñaron usando un algoritmo patentado, y posteriormente se ensayaron. La transcripción *in vitro* y la poliadenilación del ARNm y formación del casquete se produjeron a partir de ADN plásmido de ZFN, tal como se describe en la información del producto CompoZr® Knockout Zinc Finger Nucleases (ZFN). En resumen, el ADN de ZFN plásmido se alineó, y purificó mediante extracción del ADN con fenol/cloroformo. Para formar el casquete del ADN alineado se usó un MessageMax™ T7 ARCA-Capped Message Transcription Kit (Cell Script Inc.). Para agregar una cola poli(A) se usó un Poly(A) Polymerase Tailing Kit (EpiCenter). El ARNm de ZFN se purificó usando un juego MEGAclean™ (Ambion).

Se usó la línea celular CHOZN (gs -/-) que expresa IgG humano anti-conejo recombinante. Todos los medios de cultivo de células, suplementos y otros reactivos usados se adquirieron de Sigma-Aldrich, salvo que se especifique lo contrario. Antes de la transfección, las células se mantuvieron como cultivos de suspensión en EXCELL® CHO CD Fusión (Sigma-Aldrich) suplementado con L-glutamina 6mM. Las células se sembraron a una concentración de $0,5 \times 10^6$ células/ml en tubos de bioreactor un día antes de la transfección. Para cada transfección, se usaron 1×10^6 células en 150 µl de medio de crecimiento y se usaron 5 µg de cada ARNm de ZFN. Las transfecciones se llevaron a cabo mediante electroporación a 140 V y 950 µF en cubetas de 0,2 cm. Las células de electroporadas se introdujeron en 2 ml de medio de crecimiento en un cultivo estático en placas de 6 pocillos. Las células de control fueron transfectadas falsas.

Los días 3 y 10 post-transfección, las células se retiraron del cultivo y el ADN genómico se aisló usando el GeneElute Mammalian Genomic DNA Miniprep Kit de Sigma-Aldrich. La escisión inducida por ZFN se verificó usando un ensayo de nucleasa Cel-1, tal como se describe en la información del producto CompoZr® Knockout ZFN. Este ensayo se llevó a cabo para determinar la eficacia de la mutación de gen mediada por ZFN, tal como ha sido descrita previamente (Miller JC y otros, Nat. Biotechnol., vol. 25, pág. 778-785, (2007)). El ensayo detecta alelos del locus dirigido que se desvía del tipo salvaje como un resultado de la reparación imperfecta mediada por la unión del extremo no homólogo (NHEJ) de roturas de doble hebra de ADN inducido por ZFN. La amplificación mediante PCR de la región dirigida procedente de una agrupación de células tratadas con ZFN genera una mezcla de tipo salvaje (WT) y amplicones mutantes. La fusión y reasociación de esta mezcla produce desapareados que se forman entre heteroduplexes del WT y alelos mutantes. Mediante la nucleasa topógrafo Cel-1 se escindió un ADN "burbuja" formado en el sitio del desapareamiento, y los productos de escisión pueden resolverse mediante electroforesis de gel. Tal como se muestra en la FIG. 1, estaban presentes dos fragmentos de aproximadamente 215 bp y 100 bp en las células transfectadas de ZFN (bandas 1 y 2), pero ausentes en las células de control transfectadas falsas.

Ejemplo 3: Clonación unicelular y genotipado de células inactivadas Ggta1

Después de la confirmación de la actividad ZFN usando el ensayo de Cel-1, las células transfectadas con ZFN de Ggta1 se clonaron unicelularmente usando dilución limitante. Para ello, las células se sembraron a una concentración de 0,5 células/pocillo usando una mezcla de medio de clonación libre de suero CHO al 80%, medio acondicionado al 20%, y L-glutamina 4 mM. La clonalidad y el crecimiento se verificaron microscópicamente los días 7 y 14 posteriores a la siembra, respectivamente. Los clones con crecimiento se expandieron y genotiparon mediante PCR y secuenciación. Se aislaron un clon Ggta1 (-/-) y cuatro Ggta1 (+/-) que perforan deleciones de diversas longitudes, tal como se detalla más adelante en la Tabla 1. Todas las líneas celulares mostraron características de crecimiento similares a las de la línea celular madre a partir de la cual se obtuvieron.

