

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 003 873**

51 Int. Cl.:

**C12N 5/07** (2010.01)

**C12P 21/00** (2006.01)

**C12N 15/87** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.12.2015** **PCT/DK2015/050391**

87 Fecha y número de publicación internacional: **16.06.2016** **WO16091268**

96 Fecha de presentación y número de la solicitud europea: **11.12.2015** **E 15822904 (7)**

97 Fecha y número de publicación de la concesión europea: **04.12.2024** **EP 3230438**

54 Título: **N-glicosilación**

30 Prioridad:

**12.12.2014 US 201462091056 P**

**16.07.2015 US 201562193403 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**11.03.2025**

73 Titular/es:

**UNIVERSITY OF COPENHAGEN (50.00%)**

**Nørregade 10**

**1165 Copenhagen K, DK y**

**DANMARKS TEKNISKE UNIVERSITET (50.00%)**

72 Inventor/es:

**CLAUSEN, HENRIK;**

**YANG, ZHANG;**

**HALIM, ADNAN FEVZI;**

**BENNETT, ERIC;**

**BEHRENS, CARSTEN;**

**VESTER-CHRISTENSEN, MALENE BECH y**

**RAHMAN, SHAMIM HERBERT**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 3 003 873 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

N-glicosilación

5 Campo técnico de la invención

La presente invención se refiere a una célula de mamífero que comprende un gen que codifica un polipéptido de interés, en donde el polipéptido de interés se expresa comprendiendo uno o más patrones de modificación postraducciona

10 anticuerpos con las modificaciones tienen una citotoxicidad celular dependiente de anticuerpos (ADCC) mejorada. Estas modificaciones son útiles, por ejemplo, para mejorar las propiedades farmacocinéticas, es decir, uniendo cadenas de PEG o HEP a las proteínas. La presente invención también se refiere a métodos para producir las glicoproteínas y composiciones que comprenden las glicoproteínas, así como a la ingeniería genómica, el cultivo celular, la producción de proteínas y la glicosilación de proteínas y sus usos.

15

Antecedentes de la invención

La biología de las glicoproteínas es la clase terapéutica de más rápido crecimiento, y la mayoría de ellas sólo pueden producirse de forma recombinante en células de mamífero con capacidad de glicosilación similar a la humana. La célula de ovario de hámster chino (CHO) ha adquirido un papel destacado como célula huésped para la producción recombinante de glicoproteínas terapéuticas, principalmente porque produce N-glicanos bastante simples con ramificación y recubrimiento similares a los que se producen en algunas células humanas.

En particular, los CHO producen N-glicanos heterogéneos de tipo complejo con estructuras bi-, tri- y tetraantenarias con un núcleo de o6-fucosa (Fuc), una cantidad menor de poli-N-acetilactosamina (poli-LacNAc) principalmente en el brazo  $\alpha$ 1,6 de las estructuras tetraantenarias, y un recubrimiento exclusivo de LacNAc con ácido neuramínico (NeuAc) unido  $\alpha$ 2,3.

CHO generalmente no produce estructuras de protección inmunogénicas no humanas y en el hombre, como el ácido N-glicolilneuramínico (NeuGc) o  $\alpha$ 3Gal, aunque se ha informado de la aparición de estos tal vez como resultado de la inducción genética. La N-glicosilación puede variar para diferentes proteínas, así como para diferentes glicositos en proteínas individuales, y, por ejemplo, los anticuerpos IgG se producen con estructuras de N-glicano truncadas en el glicosito Asn297 conservado (estructuras biantenarias con núcleo o6Fuc, LacNAc limitado y protección NeuAc).

35

Una preocupación importante con CHO es la heterogeneidad sustancial en el procesamiento de N-glicano, que puede ser difícil de controlar durante el bioprocesamiento y puede plantear problemas para la bioactividad y la bioseguridad.

Por lo tanto, una actividad importante en el bioprocesamiento de productos terapéuticos se dedica al análisis de glicanos y al control de la fermentación para lograr consistencia.

En las últimas dos décadas se han dedicado esfuerzos sustanciales a la glicoingeniería genética de células CHO con el objetivo de ampliar la capacidad de glicosilación, reducir la heterogeneidad y mejorar o alterar especialmente la sialilación.

Todos estos estudios han usado esencialmente la integración aleatoria de ADNc que codifican glucosiltransferasas y han experimentado problemas de estabilidad, consistencia y previsibilidad de la capacidad de glucosilación introducida. El principal obstáculo ha sido la necesidad de depender de la sobreexpresión de glucosiltransferasas y la competencia con las enzimas endógenas expresadas debido a la falta de métodos sencillos para eliminarlas en líneas celulares.

Por lo tanto, solo unas pocas de estas células CHO modificadas con glicoingeniería han alcanzado la producción de terapias clínicas. Sin embargo, ha surgido una estrategia de glicoingeniería de éxito tras el descubrimiento de que las IgG sin el núcleo o6Fuc en el N-glicano Asn297 presentan una citotoxicidad celular dependiente de anticuerpos (ADCC) notablemente superior.

De este modo, a través de un tour-de-force usando dos rondas de recombinación homóloga, ambos alelos del gen fut8 que codifican la  $\alpha$  6fucosiltransferasa que controla la fucosilación central fueron eliminados en CHO, y al menos una IgG terapéutica producida en CHO sin el gen fut8 ahora está en uso clínico.

Más recientemente, el gen fut8 fue eliminado usando una edición genética precisa con la nucleasa de dedo de zinc (ZFN) dirigida al gen con una fracción del tiempo y de los recursos invertidos.

La aparición de tecnologías precisas de edición genética para knockout (KO) y knockin (KI) ha abierto la puerta a un nivel completamente diferente de velocidad y facilidad con el que se puede lograr una manipulación

genética estable de líneas de células huésped para eliminar e introducir genes de glicosiltransferasa, y esto sin duda tendrá un impacto en la ingeniería de fábricas de células huésped de mamíferos para la producción recombinante de terapias.

- 5 Por lo tanto, existe una necesidad de células de mamíferos, y especialmente de células CHO, que tengan patrones de glicosilación específicos con o sin sialilación.

#### Genes GTf

- 10 Las células de mamíferos tienen una gran cantidad de genes de glicosiltransferasa y se han identificado más de 200 genes distintos y se han determinado parcialmente sus propiedades y funciones catalíticas en los procesos de glicosilación (Ohtsubo y Marth 2006; Lairson, Henrissat et al. 2008; Bennett, Mandel et al. 2012; Schachter 2014). Estos genes están clasificados en familias de genes homólogos con pliegues estructurales relacionados en la base de datos CAZy ([www.cazy.org](http://www.cazy.org)). Las glicosiltransferasas codificadas catalizan diferentes etapas de la biosíntesis de glicosfingolípidos, glicoproteínas, anclajes GPI y proteoglicanos (denominados conjuntamente glicoconjugados), así como de oligosacáridos presentes en las células de mamíferos. Existe una enorme diversidad en las estructuras de los glicanos en estas moléculas, y se han desarrollado vías biosintéticas para diferentes tipos de glicanos (Kornfeld y Kornfeld 1985; Tarp y Clausen 2008; Bennett, Mandel et al. 2012; Schachter 2014), aunque nuestra comprensión de qué enzima(s) glicosiltransferasa(s) catalizan un enlace particular en la biosíntesis del diverso conjunto de glicoconjugados producidos en una célula de mamífero no es completa. La glicosilación en las células es un proceso no impulsado por plantillas que depende de una serie de factores, muchos de los cuales son desconocidos para producir los diferentes glicoconjugados y estructuras de glicanos con un alto grado de fidelidad y expresión y regulación diferenciales en las células. Estos factores pueden incluir: la expresión de las proteínas glicosiltransferasas; la topología subcelular y la retención en la vía secretora RE-Golgi, la síntesis, el transporte y la disponibilidad de donantes de nucleótidos de azúcar en la vía secretora; la disponibilidad de sustratos aceptores; las glicosiltransferasas competidoras; la divergencia y/o enmascaramiento de las vías de glicosilación que afectan la disponibilidad de sustratos aceptores y/o dan como resultado estructuras diferentes; y las condiciones generales de crecimiento y el estado nutricional de las células.

#### 30 Isoenzimas GTf

- Varios genes de glicosiltransferasa tienen un alto grado de similitud de secuencia y se han clasificado en subfamilias que codifican isoenzimas putativas estrechamente relacionadas, que se ha demostrado o se ha predicho que desempeñen funciones relacionadas o similares en la biosíntesis de glicanos en células (Tsuji, Datta et al. 1996; Amado, Almeida et al. 1999; Narimatsu 2006; Bennett, Mandel et al. 2012). Ejemplos de tales subfamilias incluyen las polipéptido GalNAc-transferasas (GalNAc-T1 a 20),  $\alpha$ 2,3sialiltransferasas (ST3Gal-I a VI),  $\alpha$ 2,6sialiltransferasas (ST6Gal-I y II),  $\alpha$ 2,6sialiltransferasas (ST6GalNAc-I a VI),  $\alpha$ 2,8sialiltransferasas (ST8Sia-I a VI),  $\beta$ 4galactosiltransferasas (B4Gal-T1 a 7),  $\beta$ 3galactosiltransferasas (B3Gal-T1 a 6),  $\beta$ 3GlcNAc-transferasas (B3GnT1 a 5),  $\beta$ 6GlcNAc-transferasas (C2GnT1-7 e IGnT2A a C),  $\beta$ 4GalNAc-transferasas (B4GalNAc-T1 a 4),  $\beta$ 3glucuroniltransferasa (B3GlcUA-1 a 3),  $\alpha$ 3/4-fucosiltransferasas (FUT1 a 11), O-fucosiltransferasas (O-POFUT 1 y 2), O-glucosiltransferasas (O-GLUT 1 y 2),  $\beta$ 4GlcNAc-transferasas (MGAT4A a C),  $\beta$ 6GlcNAc-transferasas (MGAT5 y 5B) y hialuronano sintasas (HAS1 a 3) (Hansen, Lind-Thomsen et al. 2014) (Consulte la TABLA 1 para obtener una descripción general de los genes y las proteínas, incluida la nomenclatura usada).

#### Funciones de la subfamilia GTf

- 50 Estas subfamilias de isoenzimas están generalmente mal caracterizadas y las funciones de las isoenzimas individuales no están claras. En la mayoría de los casos, la función de las isoformas se predice a partir del análisis enzimático in vitro con sustratos artificiales y estas predicciones a menudo han resultado erróneas o parcialmente incorrectas (Marcos, Pinho et al. 2004). Las isoenzimas pueden tener funciones diferentes o parcialmente superpuestas o pueden proporcionar respaldo parcial o completo en la biosíntesis de estructuras de glicanos en células en ausencia de una o más glicosiltransferasas relacionadas. Por lo tanto, no es posible predecir de manera confiable cómo la deficiencia de un gen particular en estas subfamilias afectará las vías de glicosilación y las estructuras de glicanos producidas en los diferentes glicoconjugados en una célula y el estado in vitro puede no reflejar el estado in vivo. Es más, incluso genes de glicosiltransferasa más aislados que no se encuentran entre dichas subfamilias, pueden tener funciones desconocidas en la glicosilación o capacidades para tales funciones desconocidas en ausencia de otros genes de glicosiltransferasa.

- 60 Conocimiento a partir de animales knockout - escasa predictibilidad del fenotipo Esto ha sido particularmente evidente en estudios de ratones knockout con deficiencia en genes de glicosiltransferasas que son miembros de subfamilias homólogas, en los que se han observado cambios sorprendentes o ausencia de cambios aparentes en la glicosilación (Angata, Lee et al. 2006).

65

Por ejemplo, los ratones con eliminación dirigida del gen  $\beta 4\text{GalT1}$  que codifica la isoenzima  $\beta 4\text{Gal-T1}$  (uno de los 7 miembros de la familia  $\beta\text{Galactosiltransferasa}$ ) demostraron una pérdida parcial de la  $\beta 4\text{galactosilación}$  de la proteína, con un cambio de cadenas de tipo 2 ( $\text{Gal } \beta 1\text{-4GlcNAc } \beta 3\text{-R}$ ) a cadena de tipo 1 ( $\text{Gal } \beta 1\text{-3GlcNAc } \beta 3\text{-R}$ ) (Kotani, Asano et al. 2004). Se observó un patrón de sialilación correspondiente con represión de la  $\alpha 2,6$  dominante y una mayor sialilación  $\alpha 2,3$  característica de las cadenas tipo 1. Fenotípicamente, los ratones sobrevivieron hasta el término pero mostraron retraso en el crecimiento y diferenciación problemática de los epitelios, así como insuficiencias endocrinas (Asano, Furukawa et al. 1997).

La misma falta de fenotipo global se observa cuando se eliminan miembros individuales de la gran familia de sialiltransferasas responsables de la sialilación de glicanos (Angata, Lee et al. 2006). Por ejemplo, los ratones deficientes en la sialilación de los O-glicanos del núcleo 1 catalizada por la enzima ST3Gal-I son groseramente normales, pero presentan una marcada disminución del número de linfocitos T CD8+ en los compartimentos periféricos. ST3Gal-I es uno de los otros seis miembros de la familia de sialiltransferasas ST3Gal, y los ratones deficientes en otros miembros de la familia también presentan en su mayoría carencias de fenotipos macroscópicos. Una excepción son los ratones que carecen de ST3Gal-IV y que presentan un trastorno hemorrágico grave debido a la falta de sialilación y residuos de galactosa expuestos en el factor von Willebrand y, por lo tanto, su eliminación de la circulación. Además, estos ratones presentan un defecto de adhesión leucocitaria con activación disminuida de los ligandos P y E-Selectina. De manera similar, la eliminación de polipéptidos individuales GalNAc-Ts de la familia de genes GALNT que inicia la O-glicosilación de tipo GalNAc solo produce fenotipos muy discretos que a menudo son difíciles de predecir (Bennett, Mandel et al. 2012). Lo mismo ocurre con la inactivación dirigida de varias GlcNAc-transferasas que catalizan la ramificación de los glicanos O-ligados. Por ejemplo, la pérdida de Core2 GlcNAc-T1 causa solo unas pocas anomalías evidentes, aunque se observa un efecto en las funciones de E, P y L-selectinas. Inesperadamente, se encontró un alargamiento compensatorio de los glicanos del núcleo 1 en ratones knock out para Core2 GlcNAc-T1, rescatando el fenotipo en las vénulas del endotelio alto (Stone, Ismail et al. 2009).

Sin embargo, en algunos casos la eliminación conduce a la pérdida completa de una capacidad de glicosilación particular. En el caso de la O-glicosilación de la sintasa core1, la eliminación de C1GALT1 parece resultar en la pérdida completa de los O-glicanos core1 alargados (Wang, Ju et al. 2010), lo que demuestra que esta función particular de glicosilación solo puede ser llevada a cabo por una enzima. Sin embargo, el análisis detallado de la función de dichos genes aislados a menudo puede resultar difícil si la deficiencia conduce a una letalidad embrionaria temprana, como es el caso de la eliminación de C1GALT1. Otro ejemplo similar es la eliminación de la sintasa GlcCer (UGCG), que provoca la eliminación de todos los glicoesfingolípidos complejos basados en LacCer y la letalidad embrionaria (Jennemann, Sandhoff et al. 2005). Esto también ocurre en la producción de N-glicanos. En varios animales knockout con falta total o casi total de glicanos ligados a N se observan fenotipos macroscópicos sin viabilidad o con viabilidad limitada (Stone, Ismail et al. 2009). El más pronunciado es el efecto por pérdida de la UDP-GlcNAc: dolicol fosfato N-acetilglucosamina-1-fosfato transferasa (GPT) esencial para la iniciación del precursor de oligosacáridos en la N-glicosilación, lo que provoca una falta completa de N-glicosilación y falta de viabilidad más allá de la etapa embrionaria E5.5 de la gestación. Se observan fenotipos graves similares por la pérdida de la glicosiltransferasa GnT1 codificada por *Mgat1*<sup>-/-</sup> que cataliza el paso esencial en la conversión de manosa alta en estructuras híbridas y complejas (Ioffe y Stanley 1994; Schachter 2014). Los ratones knockout se vuelven inviables a mitad de la gestación (E9.5-E10.5) con anomalías morfológicas que incluyen defectos en la formación y vascularización del tubo neural.

La eliminación de otras enzimas que funcionan en la ramificación de los glicanos ligados a N demostró fenotipos murinos de gravedad variable. La eliminación de GlcNAc-TII que cataliza la formación de N-glicanos complejos en el paso posterior a GlcNAc-I y  $\alpha\text{Mannosidasas}$ , demuestran una viabilidad limitada en ratones, aunque se han observado diferencias dependientes del fondo genético (Freeze 2001). Curiosamente, se ha observado un aumento inesperado del nivel de estructuras de Lewis mediante un aumento compensatorio en las estructuras bisectriz de GlcNAc catalizado por GlcNAc-TIII codificado por *Mgat3* (Wang, Schachter et al. 2002). A diferencia de la eliminación de *Mgat1* y *Mgat2*, la eliminación de *Mgat3* y *Mgat4a* que codifican GlcNAc-TIVa son viables sin fenotipos evidentes. El fenotipo de los ratones *Mgat4b* KO actualmente no está claro. Varios estudios han examinado el efecto de la eliminación de *Mgat5* que codifica la transferasa GlcNAc-V. Aunque no se observa ningún fenotipo evidente en ratones más jóvenes, ha quedado claro que los glicanos tetraantennarios catalizados por GlcNAc-V son importantes para la regulación de la activación de las células T, el reclutamiento de leucocitos en la inflamación, así como otras vías de señalización. Así pues, está claro que la eliminación completa de los elementos centrales de determinados glicanos provoca fenotipos graves, mientras que la eliminación de las enzimas de ramificación y recorte suele producir cambios impredecibles en las estructuras de los glicanos y consecuencias funcionales.

El cruce de ratones knockout ha dado como resultado ratones deficientes en hasta dos genes de glicosiltransferasa. Por ejemplo, la cría de ratones deficientes en las sialiltransferasas ST8Sia-I (GD3 sintasa) y la  $\beta 4\text{GlcNAcT}$  (GD2 sintasa), ambas importantes para la biosíntesis de glicoesfingolípidos, provocó una carencia total de todos los tipos de glicoesfingolípidos complejos. Estos ratones doblemente mutantes todavía eran viables pero tenían una alta tasa de mortalidad debido a la extrema sensibilidad a las convulsiones

inducidas audiogénicas (Kawai, Allende et al. 2001). Otros ejemplos de eliminación dirigida de dos glicosiltransferasas son la pérdida de Fut1 y Fut2. Estos animales de doble eliminación se presentan sin fenotipos claros. Además, se han generado ratones con inactivación dirigida tanto de Fut4 como de Fut7. Estos ratones nuevamente no muestran ningún fenotipo obvio, excepto una falta total de actividad de los ligandos de selectina E, P y L en los leucocitos.

#### Conocimiento de las líneas celulares knockout

En contraste con nuestro conocimiento de los efectos de la eliminación de genes de glicosiltransferasa en animales y organismos modelo, existe muy poca información sobre los efectos de la eliminación de genes de glicosiltransferasa en líneas celulares de mamíferos. En las líneas celulares humanas sólo se han identificado unos pocos mutantes espontáneos de genes de glicosiltransferasa. Por ejemplo, la línea celular de cáncer de colon LSC derivada de LS174T tiene una mutación en la chaperona COSMC que conduce a una sintasa central C1GalT mal plegada y no funcional (Ju, Lanneau et al. 2008). Originalmente se propuso que COSMC codificara una glicosiltransferasa sintetasa core1 debido a su homología con el gen C1GALT1, sin embargo, ahora se cree que es una chaperona privada para C1GalT necesaria para el plegamiento (Wang, Ju et al. 2010). El gen COSMC también está mutado en la línea celular linfoblastoide humana Jurkat (Ju, Lanneau et al. 2008; Steentoft, Vakhrushev et al. 2011). En el caso de las líneas celulares no humanas, básicamente sólo existe información de líneas celulares de ovario de hámster chino (CHO), aunque se ha encontrado una línea celular de cáncer murino similar con mutación espontánea en el gen COSMC. Una serie de líneas de células CHO (líneas de células Lee) con deficiencia en un gen de glicosiltransferasa se generaron originalmente mediante mutagénesis aleatoria seguida de selección de lectina, y luego se demostró que los clones mutantes resistentes a la lectina aislados tenían mutaciones definidas en los genes de glicosiltransferasa MGAT1 (Lec), MGAT5 (Lec4) y B4GALT1 (Lec20) (Patnaik y Stanley 2006). También se ha descubierto que varias líneas CHO mutantes tienen mutaciones en genes no glicosiltransferasas que afectan la glicosilación, como la sintasa CMP-NANA (Patnaik y Stanley 2006) y la epimerasa C4-Glc/GlcNAc (Kingsley, Kozarsky et al. 1986). Se ha descubierto que la mayoría de las líneas CHO con mutaciones en los genes de glicosiltransferasa tienen pérdidas parciales de la actividad de glicosilación relevante, aunque se descubrió que la pérdida del gen MGAT5 anula completamente la síntesis de N-glicanos tetraantennarios (North, Huang et al. 2010). Esto último sugiere que solo existe un gen de glicosiltransferasa que codifica la actividad de la  $\beta$ 6GlcNAc-transferasa que conduce a la ramificación tetraantennaria de N-glicano en CHO y que el gen homólogo MGAT5C no sirve como respaldo funcional. Se ha demostrado que el gen MGAT5C está involucrado en la ramificación de los O-glicanos O-Man, aunque se podría predecir o esperar alguna función superpuesta basándose en las similitudes de secuencia de los dos genes MGAT5.

#### Eliminación de genes de glicosilación en líneas celulares

La información limitada sobre los efectos de la eliminación de genes de glicosiltransferasa en líneas celulares se debe en parte a dificultades pasadas para realizar eliminaciones en líneas celulares antes del reciente advenimiento de tecnologías precisas de edición genética (Steentoft, Bennett et al. 2014). Así, hasta hace poco tiempo, esencialmente sólo se había eliminado un gen de glicosiltransferasa, FUT8, mediante un enfoque dirigido usando dos rondas de recombinación homóloga en un esfuerzo de gran envergadura. La disrupción genética convencional por recombinación homóloga es típicamente un proceso muy laborioso como lo evidencia esta eliminación de FUT8 en CHO, ya que se examinaron más de 100.000 líneas celulares clonales para identificar unos pocos clones FUT8-/- en crecimiento (Yamane-Ohnuki, Kinoshita et al. 2004) (US 7214775). Con la llegada de la estrategia de selección de genes mediante nucleasas de dedos de zinc (ZFN), la alteración de genes resulta menos laboriosa. Esto se demostró mediante la eliminación del gen FUT8 en una línea de células CHO, donde otros dos genes adicionales no relacionados con la glicosilación también fueron atacados de manera efectiva (Malphettes, Freyvert et al. 2010).

Más recientemente, ha surgido la estrategia de edición CRISPR/Cas9 y esta estrategia de edición también se utilizó para eliminar el gen FUT8 (Ronda, Pedersen et al. 2014). Además, la eliminación del gen *mgat1* se llevó a cabo mediante la orientación de ZFN en CHO, lo que hizo que la línea CHO fuera incapaz de producir N-glicanos de tipo complejo, y todas las N-glicoproteínas se produjeron solo con estructuras de tipo alto en manosa (Sealover, Davis et al. 2013) (US20140349341). Finalmente, se han usado ZFN para inactivar el gen *Ggta1* (US 20130004992), que codifica un  $\beta$ 3Gal-T que forma el epítipo xenoantígeno Gal $\alpha$ 1-3Gal $\beta$ 1-4GlcNAc que es altamente inmunogénico en el hombre porque el gen GGTA1 humano es un pseudogén y, por lo tanto, esta glicooestructura particular no ocurre en el hombre. En todos los casos de eliminación dirigida de genes de glicosiltransferasa, se había demostrado previamente que los genes objetivo desempeñaban funciones únicas no redundantes, y se podía esperar con una fiabilidad razonable que la eliminación daría lugar a células CHO sin la capacidad de glicosilación particular realizada por la enzima codificada. Sin embargo, como se discutió anteriormente, la mayoría de los glucógenos existen en subfamilias de homólogos cercanos y la función celular y el papel en la glicosilación de las isoenzimas codificadas de estos parálogos son en gran parte desconocidos e impredecibles. Además, la pérdida de cualquier capacidad de glicosilación puede proporcionar sustratos aceptores para glicosiltransferasas no relacionadas e impredecibles que producen estructuras de glucano diferentes a las predichas (Holmes, Ostrander et al. 1987).

Por lo tanto, queda claro que la ingeniería genética de los genes de glicosilación en células de mamíferos y animales es propensa a una incertidumbre sustancial y, por lo tanto, la identificación de los objetivos de ingeniería óptimos para una proteína dada requerirá amplios esfuerzos experimentales. Se puede encontrar alguna guía a partir de los resultados de las especificidades de sustrato determinadas in vitro de las enzimas expresadas recombinantes, que con mayor frecuencia se realizan con formas secretadas truncadas de las enzimas. Si bien estos ensayos proporcionan información general del azúcar donante y los sustratos aceptores más simples, a menudo no brindan especificidades intrincadas del sustrato aceptor con respecto a los aceptores más complejos y el tipo de glicoconjugado en el que trabajan. Se ha demostrado que varias isoenzimas funcionan con pequeños sustratos aceptores artificiales en ensayos in vitro o en algunos casos con grandes bibliotecas de aceptores de glicanos. Sin embargo, las predicciones de funciones in vivo en células basadas en estos estudios a menudo han sido erróneas, como se encontró, por ejemplo, en el caso de ST6GalNAc-II (Marcos, Bennett et al. 2011).

Vale la pena señalar en este contexto que la eliminación de genes de glicosiltransferasa usando diferentes estrategias de silenciamiento de ARNpi ha producido en gran medida información altamente dudosa. En general, la eliminación de genes en vías metabólicas donde las enzimas codificadas a menudo no son limitantes requiere una eliminación altamente eficiente para producir efectos discernibles. Además, la tecnología de ARNpi no es 100% eficiente, lo que inevitablemente resulta en cierta heterogeneidad de las glicoestructuras. Para la producción de productos biofarmacéuticos, incluso un porcentaje bajo de una glicoforma no deseada requiere un análisis y un control exhaustivos.

Además, las glicosiltransferasas son enzimas residentes del RE y del Golgi con un recambio lento, y es muy difícil reducir suficientemente los niveles de enzimas para afectar la glicosilación en un grado detectable, y es esencialmente imposible anular por completo capacidades de glicosilación específicas con estrategias de ARNpi en células de mamíferos.

No obstante, algunos estudios han informado de los efectos del derribo de las capacidades de glicosilación con los consiguientes fenómenos biológicos en las células a pesar de los niveles residuales de enzimas (a menudo confirmados por western blots). Aunque algunos de estos estudios pueden haber observado efectos relacionados con estrategias de silenciamiento no dirigidas, está claro que sólo la eliminación completa de los genes de glicosilación relevantes puede confirmar estos resultados de forma inequívoca, y han aparecido ejemplos de resultados erróneos evidentes. La eliminación del gen GALNT6 involucrado en la O-glicosilación de proteínas en una línea celular de cáncer de mama resultó en la pérdida de la expresión superficial de la mucina MUC1 asociada al cáncer (Bennett, Mandel et al. 2012). Sin embargo, MUC1 se expresa en la superficie de muchas células, incluidas las CHO, que no expresan GALNT6, y por lo tanto no es probable que la pérdida parcial de la enzima codificada GalNAc-T6 tenga efectos tan profundos en la expresión de MUC1. Así pues, está claro que las estrategias de cribado de silenciamiento no son fiables para sondear las funciones celulares in vivo de los genes glicosiltransferasa con el fin de identificar su papel en la glicosilación y evaluar posibles redundancias funcionales parciales o completas en distintas vías de glicosilación.

#### Sobreexpresión de genes de glicosilación en líneas celulares

Además, cabe destacar en este contexto que la sobreexpresión transitoria o estable de un gen de glicosiltransferasa en una célula con mayor frecuencia da como resultado solo cambios parciales en las vías de glicosilación en las que está involucrada la enzima codificada. Varios estudios han intentado sobreexpresar, por ejemplo, la enzima core2 C2GnT1 en CHO para producir O-glicanos ramificados core2, la sialiltransferasa ST6Gal-I para producir ácido siálico de enlace  $\alpha$ 2,6 en las N-glicoproteínas (E! Mai, Donadio-Andrei et al. 2013) y la sialiltransferasa ST6GalNAc-1 para producir ácido siálico unido a  $\alpha$ 2,6 en O-glicoproteínas que forman el glicano STn asociado al cáncer (Sewell, Backstrom et al. 2006). Sin embargo, en todos estos estudios se han obtenido características de glicosilación heterogéneas y a menudo inestables en líneas celulares transfectadas. Se presume que esto se debe en parte a actividades competitivas de la glicosiltransferasa endógena, ya sea que actúen con los mismos sustratos o con sustratos de vías divergentes. Otros factores también pueden explicar las características de glicosilación heterogénea. Esto incluye la disponibilidad y competencia por sustratos donantes y la posible interferencia con otros mecanismos requeridos para la maquinaria de glicosilación endógena, incluyendo, por ejemplo, la topología subcelular y la organización de las glicosiltransferasas, las funciones coordinadas de las glicosiltransferasas, así como la disponibilidad de cofactores, chaperonas y otros factores desconocidos necesarios para una glicosilación eficiente.

#### Función de la glicosilación en las proteínas

La glicosilación de proteínas es un proceso esencial y altamente conservado en todas las células de mamíferos, y el repertorio de genes de glicosiltransferasa disponibles en las células de mamíferos es casi idéntico con relaciones ortólogas bien definidas aclaradas (Hansen, Lind-Thomsen et al. 2014). Sólo unos pocos genes de glicosiltransferasa han sido inactivados a través de la evolución en el hombre, como el gen del xenoantígeno  $\alpha$ 3galactosiltransferasa, GGTA1, y el gen  $\alpha$ 3GalNAc-transferasa de Forssman, mientras que el ser humano ha

adquirido algunas nuevas actividades de glicosiltransferasa mediante la introducción de polimorfismos, como el gen  $\alpha 3\text{Gal/GalNAc-transferasa}$  del grupo sanguíneo ABO. A pesar de estas pequeñas diferencias en las glicosiltransferasas que actúan en el recubrimiento final de las estructuras de glicano, la mayoría de las vías de glicosilación en las células de mamíferos son idénticas y pueden inferirse de una célula de mamíferos a otra con respecto a las funciones de las glicosiltransferasas expresadas. Las proteínas que ingresan a la vía secretora RE-Golgi en células de mamíferos pueden sufrir N-glicosilación, así como una variedad de diferentes tipos de O-glicosilación (Hansen, Lind-Thomsen et al. 2014). Las proteínas también pueden escindirse y unirse mediante un ancla GPI. La glicosilación afecta a las proteínas de varias maneras. La glicosilación puede dirigir el plegamiento con control de calidad para la salida del RE y el transporte y la secreción, y la glicosilación puede dirigir la conformación y la función de las proteínas, así como la sensibilidad para el procesamiento proteolítico y/o la degradación. La glicosilación es especialmente importante para las terapias con proteínas recombinantes producidas en células de mamíferos, y los glicanos pueden afectar a la producción de proteínas, proporcionar solubilidad y estabilidad a las proteínas, afectar a la bioactividad, biodistribución y/o semivida circulatoria de las proteínas, y dirigir las proteínas a receptores específicos de unión a carbohidratos. Por lo tanto, es importante considerar la capacidad de glicosilación de las células huésped usadas para la producción de terapias recombinantes para la producción de terapias efectivas.

#### Producción recombinante de terapias

Las glicoproteínas biológicas son la clase terapéutica de más rápido crecimiento (Walsh 2010), y la mayoría de ellas sólo pueden producirse de forma recombinante en células de mamífero con capacidad de glicosilación similar a la humana. La célula de ovario de hámster chino ha adquirido un papel destacado como célula huésped para la producción recombinante de glicoproteínas terapéuticas, principalmente porque produce N-glicanos bastante simples con ramificación y recubrimiento similares a los que se producen en algunas células humanas (Walsh 2014). En particular, los CHO producen N-glicanos heterogéneos de tipo complejo con estructuras bi-, tri- y tetraantenarias con un núcleo de  $\alpha 6$ -fucosa, una cantidad menor de poli-N-acetilactosamina (poli-LacNAc) principalmente en el brazo  $\alpha 1,6$  de las estructuras tetraantenarias, y un recubrimiento exclusivo de LacNAc con ácido neuramínico  $\alpha 2,3$  (NeuAc) (North, Huang et al. 2010). CHO generalmente no produce estructuras de protección inmunogénicas no humanas y en el hombre, como el ácido N-glicolilneuramínico (NeuGc) o  $\alpha 3\text{Gal}$ , aunque la aparición de estas puede ser resultado de la inducción genética (Bosques, Collins et al. 2010; North, Huang et al. 2010). La N-glicosilación puede variar para diferentes proteínas así como para diferentes glicositos en proteínas individuales, y por ejemplo, los anticuerpos IgG se producen con estructuras de N-glicano truncadas. Una de las principales preocupaciones del CHO es la heterogeneidad sustancial en el procesamiento del N-glicano, que puede ser difícil de controlar durante el bioprocesamiento y plantear problemas de bioactividad y bioseguridad. Por ello, una actividad importante en el bioprocesamiento de terapéuticos se dedica al análisis de glicanos y al control de la fermentación para lograr la coherencia (Walsh 2010).

Las líneas celulares CHO genéricas producen N-glicanos con una mezcla variable de estructuras bi, tri y tetraantenarias con un alto grado de  $\alpha 6$ fucosilación, y un grado menor de estructuras poli-LacNAc y recubrimiento con  $\alpha 2,3\text{NeuAc}$  en la mayoría de las glicoproteínas (Sasaki, Bothner et al. 1987). Además, algunos sitios de N-glicosilación en proteínas pueden tener una estequiometría incompleta. Para IgG, CHO produce solo N-glicanos biantenarios con baja galactosilación y sialilación en el sitio conservado Asn297.

Los O-glicanos del tipo GalNAc se producen con la estructura core1 con  $\alpha 2,3\text{NeuAc}$  y sialilación parcial  $\alpha 2,6$  (a GalNAc) (Fukuda, Sasaki et al. 1989), y también se encuentran estequiometrías variables. Se ha producido una línea de células CHO con COSMC inactivado y esta línea celular produce únicamente estructuras de O-glicanos GalNAc (Yang, Halim et al. 2014). No se han estudiado los O-glicanos de otros tipos y solo se sabe que los O-glicanos del tipo Fuc son elongados por LacNAc y cubiertos con  $\alpha 2,3\text{NeuAc}$ .

No hay líneas celulares de mamíferos disponibles que incluyan CHO que puedan producir una única especie de antena de N-glicano homogénea, ni N-glicanos biantenarios, N-glicanos triantenarios o N-glicanos tetraantenarios. También faltan células que no produzcan pequeñas cantidades de estructuras poli-LacNAc heterogéneas en las ramas antenales. Además, no se dispone de células capaces de producir N-glicanos con recubrimiento  $\alpha 2,6\text{NeuAc}$  homogéneo, y la mayoría de las proteínas plasmáticas humanas están recubiertas por  $\alpha 2,6\text{NeuAc}$  y no por  $\alpha 2,3\text{NeuAc}$  como las producidas por CHO y muchas otras líneas celulares de mamíferos.

La ingeniería genética de células CHO con respecto a la eliminación de genes de glicosiltransferasa se limita a genes aislados individuales. No hay ejemplos de inactivación apilada de múltiples genes de glicosiltransferasa en células de mamíferos, incluidas las líneas CHO. Además, no ha habido ejemplos de inactivación apilada de múltiples genes de glicosiltransferasas que codifiquen isoenzimas con posibles funciones relacionadas en la glicosilación de proteínas. Dado que muchos pasos en la glicosilación de proteínas involucran familias de isoenzimas y sus funciones in vivo en las células son en gran medida desconocidas, además no es posible predecir qué genes inactivar y/o introducir para obtener muchas de las capacidades de glicosilación preferidas para las terapias con glicoproteínas recombinantes.

En las dos últimas décadas se han dedicado importantes esfuerzos a la glicoingeniería genética de células CHO con el objetivo de ampliar la capacidad de glucosilación, reducir la heterogeneidad y mejorar o alterar especialmente la sialilación (Sinclair y Elliott 2005; Walsh y Jefferis 2006). Estos estudios se basan en décadas de trabajo para descifrar las vías biosintéticas y los genes involucrados en la N-glicosilación (Kornfeld y Kornfeld 1985; Schachter 1991), y esencialmente todos han usado la integración aleatoria de ADNc que codifican glicosiltransferasas y han experimentado problemas con la estabilidad, la consistencia y la previsibilidad de la capacidad de glicosilación introducida (Kramer, Klausning et al. 2010). El principal obstáculo ha sido la necesidad de depender de la sobreexpresión de glicosiltransferasas con la consiguiente competencia por parte de las enzimas expresadas endógenamente. Ha sido difícil realizar la eliminación dirigida de genes en líneas celulares en estrategias de silenciamiento, pero en el caso de los genes glicosiltransferasa se ha logrado el éxito con una estrategia basada en la mutagénesis aleatoria en CHO seguida de una selección resistente a la lectina (Patnaik y Stanley 2006). Sin embargo, estas células mutantes CHO en general no tienen características de glicosilación preferibles para terapias recombinantes, también pueden contener otras mutaciones desconocidas en el genoma CHO y ninguna de ellas se ha usado para la producción de terapias clínicas.

El descubrimiento de la importancia de la  $\alpha$ 6Fucosilación central del N-glicano Asn297 en las IgG para la citotoxicidad celular dependiente de anticuerpos (ADCC) (Shinkawa, Nakamura et al. 2003), motivó la necesidad de producir células CHO con esta capacidad de glicosilación y las líneas CHO con el gen *fut8* inactivado ahora están disponibles para la producción de anticuerpos IgG. Además, se han generado líneas CHO con inactivación del gen *ggta1* o del gen *mgat1*.

Sin embargo, hasta ahora la ingeniería de CHO no ha podido abordar los problemas más comunes de heterogeneidad en la glicosilación, sialilación insuficiente e introducción de nuevas capacidades de glicosilación como una sialilación más parecida a la humana ( $\alpha$ 2,6NeuAc). Actualmente, la industria de bioprocesamiento para la producción recombinante de terapias de glicoproteínas solo tiene una línea celular CHO genérica con su capacidad de glicosilación natural disponible para la expresión, con la excepción de los ejemplos mencionados anteriormente, de los cuales solo la línea knockout *fut8* es valiosa para la terapéutica clásica.

KANDA YUTAKA ET AL: "Comparison of cell lines for stable production of fucose-negative antibodies with enhanced ADCC ", BIOTECHNOLOGY AND BIOENGINEERING, WILEY & SONS, HOBOKEN, NJ, US, vol. 94, no. 4, 1 de julio de 2006 (2006-07-01), páginas 680-688, XP002415543, ISSN: 0006-3592, DOI: 10.1002/BIT.20880 describe células CHO knock out de FUT8, y su uso para la producción de anticuerpos terapéuticos fucosa-negativos (resumen; página 681).

WONG ATHENA WET AL: "Enhancement of DNA uptake in FUT8- deleted CHO cells for transient production of afucosylated antibodies.", BIOTECHNOLOGY AND BIOENGINEERING 1AGO 2010, vol. 106, núm. 5, 1 de agosto de 2010 (2010-08-01), páginas 751-763, XP002756439, ISSN: 1097-0290 describe células CHO con FUT8 eliminado para la producción de anticuerpos afucosilados (resumen; páginas 753-755).

El documento WO 2009/009086 A2 (SANGAMO BIOSCIENCES INC [EE. UU.]; COLLINGWOOD TREVOR [EE. UU.]; COST GREGORY J) 15 de enero de 2009 (15 de enero de 2009) describe líneas celulares que comprenden un gen FUT8 alterado, por ejemplo, un gen FUT8 eliminado, usado para la producción de proteínas recombinantes, tales como, por ejemplo, anticuerpos. La línea celular puede ser CHO, NSO, HEK293 o PER (resumen; párrafos 7, 15-19, 31-33, 75, 122-125; ejemplo 4).

A. BUFFONE ET AL: "Silencing 1,3-Fucosyltransferases in Human Leukocytes Reveals a Role for FUT9 Enzyme during E-selectin-mediated Cell Adhesion", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 288, n.º 3, 28 de noviembre de 2012 (2012-11-28), páginas 1620-1633, XP55264669, EE. UU., ISSN: 0021-9258, DOI:10.1074/jbc.M112.400929 describe células knock out para FUT4, FUT7 y/o FUT9 (resumen; páginas 1621, 21628-1629).

S. SHI TE AL: "Inactivation of the Mgat1 Gene in Oocytes Impairs Oogenesis, but Embryos Lacking Complex and Hybrid N-Glycans Develop and Implant", MOLECULAR AND CELLULAR BIOLOGY, vol. 24, no. 22, 15 de noviembre de 2004 (2004-11-15), páginas 9920-9929, XP055264616, US, ISSN: 0270-7306, DOI: 10.1128/MCB.24.22.9920-9929.2004 describe delecciones del gen Mgat1 en ovocitos para identificar N-glicanos complejos (resumen; página 9922).

BHATTACHARYA RIDDHI ET AL: "Biological consequences of inactivating the Mgat3 gene that encodes GlcNAc-TII", GLYCOBIOLOGY, vol. 10, no. 10, octubre de 2000 (2000-10), página 1081, XP9189462, &5.<sup>a</sup> CONFERENCIA ANUAL DE LA SOCIEDAD DE GLYCOBIOLOGY; BOSTON, MASSACHUSETTS, EE. UU.; 8-11 DE NOVIEMBRE DE 2000, ISSN: 0959-6658 describe células que comprenden un gen Mgat3 inactivado.

El documento WO 2013/106515 A1 (SIGMA ALDRICH CO LC [EE. UU.]; LNI NAN [EE. UU.]; SEALOVER NATALIE [EE. UU.]; GEORGE) 18 de julio de 2013 (18 de julio de 2013) describe líneas celulares deficientes

en ni Mgat1. La línea celular puede ser, por ejemplo, una línea celular CHO, BHK, HEK293 o NS0. El gen podría estar inactivado. La inactivación del gen Mgat1 se obtiene mediante la acción de una endonucleasa dirigida ZFN, TALEN,... Métodos para preparar las líneas celulares deficientes en Mgat1 y para producir glicoproteínas recombinantes usando dichas líneas celulares (resumen; párrafos 4-6, 17, 18, 30-44, 79-81; ejemplos 1, 3, 7).

#### Compendio de la invención

La presente invención se refiere a una célula de mamífero en la que mgat4A, mgat4B y mgat5 están eliminados en la célula.

En otra realización de la presente invención, la célula se deriva del ovario de hámster chino o del riñón humano.

En una realización adicional de la presente invención, la célula se selecciona del grupo que consiste en CHO, NS0, SP2/0, YB2/0, CHO-K1, CHO-DXB11, CHO-DG44, CHO-S, HEK293, HUVEC, HKB, PER-C6, NS0 o derivados de cualquiera de estas células. En una realización de la presente invención la célula es una célula CHO.

En otra realización de la presente invención, la célula codifica además una proteína exógena de interés.

En otra realización más de la presente invención, la proteína de interés es un anticuerpo, un fragmento de anticuerpo, tal como un fragmento Fab, un dominio Fc de un anticuerpo o un polipéptido.

En otra realización más de la presente invención, la proteína de interés es un factor de coagulación tal como el factor de coagulación II (FII), el factor de coagulación V (FV), el factor de coagulación VII (FVIIa), el factor de coagulación VIII (FVIII), el factor de coagulación IX (FIX), el factor de coagulación X (FX) o el factor de coagulación XIII (FXIII).

En una realización adicional de la presente invención, el anticuerpo es un anticuerpo IgG.

En una realización de la presente invención se encuentra el polipéptido eritropoyetina (EPO), una proteína involucrada en la hemostasia, incluyendo un factor de coagulación.

En una realización de la presente invención, la glicosilación se hace más homogénea.

En otra realización de la presente invención la glicosilación es no sialilada.

En otra realización de la presente invención, la glicosilación no es galactosilada.

En una realización de la presente invención comprende la glicosilación de N-glicanos biantenarios.

En otra realización de la presente invención, la glicosilación no comprende poli-LacNAc.

En otra realización más de la presente invención la glicosilación es cualquier combinación sin fucosa.

En una realización adicional de la presente invención se ha eliminado FUT8.

En una realización de la presente invención, se ha eliminado B4galt1 permitiendo la generación de N-glicanos más homogéneos sin galactosa.

En otra realización de la presente invención se han eliminado B4galt1 y fut8 permitiendo la generación de N-glicanos más homogéneos sin galactosa y fucosa.

Un aspecto de la presente invención se refiere a un método para producir una glicoproteína según la presente invención.

En una realización de la presente invención, las células han sido modificadas mediante la eliminación del gen de la glicosiltransferasa y/o la inserción de una secuencia de ADN exógena que codifica una glicosiltransferasa.

Otro aspecto de la presente invención se refiere a una glicoproteína según la presente invención, que es un conjugado de glicoproteína homogénea que comprende un polímero seleccionado entre PEG, HEP, XTEN, PSA, HES.

#### Breve descripción de las figuras

La Figura 1 muestra la selección de ZFN en CHO para definir genes clave de glicosiltransferasa involucrados en la N-glicosilación usando EPO humana como reportero. (a) Representación gráfica de un N-glicano tetraantenario común con poli-LacNAc en la antena  $\beta 6$  y protección con ácidos siálicos. Las flechas indican genes de glicosiltransferasa con funciones potenciales en la biosíntesis de cada paso. Se indican las designaciones de los monosacáridos según el Consorcio de Glicómica Funcional. (b) Perfil glicolítico de EPO expresado en células CHO GS WT. Espectros MALDI-TOF de N-glicanos permétrificados liberados por la PNGasa F con etiquetado de cuatro especies principales. (c) Glicoprofilado de EPO expresada en células CHO GS con KO de genes implicados en la biosíntesis tri- y tetraantenaria, mostrando que el triple KO de mgat4A/4B/5 da lugar a N-glicanos biantenarios homogéneos con una cantidad menor de poli-LacNAc (panel inferior). (d) Glicoprofilado con KO de genes  $\beta 4$ galactosiltransferasa implicados en la biosíntesis de LacNAc, mostrando que sólo el KO de B4galt1 por sí solo produjo una pérdida sustancial aunque parcial de galactosilación, mientras que sólo el KO apilado de B4galt1/3 resultó en una pérdida casi completa de galactosilación. (e) Glicoprofilado con KO de dos genes  $\beta 3$ GlcNAc-transferasa con papeles reivindicados en la biosíntesis de poli-LacNAc, mostrando que KO de B3gnt2 resulta en pérdida completa de poli-LacNAc. (f) Glicoprofilado con KO de tres genes  $\alpha 2,3$ sialiltransferasa con supuestas funciones en la N-glicosilación, mostrando que sólo el doble KO de st3gal4/6 resultó en la pérdida completa de la cobertura de ácido siálico.

La Figura 2 muestra el diseño de novo de la capacidad de sialilación en CHO GS mediante knockin dirigido por ZFN. (a) Glycoprofilado de EPO expresada en WT, doble st3gal4/6 KO, y st3gal4/6 KO con ST6GAL1 KI, mostrando que el rango de estructuras N-glicano tetraantenarias producidas en CHO WT recuperan la sialilación en el mismo grado tras KI de ST6GAL1. (b) Glicoprofilado de EPO con triple mgat4A/4B/5 y B3gnt2 KO, mostrando la pérdida completa de poli-LacNAc en N-glicanos biantenarios. (c) Glicoprofilado de mgat4A/4B/5 y st3gal4/6 KO sin y con KI de ST6GAL1, mostrando la pérdida completa del recubrimiento de ácido siálico en los N-glicanos biantenarios y el recubrimiento completo de novo por ácido  $\alpha 2,6$  siálico en los N-glicanos biantenarios, respectivamente.

La Figura 3 muestra una representación gráfica de la matriz de diseño para la ingeniería genética de la capacidad de N-glicosilación de CHO para producir N-glicanos biantenarios homogéneos con recubrimiento de  $\alpha 2,3$ NeuAc, sin sialilación, y con recubrimiento de  $\alpha 2,3$ NeuAc en EPO. Un KO adicional de fut8 provocaría la pérdida de la fucosilación del núcleo.

La Figura 4 muestra el perfil de expresión de genes de glicosiltransferasa seleccionados por ARNsec en CHO. Se realizó un análisis de ARNsec en dos líneas CHO comunes (CHO-K1, CHO-GS) y dos clones triples mgat4A/4B/5 KO CHO GS independientes (cl#1, cl#2). Se incluye el análisis de ARNsec informado de CHO K1, que fue en gran medida idéntico a nuestros análisis, excepto que los niveles de expresión relativa fueron más altos. Es importante destacar que algunos genes, entre ellos mgat4b, que no se expresaban en CHO K1, se expresaban en el presente estudio, y mgat4b resultó ser esencial para los experimentos de glicoingeniería descritos aquí. Otra observación importante fue que los perfiles de expresión de los glucógenos en los dos análisis de clones triple KO eran idénticos a los de las líneas CHO parentales, lo que aporta pruebas de que la edición génica precisa no altera la expresión de otros glucógenos ni siquiera por razones funcionales compensatorias. Para una evaluación más concluyente de esto, necesitamos claramente obtener datos más amplios de RNAseq para todos los clones mutantes.

La Figura 5 muestra la inmunocitología de clones knockout de CHO con mutaciones relacionadas con la ramificación de N-glicanos y la biosíntesis de poli-LacNAc. (a) Se utilizó la lectina L-PHA para sondear específicamente la ramificación  $\beta 6$  de los N-glicanos y sólo el KO de mgat5 produjo (mgat4A/4b/5) alteraciones en el etiquetado L-PHA, mientras que la lectina de control ConA que une todos los tipos de N-glicanos no se vio afectada. Esto confirma que MGAT4A y 4B forman la rama  $\beta 4$ , mientras que MGAT5 forma la rama  $\beta 6$  y, por lo tanto, respalda la interpretación de los datos de espectrometría de masas presentados en la Figura 1c con respecto a las asignaciones de ramas. (b) Se utilizó una estrategia similar para investigar el poli-LacNAc con el LEL, donde sólo la KO de B3gnt2 resultó en la pérdida del etiquetado. Estos datos también respaldan la interpretación de los datos de espectrometría de masas presentados en la Figura 2 con respecto a la asignación de LacNAc a estructuras biantenarias.

La Figura 6 muestra la inmunocitología de clones knockout de CHO con mutaciones relacionadas con la formación de LacNAc. (a) Se utilizó el MAb 1B2 para marcar la presencia de LacNAc tras la eliminación de los ácidos siálicos mediante pretratamiento con neuraminidasa. El análisis de KO de B4galt1-4 en CHO GS WT con ramificación heterogénea mostró que KO de B4galt1 redujo el etiquetado, mientras que sólo KO apilado de B4galt1/3 abolió completamente el etiquetado. (b) El mismo análisis de KO de B4galt1-4 en CHO GS diseñado con N-glicanos biantenarios homogéneos (KO de mgat4A/4B/5) mostró que KO de B4galt1 solo abolió esencialmente el etiquetado de MAb 1B2, lo que indica que  $\beta 4$ Gal-T1 es la enzima principal responsable de la galactosilación de N-glicanos biantenarios. Se espera que este hallazgo tenga implicaciones para la producción recombinante de IgG terapéuticas en CHO, ya que estas proteínas en general se producen con N-glicanos biantenarios con galactosilación limitada.

La Figura 7 muestra la inmunocitología de los clones knockout y knockin de CHO relacionados con la sialilación. (a) El MAb 1B2 se utilizó para marcar la exposición de LacNAc como resultado de la pérdida de la capacidad de sialilación, y esto demostró que la KO de st3gal3 y st3gal6 no afectó sustancialmente el marcado, mientras que la KO de st3gal4 y la KO apilada de st3gal4/6 dieron como resultado un marcado fuerte. Esto sugiere que la enzima ST3Gal-IV es la principal sialiltransferasa responsable de la  $\alpha$ 2,3-sialilación de los N-glicanos. Además, la introducción de novo de ST6Gal-1 humana en el clon KO st3gal4/6 apilado abolió completamente el etiquetado MAb 1B2, demostrando una sialilación eficiente por parte de esta enzima. (b) Para confirmar que la introducción de novo de ST6Gal-1 humana resultó en  $\alpha$ 2,6sialilación, usamos la lectina SNA, que no marcó las células CHO WT como se esperaba y sólo marcó tras la introducción de ST6GAL1. Este resultado respalda las interpretaciones de espectrometría de masas presentadas en la Figura 2.

La Figura 8 muestra el análisis de la knockin dirigida de ST6GAL1 por PCR de unión. (a) Utilizamos una estrategia de KI dirigida de Obligare modificada usando dos sitios de unión ZFN invertidos que flanquean el gen ST6GAL1 en el plásmido donante. La PCR de unión 5' y 3' confirmó la integración dirigida en el sitio Safeharbor#1 en el clon CHO con st3gal4/6 KO. (b) La PCR de unión 5' y 3' confirmó la integración dirigida en el clon CHO con st3gal4/6/mgat4A/4B/5 KO. El estado del número de copias alélicas de integración se determinó mediante PCR alélica WT. La presencia de la banda deseada en la PCR alélica WT indica la presencia del sitio Safeharbor#1 sin la integración del KI dirigido del gen de interés en al menos uno de los alelos. Los resultados mostraron la integración bialélica de st6gal1 al sitio Safeharbor#1 en st3gal4/6 KO y la integración monoalélica a mgat4A/4B/5/st3gal4/6 KO, respectivamente.

La Figura 9 muestra el análisis de Coomassie mediante SDS-PAGE de EPO recombinante purificada expresada en clones CHO. Se cargó aproximadamente 1  $\mu$ g de EPO purificada en función de OD.

La Figura 10 muestra el glicoprofilado de la EPO producida en CHO-GS WT, B3gnt2/mgat4A/4B/5, y st3gal4/6/mgat4A/4B/5 KO con KI de ST6Gal-I, mostrando la pérdida completa de poli-LacNAc en los N-glicanos biantenarios, y el recubrimiento completo de novo por ácido  $\alpha$ 2,6 siálico en los N-glicanos biantenarios sin poli-LacNAc. Se utilizó el glicoprofilado ESI-MS de iones negativos para validar los datos del perfil de glicanos basado en MALDI-TOF.

La Figura 11 muestra ESI-MS/MS de iones negativos de los iones precursores como m/z 1183,92 (N-glicano biantenario sialilado con fucosa central) de EPO producido en CHO con (A) mgat4A/4B/5 KO y con (B) tanto mgat4A/4B/5 como st3gal4/6 KO y KI de ST6Gal-I. La presencia de los iones fragmen diagnósticos a m/z 306,12 indica una sialilación terminal  $\alpha$ 2,6.

Figura 12. Una representación gráfica de las familias CAZy GT humanas involucradas en la biosíntesis de los principales glicoconjugados de mamíferos y sus diferentes vías de glicosilación. Se representan las estructuras de glicanos comunes de mamíferos para (A) O-glicanos que incluyen O-GalNAc, O-GlcNAc, O-Fuc, O-Glc, O-Man, O-Xyl; (B) N-glicanos; (C) anclajes GPI; (D) C-glicanos; (E) glicoesfingolípidos; (F) hialoronano y (G) hidroxilisina. Se indican las denominaciones de los monosacáridos según el Consorcio de Glicómica Funcional.

Figura 13 Una representación gráfica de las diferentes estructuras de glicanos y vías de glicosilación encontradas en líneas celulares CHO con designaciones de genes de glicosiltransferasa expresados en líneas celulares CHO y su papel previsto en los pasos bioestrictos. Los genes están asignados a las principales funciones confirmadas o putativas en las vías de O-glicosilación (O-GalNAc (A1), O-Man (A2) y O-Fuc, O-Glc, O-GlcNAc, O-Xyl (A3)) y N-glicosilación relevantes para la terapéutica de las glicoproteínas recombinantes en la actualidad. (A1-4) Vía de glicosilación O-GalNAc con estructuras extendidas de núcleo 1, núcleo 2, núcleo 3 y núcleo 4; (B) Vía de N-glicosilación; (C) Vías de biosíntesis de glicoesfingolípidos; (D) C-glicano, y (E) GPI-anclajes. Se anotan las familias CAZy de glicosiltransferasas (genes GT) y se muestran los genes GT expresados en líneas celulares CHO. Las estructuras de glicano no sintetizadas en CHO están enmarcadas. Se indican las denominaciones de los monosacáridos según el Consorcio de Glicómica Funcional.

Figura 14 Representaciones gráficas de todos los genes que codifican isoenzimas con potencial para regular la ramificación, elongación y protección terminal de la N-glicosilación expresadas en CHO. La representación muestra la pantalla de eliminación de genes de glicosiltransferasa involucrados en la N-glicosilación en CHO usando EPO humana como glicoproteína reportera expresada recombinante. (A) Se muestra el N-glicano tetrantenario común con poli-LacNAc en la antena  $\beta$ 6 y cubierto por ácidos siálicos, los genes knockout están enmarcados y los genes no expresados en las células CHO están anotados. Obsérvese que la ST6Gal-I humana (ST6GAL1) se ha introducido mediante knock-in mediado por ZFN. Se indican las denominaciones de los monosacáridos según el Consorcio de Glicómica Funcional (CFG). (B) Representación esquemática de las regiones diana reales de las nucleasas en relación con la estructura de dominio general predicha de las proteínas glicosiltransferasas de tipo II (panel superior) y el exón diana para cada gen numerando los genes respectivos a partir del primer exón codificante (panel inferior). La representación de la estructura del exón solo incluye los primeros exones seleccionados y no incluye todos los exones de todos los genes.

Figura 15 Glicoperfil de EPO expresado en células CHO WT. Los espectros MALDI-TOF de PNGasa F liberaron N-glicanos permetilados con estructuras predichas para las cuatro especies principales.

Figura 16 Glicoperfilado de EPO expresado en células CHO con KO de genes involucrados en la biosíntesis de N-glicano de tipo complejo y la formación de antenas, que muestra que el doble KO de mgat4B/5 (B) y el triple KO de mgat4A/4B/5 (B) dan como resultado N-glicanos biantenarios homogéneos con una cantidad menor de poli-LacNAc. Dado que el KO simple de mgat4A (A) tuvo efectos menores comparado con WT es probable que cantidades menores de estructuras triantenarias estén presentes en el doble KO mgat4B/5.

Figura 17 Glicoprofilado con KO de los genes B4galt1/2/3/4 implicados en la biosíntesis de LacNAc, mostrando que sólo el KO de B4galt1 solo produjo una pérdida sustancial aunque parcial de galactosilación, mientras que sólo el KO apilado de B4galt1/3 produjo una pérdida casi completa de galactosilación, mientras que el KO apilado de B4galt1/3 produjo una pérdida casi completa de galactosilación. El KO apilado de B4galt1/2/3 resultó en una pérdida esencialmente completa de la galactosilación.

Figura 18. Perfil glicosilado con KO de tres genes de  $\beta$ 3GlcNAc-transferasa que muestra que la KO de B3gnt2 produce una pérdida completa de poli-LacNAc, mientras que la KO de B3gnt1 y B3gnt1 no tuvo efectos. El PANEL inferior muestra que la KO de B3gnt2 en combinación con mgat4A/4B/5 da como resultado una pérdida completa de poli-LacNAc en estructuras biantenarias.

Figura 19 Glicoprofilado con KO de tres genes  $\alpha$ 2,3sialiltransferasa con funciones reivindicadas en la N-glicosilación, mostrando que sólo el doble KO de st3gal4/6 dio lugar a la pérdida completa del recubrimiento de ácido siálico.

Figura 20 Citología de inmunofluorescencia con tinción de lectina ConA y L-PHA que muestra pérdida de L-PHA con eliminación de mgat5.

Figura 21 Citología de inmunofluorescencia con un anticuerpo monoclonal (clon 1B2) contra LacNAc. El panel (A) muestra sorprendentemente que sólo la KO apilada de B4galt1/3, y no de B4galt1/2/4, eliminó la galactosilación. El panel (B) muestra que la KO de B4galt1 en CHO-GS con KO de mgat4A/4B/5 eliminó la inmunorreactividad para LacNAc mientras que la KO de B4galt2, 3 y 4 en el mismo CHO-GS no tuvo efectos sustanciales.

Figura 22 La expresión recombinante de una IgG terapéutica en CHO con KO de B4galt1 dio lugar a N-glicanos biantenarios homogéneos sin galactosilación. Glicoperfilado de IgG producido en CHO WT y B4galt1 KO que muestra una pérdida esencialmente completa de la galactosilación y sialilación incompletas características del N-glicano conservado en Asn297. La KO de B4galt1 en combinación con fut8 resultó en N-glicanos sin galactosilación y fucosa.

Figura 23 La citología de inmunofluorescencia con lectina LEL muestra que B3gnt2 es el gen clave que controla la iniciación de LacNAc en células CHO.

Figura 24 Citología de inmunofluorescencia con anticuerpo monoclonal contra LacNAc (1B2). La eliminación de St3gal3, St3gal4 y St3gal6 individualmente mediante inmunocitología para evaluar la exposición de LacNAc muestra que sólo el KO de St3gal4 produjo una exposición sustancial de LacNAc.

Figura 25 Glicoprofilado de EPO expresada en CHO con KO de mgat4a/4b/5 y knockin mediado por ZFN de ST6GAL1 humano. El panel A muestra el perfil de EPO con recubrimiento homogéneo de  $\alpha$ 2,6NeuAc de la gama completa de estructuras antenales de N-glicano. También se muestra la KI de ST6GAL1 en CHO con KO de mgat4a/4b/5/B3gnt2, que produjo estructuras tetraantenarias más homogéneas. El panel B muestra el perfil con KO adicional de mgat4a/4b/5 donde la EPO se produce con N-glicanos biantenarios homogéneos cubiertos por  $\alpha$ 2,6NeuAc. También se muestran KI de ST3Gal-4 humana en combinación con KO de st3gal4/6/mgat4a/4b/5/B3gnt2 donde la EPO se produce con N-glicanos biantenarios homogéneos sin poli-LacNAc y tapados por  $\alpha$ 2,3NeuAc.

Figura 26 Ion negativo ESI-MS/MS de los iones precursores a m/z 1183.92 (N-glicano biantenario sialilado con núcleo de fucosa) de EPO producido en CHO con (A) Mgat4a/4b/5 KO y con (B) ambos Mgat4a/4b/5 y st3gal4/6 KO y KI de ST6GAL1. La presencia de los iones del fragmento de diagnóstico en m/z 306,12 demuestra la presencia de sialilación terminal  $\alpha$ 2,6.

Figura 27 Análisis de KI dirigido de ST6GAL1 mediante PCR de unión. El panel A muestra que se utilizó una estrategia de KI dirigida a ObLiGaRe modificada que usa dos sitios de unión ZFN invertidos que flanquean el marco de lectura abierto completo ST6GAL1 en el plásmido donante. La PCR de unión 5' y 3' confirmó la integración dirigida en el sitio Safeharbor#1 en el clon CHO con st3gal4/6 KO. (b) La PCR de unión 5' y 3' confirmó la integración dirigida en el clon CHO con st3gal4/6/mgat4a/4b/5 KO. El estado del número de copias alélicas de integración se determinó mediante PCR alélica WT. La presencia de la banda deseada en la PCR

alélica WT indica la presencia del sitio Safeharbor#1 sin la integración del KI dirigido del gen de interés en al menos uno de los alelos. Los resultados mostraron la integración bialélica de ST6Gal-I en el sitio Safeharbor#1 en el clon CHO st3gal4/6 KO y la integración monoalélica en el clon mgat4a/4b/5/St3gal4/6 KO. La estrategia ObLiGaRe KI fue muy eficaz, con aproximadamente un 30-50% de células clonadas individuales que expresaban ST6GAL1, según se evaluó mediante inmunocitología con anticuerpos y lectinas.

Figura 28 Representación gráfica de la deconstrucción genética de la N-glicosilación en CHO establecida usando EPO como reportero de N-glicoproteína. Varias líneas KO CHO tienen potencial para la producción de glicoproteínas terapéuticas con una glicosilación más homogénea (por ejemplo, N-glicanos biantenarios  $\alpha$ 2,3-sialilados para la EPO, y biantenarios no galactosilados/sialilados con y sin núcleo Fuc para la IgG. Se muestran ejemplos de reconstrucción basados en KI que producen  $\alpha$ 2,6-sialilación homogénea tras KO de capacidades de  $\alpha$ 2,3-sialilación con heterogeneidad de ramificación de N-glicanos WT o estructura biantenaria homogénea,  $\alpha$ 2,3-sialilación homogénea tras KO de sialilación endógena, y estructuras tetraantenarias homogéneas con  $\alpha$ 2,6-sialilación tras KO de capacidades de ramificación y sialilación endógenas. La reconstrucción puede abordar la homogeneidad como se muestra, pero también la ramificación, la elongación de poli-LacNAc y cualquier tipo de protección como se indica.

Figura 29 Glicoperfilado de EPO expresado en CHO con KO de Cosmc para eliminar la elongación de O-GalNAc con Gal y NeuAc mostrando una capacidad mejorada para el recubrimiento de ácido siálico de N-glicanos en los tres N-glicositos presentes en EPO, en comparación con WT como se muestra en la Figura 15.

Figura 30 Glicoperfilado de EPO expresado en CHO con KO de B4galT7 para eliminar la elongación de O-Xyl con Gal (PANEL A) y Pomgnt1 para eliminar la elongación de O-Man con Gal y NeuAc (PANEL B) mostrando una capacidad mejorada para el recubrimiento de ácido siálico de N-glicanos en los tres N-glicositos presentes en EPO. El PANEL C ilustra el glicoprofilado de EPO expresado en doble KO de B4galT7 y pomgnt1, y el Panel D ilustra el glicoprofilado de EPO expresado en triple KO de B4galT7, pomgnt1, y COSMC mostrando la misma capacidad mejorada para el tapado de ácido siálico de N-glicanos en EPO.

Figura 31 Glicoperfilado de EPO expresado en CHO con KO de B4galT4 para eliminar un parólogo de galactosiltransferasa no usado para N-glicanos y esta vía de glicosilación muestra una capacidad mejorada para el recubrimiento de N-glicanos con ácido siálico en los tres N-glicositos presentes en EPO.

Figura 32 Glicoperfilado de EPO expresado en CHO con KO de mgat2 que muestra pérdida de estructuras de N-glicano tetra y triantenarios y aparición de N-glicanos mono y biantenarios con protección de galactosa y NeuAc (PANEL B), y sin protección de ácido siálico con KO de st3gal4/6 (PANEL C). La eliminación adicional de mgat4A/4B/5 dio como resultado estructuras monoantenarias con cantidades menores de poli-LacNAc (PANELES D y E). La reintroducción de mgat4A y mgat5 humanos en combinación con ST6Gal-I resultó en la pérdida de N-glicanos monoantenarios y la restauración de estructuras tri y tetraantenarias (PANELES H e I).

