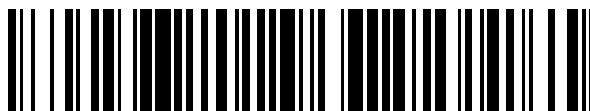


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 589 655**

51 Int. Cl.:

C12P 21/06 (2006.01)

C07K 14/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.10.2010** **PCT/US2010/052140**

87 Fecha y número de publicación internacional: **21.04.2011** **WO11046855**

96 Fecha de presentación y número de la solicitud europea: **11.10.2010** **E 10823888 (2)**

97 Fecha y número de publicación de la concesión europea: **06.07.2016** **EP 2488657**

54 Título: **Método para producir proteínas en Pichia pastoris que carecen de actividad de unión cruzada detectable a anticuerpos frente a antígenos de célula hospedadora**

30 Prioridad:

16.10.2009 US 252312 P

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

15.11.2016

73 Titular/es:

MERCK SHARP & DOHME CORP. (100.0%)
126 East Lincoln Avenue
Rahway, NJ 07065-0907, US

72 Inventor/es:

BOBROWICZ, PIOTR;
GOMATHINAYAGAM, SUJATHA;
HAMILTON, STEPHEN;
LI, HUIJUAN;
SETHURAMAN, NATARAJAN;
STADHEIM, TERRANCE, A. y
WILDT, STEFAN

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 589 655 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

Método para producir proteínas en *Pichia pastoris* que carecen de actividad de unión cruzada detectable a anticuerpos frente a antígenos de célula hospedadora

5

Antecedentes de la invención

(1) Campo de la invención

La presente invención se refiere a métodos para producir proteína y glicoproteínas en *Pichia pastoris* que carecen de actividad de unión cruzada detectable a anticuerpos formados frente a antígenos de célula hospedadora. En particular, la presente invención se refiere al uso de cepas de *Pichia pastoris* recombinantes que no presentan una actividad de β -manosiltransferasa 2 con respecto a un *N*-glicano u *O*-glicano y no presentan al menos una actividad seleccionada entre el grupo que consiste en actividad de β -manosiltransferasa 1, 3 y 4 con respecto a un *N*-glicano u *O*-glicano. Estas cepas de *Pichia pastoris* recombinantes pueden producir proteínas y glicoproteínas que carecen de restos de β -manosa resistentes a α -manosidasa detectables en las mismas. La presente invención también se refiere a métodos para producir eritropoyetina humana bi-sialilada en *Pichia pastoris* que carece de actividad de unión cruzada a anticuerpos frente a antígenos de célula hospedadora.

(2) Descripción de la técnica relacionada

La capacidad de producir proteínas humanas recombinantes ha dado lugar a grandes avances en atención de la salud humana y sigue siendo un área activa de descubrimiento de fármacos. Muchas proteínas terapéuticas requieren la adición postraducciona de glicanos a restos de asparagina específicos (*N*-glicosilación) de la proteína para asegurar la actividad de estructura-función adecuada y la posterior estabilidad en suero humano. Para uso terapéutico en seres humanos, las glicoproteínas requieren una *N*-glicosilación similar a la humana. Las líneas de células de mamíferos (por ejemplo, células CHO, células de retina humana) que puede imitar el procesamiento de glicoproteína similar a la humana tienen varios inconvenientes que incluyen títulos bajos de proteínas, tiempos de fermentación largos, productos heterogéneos y contención viral continuada. Por lo tanto, es deseable usar un sistema de expresión que no solamente produzca títulos elevados de proteína con tiempos de fermentación cortos, sino que también pueda producir glicoproteínas similares a las humanas.

Los hospedadores fúngicos tales como la levadura metilotrófica *Pichia pastoris* tienen claras ventajas para la expresión de proteína terapéutica, por ejemplo, no secretan cantidades elevadas de proteínas endógenas, están disponibles promotores inducibles fuertes para la producción de proteínas heterólogas, se pueden cultivar en medios químicos definidos y sin el uso de sueros de animal y pueden producir títulos elevados de proteínas recombinantes (Cregg *et al.*, FEMS Microbiol. Rev. 24: 45-66 (2000)). Sin embargo, las proteínas glicosiladas expresadas en *P. pastoris* por lo general contienen azúcares de manosa adicionales que dan como resultado glicanos de "alta manosa", así como grupos manosilfosfato que transmiten una carga negativa sobre las glicoproteínas. Las glicoproteínas con cualquiera de glicanos de alta manosa o mananos cargados presentan el riesgo de provocar una respuesta inmunológica no deseada en los seres humanos (Takeuchi, Trends in Glycosci. Glycotechnol. 9: S29-S35 (1997); Rosenfeld y Ballou, J. Biol. Chem. 249: 2319-2321 (1974)). Por consiguiente, es deseable producir glicoproteínas terapéuticas en células hospedadoras fúngicas en las que el patrón de glicosilación en la glicoproteína sea idéntico o similar a que se produce en glicoproteínas producidas en seres humanos y que no tengan β -manosilación detectable.

Como se pone en evidencia mediante la presencia de anticuerpos protectores en individuos no infectados, es probable que los mananos unidos en β sean inmunogénicos o que afecten de forma adversa al individuo al que se le administra una proteína o glicoproteína terapéutica que comprende mananos unidos en β . Además, los grupos de manosa expuestos en las proteínas terapéuticas se eliminan rápidamente por los receptores de manosa de las células de macrófago, lo que da como resultado una baja eficacia del fármaco. Por lo tanto, la presencia de restos de manosa unidos en β en glicanos unidos a N u O de proteínas terapéuticas heterólogas expresadas en un hospedador fúngico, por ejemplo, *P. pastoris*, no es deseable debido a su potencial inmunógeno y su capacidad para unirse a factores de eliminación.

Se ha informado que las glicoproteínas preparadas en *P. pastoris* contienen restos de manosa unidos en β . En 2003, Trimble *et al.* (Glycobiol. 14: 265-274, Epub 23 de diciembre) informaron de la presencia de restos de manosa unidos en β -1,2 en la lipasa estimulada por sales biliares humana recombinante (hBSSL) expresada en *P. pastoris*. Los genes que codifican varias β -manosiltransferasas se han identificado en *Pichia pastoris* y *Candida albicans* (Véase el documento de Patente de Estados Unidos n.º 7.465.577 y Mille *et al.*, J. Biol Chem 283: 9724-9736 (2008)).

A la vista de lo mencionado anteriormente, hay una necesidad de proporcionar métodos para preparar proteínas o glicoproteínas terapéuticas recombinantes en levadura metilotrófica tal como *Pichia pastoris* que carezcan de epítomos que pudieran provocar una reacción adversa en un individuo al que se le administra la proteína o glicoproteína terapéutica recombinante. Un método para determinar si una proteína o glicoproteína terapéutica recombinante proporciona un riesgo de provocar una reacción adversa cuando se administra a un individuo es poner

en contacto la proteína o glicoproteína terapéutica recombinante con un anticuerpo preparado frente a los antígenos totales de la célula hospedadora. Esto es de particular preocupación para las proteínas o glicoproteínas destinadas a la administración crónica. La falta de unión cruzada al anticuerpo indica que la proteína o glicoproteína terapéutica recombinante carece de actividad de unión cruzada detectable al anticuerpo y es poco probable que provoque una reacción adversa cuando se administra a un individuo. Por lo tanto, existe una necesidad de métodos para producir una proteína o glicoproteína terapéutica recombinante que carezca de actividad de unión cruzada detectable al anticuerpo y que sea poco probable que provoque una reacción adversa cuando se administra a un individuo.

Breve resumen de la invención

La invención se expone en las reivindicaciones.

En un aspecto, la presente invención proporciona un método para producir una glicoproteína recombinante en *Pichia pastoris* que carece de actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora, que comprende:

- (a) proporcionar una célula hospedadora de *Pichia pastoris* recombinante que no presente ninguna actividad de β -manosiltransferasa 1, 2, 3 y 4 con respecto a un *N*-glicano u *O*-glicano, en la que los genes de β -manosiltransferasa 1, 2 y 3 han experimentado delección o interrupción, en el que la célula hospedadora incluye una molécula de ácido nucleico que codifica la glicoproteína recombinante
- (b) cultivar la célula hospedadora en un medio en condiciones eficaces para expresar la glicoproteína recombinante; y
- (c) recuperar la glicoproteína recombinante del medio para producir la glicoproteína recombinante que carece de actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora.

En el presente documento se describe el método en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en un ELISA de sándwich.

En el presente documento se describe el método en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en una transferencia de Western.

En el presente documento se describe el método en el que la glicoproteína recombinante es una glicoproteína terapéutica.

En el presente documento se describe el método en el que la glicoproteína terapéutica se selecciona entre el grupo que consiste en eritropoyetina (EPO); citoquinas tales como interferón α , interferón β , interferón γ , e interferón ω ; y factor estimulante de colonias de granulocitos (G-CSF); GM-CSF; factores de coagulación tales como factor VIII, factor IX y proteína C humana; antitrombina III; trombina; cadena α de receptor de IgE soluble; inmunoglobulinas tales como IgG, fragmentos de IgG, fusiones de IgG, e IgM; inmunoadhesiones y otras proteínas de fusión de Fc tales como proteínas de fusión a receptor de TNF-Fc solubles; proteínas de fusión de RAGE-Fc; interleuquinas; uroquinasa; quimasa; e inhibidor de urea y tripsina; proteína de unión a IGF; factor de crecimiento epidérmico; factor de liberación de hormona de crecimiento; proteína de fusión de anexina V; angiostatina; factor 2 de crecimiento endotelial vascular; factor 1 inhibitorio de precursor mieloide; osteoprotegerina; α -1-antitripsina; α -feto proteínas; DNasa II; kringle 3 de plasminógeno humano; glucocerebrosidasa; proteína 1 de unión del TNF; hormona foliculoestimulante; antígeno 4 - Ig asociado a linfocitos T citotóxicos; activador transmembrana y modulador de calcio y ligando de ciclofilina; proteína 1 de tipo glucagón; y agonista de receptores de IL-2

En el presente documento se describe el método en el que la célula hospedadora está modificada por ingeniería genética para producir glicoproteínas que tengan *N*-glicanos similares a los humanos.

En el presente documento se describe el método en el que la célula hospedadora está modificada por ingeniería genética para producir glicoproteínas que tengan predominantemente un *N*-glicano seleccionado entre $\text{Man}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_5\text{GlcNAc}_2$, $\text{GalGlcNAcMan}_5\text{GlcNAc}_2$, $\text{NANAGalGlcNAcMan}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_3\text{GlcNAc}_2$, $\text{GlcNAc}_{(1-4)}\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_{(1-4)}\text{GlcNAc}_{(1-4)}\text{Man}_3\text{GlcNAc}_2$ y $\text{NANA}_{(1-4)}\text{Gal}_{(1-4)}\text{GlcNAc}_{(1-4)}\text{Man}_3\text{GlcNAc}_2$.

En otro aspecto, se describe el método de la presente invención para producir una eritropoyetina humana madura en *Pichia pastoris* que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora, que comprende:

- (a) proporcionar una célula hospedadora de *Pichia pastoris* recombinante modificada por ingeniería genética para producir *N*-glicanos biantenarios terminados en ácido siálico y no presenta ninguna actividad de β -manosiltransferasa 1, 2, 3 y 4 con respecto a un *N*-glicano u *O*-glicano, en la que los genes de β -manosiltransferasa 1, 2 y 3 genes han experimentado delección o interrupción, en el que la célula hospedadora incluye dos o más moléculas de ácido nucleico, cada una codificando una proteína de fusión que comprende una

eritropoyetina humana madura fusionada con un péptido señal que se dirige al RE y que se retira cuando la proteína de fusión está en el RE;

(b) cultivar la célula hospedadora en un medio en condiciones eficaces para expresar y procesar la primera y segunda proteínas de fusión; y

(c) recuperar la eritropoyetina humana madura del medio para producir la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora.

En el presente documento se describe el método en el que el péptido señal es un péptido señal α MATpre de *S. cerevisiae* o un péptido señal de lisozima de pollo.

En el presente documento se describe el método en el que al menos una molécula de ácido nucleico codifica una proteína de fusión en la que la eritropoyetina está fusionada con el péptido señal α MATpre de *S. cerevisiae* y al menos una molécula de ácido nucleico codifica una proteína de fusión en la que la eritropoyetina está fusionada con el péptido señal α MATpre de *S. cerevisiae* o un péptido señal de lisozima de pollo.

En el presente documento se describe el método en el que los codones de la secuencia de ácidos nucleicos de la molécula de ácido nucleico que codifica la eritropoyetina están optimizados para su expresión en *Pichia pastoris*.

En el presente documento se describe el método en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en un ELISA de sándwich.

En el presente documento se describe el método en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en una transferencia de Western.

En el presente documento se describe el método en el que la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora del medio incluye una etapa de cromatografía de intercambio catiónico.

En el presente documento se describe el método en el que la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora del medio incluye una etapa de cromatografía con hidroxipatita.

En el presente documento se describe el método en el que la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora del medio incluye una etapa de cromatografía de intercambio aniónico.

Definiciones

Como se usa en el presente documento, los términos "*N*-glicano" y "glicoforma" se usan indistintamente y se refieren a un oligosacárido unido a N, por ejemplo, uno que está unido por un enlace de asparagina-*N*-acetilglucosamina a un resto de asparagina de un polipéptido. Las glicoproteínas unidas a N contienen un resto de *N*-acetilglucosamina unido al nitrógeno de la amida de un resto de asparagina en la proteína. Los azúcares predominantes encontrados en las glicoproteínas son glucosa, galactosa, manosa, fucosa, *N*-acetilgalactosamina (GalNAc), *N*-acetilglucosamina (GlcNAc) y ácidos siálico (por ejemplo, ácido *N*-acetil-neuramínico (NANA)). El procesamiento de los grupos del azúcar se produce de forma cotranslacional en el lumen del RE y continúa de forma post-translacional en el aparato de Golgi para glicoproteínas unidas a N.

Los *N*-glicanos tienen un núcleo de pentasacárido común de Man₃GlcNAc₂ ("Man" se refiere a manosa; "Glc" se refiere a glucosa; y "NAc" se refiere a *N*-acetilo; GlcNAc se refiere a *N*-acetilglucosamina). Los *N*-glicanos se diferencian con respecto al número de ramificaciones (antenas) que comprenden azúcares periféricos (por ejemplo, GlcNAc, galactosa, fucosa y ácido siálico) que se añaden a la estructura del núcleo de Man₃GlcNAc₂ ("Man₃") que también se denomina el "núcleo de trimanosa", el "núcleo de pentasacárido" o el "núcleo de paucimanosa". Los *N*-glicanos se clasifican de acuerdo con sus componentes ramificados (por ejemplo, alta manosa, complejo o híbrido). Un *N*-glicano de tipo "alta manosa" tienen cinco o más restos de manosa. Por lo general un *N*-glicano de tipo "complejo" tiene al menos un GlcNAc unido a la rama de 1,3 manosa y al menos un GlcNAc unido a la rama de 1,6 manosa de un núcleo de "trimanosa". Los *N*-glicanos complejos también pueden tener restos de galactosa ("Gal") o *N*-acetilgalactosamina ("GalNAc") que se están modificados opcionalmente con ácido siálico o derivados (por ejemplo, "NANA" o "NeuAc", en el que "Neu" se refiere a ácido neuramínico y "Ac" se refiere a acetilo). Los *N*-glicanos complejos también pueden tener sustituciones intracadena que comprenden GlcNAc de "bisección" y fucosa de núcleo ("Fuc"). Los *N*-glicanos complejos también pueden tener múltiples antenas en el "núcleo de trimanosa", denominados a menudo "glicanos antenarios múltiples". Un *N*-glicano "híbrido" tiene al menos un GlcNAc en el extremo de la rama de 1,3 manosa del núcleo de trimanosa y cero o más manosas en la rama 1,6

manosa del núcleo de trimanosa. Los diversos *N*-glicanos también se denominan "glicoformas".

Las abreviaturas usadas en el presente documento son de uso común en la técnica, véanse, por ejemplo, abreviaturas de azúcares, mencionadas anteriormente. Otras abreviaturas comunes incluyen "PNGasa", o

5 "glicanasa" o "glucosidasa" que todas se refieren al péptido *N*-glicosidasa F (EC 3.2.2.18).

La expresión "célula hospedadora recombinante" ("célula hospedadora de expresión", "sistema hospedador de expresión", "sistema de expresión" o simplemente "célula hospedadora"), como se usa en el presente documento, pretende hacer referencia a una célula en la que se ha introducido un vector recombinante. Se debería entender que

10 tales términos pretenden hacer referencia no solamente a la célula objeto en particular sino también a la progenie de una célula de este tipo. Dado que se pueden producir ciertas modificaciones en las generaciones con éxito debido a cualquier mutación por influencias ambientales, tal progenie puede no ser, de hecho, idéntica a la célula precursora, pero también está incluidas dentro del alcance de la expresión "célula hospedadora" como se usa en el presente documento. Una célula hospedadora recombinante puede ser una célula aislada o línea celular que crece en cultivo

15 o puede ser una célula que reside en un tejido u organismo vivo. Las células hospedadoras preferentes son levaduras y hongos.

Una célula hospedadora que "no presenta" una actividad enzimática se refiere a una célula hospedadora en la que la actividad enzimática se ha anulado o interrumpido. Por ejemplo, la actividad enzimática se puede anular o

20 interrumpir por delección o interrupción del gen que codifica la actividad enzimática (incluida delección o interrupción de las secuencias reguladoras cadena arriba o cadena abajo que controlan la expresión del gen; la actividad enzimática se puede anular o interrumpir por mutación del gen que codifica la actividad enzimática para hacer que el gen que codifica la actividad enzimática no sea funcional; la actividad enzimática se puede anular o interrumpir mediante el uso de un inhibidor químico, peptídico o proteico de la actividad enzimática; la actividad enzimática se

25 puede anular o interrumpir mediante el uso de inhibidores de expresión basados en ácido nucleico tales como ADN y ARNs antisentido; y, la actividad enzimática se puede anular o interrumpir mediante el uso de inhibidores de transcripción o inhibidores de la expresión o actividad de factores de regulación que controlan o regulan la expresión del gen que codifica la actividad enzimática.

30 Cuando se hace referencia a "porcentaje molar" de un glicano presente en una preparación de una glicoproteína, la expresión se refiere al porcentaje molar de un glicano en particular presente en la combinación de oligosacáridos unidos a *N* liberados cuando la preparación de proteína se trata con PNG'asa y a continuación se cuantifican mediante un método que no se ve afectado por la composición de la glicoforma, (por ejemplo, etiquetado de una combinación de glicano liberado por PNG'asa con una etiqueta fluorescente tal como 2-aminobenzamida y a

35 continuación se separan mediante cromatografía líquida de alto rendimiento o electroforesis de capilaridad y a continuación los glicanos se cuantifican mediante la intensidad de fluorescencia). Por ejemplo, un 50 por ciento en moles de NANA₂Gal₂GlcNAc₂Man₃GlcNAc₂ se refiere a que un 50 por ciento de los glicanos liberados son NANA₂Gal₂GlcNAc₂Man₃GlcNAc₂ y el 50 por ciento restante está formado por otros oligosacáridos unidos a *N*. En realizaciones, el porcentaje molar de un glicano en particular en una preparación de glicoproteína estará entre un

40 20 % y un 100 %, preferentemente por encima de un 25 %, 30 %, 35 %, 40 % o un 45 %, más preferentemente por encima de un 50 %, 55 %, 60 %, 65 % o un 70 % y lo más preferentemente por encima de un 75 %, 80 % 85 %, 90 % o un 95 %.

Como se usa en el presente documento, se entenderá que el término "predominantemente" o variaciones tales como

45 "el predominante" o "que es predominante" hacen referencia a las especies de glicano que tiene el porcentaje molar más elevado (%) de *N*-glicanos totales después de haber tratado la glicoproteína con PNGasa y los glicanos liberados se analizan mediante espectroscopía de masas, por ejemplo, MALDI-TOF MS. En otras palabras, el término "predominantemente" se define como una entidad individual, de modo que una glicoforma específica, está presente en un porcentaje molar más elevado que cualquier otra entidad individual. Por ejemplo, si una composición

50 consiste en la especie A en un 40 por ciento en moles, la especie B en un 35 por ciento en moles y la especie C en un 25 por ciento en moles, la composición comprende predominantemente la especie A.

La expresión "cantidad terapéuticamente eficaz" se refiere a una cantidad de la eritropoyetina recombinante de la invención que da un aumento en el hematocrito que proporciona beneficios a un paciente. La cantidad variará de un

55 individuo a otro y dependerá de un número de factores, incluyendo la condición física general del paciente y la causa subyacente de la anemia. Por ejemplo, una cantidad terapéuticamente eficaz de eritropoyetina de la presente invención para un paciente que padece una insuficiencia renal crónica puede estar en el intervalo de 20 a 300 unidades/kg o 0,5 ug/kg a 500 ug/kg basándose en la indicación terapéutica. El término "unidad" se refiere a unidades conocidas normalmente en la técnica para evaluar la actividad de las composiciones de eritropoyetina. Un

60 miligramo de eritropoyetina pura es proximadamente equivalente a 150.000 unidades. Un programa de dosificación puede ser de aproximadamente tres veces a la semana a aproximadamente una vez cada cuatro o seis semanas. La programación real dependerá de una serie de factores que incluyen el tipo de eritropoyetina administrada a un paciente (EPO o EPO PEGilada) y la respuesta del paciente individual. Los intervalos de dosis más elevados por lo general no se usan en aplicaciones en anemia pero pueden ser útiles en otras aplicaciones terapéuticas. El medio

65 para conseguir y establecer una dosis apropiada de eritropoyetina para un paciente se conoce bien y se pone en práctica normalmente en la técnica.

Las variaciones en la cantidad dada y programa de dosificación de paciente a paciente se incluyen por referencia al término "aproximadamente" en conjunto con una cantidad o programa. La cantidad de eritropoyetina usada para terapia da una tasa aceptable de aumento de hematocrito y mantiene el hematocrito en un nivel beneficioso (por ejemplo, normalmente de al menos aproximadamente un 30 % y por lo general en un intervalo de un 30 % a un 36 %). Una cantidad terapéuticamente eficaz de las presentes composiciones la puede discernir fácilmente un experto en la materia usando materiales y procedimientos disponibles al público. Además, se puede administrar hierro al paciente para mantener un aumento de la eritropoyesis durante la terapia. La cantidad a administrar se puede determinar fácilmente con métodos usados normalmente por los expertos en la materia.

10 Breve descripción de las figuras

La **Figura 1 A-J** muestra la genealogía de la cepa **YGLY3159** de *P. pastoris* (**Figura 1E**) y las cepas **YGLY7113** a **YGLY7122** (**Figura 1I**) que comienzan a partir de la cepa **NRRL-Y11430** de tipo silvestre (**Figura 1A**).

La **Figura 2** muestra un mapa del plásmido pGLY6. El plásmido pGLY6 es un vector de integración que se dirige al locus de *URA5* y contiene una molécula de ácido nucleico que comprende el gen de invertasa de o unidad transcripción (*ScSUC2*) de *S. cerevisiae* flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *URA5* de *P. pastoris* (PpURA5-5') y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *URA5* de *P. pastoris* (PpURA5-3').

La **Figura 3** muestra un mapa del plásmido pGLY40. El plásmido pGLY40 es un vector de integración que se dirige al locus de *OCH1* y contiene una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ* (repetición de *lacZ*) que a su vez está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *OCH1* (PpOCH1-5') y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *OCH1* (PpOCH1-3').

La **Figura 4** muestra un mapa del plásmido pGLY43a. El plásmido pGLY43a es un vector de integración que se dirige al locus de *BMT2* y contiene una molécula de ácido nucleico que comprende el gen transportador (UDP-GlcNAc) o unidad de transcripción (KIGlcNAc Transp.) de UDP-*N*-acetilglucosamina de *K. lactis* adyacente a una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ* (repetición de *lacZ*). Los genes adyacentes están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *BMT2* (PpPBS2-5') y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *BMT2* (PpPBS2-3').

La **Figura 5** muestra un mapa del plásmido pGLY48. El plásmido pGLY48 es un vector de integración que se dirige al locus de *MNN4L1* y contiene un casete de expresión que comprende una molécula de ácido nucleico que codifica el homólogo de ratón del marco de lectura abierto (ORF) del transportador de UDP-GlcNAc (Transp. de MmGlcNAc) unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GAPDH* de *P. pastoris* (Prom PpGAPDH) y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de *CYC* (*ScCYC* TT) de *S. cerevisiae* adyacente a una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por repeticiones de *lacZ* (repetición de *lacZ*) y en el que los casetes de expresión juntos están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *MNN4L1* (PpMNN4L1-5') de *P. pastoris* y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *MNN4L1* (PpMNN4L1-3').

La **Figura 6** muestra un mapa del plásmido pGLY45. El plásmido pGLY45 es un vector de integración que se dirige a los loci *PNO1/MNN4* y una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ* (repetición de *lacZ*) que a su vez está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *PNO1* (PpPNO1-5') y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *MNN4* (PpMNN4-3').

La **Figura 7** muestra un mapa del plásmido pGLY247. El plásmido pGLY247 es un vector de integración que se dirige al locus de *MET16* y contiene una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ* (repetición de *lacZ*) que a su vez está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *MET16* (PpMET16-5') y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *MET16* (PpMET16-3').

La **Figura 8** muestra un mapa del plásmido pGLY248. El plásmido pGLY248 es un vector de integración que se dirige al locus de *URA5* y contiene una molécula de ácido nucleico que comprende el gen *MET16* o unidad de transcripción (PpMET16) de *P. pastoris* flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *URA5* (PpURA5-5') y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la

posición 3' del gen *URA5* (PpURA5-3').

La **Figura 9** muestra un mapa del plásmido pGLY582. El plásmido pGLY582 es un vector de integración que se dirige al locus de *HIS1* y contiene en tándem cuatro casetes de expresión que codifican (1) la epimerasa de UDP-glucosa (*ScGAL10*) de *S. cerevisiae*, (2) el dominio catalítico de la galactosiltransferasa I (*hGalT*) humana fusionado en el extremo N-terminal al péptido líder *KRE2-s* de *S. cerevisiae* (33), (3) el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por repeticiones de *lacZ* (repetición de *lacZ*) y (4) el transportador de UDP-galactosa (*DmUGT*) de *D. melanogaster*. Todos flanqueados por la región en la posición 5' del gen *HIS1* (PpHIS1-5') y la región en la posición 3' del gen *HIS1* (PpHIS1-3'). PMA1 es el promotor *PMA1* de *P. pastoris*; PpPMA1 TT es la secuencia de terminación de *PMA1* de *P. pastoris*; GAPDH es el promotor *GADPH* de *P. pastoris* y ScCYC TT es la secuencia de terminación de *CYC* de *S. cerevisiae*; Prom PpOCH1 es el promotor *OCH1* de *P. pastoris* y PpALG12 TT es la secuencia de terminación de *ALG12* de *P. pastoris*.

La **Figura 10** muestra un mapa del plásmido pGLY167b. El plásmido pGLY167b es un vector de integración que se dirige al locus de *ARG1* y contiene en tándem tres casetes de expresión que codifican (1) el dominio catalítico de la manosidasa II de *D. melanogaster* (optimizado por codón) fusionado en el extremo N-terminal al péptido líder *MNN2* (CO-KD53) de *S. cerevisiae*, (2) el gen *HIS1* o unidad de transcripción de *P. pastoris* y (3) el dominio catalítico de la N-acetilglucosamina (GlcNAc) transferasa II de rata (optimizado por codón) fusionado en el extremo N-terminal al péptido líder *MNN2* (CO-TC54) de *S. cerevisiae*. Todos flanqueados por la región en la posición 5' del gen *ARG1* (PpARG1-5') y la región en la posición 3' del gen *ARG1* (PpARG1-3'). PpPMA1 prom es el promotor *PMA1* de *P. pastoris*; PpPMA1 TT es la secuencia de terminación de *PMA1* de *P. pastoris*; PpGAPDH es el promotor *GADPH* de *P. pastoris*; ScCYC TT es la secuencia de terminación de *CYC* de *S. cerevisiae*; Prom PpOCH1 es el promotor *OCH1* de *P. pastoris*; y PpALG12 TT es la secuencia de terminación de *ALG12* de *P. pastoris*.

La **Figura 11** muestra un mapa del plásmido pGLY1430. El plásmido pGLY1430 es un vector de integración de KINKO que se dirige al locus de *ADE1* sin interrupción de la expresión del locus de y contiene en tándem cuatro casetes de expresión que codifican (1) el dominio catalítico de la GlcNAc transferasa I humana (optimizado por codón) fusionado en el extremo N-terminal al péptido líder *SEC12* (CO-NA10) de *P. pastoris*, (2) homólogo de ratón del transportador de UDP-GlcNAc (MmTr), (3) el dominio catalítico de la manosidasa IA de ratón (FB) fusionado en el extremo N-terminal al péptido líder *SEC12* (FB8) de *S. cerevisiae* y (4) el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por repeticiones de *lacZ* (*lacZ*). Todos flanqueados por la región en la posición 5' del gen *ADE1* y ORF (*ADE1* 5' y ORF) y la región en la posición 3' del gen *ADE1* (PpADE1-3'). PpPMA1 prom es el promotor *PMA1* de *P. pastoris*; PpPMA1 TT es la secuencia de terminación de *PMA1* de *P. pastoris*; SEC4 es el promotor *SEC4* de *P. pastoris*; OCH1 TT es la secuencia de terminación *OCH1* de *P. pastoris*; ScCYC TT es la secuencia de terminación de *CYC* de *S. cerevisiae*; Prom PpOCH1 es el promotor *OCH1* de *P. pastoris*; PpALG3 TT es la secuencia de terminación *ALG3* de *P. pastoris*; y PpGAPDH es el promotor *GADPH* de *P. pastoris*.

La **Figura 12** muestra un mapa del plásmido pGFI165. El plásmido pGFI165 es un vector de integración de KINKO que se dirige al locus de *PRO1* sin interrupción de la expresión del locus de y contiene casetes de expresión que codifican (1) el dominio catalítico de la α -1,2-manosidasa de *T. reesei* fusionado en el extremo N-terminal al péptido señal α MATpre (α MATTrMan) de *S. cerevisiae* para dirigir la proteína quimérica a la ruta secretoria y secreción desde la célula y (2) el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por repeticiones de *lacZ* (repetición de *lacZ*). Todos flanqueados por la región en la posición 5' del gen *PRO1* y ORF (5'PRO1orf) y la región en la posición 3' del gen *PRO1* (3'PRO). ScCYC TT es la secuencia de terminación de *CYC* de *S. cerevisiae*; PpALG3 TT es la secuencia de terminación de *ALG3* de *P. pastoris*; y PpGAPDH es el promotor *GADPH* de *P. pastoris*.

La **Figura 13** muestra un mapa del plásmido pGLY2088. El plásmido pGLY2088 es un vector de integración que se dirige al locus de *TRP2* o al locus de *AOX1* y contiene casetes de expresión que codifican (1) eritropoyetina (co-hEPO) humana madura optimizada por codón fusionada en el extremo N-terminal a un péptido señal α MATpre (alfa MF-pre) de *S. cerevisiae* para dirigir la proteína quimérica a la ruta secretoria y secreción desde la célula y (2) la proteína de resistencia a zeocina (ZeocinaR). Los casetes están flanqueados en un extremo con el promotor *AOX1* (Prom de PpAOX1) de *P. pastoris* y en el otro extremo con el gen *TRP2* o unidad de transcripción (PpTRP2) de *P. pastoris*. ScCYC TT es la secuencia de terminación de *CYC* de *S. cerevisiae* y ScTEF Prom es el promotor *TEF1* de *S. cerevisiae*.

La **Figura 14** muestra un mapa del plásmido pGLY2456. El plásmido pGLY2456 es un vector de integración de KINKO que se dirige al locus de *TRP2* sin interrupción de la expresión del locus de y contiene seis casetes de expresión que codifican (1) el transportador de CMP-ácido siálico de ratón optimizado por codón (Transp CO mCMP-Sia), (2) la UDP-GlcNAc 2-epimerasa/N-acetilmanosamina quinasa humana optimizada por codón (CO hGNE), (3) el gen *ARG1* o unidad de transcripción de *Pichia pastoris*, (4) la CMP-ácido siálico sintasa humana optimizada por codón (CO hCMP-NANA S), (5) la N-acetilneuraminato-9-fosfato sintasa humana optimizada por codón (CO hSIAP S) y, (6) el dominio catalítico de α -2,6-sialiltransferasa de ratón optimizado por codón fusionado en el extremo N-terminal al péptido líder *KRE2* (comST6-33) de *S. cerevisiae*. Todos flanqueados por la región en la posición 5' del gen *TRP2* y ORF (PpTRP2 5') y la región en la posición 3' del gen *TRP2* (PpTRP2-3'). PpPMA1 prom es el promotor *PMA1* de *P. pastoris*; PpPMA1 TT es la secuencia de terminación de *PMA1* de *P. pastoris*; CYC TT es la secuencia de terminación de *CYC* de *S. cerevisiae*; PpTEF Prom es el promotor *TEF1* de *P. pastoris*; PpTEF TT es la secuencia de terminación de *TEF1* de *P. pastoris*; PpALG3 TT es la secuencia de terminación de *ALG3* de *P. pastoris*; y pGAP es el promotor *GADPH* de *P. pastoris*.

La **Figura 15** muestra un mapa del plásmido pGLY3411 (pSH1092). El plásmido pGLY3411 (pSH1092) es un vector de integración que contiene el casete de expresión que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por repeticiones de *lacZ* (repetición de *lacZ*) flanqueado en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT4* (PpPBS4 5') de *P. pastoris* y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT4* (PpPBS4 3') de *P. pastoris*.

La **Figura 16** muestra un mapa del plásmido pGLY3430 (pSH1115). El plásmido pGLY3430 (pSH1115) es un vector de integración que contiene un casete de expresión que comprende una molécula de ácido nucleico que codifica el ORF de resistencia a Nourseotricina (NAT) unido de forma operativa al promotor *TEF1* de *Ashbya gossypii* (PTEF) y la secuencia de terminación de *TEF1* (TTEF) de *Ashbya gossypii* flanqueada en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT1* (PBS1 5') de *P. pastoris* y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT1* (PBS1 3') de *P. pastoris*.

La **Figura 17** muestra un mapa del plásmido pGLY4472 (pSH1186). El plásmido pGLY4472 (pSH1186) contiene una molécula de ácido nucleico que codifica el ORF del gen de fosfotransferasa de la higromicina B (Hyg) de *E. coli* unido de forma operativa al promotor *TEF1* de *Ashbya gossypii* (pTEF) y la secuencia de terminación de *TEF1* (TRFtt) de *Ashbya gossypii* flanqueada en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT3* (PpPBS3 5') de *P. pastoris* y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT3* (PpPBS3 3') de *P. pastoris*.