Tabla 1. Caracterización genotípica de clones inactivados Ggta1

Clon ID	Alelo1	Alelo 2	Genotipo
BF10	Delección de 14 bp	No detectada	Bialélico inactivado
CG3	Delección de 7 bp	Tipo salvaje	Heterozigoto
DA4	Delección de 2 bp	Tipo salvaje	Heterozigoto
DC7	Delección de 11 bp	Tipo salvaje	Heterozigoto
DG5	Delección de 6 bp + delección de 2 bp	Tipo salvaje	Heterozigoto

Ejemplo 4: Generación de células CHO que comprenden un locus Cmah inactivado

Se diseñó un par de ZFNs para identificar una región con exón 5 del del locus Cmah en células CHO (UniProtKB/Swiss-Prot No. de Acceso: Q9WV23; hámster chino). La línea celular CHO K1 se transfectó con 20 µg de ARN que codifica los ZFNs usando procedimientos y métodos convencionales similares a los descritos en el Ejemplo 2. Las células de control se transfectaron con ARN que codifica GFP.

La eficacia de las roturas cromosómicas de doble hebra inducida por ZFN se determinó usando el ensayo de nucleasa Cel-1. Tal como se muestra en la FIG. 2, los ZFNs de Cmah escindieron la diana Cmah en células CHO. La frecuencia de escisión mediada por ZFN puede estimarse comparando la intensidad relativa de los productos de escisión con la intensidad relativa de la banda madre. La frecuencia de escisión, calculada mediante el software de ImageJ, fue de aproximadamente 11%.

Ejemplo 5: Clonación unicelular y genotipado de células inactivadas Cmah

Las células transfectadas con ZFN de Cmah se clonaron unicelularmente usando dilución limitante (tal como se ha descrito anteriormente) o mediante clasificación de células activadas por fluorescencia (FACS). Los clones con crecimiento se expandieron y genotiparon mediante PCR y la secuenciación mediante genotipado reveló que la totalidad de los 20 clones procedentes de este ronda de trabajo fueron Cmah (+/-) que perforan deleciones e inserciones de varias longitudes. A continuación, se agruparon siete clones Cmah (+/-), y se re-transfectaron con ARN con ZFN de Cmah, y las células individuales se clonaron mediante dilución limitante. Mediante PCR y secuenciación se verificaron seis clones procedentes de la segunda ronda de trabajo que eran Cmah (-/-). La Tabla 2 lista loci Cmah inactivados que perforan deleciones e inserciones de varias longitudes. Igualmente, en la Tabla 2 están listados los genotipos de estas líneas celulares inactivadas bialélicas de Cmah. Todas las líneas celulares mostraron características de crecimiento que fueron similares a las de la línea celular madre a partir de la cual se obtuvieron.

Tabla 2. Caracterización genotípica de clones Cmah (-/-)

Clon ID	Alelo 1	Alelo 2	Genotipo
AB10	Inserción de 5 bp	Delección de 12 bp	Bialélico inactivado
AB2	Delección de 17 bp	Delección de 31 bp	Bialélico inactivado
BA4	Delección de 17 bp	Delección de 33 bp	Bialélico inactivado
BD10	Delección de 12 bp	Delección de 12 bp	Bialélico inactivado
BF3	Inserción de 20 bp	Delección de 4 bp	Bialélico inactivado
BG11	Inserción de 4 bp	Delección de 17 bp	Bialélico inactivado

Ejemplo 6: Generación de células inactivadas dobles Cmah/Ggta1

Se transfectó una línea celular clonal con un genotipo Cmah (-/-) confirmado (es decir, AB2) con ZFNs dirigidos a Ggta1 esencialmente tal como se ha descrito anteriormente en el Ejemplo 2. La actividad de ZFN se confirmó usando el ensayo de nucleasa Cel-1. Tal como se muestra en la **FIG. 3**, los productos de escisión se detectaron en las células transfectadas con ZFN, pero no en las células transfectadas falsas.

Las células fueron clonadas unicelularmente usando dilución limitante esencialmente tal como se ha descrito anteriormente en el Ejemplo 3. La PCR anidada se realizó sobre 192 clones, y se identificaron 26 clones como clones inactivados dobles potenciales en base al tamaño del producto PCR. El ADN genómico se aisló a partir de los clones inactivados dobles potenciales, se amplificó mediante PCR, y se secuenció. Cuatro clones tenían cobertura de secuencia 2x, lo que confirmó la delección bialélica de Ggta1 en la base genética inactivada Cmah. La Tabla 3 presenta los genotipos de los clones inactivados dobles Cmah (-/-)/(Ggta1 (-/-)).