Figura 33 Análisis en gel SDS-PAGE de las reacciones de glicoPEGilación de EPO-Myc-His6 de tipo salvaje recombinante y EPO-Myc-His6 recombinante con N-glicanos monoantenarios. Los carriles contenían: (1) HiMark Standard, (2) ST3Gal3, (3) Sialidasa, (4) wt hEPO, wt EPO-Myc-His6 glicoPEGilado con ST3Gal-III y (5) 1x, (6) 5x, (7) 10x, (8) 25x, (9) 50x reactivo 10kDa-PSC respectivamente, (10) variante de glicosilación EPO-Myc-His6 y variante de glicosilación EPO-Myc-His6 glicoPEGilada con ST3Gal-III y (11) 1x, (12) 5x, (13) 10x, (14) 25x, (15) 50x reactivo 10kDa-PSC. Las bandas superiores a aproximadamente 160 kD indican glicoPEGilación con más de dos cadenas de PEG en uno o más N-glicanos.

Figura 34 Cuantificación del número de cadenas de PEG unidas mediante densitometría de geles SDS-PAGE. (PANEL A) Perfil del producto de la reacción enzimática de glicoPEGilación de 10 kDa con EPO-Myc-His6 producida en líneas celulares CHO ordinarias. (PANEL B) Perfil de producto de la reacción enzimática de 10kDa-GlycoPEGylation con EPO-Myc-His6 producida en líneas celulares Mgat2/4a/4b/5 KO CHO. Los perfiles A y B se obtuvieron mediante mediciones densitométricas en el carril 9 y el carril 15 de la Figura 33, respectivamente.

Figura 35 Glicoperfilado de EPO expresado en células CHO con KO de B3gnt2 apilado con KO de Cosmc/Pomgnt1/B4galT7 (este último triple KO se ve en la Fig30D) (panel superior), doble KO de B4galT5 y B4galT6 (segundo panel), doble KO de Mgat2 y Mgat4B (tercer panel) y doble KO de Mgat2 y Mgat5 (panel inferior). En comparación con el fondo (ver Fig. 30D), la KO de B3gnt2 elimina poliLacNAc. El KO de B4galT5 y B4galT6 da más LacNAc en comparación con la EPO wt (Fig 31). Los dos últimos paneles con doble KO de Mgat2/4b o Mgat2/5 muestran que se puede obtener monoantenario con doble KO, por lo que no es necesario el KO de cuatro genes Mgat2/4A/4B/5 (comparar con la Fig. 32D).

Figura 36 Glicoperfilado de IgG expresado en células CHO con KO de Mgat2 (arriba), doble KO de St3Gal4/6 (medio) o KO de B4galT1/3 apilado sobre Cosmc/Fut8 (panel inferior). El KO de Mgat2 da como resultado una estructura mono antenal homogénea y un mayor grado de sialilación que la IgG wt (comparar con Fig 22). St3gal4/6 KO da una mayor galactosilación (comparar con Fig22). La doble KO de B4galT1 y B4galT3 muestra

N-glicanos más homogéneos sin galactosa que los obtenidos con una sola *ko* de B4galt1 (comparar con la Fig. 22, panel inferior).

La presente invención se describirá con más detalle a continuación.

5

#### Descripción detallada de la invención

Las cadenas de azúcar de las glicoproteínas se dividen a grandes rasgos en dos tipos, a saber, una cadena de azúcar que se une a la asparagina (cadena de azúcar ligada a un N-glicosido) y una cadena de azúcar que se une a otro aminoácido como la serina, la treonina (cadena de azúcar ligada a un O-glicosido), en función de la forma de unión a la fracción proteica.

10

El extremo de la cadena de azúcar que se une a la asparagina se denomina extremo reductor, y el lado opuesto, extremo no reductor. Se sabe que la cadena de azúcares ligada a N-glucósidos incluye un tipo con alto contenido en manosa en el que la manosa sola se une al extremo no reductor de la estructura central; un tipo complejo en el que el extremo no reductor de la estructura central tiene al menos una rama paralela de galactosa-N-acetilglucosamina (en lo sucesivo denominada "Gal-GlcNAc") y el extremo no reductor de Gal-GlcNAc tiene una estructura de ácido siálico, bisección de N-acetilglucosamina o similar; un tipo híbrido en el que el extremo no reductor de la estructura central tiene ramas tanto del tipo de alta manosa como del tipo complejo.

15

20

En general, la mayoría de los anticuerpos humanizados cuya aplicación en medicamentos se está considerando se preparan usando técnicas de recombinación genética y se producen usando células CHO derivadas de tejido de ovario de hámster chino como célula huésped. Como se describió anteriormente, la estructura de la cadena de azúcar juega un papel importante para la estructura, función, tamaño, vida media circulatoria y comportamiento farmacocinético de los fármacos glicoproteicos. Además, la estructura del azúcar desempeña un papel notablemente importante en la función efectora de los anticuerpos y se observan diferencias en la estructura de la cadena de azúcar de las glicoproteínas expresadas por las células huésped, por lo que se desea el desarrollo de una célula huésped que pueda usarse para la producción de un anticuerpo que tenga una mayor función efectora.

25

30

Para apoyar la producción de un número creciente de glicoproteínas biofarmacéuticas es necesario desarrollar nuevas líneas celulares de mamíferos y, preferiblemente, líneas celulares derivadas de CHO con diferentes capacidades y características de glicosilación.

35

Además, es necesario desarrollar matrices de diseño para vías de glicosilación individuales que impliquen la inactivación de genes y/o la introducción de genes estables que permitan generar líneas celulares de mamíferos a medida con capacidades y características de glicosilación diferentes y más homogéneas.

40

45

Además, existe la necesidad de desarrollar líneas celulares de mamíferos y, preferentemente, líneas celulares derivadas de CHO con inactivación de dos o más genes de glicosiltransferasa que puedan producir glicoproteínas recombinantes con una glicosilación diferente a la de su homólogo natural. Más particularmente, existe una necesidad de desarrollar dichas líneas celulares con inactivación de dos o más genes de glicosiltransferasa que codifican isoenzimas con funciones relacionadas en la misma vía de glicosilación, siendo por ejemplo la vía de N-glicosilación. Además, existe la necesidad de desarrollar tales líneas celulares con inactivación de dos o más genes de glicosiltransferasas que codifican enzimas con funciones no relacionadas en la misma vía de glicosilación.

50

Además, existe la necesidad de introducir una o más glicosiltransferasas en líneas celulares y, preferentemente, en líneas celulares derivadas de CHO para obtener capacidades de glicosilación deseables, incluidas capacidades homogéneas y novedosas. Esta introducción de glicosiltransferasas puede combinarse con la inactivación de uno o más de los genes de glicosiltransferasas endógenas.

55

Un objeto de la presente invención es proporcionar una célula modificada con capacidad estable diseñada genéticamente para la producción de una proteína terapéutica, y que produce dicha proteína con cadenas de glicano con estructuras definidas, y/o estructuras de glicano más homogéneas, y/o con bioactividad mejorada.

60

Esta invención divulga un sistema de inactivación y deconstrucción de genes de glicosiltransferasa en CHO que proporciona una matriz de diseño para diseñar una célula con una multitud de capacidades de glicosilación bien definidas. Además, la invención proporciona diseños de reconstrucción para la ingeniería de novo de una multitud de capacidades de glicosilación deseadas y/o novedosas bien definidas mediante la combinación de uno o más eventos de inactivación e introducción de genes de glicosiltransferasas en una célula para mejorar la producción de terapéuticas de glicoproteínas recombinantes.

Esta invención divulga un mamífero en el que mgat4A, mgat4B y mgat5 están eliminados en la célula, y con nuevas capacidades de glicosilación que permiten mejoras para la producción recombinante de proteínas terapéuticas.

- 5 En un aspecto, se proporcionan diseños de orientación de ZFN para la inactivación de genes de glicosiltransferasa.

En otro aspecto, se proporcionan diseños de TALEN para la inactivación de genes de glicosiltransferasas.

- 10 En otro aspecto más, se proporciona orientación basada en CRISPR/Cas9 para la inactivación de genes de glicosiltransferasa.

- 15 Los presentes inventores han empleado por primera vez una pantalla de knockout mediada por ZFN en CHO para explorar el potencial de la ingeniería de N-glicosilación de glicoproteínas recombinantes. Muchos pasos en la vía biosintética del N-glicano son potencialmente catalizados por múltiples isoenzimas, lo que hace que la ingeniería genética sea impredecible (Fig. 1a). Es necesario analizar las funciones *in vivo* de cada una de estas isoenzimas para construir una matriz de opciones de diseño.

- 20 Los presentes inventores han examinado los genes implicados en esta maquinaria altamente compleja (véanse los ejemplos) y han identificado efectos.

Los genes pueden ser eliminados (KO) o incluidos (KI).

- 25 En un aspecto de la presente invención, se elimina mgat4A/4B en la célula para eliminar N-glicanos tetraantennarios ramificados en  $\beta$ 4 y se elimina mgat5 en la célula para eliminar estructuras tetraantennarias ramificadas en  $\beta$ 6.

- 30 En otro aspecto más de la presente invención, mgat5 también se elimina en la célula, lo que conduce a la pérdida del marcaje de lectina L-PHA.

En un aspecto adicional de la presente invención, mgat4A, mgat4B y mgat5 se eliminan en la célula, lo que conduce a N-glicanos biantennarios homogéneos.

- 35 En un aspecto adicional de la presente invención, B3gnt2 también se elimina en la célula, lo que conduce a la eliminación de poli-LacNAc.

En otro aspecto de la presente invención, st3gal4, st3gal6, mgat4A, mgat4B y mgat5 se eliminan en la célula, lo que produce N-glicanos biantennarios sin sialilación, pero aumento de poli-LacNAc.

- 40 En otro aspecto de la presente invención, st3gal3, st3gal4 y st3gal6 también se eliminan en la célula, lo que conduce a una falta total de ácido siálico.

- 45 En un aspecto adicional de la presente invención, st3gal4, st3gal6, mgat4A, mgat4B, mgat5 se eliminan en la célula, lo que conduce a N-glicanos biantennarios homogéneos cubiertos por  $\alpha$ 2,6NeuAc.

La introducción de ST6GAL1 (knock in, KI) con KO de st3gal4/6 produce la gama completa de estructuras antenales de N-glicano con un grado normal de recubrimiento de NeuAc (Fig. 2a).

- 50 ST6GAL1 introducido con KO adicional de mgat4A/4B/5 produce glicoproteínas con N-glicanos biantennarios homogéneos cubiertos por  $\alpha$ 2,6NeuAc (Fig. 2b).

La introducción de nuevo de ST6GAL1 anula las pequeñas cantidades de poli-LacNAc formadas en las estructuras biantennarias cuando se elimina la sialilación (Fig. 2b).

- 55 Por lo tanto, el KI de ST6GAL1 se puede usar en combinación con uno o más de los knockouts descritos en este documento.

En el presente contexto, el NeuAc también se conoce como ácido neuramínico.

- 60 Por lo tanto, un aspecto de la presente invención se refiere a una proteína expresada en una célula, en donde el patrón de modificación postraducciona comprende N-glicanos biantennarios homogéneos con o sin protección de  $\alpha$ 2,3NeuAc o con y sin protección de  $\alpha$ 2,6NeuAc.

Por eliminación se entiende la alteración total o parcial de la función del gen de interés.

65

En un aspecto de la presente invención, uno o más de los genes mencionados anteriormente se eliminan usando nucleasas de dedo de zinc ZFN. Los ZFN pueden usarse para inactivar un gen FUT8 o cualquiera de los otros genes aquí descritos. Las ZFN comprenden una proteína de dedo de zinc (ZFP) y un dominio de nucleasa (escisión).

5

En otra realización, la presente invención proporciona líneas celulares CHO con diferentes capacidades de N-glicosilación bien definidas que permiten la producción recombinante de glicoproteínas terapéuticas con N-glicanos compuestos de N-glicanos biantenarios con o sin poli-LacNAc, y con o sin recubrimiento  $\alpha 2,3\text{NeuAc}$ .

10 En otra realización, esta invención proporciona líneas celulares con diferentes capacidades de N-glicosilación bien definidas que permiten la producción recombinante de glicoproteínas terapéuticas con N-glicanos compuestos de N-glicanos biantenarios, triantenarios o tetraantenarios con o sin poli-LacNAc, y con o sin recubrimiento  $\alpha 2,6\text{NeuAc}$ .

15 En otra realización más, esta invención proporciona líneas celulares que permiten la producción de glicoproteínas con N-glicanos homogéneos sin protección con ácido siálico, lo que permite la modificación enzimática directa de N-glicanos por sialiltransferasas in vitro después de la producción.

20 En otra realización más, la presente invención proporciona líneas celulares que permiten la producción de glicoproteínas con N-glicanos homogéneos de estado biantenario sin recubrimiento de ácido siálico, lo que permite la modificación enzimática directa de los N-glicanos por sialiltransferasas en postproducción in vitro.

25 En otra realización, la presente invención proporciona líneas celulares que permiten la producción de glicoproteínas con N-glicanos homogéneos de estado monoantenario con o sin recubrimiento de ácido siálico, lo que permite la modificación enzimática directa de un N-glicano por sialiltransferasas en postproducción in vitro.

30 En otro aspecto más, también se proporciona una célula aislada que comprende cualquiera de las proteínas y/o polinucleótidos como se describe en este documento. En ciertas realizaciones, uno o más genes de glicosiltransferasa están inactivados (parcial o totalmente) en la célula. Cualquiera de las células descritas en el presente documento puede incluir genes adicionales que han sido inactivados, por ejemplo, usando nucleasas de dedos de zinc, TALEN y/o CRISPR/Cas9 diseñadas para unirse a un sitio diana en el gen seleccionado. En ciertas realizaciones, se proporcionan en el presente documento células o líneas celulares en las que se han inactivado dos o más genes de glucosiltransferasas, y células o líneas celulares en las que se han inactivado uno o más genes de glucosiltransferasas y se han introducido una o más glucosiltransferasas.

35 En otra realización más, esta invención proporciona líneas celulares con inactivación de vías de glicosilación indeseables para la producción optimizada de glicoproteínas con características de glicosilación deseables. En ciertas realizaciones, la inactivación de uno o más pasos biosintéticos en las vías de O-glicosilación de proteínas (O-GalNAc, O-Xyl, O-Man y glicolípido) se obtiene individualmente o en combinación para mejorar la reserva de nucleótidos de azúcar disponible para las características de glicosilación deseables.

40 En una realización, esta invención proporciona una célula con inactivación mgat2, mgat4A, mgat4B, y mgat5, producen eritropoyetina con N-glicanos que tienen estructuras monoantenarias homogéneas, y son adecuadas para una glicomodificación más homogénea.

45 Por ejemplo, las líneas celulares descritas en el presente documento que tienen inactivados los genes mgat2, mgat4A, mgat4B, mgat5, st3gal4 y 6, producen eritropoyetina con N-glicanos que tienen estructuras monoantenarias homogéneas sin recubrimiento de ácido siálico, y son adecuadas para una glicomodificación más homogénea.

50 Inactivación de glicosiltransferasas innecesarias para mejorar las vías de glicosilación deseables. En otra realización más, esta invención proporciona líneas celulares con inactivación de enzimas de glicosilación indeseables para mejorar la capacidad y/o fidelidad de las características de glicosilación deseables. En ciertas realizaciones, se proporcionan líneas celulares con activación de una o más sialiltransferasas, y/o galactosiltransferasas, y/o GlcNAc-transferasas, que no funcionan en una vía de glicosilación deseable, tal como por ejemplo la N-glicosilación.

55 En una realización, esta invención proporciona una célula con inactivación de una  $\beta 4$ galactosiltransferasa ( $\beta 4\text{Gal-T7}$ ) con poca o ninguna función en la vía de N-glicosilación, y capacidad mejorada para galactosilación, poliLacNAc y tapado con ácido siálico de N-glicanos.

60 En una realización, esta invención proporciona una célula con inactivación de una sialiltransferasa con poca o ninguna función en la vía de la N-glicosilación, y capacidad mejorada para la galactosilación y el recubrimiento con ácido siálico de N-glicanos.

65

En una realización, esta invención proporciona una célula con inactivación de una  $\beta$ 3galactosiltransferasa (C1GalT1) con poca o ninguna función en la vía de N-glicosilación, y capacidad mejorada para la galactosilación y el tapado con ácido siálico de N-glicanos.

- 5 En una realización, esta invención proporciona una célula con inactivación de una  $\beta$ 4galactosiltransferasa con poca o ninguna función en la vía de N-glicosilación, y capacidad mejorada para la galactosilación y el recubrimiento con ácido siálico de N-glicanos.

- 10 En una realización, esta invención proporciona una célula con inactivación de una  $\beta$ 2GlcNAc-transferasa (POMGnT1) con poca o ninguna función en la vía de N-glicosilación, y capacidad mejorada para la galactosilación y el tapado con ácido siálico de N-glicanos.

- 15 En una realización, esta invención proporciona una célula con inactivación de una  $\beta$ 3galactosiltransferasa (C1GalT1), y/o una  $\beta$ 2GlcNAc-transferasa (POMGnT1), y/o una  $\beta$ 4galactosiltransferasa ( $\beta$ 4Gal-T7) con poca o ninguna función en la vía de N-glicosilación, y capacidad mejorada para la galactosilación y el recubrimiento con ácido siálico de N-glicanos.

- 20 En una realización, esta invención proporciona una célula con inactivación de una  $\beta$ 4galactosiltransferasa ( $\beta$ 4Gal-T4) con poca o ninguna función en la vía de N-glicosilación, y capacidad mejorada para la galactosilación y el recubrimiento con ácido siálico de N-glicanos.

- 25 Además, se proporcionan métodos de uso de las proteínas zinc finger y sus fusiones en métodos de inactivación de genes glicosiltransferasa en una célula o línea celular. En ciertas realizaciones, la inactivación de uno o más genes de glicosiltransferasa da lugar a una línea celular que puede producir proteínas recombinantes a niveles más altos o en la que una o más actividades (funciones) de las proteínas se incrementan en comparación con las proteínas producidas en células en las que el gen o genes no están inactivados.

- 30 Por lo tanto, en otro aspecto, aquí se proporcionan métodos para inactivar los genes de glicosiltransferasa celular al menos mgat4A, mgat4B y mgat5 (p. ej. también, los genes endógenos mgat1, mgat2, mgat3, mgat4A, mgat4B, mgat4C, mgat5, mgat5B, B4galt1, B4galt2, B4galt3, B4galt4, B3gnt1, B3gnt2, B3gnt8, st3gal3, st3gal4, st3gal6, fut8) en una célula, mediante el uso de métodos que comprenden la perturbación del genoma, la edición de genes y/o la capacidad de disrupción génica, tales como los sistemas de vectores de ácidos nucleicos relacionados con las repeticiones palindrómicas cortas agrupadas y regularmente espaciadas (CRISPR) y sus componentes, sistemas de vectores de ácidos nucleicos que codifican proteínas de fusión que comprenden dominios de unión al ADN de dedos de cinc (ZF) y al menos un dominio de escisión o al menos un semidominio de escisión (ZFN) y/o sistemas de vectores de ácidos nucleicos que codifican un primer monómero de endonucleasa efectora de tipo activador de la transcripción (TAL) y un ácido nucleico que codifica un segundo dominio de escisión o al menos un semidominio de escisión (TALEN). La introducción en una  
40 célula de cualquiera de los agentes de escisión de ácido nucleico mencionados anteriormente (CRISPR, TALEN, ZFN) son capaces de escindir específicamente un sitio diana del gen de la glucosiltransferasa como resultado de la introducción celular de: (1) un ácido nucleico que codifica un par de proteínas de unión a la diana del gen de la glucosiltransferasa ZF o TAL, cada una fusionada a dicha fracción de corte de ácido nucleico, en la que al menos uno de dichos polipéptidos ZF o TAL es capaz de unirse específicamente a una  
45 secuencia de nucleótidos situada aguas arriba de dicho sitio de corte diana, y la otra proteína ZF o TAL es capaz de unirse específicamente a una secuencia nucleotídica situada aguas abajo del sitio de corte de la diana, por lo que cada una de las proteínas zinc finger se une independientemente a la diana de ácido nucleico y la rodea, seguido de la disrupción del ácido nucleico diana por rotura de doble cadena mediada por las moléculas de corte de la endonucleasa fusionada, (2) una secuencia nucleotídica que codifica un ARN guía del sistema CRISPR-Cas que hibrida la secuencia diana del gen de la glucosiltransferasa, y b) una segunda  
50 secuencia de nucleótidos que codifica una proteína Cas9 de tipo II, en la que los componentes (a) y (b) están situados en el mismo vector o en vectores diferentes del sistema, en donde el ARN guía está compuesto por un ARN quimérico e incluye una secuencia guía y una secuencia trans-activadora cr (tracr), por lo que el ARN guía se dirige a la secuencia diana del gen de la glucosiltransferasa y la proteína Cas9 escinde el sitio diana  
55 del gen de la glucosiltransferasa.

- 60 En otra realización más, esta invención proporciona líneas celulares con inactivación de uno o más genes de glicosiltransferasas y con introducción estable de una o más glicosiltransferasas para mejorar la fidelidad de características de glicosilación deseables y/o introducir características de glicosilación mejoradas y/o características de glicosilación novedosas.

- 65 En ciertas realizaciones se proporcionan líneas celulares con inactivación de una o más sialiltransferasas, y/o galactosiltransferasas, y/o glucosiltransferasas, y/o GlcNAc-transferasas, y/o GalNAc-transferasas, y/o xilotransferasas, y/o glucuronosiltransferasas, mannosiltransferasas, y/o fucosiltransferasas, y en las que se han introducido de forma estable una o más glucosiltransferasas.

Por ejemplo, las líneas celulares como las descritas en este documento que tienen genes mgat4A, mgat4B, mgat5, st3gal4 y 6 inactivados y en las que se ha introducido la sialiltransferasa ST6Gal-I humana, producen eritropoyetina con N-glicanos biantenarios cubiertos de manera homogénea por  $\alpha$ 2,6NeuAc y son más similares a las humanas.

5

Por ejemplo, las líneas celulares descritas en el presente documento que tienen inactivados los genes mgat4A, mgat4B, mgat5, st3gal4 y 6 y en las que se han introducido los genes MGAT4A y ST6Gal-I humanos, producen eritropoyetina con N-glicanos triantenarios tapados homogéneamente por  $\alpha$ 2,6NeuAc y son más parecidas a las humanas.

10

Por ejemplo, las líneas celulares descritas en el presente documento que tienen inactivados los genes mgat4A, mgat4B, mgat5, st3gal4 y 6 y en las que se han introducido MGAT4A, MGAT5 y ST6Gal-I, producen eritropoyetina con N-glicanos tetraantenarios tapados homogéneamente por  $\alpha$ 2,6NeuAc y son más parecidas a las humanas.

15

Hay varios métodos disponibles para introducir genes exógenos, como los genes de la glucosiltransferasa, en células de mamíferos y seleccionar líneas celulares clonales estables que alberguen y expresen el gen de interés. Típicamente, el gen de interés se co-transfecta con un gen marcador de selección que favorece las líneas celulares que expresan el marcador de selección bajo ciertas condiciones definidas de cultivo de medios.

20

El medio podría contener un inhibidor de la proteína marcadora de selección o la composición del medio podría estresar el metabolismo celular y requerir así una mayor expresión del marcador de selección. El gen marcador de selección puede estar presente en el mismo plásmido que el gen de interés o en otro plásmido.

25

En una realización, la introducción de una o más glicosiltransferasas exógenas se lleva a cabo mediante transfección plasmídica con un plásmido que codifica la expresión impulsada por un promotor constitutivo tanto del gen de la glicosiltransferasa como de un marcador antibiótico seleccionable, donde el marcador seleccionable también podría representar un gen esencial no presente en la célula huésped, como el sistema GS (Sigma/Lonza), y/o plásmidos separados que codifican el gen de la glicosiltransferasa impulsado por un promotor constitutivo o el marcador seleccionable. Por ejemplo, se han transfectado plásmidos que codifican ST6GalNAc-I y Zeocin en células y se han seleccionado líneas estables que expresan ST6Gal-I en función de la resistencia a la zeocina.

30

En otra realización, la introducción de una o más glicosiltransferasas exógenas se realiza mediante inserción mediada por nucleasa dirigida al sitio.

35

En una realización, un método para expresar de manera estable al menos un producto de una secuencia de ácido nucleico exógeno en una célula mediante la introducción de roturas de doble cadena en el locus genómico PPP1R12C o Safe Harbor#1 usando agentes de escisión de ácido nucleico ZFN y una secuencia de ácido nucleico exógena que, mediante una manera dependiente de la homología o a través de salientes de corte de ZFN flanqueantes compatibles, se inserta en el sitio de escisión y se expresa. Los sitios Safe Harbor son sitios del genoma que al ser manipulados no tienen consecuencias celulares o fenotípicas evidentes. Además de los sitios mencionados anteriormente, se han identificado varios otros sitios como Safe Harbor#2, CCR5 y Rosa26. Además de la tecnología ZFN, las herramientas TALEN y CRISPR también pueden permitir la integración de secuencias exógenas en líneas celulares o genomas de manera precisa. Para ello se debe evaluar: i) hasta qué punto se produce el silenciamiento epigenético y ii) cuál debe ser el nivel de expresión deseado del gen de interés. En los ejemplos incluidos en este documento, la integración específica del sitio de glicosiltransferasas humanas, por ejemplo, ST6Gal-I, MGAT4A o MGAT5, usando la estrategia de inserción ObLiGaRe, se basa en un diseño de vector flanqueado por aislador impulsado por la expresión de CMV.

40

45

50

En otro aspecto más, la divulgación proporciona un método para producir una proteína recombinante de interés en una célula huésped, comprendiendo el método los pasos de: (a) proporcionar una célula hospedadora que comprende dos o más genes de glucosiltransferasa endógenos; (b) inactivar los genes de glucosiltransferasa endógenos de la célula hospedadora mediante cualquiera de los métodos descritos en el presente documento; e (c) introducir un vector de expresión que comprende un transgén, comprendiendo el transgén una secuencia que codifica una proteína de interés, en la célula hospedadora, produciendo así la proteína recombinante. En ciertas realizaciones, la proteína de interés comprende, por ejemplo, eritropoyetina o un anticuerpo, por ejemplo, un anticuerpo monoclonal.

55

60

En otro aspecto más, la divulgación proporciona un método para producir una proteína recombinante de interés en una célula, comprendiendo el método los pasos de: (a) proporcionar una célula según la presente invención; (b) inactivar el gen o genes de glucosiltransferasa endógenos de la célula huésped; (c) introducir uno o más genes de glucosiltransferasa en la célula mediante cualquiera de los métodos descritos en el presente documento; e (d) introducir un vector de expresión que comprende un transgén, comprendiendo el transgén una secuencia que codifica una proteína de interés, en la célula, produciendo así la proteína recombinante. En ciertas realizaciones, la proteína de interés comprende, por ejemplo, eritropoyetina o un anticuerpo, por ejemplo, un anticuerpo monoclonal.

65

Otro aspecto de la divulgación abarca un método para producir una proteína recombinante con un patrón de glicosilación más homogéneo y/o similar al humano y/o novedoso y/o funcionalmente beneficioso. El método comprende expresar la proteína en una línea celular de mamífero según la presente invención. En una realización específica, la línea celular es una línea celular de ovario de hámster chino (CHO). En una realización, la línea celular comprende secuencias cromosómicas inactivadas que codifican cualquier glicosiltransferasa endógena. En una realización, las secuencias cromosómicas inactivadas que codifican cualquier glicosiltransferasa endógena son monoalélicas y la línea celular produce una cantidad reducida de dichas glicosiltransferasas. En otra realización, las secuencias cromosómicas inactivadas que codifican cualquier glucosiltransferasa endógena son bialélicas, y la línea celular no produce dichas glucosiltransferasas mensurables. En otra realización, la proteína recombinante tiene un patrón de glicosilación más homogéneo y/o similar al humano y/o novedoso y/o funcionalmente beneficioso. En una realización, la proteína recombinante tiene al menos una propiedad mejorada en relación con una proteína recombinante similar producida por una línea celular comparable no deficiente en dichas glicosiltransferasas endógenas, por ejemplo, inmunogenicidad reducida, biodisponibilidad aumentada, eficacia aumentada, estabilidad aumentada, solubilidad aumentada, semivida mejorada, aclaramiento mejorado, farmacocinética mejorada y combinaciones de las mismas. La proteína recombinante puede ser cualquier proteína, incluida una proteína terapéutica. Las proteínas ejemplares incluyen aquellas seleccionadas entre, pero no limitadas a, un anticuerpo, un fragmento de anticuerpo, un factor de crecimiento, una citocina, una hormona, una enzima lisosomal, un factor de coagulación, una gonadotropina y un fragmento funcional o variantes de los mismos.

La divulgación también se puede usar para identificar genes objetivo para modificación y usar este conocimiento para glicoingeniería de una línea celular de mamífero existente previamente transfectada con ADN que codifica la proteína de interés.

En cualquiera de las células y métodos descritos en este documento, la célula o línea celular puede ser una COS, CHO (por ejemplo, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (por ejemplo, HEK293-F, HEK293-H, HEK293-T) y PERC6.

Las líneas celulares descritas en el presente documento también pueden usarse para producir otras N-glicoproteínas, incluyendo sin ánimo de limitación, por ejemplo,  $\alpha$ 1-antitripsina, gonadotropinas, proteínas enzimáticas lisosomales (por ejemplo, glucocerebrosidasa, alfa-galactosidasa, alfa-glucosidasa, sulfatasas, glucuronidasa, iduronidasa). Los genes de glicosiltransferasa humana conocidos se ensamblan en familias de genes homólogos en la base de datos CAZy y estas familias se asignan además a diferentes vías de glicosilación en Hansen et al. (Hansen, Lind-Thomsen et al. 2014). La TABLA 1 enumera todos los genes de glicosiltransferasa humana en las familias CAZy GT con identificaciones de genes de NCBI y asignación de funciones confirmadas o putativas en la biosíntesis de diferentes glicoconjugados de mamíferos (N-glicanos, O-GalNAc, O-GlcNAc, O-Glc, O-Gal, O-Fuc, O-Xyl, O-Man, C-Man, glicosfingolípidos, hialuronano y anclajes GPI). La FIGURA 1 representa gráficamente además los roles confirmados y putativos de las familias CAZy GT humanas en la biosíntesis de diferentes glicoconjugados.

Tabla 1. Genes GTf humanos

| Familia CAZy <sup>(1)</sup> | Identificación genética <sup>(2)</sup> | Símbolo <sup>(3)</sup> | Descripción <sup>(4)</sup>                                           | Ubicación en el mapa <sup>(5)</sup> |
|-----------------------------|----------------------------------------|------------------------|----------------------------------------------------------------------|-------------------------------------|
| GT32                        | 53947                                  | <i>A4GALT</i>          | $\alpha$ 1,4-galactosiltransferasa                                   | 22q13.2                             |
| GT32                        | 51146                                  | <i>A4GNT</i>           | $\alpha$ 1,4-N-acetilglucosaminiltransferasa                         | 3p14.3                              |
| GT6                         | 28                                     | <i>ABO</i>             | Grupo sanguíneo ABO                                                  | 9q34.2                              |
| GT33                        | 56052                                  | <i>ALG1</i>            | quitobiosildifosfodolichol manosiltransferasa $\beta$ -              | 16p13.3                             |
| GT59                        | 84920                                  | <i>ALG10</i>           | $\alpha$ 1,2-glucosiltransferasa                                     | 12p11.1                             |
| GT59                        | 144245                                 | <i>ALG10B</i>          | $\alpha$ 1,2-glucosiltransferasa                                     | 12q12                               |
| GT4                         | 440138                                 | <i>ALG11</i>           | $\alpha$ 1,2-manosiltransferasa                                      | 13t 14.2                            |
| GT22                        | 79087                                  | <i>ALG12</i>           | $\alpha$ 1,6-manosiltransferasa                                      | 22t13.33                            |
| GT1                         | 79868                                  | <i>ALG13</i>           | Subunidad de la UDP-N-acetilglucosaminiltransferasa                  | Xq23                                |
| GT1                         | 199857                                 | <i>ALG14</i>           | Subunidad de la UDP-N-acetilglucosaminiltransferasa                  | 1p21.3                              |
| GT33                        | 200810                                 | <i>ALG1L</i>           | similar a quitobiosildifosfodolichol manosiltransferasa $\beta$ -    | 3q21.2                              |
| GT33                        | 644974                                 | <i>ALG1L2</i>          | similar a quitobiosildifosfodolichol manosiltransferasa- 2 $\beta$ - | 3q22.1                              |

ES 3 003 873 T3

| Familia CAZy <sup>(1)</sup> | Identificación genética <sup>(2)</sup> | Símbolo <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                                                           | Ubicación en el mapa <sup>(5)</sup> |
|-----------------------------|----------------------------------------|------------------------|----------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| GT4                         | 85365                                  | <i>ALG2</i>            | $\alpha$ 1,3/1,6-manosiltransferasa                                                                                  | 9q22.33                             |
| GT58                        | 10195                                  | <i>ALG3</i>            | $\alpha$ 1,3-manosiltransferasa                                                                                      | 3q27.1                              |
| GT2                         | 29880                                  | <i>ALG5</i>            | dolicil-fosfato $\beta$ -glucosiltransferasa                                                                         | 13q13.3                             |
| GT57                        | 29929                                  | <i>Alg6</i>            | $\alpha$ 1,3-glucosiltransferasa                                                                                     | 1p31.3                              |
| GT57                        | 79053                                  | <i>ALG8</i>            | $\alpha$ 1,3-glucosiltransferasa                                                                                     | 11q14.1                             |
| GT22                        | 79796                                  | <i>ALG9</i>            | $\alpha$ 1,2-manosiltransferasa                                                                                      | 11q23                               |
| GT31                        | 8706                                   | <i>B3GALNT1</i>        | $\beta$ 1,3-N-acetilgalactosaminiltransferasa 1                                                                      | 3q25                                |
| GT31                        | 148789                                 | <i>B3GALNT2</i>        | $\beta$ 1,3-N-acetilgalactosaminiltransferasa 2                                                                      | 1q42.3                              |
| GT31                        | 8708                                   | <i>B3GALT1</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 1                                             | 2q24.3                              |
| GT31                        | 8707                                   | <i>B3GALT2</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 2                                             | 1q31                                |
| GT31                        | 8705                                   | <i>B3GALT4</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 4                                             | 6p21.3                              |
| GT31                        | 10317                                  | <i>B3GALT5</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 5                                             | 21q22.3                             |
| GT31                        | 126792                                 | <i>B3GALT6</i>         | UDP-Gal: $\beta$ Gal- $\beta$ 1,3-galactosiltransferasa polipéptido 6 similar a la $\beta$ 1,3-galactosiltransferasa | 1p36.33                             |
| GT31                        | 145173                                 | <i>B3GALT</i>          | similar a la $\beta$ 1,3-galactosiltransferasa                                                                       | 13q12.3                             |
| GT43                        | 27087                                  | <i>B3GAT1</i>          | $\beta$ 1,3-glucuroniltransferasa 1                                                                                  | 11q25                               |
| GT43                        | 135152                                 | <i>B3GAT2</i>          | $\beta$ 1,3-glucuroniltransferasa 2                                                                                  | 6q13                                |
| GT43                        | 26229                                  | <i>B3GAT3</i>          | $\beta$ 1,3-glucuroniltransferasa 3                                                                                  | 11q12.3                             |
| GT49                        | 11041                                  | <i>B3GNT1</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 1                                                | 11q13.2                             |
| GT31                        | 10678                                  | <i>B3GNT2</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 2                                                | 2p15                                |
| GT31                        | 10331                                  | <i>B3GNT3</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 3                                                | 19p13.1                             |
| GT31                        | 79369                                  | <i>B3GNT4</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 4                                                | 12q24                               |
| GT31                        | 84002                                  | <i>B3GNT5</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 5                                                | 3q28                                |
| GT31                        | 192134                                 | <i>B3GNT6</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 6                                                | 11q13.4                             |
| GT31                        | 93010                                  | <i>B3GNT7</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 7                                                | 2q37.1                              |
| GT31                        | 374907                                 | <i>B3GNT8</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 8                                                | 19q13.2                             |
| GT31                        | 84752                                  | <i>B3GNT9</i>          | UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 9                                                | 16q22.1                             |
| GT2                         | 146712                                 | <i>B3GNTL1</i>         | similar a UDP-GlcNAc: $\beta$ Gal $\beta$ 1,3-N-acetilglucosaminiltransferasa 1                                      | 17q25.3                             |
| GT12                        | 2583                                   | <i>B4GALNT1</i>        | $\beta$ 1,4-N-acetil-galactosaminiltransferasa 1                                                                     | 12q13.3                             |
| GT12                        | 124872                                 | <i>B4GALNT2</i>        | $\beta$ 1,4-N-acetil-galactosaminiltransferasa 2                                                                     | 17q21.32                            |
| GT7                         | 283358                                 | <i>B4GALNT3</i>        | $\beta$ 1,4-N-acetil-galactosaminiltransferasa 3                                                                     | 12p13.33                            |
| GT7                         | 338707                                 | <i>B4GALNT4</i>        | $\beta$ 1,4-N-acetil-galactosaminiltransferasa 4                                                                     | 11p15.5                             |
| GT7                         | 2683                                   | <i>B4GALT1</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 1                                             | 9p13                                |
| GT7                         | 8704                                   | <i>B4GALT2</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 2                                             | 1p34-p33                            |
| GT7                         | 8703                                   | <i>B4GALT3</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 3                                             | 1q21-q23                            |
| GT7                         | 8702                                   | <i>B4GALT4</i>         | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 4                                             | 3q13.3                              |

| Familia CAZy <sup>(1)</sup> | Identificación genética <sup>(2)</sup> | Símbolo <sup>(3)</sup> | Descripción <sup>(4)</sup>                                            | Ubicación en el mapa <sup>(5)</sup> |
|-----------------------------|----------------------------------------|------------------------|-----------------------------------------------------------------------|-------------------------------------|
| GT7                         | 9334                                   | <i>B4GALT5</i>         | UDP-Gal:βGlcNAc galactosiltransferasa, polipéptido 5                  | 20q13.1-q13.2                       |
| GT7                         | 9331                                   | <i>B4GALT6</i>         | UDP-Gal:βGlcNAc galactosiltransferasa, polipéptido 6                  | 18q11                               |
| GT7                         | 11285                                  | <i>B4GALT7</i>         | xilosilproteína β1,4-galactosiltransferasa, polipéptido 7             | 5q35.2-q35.3                        |
| GT31                        | 56913                                  | <i>C1GALT1</i>         | núcleo 1 sintasa, galactosiltransferasa 1                             | 7p21.3                              |
| GT31                        | 29071                                  | <i>C1GALT1C1</i>       | Chaperona 1 específica de C1GALT1                                     | Xq24                                |
| GT25                        | 51148                                  | <i>CERCAM</i>          | molécula de adhesión de células endoteliales cerebrales               | 9q34.11                             |
| GT7/31                      | 79586                                  | <i>CHPF</i>            | factor polimerizador de condroitina                                   | 2q35                                |
| GT7/31                      | 54480                                  | <i>CHPF2</i>           | factor 2 polimerizador de condroitina                                 | 7q36.1                              |
| GT7/31                      | 22856                                  | <i>CHSY1</i>           | condroitín sulfato sintasa 1                                          | 15q26.3                             |
| GT7/31                      | 337876                                 | <i>CHSY3</i>           | condroitín sulfato sintasa 3                                          | 5q23.3                              |
| GT25                        | 79709                                  | <i>COLGALT1</i>        | colágeno β(1-O)galactosiltransferasa 1                                | 19p13.11                            |
| GT25                        | 23127                                  | <i>COLGALT2</i>        | colágeno β(1-O)galactosiltransferasa 2                                | 1q25                                |
| GT7                         | 55790                                  | <i>CSGALNACT1</i>      | sulfato de condroitina N-acetilgalactosaminiltransferasa 1            | 8p21.3                              |
| GT7                         | 55454                                  | <i>CSGALNACT2</i>      | sulfato de condroitina N-acetilgalactosaminiltransferasa 2            | 10q11.21                            |
| GT2                         | 8813                                   | <i>DPM1</i>            | polipéptido 1 de manosiltransferasa de dolícilfosfato                 | 20q13.13                            |
| GTnc                        | 23333                                  | <i>DPY19L1</i>         | similar a dpy-19 1 (C. elegans)                                       | 7p14.3-p14.2                        |
| GTnc                        | 283417                                 | <i>DPY19L2</i>         | similar a dpy-19 2 (C. elegans)                                       | 12q14.2                             |
| GTnc                        | 147991                                 | <i>DPY19L3</i>         | similar a dpy-19 3 (C. elegans)                                       | 19q13.11                            |
| GTnc                        | 286148                                 | <i>DPY19L4</i>         | similar a dpy-19 4 (C. elegans)                                       | 8q22.1                              |
| GT61                        | 285203                                 | <i>EOGT</i>            | N-acetilglucosamina transferasa ligada a O específica del dominio EGF | 3p14.1                              |
| GT47/64                     | 2131                                   | <i>Ext1</i>            | exostosina glicosiltransferasa 1                                      | 8q24.11                             |
| GT47/64                     | 2132                                   | <i>EXT2</i>            | exostosina glicosiltransferasa 2                                      | 11p12-p11                           |
| GT47/64                     | 2134                                   | <i>EXTL1</i>           | glicosiltransferasa 1 similar a la exostosina                         | 1p36.1                              |
| GT64                        | 2135                                   | <i>Extl2</i>           | glicosiltransferasa 2 similar a la exostosina                         | 1p21                                |
| GT47/64                     | 2137                                   | <i>EXTL3</i>           | glicosiltransferasa 3 similar a la exostosina                         | 8p21                                |
| GTnc                        | 79147                                  | <i>FKRP</i>            | proteína relacionada con la fukutina                                  | 19q13.32                            |
| GTnc                        | 2218                                   | <i>FKTN</i>            | Fukutín                                                               | 9q31-q33                            |
| GT11                        | 2523                                   | <i>FUT1</i>            | fucosiltransferasa 1, grupo sanguíneo H                               | 19q13.3                             |
| GT10                        | 84750                                  | <i>FUT10</i>           | fucosiltransferasa 10, α1,3                                           | 8p12                                |
| GT10                        | 170384                                 | <i>FUT11</i>           | fucosiltransferasa 11, α1,3                                           | 10q22.2                             |
| GT11                        | 2524                                   | <i>FUT2</i>            | estado secretor de fucosiltransferasa 2 incluye                       | 19q13.3                             |
| GT10                        | 2525                                   | <i>FUT3</i>            | fucosiltransferasa 3, grupo sanguíneo Lewis                           | 19p13.3                             |
| GT10                        | 2526                                   | <i>FUT4</i>            | fucosiltransferasa 4, α1,3                                            | 11q21                               |
| GT10                        | 2527                                   | <i>FUT5</i>            | fucosiltransferasa, mielóide específica                               | 19p13.3                             |
| GT10                        | 2528                                   | <i>FUT6</i>            | fucosiltransferasa 5, α1,3                                            | 19p13.3                             |
| GT10                        | 2529                                   | <i>FUT7</i>            | fucosiltransferasa 6, α1,3                                            | 19p13.3                             |
| GT10                        |                                        |                        | fucosiltransferasa 7, α1,3                                            | 9q34.3                              |

| Familia<br>CAZy <sup>(1)</sup> | Identificación<br>genética <sup>(2)</sup> | Símbolo <sup>(3)</sup> | Descripción <sup>(4)</sup>                                        | Ubicación<br>en el mapa <sup>(5)</sup> |
|--------------------------------|-------------------------------------------|------------------------|-------------------------------------------------------------------|----------------------------------------|
| GT23                           | 2530                                      | <i>FUT8</i>            | fucosiltransferasa 8, α1,6                                        | 14q24.3                                |
| GT10                           | 10690                                     | <i>FUT9</i>            | fucosiltransferasa 9, α1,3                                        | 6q16                                   |
| GT27                           | 2589                                      | <i>GALNT1</i>          | polipéptido acetilgalactosaminiltransferasa 1 N-                  | 18q 12.1                               |
| GT27                           | 55568                                     | <i>GALNT10</i>         | polipéptido acetilgalactosaminiltransferasa 10 N-                 | 5q33.2                                 |
| GT27                           | 63917                                     | <i>GALNT11</i>         | polipéptido acetilgalactosaminiltransferasa 11 N-                 | 7q36.1                                 |
| GT27                           | 79695                                     | <i>GALNT12</i>         | polipéptido acetilgalactosaminiltransferasa 12 N-                 | 9q22.33                                |
| GT27                           | 114805                                    | <i>GALNT13</i>         | polipéptido acetilgalactosaminiltransferasa 13 N-                 | 2q24.1                                 |
| GT27                           | 79623                                     | <i>GALNT14</i>         | polipéptido acetilgalactosaminiltransferasa 14 N-                 | 2p23.1                                 |
| GT27                           | 117248                                    | <i>GALNT15</i>         | polipéptido acetilgalactosaminiltransferasa 15 N-                 | 3p25.1                                 |
| GT27                           | 57452                                     | <i>GALNT16</i>         | polipéptido acetilgalactosaminiltransferasa 16 N-                 | 14q24.1                                |
| GT27                           | 374378                                    | <i>Galnt18</i>         | polipéptido acetilgalactosaminiltransferasa 18 N-                 | 11p15.3                                |
| GT27                           | 2590                                      | <i>GALNT2</i>          | polipéptido acetilgalactosaminiltransferasa 2 N-                  | 1q41-q42                               |
| GT27                           | 2591                                      | <i>GALNT3</i>          | polipéptido acetilgalactosaminiltransferasa 3 N-                  | 2q24-q31                               |
| GT27                           | 8693                                      | <i>GALNT4</i>          | polipéptido acetilgalactosaminiltransferasa 4 N-                  | 12q21.33                               |
| GT27                           | 11227                                     | <i>GALNT5</i>          | polipéptido acetilgalactosaminiltransferasa 5 N-                  | 2q24.1                                 |
| GT27                           | 11226                                     | <i>GALNT6</i>          | polipéptido acetilgalactosaminiltransferasa 6 N-                  | 12q13                                  |
| GT27                           | 51809                                     | <i>GALNT7</i>          | polipéptido acetilgalactosaminiltransferasa 7 N-                  | 4q31.1                                 |
| GT27                           | 26290                                     | <i>GALNT8</i>          | polipéptido acetilgalactosaminiltransferasa 8 N-                  | 12p13.3                                |
| GT27                           | 50614                                     | <i>GALNT9</i>          | polipéptido acetilgalactosaminiltransferasa 9 N-                  | 12q24.33                               |
| GT27                           | 168391                                    | <i>GALNTL5</i>         | similar al polipéptido acetilgalactosaminiltransferasa 5 N-       | 7q36.1                                 |
| GT27                           | 442117                                    | <i>GALNTL6</i>         | similar al polipéptido acetilgalactosaminiltransferasa 6 N-       | 4q34.1                                 |
| GT6                            | 26301                                     | <i>GBGT1</i>           | globósido α1,3-N-acetilgalactosaminiltransferasa 1                | 9q34.13-q34.3                          |
| GT14                           | 2650                                      | <i>GCNT1</i>           | glucosaminil (N-acetil) transferasa 1, núcleo 2                   | 9q13                                   |
| GT14                           | 2651                                      | <i>GCNT2</i>           | glucosaminil (N-acetil) transferasa 2, enzima ramificadora I      | 6p24.2                                 |
| GT14                           | 9245                                      | <i>GCNT3</i>           | glucosaminil (N-acetil) transferasa 3, tipo mucina                | 15q21.3                                |
| GT14                           | 51301                                     | <i>GCNT4</i>           | glucosaminil (N-acetil) transferasa 4, núcleo 2                   | 5q12                                   |
| GT14                           | 644378                                    | <i>GCNT6</i>           | glucosaminil (N-acetil) transferasa 6                             | 6p24.2                                 |
| GT14                           | 140687                                    | <i>GCNT7</i>           | miembro 7 de la familia de la glucosaminil (N-acetil) transferasa | 20q13.2                                |
| GT6                            | 2681                                      | <i>GGTA1P</i>          | glicoproteína, pseudogén de α-galactosiltransferasa 1             | 9q33.2                                 |
| GT4                            | 144423                                    | <i>GLT1D1</i>          | dominio 1 que contiene glicosiltransferasa 1                      | 12q24.33                               |
| GT6                            | 360203                                    | <i>GLT6D1</i>          | dominio 1 que contiene glicosiltransferasa 6                      | 9q34.3                                 |

| Familia CAZy <sup>(1)</sup> | Identificación genética <sup>(2)</sup> | Símbolo <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                    | Ubicación en el mapa <sup>(5)</sup> |
|-----------------------------|----------------------------------------|------------------------|-------------------------------------------------------------------------------|-------------------------------------|
| GT8                         | 55830                                  | <i>GLT8D1</i>          | dominio 1 que contiene glicosiltransferasa 8                                  | 3p21.1                              |
| GT8                         | 83468                                  | <i>GLT8D2</i>          | dominio 2 que contiene glicosiltransferasa 8                                  | 12q                                 |
| GT4                         | 79712                                  | <i>GTDC1</i>           | dominio 1 similar que contiene glicosiltransferasa                            | 2q22.3                              |
| GT8                         | 283464                                 | <i>GXYLT1</i>          | glucósido xilosiltransferasa 1                                                | 12q12                               |
| GT8                         | 727936                                 | <i>GXYLT2</i>          | xilósido xilosiltransferasa 2                                                 | 3p13                                |
| GT8                         | 2992                                   | <i>GYG1</i>            | glucogenina 1                                                                 | 3q24-q25.1                          |
| GT8                         | 8908                                   | <i>GYG2</i>            | glucogenina 2                                                                 | Xp22.3                              |
| GT8/49                      | 120071                                 | <i>GYLTL1B</i>         | similar a la glicosiltransferasa 1B, LARGE2                                   | 11p11.2                             |
| GT3                         | 2997                                   | <i>GYS1</i>            | glucógeno sintasa 1 (músculo)                                                 | 19q13.3                             |
| GT3                         | 2998                                   | <i>GYS2</i>            | glucógeno sintasa 2 (hígado)                                                  | 12p12.2                             |
| GT2                         | 3036                                   | <i>HAS1</i>            | hialuronano sintasa 1                                                         | 19q13.4                             |
| GT2                         | 3037                                   | <i>HAS2</i>            | hialuronano sintasa 2                                                         | 8q24.12                             |
| GT2                         | 3038                                   | <i>HAS3</i>            | hialuronano sintasa 3                                                         | 16q22.1                             |
| GT90                        | 79070                                  | <i>KDEL1C</i>          | KDEL (Lys-Asp-Glu-Leu) que contiene 1 glucosiltransferasa similar             | 13q33                               |
| GT8/49                      | 9215                                   | <i>LARGE</i>           | $\beta$ 1,3-xilosiltransferasa                                                | 22q12.3                             |
| GT31                        | 3955                                   | <i>LFNG</i>            | O-fucosilpéptido acetilglucosaminiltransferasa                                | 3- $\beta$ N- 7p22.2                |
| GT31                        | 4242                                   | <i>MFNG</i>            | O-fucosilpéptido acetilglucosaminiltransferasa                                | 3- $\beta$ N- 22q12                 |
| GT13                        | 4245                                   | <i>MGAT1</i>           | manosil $\alpha$ 1,3-glicoproteína acetilglucosaminiltransferasa              | $\beta$ 1,2N- 5q35                  |
| GT16                        | 4247                                   | <i>MGAT2</i>           | manosil $\alpha$ 1,6-glicoproteína acetilglucosaminiltransferasa              | $\beta$ 1,2N- 14q21                 |
| GT17                        | 4248                                   | <i>MGAT3</i>           | manosil $\alpha$ 1,4-glicoproteína acetilglucosaminiltransferasa              | $\beta$ 1,4N- 22q13.1               |
| GT54                        | 11320                                  | <i>MGAT4A</i>          | manosil $\alpha$ 1,3-glicoproteína acetilglucosaminiltransferasa              | $\beta$ 1,4N- 2q12                  |
| GT54                        | 11282                                  | <i>MGAT4B</i>          | manosil $\alpha$ 1,3-glicoproteína acetilglucosaminiltransferasa              | $\beta$ 1,4N- 5q35                  |
| GT54                        | 25834                                  | <i>MGAT4C</i>          | manosil $\alpha$ 1,3-glicoproteína acetilglucosaminiltransferasa              | $\beta$ 1,4N- 12q21                 |
| GT18                        | 4249                                   | <i>MGAT5</i>           | manosil $\alpha$ 1,6-glicoproteína acetilglucosaminiltransferasa              | $\beta$ 1,6N- 2q21.3                |
| GT18                        | 146664                                 | <i>MGAT5B</i>          | manosil $\alpha$ 1,6-glicoproteína acetilglucosaminiltransferasa, isoenzima B | $\beta$ 1,6N- 17q25.2               |
| GT41                        | 8473                                   | <i>OGT</i>             | N-acetilglucosamina O-ligada (GlcNAc) transferasa                             | Xq13                                |
| GT4                         | 5277                                   | <i>PIGA</i>            | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase A            | Xp22.1                              |
| GT22                        | 9488                                   | <i>PIGB</i>            | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase B            | 15q21.3                             |
| GT50                        | 93183                                  | <i>PIGM</i>            | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase M            | 1q23.2                              |
| GT76                        | 55650                                  | <i>PIGV</i>            | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase V            | 1p36.11                             |
| GT22                        | 80235                                  | <i>PIGZ</i>            | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase Z            | 3q29                                |
| GTnc                        | 8985                                   | <i>PLOD3</i>           | procolágeno-lisina, 2-oxoglutarato 5-dioxigenasa 3                            | 7q22                                |
| GT65                        | 23509                                  | <i>POFUT1</i>          | O-fucosiltransferasa 1                                                        | 20q11                               |
| GT68                        | 23275                                  | <i>POFUT2</i>          | O-fucosiltransferasa 2                                                        | 21q22.3                             |

| Familia CAZy <sup>(1)</sup> | Identificación genética <sup>(2)</sup> | Símbolo <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                                                           | Ubicación en el mapa <sup>(5)</sup> |
|-----------------------------|----------------------------------------|------------------------|----------------------------------------------------------------------------------------------------------------------|-------------------------------------|
| GT90                        | 56983                                  | <i>POGLUT1</i>         | O-glucosiltransferasa 1                                                                                              | 3q13.33                             |
| GT13                        | 55624                                  | <i>POMGNT1</i>         | O-ligada-manosa $\beta$ 1,2-N-acetilglucosaminiltransferasa 1                                                        | 1p34.1                              |
| GT61                        | 84892                                  | <i>POMGNT2</i>         | O-ligada-manosa $\beta$ 1,4-N-acetilglucosaminiltransferasa 2                                                        | 3p22.1                              |
| GT39                        | 10585                                  | <i>POMT1</i>           | O-manosil-transferasa 1                                                                                              | 9q34.1                              |
| GT39                        | 29954                                  | <i>POMT2</i>           | O-manosil-transferasa 2                                                                                              | 14q24                               |
| GT35                        | 5834                                   | <i>PYGB</i>            | fosforilasa, glucógeno; cerebro                                                                                      | 20p11.21                            |
| GT35                        | 5836                                   | <i>PYGL</i>            | fosforilasa, glucógeno, hígado                                                                                       | 14q21-q22                           |
| GT35                        | 5837                                   | <i>PYGM</i>            | fosforilasa, glucógeno, músculo                                                                                      | 11q12-q13.2                         |
| GT31                        | 5986                                   | <i>RFNG</i>            | O-fucosilpéptido 3- $\beta$ -N-acetilglucosaminiltransferasa                                                         | 17q25                               |
| GT29                        | 6482                                   | <i>ST3GAL1</i>         | $\beta$ -galactósido $\alpha$ -2,3-sialiltransferasa 1                                                               | 8q24.22                             |
| GT29                        | 6483                                   | <i>ST3GAL2</i>         | $\beta$ -galactósido $\alpha$ -2,3-sialiltransferasa 2                                                               | 16q22.1                             |
| GT29                        | 6487                                   | <i>ST3GAL3</i>         | $\beta$ -galactósido $\alpha$ -2,3-sialiltransferasa 3                                                               | 1p34.1                              |
| GT29                        | 6484                                   | <i>ST3GAL4</i>         | $\beta$ -galactósido $\alpha$ -2,3-sialiltransferasa 4                                                               | 11q24.2                             |
| GT29                        | 8869                                   | <i>ST3GAL5</i>         | $\beta$ -galactósido $\alpha$ -2,3-sialiltransferasa 5                                                               | 2p11.2                              |
| GT29                        | 10402                                  | <i>ST3GAL6</i>         | $\beta$ -galactósido $\alpha$ -2,3-sialiltransferasa 6                                                               | 3q12.1                              |
| GT29                        | 6480                                   | <i>ST6GAL1</i>         | $\beta$ -galactosamida $\alpha$ -2,6-sialiltransferasa 1                                                             | 3q27-q28                            |
| GT29                        | 84620                                  | <i>ST6GAL2</i>         | $\beta$ -galactosamida $\alpha$ -2,6-sialiltransferasa 2                                                             | 2q11.2-q12.1                        |
| GT29                        | 55808                                  | <i>ST6GALNAC1</i>      | $\alpha$ -N-acetil-neuraminil-2,3- $\beta$ -galactosil-1,3-N-acetilgalactosaminida $\alpha$ -2,6-sialiltransferasa 1 | 17q25.1                             |
| GT29                        | 10610                                  | <i>ST6GALNAC2</i>      | $\alpha$ -N-acetil-neuraminil-2,3- $\beta$ -galactosil-1,3-N-acetilgalactosaminida $\alpha$ -2,6-sialiltransferasa 2 | 17q25.1                             |
| GT29                        | 256435                                 | <i>ST6GALNAC3</i>      | $\alpha$ -N-acetil-neuraminil-2,3- $\beta$ -galactosil-1,3-N-acetilgalactosaminida $\alpha$ -2,6-sialiltransferasa 3 | 1p31.1                              |
| GT29                        | 27090                                  | <i>ST6GALNAC4</i>      | $\alpha$ -N-acetil-neuraminil-2,3- $\beta$ -galactosil-1,3-N-acetilgalactosaminida $\alpha$ -2,6-sialiltransferasa 4 | 9q34                                |
| GT29                        | 81849                                  | <i>ST6GALNAC5</i>      | $\alpha$ -N-acetil-neuraminil-2,3- $\beta$ -galactosil-1,3-N-acetilgalactosaminida $\alpha$ -2,6-sialiltransferasa 5 | 1p31.1                              |
| GT29                        | 30815                                  | <i>ST6GALNAC6</i>      | $\alpha$ -N-acetil-neuraminil-2,3- $\beta$ -galactosil-1,3-N-acetilgalactosaminida $\alpha$ -2,6-sialiltransferasa 6 | 9q34.11                             |
| GT29                        | 6489                                   | <i>ST8SIA1</i>         | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 1                                                             | 12p12.1-p11.2                       |
| GT29                        | 8128                                   | <i>ST8SIA2</i>         | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 2                                                             | 15q26                               |
| GT29                        | 51046                                  | <i>ST8SIA3</i>         | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 3                                                             | 18q21.31                            |
| GT29                        | 7903                                   | <i>ST8SIA4</i>         | $\alpha$ -N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 4                                                     | 5q21                                |
| GT29                        | 29906                                  | <i>ST8SIA5</i>         | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 5                                                             | 18q21.1                             |
| GT29                        | 338596                                 | <i>ST8SIA6</i>         | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 6                                                             | 10p12.33                            |
| GT66                        | 3703                                   | <i>STT3A</i>           | subunidad del complejo oligosacariltransferasa (catalítica)                                                          | 11q23.3                             |
| GT66                        | 201595                                 | <i>STT3B</i>           | subunidad del complejo oligosacariltransferasa (catalítica)                                                          | 3p23                                |
| GTnc                        | 10329                                  | <i>TMEM5</i>           | proteína transmembrana 5                                                                                             | 12q14.2                             |

| Familia CAZY <sup>(1)</sup> | Identificación genética <sup>(2)</sup> | Símbolo <sup>(3)</sup> | Descripción <sup>(4)</sup>                               | Ubicación en el mapa <sup>(5)</sup> |
|-----------------------------|----------------------------------------|------------------------|----------------------------------------------------------|-------------------------------------|
| GT21                        | 7357                                   | <i>ST8SIA</i>          | UDP-glucosa ceramida glucosiltransferasa                 | 9q31                                |
| GT24                        | 56886                                  | <i>UGGT1</i>           | UDP-glucosa glucoproteína glucosiltransferasa 1          | 2q14.3                              |
| GT24                        | 55757                                  | <i>UGGT2</i>           | UDP-glucosa glicoproteína glucosiltransferasa 2          | 13q32.1                             |
| GT1                         | 54658                                  | <i>UGT1A1</i>          | Familia de UDP glucuronosiltransferasa 1, polipéptido A1 | 2q37                                |
| GT1                         | 54575                                  | <i>UGT1A10</i>         | Familia UDP glucuronosiltransferasa 1, polipéptido A10   | 2q37                                |
| GT1                         | 54659                                  | <i>UGT1A3</i>          | Familia UDP glucuronosiltransferasa 1, polipéptido A3    | 2q37                                |
| GT1                         | 54657                                  | <i>UGT1A4</i>          | Familia UDP glucuronosiltransferasa 1, polipéptido A4    | 2q37                                |
| GT1                         | 54579                                  | <i>UGT1A5</i>          | Familia UDP glucuronosiltransferasa 1, polipéptido A5    | 2q37                                |
| GT1                         | 54578                                  | <i>UGT1A6</i>          | Familia UDP glucuronosiltransferasa 1, polipéptido A6    | 2q37                                |
| GT1                         | 54577                                  | <i>UGT1A7</i>          | Familia UDP glucuronosiltransferasa 1, polipéptido A7    | 2q37                                |
| GT1                         | 54576                                  | <i>UGT1A8</i>          | Familia UDP glucuronosiltransferasa 1, polipéptido A8    | 2q37                                |
| GT1                         | 54600                                  | <i>UGT1A9</i>          | Familia UDP glucuronosiltransferasa 1, polipéptido A9    | 2q37                                |
| GT1                         | 10941                                  | <i>UGT2A1</i>          | Familia UDP glucuronosiltransferasa 2, polipéptido A1    | 4q13                                |
| GT1                         | 79799                                  | <i>UGT2A3</i>          | Familia UDP glucuronosiltransferasa 2, polipéptido A3    | 4q13.2                              |
| GT1                         | 7365                                   | <i>UGT2B10</i>         | Familia UDP glucuronosiltransferasa 2, polipéptido B10   | 4q13.2                              |
| GT1                         | 10720                                  | <i>UGT2B11</i>         | Familia UDP glucuronosiltransferasa 2, polipéptido B11   | 4q13.2                              |
| GT1                         | 7366                                   | <i>UGT2B15</i>         | Familia UDP glucuronosiltransferasa 2, polipéptido B15   | 4q13                                |
| GT1                         | 7367                                   | <i>UGT2B17</i>         | Familia UDP glucuronosiltransferasa 2, polipéptido B17   | 4q13                                |
| GT1                         | 54490                                  | <i>UGT2B28</i>         | Familia UDP glucuronosiltransferasa 2, polipéptido B28   | 4q13.2                              |
| GT1                         | 7363                                   | <i>UGT2B4</i>          | Familia UDP glucuronosiltransferasa 2, polipéptido B4    | 4q13                                |
| GT1                         | 7364                                   | <i>UGT2B7</i>          | Familia UDP glucuronosiltransferasa 2, polipéptido B7    | 4q 13                               |
| GT1                         | 133688                                 | <i>UGT3A1</i>          | Familia UDP glicosiltransferasa 3, polipéptido A1        | 5p13.2                              |
| GT1                         | 167127                                 | <i>UGT3A2</i>          | Familia UDP glicosiltransferasa 3, polipéptido A2        | 5p13.2                              |
| GT1                         | 7368                                   | <i>UGT8</i>            | UDP glicosiltransferasa 8                                | 4q26                                |
| GT27                        | 64409                                  | <i>WBSCR17</i>         | Región cromosómica 17 del síndrome de Williams-Beuren    | 7q11.23                             |
| GT8                         | 152002                                 | <i>XXYL1</i>           | xilósido xilosiltransferasa 1                            | 3q29                                |
| GT14                        | 64131                                  | <i>XYLT1</i>           | xilosiltransferasa I                                     | 16p12.3                             |
| GT14                        | 64132                                  | <i>XYLT2</i>           | xilosiltransferasa II                                    | 17q21.33                            |

<sup>(1)</sup> Sistema de clasificación GT (Lombard et al. (2013). Nucl Acid Res 42:D1P:D490-D495)<sup>(2)</sup> símbolo del gen<sup>(3)</sup> ID del gen GenBank<sup>(4)</sup> Nombre de la proteína UniProt<sup>(5)</sup> Posición cromosómica humana (hg19)

La mayoría de los genes de glicosiltransferasa de CHO ortólogos correspondientes se asignaron previamente en relación con la secuenciación reciente del genoma de CHO (Xu, Nagarajan et al. 2011), pero algunos genes se asignaron incorrectamente o se omitieron. El conjunto actual de genes de glicosiltransferasa ortólogos en humanos y CHO se enumeran en la TABLA 2.