La **Figura 18** muestra un mapa del plásmido pGLY2057. El plásmido pGLY2057 es un vector de integración que se dirige al locus de *ADE2* y contiene un casete de expresión que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por repeticiones de *lacZ* (repetición de *lacZ*). El casete de expresión está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *ADE2* (PpADE2-5') y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *ADE2* (PpADE2-3').

La **Figura 19** muestra un mapa del plásmido pGLY2680. El plásmido pGLY2680 es un vector de integración que se puede dirigir al locus de *TRP2* o *AOX1* y contiene casetes de expresión que codifican (1) eritropoyetina (co-hEPO) humana madura optimizada por codón fusionada en el extremo N-terminal a péptido señal de lisozima de pollo (Lisozima de pollo ss) y (2) el gen *ADE2* de *P. pastoris* sin un promotor (PpADE2). Los casetes están flanqueados en un extremo con el promotor *AOX1* (Prom de PpAOX1) de *P. pastoris* y en el otro extremo con el gen *TRP2* o unidad de transcripción (PpTRP2) de *P. pastoris*. ScCYC TT es la secuencia de terminación de *CYC* de *S. cerevisiae*.

La **Figura 20** muestra un mapa del plásmido pGLY2713. El plásmido pGLY2713 es un vector de integración que contiene el ORF de *PNO1* (ORF de PpPNO1) de *P. pastoris* adyacente al casete de expresión que comprende el gen *URA5* o unidad de transcripción (PpURA5) de *P. pastoris* flanqueado por repeticiones de *lacZ* (repetición de *lacZ*) y flanqueado en un lado con la secuencia de nucleótidos de la región en la posición 5' del gen *PEP4* (PpPEP4 5') de *P. pastoris* y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *PEP4* (PpPEP4 3') de *P. pastoris*.

La **Figura 21** muestra un diagrama esquemático que ilustra el flujo del proceso de fermentación.

La **Figura 22** muestra que la rhEPO producida en la cepa **YGLY3159** tiene actividad de unión cruzada a anticuerpos anti-HCA. El panel de la parte izquierda muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie que muestra la posición de la rhEPO y el panel de la parte derecha muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA de conejo (conejo purificado con SL rProA: 9161) a una dilución a 1:3.000. El anticuerpo anti-HCA unido se detectó usando anticuerpo anti-conejo de cabra conjugado a peroxidasa de rábano picante (HRP) a una dilución a 1:5.000 en PBS. La detección del anticuerpo secundario unido usó el sustrato 3,3 diaminobencidina (DAB).

La **Figura 23** muestra que la actividad de unión cruzada de la rhEPO producida en la cepa **YGLY3159** a anticuerpos anti-HCA no se detecta cuando la rhEPO se desglicosila usando PNAGasa F. El panel de la parte izquierda muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie que muestra la posición de las formas glicosilada y desglicosilada de rhEPO y el panel de la parte derecha muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22**.

La **Figura 24** muestra que un anticuerpo recombinante (rhIgG) producido en *P. pastoris* de tipo silvestre y una cepa GS2.0 de *P. pastoris* glicomodificada por ingeniería en la que el gen *BMT2* se ha interrumpido o sometido a delección presentaba actividad de unión cruzada a anticuerpos anti-HCA. El panel de la parte izquierda muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie y el panel de la parte derecha muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22**. GS 2.0 es una cepa de *P. pastoris* que produce glico proteínas que tienen predominantemente *N*-glicanos de Man₅GlcNAc₂. La cepa GS 2.0 mostrada producía rhIgG con aproximadamente un 5 % de *N*-glicanos de Man₅GlcNAc₂. WT es *P. pastoris* de tipo silvestre.

La **Figura 25** compara la actividad de unión cruzada de rhEPO producida en la cepa **YGLY3159** a otras proteínas glicosiladas que contienen patrones de glicosilación complejos pero que no se producían en *P. pastoris* para anticuerpo anti-HCA. El panel de la parte superior muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie que muestra la posición de las formas glicosilada y desglicosilada de rhEPO producidas en *P. pastoris* y de fetuína humana recombinante, asialofetuína (fetuína humana con restos de ácido siálico terminales retirados), albúmina de suero humano (HSA) y LEUKINE recombinante producidos en *S. cerevisiae* y el panel de la parte inferior muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22**. Las combinaciones de S30S son rhEPO purificada mediante cromatografía de intercambio catiónico.

La **Figura 26** muestra que la rhEPO producida en la cepa **YGLY3159** y purificada mediante cromatografía de hidroxipatita todavía tiene actividad de unión cruzada a anticuerpos anti-HCA. El panel de la parte izquierda muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de las combinaciones 1, 2 y 3 de dilución mediante cromatografía que muestran la posición de la rhEPO (reducida o no reducida) y el panel de la

parte derecha muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22**. A continuación, en los paneles se muestran los resultados de un análisis de HPLC de *N*-glicanos en las combinaciones 1, 2 y 3.

La **Figura 27A** muestra un cromatograma de purificación de rhEPO mediante cromatografía aniónica de Q

SEPHAROSE FF producida en la cepa **YGLY3159** a partir de la combinación 1 de hidroxipatita.

La **Figura 27B** muestra un ELISA de sándwich que muestra que la combinación de Q SEPHAROSE FF que contiene rhEPO a partir de la cromatografía aniónica de Q SEPHAROSE FF no tiene actividad de unión cruzada detectable a anticuerpos anti-HCA mientras que el flujo continuo contenía actividad de unión cruzada a anticuerpos anti-HCA. El anticuerpo de captura era anticuerpo anti-hEPO y la actividad de unión cruzada se detectó con anticuerpo anti-HCA de conejo a una dilución de partida de 1:800 en PBS que a continuación se diluyó en serie a 1:1 en PBS a través de una fila de terminación con el 11^o pocillo a una dilución de 1:819.200 (pocillo 12: control negativo). El anticuerpo anti-HCA unido se detectó usando anticuerpo anti-conejo de cabra conjugado con fosfatasa alcalina (AP) a una dilución de 1:10.000 dilution en PBS. La detección del anticuerpo secundario unido usó fosfato de 4-Metilumbeliferilo (4-MUPS).

La **Figura 28** muestra que la rhEPO producida en las cepas **YGLY6661** ($\Delta bmt2$, $\Delta bmt4$ y $\Delta bmt1$) e **YGLY7013** ($\Delta bmt2$ y $\Delta bmt4$) y capturada mediante cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) aún tiene actividad de unión cruzada a anticuerpos anti-HCA. El panel de la parte izquierda muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de las combinaciones de Azul con (+) y sin (-) tratamiento con PNGasa F. El panel central muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-hEPO conjugados con HRP a una dilución a 1:1.000 y DAB como el sustrato. El panel de la derecha muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22**.

La **Figura 29** muestra que un ELISA de sándwich para detectar la actividad de unión cruzada a anticuerpos anti-HCA de rhEPO producidos en las cepas **YGLY6661** ($\Delta bmt2$, $\Delta bmt4$ y $\Delta bmt1$) e **YGLY7013** ($\Delta bmt2$ y $\Delta bmt4$) y capturada mediante cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) aún tiene actividad de unión cruzada a anticuerpos anti-HCA. El ELISA se realizó como en la **Figura 27B**.

La **Figura 30** muestra ELISA de sándwich para detectar la actividad de unión cruzada a anticuerpos anti-HCA de rhEPO producidos en las cepas **YGLY6661** ($\Delta bmt2$, $\Delta bmt4$ y $\Delta bmt1$) e **YGLY7013** ($\Delta bmt2$ y $\Delta bmt4$), capturados mediante cromatografía de Azul Sepaharose 6 FF y purificados mediante cromatografía de hidroxipatita (combinación 1 de HA). La rhEPO en la combinación 1 de HA de la cepa **YGLY6661** no tenía actividad de unión cruzada a anticuerpos anti-HCA detectable. Los ELISA se realizaron como en la **Figura 27B**.

La **Figura 31** muestra que la rhEPO producida en la cepa **YGLY6661** ($\Delta bmt2$, $\Delta bmt4$ y $\Delta bmt1$) y capturada mediante cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) aún tiene actividad de unión cruzada a anticuerpos anti-HCA. El panel de la parte izquierda muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de las combinaciones de Azul con (+) y sin (-) tratamiento con PNGasa F. El panel central muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-hEPO conjugados con HRP a una dilución a 1:1.000 y DAB como el sustrato. El panel de la derecha muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22**.

La **Figura 32A** muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de las combinaciones de captura de Azul Sepaharose 6 FF (combinaciones de Azul) preparado a partir de las cepas **YGLY7361-7366** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$) con (+) y sin (-) tratamiento con PNGasa F. Las cepas se cultivaron en fermentadores SixFors de 500 ml.

La **Figura 32B** muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de las combinaciones de captura de Azul Sepaharose 6 FF (combinaciones de Azul) preparadas a partir de las cepas **YGLY7393-7398** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$) con (+) y sin (-) tratamiento con PNGasa F. Las cepas se cultivaron en fermentadores SixFors de 500 ml.

La **Figura 33** muestra los resultados de ELISA de sándwich para detectar la actividad de unión cruzada a anticuerpos anti-HCA de rhEPO producidos en las cepas **YGLY7361-7366** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$) e **YGLY7393-7398** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$) y capturados mediante cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul). Solamente la rhEPO en las combinaciones de Azul de la cepa **YGLY7363** e **YGLY7365** tenían actividad de unión cruzada detectable a anticuerpos anti-HCA. Los ELISA se realizaron como en la **Figura 27B**.

La **Figura 34** muestra en forma de gráfico los resultados del análisis de HPLC de los *N*-glicanos en la rhEPO en las combinaciones de Azul preparadas a partir de las cepas **YGLY7361-7366** e **YGLY7393-7398** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$). "Bi" se refiere a *N*-glicanos en donde ambas ramas del *N*-glicano biantenarico están sialiladas. "Mono" se refiere a *N*-glicanos en donde solamente una rama del *N*-glicano biantenarico están sialiladas. "Neutro" se refiere a *N*-glicanos que no están sialilados.

La **Figura 35A** muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de las combinaciones de cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) y purificación con hidroxipatita (combinación de HA 1s) preparadas a partir de las cepas **YGLY7362**, **YGLY7366**, **YGLY7396**, e **YGLY7398** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$), e **YGLY3159** ($\Delta bmt2$).

La **Figura 35B** muestra una transferencia de Western de un gel de SDS-PAGE al 4-20 % de las combinaciones de cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) y purificación con hidroxipatita

(combinación de HA 1s) preparadas a partir de las cepas **YGLY7362**, **YGLY7366**, **YGLY7396**, e **YGLY7398** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$), e **YGLY3159** ($\Delta bmt2$) y sondeadas con anticuerpos anti-HCA como en la **Figura 22**.

La **Figura 36** muestra que la rhEPO producida en la cepa **YGLY7398** ($\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$) y capturada mediante cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) y purificada mediante cromatografía de hidroxipatita (combinación de HA 1s) no tenía actividad de unión cruzada a anticuerpos anti-HCA detectable. El panel de la parte izquierda muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de la combinación de Azul y la combinación 1 de HA preparadas a partir de la cepa **YGLY7398** en comparación con la rhEPO preparada a partir de la cepa **YGLY3159**. El panel central muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22**. El panel central muestra una transferencia de Western de un gel similar sondeado con anticuerpos anti-HCA como en la **Figura 22** excepto porque los anticuerpos anti-HCA eran de otra preparación de anticuerpo (conejo policlonal de GiF::6316 a 1:2.000).

La **Figura 37** muestra los resultados de ELISA de sándwich para detectar la actividad de unión cruzada a anticuerpos anti-HCA de la rhEPO producida en las cepas **YGLY7113-7122** ($\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$) y capturada mediante cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul). La cepa **YGLY7118** presentaba una actividad de unión cruzada a anticuerpos anti-HCA detectable muy baja. Ninguna de las otras cepas mostraba ninguna actividad de unión cruzada a anticuerpos anti-HCA detectable. Los ELISA se realizaron como en la **Figura 27B**.

La **Figura 38** muestra en forma de gráfico los resultados del análisis de HPLC de los *N*-glicanos en la rhEPO en las combinaciones de Azul preparadas a partir de las cepas **YGLY7113-7122** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$). "Bi" se refiere a *N*-glicanos en donde ambas ramas del *N*-glicano biantenarico están sialiladas. "Mono" se refiere a *N*-glicanos en donde solamente una rama del *N*-glicano biantenarico están sialiladas. "Neutro" se refiere a *N*-glicanos que no están sialilados.

La **Figura 39A** muestra un gel de SDS-PAGE al 4-20 % teñido con Azul de Commassie de las combinaciones de cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) y purificación con hidroxipatita (combinación de HA 1s) preparado a partir de las cepas **YGLY7115**, **YGLY7117**, **YGLY7394**, **YGLY7395**, e **YGLY7120** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$), e **YGLY3159** ($\Delta bmt2$).

La **Figura 39B** muestra una transferencia de Western de un gel de SDS-PAGE al 4-20 % de las combinaciones de cromatografía de Azul SEPHAROSE 6 FF (combinaciones de Azul) y purificación con hidroxipatita (combinación de HA 1s) preparado a partir de las cepas **YGLY7115**, **YGLY7117**, **YGLY7394**, **YGLY7395**, e **YGLY7120** (todas $\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$), e **YGLY3159** ($\Delta bmt2$) y sondeado con anticuerpos anti-HCA como en la **Figura 22**.

La **Figura 40A** muestra una traza de HPLC de los *N*-glicanos de la rhEPO producida en **YGLY3159** ($\Delta bmt2$) y purificada mediante cromatografía en columna de hidroxipatita (es decir, análisis de la combinación 1 de HA).

La **Figura 40B** muestra una traza de HPLC de los *N*-glicanos de la rhEPO producida **YGLY7117** ($\Delta bmt2$, $\Delta bmt4$, $\Delta bmt1$ y $\Delta bmt3$) y purificada mediante cromatografía en columna de hidroxipatita (es decir, análisis de la combinación 1 de HA).

Descripción detallada

La presente invención proporciona métodos para producir proteínas y glicoproteínas en *Pichia pastoris* que carecen de unión cruzada detectable a anticuerpos preparados frente a antígenos de célula hospedadora. Los antígenos de célula hospedadora también pueden incluir proteína de célula hospedadora residual y contaminantes de la pared celular que pueden llevar a composiciones de proteína recombinante que pueden ser inmunogénicas y que pueden alterar la eficacia o seguridad terapéutica de una proteína terapéutica. Una composición que tiene reactividad cruzada con anticuerpos preparados frente a antígenos de célula hospedadora se refiere a que la composición contiene algún material de célula hospedadora contaminante, normalmente *N*-glicanos con restos de fosfomanosa o restos de β -manosa o similares. Las cepas de tipo silvestre de *Pichia pastoris* producirán glicoproteínas que tengan estas estructuras de *N*-glicano. Se podría esperar que las preparaciones de anticuerpo preparadas frente a proteínas de célula hospedadora totales incluyeran anticuerpos frente a estas estructuras. Las proteínas que no contienen *N*-glicanos, sin embargo, también podrían incluir material contaminante (proteínas o similares) que reaccionaran de forma cruzada con anticuerpos preparados frente a la célula hospedadora.

Los métodos y las células hospedadoras permiten la producción de proteínas y glicoproteínas terapéuticas recombinantes con una reducción del riesgo de provocar una reacción adversa en un individuo al que se le administran las proteínas y glicoproteínas terapéuticas recombinantes en comparación con las mismas que se están produciendo en cepas no modificadas como se desvela en el presente documento. Una reacción adversa incluye provocar una respuesta inmunológica no deseada en el individuo o una unión no deseada o inapropiada a, congregación en, o interacción con un sitio en el individuo que en general influye de forma adversa en la salud del individuo. El riesgo de provocar una reacción adversa en un individuo al que se le está administrando la proteína o glicoproteína terapéutica es de preocupación en particular para proteínas o glicoproteínas destinadas a su administración al individuo de forma crónica (por ejemplo, terapias destinadas a su realización durante un periodo de tiempo prolongado). Las proteínas o glicoproteínas terapéuticas recombinantes producidas de acuerdo con los métodos en el presente documento no tienen actividad de unión cruzada detectable anticuerpos frente a antígenos de célula hospedadora y por lo tanto, presentan una reducción del riesgo de provocar una reacción adversa en un

individuo al que se le administran las proteínas o glicoproteínas recombinantes. Los métodos y las células hospedadoras también son útiles para producir proteínas o glicoproteínas recombinantes que tengan un potencial más bajo unión a factores de eliminación.

- 5 Los inventores han encontrado que las glicoproteínas en particular que se producen en algunas cepas de *Pichia pastoris* pueden tener *N*- u *O*-glicanos en las mismas en las que uno o más de los restos de manosa en las mismas están en una unión en β 1,2. Las glicoproteínas destinadas para usos terapéuticos y que tienen uno o más restos de manosa unidos en β 1,2 en las mismas proporcionan un riesgo de ser capaces de provocar una respuesta inmunológica no deseada en el individuo al que se le está administrando la glicoproteína. Estos restos de manosa unidos en β se pueden detectar usando anticuerpos preparados frente a antígenos de célula hospedadora totales. Dado que no se puede predecir qué glicoproteínas terapéuticas tendrán *N*- u *O*-glicanos que comprenden uno o más restos de manosa unidos en β 1,2 y si una glicoproteína terapéutica que tiene *N*- u *O*-glicanos que comprenden restos de manosa unidos en β 1,2 en las mismas producirá una respuesta inmunológica no deseada en el individuo que recibe la glicoproteína, es deseable producir glicoproteínas terapéuticas en cepas de *Pichia pastoris* que se hayan modificado por ingeniería genética para que carezcan de unión cruzada detectable a anticuerpos preparados frente a antígenos de célula hospedadora. Tales cepas se pueden producir por una delección o interrupción de las actividades de al menos tres de las cuatro β -manosiltransferasas (*Bmt*p) conocidas en la familia de genes de β -manosiltransferasa (*BMT*) de *Pichia pastoris*. Como se muestra en el presente documento, las cepas de *Pichia pastoris* que incluyen una delección o interrupción de al menos tres de estos genes de *BMT* proporciona una cepa de *Pichia pastoris* que puede producir proteínas o glicoproteínas que carecen de unión cruzada detectable a anticuerpos preparados frente a antígenos de célula hospedadora. Estas cepas son útiles para producir proteínas y glicoproteínas terapéuticas. Se ha demostrado que la presencia de estructuras de β -manosa en *N*- y/o *O*-glicanos provoca una respuesta inmunológica.
- 25 La identificación de los genes de β -manosiltransferasa en *Pichia pastoris* y *Candida albicans* se informó en el documento de Patente de Estados Unidos n.º 7.465.577 y que en Mille *et al.*, J. Biol. Chem. 283: 9724-9736 (2008), que desvelan que la β -manosilación se realizaba con una β -manosiltransferasa que se denominó *AMR2* o *BMT2* y que la interrupción o delección del gen en *Pichia pastoris* daba como resultado una célula hospedadora recombinante que era capaz de producir glicoproteínas con reducción de la β -manosilación. El documento de patente también desvelaba tres homólogos del gen, *BMT1*, *BMT3* y *BMT4*. Sin embargo, cuando se investiga la fuente de actividad y unión cruzada de algunas preparaciones de glicoproteína a anticuerpos preparados frente a antígenos de célula hospedadora, los inventores descubrieron que la actividad de unión cruzada era una consecuencia de la β -manosilación residual persistente en algunas cepas de células hospedadoras de *P. pastoris* recombinantes en las que el gen *BMT2* se había interrumpido o experimentado delección. Por lo tanto, las glicoproteínas heterólogas producidas en estas células hospedadoras recombinantes tienen *N*-glicanos que aún contienen restos de β -manosa. Estos restos de β -manosa eran detectables en ELISA y transferencias de Western de las glicoproteínas heterólogas obtenidas a partir de cultivos de estas células hospedadoras recombinantes sondeadas con anticuerpos preparados frente a antígenos de célula hospedadora (HCA). Los anticuerpos anti-HCA son anticuerpos policlonales dirigidos frente a cepas de *Pichia pastoris* de tipo silvestre o una cepa de NORF: una célula hospedadora recombinante que se construye de la misma manera que la célula hospedadora recombinante que produce la glicoproteína heteróloga excepto porque se ha omitido el marco de lectura abierto (ORF) que codifica la proteína heteróloga. Para glicoproteínas terapéuticas producidas en *Pichia pastoris*, estos restos de β -manosa residuales presentan el riesgo de provocar una respuesta inmunológica en algunos individuos que reciben la proteína terapéutica en un tratamiento para una enfermedad o trastorno. La presente invención proporciona un método para producir glicoproteínas en *Pichia pastoris* que no contienen ninguna β -manosilación detectable y como tal no se unen de forma cruzada a anticuerpos preparados frente a antígenos de célula hospedadora.

- BMT1*, *BMT2* y *BMT3* demuestran un alto grado de homología de secuencia es mientras que *BMT4* es homólogo en una menor medida y se cree que es una protección de la alfa-manosiltransferasa. Sin embargo, parece que los cuatro miembros de la familia *BMT* están implicados en la síntesis de *N*- y/o *O*-glicanos que tienen estructuras de manosa unidas en β . Aunque una MALDI-TOF de *N*-glicanos de una proteína de ensayo producida en una cepa de *Pichia pastoris* en la que el gen *BMT2* ha experimentado delección podría fallar en la detección de la β -manosilación, los ensayos basados en anticuerpos sensibles en el presente documento eran capaces de detectar la β -manosilación en cepas de $\Delta bmt2$. Por lo tanto, los métodos de detección basados en anticuerpo anti-HCA enseñados en el presente documento presentaban que la delección o interrupción también de los genes *BMT1* y *BMT3* y opcionalmente del gen *BMT4* era necesaria para retirar todas las estructuras de β -manosa detectables. La delección o interrupción de los genes que codifican las tres β -manosiltransferasas se puede conseguir mediante (1) supresión genética completa o parcial del gen (incluyendo las secuencias promotoras, marco de lectura abierto (ORF) y/o las secuencias terminadoras de la transcripción); (2) introducción de un desplazamiento en marco del ORF; (3) inactivación o regulación del promotor; (4) supresión genética del mensaje mediante el ARNsi o ARN antisentido; (5) o el uso de inhibidores químicos. El resultado es la producción de una célula hospedadora que es capaz de producir una glicoproteína que carece de actividad de unión cruzada a anticuerpos anti-HCA detectable.

- Para ejemplificar los métodos para producir una glicoproteína que carezca de actividad de unión cruzada a anticuerpos anti-HCA detectable, una cepa de *Pichia pastoris*, que se había modificado por ingeniería genética para que careciera de expresión o actividad de *BMT2* y que fuera capaz de producir eritropoyetina humana madura

recombinante (EPO) con *N*-glicanos biantenarios terminados en ácido siálico, se modificó por ingeniería genética adicionalmente para que careciera de expresión de los genes *BMT1* y/o *BMT3* y/o *BMT4*. La cepa en la que solamente se había interrumpido la expresión del gen *BMT2* produjo EPO humana madura recombinante con cierta actividad de unión cruzada a anticuerpos anti-HCA detectable. Se encontró que la actividad de unión cruzada detectable se debía a la presencia de restos de manosa unidos en β en la molécula de EPO (véanse las Figuras 22-27B, Ejemplo 6). Cuando los genes que codifican *BMT1* y *BMT4* se interrumpían o experimentaban delección en la cepa, la EPO producida aún tenía actividad de unión cruzada a anticuerpos anti-HCA detectable (véanse las Figuras 28-31). Sin embargo, cuando los genes *BMT1*, *BMT2*, *BMT3* y *BMT4* se interrumpían o experimentaban delección, la mayoría de las cepas producían EPO humana recombinante glicosilada que carecía de actividad de unión cruzada a anticuerpos anti-HCA detectable y por lo tanto carecía de restos de β -manosa detectables (véanse las Figuras 33 y 35B por ejemplo).

Por lo tanto, la presente invención también proporciona un método para producir una proteína o glicoproteína recombinante que carezca de actividad de unión cruzada detectable a anticuerpos preparados frente a antígenos de célula hospedadora que implica la construcción de células hospedadoras destinadas a su uso para producir la proteína recombinante para que no presente adicionalmente ninguna actividad en 1-4 de β -manosiltransferasa debida a delección o interrupción de los genes 1, 2 y 3 de la β -manosiltransferasa. Otras combinaciones de delecciones se describen a continuación a modo de ilustración pero no constituyen la invención. A modo de ejemplo, se construye una célula hospedadora que no presenta actividad de β -manosiltransferasa 2 con respecto a un *N*-glicano u *O*-glicano. La célula hospedadora que carece de presentación de actividad de β -manosiltransferasa 2 se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la cepa. Una cepa de NORF es una cepa que es la misma que la cepa hospedadora excepto porque carece del marco de lectura abierto que codifica la glicoproteína recombinante. Si la proteína o glicoproteína recombinante producida por la célula hospedadora carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora, entonces la célula hospedadora es útil para producir la proteína o glicoproteína recombinante que carece de actividad de unión cruzada a los anticuerpos frente a antígenos de célula hospedadora.

Sin embargo, si se detecta la actividad de unión cruzada detectable, entonces la célula hospedadora se manipula adicionalmente para que no presente actividad de β -manosiltransferasa 1, β -manosiltransferasa 3, o β -manosiltransferasa 4y con respecto a un *N*-glicano u *O*-glicano. Por ejemplo, la célula hospedadora que carece de actividad de β -manosiltransferasa 2 se manipula adicionalmente para que carezca de actividad de β -manosiltransferasa 1. La célula hospedadora se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la cepa. Si la proteína o glicoproteína recombinante producida por la célula hospedadora carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora, entonces la célula hospedadora es útil para producir la proteína o glicoproteína recombinante que carece de actividad de unión cruzada a los anticuerpos frente a antígenos de célula hospedadora.

Sin embargo, si se detecta la actividad de unión cruzada detectable, entonces la célula hospedadora se manipula adicionalmente para que no presente actividad de β -manosiltransferasa 3 o actividad de β -manosiltransferasa 4. Por ejemplo, la célula hospedadora que carece de actividad de β -manosiltransferasa 2 y actividad de β -manosiltransferasa 1 se manipula adicionalmente para que carezca de actividad de β -manosiltransferasa 3 con respecto a un *N*-glicano u *O*-glicano. La célula hospedadora se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la cepa. Si la proteína o glicoproteína recombinante producida por la célula hospedadora carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora, entonces la célula hospedadora es útil para producir la proteína o glicoproteína recombinante que carece de actividad de unión cruzada a los anticuerpos frente a antígenos de célula hospedadora.

Sin embargo, si se detecta la actividad de unión cruzada detectable, entonces la cepa se manipula adicionalmente para que no presente actividad de β -manosiltransferasa 4 con respecto a un *N*-glicano u *O*-glicano. La célula hospedadora se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la cepa para confirmar que la proteína o glicoproteína recombinante carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora.

A modo de un ejemplo adicional, se construye una célula hospedadora de *Pichia pastoris* en la que diversas combinaciones de genes de *BMT* se someten a delección o se interrumpen. A modo de ejemplo, se construye una célula hospedadora de *Pichia pastoris* que tiene una interrupción o delección del gen *BMT2*. La célula hospedadora $\Delta bmt2$ se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la cepa. Una cepa NORF es una cepa que es la misma que la cepa hospedadora excepto porque carece del marco de lectura abierto que codifica la glicoproteína recombinante. Si la proteína o glicoproteína recombinante producida por la célula hospedadora $\Delta bmt2$ carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora, entonces la delección o interrupción de *BMT2* es suficiente para permitir que la célula hospedadora

para produzca la proteína o glicoproteína recombinante que carece de actividad de unión cruzada a los anticuerpos frente a antígenos de célula hospedadora.

Sin embargo, si se detecta la actividad de unión cruzada detectable, entonces la célula hospedadora se manipula adicionalmente para que tenga una delección de los genes *BMT1*, *BMT3*, o *BMT4*. Por ejemplo, la célula hospedadora que tiene una interrupción o delección del gen *BMT2* se manipula adicionalmente para que tenga una interrupción o interrupción del gen *BMT1*. La célula hospedadora $\Delta bmt2 \Delta bmt1$ se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la cepa. Si la proteína o glicoproteína recombinante producida por la célula hospedadora $\Delta bmt2 \Delta bmt1$ carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora, entonces las delecciones o interrupciones de *BMT1* y *BMT2* son suficientes para permitir que la célula hospedadora produzca la proteína o glicoproteína recombinante que carece de actividad de unión cruzada a los anticuerpos frente a antígenos de célula hospedadora.

Sin embargo, si se detecta la actividad de unión cruzada detectable, entonces la célula hospedadora se manipula adicionalmente para que tenga una delección de los genes *BMT3* o *BMT4*. Por ejemplo, la célula hospedadora que tiene una interrupción o delección del gen *BMT1* y *BMT2* se manipula adicionalmente para que tenga una delección o interrupción del gen *BMT3*. La célula hospedadora $\Delta bmt2 \Delta bmt1 \Delta bmt3$ se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la célula hospedadora. Si la proteína o glicoproteína recombinante producida por la célula hospedadora $\Delta bmt2 \Delta bmt1 \Delta bmt3$ carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora, entonces las delecciones e interrupciones de *BMT1*, *BMT2* y *BMT3* son suficientes para permitir que la célula hospedadora produzca la proteína o glicoproteína recombinante que carece de actividad de unión cruzada a los anticuerpos frente a antígenos de célula hospedadora.

Sin embargo, si se detecta la actividad de unión cruzada detectable, entonces la célula hospedadora se manipula adicionalmente para que tenga una delección del gen *BMT4*. La célula hospedadora $\Delta bmt2 \Delta bmt1 \Delta bmt3 \Delta bmt4$ se usa para producir la proteína o glicoproteína recombinante, que a continuación se evalúa mediante transferencia de Western o ELISA usando un anticuerpo que se ha preparado frente a una versión NORF de la cepa para confirmar que la proteína o glicoproteína recombinante carece de unión detectable al anticuerpo preparado frente a antígenos de célula hospedadora.

La presente invención también proporciona una célula hospedadora de *Pichia pastoris* recombinante en la que el gen *BMT2* y el gen *BMT1* y el gen *BMT3* han experimentado delección o interrupción y que incluye una molécula de ácido nucleico que codifica la proteína o glicoproteína recombinante. En otras realizaciones, el gen *BMT2*, el gen *BMT1* y el gen *BMT3* han experimentado delección o interrupción. En un aspecto adicional, la presente invención proporciona una célula hospedadora de *Pichia pastoris* recombinante en la que el gen *BMT1*, el gen *BMT2*, el gen *BMT3* y el gen *BMT4* han experimentado delección o interrupción y que incluye una molécula de ácido nucleico que codifica la proteína o glicoproteína recombinante.

La presente invención también proporciona un método general para producir una proteína o glicoproteína recombinante que carece de actividad de unión cruzada detectable a anticuerpos de antígeno anti-célula hospedadora que comprende proporcionar una célula hospedadora de *Pichia pastoris* recombinante en la que el gen *BMT2* y el gen *BMT1* y el gen *BMT3* han experimentado delección o interrupción y que incluye una molécula de ácido nucleico que codifica la proteína o glicoproteína recombinante; cultivar la célula hospedadora en un medio en condiciones eficaces para expresar la proteína o glicoproteína recombinante; y recuperar la proteína o glicoproteína recombinante del medio para producir la proteína o glicoproteína recombinante que carece de actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora. En otras realizaciones, el gen *BMT2*, el gen *BMT1* y el gen *BMT3* han experimentado delección o interrupción.

En un aspecto adicional, la presente invención proporciona un método general para producir una proteína o glicoproteína recombinante que carece de actividad de unión cruzada detectable a anticuerpos de antígeno anti-célula hospedadora que comprende proporcionar una célula hospedadora de *Pichia pastoris* recombinante en la que el gen *BMT1*, el gen *BMT2*, el gen *BMT3* y el gen *BMT4* han experimentado delección o interrupción y que incluye una molécula de ácido nucleico que codifica la proteína o glicoproteína recombinante; cultivar la célula hospedadora en un medio en condiciones eficaces para expresar la proteína o glicoproteína recombinante; y recuperar la proteína o glicoproteína recombinante del medio para producir la proteína o glicoproteína recombinante que carece de actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora.

La presente invención también proporciona una célula hospedadora de *Pichia pastoris* recombinante en la que el gen *BMT2* y el gen *BMT1* y el gen *BMT3* han experimentado delección o interrupción y que incluye una molécula de ácido nucleico que codifica la proteína o glicoproteína recombinante. En otras realizaciones, el gen *BMT2*, el gen *BMT1* y el gen *BMT3* han experimentado delección o interrupción. En un aspecto adicional, la presente invención proporciona una célula hospedadora de *Pichia pastoris* recombinante en la que el gen *BMT1*, el gen *BMT2*, el gen *BMT3* y el gen *BMT4* han experimentado delección o interrupción y que incluye una molécula de ácido nucleico que codifica la proteína o glicoproteína recombinante.

En general, la proteína o glicoproteína recombinante es una glicoproteína terapéutica. Los ejemplos de glicoproteínas terapéuticas contempladas, incluyen, pero no se limitan a, eritropoyetina (EPO); citoquinas tales como interferón α , interferón β , interferón γ , e interferón ω ; y factor estimulante de colonias de granulocitos (G-CSF); GM-CSF; factores de coagulación tales como factor VIII, factor IX y proteína C humana; antitrombina III; trombina; cadena α de receptor de IgE soluble; inmunoglobulinas tales como IgG, fragmentos de IgG, fusiones de IgG, e IgM; inmunoadhesiones y otras proteínas de fusión de Fc tales como proteínas de fusión a receptor de TNF-Fc solubles; proteínas de fusión de RAGE-Fc; interleuquinas; uroquinasa; quimasa; e inhibidor de urea y tripsina; proteína de unión a IGF; factor de crecimiento epidérmico; factor de liberación de hormona de crecimiento; proteína de fusión de anexina V; angiostatina; factor 2 de crecimiento endotelial vascular; factor 1 inhibitorio de precursor mieloide; osteoprotegerina; α -1-antitripsina; α -feto proteínas; DNasa II; kringle 3 de plasminógeno humano; glucocerebrosidasa; proteína 1 de unión del TNF; hormona foliculoestimulante; antígeno 4 - Ig asociado a linfocitos T citotóxicos; activador transmembrana y modulador de calcio y ligando de ciclofilina; proteína 1 de tipo glucagón; y agonista de receptores de IL-2.

En aspectos en particular de la invención, la molécula de ácido nucleico que codifica la proteína o glicoproteína recombinante está optimizada por codón para aumentar la expresión de la proteína o glicoproteína recombinante en la célula hospedadora. Por ejemplo, como se muestra en los ejemplos, la molécula de ácido nucleico que codifica la forma madura humana de la eritropoyetina se optimizó con codón para aumento de la expresión de la eritropoyetina en una levadura metilotrófica tal como una cepa de *Pichia pastoris* que se había modificado ingeniería genética para producir una variante de eritropoyetina que comprendía N-glicanos biantenarios en donde la glicofoma predominante comprendía ambas antenas sialiladas de forma terminal.