Tabla 3. Caracterización genotípica de clones inactivados dobles Cmah/Ggta1

Clon #	Alelos Cmah	Alelos ggta1	Genotipo
8	Delección de 1,17 bp Delección de 2,31 bp	Delección de 1,4 bp Delección de 2,4 bp	Bialélico inactivado doble
19	Delección de 1,17 bp Delección de 2,31 bp	Delección de 1,3 bp Delección de 2,3 bp	Bialélico inactivado doble
17	Delección de 1,17 bp Delección de 2,31 bp	Delección de 1,3 bp Delección de 2,4 bp	Bialélico inactivado doble
1	Delección de 1,17 bp Delección de 2,31 bp	Delección de 1,23 bp Delección de 2,23 bp	Bialélico inactivado doble

REIVINDICACIONES

- 5 **1.** Una línea celular de mamífero no humano que comprende una secuencia cromosómica inactivada que codifica ácido citidino monofosfato-N-acetilneuramínico hidroxilasa (Cmah), y una secuencia cromosómica inactivada que codifica glucoproteína alfa-1,3-galactosiltransferasa (Ggta1), opcionalmente en la que la línea celular es una línea celular de ovario de hamster chino (CHO).
- 2.** La línea celular de mamífero no humano de la reivindicación 1, en la que cada una de las secuencias cromosómicas inactivadas comprende una delección de al menos un nucleótido, una inserción de al menos un nucleótido, una substitución de al menos un nucleótido, o una combinación de los mismos.
- 10 **3.** La línea celular de mamífero no humano de la reivindicación 1 o la reivindicación 2, en la que la secuencia cromosómica inactivada que codifica Cmah comprende una secuencia introducida no exógenamente, y/o en la que la secuencia cromosómica inactivada que codifica Cmah es monoalélica y la línea celular produce una cantidad reducida de Cmah, o en la que la secuencia cromosómica inactivada que codifica Cmah es bialélica y la línea celular no produce Cmah.
- 15 **4.** La línea celular de mamífero no humano de la reivindicación 1 o la reivindicación 2, en la que la secuencia cromosómica inactivada que codifica Ggta1 comprende una secuencia introducida no exógenamente, y/o en la que la secuencia cromosómica inactivada que codifica Ggta1 es monoalélica y la línea celular produce una cantidad reducida de Ggta1, o en la que la secuencia cromosómica inactivada que codifica Ggta1 es bialélica y la línea celular no produce Ggta1.
- 20 **5.** La línea celular de mamífero no humano de una cualquiera de las reivindicaciones 1 a 4, en la que las secuencias cromosómicas están inactivadas con una endonucleasa dirigida.
- 6.** La línea celular de mamífero no humano de una cualquiera de las reivindicaciones 1-5, en la que la línea celular produce al menos una proteína que tiene un contenido reducido o carece de restos de ácido N-glicolilneuramínico (Neu5Gc) y un contenido reducido o carece de restos de galactosa-alfa-1,3-galactosa (alga-Gal), comparado con la misma proteína recombinante producida por una línea celular comparable no deficiente en Cmah o Ggta1.
- 25 **7.** La línea celular de mamífero no humano de una cualquiera de las reivindicaciones 1-6, en la que la línea celular es una línea celular CHO, todas las copias de la secuencia cromosómica que codifica Cmah y todas las copias de la secuencia cromosómica que codifica Ggta1 están inactivadas, y la línea celular no produce Cmah ni Ggta1.
- 8.** La línea celular de mamífero no humano de la reivindicación 7, en la que la célula CHO es una célula CHO/GS⁺.
- 30 **9.** Un procedimiento de producción de una proteína recombinante con un patrón de glucosilación de tipo humano, comprendiendo el procedimiento la expresión de la proteína recombinante en una línea celular de mamífero no humano de cualquier reivindicación precedente.
- 10.** El procedimiento de la reivindicación 9, en el que la proteína recombinante tiene al menos una propiedad que se mejora con relación a la misma proteína recombinante producida mediante una línea celular comparable no deficiente en Cmah y Ggta1, opcionalmente elegida entre inmunogenicidad reducida, biodisponibilidad incrementada, eficacia incrementada, estabilidad incrementada, solubilidad incrementada, vida media mejorada, aclaramiento mejorado, y farmacocinéticas mejoradas.
- 35 **11.** El procedimiento de la reivindicación 9 ó 10, en el que la proteína recombinante se selecciona entre un anticuerpo, un fragmento de anticuerpo, un factor de crecimiento, una citocina, una hormona, y un factor de coagulación.