Tabla 2. Genes GTfs del hámster chino (*Cricetulus griseus*) Familia CAZy<sup>(1)</sup> Símbolo<sup>(2)</sup> ID del gen<sup>(3)</sup> Descripción<sup>(4)</sup>

| Familia CAZy <sup>(1)</sup> | Símbolo <sup>(2)</sup> | Identificación genética <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                       |
|-----------------------------|------------------------|----------------------------------------|----------------------------------------------------------------------------------|
| GT1                         | <i>Alg13</i>           | 100754023                              | Subunidad de la UDP-N-acetilglucosaminiltransferasa                              |
| GT1                         | <i>Alg14</i>           | 100773644                              | Subunidad de la UDP-N-acetilglucosaminiltransferasa                              |
| GT1                         | <i>Ugt1a1</i>          | 100755423                              | UDP-glucuronosiltransferasa 1-6                                                  |
| GT1                         | <i>Ugt2a1</i>          | 100762963                              | Familia 2 de la UDP glucuronosiltransferasa, polipéptido A1, locus complejo      |
| GT1                         | <i>Ugt2a3</i>          | 100762673                              | UDP-glucuronosiltransferasa 2A3                                                  |
| GT2                         | <i>Alg5</i>            | 100769679                              | ALG5, dolícil-fosfato $\beta$ -glucosiltransferasa                               |
| GT2                         | <i>B3gnt1</i>          | 100751193                              | similar a UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 1 |
| GT2                         | <i>Dpm1</i>            | 100689420                              | dolícil-fosfato mannosiltransferasa polipéptido 1, subunidad catalítica          |
| GT2                         | <i>Has2</i>            | 100751055                              | hialuronano sintasa 2                                                            |
| GT2                         | <i>Has2</i>            | 100751055                              | hialuronano sintasa 2                                                            |
| GT2                         | <i>Has3</i>            | 100757895                              | hialuronano sintasa 3                                                            |
| GT3                         | <i>Gys1</i>            | 100769788                              | glucógeno sintasa 1 (músculo)                                                    |
| GT3                         | <i>Gys2</i>            | 100770628                              | glucógeno sintasa 2 (hígado)                                                     |
| GT4                         | <i>Alg11</i>           | 100771009                              | $\alpha$ -1,2-manosiltransferasa                                                 |
| GT4                         | <i>Alg2</i>            | 100768412                              | $\alpha$ -1,3/1,6-manosiltransferasa                                             |
| GT4                         | <i>Glt1d1</i>          | 100761757                              | dominio 1 que contiene glicosiltransferasa 1                                     |
| GT4                         | <i>Gtdc1</i>           | 100752161                              | dominio 1 similar que contiene glicosiltransferasa                               |
| GT4                         | <i>Piga</i>            | 100773712                              | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase A               |
| GT6                         | <i>Abo</i>             | 100772592                              | Grupo sanguíneo ABO                                                              |
| GT6                         | <i>Gbgt1</i>           | 100771256                              | globoside 1 $\beta$ 1,3-N-acetil-galactosaminiltransferasa                       |
| GT7                         | <i>B4galnt3</i>        | 100756528                              | $\beta$ -1,4-N-acetil-galactosaminiltransferasa 3                                |
| GT7                         | <i>B4galnt4</i>        | 100758404                              | $\beta$ -1,4-N-acetil-galactosaminiltransferasa 4                                |
| GT7                         | <i>B4galt1</i>         | 100689430                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 1         |
| GT7                         | <i>B4galt2</i>         | 100689434                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 2         |
| GT7                         | <i>B4galt3</i>         | 100689346                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 3         |
| GT7                         | <i>B4galnt4</i>        | 100689435                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 4         |
| GT7                         | <i>B4galt5</i>         | 100689347                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 5         |
| GT7                         | <i>B4galt6</i>         | 100689438                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,4-galactosiltransferasa, polipéptido 6         |
| GT7                         | <i>B4galt7</i>         | 100769652                              | xilosilproteína $\beta$ 1,4-galactosiltransferasa, polipéptido 7                 |
| GT7/GT3 1                   | <i>Chpf</i>            | 100765856                              | factor polimerizador de condroitina                                              |
| GT7/GT3 1                   | <i>Chsy1</i>           | 100770803                              | condroitín sulfato sintasa 1                                                     |
| GT7/GT3 1                   | <i>Chsy3</i>           | 100767985                              | condroitín sulfato sintasa 3                                                     |
| GT7                         | <i>Csgalnact1</i>      | 100754969                              | sulfato de condroitina N-acetilgalactosaminiltransferasa 1                       |
| GT7                         | <i>Csgalnact2</i>      | 100770830                              | sulfato de condroitina N-acetilgalactosaminiltransferasa 2                       |
| GT8                         | <i>Glt8d1</i>          | 100762757                              | dominio 1 que contiene glicosiltransferasa 8                                     |
| GT8                         | <i>Glt8d2</i>          | 100772458                              | dominio 2 que contiene glicosiltransferasa 8                                     |
| GT8                         | <i>Gxylt1</i>          | 100760543                              | glucósido xilosiltransferasa 1                                                   |
| GT8                         | <i>Gxylt2</i>          | 100758512                              | xilósido xilosiltransferasa 2                                                    |
| GT8                         | <i>Gyg1</i>            | 100751671                              | glucogenina 1                                                                    |
| GT8/49                      | <i>Gylt1b</i>          | 100750932                              | similar a la glicosiltransferasa 1B                                              |
| GT8/49                      | <i>Grande</i>          | 100765249                              | similar a laglicosiltransferasa                                                  |
| GT8                         | <i>Xxylt1</i>          | 100760146                              | xilósido xilosiltransferasa 1                                                    |

| Familia CAZy <sup>(1)</sup> | Símbolo <sup>(2)</sup> | Identificación genética <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                         |
|-----------------------------|------------------------|----------------------------------------|------------------------------------------------------------------------------------|
| GT10                        | <i>Fut10</i>           | 100760260                              | fucosiltransferasa 10 ( $\alpha$ (1,3) fucosiltransferasa)                         |
| GT10                        | <i>Fut11</i>           | 100758336                              | fucosiltransferasa 11 ( $\alpha$ (1,3) fucosiltransferasa)                         |
| GT10                        | <i>Fut4</i>            | 100754814                              | fucosiltransferasa 4 ( $\alpha$ (1,3) fucosiltransferasa, meloide específica)      |
| GT10                        | <i>Fut6a</i>           | 100689084                              | $\alpha$ 1,3 fucosiltransferasa 6A                                                 |
| GT10                        | <i>Fut6b</i>           | 100689083                              | $\alpha$ 1,3 fucosiltransferasa 6B                                                 |
| GT10                        | <i>Fut7</i>            | 100772214                              | $\alpha$ 1,3 fucosiltransferasa                                                    |
| GT10                        | <i>Fut9</i>            | 100689036                              | $\alpha$ 1,3 fucosiltransferasa                                                    |
| GT11                        | <i>Fut1</i>            | 100757047                              | galactósido 2o-L-fucosiltransferasa, grupo sanguíneo H                             |
| GT11                        | <i>Fut2</i>            | 100751185                              | fucosiltransferasa 2                                                               |
| GT12                        | <i>B4galnt1</i>        | 100764682                              | $\beta$ 1,4-N-acetil-galactosaminiltransferasa 1                                   |
| GT12                        | <i>B4galnt2</i>        | 100760696                              | $\beta$ 1,4-N-acetil-galactosaminiltransferasa 2                                   |
| GT13                        | <i>Mgat1</i>           | 100682529                              | manosil ( $\alpha$ 1,3)-glicoproteína $\beta$ 1,2N-acetilglucosaminiltransferasa   |
| GT13                        | <i>Pomgnt1</i>         | 100772511                              | proteína O-ligada manosa N-acetilglucosaminiltransferasa 1 ( $\beta$ 1,2-)         |
| GT14                        | <i>Gent1</i>           | 100767124                              | glucosaminil (N-acetil) transferasa 1, núcleo 2                                    |
| GT14                        | <i>Gen t3</i>          | 100774815                              | glucosaminil (N-acetil) transferasa 3, tipo mucina                                 |
| GT14                        | <i>Gen t4</i>          | 100757239                              | glucosaminil (N-acetil) transferasa 4, núcleo 2                                    |
| GT14                        | <i>Gcnt7</i>           | 100760777                              | miembro 7 de la familia de la glucosaminil (N-acetil) transferasa                  |
| GT14                        | <i>Xylt1</i>           | 100756878                              | xilosiltransferasa I                                                               |
| GT14                        | <i>Xylt2</i>           | 100759604                              | xilosiltransferasa II                                                              |
| GT16                        | <i>Mgat2</i>           | 100753385                              | manosil ( $\alpha$ 1,6)-glicoproteína $\beta$ -1,2-N-acetilglucosaminiltransferasa |
| GT17                        | <i>Mgat3</i>           | 100689076                              | manosil ( $\beta$ 1,4)-glicoproteína $\beta$ -1,4-N-acetilglucosaminiltransferasa  |
| GT18                        | <i>Mgat5</i>           | 100760162                              | manosil ( $\beta$ 1,6)-glicoproteína $\beta$ -1,6-N-acetilglucosaminiltransferasa  |
| GT18                        | <i>Mgat5b</i>          | 100771275                              | manosil ( $\alpha$ 1,6)-glicoproteína $\beta$ -1,6-N-acetilglucosaminiltransferasa |
| GT21                        | <i>Ugcg</i>            | 100689432                              | UDP-glucosa ceramida glucosiltransferasa                                           |
| GT22                        | <i>Alg12</i>           | 100770096                              | $\alpha$ -1,6-manosiltransferasa                                                   |
| GT22                        | <i>Alg9</i>            | 100755062                              | $\alpha$ -1,2-manosiltransferasa                                                   |
| GT22                        | <i>Pigb</i>            | 100768002                              | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase B                 |
| GT22                        | <i>Pigz</i>            | 100750622                              | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase Z                 |
| GT23                        | <i>Fut8</i>            | 100751648                              | $\alpha$ -1,6-manosiltransferasa)                                                  |
| GT24                        | <i>Uggt1</i>           | 100773968                              | UDP-glucosa glucoproteína glucosiltransferasa 1                                    |
| GT24                        | <i>Uggt2</i>           | 100762273                              | UDP-glucosa glicoproteína glucosiltransferasa 2                                    |
| GT25                        | <i>Cercam</i>          | 100765284                              | molécula de adhesión de células endoteliales cerebrales                            |
| GT25                        | <i>Colgalt1</i>        | 100774081                              | colágeno $\beta$ (1-O)galactosiltransferasa 1                                      |
| GT25                        | <i>Colgalt2</i>        | 100764465                              | colágeno $\beta$ (1-O)galactosiltransferasa 2                                      |
| GT27                        | <i>Galnt1</i>          | 100763868                              | polipéptido N-acetilgalactosaminiltransferasa 1                                    |
| GT27                        | <i>Galnt10</i>         | 100768523                              | polipéptido N-acetilgalactosaminiltransferasa 10                                   |
| GT27                        | <i>Galnt11</i>         | 100758359                              | polipéptido N-acetilgalactosaminiltransferasa 11                                   |
| GT27                        | <i>Galnt12</i>         | 100772146                              | polipéptido N-acetilgalactosaminiltransferasa 12                                   |

| Familia CAZy <sup>(1)</sup> | Símbolo <sup>(2)</sup> | Identificación genética <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                                        |
|-----------------------------|------------------------|----------------------------------------|---------------------------------------------------------------------------------------------------|
| GT27                        | <i>Galnt13</i>         | 100768831                              | polipéptido N-<br>acetilgalactosaminiltransferasa 13                                              |
| GT27                        | <i>Galnt14</i>         | 100759781                              | polipéptido N-<br>acetilgalactosaminiltransferasa 14                                              |
| GT27                        | <i>Galnt15</i>         | 100766936                              | polipéptido N-<br>acetilgalactosaminiltransferasa 15                                              |
| GT27                        | <i>Galnt16</i>         | 100764829                              | polipéptido N-<br>acetilgalactosaminiltransferasa 16                                              |
| GT27                        | <i>Galnt18</i>         | 100767393                              | polipéptido N-<br>acetilgalactosaminiltransferasa 18                                              |
| GT27                        | <i>Galnt2</i>          | 100767525                              | polipéptido N-<br>acetilgalactosaminiltransferasa 2                                               |
| GT27                        | <i>Galnt3</i>          | 100753226                              | polipéptido N-<br>acetilgalactosaminiltransferasa 3                                               |
| GT27                        | <i>Galnt4</i>          | 100765247                              | polipéptido N-<br>acetilgalactosaminiltransferasa 4                                               |
| GT27                        | <i>Galnt5</i>          | 100757219                              | polipéptido N-<br>acetilgalactosaminiltransferasa 5                                               |
| GT27                        | <i>Galnt6</i>          | 100751126                              | polipéptido N-<br>acetilgalactosaminiltransferasa 6                                               |
| GT27                        | <i>Galnt6</i>          | 100751126                              | polipéptido N-<br>acetilgalactosaminiltransferasa 6                                               |
| GT27                        | <i>Galnt7</i>          | 100762043                              | polipéptido N-<br>acetilgalactosaminiltransferasa 7                                               |
| GT27                        | <i>Galnt8</i>          | 100764127                              | polipéptido N-<br>acetilgalactosaminiltransferasa 8                                               |
| GT27                        | <i>Galnt9</i>          | 100773797                              | polipéptido N-<br>acetilgalactosaminiltransferasa 9                                               |
| GT27                        | <i>Galnt15</i>         | 100767659                              | similar al polipéptido N-<br>acetilgalactosaminiltransferasa 5                                    |
| GT27                        | <i>Galnt16</i>         | 100761752                              | similar al polipéptido N-<br>acetilgalactosaminiltransferasa 6                                    |
| GT27                        | <i>Wbser17</i>         | 100750837                              | Región cromosómica 17 del síndrome de Williams-Beuren                                             |
| GT29                        | <i>St3gal1</i>         | 100754088                              | β-galactósido α-2,3-sialiltransferasa 1                                                           |
| GT29                        | <i>St3gal1</i>         | 100754088                              | β-galactósido α-2,3-sialiltransferasa 1                                                           |
| GT29                        | <i>St3gal2</i>         | 100767717                              | β-galactósido α-2,3-sialiltransferasa 2                                                           |
| GT29                        | <i>St3gal2</i>         | 100767717                              | β-galactósido α-2,3-sialiltransferasa 2                                                           |
| GT29                        | <i>St3gal3</i>         | 100689187                              | β-galactósido α-2,3-sialiltransferasa 3                                                           |
| GT29                        | <i>St3gal4</i>         | 100689440                              | β-galactósido α-2,3-sialiltransferasa 4                                                           |
| GT29                        | <i>St3gal5</i>         | 100754838                              | β-galactósido α-2,3-sialiltransferasa 5                                                           |
| GT29                        | <i>St3gal6</i>         | 100771326                              | β-galactósido α-2,3-sialiltransferasa 6                                                           |
| GT29                        | <i>St6gal1</i>         | 100689389                              | β-galactosamida α-2,6-sialiltransferasa 1                                                         |
| GT29                        | <i>St6gal2</i>         | 100763756                              | β-galactosamida α-2,6-sialiltransferasa 2                                                         |
| GT29                        | <i>St6galnac1</i>      | 100763224                              | o-N-acetilgalactosaminida α-2,6-sialiltransferasa 1                                               |
| GT29                        | <i>St6galnac3</i>      | 100762285                              | (α-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida α-2,6-sialiltransferasa 3 |
| GT29                        | <i>St6galnac4</i>      | 100759065                              | (α-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida α-2,6-sialiltransferasa 4 |
| GT29                        | <i>St6galnac5</i>      | 100759381                              | (α-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida α-2,6-sialiltransferasa 5 |
| GT29                        | <i>St6galnac6</i>      | 100757138                              | (α-N-acetil-neuraminil-2,3-beta-galactosil-1,3)-N-acetilgalactosaminida α-2,6-sialiltransferasa 6 |
| GT29                        | <i>St8sia1</i>         | 100768920                              | o-N-acetil-neuraminida α-2,8-sialiltransferasa 1                                                  |
| GT29                        | <i>St8sia2</i>         | 100759559                              | o-N-acetil-neuraminida α-2,8-sialiltransferasa 2                                                  |

| Familia CAZy <sup>(1)</sup> | Símbolo <sup>(2)</sup> | Identificación genética <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                   |
|-----------------------------|------------------------|----------------------------------------|------------------------------------------------------------------------------|
| GT29                        | <i>St8sia3</i>         | 100764454                              | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 3                     |
| GT29                        | <i>St8sia4</i>         | 100689217                              | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 4                     |
| GT29                        | <i>St8sia5</i>         | 100750766                              | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 5                     |
| GT29                        | <i>St8sia6</i>         | 100774188                              | o-N-acetil-neuraminida $\alpha$ -2,8-sialiltransferasa 6                     |
| GT31                        | <i>B3galnt1</i>        | 100756438                              | $\beta$ -1,3-N-acetilgalactosaminiltransferasa 1 (grupo sanguíneo globósido) |
| GT31                        | <i>B3galnt2</i>        | 100768564                              | $\beta$ 1,3-N-acetilgalactosaminiltransferasa 2                              |
| GT31                        | <i>B3galt1</i>         | 100761765                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 1     |
| GT31                        | <i>B3galt2</i>         | 100766715                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 2     |
| GT31                        | <i>B3galt4</i>         | 100751197                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 4     |
| GT31                        | <i>B3galt5</i>         | 100755714                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 5     |
| GT31                        | <i>B3galt6</i>         | 100767598                              | UDP-Gal: $\beta$ GlcNAc $\beta$ 1,3-galactosiltransferasa, polipéptido 6     |
| GT31                        | <i>B3galtl</i>         | 100759864                              | Similar a la $\beta$ 1,3-galactosiltransferasa                               |
| GT31                        | <i>B3gnt2</i>          | 100766691                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 2       |
| GT31                        | <i>B3gnt3</i>          | 100773213                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 3       |
| GT31                        | <i>B3gnt4</i>          | 100769077                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 4       |
| GT31                        | <i>B3gnt5</i>          | 100757919                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 5       |
| GT31                        | <i>B3gnt6</i>          | 100768656                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 6       |
| GT31                        | <i>B3gnt7</i>          | 100760456                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 7       |
| GT31                        | <i>B3gnt8</i>          | 103160327                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 8       |
| GT31                        | <i>B3gnt9</i>          | 103159649                              | UDP-GlcNAc: $\beta$ Gal $\beta$ -1,3-N-acetilglucosaminiltransferasa 9       |
| GT31                        | <i>C1galt1</i>         | 100761169                              | glicoproteína-N-acetilgalactosamina $3$ - $\beta$ -galactosiltransferasa, 1  |
| GT31                        | <i>C1galt1c1</i>       | 100751243                              | Chaperona 1 específica de C1GALT1                                            |
| GT31                        | <i>Lfng</i>            | 100762397                              | O-fucosilpéptido $3$ - $\beta$ -N-acetilglucosaminiltransferasa              |
| GT31                        | <i>Mfng</i>            | 100762875                              | O-fucosilpéptido $3$ - $\beta$ -N-acetilglucosaminiltransferasa              |
| GT31                        | <i>Rfng</i>            | 100771257                              | O-fucosilpéptido $3$ - $\beta$ -N-acetilglucosaminiltransferasa              |
| GT32                        | <i>A4galt</i>          | 100770462                              | $\alpha$ -1,4-galactosiltransferasa                                          |
| GT32                        | <i>A4gnt</i>           | 100771969                              | $\alpha$ -1,4-N-acetilglucosaminiltransferasa                                |
| GT33                        | <i>Alg1</i>            | 100773731                              | quitobiosildifosfodolico $\beta$ -manosiltransferasa                         |
| GT35                        | <i>Pygb</i>            | 100769186                              | fosforilasa, glucógeno; cerebro                                              |
| GT35                        | <i>Pygm</i>            | 100757350                              | fosforilasa, glucógeno, músculo                                              |
| GT39                        | <i>Pomt1</i>           | 100755033                              | proteína-O-manosiltransferasa 1                                              |
| GT39                        | <i>Pomt2</i>           | 100764752                              | proteína-O-manosiltransferasa 2                                              |
| GT41                        | <i>Ogt</i>             | 100768670                              | N-acetilglucosamina O-ligada (GlcNAc) transferasa                            |
| GT43                        | <i>B3gat1</i>          | 100750701                              | $\beta$ -1,3-glucuroniltransferasa $1$ (glucuronosiltransferasa P)           |
| GT43                        | <i>B3gat2</i>          | 100756696                              | $\beta$ -1,3-glucuroniltransferasa $2$ (glucuronosiltransferasa S)           |

| Familia CAZy <sup>(1)</sup> | Símbolo <sup>(2)</sup> | Identificación genética <sup>(3)</sup> | Descripción <sup>(4)</sup>                                                      |
|-----------------------------|------------------------|----------------------------------------|---------------------------------------------------------------------------------|
| GT43                        | <i>B3gat3</i>          | 100689419                              | β-1,3-glucuroniltransferasa (glucuronosiltransferasa I) 3                       |
| GT47/64                     | <i>Ext1</i>            | 100689334                              | exostosina glicosiltransferasa 1                                                |
| GT47/64                     | <i>Ext2</i>            | 100751585                              | exostosina glicosiltransferasa 2                                                |
| GT47/64                     | <i>Extl1</i>           | 100770000                              | glicosiltransferasa 1 similar a la exostosina                                   |
| GT47/64                     | <i>Extl3</i>           | 100751999                              | glicosiltransferasa 3 similar a la exostosina                                   |
| GT49                        | <i>B3gnt1</i>          | 100762253                              | UDP-GlcNAc:βGal β-1,3-N-acetilglucosaminiltransferasa 1                         |
| GT50                        | <i>Pigm</i>            | 100764842                              | biosíntesis de anclajes de glicanos de fosfatidilinositol, clase M              |
| GT54                        | <i>Mgat4a</i>          | 100766200                              | manosil (α1,3)-glicoproteína β-1,4-N-acetilglucosaminiltransferasa              |
| GT54                        | <i>Mgat4b</i>          | 100768637                              | manosil (α1,3)-glicoproteína β-1,4-N-acetilglucosaminiltransferasa              |
| GT54                        | <i>Mgat4c</i>          | 100760589                              | manosil (α-1,3)-glicoproteína β-1,4-N-acetilglucosaminiltransferasa             |
| GT57                        | <i>Alg6</i>            | 100753783                              | α-1,3-glucosiltransferasa                                                       |
| GT57                        | <i>Alg8</i>            | 100766150                              | α-1,3-glucosiltransferasa                                                       |
| GT58                        | <i>Alg3</i>            | 100772003                              | α-1,3- manosiltransferasa                                                       |
| GT61                        | <i>Eogt</i>            | 100757071                              | N-acetilglucosamina O-ligada (GlcNAc) transferasa específica del dominio EGF    |
| GT61                        | <i>Pomgnt2</i>         | 100770708                              | proteína manosa ligada a O N-acetilglucosaminiltransferasa 2 (β1,4-)            |
| GT64                        | <i>Extl2</i>           | 100761218                              | glicosiltransferasa 2 similar a la exostosina                                   |
| GT65                        | <i>Pofut1</i>          | 100753417                              | proteína O-fucosiltransferasa 1                                                 |
| GT66                        | <i>Stt3a</i>           | 100751391                              | subunidad del complejo oligosacariltransferasa (catalítica)                     |
| GT66                        | <i>Stt3b</i>           | 100752084                              | subunidad del complejo oligosacariltransferasa (catalítica)                     |
| GT68                        | <i>Pofut2</i>          | 100765129                              | proteína O-fucosiltransferasa 2                                                 |
| GT90                        | <i>Kdelc1</i>          | 100763723                              | KDEL (Lys-Asp-Glu-Leu) que contiene 1                                           |
| GT90                        | <i>Poglut1</i>         | 100760325                              | proteína O-glucosiltransferasa 1                                                |
| GTnc                        | <i>Dpy19/1</i>         | 100765722                              | similar a dpy-19 1 (C. elegans)                                                 |
| GTnc                        | <i>Dpy19/2</i>         | 100752353                              | similar a dpy-19 2 (C. elegans)                                                 |
| GTnc                        | <i>Dpy19/3</i>         | 100756435                              | similar a dpy-19 3 (C. elegans)                                                 |
| GTnc                        | <i>Dpy19/4</i>         | 100759810                              | similar a dpy-19 4 (C. elegans)                                                 |
| GTnc                        | <i>Fktn</i>            | 100761516                              | fukutin                                                                         |
| GTnc                        | <i>Plod3</i>           | 100768993                              | procolágeno-lisina, 2-oxoglutarato 5-dioxigenasa 3                              |
| GTnc                        | <i>Tmem5</i>           | 100757873                              | proteína transmembrana 5                                                        |
| -                           | <i>Dpagt1</i>          | 100689054                              | dolicil-fosfato (UDP-N-acetilglucosamina) N-acetilglucosaminafosfotransferasa 1 |

<sup>(1)</sup> Sistema de clasificación GT (Lombard et al. (2013). Nucl Acid Res 42:D1P:D490-D495)<sup>(2)</sup> símbolo del gen HGNC<sup>(3)</sup> ID del gen GenBank<sup>(4)</sup> Nombre de la proteína UniProt

De manera similar, los datos del transcriptoma para un clon CHO-K1 informados en Xu et al. (Xu, Nagarajan et al. 2011) no incluyeron varios genes de glicosiltransferasa. Por lo tanto, realizamos un análisis de secuenciación de ARN de un panel de líneas de CHO, incluidas aquellas con knockout de genes de glicosiltransferasas mediado por ZFN, para evaluar la expresión de todos los genes de glicosiltransferasas en CHO. La TABLA 3 proporciona una lista de los niveles de expresión de todos los genes de glicosiltransferasa de CHO asignados, y se encontraron algunas diferencias cualitativas en los niveles de expresión en comparación con los informados previamente (Xu, Nagarajan et al. 2011).

10

Tabla 3. Comparación de los niveles de expresión de GTfs a partir de datos de ARN\_seq para CHO-GS y CHO-K1

| Familia CAZy | Genes GTf (hGTfs de la Tabla 1)* | CHO-GS_FPKM ** (4) | Profundidad de CHO-K1 ARN_seq (Xu et al 2011)**(4) |
|--------------|----------------------------------|--------------------|----------------------------------------------------|
| GT1          | <i>Alg13</i>                     | 13                 | na                                                 |

| Familia CAZy | Genes GTf (hGTfs de la Tabla 1)* | CHO-GS_FPKM ** (4) | Profundidad de CHO-K1 ARN_seq (Xu et al 2011)**(4) |
|--------------|----------------------------------|--------------------|----------------------------------------------------|
| GT1          | <i>Alg14</i>                     | 27                 | 55                                                 |
| GT1          | <i>Ugt1a1-10</i> <sup>(21)</sup> | 63                 | 96                                                 |
| GT1          | <i>Ugt2a1</i> <sup>(1)</sup>     | 0                  | 0                                                  |
| GT1          | <i>Ugt2a3</i> <sup>(1)</sup>     | 0                  | na                                                 |
| GT1          | <i>Ugt2b10</i> <sup>(1)</sup>    | na                 | 0                                                  |
| GT1          | <i>Ugt2b11</i> <sup>(1)</sup>    | na                 | na                                                 |
| GT1          | <i>Ugt2b15</i> <sup>(1)</sup>    | na                 | na                                                 |
| GT1          | <i>Ugt2b17</i> <sup>(1)</sup>    | na                 | na                                                 |
| GT1          | <i>Ugt2b28</i> <sup>(1)</sup>    | na                 | 0                                                  |
| GT1          | <i>Ugt2b4</i> <sup>(1)</sup>     | na                 | na                                                 |
| GT1          | <i>Ugt2b7</i> <sup>(1)</sup>     | na                 | na                                                 |
| GT1          | <i>Ugt3a1</i> <sup>(1)</sup>     | na                 | na                                                 |
| GT1          | <i>Ugt3a2</i> <sup>(1)</sup>     | na                 | na                                                 |
| GT1          | <i>Ugt8</i>                      | na                 | 0                                                  |
| GT2          | <i>Alg5</i>                      | 53                 | 71                                                 |
| GT2          | <i>B3gnt1</i>                    | 11                 | na                                                 |
| GT2          | <i>Dpm1</i>                      | 85                 | 78                                                 |
| GT2          | <i>Has1</i>                      | na                 | na                                                 |
| GT2          | <i>Has2</i>                      | 0                  | 0                                                  |
| GT2          | <i>Has3</i>                      | 0                  | 2                                                  |
| GT3          | <i>Gys1</i>                      | 47                 | na                                                 |
| GT3          | <i>Gys2</i>                      | 0                  | na                                                 |
| GT4          | <i>Alg11</i>                     | 16                 | 20                                                 |
| GT4          | <i>Alg2</i>                      | 19                 | 52                                                 |
| GT4          | <i>Glt1d1</i>                    | 0                  | na                                                 |
| GT4          | <i>Gtde1</i>                     | 10                 | na                                                 |
| GT4          | <i>Piga</i>                      | 8                  | 9                                                  |
| GT6          | <i>Abo</i>                       | 0                  | na                                                 |
| GT6          | <i>Gbgt1</i>                     | 4                  | na                                                 |
| GT6          | <i>Ggt1p</i>                     | na                 | na                                                 |
| GT6          | <i>Glt6d1</i>                    | na                 | na                                                 |
| GT7          | <i>B4galInt3</i>                 | 0                  | 0                                                  |
| GT7          | <i>B4galInt4</i>                 | 0                  | 0                                                  |
| GT7          | <i>B4galt1</i>                   | 15                 | 36                                                 |
| GT7          | <i>B4galt2</i>                   | 38                 | 41                                                 |
| GT7          | <i>B4galt3</i>                   | 25                 | 74                                                 |
| GT7          | <i>B4galInt4</i>                 | 19                 | 28                                                 |

| Familia CAZy | Genes GTf (hGTfs de la Tabla 1)* | CHO-GS_FPKM ** (4) | Profundidad de CHO-K1 ARN_seq (Xu et al 2011)**(4) |
|--------------|----------------------------------|--------------------|----------------------------------------------------|
| GT7          | <i>B4galt5</i>                   | 14                 | 71                                                 |
| GT7          | <i>B4galt6</i>                   | 8                  | 71                                                 |
| GT7          | <i>B4galt7</i>                   | 31                 | 169                                                |
| GT7/31       | <i>Chpf</i>                      | 53                 | 332                                                |
| GT7/31       | <i>Chpf2</i>                     | na                 | 130                                                |
| GT7/31       | <i>Chsy1</i>                     | 26                 | 0                                                  |
| GT7/31       | <i>Chys3</i>                     | na                 | na                                                 |
| GT7          | <i>Csgalnact1</i>                | 0                  | 0                                                  |
| GT7          | <i>Csgalnact2</i>                | 9                  | 0                                                  |
| GT8          | <i>Glt8d1</i>                    | 20                 | na                                                 |
| GT8          | <i>Glt8d2</i>                    | 0                  | na                                                 |
| GT8          | <i>Gxylt1</i>                    | 10                 | na                                                 |
| GT8          | <i>Gxylt2</i>                    | 0                  | na                                                 |
| GT8          | <i>Gyg</i>                       | 72                 | na                                                 |
| GT8          | <i>Gyg2</i>                      | na                 | na                                                 |
| GT8/49       | <i>Gylt1b</i>                    | 0                  | 0                                                  |
| GT8/49       | <i>Grande</i>                    | 14                 | 22                                                 |
| GT8          | <i>Xxylt1</i>                    | 70                 | na                                                 |
| GT10         | <i>Fut10</i>                     | 0                  | 0                                                  |
| GT10         | <i>Fut11</i>                     | 5                  | 0                                                  |
| GT10         | <i>Fut4</i>                      | 0                  | 0                                                  |
| GT10         | <i>Fut5</i>                      | na                 | 0                                                  |
| GT10         | <i>Fut6a</i>                     | 0                  | 0                                                  |
| GT10         | <i>Fut6b</i>                     | 0                  | 0                                                  |
| GT10         | <i>Fut7</i>                      | 0                  | 0                                                  |
| GT10         | <i>Fut9</i>                      | 0                  | 0                                                  |
| GT11         | <i>Fut1</i>                      | 0                  | 0                                                  |
| GT11         | <i>Fut2</i>                      | 0                  | 0                                                  |
| GT12         | <i>B4galnt1</i>                  | 0                  | 0                                                  |
| GT12         | <i>B4galnt2</i>                  | 0                  | 0                                                  |
| GT13         | <i>Mgat1</i>                     | 24                 | 61                                                 |
| GT13         | <i>Pomgnt1</i>                   | 64                 | 101                                                |
| GT14         | <i>Gcnt1</i>                     | 0                  | 0                                                  |
| GT14         | <i>Gcnt2</i>                     | 0                  | na                                                 |
| GT14         | <i>Gcnt3</i>                     | 0                  | 0                                                  |
| GT14         | <i>Gcnt4</i>                     | 0                  | 0                                                  |
| GT14         | <i>Gcnt6</i>                     | na                 | na                                                 |

| Familia CAZy | Genes GTf (hGTfs de la Tabla 1)* | CHO-GS_FPKM ** (4) | Profundidad de CHO-K1 ARN_seq (Xu et al 2011)**(4) |
|--------------|----------------------------------|--------------------|----------------------------------------------------|
| GT14         | <i>Gcnt7</i>                     | 0                  | na                                                 |
| GT14         | <i>Xylt1</i>                     | 0                  | 0                                                  |
| GT14         | <i>Xylt2</i>                     | 11                 | 6                                                  |
| GT16         | <i>Mgat2</i>                     | 29                 | 138                                                |
| GT17         | <i>Mgat3</i>                     | 0                  | 0                                                  |
| GT18         | <i>Mgat5</i>                     | 13                 | 20                                                 |
| GT18         | <i>Mgat5b</i>                    | 0                  | 0                                                  |
| GT21         | <i>Ugcg</i>                      | 40                 | na                                                 |
| GT22         | <i>Alg12</i>                     | 34                 | 69                                                 |
| GT22         | <i>Alg9</i>                      | 25                 | 98                                                 |
| GT22         | <i>Pigb</i>                      | 24                 | 75                                                 |
| GT22         | <i>Pigz</i>                      | 0                  | na                                                 |
| GT23         | <i>Fut8</i>                      | 18                 | 166                                                |
| GT24         | <i>Uggt1</i>                     | 27                 | 0                                                  |
| GT24         | <i>Uggt2</i>                     | 6                  | 95                                                 |
| GT25         | <i>Cercam</i>                    | 7                  | na                                                 |
| GT25         | <i>Glt25d1/Colgalt1</i>          | 129                | 0                                                  |
| GT25         | <i>Glt25d2/Colgalt2</i>          | 0                  | 0                                                  |
| GT27         | <i>Galnt1</i>                    | 39                 | 0                                                  |
| GT27         | <i>Galnt10</i>                   | 9                  | na                                                 |
| GT27         | <i>Galnt11</i>                   | 32                 | 63                                                 |
| GT27         | <i>Galnt12</i>                   | 0                  | 0                                                  |
| GT27         | <i>Galnt13</i>                   | 2                  | 0                                                  |
| GT27         | <i>Galnt14</i>                   | 0                  | 0                                                  |
| GT27         | <i>Galnt15/I2</i>                | 0                  | 0                                                  |
| GT27         | <i>Galnt16/I1</i>                | 0                  | 0                                                  |
| GT27         | <i>Galnt18</i>                   | 0                  | 0                                                  |
| GT27         | <i>Galnt19/I3(Wbscr17)</i>       | 0                  | 63                                                 |
| GT27         | <i>Galnt2</i>                    | 60                 | 324                                                |
| GT27         | <i>Galnt3</i>                    | 0                  | 0                                                  |
| GT27         | <i>Galnt4</i>                    | 3                  | na                                                 |
| GT27         | <i>Galnt5</i>                    | 0                  | 0                                                  |
| GT27         | <i>Galnt6</i>                    | 0                  | 0                                                  |
| GT27         | <i>Galnt7</i>                    | 30                 | 44                                                 |
| GT27         | <i>Galnt8</i>                    | 0                  | 0                                                  |
| GT27         | <i>Galnt9</i>                    | 0                  | 0                                                  |
| GT27         | <i>Galnt20/ I5</i>               | 0                  | 0                                                  |

| Familia CAZy | Genes GTf (hGTfs de la Tabla 1)* | CHO-GS_FPKM ** (4) | Profundidad de CHO-K1 ARN_seq (Xu et al 2011)**(4) |
|--------------|----------------------------------|--------------------|----------------------------------------------------|
| GT27         | <i>Galnt17/I6</i>                | 0                  | na                                                 |
| GT29         | <i>St3gal1</i>                   | 54                 | 195                                                |
| GT29         | <i>St3gal2</i>                   | 17                 | 28                                                 |
| GT29         | <i>St3gal3</i>                   | 13                 | 75                                                 |
| GT29         | <i>St3gal4</i>                   | 44                 | 33                                                 |
| GT29         | <i>St3gal5</i>                   | 21                 | 44                                                 |
| GT29         | <i>St3gal6</i>                   | 13                 | 29                                                 |
| GT29         | <i>St6gal1</i>                   | 0                  | 0                                                  |
| GT29         | <i>St6gal2</i>                   | 0                  | na                                                 |
| GT29         | <i>St6galnac1</i>                | 0                  | 0                                                  |
| GT29         | <i>St6galnac2</i>                | 0                  | 0                                                  |
| GT29         | <i>St6galnac3</i>                | 0                  | 0                                                  |
| GT29         | <i>St6galnac4</i>                | 24                 | 67                                                 |
| GT29         | <i>St6galnac5</i>                | 0                  | 0                                                  |
| GT29         | <i>St6galnac6</i>                | 18                 | 25                                                 |
| GT29         | <i>St8sia1</i>                   | 0                  | 0                                                  |
| GT29         | <i>St8sia2</i>                   | 0                  | 0                                                  |
| GT29         | <i>St8sia3</i>                   | 0                  | 0                                                  |
| GT29         | <i>St8sia4</i>                   | 3                  | 0                                                  |
| GT29         | <i>St8sia5</i>                   | 0                  | 0                                                  |
| GT29         | <i>St8sia6</i>                   | 0                  | 0                                                  |
| GT31         | <i>B3galnt1</i>                  | 12                 | 37                                                 |
| GT31         | <i>B3galnt2</i>                  | 30                 | 0                                                  |
| GT31         | <i>B3galt1</i>                   | 1                  | 0                                                  |
| GT31         | <i>B3galt2</i>                   | 0                  | 0                                                  |
| GT31         | <i>B3galt4</i>                   | 2                  | 16                                                 |
| GT31         | <i>B3galt5</i>                   | 0                  | 0                                                  |
| GT31         | <i>B3galt6</i>                   | 5                  | 42                                                 |
| GT31         | <i>B3galtl</i>                   | 19                 | na                                                 |
| GT31         | <i>B3gnt2</i>                    | 42                 | 189                                                |
| GT31         | <i>B3gnt3</i>                    | 0                  | 0                                                  |
| GT31         | <i>B3gnt4</i>                    | 0                  | 0                                                  |
| GT31         | <i>B3gnt5</i>                    | 0                  | 0                                                  |
| GT31         | <i>B3gnt6</i>                    | 0                  | 0                                                  |
| GT31         | <i>B3gnt7</i>                    | 0                  | 0                                                  |
| GT31         | <i>B3gnt9</i>                    | na                 | na                                                 |
| GT31         | <i>B3gntl1</i>                   | 11                 | 0                                                  |

| Familia CAZy | Genes GTf (hGTfs de la Tabla 1)* | CHO-GS_FPKM ** (4) | Profundidad de CHO-K1 ARN_seq (Xu et al 2011)**(4) |
|--------------|----------------------------------|--------------------|----------------------------------------------------|
| GT31         | <i>C1galt1</i>                   | 15                 | 26                                                 |
| GT31         | <i>C1galt1c1</i>                 | 26                 | 120                                                |
| GT31         | <i>Lfng</i>                      | 5                  | 24                                                 |
| GT31         | <i>Mfng</i>                      | 0                  | 0                                                  |
| GT31         | <i>Rfng</i>                      | 9                  | 158                                                |
| GT32         | <i>A4galt</i>                    | 0                  | 0                                                  |
| GT32         | <i>A4gnt</i>                     | 0                  | 0                                                  |
| GT33         | <i>Alg1</i>                      | 13                 | 41                                                 |
| GT33         | <i>Alg1l</i>                     | na                 | na                                                 |
| GT33         | <i>Alg112</i>                    | na                 | na                                                 |
| GT35         | <i>Pygb</i>                      | 61                 | na                                                 |
| GT35         | <i>Pygl</i>                      | na                 | na                                                 |
| GT35         | <i>Pygm</i>                      | 1                  | na                                                 |
| GT39         | <i>Pomt1</i>                     | 15                 | 0                                                  |
| GT39         | <i>Pomt2</i>                     | 9                  | 0                                                  |
| GT41         | <i>Ogt</i>                       | 14                 | 39                                                 |
| GT43         | <i>B3gat1</i>                    | 0                  | 0                                                  |
| GT43         | <i>B3gat2</i>                    | 0                  | 0                                                  |
| GT43         | <i>B3gat3</i>                    | 23                 | 31                                                 |
| GT47/64      | <i>Ext1</i>                      | 21                 | 83                                                 |
| GT47/64      | <i>Ext2</i>                      | 43                 | 136                                                |
| GT47/64      | <i>Extl1</i>                     | 9                  | 2                                                  |
| GT47/64      | <i>Extl3</i>                     | 24                 | 79                                                 |
| GT49         | <i>B3gnt1</i>                    | 32                 | 0                                                  |
| GT50         | <i>Pigm</i>                      | 6                  | 54                                                 |
| GT54         | <i>Mgat4a</i>                    | 0                  | 0                                                  |
| GT54         | <i>Mgat4b</i>                    | 37                 | 0                                                  |
| GT54         | <i>Mgat4c</i>                    | 0                  | na                                                 |
| GT57         | <i>Alg6</i>                      | 19                 | 22                                                 |
| GT57         | <i>Alg8</i>                      | 18                 | 22                                                 |
| GT58         | <i>Alg3</i>                      | 15                 | 38                                                 |
| GT59         | <i>Alg10</i>                     | na                 | 0                                                  |
| GT59         | <i>Alg10b</i>                    | na                 | 0                                                  |
| GT61         | <i>Eogt</i>                      | 6                  | 0                                                  |
| GT61         | <i>Pomgnt2</i>                   | 10                 | na                                                 |
| GT64         | <i>Extl2</i>                     | 31                 | 179                                                |
| GT65         | <i>Pofut1</i>                    | 44                 | 126                                                |

| Familia CAZy | Genes GTf (hGTfs de la Tabla 1)* | CHO-GS_FPKM ** (4) | Profundidad de CHO-K1 ARN_seq (Xu et al 2011)**(4) |
|--------------|----------------------------------|--------------------|----------------------------------------------------|
| GT66         | <i>Stt3a</i>                     | 71                 | 0                                                  |
| GT66         | <i>Stt3b</i>                     | 85                 | 0                                                  |
| GT68         | <i>Pofut2</i>                    | 68                 | 15                                                 |
| GT76         | <i>Pigv</i>                      | na                 | na                                                 |
| GT90         | <i>Kdelc1</i>                    | 26                 | na                                                 |
| GT90         | <i>Poglut1</i>                   | 22                 | na                                                 |
| -            | <i>Dpagt1</i> <sup>(3)</sup>     | 40                 | 59                                                 |
| GTnc         | <i>Dpy19l1</i>                   | 25                 | na                                                 |
| GTnc         | <i>Dpy19l2</i>                   | 0                  | na                                                 |
| GTnc         | <i>Dpy19l3</i>                   | 11                 | na                                                 |
| GTnc         | <i>Dpy19l4</i>                   | 5                  | na                                                 |
| GTnc         | <i>Fkrp</i>                      | na                 | na                                                 |
| GTnc         | <i>Fktn</i>                      | 16                 | na                                                 |
| GTnc         | <i>Plod3</i>                     | 61                 | na                                                 |
| GTnc         | <i>Tmem5</i>                     | 49                 | na                                                 |

\* un total de 208 hGTfs se reducen a 200 más *Dpagt1* (un gen no GTf) resultando en un total de 201 genes incluidos en la comparación de datos ARN\_seq\*\* los niveles de expresión no son directamente comparables con los datos CHO-GS son FRKP y para los datos CHO-K1 (Xu et al. 2011) la profundidad de lecturas se toma como una medida para la expresión de , na indica genes no analizados debido a anotaciones faltantes<sup>(1)</sup> los genes *Ugt1a* y *Ugt2a/b* son muy heterogéneos entre primates y roedores, y generalmente codifican el conjunto más grande y heterogéneo de genes de glucuroniltransferasa de los genomas de roedores<sup>(2)</sup> los *Ugt1a1-10* representan un locus con 9 genes, estos se han combinado en un gen en el análisis de datos ARN\_seq<sup>(3)</sup> *Dpagt1* no está incluido en la lista de GTfs humanos<sup>(4)</sup> na - no analizado

- Para explorar las funciones de genes de glicosiltransferasa individuales supuestamente involucrados en la N-glicosilación, la presente invención empleó una prueba de KO mediada por ZFN en una línea celular CHO-GS (Sigma), y también en otras líneas celulares. La presente invención diseñó una pantalla KO para diseccionar las funciones in vivo de todos los genes que codifican isoenzimas con potencial para regular la ramificación de N-glicosilación, la elongación y el taponamiento terminal expresados en CHO (Figura 14).
- La presente invención sondeó los efectos de la pantalla de knockout sobre la capacidad de N-glicosilación mediante la expresión recombinante de eritropoyetina humana (EPO) en líneas celulares mutantes CHO y el análisis de la glicosilación de EPO purificada mediante la liberación de N-glicanos y la elaboración de perfiles usando MALDI-TOF. La EPO se utilizó como indicador de la capacidad de N-glicosilación porque es una de las N-glicoproteínas mejor caracterizadas producidas en CHO, con tres N-glicanos con estructura principalmente tetraantenaria, bajo nivel de poli-LacNAc, y un recubrimiento de ácido siálico 2,3 (Figura 15) (Sasaki, Ochi et al. 1988). La EPO expresada en la línea de producción CHO-GS no modificada cultivada en suspensión en un medio libre de proteínas estaba glicosilada, esencialmente idéntica a los informes anteriores de EPO producida en CHO-K1 (Figura 15). Además, el perfil de glicosilación era esencialmente idéntico al de los productos terapéuticos de EPO producidos en CHO y reclamados en las solicitudes de medicamentos (Sasaki, Bothner et al. 1987).
- La presente invención utilizó ZFN, TALEN y CRISPR/Cas9 para apuntar y eliminar 19 genes de glicosiltransferasa (Figura 14A) involucrados en la ramificación de N-glicano (Mgat1/2/3/4A/4B/4C/5/5B) (FIG 5, FIG 21B), galactosilación (B4galt1/2/3/4 (Figura 17), elongación de poli-LacNAc (B3gnt1/2/8) (Figura 18), taponamiento terminal por sialilación (st3gal3/4/6) (Figura 19), y fucosilación del núcleo (fut8) (Figura 22).
- Estado de antena de N-glicano: primero se exploró el control biosintético del estado de antena de N-glicano. MGAT1 y 2 controlan cada uno la formación de una de las dos ramas  $\beta$ 2 en los N-glicanos biantenarios, mientras que se predice una posible redundancia funcional parcial para la formación de ramas tri- y tetraantenarias por MGAT4A/4B/4C (rama  $\beta$ 4) y MGAT5/5B (rama  $\beta$ 6), respectivamente. Se encontró que solo MGAT4B y 5 se expresaban sustancialmente en CHO según lo evaluado por los resultados de ARNsec (TABLA

3), sin embargo, sorprendentemente, apuntar al gen *Mgat4a* tuvo un efecto claro aunque menor, y solo apuntar a *Mgat4a/4b* eliminó por completo los N-glicanos tetraantenarios ramificados  $\beta 4$  (Figura 16A-B).

5 Por el contrario, la selección de *Mgat5* como diana eliminó el etiquetado de células con lectina L-PHA (Figura 20) y las estructuras tetraantenarias ramificadas  $\beta 6$  (FIG. 5), lo que concuerda con los estudios del mutante *Lec4* (Patnaik y Stanley 2006; North, Huang et al. 2010). Además, el apilamiento de KO de *Mgat4a/4b/5* produjo N-glicanos biantenarios casi homogéneos con una cantidad menor de poli-LacNAc (Figura 16B), que también se encuentra como un componente menor en células de tipo salvaje (WT) (North, Huang et al. 2010).

10 Control de LacNAc: las células CHO expresan las siete  $\beta 4$ galactosiltransferasas conocidas (TABLA 3), y nuestro conocimiento de las funciones in vivo de estas isoenzimas es deficiente. Sólo se espera que los *B4galt1-4* desempeñen funciones en la N-glicosilación y se ha sugerido que tienen preferencias por diferentes ramas de N-glicanos. La presente invención primero examinó el KO individual de *B4galt1-4* en CHO-GS WT y confirmó que *B4galt1* parecía tener el papel principal en la formación de LacNAc usando EPO como molécula  
15 reportera (Figura 17). Por tanto, la invención sólo podría evaluar las contribuciones de las otras isoformas en combinaciones apiladas con KO de *B4galt1*. El sondeo de las combinaciones apiladas mediante inmunocitología con un anticuerpo contra LacNAc mostró sorprendentemente que sólo la KO apilada de *B4galt1/3*, y no de *B4galt1/2/4*, parecía abolir la galactosilación (FIG 10). De acuerdo con esto, la invención encontró que la EPO expresada en clones KO apilados *B4galt1/3* pero no *B4galt1/2* tuvo una reducción  
20 sustancial (>90%) en la galactosilación (Figura 17).

La invención probó además la KO apilada de *B4galt* individuales en líneas celulares KO CHO-GS *Mgat4a/4b/5* con N-glicanos biantenarios homogéneos, y encontró que la KO de *B4galt1* eliminó la inmunorreactividad para LacNAc mientras que la KO de *B4galt2*, 3 y 4 no tuvo efectos sustanciales (Figura 21A-B). Esto sugirió que la  
25 isoforma de galactosiltransferasa codificada por *B4galt1* es la principal o única isoenzima capaz de transferir Gal a los N-glicanos biantenarios, mientras que las otras isoformas también funcionan con los N-glicanos tri- y tetraantenarios. Para probar esto más a fondo, la invención incluyó una IgG terapéutica humana como una molécula reportera expresada recombinante (Figura 22). Las IgG humanas suelen tener sólo un N-glicano en la región conservada en Asn297, y este sitio generalmente sólo está glicosilado con N-glicanos biantenarios  
30 con grados variables y heterogéneos de galactosilación y sialilación. Algunas moléculas de IgG tienen sitios de N-glicano adicionales en las regiones variables y estos suelen estar glicosilados con estructuras de N-glicanos tri y tetraantenarios más complejas con galactosilación eficiente y protección con ácidos siálicos. Es importante destacar que la expresión recombinante de una IgG en CHO-GS con KO de *B4galt1* resultó en N-glicanos biantenarios homogéneos sin galactosilación (Figura 22). Dado que CHO normalmente solo produce IgG con  
35 estructuras biantenarias, se deduce que solo se requiere la eliminación de *B4galt1* para lograr esto. Esto demuestra que la heterogeneidad inherente encontrada en la N-glicosilación de los anticuerpos monoclonales terapéuticos se puede eliminar mediante el uso de esta línea celular CHO modificada.

Dado que los CHO producen un alto grado de a6Fuc en los N-glicanos y que la eliminación de esta glicosilación  
40 puede lograrse mediante el knockout de *FUT8*, está claro que las combinaciones de knockout de *Mgat4a/4b/5*, *B4galt1* y *Fut8*, así como el KO apilado de *B4galt1* y *Fut8*, generarán clones CHO que producirán N-glicanos homogéneos sin a6Fuc. Estas líneas celulares de producción serán importantes para producir anticuerpos IgG terapéuticos con ADCC mejorado.

45 Control de poli-LacNAc - Se ha encontrado consistentemente que el poli-LacNAc en los N-glicanos es incompleto y heterogéneo (Fukuda, Sasaki et al. 1989; Takeuchi, Inoue et al. 1989; North, Huang et al. 2010). Las células CHO generalmente producen bajas cantidades de poli-LacNAc en N-glicanos y principalmente en estructuras tetraantenarias en la rama  $\beta 6$  controlada por *MGAT5* (North, Huang et al. 2010). La biosíntesis de poli-LacNAc en N-glicanos es poco conocida y tres genes, *B3gnt1*, *B3gnt2* y *B3gnt8*, son candidatos (Narimatsu  
50 2006). Utilizando KO de un solo gen de *B3gnt1*, *B3gnt2* y *B3gnt8* con inmunocitología de lectina LEL, la invención actual identificó *B3gnt2* como el gen clave que controla la iniciación de LacNAc en células CHO (Figura 23), y la EPO expresada en CHO con KO de *B3gnt2* pero no de *B3gnt1* carecía de poli-LacNAc (Figura 18). Además, la KO de *B3gnt2* en CHO con KO adicional de *Mgat4a/4b/5* resultó en EPO con N-glicanos biantenarios homogéneos sin poli-LacNAc (Figura 18). Fue sorprendente que *B3GNT1* no desempeñara un  
55 papel, ya que el gen se identificó originalmente mediante clonación de expresión como la poli-LacNAc sintasa (Sasaki, Kurata-Miura et al. 1997), aunque más tarde se demostró que este gen también funcionaba en la O-manosilación (Bao, Kobayashi et al. 2009). Sorprendentemente, la KO de *B3gnt8* no debilitó la inmunocitología de la lectina LEL. Se demostró previamente que *B3gnt8* forma heterodímero in vitro con *B3gnt2* y activa la actividad enzimática de *B3gnt2*.

60 Control de la sialilación: los CHO producen exclusivamente un recubrimiento NeuAc con enlaces  $\alpha 2,3$  de los N-glicanos (Watson, Bhide et al. 1994). Unos pocos informes han sugerido la presencia de cantidades menores de  $\alpha 2,6$ NeuAc así como de NeuGc, y los genes para la síntesis tanto del enlace  $\alpha 2,6$  como del CMP-NeuGc se encuentran en CHO pero éstos no se expresan normalmente (TABLA 3). Sin embargo, uno o ambos genes  
65 pueden ser inactivados por la presente invención y en informes anteriores (patente de EE. UU US20130004992). Las seis sialiltransferasas *st3gal1-6* conocidas se expresan en células CHO (TABLA 3) y las

funciones de cada una de ellas en la sialilación de N-glicanos son poco conocidas (Tsuji, Datta et al. 1996). Los genes de sialiltransferasa st3gal3, st3gal4 y st3gal6 se analizaron primero individualmente mediante inmunocitología para evaluar la exposición de LacNAc, y se descubrió que solo la KO de st3gal4 produjo una exposición sustancial de LacNAc (Figura 24). El análisis de EPO expresada en clones KO individuales y apilados mostró que la pérdida de st3gal4 y 6 produjo una exposición sustancial. En CHO WT, la KO de st3gal4/6 resultó en EPO con N-glicanos tetraantennarios heterogéneos sin sialilación (Figura 19).

Utilizando la presente invención se demostró que la KO de St3gal4/6 en combinación con Mgat4a/4b/5 produjo de manera interesante N-glicanos biantennarios sin sialilación pero con aumento en poli-LacNAc (Figura 25). Por lo tanto, CHO con st3gal4/6 KO puede producir EPO con extremos LacNAc sin protección y esto abre por primera vez la puerta a la ingeniería de novo de glicoproteínas recombinantes con protección a2,6NeuAc sin competencia por sialilación a2,3 endógena (El Mai, Donadio-Andrei et al. 2013).

N-glicanos monoantennarios: usando la presente invención también investigamos el efecto de mgat2 KO solo y en combinación con Mgat4a/4b/5 y st3gal4/6 (Figura 32). El KO de mgat2 solo produjo una mezcla de estructuras de N-glicanos mono y biantennarios, y en combinación posterior con Mgat4a/4b/5 N-glicanos monoantennarios homogéneos. Para confirmar que el N-glicano monoantennario estaba unido a la rama  $\beta$ 1-2  $\alpha$ 1-3Man controlada por mgat1, reintroducimos mediante knockin MGAT4A humano así como MGAT4A y 5 en combinación con ST6Gal-I, lo que dio como resultado N-glicanos tri y tetraantennarios, respectivamente (FIG 32).

#### Reconstrucción de la glicosilación homogénea

La presente invención proporciona además estrategias para desarrollar líneas celulares de mamíferos con capacidades de glicosilación definidas y/o más homogéneas que sirvan como plantilla para la ingeniería de novo de capacidades de glicosilación deseables mediante la introducción de una o más glicosiltransferasas usando inserciones génicas dirigidas al sitio y/o transfección aleatoria clásica de cADN y/o constructos genómicos. La estrategia implica la inactivación de los genes de glicosiltransferasa para obtener una capacidad de glicosilación homogénea en un tipo particular deseable de glicosilación, por ejemplo, usando la matriz de deconstrucción desarrollada en este documento para la N-glicosilación, y por ejemplo, pero sin limitarse a, la inactivación de los genes de sialiltransferasa St3gal4 y 6 en una célula y la obtención de una célula sin protección de ácido siálico de los N-glicanos. En dicha célula, la introducción de novo de una o más nuevas capacidades de glicosilación que usan el producto de glicano truncado más homogéneo obtenido por uno o más eventos de inactivación del gen de glicosiltransferasa, proporcionará una glicosilación no competitiva y una glicosilación más homogénea por las glicosiltransferasas introducidas de novo. Por ejemplo, pero no limitado a, la introducción de una a2,6sialiltransferasa como ST6GAL1 en una célula de mamífero con genes de sialiltransferasa St3gal4 y St3gal6 inactivados. El principio general de la estrategia es simplificar la glicosilación de una vía particular, por ejemplo la N-glicosilación, hasta un punto en el que se produzcan estructuras de glicanos razonablemente homogéneas en la célula de mamífero en la que se han introducido uno o más eventos de inactivación del gen de la glicosiltransferasa en un proceso de deconstrucción como el proporcionado en la presente invención para la N-glicosilación. Tomar dicha célula de mamífero con capacidad de glicosilación deconstruida y simplificada e introducir de novo capacidades de glicosilación deseables que se basen en las estructuras de glicano producidas por la deconstrucción.

#### Reconstrucción de la sialilación de N-glicano mediante knockin dirigido: diseño de novo de la protección de a2,6NeuAc

La mayoría de las glicoproteínas en circulación en el hombre están protegidas con a2,6NeuAc (El Mai, Donadio-Andrei et al. 2013), mientras que esencialmente todos los productos terapéuticos recombinantes se producen con a2,3NeuAc en CHO, HEK293 y otras líneas celulares de producción actuales. Por lo tanto, es claramente deseable obtener líneas celulares de mamíferos como CHO con capacidad de producción de glicoproteínas más parecidas a las humanas con un recubrimiento homogéneo de a2,6NeuAc. Varios grupos de investigación han abordado este problema e intentaron producir células CHO estables capaces de producir N-glicanos homogéneos con protección 2,6NeuAc sin éxito (El Mai, Donadio-Andrei et al. 2013).

Para demostrar las perspectivas de reconstrucción y glicoingeniería de novo posibilitadas por nuestra pantalla de deconstrucción KO, se generaron clones CHO capaces de producir una protección homogénea de a2,6NeuAc (FIG 25). Para evitar problemas pasados con la integración aleatoria de plásmidos (El Mai, Donadio-Andrei et al. 2013), se utilizó ZFN dirigido a knockin (KI) de ST6GAL1 con una estrategia ObLiGaRe modificada en un sitio de integración SafeHarbor (Maresca, Lin et al. 2013). Esta estrategia tampoco implica ni necesita la selección o el mantenimiento de antibióticos para la clonación y la estabilidad, lo cual es importante para el potencial de producción de terapias recombinantes para uso clínico. La ST6GAL1 humana introducida en CHO con KO de St3gal4/6 produjo la gama completa de estructuras antenarias de N-glicanos con un grado normal de recubrimiento NeuAc (FIG 25), y cuando se introdujo en CHO con KO adicional de Mgat4a/4b/5, se produjo EPO con N-glicanos biantennarios homogéneos recubiertos por a2,6NeuAc (FIG 25). Que la sialilación

introducida era de hecho NeuAc  $\alpha$ 2,6-enlazado y no  $\alpha$ 2,3-enlazado se confirmó mediante ESI-MS/MS de iones negativos como se describe en la FIGURA 15.

- 5 Sorprendentemente, la introducción de novo de ST6Gal-I anuló las cantidades menores de poli-LacNAc formadas en estructuras biantenarias en CHO con KO de Mgat4a/4b/5 (FIG 25). La eliminación de B3gnt2 fue necesaria para eliminar poli-LacNAc en CHO WT y en CHO con KO de Mgat4a/4b/5 (FIG 18). Este ejemplo demuestra claramente la intrincada interacción entre las glicosiltransferasas que usan el mismo sustrato aceptor y cómo las diferentes reacciones de glicosilación pueden competir entre sí de formas impredecibles. Así, es posible introducir la capacidad de sialilación de  $\alpha$ 2,6NeuAc sin  $\alpha$ 2,3sialilación competitiva, y reconstruir una célula CHO capaz de producir N-glicoproteínas sialiladas con recubrimiento homogéneo de  $\alpha$ 2,6NeuAc ya sea con el repertorio normal de ramificación de N-glicanos o con N-glicanos biantenarios homogéneos.

Construcción de N-glicanos monoantennarios adecuados para la glicomodificación enzimática homogénea

- 15 La glicopegilación enzimática y otras modificaciones de proteínas es un método bien establecido para la unión de sitios específicos de cadenas de polietilenglicol (PEG) u otros compuestos para prolongar la vida media de los fármacos (DeFrees, Wang et al. 2006). La glicomodificación específica del sitio de N-glicanos usando una  $\alpha$ 2,3-sialiltransferasa para introducir la modificación a través de un sustrato CMP-NeuAc modificado se basa en i) una glicoproteína expresada recombinante purificada con uno o más N-glicanos; ii) la eliminación de los ácidos siálicos que cubren los N-glicanos en la glicoproteína usando una neuraminidasa recombinante; y iii) la transferencia del ácido siálico modificado a los residuos de galactosa expuestos en los N-glicanos usando una sialiltransferasa recombinante y el sustrato donante modificado. Los N-glicanos de las glicoproteínas recombinantes tienen distintos grados de estructuras antenales bi, tri y tetra y la desialilación de las glicoproteínas recombinantes expone múltiples residuos de galactosa de cada estructura antenal, lo que da como resultado una heterogeneidad sustancial en el proceso de glicomodificación enzimática. Para superar el paso de desialilación, produjimos una serie de diseños de ingeniería de CHO que producen N-glicanos sin protección con ácido siálico (FIG. 19, 25 y 32), y estos permitirían la glicomodificación enzimática directa sin pretratamiento con neuraminidasa para eliminar los ácidos siálicos. Para superar la heterogeneidad inducida por los múltiples sitios aceptores de galactosa expuestos para la transferencia enzimática en cada N-glicano multiantennario apuntamos a mgat2 para bloquear la formación de estructuras biantenarias, pero el KO de mgat2 solo no generó estructuras monoantennarias homogéneas como se esperaba originalmente (FIG 32). En cambio, aproximadamente el 50% seguía siendo biantennario, pero combinado con el KO de mgat4A/4B/5 se obtuvieron estructuras monoantennarias bastante homogéneas. La única heterogeneidad detectable estuvo asociada con poli-LacNAc, que no proporciona exposición de múltiples residuos de galactosa (FIG 18).

- 35 Una estructura monoantennaria bastante homogénea también puede lograrse mediante los siguientes KO dobles: KO de Mgat2/4B o KO de Mgat2/5. (Figura 35).

- 40 Para el anticuerpo IgG se puede obtener una estructura mono antenal homogénea y un mayor grado de sialilación por KO de Mgat2 (Figura 36).

Eliminación de vías de glicosilación innecesarias

- 45 Los métodos de la invención permitieron investigar si la inactivación de diferentes vías de O-glicosilación afectaría la viabilidad celular y las eficiencias de glicosilación de vías de glicosilación no dirigidas. Primero, anulamos la vía de elongación de Core1 de la glicosilación de O-GalNAc como se describió previamente (Yang, Halim et al. 2014) mediante la inactivación de la chaperona Cosmc para la sintasa core1, C1Gal-T1, que conduce a células CHO capaces de producir EPO con glicosilación de O-GalNAc truncada limitada a un único residuo de GalNAc sin sialilación en el único O-glicosito. La CHO WT produce O-glicanos con cubierta de ácido siálico  $\alpha$ 2,3 y  $\alpha$ 2,6, y la glicosilación O-GalNAc es uno de los tipos más abundantes de glicosilación de proteínas en una célula y, por tanto, exigente en cuanto a la reserva de nucleótidos de azúcar. En consonancia con esto, encontramos que la EPO producida en CHO con inactivación de Cosmc y elongación de O-glicano, produjo EPO con un recubrimiento de ácido siálico mejorado de los N-glicanos en los tres N-glicositos presentes en la EPO (FIG 29).

- 55 Aplicando los métodos de la presente invención, la vía de elongación de la glicosilación de O-Man fue dirigida por la inactivación de Pomgnt1, que codifica la  $\beta$ 2GlcNAc-transferasa que sintetiza la principal vía de elongación de los glicanos de O-Man (ref. PNAS). Las proteínas como el distroglicano expresadas en células de mamíferos, incluido CHO (Yang, Halim et al. 2014), tienen una densa cobertura del tetrasacárido O-Man (NeuAc $\alpha$ 2,3Gal $\beta$ 1,4GlcNAc $\beta$ 1,2Man $\alpha$ ). Cuando se expresó EPO en una célula CHO con inactivación de Pomgnt1, se mejoró el grado de protección con galactosa, poliLacNac y ácido siálico (FIG 30).

- 65 La presente invención se aplicó para inactivar la vía de elongación de la glicosilación de O-Xyl por KO de B4galt7, que codifica la  $\beta$ 4Galactosiltransferasa que sintetiza el segundo paso en la región de enlace requerida para la biosíntesis de proteoglicanos (Almeida, Levery et al. 1999). Las células CHO producen proteoglicanos y tienen capacidades eficientes para producirlos, por ejemplo, a partir de sacáridos suministrados

exógenamente. Cuando la EPO se expresó en una célula CHO con inactivación de B4galt7, aumentó el grado de recubrimiento de galactosa, poli-LacNac y ácido siálico (FIG. 30).

- La invención actual también se utilizó para inactivar combinaciones de dos o tres genes que controlan las vías de glicosilación de O-GalNAc, O-Man y O-Xyl. Específicamente, se generaron células CHO con inactivación de B4galt7 y Pomgnt1, y con inactivación de B4galt7, Pomgnt1 y Cosmc. Cuando se expresó EPO en dichas líneas de células CHO con inactivación de dos o tres tipos de vías de O-glicosilación, se mejoró el grado de protección con galactosa y ácido siálico.
- Estos resultados demuestran claramente que la capacidad de glicosilación más homogénea de proteínas expresadas recombinantemente en células de mamíferos como CHO puede mejorarse mediante la inactivación de vías de glicosilación innecesarias que usan las reservas comunes de donantes de nucleótidos de azúcar y otras partes de las capacidades metabólicas y de glicosilación generales de las células. Los expertos en la materia saben que la inactivación de otros genes enumerados en las TABLAS 1 y 2 en las vías de glicosilación anteriores u otras vías de glicosilación tendrá efectos similares y pueden ser eventos de ingeniería deseables solos o en combinación con cualquiera de las ingenierías genéticas realizadas en este documento o concebibles a partir de las vías de glicosilación descritas en la FIG. 13.

- El principio general para la selección de estrategias de inactivación óptimas con el objetivo de dirigir el uso de nucleótidos de azúcar en una célula huésped de mamífero para vías de glicosilación que son relevantes para la producción recombinante de una glicoproteína dada, es bloquear vías de glicosilación no relevantes en un paso temprano en la vía biosintética donde una única inactivación génica da como resultado estructuras de glicano truncadas preferiblemente sin elongación y recubrimiento por ácido siálico, y/o galactosa, y/o GlcNAc. Se proporcionan ejemplos para la inactivación de tres tipos distintos de O-glicosilación apuntando al segundo paso en su biosíntesis para evitar y/o limitar el uso de CMP-NeuAc, UDP-Gal y UDP-GlcNAc para la elongación de estas vías. Es evidente que muchos otros genes enumerados en la Tabla 2 con pasos biosintéticos asignados para las diferentes vías de glicosilación en la célula CHO de mamíferos, pueden ser objeto de inactivación para lograr los mismos efectos o efectos similares en el uso de nucleótidos de azúcar donantes teniendo en cuenta los efectos que la inactivación tiene en la vía de glicosilación y las estructuras de glicanos alteradas.