Las células hospedadoras adecuadas incluyen cualquier célula hospedadora que incluya homólogos de los genes *BMT1*, *BMT2*, *BMT3* y/o *BMT4* de *Pichia pastoris*. En la actualidad, los ejemplos de tales células hospedadoras incluyen *Candida albicans* y la levadura metilotrófica *Pichia pastoris*. La descripción describe que la célula hospedadora es una levadura metilotrófica tal como *Pichia pastoris* y mutantes de la misma y variantes de la misma modificadas por ingeniería genética. La invención sin embargo se limita a *Pichia pastoris*. La levadura metilotrófica tal como *Pichia pastoris* se puede modificar genéticamente para que exprese glicoproteínas en las que el patrón de glicosilación es similar al humano o humanizado. De esta manera, se pueden producir composiciones de glicoproteína en las que la glicofoma deseada específica es predominante en la composición. Esto se puede conseguir mediante la eliminación de enzimas de glicosilación endógenas seleccionadas y/o mediante modificación por ingeniería genética de las células hospedadoras y/o proporcionando enzimas exógenas para que imiten toda o parte de la ruta de glicosilación de mamífero como se describe en el documento US 2004/0018590. Si se desea, se puede realizar ingeniería genética adicional de la glicosilación, de modo que la glicoproteína se puede producir con o sin fucosilación del núcleo. El uso de células eucariotas inferiores es adicionalmente ventajoso porque éstas células son capaces de producir composiciones de glicoproteína altamente homogéneas, de modo que la glicofoma predominante de la glicoproteína puede estar presente en un porcentaje molar superior a un treinta por ciento de la glicoproteína en la composición. La descripción describe que la glicofoma predominante puede estar presente en un porcentaje molar superior a un cuarenta por ciento en moles, cincuenta por ciento en moles, sesenta por ciento en moles, setenta por ciento en moles y, lo más preferentemente, superior a un ochenta por ciento en moles de la glicoproteína presente en la composición. Esto se puede conseguir mediante la eliminación de enzimas de glicosilación endógenas y/o proporcionando enzimas exógenas como se describe en Gerngross *et al.*, documento de Patente de Estados Unidos n.º 7.029.872 y documento de Patente de Estados Unidos n.º 7.449.308. Por ejemplo, una célula hospedadora de seleccionar modificar por ingeniería para que esté suprimida en actividades de 1,6-manosil transferasa, que de otro modo podrían añadir restos de manosa en el N-glicano en una glicoproteína.

Por lo tanto, el control de la O-glicosilación puede ser útil para producir lipoproteínas en particular en las células hospedadoras desveladas en el presente documento con un rendimiento total mejor o con un rendimiento de la glicoproteína ensamblada de forma apropiada. Parece que la reducción o eliminación de la O-glicosilación tiene un efecto beneficioso en el ensamblaje y transporte de glicoproteínas tales como anticuerpos completos ya que atraviesan la ruta secretoria y son transportadas a la superficie celular. Por lo tanto, en células en las que la O-glicosilación está controlada, el rendimiento de las glicoproteínas ensambladas de forma apropiada tales como fragmentos de anticuerpo aumenta con respecto al rendimiento obtenido en las células hospedadoras en las que la O-glicosilación no está controlada.

Por lo tanto, los métodos desvelados el presente documento pueden usar cualquier célula hospedadora que se haya modificado genéticamente para producir glicoproteínas en las que el N-glicano predominante se selecciona entre el grupo que consiste en N-glicanos complejos, N-glicanos híbridos y N-glicanos de alta manosa en el que los N-glicanos complejos se seleccionan entre el grupo que consiste en $\text{Man}_3\text{GlcNAc}_2$, $\text{GlcNAc}_{(1-4)}\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_{(1-4)}\text{GlcNAc}_{(1-4)}\text{Man}_3\text{GlcNAc}_2$ y $\text{NANA}_{(1-4)}\text{Gal}_{(1-4)}\text{Man}_3\text{GlcNAc}_2$; los N-glicanos híbridos se seleccionan entre el grupo que consiste en $\text{Man}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_5\text{GlcNAc}_2$, $\text{GalGlcNAcMan}_5\text{GlcNAc}_2$ y $\text{NANAGlcNAcMan}_5\text{GlcNAc}_2$; y los N-glicanos de alta manosa se seleccionan entre el grupo que consiste en $\text{Man}_6\text{GlcNAc}_2$, $\text{Man}_7\text{GlcNAc}_2$, $\text{Man}_8\text{GlcNAc}_2$ y $\text{Man}_9\text{GlcNAc}_2$. Los ejemplos de estructuras de N-glicano incluyen, pero no se limitan a, $\text{Man}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_3\text{GlcNAc}_2$, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$, $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$, $\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$, $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_2\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_2\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_3\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_3\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}_4\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$,

NANAGal₂GlcNAc₂Man₃GlcNAc₂, NANA₂Gal₂GlcNAc₂Man₃GlcNAc₂, NANA₃Gal₃GlcNAc₃Man₃GlcNAc₂ y
NANA₄Gal₄GlcNAc₄Man₃GlcNAc₂.

Los marcadores seleccionables de levadura que se pueden usar para construir las células hospedadoras recombinantes incluyen marcadores de resistencia a fármaco y funciones genéticas que permiten que la célula hospedadora de levadura sintetice nutrientes celulares esenciales, por ejemplo aminoácidos. Los marcadores de resistencia a fármaco que se usan normalmente en levadura incluyen cloramfenicol, kanamicina, metotrexato, G418 (geneticina), Zeocina y similares. Las funciones genéticas que permiten que la célula hospedadora sintetice nutrientes celulares esenciales se usan con cepas de levadura disponibles que tienen mutaciones auxotróficas en la función genómica correspondiente. Los marcadores seleccionables de levadura comunes proporcionan funciones genéticas para sintetizar leucina (*LEU2*), triptófano (*TRP1* y *TRP2*), prolina (*PRO1*), uracilo (*URA3*, *URA5*, *URA6*), histidina (*HIS3*), lisina (*LYS2*), adenina (*ADE1* o *ADE2*) y similares. Otros marcadores seleccionables de levadura incluyen el gen *ARR3* de *S. cerevisiae*, que confiere resistencia a arsenito a células de levadura que se cultivan en presencia de arsenito (Bobrowicz *et al.*, Yeast, 13: 819-828 (1997); Wysocki *et al.*, J. Biol. Chem. 272: 30061-30066 (1997)). Un número de sitios de integración adecuados incluyen los enumerados en el documento de Patente de Estados Unidos n.º 7.479.389 e incluyen homólogos para loci conocidos para *Saccharomyces cerevisiae* y otras levaduras u hongos. Los métodos para integrar vectores en levaduras se conocen bien (véase por ejemplo, documento de Patente de Estados Unidos n.º 7.479.389, documento de Patente de Estados Unidos n.º 7.514.253, Solicitud Publicada de Estados Unidos n.º 2009012400 y documento WO2009/085135). Los ejemplos de sitios de inserción incluyen, pero no se limitan a, genes *ADE* de *Pichia*; genes *TRP* de *Pichia* (incluyendo *TRP1* a *TRP2*); genes *MCA* de *Pichia*; genes *CYM* de *Pichia*; genes *PEP* de *Pichia*; genes *PRB* de *Pichia*; y genes *LEU* de *Pichia*. Los genes *ADE1* y *ARG4* de *Pichia* se han descrito en Lin Cereghino *et al.*, Gene 263: 159-169 (2001) y en el documento de Patente de Estados Unidos n.º 4.818.700, los genes *HIS3* y *TRP1* se han descrito en Cosano *et al.*, Yeast 14: 861-867 (1998), *HIS4* se ha descrito en GenBank con el n.º de Registro X56180.

También se describió un método para producir una eritropoyetina humana madura en levadura metilotrófica tal como *Pichia pastoris* que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora. El método comprende proporcionar una célula hospedadora de *Pichia pastoris* recombinante modificada por ingeniería genética para producir *N*-glicanos biantenarios terminados en ácido siálico y en la que al menos los genes *BMT1*, *BMT2* y *BMT3* han experimentado delección o interrupción y que incluye dos o más moléculas de ácido nucleico, cada una codificando una proteína de fusión que comprende una eritropoyetina humana madura EPO fusionada a un péptido señal se dirige al aparato del RE y que se retira cuando la proteína de fusión está en el RE de cultivo de la célula hospedadora en un medio en condiciones eficaces para expresar y procesar la primera y segunda proteínas de fusión; y recuperar la eritropoyetina humana madura del medio para producir la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora.

En aspectos en particular, la molécula de ácido nucleico que codifica la eritropoyetina humana madura está optimizada por codón a la expresión óptima en la levadura metilotrófica tal como *Pichia pastoris*. Como se muestra en los ejemplos, la eritropoyetina humana madura está codificada como una proteína de fusión en la que la EPO está fusionada en el extremo N-terminal de la forma madura de la eritropoyetina par con el extremo C-terminal de un péptido señal y dirige la proteína de fusión a las ruta secretoria para procesamiento, incluyendo glicosilación. Los ejemplos de péptidos señal incluyen, pero no se limitan a, el péptido señal α MATpre de *S. cerevisiae* o un péptido señal de lisozima de pollo. En lugar de las que se desvelan en el presente documento, se pueden usar otras secuencias señal, por ejemplo, el péptido señal de α -amilasa de *Aspergillus niger* y péptido señal de albúmina de suero humano (HSA). En una realización, una primera molécula de ácido nucleico codifica una proteína de fusión en la que la eritropoyetina madura está condensada con el péptido señal α MATpre de *S. cerevisiae* y la segunda molécula de ácido nucleico codifica una proteína de fusión en la que la eritropoyetina madura está condensada con el péptido señal α MATpre de *S. cerevisiae* un péptido señal de lisozima de pollo. El péptido señal se puede fusionar con la eritropoyetina humana madura mediante un péptido conector que puede contener uno o más sitios de escisión de proteasa.

La célula hospedadora está modificada con ingeniería genética para producir *N*-glicanos biantenarios terminados en ácido siálico y en la que al menos los genes *BMT1*, *BMT2* y *BMT3* han experimentado delección o interrupción.

La detección de actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se puede determinar en un ELISA de sándwich o en una transferencia de Western.

En aspectos adicionales, la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora incluye una etapa de cromatografía de intercambio catiónico y/o una etapa de cromatografía con hidroxipatita y/o una etapa de cromatografía de intercambio aniónico. En una realización, la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora comprende una etapa de cromatografía de intercambio catiónico seguido

de una etapa de cromatografía con hidroxipatita. Opcionalmente, la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora incluye una etapa de cromatografiar aniónica.

5 EJEMPLO 1

La cepa **YGLY3159** de *Pichia pastoris* modificada por ingeniería genética es una cepa que produce eritropoyetina humana recombinante con *N*-glicanos sialilados (rhEPO). La construcción de la cepa se ha descrito en la Solicitud Publicada de Estados Unidos n.º 20080139470 y se ilustra de forma esquemática en la **Figura 1**. En resumen, la cepa se construyó como sigue a continuación.

La cepa **YGLY3159** se construyó a partir de la cepa **NRRL-Y 11430** de *Pichia pastoris* de tipo silvestre usando métodos descritos anteriormente (véase por ejemplo, documento de Patente de Estados Unidos n.º 7.449.308; documento de Patente de Estados Unidos n.º 7.479.389; Solicitud Publicada de Estados Unidos n.º 20090124000; Solicitud de PCT Publicada n.º WO2009085135; Nett y Gerngross, Yeast 20: 1279 (2003); Choi *et al.*, Proc. Natl. Acad. Sci. USA 100: 5022 (2003); Hamilton *et al.*, Science 301: 1244 (2003)). Todos los plásmidos se prepararon en un plásmido pUC19 usando procedimientos convencionales de biología molecular. Para secuencias de nucleótidos que se optimizaron para expresión en *P. pastoris*, las secuencias de nucleótidos nativas se analizaron con el software GENEOPTIMIZER (GeneArt, Regensburg, Alemania) y los resultados se usaron para generar secuencias de nucleótidos en las que los codones se optimizaron para expresión de *P. pastoris*. Las cepas de levadura se transformaron mediante electroporación (usando técnicas convencionales tal como lo recomienda el fabricante del electroporador BioRad).

El plásmido pGLY6 (**Figura 2**) es un vector de integración que se dirige al locus de *URA5* y contiene una molécula de ácido nucleico que comprende el gen de invertasa o unidad de transcripción de *S. cerevisiae* (*ScSUC2*; SEQ ID NO: 1) flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *URA5* de *P. pastoris* (SEQ ID NO: 59) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *URA5* de *P. pastoris* (SEQ ID NO: 60). El plásmido pGLY6 se linealizó y el plásmido linealizado se transformó en la cepa **NRRL-Y 11430** de tipo silvestre para producir una serie de cepas en las que el gen *ScSUC2* estaba insertado en el locus de *URA5* mediante recombinación homóloga de doble cruce. La cepa **YGLY1-3** se seleccionó a partir de las cepas producidas y es auxotrófica para uracilo.

El plásmido pGLY40 (**Figura 3**) es un vector de integración que se dirige al locus de *OCH1* y contiene una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción de *P. pastoris* (SEQ ID NO: 61) flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ* (SEQ ID NO: 62) que a su vez está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *OCH1* (SEQ ID NO: 64) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *OCH1* (SEQ ID NO: 65). El plásmido pGLY40 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY1-3** para producir una serie de cepas en las que el gen *URA5* flanqueado por las repeticiones de *lacZ* se ha insertado en el locus de *OCH1* mediante recombinación homóloga de doble cruce. La cepa **YGLY2-3** se seleccionó a partir de las cepas producidas y es prototrófica para *URA5*. Se hizo contraselección de la cepa **YGLY2-3** en presencia de ácido 5-fluoroorótico (5-FOA) para producir una serie de cepas en las que el gen *URA5* se ha perdido y solamente las repeticiones de *lacZ* permanecen en el locus de *OCH1*. Esto hace que la cepa sea auxotrófica para uracilo. La cepa **YGLY4-3** se seleccionó.

El plásmido pGLY43a (**Figura 4**) es un vector de integración que se dirige al locus de *BMT2* y contiene una molécula de ácido nucleico que comprende el gen transportador de UDP-N-acetilglucosamina (UDP-GlcNAc) o unidad de transcripción de *K. lactis* (*KIMNN2-2*, SEQ ID NO: 3) adyacente a una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ*. Los genes adyacentes están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *BMT2* (SEQ ID NO: 66) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *BMT2* (SEQ ID NO: 67). El plásmido pGLY43a se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY4-3** para producir una serie de cepas en las que el gen *KIMNN2-2* y el gen *URA5* flanqueado por las repeticiones de *lacZ* se ha insertado en el locus de *BMT2* mediante recombinación homóloga de doble cruce. El gen *BMT2* se ha desvelado en Mille *et al.*, J. Biol. Chem. 283: 9724-9736 (2008) tienen el documento de Patente de Estados Unidos n.º 7.465.557. La cepa **YGLY6-3** se seleccionó a partir de las cepas producidas y es prototrófica para uracilo. Se hizo contraselección de la cepa **YGLY6-3** en presencia de 5-FOA para producir cepas en las que el gen *URA5* se ha perdido y solamente permanecen las repeticiones de *lacZ*. Esto hace que la cepa sea auxotrófica para uracilo. La cepa **YGLY8-3** se seleccionó.

El plásmido pGLY48 (**Figura 5**) es un vector de integración que se dirige al locus de *MNN4L1* y contiene un casete de expresión que comprende una molécula de ácido nucleico que codifica el homólogo de ratón del marco de

lectura abierto (ORF) del transportador de UDP-GlcNAc (SEQ ID NO: 17) unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GAPDH* de *P. pastoris* (SEQ ID NO: 53) y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende las secuencias de terminación *CYC* de *S. cerevisiae* (SEQ ID NO: 56) adyacentes a una molécula de ácido nucleico que comprende el gen *URA5* de *P. pastoris* flanqueado por repeticiones de *lacZ* y en el que los casetes de expresión juntos están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *MNN4L1* de *P. pastoris* (SEQ ID NO: 76) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *MNN4L1* (SEQ ID NO: 77). El plásmido pGLY48 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY8-3** para producir una serie de cepas en las que el casete de expresión que codifica el transportador UDP-GlcNAc de ratón y el gen *URA5* se ha insertado en el locus de *MNN4L1* mediante recombinación homóloga de doble cruce. El gen *MNN4L1* (también denominado *MNN4B*) se ha desvelado en el documento de Patente de Estados Unidos n.º 7.259.007. La cepa **YGLY10-3** se seleccionó a partir de las cepas producidas y a continuación se hizo contraselección en presencia de 5-FOA para producir una serie de cepas en las que el gen *URA5* se ha perdido y solamente permanecen las repeticiones de *lacZ*. La cepa **YGLY12-3** se seleccionó.

El plásmido pGLY45 (**Figura 6**) es un vector de integración que se dirige a los loci *PNO1/MNN4* y contiene una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ* que a su vez está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *PNO1* (SEQ ID NO: 74) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *MNN4* (SEQ ID NO: 75). El plásmido pGLY45 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY12-3** para producir una serie de cepas en las que el gen *URA5* flanqueado por las repeticiones de *lacZ* se ha insertado en los loci *MNN4* mediante recombinación homóloga de doble cruce. El gen *PNO1* se ha desvelado en el documento de Patente de Estados Unidos n.º 7.198.921 y el gen *MNN4* (también denominado *MNN4B*) se ha desvelado en el documento de Patente de Estados Unidos n.º 7.259.007. La cepa **YGLY14-3** se seleccionó a partir de las cepas producidas y a continuación se hizo contraselección en presencia de 5-FOA para producir una serie de cepas en las que el gen *URA5* se ha perdido y solamente permanecen las repeticiones de *lacZ*. La cepa **YGLY16-3** se seleccionó.

El plásmido pGLY247 (**Figura 7**) es un vector de integración que se dirige al locus de *MET16* y contiene una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ* que a su vez está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *MET16* (SEQ ID NO: 84) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *MET16* (SEQ ID NO: 85). El plásmido pGLY247 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY16-3** para producir una serie de cepas en las que el *URA5* flanqueado por las repeticiones de *lacZ* se ha insertado en el locus de *MET16* mediante recombinación homóloga de doble cruce. La cepa **YGLY20-3** se seleccionó.

El plásmido pGLY248 (**Figura 8**) es un vector de integración que se dirige al locus de *URA5* y contiene una molécula de ácido nucleico que comprende el gen de *MET16* *P. pastoris* (SEQ ID NO: 86) flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *URA5* (SEQ ID NO: 59) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *URA5* (SEQ ID NO: 60). El plásmido pGLY248 se linealizó y el plásmido linealizado se transformó en la cepa **YGLY20-3** para producir una serie de cepas en las que el gen *ScSUC2* insertado en el locus de *URA5* se ha reemplazado por el gen *MET16* mediante recombinación homóloga de doble cruce. La cepa **YGLY24-3** se seleccionó y a continuación se hizo contraselección en presencia de 5-FOA para producir una serie de cepas en las que el gen *URA5* insertado en el locus de *MET16* se ha perdido y solamente permanecen las repeticiones de *lacZ*. La cepa **YGLY24-3** se seleccionó.

El plásmido pGLY582 (**Figura 9**) es un vector de integración que se dirige al locus de *HIS1* y contiene en tándem cuatro casetes de expresión que codifican (1) la epimerasa de UDP-glucosa (*ScGAL10*) de *S. cerevisiae*, (2) el dominio catalítico de la galactosiltransferasa I (*hGalT*) humana fusionado en el extremo N-terminal al péptido líder *KRE2-s* de *S. cerevisiae* (33) para dirigir la enzima quimérica al RE o aparato de Golgi, (3) el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por repeticiones de *lacZ* y (4) el transportador de UDP-galactosa (*DmUGT*) de *D. melanogaster*. El casete de expresión que codifica el *ScGAL10* comprende una molécula de ácido nucleico que codifica el ORF de *ScGAL10* (SEQ ID NO: 21) unida de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *PMA1* de *P. pastoris* (SEQ ID NO: 45) y unida de forma operativa en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *PMA1* de *P. pastoris* (SEQ ID NO: 46). El casete de expresión que codifica la galactosiltransferasa I quimérica comprende una molécula de ácido nucleico que codifica el dominio catalítico *hGalT* optimizado por codón para expresión en *P. pastoris* (SEQ ID NO: 23) fusionado en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el líder 33 de *KRE2-s* (SEQ ID NO: 13), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GAPDH* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la

transcripción *CYC* de *S. cerevisiae*. El casete de expresión de *URA5* comprende una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ*. El casete de expresión que codifica el *DMUGT* comprende una molécula de ácido nucleico que codifica el ORF de *DmUGT* (SEQ ID NO: 19) unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *OCH1* de *P. pastoris* (SEQ ID NO: 47) y unida de forma operativa en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *ALG12* de *P. pastoris* (SEQ ID NO: 48). Los cuatro casetes en tándem están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *HIS1* (SEQ ID NO: 87) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *HIS1* (SEQ ID NO: 88). El plásmido pGLY582 se linealizó y el plásmido linealizado se transformó en la cepa **YGLY24-3** para producir una serie de cepas en las que los cuatro casetes de expresión en tándem se han insertado en el locus de *HIS1* mediante recombinación homóloga. La cepa **YGLY58** se seleccionó y es auxotrófica para histidina y prototrófica para uridina.

El plásmido pGLY167b (**Figura 10**) es un vector de integración que se dirige al locus de *ARG1* y contiene en tándem tres casetes de expresión que codifican (1) el dominio catalítico de la manosidasa II de *D. melanogaster* (KD) fusionado en el extremo N-terminal al péptido líder *MNN2* de *S. cerevisiae* (53) para dirigir la enzima quimérica al RE o aparato de Golgi, (2) el gen *HIS1* o unidad de transcripción de *P. pastoris* y (3) el dominio catalítico de la N-acetilglucosamina (GlcNAc) transferasa II de rata (TC) fusionado en el extremo N-terminal al péptido líder *MNN2* de *S. cerevisiae* (54) para dirigir la enzima quimérica al RE o aparato de Golgi. El casete de expresión que codifica el KD53 comprende una molécula de ácido nucleico que codifica el dominio catalítico de la manosidasa II de *D. melanogaster* optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 33) fusionada en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el líder 53 de *MNN2* (SEQ ID NO: 5), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GAPDH* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *CYC* de *S. cerevisiae*. El casete de expresión de *HIS1* comprende una molécula de ácido nucleico que comprende el gen *HIS1* o unidad de transcripción de *P. pastoris* (SEQ ID NO: 89). El casete de expresión que codifica el TC54 comprende una molécula de ácido nucleico que codifica el dominio catalítico de la NGlcNAc transferasa II de rata optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 31) fusionado en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el líder 54 de *MNN2* (SEQ ID NO: 7), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *PMA1* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *PMA1* de *P. pastoris*. Los tres casetes en tándem están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *ARG1* gene (SEQ ID NO: 79) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *ARG1* (SEQ ID NO: 80). El plásmido pGLY167b se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY58** para producir un número de cepas (en las que los tres cassettes de expresión en tándem se han insertado en el locus de *ARG1* mediante recombinación homóloga de doble cruce. La cepa **YGLY73** se seleccionó a partir de las cepas producidas y es auxotrófica para arginina y prototrófica para uridina e histidina. A continuación se hizo contraselección de la cepa en presencia de 5-FOA para producir un número de cepas que ahora son auxotróficas para uridina. La cepa **YGLY1272** se seleccionó.

El plásmido pGLY1430 (**Figura 11**) es un vector de integración de KINKO que se dirige al locus de *ADE1* sin interrupción de la expresión del locus y contiene en tándem cuatro casetes de expresión que codifican (1) el dominio catalítico de la GlcNAc transferasa I humana (NA) fusionado en el extremo N-terminal al péptido líder *SEC12* de *P. pastoris* (10) para dirigir la enzima quimérica al RE o aparato de Golgi, (2) homólogo de ratón del transportador de UDP-GlcNAc (MmTr), (3) el dominio catalítico de la manosidasa IA de ratón (FB) fusionado en el extremo N-terminal al péptido líder *SEC12* (FB8) de *S. cerevisiae* (8) para dirigir la enzima quimérica al RE o aparato de Golgi y (4) el gen *URA5* o unidad de transcripción de *P. pastoris*. Los vectores de integración de KINKO (Con Supresión Genética con poca o Ninguna Supresión Genética) permiten la inserción de ADN heterólogo en un locus dirigido sin intervención de la expresión del gen en el locus dirigido y se han descrito en la Solicitud Publicada de Estados Unidos n.º 20090124000. El casete de expresión que codifica el NA10 comprende una molécula de ácido nucleico que codifica el dominio catalítico de la GlcNAc transferasa I humana optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 25) fusionado en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el líder 10 de *SEC12* (SEQ ID NO: 11), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *PMA1* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *PMA1* de *P. pastoris*. El casete de expresión que codifica MmTr comprende una molécula de ácido nucleico que codifica el homólogo de ratón del ORF transportador de UDP-GlcNAc unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *SEC4* de *P. pastoris* (SEQ ID NO: 49) y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende las secuencias de terminación *OCH1* de *P. pastoris* (SEQ ID NO: 50). El casete de expresión que codifica el FB8 comprende una molécula de ácido nucleico que codifica el dominio catalítico de la manosidasa IA de ratón (SEQ ID NO: 27) fusionado en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el líder 8 de *SEC12-m* (SEQ ID NO: 15), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GADPH* de *P.*

pastoris y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *CYC* de *S. cerevisiae*. El casete de expresión de *URA5* comprende una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ*. Los cuatro casetes en tándem están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' y el ORF completo del gen *ADE1* (SEQ ID NO: 82) seguido de una secuencia de terminación *ALG3* de *P. pastoris* (SEQ ID NO: 54) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *ADE1* (SEQ ID NO: 83). El plásmido pGLY1430 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY1272** para producir una serie de cepas en las que los cuatro casetes de expresión en tándem se han insertado en el locus de *ADE1* inmediatamente después del ORF de *ADE1* mediante recombinación homóloga de doble cruce. La cepa **YGLY1305** se seleccionó a partir de las cepas producidas y es auxotrófica para arginina y ahora es prototrófica para uridina, histidina y adenina. A continuación se hizo contraselección de la cepa en presencia de 5-FOA para producir un número de cepas que ahora son auxotróficas para uridina. La cepa **YGLY1461** se seleccionó y es capaz de formar glicoproteínas que tienen predominantemente N-glicanos terminados en galactosa.

El plásmido pGFI165 (**Figura 12**) es un vector de integración de KINKO que se dirige al locus de *PRO1* sin interrupción de la expresión del locus y contiene casetes de expresión que codifican (1) el dominio catalítico de la α -1,2-manosidasa de *T. reesei* fusionado en el extremo N-terminal al péptido señal α MATpre (α MATTrMan) de *S. cerevisiae* para dirigir la proteína quimérica a la ruta secretoria y secreción desde la célula y (2) el gen *URA5* o unidad de transcripción de *P. pastoris*. El casete de expresión que codifica el α MATTrMan comprende una molécula de ácido nucleico que codifica el dominio catalítico de *T. reesei* (SEQ ID NO: 29) fusionado en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el péptido señal α MATpre de *S. cerevisiae* (SEQ ID NO: 9), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GAPDH* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *CYC* de *S. cerevisiae*. El casete de expresión de *URA5* comprende una molécula de ácido nucleico que comprende el gen *URA5* o unidad de transcripción de *P. pastoris* flanqueado por moléculas de ácido nucleico que comprenden repeticiones de *lacZ*. Los dos casetes en tándem están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' y el ORF completo del gen *PRO1* (SEQ ID NO: 90) seguido de una secuencia de terminación *ALG3* de *P. pastoris* y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *PRO1* (SEQ ID NO: 91). El plásmido pGFI165 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY1461** para producir una serie de cepas en las que los dos casetes de expresión se han insertado en el locus de *PRO1* inmediatamente después del ORF de *PRO1* mediante recombinación homóloga de doble cruce. La cepa **YGLY1703** se seleccionó a partir de las cepas producidas y es auxotrófica para arginina y prototrófica para uridina, histidina, adenina y prolina. Esta cepa es capaz de producir glicoproteínas que tienen O-glicosilación reducida (véase la Solicitud Publicada de Estados Unidos n.º 20090170159).

El plásmido pGLY2088 (**Figura 13**) es un vector de integración que se dirige al locus de *TARP2* o al locus *AOX1* y contiene casetes de expresión que codifican (1) eritropoyetina humana madura (EPO) fusionada en el extremo N-terminal a un péptido señal α MATpre de *S. cerevisiae* (alfa MF-pre) para dirigir la proteína quimérica a la ruta secretoria y secreción desde la célula y (2) la proteína de resistencia a zeocina (Sh ble o Zeocina^R). El casete de expresión que codifica la EPO comprende una molécula de ácido nucleico que codifica la eritropoyetina EPO humana madura optimizada por codón para su expresión en *P. pastoris* (SEQ ID NO: 92) fusionada en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el péptido señal α MATpre de *S. cerevisiae*, que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *AOX1* de *P. pastoris* (SEQ ID NO: 55) y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *CYC* de *S. cerevisiae*. El casete de expresión de Zeocina^R comprende una molécula de ácido nucleico que codifica el ORF de Sh ble (SEQ ID NO: 58) que se une de forma operativa al extremo en la posición 5' al promotor *TEF1* de *S. cerevisiae* (SEQ ID NO: 57) y en el extremo en la posición 3' a la secuencia de terminación de *CYC* de *S. cerevisiae*. Los dos casetes en tándem están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos que comprende el gen *TRP2* (SEQ ID NO: 78). El plásmido pGLY2088 se linealizó en el sitio *PmeI* y se transformó en la cepa **YGLY1703** para producir una serie de cepas en las que los dos casetes de expresión se han insertado en el locus de *AOX1* mediante recombinación homóloga de cruce individual incorporado, que da como resultado múltiples copias del casete de expresión de EPO insertado en el locus de *AOX1* de sin interrumpir el locus de *AOX1*. La cepa **YGLY2849** se seleccionó a partir de las cepas producidas y es auxotrófica para arginina y ahora es prototrófica para uridina, histidina, adenina y prolina. La cepa contiene aproximadamente de tres a cuatro copias del casete de expresión de EPO tal como se determina mediante la medición de la intensidad de los datos de secuenciación del ADN aislado de la cepa. Durante el procesamiento de la EPO quimérica en el RE y en el aparato de Golgi, el péptido líder se retira. Por lo tanto, la rhEPO producida es la forma madura de la EPO.

El plásmido pGLY2456 (**Figura 14**) es un vector de integración de KINKO que se dirige al locus de *TARP2* sin interrupción de la expresión del locus y contiene seis casetes de expresión que codifican (1) el transportador de CMP-ácido siálico de ratón (Transp de mCMP-Sia), (2) la UDP-GlcNAc 2-epimerasa/N-acetilmanosamina quinasa

humana (hGNE), (3) el gen *ARG1* o unidad de transcripción de *Pichia pastoris*, (4) la CMP-ácido siálico sintasa humana (hCMP-NANA), (5) la N-acetilneuraminato-9-fosfato sintasa humana (hSIAP S), (6) el dominio catalítico de a-2,6-sialiltransferasa de ratón (mST6) fusionado en el extremo N-terminal al péptido líder *KRE2* de *S. cerevisiae* (33) para dirigir la enzima quimérica al RE o aparato de Golgi y el gen *ARG1* o unidad de transcripción de *P. pastoris*. El casete de expresión que codifica el transportador de CMP-ácido siálico de ratón comprende una molécula de ácido nucleico que codifica el ORF del Transp de mCMP Sia optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 35), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *PMA1* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *PMA1* de *P. pastoris*. El casete de expresión que codifica la UDP-GlcNAc 2-epimerasa/N-acetilmanosamina quinasa humana comprende una molécula de ácido nucleico que codifica el ORF de hGNE optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 37), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GAPDH* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *CYC* de *S. cerevisiae*. El casete de expresión que codifica el gen *ARG1* de *P. pastoris* comprende (SEQ ID NO: 81). El casete de expresión que codifica la CMP-ácido siálico sintasa humana comprende una molécula de ácido nucleico que codifica el ORF de hCMP-NANA S optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 39), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *GPDH* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *CYC* de *S. cerevisiae*. El casete de expresión que codifica la N-acetilneuraminato-9-fosfato sintasa humana comprende una molécula de ácido nucleico que codifica el ORF de hSIAP S optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 41), que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *PMA1* de *P. pastoris* y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción *PMA1* de *P. pastoris*. El casete de expresión que codifica la a-2,6-sialiltransferasa de ratón quimérica comprende una molécula de ácido nucleico que codifica el dominio catalítico de mST6 optimizado por codón para su expresión en *P. pastoris* (SEQ ID NO: 43) fusionado en el extremo en la posición 5' a una molécula de ácido nucleico que codifica el péptido señal *KRE2* de *S. cerevisiae*, que está unido de forma operativa en el extremo en la posición 5' a una molécula de ácido nucleico que comprende el promotor *TEF* de *P. pastoris* (SEQ ID NO: 51) y en el extremo en la posición 3' a una molécula de ácido nucleico que comprende la secuencia de terminación de la transcripción de *TEF* de *P. pastoris* (SEQ ID NO: 52). Los seis casetes en tándem están flanqueados en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen ORF que codifica Trp2p que termina en el codón de parada (SEQ ID NO: 98) seguido de una secuencia de terminación *ALG3* de *P. pastoris* y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *TRP2* (SEQ ID NO: 99). El plásmido pGLY2456 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY2849** para producir una serie de cepas en las que los seis casetes de expresión se han insertado en el locus de *TARP2* inmediatamente después del ORF de *TARP2* mediante recombinación homóloga de doble cruce. La cepa **YGLY3159** se seleccionó a partir de las cepas producidas y ahora es prototrófica para uridina, histidina, adenina, prolina, arginina y triptófano. La cepa es resistente a Zeocina y contiene aproximadamente de tres a cuatro copias del casete de expresión de EPO. La cepa producía rhEPO; sin embargo, usando los métodos en el Ejemplo 5, la rhEPO tenía reactividad cruzada de unión a anticuerpos preparados frente a HCA (véase el Ejemplo 6).

Aunque los diversos casetes de expresión se integraron en loci en particular del genoma de *Pichia pastoris* en los ejemplos en el presente documento, se entiende que la operación de la invención es independiente de los loci usados para integración. Los loci distintos de los desvelados en el presente documento se pueden usar para integración de los casetes de expresión. Los sitios de integración adecuados incluyen los enumerados en la Solicitud Publicada de Estados Unidos n.º 20070072262 e incluyen homólogos para loci conocidos de *Saccharomyces cerevisiae* y otras levaduras u hongos.

EJEMPLO 2

La cepa **YGLY3159** en el Ejemplo 1 se modificó por ingeniería genética adicionalmente para interrumpir los genes *BMT1*, *BMT3* y *BMT4* como sigue a continuación.

Se hizo contraselección de la cepa **YGLY3159** en presencia de 5-FOA para producir la cepa **YGLY3225**, que ahora es auxotrófica para uridina.