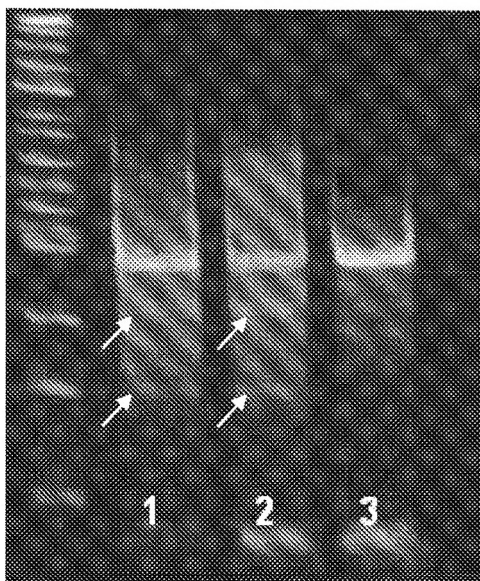


FIG. 1

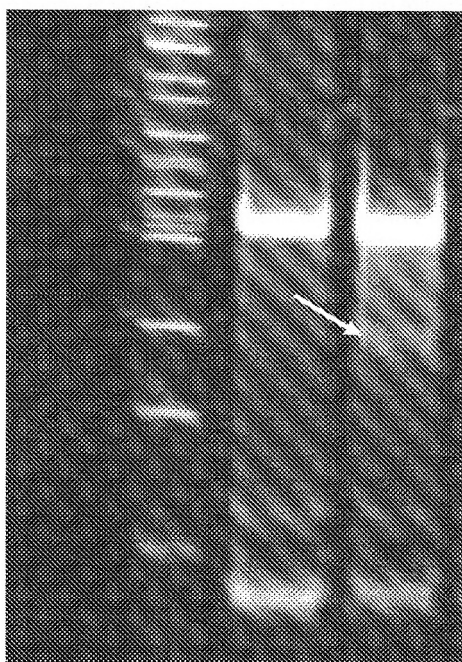


FIG. 2

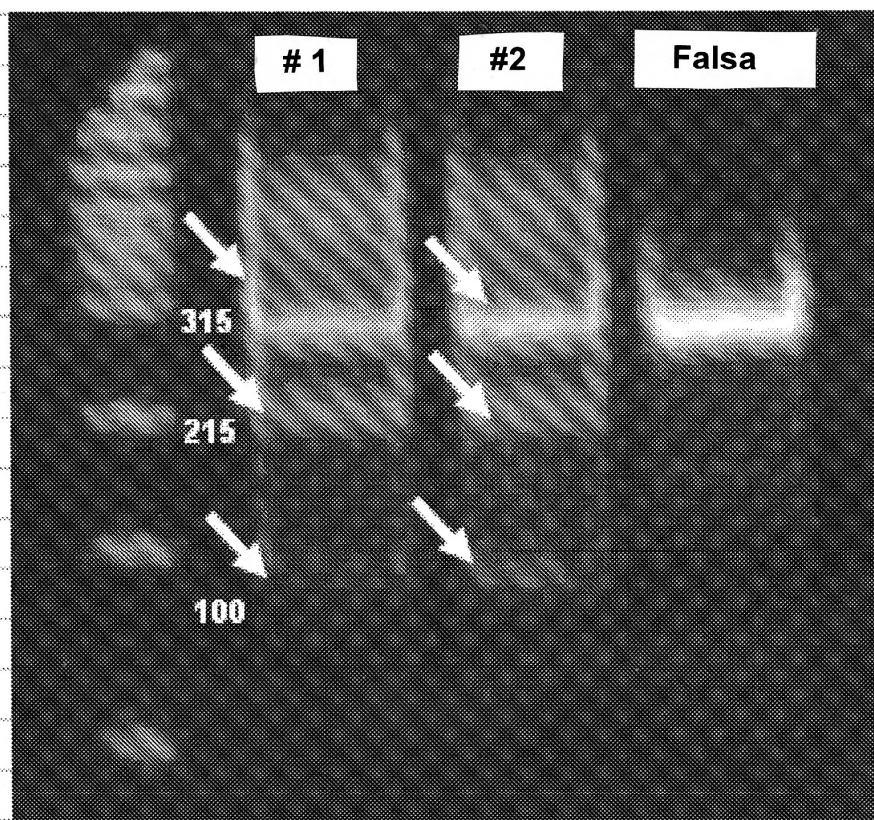


FIG. 3