#### Eliminación de glicosiltransferasas innecesarias

- Las glicosiltransferasas residentes del RE-Golgi se mantienen para funcionar en la vía secretora mediante una variedad de funciones parcialmente conocidas que incluyen el reconocimiento de parentesco, el tipo y la longitud de la región transmembrana, las interacciones a través de la región del tallo (Colley 1997) y a través del transporte retrógrado de vesículas COP-I presumiblemente dependiente de las colas citosólicas y las TM de las proteínas glicosiltransferasas, por ejemplo, GOLPH3 (Eckert, Reckmann et al. 2014). Cuando las glicosiltransferasas no se retienen en su posición natural en la vía secretora, pueden funcionar de manera diferente y adversa en la glicosilación, y cuando las glicosiltransferasas pierden su retención, por ejemplo, por escisión proteolítica en la región del tallo, se secretan y liberan con poca o ninguna influencia en los procesos de glicosilación en las células. Teniendo en cuenta las numerosas glicosiltransferasas expresadas en cualquier célula de mamífero, y por ejemplo al menos más de 60 en las células CHO (TABLA 3), planteamos la hipótesis de que habría mecanismos comunes de retención de grupos y/o clases de enzimas. Se podría constituir un grupo o clase de enzimas sin intención de limitaciones, por ejemplo, enzimas que trabajan en la misma vía y/o en pasos biosintéticos consecutivos, y/o enzimas retenidas en la misma topología subcelular, y/o enzimas que tienen una secuencia de aminoácidos y/o señales de retención estructural similares. Estos últimos podrían, sin intención de limitaciones, encontrarse, por ejemplo, en isoenzimas homólogas dado que tienen secuencias similares, propiedades catalíticas relacionadas y, al menos en algunos casos, se ha demostrado que tienen topologías subcelulares similares (Rottger, White et al. 1998). Un grupo de este tipo de isoenzimas son las  $\beta$ 4galactosiltransferasas, de las cuales especialmente B4GAL-T1, T2, T3 y T4 tienen el mayor grado de similitud de secuencia y funciones relacionadas (Amado, Almeida et al. 1999; Togayachi, Sato et al. 2006). La comprensión de las funciones in vivo de estas cuatro enzimas es deficiente y se ha demostrado que todas producen el enlace Gal $\beta$ 1-4GlcNAc in vitro. Como se muestra aquí,  $\beta$ 4Galt1 y 3 parecen ser las principales  $\beta$ 4galactosiltransferasas involucradas en la N-glicosilación, y estudios previos han demostrado que B4GALT4 tiene funciones principalmente en la biosíntesis de glicolípidos y/o en glicanos sulfatados (Schwientek, Almeida et al. 1998). Por lo tanto, generamos células CHO con inactivación de B4galt4 en CHO WT, así como también combinadas con inactivación de mgat4a/4b/5. La expresión de EPO en células CHO con inactivación de B4galt4 sola resultó en un aumento de N-glicanos tetraantenarios con un recubrimiento de ácido siálico más completo (FIG. 17). La KO de B3gnt2 dio lugar a glicoestructuras de EPO sin poli LacNac (Fig 35) mientras que la KO de B4galt5/6 dio más poliLacNac (Fig 35). En esta línea, la IgG producida en células CHO en las que se ha eliminado St3gal4/6 presenta una galactosilación más alta (Fig. 36). Estos resultados demuestran claramente que la capacidad para una glicosilación más homogénea de proteínas expresadas recombinantemente en células de mamíferos como CHO puede mejorarse mediante la inactivación de glicosiltransferasas innecesarias que pueden competir por mecanismos que dirigen la residencia RE-Golgi y/o

usar las reservas comunes de donantes de nucleótidos de azúcar y otras partes de las capacidades metabólicas y de glicosilación generales en una célula de mamífero.

- Estos resultados demuestran claramente que la capacidad de glicosilación más homogénea de proteínas expresadas recombinantemente en células de mamíferos como CHO puede mejorarse mediante la inactivación de vías de glicosilación innecesarias que usan las reservas comunes de donantes de nucleótidos de azúcar y otras partes de las capacidades metabólicas y de glicosilación generales de las células. Los expertos en la materia saben que la inactivación de otros genes enumerados en las TABLAS 1 y 2 en las vías de glicosilación anteriores u otras vías de glicosilación tendrá efectos similares y pueden ser eventos de ingeniería deseables solos o en combinación con cualquiera de las ingenierías genéticas realizadas en este documento o concebibles a partir de las vías de glicosilación esbozadas en la FIG 13.

#### Eliminación de la capacidad de N-glicosilación postraducciona

- El complejo oligosacaryltransferasa que inicia la N-glicosilación de proteínas consta de ocho proteínas, de las cuales una subunidad involucra dos genes parálogos, STT3A y STT3B. Los estudios han sugerido que un complejo de oligosacaryltransferasa que contiene la subunidad STT3A funciona principalmente en la N-glicosilación cotraducciona, mientras que un complejo de oligosacaryltransferasa que contiene la subunidad STT3B funciona principalmente en la N-glicosilación postraducciona. Planteamos la hipótesis de que la fidelidad y la estequiometría de la N-glicosilación cotraducciona son naturalmente mejores que las de la N-glicosilación postraducciona. Muchas glicoproteínas expresadas recombinantemente tienen N-glicanos mal usados, lo que resulta en una ocupación incompleta del N-glicano (estequiometría) y/o en el procesamiento del N-glicano (estructura). Por lo tanto, a continuación generamos células CHO con inactivación de Stt3b así como células CHO con inactivación combinada de los genes Mgat4A/4B/5. Estas células CHO tienen una capacidad mejorada para producir glicoproteínas homogéneas con un mayor grado de estequiometría de N-glicano y/o estructuras de glicanos cuando se expresan proteínas con N-glicositos naturalmente poco usados.

#### Estrategia de eliminación selectiva

- Es claro para los expertos en la técnica que la inactivación de un gen de glicosiltransferasa puede tener una multitud de resultados y efectos sobre la transcripción y/o el producto proteico traducido a partir de este. Los experimentos de inactivación dirigida realizados aquí implicaron PCR y secuenciación de las alteraciones introducidas en los genes, así como análisis de ARNsec de clones para determinar si se formó una transcripción y si posibles variaciones de empalme novedosas introdujeron nuevas estructuras de proteínas. Además, existen métodos para determinar la presencia de proteínas a partir de dichas transcripciones, que incluyen espectrometría de masas y análisis Western blot SDS-PAGE con anticuerpos relevantes que detectan la región más N-terminal de los productos proteicos.

- Los constructos de selección se diseñaron generalmente para seleccionar el primer 1/3 del marco de lectura abierto (ORF) de las regiones codificantes, pero también se seleccionaron otras regiones FIG 14B. Para la mayoría de los clones seleccionados, se seleccionaron mutaciones fuera del marco (Indels) que introdujeron codones de terminación o sin sentido con el mismo exón. Se esperaba que esto produjera proteínas truncadas, si es que hay alguna, y sin el dominio catalítico y, por lo tanto, sin actividad enzimática. La mayoría de los clones KO presentaban inserciones y/o deleciones (indels) en el rango de  $\pm 20$ bps, y la mayoría de los genes diana estaban presentes con dos alelos, mientras que algunos (mgat4B y mgat5) estaban presentes con 1 o 3 alelos, respectivamente (TABLA 4). En algunos casos se encontraron deleciones más grandes que alteraron uno o más exones y los límites exón/intrón, lo que también dio como resultado proteínas truncadas.

- La mayoría de las glicosiltransferasas del RE-Golgi comparten la estructura transmembrana de tipo 2 común con una cola citosólica corta que puede dirigir el tráfico retrógrado y el tiempo de residencia, un péptido señal no escindido que contiene un dominio de hélice  $\alpha$  transmembrana hidrofóbico para la retención en la membrana del RE-Golgi, una región de tallo o vástago de longitud variable que se cree que desplaza el dominio catalítico hacia el lumen del RE-Golgi y un dominio catalítico C-terminal necesario para la función enzimática (Colley 1997). Solo las GalNAc-transferasas polipeptídicas tienen un dominio lectina C-terminal adicional (Bennett, Mandel et al. 2012). La organización genómica de los genes de la glicosiltransferasa varía sustancialmente: algunos genes tienen un solo exón codificante y otros más de 10-15 exones codificantes, aunque pocos genes de la glicosiltransferasa producen diferentes variantes de empalme que codifican diferentes productos proteicos.

- La inactivación de los genes de glicosiltransferasa en las regiones de codificación temprana puede tener una multitud de efectos sobre los productos de transcripción y proteína si estos se realizan: i) una o más transcripciones pueden ser inestables y degradarse rápidamente dando como resultado poca o ninguna transcripción y/o proteína; ii) una o más transcripciones pueden ser estables pero no traducidas o solo mal traducidas dando como resultado poca o ninguna traducción de proteína; iii) una o más transcripciones pueden ser estables y traducidas dando como resultado la traducción de proteína; iv) una o más transcripciones pueden dar como resultado la traducción de proteína pero los productos de proteína se degradan debido a, por ejemplo,

truncamientos y/o plegamiento incorrecto; v) una o más transcripciones pueden dar como resultado la traducción de proteína y productos de proteína estables que están truncados y enzimáticamente inactivos; y vi) una o más transcripciones pueden dar como resultado la traducción de proteína y productos de proteína estables que están truncados pero tienen actividad enzimática.

Es evidente para los expertos en la materia que la inactivación de genes que conduce a productos proteicos con actividad enzimática es indeseable, y estos eventos se detectan fácilmente mediante los métodos usados en la presente invención, por ejemplo, pero sin limitarse a ellos, el marcado de lectina/anticuerpo y el perfil glicolítico de proteínas expresadas en líneas celulares mutantes.

Sin embargo, es deseable eliminar los posibles productos proteicos truncados que pueden expresarse a partir de transcripciones mutadas, ya que estos pueden tener efectos indeseables sobre la capacidad de glicosilación. Por lo tanto, los productos proteicos truncados de los genes de la glicosiltransferasa tipo 2 que contienen parte o parte entera del dominio citosólico y/o la señal de retención transmembrana y/o la región madre y/o parte del dominio catalítico pueden ejercer una serie de efectos sobre la glicosilación en las células. Por ejemplo, la cola citosólica puede competir por el tráfico retrógrado COP-I de las proteínas residentes en el Golgi (Eckert, Reckmann et al. 2014), el dominio transmembrana puede competir por la localización en el RE-Golgi y posibles asociaciones y/o agregaciones normales de proteínas, y la región del tallo, así como parte de un dominio catalítico inactivo, pueden tener funciones similares o parte de funciones en asociaciones y/o agregaciones normales de proteínas en el RE-Golgi. Estas funciones y otras desconocidas pueden afectar a vías específicas de glicosilación en las que participan las enzimas, a funciones específicas de las isoenzimas o, de forma más general, a las capacidades de glicosilación de una célula. Si bien estas funciones y efectos son desconocidos e impredecibles hoy en día, es una parte inherente de la presente invención que la selección de clones de células de mamíferos con genes de glicosiltransferasa inactivados incluye la selección de eventos de edición que no producen productos proteicos truncados.

Es evidente que la pantalla de deconstrucción de KO realizada identifica los glucógenos clave que controlan los pasos decisivos en la N-glicosilación de proteínas en CHO y proporciona una matriz de diseño para la deconstrucción genética de la capacidad de N-glicosilación a estructuras homogéneas deseables que pueden servir como puntos de partida para la reconstrucción de capacidades de glicosilación mejoradas, novedosas y/o más homogéneas mediante la introducción estable de uno o más genes de glicosiltransferasa. La FIGURA 28 proporciona ejemplos generales de andamios de N-glicano que pueden producirse mediante la matriz de diseño de deconstrucción, que sirve como plataformas para la reconstrucción de capacidades de glicosilación más homogéneas deseables. El valor de esta estrategia de deconstrucción y reconstrucción combinadas queda claramente ejemplificado aquí por el diseño de células CHO con capacidad de producción de EPO con N-glicanos biantenarios homogéneos con  $\alpha 2,3$  y, por primera vez, N-glicanos homogéneos tapados con ácido  $\alpha 2,6$ -siálico. Estas líneas CHO complementan y amplían claramente los esfuerzos realizados previamente para establecer una levadura glicoingenierizada con capacidad para la producción de EPO con N-glicanos biantenarios homogéneos, así como la síntesis química de gran éxito de EPO con N-glicanos biantenarios (Hamilton, Davidson et al. 2006; Wang, Dong et al. 2013). Más específicamente, las líneas CHO producidas permiten la producción de EPO y otras glicoproteínas en una célula huésped de mamífero probada con capacidades de plegamiento y procesamiento de proteínas bien establecidas y un perfil de seguridad de décadas de duración para terapias humanas.

Es además evidente a partir de la pantalla de deconstrucción de KO realizada aquí que la célula CHO tiene una notable plasticidad y tolerancia a la glicoingeniería. Las capacidades de glicosilación logradas por las líneas KO en CHO resultaron ser estables y consistentes en la producción de glicoproteínas recombinantes, y se espera que estas líneas CHO de ingeniería CHO tengan un rendimiento similar a las CHO WT en el bioprocesamiento. Además, la reconstrucción de las capacidades de glicosilación por KI ejemplificada por la introducción de ST6Gal-1 sin  $\alpha 2,3$ -sialiltransferasas competidoras demuestra que se puede lograr una glicoingeniería estable y consistente mediante la estrategia de deconstrucción y reconstrucción inventada aquí, y esto debería ser aplicable a otras capacidades de glicosilación deseables. Esto incluye la reconstrucción de las capacidades de glicosilación endógena de CHO WT, que son inconsistentes y heterogéneas, como por ejemplo la formación de ramas multiantenarias de N-glicano (tri y tetraantenarias) y cadenas de poli-LacNAc. Otros ejemplos incluyen la mejora de la capacidad de proteger los N-glicanos de la IgG. Además, se espera que la matriz de diseño de ingeniería sea transferible a cualquier línea anfitriona de CHO, así como a las líneas de producción establecidas. Además, se espera que la matriz de diseño de ingeniería sea transferible a cualquier línea celular de mamíferos con la debida expansión dependiendo de la existencia de capacidades de glicosilación más complejas. El panel de células CHO aquí generado permitirá la disección y el diseño de reconstrucción de novo de la N-glicosilación para cualquier terapéutica proteica. Además, este panel de células CHO puede usarse para producir glicoproteínas terapéuticas con una serie de glicoconformidades diferentes, lo que por primera vez permitirá evaluar experimentalmente las propiedades biológicas de glicoconformidades definidas y hacer una selección racional de los diseños óptimos. El panel de células CHO producido aquí permite la prueba directa de los efectos funcionales de estructuras específicas de N-glicano en cualquier glicoproteína terapéutica, lo que proporcionará orientación para el diseño óptimo de un fármaco de glicoproteína. De este modo, usando el panel de células CHO, se puede abordar de manera sistemática el

- efecto de la ramificación de N-glicano, la extensión de poli-LacNAc y el tipo de sialilación. Estos efectos incluyen, por ejemplo, la secreción de proteínas, la estabilidad, el rendimiento, la semivida circulatoria en sangre, la biodistribución, la bioactividad y las propiedades farmacocinéticas generales de relevancia para los fármacos terapéuticos. El panel CHO también permitirá, por ejemplo, el diseño de terapias de glicoproteínas de novo, donde la introducción de N-glicanos se usa para mejorar la vida media circulatoria sin interferir con la actividad biológica. En este caso, el efecto de la posición, el tamaño y la estructura de los N-glicanos introducidos puede cartografiarse en detalle para diseñar nuevos productos biológicos de glicoproteínas con propiedades farmacocinéticas óptimas. La presente invención proporciona así una nueva era para el mayor productor de terapias con glicoproteínas: la célula CHO.
- En una realización adicional de la presente invención, se han eliminado mgat4A, mgat4B, mgat5 y FUT8.
- En una realización de la presente invención, la célula o línea celular es una célula de mamífero.
- En otra realización de la presente invención, la célula se deriva del ovario de hámster chino o del riñón humano.
- En otra realización más de la presente invención, la célula se selecciona del grupo que consiste en CHO, NS0, SP2/0, YB2/0, CHO-K1, CHO-DXB11, CHO-DG44, CHO-S, HEK293, HUVEC, HKB, PER-C6, NS0, o derivados de cualquiera de estas células.
- En una realización de la presente invención es la célula una célula CHO.
- En otra realización de la presente invención codifica la célula además una proteína exógena de interés.
- En otra realización de la presente invención, la proteína de interés es un anticuerpo, un fragmento de anticuerpo o un polipéptido.
- En otra realización de la presente invención, el anticuerpo es un anticuerpo IgG.
- En una realización de la presente invención es el polipéptido EPO, una proteína implicada en la hemostasia, incluyendo un factor de coagulación.
- En una realización de la presente invención la glicosilación se hace más homogénea.
- En otra realización de la presente invención la glicosilación es no sialilada.
- En otra realización de la presente invención, la glicosilación no es galactosilada.
- En una realización de la presente invención comprende la glicosilación N-glicanos biantenarios.
- En otra realización de la presente invención, la glicosilación no comprende poli-LacNAc.
- En otra realización de la presente invención es la glicosilación cualquier combinación sin fucosa.
- En una realización de la presente invención se ha eliminado B4galt1 permitiendo la generación de N-glicanos más homogéneos sin galactosa.
- En otra realización de la presente invención se han eliminado B4galt1 y fut8 permitiendo la generación de N-glicanos más homogéneos sin galactosa y fucosa.
- En otra realización más de la presente invención, se han eliminado B4galt1 y B4galt3, lo que da como resultado N-glicanos más homogéneos sin galactosa que los obtenidos con una única eliminación de B4galt1 o B4galt3 (Figura 36).
- En una realización de la presente invención, la célula tiene knock out de uno o más de los genes de glicosiltransferasa seleccionados del grupo que consiste en mgat1, mgat2, mgat4A, mgat4B, mgat4C, mgat5 y mgat5B.
- En otra realización de la presente invención, la célula tiene knockout de uno o más de los genes glicosiltransferasa seleccionados del grupo que consiste en mgat4A, mgat4B, y mgat4C.
- En una realización adicional de la presente invención, la célula tiene knock out de mgat4A, mgat4B y mgat4C.
- En otra realización de la presente invención, la célula comprende la eliminación de mgat5 y/o mgat5B.
- En otra realización de la presente invención, la célula comprende knockout de mgat5 y mgat5B.

En una realización de la presente invención, la célula comprende la eliminación de uno o más de los genes de galatosilación seleccionados del grupo que consiste en B4gal1, B4gal2, B4gal3 y B4gal.

- 5 En otra realización de la presente invención, también se ha eliminado B4gal1 o B4gal3, o B4gal1 y B4gal3.

En una realización adicional de la presente invención, la célula también tiene knockout de uno o más de los genes de elongación de poli-LacNAc seleccionados del grupo que consiste en B3gnt1, B3gnt2 y B3gnt8.

- 10 En otra realización más de la presente invención, también se ha eliminado B3gnt2.

En una realización de la presente invención, también se han eliminado B3gnt2, mgat4A, mgat4B y mgat5.

- 15 En otra realización de la presente invención, la célula también tiene knockout de uno o más de los genes de sialiltransferasa seleccionados del grupo que consiste en ST3gal3, ST3gal4 y ST3gal6.

En otra realización más de la presente invención, también se han eliminado ST3gal3, ST3gal4 y ST3gal6.

- 20 En una realización adicional de la presente invención, también se han eliminado ST3gal4 y ST3gal6.

En una realización de la presente invención, también se han eliminado ST3gal4, ST3gal6, mgat4A, mgat4B y mgat5.

- 25 En otra realización de la presente invención, también se ha eliminado ST6gal1 y se han eliminado ST3gal4 y ST3gal6.

En otra realización más de la presente invención, se ha introducido ST6gal1 y se han eliminado ST3gal4, ST3gal6, mgat4A, mgat4B y mgat5.

- 30 En una realización de la presente invención, se han eliminado mgat2, ST3gal4 y ST3gal6.

En otra realización de la presente invención, mgat2, mgat4A, mgat4B y mgat5 han sido eliminados.

- 35 En otra realización de la presente invención, mgat2, ST3gal4, ST3gal6, mgat4A, mgat4B y mgat5 han sido eliminados.

Un aspecto de la presente invención se refiere a una glicoproteína según la presente invención, que es un glicoconjugado homogéneo producido a partir de una glicoproteína que tiene un perfil de glicanos simplificado.

- 40 Un objeto de la presente invención es proporcionar una célula capaz de expresar un gen que codifica un polipéptido de interés, en donde el polipéptido de interés se expresa comprendiendo uno o más patrones de modificación postraduccional

En una realización de la presente invención, el patrón de modificación postraduccional es una glicosilación.

- 45 La glicoforma óptima de una glicoproteína puede identificarse mediante el siguiente proceso: (i) producción de una pluralidad de glicoformas diferentes de dicha glicoproteína mediante expresión en líneas celulares que alberguen al menos una capacidad de glicosilación novedosa, por ejemplo, albergando dos o más modificaciones de los niveles de expresión del gen GT, y (ii) determinación de la actividad de las diferentes glicoformas en comparación con una glicoproteína de referencia en (un) bioensayo(s) adecuado(s); y (iii) selección de la glicoforma con la actividad mayor o más elevada y determinación de la huella dactilar del genotipo de la célula de producción que esté correlacionada con el nivel de actividad mayor o más elevado de dicha glicoproteína.

- 55 El proceso descrito anteriormente permite la identificación de la glicoforma óptima de una glicoproteína. Para los expertos en la materia, mediante el uso de la huella genotípica identificada en (iii), se puede generar una línea celular diseñada eficiente con el genotipo óptimo para producir glicoproteína con dicha glicoforma.

- 60 Otro aspecto de la presente invención se refiere a una glicoproteína según la presente invención, que es un conjugado de glicoproteína homogéneo que comprende un polímero seleccionado entre, PEG, HEP, XTEN, PSA, HES.

- 65 Un aspecto adicional de la presente invención se refiere a una glicoproteína según la presente invención, que es un conjugado proteico PEGilado homogéneo producido a partir de una variante proteica según la invención que tiene un perfil de glicano simplificado.

Otro aspecto de la presente invención se refiere a una glicoproteína según la presente invención, que es un conjugado de EPO PEGilado homogéneo producido a partir de una variante de EPO que tiene un perfil de glicano simplificado.

5 En una realización de la presente invención, Cosmc también se introduce o se elimina.

En una realización de la presente invención, el resultado de la glicoestructura tiene uno o más de los siguientes cambios seleccionados del grupo que consiste en una estructura de glicano más simple, un producto más homogéneo, más ácidos siálicos por molécula, no sialilado, no galactosilado, bi-antenario más homogéneo, monoantenario más homogéneo, triantenario más homogéneo, más homogéneo sin poli-LacNAc, mayor  
10 productividad en el cultivo celular, nuevas capacidades de glicosilación homogénea estable, más glicoestructura humana, glicosilación más homogénea y una mejor focalización de ADCC de IgG, nivel de fucosa modificado, sin fucosa, sustrato mejorado para generar glicoconjugados.

15 En otro aspecto de la presente invención, uno o más de los genes mencionados anteriormente se eliminan usando nucleasas efectoras similares a activadores de la transcripción (TALEN).

Las TALEN son enzimas de restricción artificiales generadas mediante la fusión de un dominio de unión de ADN efector TAL a un dominio de escisión de ADN.

20 En otro aspecto más de la presente invención, uno o más de los genes mencionados anteriormente se eliminan usando CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas).

Los CRISPR son loci de ADN que contienen repeticiones cortas de secuencias de bases. Cada repetición es  
25 seguida por segmentos cortos de "ADN espaciador" de exposiciones previas a un virus.

Los CRISPR a menudo se asocian con genes cas que codifican proteínas relacionadas con los CRISPR. El sistema CRISPR/Cas es un sistema inmune procarionota que confiere resistencia a elementos genéticos extraños como plásmidos y fagos y proporciona una forma de inmunidad adquirida.

30 Los espaciadores CRISPR reconocen y cortan estos elementos genéticos exógenos de una manera análoga al ARNpi en organismos eucariotas.

El sistema CRISPR/Cas se usa para la edición genética (añadir, alterar o cambiar la secuencia de genes  
35 específicos) y la regulación genética en especies a lo largo del árbol de la vida. Al introducir la proteína Cas9 y los ARN guía adecuados en una célula, se puede cortar el genoma del organismo en cualquier ubicación deseada.

En una realización adicional de la presente invención, la célula de mamífero tiene o no actividad de  $\alpha$ -1,6-  
40 fucosiltransferasa.

La célula de la presente invención puede ser una célula que no comprende el gen de interés a expresar o la célula puede comprender el gen de interés a expresar. La célula que no contiene el gen que interesa expresar se suele denominar "célula desnuda".

45 La célula huésped de la presente invención puede ser cualquier huésped, siempre que pueda expresar una molécula de anticuerpo. Los ejemplos incluyen una célula de levadura, una célula animal, una célula de insecto, una célula vegetal y similares.

50 En una realización de la presente invención, la célula se selecciona del grupo que consiste en CHO, NS0, SP2/0, YB2/0, YB2/3HL.P2.G11.16Ag.20, NSO, SP2/0-Ag14, célula BHK derivada de un tejido de riñón de hámster sirio, célula de hibridoma productora de anticuerpos, línea celular de leucemia humana (célula de Namalwa), una célula madre embrionaria y un óvulo fertilizado.

55 En una realización preferida de la presente invención, la célula es una célula CHO.

La célula puede ser una célula aislada o en cultivo celular o una línea celular.

La proteína de interés puede ser varios tipos de proteína, y en particular proteínas que se benefician de ser  
60 expresadas como glicoproteínas, y que pueden ser producidas por las células de la presente invención.

En una realización preferida, la proteína es eritropoyetina (EPO).

En otra realización está la proteína  $\alpha$ 1-antitripsina.

65 En un aspecto de la presente invención, la proteína es un factor sanguíneo recombinante.

En una realización de la presente invención, el factor sanguíneo recombinante se selecciona del grupo que consiste en uno o más de los factores VIII, IX, subunidad A del factor XIII, trombina y VIIa.

5 En otra realización de la presente invención, la proteína de interés es un factor de coagulación, como el factor de coagulación II (FII), el factor de coagulación V (FV), el factor de coagulación VII (FVIIa), el factor de coagulación VIII (FVIII), el factor de coagulación IX (FIX), el factor de coagulación X (FX) o el factor de coagulación XIII (FXIII).

10 En un aspecto de la presente invención es la proteína hormona de crecimiento humano.

En otra realización de la presente invención, la proteína de interés es un anticuerpo, un fragmento de anticuerpo, como un fragmento Fab, un dominio Fc de un anticuerpo o un polipéptido.

15 Un aspecto adicional de la presente invención se refiere a un método para producir una glicoproteína modificada enzimáticamente, que comprende el paso de generar una célula con propiedades de glicosilación específicas, y la modificación enzimática de interés.

En una realización de la presente invención se realiza la modificación enzimática de un polímero.

20 En otra realización de la presente invención es la modificación enzimática una conjugación a otra proteína.

Otro aspecto de la presente invención se refiere a una glicoproteína según la presente invención, que es un conjugado de glicoproteína homogéneo que comprende un polímero.

25 La glicoproteína puede modificarse enzimáticamente con el polímero, y la glicoproteína puede ser producida por la célula de acuerdo con la presente invención.

30 En un aspecto de la presente invención, la proteína es un trombolítico recombinante, un anticoagulante u otro producto relacionado con la sangre.

35 En una realización de la presente invención es el trombolítico recombinante, anticoagulante u otro producto relacionado con la sangre seleccionado del grupo que consiste en uno o más de activador tisular del plasminógeno (tPA), hirudina, antitrombina, plasmina, inhibidor de la calicreína plasmática y proteína C activada.

En un aspecto de la presente invención es la proteína una hormona recombinante.

40 En una realización de la presente invención, la hormona recombinante se selecciona del grupo que consiste en uno o más de los siguientes: insulina, insulina degludec, hormona de crecimiento humana, somatropina, pegvisomant, hormona folículo estimulante, folitropina alfa, corifolitropina alfa, folitropina beta, metreleptina, liraglutida, hormona paratiroidea, lutropina, teriparatida, nesiritida y glucagón.

En un aspecto de la presente invención es la proteína una hormona de crecimiento recombinante.

45 En una realización de la presente invención, la hormona de crecimiento recombinante se selecciona del grupo que consiste en uno o más de EPO, filgrastim, sargramostim, mecaserim y palifermina.

50 En un aspecto de la presente invención es la proteína un interferón recombinante, interleucina o factor de necrosis tumoral.

55 En una realización de la presente invención es el interferón recombinante, interleucina o factor de necrosis tumoral seleccionado del grupo que consiste en uno o más de interferón alfa, PEGinterferón alfa, PEGinterferón alfa-2a, PEGinterferón alfa, interferón beta-1b, alglucosidasa alfa y laronidasa.

En un aspecto de la presente invención, la proteína es un antígeno o componente de vacuna.

60 En un aspecto de la presente invención, la proteína de la presente invención es relevante para una enfermedad o trastorno seleccionado del grupo que consiste en uno o más de los siguientes: hemofilia A, hemofilia B, infarto agudo de miocardio, trombocitopenia asociada a heparina, trombosis venosa, adhesión vitreomacular sintomática/tracción vitreomacular, angioedema agudo, deficiencia hereditaria de antitrombina, angioedema hereditario, sepsis, diabetes, diabetes mellitus, retraso del crecimiento/deficiencia de la hormona del crecimiento, infertilidad/subfertilidad, diabetes tipo 2, osteoporosis, hipoglucemia, enfermedad de Paget, cáncer, anemia, neutropenia, hepatitis C e hiperuricemia, artritis reumatoide, hepatitis A y B, artritis, colitis, enfermedad de Crohn, psoriasis, espondilitis anquilosante, colitis ulcerosa.

En una realización de la presente invención, la proteína es un anticuerpo.

En una realización preferida de la presente invención, el anticuerpo es un anticuerpo IgG.

- 5 En la presente invención, la molécula de anticuerpo incluye cualquier molécula, siempre que comprenda la región Fc de un anticuerpo. Los ejemplos incluyen un anticuerpo, un fragmento de anticuerpo, una proteína de fusión que comprende una región Fc y similares.

- 10 El anticuerpo es una proteína que se produce en el cuerpo vivo por reacción inmune como resultado de la estimulación del antígeno exógeno y tiene una actividad para unirse específicamente al antígeno. Los ejemplos del anticuerpo incluyen un anticuerpo secretado por una célula de hibridoma preparada a partir de una célula de bazo de un animal inmunizado con un antígeno; un anticuerpo preparado mediante una técnica de recombinación genética, es decir, un anticuerpo obtenido introduciendo un vector de expresión de anticuerpo insertado en un gen de anticuerpo en una célula huésped; y similares. Los ejemplos específicos incluyen un anticuerpo producido por un hibridoma, un anticuerpo humanizado, un anticuerpo humano y similares.

- 15 Un hibridoma es una célula que se obtiene por fusión celular entre una célula B obtenida mediante la inmunización de un mamífero distinto de un ser humano con un antígeno y una célula de mieloma derivada de ratón o similar y puede producir un anticuerpo monoclonal que tiene la especificidad de antígeno deseada.

- 20 Los ejemplos del anticuerpo humanizado incluyen un anticuerpo quimérico humano, un anticuerpo humano injertado con CDR y similares.

- 25 Un anticuerpo quimérico humano es un anticuerpo que comprende una región variable de cadena pesada de anticuerpo (en lo sucesivo denominada "HV" o "VH", siendo la cadena pesada "cadena H") y una región variable de cadena ligera de anticuerpo (en lo sucesivo denominada "LV" o "VL", siendo la cadena ligera "cadena L"), ambas de un animal distinto del humano, una región constante de cadena pesada de anticuerpo humano (en lo sucesivo denominada también "CH") y una región constante de cadena ligera de anticuerpo humano (en lo sucesivo denominada también "CL"). Como animal distinto del ser humano, se puede usar cualquier animal como ratón, rata, hámster, conejo o similar, siempre que se pueda preparar un hibridoma a partir de ellos.

- 30 El anticuerpo quimérico humano se puede producir obteniendo ADNc que codifica VH y VL de un hibridoma productor de anticuerpos monoclonales, insertándolos en un vector de expresión para una célula huésped que tiene genes que codifican el anticuerpo humano CH y el anticuerpo humano CL para construir de ese modo un vector de expresión de anticuerpo quimérico humano, y luego introduciendo el vector en una célula huésped para expresar el anticuerpo.

- 35 Como CH del anticuerpo quimérico humano, se puede usar cualquier CH, siempre que pertenezca a la inmunoglobulina humana (en adelante denominada "hlg"). Pero son preferibles los que pertenecen a la clase hlgG y se puede usar cualquiera de las subclases pertenecientes a la clase hlgG, como hlgG1, hlgG2, hlgG3 y hlgG4. Además, como CL del anticuerpo quimérico humano, se puede usar cualquier CL, siempre que pertenezca a la clase hlg, y también se pueden usar aquellos que pertenecen a la clase  $\kappa$  o a la clase  $\lambda$ .

- 40 Un anticuerpo injertado con CDR humano es un anticuerpo en el que las secuencias de aminoácidos de CDR de VH y VL de un anticuerpo derivado de un animal distinto de un humano se injertan en posiciones apropiadas de VH y VL de un anticuerpo humano.

- 45 El anticuerpo humano con CDR injertada puede producirse construyendo ADNc que codifiquen regiones V en las que las CDR de VH y VL de un anticuerpo derivado de un animal que no sea humano se injerten en CDR de VH y VL de un anticuerpo humano, insertándolos en un vector de expresión para una célula huésped que tenga genes que codifiquen el anticuerpo humano CH y el anticuerpo humano CL para construir así un vector de expresión de anticuerpo humano injertado con CDR, e introduciendo después el vector de expresión en una célula huésped para expresar el anticuerpo humano injertado con CDR.

- 50 Como CH del anticuerpo humano injertado con CDR, puede usarse cualquier CH, siempre que pertenezca a la clase hlg, pero son preferibles los de la clase hlgG y puede usarse cualquiera de las subclases pertenecientes a la clase hlgG, como hlgG1, hlgG2, hlgG3 y hlgG4. Además, como CL del anticuerpo humano injertado con CDR, se puede usar cualquier CL, siempre que pertenezca a la clase hlg, y también se pueden usar los que pertenezcan a la clase  $\kappa$  o a la clase  $\lambda$ .

- 55 Un anticuerpo humano es originalmente un anticuerpo que existe naturalmente en el cuerpo humano, pero también incluye anticuerpos obtenidos de una biblioteca de fagos de anticuerpos humanos, un animal transgénico productor de anticuerpos humanos y una planta transgénica productora de anticuerpos humanos, que se preparan basándose en el reciente avance en ingeniería genética, ingeniería celular e ingeniería del desarrollo.

Con respecto al anticuerpo existente en el cuerpo humano, un linfocito capaz de producir el anticuerpo puede cultivarse aislando un linfocito de sangre periférica humana, inmortalizándolo mediante su infección con el virus EB o similar y luego clonándolo, y el anticuerpo puede purificarse a partir del cultivo.

- 5 La biblioteca de fagos de anticuerpos humanos es una biblioteca en la que fragmentos de anticuerpos tales como Fab, anticuerpos de cadena sencilla y similares se expresan en la superficie del fago insertando un gen que codifica un anticuerpo preparado a partir de una célula B humana en un gen de fago. Un fago que exprese un fragmento de anticuerpo con la actividad de unión a antígeno deseada puede recuperarse de la biblioteca, usando como marcador su actividad de unión a un sustrato inmovilizado con antígeno. El fragmento de anticuerpo puede convertirse además en una molécula de anticuerpo humano que comprende dos cadenas B completas y dos cadenas L completas mediante técnicas de ingeniería genética.

- 15 Un animal no humano transgénico productor de anticuerpos humanos es un animal en el que se introduce un gen de anticuerpo humano en las células. Concretamente, se puede preparar un animal transgénico productor de anticuerpos humanos introduciendo un gen de anticuerpos humanos en una célula ES de un ratón, trasplantando la célula ES a un embrión en fase inicial de otro ratón y desarrollándolo a continuación. Introduciendo un gen de anticuerpo quimérico humano en un óvulo fertilizado y desarrollándolo, también se puede preparar el animal transgénico. En lo que respecta al método de preparación de un anticuerpo humano a partir del animal transgénico productor de anticuerpos humanos, el anticuerpo humano puede producirse y acumularse en un cultivo obteniendo un hibridoma productor de anticuerpos humanos mediante un método de preparación de hibridomas realizado habitualmente en mamíferos distintos del humano y cultivándolo a continuación.

- 25 Los ejemplos del animal no humano transgénico incluyen ganado, ovejas, cabras, cerdos, caballos, ratones, ratas, aves, monos, conejos y similares.

- Otro aspecto de la presente invención se refiere a un método para producir una composición de anticuerpos, que comprende cultivar la célula de mamífero según la presente invención en un medio para producir y acumular una composición de anticuerpos en el cultivo; y recuperar la composición de anticuerpos del cultivo.

- 30 Además, en la presente invención, es preferible que el anticuerpo sea un anticuerpo que reconoce un antígeno relacionado con un tumor, un anticuerpo que reconoce un antígeno relacionado con una alergia o inflamación, un anticuerpo que reconoce un antígeno relacionado con una enfermedad de órganos circulatorios, un anticuerpo que reconoce un antígeno relacionado con una enfermedad autoinmune o un anticuerpo que reconoce un antígeno relacionado con una infección viral o bacteriana, y es preferible un anticuerpo humano que pertenece a la clase IgG.

- 40 Un fragmento de anticuerpo es un fragmento que comprende la región Fc de un anticuerpo. Los ejemplos del fragmento de anticuerpo incluyen un monómero de cadena H, un dímero de cadena H y similares.

- Una proteína de fusión que comprende una región Fc es una composición en la que un anticuerpo que comprende la región Fc de un anticuerpo o el fragmento de anticuerpo se fusiona con una proteína tal como una enzima, una citocina o similar.

- 45 Una realización de la presente invención se refiere a una composición que comprende el anticuerpo de la presente invención, también denominada composición de anticuerpo.

- En una realización de la presente invención, el anticuerpo o la composición de anticuerpos muestra una alta actividad de ADCC.

- 50 En la presente invención, la actividad ADCC es una actividad citotóxica en la que un anticuerpo unido a un antígeno de la superficie celular en una célula tumoral en el cuerpo vivo activa una célula efectora a través de un receptor Fc existente en la región Fc del anticuerpo y la superficie de la célula efectora y de ese modo obstruye la célula tumoral y similares.

- 55 La composición de anticuerpos de la presente invención tiene una potente actividad citotóxica mediada por células dependiente de anticuerpos (ADCC). Un anticuerpo que tenga una potente actividad citotóxica mediada por células dependientes de anticuerpos es útil para prevenir y tratar diversas enfermedades, como cánceres, enfermedades inflamatorias, enfermedades inmunitarias como enfermedades autoinmunitarias, alergias y similares, enfermedades de los órganos circulatorios e infecciones víricas o bacterianas.

- 60 En el caso de los cánceres, es decir, los tumores malignos, las células cancerosas crecen. Los agentes antitumorales generales inhiben el crecimiento de las células cancerosas. Por el contrario, un anticuerpo que tenga una potente actividad citotóxica mediada por células dependientes de anticuerpos puede tratar cánceres lesionando las células cancerosas mediante su efecto de destrucción celular y, por lo tanto, es más eficaz como agente terapéutico que los agentes antitumorales generales.

En las enfermedades inmunitarias como las enfermedades inflamatorias, las enfermedades autoinmunitarias, las alergias y similares, las reacciones in vivo de las enfermedades son inducidas por la liberación de una molécula mediadora por los inmunocitos, de modo que la reacción de alergia puede inhibirse eliminando los  
5 inmunocitos mediante un anticuerpo que tenga una potente actividad citotóxica mediada por células dependientes de anticuerpos.

Entre los ejemplos de enfermedades del órgano circulatorio se incluyen la arteriosclerosis y similares. En la actualidad, la arteriosclerosis se trata mediante catéter con balón, pero las enfermedades de los órganos  
10 circulatorios se pueden prevenir y tratar inhibiendo el crecimiento de células arteriales en la restricción después del tratamiento usando un anticuerpo que tenga una potente actividad citotóxica mediada por células dependiente de anticuerpos.

Diversas enfermedades, incluidas las infecciones virales y bacterianas, se pueden prevenir y tratar inhibiendo la proliferación de células infectadas con un virus o bacteria usando un anticuerpo que tenga una potente  
15 actividad citotóxica mediada por células dependiente de anticuerpos.

Las glicoproteínas de la presente invención pueden por tanto usarse para tratar enfermedades inmunes, cáncer, infecciones virales o bacterianas u otras enfermedades o trastornos mencionados anteriormente.  
20

El medicamento que comprende la composición glicoproteica de la presente invención puede administrarse como agente terapéutico solo, pero generalmente, es preferible proporcionarlo como una formulación farmacéutica producida por un método apropiado bien conocido en el campo técnico de la farmacia de  
25 fabricación, mezclándolo con al menos un portador farmacéuticamente aceptable.

Es deseable seleccionar una vía de administración que sea más eficaz en el tratamiento. Los ejemplos incluyen la administración oral y la administración parenteral, como bucal, traqueal, rectal, subcutánea, intramuscular, intravenosa o similares. En una preparación de anticuerpos, es preferible la administración intravenosa.  
30

La forma de dosificación incluye aerosoles, cápsulas, tabletas, gránulos, jarabes, emulsiones, supositorios, inyecciones, ungüentos, cintas y similares.

Los ejemplos de la preparación farmacéutica adecuada para administración oral incluyen emulsiones, jarabes, cápsulas, tabletas, polvos, gránulos y similares.  
35

Se pueden producir preparaciones líquidas, tales como emulsiones y jarabes, usando, como aditivos, agua; sacáridos, tales como sacarosa, sorbitol, fructosa, etc.; glicoles, tales como polietilenglicol, propilenglicol, etc.; aceites, tales como aceite de sésamo, aceite de oliva, aceite de soja, etc.; antisépticos, tales como ésteres de ácido p-hidroxibenzoico, etc.; sabores, tales como sabor a fresa, menta, etc.; y similares.  
40

Se pueden producir cápsulas, comprimidos, polvos, gránulos y similares usando, como aditivo, cargas, tales como lactosa, glucosa, sacarosa, manitol, etc.; agentes desintegrantes, tales como almidón, alginato de sodio, etc.; lubricantes, tales como estearato de magnesio, talco, etc.; aglutinantes, tales como alcohol polivinílico, hidroxipropilcelulosa, gelatina, etc.; surfactantes, tales como ésteres de ácidos grasos, etc.; plastificantes, tales  
45 como glicerina, etc.; y similares.

Los ejemplos de la preparación farmacéutica adecuada para la administración parenteral incluyen inyecciones, supositorios, aerosoles y similares.

Las inyecciones se pueden preparar usando un vehículo, tal como una solución salina, una solución de glucosa, una mezcla de ambas o similares. Además, se pueden preparar inyecciones en polvo liofilizando la composición de anticuerpos de la forma habitual y añadiéndole cloruro de sodio.  
50

Los supositorios se pueden preparar usando un vehículo tal como manteca de cacao, grasa hidrogenada, ácido carboxílico o similares.  
55

Además, se pueden preparar aerosoles usando la composición de anticuerpos como tal o usando un portador que no estimule la membrana mucosa bucal o de las vías respiratorias del paciente y que pueda facilitar la absorción de la composición de anticuerpos dispersándola como partículas finas.  
60

Los ejemplos del portador incluyen lactosa, glicerol y similares. Dependiendo de las propiedades de la composición del anticuerpo y del portador, es posible producir preparaciones farmacéuticas tales como aerosoles, polvos secos y similares. Por otra parte, los componentes ejemplificados como aditivos para preparados orales también pueden añadirse a los preparados parenterales.  
65

Aunque la dosis clínica o la frecuencia de administración varía dependiendo del efecto terapéutico objetivo, el método de administración, el período de tratamiento, la edad, el peso corporal y similares, normalmente es de 10 µg/kg a 20 mg/kg por día y por adulto.

- 5 Cabe señalar que las realizaciones y características descritas en el contexto de uno de los aspectos de la presente invención también se aplican a los demás aspectos de la invención.

#### Definiciones

#### 10 Glicobiología general

Los principios y definiciones básicos de la glicobiología se describen en Varki et al. *Essentials of Glycobiology*, 2.<sup>a</sup> edición, 2009.

- 15 "N-glicosilación" se refiere a la unión del oligosacárido de la molécula de azúcar conocido como glicano a un residuo de átomo de nitrógeno de una proteína.

La "O-glicosilación" se refiere a la unión de una molécula de azúcar a un átomo de oxígeno en un residuo de aminoácido de una proteína.

- 20 "Galactosilación" significa adición enzimática de un residuo de galactosa a lípidos, carbohidratos o proteínas.

La "sialilación" es la adición enzimática de un residuo de ácido neuramínico.

- 25 El "ácido neuramínico" es un monosacárido de 9 carbonos, un derivado de una cetona.

El glicano N-ligado "monoantenario" es un N-glicano diseñado que consiste en el núcleo del N-glicano (Man $\alpha$ 1-6(Man $\alpha$ 1-3)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-Asn-X-Ser/Thr) que se alargó con un único residuo de GlcNAc unido a C-2 y de la manosa  $\alpha$ 1-3. El único residuo GlcNAc puede alargarse aún más, por ejemplo, con residuos Gal o Gal y NeuAc.

- 30 El glicano N-ligado "biantenario" es el más simple de los complejos. Los glicanos N-ligados consisten en un núcleo de N-glicano (Man $\alpha$ 1-6(Man $\alpha$ 1-3)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-Asn-X-Ser/Thr) alargado con dos residuos de GlcNAc unidos a C-2 y de la manosa  $\alpha$ 1-3 y la manosa  $\alpha$ 1-6. Esta estructura central puede entonces alargarse o modificarse mediante diversas estructuras de glicanos.

Los glicanos N-ligados "triantenarios" se forman cuando se añade un residuo GlcNAc adicional al C-4 de la manosa del núcleo  $\alpha$ 1-3 o al C-6 de la manosa del núcleo  $\alpha$ 1-6 de la estructura del núcleo bi-antenario. A continuación, esta estructura puede alargarse o modificarse mediante diversas estructuras de glicanos.

- 40 Los glicanos N-ligados "tetraantenarios" se forman cuando se añaden dos residuos GlcNAc adicionales al C-4 de la manosa del núcleo  $\alpha$ 1-3 o al C-6 de la manosa del núcleo  $\alpha$ 1-6 de la estructura del núcleo bi-antenario. Esta estructura central puede entonces alargarse o modificarse mediante diversas estructuras de glicanos.

- 45 "Poly-LacNAc" poli-N-acetil-lactosamina ([Gal $\beta$ 1-4GlcNAc]<sub>n</sub>; n  $\geq$  2.)

"Glicoperfilado" significa la caracterización de las estructuras de glicano residentes en una molécula o célula biológica.

- 50 La "vía de glicosilación" se refiere al ensamblaje de monosacáridos en un grupo de estructuras de carbohidratos complejos relacionadas mediante la acción gradual de enzimas, conocidas como glicosiltransferasas. Las vías de glicosilación en las células de mamíferos se clasifican en glicosilación proteica ligada al N, diferentes glicosilaciones proteicas ligadas al O (O-GalNAc, O-GlcNAc, O-Fuc, O-Glc, O-Xyl, O-Gal), series difeernéticas de glicosfingolípidos y anclajes GPI.

- 55 "Paso biosintético" significa la adición de un monosacárido a una estructura de glicano.

Las "glicosiltransferasas" son enzimas que catalizan la formación del enlace glucosídico para formar un glicósido. Estas enzimas usan fosfatos de azúcar "activados" como donadores de glicosilo y catalizan la transferencia del grupo glicosilo a un grupo nucleofílico, generalmente un alcohol. El producto de la transferencia de glicosilo puede ser un O-, N-, S- o C-glicósido; el glicósido puede ser parte de un monosacárido, oligosacárido o polisacárido.

- 60 "Capacidad de glicosilación" significa la capacidad de producir una cantidad de una estructura de glicano específica por una célula dada o un proceso de glicosilación dado.

"HEP" es una abreviatura de heparosano. El heparosano (HEP) es un polímero de azúcar natural que comprende repeticiones de (-GlcUA-1,4-GlcNAc-1,4-). Pertenece a la familia de los polisacáridos glicosaminoglicanos y es un polímero con carga negativa a pH fisiológico. Se encuentra en la cápsula de ciertas bacterias, pero también en vertebrados superiores, donde sirve como precursor de los polímeros naturales heparina y heparán sulfato.

"PEG" es una abreviatura de polietilenglicol. Los polímeros de PEG se usan generalmente para aumentar la vida media de las moléculas terapéuticas. Con su alto volumen hidrodinámico, los polímeros de PEG aumentan el tamaño efectivo de los fármacos y, por lo tanto, retardan su eliminación del torrente sanguíneo a través de la filtración renal o protegiendo el fármaco proteico hacia los receptores de eliminación y la degradación proteolítica.

"XTEN" son secuencias de péptidos bien definidas basadas en un subconjunto limitado de aminoácidos y ensambladas en secuencias repetidas más largas. XTEN también aumenta el tamaño de las moléculas terapéuticas y prolonga su presencia en el torrente sanguíneo. La "glicoPEGilación" es el proceso de unir covalentemente PEG a los glicanos de una proteína de interés.

La "glicoPEGilación enzimática" es el proceso de unir covalentemente PEG a los glicanos de una proteína de interés usando una glicosiltransferasa y grupos de transferencia de glicosilo PEGilados adecuados, como moléculas NeuAc-CMP PEGiladas.

"Glicoconjugado" es una macromolécula que contiene monosacáridos unidos covalentemente a proteínas o lípidos.

"Estructura de glicano más simple" es una estructura de glicano que contiene menos monosacáridos y/o que tiene menor masa y/o que tiene menos antenas.

"Glicosilación similar a la humana" significa tener estructuras de glicano similares a las de las células humanas. Los ejemplos incluyen más ácidos siálicos con enlace  $\alpha$ 2,6 (más enzima  $\alpha$ 2,6 sialiltransferasa) y/o menos ácidos siálicos con enlace  $\alpha$ 2,3 y/o más ácido N-acetilneuramínico (Neu5Ac) y/o menos ácido N-glicolilneuramínico (Neu5Gc).

"Escisión" se refiere a la rotura de la cadena principal covalente de una molécula de ADN. La escisión puede iniciarse mediante una variedad de métodos que incluyen, entre otros, la hidrólisis enzimática o química de un enlace fosfodiéster. Son posibles tanto la escisión de cadena simple como de escisión de cadena doble, y la escisión de cadena doble puede ocurrir como resultado de dos eventos de escisión de cadena doble distintos. La escisión del ADN puede dar lugar a la producción de extremos romos o escalonados.

En ciertas realizaciones, los polipéptidos de fusión se usan para la escisión dirigida de ADN bicatenario.

"Deconstrucción" significa obtener células que produzcan estructuras de glicano más simples mediante la eliminación simple o apilada de glicosiltransferasas. La deconstrucción de una vía de glicosilación significa el KO de las glicosiltransferasas implicadas en cada paso de la biosíntesis y la identificación de las glicosiltransferasas que controlan cada paso biosintético.

"Reconstrucción" se refiere a la inserción de genes exógenos en células o a la activación de genes endógenos silenciosos para obtener estructuras de glicanos más complejas. Normalmente, las células objetivo producen estructuras de glicano simples como resultado de la deconstrucción.

"Perfil de glicanos modificado" se refiere al cambio en el número, tipo o posición de oligosacáridos en los glicanos de una glicoproteína determinada.

Más "glicosilación homogénea" o "capacidades de glicosilación estable homogénea" significa que la proporción de estructuras de glicanos idénticas observadas al glicoprofilizar una proteína dada expresada en una célula es mayor que la proporción de estructuras de glicanos idénticas observadas al glicoprofilizar la misma proteína expresada en otra célula. Se puede obtener una glicosilación más homogénea mediante la inserción y/o eliminación de uno o más genes de glicosiltransferasa en una célula. Un ejemplo de tal modificación podría ser la eliminación de B4galt1 para generar N-glicanos más homogéneos sin galactosa. La eliminación adicional de fut8 permite la producción de N-glicanos más homogéneos sin galactosa ni fucosa. En general, este tipo de modificaciones knock-in y/o knockout puede dar lugar a cambios del resultado de la glicoestructura de la célula modificada, como la lista no limitativa de estructura de glicano más simple, producto más homogéneo, más ácidos siálicos por molécula, no sialilado, no galactosilado, biantenarior más homogéneo, monoantenarior más homogéneo, triantenarior más homogéneo, glicosilación más homogénea sin poli-LacNAc, mayor productividad en cultivo celular, nuevas capacidades estables de glicosilación homogénea, glicoestructura más humana, glicosilación más homogénea y mejor orientación ADCC de IgG, nivel de fucosa modificado, sin fucosa y sustrato mejorado para generar glicoconjugados.

ADN general, biol. molecular. Cualquiera de las diversas técnicas usadas para separar y recombinar segmentos de ADN o genes, comúnmente mediante el uso de una enzima de restricción para cortar un fragmento de ADN del ADN donante e insertarlo en un plásmido o ADN viral. Utilizando estas técnicas, el ADN que codifica una proteína de interés se recombina/clona (usando PCR y/o enzimas de restricción y ADN ligasas o métodos independientes de la ligación, como la clonación USER) en un plásmido (conocido como vector de expresión), que posteriormente puede introducirse en una célula mediante transfección usando diversos métodos de transfección, como la transfección con fosfato cálcico, la electroporación, la microinyección y la transfección con liposomas. Se puede encontrar una descripción general, información complementaria y métodos para construir secuencias de ADN sintéticas, inserción en vectores plasmídicos y posterior transfección en células en Ausubel et al, 2003 y/o Sambrook & Russell, 2001.

"Gen" se refiere a una región de ADN (incluyendo exones e intrones) que codifica un producto génico, así como a todas las regiones de ADN que regulan la producción del producto génico, ya sea que dichas secuencias reguladoras estén adyacentes a secuencias codificantes y/o transcritas o situadas lejos del gen cuya función regulan. En consecuencia, un gen incluye, pero no se limita necesariamente a, secuencias promotoras, terminadoras, secuencias reguladoras de la traducción, como sitios de unión a ribosomas y sitios de entrada de ribosomas internos, potenciadores, silenciadores, aisladores, elementos delimitadores, orígenes de replicación, sitios de unión a matrices y regiones de control de locus.

"Modificaciones genéticas dirigidas", "edición genética" o "edición del genoma"

La edición genética o edición del genoma se refiere a un proceso mediante el cual se cambia una secuencia cromosómica específica. La secuencia cromosómica editada puede comprender una inserción de al menos un nucleótido, una eliminación de al menos un nucleótido y/o una sustitución de al menos un nucleótido. En general, la edición del genoma inserta, sustituye o elimina ácidos nucleicos de un genoma usando nucleasas artificiales como las nucleasas de dedos de zinc (ZFN), las nucleasas efectoras similares a activadores de la transcripción (TALEN), el sistema CRISPR/Cas y las endonucleasas homing reingenierizadas con meganucleasas. Los principios de edición del genoma se describen en Steentoft et al y los métodos de edición de genes se describen en las referencias de los mismos y también se usan ampliamente y, por tanto, son conocidos por los expertos en la materia.

La secuencia/gen/proteína "endógena" se refiere a una secuencia cromosómica, un gen o una proteína que es nativa de la célula o que se origina dentro de la célula u organismo analizado.

Secuencia o gen "exógeno" se refiere a una secuencia cromosómica que no es nativa de la célula, o una secuencia cromosómica cuya ubicación cromosómica nativa está en una ubicación diferente en un cromosoma o se origina fuera de la célula u organismo analizado.

"Secuencia cromosómica inactivada" se refiere a la secuencia genómica que ha sido editada dando lugar a la pérdida de función de un determinado producto génico. Se dice que el gen ha sido eliminado.

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

Los términos "ácido nucleico" y "polinucleótido" se refieren a un polímero de desoxirribonucleótido o ribonucleótido, en conformación lineal o circular, y en forma monocatenaria o bicatenaria. A efectos de la presente divulgación, estos términos no deben interpretarse como limitativos con respecto a la longitud de un polímero. Las expresiones pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en los restos de base, azúcar y/o fosfato (por ejemplo, estructuras de fosforotioato). En general, un análogo de un nucleótido en particular tiene la misma especificidad de emparejamiento de bases; es decir, un análogo de A se emparejará con T.

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 de purina o pirimidina modificada o una fracción de ribosa modificada. Un análogo de nucleótido puede ser un nucleótido de origen natural (por ejemplo, inosina) o un nucleótido de origen no natural.

Los términos "polipéptido" y "proteína" se usan indistintamente para referirse a un polímero de residuos de aminoácidos. Estos términos también pueden referirse a variantes glicosiladas del "polipéptido" o "proteína", también denominados "glicoproteína", "polipéptido", "proteína" y "glicoproteína" se usan indistintamente a lo largo de esta divulgación.

El término "recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos. A efectos de esta divulgación, "recombinación homóloga" se refiere a la forma especializada de dicho intercambio que tiene lugar, por ejemplo, durante la reparación de roturas de doble cadena en las

células. Este proceso requiere similitud de secuencia entre los dos polinucleótidos, usa una molécula "donante" o "de intercambio" para la reparación de la molécula "diana" (es decir, la que ha sufrido la rotura de la doble cadena) y se conoce como "conversión génica no cruzada" o "conversión génica de tracto corto", ya que conduce a la transferencia de información genética del donante a la diana. Sin ceñirse a ninguna teoría en particular, dicha transferencia puede implicar la corrección de los desajustes del ADN heterodúplex que se forma entre la diana rota y el donante, y/o el "recocido de hebra dependiente de la síntesis", en el que el donante se usa para resintetizar la información genética que pasará a formar parte de la diana, y/o procesos relacionados. Dicha recombinación homóloga especializada a menudo da lugar a una alteración de la secuencia de la molécula diana de forma que parte o toda la secuencia del polinucleótido donante se incorpora al polinucleótido diana.

Como se usa en este documento, los términos "sitio objetivo" o "secuencia objetivo" se refieren a una secuencia de ácido nucleico que define una porción de una secuencia cromosómica que se va a editar y a la cual una endonucleasa de orientación está diseñada para reconocer, unirse y escindir.

La "integración dirigida" es el método mediante el cual los elementos de ácido nucleico exógeno se integran específicamente en loci definidos del genoma celular. Las roturas bicatenarias específicas se introducen en el genoma mediante nucleasas de edición genómica que permiten la integración del elemento de ácido nucleico donante administrado exógenamente en el sitio de rotura bicatenaria. De este modo, el elemento de ácido nucleico donante suministrado exógenamente se integra de forma estable en el locus definido del genoma celular.

Las Técnicas de identidad de secuencias para determinar la identidad de secuencias de ácidos nucleicos y aminoácidos son conocidas en la técnica. Típicamente, tales técnicas incluyen la determinación de la secuencia de nucleótidos del ARNm de un gen y/o la determinación de la secuencia de aminoácidos codificada por el mismo, y la comparación de estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también se pueden determinar y comparar de esta manera. En general, la identidad se refiere a una correspondencia exacta de nucleótido a nucleótido o de aminoácido a aminoácido de dos polinucleótidos o secuencias polipeptídicas, respectivamente. Se pueden comparar dos o más secuencias (polinucleótidos o aminoácidos) determinando su porcentaje de identidad. El porcentaje de identidad de dos secuencias, ya sean secuencias de ácidos nucleicos o 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. El algoritmo de homología local de Smith y Waterman, *Advances in Applied Mathematics* 2:482- 489 (1981 ) proporciona un alineamiento aproximado de las secuencias de ácidos nucleicos. Este algoritmo se puede aplicar a secuencias de aminoácidos usando la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*, MO Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation,

Washington, DC, EE. UU., y normalizado por Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). Una implementación ejemplar de este algoritmo para determinar el porcentaje de identidad de una secuencia la proporciona el Genetics Computer Group (Madison, Wis.) en el programa aplicación "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 usarse con los siguientes parámetros por defecto: código genético=estándar; filtro=ninguno; cadena=ambas; corte =60; expect=10; matriz=BLOSUM62; descripciones=50 secuencias; ordenar por=PUNTUACIÓN ALTA; bases de datos=no redundantes, GenBank+EMBL+DDBJ+PDB+traducciones CDS de GenBank+Swiss proteína+Spupdate+PIR. Encontrará información detallada sobre estos programas en la página web de GenBank. Con respecto a las secuencias descritas en el presente documento, el rango de grados deseados de identidad de secuencia es aproximadamente del 80% al 100% y cualquier valor entero entre ambos. Típicamente los porcentajes de identidad entre secuencias son de al menos 70-75%, preferiblemente 80- 82%, más preferiblemente 85-90%, aún más preferiblemente 92%, aún más preferiblemente 95%, y más preferiblemente 98% de identidad de secuencia.

Todas las referencias de patentes y no patentes citadas en la presente solicitud se incorporan por referencia en su totalidad.

La invención se describirá ahora con más detalle en los siguientes ejemplos no limitativos.

#### Ejemplos

El propósito de los siguientes ejemplos se da como una ilustración de varias realizaciones de la invención y, por lo tanto, no pretenden limitar la presente invención de ninguna manera. Junto con los presentes ejemplos, los métodos aquí descritos son actualmente representativos de realizaciones preferidas, son ejemplares y no pretenden limitar el alcance de la invención. A los expertos en la materia se les ocurrirán cambios en la misma y otros usos que estén comprendidos dentro del espíritu de la invención tal como se define en el alcance de las reivindicaciones.

## Ejemplo 1: deconstrucción de la n-glicosilación en células cho

## Párrafo introductorio

5

La célula de ovario de hámster chino (CHO) es la fábrica de células elegida para la producción recombinante de terapias biológicas. CHO es capaz de realizar muchas modificaciones postraduccionales necesarias para la producción de proteínas bioactivas, y tiene una capacidad genérica para la glicosilación de proteínas compatibles con el uso humano. La glicosilación tiene una amplia gama de efectos sobre la producción, la consistencia del producto, la solubilidad, la estabilidad, la bioactividad y la vida media circulatoria de las terapias proteicas, y la disponibilidad de una única plataforma genérica de producción de CHO es un factor limitante para el diseño y el desarrollo de productos biológicos. Este ejemplo muestra un exhaustivo cribado knockout de genes de glicosiltransferasas implicadas en la N-glicosilación en CHO que define las enzimas clave que controlan la ramificación, elongación y sialilación de las antenas. Se estableció un panel de líneas CHO de ingeniería estable con capacidades de glicosilación de diseño, como glicanos biantenarios homogéneos con y sin recubrimiento de ácido siálico  $\alpha 2,3$  y  $\alpha 2,6$ . La pantalla demuestra una gran plasticidad para la glicoingeniería de CHO y proporciona vías claras para diseños personalizados.

10

15

## Introducción

20

Los productos biológicos a base de glicoproteínas son la clase terapéutica de más rápido crecimiento, y la mayoría de ellos sólo pueden producirse de forma recombinante en células de mamífero con capacidad de glicosilación similar a la humana. La célula de ovario de hámster chino (CHO) ha adquirido un papel destacado como célula huésped para la producción recombinante de glicoproteínas terapéuticas, principalmente porque produce N-glicanos bastante simples con ramificación y recubrimiento similares a los que se producen en algunas células humanas.

25

En particular, los CHO producen N-glicanos heterogéneos de tipo complejo con estructuras bi-, tri- y tetraantenarias con un núcleo de  $\alpha$  6-fucosa (Fuc), una cantidad menor de poli-N-acetilactosamina (poli-LacNAc) principalmente en el brazo  $\alpha 1,6$  de las estructuras tetraantenarias, y un recubrimiento exclusivo de LacNAc con ácido neuramínico (NeuAc) enlazado  $\alpha 2,3$  (véase la Fig. 1a). Por lo general, el CHO no produce las estructuras de recubrimiento inmunogénicas no humanas y en el hombre, como el ácido N-glicolilneuramínico (NeuGc) o el  $\alpha 3\text{Gal}$ , aunque se ha informado de la aparición de éstas quizás como resultado de la inducción génica. La N-glicosilación puede variar para diferentes proteínas, así como para diferentes glicositos en proteínas individuales y, por ejemplo, los anticuerpos IgG se producen con estructuras truncadas de N-glicanos en el glicosito conservado Asn297 (estructuras biantenarias con núcleo  $\alpha 6\text{Fuc}$ , LacNAc limitado y recubrimiento NeuAc). Una de las principales preocupaciones del CHO es la heterogeneidad sustancial en el procesamiento del N-glicano, que puede ser difícil de controlar durante el bioprocesamiento y plantear problemas de bioactividad y bioseguridad. Así pues, una de las principales actividades del bioprocesamiento de productos terapéuticos se dedica al análisis de los glicanos y al control de la fermentación para lograr la consistencia.

30

35

40

En las dos últimas décadas se han dedicado importantes esfuerzos a la glicoingeniería genética de células CHO con el objetivo de ampliar la capacidad de glucosilación, reducir la heterogeneidad y mejorar o alterar especialmente la sialilación.

45

Todos estos estudios han usado esencialmente la integración aleatoria de ADNc que codifican glucosiltransferasas y han experimentado problemas de estabilidad, consistencia y previsibilidad de la capacidad de glucosilación introducida. El principal obstáculo ha sido la necesidad de depender de la sobreexpresión de glicosiltransferasas y la competencia con las enzimas endógenas expresadas debido a la falta de métodos sencillos para eliminarlas en líneas celulares. Por lo tanto, hasta donde sabemos, este tipo de células CHO de glicoingeniería no han llegado a la producción de terapias clínicas. Sin embargo, ha surgido una estrategia de glicoingeniería de éxito tras el descubrimiento de que las IgG sin núcleo  $\alpha 6\text{Fuc}$  en el N-glicano Asn297 presentan una citotoxicidad celular dependiente de anticuerpos (ADCC) notablemente superior.