El plásmido pGLY3411 (pSH1092) (**Figura 15**) es un vector de integración que contiene el casete de expresión que comprende el gen *URA5* de *P. pastoris* flanqueado por repeticiones de *lacZ* flanqueado en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT4* (PpPBS4 5') de *P. pastoris* (SEQ ID NO: 72) y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT4* de *P. pastoris* (SEQ ID NO: 73). El plásmido pGLY3411 se linealizó y el plásmido linealizado se transformó en **YGLY3159** para producir una serie de cepas en las que el casete de expresión de *URA5* se ha insertado en el locus de *BMT4* de mediante recombinación homóloga de doble cruce. La cepa **YGLY4439** se seleccionó a partir de las cepas producidas y es prototrófica para uracilo, adenina, histidina, prolina, arginina y triptófano. La cepa es resistente a Zeocina y contiene aproximadamente de tres a cuatro copias del casete de expresión de rhEPO. La cepa tiene una interrupción o delección de los genes *BMT2* y *BMT4*.

El plásmido pGLY3430 (pSH1115) (**Figura 16**) es un vector de integración que contiene una molécula de ácido nucleico que codifica el ORF de resistencia a Nourseotricina (NATR) (originalmente de pAG25 de ERO-SCARF, Scientific Research and Development GmbH, Daimlerstrasse 13a, D-61352 Bad Homburg, Alemania, véase Goldstein *et al.*, Yeast 15: 1541 (1999)), ORF (SEQ ID NO: 102) unido de forma operativa al promotor *TEF1* de *Ashbya gossypii* (SEQ ID NO: 105) y secuencias de terminación de *TEF1* de *Ashbya gossypii* (SEQ ID NO: 106) flanqueado en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT1* de *P. pastoris* (SEQ ID NO: 68) y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT1* de *P. pastoris* (SEQ ID NO: 69). El plásmido pGLY3430 se linealizó y el plásmido linealizado se transformó en la cepa **YGLY4439** para producir una serie de cepas en las que el casete de expresión de NAT^R se ha insertado en el locus de *BMT1* mediante recombinación homóloga de doble cruce. La cepa **YGLY6661** se seleccionó a partir de las cepas producidas y es prototrófica para uracilo, adenina, histidina, prolina, arginina y triptófano. La cepa es resistente a Zeocina y Nourseotricina y contiene aproximadamente de tres a cuatro copias del casete de expresión de EPO. La cepa tiene una interrupción o delección de los genes *BMT1*, *BMT2* y *BMT4*. La cepa **YGLY7013** también se seleccionó; sin embargo, esta cepa presentaba solamente una interrupción parcial del gen *BMT1*. Estafeta se diseñó como con una interrupción o delección de los genes *BMT1*, *BMT2* y *BMT4*.

El plásmido pGLY4472 (pSH1186) (**Figura 17**) es un vector de integración que contiene una molécula de ácido nucleico que codifica el gen de higromicina B fosfotransferasa de *E. coli* (Hyg^R) ORF (SEQ ID NO: 103) unido de forma operativa al promotor *TEF1* de *Ashbya gossypii* (SEQ ID NO: 105) y secuencias de terminación de *TEF1* de *Ashbya gossypii* (SEQ ID NO: 106) flanqueado en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT3* de *P. pastoris* (SEQ ID NO: 70) y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT3* de *P. pastoris* (SEQ ID NO: 71). El plásmido pGLY3430 se linealizó y el plásmido linealizado se transformó en la cepa **YGLY6661** para producir una serie de cepas en las que el casete de expresión de Hyg^R se ha insertado en el locus de *BMT3* mediante recombinación homóloga de doble cruce. Las cepas **YGLY7361** a **YGLY7366** y las cepas **YGLY7393** a **YGLY7398** se seleccionaron a partir de las cepas producidas y son prototróficas para uracilo, adenina, histidina, prolina, arginina y triptófano. Las cepas son resistentes a Zeocina, Nourseotricina, e Higromicina y contienen aproximadamente de tres a cuatro copias del casete de expresión de EPO. Las cepas tienen interrupciones o delecciones de los genes *BMT1*, *BMT2*, *BMT3* y *BMT4* y producen rhEPO que carece de unión de reactividad cruzada a anticuerpos preparados frente a antígenos de célula hospedadora (HCA).

EJEMPLO 3

La cepa **YGLY3159** en el Ejemplo 1 se modificó por ingeniería genética adicionalmente para producir cepas en las que los genes *BMT1*, *BMT3* y *BMT4* se han interrumpido o experimentado delección y para incluir varias copias de un casete de expresión que codifica EPO humana madura fusionada con el péptido líder de lisozima de pollo. En resumen, la construcción de estas cepas a partir de **YGLY3159** se muestra en la **Figura 1** y se describe brevemente como sigue a continuación.

Se hizo contraselección de la cepa **YGLY3159** en presencia de 5-FOA para producir la cepa **YGLY3225**, que ahora es auxotrófica para uridina.

El plásmido pGLY2057 (**Figura 18**) es un vector de integración que se dirige al locus de *ADE2* y contiene un casete de expresión que codifica el gen *URA5* de *P. pastoris* flanqueado por repeticiones de *lacZ*. El casete de expresión está flanqueado en un lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 5' del gen *ADE2* (SEQ ID NO: 100) y en el otro lado por una molécula de ácido nucleico que comprende una secuencia de nucleótidos de la región en la posición 3' del gen *ADE2* (SEQ ID NO: 101). El plásmido pGLY2057 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY3225** para producir una serie de cepas en las que el casete de *URA5* se ha insertado en el locus de *ADE2* mediante recombinación homóloga de doble cruce. La cepa **YGLY3229** se seleccionó a partir de las cepas producidas y es auxotrófica para adenina y prototrófica para uridina, histidina, prolina, arginina y triptófano. La cepa es resistente a Zeocina y contiene aproximadamente de tres a cuatro copias del casete de expresión de EPO.

El plásmido pGLY2680 (**Figura 19**) es un vector de integración que se puede dirigir al locus de *TRP2* o *AOX1* y contiene casetes de expresión que codifican (1) una EPO quimérica que comprende la eritropoyetina madura humana (EPO) fusionada en el extremo N-terminal con un péptido señal de lisozima de pollo para dirigir la proteína quimérica a la ruta secretoria y secreción desde la célula y (2) el gen *ADE2* de *P. pastoris* sin un promotor. El gen *ADE2* se transcribe escasamente desde un promotor críptico. Por lo tanto, la selección de cepas de levadura *ade2Δ* transformarlas con el vector en medio no complementado con adenina requiere múltiples copias del vector a integrar en el genoma para hacer que sea prototrófico recombinante para adenina. Dado que el vector incluye adicionalmente el casete de expresión de EPO, la levadura recombinante también incluirá múltiples copias del casete de EPO integradas en el genoma. Este vector y método se ha descrito en la Solicitud de PCT Publicada WO2009085135. La secuencia de ADN que codifica el péptido señal de lisozima de pollo se muestra en la SEQ ID NO: 94, el ORF optimizado por codón que codifica la EPO madura humana se muestra en la SEQ ID NO: 92 y el gen *ADE2* de *P. pastoris* sin su promotor pero incluyendo sus secuencias de terminación se muestra en la SEQ ID NO: 96. La EPO quimérica está unida de forma operativa al promotor de *AOX1* y secuencias de terminación CYC de *S. cerevisiae*. Los dos casetes en tándem están flanqueados en un lado por una molécula de ácido nucleico que

comprende una secuencia de nucleótidos que comprende el gen *TRP2*.

El plásmido pGLY2680 se linealizó en el sitio *PmeI* y se transformó en **YGLY3229** para producir una serie de cepas en las que los dos casetes de expresión se han insertado en el locus de *AOX1* mediante recombinación homóloga de cruce individual incorporado, que da como resultado múltiples copias del casete de expresión de EPO insertado en el locus de *AOX1* sin interrumpir el locus de *AOX1*. La cepa **YGLY4209** se seleccionó a partir de las cepas producidas. En esta cepa hay aproximadamente 5-7 copias del casete de expresión de EPO tal como se determina mediante la medición de la intensidad de los datos de secuenciación del ADN aislado de la cepa insertado en el locus. La cepa es prototrófica para adenina, uridina, histidina, prolina, arginina y triptófano. La cepa contiene en total de aproximadamente ocho a once copias de casetes de expresión de EPO. Durante el procesamiento de la EPO quimérica en el RE y en el aparato de Golgi, el péptido líder se retira. Por lo tanto, la rhEPO producida es la forma madura de la EPO.

Se hizo contraselección de la cepa **YGLY4209** en presencia de 5'-FOA para producir un número de cepas que eran auxotróficas para uracilo. A partir de los transformantes producidos, la cepa **YGLY4244** se seleccionó.

El plásmido pGLY2713 (**Figura 20**), un vector de integración que se dirige al gen *PEP4* de *P. pastoris* (SEQ ID NO: 104), contiene el ORF de *PNO1* de *P. pastoris* adyacente al casete de expresión que comprende el gen *URA5* de *P. pastoris* flanqueado por repeticiones de *lacZ* y flanqueado en un lado con la secuencia de nucleótidos de la región en la posición 5' del gen *PEP4* de *P. pastoris* y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *PEP4* de *P. pastoris*. El plásmido pGLY2713 se linealizó con *SfiI* y el plásmido linealizado se transformó en la cepa **YGLY4244** para producir una serie de cepas en las que el ORF de *PNO1* y el casete de expresión de *URA5* se han insertado en el locus de *PEP4* mediante recombinación homóloga de doble cruce. La cepa **YGLY5053** se seleccionó a partir de las cepas producidas y se hizo contraselección en presencia de 5-FOA para producir una serie de cepas en las que el *URA5* se ha perdido del genoma. La cepa **YGLY5597** se seleccionó a partir de las cepas producidas y es prototrófica para adenina, histidina, prolina, arginina y triptófano. La cepa es resistente a Zeocina y contiene de aproximadamente ocho a once copias del casete de expresión de rhEPO.

El plásmido pGLY3411 (pSH1092) (**Figura 15**) es un vector de integración que contiene el casete de expresión que comprende el gen *URA5* de *P. pastoris* flanqueado por repeticiones de *lacZ* flanqueado en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT4* (PpPBS4 5') de *P. pastoris* (SEQ ID NO: 72) y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT4* de *P. pastoris* (SEQ ID NO: 73). El plásmido pGLY3411 se linealizó y el plásmido linealizado se transformó en la cepa **YGLY5597** para producir una serie de cepas en las que el casete de expresión de *URA5* se ha insertado en el locus de *BMT4* mediante recombinación homóloga de doble cruce. La cepa **YGLY5618** se seleccionó a partir de las cepas producidas y es prototrófica para uracilo, adenina, histidina, prolina, arginina y triptófano. La cepa es resistente a Zeocina y Nourseotricina y contiene de aproximadamente ocho a once copias del casete de expresión de rhEPO. La cepa tiene interrupciones de los genes *BMT2* y *BMT4*.

El plásmido pGLY3430 (pSH1115) (**Figura 16**) es un vector de integración que contiene una molécula de ácido nucleico que codifica el ORF de resistencia a Nourseotricina (NATR) (originalmente de pAG25 de ERO-SCARF, Scientific Research and Development GmbH, Daimlerstrasse 13a, D-61352 Bad Homburg, Alemania, véase Goldstein *et al.*, Yeast 15: 1541 (1999)) ORF (SEQ ID NO: 102) unido de forma operativa al promotor *TEF1* de *Ashbya gossypii* y secuencias de terminación de *TEF1* de *Ashbya gossypii* flanqueado en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT1* de *P. pastoris* (SEQ ID NO: 68) y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT1* de *P. pastoris* (SEQ ID NO: 69). El plásmido pGLY3430 se linealizó y el plásmido linealizado se transformó en la cepa **YGLY5618** para producir una serie de cepas en las que el casete de expresión de NAT^R se ha insertado en el locus de *BMT1* mediante recombinación homóloga de doble cruce. La cepa **YGLY7110** se seleccionó a partir de las cepas producidas y es prototrófica para uracilo, adenina, histidina, prolina, arginina y triptófano. La cepa es resistente a Zeocina y Nourseotricina y contiene de aproximadamente ocho a once copias del casete de expresión de rhEPO. La cepa tiene interrupciones de los genes *BMT1*, *BMT2* y *BMT4*.

El plásmido pGLY4472 (pSH1186) (**Figura 17**) es un vector de integración que contiene una molécula de ácido nucleico que codifica el ORF del gen de higromicina B fosfotransferasa de *E. coli* (Hyg^R) (SEQ ID NO: 103) unido de forma operativa al promotor *TEF1* de *Ashbya gossypii* y secuencias de terminación de *TEF1* de *Ashbya gossypii* flanqueado en un lado con la secuencia de nucleótidos en la posición 5' del gen *BMT3* de *P. pastoris* (SEQ ID NO: 70) y en el otro lado con la secuencia de nucleótidos en la posición 3' del gen *BMT3* de *P. pastoris* (SEQ ID NO: 71). El plásmido pGLY3430 se linealizó y el plásmido linealizado se transformó en la cepa **YGLY7110** para producir una serie de cepas en las que el casete de expresión de Hyg^R se ha insertado en el locus de *BMT3* mediante recombinación homóloga de doble cruce. Las cepas **YGLY7113** a **YGLY7122** se seleccionaron a partir de las cepas producidas y son prototróficas para uracilo, adenina, histidina, prolina, arginina y triptófano. Las cepas son resistentes a Zeocina, Nourseotricina, e Higromicina y contienen de aproximadamente ocho a once copias del casete de expresión de EPO. Las cepas tienen interrupciones de los genes *BMT1*, *BMT2*, *BMT3* y *BMT4* y producen rhEPO que carece de unión de reactividad cruzada detectable a anticuerpos preparados frente a HCA.

EJEMPLO 4

Varias de las cepas de los Ejemplos 1 a 3 se usaron para producir rhEPO como se describe a continuación y se muestran esquemáticamente en la **Figura 21**. En resumen, la producción comienza mediante inoculación de
 5 matraces de agitación que contienen medios de cultivo con células del banco de células de trabajo y evoluciona a través de una serie de inoculaciones, incurvaciones y transferencias de los cultivos en expansión en recipientes de tamaño creciente hasta que hay suficiente biomasa disponible para inocular el biorreactor de producción. El glicerol es la fuente primaria de carbono durante la fase discontinua, a continuación el crecimiento del cultivo se mantiene a través de alimentación de glicerol y sales. Cuando el glicerol se agota, las células se inducen para que expresen la
 10 proteína rhEPO mediante intercambio a una alimentación de metanol. Los inhibidores se añaden a inducción para minimizar la O-glicosilación (por ejemplo, PMTi 3, Ácido 5-[[3-(1-feniletoksi)-4-(2-feniletoksi)]fenil]metilen]-4-oxo-2-tioxo-3-tiazolidinaacético, (véase la Solicitud de PCT Publicada n.º WO 2007061631)) y para minimizar la proteólisis. Los inhibidores de la proteólisis se añaden de nuevo al final de la fase para minimizar la proteólisis. El cultivo se enfría a aproximadamente 4 °C y se cosecha.

El cultivo de las cepas a escala de laboratorio se realizó en fermentadores SixFors de 500 ml y de 3 l usando en general los siguientes procedimientos. Se realizan Identificaciones Sistemáticas en Biorreactor (SIXFORS) en recipientes de 0,5 l (sistema de multif fermentación Sixfors, ATR Biotech Laurel, MD) en las condiciones siguientes: pH a 6,5, 24 °C, 0,3 SLPM y una velocidad inicial del agitador de 550 rpm con aun volumen de trabajo inicial de
 20 350 ml (330 ml de medio BMGY y 20 ml de inóculo). El software del multif fermentador IRIS (ATR Biotech Laurel, MD) se usa para aumentar de forma lineal la velocidad del agitador de 550 rpm a 1200 rpm durante 10 horas, una hora después de la inoculación. Los cultivos de semilla (200 ml de BMGY en un matraz de 1 l con tabique deflector) se inoculan directamente de placas de agar. Los Matraces de semilla se incubaron durante 72 horas a 24 °C para alcanzar densidades ópticas (DO_{600}) entre 95 y 100. Los fermentadores se inoculan con cultivos en matraz en fase
 25 estacionaria de 200 ml que se concentraron hasta 20 ml mediante centrifugación. La fase discontinua terminó al finalizar la fermentación inicial de carga de glicerol (18-24 h) y va seguida de una segunda fase discontinua que comienza mediante la adición de 17 ml de solución de alimentación de glicerol (50 % [p/p] de glicerol, 5 mg/l de Biotina, 12,5 ml/l de sales de PMTi (65 g/l de $FeSO_4 \cdot 7H_2O$, 20 g/l de $ZnCl_2$, 9 g/l de H_2SO_4 , 6 g/l de $CuSO_4 \cdot 5H_2O$, 5 g/l de H_2SO_4 , 3 g/l de $MnSO_4 \cdot 7H_2O$, 500 mg/l de $CoCl_2 \cdot 6H_2O$, 200 mg/l de $NaMoO_4 \cdot 2H_2O$, 200 mg/l de biotina, 80 mg/l de NaI, 20 mg/l de H_3BO_4)). Después de finalizar la segunda fase discontinua, tal como señalizar con una
 30 adición en oxígeno disuelto, la fase de inducción comienza por alimentación de una solución de alimentación de metanol (100 % de MeOH, 5 mg/l de biotina, 12,5 ml/l de PMTi) a 0,6 queg/h durante 32-40 horas. El cultivo se cosecha mediante centrifugación.

Los cultivos en biorreactor (3 l) se realizan en biorreactores de vidrio de 3 l (Applikon, Foster City, CA) y 15 l (Applikon, Foster City, CA) y un autoclave de acero inoxidable de 40 l (Applikon, Foster City, CA) en lugar del biorreactor. Los cultivos de semilla se preparan por inoculación de medios BMGY directamente con viales de reserva congelados a una proporción volumétrica de un 1 %. Los matraces de semilla se incuban a 24 °C durante 48 horas para obtener una densidad óptica (DO_{600}) de 20 ± 5 para asegurar que las células están creciendo de manera
 40 exponencial después de la transferencia. El medio de cultivo contenía 40 g de glicerol, 18,2 g de sorbitol, 2,3 g de K_2HPO_4 , 11,9 g de KH_2PO_4 , 10 g de extracto de levadura (BD, Franklin Lakes, NJ), 20 g de peptona (BD, Franklin Lakes, NJ), 4×10^{-3} g de biotina y 13,4 g de Base Nitrogenada de Levadura (BD, Franklin Lakes, NJ) por litro. El biorreactor se inocula con una proporción volumétrica de un 10 % de semilla con respecto al medio inicial. Los cultivos se realizan en modo semicontinuo en las condiciones siguientes: temperatura ajustada a $24 \pm 0,5$ °C, pH controlado a $6,5 \pm 0,1$ con NH_4OH , el oxígeno disuelto se mantuvo a $1,7 \pm 0,1$ mg/l con una tasa de agitación en cascada en la adición de O_2 . El caudal de aire se mantiene a 0,7 wm. Después de agotar la carga inicial de glicerol (40 g/l), una solución de glicerol al 50 % que contiene 12,5 ml/l de sales de PTM1 se alimenta de forma exponencial a un 50 % de la tasa de crecimiento máximo durante ocho horas hasta que se consiguen 250 g/l de peso celular húmedo. La inducción comienza después de una fase de privación de 30 minutos cuando el metanol se alimento de
 50 forma exponencial para mantener una tasa de crecimiento específico de $0,01 h^{-1}$. Cuando se alcanza una tasa de absorción de oxígeno de 150 mM/l/h, la tasa de alimentación de metanol se mantiene constante para evitar la limitación de oxígeno. El cultivo se cosecha mediante centrifugación.

Después del aclarado por centrifugación y microfiltración, el filtrado se concentra 10X mediante ultrafiltración y la proteína rhEPO se purifica a través de una secuencia de dos etapas de cromatografía usando una afinidad por colorante azul e hidroxipatita.

El aclarado principal se realiza por centrifugación. El de cultivo de células completas se transfiere en frascos de centrifugación de 1000 ml y se centrifuga a 4 °C durante 15 minutos a $13.000 \times g$. Una etapa de ultrafiltración se puede usar para fermentadores más grandes (de 10 l a 40 l y más grandes). Esta etapa se puede realizar usando láminas planas de Sartorius con un tamaño de poro de 10 K hasta una concentración de cinco veces.

Una etapa de captura se realiza usando Cromatografía de Flujo Rápido de Azul SEPHAROSE 6 (Pseudo-Afinidad). Una columna de de Flujo Rápido (FF) de Azul SEPHAROSE 6 (GE Healthcare) se equilibra con MOPS 50 mM, pH 7,0. El sobrenadante del cultivo se ajusta a NaCl 100 mM y se pasa a través de un filtro sin salida (Whatman, Polycap TC) antes de cargarlo a la columna. El tiempo de permanencia se mantiene durante aproximadamente 10

minutos con un lavado con 3 volúmenes de columna (CV) después de la carga. La elución es la etapa de elución de 4 CV con NaCl 1 M en MOPS 50 mM, pH 7,0. La EPO eluye al NaCl 1 M.

Una etapa intermedia se realiza usando cromatografía de hidroxapatita (HA). Una hidroxapatita de cerámica Macro-prep de Tipo I de 40 µm (Bio-Rad) se usa después de la etapa de captura. Esta columna se equilibra con solución de equilibrio: MOPS 50 mM, pH 7,0 que contiene NaCl 1 M y CaCl₂ 10 mM. Se añade CaCl₂ aproximadamente 10 mM a la rhEPO combinada de la columna de azul antes de la carga. El lavado de la columna se realiza con 3 CV de solución de equilibrio seguido de etapa de elución de 10 CV a fosfato Na 12,5 mM en MOPS, pH 7,0 para proporcionar la combinación 1 de HA que contiene la rhEPO.

Una etapa de cromatografía de intercambio catiónico se puede usar para purificar adicionalmente la rhEPO. La muestra combinada después de la etapa de cromatografía de hidroxapatita (por ejemplo, combinación 1 de HA) se dializa frente a acetato Na 50 mM, pH 5,0 durante una noche a 4 °C y una columna Source 30S o columna de intercambio catiónico Poros (GE Healthcare) se equilibra con el mismo tampón. La muestra dializada se aplica a la columna y un gradiente lineal de 10 CV de NaCl de 0 a 750 mM se aplica con rhEPO eluyendo con NaCl entre 350 mM y 500 mM para proporcionar la rhEPO.

El extremo N terminal de la molécula de rhEPO purificada se puede conjugar con polietilenglicol lineal de 40 kDa (PEG) a través de aminación reductora (PEGilación). El PEG activado se añade a la muestra de rhEPO (conc. a aproximadamente 1 mg/ml) en tampón de acetato sódico 50 mM a pH 5,2 a una proporción de proteína:PEG de 1:10. La reacción se realiza a temperatura ambiente en condiciones reductoras mediante la adición de cianoborohidruro sódico 10 mM a la mezcla de reacción con agitación durante una noche. La reacción se detiene mediante la adición de Tris 10 mM, pH 6,0.

El producto de rhEPO mono-PEGilado se purifica usando una etapa de cromatografía de intercambio catiónico antes de diafiltración en el tampón de formulación final (fosfato sódico 20 mM, cloruro sódico 120 mM, Polisorbato 20 al 0,005 % (p/v), pH 7,0).

El producto final se diluye hasta una concentración adecuada para su carga y se filtra de forma estéril en el envase de almacenamiento de la sustancia farmacológica. La rhEPO PEGilada se puede almacenar a 2-8 °C hasta su carga, momento en el que se carga de forma aséptica en viales de vidrio que a continuación se cierran herméticamente con un tapón de caucho y protección de aluminio.

Se conocen y se pueden usar formulaciones comerciales de proteínas. Los ejemplos incluyen, pero no se limitan a, ARANESP®: solución de Polisorbato: Cada 1 ml contiene 0,05 mg de polisorbato 80 y se formula a pH 6,2 ± 0,2 con 2,12 mg de fosfato sódico monohidratado monobásico, 0,66 mg de fosfato sódico dibásico anhidro y 8,18 mg de cloruro sódico en agua para inyección, USP (para 1 ml). Solución de albúmina: Cada 1 ml contiene 2,5 mg de albúmina (humana) y se formula a pH 6,0 ± 0,3 con 2,23 mg de fosfato sódico monohidratado monobásico, 0,53 mg de fosfato sódico dibásico anhidro y 8,18 mg de cloruro sódico en agua para inyección, USP (hasta 1 ml). EPOGEN® se formula como un líquido incoloro, estéril en una solución isotónica tamponada con cloruro sódico/citrato sódico o una solución tamponada de cloruro sódico/fosfato sódico para administración intravenosa (IV) o subcutánea (SC). Vial sin Conservantes, de Una Sola dosis: Cada 1 ml de solución contiene 2000, 3000, 4000 o 10.000 Unidades de Epoetina alfa, 2,5 mg de Albúmina (Humana), 5,8 mg de citrato sódico, 5,8 mg de cloruro sódico y 0,06 mg de ácido cítrico en agua para inyección, USP (pH 6,9 ± 0,3). Esta formulación no contiene conservantes. Los viales conservados contienen un 1 % de alcohol bencílico.

EJEMPLO 5

Los métodos usados para analizar la presencia o ausencia de antígeno de célula hospedadora (HCA) incluían análisis de transferencia de Western y ensayo de inmunoabsorción unida a enzima de sándwich (ELISA).

El anticuerpo de antígeno de célula hospedadora (HCA) se preparó en conejos usando el sobrenadante de cultivos de la cepa NORF. La cepa NORF es genéticamente la misma que YGLY3159 excepto porque carece del ORF que codifica la EPO madura humana. El sobrenadante de fermentación de la cepa NORF preparado en adyuvante completo de Freund se inyectó en conejos, que a continuación se estimularon tres veces con sobrenadante de fermentación preparado en adyuvante incompleto de Freund. Después de 45 días, se extrajo sangre de los conejos y los anticuerpos policlonales para HCA se prepararon usando métodos convencionales, por ejemplo, 9161 F072208-S de IgG policlonal de conejo, que era Proteína A de SLr purificada y suero de conejo policlonal GiF2::suero de conejo completo 6316. El anticuerpo GiF2 no era la proteína A purificada.

Las transferencias de Western para detectar de HCA de *P. pastoris* se realizaron como sigue a continuación. Las muestras que contenían rhEPO PEGilada o no PEGilada purificada se redujeron en tampón de carga de muestra, del que a continuación se aplicó 1 µl a los pocillos de geles de poliacrilamida SDS Tris-HCl al 4-20 % (SDS-PAGE al 4-20 %) (Bio RAD) y se sometió a electroforesis a 150 V durante aproximadamente 60 minutos. Las proteínas resueltas se electrotransfirieron a membranas de nitrocelulosa a 100 V durante aproximadamente 60 minutos. Después de la transferencia, las membranas se bloquearon durante una hora con solución de bloqueo al 1 % (Roche

Diagnostics). Después del bloqueo, las membranas se sondearon con el anticuerpo policlonal anti-HCA de conejo (anticuerpo primario) diluido a 1:3000. A continuación, las membranas se lavaron y la detección del anticuerpo anti-HCA de conejo se realizó con el anticuerpo secundario, IgG de cabra-anti-conejo (H+L) (Pierce n.º 31460, Lote n.º H51015156), conjugado con peroxidasa de rábano picante (HRP), a una dilución de 1:5000. Después de lavar las membranas, la detección del anticuerpo secundario unido se realizó usando 3,3' Diaminobencidina (DAB). Para detectar la proteína EPO, el anticuerpo primario era anticuerpo conjugado con EPO (B-4) HRP usado a una dilución de 1:1000 (SC5290 Lote n.º A0507, Santa Cruz Biotechnology). No se usó un anticuerpo secundario. De forma rutinaria, las muestras de EPO se sometieron a electroforesis en paralelo con muestras de rhEPO que se habían desglicosilado con tratamiento con PNGasaF. La desglicosilación se realizó con muestras de 50 µl a las que se añadió 1 µl de enzima PNGasaF a 500 unidades/µl. Después de incubación a 37 °C durante dos horas, las muestras se redujeron en tampón de carga de muestra y se retiraron alícuotas de 1 µl y se aplicaron a los geles de SDS como se ha mencionado anteriormente.

Los ELISA de sándwich para detectar HCA de *P. pastoris* se realizaron como sigue a continuación. Los pocillos de placas para ELISA de 96 pocillos se revistieron con 1 µg/pocillo de anticuerpo monoclonal anti-hEPO de ratón. A continuación, los pocillos se bloquearon durante 30 minutos con solución salina tamponada con fosfato (PBS). A los pocillos se añadieron aproximadamente 100 µl de muestras que contenían rhEPO no PEGilada purificada concentrada a aproximadamente 200 ng/ml. La detección primaria usó el anticuerpo policlonal anti-HCA de conejo a una dilución de partida de 1:800 en PBS que a continuación se diluyó en serie a 1:1 en PBS a través a de la fila de terminación con el 11º pocillo a una dilución de 1:819.200. El 12º pocillo sirvió como un control negativo. El patrón para el ELISA fue rhEPO purificada de **YGLY3159**. Después de 60 minutos, los pocillos se lavaron con PBS tres veces. La detección del anticuerpo anti-HCA de conejo usó anticuerpo anti-conejo de cabra conjugado con fosfatasa alcalina (AP) a una dilución de 1:10.000 en PBS. Después de 60 minutos, los pocillos se lavaron tres veces con PBS y la detección del anticuerpo secundario unido usó fosfato de 4-Metilumbeliferilo (4-MUPS). Las placas para ELISA se leyeron usando un lector de microplacas de múltiple detección Tecan Genios a una longitud de onda de excitación de 340 nm y una longitud de onda de emisión de 465 nm.

EJEMPLO 6

Este ejemplo muestra que **YGLY3159** produce rhEPO con actividad de unión cruzada (CBA) con anticuerpo anti-HCA y que la actividad de unión cruzada se debía a la presencia de restos de β -1,2-manosa (resistente a α -1,2-manosidasa) en al menos una parte de los *N*-glicanos en la rhEPO incluso cuando la rhEPO se había producido en la cepa en la que el gen *BMT2* de la β -1,2-manosiltransferasa había experimentado delección o interrupción.

La rhEPO se recuperó con una separación cromatográfica de tres etapas del sobrenadante de fermentación de producción de *P. pastoris* modificada por glico-ingeniería. La cepa **YGLY3159** presentaba una pureza de proteínas de aproximadamente un 95 % tal como se determina con SDS-PAGE, RP-HPLC y SEC-HPLC. La rhEPO mono-PEGilada se separó mediante una etapa de cromatografía de intercambio catiónico a partir de sus conjugados hiper y no PEGilados con una pureza de aproximadamente un 96 % tal como se determina con gel de SDS-PAGE. Sin embargo, el anticuerpo frente a HCA de la cepa **YGLY3159** detectaba una glicoproteína en las preparaciones de rhEPO producidas a partir de la cepa que comigraba con rhEPO en las transferencias de Western. La **Figura 22** muestra que el anticuerpo anti-HCA identificaba una proteína que comigra con rhEPO en geles SDS-PAGE al 4-20 %. La retirada de ácido siálico de rhEPO no anulaba la actividad de unión cruzada; sin embargo, la retirada de todo el *N*-glicano de la rhEPO usando PNGasa F producía una forma desglicosilada de rhEPO que no era detectable en las transferencias de Western sondeadas con anticuerpo anti-HCA. Esto se muestra en la **Figura 23** que muestra que solamente la forma desglicosilada de rhEPO carecía de actividad de unión cruzada con el anticuerpo anti-HCA.

Para determinar si la actividad de unión cruzada era específica de rhEPO o se podía identificar en preparaciones de glicoproteína de otras cepas de *P. pastoris* recombinantes, se aislaron unas glicoproteínas producidas en otras cepas, se resolvieron con geles de SDS-PAGE al 4-20 % y los geles se transfirieron a membranas de nitrocelulosa. En el caso de un anticuerpo completo humano recombinante (rhIgG) producido en una *P. pastoris* recombinante, la actividad de unión cruzada se detectó en preparaciones de proteínas producidas en *P. pastoris* de tipo silvestre (hiperanosilada a partir de la región tanto *N* como *O*-glicosilada) y en una cepa GS2.0 recombinante que produce predominantemente *N*-glicanos de Man₅GlcNAc₂ pero que también contenía *N*-glicanos de Man₉GlcNAc₂ detectables que eran resistentes a la α -1,2-manosidasa (**Figura 24**, flecha). Sin embargo, las preparaciones de rhIgG de *P. pastoris* de tipo silvestre contenían actividad de unión cruzada con un peso molecular aparente superior al de rhIgG lo que sugiere que las preparaciones contenían contaminación con glicoproteínas de célula hospedadora. La actividad de unión cruzada no se eliminó mediante digestión con PNGasa F (rodeado con un círculo en la **Figura 24**).

La **Figura 25** muestra que la rhEPO glicosilada producida en **YGLY3159** tenía actividad de unión cruzada al anticuerpo anti-HCA pero que el panel de la parte superior muestra rpo anti-HCA. Las fetuínas son glicoproteínas de sangre altamente glicosiladas que se forman en el hígado y se secretan en el torrente sanguíneo. Pertenecen a un gran grupo de proteínas de unión que median el transporte y la disponibilidad de una gran diversidad de sustancias de carga en el torrente sanguíneo. El ejemplo representativo mejor conocido de estas proteínas vehículo es la albúmina de suero, la proteína más abundante en el plasma sanguíneo de animales adultos. La fetuína es más

abundante en la sangre fetal, de ahí el nombre de "fetuína" (del latín, feto). El suero bovino fetal contiene más fetuína que la albúmina mientras que el suero de adulto contiene más albúmina que fetuína. Las asialofetuínas son fetuínas en las que el ácido siálico terminal de los *N*- y *O*-glicanos se retira por hidrólisis suave o tratamiento con neuraminidasa. En la actualidad, no hay informes de manosas unidas en β en *S cerevisiae*. La HSA no es una

Los datos a escala de laboratorio Lab demostraban que la purificación en la etapa intermedia de cromatografía de rhEPO de la combinación de captura de Azul SEPHAROSE 6 FF usando una resina de 40 μ m de tipo I de hidroxiapatita (HA) puede separar rhEPO que tiene una actividad de unión cruzada casi indetectable (combinación 1 de HA) de rhEPO que presentaba *N*-glicanos de tipo alta manosa (combinaciones 2 y 3 de HA). La combinación 1 de HA contenía aproximadamente un 90,40 % de *N*-glicanos bisialilados (la forma de *N*-glicano deseada) y menos de un 3,5 % de *N*-glicanos neutros. Por el contrario, la elución en gradiente lineal de de fosfato sódico de 0 a 100 mM mostraba que las últimas fracciones de elución (combinaciones 2 y 3 de HA) contenían *N*-glicanos de tipo alta manosa y aumentaban la actividad de unión cruzada a anticuerpo anti-HCA en transferencias de Western. Thesto se puede observar en el análisis de *N*-glicano por HPLC y transferencias de Western de geles de SDS-PAGE al 4-20 % mostrados en la Figura 26.