50

55

Así, mediante un tour-de-force usando dos rondas de recombinación homóloga, ambos alelos del gen *fut8* que codifica la  $\alpha$  6fucosiltransferasa que controla la fucosilación del núcleo fueron eliminados en CHO, y al menos una IgG terapéutica producida en CHO sin el gen *fut8* está ahora en uso clínico. Más recientemente, el gen *fut8* se eliminó mediante la edición precisa de genes con la nucleasa de dedo de zinc (ZFN) con una fracción del tiempo y los recursos invertidos. La aparición de tecnologías precisas de edición génica para knockout (KO) y knockin (KI) ha abierto la puerta a un nivel totalmente distinto de rapidez y facilidad con la que se puede lograr la manipulación genética estable de líneas celulares huésped para eliminar e introducir genes de glucosiltransferasas, lo que sin duda repercutirá en la ingeniería de fábricas de células huésped de mamíferos para la producción recombinante de terapéuticos.

60

65

En este caso, hemos empleado por primera vez una pantalla de knockout mediada por ZFN en CHO para explorar el potencial de la ingeniería de la N-glicosilación de glicoproteínas recombinantes. Muchos pasos en la vía biosintética del N-glicano son potencialmente catalizados por múltiples isoenzimas, lo que hace que la ingeniería genética sea impredecible (Fig. 1a). Es necesario analizar las funciones in vivo de cada una de estas isoenzimas para construir una matriz de opciones de diseño. El cribado KO incluyó todos los genes que codifican isoenzimas con potencial para regular la ramificación de la N-glicosilación, la elongación y el tapado terminal expresados en CHO. La expresión de los genes relevantes en CHO GS se evaluó mediante secuenciación cuantitativa de ARN de nueva generación (ARNsec) (Figura 4). La estrategia sólo ha sido posible gracias a la reciente secuenciación del genoma de CHO y al análisis del transcriptoma CHO-K1. Nos centramos en genes individuales y en las combinaciones apiladas más relevantes.

Utilizamos un total de 15 ZFNs dirigidos a genes de glicosiltransferasas (Fig. 1a) implicados en la ramificación antenal (mgat4A/4B/4C/5/5B) (Fig. 1c), galactosilación (B4galt1/2/3/4) (Fig. 1d), elongación de poli-LacNAc (B3gnt1/2/8) (Fig. 1e), y tapado terminal por sialilación (st3gal3/4/6) (Fig. 1f). Utilizamos métodos desarrollados recientemente para enriquecer clones KO mediante FACS (ZFN marcados con GFP/Crimson) y detección de alto rendimiento mediante una estrategia de etiquetado de amplicones (IDAA). La mayoría de los clones KO presentaban inserciones y/o deleciones (indels) en el rango de  $\pm 20$ bps, y la mayoría de los genes diana estaban presentes con dos alelos, mientras que algunos (mgat4B y mgat5) estaban presentes con 1 o 3 alelos, respectivamente. Para monitorizar los efectos del cribado de KO de glicogenes usamos la expresión recombinante estable de eritropoyetina humana (EPO), ya que esta proteína tiene tres N-glicanos con estructuras principalmente tetraantenarias, bajo nivel de poli-LacNAc, y un recubrimiento de ácido siálico a2,3 (Fig. 1b), y es una de las N-glicoproteínas mejor caracterizadas producidas en CHO. La ingeniería se realizó en la línea de producción CHO-GS (Sigma) cultivada en suspensión en medio libre de proteínas, y se encontró que la EPO expresada en la célula de tipo salvaje estaba glicosilada de forma similar a lo que se ha informado en el pasado con otras líneas CHO (Fig. 1b).

No se observaron características morfológicas o fenotípicas de crecimiento consistentes aparentes, aparte de variaciones sutiles en los clones. En promedio, se examinaron entre 5 y 50 clones de cada uno mediante ELISA y se seleccionaron los clones con mayor expresión. Los rendimientos finales de purificación varían de 0,2 mg/L a 6,2 mg/L, pero esta variación se debe principalmente, si no totalmente, al azar y al número limitado de clones examinados para cada clon KO. El clon CHO WT seleccionado produjo rendimientos finales de 0,9 mgs/L, mientras que, por ejemplo, el clon KO mgat4A/4B/5/B3gnt2 con cuatro genes eliminados produjo 3,4 mgs/L.

#### Control del estado de las antenas de N-glicano

MGAT1 y 2 controlan cada uno la formación de una de las dos ramas  $\beta 2$  en N-glicanos biantenarios, mientras que se predice una posible redundancia funcional parcial para la formación de ramas tri- y tetraantenarias por MGAT4A/4B/4C (rama  $\beta 4$ ) y MGAT5/5B (rama  $\beta 6$ ), respectivamente. Sólo mgat4B y 5 parecen expresarse en CHO (Fig. 4). La focalización de mgat4A tuvo un efecto menor, mientras que la focalización de mgat4A/4B casi eliminó los N-glicanos tetraantenarios ramificados  $\beta 4$  (Fig. 1c). Por el contrario, la focalización de mgat5 eliminó por completo las estructuras tetraantenarias ramificadas  $\beta 6$  (Fig. 1c), así como el marcaje de las células con lectina L-PHA (Fig. 5), lo que concuerda con estudios previos de mutantes CHO resistentes a la lectina que carecían de mgat5. El apilamiento KO de mgat4A/4B y 5 produjo N-glicanos biantenarios homogéneos en EPO (Fig. 1c). Sin embargo, se encontró una cantidad menor de poli-LacNAc en los N-glicanos biantenarios, que también se encuentra como componente menor en las células de tipo salvaje.

#### Control de LacNAc

Los CHO expresan las siete  $\beta 4$ galactosiltransferasas conocidas y nuestro conocimiento de las funciones in vivo de estas isoenzimas es deficiente, pero se espera que solo B4Gal-T1-4 desempeñe funciones en la N-glicosilación. Primero examinamos clones KO individuales para B4galt1-4 y descubrimos que B4Gal-T1 parecía tener el papel principal (Fig. 1d). Así, las contribuciones de las otras isoformas sólo pudieron evaluarse en combinaciones apiladas con KO de B4galt1. Probamos las combinaciones apiladas mediante inmunocitología con un anticuerpo contra LacNAc mostrando que sólo el KO apilado de B4galt1/3, pero no de B4galt1/2/4, parecía abolir la galactosilación (Fig. Suplementaria 3). De acuerdo con esto, encontramos que la EPO expresada en clones KO apilados B4galt1/3 pero no B4galt1/2 tenía una reducción sustancial (>90%) de la galactosilación en N-glicanos (Fig. 1d). Sin embargo, la pérdida completa de la capacidad de galactosilación en EPO puede requerir la KO de tres o más genes.

#### Control de poli-LacNAc

Si bien el poli-LacNAc en los N-glicanos puede conferir propiedades de vida media circulatoria superiores a, por ejemplo, la EPO, esta modificación siempre es incompleta y una de las más heterogéneas y difíciles de controlar para la consistencia del producto. Las células CHO generalmente producen bajas cantidades de poli-LacNAc en N-glicanos y principalmente en estructuras tetraantenarias en el brazo  $\beta 6$  controlado por MGAT5. La biosíntesis de poli-LacNAc en N-glicanos es poco conocida y tres genes pueden estar implicados, B3gnt1,

B3gnt2 y B3gnt8. Evaluamos la KO de genes individuales de B3gnt1, B3gnt2, y B3gnt8 con inmunocitología lectina usando la lectina LEL que reconoce poli-LacNAc, e identificamos B3gnt2 como el gen clave que controla la iniciación de LacNAc en células CHO (Fig. 5B). De acuerdo con esto, la EPO expresada en clones CHO con KO de B3gnt2 y no de B3gnt1, resultó en la eliminación de la fracción menor de poli-LacNAc en la EPO (Fig. 1e).

#### Control de la sialilación

Los CHO producen casi exclusivamente una protección NeuAc de los N-glicanos unida mediante enlaces  $\alpha 2,3$ . Las seis ST3Gal transferasas conocidas se expresan en células CHO (Fig. 4) y las funciones de cada una de ellas en la sialilación de los N-glicanos no se conocen bien, pero las candidatas son principalmente las isoformas ST3Gal-III, ST3Gal-IV y ST3Gal-VI. Primero nos enfocamos en estos tres genes individualmente usando inmunocitología para evaluar la exposición de LacNAc y descubrimos que solo la KO de st3gal4 produjo efectos sustanciales (Fig. 7). El análisis de la EPO expresada en clones KO simples y apilados mostró que la pérdida de st3gal4 y 6 produjo efectos sustanciales, mientras que encontramos dos patrones en clones con KO de st3gal4/6. En el fondo WT, KO de st3gal4/6 produjo EPO con N-glicanos tetraantennarios heterogéneos sin rastros de sialilación (Fig. 1f). El KO de st3gal4/6 en combinación con mgat4A/4B/5 produjo N-glicanos biantennarios sin sialilación pero con aumento de poli-LacNAc (Fig. 2b), o en un clon sin aumento de poli-LacNAc pero con cantidades menores de un N-glicano biantennario mono-sialilado. Por lo tanto, es posible que sea necesario apuntar a st3gal3/4/6 para lograr una ausencia total de ácido siálico. A pesar de ello, los clones CHO sin st3gal4 y 6 produjeron EPO con terminales LacNAc sin tapón, lo que abre la puerta a la ingeniería de novo de glicoproteínas recombinantes con tapón  $\alpha 2,6$ NeuAc.

#### Glicoingeniería personalizada de CHO para producir N-glicanos protegidos con $\alpha 2,6$ NeuAc

La mayoría de las glicoproteínas en circulación están cubiertas con  $\alpha 2,6$ NeuAc, mientras que prácticamente todas las terapias recombinantes se producen con  $\alpha 2,3$ NeuAc. Desde hace mucho tiempo ha sido de interés producir células CHO capaces de producir N-glicanos homogéneos con protección  $\alpha 2,6$ NeuAc. Para demostrar las perspectivas de glicoingeniería de novo proporcionadas por nuestro análisis KO, generamos clones CHO capaces de producir una protección homogénea de  $\alpha 2,6$ NeuAc (Fig. 2). Para evitar problemas anteriores con la integración aleatoria de plásmidos, usamos el knockin (KI) dirigido por ZFN de ST6Gal-I con una estrategia Obligare modificada en un sitio de integración SafeHarbor. La ST6GAL1 humana introducida en CHO con KO de st3gal4/6 produjo la gama completa de estructuras antenarias de N-glicanos con un grado normal de recubrimiento NeuAc (Fig. 2a), y cuando se introdujo en CHO con KO adicional de mgat4A/4B/5 se produjo EPO con N-glicanos biantennarios homogéneos recubiertos por  $\alpha 2,6$ NeuAc (Fig. 2b). Curiosamente, la introducción de novo de ST6GAL1 también abrogó las cantidades menores de poli-LacNAc formadas en estructuras biantennarias cuando se eliminó la sialilación (Fig. 2b), demostrando que la KO de B3gnt2 también era innecesaria. Así, pudimos introducir la capacidad de sialilación de  $\alpha 2,6$ NeuAc y producir una célula CHO capaz de producir N-glicoproteínas con un recubrimiento homogéneo de  $\alpha 2,6$ NeuAc. La estrategia Obligare KI fue altamente eficiente con aproximadamente 30-50% de células clonadas individuales que expresaban ST6Gal-I según lo evaluado por inmunocitología de anticuerpos y lectina (Fig. 7), mientras que confiar en la recombinación homóloga mediada por ZFN no produjo una integración dirigida exitosa en células CHO. Los clones ST6GAL1 KI fueron estables durante más de 20 pasajes según se evaluó mediante inmunocitología con lectina SNA y MAB 1B2 (no se muestra).

En resumen, la prueba KO realizada identifica los glucógenos clave que controlan los pasos decisivos en la N-glicosilación de proteínas en CHO (Fig. 3), y demuestra una notable plasticidad en la tolerancia a la glicosilación sin evidencia de cambios compensatorios inesperados en la glicosilación. Proporcionamos estrategias de diseño para la generación de células huésped CHO con N-glicanos biantennarios homogéneos con y sin sialilación que servirán como plataformas para la creación de novo de las capacidades de glicosilación deseadas. La matriz proporciona una guía general para el diseño personalizado de la capacidad de N-glicosilación de CHO, que probablemente pueda usarse para cualquier línea huésped de CHO e incluso para líneas de producción establecidas. Demostramos la viabilidad de la ingeniería de novo de capacidades de glicosilación homogénea introduciendo la  $\alpha 2,6$ sialilación en CHO mediante una estrategia KI precisa en líneas CHO con KO de genes  $\alpha 2,3$ sialiltransferasa endógenos. El panel de células CHO generado permitirá la disección y el diseño de novo de la N-glicosilación para cualquier terapia proteica, y la producción de una amplia gama de glicoformas de proteínas terapéuticas permitirá una evaluación sencilla y la selección del diseño óptimo. El advenimiento de la ingeniería genética precisa claramente promete una nueva era para el mayor productor de terapias con glicoproteínas de la actualidad: la célula CHO.

#### Métodos

Orientación de ZFN a glucógenos en células CHO. Toda la focalización genética se realizó en el clon CHOZN GS/- (glutamina sintasa) producido por ZFN KO y proporcionado por Sigma-Aldrich, St. Louis, MO, o en CHO-K1 obtenido de ATCC. Todos los medios para CHO, suplementos y otros reactivos usados se obtuvieron de Sigma-Aldrich, a menos que se especifique lo contrario. Las células CHO GS se mantuvieron como cultivos en

suspensión en medio sin suero EX-CELL CHO CD Fusion, suplementado con 4 mM de L-glutamina. Las células se sembraron a  $0,5 \times 10^6$  células/mL en un matraz T25 (NUNC, Dinamarca) un día antes de la transfección. Se usaron  $2 \times 10^6$  células y 2 µg de ADN plasmídico libre de endotoxinas de cada ZFN (Sigma, EE. UU.) para la transfección. Cada ZFN fue etiquetado con GFP y Crimson mediante un enlazador 2A. Las transfecciones se realizaron mediante electroporación usando el kit Amaxa V y el programa U24 con Amaxa Nucleofector 2B (Lonza, Suiza). Posteriormente, las células electroporadas se colocaron en 3 mL de medio de crecimiento en una placa de 6 pocillos. Las células se trasladaron a 30 °C para un choque de frío de 24 horas. 72 horas después de la nucleofección, la reserva de células con el 10-15 % más marcado tanto para GFP como para Crimson se enriqueció mediante FACS. Las células se volvieron a clasificar por FACS una semana después para obtener clones individuales en placas de 96 pocillos de fondo redondo. Los clones KO se identificaron mediante el análisis de delección por inserción (IDAA) descrito recientemente, así como, cuando fue posible, mediante inmunocitología con lectinas y anticuerpos monoclonales apropiados. Los clones seleccionados fueron verificados adicionalmente por TOPO para la caracterización detallada de las mutaciones introducidas. La estrategia permitió una rápida selección y detección de clones KO con mutaciones de cambio de marco y, en promedio, seleccionamos entre 2 y 5 clones de cada evento objetivo. El análisis de ARNsec de los clones CHO fue realizado por BGI.

Expresión recombinante de EPO humana en células CHO. Genewiz, EE. UU., sintetizó un constructo de expresión que contiene la secuencia codificante completa de EPO humana clonada en pcDNA3.1/myc-His (etiquetas C-terminales). La EPO es una proteína de 166 aminoácidos con N-glicanos en Asn24, Asn38 y Asn83. Las células CHO se cotransfectaron con plásmidos EPO y GFP, seguido de un FACS para enriquecer los clones que expresaban GFP. Los transfectantes estables se seleccionaron en 0,4 mg/mL de Zeocin (Invitrogen). Los clones estables positivos se examinaron mediante ensayos inmunoabsorbentes ligados a enzimas (ELISA) directos usando el anticuerpo monoclonal anti-His (C-Term)-HRP (Invitrogen). La EPO humana recombinante marcada con His se purificó mediante purificación por afinidad de níquel (Invitrogen, EE. UU.). Los medios se mezclaron 3:1 (v/v) en tampón de unión 4x (200 mM Tris, pH8,0, 1,2 M NaCl) y se aplicaron a 0,3 ml de agarosa NiNTA empacada (Invitrogen), preequilibrada en tampón de unión (50 mM Tris, pH 8,0, 300 mM NaCl). La columna se lavó con tampón de unión y, a continuación, la proteína unida se eluyó con tampón de unión con 250 mM adicionales de imidazol. Las fracciones que contienen EPO se determinaron mediante SDS-PAGE y se purificaron adicionalmente en una purificación por HPLC de fase inversa con una columna Jupiter C4 (5 µm, 300 Å, columna 250 x 4,6 mm) (Phenomenex), usando ácido trifluoroacético (TFA) al 0,1 % y un gradiente de acetonitrilo al 10-100 %. La pureza se evaluó mediante tinción de Coomassie de geles SDS-PAGE y las proteínas se cuantificaron mediante el kit de ensayo de proteínas BCA (Thermo Scientific).

Perfil de N-glicano de EPO recombinante en clones CHO. Aproximadamente 25 µg de EPO purificada se digirieron con 2 U de PNGasa F (Roche Diagnostics, Mannheim, Alemania) en 50 mM de bicarbonato amónico a 37°C durante toda la noche. Los N-glicanos liberados se separaron de la rhEPO usando Stagetips empacados internamente (Empore disk-C18, 3M) y se incubaron en acetato de amonio 100 mM, pH 5,0 durante 60 minutos a 22 °C. Los N-glicanos se secaron en viales de vidrio y se permitilaron como se describió anteriormente. Los N-glicanos permitilados se extrajeron con cloroformo y se desalinizaron lavando la fase orgánica repetidamente con agua desionizada. La fase orgánica se evaporó sobre una corriente de gas nitrógeno y los N-glicanos permitilados se disolvieron en 20 µL de metanol al 50%. Finalmente, 1 µL de muestra se cocrystalizó con 1 µL de matriz compuesta de ácido 2,5-dihidroxi-benzoico (10 mg/mL) en acetonitrilo al 70% (v/v), ácido trifluoroacético al 0,1% (v/v) y acetato de sodio 2,5 mM. Los N-glicanos permitilados se analizaron mediante el modo positivo MALDI-TOF (Autoflex Speed, BrukerDaltronics, Bremen) operado en el modo reflector con adquisición de datos en el rango de masas m/z de 1000 a 7000.

## Ejemplo 2

Determinación del repertorio de glicosiltransferasas expresadas en una línea celular CHO de mamífero.

CHO-K1 se obtuvo de ATCC y CHO-GS (clon CHOZN GS<sup>-/-</sup> (glutamina sintasa) producido por ZFN KO) se obtuvo de Sigma-Aldrich, St. Louis, MO. Todos los medios para CHO, suplementos y otros reactivos usados se obtuvieron de Sigma-Aldrich, a menos que se especifique lo contrario. Los CHO-GS ellos se mantuvieron como cultivos en suspensión en medios EX-CELL CHO CD Fusion libres de suero y componentes animales, suplementados con 4mM de L-glutamina. Las células CHO se sembraron a  $0,25 \times 10^6$  células/mL en una placa de 6 pocillos y se recolectaron en fase exponencial 48 h después de la inoculación para la extracción de ARN total con el minikit RNeasy (Qiagen). La integridad y la calidad del ARN se comprobaron mediante el Bioanalizador 2100 (Agilent Technologies). La construcción de la biblioteca y la secuenciación de próxima generación se realizaron usando el sistema Illumina HiSeq 2000 (Illumina, EE. UU.) en condiciones estándar recomendadas por el proveedor de servicios ARNsec. Los datos alineados se usaron para calcular la distribución de las lecturas en los genes de referencia CHO y se realizó un análisis de cobertura. Sólo los resultados de alineación que pasaron el control de calidad se usaron para el análisis de expresión genética posterior.

Se realizó un análisis de ARNsec en dos líneas CHO comunes (CHO-K1, CHO-GS) y dos clones triples *mgat4a/4b/5* KO CHO-GS independientes (ZFN91-1C8, ZFN91-2A6) (TABLA 3). El análisis de ARNsec informado de CHO-K1 se incluyó como referencia y fue en gran medida similar a nuestros análisis, excepto que los niveles de expresión relativa fueron más altos. Cabe destacar que algunos genes, como *galnt1* y *mgat4b*, que no se expresaban en CHO-K1 (Xu, Nagarajan et al. 2011), sí lo hacían. Además, en el caso de *mgat4b* se descubrió que este gen era esencial para los experimentos de glicoingeniería llevados a cabo aquí.

Una observación importante del análisis de los transcriptomas de CHO fue que los perfiles de expresión de los glucógenos en los análisis de los dos clones triples KO fueron idénticos a los de las líneas CHO parentales, lo que demuestra que la edición del gen ko dirigida realizada en las líneas celulares CHO no alteró la expresión de otros genes de glicosiltransferasa.

#### Ejemplo 3

##### Inactivación genética de genes de glicosiltransferasa en la vía de N-glicosilación (Deconstrucción).

Toda la inactivación dirigida del gen de la glicosiltransferasa se realizó en CHO-GS y/o en CHO-K1 y las células se cultivaron como se describe en el Ejemplo 2. Las células se sembraron a  $0,5 \times 10^6$  células/mL en un matraz T25 (NUNC, Dinamarca) un día antes de la transfección. Se usaron  $2 \times 10^6$  células y 2 µg de ADN plasmídico libre de endotoxinas de cada ZFN (Sigma, EE. UU.) para la transfección. Cada ZFN fue marcado con GFP y Crimson mediante un enlazador 2A (Duda, Lonowski et al. 2014). Las transfecciones se realizaron mediante electroporación usando el kit Amaxa V y el programa U24 con Amaxa Nucleofector 2B (Lonza, Suiza). Posteriormente, las células electroporadas se sembraron en 3 mL de medio de cultivo en una placa de 6 pocillos. Las células se trasladaron a 30°C para un choque frío de 24 horas. 72 h después de la nucleofección, el 10-15% de la reserva de células más marcadas con GFP y Crimson se enriquecieron mediante FACS. Utilizamos métodos desarrollados recientemente para enriquecer clones KO mediante FACS (ZFN marcados con GFP/Crimson) (Duda, Lonowski et al. 2014) y una detección de alto rendimiento mediante una estrategia de etiquetado de amplicones (IDAA) (Zhang, Steentoft et al. enviado). Las células se volvieron a clasificar por FACS una semana después para obtener clones individuales en placas de 96 pocillos de fondo redondo. Los clones KO se identificaron mediante IDAA y, cuando fue posible, también mediante inmunocitología con lectinas y anticuerpos monoclonales apropiados. Los clones seleccionados se verificaron posteriormente mediante la clonación TOPO de productos de PCR (Invitrogen, EE. UU.) para la caracterización detallada por secuenciación Sanger de las mutaciones introducidas en clones TOPO individuales. La estrategia permitió una rápida selección y detección de clones KO con mutaciones de inactivación apropiadas como se describe aquí, y en promedio seleccionamos entre 2 y 5 clones de cada evento objetivo.

La mayoría de los clones KO presentaban inserciones y/o deleciones fuera de marco (indels) en el rango de  $\pm 20$ bps, y la mayoría de los genes diana estaban presentes con dos alelos, mientras que algunos (*mgat4B* y *mgat5*) estaban presentes con 1 o 3 alelos, respectivamente (TABLA 4).

La TABLA 4 enumera todos los eventos de inactivación del gen de la glicosiltransferasa individuales y apilados y el orden (línea de ascendencia) de apilamiento de la inactivación del gen de la glicosiltransferasa en las células para producir una matriz de deconstrucción de la vía de N-glicosilación.

#### Ejemplo 4

##### Inserción de genes de glicosiltransferasas en la vía de la N-glicosilación (Reconstrucción).

La integración específica del objetivo (knockin/KI) se dirigió hacia el locus CHO Safe Harbor #1 (S.Bahr et al., BMC Proceedings 2013, 7(Suppl 6):P3.doi:10.1186/1753-6561-7-S6-P3). Se utilizó una estrategia ObLiGaRe modificada (33) donde dos sitios de unión ZFN invertidos flanquean el gen del plásmido donante de interés que se va a insertar. En primer lugar se sintetizó un vector lanzadera denominado ObLiGaRe-2X-Ins (TABLA 5). ObLiGaRe-2X-Ins se diseñó de tal manera que cualquier secuencia de ADNc que codifique una proteína de interés puede insertarse en un sitio de clonación múltiple en el que la transcripción está dirigida por el promotor IE del CMV y un terminador BgH. Para minimizar el silenciamiento epigenético, se insertaron dos elementos aislantes que flanquean la unidad de transcripción. Además, se incluyó una "plataforma de aterrizaje" AAVS1 en el extremo 3' de ObLiGaRe-2X-Ins justo aguas arriba del sitio de unión ZFN invertido 3'. Se insertó un marco de lectura abierto ST6GAL1 completo direccionalmente en ObLiGaRe-2X-Ins generando ObLiGaRe-2X-Ins-ST6GAL1 (TABLA 5). La transfección y la clasificación de los clones se realizaron como se describe en los EJEMPLOS 2 y 3. Los clones se examinaron inicialmente mediante tinción positiva con lectina SNA y los clones seleccionados se analizaron posteriormente mediante PCR de unión 5' y 3' para confirmar la correcta integración dirigida en el sitio Safeharbor#1 de la célula CHO. *st3gal4/6* KO. El número de copias alélicas de integración se determinó mediante PCR alélica WT. Posteriormente, el marco de lectura abierto MGAT4A humano completo se insertó direccionalmente en el segundo alelo de Safe Harbor#1 en el clon ST6GAL1 KI, seguido por la inserción de MGAT5 humano completo en la "plataforma de aterrizaje" AAVS1, que se incluyó en el extremo 3' de ObLiGaRe-2X-Ins justo antes del sitio de unión ZFN invertido 3' (TABLA 5).

## Ejemplo 5

Producción recombinante de EPO e IgG humanas en células CHO de mamíferos editadas genéticamente.

- 5 Los constructos de expresión que contienen la secuencia codificante completa de EPO humana y una IgG terapéutica clonada en pcDNA3.1/myc-His (etiquetas C-terminales) y pBUDCE4.1 (Invitrogen), respectivamente, fueron sintetizados por Genewiz, EE. UU. La EPO es una proteína de 166 aminoácidos con N-glicanos en Asn24, Asn38 y Asn83, y la IgG tiene un único N-glicano en Asn297, respectivamente. Las  
10 células CHO Wt o KO fueron transfectadas con EPO o IgG y mantenidas a 37 °C como cultivos en suspensión en medio EX-CELL CHO CD Fusion libre de suero, suplementado con 4mM de L-glutamina y 400ug/mL de Zeocin en incubadora con 5% de CO<sub>2</sub>. Los clones se ampliaron y se cultivaron en matraces T75 a una densidad de siembra de 0,5×10<sup>6</sup> células/mL y se cosecharon 72 o 96 horas después. La viabilidad celular fue superior al 90% en todos los clones que expresaban reporteros y la densidad de células viables estuvo entre 2 y 3×10<sup>6</sup>  
15 células/mL en el momento de la cosecha. Las células se eliminaron mediante centrifugación a 300 g durante 5 minutos y el sobrenadante se almacenó a -80 °C. La EPO humana recombinante marcada con His se purificó mediante purificación por afinidad de níquel (Invitrogen, EE. UU.). Los medios se mezclaron 3:1 (v/v) en tampón de unión 4x (200 mM Tris, pH 8,0, 1,2 M NaCl) y se aplicaron a 0,3 mL de agarosa NINTA empaquetada (Invitrogen), preequilibrada en tampón de unión (50 mM Tris, pH 8,0, 300 mM NaCl). La columna se lavó con  
20 tampón de unión y, a continuación, la proteína unida se eluyó con tampón de unión con 250 mM adicionales de imidazol. Las fracciones que contienen EPO se determinaron mediante SDS-PAGE y se purificaron adicionalmente en una purificación por HPLC de fase inversa con una columna Jupiter C4 (5 µm, 300 Å, columna 250 x 4,6 mm) (Phenomenex), usando ácido trifluoroacético (TFA) al 0,1 % y un gradiente de acetonitrilo al 10-100 %. La IgG se purificó mediante HiTrap™ Protein G HP (GE Healthcare, EE. UU.)  
25 preequilibrada/lavada en PBS y eluida con 0,1 M de glicina (pH 2,7). La pureza de las proteínas se evaluó mediante la tinción de Coomassie de los geles de SDS-PAGE, y las proteínas se cuantificaron mediante el kit de ensayo de proteínas BCA (Thermo Scientific, Rockford, EE. UU.).

## Ejemplo 6

Perfil de N-glicano de EPO e IgG purificadas.

- Aproximadamente 25 µg de EPO purificada se digirieron con 2 U de PNGasa F (Roche Diagnostics, Mannheim, Alemania) en 50 mM de bicarbonato amónico a 37°C durante toda la noche. Para la IgG, 35 µg de muestras  
35 se digirieron inicialmente (durante la noche, 37°C) con tripsina en bicarbonato de amonio 50 mM, y después se calentaron a 95°C durante 15 min y se enfriaron a RT antes de la digestión con PNGasa F como se ha indicado anteriormente. Los N-glicanos liberados se separaron de la EPO intacta o del digerido trípico de IgG usando Stagetips envasados en la propia empresa (disco Empore-C18, 3M) y se incubaron en acetato de amonio 100 mM, pH 5,0 durante 60 min a 22°C. Los N-glicanos se dividieron en tres alícuotas iguales. Un tercio  
40 se secó en viales de vidrio y se permitió como se describió anteriormente (Ciucanu y Costello 2003) y el resto se guardó para el análisis de N-glicano nativo. Los N-glicanos permitidos se extrajeron con cloroformo y se desalinizaron lavando la fase orgánica repetidamente con agua desionizada. La fase orgánica se evaporó sobre una corriente de gas nitrógeno y los N-glicanos permitidos se disolvieron en 20 µL de metanol al 50% (v/v). Finalmente, 1 µL de muestra se cocristalizó con 1 µL de matriz compuesta de ácido 2,5-dihidroxibenzoico (10  
45 mg/mL) en acetonitrilo al 70% (v/v), ácido trifluoroacético al 0,1% (v/v) y acetato de sodio 2,5 mM. Los N-glicanos permitidos se analizaron mediante MALDI-TOF en modo positivo (Autoflex Speed, BrukerDaltonics, Bremen) operado en modo reflector con adquisición de datos (2000 disparos/punto) en el rango de masas *m/z* 1000-7000. El análisis de la composición de glicanos de todos los espectros se realizó mediante la plataforma SysBioWare y se presenta en la Tabla complementaria 4 (Vakhrushev, Dadimov et al. 2009). Los N-glicanos  
50 nativos liberados se reconstituyeron en MeOH/H<sub>2</sub>O (1/1; v/v, que contiene 50 mM de bicarbonato de amonio) a una concentración de 0,5 pmol/µL y se analizaron en un OrbiTrap Fusion MS (Thermo San Hose, EE. UU.). Las muestras se infundieron directamente a través de EASY-Spray y se analizaron mediante modo de iones negativos. El caudal se ajustó a 500 nL/min, el voltaje de pulverización a 1900-2100 V y la temperatura del tubo de transferencia de iones a 275 °C. Se seleccionaron iones precursores (ancho de aislamiento, *m/z* 3) y se sometieron a fragmentación HCD al 35 % de energía de colisión normalizada (estado de carga predeterminado, *z*=2). Todos los espectros se registraron con una resolución de 120.000.

## Ejemplo 7

- 60 Inactivación de genes de glicosiltransferasa mediante edición genética multiexónica precisa con oligodesoxinucleótido monocatenario (ssODN)/ZFN en tándem.

- En este ejemplo, la inactivación del gen se realiza eliminando los exones que codifican la cola citoplasmática N-terminal, la región transmembrana y partes de la región madre del gen de la glucosiltransferasa diana. La  
65 eliminación de la región objetivo está mediada por el uso del enfoque de oligo desoxirribonucleótido monocatenario en tándem (ssODN)/ZFN (Duda, Lonowski et al. 2014). El sitio objetivo de ZFN se selecciona

en o lo más cerca posible de la secuencia que codifica la región transmembrana de la glicosiltransferasa. Como ejemplo, se selecciona CHO *mgat4a* usando ZFN ComposR de CHO *mgat4a* (Sigma) y un ssODN que posee secuencias de brazo de homología de 60bp que flanquean la región UTR 5', incluido el sitio de inicio de la traducción y el sitio diana ZFN del exón 2 de *mgat4a* que flanquea 3' (ssODN *mgata*

- 5 :gagagagattggtctgtgtcatcaccaacgtatgaaccagtgtgatggtgaaatgagtcGCCGTGCAGGA  
 GCAGCAACTAACGGAAGTAACACTGCATTGACTACATTTTCAGGTAC) (la región UTR 5' se muestra en minúsculas, el sitio de corte ZFN parcial en minúsculas cursivas y el sitio de unión ZFN de los exones 2 que flanquean 3' en mayúsculas), eliminando así la cola citoplasmática que codifica el exón 1-2, el dominio transmembrana y el tallo de *mgat4a*. La inactivación genética se realiza en CHO-GS como se describe en el Ejemplo 2, incluido *mgata* ssODN. Las transfecciones se realizaron mediante electroporación usando el kit Amaxa V y el programa U24 con Amaxa Nucleofector 2B (Lonza, Suiza). Posteriormente, las células electroporadas se colocaron en 3 mL de medio de crecimiento en una placa de 6 pocillos. Las células se trasladaron a 30°C para un choque frío de 24 horas. 72 h después de la nucleofección, el 10-15% de la reserva de células más marcadas con GFP y Crimson se enriquecieron mediante FACS. Las células se volvieron a clasificar por FACS una semana después para obtener clones individuales en placas de 96 pocillos de fondo redondo. El enriquecimiento de clones KO se realizó mediante FACS (ZFN marcadas con GFP/Crimson). Los clones se analizan mediante PCR usando los cebadores Mgataex1F (5'-TATCCACTGTGTTGCTTGCTG-3')/Mgataex2R (5'-Actgctctccagaggtcctg d-3'), detectándose únicamente los clones correctamente eliminados. La orientación mono/bialélica se determinó mediante detección de presencia de alelos wt mediante PCR usando los cebadores Mgataex2F (5'-gaacgcctcgaatagctgaacatagg-3')/Mgataex2R. Los resultados de PCR genética se validan mediante análisis Southern blot usando una sonda específica del intrón 2 que detecta la eliminación correcta de la región objetivo del intrón 1 de 37 KBp que contiene sitios KpnI de diagnóstico.

#### Ejemplo 8

- 25 Inactivación de genes de glicosiltransferasa mediante edición genética dual CRISPR/Cas9 de múltiples exones.

- 30 La inactivación genética se asegura mediante la eliminación de los exones que codifican la cola citoplasmática/transmembrana y partes de las secuencias codificadas del tallo del gen deseado que se necesita inactivar. La eliminación de la región objetivo está mediada por el uso de un par de CRISPR/Cas9-gARN dirigidos a las regiones flanqueantes que codifican la secuencia señal *Mgat4a*, el anclaje transmembrana y las regiones madre. CHO *Mgat4a* se dirige usando *Mgat4aEx1gARN* de CHO *Mgat4a* CRISPR/Cas9 gARN (5'-GGTATACCACATGGCAAAATGGG-3') específico para el exón1 y *Mgat4aEx2gARN* (5'-GTCCACAGTTTCGCCGTGCAGG-3') específico para el exón2. Ambos ARNg se clonaron en el sitio diana de ARNg BbsI (NEB, EE. UU.) de px458 (Addgene, EE. UU.) que codifica el Cas9 marcado con GFP de *S.pyogenes* y el casete de ARNg impulsado por el promotor U6, generando px458-CRISPR-*Mgat4aex1gRNA* y px458-CRISPR-*Mgat4aex2gRNA*. Se realizó una edición genética precisa mediante nucleofección usando 2 ug de estas herramientas de edición CRISPR/Cas9-*Mgat4a* y validación de selección de objetivos en células CHO-GS como se describe en los Ejemplos 3 y 7.

#### Ejemplo 9

- 45 **GlicoPEGilación enzimática de glicoproteínas N-glicanos monoantennarios.** La KO de *mgat2* en combinación con *mgat4a/4b/5* en células CHO dio como resultado N-glicanos monoantennarios homogéneos cuando se expresó EPO (FIG. 32). La reintroducción de MGAT4A humano así como de MGAT4A y 5 en combinación por knockin dio como resultado N-glicanos tri y tetraantennarios, respectivamente (FIG 32), lo que demuestra que el N-glicano monoantennario obtenido está vinculado  $\beta$ 1-2 a la rama  $\alpha$ 1-3Man controlada por *mgatl*. Se probó la glicopegilación enzimática o la transferencia de otros compuestos por sialiltransferasas y CMP-(PEG)NeuAc o un donante modificado relevante usando EPO como ejemplo (FIGURA 33-34).

- 55 Material y métodos Las glicovariantes de EPO-Myc-His6 se produjeron en CHO con KO *mgat2/4a/4b/5* y se purificaron en una columna de níquel-NTA y el tampón se intercambió en un tampón Tris 100 mM, NaCl 150 mM, Na<sub>3</sub> al 0,005 %, pH 7,2, como se describe en el Ejemplo 4. El EPO-Myc-His6 de tipo salvaje con perfil de glicosilación normal se produjo en células CHO no modificadas y se purificó de manera similar en una columna de níquel-NTA antes del intercambio del tampón en un tampón Tris 100 mM, NaCl 150 mM, Na<sub>3</sub> al 0,005 %, pH 7,2. Las concentraciones de proteínas se determinaron mediante el espectrofotómetro NanoDrop Lite (Thermo Scientific™). La sialiltransferasa (ST3Gal-III; 1,35 U/mL) en 20 mM HEPES, 120 mM NaCl, 50 % glicerol, pH 7,0 se obtuvo de FujiFilm Diosynth Biotechnologies, Japón. La sialidasa (*Artherobacter urifaciens*) se produjo en Novo Nordisk A/S y se utilizó como una solución de 200 U/mL en 20 mM HEPES, 120 mM NaCl, 50 % glicerol, pH 7,0. La neuraminidasa unida mediante soporte (*Clostridium perfringens*) en agarosa se obtuvo de Sigma (Cat. # N5254-10UN). La resina se lavó exhaustivamente en agua MilliQ antes de su uso. El 10k-PSC (citidina-5'-monofosforil-[5-(N-10kDa-metoxipolioxietileno-oxicarboxamido)glicilamido-3,5-didesoxi-D-glicero-n-galacto-2-nonulopiranosuronato], lote: 3605-A-R0-01-61-1) se obtuvo de Albany Molecular Research, Inc, Albany, Nueva York 12203, EE. UU. El 10k-PSC se produjo como se describe en el documento

WO03031464. El tampón tris usado en los ejemplos siguientes contenía 50 mM de tris-HCl, 150 mM de NaCl, pH 7,6.

La electroforesis en gel SDS-PAGE se realizó en geles prefabricados NuPage (NuPage 7 % Tris-Acetate Gel, 15 pocillos, Invitrogen cat. No. EA03555BOX) usando tampón de ejecución NuPage Tris-Acetate SDS (Invitrogen cat. No. # LA0041). A las muestras (0,15 mg/mL) se les añadió agente reductor de muestras NuPage (Invitrogen cat. No. NP0004) y tampón de muestra LDS (Invitrogen cat. No. NP0008) y se calentaron a 70 °C durante 10 minutos antes de cargar el gel. Se utilizó HiMark HMW (Invitrogen cat. No. LC6060) como estándar. La electroforesis se realizó a 150 V, 120 mA y 25 W durante 60 min. Los geles se tiñeron con SimplyBlue SafeStain (Invitrogen cat. No. LC5688) de acuerdo con los protocolos de Invitrogen. Los geles teñidos con Coomassie se analizaron mediante densitometría usando el software GelQuant y los métodos descritos previamente (Rehbein y Schwalbe 2015) (FIG 34).

El análisis por HPLC de los conjugados de EPO-Myc-His6 glicoPEGilados de la invención se realizó de la siguiente manera. Las muestras se analizaron en forma no reducida. Se utilizó una columna Zorbax 300SB-C3 (4,6 x 50 mm; 3,5 µm Agilent, Cat. No.: 865973-909). La columna se operó a 30 °C. Se inyectaron 2,5 µg de muestra y la columna se eluyó con un sistema de disolventes agua (A) - acetonitrilo (B) que contenía 0,1 % de ácido trifluoroacético. El programa de gradiente fue el siguiente: 0 min (25% B); 4 min (25% B); 14 min (46% B); 35 min (52% B); 40 min (90% B); 40,1 min (25% B).

La glicoPEGilación enzimática de las variantes EPO-Myc-His6 se llevó a cabo mediante un método de un solo paso usando sialidasa en combinación con ST3Gal-III, o mediante un método secuencial en el que la variante EPO-Myc-His6 se hizo reaccionar primero con gel de agarosa-neuraminidasa durante 3 h a 25 °C, y después de transferir el sobrenadante a un vial nuevo se hizo reaccionar con el reactivo ST3Gal-III y PSC. El método de una sola olla se basa en el hecho de que la sialidasa de *Artherobacter urifaciens* no puede desialilar los sialósidos PEGilados. El uso de glicoproteínas recombinantes con N-glicanos monoantennarios sin protección de ácido siálico tal como se inventó elimina la necesidad de desialilación previa.

GlicoPEGilación de eritropoyetina humana recombinante: la EPO-Myc-His6 de tipo salvaje recombinante producida en líneas celulares CHO se disolvió en 100 mM de Tris, 150 mM de NaCl, 0,005 % de NaN<sub>3</sub> (pH 7,2) a una concentración de 1,17 mg/mL. El reactivo 10k-PSC (2,0 mg) se disolvió en 100 µL de Tris-HCl 50 mM, NaCl 150 mM (pH 7,6) a una concentración de 20 mg/mL. Tanto ST3Gal-III como AUS se disolvieron en 20 mM HEPES, 120 mM NaCl, 50 % glicerol, pH 7,0 a una concentración volumétrica de 1,35 U/mL y 200 U/mL respectivamente. Los reactivos se mezclaron según el esquema siguiente:

|                  | Unidad         | Exp.1 | Exp.2 | Exp.3 | Exp.4 | Exp.5 |
|------------------|----------------|-------|-------|-------|-------|-------|
| peso EPO         | ug             | 10    | 10    | 10    | 10    | 10    |
| 10k-PSC          | Ec.            | 1     | 5     | 10    | 25    | 50    |
| ST3Ga13          | mU/ug peso EPO | 0,196 | 0,196 | 0,196 | 0,196 | 0,196 |
| Sialidasa        | U/mL           | 5     | 5     | 5     | 5     | 5     |
| Tampón Tris      | uL             | 9,0   | 8,4   | 7,6   | 5,3   | 1,3   |
| Volumen final    | uL             | 28,45 | 28,45 | 28,45 | 28,45 | 28,45 |
| Final [peso EPO] | mg/mL          | 0,35  | 0,35  | 0,35  | 0,35  | 0,35  |

Las reacciones se incubaron durante 18 h a 32 °C. Luego, las muestras se analizaron mediante HPLC y SDS-PAGE (FIG 33, carriles 5-9).

GlicoPEGilación de eritropoyetina producida en células CHO con KO de *mgat2/4A/4B/5* - La variante de glicosilación wtEPO-Myc-His6 recombinante producida en la línea celular CHO knockout *Mgat2/4a/4b/5* se disolvió en 100 mM de Tris, 150 mM de NaCl, pH 7,2 a una concentración de 0,585 mg/mL. El reactivo 10k-PSC (2,0 mg) se disolvió en 100 µL de Tris-HCl 50 mM, NaCl 150 mM, pH 7,6 a una concentración de 20 mg/mL. Tanto ST3Gal-III como AUS se disolvieron en 20 mM HEPES, 120 mM NaCl, 50 % glicerol, pH 7,0 a una concentración volumétrica de 1,35 U/mL y 200 U/mL respectivamente. Los reactivos se mezclaron según el esquema siguiente:

|               | Unidad    | Exp.1 | Exp.2 | Exp.3 | Exp.4 | Exp.5 |
|---------------|-----------|-------|-------|-------|-------|-------|
| EPO-Myc-His6  | ug        | 10    | 10    | 10    | 10    | 10    |
| 10k-PSC       | Ec.       | 1     | 5     | 10    | 25    | 50    |
| ST3Gal3       | mU/ug EPO | 0,196 | 0,196 | 0,196 | 0,196 | 0,196 |
| Sialidasa     | U/mL      | 5     | 5     | 5     | 5     | 5     |
| Tampón Tris   | uL        | 9,0   | 8,4   | 7,6   | 5,3   | 1,3   |
| Volumen final | uL        | 28,45 | 28,45 | 28,45 | 28,45 | 28,45 |
| Final [EPO]   | mg/mL     | 0,35  | 0,35  | 0,35  | 0,35  | 0,35  |

## ES 3 003 873 T3

Las reacciones se incubaron durante 18 h a 32 °C. A continuación, se analizaron las muestras mediante HPLC y SDS-PAGE (FIG. 22, carriles 11-15).

- 5 La reacción de glicoPEGilación de dos pasos de la eritropoyetina humana recombinante - EPO-Myc-His6 (272 ug) en 100 mM Tris, 150 mM NaCl, 0,005 % NaN<sub>3</sub>, pH 7,2 (0,214 mg/mL) se trató con sialidasa inmovilizada prelavada (suspensión de 150 uL, 150 mU) durante 3 h a temperatura ambiente. Luego se eliminó la resina mediante filtración. El filtrado se analizó por LC-MS para confirmar la formación de asialo EPO-Myc-His6 en forma de N-glicano monoantenario (masa calculada 25464,97 Da, análisis 25465,30 Da).

| Cálculo de masa media | Encontrado  | Asignación                                                                                                   |
|-----------------------|-------------|--------------------------------------------------------------------------------------------------------------|
| 26921,23 Da           | 26921,53 Da | EPO-myc-His6 pentasialilado que contiene 3 N-glicanos monoantenarios y un O-glicano                          |
| 26629,98 Da           | 26630,34 Da | EPO-myc-His6 tetrasialilado que contiene 3 N-glicanos monoantenarios y un O-glicano                          |
| 26338,72 Da           | 26339,16 Da | EPO-myc-His6 trisialilado que contiene 3 N-glicanos monoantenarios y un O-glicano                            |
| 25464,97 Da           | 25465,30 Da | EPO-myc-His6 desialilado (tratado con neuraminidasa) que contiene 3 N-glicanos monoantenarios y un O-glicano |

10

Luego se añadió al filtrado reactivo 10k-PSC disuelto en 100 uL de Tris-HCl 50 mM, NaCl 150 mM, pH 7,6. Tanto ST3Gal-III como AUS se disolvieron en 20 mM HEPES, 120 mM NaCl, 50 % glicerol, pH 7,0 a una concentración volumétrica de 1,35 U/mL y 200 U/mL respectivamente. Los reactivos se mezclaron según el

15

esquema siguiente:

|                     | Unidad    | Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 | Exp. 5 |
|---------------------|-----------|--------|--------|--------|--------|--------|
| asialo EPO-Myc-His6 | ug        | 20     | 20     | 20     | 20     | 20     |
| 10k-PSC             | Ec.       | 1      | 3      | 5      | 10     | 20     |
| St3Gal3             | mU/ug EPO | 0,196  | 0,196  | 0,196  | 0,196  | 0,196  |
| Tampón Tris         | uL        | 14,0   | 12,5   | 11,0   | 7,4    | 0      |
| Volumen final       | uL        | 117,6  | 117,6  | 117,6  | 117,6  | 117,6  |
| EPO Final           | mg/mL     | 0,17   | 0,17   | 0,17   | 0,17   | 0,17   |

20

Las reacciones se incubaron durante 20 h a 32 °C. Luego, las muestras se analizaron mediante HPLC y SDS-PAGE.

Más tablas

Tabla 4. Diseño de ZFN y análisis de secuencia de clones CHO glicoingeniería.

25

| Clon                  | Insertar y retirar | Alineación                                                                                                   |
|-----------------------|--------------------|--------------------------------------------------------------------------------------------------------------|
| KO: <i>mgat4A</i>     | ZFN                | TTCTGAGTTGAATGCCATTGTCCAACAgtt-----                                                                          |
| <i>mgat4A-WT</i>      |                    | tcGCCGTGCAGGAGCAGCAACTAACGGAAG                                                                               |
| <i>mgat4A-allele1</i> | -5bp               | TTCTGAGTTGAATGCCATTGTCCAACAgtt-----                                                                          |
| <i>mgat4A-allele2</i> | +13bp              | CCGTGCAGGAGCAGCAACTAACGGAAG<br>TTCTGAGTTGAATGCCATTGTCCAACAgttGAATTCTAGATGAtcGCCGTGC<br>AGGAGCAGCAACTAACGGAAG |
| KO: <i>mgat4A/4B</i>  | ZFN                | TTCTGAGTTGAATGCCATTGTCCAACAgtt-----                                                                          |
| <i>mgat4A-WT</i>      |                    | tcGCCGTGCAGGAGCAGCAACTAACGGAAG                                                                               |
| <i>mgat4A-allele1</i> | -5bp               | TTCTGAGTTGAATGCCATTGTCCAACAgtt-----                                                                          |
| <i>mgat4A-allele2</i> | +13bp              | CCGTGCAGGAGCAGCAACTAACGGAAG<br>TTCTGAGTTGAATGCCATTGTCCAACAgttGAATTCTAGATGAtcGCCGTGC<br>AGGAGCAGCAACTAACGGAAG |
| <i>mgat4B-WT</i>      | ZFN                | GCCCTCCAGCAGCCCTCTqaggaCTGGATGATCCTGGAGTT                                                                    |
| <i>mgat4B-allele1</i> | -7bp               | GCCCTCCAGCAGCCCTCTg-----GATGATCCTGGAGTT                                                                      |

|                                            |       |                                                      |
|--------------------------------------------|-------|------------------------------------------------------|
| <i>KO:mgat5</i><br><i>mgat5-WT</i>         | ZFN   | GGGGGATGATGCTTCTGCACTTCACCATCCAg----                 |
|                                            |       | cagcgGACTCAGCCTGAGAGCAGCTCCATGTT                     |
| <i>mgat5-alle1</i>                         | +14bp | GGGGGATGATGCTTCTGCACTTCACCATCCAgcaTGAATTCTAGATGAgcgG |
|                                            |       | ACTCAGCCTGAGAGCAGCTCCATGTT                           |
| <i>mgat5-alle2</i>                         | -1bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAg----                 |
|                                            |       | agcgGACTCAGCCTGAGAGCAGCTCCATGTT                      |
| <i>mgat5-alle3</i>                         | +61bp | GGGGGATGATGCTTCTGCACTTCACCATCCAgcag----- (+61bp) --  |
|                                            |       | cgGACTCAGCCTGAGAGCAGCTCCATGTT                        |
| <i>KO: mgat4A/4B/5</i><br><i>mgat4A-WT</i> | ZFN   | TTCTGAGTTGAATGCCATTGTCCAACAgtt-----                  |
|                                            |       | tcGCCGTGCAGGAGCAGCAACTAACGGAAG                       |
| <i>mgat4A-alle1</i>                        | -5bp  | TTCTGAGTTGAATGCCATTGTCCAACAg-----                    |
|                                            |       | CCGTGCAGGAGCAGCAACTAACGGAAG                          |
| <i>mgat4A-alle2</i>                        | +13bp | TTCTGAGTTGAATGCCATTGTCCAACAgttGAATTCTAGATGAtcGCCGTGC |
|                                            |       | AGGAGCAGCAACTAACGGAAG                                |
| <i>mgat4B-WT</i>                           | ZFN   | GCCCTCCAGCAGCCCTCTqaggaCTGGATGATCCTGGAGTT            |
| <i>mgat4B-alle1</i>                        | -7bp  | GCCCTCCAGCAGCCCTCTg-----GATGATCCTGGAGTT              |
| <i>mgat5-WT</i>                            | ZFN   | GGGGGATGATGCTTCTGCACTTCACCATCCAg----                 |
|                                            |       | cagcgGACTCAGCCTGAGAGCAGCTCCATGTT                     |
| <i>mgat5-alle1</i>                         | +4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAgcagcCAGCgGACTCAGCCTG |
|                                            |       | AGAGCAGCTCCATGTT                                     |
| <i>mgat5-alle2</i>                         | -4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAg----                 |
|                                            |       | gGACTCAGCCTGAGAGCAGCTCCATGTT                         |
| <i>mgat5-alle3</i>                         | -16bp | GGGGGATGATGCTTCTGCACTTC-----                         |
|                                            |       | CTCAGCCTGAGAGCAGCTCCATGTT                            |
| <i>KO:B3gnt1</i><br><i>B3gnt1-WT</i>       | ZFN   | TGCAGCTGCTCTACCTGTC-----gctgcTCTCCGGACTGCACG         |
| <i>B3gnt1-alle1</i>                        | +11bp | TGCAGCTGCTCTACCTGTGcgagcTGCTCTACCTGTCTCCCGGA         |
|                                            |       | CTGCACG                                              |
| <i>B3gnt1-alle2</i>                        | -5bp  | TGCAGCTGCTCTACCTGTC-----TCTCCGGACTGCACG              |
| <i>KO:B3gnt2</i><br><i>B3gnt2-WT</i>       | ZFN   | TTCAGCCCTTCCcgggcGTAAGTGGAAACAGAGAGCA                |
| <i>B3gnt2-alle1</i>                        | -1bp  | TTCAGCCCTTCC-gggcGTAAGTGGAAACAGAGAGCA                |
| <i>B3gnt2-alle2</i>                        | -4bp  | TTCAGCCCTTCCcg----TACTGGAACAGAGAGCA                  |
| <i>KO:B4gal tl</i><br><i>B4gal tl-WT</i>   | ZFN   | TGCATCCGGTCCTACAGCqccagc----AACTGGACTATGGTA          |
| <i>B4gal tl-alle1</i>                      | +4bp  | TGCATCCGGTCCTACAGCqccagcCAGCAACTGGACTATGGTA          |
| <i>KO:B4galt2</i><br><i>B4galt2-WT</i>     | ZFN   | CAGCCCCGCCACTTTqcc-----atcGCCATGGACAAGTTTGGCT        |
| <i>B4galt2-alle1</i>                       | +73bp | CAGCCCCGCCACTTTgcc--(+73bp)--                        |
|                                            |       | atcGCCATGGACAAGTTTGGCT                               |
| <i>KO:B4galt3</i><br><i>B4galt3-WT</i>     | ZFN   | CTAGCCCTCAAGTCAGGAtgt-----tqCGGAGGCTGCTGGAGAGG       |
| <i>B4galt3-alle1</i>                       | +5bp  | CTAGCCCTCAAGTCAGGAtgtCGTGTtgCGGAGGCTGCTGGAGAGG       |
| <i>B4galt3-alle2</i>                       | +2bp  | CTAGCCCTCAAGTCAGGAtgt---tgCCCGAGGCTGCTGGAGAGG        |
| <i>KO:B4galt4</i><br><i>B4galt4-WT</i>     | ZFN   | AACTGGGACTGCTTTat----attcCACGATGTGGACCTGGTG          |
| <i>B4galt4-alle1</i>                       | +1bp  | AACTGGGACTGCTTTat---TattcCACGATGTGGACCTGGTG          |
| <i>B4galt4-alle2</i>                       | +4bp  | AACTGGGACTGCTTTatattTATTcCACGATGTGGACCTGGTG          |

5

|                              |       |                                                                                   |
|------------------------------|-------|-----------------------------------------------------------------------------------|
| <i>KO:B4galt1/2</i>          |       |                                                                                   |
| <i>B4gal tl-WT</i>           | ZFN   | <u>TGCATCCGGTCCTACAGCqccagc</u> ---- <u>AACTGGACTATGGTA</u>                       |
| <i>B4gal tl-allele1</i>      | +4bp  | TGCATCCGGTCCTACAGCgccagcCAGCAACTGGACTATGGTA                                       |
| <i>B4galt2-WT</i>            | ZFN   | <u>CAGCCCCGCCACTTTqccatcGCCATGGACAAGTTTGGCT</u>                                   |
| <i>B4galt2-allele1</i>       | -8bp  | CAGCCCCGCCAC-----cGCCATGGACAAGTTTGGCT                                             |
| <i>B4galt2-allele2</i>       | -14bp | CAGCCC-----cGCCATGGACAAGTTTGGCT                                                   |
| <i>KO:st3Gal3</i>            |       |                                                                                   |
| <i>st3Gal3-WT</i>            | ZFN   | <u>CTCTCTCTTTGTCCCTTGCTggcttCAAATGGCAGGACTTCAAG</u>                               |
| <i>st3Gal3-allele1</i>       | -5bp  | CTCTCTCTTTGTCCCTTGCTt----CAAATGGCAGGACTTCAAG                                      |
| <i>st3Gal3-allele2</i>       | -1bp  | CTCTCTCTTTGTCCCTTGCTggcttCAAATGGCAGGACTTCAAG                                      |
| <i>KO:st3Gal4</i>            |       |                                                                                   |
| <i>st3Gal4-WT</i>            | ZFN   | <u>GGCAGCCTCCAGTGTCTGTCgttgt</u> ---- <u>qTTGTGGTGGGGAATGGGC</u>                  |
| <i>st3Gal4-allele1</i>       | +4bp  | GGCAGCCTCCAGTGTCTGTCgttgtTTGTgTTGTGGTGGGGAATGGGC                                  |
| <i>st3Gal4-allele2</i>       | -4bp  | GGCAGCCTCCAGTGTCTGTCg-----gTTGTGGTGGGGAATGGGC                                     |
| <i>KO:st3Gal6</i>            |       |                                                                                   |
| <i>st3Gal6-WT</i>            | ZFN   | <u>CGGTACCTCTGATTTTGCTttgccCTATGGACAAGGCC</u>                                     |
| <i>st3Gal6-allele1</i>       | -4bp  | CGGTACCTCTGATTTTGCTt----CTATGGGACAAGGCC                                           |
| <i>st3Gal6-allele2</i>       | -22bp | CGGTACCTCTGA-----AGGCC                                                            |
| <i>KO:st3Gal3/4</i>          |       |                                                                                   |
| <i>st3Gal3-WT</i>            | ZFN   | <u>CTCTCTCTTTGTCCCTTGCTggcttCAAATGGCAGGACTTCAAG</u>                               |
| <i>st3Gal3-allele1</i>       | -5bp  | CTCTCTCTTTGTCCCTTGCTt----CAAATGGCAGGACTTCAAG                                      |
| <i>st3Gal3-allele2</i>       | -1bp  | CTCTCTCTTTGTCCCTTGCTggcttCAAATGGCAGGACTTCAAG                                      |
| <i>st3Gal4-WT</i>            | ZFN   | <u>GGCAGCCTCCAGTGTCTGTCgttgt</u> ---- <u>gTTGTGGTGGGGAATGGGC</u>                  |
| <i>st3Gal4-allele1</i>       | -4bp  | GGCAGCCTCCAGTGTCTGTCg-----gTTGTGGTGGGGAATGGGC                                     |
| <i>st3Gal4-allele2</i>       | +5bp  | GGCAGCCTCCAGTGTCTGTCgACACGttgtgTTGTGGTGGGGAATGGGC                                 |
| <i>KO:st3Gal4/6</i>          |       |                                                                                   |
| <i>st3Gal4-WT</i>            | ZFN   | <u>GGCAGCCTCCAGTGTCTGTCgttgt</u> ---- <u>qTTGTGGTGGGGAATGGGC</u>                  |
| <i>st3Gal4-allele1</i>       | +4bp  | GGCAGCCTCCAGTGTCTGTCgttgtTTGTgTTGTGGTGGGGAATGGGC                                  |
| <i>st3Gal4-allele2</i>       | -4bp  | GGCAGCCTCCAGTGTCTGTCg-----gTTGTGGTGGGGAATGGGC                                     |
| <i>st3Gal6-WT</i>            | ZFN   | <u>CGGTACCTCTGATTTTGCT-ttgccCTATGGGACAAGGCC</u>                                   |
| <i>st3Gal6-allele1</i>       | +1bp  | CGGTACCTCTGATTTTGCTttgccCTATGGGACAAGGCC                                           |
| <i>st3Gal6-allele2</i>       | -7bp  | CGGTACCTCTGATTTTG-----CTATGGGACAAGGCC                                             |
| <i>KO:B3gnt2/mgat4A/4B/5</i> |       |                                                                                   |
| <i>mgat4A-WT</i>             | ZFN   | <u>TTCTGAGTTGAATGCCATTGTCCAACAgtt</u> -----<br>tcGCCGTGCAGGAGCAGCAACTAACGGAAG     |
| <i>mgat4A-allele1</i>        | -5bp  | TTCTGAGTTGAATGCCATTGTCCAACAg-----<br>CCGTGCAGGAGCAGCAACTAACGGAAG                  |
| <i>mgat4A-allele2</i>        | +13bp | TTCTGAGTTGAATGCCATTGTCCAACAgttGAATTCTAGATGATcGCCGTGC<br>AGGAGCAGCAACTAACGGAAG     |
| <i>mgat4B-WT</i>             | ZFN   | <u>GCCCTCCAGCAGCCCCCTCTgaggaCTGGATGATCCTGGAGTT</u>                                |
| <i>mgat4B-allele1</i>        | -7bp  | GCCCTCCAGCAGCCCCCTCTg-----GATGATCCTGGAGTT                                         |
| <i>mgat5-WT</i>              | ZFN   | <u>GGGGGATGATGCTTCTGCACTTCACCATCCA</u> g-----<br>cagcgGACTCAGCCTGAGAGCAGCTCCATGTT |
| <i>mgat5-allele1</i>         | +4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAgcagcCAGCgGACTCAGCCTG<br>AGAGCAGCTCCATGTT          |
| <i>mgat5-allele2</i>         | -4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAg-----<br>gGACTCAGCCTGAGAGCAGCTCCATGTT             |
| <i>mgat5-allele3</i>         | -16bp | GGGGGATGATGCTTCTGCACTTC-----<br>CTCAGCCTGAGAGCAGCTCCATGTT                         |
| <i>B3gnt2-WT</i>             | ZFN   | <u>TTCAGCCCTTCC--cgggcGTA</u> CTGGAACAGAGAGCA                                     |
| <i>B3gnt2-allele1</i>        | +2bp  | TTCAGCCCTTCCCCcgggcGTACTGGAACAGAGAGCA                                             |
| <i>B3gnt2-allele2</i>        | +1bp  | TTCAGCCTTCCC-cgggcGTACTGGAACAGAGAGCA                                              |

|                                            |       |                                                                               |
|--------------------------------------------|-------|-------------------------------------------------------------------------------|
| <i>KO:st3Gal4/6/mgat4A/4B/5</i>            |       |                                                                               |
| <i>mgat4A-WT</i>                           | ZFN   | TTCTGAGTTGAATGCCATTGTCCAACAgtt-----<br>tcGCCGTGCAGGAGCAGCAACTAACGGAAG         |
| <i>mgat4A-allele1</i>                      | -5bp  | TTCTGAGTTGAATGCCATTGTCCAACAg-----<br>CCGTGCAGGAGCAGCAACTAACGGAAG              |
| <i>mgat4A-allele2</i>                      | +13bp | TTCTGAGTTGAATGCCATTGTCCAACAgttGAATTCTAGATGAtcGCCGTGC<br>AGGAGCAGCAACTAACGGAAG |
| <i>mgat4B-WT</i>                           | ZFN   | <u>GCCCTCCAGCAGCCCCCTCTgaggaCTGGATGATCCTGGAGTT</u>                            |
| <i>mgat4B-allele1</i>                      | -7bp  | GCCCTCCAGCAGCCCCCTCTg-----GATGATCCTGGAGTT                                     |
| <i>mgat5-WT</i>                            | ZFN   | GGGGGATGATGCTTCTGCACTTCACCATCCAg-----<br>cagcgGACTCAGCCTGAGAGCAGCTCCATGTT     |
| <i>mgat5-allele1</i>                       | +4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAgcagcCAGCgGACTCAGCCTG<br>AGAGCAGCTCCATGTT      |
| <i>mgat5-allele2</i>                       | -4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAg-----<br>gGACTCAGCCTGAGAGCAGCTCCATGTT         |
| <i>mgat5-allele3</i>                       | -16bp | GGGGGATGATGCTTCTGCACTTC-----<br>CTCAGCCTGAGAGCAGCTCCATGTT                     |
| <i>st3Gal4-WT</i>                          | ZFN   | <u>GGCAGCCTCCAGTGTCGTCgttggt---gTTGTGGTGGGGAATGGGC</u>                        |
| <i>st3Gal4-allele1</i>                     | -5bp  | GGCAGCCTCCAGTGTCGTCgttggt-----gGTGGGGAATGGGC                                  |
| <i>st3Gal4-allele2</i>                     | +4bp  | GGCAGCCTCCAGTGTCGTCgTTGTttgtgTTGTGGTGGGGAATGGGC                               |
| <i>st3Gal6-WT</i>                          | ZFN   | <u>CGGTACCTCTGATTTTGCTttgccCTATGGACAAGGCC</u>                                 |
| <i>st3Gal6-allele1</i>                     | -4bp  | CGGTACCTCTGATTTTGCTt---CTATGGGACAAGGCC                                        |
| <i>st3Gal6-allele2</i>                     | -2bp  | CGGTACCTCTGATTTTGCTttg--CTATGGGACAAGGCC                                       |
| <i>KI:ST6GAL1/KO:st3gal4/6</i>             |       |                                                                               |
| <i>st3Gal4-WT</i>                          | ZFN   | <u>GGCAGCCTCCAGTGTCGTCgttggt---gTTGTGGTGGGGAATGGGC</u>                        |
| <i>st3Gal4-allele1</i>                     | +4bp  | GGCAGCCTCCAGTGTCGTCgttggtTTGTgTTGTGGTGGGGAATGGGC                              |
| <i>st3Gal4-allele2</i>                     | -4bp  | GGCAGCCTCCAGTGTCGTCg-----gTTGTGGTGGGGAATGGGC                                  |
| <i>st3Gal6-WT</i>                          | ZFN   | <u>CGGTACCTCTGATTTTGCT-ttgccCTATGGGACAAGGCC</u>                               |
| <i>st3Gal6-allele1</i>                     | +1bp  | CGGTACCTCTGATTTTGCTttgccCTATGGGACAAGGCC                                       |
| <i>st3Gal6-allele2</i>                     | -7bp  | CGGTACCTCTGATTTTG-----CTATGGGACAAGGCC                                         |
| <i>KI:ST6GAL1/KO:st3gal4/6/mgat4A/4B/5</i> |       |                                                                               |
| <i>mgat4A-WT</i>                           | ZFN   | TTCTGAGTTGAATGCCATTGTCCAACAgtt-----<br>tcGCCGTGCAGGAGCAGCAACTAACGGAAG         |
| <i>mgat4A-allele1</i>                      | -5bp  | TTCTGAGTTGAATGCCATTGTCCAACAg-----<br>CCGTGCAGGAGCAGCAACTAACGGAAG              |
| <i>mgat4A-allele2</i>                      | +13bp | TTCTGAGTTGAATGCCATTGTCCAACAgttGAATTCTAGATGAtcGCCGTGC<br>AGGAGCAGCAACTAACGGAAG |
| <i>mgat4B-WT</i>                           | ZFN   | <u>GCCCTCCAGCAGCCCCCTCTgaggaCTGGATGATCCTGGAGTT</u>                            |
| <i>mgat4B-allele1</i>                      | -7bp  | GCCCTCCAGCAGCCCCCTCTg-----GATGATCCTGGAGTT                                     |
| <i>mgat5-WT</i>                            | ZFN   | GGGGGATGATGCTTCTGCACTTCACCATCCAg-----<br>cagcgGACTCAGCCTGAGAGCAGCTCCATGTT     |
| <i>mgat5-allele1</i>                       | +4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAgcagcCAGCgGACTCAGCCTG<br>AGAGCAGCTCCATGTT      |
| <i>mgat5-allele2</i>                       | -4bp  | GGGGGATGATGCTTCTGCACTTCACCATCCAg-----<br>gGACTCAGCCTGAGAGCAGCTCCATGTT         |
| <i>mgat5-allele3</i>                       | -16bp | GGGGGATGATGCTTCTGCACTTC-----<br>CTCAGCCTGAGAGCAGCTCCATGTT                     |
| <i>st3Gal4-WT</i>                          | ZFN   | <u>GGCAGCCTCCAGTGTCGTCgttggt---gTTGTGGTGGGGAATGGGC</u>                        |
| <i>st3Gal4-allele1</i>                     | -5bp  | GGCAGCCTCCAGTGTCGTCgttggt-----gGTGGGGAATGGGC                                  |
| <i>st3Gal4-allele2</i>                     | +4bp  | GGCAGCCTCCAGTGTCGTCgTTGTttgtgTTGTGGTGGGGAATGGGC                               |
| <i>st3Gal6-WT</i>                          | ZFN   | <u>CGGTACCTCTGATTTTGCTttgccCTATGGACAAGGCC</u>                                 |
| <i>st3Gal6-allele1</i>                     | -4bp  | CGGTACCTCTGATTTTGCTt---CTATGGGACAAGGCC                                        |
| <i>st3Gal6-allele2</i>                     | -2bp  | CGGTACCTCTGATTTTGCTttg--CTATGGGACAAGGCC                                       |

|                                                          |       |                                                                                                     |
|----------------------------------------------------------|-------|-----------------------------------------------------------------------------------------------------|
| <i>KO:mgat2</i><br><i>mgat2-WT</i>                       | TALEN | <u>TCCTTTGTCGCCCATTGCT</u> <i>gctccagaggacgaagCCGCAGGCGGCCACCAC</i><br><u>GA</u>                    |
| <i>mgat2-allele1</i>                                     | -4bp  | TCCTTTGTCGCCCATTGCT <i>gctcc----</i><br><i>gacgaagccgaggcgccaccacga</i>                             |
| <i>KO: mgat2/stgal4/6</i><br><i>mgat2-WT</i>             | TALEN | <u>TCCTTTGTCGCCCATTGCT</u> <i>gctccagaggacgaagCCGCAGGCGGCCACCAC</i><br><u>GA</u>                    |
| <i>mgat2-allele1</i>                                     | -4bp  | TCCTTTGTCGCCCATTGCT <i>gctccag----</i><br><i>cgaagCCGCAGGCGGCCACCACGA</i>                           |
| <i>st3Gal4-WT</i>                                        | ZFN   | <u>GGCAGCCTCCAGTGTCTGTC</u> <i>qttgt----</i> <u>gTTGTGGTGGGGAATGGGC</u>                             |
| <i>st3Gal4-allele1</i>                                   | +4bp  | <u>GGCAGCCTCCAGTGTCTGTC</u> <i>qttgtTTGTgTTGTGGTGGGGAATGGGC</i>                                     |
| <i>st3Gal4-allele2</i>                                   | -4bp  | <u>GGCAGCCTCCAGTGTCTGTC</u> <i>g-----</i> <u>gTTGTGGTGGGGAATGGGC</u>                                |
| <i>st3Gal6-WT</i>                                        | ZFN   | <u>CGGTACCTCTGATTTT</u> <u>GCT</u> <i>-ttgccCTATGGGACAAGGCC</i>                                     |
| <i>st3Gal6-allele1</i>                                   | +1bp  | <u>CGGTACCTCTGATTTT</u> <u>GCT</u> <i>ttgccCTATGGGACAAGGCC</i>                                      |
| <i>st3Gal 6- allele2</i>                                 | -7bp  | <u>CGGTACCTCTGATTTT</u> <u>G-----</u> <u>CTATGGGACAAGGCC</u>                                        |
| <i>KO: mgat2/mgat4A/4B/5</i><br><i>mgat2-WT</i>          | TALEN | <u>TCCTTTGTCGCCCATTGCT</u> <i>gctccagaggacgaagCCGCAGGCGGCCACCAC</i><br><u>GA</u>                    |
| <i>mgat2-allele1</i>                                     | -4bp  | TCCTTTGTCGCCCATTGCT <i>gctccag----</i><br><i>cgaagCCGCAGGCGGCCACCACGA</i>                           |
| <i>mgat4A-WT</i>                                         | ZFN   | <u>TTCTGAGTTGAATGCCATTGTCCAACA</u> <i>gtt----</i><br><i>tcGCCGTGCAGGAGCAGCAACTAACGGAAG</i>          |
| <i>mgat4A-allele1</i>                                    | -5bp  | <u>TTCTGAGTTGAATGCCATTGTCCAACA</u> <i>g-----</i><br><i>CCGTGCAGGAGCAGCAACTAACGGAAG</i>              |
| <i>mgat4A-allele2</i>                                    | +13bp | <u>TTCTGAGTTGAATGCCATTGTCCAACA</u> <i>gttGAATTCTAGATGatcGCCGTGC</i><br><i>AGGAGCAGCAACTAACGGAAG</i> |
| <i>mgat4B-WT</i>                                         | ZFN   | <u>GCCCTCCAGCAGCCCTCT</u> <i>gaggaCTGGATGATCCTGGAGTT</i>                                            |
| <i>mgat4B-allele1</i>                                    | -7bp  | <u>GCCCTCCAGCAGCCCTCT</u> <i>g-----</i> <u>GATGATCCTGGAGTT</u>                                      |
| <i>mgat5-WT</i>                                          | ZFN   | <u>GGGGGATGATGCTTCTGCACTTCACCATCCA</u> <i>g-----</i><br><i>cagcgGACTCAGCCTGAGAGCAGCTCCATGTT</i>     |
| <i>mgat5-allele1</i>                                     | +4bp  | <u>GGGGGATGATGCTTCTGCACTTCACCATCCA</u> <i>gagcagcCAGCgGACTCAGCCTG</i><br><i>AGAGCAGCTCCATGTT</i>    |
| <i>mgat5-allele2</i>                                     | -4bp  | <u>GGGGGATGATGCTTCTGCACTTCACCATCCA</u> <i>g-----</i><br><i>gGACTCAGCCTGAGAGCAGCTCCATGTT</i>         |
| <i>mgat5-allele3</i>                                     | -16bp | <u>GGGGGATGATGCTTCTGCACTTC-----</u><br><i>CTCAGCCTGAGAGCAGCTCCATGTT</i>                             |
| <i>KO:mgat2/st3gal4/6/mgat4A/4B/5</i><br><i>mgat2-WT</i> | TALEN | <u>TCCTTTGTCGCCCATTGCT</u> <i>gctccagaggacgaagCCGCAGGCGGCCACCAC</i><br><u>GA</u>                    |
| <i>mgat2- allele1</i>                                    | -5bp  | TCCTTTGTCGCCCATTGCT <i>gctcca-----</i><br><i>cgaagCCGCAGGCGGCCACCACGA</i>                           |
| <i>mgat4A-WT</i>                                         | ZFN   | <u>TTCTGAGTTGAATGCCATTGTCCAACA</u> <i>gtt----</i><br><i>tcGCCGTGCAGGAGCAGCAACTAACGGAAG</i>          |
| <i>mgat4A-allele1</i>                                    | -5bp  | <u>TTCTGAGTTGAATGCCATTGTCCAACA</u> <i>g-----</i><br><i>CCGTGCAGGAGCAGCAACTAACGGAAG</i>              |
| <i>mgat4A-allele2</i>                                    | +13bp | <u>TTCTGAGTTGAATGCCATTGTCCAACA</u> <i>gttGAATTCTAGATGatcGCCGTGC</i><br><i>AGGAGCAGCAACTAACGGAAG</i> |
| <i>mgat4B-WT</i>                                         | ZFN   | <u>GCCCTCCAGCAGCCCTCT</u> <i>gaggaCTGGATGATCCTGGAGTT</i>                                            |

|                          |        |                                                                                        |
|--------------------------|--------|----------------------------------------------------------------------------------------|
| <i>mgat4B- allele1</i>   | -7bp   | <u>GCCCTCCAGCAGCCCCCTCTg-----GATGATCCTGGAGTT</u>                                       |
| <i>mgat5-WT</i>          | ZFN    | GGGGGATGATGCTTCTGCACTTCACCATCCA <del>g----</del><br>cagcgGACTCAGCCTGAGAGCAGCTCCATGTT   |
| <i>mgat5-allele1</i>     | +4bp   | GGGGGATGATGCTTCTGCACTTCACCATCCA <del>g</del> cagcgCAGCgGACTCAGCCTG<br>AGAGCAGCTCCATGTT |
| <i>mgat5-allele2</i>     | -4bp   | GGGGGATGATGCTTCTGCACTTCACCATCCA <del>g-----</del><br>gGACTCAGCCTGAGAGCAGCTCCATGTT      |
| <i>mgat5-allele3</i>     | -16bp  | GGGGGATGATGCTTCTGCACTTC-----<br>CTCAGCCTGAGAGCAGCTCCATGTT                              |
| <i>st3Gal4-WT</i>        | ZFN    | <u>GGCAGCCTCCAGTGTCGTCgttgt---gTTGTGGTGGGGAATGGGC</u>                                  |
| <i>st3Gal4-allele1</i>   | -5bp   | GGCAGCCTCCAGTGTCGTCgttgt-----gGTGGGGAATGGGC                                            |
| <i>st3Gal4-allele2</i>   | +4bp   | GGCAGCCTCCAGTGTCGTCgTTGTtTgtgTTGTGGTGGGGAATGGGC                                        |
| <i>st3Gal6-WT</i>        | ZFN    | <u>CGGTACCTCTGATTTTGCTtggccCTATGGACAAGGCC</u>                                          |
| <i>st3Gal6-allele1</i>   | -4bp   | CGGTACCTCTGATTTTGCTt---CTATGGGACAAGGCC                                                 |
| <i>st3Gal6-allele2</i>   | -2bp   | CGGTACCTCTGATTTTGCTtgg--CTATGGGACAAGGCC                                                |
| <i>KO:futB</i>           |        |                                                                                        |
| <i>fut8-WT</i>           | CRISPR | CAAATACTTGATCCGTCCACAACCT-TGGCTGGAAAGGGAA                                              |
| <i>fut8-allele1</i>      | -1bp   | CAAATACTTGATCCGTCCACAACC--TGGCTGGAAAGGGAA                                              |
| <i>fut8-allele2</i>      | +1bp   | CAAATACTTGATCCGTCCACAACCTTTGGCTGGAAAGGGAA                                              |
| <i>KO:futB/B4galt1</i>   |        |                                                                                        |
| <i>fut8-WT</i>           | CRISPR | CAAATACTTGATCCGTCCACAACCT-TGGCTGGAAAGGGAA                                              |
| <i>fut8-allele1</i>      | -4bp   | CAAATACTTGATCCGTCCACAA-----GGCTGGAAAGGGAA                                              |
| <i>fut8-allele2</i>      | +1bp   | CAAATACTTGATCCGTCCACAACCTTTGGCTGGAAAGGGAA                                              |
| <i>B4galtl-WT</i>        | ZFN    | <u>TGCATCCGGTCCTACAGCgcccagc---AACTGGACTATGGTA</u>                                     |
| <i>B4gal tl- allele1</i> | +4bp   | TGCATCCGGTCCTACAGCgcccagcCAGCAACTGGACTATGGTA                                           |
| <i>KO:mgat3</i>          |        |                                                                                        |
| <i>mgat3-WT</i>          | ZFN    | <u>TTCCTGGACCACTTCCCAcccggg-----GGCCGGCAGGATGGC</u>                                    |
| <i>mgat3-allele1</i>     | -21bp  | TTCCTGGACCACT-----ATGGC                                                                |
| <i>mgat3-allele2</i>     | +293bp | TTCCTGGACCACTGATACcc--+293bp--cgggGGCCGGCAGGATGGC                                      |
| <i>KO:mgat4C</i>         |        |                                                                                        |
| <i>mgat4C-WT</i>         | ZFN    | <u>ATACTTCAGACTATtatgtAATGCTCGAAGATGATGTT</u>                                          |
| <i>mgat4C-allele1</i>    | -13bp  | ATACTTCAGACT-----CGAAGATGATGTT                                                         |
| <i>mgat4C-allele2</i>    | -8bp   | ATACTTCAGACTAT-----GCTCGAAGATGATGTT                                                    |
| <i>KO:mgat5B</i>         |        |                                                                                        |
| <i>mgat5B-WT</i>         | ZFN    | <u>CGTGGCGCCCTCCGCAAGatgagtGACCTGCTGGAGCTG</u>                                         |
| <i>mgat5B-allele1</i>    | -7bp   | CAGCTCCAGCAGGT-----CTTGCGGAGGGCGCCACG                                                  |
| <i>mgat5B-allele2</i>    | -5bp   | CAGCTCCAGCAGGT-----ATCTTGCGGAGGGCGCCACG                                                |
| <i>KO:B3gnt8</i>         |        |                                                                                        |
| <i>B3gnt8-WT</i>         | TALEN  | <u>TGGTCCAGAGATAGCTAATgaagcttctagggtgGAGAAGCTGGGGCTGCTG</u><br><u>A</u>                |
| <i>B3gnt8-allele1</i>    | -17bp  | TGGTCCAGAGG-----<br>AGGGTGGAGAAGCTGGGGCTGCTGA                                          |
| <i>KO:mgat1</i>          |        |                                                                                        |
| <i>mgat1-WT</i>          | ZFN    | <u>AACAAGTTCAAGTTCCcagcaGCTGTGGTAGTGGAGGAC</u>                                         |
| <i>mgat1-allele1</i>     | -2bp   | <u>AACAAGTTCAAGTTCCc--gcaGCTGTGGTAGTGGAGGAC</u>                                        |

5

Región objetivo del gen subrayada.  
Tabla 5

tcgcgcgtttcgggtgatgacgggtgaaaacctctgacacatgcagctcccgagacgggtcacagcttgtctgta  
agcggatgccgggagcagacaagcccgctcagggcgcgctcagcgggtgttggcgggtgtcggggctggcttaac  
tatgcggcatcagagcagattgtactgagagtgcaccatatgcgggtgtgaaataccgcacagatgcgtaagga  
gaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcgggtcggggcctc  
ttcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagtgtggtaacgccagggttttcc  
cagtcacgacgttgtaaaaacgacggccagtgaattcgagctcgggtaccaagcttgggtgcatgctgtccggag  
tctcagcgttataccagaagtgcacctgggtcggggaagactatagtgtcacctaaatctctagagcccttcat  
taggcgcgccaatcccattgcaaatctctacaaaaggagtgtttcccaactgctctatcaagaggaatgttgca  
caactgtgacctgaatgcaaacatcacacgcgccagcagagaggaagaagagaggcttccctgacccgggaatcg  
aaccggggccgcgcggtgagagcgccgaatcctaaccactagaccaccaggaggacgcgcgcaaaagctcaat  
gagctataattatcccccttgaaaaacctacaaaaacagtgtttcaaaactgctctgtgaaaagggacctttgc  
tagcacgcggcgccaggcaaaacgtgggcacgctgcggttggccgggaatcgaacccgggtcaactgcttgaa  
ggcagctatgctcaccactataccaccaacgcgcacacgcgcagcagattctacgggaagagtgtttcaaaa  
ctgctctatcaagagaaatgttccaccttgtgtgtggaatgcagccatcacacgcgtccatgaaagggcttaa  
ttaagatatcggttaaacgtcgacctgcagaggcggcgataactagctgatcgcggaatcctgtccctagg  
ccaccactgtgggggtgcccttcattaggcgcgccaatcccattgcaaatctctacaaaaggagtgtttcccaa  
ctgctctatcaagaggaatgttgcacactgtgacctgaatgcaaacatcacacgcgccagcagagaggaaga  
gagaggcttccctgacccgggaatcgaacccggggcgcgcggtgagagcgccgaatcctaaccactagaccac  
caggggagcacgcgcgcaaaagctcaatgagctataattatcccccttgaaaaacctacaaaaacagtgtttcaaaa  
ctgctctgtgaaaagggacctttgctagcacgcggcgccaggcaaaacgtgggcacgctgcggttggccgggaa  
tcgaacccgggtcaactgcttggaaggcagctatgctcaccactataccaccaacgcgcacacgcgcagcag  
attctacgggaagagtgtttcaaaaactgctctatcaagagaaatgttccaccttgtgtgtggaatgcagccat  
cacacgcgtccatgaaagggcggttgcattgctgtccggagtctcagcgttataccagaagtgaacctgggtcgg  
ggaagaaaagcttcatgggtcatagctgtttcctgtgtgaaattgttatccgctcacaaattccacacaacatac  
gagccggaagcataaagtgtaaagcctgggggtgcctaagtgtgagctaactcacattaattgogttgcgctc  
actgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgccaacgcgcggggagaggc  
ggtttgctatttgggcgctcttccgcttcctcgctcactgactcgctgcgctcggctcggttcggctgcggcgag  
cggtatcagctcactcaaaggcggttaatacgggttatccacagaatcaggggataacgcaggaaagaacatgtg  
agcaaaaggccagcaaaaggccagggaacgtaaaaaggccggttgcgtggcggtttttccataggctccgcccc  
cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg  
cgtttccccctggaagctccctcgctgcgctctcctgttccgacctgcgcttacggatacctgtccgcctt  
tctcccttcgggaagcggtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggctcgttcgc  
tccaagctgggctgtgtgcacgaaccccccggttcagcccgacgctgcgccttatccggtaactatcgctcttg

agtccaaccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggta  
 tgtaggcgggtgctacagagttccttgaagtgggtggcctaactacggctacactagaagaacagtatcttgggtatc  
 tgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccacgctg  
 gtagcgggtggttttttggtttgaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttgat  
 cttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttgggtcatgagattatcaaaa  
 aggatcttcacctagatccttttaaatataaaatgaagtttttaaatcaatctaaagtatatatgagtaaactt  
 ggtctgacagttagaaaaactcatcgagcatcaaataaaaactgcaatttattcatatcaggattatcaatacc  
 atatcttttgaaaaagcggtttctgtaataagggagaaaaactcacgaggcagttccataggatggcaagatcc  
 tgggtatcggctctgcgattccgactcgtccaacatcaatacaacctattaatttccctcgtcaaaaaataaggt  
 tatcaagtgagaaatcaccatgagtgacgactgaatccgggtgagaatggcaaaagtttatgcatttctttcca  
 gacttgttcaacaggccagccattacgctcgtcatcaaaatcactcgcacatcaacaaaaccggttatcattcgt  
 gattgcgcctgagcgcagacgaaatacgcgatcgcgtgttaaaaggacaattacaaacaggaatcgaatgcaacc  
 ggcgaggaacactgccagcgcacatcaacaatatcttcacctgaatcaggatattcttctaataacctggaatgc  
 tgttttccagggtatcgagtggtgagtaaccatgcacatcaggagtagcgataaaatgcttgatgggtcgga  
 agaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgc  
 catgtttcagaaacaactctggcgcatcgggcttccatacaatcgatagattgtcgcacctgattgcccgc  
 attatcgcgagcccatcttatacccatataaatcagcatccatggttggaatttaatcgcggcctagagcaagac  
 gtttcccggtgaatatgggtcatactcttctttttcaatattattgaagcatttatcagggttattgtctca  
 tgagcggatacatatttgatgtatttagaaaaataaacaataaggggttcgcgcacatttccccgaaaagt  
 gccacctgacgtctaagaaaccattattatcatgacattaacctataaaaaataggcgtatcacgaggcccttt  
 cgtc

**REIVINDICACIONES**

1. Una célula de mamífero en donde mgat4A, mgat4B y mgat5 están eliminados en la célula.
- 5 2. La célula según la reivindicación 1, en donde los genes han sido inactivados usando dedos de zinc, TALENS, o CRISPR/Cas 9.
3. La célula según cualquiera de las reivindicaciones 1-2, con la eliminación de mgat4A, mgat4B, y mgat4C.
- 10 4. La célula según cualquiera de las reivindicaciones 1-3, que comprende la eliminación de mgat5 y mgat5B.
5. La célula según una de las reivindicaciones 1-4, en donde se ha eliminado FUT8.
- 15 6. La célula según una de las reivindicaciones 1-5, en donde se ha eliminado B3gnt2.
7. La célula según cualquiera de las reivindicaciones 1-6, en donde se han eliminado mgat2, mgat4A, mgat4B y mgat5.
- 20 8. La célula según cualquiera de las reivindicaciones 1-7, en donde se han eliminado ST3gal3, ST3gal4 y ST3gal6.
9. La célula según una de las reivindicaciones 1-8, en donde se ha activado FUT8.ST6Gal1.
- 25 10. La célula según una cualquiera de las reivindicaciones 1-9, en donde la célula se selecciona del grupo que consiste en CHO, NS0, SP2/0, YB2/0, CHO-K1, CHO-DXB11, CHO-DG44, CHO-S, HEK293, HUVEC, HKB, PER-C6, NS0, o derivados de cualquiera de estas células.
11. La célula según cualquiera de las reivindicaciones 1-10, en donde la célula es una célula CHO.
- 30 12. La célula según una cualquiera de las reivindicaciones 1-11, codifica además una proteína exógena de interés, que es un polipéptido seleccionado del grupo que consiste en EPO, un anticuerpo IgG, una proteína implicada en la hemostasia, o un factor de coagulación.
- 35 13. Un método para producir una proteína recombinante con un patrón de glicosilación más homogéneo, el método comprende:
  - expresar la proteína en una célula de mamífero según cualquiera de las reivindicaciones anteriores.

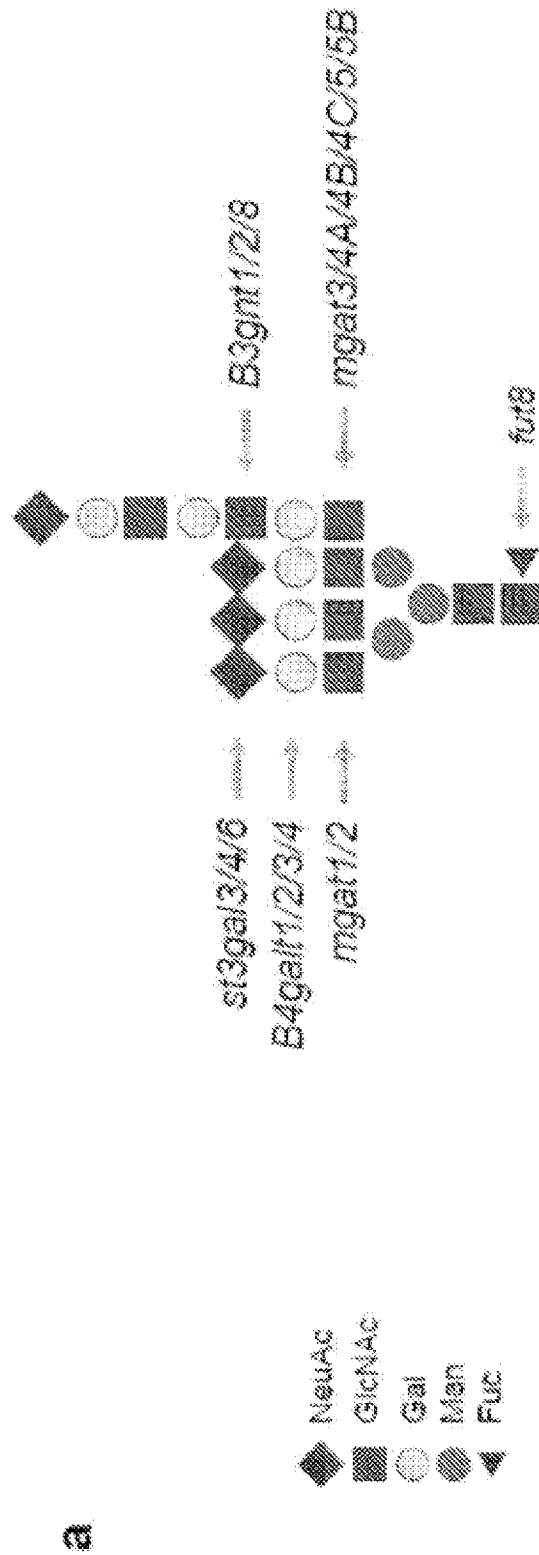


Fig. 1A

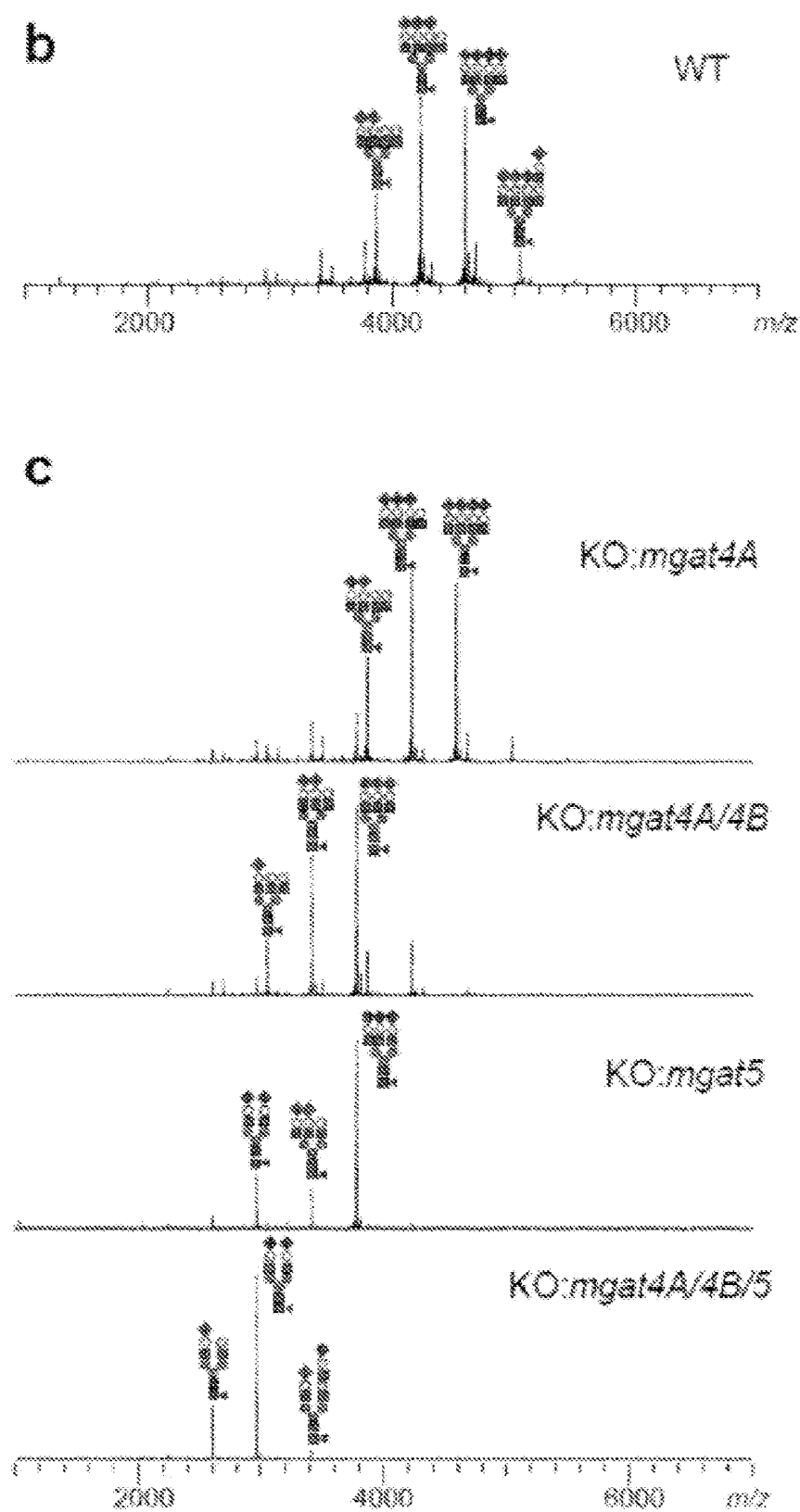


Fig. 1B-C

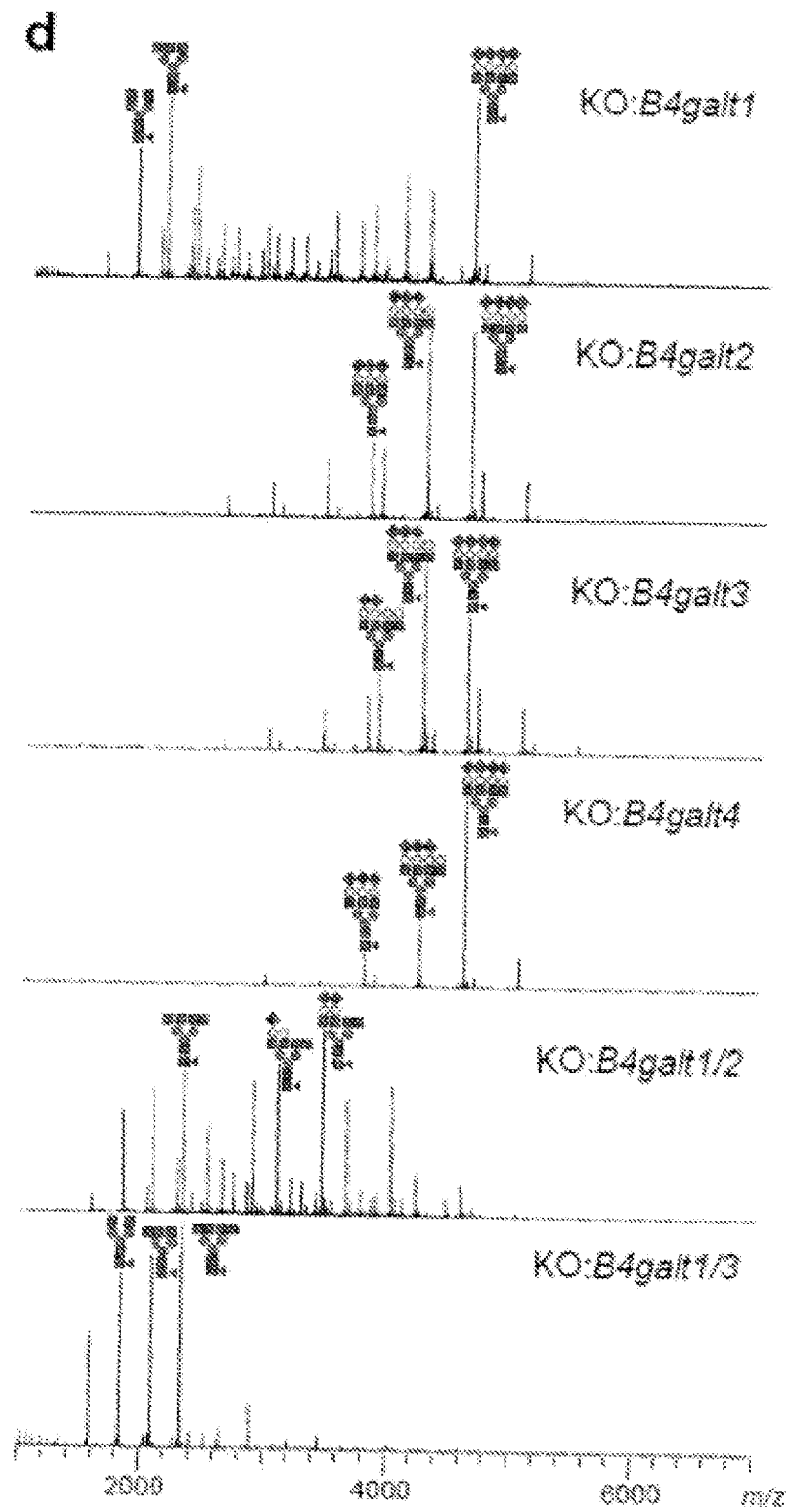


Fig. 1D

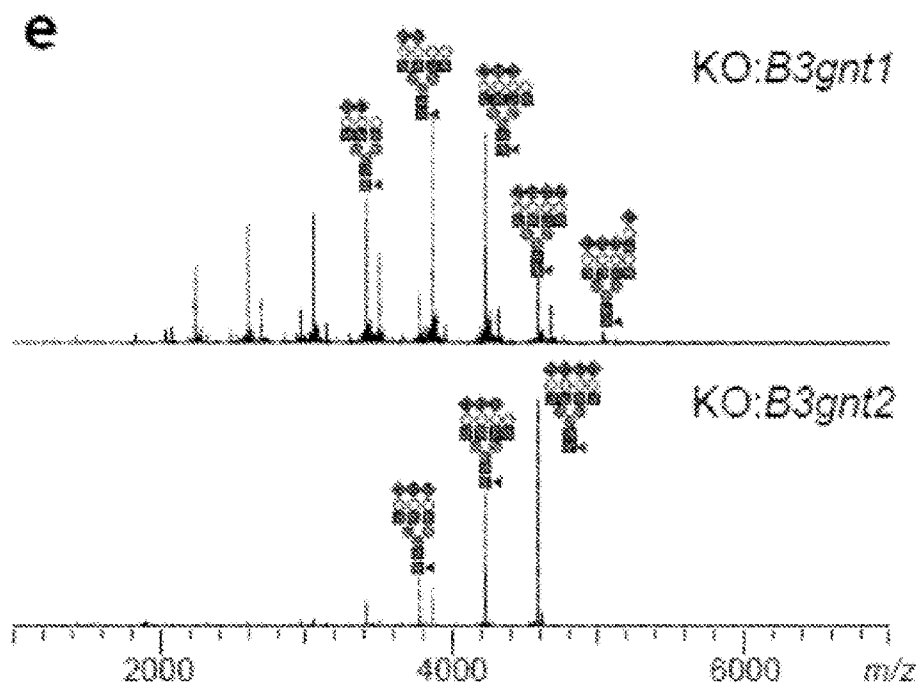


Fig. 1E

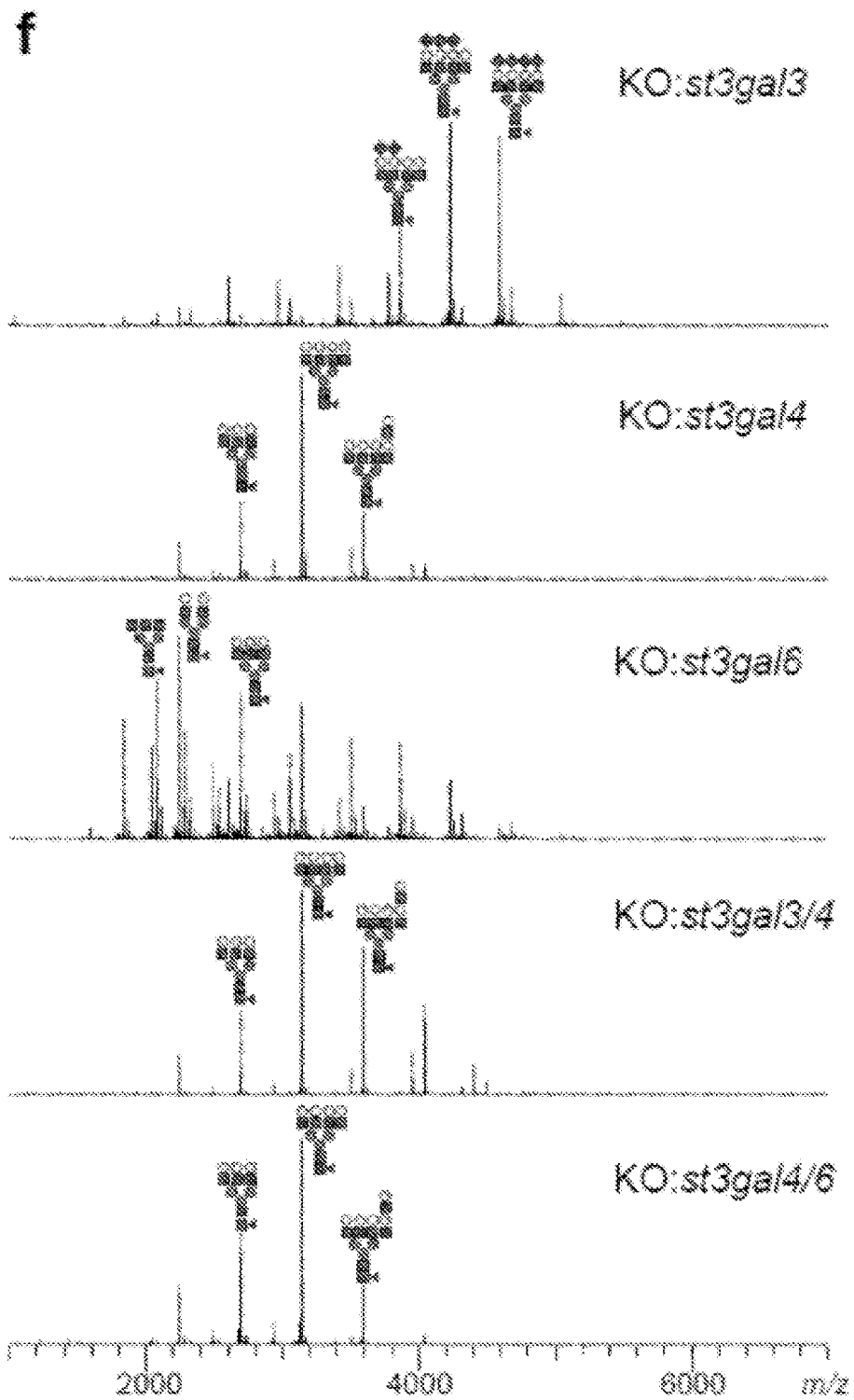


Fig. 1F

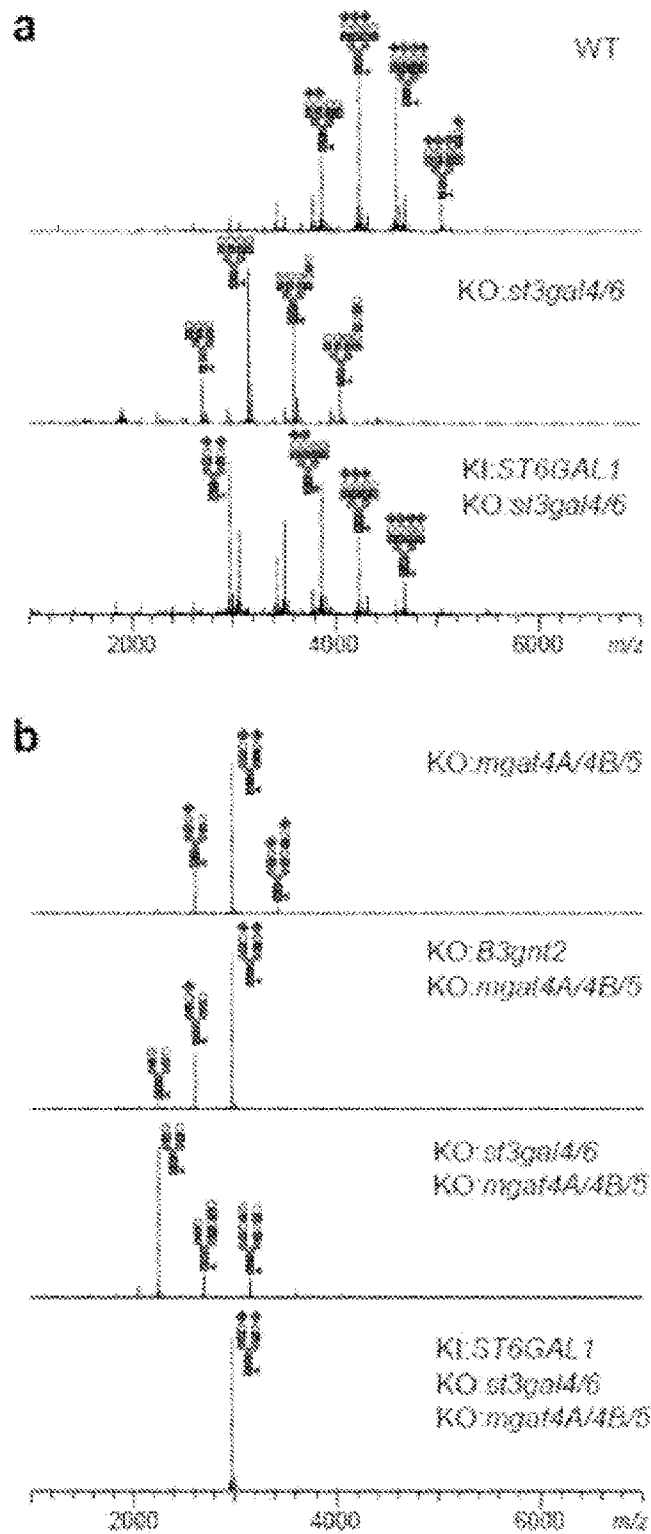


Fig. 2

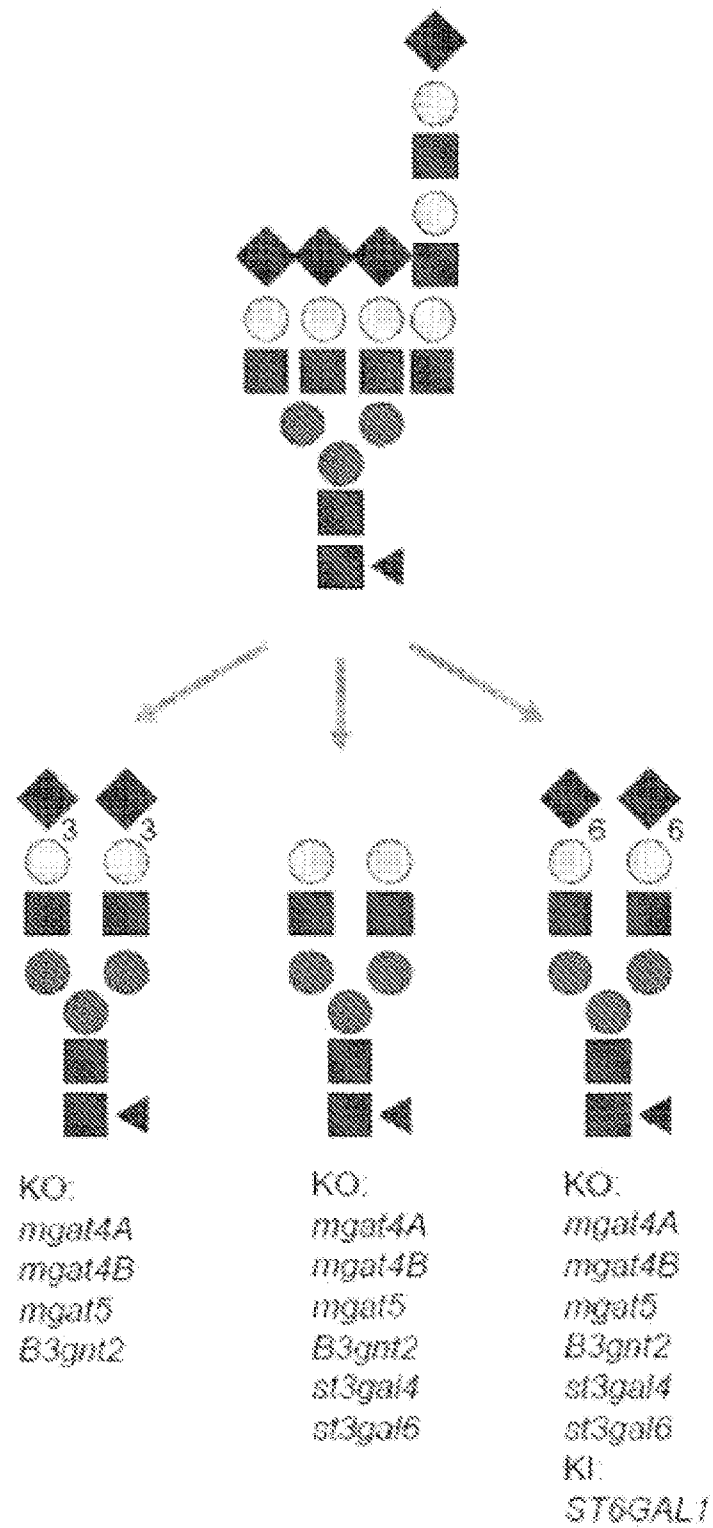


Fig. 3

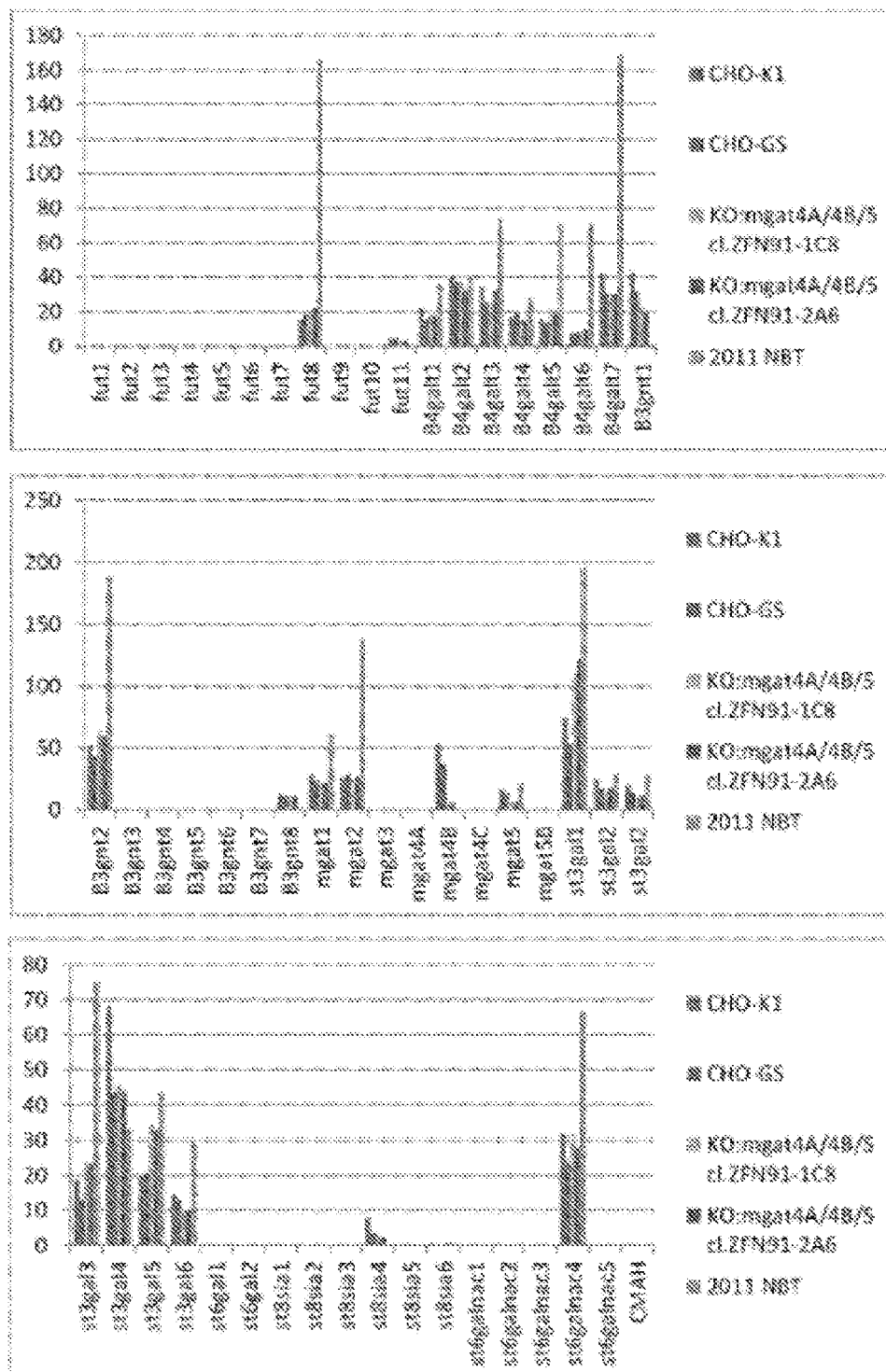


Fig. 4

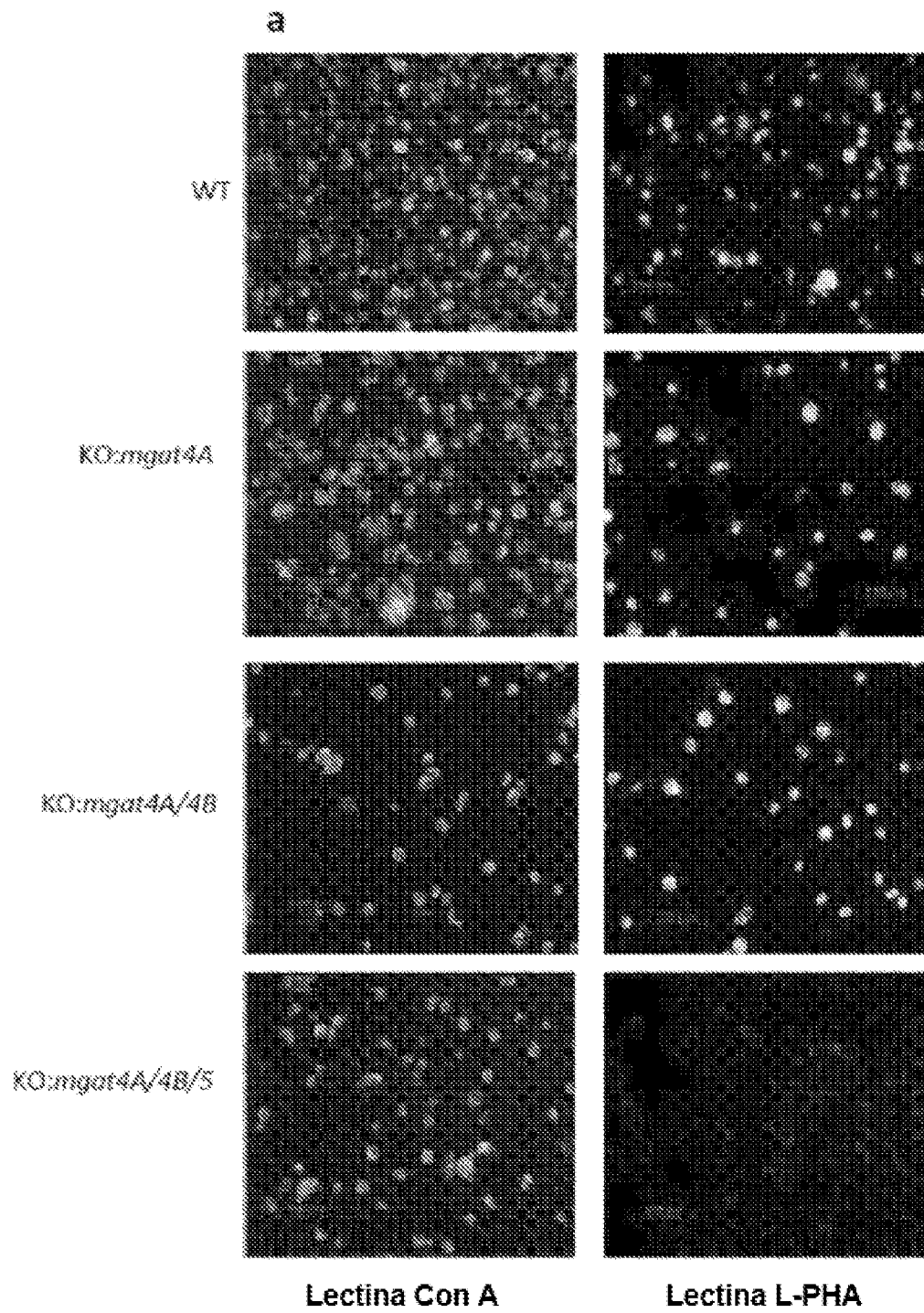


Fig. 5A

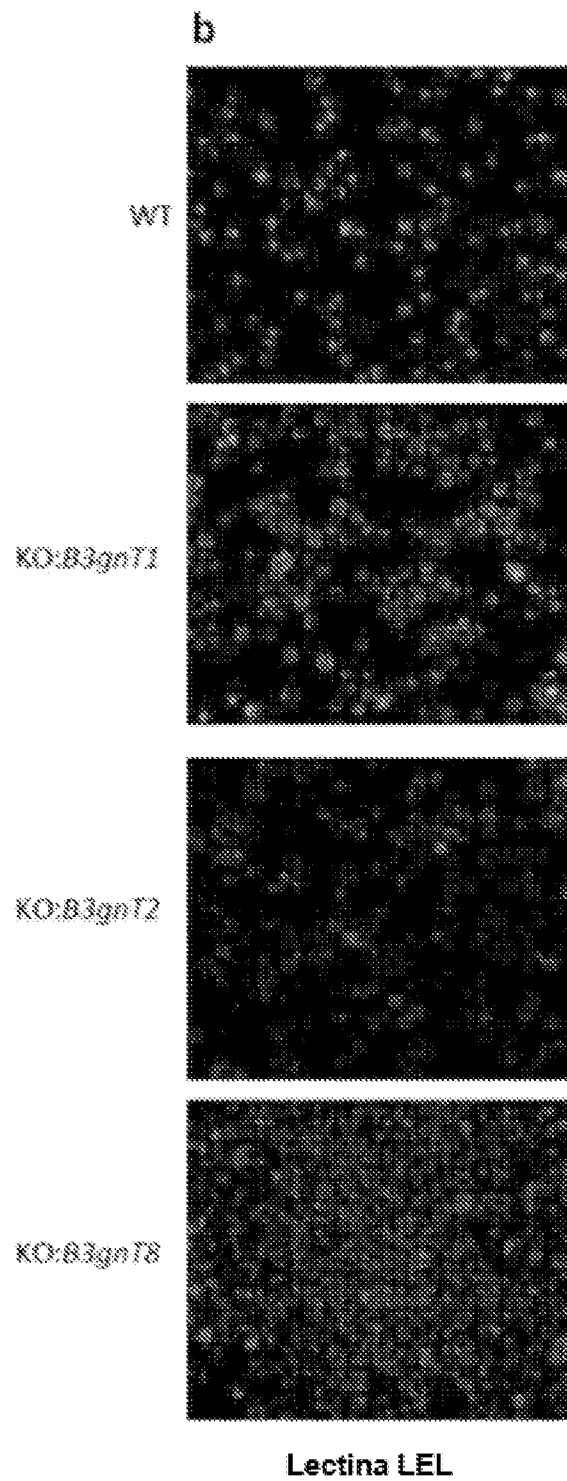


Fig. 5B

61

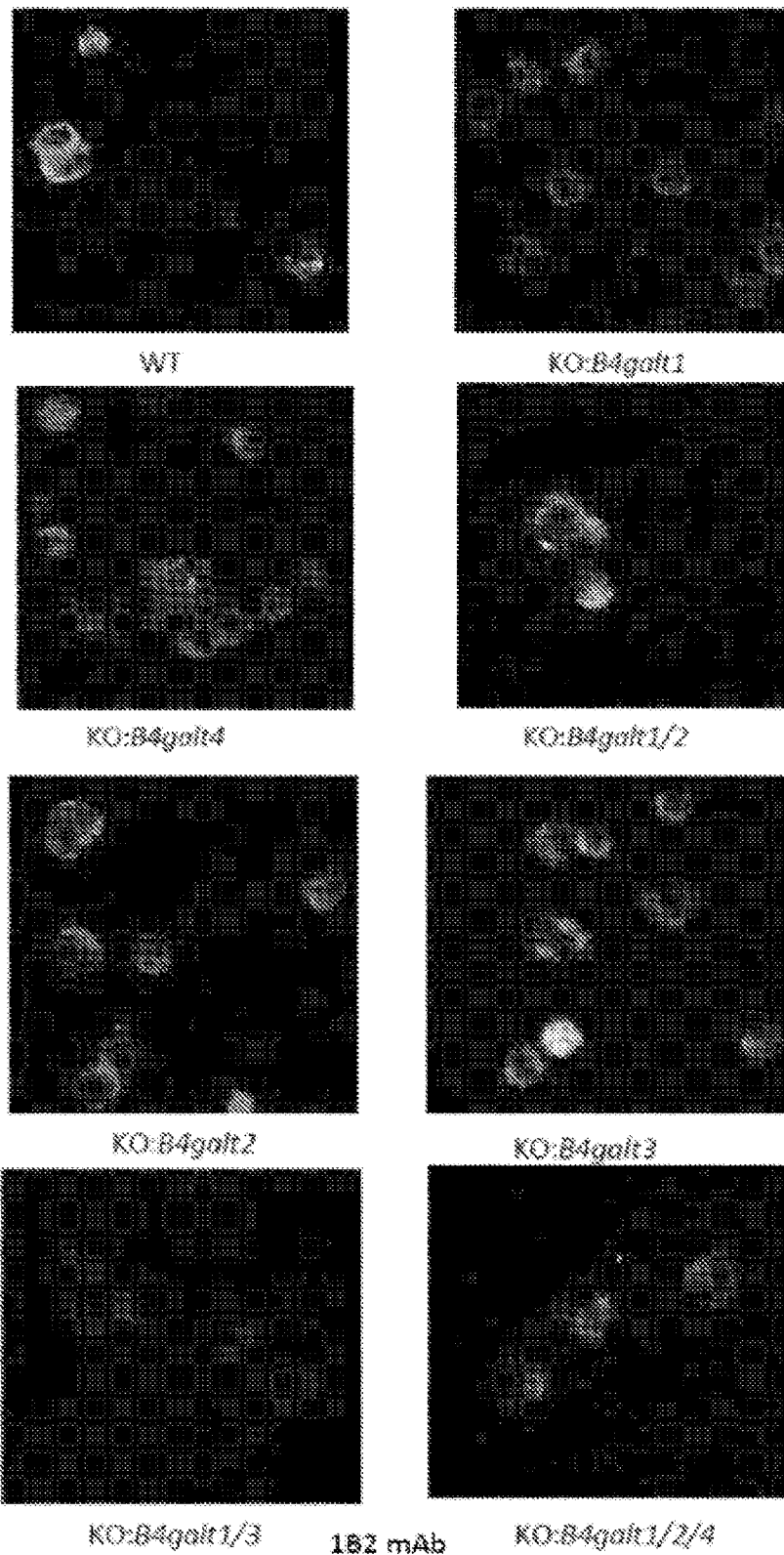
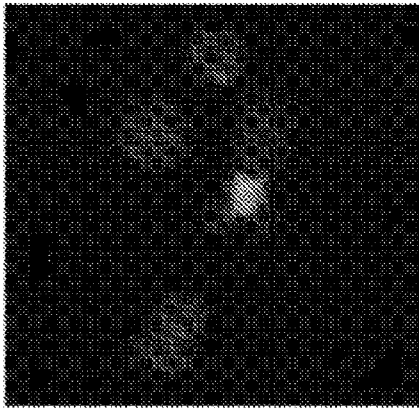
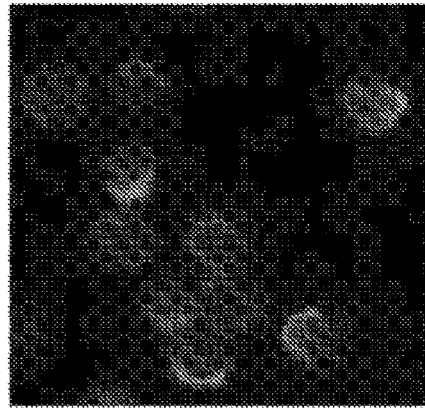


Fig. 6A

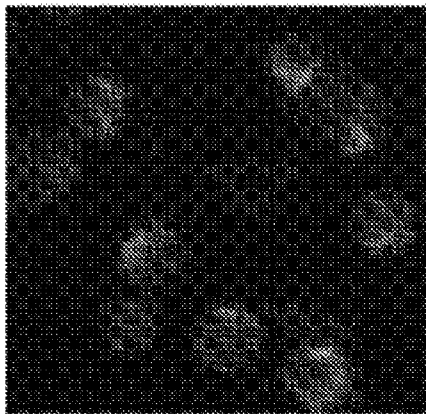
**b**



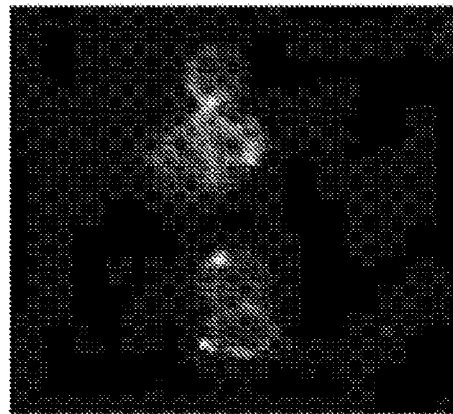
KO:B4galt1/mgat4A/4B/5



KO:B4galt2/mgat4A/4B/5



KO:B4galt3/mgat4A/4B/5



KO:B4galt4/mgat4A/4B/5

1B2 mAb

Fig. 6B

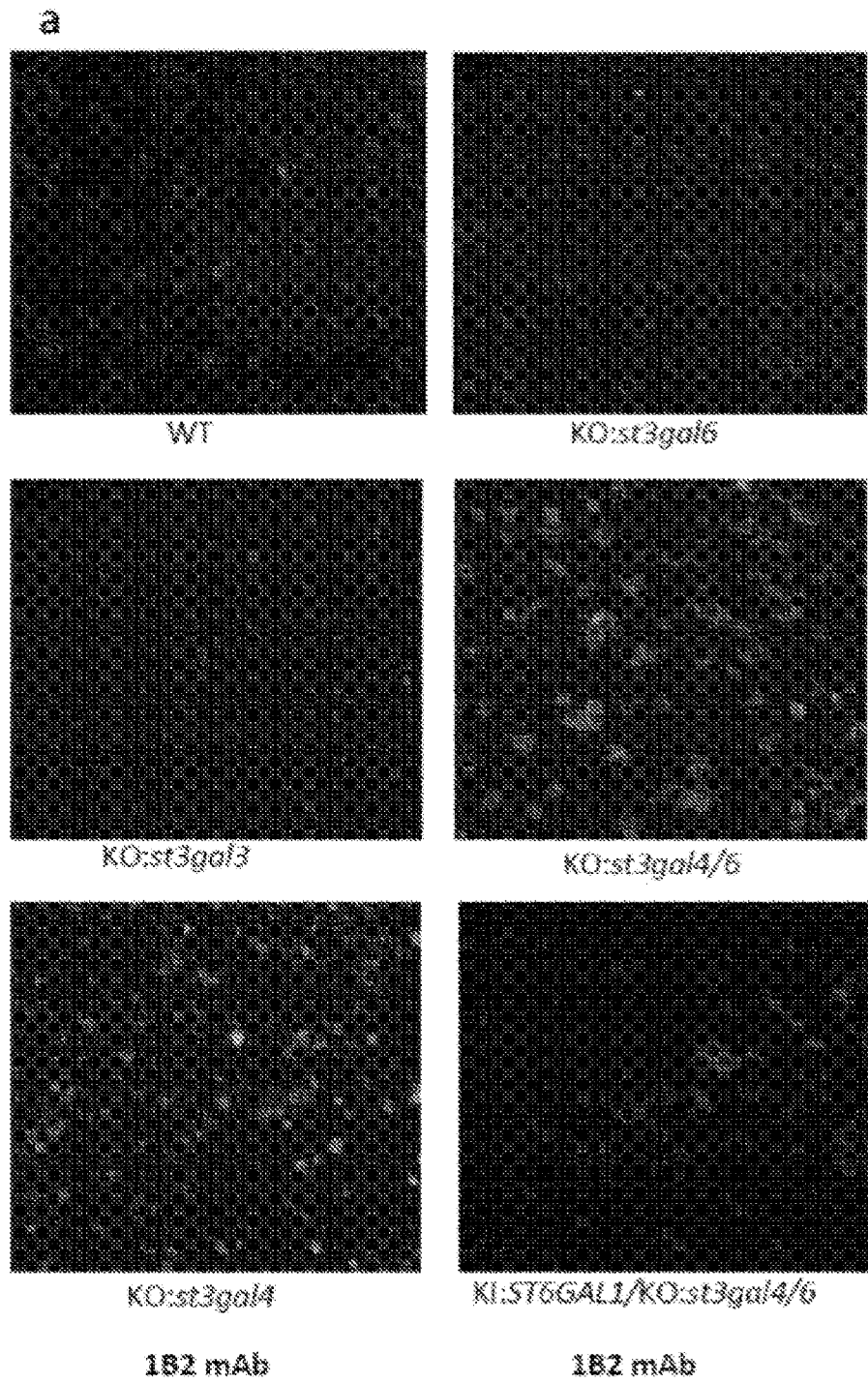
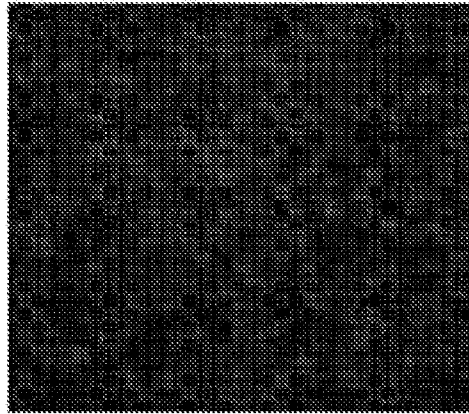
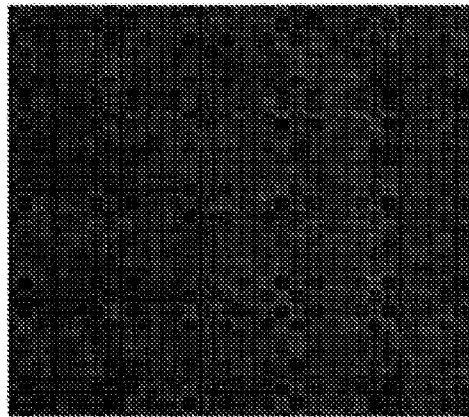