También se sometió a ensayo la cromatografía en columna usando las resinas aniónicas Q SEPHAROSE FF o Source 30Q. Las combinaciones 1-3 de HA se combinaron y se dializaron frente a acetato de Na 50 mM, pH 5,0 durante una noche a 4 °C. La muestra dializada se aplicó a la columna y se aplicó un gradiente lineal de 10 CV de NaCl de 0 a 750 mM con rhEPO eluyendo con NaCl entre 350 y 500 mM para proporcionar la rhEPO. La **Figura 27A** muestra un ejemplo de una purificación con Q SEPHAROSE FF de rhEPO. Los datos mostraban que todos los glicanos de tipo alta manosa ($\text{Man}_{6,7,8,9} > 9$, en su mayoría resistente a α 1,2 manosidasa) que muestran una actividad de unión cruzada correspondiente más elevada no se unían a las resinas de intercambio aniónico cuando el material unido y sin unir se analizaba en un ELISA de sándwich (**Figura 27B**). La Tabla 1 muestra los resultados del análisis de HPLC del contenido de *N*-glicano de la rhEPO en la fracción unida (combinación 1 de Q SEPHAROSE FF) y la fracción de flujo continuo (Flujo Continuo de Q SEPHAROSE FF). La Tabla 2 muestra el contenido de *N*-glicano de los *N*-glicanos neutros mostrados en la Tabla 1.

Tabla 1			
Q Sepharose FF - Análisis por HPLC de Purificación de <i>N</i> -Glicano de rhEPO			
Muestra	% Neutro	% Mono Sialilado	% Bi Sialilado
Entrada (combinaciones de HA)	11,04	13,66	75,30
Combinación de 1 Q SEPHAROSE FF	3,01	6,47	90,52
Flujo Continuo de Q SEPHAROSE FF	26,71	21,19	52,10

Tabla 2				
Q Sepharose FF - Purificación de rhEPO % de Perfil de <i>N</i> -Glicano Neutro				
Muestra	% G2	% Man_5	% Man_{6-8}	% Man_9
Entrada (combinaciones de HA)	3,1	2,8	2,4	2,74
Combinación de 1 Q SEPHAROSE FF	3,01	ND	ND	ND
Flujo Continuo de Q SEPHAROSE FF	4,2	8,0	3,9	10,61
ND - no detectado G2 - estructura de <i>N</i> -glicano es $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$				

Las figuras y las tablas muestran que la rhEPO con actividad de unión cruzada indetectable a anticuerpos anti-HCA y buena calidad de proteína y glicano por lo tanto se puede unir/eluir de resinas de intercambio aniónico. Estos datos también sugieren que la familia de genes fúngicos implicados en la biosíntesis de oligomanósidos unidos en β -1,2 (*BMT1*, *BMT2*, *BMT3*, *BMT4*) era responsable de las impurezas de unión cruzada de bajo nivel en las preparaciones de rhEPO.

Por lo tanto, cuando se visualizan como un conjunto, los resultados sugieren que la actividad de unión cruzada a anticuerpos anti-HCA no era específica para rhEPO pero se debía a *N*-glicanos resistentes a α -1,2-manosidasa en

las glicoproteínas. **YGLY3159** se había generado por supresión genética de cinco genes de glicosilación endógenos y por introducción de 15 genes heterólogos. **YGLY3159** es la cepa de supresión genética de *bmt2Δ*. Los estudios de espectroscopía de RMN sugieren que las cepas con supresión genética de *bmt2Δ* pueden producir glicoproteínas con cantidades variables de residual *N*-glicanos de β-1,2-manosa. Dado que **YGLY3159** es *bmt2Δ*, se postuló que **BMT1** y **BMT3** eran responsables del bajo nivel residual de transferencia de β-1,2-manosa en los *N*-glicanos de núcleo.

Aunque una combinación de etapas de cromatografía para purificar la rhEPO puede producir preparaciones de rhEPO libres de actividad de unión cruzada detectable a anticuerpos anti-HCA, en particular podría ser deseable modificar genéticamente las cepas hospedadoras de *P. pastoris* para reducir o eliminar la actividad de unión cruzada detectable a anticuerpos anti-HCA en las cepas. Esto reduce al mínimo el riesgo de posible contaminación de las preparaciones de rhEPO con actividad de unión cruzada debido a la variabilidad durante la purificación. Además, dado que cada etapa de purificación puede dar como resultado una pérdida de rhEPO, las cepas de *P. pastoris* modificadas genéticamente pueden reducir el número de etapas de purificación y por lo tanto reducir la cantidad de rhEPO perdida durante las etapas eliminadas. Por lo tanto, se hizo una delección o interrupción en serie de la expresión de los cuatro genes *BMT* para identificar cepas que no produjeran actividad de unión cruzada detectable a anticuerpos anti-HCA.

EJEMPLO 7

Para reducir la presencia de *N*-glicanos de tipo manosa unidos en β a niveles indetectables, los genes *BMT1* y *BMT4* se interrumpieron y la rhEPO se analizó para la presencia de *N*-glicanos resistentes a α-1,2-manosidasa.

Las cepas **YGLY6661** e **YGLY7013** se construyeron como se describe en el Ejemplo 2 y se analizaron para la presencia de *N*-glicanos resistentes a α-1,2-manosidasa usando anticuerpos anti-HCA. La cepa **YGLY7013** era *bmt2Δ* y *bmt4Δ* y la cepa **YGLY6661** era *bmt2Δ*, *bmt4Δ* y *bmt1Δ*. La rhEPO producida a partir de las cepas se sometieron a cromatografía de Azul SEPHAROSE 6 FF y las alícuotas de la combinación de captura con Azul SEPHAROSE 6 FF se trataron con PNGasa F vel non. Las alícuotas tratadas y sin tratar se sometieron a electroforesis en SDS-PAGE, los geles se transfirieron a membranas de nitrocelulosa y las membranas se sondearon con anticuerpo anti-EPO o anticuerpos anti-HCA. La **Figura 28** muestra que, en transferencias de Western de geles de SDS-PAGE al 4-20 % de alícuotas de combinaciones de captura de Azul SEPHAROSE 6 FF que producían rhEPO en cualquier cepa, aún producían *N*-glicanos resistentes a α-1,2-manosidasa que reaccionaban de forma cruzada con anticuerpos anti-HCA. La Tablas 3 y 4 muestran la distribución de especies de *N*-glicano en rhEPO en combinaciones de captura de Azul Sepharose 6 FF tanto de cultivos de fermentación como de SixFors. Como se muestra en las tablas, ambas cepas producían una cantidad sustancial de *N*-glicanos neutros de los cuales, 1 parte de la resistencia a la digestión de α1,2-mannosidasa *in vitro*.

Tabla 3							
Semana 44 - rhEPO - Combinación de Captura de Azul SEPHAROSE 6 FF - Fermentación							
Combinaciones	% BiSialilado	% Mono Sialilado	% Neutro	% Neutro			
				% G2	% M5	% M6-M8	% M9+
F074411	52,98	34,08	12,94	1,7	2,63	3,65	4,96
(YGLY6661)							
F074411 (YGLY6661) α1,2 Manosidasa	53,42	34,32	12,26	1,9	5,05	3,4	1,91
F074410 (YGLY7013)	25,10	47,00	27,90	12,99	2,22	5,67	7,02
F074410 (YGLY7013) α1,2 Manosidasa	26,34	49,03	26,34	13,14	5,39	4,68	1,42
G2 - <i>N</i> -glicanos de Gal ₂ GlcNAc ₂ Man ₃ GlcNAc ₂ M6-M8 - <i>N</i> -glicanos de Man ₆ GlcNAc ₂ a Man ₈ GlcNAc ₂ M9+ - <i>N</i> -glicanos de Man ₉ GlcNAc ₂ y de mayor tamaño							

Tabla 4								
Semana 41 - rhEPO - Combinación de Captura de Azul SEPHAROSE 6 FF - SixFors								
Combinaciones	% BiSialilado	% Mono Sialilado	% Neutro	% Neutro				
				% G2	% M5	% M6-M8	% M9	% M9+
X074128 (YGLY6661)	43,49	39,24	17,27	1,7	7,8	6,69	0,45	0,63
X074128 (YGLY6661) α1,2 Manosidasa	42,52	39,26	18,22	1,2	11,84	5,02	0,1	0,06
X074131 (YGLY7013)	66,90	18,83	14,27	1,84	8,36	2,82	0,66	0,59
X074131 (YGLY7013) α1,2 Manosidasa	64,81	19,70	15,49	1,06	13,1	0,77	0,56	0
G2 - N-glicanos de Gal ₂ GlcNAc ₂ Man ₃ GlcNAc ₂ M6-M8 - N-glicanos de Man ₆ GlcNAc ₂ a Man ₈ GlcNAc ₂ M9 - N-glicanos de Man ₉ GlcNAc ₂ M9+ - N-glicanos de Man ₉ GlcNAc ₂ y de tamaño superior								

Un ELISA de sándwich de rhEPO en las combinaciones de captura de Azul SEPHAROSE 6 FF preparadas a partir de ambas cepas en comparación con YGLY3159 mostraba que ambas cepas tenían actividad de unión cruzada a anticuerpo anti-HCA (**Figura 29**). La purificación adicional de la rhEPO mediante cromatografía de hidroxipatita (HA) y el análisis de las muestras mediante ELISA de sándwich mostraba que la combinación 1 de HA que contenía rhEPO producida a partir de YGLY6661 (*bmt2Δ*, *bmt4Δ* y *bmt1Δ*) parecía carecer de actividad de unión cruzada detectable a anticuerpo anti-HCA pero que la rhEPO producida en YGLY7013 (*bmt2Δ* y *bmt4Δ*) todavía tenía actividad de unión cruzada detectable a anticuerpo anti-HCA (**Figura 30**). Los resultados sugerían que la delección de los genes *BMT2* y *BMT1* no era suficiente para retirar toda la actividad de unión cruzada detectable. Los resultados también muestran que la cromatografía de hidroxipatita puede eliminar la actividad de unión cruzada detectable en la combinación 1 de HA. La **Figura 31** es una transferencia de Western de geles de SDS-PAGE al 4-20 % que muestran que rhEPO en otra combinación de captura de Azul SEPHAROSE 6 FF preparada a partir de la cepa YGLY6661 continuaba teniendo actividad de unión cruzada a anticuerpo anti-HCA y que la actividad y unión cruzada aún se podría hacer indetectable mediante desglucosilación de la rhEPO. El resultado indicaba que para producir rhEPO que no tenía actividad de unión cruzada a anticuerpos anti-HCA detectable, era necesario que la expresión del gen *BMT3* se anulara mediante interrupción o delección.

EJEMPLO 8

Para conseguir de forma más eficaz la eliminación de los glicanos de tipo manosa unidos en β detectable, se interrumpieron los cuatro genes *BMT* implicados en la ruta de manosiltransferasa. Las cepas YGLY7361-7366 e YGLY7393-7398 (Ejemplo 2) se evaluaron para su capacidad para producir rhEPO carente de actividad de unión cruzada detectable a anticuerpo anti-HCA.

Diversas cepas de YGLY7361-7366 e YGLY7393-7398 en las que se interrumpieron los cuatro genes *BMT* implicados en la ruta de β-manosiltransferasa se cultivaron en fermentadores SixFors de 500 ml y a continuación se procesaron para rhEPO a través de combinaciones de Azul SEPHAROSE 6 FF (combinaciones de Azul). Las alícuotas de las combinaciones de Azul se analizaron por SDS-PAGE al 4-20 %. La **Figura 32** muestra geles de SDS-PAGE al 4-20 % teñidos con azul de Commassie de las combinaciones de Azul de las diversas cepas con y sin tratamiento con PNGasa F. Los geles muestran que todas de las cuatro cepas sometidas a ensayo producían rhEPO glicosilada. Varias de las cepas se evaluaron para actividad de unión cruzada a anticuerpo anti-HCA mediante ELISA de sándwich. La **Figura 33** muestra que la mayoría de las cepas carecían de actividad de unión cruzada detectable a anticuerpo anti-HCA. Sin embargo, las cepas YGLY7363 e YGLY7365 tenían actividad de unión cruzada detectable a anticuerpo anti-HCA. La reconfirmación de YGLY7365 por PCR indicaba que esta cepa no tenía una supresión genética completa del gen *BMT3*, lo que explica la unión relativamente elevada observada con el anticuerpo anti-HCA presente en el ELISA (**Figura 33**). El análisis de N-glicano por HPLC de las cepas YGLY7361-7366 se muestra en la Tabla 5 y las cepas YGLY7393-7398 se muestran en la Tabla 6. Los datos en las tablas se presentan gráficamente en la **Figura 34**.

Tabla 5							
Semana 46a - SixFors - cepas <i>Δbmt1-4</i> - combinaciones de Azul							
Combinaciones	% BiSialilado	% Mono α1,2 Manosidasa	% α1,2 Manosidasa	% Neutro			
				% G2	% M5	% M6-M8	% M9+
X074613 (YGLY7361)	22,10	47,83	30,07	13,85	2,53	6,77	6,92
X074613 (YGLY7361) α1,2 Manosidasa	21,55	48,18	30,27	13,68	4,53	5,47	6,59
X074614 (YGLY7362)	67,36	24,69	7,95	0,6	4,36	2,8	0,19
X074614 (YGLY7362) α1,2 Manosidasa	66,21	25,25	8,54	1,1	6,7	0,68	0,06
*X074615 (YGLY7363)	49,40	39,17	11,43	0,8	4,42	5,91	0,3
X074615 (YGLY7363) α1,2 Manosidasa	48,52	39,20	12,28	0,4	7,2	4,68	ND
X074616 (YGLY7366)	55,99	33,85	10,16	0,8	3,73	4,94	0,69
X074616 (YGLY7366) α1,2 Manosidasa	55,44	34,24	10,32	1,9	7,2	1,02	0,2
*X074617 (YGLY7365)	43,22	42,10	14,68	5,37	5,88	3,03	0,4
X074617 (YGLY7365) α1,2 Manosidasa	42,70	42,40	14,90	4,5	8,4	2,0	ND
X074618 (YGLY7364)	48,18	38,44	13,38	0,7	6,56	5,76	0,36
X074618 (YGLY7364) α1,2 Manosidasa	47,52	39,75	12,73	0,4	6,74	5,09	0,5
G2 - N-glicanos de Gal ₂ GlcNAc ₂ Man ₃ GlcNAc ₂ M6-M8 - N-glicanos de Man ₆ GlcNAc ₂ a Man ₈ GlcNAc ₂ M9+ - N-glicanos de Man ₉ GlcNAc ₂ y de tamaño superior							
*Presentaban actividad de unión cruzada anticuerpo anti-HCA							

Tabla 6								
Semana 46a - SixFors - cepas <i>Δbmt1-4</i> - combinaciones de Azul								
Combinaciones	% BiSialilado	% Mono α1,2 Manosidasa	% α1,2 Manosidasa	% Neutro				
				% G2	% M5	% M6-M8	% M9	% Hib
X074637 (YGLY7393)	51,04	35,45	13,51	2,3	6,4	1,41	1,2	2,2
X074637 (YGLY7393) α1,2 Manosidasa	50,33	35,44	14,23	2,5	9,14	ND	ND	2,54
X074638 (YGLY7394)	63,56	25,65	10,79	1,1	6,6	1,9	0,4	0,79
X074638 (YGLY7396) α1,2 Manosidasa	62,88	25,75	11,37	1,2	8,96	ND	ND	1,21

X074639 (YGLY7395)	56,05	31,43	12,52	1,9	5,1	2,2	1,9	1,4
X074639 (YGLY7395) α 1,2 Manosidasa	56,27	31,81	11,92	1,9	8,43	ND	ND	1,59
X074640 (YGLY7396)	50,42	36,71	12,87	3,2	6,7	1,27	0,3	1,4
X074640 (YGLY7396) α 1,2 Manosidasa	49,94	36,86	13,20	3,2	8,12	ND	ND	1,88
X074641 (YGLY7397)	49,32	36,07	14,61	2,6	7,0	2,4	0,5	2,11
X074641 (YGLY7398) α 1,2 Manosidasa	48,72	35,86	15,42	2,7	10,24	ND	ND	2,48
X074642 (YGLY7398)	65,74	22,61	11,65	0,8	7,7	1,97	0,43	3,71
X074642 (YGLY7398)	64,99	22,87	12,14	1,0	10,02	ND	ND	1,12
α 1,2 Manosidasa								
G2 - <i>N</i> -glicanos de Gal ₂ GlcNAc ₂ Man ₃ GlcNAc ₂ M6-M8 - <i>N</i> -glicanos de Man ₆ GlcNAc ₂ a Man ₈ GlcNAc ₂ M9 - <i>N</i> -glicanos de Man ₉ GlcNAc ₂ M9+ - <i>N</i> -glicanos de Man ₉ GlcNAc ₂ y de tamaño superior Hib - <i>N</i> -glicanos híbridos								

Las cepas **YGLY7362**, **7366**, **7396** y **7398** se cultivaron en fermentadores de 3 l y se procesaron a través de cromatografía de Azul Sepharose 6 FF seguido de cromatografía de hidroxiapatita (HA). Las alícuotas tanto de las combinaciones de Azul como de las combinaciones de HA se redujeron y se analizaron por SDS-PAGE al 4-20 %.

Las combinaciones correspondientes para **YGLY3159** se incluyeron como controles positivos. La **Figura 35A** muestra un SDS-PAGE al 4-20 % teñido con azul de Commassie que muestra que tanto las combinaciones de Azul (mitad izquierda del gel) como las combinaciones de HA (mitad derecha del gel) producían rhEPO. La **Figura 35B** muestra una transferencia de Western de las mismas muestras sondeadas con anticuerpos anti-HCA. Ninguna de las cepas sometidas a ensayo presentaba actividad alguna de unión cruzada detectable a anticuerpos anti-HCA en cualquiera de la combinación de Azul o la combinación 1 de HA.

La **Figura 36** analiza la combinación de Azul y la combinación 1 de HA para rhEPO aislada de cultivos de SixFors de 500 ml de **YGLY7398** para la actividad de unión cruzada a anticuerpos anti-HCA. El panel que se encuentra en la parte más a la derecha muestra una transferencia de Western sondeada con otra preparación de anti-HCA: conejo policlonal GiF2::6316. Este anticuerpo producía los mismos resultados que los producidos usando el anticuerpo F072208-S, que se había usado para producir los ELISA y las transferencias de Western que se muestran en el presente documento. El anticuerpo 6316 muestra que la actividad de unión cruzada no es específica de anticuerpo.

Estos resultados muestran que la delección o la interrupción de los cuatro genes *BMT* puede dar como resultado cepas que no producen la actividad de unión cruzada detectable a anticuerpos anti-HCA en cualquiera de la rhEPO después de en la etapa de captura preliminar de Azul SEPHAROSE 6 FF o la etapa intermedia de hidroxiapatita usando hidroxiapatita 40 μ M de Tipo I. Estas cepas minimizan el riesgo de que se produzcan preparaciones de rhEPO que contengan actividad de unión cruzada a anticuerpos anti-HCA. Esto permite la producción de rhEPO con menos riesgo de inducción de una respuesta inmunológica adversa en el individuo que recibe la rhEPO.

EJEMPLO 9

Una comparación de la farmacocinética de la rhEPO producida en las cepas producidas en el Ejemplo 2 con los cuatro genes *BMT* interrumpidos o con delección y PEGilados se comparó con la rhEPO PEGilada producida a partir de la cepa **YGLY3159**. La comparación mostraba que la EPO PEGilada presentaba una reducción de la semivida *in vivo* y menor potencia *in vivo* (Véanse las Tablas 7 y 8). La rhEPO producida en las cepas producidas en el Ejemplo 2 sin actividad de unión cruzada detectable a anticuerpos anti-HCA tenía una farmacocinética por lo general similar a la de EPOGEN y no la farmacocinética más elevada de ARANESP. Se encontró que la farmacocinética reducida era una función de la cantidad de *N*-glicanos biantenarios bisialilados. Los niveles más elevados de *N*-glicano biantenarios bisialilados en la rhEPO se correlacionaron con una farmacocinética más elevada. Estos resultados son coherentes con los datos publicados que muestran que la semivida más larga se correlaciona con un contenido más elevado de ácido siálico en la eritropoyetina humana recombinante producida en células CHO (Egrie *et al.*, Exp. Hematol. 31: 290-299 (2003)).

Tabla 7		
PK de rhEPO de YGLY3159 (CBA) a YGLY7398 (sin CBA)		
	YGLY3159	YGLY7398
T1/2 (h)	20,9 ± 2	13 ± 2
CBA – actividad de unión cruzada		

Tabla 8		
Fuente de rhEPO	Potencia Relativa (Producción de Reticulocitos)	Intervalo de Confianza de un 95 %
YGLY3159 a YGLY7398	0,82	(0,68, 1,00)

EJEMPLO 10

- 5 Para conseguir de forma eficaz la eliminación de glicanos de tipo manosa unidos en β detectable y producir una cepa que produzca rhEPO con una farmacocinética más elevada, las cepas **YGLY7113-7122** descritas en el Ejemplo 3 se prepararon y se evaluaron por su capacidad para producir rhEPO carente de actividad de unión cruzada detectable a anticuerpo anti-HCA. Estas cepas se modificaron para que también expresaron EPO madura humana como una proteína de fusión fusionada a la secuencia líder de lisozima de pollo. Por lo tanto, estas cepas
- 10 expresan tanto EPO madura humana fusionada con el péptido señal α MATpre de *S. cerevisiae* como la EPO madura humana como una proteína de fusión fusionada a la secuencia líder de lisozima de pollo.

- Diversas cepas **YGLY7113-YGLY7122** en las que los cuatro genes *BMT* implicados en la ruta de β -manosiltransferasa se interrumpieron y se expresaron cultivados en fermentadores SixFors de 500 ml y a continuación se procesaron para rhEPO a través de combinaciones de Azul SEPHAROSE 6 FF (combinaciones de Azul). Las alícuotas de las combinaciones de Azul para varias cepas se analizaron por ELISA de sándwich usando anticuerpos anti-HCA. La **Figura 37** muestra que **YGLY7118** presentaba una actividad de unión cruzada a anticuerpo anti-HCA muy baja pero las demás cepas no presentaban actividad de unión cruzada detectable a anticuerpos anti-HCA. El análisis de N-glicano por HPLC de las cepas **YGLY7113-7117** se muestra en la Tabla 9 y las cepas **YGLY7118-7122** se muestra en la Tabla 10. Las tablas se presentan gráficamente en la **Figura 38**.

Tabla 9								
Semana 48 - SixFors - cepas $\Delta bmt1-4$ - combinaciones de Azul								
Combinaciones	% Bi-Sialilado	% Mono Sialilado	% Neutro	% Neutro				
				% G2	% M5	% M6-M8	% M9	% Híb
X074814 (YGLY7113)	70,23	9,97	19,80	0,2	9,2	6,85	2,05	1,5
X074814 (YGLY7113) α 1,2 Manosidasa	68,96	10,56	20,48	0,3	18,4	ND	ND	1,78
X074815 (YGLY7115)	62,61	14,01	23,38	0,5	7,15	10,37	3,96	1,4
X074815 (YGLY7115) α 1,2 Manosidasa	61,77	13,95	24,28	0,1	22,22	ND	ND	1,96
X074816 (YGLY7114)	67,64	8,22	24,14	0,2	4,2	11,41	6,33	2,0
X074816 (YGLY7114) α 1,2 Manosidasa	65,92	8,32	25,76	0,2	23,35	ND	ND	2,21
X074817 (YGLY7116)	66,46	8,06	25,48	4,73	5,38	6,94	7,23	1,2
X074817 (YGLY7116) α 1,2 Manosidasa	65,54	8,69	25,77	0,5	23,8	ND	ND	1,47

X074818 (YGLY7117)	70,06	11,09	18,85	0,6	8,59	6,0	2,21	1,45
X074818 (YGLY7117) α 1,2 Manosidasa	68,67	11,42	19,91	0,4	17,5	ND	ND	2,01
G2 - <i>N</i> -glicanos de Gal ₂ GlcNAc ₂ Man ₃ GlcNAc ₂ M6-M8 - <i>N</i> -glicanos de Man ₆ GlcNAc ₂ a Man ₈ GlcNAc ₂ M9 - <i>N</i> -glicanos de Man ₉ GlcNAc ₂ M9+ - <i>N</i> -glicanos de Man ₉ GlcNAc ₂ y de tamaño superior Hib - <i>N</i> -glicanos híbridos								

Tabla 10								
Semana 48 - SixFors – cepas de $\Delta bmt1-4$ – combinaciones de Azul								
Combinaciones	% Bi-	% Mono	% Neutro	% Neutro				
				% G2	% M5	% M6-M8	% M9	% Hib
X074819 (YGLY7119)	58,12	27,10	14,78	0,7	6,83	4,98	1,17	1,1
X074819 (YGLY7119) α 1,2 Manosidasa	57,03	26,87	16,10	0,45	14,55	ND	ND	1,1
X074820 (YGLY7120)	73,60	10,84	15,56	0,89	8,6	3,75	1,64	0,68
X074820 (YGLY7120) α 1,2 Manosidasa	72,43	11,13	16,44	0,7	15,63	ND	ND	0,11
X074821 (YGLY7121)	59,41	19,85	20,74	0,8	3,04	10,7	5,55	0,65
X074821 (YGLY7121) α 1,2 Manosidasa	58,39	20,00	21,6	0,4	20,17	ND	ND	1,04
X074822 (YGLY7122)	57,43	24,16	18,41	1,37	10,89	4,95	0,4	0,8
X074822 (YGLY7122) α 1,2 Manosidasa	55,77	24,44	19,79	1,8	17,28	ND	ND	0,71
X074824 (YGLY7118)	55,56	21,47	22,97	0,33	2,98	11,85	6,59	1,22
X074824 (YGLY7118) α 1,2 Manosidasa	54,68	21,67	23,65	0,4	22,5	ND	ND	0,75
G2 - <i>N</i> -glicanos de Gal ₂ GlcNAc ₂ Man ₃ GlcNAc ₂ M6-M8 - <i>N</i> -glicanos de Man ₆ GlcNAc ₂ a Man ₈ GlcNAc ₂ M9 - <i>N</i> -glicanos de Man ₉ GlcNAc ₂ M9+ - <i>N</i> -glicanos de Man ₉ GlcNAc ₂ y de tamaño superior Hib - <i>N</i> -glicanos híbridos								

Las cepas **YGLY7115**, **7117** y **7120** se cultivaron en fermentadores de 3 l y se procesaron a través de Azul SEPHAROSE 6 FF cromatografía seguido de cromatografía de hidroxipatita (HA). Las alícuotas tanto de las combinaciones de Azul como de las combinaciones de HA se redujeron y se analizaron por SDS-PAGE al 4-20 %. Las combinaciones correspondientes para **YGLY3159** se incluyeron como controles positivos. Las combinaciones correspondientes para **YGLY7395** se incluyeron como controles negativos. La **Figura 39A** muestra un SDS-PAGE al 4-20 % teñido con azul de Commassie que muestra que tanto las combinaciones de Azul (mitad izquierda del gel) como las combinaciones de HA (mitad derecha del gel) que producen rhEPO. La **Figura 39B** muestra una transferencia de Western de las mismas muestras sondeadas con anticuerpos anti-HCA. Ninguna de las cepas sometidas a ensayo presentaba actividad alguna de unión cruzada detectable a anticuerpos anti-HCA en cualquiera de la combinación de Azul o la combinación 1 de HA.

Estos resultados también muestran que la delección o la interrupción de los cuatro genes *BMT* puede dar como resultado cepas que no producen la actividad de unión cruzada detectable a anticuerpos anti-HCA en cualquiera de

la rhEPO después de la etapa preliminar de captura de Azul SEPHAROSE 6 FF o la etapa intermedia de hidroxapatita usando hidroxapatita 40 µM de Tipo I. Estas cepas minimizan el riesgo de que se realicen preparaciones de rhEPO que contengan actividad de unión cruzada a anticuerpos anti-HCA. Esto permite la producción de rhEPO con menos riesgo de inducción de una respuesta inmunológica adversa en el individuo que recibe la rhEPO.

EJEMPLO 11

Las combinaciones de azul que contienen rhEPO producida por **YGLY7117** se sometieron adicionalmente a cromatografía en columna de hidroxapatita y la rhEPO en las combinaciones de HA se analizaron para contenido de sialilación. La **Figura 40A** y la **Figura 40B** muestran trazas por HPLC de los *N*-glicanos de la rhEPO producida en **YGLY3159** en comparación con los *N*-glicanos de la rhEPO producida en **YGLY7117**, respectivamente. Las figuras también muestran que la columna de hidroxapatita retiraba contaminantes adicionales; por lo tanto, en este análisis, el contenido de sialilación de la rhEPO producida por **YGLY7117** era de aproximadamente un 99 % (los *N*-glicanos neutros eran aproximadamente un 1 %) de los que aproximadamente un 89 % era A2 o bisialilado y aproximadamente un 10 % era A1 o monosialilado.

El análisis de sialilación de rhEPO producido en **YGLY7117** después de la PEGilación de acuerdo con el proceso en el Ejemplo 3 era similar a la cantidad de sialilación antes de la PEGilación; sin embargo, la cantidad de sialilación puede variar hasta un punto limitado dependiendo por ejemplo, de qué modificaciones se realizaron para las condiciones de crecimiento, por ejemplo, composiciones del medio, tasa de alimentación, etc (Véase la Tabla 11). Por lo tanto, los métodos en el presente documento producen composiciones de rhEPO que tienen una sialilación de A2 de al menos aproximadamente un 75 % o una sialilación de A2 entre aproximadamente un 75 % y un 89 %. Por lo tanto, el contenido de ácido siálico total es al menos de 4,5 moles de ácido siálico por mol de rhEPO, de forma más específica, de aproximadamente 4,6 a 5,7 moles de ácido siálico por mol de rhEPO.

Tabla 11				
	BPP (2000 l) (n = 3)	FPP (800 l) (n = 2)	Avecia (15 l) (n = 2)	Avecia (100 l) (n = 1)
Pureza por SDS PAGE (relacionada con EPO) ($\geq 95,0\%$)	99,5 $\pm$ 0,4 %	99,4 $\pm$ 0,0 %	99,4 $\pm$ 0,1 %	99,4 %
Integridad por SDS PAGE (Mono-PEG) ($\geq 80,0\%$)	96,8 $\pm$ 0,7 %	96,0 $\pm$ 2,2 %	95,2 $\pm$ 2,0 %	97,7 %
Ácido siálico total ($\geq 4,5$ mol de SA / mol de proteína)	5,0-5,7	4,6-4,7	5,1-5,2	5,2
Glicano unido a N por CE (70-85 % de A2)	75,2-80,2 %	74,2-77,8 %	80,9-88,7 %	83,9 %
A2 - bi-sialilado CE - electroforesis por capilaridad SA - ácido siálico BPP- Planta Piloto Biológica FPP - Planta Piloto de Fermentación				

Una comparación de la farmacocinética de la rhEPO producida en la cepa **YGLY7117** producida en el Ejemplo 3 con los cuatro genes *BMT* interrumpidos o sometidos a delección y PEGilados se comparó con la rhEPO PEGilada producida a partir de la cepa **YGLY3159**. La comparación mostraba que la rhEPO PEGilada producida en la cepa **YGLY7117** presentaba una semivida *in vivo* y potencia *in vivo* similares a las de **YGLY3159** y ARANESP (Véanse las Tablas 12 y 13).

Tabla 12		
PK de rhEPO de YGLY3159 (CBA) a YGLY7117 (sin CBA)		
	YGLY3159	YGLY7117
T1/2 (h)	20,9 $\pm$ 2	20,6 $\pm$ 4

CBA - actividad de unión cruzada

Tabla 13		
Fuente de rhEPO	Potencia Relativa (Producción de Reticulocitos)	Intervalo de Confianza de un 95 %
YGLY3159 a YGLY7117	0,94	(0,77, 1,14)

SECUENCIAS

5

Las secuencias que se usaron para producir algunas de las cepas desveladas en los Ejemplos 1-11 se proporcionan en la Tabla 14.