Fig. 7A

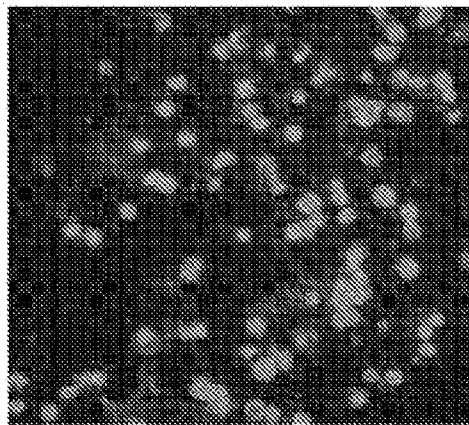
**b**



WT



*KO:st3gal4/6*



*Ki:ST6GAL1/KO:st3gal4/6*

**Lectina SNA**

**Fig. 7B**

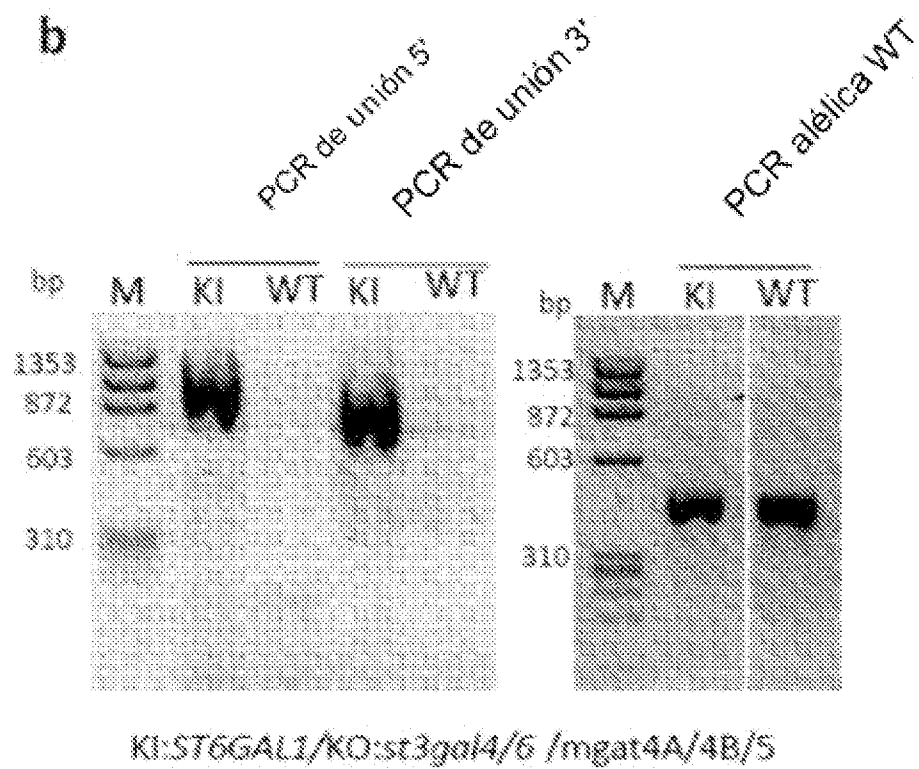
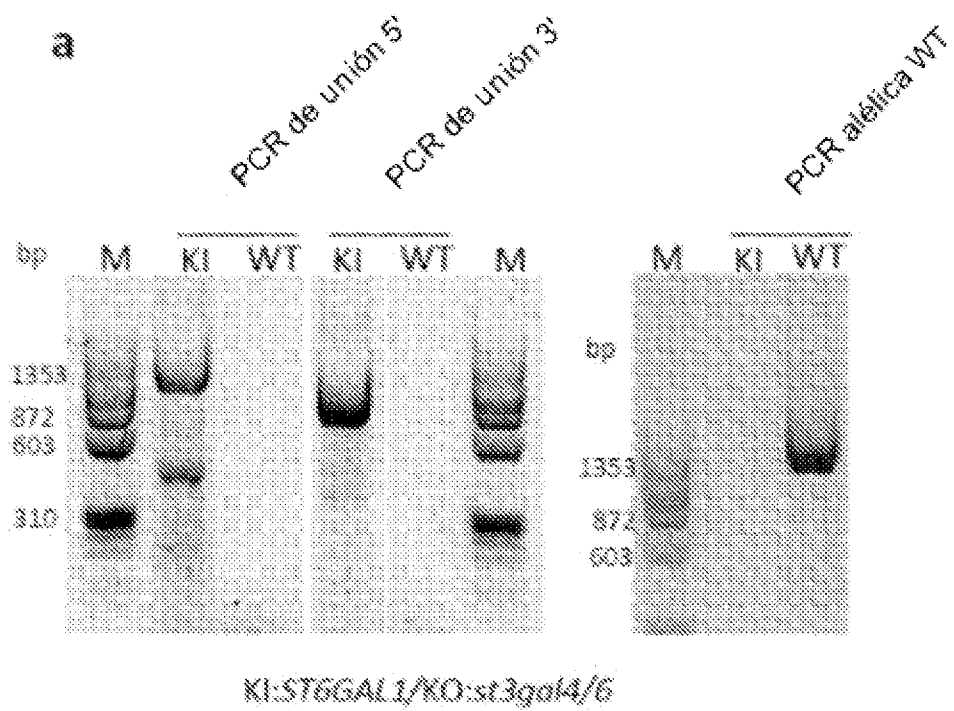


Fig. 8

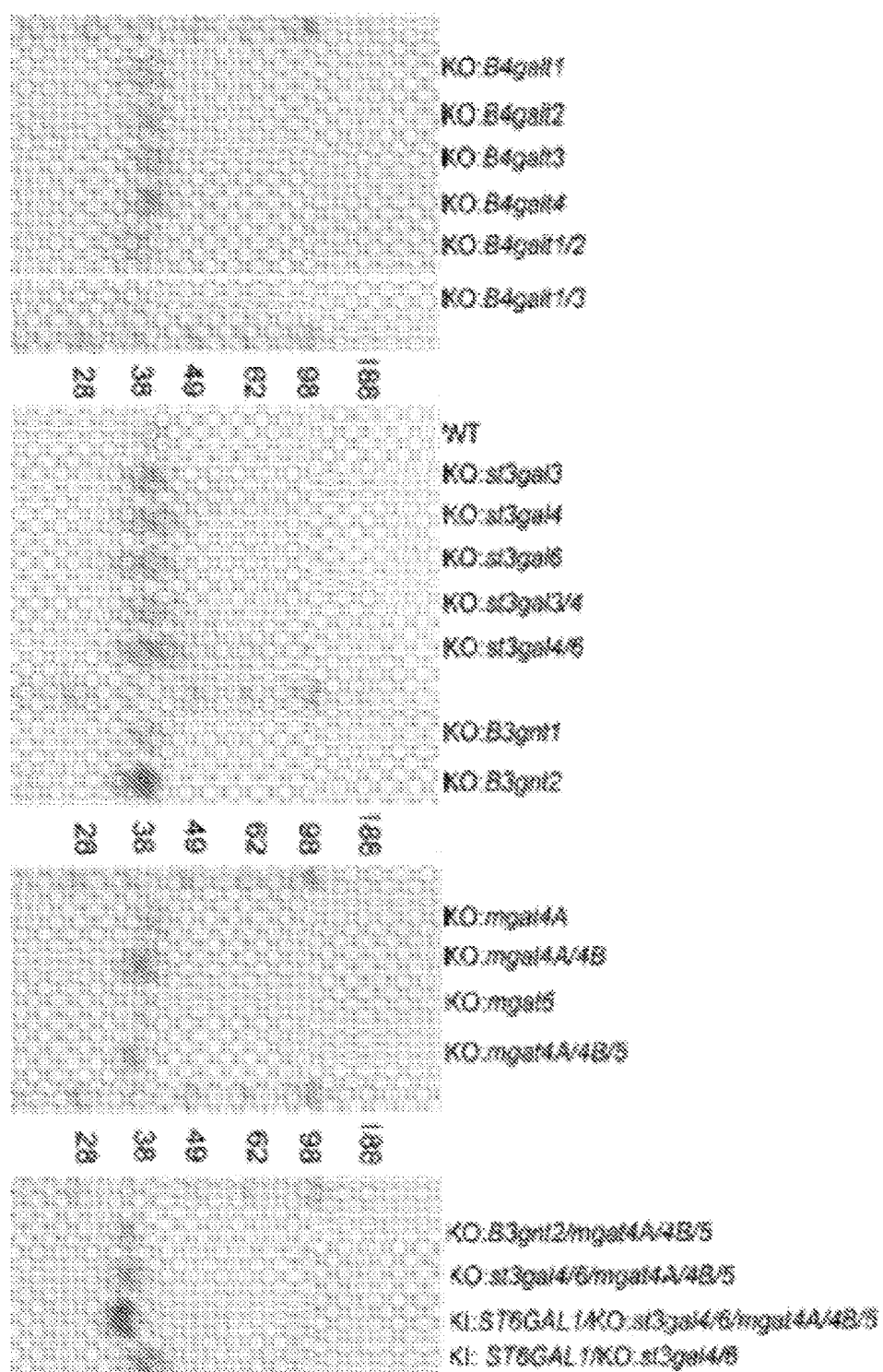


Fig. 9

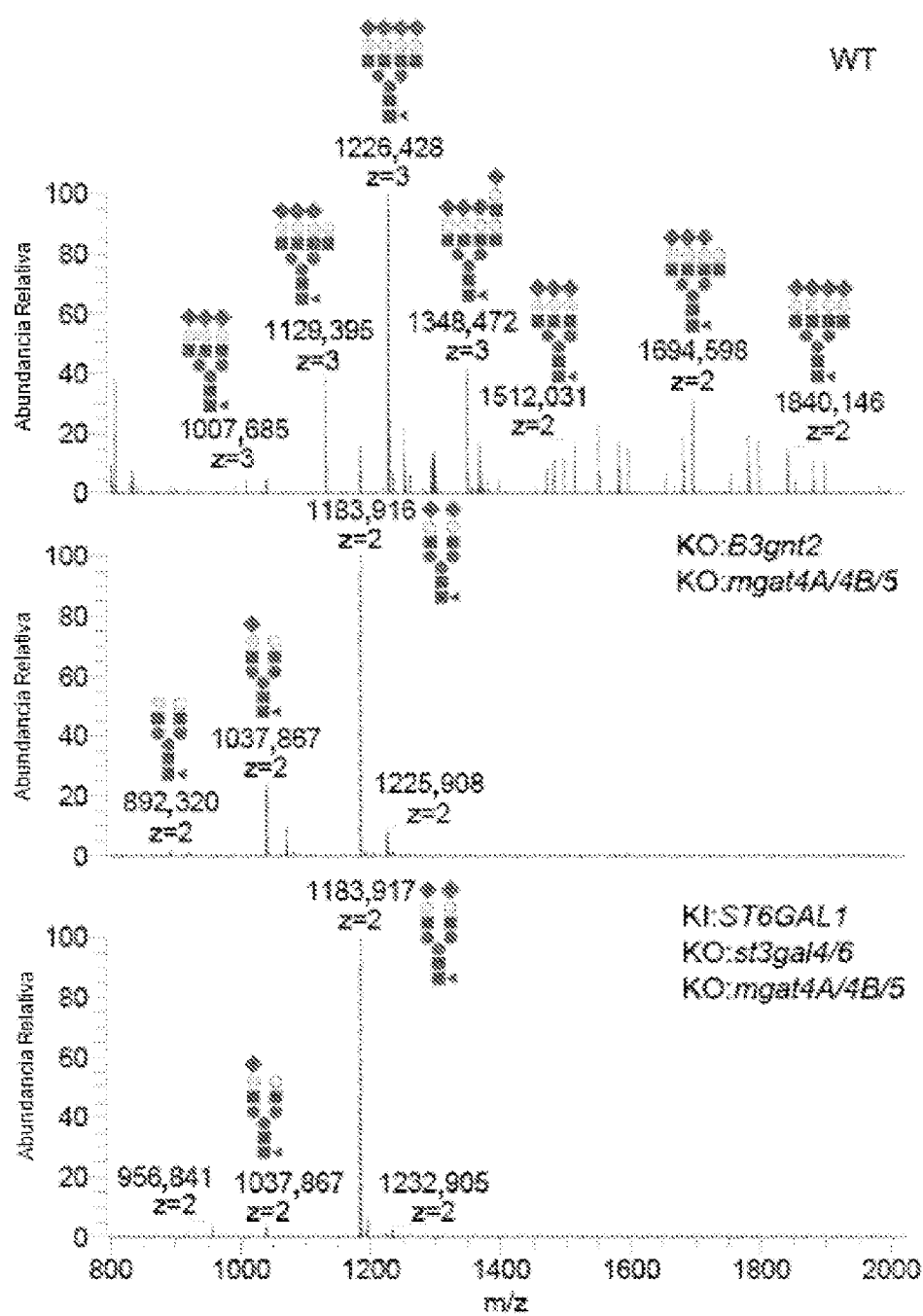
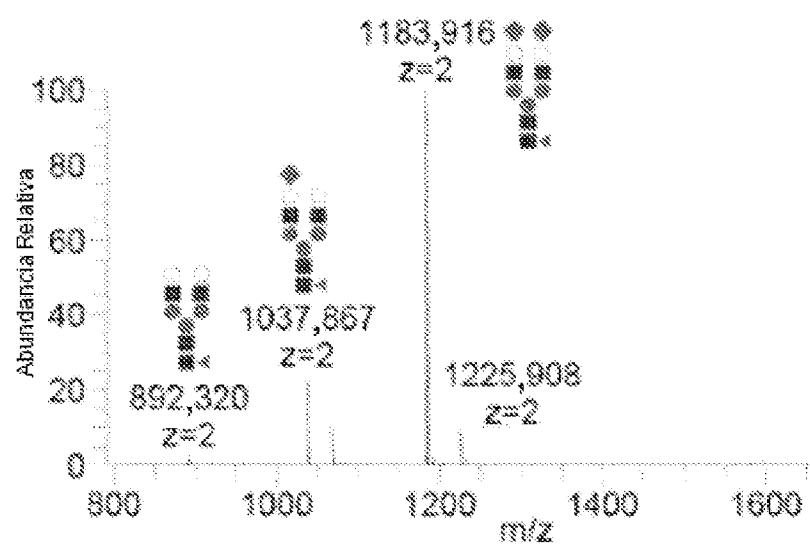


Fig. 10

KO:B3gnt2

KO:mgat4A/4B/5



HCD-MS2@1183.9

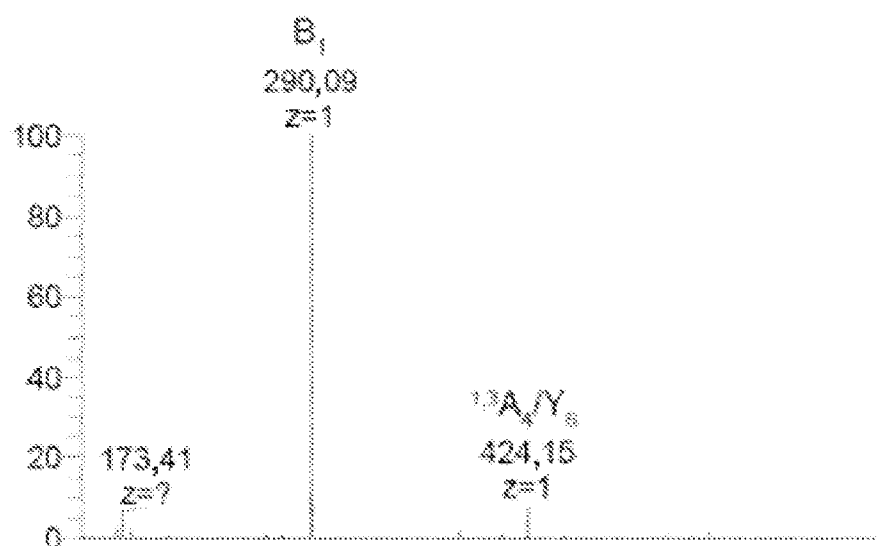


Fig. 11A

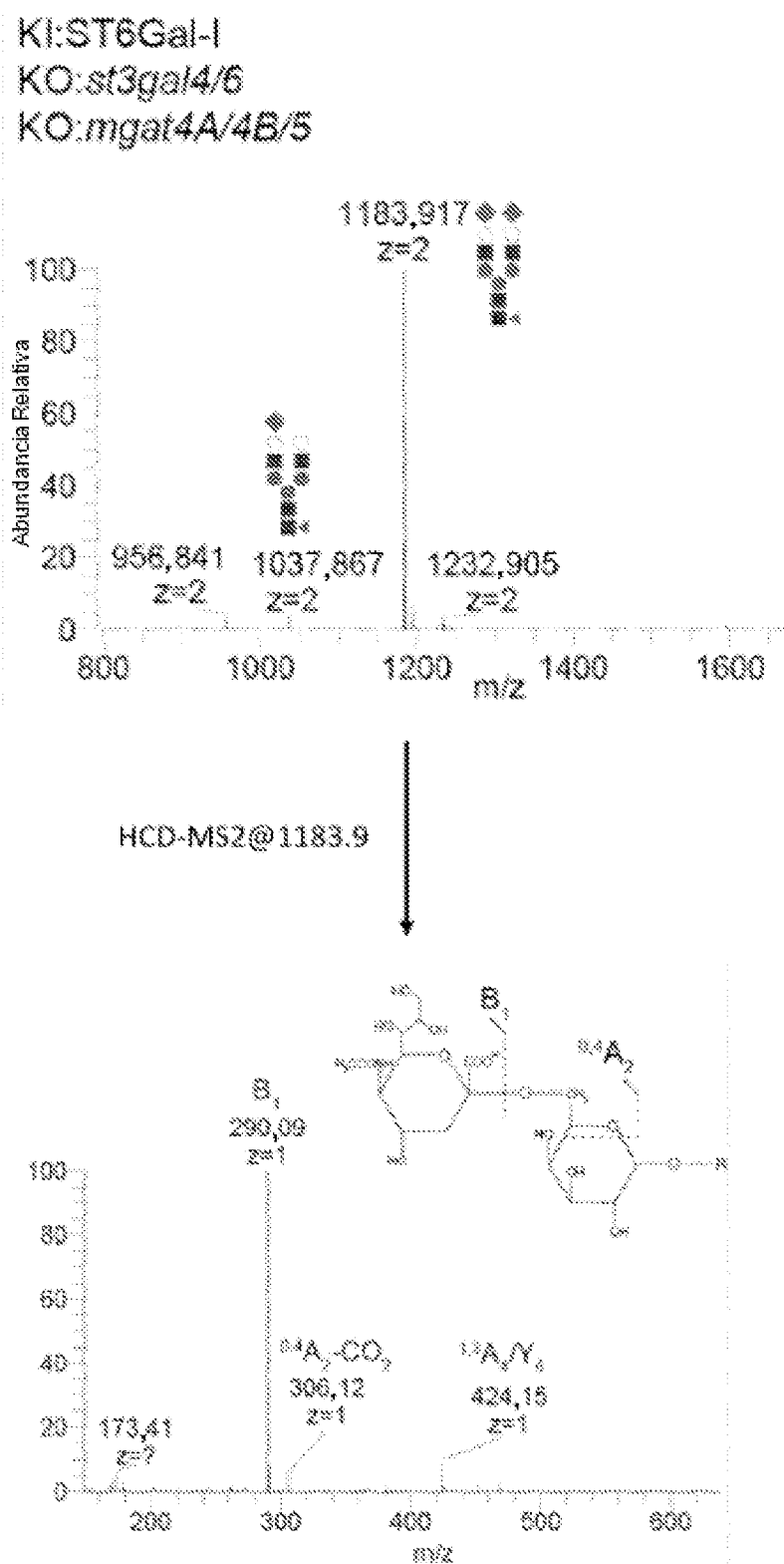


Fig. 11B

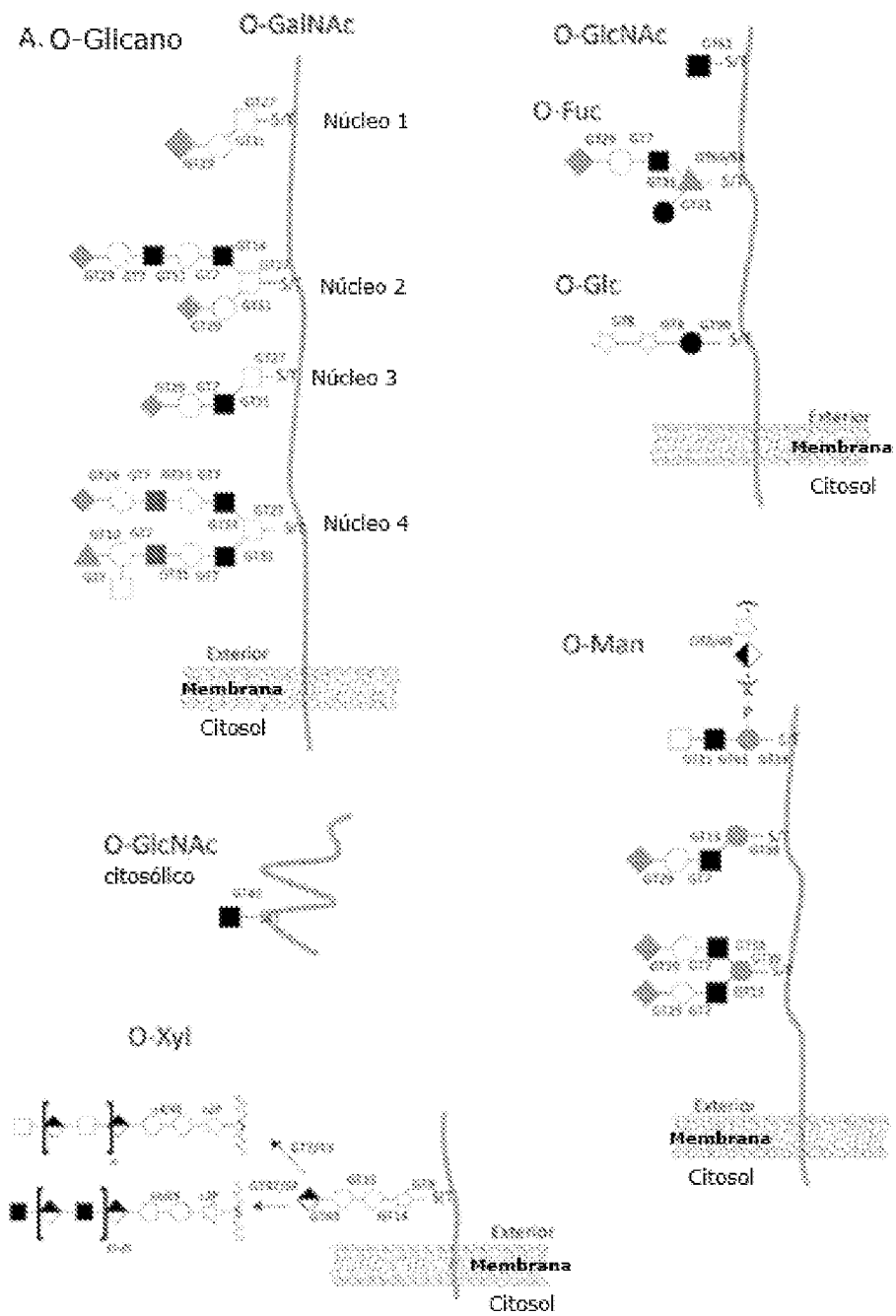


Fig. 12A



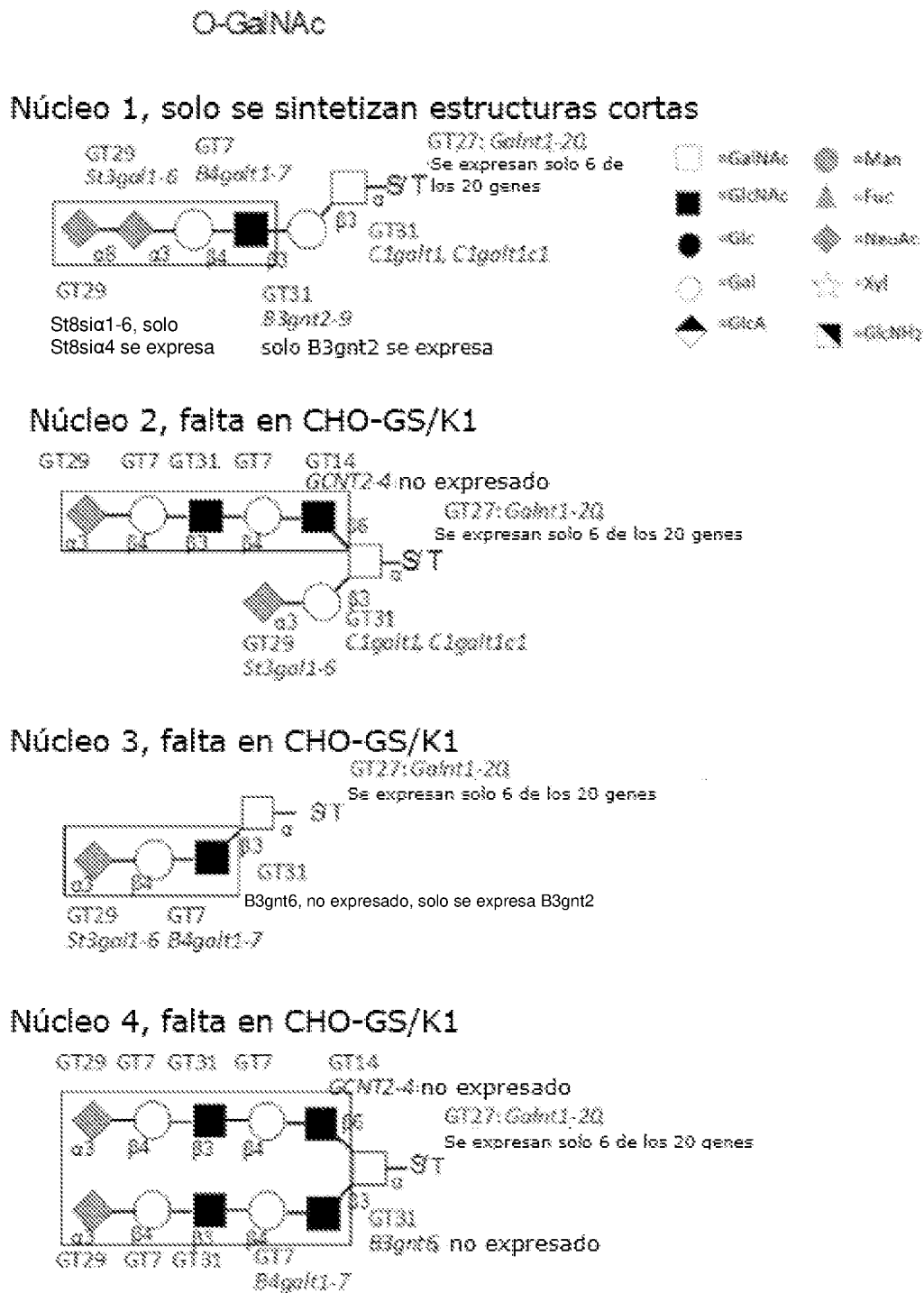


Fig. 13A1

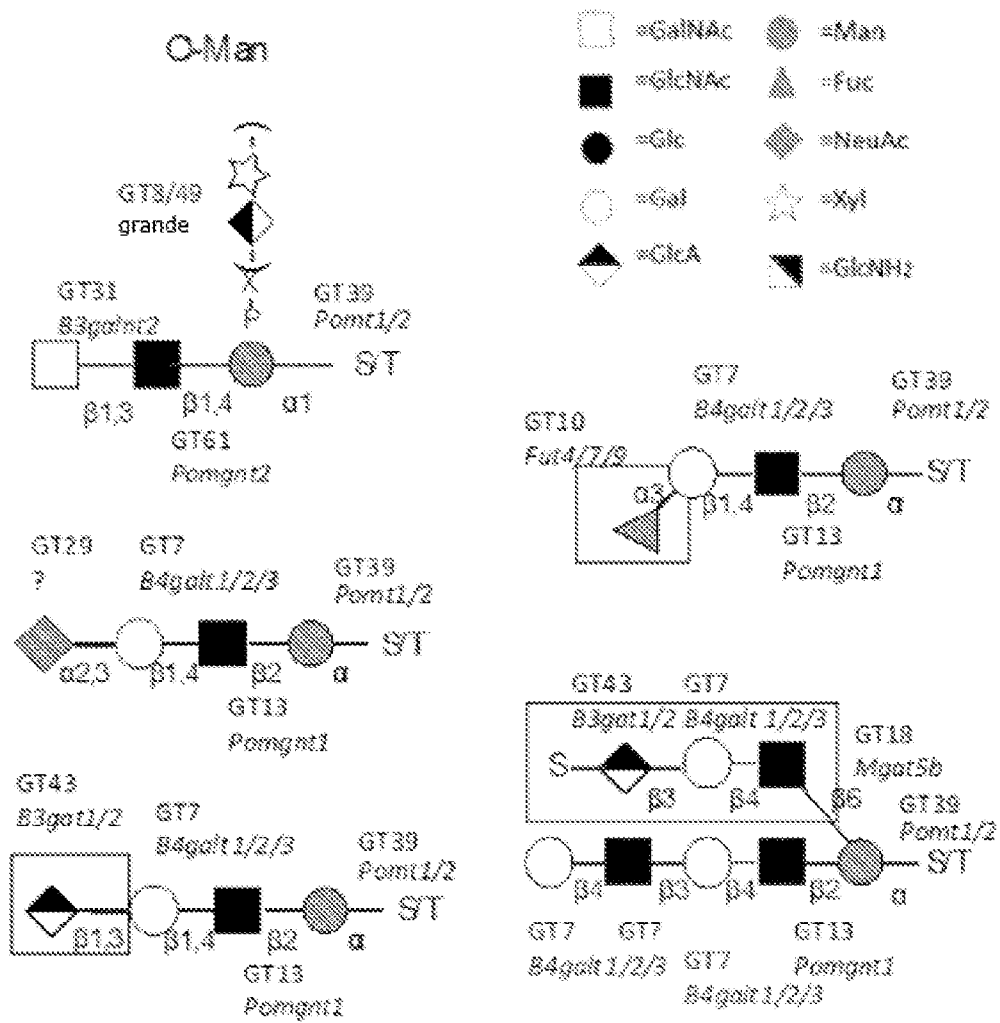


Fig. 13A2

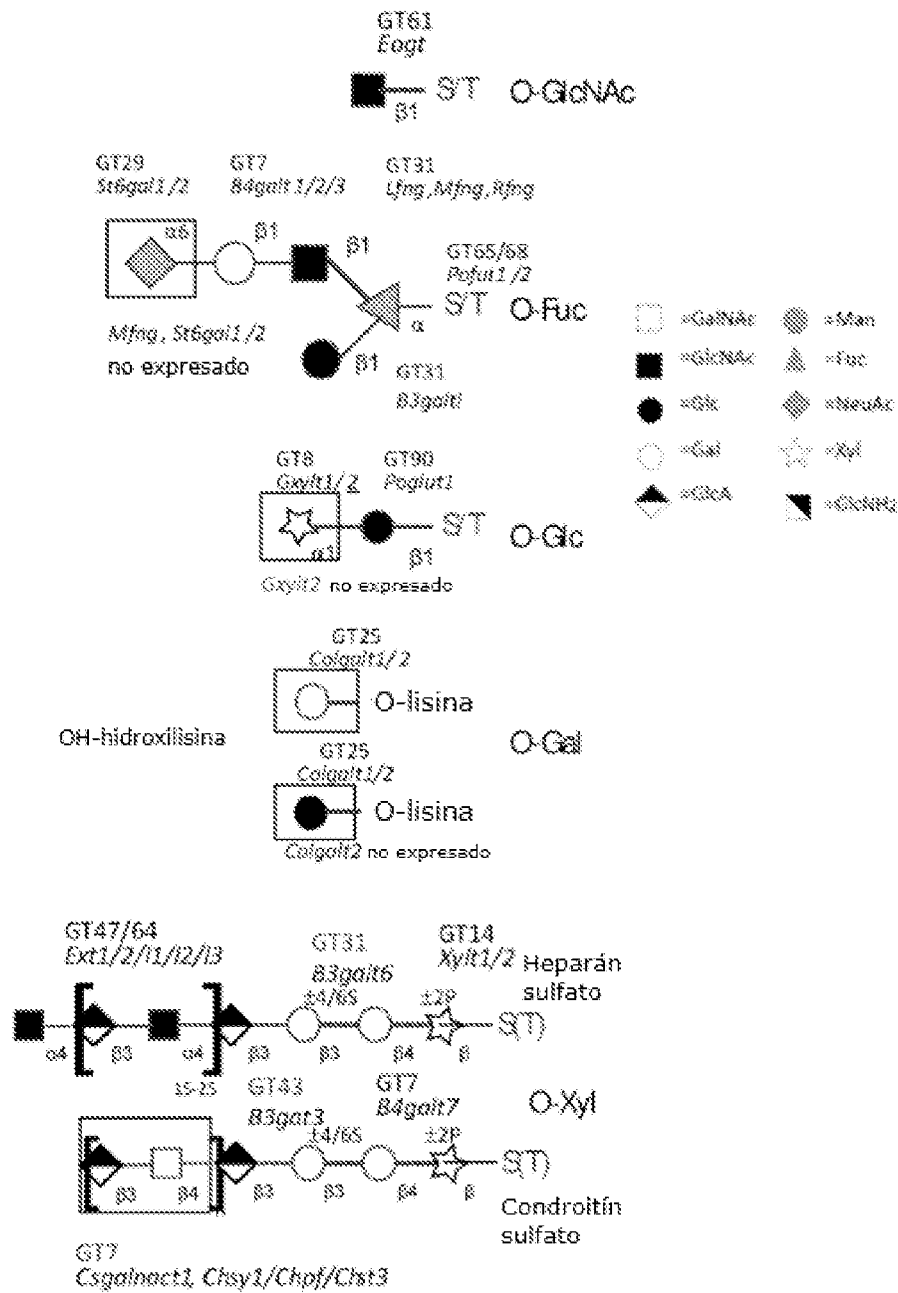


Fig. 13A3

# CHO-GS/K1 estructuras O-GalNAc

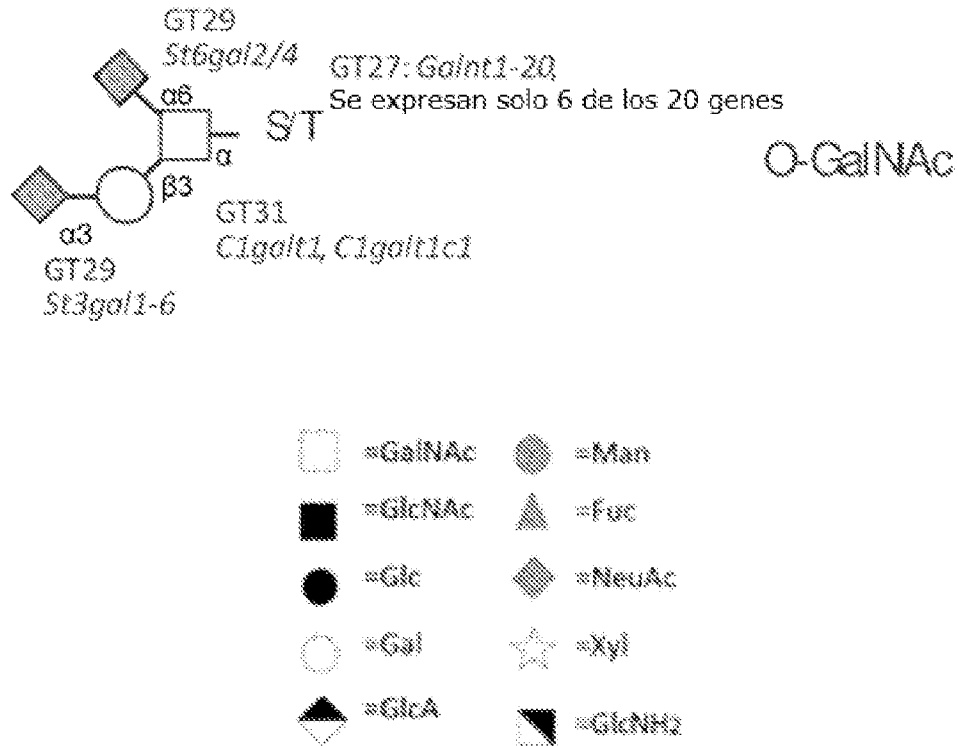


Fig. 13A4

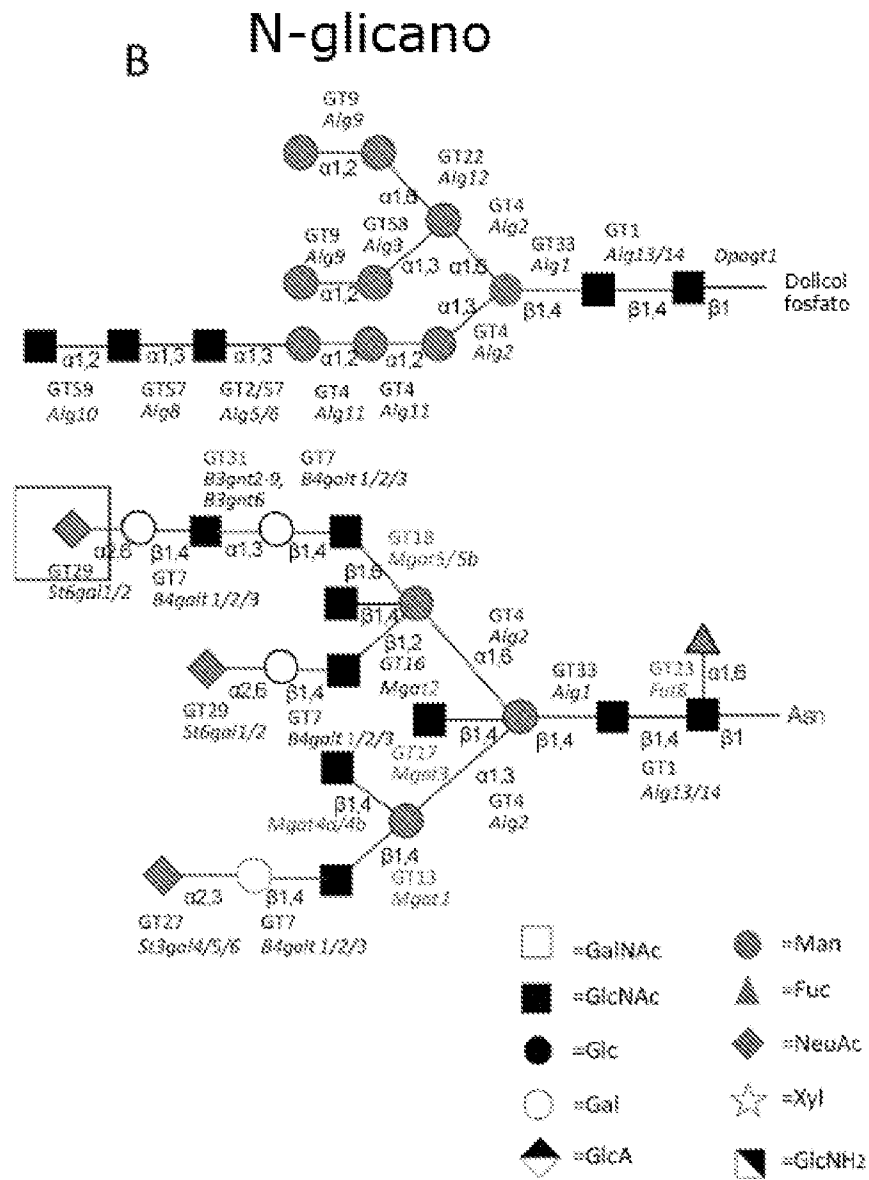


Fig. 13B

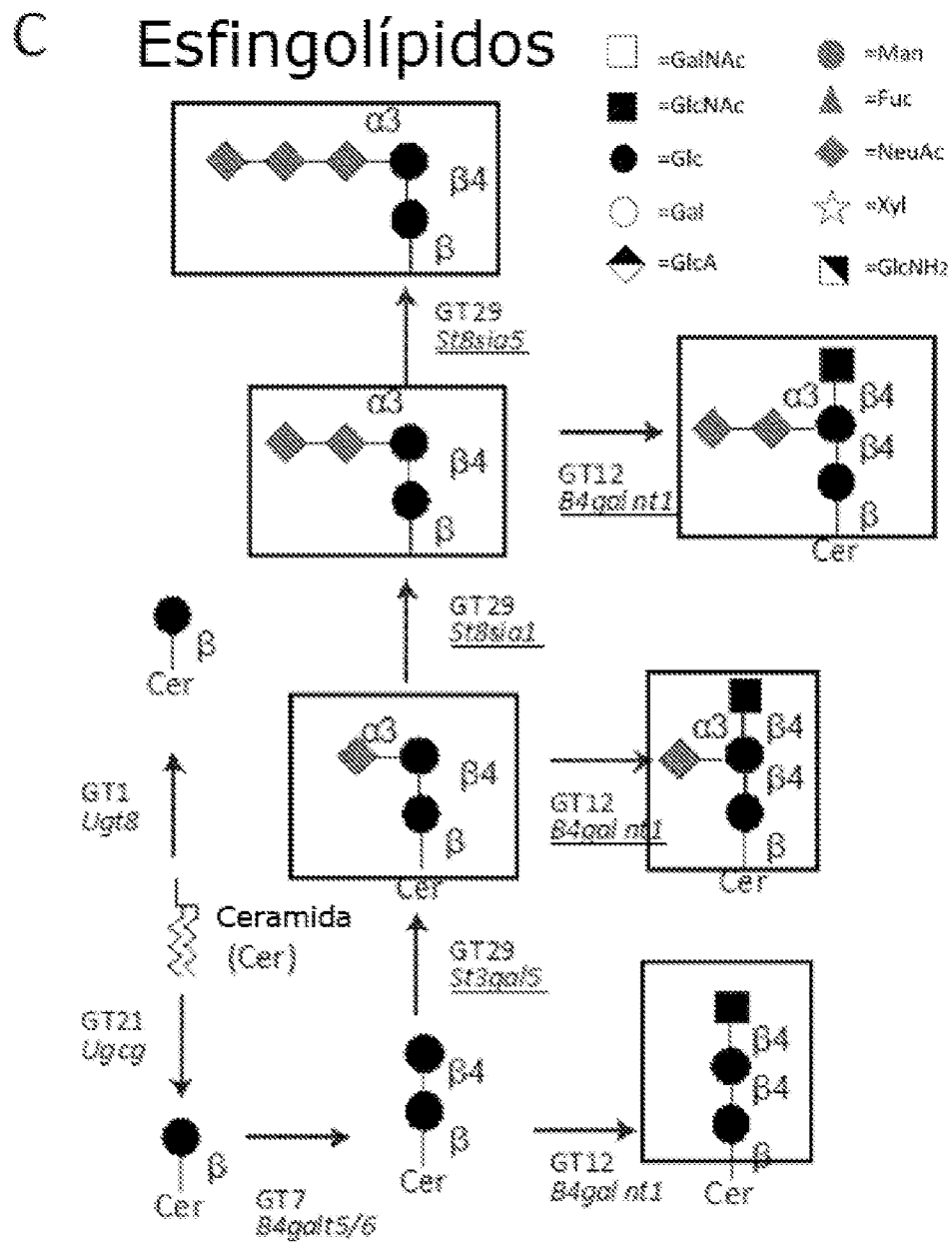
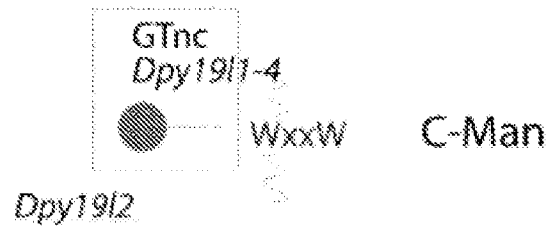


Fig. 13C

## D C-glicano



## E Anclaje GPI

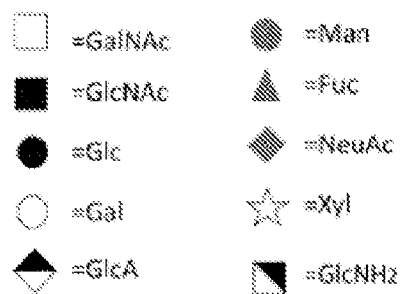
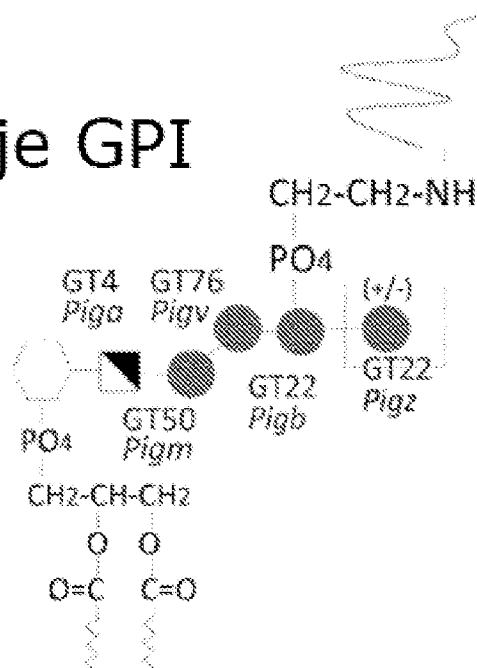


Fig. 13D-E

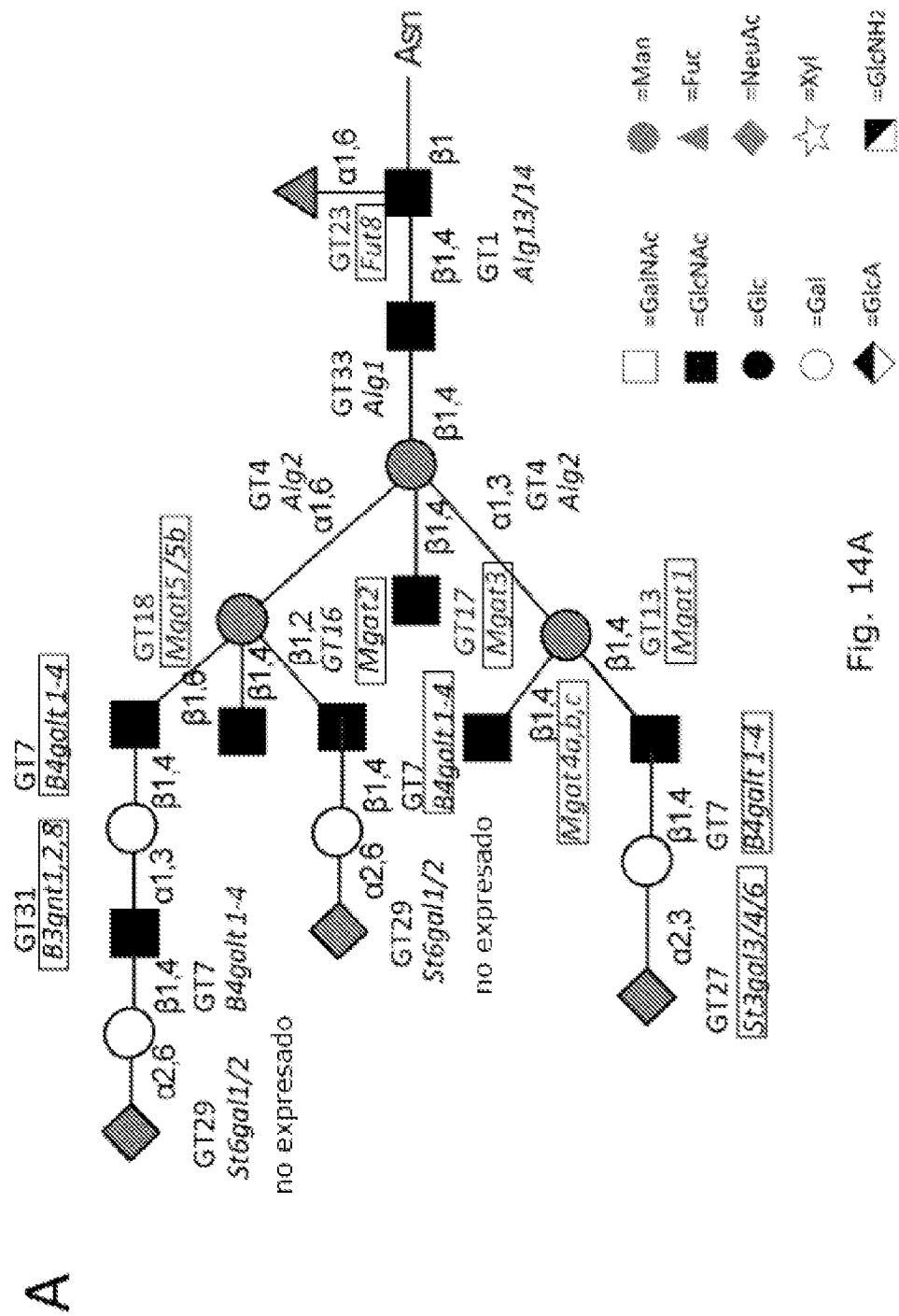


Fig. 14A

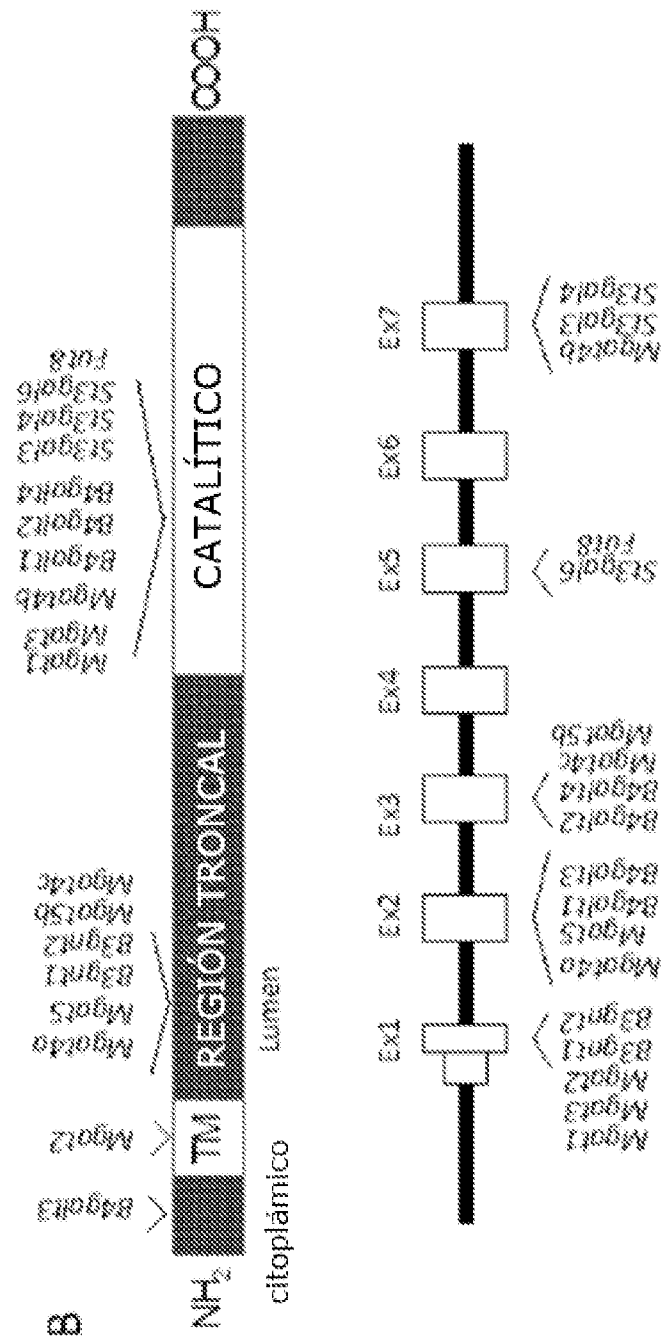


Fig. 14B

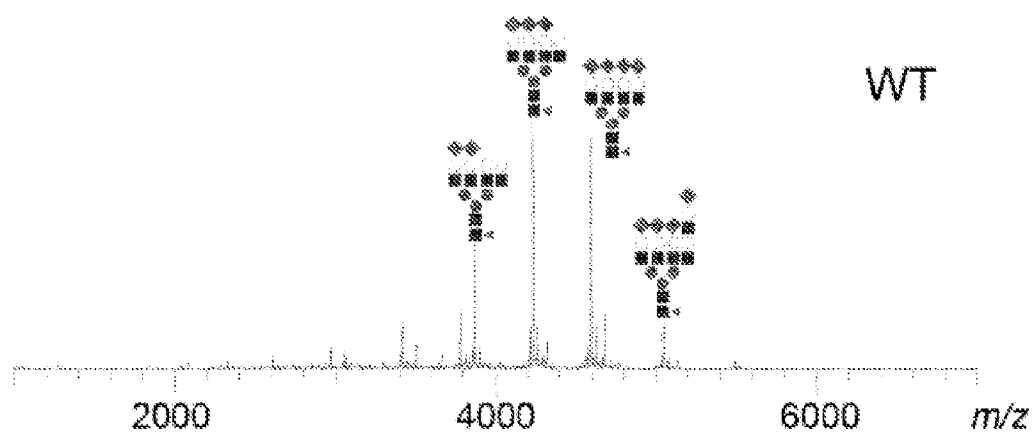


Fig. 15

**A**

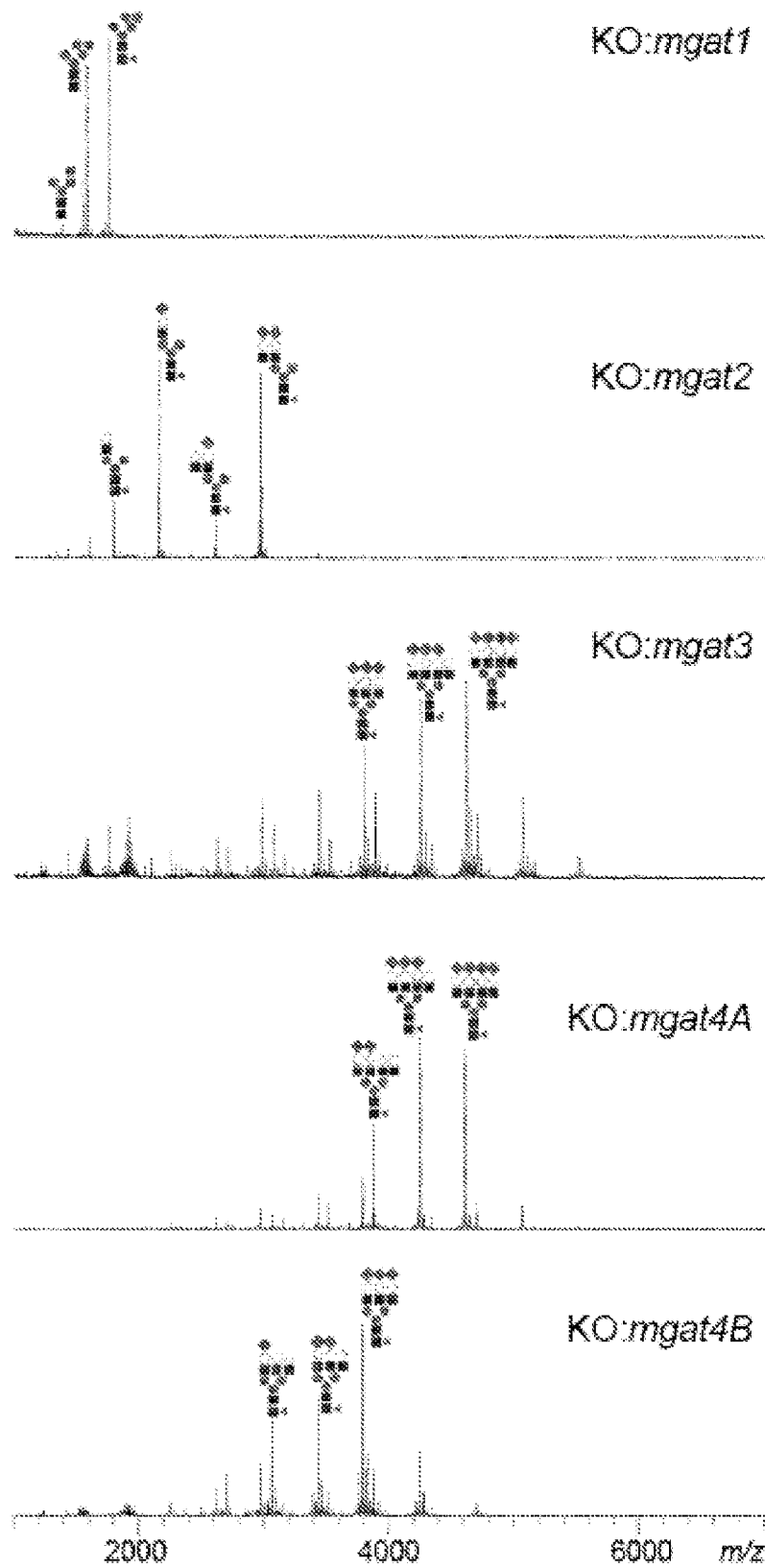


Fig. 16A

**B**

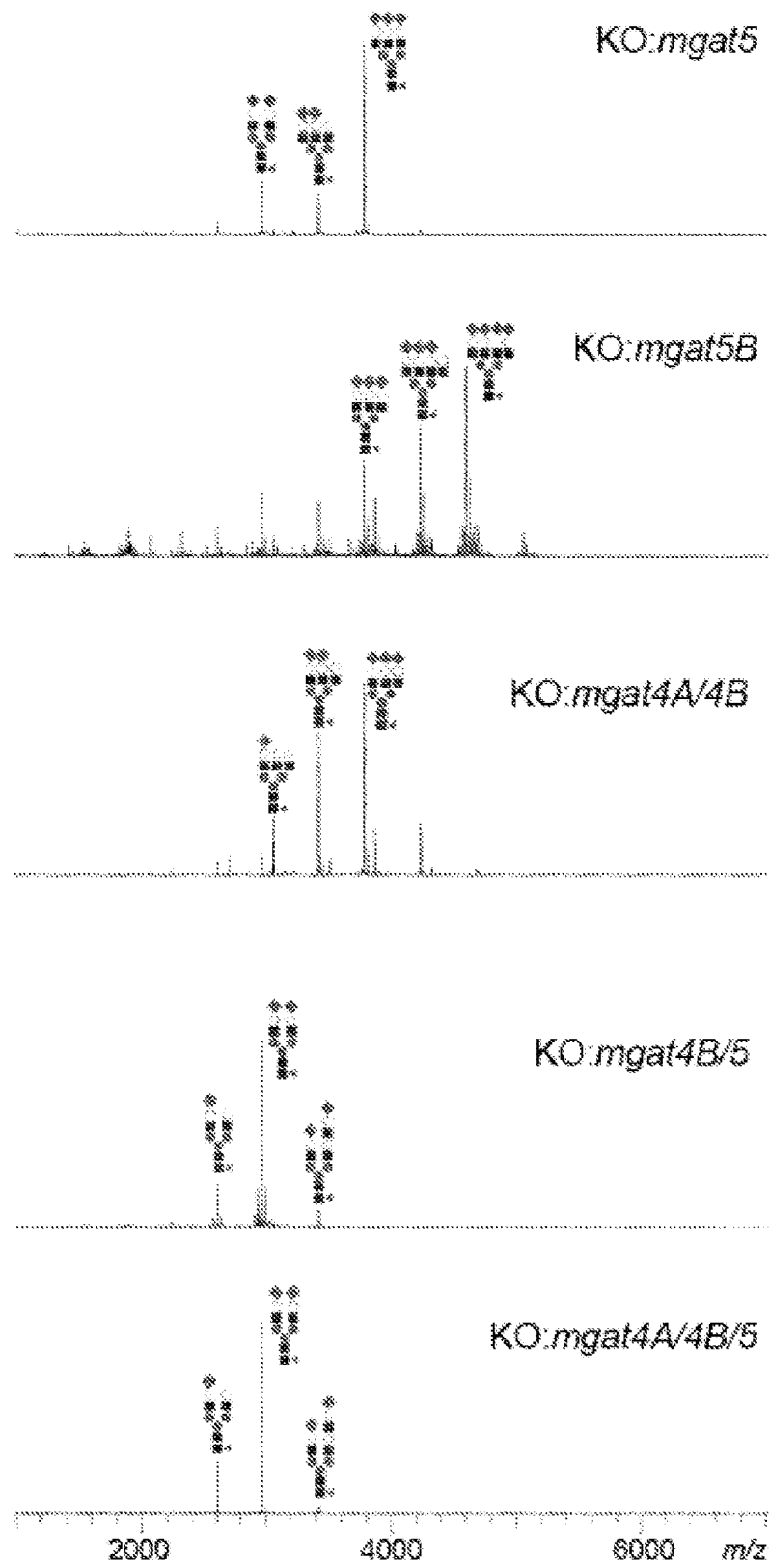


Fig. 16B

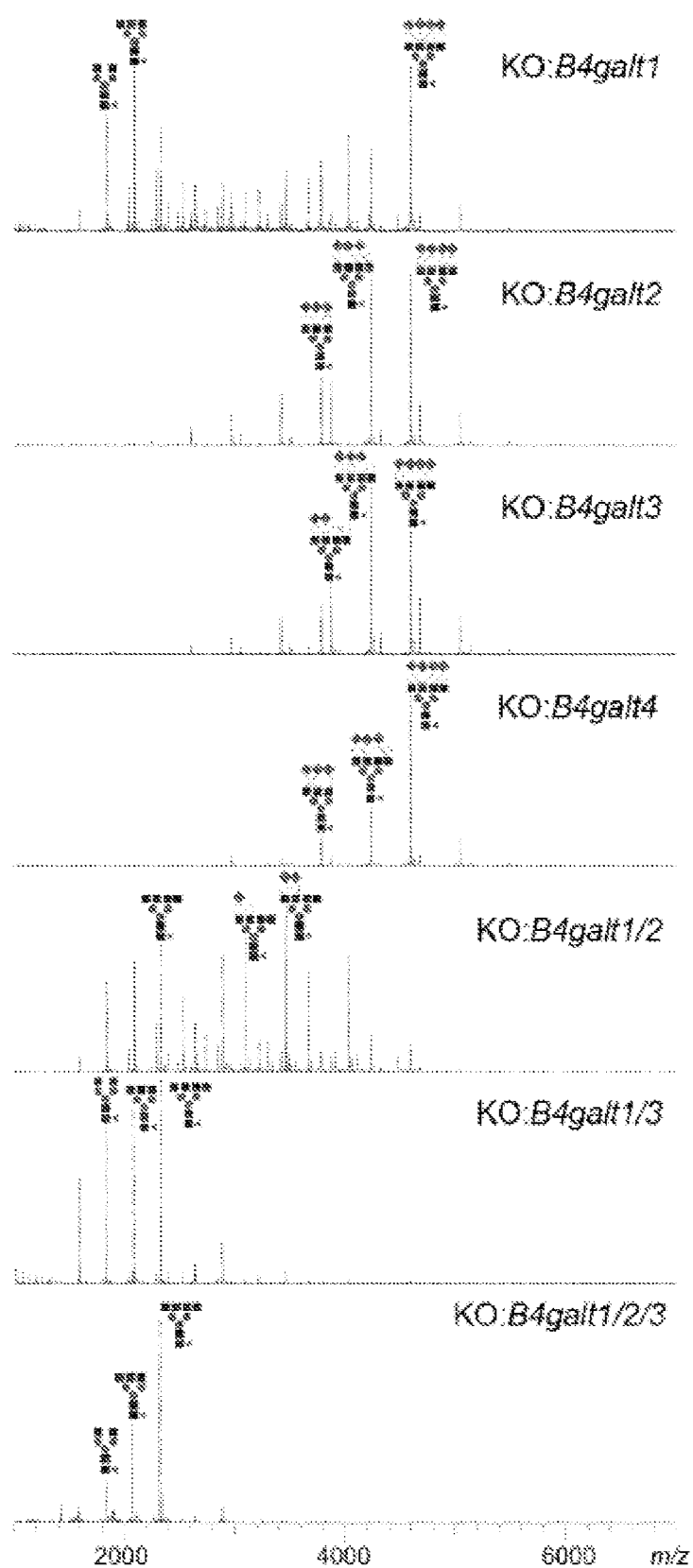


Fig. 17

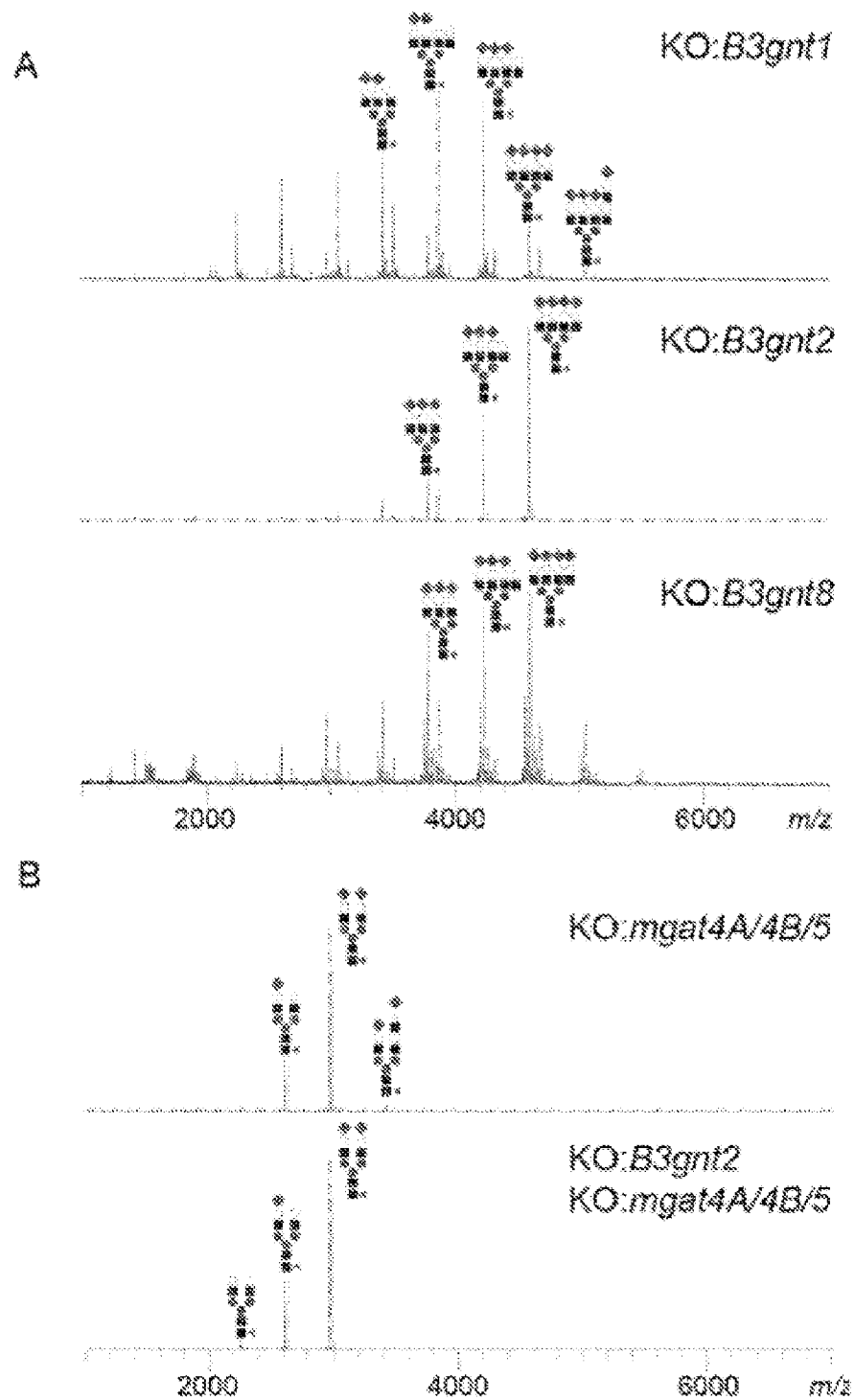


Fig. 18

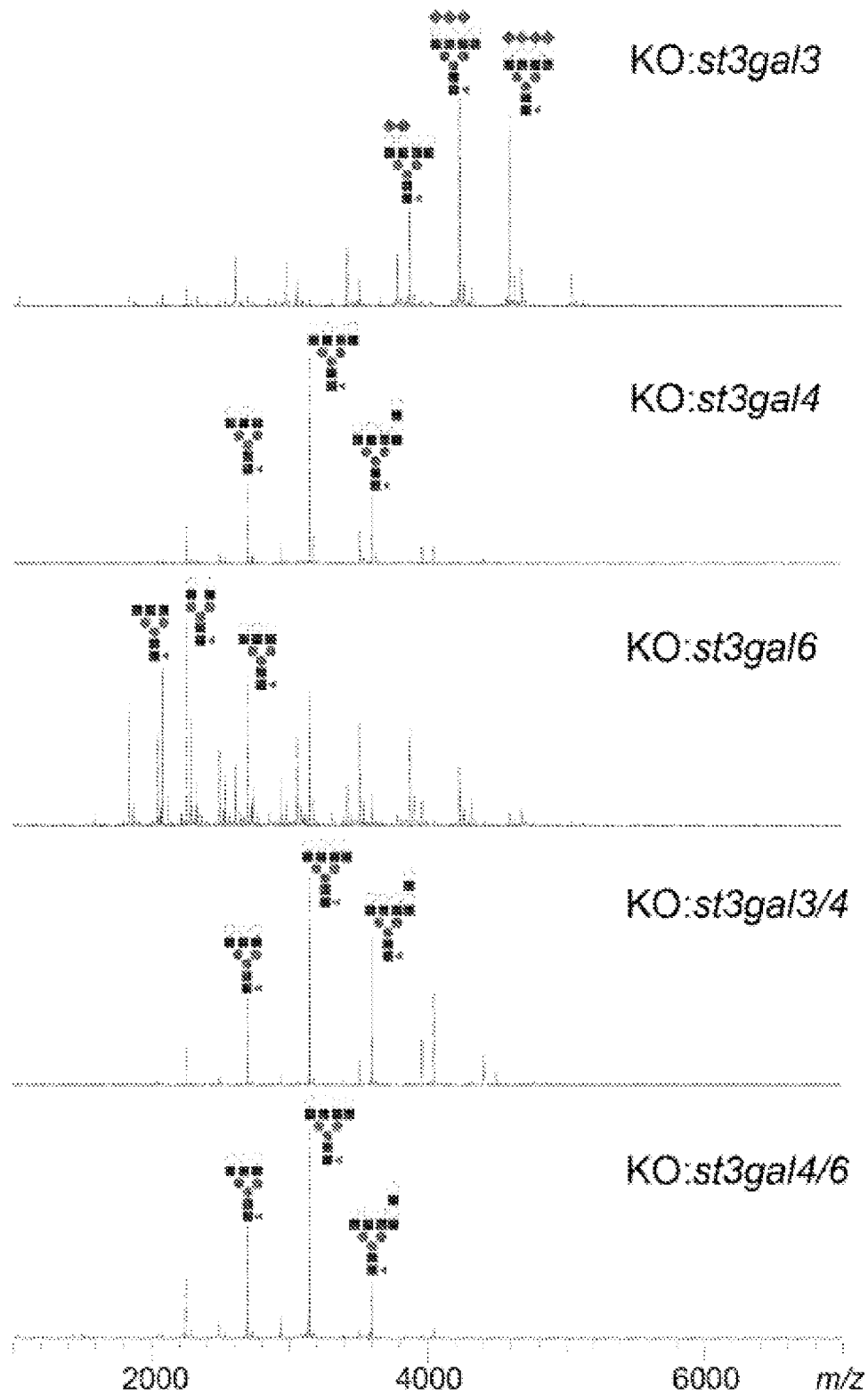


Fig. 19

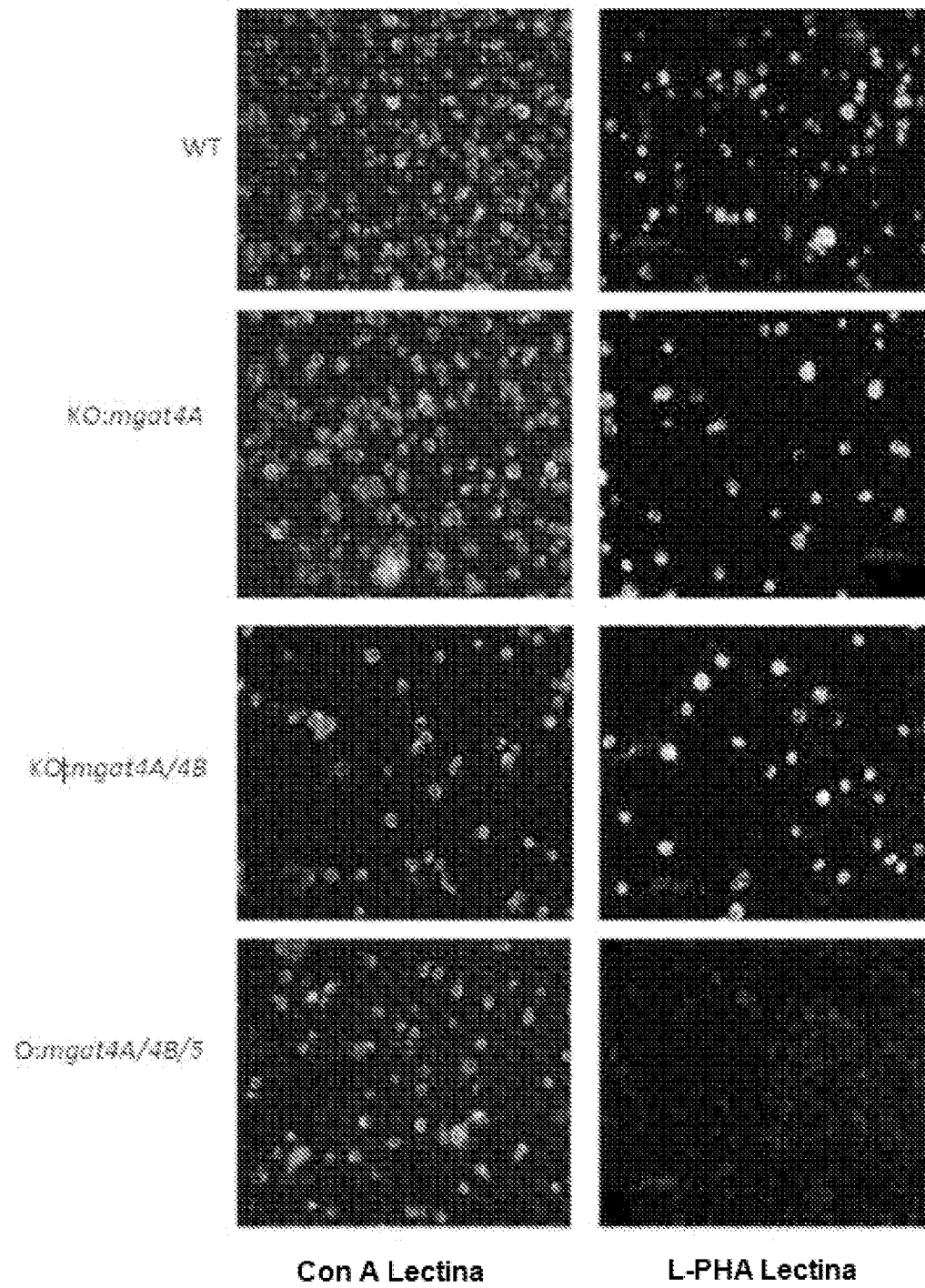


Fig. 20

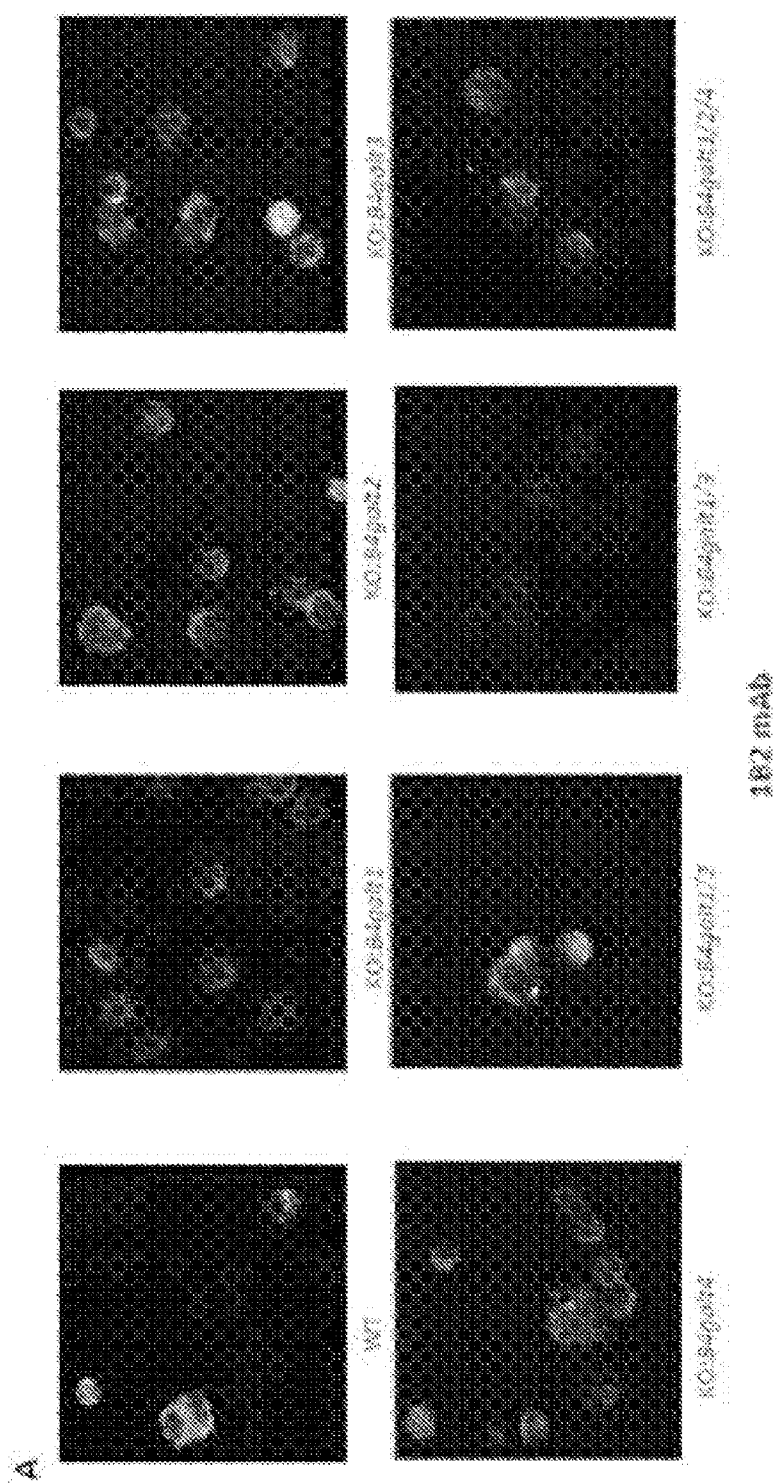


Fig. 21A

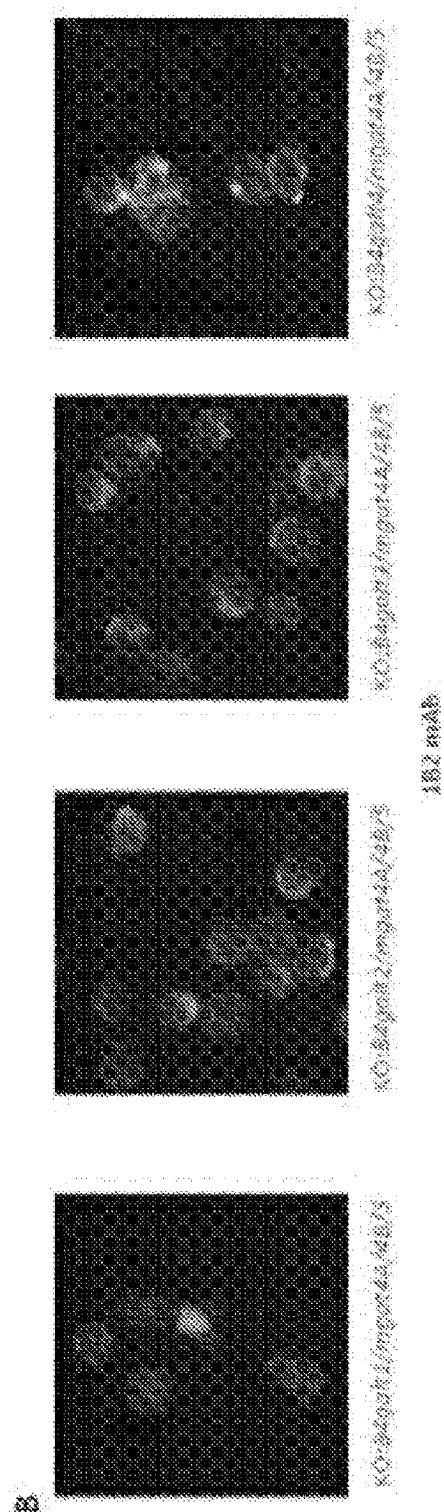


Fig. 21B

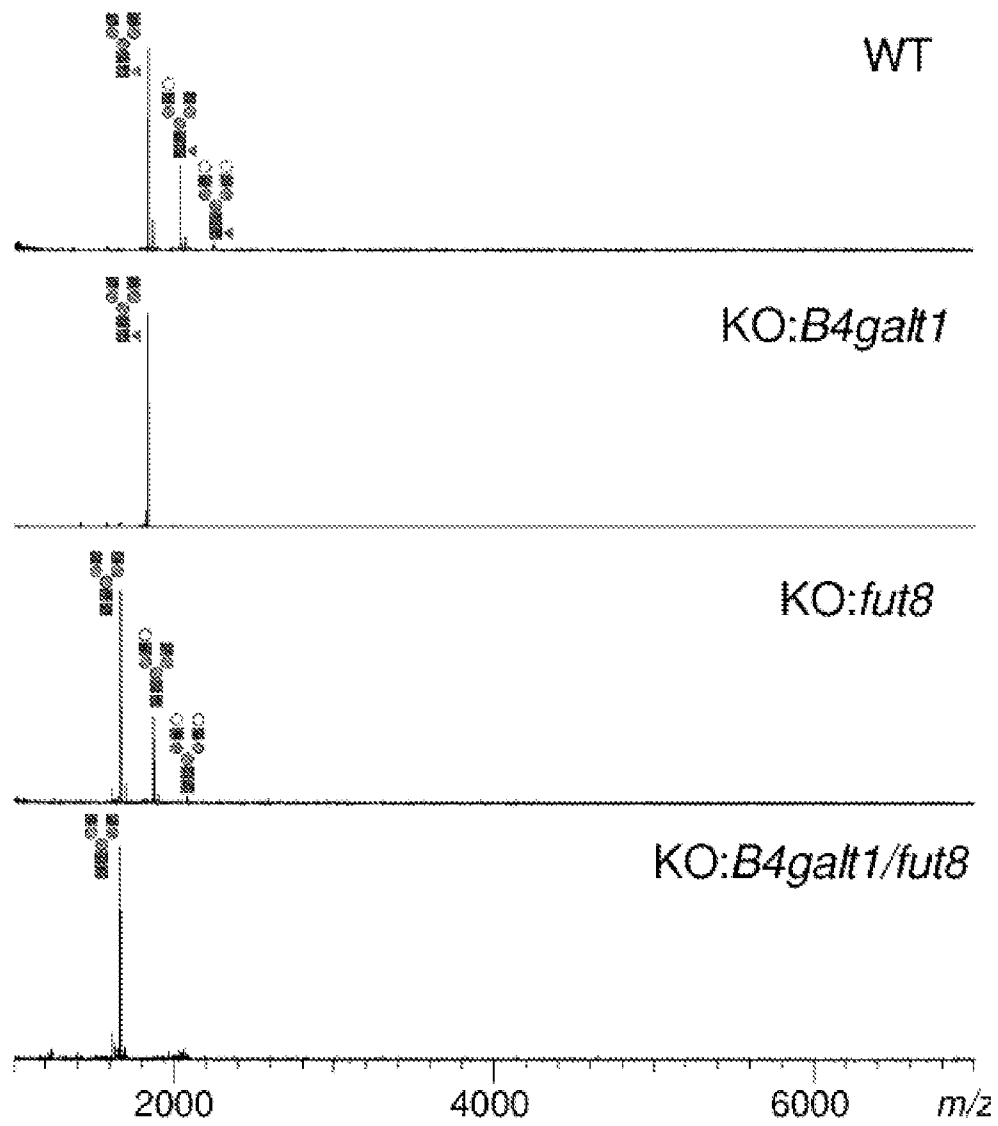


Fig. 22

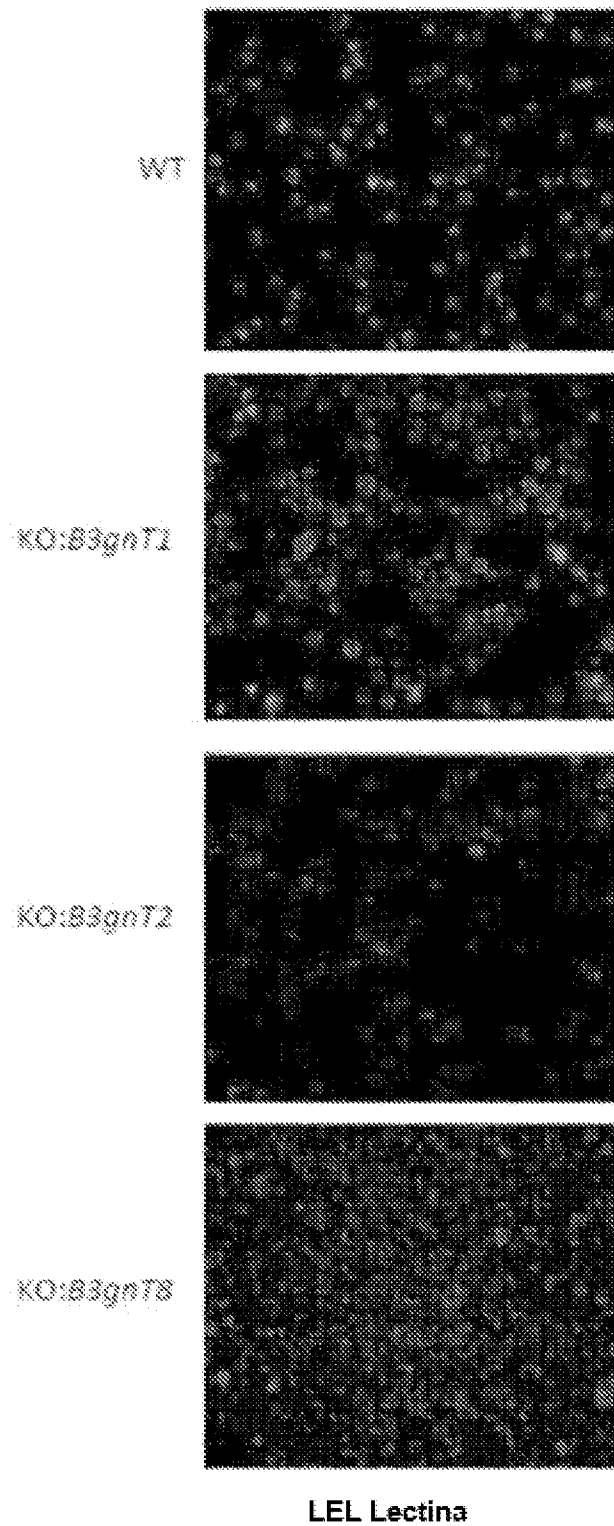


Fig. 23

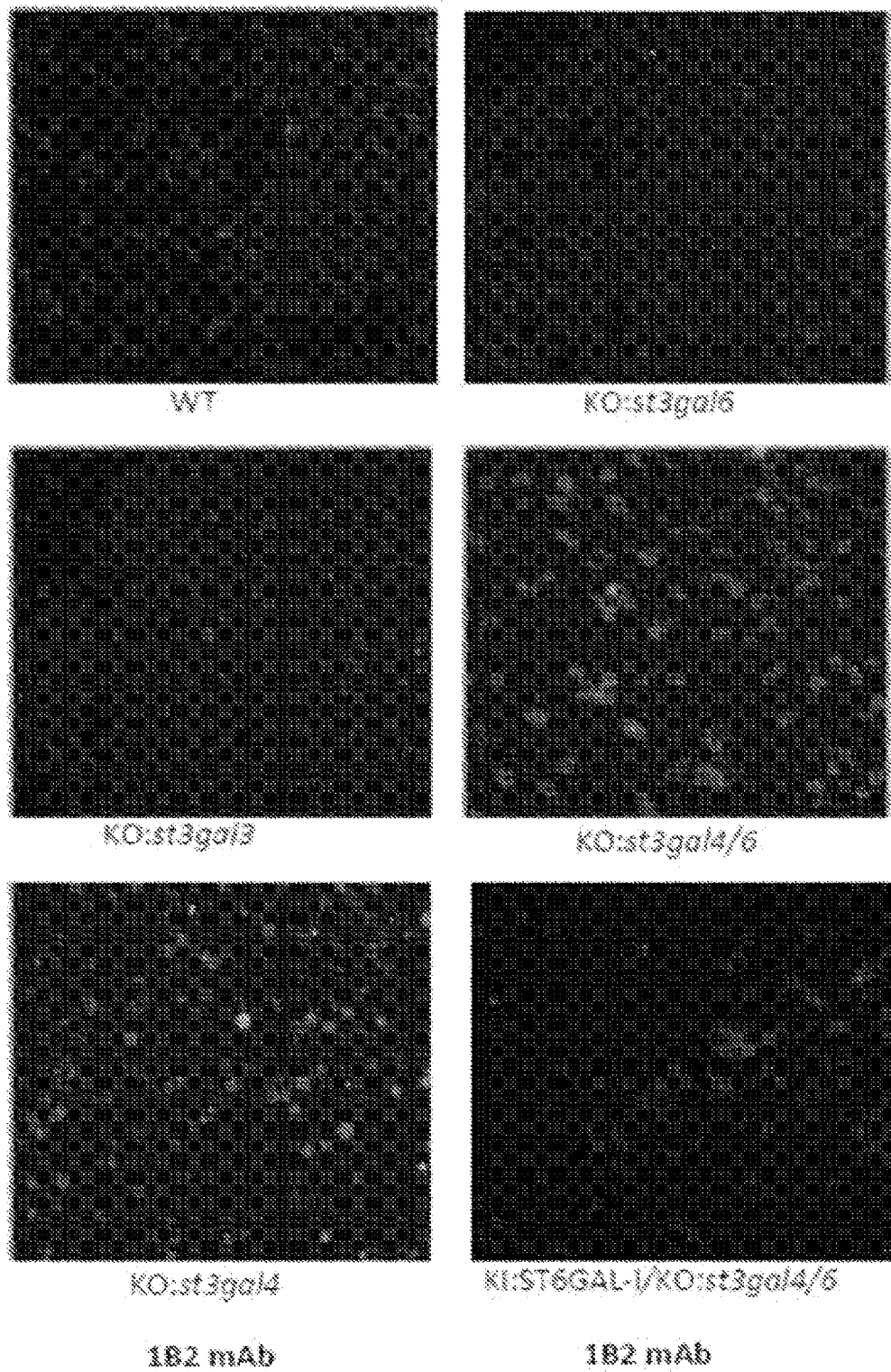


Fig. 24

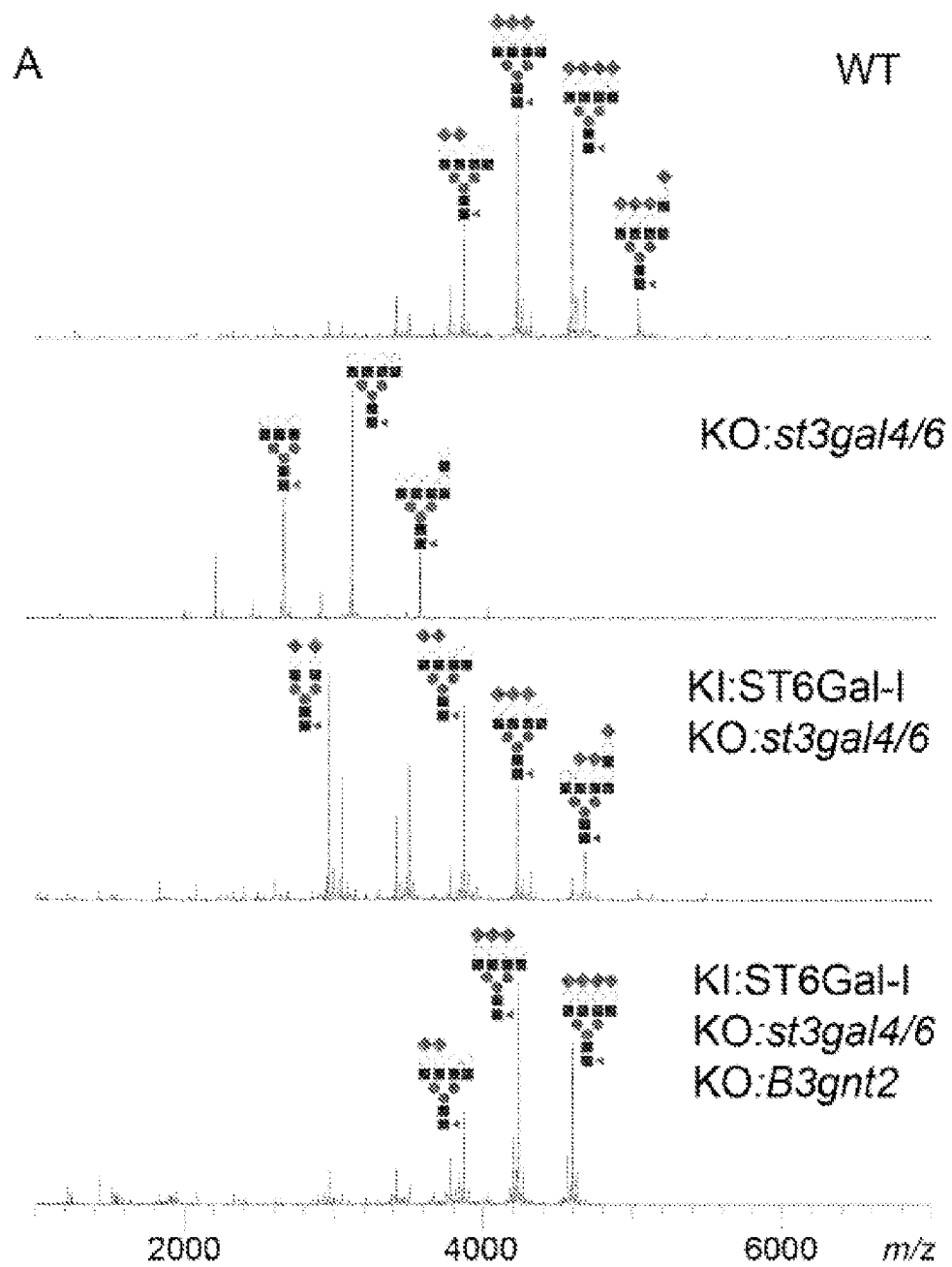


Fig. 25A

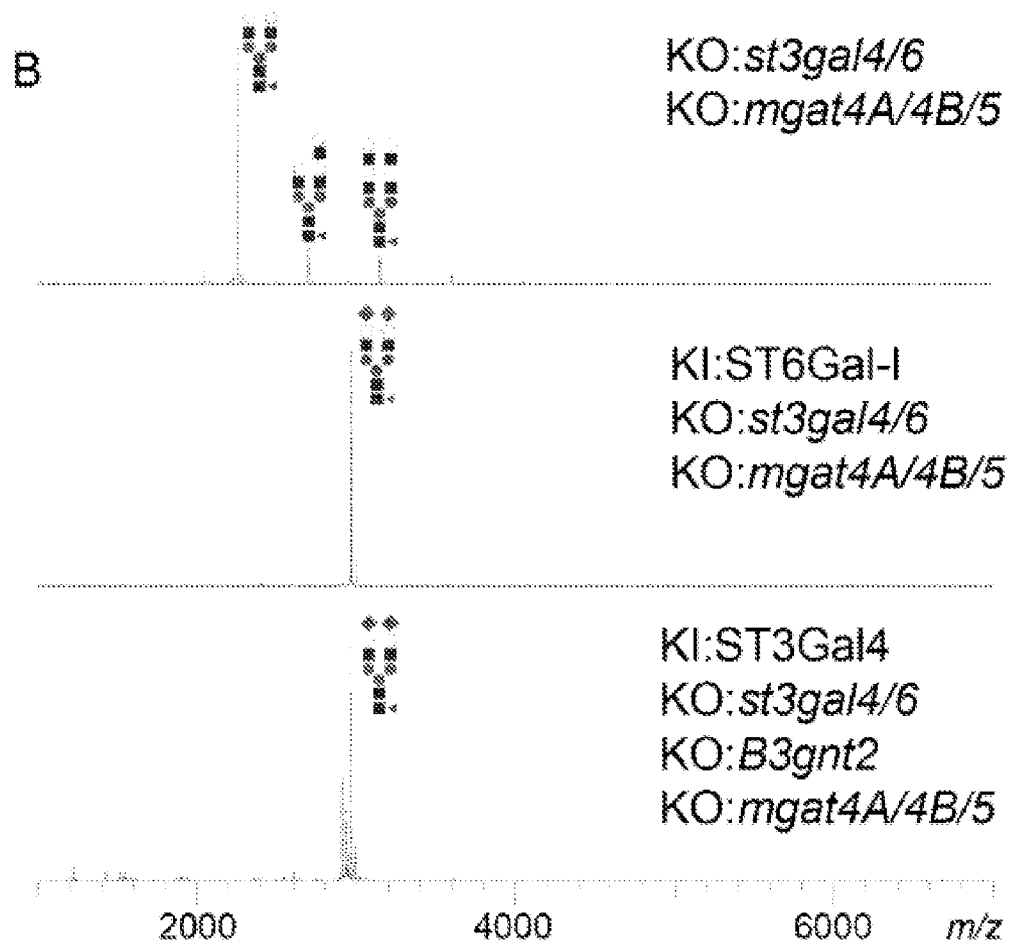


Fig. 25B

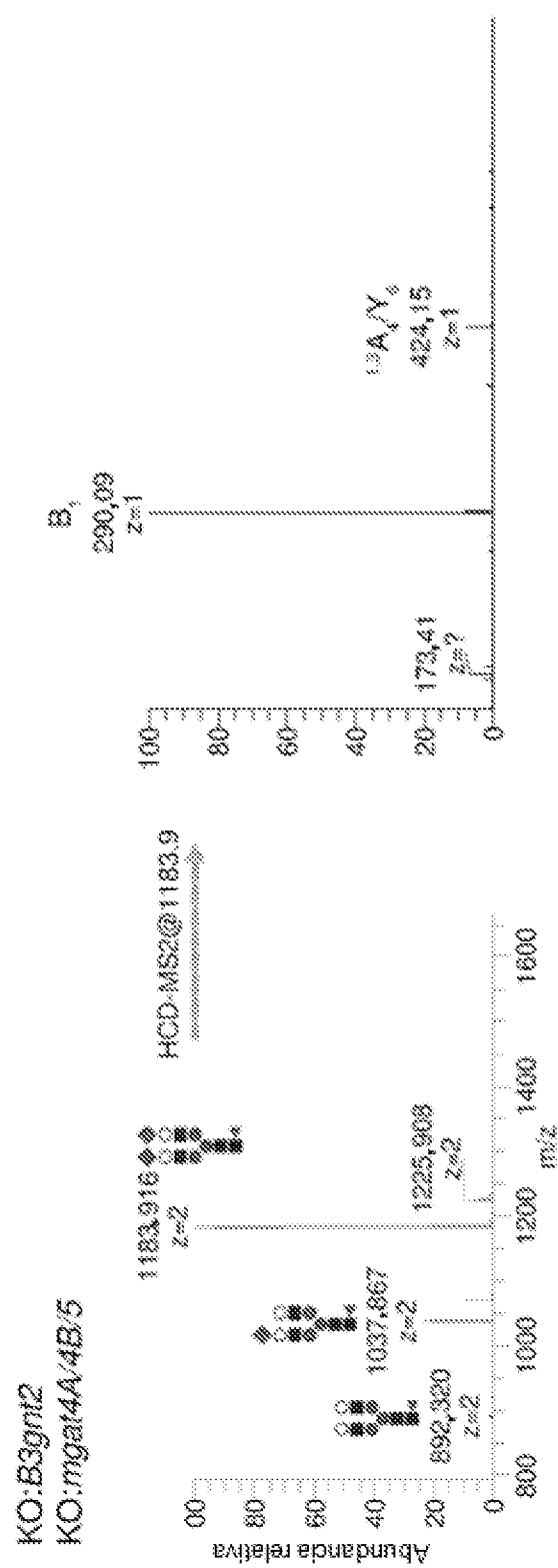


Fig. 26A

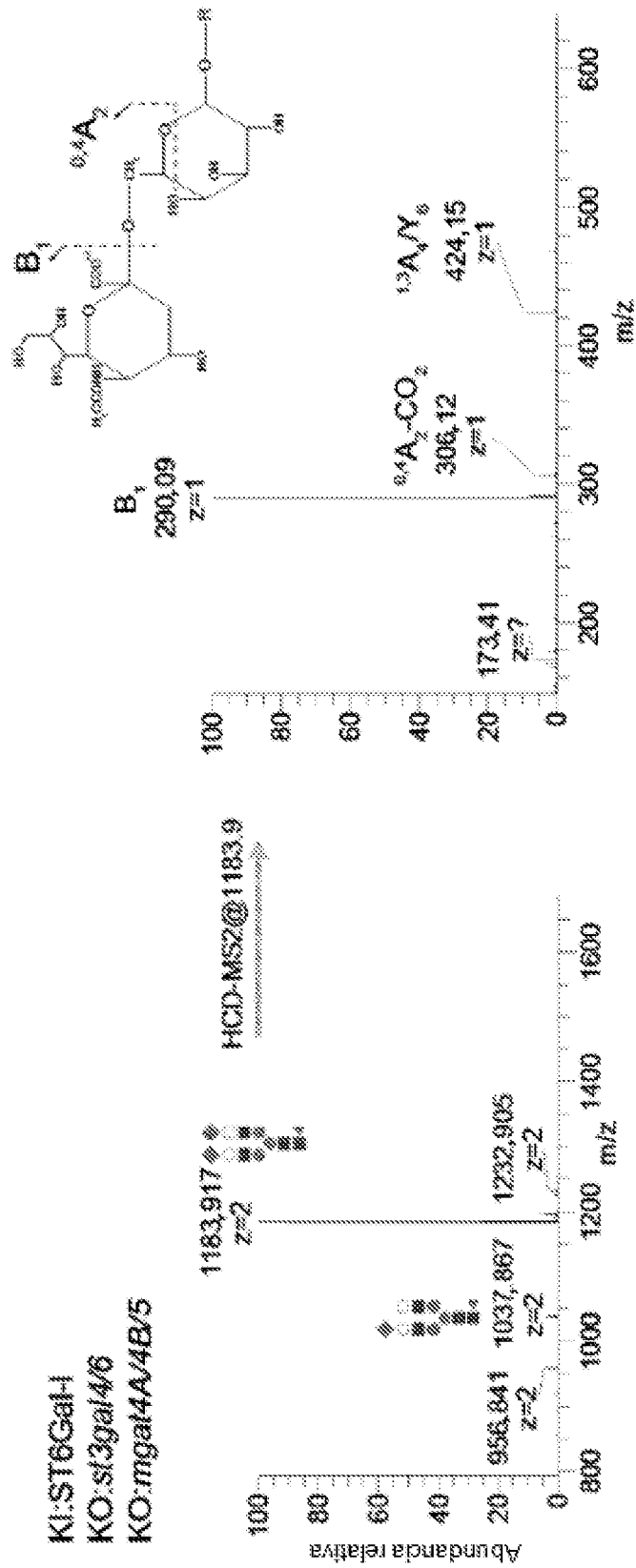


Fig. 26B

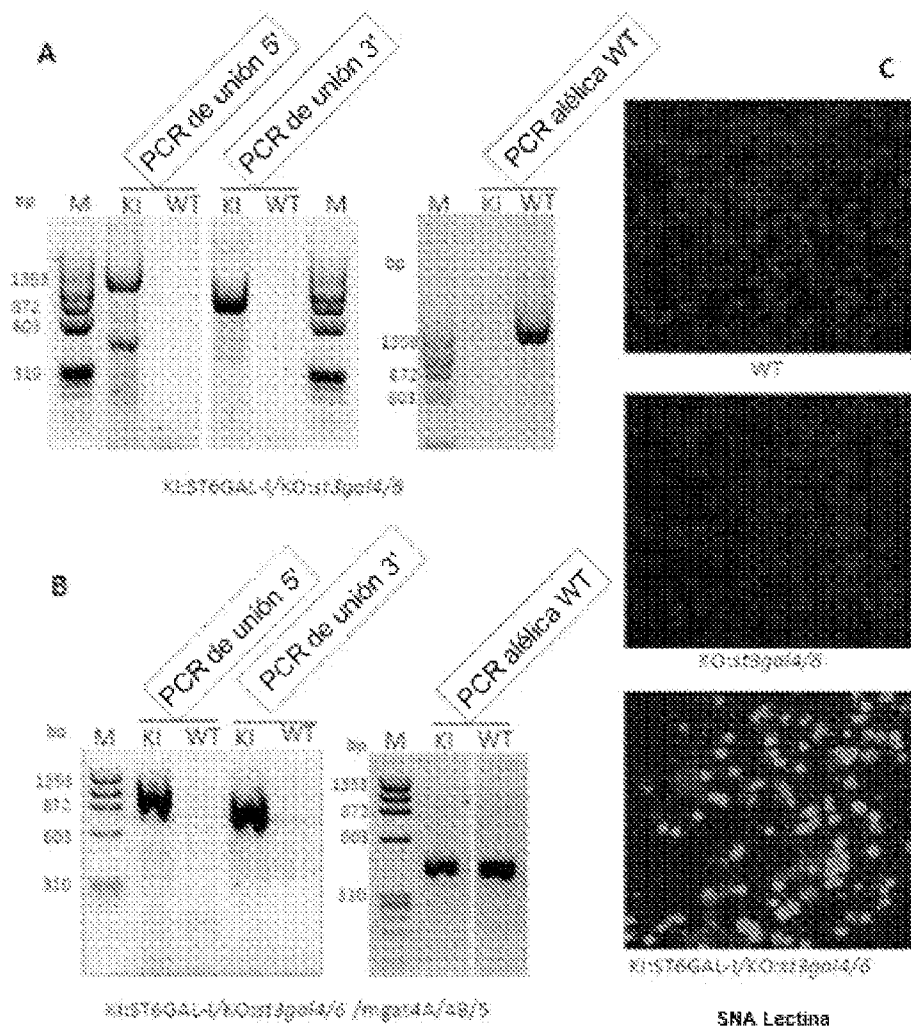


Fig. 27

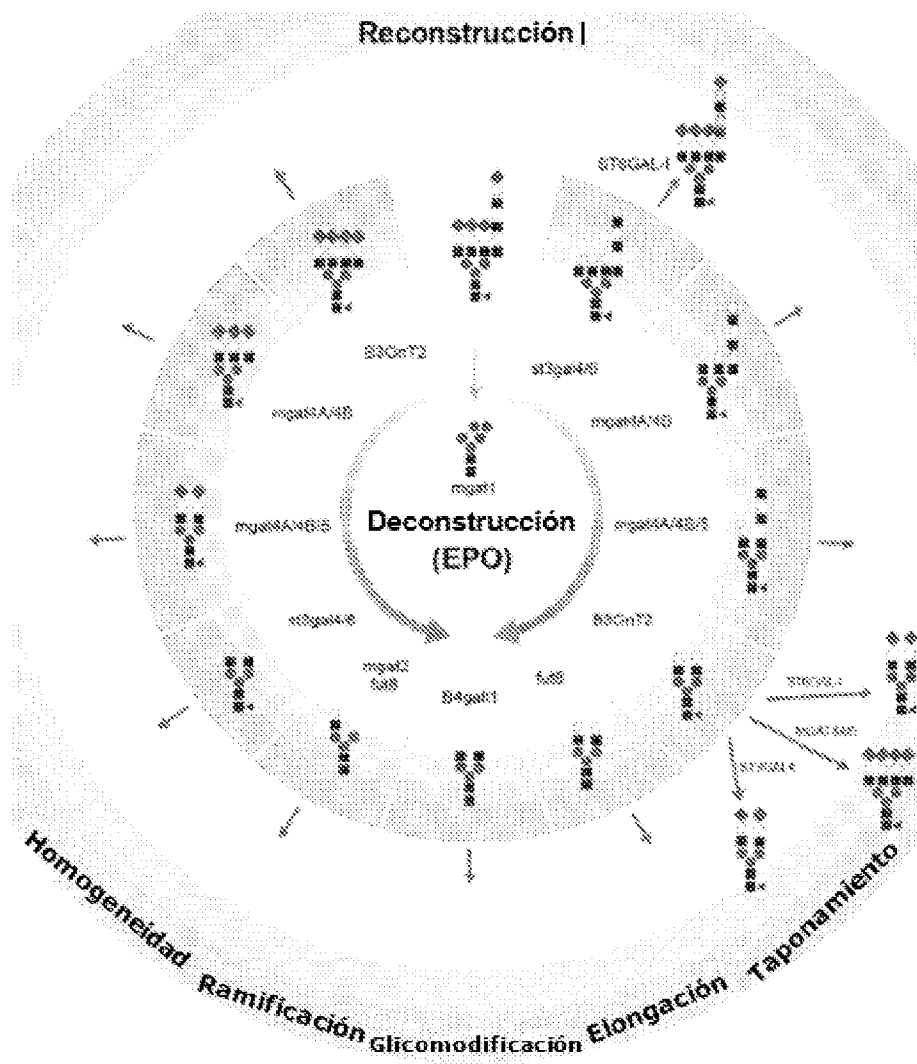


Fig. 28

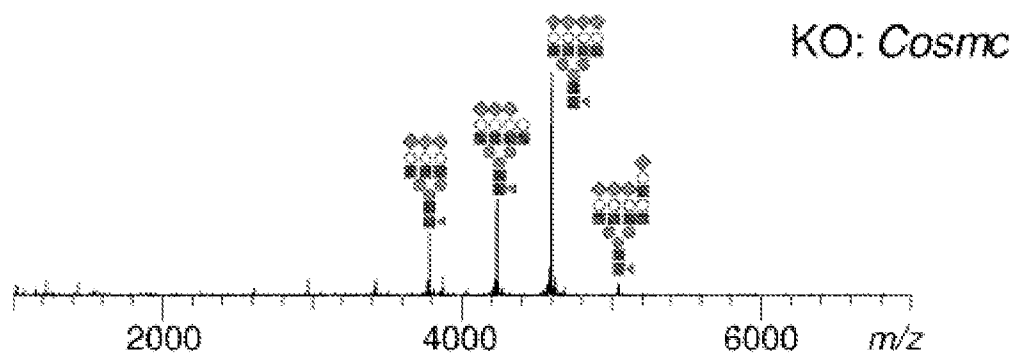


Fig. 29

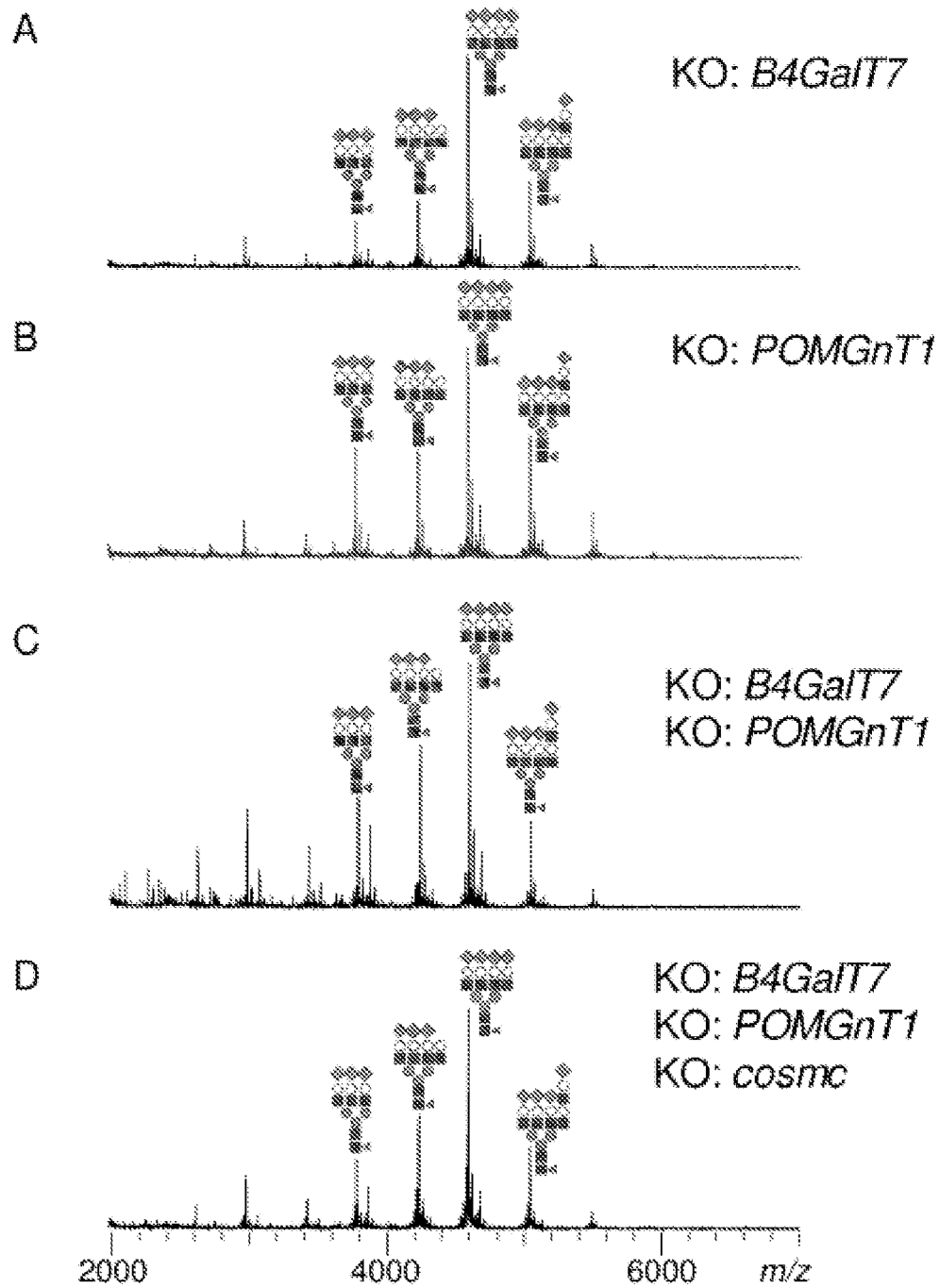


Fig. 30

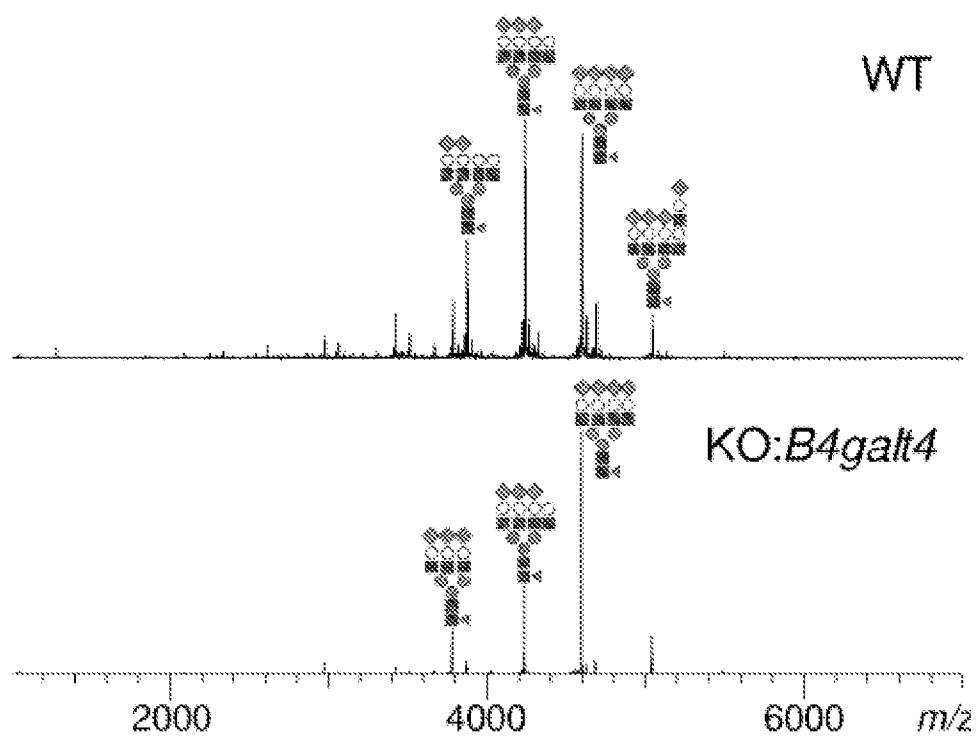


Fig. 31

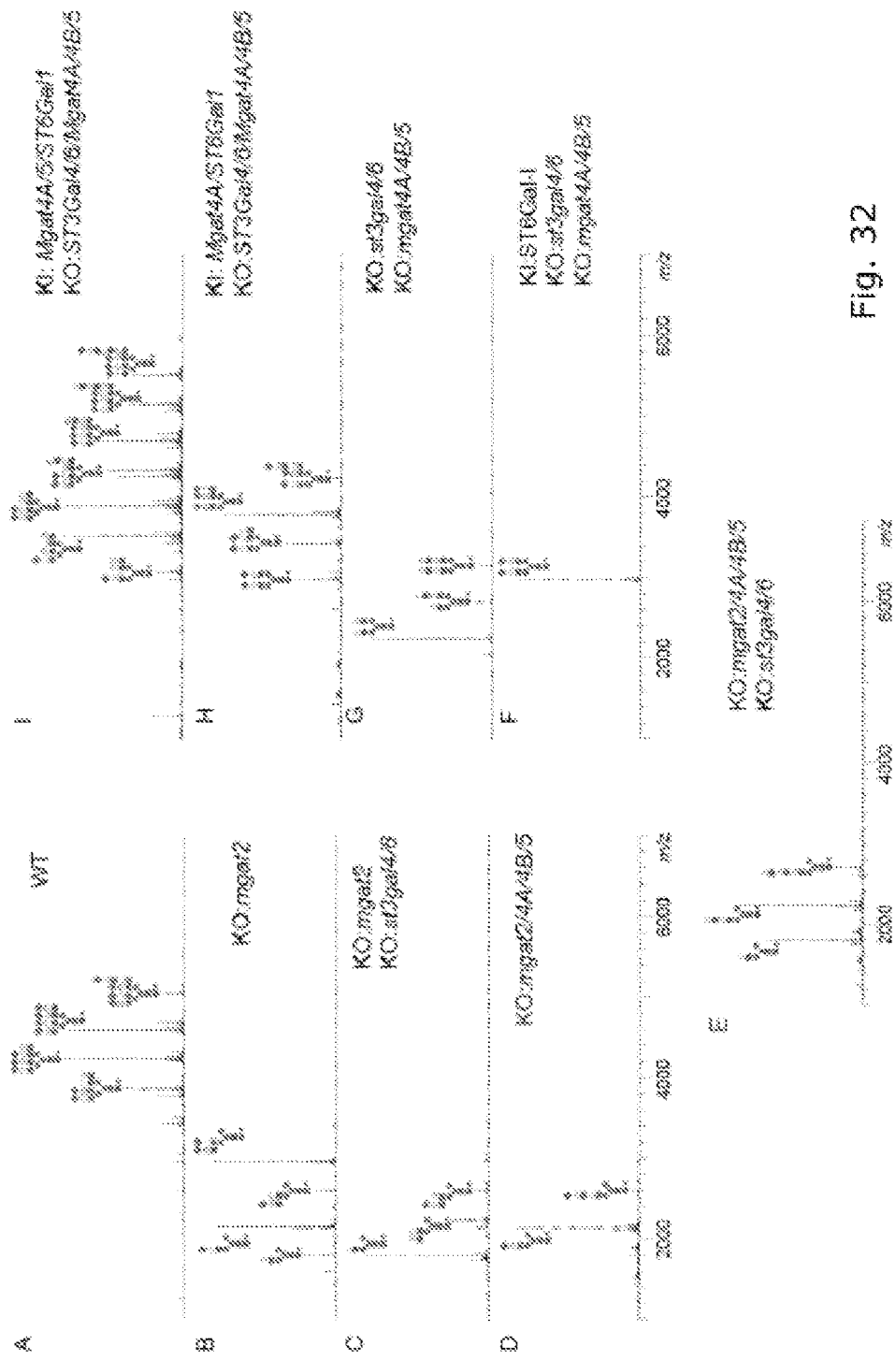


Fig. 32

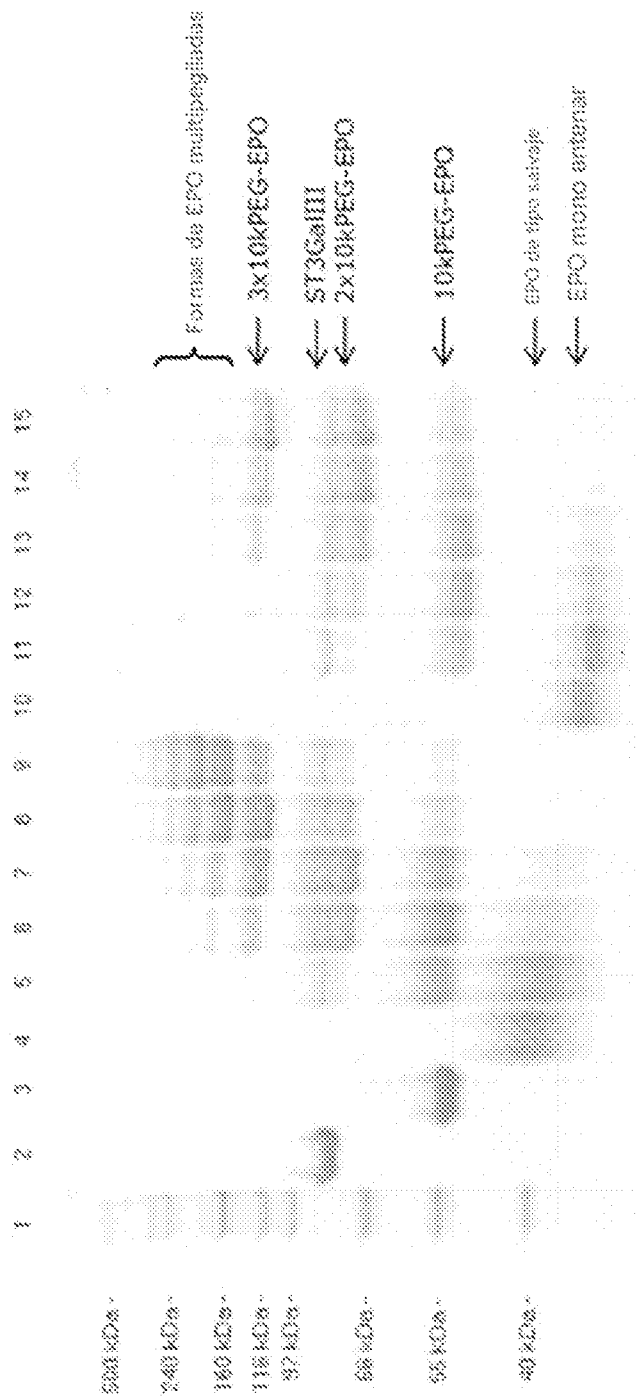


Fig. 33

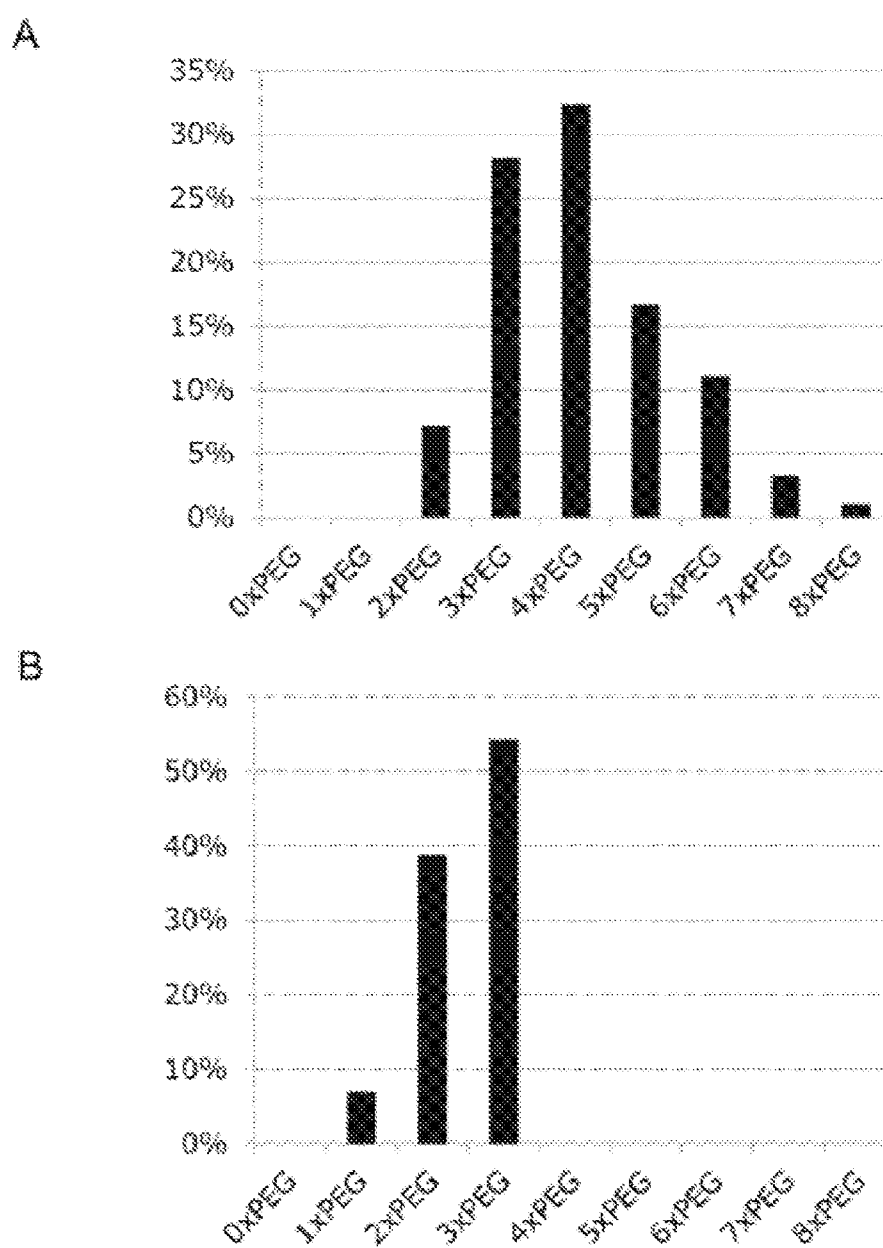


Fig. 34

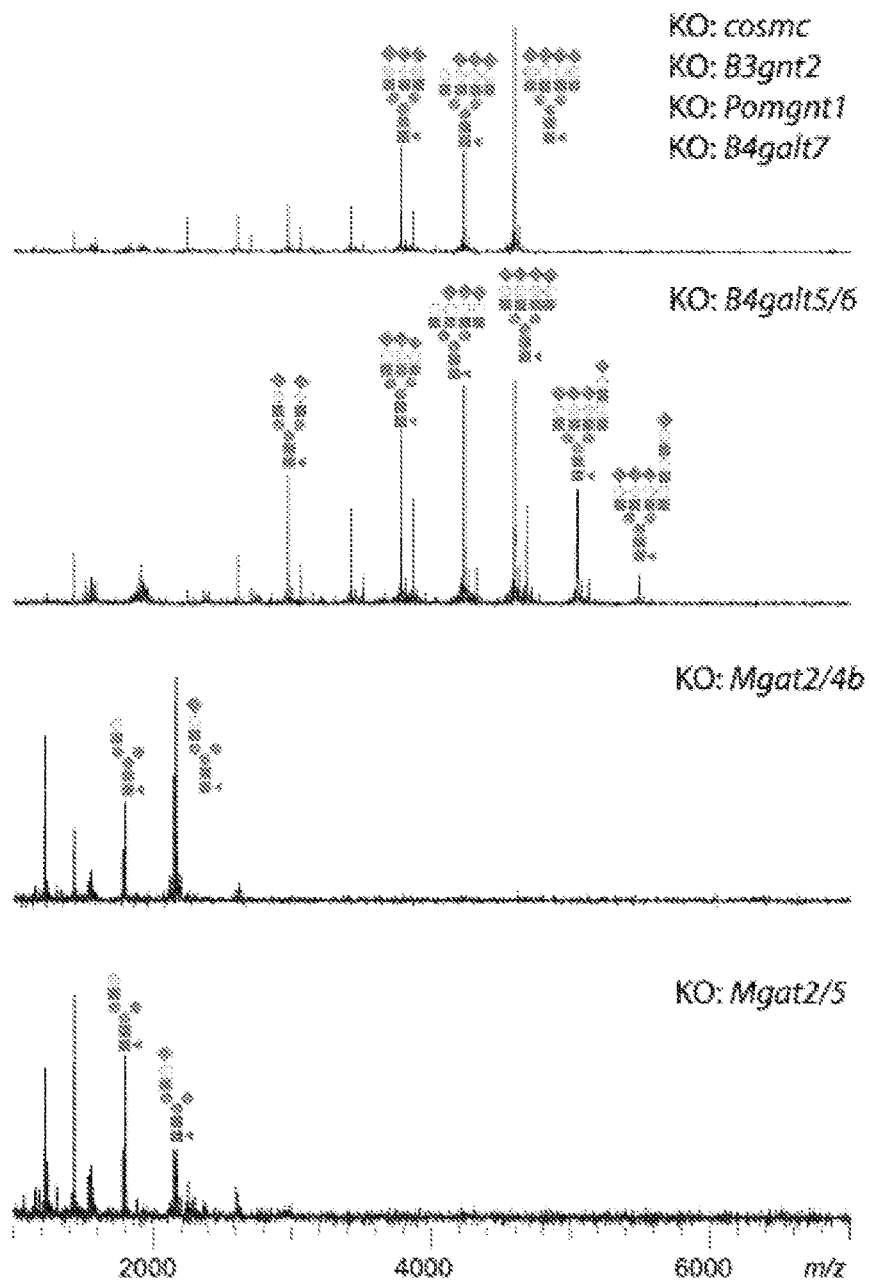


Fig. 35

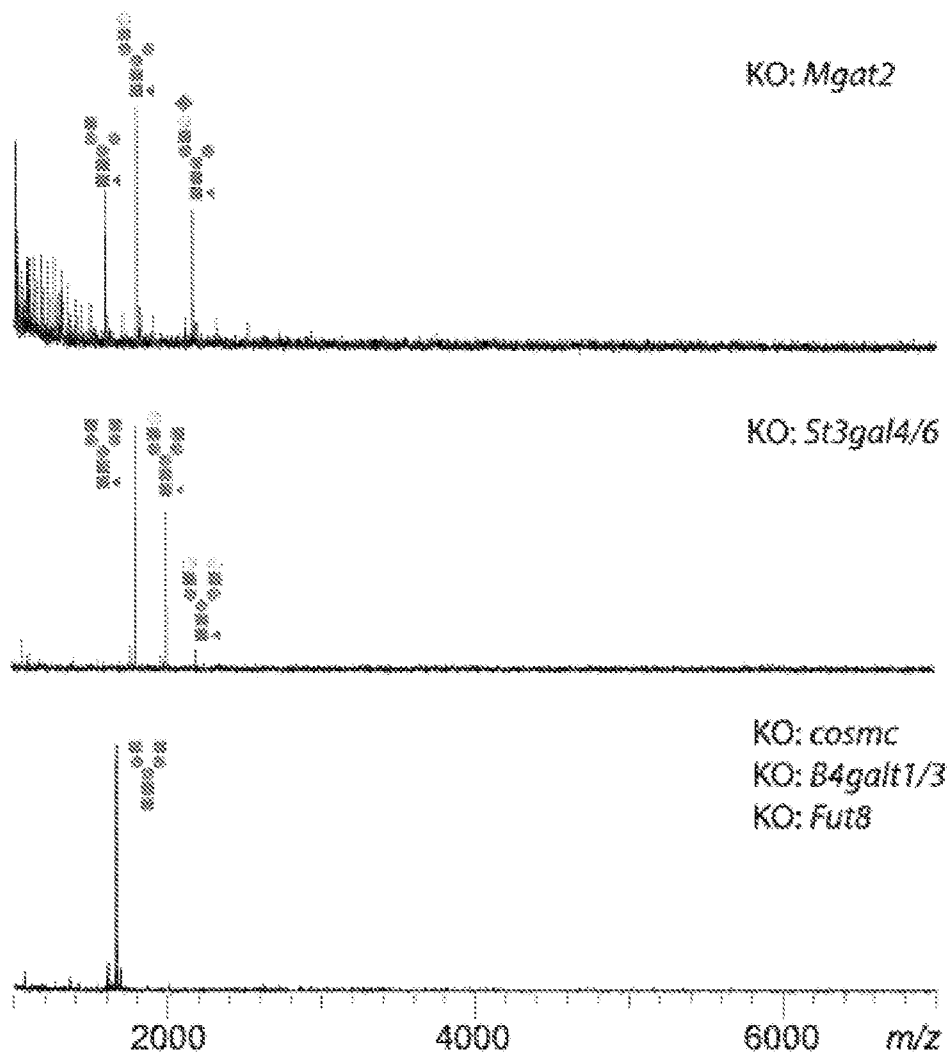


Fig. 36