Tabla 14

SEQ ID NO:	Descripción	Secuencia
1	gen de invertasa de <i>S. cerevisiae</i> (ScSUC2)	AGGCCTCGCAACAACCTATAATTGAGTTAAGTGCCTTTCCAAGCT AAAAAGTTTGAGGTTATAGGGGCTTAGCATCCACACGTCACAATC TCGGGTATCGAGTATAGTATGTAGAATTACGGCAGGAGGTTTCCC AATGAACAAAGGACAGGGGCACGGTGAGCTGTCGAAGGTATCCA TTTTATCATGTTTCGTTTGTACAAGCACGACATACTAAGACATTTA CCGTATGGGAGTTGTTGTCCTAGCGTAGTTCTCGCTCCCCCAGCA AAGCTCAAAAAAGTACGTCATTTAGAATAGTTTGTGAGCAAATTA CCAGTCGGTATGCTACGTTAGAAAAGGCCACAGTATTCTTCTACC AAAGGCGTGCTTTTGTGAACTCGATCCATTATGAGGGCTTCCAT TATCCCCGCATTTTTATTACTCTGAACAGGAATAAAAAGAAAAA ACCCAGTTTAGGAAATTATCCGGGGGCGAAGAAATACGCGTAGC GTTAATCGACCCACGTCAGGGTTTTTCCATGGAGGTTTCTGGA AAAAGTACGAGGAATGTGATTATAAATCCCTTTATGTGATGTCT AAGACTTTTAAGGTACGCGCGATGTTTGCCTATTACCATCATAGA GACGTTTCTTTTCGAGGAATGCTTAAACGACTTTGTTTGACAAAA ATGTTGCCTAAGGGCTCTATAGTAAACCATTTGGAAGAAAGATT GACGACTTTTTTTTTTTGGATTTTCGATCCTATAATCCTTCTCCTG AAAAGAAACATATAAATAGATATGTATTATTCTTCAAAACATTCT CTTGTTCTTGCTTTTTTTTTTACCATATATCTTACTTTTTTTTTTC TCTCAGAGAAACAAGCAAAACAAAAAGCTTTTCTTTTCACTAACG TATATGATGCTTTTGCAAGCTTTCCTTTTCTTTTGGCTGGTTTTG <u>CAGCCAAAATATCTGCATCAATGACAAACGAACTAGCGATAGA</u> <u>CCTTTGGTCCACTTCACACCCAACAAGGGCTGGATGAATGACCCA</u> <u>AATGGGTTGTGGTACGATGAAAAAGATGCCAAATGGCATCTGTA</u> <u>CTTTCAATACAACCCAATGACACCGTATGGGGTACGCCATTGTT</u> <u>TTGGGGCCATGCTACTTCCGATGATTGACTAATTGGGAAGATCA</u> <u>ACCCATTGCTATCGCTCCAAGCGTAACGATTACGGTGCTTTCTC</u> <u>TGGCTCCATGGTGGTTGATTACAACAACACGAGTGGGTTTTTCAA</u> <u>TGATACTATTGATCCAAGACAAAGATGCGTTGCGATTGGACTTA</u> <u>TAACACTCCTGAAAGTGAAGAGCAATACATTAGCTATTCTCTTGA</u> <u>TGGTGGTTACACTTTTACTGAATACCAAAAAGAACCCTGTTTAGC</u> <u>TGCCAACTCCACTCAATTCAGAGATCCAAAGGTGTTCTGGTATGA</u> <u>ACCTTCTCAAAAATGGATTATGACGGCTGCCAAATCACAAGACTA</u> <u>CAAAATTGAAATTTACTCCTCTGATGACTTGAAGTCTGGAAGCT</u> <u>AGAATCTGCATTTGCCAATGAAGGTTTCTTAGGCTACCAATACGA</u> <u>ATGTCCAGGTTTGATTGAAGTCCCAACTGAGCAAGATCCTTCCAA</u> <u>ATCTTATTGGGTCATGTTTATTTCTATCAACCCAGGTGCACCTGCT</u> <u>GGCGGTTCCTTCAACCAATATTTGTTGGATCCTTCAATGGTACT</u> <u>CATTTTGAAGCGTTTGACAATCAATCTAGAGTGGTAGATTTTGGT</u> <u>AAGGACTACTATGCCTTGCAAACCTTTCTTCAACACTGACCCAACC</u> <u>TACGGTTCAGCATTAGGTATTGCCTGGGCTTCAAACCTGGGAGTAC</u>

		AGTGCCTTTGTCCCAACTAACCCTATGGAGATCATCCATGTCTTTG GTCCGCAAGTTTTCTTTGAACACTGAATATCAAGCTAATCCAGAG ACTGAATTGATCAATTTGAAAGCCGAACCAATATTGAACATTAGT AATGCTGGTCCCTGGTCTCGTTTTGCTACTAACACAACCTCTA AAGGCCAATTCTTACAATGTCGATTTGAGCAACTCGACTGGTACC CTAGAGTTTGAGTTGGTTTACGCTGTTAACACCACACAAACCATA TCCAAATCCGTCTTTGCCGACTTATCACTTTGGTTCAAGGGTTTA GAAGATCCTGAAGAATATTTGAGAATGGGTTTTGAAGTCAGTGCT TCTTCCTTCTTTTTGGACCGTGGTAACTCTAAGGTCAAGTTGTCA AGGAGAACCCATATTTACAAACAGAATGTCTGTCAACAACCAAC CATTCAAGTCTGAGAACGACCTAAGTTACTATAAAGTGTACGGCC TACTGGATCAAAACATCTTGGAATTGTACTTCAACGATGGAGATG TGGTTTCTACAAATACCTACTTTCATGACCACCGGTAACGCTCTAG GATCTGTGAACATGACCCTGGTGTGCGATAAATTTGTTCTACATTG ACAAGTTCCAAGTAAGGGGAAGTAAATAAGAGTTATAAAACTTA TTGCTTTTTTATTTTTTCAAAAGCCATTCTAAAGGGCTTTAGCT AACGAGTGACGAATGTAAACTTTATGATTTCAAAGAATACCTCC AAACCATTGAAAATGTATTTTTATTTTTATTTTCTCCCGACCCAG TTACCTGGAATTTGTTCTTTATGTACTTTATATAAGTATAATTCTC TTAAAAATTTTTACTACTTTGCAATAGACATCATTTTTTTCACGTAA TAAACCCACAATCGTAATGTAGTTGCCTTACACTACTAGGATGGA CCTTTTTGCCTTTATCTGTTTTGTTACTGACACAATGAAACCGGGT AAAGTATTAGTTATGTGAAAATTTAAAAGCATTAAAGTAGAAGTAT ACCATATTGTAAAAAAGCGTTGTCTTCTACGTAAAAGTGT TCTCAAAAAGAAGTAGTGAGGGAAATGGATACCAAGCTATCTGT AACAGGAGCTAAAAAATCTCAGGGAAAAGCTTCTGGTTTGGGAA ACGGTCGAC
2	invertasa de <i>S. cerevisiae</i> (ScSUC2)	MLLQAFLLAGFAAKISASMTNETSDRPLVHFTPNKGWMNDPNGL WYDEKDAKWHLYFYQNPNDTVWGTPLFWGHATSDDLTNWEDQPI AIAPKRNDSGAFSGSMVVDYNNNTSGFFNDTIDPRQRCVAIWYNTPE SEEQYISYSLDGGYTFTYQKNPVLAA NSTQFRDPKVFWEPSQKWI MTAAKSQDYKIEIYSSDDLKSWKLESAFANEGFLGYQYECPLIEVP TEQDPSKSYWVMFISINPGAPAGGSFNQYFVGSFNHGFHFAFDNQSR VVDGFKDYALQTFNTDPTYGSALGLAWASNWEYSFVPTNPWR SSMSLVRKFSNLNTEYQANPETELINLKAEPILNISNAGPWSRFATNTT LTKANSYNVDLSNSTGTLEFELVYAVNTTQTISKSVFADLSLWFKGL EDPEEYLRMGFEVSASSFFLDRGNSKVFKVKNPYFTNRMSVNNQP FKSENDLSYYKVYGLLDQNLLEYFNDGDVVSTNTYFMTTGNALGS VNMTTGVDNLFYIDKFQVREVK
3	gen transportador de UDP-GlcNAc de <i>K. lactis</i> (KIMNN2-2)	AAACGTAACGCCTGGCACTCTATTTTCTCAAACCTTCTGGGACGGA AGAGCTAAATATTGTGTTGCTTGAACAAACCCAAAAAACA AAATGAACAACTAAACTACACCTAAATAAACCGTGTGTAAAA CGTAGTACCATATTACTAGAAAAGATCACAAGTGTATCACACATG TGCATCTCATATTACATCTTTTATCCAATCCATTCTCTCTATCCCG TCTGTTCTGTGAGATTCTTTTCCATAAAAAGAAGAAGACCCCG AATCTCACCGGTACAATGCAAACTGCTGAAAAAAGAAAGT TCACTGGATACGGGAACAGTGCCAGTAGGCTTCACCACATGGAC AAAACAATTGACGATAAAATAAGCAGGTGAGCTTCTTTTCAAGT CACGATCCCTTTATGTCTCAGAAACAATATATACAAGCTAAACCC

		TTTTGAACCAGTTCTCTCTTCATAGTTATGTTACATAAAATTGCGG GAACAAGACTCCGCTGGCTGTCAGGTACACGTTGTAACGTTTTTCG TCCGCCCAATTATTAGCACAAACATTGGCAAAAAGAAAACTGCTC GTTTTCTCTACAGGTAAATTACAATTTTTTTCAGTAATTTTCGCTG AAAAATTTAAAGGGCAGGAAAAAAGACGATCTCGACTTTGCGAT AGATGCAAGAACTGTGGTCAAACTTGAAATAGTAATTTTGCTGT GCGTGAACTAATAAATATATATATATATATATATATATATTTGTG TATTTTGTATATGTAATTGTGCACGTCTTGGCTATTGGATATAAG ATTTTCGCGGGTTGATGACATAGAGCGTGTACTACTGTAATAGTT GTATATTCAAAAGCTGCTGCGTGGAGAAAGACTAAAATAGATAA AAAGCACACATTTGACTTCGGTACCGTCAACTTAGTGGGACAGT CTTTTATATTTGGTGTAAGCTCATTCTGGTACTATTCGAAACAGA ACAGTGTTTTCTGTATTACCGTCCAATCGTTTGTCTAGATTTTGT ATTGATTTTGTGCTTAGTGTTTCGGAGGATGTTGTTCCAATGTGAT TAGTTTCGAGCACATGGTGCAAGGCAGCAATATAAAATTTGGGAA ATATTGTTACATTCACTCAATTCGTGTCTGTGACGCTAATTCAGTT GCCCAATGCTTTGGACTTCTCTCACTTTCCGTTTAGGTTGCGACCT AGACACATTCCTCTTAAGATCCATATGTTAGCTGTGTTTTTGTCT TTACCAGTTCAGTCGCCAATAACAGTGTGTTTAAATTTGACATTT CCGTTCCGATTCATATTATCATTAGATTTTCAGGTACCACTTTGAC GATGATAATAGGTTGGGCTGTTTGTAAATAAGAGGTACTCCAACT TCAGGTGCAATCTGCCATCATTATGACGCTTGGTGCGATTGTGCG ATCATTATACCGTGACAAAGAATTTCAATGGACAGTTTAAAGTT GAATACGGATTCAGTGGGTATGACCCAAAAATCTATGTTTGGTAT CTTTGTTGTGCTAGTGGCCACTGCCTTGATGTCAATTGTTGTGCTTG CTCAACGAATGGACGTATAACAAGTACGGGAAACATTGGAAAGA AACTTTGTTCTATTTCGCAATTTCTTGGCTCTACCGTTGTTTATGTTG GGGTACACAAGGCTCAGAGACGAATTCAGAGACCTCTTAATTTCC TCAGACTCAATGGATATTCCTATTGTTAAATTACCAATTGCTACG AACTTTTCATGCTAATAGCAAATAACGTGACCCAGTTCATTTGT ATCAAAGGTGTTAACATGCTAGCTAGTAACACGGATGCTTTGACA CTTTCTGTGCTGCTTCTAGTGCGTAAATTTGTTAGTCTTTTACTCA GTGTCTACATCTACAAGAACGTCCTATCCGTGACTGCATACCTAG GGACCATCACCGTGTTCTTGGGAGCTGGTTTGTATTATATGGTT CGGTCAAACTGCACTGCCTCGCTGAAACAATCCACGTCTGTATG ATACTCGTTTCAGAATTTTTTTGATTTTCTGCCGGATATGGTTTCT CATCTTTACAATCGCATTCTTAATTATACCAGAACGTAATTCAAT GATCCCAGTGACTCGTAACCTCTTATATGTCAATTTAAGC
4	transportador de UDP-GlcNAc de <i>K. lactis</i> (KIMNN2-2)	MSFVLILSLVFGGCCSNVISFEHMOVQGSNINLGNIVTFTQFVSVTLIQ LPNALDFSHPFRLRPRHIPLKIHMLAVLFFFTSSVANNSVFKFDISVPI HHIRFSGTTLTMIIGWAVCNKRYSKLQVQSAIIMTLGAIVASLYRDK EFSMDSLKLNTDSVGMTQKSMFGIFVVLVATALMSLLSLLNEWTYN KYGKHWKETLFYSHFLALPLFMLGYTRLRDEFRLDISSDSMDIPIV KLPIATKLFMLIANNVTFICIKGVNMLASNTDALTSVLLVRKFVS LLLSVYIYKNVLSVTAYLGTITVFLGAGLYSYGSVK TALPR
5	ADN codifica líder de Mnn2 (53)	ATGCTGCTTACCAAAAGGTTTTCAAAGCTGTTCAAGCTGACGTTT ATAGTTTTGATATTGTGCGGGCTGTTGCTCATTACAAACAAATAC ATGGATGAGAACACGTCG
6	líder de Mnn2 (53)	MLLTKRFSKLFKLT FIVLILCGLFVITNKYMDENTS

7	ADN codifica líder de Mnn2 (54) Los últimos 9 nucleótidos son el conector que contiene el sitio de restricción de Ascl)	ATGCTGCTTACCAAAAGGTTTTCAAAGCTGTTCAAGCTGACGTTT ATAGTTTTGATATTGTGCGGGCTGTTTCGTCATTACAAACAAATAC ATGGATGAGAACACGTCGGTCAAGGAGTACAAGGAGTACTTAGA CAGATATGTCCAGAGTTACTCCAATAAGTATTCATCTTCCTCAGA CGCCGCCAGCGCTGACGATTCAACCCCATTTGAGGGACAATGATG AGGCAGGCAATGAAAAGTTGAAAAGCTTCTACAACAACGTTTTCA ACTTTCTAATGGTTGATTTCGCCCCGGGCGCGCC
8	líder de Mnn2 (54)	MLLTKRFSKLFKLT FIVLILCGLFVITNKYMDENTSVKEYKEYLDRY VQSYSNKYSSSSD AASADDSTPLRDND EAGNEKLKSFYNNVFNFLMVDSPGRA
9	ADN codifica secuencia pre señal del Factor de Acoplamiento a <i>S. cerevisiae</i>	ATG AGA TTC CCA TCC ATC TTC ACT GCT GTT TTG TTC GCT GCT TCT TCT GCT TTG GCT
10	secuencia pre señal del Factor de Acoplamiento a <i>S. cerevisiae</i>	MRFPSIFTAVLFAASSALA
11	ADN codifica Pp SEC12 (10) Los últimos 9 nucleótidos son el conector que contiene el sitio de restricción de Ascl usado para fusión a proteínas de interés.	ATGCCCAGAAAAATATTTAACTACTTCATTTTGACTGTATTCATG GCAATTCTTGCTATTGTTTTACAATGGTCTATAGAGAATGGACAT GGGCGCGCC
12	Pp SEC12 (10)	MPRKIFNYFILTVMAILAIVLQWSIENGHGRA
13	ADN codifica	ATGGCCCTCTTTCTCAGTAAGAGACTGTTGAGATTTACCGTCATT GCAGGTGCGGTTATTGTTCTCCTCCTAACATTGAATTCACACAGT
	ScMnt1 (Kre2) (33)	AGAACTCAGCAATATATTCCGAGTTCCATCTCCGCTGCATTTGAT TTTACCTCAGGATCTATATCCCCTGAACAACAAGTCATCGGGCGC GCC
14	ScMnt1 (Kre2) (33)	MALFLSKRLLRFTVIAGAVIVLLLTLNSNSRTQQYIPSSISAAFDFTSG SISPEQQVIGRA
15	ADN codifica ScSEC12 (8) Los últimos 9 nucleótidos son el conector que contiene el sitio de restricción de Ascl usado para fusión a proteínas de interés	ATGAACACTATCCACATAATAAAAATTACCGCTTAACTACGCCAAC TACACCTCAATGAAACAAAAAATCTCTAAATTTTACCAACTTC ATCCTTATTGTGCTGCTTTCTTACATTTTACAGTTCTCCTATAAGC ACAATTTGCATTCCATGCTTTTCAATTACGCGAAGGACAATTTTCT AACGAAAAGAGACACCATCTCTTCGCCCTACGTAGTTGATGAAGA CTTACATCAAACAACTTTGTGTTGGCAACCACGGTACAAAAACATC TGTACCTAGCGTAGATTCCATAAAAGTGCATGGCGTGGGGCGCG CC
16	ScSEC12 (8)	MNTIHIIKLPLNYANYTSMKQKISKFFTNFILIVLLSYILQFSYKHNH SMLFNYAKDNFLTKRDTISSPYVVEDDLHQTTLFGNHGKTSVPSV DSIKVHGVGRA

17	ADN codifica transportador de MmSLC35 A3 UDP- GlcNAc	ATGCTCTGCCAACCTAAAATATCTTTCCTTGGGAATTTTGGTGTTC AGACTACCAGTCTGGTTCTAACGATGCGGTATTCTAGGACTTTAA AAGAGGAGGGGCGCTCGTTATCTGTCTTCTACAGCAGTGGTTGTGG CTGAATTTTGAAGATAATGGCCTGCATCTTTTAGTCTACAAAG ACAGTAAGTGTAGTGTGAGAGCACTGAATAGAGTACTGCATGAT GAAATTCTTAATAAGCCCATGGAAACCCTGAAGCTCGCTATCCCG TCAGGGATATATACTCTTCAGAACAACTTACTCTATGTGGCACTG TCAAACCTAGATGCAGCCACTTACCAGGTTACATATCAGTTGAAA ATACTTACAACAGCATTATTTTCTGTGTCTATGCTTGGTAAAAAA TTAGGTGTGTACCAGTGGCTCTCCCTAGTAATTCTGATGGCAGGA GTTGCTTTTGTACAGTGGCCTTCAGATTCTCAAGAGCTGAACTCT AAGGACCTTTCAACAGGCTCACAGTTTGTAGGCCTCATGGCAGTT CTCACAGCCTGTTTTTCAAGTGGCTTGTCTGGAGTTTATTTTGAG AAAATCTTAAAAGAAACAAAACAGTCAGTATGGTAAGGAACAT TCAACTTGGTTTCTTTGGAAGTATATTTGGATTAATGGGTGTATA CGTTTATGATGGAGAATTGGTCTCAAAGAATGGATTTTTCAGGG ATATAATCAACTGACGTGGATAGTTGTTGCTCTGCAGGCACCTGG AGGCCTTGTAATAGCTGCTGTCATCAAATATGCAGATAACATTTT AAAAGGATTTGCGACCTCCTTATCCATAATATTGTCAACAATAAT ATCTTATTTTGGTTGCAAGATTTTGTGCCAACCAAGTGTCTTTTC CTTGAGCCATCCTTGTAAATAGCAGCTACTTTCTTGTATGGTTAC GATCCCAAACCTGCAGGAAATCCCACTAAAGCATAG
18	transportador de MmSLC35 A3 UDP- GlcNAc	MSANLKYLSLGLVFQTTSLVLTMRYSRTLKEEGPRYLSSTAVVVAE FLKIMACIFLVYKDSKCSVRALNRVLHDEILNKP METLKLAIPSGIYT LQNNLLYVALSNLDAATYQVTYQLKILTTALFSVSMLGKKLGVYQ WLSLVILMAGVAFVQWPSDSQELNSKDLSTGSQFVGLMAVLTACFS
		SGFAGVYFEKILKETKQSVWIRNIQLGFFGSIFGLMGVYVYDGELVS KNGFFQGYNQLTWIVVALQALGGLVIAAVIKYADNILKGFATSLSHL STIISYFWLODFVPTSVFFLGAILVIAATFLYGYDPKPAGNPTKA
19	ADN codifica DmUGT	ATGAATAGCATACACATGAACGCCAATACGCTGAAGTACATCAG CCTGCTGACGCTGACCCTGCAGAATGCCATCCTGGGCCTCAGCAT GCGCTACGCCCCGACCCCGGCCAGGCGACATCTTCTCAGCTCCAC GGCCGTACTCATGGCAGAGTTCGCCAACTGATCAGTGCCTGTT CCTGGTCTTCAACGAGGAGGGCAAGGATGCCAGAAGTTTGTAC GCTCGCTGCACAAGACCATCATTGCGAATCCCATGGACACGCTGA AGGTGTGCGTCCCCTCGCTGGTCTATATCGTTCAAAACAATCTGC TGTACGTCTCTGCCTCCCATTTGGATGCGGCCACCTACCAGGTGA CGTACCAGCTGAAGATTCTCACACGCGCCATGTTGCGGGTTGTCA TTCTGCGCCGCAAGCTGCTGAACACGCAGTGGGGTTCGCTGCTGC TCCTGGTGTATGGGCATCGTCTGGTGCAGTTGGCCCAAACGGAG GGTCCGACGAGTGGCTCAGCCGGTGGTGCCGCAGCTGCAGCCAC GGCCGCCTCCTCTGGCGGTGCTCCCGAGCAGAACAGGATGCTCG GACTGTGGGCCGCACTGGGCGCTGCTTCTCTCCGGATTCCGCGG GCATCTACTTTGAGAAGATCCTCAAGGGTGCCGAGATCTCCGTGT GGATGCGGAATGTGCAGTTGAGTCTGCTCAGCATTCCCTTCGGCC TGCTCACCTGTTTCGTTAACGACGGCAGTAGGATCTTCGACCAGG GATTCTTCAAGGGCTACGATCTGTTTGTCTGGTACCTGGTCTGC TGCAGGCCGCGGTGGATTGATCGTTGCCGTGGTGGTCAAGTAC GCGGATAACATTCTCAAGGGCTTCGCCACCTCGCTGGCCATCATC ATCTCGTGCGTGGCCTCCATATACATCTTCGACTTCAATCTCACG CTGCAGTTCAGCTTCGGAGCTGGCCTGGTCATCGCCTCCATATTT CTCTACGGCTACGATCCGGCCAGGTGCGCGCCGAAGCCAACATATG CATGGTCTGGCGGCGATGAGGAGAAGCTGCTGCCGCGCGTCTA G

20	DmUGT	MNSIHMNANTLKYISLLTLTLQNAILGLSMRYARTRPGDIFLSSTAVL MAEFAKLITCLFLVFNEEGKDAQKFVRSLSHKTIANPMDTLKVCVPS LVYIVQNNLLYVSASHLDAATYQVTYQLKILTTAMFAVVILRRKLL NTQWGALLLLVMGIVLVQLAQTEGPTSGSAGGAAAAATAASSGGA PEQNRMLGLWAALGACFLSGFAGIYFEKILKGAEISVWMRNVQLSL LSIPFGLLTCFVNDGSRIFDQGFVKGYDLFVWYLVLLQAGGGLIVAV VVKYADNILKGFATSLAIIISCVASIYIFDENLTLQFSFGAGLVIASIFL YGYDPARSAPKPTMHGPGGDEEKLLPRV
21	ADN codifica ScGAL10	ATGACAGCTCAGTTACAAAGTGAAAGTACTTCTAAAATTGTTTTG GTTACAGGTGGTGCTGGATACATTGGTTCACACACTGTGGTAGAG CTAATTGAGAATGGATATGACTGTGTTGTTGCTGATAACCTGTGCG AATTCAACTTATGATTCTGTAGCCAGGTTAGAGGTCTTGACCAAG CATCACATTCCCTTCTATGAGGTTGATTTGTGTGACCGAAAAGGT CTGGAAAAGGTTTTCAAAGAATATAAAATTGATTTCGGTAATTCAC TTTGCTGGTTTTAAAGGCTGTAGGTGAATCTACACAAATCCCGCTG AGATACTATCACAATAACATTTTGGGAACTGTCGTTTTATTAGAG TTAATGCAACAATACAACGTTTCCAAATTTGTTTTTTCATCTTCTG CTACTGTCTATGGTGATGCTACGAGATTCCCAAATATGATTCCTA TCCCAGAAGAATGTCCCTTAGGGCCTACTAATCCGTATGGTCATA CGAAATACGCCATTGAGAATATCTTGAATGATCTTTACAATAGCG ACAAAAAAGTTGGAAGTTTGCTATCTTGCGTTATTTTAACCCAA
		TTGGCGCACATCCCTCTGGATTAATCGGAGAAGATCCGCTAGGTA TACCAAACAATTTGTTGCCATATATGGCTCAAGTAGCTGTTGGTA GGCGCGAGAAGCTTTACATCTTCGGAGACGATTATGATTCCAGAG ATGGTACCCCGATCAGGGATTATATCCACGTAGTTGATCTAGCAA AAGGTCATATTGCAGCCCTGCAATACCTAGAGGCCTACAATGAAA ATGAAGGTTTGTGTCGTGAGTGGAACCTTGGGTTCGGGTAAAGGTT CTACAGTTTTTGAAGTTTATCATGCATTCTGCAAAGCTTCTGGTAT TGATCTTCCATACAAAGTTACGGGCAGAAAGAGCAGGTGATGTTTT GAACTTGACGGCTAAACCAGATAGGGCCAAACGCGAACTGAAAT GGCAGACCGAGTTGCAGGTTGAAGACTCCTGCAAGGATTTATGG AAATGGACTACTGAGAATCCTTTTGGTTACCAGTTAAGGGGTGTC GAGGCCAGATTTTCCGCTGAAGATATGCGTTATGACGCAAGATTT GTGACTATTGGTGCCGGCACCAGATTTCAAGCCACGTTTGCCAAT TTGGGCGCCAGCATTTGTTGACCTGAAAGTGAAACGGACAATCAGTT GTTCTTGGCTATGAAAATGAGGAAGGGTATTTGAATCCTGATAGT GCTTATATAGGCGCCACGATCGGCAGGTATGCTAATCGTATTTTCG AAGGGTAAGTTTAGTTTATGCAACAAAGACTATCAGTTAACCGTT AATAACGGCGTTAATGCGAATCATAGTAGTATCGGTTCTTTCCAC AGAAAAAGATTTTGGGACCCATCATTCAAAATCCTTCAAAGGAT GTTTTTACCGCCGAGTACATGCTGATAGATAATGAGAAGGACACC GAATTTCCAGGTGATCTATTGGTAACCATAAGTATACTGTGAAC GTTGCCCAAAAAAGTTTGGAATGGTATATAAAAGGTAAATTGACT GCTGGTGAAGCGACGCCAATAAATTTAACAATCATAGTTATTTTC AATCTGAACAAGCCATATGGAGACACTATTGAGGGTACGGAGAT TATGGTGCGTTCAAAAAAATCTGTTGATGTCGACAAAAACATGAT TCCTACGGGTAATATCGTCGATAGAGAAATTGCTACCTTTAACTC TACAAAGCCAACGGTCTTAGGCCCAAAAAATCCCAAGTTTGATTG TTGTTTTGTGGTGGATGAAAATGCTAAGCCAAGTCAAATCAATAC TCTAAACAATGAATTGACGCTTATTGTCAAGGCTTTTCATCCCGA TTCCAATATTACATTAGAAGTTTTAAGTACAGAGCCAACCTATCA ATTTTATACCGGTGATTTCTTGTCTGCTGGTTACGAAGCAAGACA AGGTTTTGCAATTGAGCCTGGTAGATACATTGATGCTATCAATCA AGAGAACTGGAAAGATTGTGTAACCTTGAAAAACGGTGAAACTT ACGGGTCCAAGATTGTCTACAGATTTTCCTGA

22	ScGa110	MTAQLQSESTSKIVLVTGGAGYIGSHTVVELIENGYDCVVADNLSN STYDSVARLEVLTkHHIPFYEVdLcDRKGLEKVFKEYKIDSVIHfAG LKAvgESTQIPLRYyHNNILGTvVlLELMQQYNVSKFVfSSSATVYG DATRFPNMiPIPEECPLGPTNPYGHTKYAIENILNDLYNSDKKSWKFA ILRYFNPIGAHPSGLIGEDPLGIPNNLLPYMAQVAVGRREKLYIFGDD YDSRDGTPIRDYIHVVdLAKGHIAALQYLEAYNENEGLCREWNLGs GKGSTVFEVYHAFCKASGIDLPYKVTGRRAGDVLNLtAKPDRAKRE LKWQTELQVEDSCKDLWKWTTENPFgyQLRGVEARfSAEDMRYD ARFVTIGAGTRFQATFANLG ASIVDLKvNGQSVVLGYENEeGYLNPDSAYIGATIGRYANRISKGKF SLCNKDYQLTVNNGVnANHSSIGSFHRKRFLGPIIQNPskDVfTAey MLIDNEKDTEFPgDLLVTIQYTVNVAQKSLEmVYKGKLTAGEATPI NLtNHsYfNLNKPYGDTIEGTEIMVRSKKSVDVDKNMiptGNiVDRE IATFNSTKPTVLGPKNPQFDCCFVVDENAKPSQINTLNnelTLIVKAf HPDSNITLeVLSTePTYQfYtGDFLSAGYEARQGFaIEPGRYIDAINQ
		ENWKDCVTLKNGETYGSKIVYRFS
23	hGalT optimizado por codón (XB)	GGTAGAGATTTGTCTAGATTGCCACAGTTGGTTGGTGTttCCACT CCATTGCAAGGAGGTtCTAACTCTGCTGCTGCTATTGGTCAATCT TCCGGTGAGTTGAGAACTGGTGGAGCTAGACCACCTCCACCATTG GGAGCTTCTCTCAACCAAGACCAGGTGGTGATTCTTCTCCAGTT GTTGACTCTGGTCCAGGTCCAGCTTCTAACTTGACTTCCGTTCCA GTTCCACACACTACTGCTTTGTCTCCTTGCCAGCTTGTCAGAAGAA TCCCCATTGTTGGTTGGTCCAATGTTGATCGAGTTCAACATGCCA GTTGACTTGAGATTGGTTGCTAAGCAGAACCCAAACGTTAAGATG GGTGGTAGATACGCTCCAAGAGACTGTGTTTCCCCACACAAAGTT GCTATCATCATCCCATTCAGAAACAGACAGGAGCACTTGAAGTAC TGGTTGTACTACTTGCAcccAGTTTTGCAAAGACAGCAGTTGGAC TACGGTATCTACGTTATCAACCAGGCTGGTGACACTATTTTCAAC AGAGCTAAGTTGTTGAATGTTGGTTTCCAGGAGGCTTTGAAGGAT TACGACTACACTTGTTTCGTTTTCTCCGACGTTGACTTGATTCCAA TGAACGACCACAACGCTTACAGATGTTTCTCCAGCCAAGACACA TTTCTGTTGCTATGGACAAGTTCGGTTTTCTCCTTGCCATACGTTCA ATACTTCGGTGGTGTTTCCGCTTTGTCCAAGCAGCAGTTCTTGAC TATCAACGGTTTCCCAAACAATTACTGGGGATGGGGTGGTGAAG ATGACGACATCTTTAACAGATTGGTTTTTCAGAGGAATGTCCATCT CTAGACCAAACGCTGTTGTTGGTAGATGTAGAATGATCAGACACT CCAGAGACAAGAAGAACGAGCCAAACCCACAAAGATTTCGACAGA ATCGCTCACACTAAGGAACTATGTTGTCCGACGGATTGAACTCC TTGACTTACCAGGTTTTGGACGTTcAGAGATACCCATTGTACTACT CAGATCACTGTTGACATCGGTACTCCATCCTAG
24	domicilio catalítico de hGalT I (XB)	GRDLSRLPQLVGVSTPLQGGSNSAAaIGQSSGELRTGGARPPPLGA SSQPRPGGDSSPVVDSGPGPASNLTSVPVPHTTALSLPACPEESPLL GPMlIEFNMPVDLELVAKQNPVnKMGGRYAPRDCVSPHKVAIIPFR NRQeHLKYWLYYLHPVLQRQQLDYGIYVINQAGDTIFNRAKLLNVG FQEALKDYDYTCFVfSDVDLIPMNDHNAYRCfSQPRHISVAMDKFG FSLPYVQYFGGVsALSkkQQLTINGFPNNYWGwGGEDDDIFNRLVF RGMSISRPNAVVGRCRMIRHSRDkKNEPNPQRFDRIAHTKETMLSD GLNSLTyOVLdVORYPLYTOITVDIGTPS

25	ADN codifica dominio catalítico de GnT1 humano (NA) Optimizado por codón	TCAGTCAGTGCTCTTGATGGTGACCCAGCAAGTTTGACCAGAGAA GTGATTAGATTGGCCCAAGACGCAGAGGTGGAGTTGGAGAGACA ACGTGGACTGCTGCAGCAAATCGGAGATGCATTGTCTAGTCAAA GAGGTAGGGTGCCTACCGCAGCTCCTCCAGCACAGCCTAGAGTG CATGTGACCCCTGCACCAGCTGTGATTCCTATCTTGGTCATCGCC TGTGACAGATCTACTGTTAGAAGATGTCTGGACAAGCTGTTGCAT TACAGACCATCTGCTGAGTTGTTCCCTATCATCGTTAGTCAAGAC TGTGGTCACGAGGAGACTGCCCAAGCCATCGCCTCCTACGGATCT GCTGTCACTCACATCAGACAGCCTGACCTGTCACTATTGCTGTG CCACCAGACCACAGAAAGTTCCAAGGTTACTACAAGATCGCTAG ACACTACAGATGGGCATTGGGTCAAGTCTTCAGACAGTTTAGATT CCCTGCTGCTGTGGTGGTGGAGGATGACTTGGAGGTGGCTCCTGA CTTCTTTGAGTACTTTAGAGCAACCTATCCATTGCTGAAGGCAGA CCCATCCCTGTGGTGTGTCTCTGCCTGGAATGACAACGGTAAGGA GCAAATGGTGGACGCTTCTAGGCCTGAGCTGTTGTACAGAACCG ACTTCTTTCCTGGTCTGGGATGGTTGCTGTTGGCTGAGTTGTGGG
		CTGAGTTGGAGCCTAAGTGGCCAAAGGCATTCTGGGACGACTGG ATGAGAAGACCTGAGCAAAGACAGGGTAGAGCCTGTATCAGACC TGAGATCTCAAGAACCATGACCTTTGGTAGAAAGGGAGTGTCTCA CGGTCAATTCTTTGACCAACACTTGAAGTTTATCAAGCTGAACCA GCAATTTGTGCACTTCACCCAACTGGACCTGTCTTACTTGCAGAG AGAGGCCTATGACAGAGATTTCTAGCTAGAGTCTACGGAGCTCC TCAACTGCAAGTGGAGAAAGTGAGGACCAATGACAGAAAGGAGT TGGGAGAGGTGAGAGTGCAGTACACTGGTAGGGACTCCTTTAAG GCTTTCGCTAAGGCTCTGGGTGTCATGGATGACCTTAAGTCTGGA GTTCTAGAGCTGGTTACAGAGGTATTGTCACCTTTCAATTCAGA GGTAGAAGAGTCCACTTGGCTCCTCCACCTACTTGGGAGGGTTAT GATCCTTCTTGAATTAG
26	dominio catalítico de GnT I humano (NA)	SVSALDGDPA SLTREVIRLAQDAEVELERQ RGLLQQIGDALSSQRGR VPTA APPAQPRVHVTPAPAVIPILVIACDRSTVRRCLDKLLHYRPSAE LFPIIVSQDCGHEETAQAIASYGSAVTHIROPDLSIAVPPDHRKFQ YYKIARHYRWALGQVFRQFRFPAAVVVEDDLEVPDFFEYFRATYP LLKADPSLWCVSAWNDNGKEQMVDASRPELLYRTDFFPGLGWLLL AELWAELEPKWPKAFWDDWMRRPEQRQGRACIRPEISRTMTFGRK GVSHGQFFDQHLKFIKLNQQFVHFTQLDLSYLQREAYDRDFLARVY GAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAFKALGVMDDLK SGVPRAGYRGIVTFQFRGRRVHLAPPPTWEGYDPSWN

27	ADN codifica dominio catalítico de Mm ManI (FB)	GAGCCCCGCTGACGCCACCATCCGTGAGAAGAGGGCAAAGATCAA AGAGATGATGACCCATGCTTGGGAATAATTATAAACGCTATGCGTG GGGCTTGAACGAACTGAAACCTATATCAAAAGAAGGCCATTCAA GCAGTTTGTGGCAACATCAAAGGAGCTACAATAGTAGATGCCC TGGATACCCCTTTTCATTATGGGCATGAAGACTGAATTTCAAGAAG CTAAATCGTGGATTAAAAAATATTTAGATTTTAATGTGAATGCTG AAGTTTCTGTTTTTGAAGTCAACATACGCTTCGTCCGTGGACTGC TGTCAGCCTACTATTTGTCCGGAGAGGAGATATTTGAAAGAAAAG CAGTGGAACCTGGGGTAAAATTGCTACCTGCATTTTCATACTCCCT CTGGAATACCTGGGCATTGCTGAATATGAAAAGTGGGACGGG CGGAACCTGGCCCTGGGCCTCTGGAGGCAGCAGTATCCTGGCCGA ATTTGGAACCTGCATTTAGAGTTTATGCACCTGTCCCACTTATCA GGAGACCCAGTCTTTGCCGAAAAGGTTATGAAAATTCGAACAGT GTTGAACAACTGGACAAACCAGAAGGCCTTTATCCTAACTATCT GAACCCAGTAGTGGACAGTGGGGTCAACATCATGTGTGCGTTG GAGGACTTGGAGACAGCTTTTATGAATATTTGCTTAAGGCGTGGT TAATGTCTGACAAGACAGATCTCGAAGCCAAGAAGATGTATTTTG ATGCTGTTCAAGCCATCGAGACTCACTTGATCCGCAAGTCAAGTG GGGGACTAACGTACATCGCAGAGTGGAAGGGGGGCTCCTGGAA CACAAGATGGGCCACCTGACGTGCTTTGCAGGAGGCATGTTTGCA CTTGGGGCAGATGGAGCTCCGGAAGCCCGGGCCCAACTACCT TGAACCTCGGAGCTGAAATTGCCCGCACTTGTTCATGAATCTTATAA TCGTACATATGTGAAGTTGGGACCGGAAGCGTTTCGATTTGATGG CGGTGTGGAAGCTATTGCCACGAGGCAAAATGAAAAGTATTACA TCTTACGGCCCGAGGTCATCGAGACATACATGTACATGTGGCGAC TGACTACGACCCCAAGTACAGGACCTGGGCCTGGGAAGCCGTG GAGGCTCTAGAAAGTCACTGCAGAGTGAACGGAGGCTACTCAGG CTTACGGGATGTTTACATTGCCCGTGAGAGTTATGACGATGTCCA
		GCAAAGTTTCTTCCTGGCAGAGACACTGAAGTATTTGTAAGTGGAT ATTTCCGATGATGACCTTCTTCCACTAGAACACTGGATCTTCAA CACCGAGGCTCATCCTTTCCCTATACTCCGTGAACAGAAGAAGGA AATTGATGGCAAAGAGAAATGA
28	dominio catalítico de Mm ManI (FB)	EPADATIREKRAKIKEMMTHAWNMYKRYAWGLNELKPISKEGHSSS LFGNIKGATIVDALDTLIMGMKTEFQEAWSWIKKYLDNFVNAEVS FEVNIRFVGGLLSAYYLSGEEIFRKKAVELGVKLLPAFHTPSGIPWAL LNMKSGIGRNWPWASGGSSILAEFGTLHLEFMHLSHLSGDPVFAEK VMKIRTVLNLKDKPEGLYPNYLNPSSGQWGHVSVGGGLGDSFYE YLLKAWLMSDKTDLEAKKMYFQAVQAIETHLIRKSSGGLTYIAEWK GGLLEHKMGHLTCFAGGMFALGADGAPEARAQHYLELGAELARTC HESYNRTYVKLGPEAFRFDGGVEAIATRQNEKYIILRPEVIETMY MWRLTHDPKYRTWAVEAVEALESHCRVNGGYSLRDVYIARES DDVQQSFFLAETLKYLYLIFSDDDLLPLEHWIFNTEAHPFPILREQK EIDGKEK

29	ADN codifica dominio catalítico de Tr ManI	CGCGCCGGATCTCCCAACCCTACGAGGGCGGCAGCAGTCAAGGC CGCATTCCAGACGTCGTGGAACGCTTACCACCATTTTGCCTTTCC CCATGACGACCTCCACCCGGTCAGCAACAGCTTTGATGATGAGAG AAACGGCTGGGGCTCGTCGGCAATCGATGGCTTGGACACGGCTA TCCTCATGGGGGATGCCGACATTGTGAACACGATCCTTCAGTATG TACCGCAGATCAACTTCACCACGACTGCGGTTGCCAACCAAGGCA TCTCCGTGTTTCGAGACCAACATTTCGGTACCTCGGTGGCCTGCTTT CTGCCTATGACCTGTTGCGAGGTCCTTTCAGCTCCTTGGCGACAA ACCAGACCCTGGTAAACAGCCTTCTGAGGCAGGCTCAAACACTG GCCAACGGCCTCAAGGTTGCGTTCACTACTCCAGCGGTGTCCCG GACCCTACCGTCTTCTTCAACCCTACTGTCCGGAGAAGTGGTGCA TCTAGCAACAACGTCGCTGAAATTGGAAGCCTGGTGCTCGAGTG GACACGGTTGAGCGACCTGACGGGAAACCCGCAAGTATGCCACGC TTGCGCAGAAGGGCGAGTCGTATCTCCTGAATCCAAAGGGAAGC CCGGAGGCATGGCCTGGCCTGATTGGAACGTTTGTACGACAGAG CAACGGTACCTTTTACAGGATAGCAGCGGCAGCTGGTCCGGCCTCAT GGACAGCTTCTACGAGTACCTGATCAAGATGTACCTGTACGACCC GGTTGCGTTTGCACACTACAAGGATCGCTGGGTCTTGTCTGCCGA CTCGACCATTGCGCATCTCGCCTCTCACCCGTCGACGCGCAAGGA CTTGACCTTTTTGTCTTCTGACAAACGGACAGTCTACGTGCGCAAA CTCAGGACATTTGGCCAGTTTGGCCGGTGGCAACTTCATCTTGGG AGGCATTCTCCTGAACGAGCAAAAGTACATTGACTTTGGAATCAA GCTTGCCAGCTCGTACTTTGCCACGTACAACCAGACGGCTTCTGG AATCGGCCCCGAAGGCTTCGCGTGGGTGGACAGCGTGACGGGGCG CCGGCGGCTCGCCGCCCTCGTCCAGTCCGGGTTTCTACTCGTCGG CAGGATTCTGGGTGACGGCACCGTATTACATCCTGCGGCGCGAG ACGCTGGAGAGCTTGTACTACGCATACCGCGTCACGGGCGACTCC AAGTGGCAGGACCTGGCGTGGGAAGCGTTCAGTGCCATTGAGGA CGCATGCCGCGCCGGCAGCGCGTACTCGTCCATCAACGACGTGAC GCAGGCCAACGGCGGGGGTGCCTCTGACGATATGGAGAGCTTCT GGTTTGCCGAGGCGCTCAAGTATGCGTACCTGATCTTTGCGGAGG AGTCGGATGTGCAGGTGCAGGCCAACGGCGGGAACAAATTTGTC TTTAACACGGAGGCGCACCCCTTTAGCATCCGTTTCATCATCACGA CGGGGCGGCCACCTTGCTTAA
30	dominio catalítico de Tr Man I	RAGSPNPTRAAAVKAAFQTSWNAHYHHFAFPHDDLHPVSNSFDDERN GWGSSAIDGLDTAILMGDADIVNTILQYVPQINFNTTAVANQGISVF ETNIRYLGGLLSAYDLLRGPFFSLATNQTLVNSLLRQAQTLANGLKV AFTTPSGVPDPTVFFNPTVRRSGASSNNVAEIGSLVLEWTRLSDLTG NPQYAQLAQKGESYLLNPKGSPEAWPGLIGTFVSTSNGTQDSSGS WSGLMDSFYEYLIKMYLYDPVAFAHYKDRWVLAADSTIAHLASHP STRKDLTFLSSYNGQSTSPNSGHLASFAGGNFILGILLNEQKYIDFGI KLASSYFATYNQTASGIGPEGFAWVDSVTGAGGSPSSQSGFYSSAG FWVTAPYYILRPETLESLYAYRVTGDSKWQDLAWEAFSAIEDACR AGSAYSSINDVTQANGGASDDMESFWFAEALKYAYLIFAESDVQ VQANGGNKFVFENTEHPFSIRSSSRGGHLA

31	ADN codifica Rat GnT II (TC) Optimizado por codón	TCCTTGGTTTACCAATTGAACTTCGACCAGATGTTGAGAAACGTT GACAAGGACGGTACTTGGTCTCCTGGTGAGTTGGTTTTGGTTGTT CAGGTTCAACAACAGACCAGAGTACTTGAGATTGTTGATCGACTCC TTGAGAAAGGCTCAAGGTATCAGAGAGGTTTTGGTTATCTTCTCC CACGATTTCTGGTCTGCTGAGATCAACTCCTTGATCTCCTCCGTTG ACTTCTGTCCAGTTTTGCAGGTTTTCTTCCCATTCTCCATCCAATT GTACCCATCTGAGTTCCCAGGTTCTGATCCAAGAGACTGTCCAAG AGACTTGAAGAAGAACGCTGCTTTGAAGTTGGGTTGTATCAACGC TGAATACCCAGATTCTTTCGGTCACTACAGAGAGGCTAAGTTCTC CCAACTAAGCATCATTGGTGGTGGAAGTTGCACCTTGTTTGGGA GAGAGTTAAGGTTTTGCAGGACTACACTGGATTGATCTTGTTCCTT GGAGGAGGATCATTACTTGGCTCCAGACTTCTACCACGTTTTCAA GAAGATGTGGAAGTTGAAGCAACAAGAGTGTCCAGGTTGTGACG TTTTGTCCTTGGGAACCTTACACTACTATCAGATCCTTCTACGGTAT CGCTGACAAGGTTGACGTTAAGACTTGGAAGTCCACTGAACACA ACATGGGATTGGCTTTGACTAGAGATGCTTACCAGAAGTTGATCG AGTGTACTGACACTTTCTGTACTTACGACGACTACAACCTGGGACT GGACTTTGCAGTACTTGACTTTGGCTTGTTTGCCAAAAGTTTGGGA AGGTTTTGGTTCCACAGGCTCCAAGAATTTCCACGCTGGTGACT GTGGAATGCACCACAAGAAAACCTTGTAGACCATCCACTCAGTCCG CTCAAATTGAGTCCTTGTGTAACAACAACAAGCAGTACTTGTTC CAGAGACTTTGGTTATCGGAGAGAAGTTTCCAATGGCTGCTATTT CCCCACCAAGAAAGAATGGTGGATGGGGTGATATTAGAGACCAC GAGTTGTGTAAATCCTACAGAAGATTGCAGTAG
32	Rat GnTII (TC)	SLVYQLNFDQMLRNVDKDGTSWSPGELVLVVQVHNRPEYLRLLIDSL RKAQGIREVLVIFSHDFWSAEINSLISSVDFCPVLQVFFPFSIQLYPSEF PGSDPRDCPRDLKKNAAALKLGCINAEYPDSFGHYREAKFSQTKHHW WWKLHFVWERVKVLQDYTGILFLIEDHYLAPDFYHVFKKMWKL KQQECPCDVLSTGYTTIRSFYGIADKVDVKTWKSTEHNMGALALT RDAYQKLIETDTFCTYDDYNWDWTLQYLTLACLPKVWKVLVPQA PRIFHAGDCGMHHKKTCPSTQSAQIESLLNNNKQYLPETLVIGE KFPMAAISPPRKNGGWDIRDHELCKSYRRLO
33	ADN codifica ManII de <i>Drosophila</i> <i>melanogaster</i> optimizado por codón (KD)	AGAGACGATCCAATTAGACCTCCATTGAAGGTTGCTAGATCCCCA AGACCAGGTCAATGTCAAGATGTTGTTTCAGGACGTCCCAAACGTT GATGTCCAGATGTTGGAGTTGTACGATAGAATGTCCTTCAAGGAC ATTGATGGTGGTGTGTTGGAAGCAGGGTTGGAACATTAAGTACGA TCCATTGAAGTACAACGCTCATCACAAGTTGAAGGTTCTCGTTGT CCCACACTCCCACAACGATCCTGGTTGGATTGAGACCTTCGAGGA

		ATACTACCAGCACGACACCAAGCACATCTTGTCCAACGCTTTGAG ACATTTGCACGACAACCCAGAGATGAAGTTCATCTGGGCTGAAAT CTCCTACTTCGCTAGATTCTACCACGATTTGGGTGAGAACAAGAA GTTGCAGATGAAGTCCATCGTCAAGAACGGTCAGTTGGAATTCGT CACTGGTGGATGGGTTCATGCCAGACGAGGCTAACTCCCACTGGA GAAACGTTTTGTGTCAGTTGACCGAAGGTCAAACCTTGGTTGAAGC AATTCATGAACGTCACTCCAACCTGCTTCCTGGGCTATCGATCCAT TCGGACACTCTCCAACCTATGCCATACATTTTGCAGAAAGTCTGGTT TCAAGAATATGTTGATCCAGAGAACCCACTACTCCGTTAAGAAGG AGTTGGCTCAACAGAGACAGTTGGAGTTCTTGTGGAGACAGATCT GGGACAACAAAGGTGACACTGCTTTGTTCAACCCACATGATGCCAT TCTACTCTTACGACATTCCCTCATAACCTGTGGTCCAGATCCAAAGG TTTGTTGTGAGTTGATTTCAAAGAATGGGTTCCCTTCGGTTTGTCT TTGTCCATGGAAGGTTCCACCTAGAACTATCTCTGATCAAAATGT TGCTGCTAGATCCGATTTGTTGGTTGATCAGTGGAAGAAGAAGGC TGAGTTGTACAGAACCAACGCTCTTGTGATTCCATTGGGTGACGA CTTCAGATTCAAGCAGAACACCGAGTGGGATGTTTCAGAGAGTCA ACTACGAAAGATTGTTGGAACACATCAACTCTCAGGCTCACTTCA ATGTCCAGGCTCAGTTCCGGTACTTTGCAGGAATACTTCGATGCTG TTCACCAGGCTGAAAGAGCTGGACAAGCTGAGTTCCCAACCTTGT CTGGTGACTTCTTCACTTACGCTGATAGATCTGATAACTACTGGT CTGGTTACTACACTTCCAGACCATAACCATAAGAGAATGGACAGAG TCTTGATGCACTACGTTAGAGCTGCTGAAATGTTGTCCGCTTGGC ACTCCTGGGACGGTATGGCTAGAATCGAGGAAAGATTGGAGCAG GCTAGAAGAGAGTTGTCTTGTTCAGCACCACGACGGTATTACT GGTACTGCTAAAACTCACGTTGTCTGCTGACTACGAGCAAAGAATG CAGGAAGCTTTGAAAGCTTGTCAAATGGTCATGCAACAGTCTGTC TACAGATTGTTGACTAAGCCATCCATCTACTCTCCAGACTTCTCCT TCTCCTACTTCACTTTGGACGACTCCAGATGGCCAGGTTCTGGTG TTGAGGACTCTAGAATAACCATCATCTTGGGTGAGGATATCTTGC CATCCAAGCATGTTGTATGCACAACACCTTGCCACACTGGAGAG AGCAGTTGGTTGACTTCTACGTCTCCTCTCCATTCTGTTTCTGTTAC CGACTTGGCTAACAATCCAGTTGAGGCTCAGGTTTCTCCAGTTTG GTCTTGGCACCACGACACTTTGACTAAGACTATCCACCCACAAGG TTCCACCACCAAGTACAGAATCATCTTCAAGGCTAGAGTTCCACC AATGGGTTTGGCTACCTACGTTTGGACCATCTCCGATTCCAAGCC AGAGCACACCTCCTACGTTCCAATTTGTTGCTTAGAAAGAACCC AACTTCCTTGCCATTGGGTCAATACCCAGAGGATGTCAAGTTTCGG TGATCCAAGAGAGATCTCCTTGAGAGTTGGTAACGGTCCAACCTT GGCTTTCTCTGAGCAGGGTTTGTGGAAGTCCATTCAAGTTGACTCA GGATTCTCCACATGTTCCAGTTCACCTCAAGTTCTTGAAGTACGG TGTTAGATCTCATGGTGATAGATCTGGTGCTTACTTGTCTTGCC AAATGGTCCAGCTTCTCCAGTCGAGTTGGGTGAGCAGTTGCTT GGTCATAAGGGTAAATTGGAGTCTTCCGTTTCTGTTGGTTTGCC ATCTGTCGTTACACAGACCATCATGAGAGGTGGTGCTCCAGAGAT TAGAAATTTGGTCGATATTGGTTCTTTGGACAACACTGAGATCGT CATGAGATTGGAGACTCATATCGACTCTGGTGATATCTTCTACAC TGATTTGAATGGATTGCAATTCATCAAGAGGAGAAGATTGGACA AGTTGCCATTGCAGGCTAACTACTACCCAATTCCATCTGGTATGT
--	--	--

		TCATTGAGGATGCTAATACCAGATTGACTTTGTTGACCGGTCAAC CATTGGGTGGATCTTCTTTGGCTTCTGGTGAGTTGGAGATTATGC AAGATAGAAGATTGGCTTCTGATGATGAAAGAGGTTTGGGTCAG GGTGT TTTGGACAACAAGCCAGTTTTGCATATTTACAGATTGGTC TTGGAGAAGGTTAACAACCTGTGTCAGACCATCTAAGTTGCATCCA GCTGGTTACTTGACTTCTGCTGCTCACAAAGCTTCTCAGTCTTTGT TGGATCCATTGGACAAGTTCATCTTCGCTGAAAATGAGTGGATCG GTGCTCAGGGTCAATTCGGTGGTGATCATCCATCTGCTAGAGAGG ATTTGGATGTCTCTGTCATGAGAAGATTGACCAAGTCTTCTGCTA AAACCCAGAGAGTTGGTTACGTTTTGCACAGAACCAATTTGATGC AATGTGGTACTCCAGAGGAGCATACTCAGAAAGTTGGATGTCTGTC ACTTGTGCCAAATGTTGCTAGATGTGAGAGAACTACCTTGACTT TCTTGCAGAAATTTGGAGCACTTGGATGGTATGGTTGCTCCAGAAG TTTGTCCAATGGAAACCGCTGCTTACGTCTCTTCTCACTCTTCTTG A
34	dominio catalítico de ManII de <i>Drosophila melanogaster</i> (KD)	RDDPIRPPLKVARSPRPGQCQDVVQDVPNVQVQMLELYDRMSFKDI DGGVWKQGWNIKYDPLKYNAHHKLKVFPVPHSHNDPGWIQTFFEE YYQHDTKHILSNALRHLHDNPEMKFIWAEISYFARFYHDLGENKKL QMKSIKNGQLEFVTGGWVMPDEANSHWRNVLLQLTEGQTLWLKQ FMNVTPTASWAIDPFGHSPTMPYILQKSGFKNMLIQRTHYSVKKELA QQRQLEFLWRQIWDNKGDTALFTHMMPFYSDIPHTCGPDPKVCC QFDFKRMGSFGLSCPWKVPPRTISDQNVAAARSDDLVDQWKKKAELY RTNVLLIPLGDDFRFKQNTIEWDVQRVNYERLFEHINSQAHFNVQAQ FGTLQEYFDAVHQAERAGQAEFPTLSGDFFTYADRSDNYWSGYYS RPYHKRMDRVLHMYVRAAEMLSAWHSWDGMARIEERLEQARREL SLFQHHDGITGTAKTHVVVDYEQRMQEALKACQMVMQSQSVYRLL TKPSIYSPDFSFSYFTLDDSRWPGSGVEDSRTTIILGEDILPSKHVVMH NTLPHWREQLVDFYVSSPFVSVTDLANNPVEAQVSPVWSWHHDTL TKTIHPQGSTTKYRIIFKARVPPMGLATYVLTISDSKPEHTSYASNLL LRKNPTSLPLGQYPEDVKFGDPREISLRVGNPGLAFSEQGLLKSQIL TQDSPHVPVHFVKFLKYGVRSRGDRSGAYLFLPNGPASPVELGQPVV LVTKGKLESSVSVGLPSVVHQTIMRGGAPEIRNLVDIGSLDNTEIVM RLETHIDSGDIFYTDLNGLQFIKRRRLDKLPLQANYPIPSGMFIEDA NTRLTLLTGQPLGGSSLASGELEIMQDRRLASDDERGLQGQVLDNK PVLHIYRLVLEKVNNCVRPSKLHPAGYLTSAAHKASQSLDPLDKFI FAENEWIGAQQQFGGDHPSAREDLDVSVMRRLTKSSAKTQRVGYV LHRTNLMQCGTPEEHTQKLDVCHLLPNVARCERTTLTFLQNLHLD GMVAPEVCPMETAAYVSSHSS
35	Transportador de CMP-ácido siálico de ratón (MmCST) Optimizado por codón	ATGGCTCCAGCTAGAGAAAACGTTTCCTTGTCTTCAAGTTGTAC TGTTTGGCTGTTATGACTTTGGTTGCTGCTGCTTACACTGTTGCTT TGAGATACACTAGAACTACTGCTGAGGAGTTGTACTTCTCCACTA CTGCTGTTTGTATCACTGAGGTTATCAAGTTGTTGATCTCCGTTG GTTTGTGCTAAGGAGACTGGTTCTTTGGGAAGATTCAAGGCTT CCTTGTCGAAAACGTTTTGGGTTCCCCAAAGGAGTTGGCTAAGT TGTCTGTTCCATCCTTGGTTTACGCTGTTTCAGAACAAACATGGCTTT CTTGGCTTTGTCTAACTTGGACGCTGCTGTTTACCAAGTTACTTAC CAGTTGAAGATCCCATGTACTGCTTTGTGTACTGTTTTGATGTTG AACAGAACATTGTCCAAGTTGCAGTGGATCTCCGTTTTCATGTTG TGTGGTGGTGTACTTTGGTTCAGTGGAAGCCAGCTCAAGCTTCC

		AAAGTTGTTGTTGCTCAGAACCCATTGTTGGGTTTCGGTGCTATT GCTATCGCTGTTTTGTGTTCCGGTTTCGCTGGTGTCTTACTTCGAGA AGGTTTTGAAGTCCTCCGACACTTCTTTGTGGGTTAGAAACATCC AGATGTACTTGTCCGGTATCGTTGTTACTTTGGCTGGTACTTACTT GTCTGACGGTGCTGAGATTCAAGAGAAGGGATTCTTCTACGGTTA CACTTACTATGTTTTGGTTCGTTATCTTCTTGGCTTCCGTTGGTGGT TTGTACACTTCCGTTGTTGTTAAGTACACTGACAACATCATGAAG GGATTCTCTGCTGCTGCTGCTATTGTTTTGTCCACTATCGCTTCCG TTTTGTTGTTCCGATTGCAGATCACATTGTCCTTTGCTTTGGGAGC TTTGTGGTTTTGTGTTTCCATCTACTTGTACGGATTGCCAAGACAA GACACTACTTCCATTACAGCAAGAGGCTACTTCCAAGGAGAGAATC ATCGGTGTTTAGTAG
36	transportador de CMP-ácido siálico de ratón (MmCST)	MAPARENVSLFFKLYCLAVMTLVAAAYTVALRYTRTTAEELYFSTT AVCITEVIKLLISVGLLAKETGSLGRFKASLSENVLGSPKELAKLSVPS LVYAVQNNMAFLALSNLDAAVYQVTYQLKIPCTALCTVLMNLRTL SKLQWISVFMLCGGVTLVQWKPQAQASKVVVAQNPLLGFGAIAI AVL CSGFAGVYFEKVLKSSDTSLWVRNIQMYLSGIVVTLAGTYLSDGAEI QEKGFFYGYTTYVWFVIFLASVGGLYTSVVVKYTDNIMKGFSAAAA IVLSTIASVLLFGLQITLSFALGALLVCVSIYLYGLPRQDTSIQQEAT SKERIIGV
37	2-epimerasa/ N- acetilmanosamina quinasa Humana (HsGNE) de UDP- GlcNAc optimizada por codón	ATGGA AAAAGAACGGTAACAACAGAAAGTTGAGAGTTTGTGTTGC TACTTGTAACAGAGCTGACTACTCCAAGTTGGCTCCAATCATGTT CGGTATCAAGACTGAGCCAGAGTTCTTCGAGTTGGACGTTGTTGT TTTGGGTTCCCACTTGATTGATGACTACGGTAACACTTACAGAAT GATCGAGCAGGACGACTTCGACATCAACACTAGATTGCACACTAT TGTTAGAGGAGAGGACGAAGCTGCTATGGTTGAATCTGTTGGATT GGCTTTGGTTAAGTTGCCAGACGTTTTGAACAGATTGAAGCCAGA CATCATGATTGTTACGGTGACAGATTTCGATGCTTTGGCTTTGGC TACTTCCGCTGCTTTGATGAACATTAGAATCTTGCACATCGAGGG TGGTGAAGTTTCTGGTACTATCGACGACTCCATCAGACACGCTAT CACTAAGTTGGCTCACTACCATGTTTGTGTTACTAGATCCGCTGA GCAACACTTGATTTCATGTGTGAGGACCACGACAGAATTTTGT GGCTGGTTGTCCATCTTACGACAAGTTGTTGTCCGCTAAGAACAA GGACTACATGTCCATCATCAGAATGTGGTTGGGTGACGACGTTAA GTCTAAGGACTACATCGTTGCTTTGCAGCACCCAGTTACTACTGA CATCAAGCACTCCATCAAGATGTTTCGAGTTGACTTTGGACGCTTT GATCTCCTTCAACAAGAGAACTTTGGTTTTGTTCCCAAACATTGA CGCTGGTTCCAAAGAGATGGTTAGAGTTATGAGAAAGAAGGGTA TCGAACACCACCCAACTTCAGAGCTGTTAAGCACGTTCCATTTCG ACCAATTCATCCAGTTGGTTGCTCATGCTGGTTGTATGATCGGTA ACTCCTCCTGTGGTGTTAGAGAAGTTGGTGCTTTCCGTTACTCCAG TTATCAACTTGGGTACTAGACAGATCGGTAGAGAGACTGGAGAA AACGTTTTGCATGTTAGAGATGCTGACACTCAGGACAAGATTTTG CAGGCTTTGCACTTGCAATTCGGAAGCAGTACCCATGTTCCAAA ATCTACGGTGACGGTAACGCTGTTCCAAGAATCTTGAAGTTTTTG AAGTCCATCGACTTGCAAGAGCCATTGCAGAAGAAGTTCTGTTTC CCACCAGTTAAGGAGAACATCTCCAGGACATTGACCACATCTTG GAGACATTGTCCGCTTTGGCTGTTGATTTGGGTGGAACCTTGA AGAGTTGCTATCGTTTCCATGAAGGGAGAGATCGTTAAGAAGTAC

		ACTCAGTTCAACCCAAAGACTTACGAGGAGAGAATCAACTTGATC TTGCAGATGTGTGTTGAAGCTGCTGCTGAGGCTGTTAAGTTGAAC TGTAGAATCTTGGGTGTTGGTATCTCTACTGGTGGTAGAGTTAAT CCAAGAGAGGGTATCGTTTTGCACTCCACTAAGTTGATTGAGGAG TGGAACCTCGTTGATTTGAGAACTCCATTGTCCGACACATTGCAC TTGCCAGTTTGGGTGACAACGACGGTAATTGTGCTGCTTTGGCT GAGAGAAAGTTCGGTCAAGGAAAGGGATTGGAGAACTTCGTTAC TTTGATCACTGGTACTGGTATTGGTGGTGGTATCATTACCCAGCA CGAGTTGATTACGGTTCTTCTTCTGTGCTGCTGAATTGGGACA CTTGGTTGTTTCTTTGGACGGTCCAGACTGTTCTTGTGGTTCCAC GGTTGTATTGAAGCTTACGCATCAGGAATGGCATTGCAGAGAGA GGCTAAGAAGTTGCACGACGAGGACTTGTGTTGTTGAGGGAA TGTCTGTTCCAAAGGACGAGGCTGTTGGTGGTTCATTGATCC AGGCTGCTAAGTTGGGTAATGCTAAGGCTCAGTCCATCTTGAGAA CTGCTGGTACTGCTTTGGGATTGGGTGTTGTTAATATCTTGCACA CTATGAACCCATCCTTGGTTATCTTGTCCGGTGTGTTGGCTTCTCA CTACATCCACATCGTTAAGGACGTTATCAGACAGCAAGCTTTGTC CTCCGTTCAAGACGTTGATGTTGTTGTTTCCGACTTGGTTGACCC AGCTTTGTTGGGTGCTGCTTCCATGGTTTTGGACTACACTACTAG AAGAATCTACTAATAG
38	2-epimerasa/ N-acetilmanosamina quinasa UDP-GlcNAc Humana (HsGNE)	MEKNGNNRKL RVCVATCNRADYSKLAPIMFGIKTEPEFFELD VVVL GSHLIDDYGN TYRMIEQDDFDINTRLHTIVRGEDEAAMVESVGLAL VKLPDVLNRLKPDIMIVHGDRFDALALATSAALMNIRILHIEGGEVS GTIDDSIRHAITKLAHYHVCCTRSAEQHLISMCEHDHRI LLAGCP SYD KLLSAKNKDYMSIIRMWLGDDVKSKDYIVALQHPVTTDIKHSIKMF ELTLDALISFNKRTLVLFPNIDAGSKEMVRVMRKKGIEHHPNFRAVK HVPFDQFIQLVAHAGCMIGNSSCGVREVGAFGTPVINLGTROIGRET GENVLHVRDADTQDKILQALHLQFGKQYPCSKIYGDGNAVPRILKF LKSIDLQEPLQKKFCFPPVKENISQDIDHILETLSALAVDLGGTNLRV AIVSMKGEIVKKYTQFNPKTYEERINLILQMCVEAAAEAVKLNCRIL GVGISTGGRVNPREGIVLHSTKLIQEWNSVDLRTPLSDTLHLPVWVD NDGNCAALAERKFGQGKLENFVTLITGTGIGGGIIHQHELIHGSSFC AAELGHLVVSLDGPDCSCGSHGCIEAYASGMALQREAKKLHDEDL LVEGMSVPKDEAVGALHLIQA AKLGNAKAQSILRTAGTALGLGVVN ILHTMNP SLVILSGVLASHYIHVKDVIRQQALSSVQD VDVVVS DLV DPALLGAASMVL DYTTTRIY
39	CMP-ácido siálico sintasa Humana (HsCSS) optimizada por codón	ATGGACTCTGTTGAAAAGGGTGCTGCTACTTCTGTTTCCAACCCA AGAGGTAGACCATCCAGAGGTAGACCTCCTAAGTTGCAGAGAAA CTCCAGAGGTGGTCAAGGTAGAGGTGTTGAAAAGCCACCACACT TGGCTGCTTTGATCTTGGCTAGAGGAGGTTCTAAGGGTATCCCAT TGAAGAACATCAAGCACTTGGCTGGTGTTCATTGATTGGATGGG TTTTGAGAGCTGCTTTGGACTCTGGTGCTTTCCAATCTGTTTGGGT TTCCACTGACCACGACGAGATTGAGAACGTTGCTAAGCAATTCGG TGCTCAGGTTACAGAAAGATCCTCTGAGGTTTCCAAGGACTCTTC TACTTCCTTGGACGCTATCATCGAGTTCTTGAAC TACCACAACGA GGTTGACATCGTTGGTAACATCCAAGCTACTTCCCCATGTTTGCA CCCAACTGACTTGCAAAAAGTTGCTGAGATGATCAGAGAAGAGG GTTACGACTCCGTTTTCTCCGTTGTTAGAAGGCACCAGTTCAGAT GGTCCGAGATTCAGAAGGGTGTTAGAGAGGTTACAGAGCCATTG

		AACTTGAACCCAGCTAAAAGACCAAGAAGGCAGGATTGGGACGG TGAATTGTACGAAAACGGTTCCTTCTACTTCGCTAAGAGACACTT GATCGAGATGGGATACTTGCAAGGTGGAAAGATGGCTTACTACG AGATGAGAGCTGAACACTCCGTTGACATCGACGTTGATATCGACT GGCCAATTGCTGAGCAGAGAGTTTTGAGATACGGTACTTCGGAA AGGAGAAGTTGAAGGAGATCAAGTTGTTGGTTTGTAAACATCGAC GGTTGTTTGACTAACGGTCACATCTACGTTTCTGGTGACCAGAAG GAGATTATCTCCTACGACGTTAAGGACGCTATTGGTATCTCCTTG TTGAAGAAGTCCGGTATCGAAGTTAGATTGATCTCCGAGAGAGCT TGTTCCAAGCAAACATTGTCCTCTTTGAAGTTGGACTGTAAGATG GAGGTTTCCGTTTCTGACAAGTTGGCTGTTGTTGACGAATGGAGA AAGGAGATGGGTTTGTGTTGGAAGGAAGTTGCTTACTTGGGTAA CGAAGTTTCTGACGAGGAGTGTTTGAAGAGAGTTGGTTTGTCTGG TGCTCCAGCTGATGCTTGTTCCACTGCTCAAAGGCTGTTGGTTA CATCTGTAAGTGTAACGGTGGTAGAGGTGCTATTAGAGAGTTCGC TGAGCACATCTGTTTGTGATGGAGAAAGTTAATAACTCCTGTCA GAAGTAGTAG
40	CMP-ácido siálico sintasa Humana (HsCSS)	MDSVEKGAATSVSNPRGRPSRGRPPKLQRNSRGGQGRGVEKPPHLA ALILARGGSKGIPLKNIKHLAGVPLIGWVLRAALDSGAFQSVWVSTD HDEIENVAKQFGAQVHRRSSEVSKDSSTSLDAIIEFLNYHNEVDIVG NIQATSPCLHPTDLQKVAEMIREEGYDSVFSVRRHQFRWSEIQKGV REVTEPLNLPKRPRRQDWDGELYENGFSFYFAKRHLIEMGYLQGG KMAYYEMRAEHSVDIDVDIDWPIDAEQRLRYGYFGKEKLKEIKLLV CNIDGCLTNGHIYVSGDQKEIISYDVKDAIGISLLKKSIEVRLISERA CSKQTLSSLKLDCKMEVSVSDKLAVVDEWRKEMGLCWKEVAYLG NEVSDEECLKRVGLSGAPADACSTAQKAVGYICKNGGRGAIREFA EHICLLMEKVNNSCQK
41	N-acetilneuraminato- 9-fosfato sintasa Humana (HsSPS) optimizada por codón	ATGCCATTGGAATTGGAGTTGTGTCCTGGTAGATGGGTTGGTGGT CAACACCCATGTTTCATCATCGCTGAGATCGGTCAAACACCAAA GGAGACTTGGACGTTGCTAAGAGAATGATCAGAATGGCTAAGGA ATGTGGTGCTGACTGTGCTAAGTTCCAGAAGTCCGAGTTGGAGTT CAAGTTCAACAGAAAGGCTTTGGAAAGACCATACTTCCAAGC ACTCTTGGGGAAAGACTTACGGAGAACACAAGAGACACTTGGAG TTCTCTCACGACCAATACAGAGAGTTGCAGAGATACGCTGAGGA AGTTGGTATCTTCTTCACTGCTTCTGGAATGGACGAAATGGCTGT TGAGTTCTTGACGAGTTGAACGTTCCATTCTTCAAAGTTGGTTC CGGTGACACTAACAACCTTCCCATACTTGGAAAAGACTGCTAAGAA AGGTAGACCAATGGTTATCTCCTCTGGAATGCAGTCTATGGACAC TATGAAGCAGGTTTACCAGATCGTTAAGCCATTGAACCCAACTT TTGTTTCTTGCAGTGTACTTCCGCTTACCCATTGCAACCAGAGGA CGTTAATTTGAGAGTTATCTCCGAGTACCAGAAGTTGTTCCAGA CATCCCAATTGGTTACTCTGGTCACGAGACTGGTATTGCTATTTT CGTTGCTGCTGTTGCTTTGGGTGCTAAGGTTTTGGAGAGACACAT CACTTTGGACAAGACTTGAAGGGTTCTGATCACTCTGCTTCTTT GGAACCTGGTGAGTTGGCTGAAGTTGTTAGATCAGTTAGATTGGT TGAGAGAGCTTTGGGTTCCCCAACTAAGCAATTGTTGCCATGTGA GATGGCTTGTAACGAGAAGTTGGGAAAGTCCGTTGTTGCTAAGG TTAAGATCCCAGAGGGTACTATCTTGACTATGGACATGTTGACTG TTAAAGTTGGAGAGCCAAAGGGTTACCCACCAGAGGACATCTTTA
		ACTTGGTTGGTAAAAAGGTTTTGGTTACTGTTGAGGAGGACGACA CTATTATGGAGGAGTTGGTTGACAACCACGGAAAGAAGATCAAG TCCTAG

42	N-acetilneuraminato-9-fosfato sintasa Humana (HsSPS)	MPLELELCPRWVGGQHPCFIIAEIGQNHQGDLDVAKRMIRMAKEC GADCAKFQKSELEFKFNRKALERPYTSKHSWGKTYGEHKRHLEFSH DQYRELQRYAEEVGIFFTASGMDEMAVEFLHELNVPPFKVGSBDTN NFPYLEKTAKKGRPMVISSGMQSMMDTMKQVYQIVKPLNPFCFLQC TSAYPLQPEDVNLRVISEYQKLFPDIPIGYSGHETGIAISVAVALGA KVLERHITLDKTWKGS DHSASLEPGELAE LVR SVRLVERALGSPTKQ LLPCEMACNEKLGKSVVAKVKIPEGTILTM DMLTVKVGE PKGY PPE DIFNLVGKKVLVTVEEDDTIMEELVDNHGKKIKS
43	dominio catalítico de alfa-2,6-sialil transferasa de ratón (MmmST6) optimizado por codón	GTTTTTCAAATGCCAAAGTCCCAGGAGAAAGTTGCTGTTGGTCCA GCTCCACAAGCTGTTTTCTCCAACCTCCAAGCAAGATCCAAAGGAG GGTGTTCAAATCCTTGTCCTACCCAAGAGTTACTGCTAAGGTAAAG CCACAACCATCCTTGCAAGTTTGGGACAAGGACTCCACTTACTCC AAGTTGAACCCAAGATTGTTGAAGATTTGGAGAACTACTTGAAC ATGAACAAGTACAAGGTTTCTACAAGGGTCCAGGTCCAGGTGTT AAGTTCTCCGTTGAGGCTTTGAGATGTCACTTGAGAGACCACGTT AACGTTTCCATGATCGAGGCTACTGACTTCCCATTCAACACTACT GAATGGGAGGGATACTTGCCAAAGGAGAACTTCAGAACTAAGGC TGGTCCATGGCATAAGTGTGCTGTTGTTTCTTCTGCTGGTTCCTTG AAGAACTCCCAGTTGGGTAGAGAAATTGACAACCACGACGCTGT TTTGAGATTCAACGGTGCTCCAACCTGACAACCTCCAGCAGGATGT TGGTACTAAGACTACTATCAGATTGGTTAACTCCCAATTGGTTAC TACTGAGAAGAGATTCTTGAAGGACTCCTTGTACACTGAGGGAAT CTTGATTTTGTGGGACCCATCTGTTTACCACGCTGACATTCCACA ATGGTATCAGAAGCCAGACTACAACCTTCTTCGAGACTTACAAGTC CTACAGAAGATTGCACCCATCCCAGCCATTCTACATCTTGAAGCC ACAAATGCCATGGGAATTGTGGGACATCATCCAGGAAATTTCCCC AGACTTGATCCAACCAAACCCACCATCTTCTGGAATGTTGGGTAT CATCATCATGATGACTTTGTGTGACCAGGTTGACATCTACGAGTT CTTGCCATCCAAGAGAAAGACTGATGTTTGTACTACCACCAGAA GTTCTTCGACTCCGCTTGTACTATGGGAGCTTACCACCCATTGTT GTTGAGAGAAGAACATGGTTAAGCACTTGAACGAAGGTACTGACG AGGACATCTACTTGTTCGAAAGGCTACTTTGTCCGGTTTCAGAA ACAACAGATGTTAG
44	dominio catalítico de alfa-2,6-sialil transferasa de ratón (MmmST6)	VFQMPKSQEKVAVGPAPQAVFSNSKQDPKEGVQILSYPRVTAKVKP QPSLQVWDKDYSTYSKLNPRLKIWRNYLNMNKYKVSYKGP GPGVK FSVEALRCHLRDHVNVSMIEATDFPFNTTEWEGYLPKENFRTKAGP WHKCAVVSSAGSLKNSQLGREIDNHDAVLRFN GAPTDFNQDVG KTTIRLVNSQLVTTEKRFLKDSLYTEGILILWDPSVYHADIPQWYQK PDYNFFETYKS YRRLHPSQPFYILKPQMPWELWDIIQEISPDLIQPNPPSSGMLGIIMM TLCDQVDIYEFLPSKRKTDCVYHQKFFDSACTMGAYHPLLEKNNM VKHLNEGTDEDIYLF GKATLSGFRNNRC
45	Secuencia del promotor de PpPMA1:	AAATGCGTACCTCTTCTACGAGATTCAAGCGAATGAGAATAATGT AATATGCAAGATCAGAAAGAATGAAAGGAGTTGAAAAAAAAAAC CGTTGCGTTTTTGACCTTGAATGGGGTGGAGGTTTCCATTCAAAGT AAAGCCTGTGTCTTGGTATTTTCGGCGGCACAAGAAATCGTAATT

		TTCATCTTCTAAACGATGAAGATCGCAGCCCAACCTGTATGTAGT TAACCGGTCGGAATTATAAGAAAGATTTTCGATCAACAAACCCTA GCAAATAGAAAGCAGGGTTACAACCTTTAAACCGAAGTCACAAAC GATAAACCACTCAGCTCCACCCAAATTCATTCCCACTAGCAGAA AGGAATTATTTAATCCCTCAGGAAACCTCGATGATTCTCCCGTTC TTCCATGGGCGGGTATCGCAAAATGAGGAATTTTCAAATTTCTC TATTGTCAAGACTGTTTATTATCTAAGAAATAGCCCAATCCGAAG CTCAGTTTTGAAAAAATCACTTCCGCGTTTCTTTTTTACAGCCCGA TGAATATCCAAATTTGGAATATGGATTACTCTATCGGGACTGCAG ATAATATGACAACAACGCAGATTACATTTTAGGTAAGGCATAAAC ACCAGCCAGAAATGAAACGCCCACTAGCCATGGTCAATAGTCC AATGAATTCAGATAGCTATGGTCTAAAAGCTGATGTTTTTTATTG GGTAATGGCGAAGAGTCCAGTACGACTTCCAGCAGAGCTGAGAT GGCCATTTTTGGGGGTATTAGTAACTTTTTGAGCTCTTTTCACTTC GATGAAGTGTCCCATTCGGGATATAATCGGATCGCGTCGTTTTCT CGAAAAATACAGCTTAGCGTCGTCCGCTTGTTGTAAGCAGCACC ACATTCCTAATCTCTTATATAAAACAAAACAACCCAAATTATCAGT GCTGTTTTCCACCAGATATAAGTTTCTTTTCTTCCGCTTTTTG ATTTTTATCTCTTTCCTTTAAAAACTTCTTTACCTTAAAGGGCGG CC
46	Secuencia del terminador de PpPMA1:	TAAGCTTCACGATTTGTGTTCCAGTTTATCCCCCTTTATATACCG TTAACCTTTCCCTGTTGAGCTGACTGTTGTTGTATTACCGCAATT TTTCCAAGTTTGCCATGCTTTTCGTGTTATTTGACCGATGTCTTTT TTCCCAAATCAAATATATTTGTTACCATTTAAACCAAGTTATCTT TTGTATTAAGAGTCTAAGTTTGTTCACAGGCTTCATGTGAGAGTG ATAACCATCCAGACTATGATTCTTGTTTTTATTGGGTTTGTGTTGT GTGATACATCTGAGTTGTGATTCTGTAAGTATGTCAGTCTATCTA GATTTTTAATAGTTAATTGGTAATCAATGACTTGTTTGTTTAACT TTTAAATTGTGGGTCGTATCCACGCGTTTAGTATAGCTGTTTCATG GCTGTTAGAGGAGGGCGATGTTTATATACAGAGGACAAGAATGA GGAGGCGGCGTGATTTTTTAAAATGGAGACGCGACTCCTGTACAC CTTATCGGTTGG
47	Secuencia del promotor de PpOCH1:	TGGACACAGGAGACTCAGAAACAGACACAGAGCGTTCTGAGTCC TGGTGCTCCTGACGTAGGCCTAGAACAGGAATTATTGGCTTTATT TGTTTGTCATTTTCATAGGCTTGGGGTAATAGATAGATGACAGAG AAATAGAGAAGACCTAATATTTTTTGTTCATGGCAAATCGCGGGT TCGCGGTCGGGTCACACACGGAGAAGTAATGAGAAGAGCTGGTA ATCTGGGGTAAAAGGGTTCAAAAAGAAGGTCGCTGGTAGGGATG CAATACAAGGTTGTCTTGGAGTTTACATTGACCAGATGATTTGGC TTTTTCTCTGTTCAATTCACATTTTTCAGCGAGAATCGGATTGACG GAGAAATGGCGGGGTGTGGGGTGGATAGATGGCAGAAATGCTCG CAATCACCGCGAAAGAAAGACTTTATGGAATAGAATACTGGGT GGTGTAAGGATTACATAGCTAGTCCAATGGAGTCCGTTGGAAAG GTAAGAAGAAGCTAAAACCGGCTAAGTAAGTGGGAAGAATGAT CAGACTTTGATTTGATGAGGTCTGAAAATACTCTGCTGCTTTTTTC AGTTGCTTTTTCCCTGCAACCTATCATTTTCTTTTCATAAGCCTG CCTTTTCTGTTTTCACTTATATGAGTCCGCGGAGACTTCCCCAAA TTCTCTCCTGGAACATTCTCTATCGCTCTCCTTCCAAGTTGCGCCC CCTGGCACTGCCTAGTAATATTACCACGCGACTTATATTCAAGTTC
		CACAATTTCCAGTGTTTCGTAGCAAATATCATCAGCC

48	Secuencia del terminador de PpALG12:	AATATATACCTCATTTGTTCAATTTGGTGTAAGAGTGTGGCGGA TAGACTTCTTGTAATCAGGAAAGCTACAATTCCAATTGCTGCAA AAAATACCAATGCCCATAAACCAGTATGAGCGGTGCCTTCGACG GATTGCTTACTTTCCGACCCTTTGTCGTTTGATTCTTCTGCCTTTG GTGAGTCAGTTTGTTCGACTTTATATCTGACTCATCAACTTCCTT TACGGTTGCGTTTTTAATCATAATTTTAGCCGTTGGCTTATTATCC CTTGAGTTGGTAGGAGTTTTGATGATGCTG
49	Secuencia del promotor de PpSEC4:	GAAGTAAAGTTGGCGAAACTTTGGGAACCTTTGGTTAAACTTTG TAATTTTTGTGCTACCCATTAGGCAGAATCTGCATCTTGGGAGG GGGATGTGGTGGCGTTCTGAGATGTACGCGAAGAATGAAGAGCC AGTGGTAACAACAGGCCTAGAGAGATACGGGCATAATGGGTATA ACCTACAAGTTAAGAATGTAGCAGCCCTGGAAACCAGATTGAAA CGAAAAACGAAATCATTTAACTGTAGGATGTTTTGGCTCATTGT CTGGAAGGCTGGCTGTTTATTGCCCTGTTCTTTGCATGGGAATAA GCTATTATATCCCTCACATAATCCCAGAAAATAGATTGAAGCAAC GCGAAATCCTTACGTATCGAAGTAGCCTTCTTACACATTCACGTT GTACGGATAAGAAAACACTCAACGAACAATC
50	Secuencia del terminador de PpOCH1:	AATAGATATAGCGAGATTAGAGAATGAATACCTTCTTCTAAGCGA TCGTCCGTCATCATAGAATATCATGGACTGTATAGTTTTTTTTTG TACATATAATGATTAAACGGTCATCCAACATCTCGTTGACAGATC TCTCAGTACGCGAAATCCCTGACTATCAAAGCAAGAACCGATGAA GAAAAAACAACAGTAACCCAAACACCACAACAAACACTTTATCT TCTCCCCCCCCAACCAATCATCAAAGAGATGTGCGAACACAAAC ACCAAGAAGCAAAAACCTAACCCCATATAAAAACATCCTGGTAGA TAATGCTGGTAACCCGCTCTCCTTCCATATTCTGGGCTACTTCAG AAGTCTGACCGGTCTCAGTTGATCAACATGATCCTCGAAATGG
51	Secuencia del promotor de PpTEF1	TTAAGGTTTGGAAACAACACTAAACTACCTTGCGGTACTACCATTG ACACTACACATCCTTAATTCCAATCCTGTCTGGCCTCCTTACCTT TTAACCATCTTGCCCATTCCTCAACTCGTGTGAGATTGCGTATCAAG TGAAAAAATAAATTTTAAATCTTTAACCAATCAGGTAATAAC TGTCGCCTCTTTTATCTGCCGCACTGCATGAGGTGTCCCCTTAGT GGGAAAGAGTACTGAGCCAACCCTGGAGGACAGCAAGGGGAAAAA TACCTACAACCTTGCTTCATAATGGTCGTAAAAACAATCCTTGTCG GATATAAGTGTGTAGACTGTCCCTTATCCTCTGCGATGTTCTTCC TCTCAAAGTTTGCGATTCTCTCTATCAGAATTGCCATCAAGAGA CTCAGGACTAATTTGCGAGTCCCACACGCACTCGTACATGATTGG CTGAAATTTCCCTAAAGAATTTCTTTTTTACGAAAATTTTTTTTTT ACACAAGATTTTACGAGATATAAAATGGAGAGCAGGACCTCCG CTGTGACTCTTCTTTTTTTTCTTTTATTCTCACTACATATTTAG TTATTCGCCAAC
52	Secuencia del terminador de PpTEF1:	ATTGCTTGAAGCTTTAATTTATTTTATTAACATAATAATAACAA GCATGATATATTTGTATTTTGTTCGTTAACATTGATGTTTTCTTCA TTTACTGTTATTGTTTGTAACTTTGATCGATTATCTTTTCTACTTT ACTGTAATATGGCTGGCGGGTGAGCCTTGAACCTCCCTGTATTACT TTACCTTGCTATTACTTAATCTATTGACTAGCAGCGACCTCTTCAA
		CCGAAGGGCAAGTACACAGCAAGTTCATGTCTCCGTAAGTGTGAT CAACCCTGGAAACAGTGGGCCATGTC

53	Secuencia del promotor de PpGAPDH:	TTTTGTAGAAATGTCTTGGTGTCTCGTCCAATCAGGTAGCCAT CTCTGAAATATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGC AACGTAAAATTCTCCGGGGTAAAACTTAAATGTGGAGTAATGGA ACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGCCCGTTACC GTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCC CCTTGCAGCAATGCTCTTCCAGCATTACGTTGCGGGTAAAACGG AGGTCGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCG GCCGTCGCTGGCAATAATAGCGGGCGGACGCATGTCATGAGATT ATTGGAACACCAGCAATCGAATATAAAAGGCGAACACCTTTCCC AATTTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTG TCCCTATTTCAATCAATTGAACAACCTATCAAAACACA
54	Secuencia del terminador de PpALG3:	ATTTACAATTAGTAATATTAAGGTGGTAAAAACATTTCGTAGAATT GAAATGAATTAATATAGTATGACAATGGTTCATGTCTATAAATCT CCGGCTTCGGTACCTTCTCCCAATTGAATACATTGTCAAAATGA ATGGTTGAACTATTAGGTTTCGCCAGTTTCGTTATTAAGAAAAGT TTAAATCAAATTCCATATCATCGGTTCCAGTGGGAGGACCAGTT CCATCGCCAAAATCCTGTAAGAATCCATTGTCAGAACCTGTAAAG TCAGTTTGAGATGAAATTTTCCGGTCTTTGTTGACTTGGAAGCT TCGTTAAGGTTAGGTGAAACAGTTTGATCAACCAGCGGCTCCCGT TTTCGTCGCTTAGTAG
55	Secuencia del promotor de PpAOX 1 y locus de integración:	AACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGA CATCCACAGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGA GGGGATACACTAGCAGCAGACCGTTGCAAACGCAGGACCTCCAC TCCTCTTCTCCTCAACACCCACTTTTGCCATCGAAAAACCAGCCC AGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATT AGGCTACTAACACCATGACTTTATTAGCCTGTCTATCCTGGCCCC CCTGGCGAGGTTTCATGTTTGTTATTTCCGAATGCAACAAGCTCC GCATTACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGTG GGGTCAAATAGTTTCATGTTCCCAAAATGGCCCAAACTGACAGT TTAAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCAT CCAAGATGAACTAAGTTTGGTTTCGTTGAAATGCTAACGGCCAGTT GGTCAAAAAGAACTTCCAAAAGTCGGCATAACGTTTGTCTTGT TGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAAATGCTTA GCGCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGA AACGCAATGGGGAAACACCCGCTTTTGGATGATTATGCATTGT CTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGAT AGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTT GACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTT TTTTTTATCATCATTATTAGCTTACTTTCATAATTGCGACTGGTT CCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGACAACCTTG AGAAGATCAAAAACAACCTAATTATTTCGAAACG
56	Secuencia del terminador de ScCYC1:	ACAGGCCCTTTTCTTTGTGCGATATCATGTAATTAGTTATGTCAC GCTTACATTACGCCCTCCTCCACATCCGCTCTAACCGAAAAGG AAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTTT AATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAAATTTT CTTTTTTTCTGTACAAACGCGTGTACGCATGTAACATTATACTGA AAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCTTAAATT GCAAGCTGCCGGCTCTTAAG

57	Secuencia del promotor de ScTEF1:	GATCCCCCACACACCATAGCTTCAAAATGTTTCTACTCCTTTTTTA CTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCA AAACACCCAAGCACAGCATACTAAATTTCCCTCTTTCTTCTCTA GGGTGTCGTTAATTACCCGTAATAAGGTTTGAAAAAGAAAAA GAGACCGCCTCGTTTCTTTTCTTCGTCGAAAAAGGCAATAAAAA TTTTATCACGTTTCTTTTCTTGAAAAATTTTTTTTTTGATTTTT TCTCTTTCGATGACCTCCCATGATATTTAAGTTAATAAACGGTCT TCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAAC TTTTTACTTCTTGCTCATTAGAAAAGAAAGCATAGCAATCTAATCT AAGTTTTAATTACAAA
58	Secuencia del ORF Sh ble (marcador de resistencia a Zeocina):	ATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTACCGCGCGCGAC GTCGCCGGAGCGGTGAGTTCTGGACCGACCGGCTCGGGTTCTCC CGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGA CGTGACCCGTGTTATCAGCGCGGTCCAGGACCAGGTGGTGCCGG ACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGACGAGCTG TACGCCGAGTGGTCCGAGGTGCTGTCCACGAACCTCCGGGACGC CTCCGGGCCCGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGC GGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTC GTGGCCGAGGAGCAGGACTGA
59	Secuencia de la Región en la posición 5' usada para supresión genética de PpURA5:	ATCGGCCTTTGTTGATGCAAGTTTTACGTGGATCATGGACTAAGG AGTTTTATTTGGACCAAGTTCATCGTCCTAGACATTACGGAAAGG GTTCTGCTCCTCTTTTTGGAAACTTTTTGGAACCTCTGAGTATGAC AGCTTGGTGGATTGTACCCATGGTATGGCTTCTGTGAATTTCTA TTTTTCTACATTGGATTACCAATCAAACAAATAGTCGCCAT GGCTTTTTGGCTTTTGGGTCTATTTGTTTGGACCTTCTTGGAATAT GCTTTGCATAGATTTTGTTCCTTGGACTACTATCTTCCAGAGA ATCAAATTGCATTTACCATTCTTTCTTATTGCATGGGATACACCA CTATTTACCAATGGATAAATACAGATTGGTGATGCCACCTACACT TTTCATTGTACTTTGCTACCCAATCAAGACGCTCGTCTTTTCTGTT CTACCATATTACATGGCTTGTCTGGATTGTCAGGTGGATTCTCTG GGCTATATCATGTATGATGTCACTCATTACGTTCTGCATCACTCC AAGCTGCCTCGTTATTTCCAAGAGTTGAAGAAATATCATTTGGAA CATCACTACAAGAATTACGAGTTAGGCTTTGGTGTCACTTCCAAA TTCTGGGACAAAGTCTTTGGGACTTATCTGGGTCCAGACGATGTG TATCAAAAAGACAAATTAGAGTATTTATAAAGTTATGTAAGCAAAT AGGGGCTAATAGGGAAAGAAAAATTTTGGTTCTTTATCAGAGCTG GCTCGCGCGCAGTGTTTTTCTGTGCTCCTTTGTAATAGTCATTTTTG ACTACTGTTTCAAGATTGAAATCACATTGAAGATGTCACTCGAGGGG TACCAAAAAAGGTTTTTGGATGCTGCAGTGGCTTCGC
60	Secuencia de la Región en la posición 3' usada para supresión genética de PpURA5:	GGTCTTTTCAACAAAGCTCCATTAGTGAGTCAGCTGGCTGAATCT

		TATGCACAGGCCATCATTAAACAGCAACCTGGAGATAGACGTTGTA TTTGGACCAGCTTATAAAGGTATTCTTTGGCTGCTATTACCGTG TTGAAGTTGTACGAGCTCGGCGGCAAAAAATACGAAAATGTCGG ATATGCGTTCAATAGAAAAGAAAAGAAAGACCACGGAGAAGGTG GAAGCATCGTTGGAGAAAGTCTAAAGAATAAAAGAGTACTGATT ATCGATGATGTGATGACTGCAGGTACTGCTATCAACGAAGCATT GCTATAATTGGAGCTGAAGGTGGGAGAGTTGAAGGTAGTATTAT TGCCCTAGATAGAATGGAGACTACAGGAGATGACTCAAATACCA GTGCTACCCAGGCTGTTAGTCAGAGATATGGTACCCCTGTCTTGA GTATAGTGACATTGGACCATAATTGTGGCCCATTTGGGCGAAACTT TCACAGCAGACGAGAAAATCTCAAATGGAAACGTATAGAAAAAAG TATTTGCCCAAATAAGTATGAATCTGCTTCGAATGAATGAATTAA TCCAATTATCTTCTCACCATTATTTTCTTCTGTTTCGGAGCTTTGG GCACGGCGGCGGGTGGTGCGGGCTCAGGTTCCCTTTCATAAACA GATTTAGTACTTGGATGCTTAATAGTGAATGGCGAATGCAAAGGA ACAATTTTCGTTTCATCTTTAACCCTTTCCTCGGGGTACACGTTCTG GAATGTACCCGCCCTGTTGCAACTCAGGTGGACCGGGCAATTCTT GAACTTTCTGTAAACGTTGTTGGATGTTCAACCAGAAATTGTCCTA CCAAGTATTAGTTTCTTTTGGTCTTATATTGTTTCATCGAGATA CTTCCCACTCTCCTTGATAGCCACTCTCACTCTTCTGGATTACCA AAATCTTGAGGATGAGTCTTTTCAGGCTCCAGGATGCAAGGTATA TCCAAGTACCTGCAAGCATCTAATATTGTCTTTGCCAGGGGGTTC TCCACACCATACTCTTTTGGCGCATGC
61	Secuencia del marcador auxotrófico de PpURA5:	TCTAGAGGGACTTATCTGGGTCCAGACGATGTGTATCAAAAGACA AATTAGAGTATTTATAAAGTTATGTAAGCAAATAGGGGCTAATAG GGAAAGAAAAATTTTGGTCTTTATCAGAGCTGGCTCGCGCGCAG TGTTTTTCGTGCTCCTTTGTAATAGTCATTTTGTACTACTGTTTCA ATTGAAATCACATTGAAGATGTCACTGGAGGGGTACCAAAAAAG GTTTTTGGATGCTGCAGTGGCTTCGCAGGCCTTGAAGTTTGAAC TTTCACCTTGAAAAGTGGAAGACAGTCTCCATACTTCTTTAACAT GGGTCTTTTCAACAAAGCTCCATTAGTGAGTCAGCTGGCTGAATC TTATGCTCAGGCCATCATTAAACAGCAACCTGGAGATAGACGTTGT ATTTGGACCAGCTTATAAAGGTATTCTTTGGCTGCTATTACCGT GTTGAAGTTGTACGAGCTGGGCGGCAAAAAATACGAAAATGTCTG GATATGCGTTCAATAGAAAAGAAAAGAAAGACCACGGAGAAGGT GGAAGCATCGTTGGAGAAAGTCTAAAGAATAAAAGAGTACTGAT TATCGATGATGTGATGACTGCAGGTACTGCTATCAACGAAGCATT TGCTATAATTGGAGCTGAAGGTGGGAGAGTTGAAGGTTGTATTAT TGCCCTAGATAGAATGGAGACTACAGGAGATGACTCAAATACCA GTGCTACCCAGGCTGTTAGTCAGAGATATGGTACCCCTGTCTTGA GTATAGTGACATTGGACCATAATTGTGGCCCATTTGGGCGAAACTT TCACAGCAGACGAGAAAATCTCAAATGGAAACGTATAGAAAAAAG TATTTGCCCAAATAAGTATGAATCTGCTTCGAATGAATGAATTAA TCCAATTATCTTCTCACCATTATTTTCTTCTGTTTCGGAGCTTTGG GCACGGCGGCGGATCC
62	Secuencia de la parte	CCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCG GTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAAACAGTTGATTG

	del gen Ec lacZ que se usó para la construcción del PpURA5 blaster (marcador auxotrófico reciclable)	AACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGG CTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAGA AGCCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTGGCGGAAA ACCTCAGTGTGACGCTCCCCGCCGCGTCCCACGCCATCCCGCATC TGACCACCAGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGC GTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGA TTGGCGATAAAAAACAACCTGCTGACGCCGCTGCGCGATCAGTTCA CCCGTGACCCGCTGGATAACGACATTGGCGTAAGTGAAGCGACC CGCATTGACCCTAACGCCTGGGTGCGAACGCTGGAAGGCGGCGGG CCATTACCAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATA CACTTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGCAGC ATCAGGGGAAAACCTTATTTATCAGCCGAAAACCTACCGGATTG ATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCG AGCGATACACCGCATCCGGCGCGGATTGGCCTGAACTGCCAG
63	Secuencia de aminoácidos de PpURA5	MSLEGYQKRFLDAAVASQALKFGTFTLKSGRQSPYFFNMGLFNKAP LVSQLAESYAQAIIINSNLEIDVVFGPAYKGIPLAAITVLKLYELGGKK YENVGYAFNRKEKKDHGEGGSIVGESLKNKRVLIIDDVMTAGTAIN EAFAIIGAEGGRVEGCHALDRMETTGDDSNSTATQAVSQRYGTPVL SIVTLDHIVAHLGETFTADEKSOMETYRKKYLPKZ
64	Secuencia de la Región en la posición 5' usada para supresión genética de PpOCH1:	AAAACCTTTTTTCTTATTCAAACACAAGGCATTGCTTCAACACGT GTGCGTATCCTTAACACAGATACTCCATACTTCTAATAATGTGAT AGACGAATACAAAGATGTTCACTCTGTGTTGTGTCTACAAGCATT TCTTATTCTGATTGGGGATATTCTAGTTACAGCACTAAACAACCTG GCGATACAAACTTAAATTAATAATCCGAATCTAGAAAATGAACT TTTGGATGGTCCGCCTGTTGGTTGGATAAATCAATACCGATTAAA TGGATTCTATTCCAATGAGAGAGTAATCCAAGACACTCTGATGTC AATAATCATTTGCTTGCAACAACAAACCCGTCATCTAATCAAAGG GTTTGATGAGGCTTACCTTCAATTGCAGATAAACTCATTGCTGTC CACTGCTGTATTATGTGAGAATATGGGTGATGAATCTGGTCTTCT CCACTCAGCTAACATGGCTGTTTGGGCAAAGGTGGTACAATTATA CGGAGATCAGGCAATAGTGAAATTGTTGAATATGGCTACTGGAC GATGCTTCAAGGATGTACGTCTAGTAGGAGCCGTGGGAAGATTG CTGGCAGAACCAGTTGGCACGTGCGAACAATCCCCAAGAAATGA AATAAGTGAAAACGTAACGTCAAAGACAGCAATGGAGTCAATAT TGATAACACCACTGGCAGAGCGGTTCTGACGTGTTTTGGAGCCG ATATGAGGCTCAGCGTGCTAACAGCACGATTGACAAGAAGACTC TCGAGTGACAGTAGGTTGAGTAAAGTATTTCGCTTAGATTCCCAAC CTTCGTTTTATTCTTTCTGAGACAAAGAAGCTGCATGCGAACATA GGGACAACTTTTATAAATCCAATTGTCAAACCAACGTAAAACCT CTGGCACCATTITCAACATATATTTGTGAAGCAGTACGCAATATC GATAAATACTCACCGTTGTTTGTAAACAGCCCCAACTTGCATACGC CTTCTAATGACCTCAAATGGATAAGCCGCAGCTTGTGCTAACATA CCAGCAGCACCGCCCCGCGGTGAGCTGCGCCACACATATAAAGG CAATCTACGATCATGGGAGGAATTAGTTTTGACCGTCAGGTCTTC AAGAGTTTTGAACTCTTCTTCTTGAAGTGTGTAACCTTTTAAATGA CGGGATCTAAATACGTCATGGATGAGATCATGTGTGTAAAACTG ACTCCAGCATATGGAATCATTCCAAAGATTGTAGGAGCGAACCCA

		CGATAAAAGTTTCCCAACCTTGCCAAAGTGTCTAATGCTGTGACT TGAAATCTGGGTTCTCGTTGAAGACCCTGCGTACTATGCCCAA AACTTTCCTCCACGAGCCCTATTAACCTTCTCTATGAGTTTCAAATG CCAAACGGACACGGATTAGGTCCAATGGGTAAGTGAAAAACACA GAGCAAACCCAGCTAATGAGCCGGCCAGTAACCGTCTTGGAGC TGTTTCATAAGAGTCAATTAGGGATCAATAACGTTCTAATCTGTT ATAACATACAAATTTTATGGCTGCATAGGGAAAAATTCTCAACAG GGTAGCCGAATGACCCTGATATAGACCTGCGACACCATCATACCC ATAGATCTGCCTGACAGCCTTAAAGAGCCCGCTAAAAGACCCGG AAAACCGAGAGAACTCTGGATTAGCAGTCTGAAAAAGAATCTTC ACTCTGTCTAGTGGAGCAATTAATGTCTTAGCGGCCTTCCTGCT ACTCCGCCAGCTACTCCTGAATAGATCACATAGCAGAAAGACTGC TTGTTCGATGACCTTGGGGTTATTTAGCTTCAAGGGCAATTTTTGG GACATTTTGGACACAGGAGACTCAGAAACAGACACAGAGCGTTC TGAGTCCTGGTGTCTCTGACGTAGGCCTAGAACAGGAATTATTGG CTTTATTTGTTTGTCCATTTTCATAGCCTTGGGGTAATAGATAGAT GACAGAGAAATAGAGAAGACCTAATATTTTTTGTTCATGGCAAAT CGCGGGTTCGCGGTCGGGTCACACACGGAGAAGTAATGAGAAGA GCTGGTAATCTGGGGTAAAAGGGTTCAAAAAGAAGGTCGCTGGT AGGGATGCAATACAAGGTTGTCTTGGAGTTTACATTGACCAGATG ATTTGGCTTTTTCTCTGTTCAATTCACATTTTTCAGCGAGAATCGG ATTGACGGAGAAATGGCGGGGTGTGGGGTGGATAGATGGCAGAA ATGCTCGCAATCACCGCGAAAGAAAGACTTTATGGAATAGAAT ACTGGGTGGTGTAAAGGATTACATAGCTAGTCCAATGGAGTCCGTT GGAAAGGTAAGAAGAAGCTAAAACCGGCTAAGTAACTAGGGAAG AATGATCAGACTTTGATTTGATGAGGTCTGAAAATACTCTGCTGC TTTTTCAGTTGCTTTTTCCCTGCAACCTATCATTTTCCTTTTCATAA GCCTGCCTTTTCTGTTTTCACTTATATGAGTTCGCGGAGACTTCC CCAAATTCTCTCCTGGAACATTCTCTATCGCTCTCCTTCCAAGTTG CGCCCCCTGGCACTGCCTAGTAATATTACCACGCGACTTATATTC AGTTCCACAATTTCCAGTGTTTCGTAGCAAATATCATCAGCCATGG CGAAGGCAGATGGCAGTTTGCTCTACTATAATCCTCACAATCCAC CCAGAAGGTATTACTTCTACATGGCTATATTGCGCGTTTCTGTCTAT TTGCGTTTTGTACGGACCCTCACAACAATTATCATCTCCAAAAAT AGACTATGATCCATTGACGCTCCGATCACTTGATTTGAAGACTTT GGAAGCTCCTTCACAGTTGAGTCCAGGCACCGTAGAAGATAATCT TCG
65	Secuencia de la Región en la posición 3' usada para supresión genética de PpOCH1:	AAAGCTAGAGTAAAATAGATATAGCGAGATTAGAGAATGAATAC CTTCTTCTAAGCGATCGTCCGTCATCATAGAATATCATGGACTGT ATAGTTTTTTTTTTGTACATATAATGATTAAACGGTCATCCAACAT CTCGTTGACAGATCTCTCAGTACGCGAAATCCCTGACTATCAAAG CAAGAACCGATGAAGAAAAAACAACAGTAACCCAAACACCACA ACAAACACTTTATCTTCTCCCCCAACACCAATCATCAAAGAGA TGTCGGAACCAACACCAAGAAGCAAAAACCTAACCCCATATAAA AACATCCTGGTAGATAATGCTGGTAACCCGCTCTCCTTCCATATT CTGGGCTACTTCACGAAGTCTGACCGGTCTCAGTTGATCAACATG ATCCTCGAAATGGGTGGCAAGATCGTTCCAGACCTGCCTCCTCTG GTAGATGGAGTGTTGTTTTTGACAGGGGATTACAAGTCTATTGAT

		GAAGATACCCTAAAGCAACTGGGGGACGTTCCAATATACAGAGA CTCCTTCATCTACCAGTGTTTTGTGCACAAGACATCTCTTCCCATT GACACTTTCCGAATTGACAAGAACGTCGACTTGGCTCAAGATTTG ATCAATAGGGCCCTTCAAGAGTCTGTGGATCATGTCACCTTCTGCC AGCACAGCTGCAGCTGCTGCTGTTGTTGTCGCTACCAACGGCCTG TCTTCTAAACCAGACGCTCGTACTAGCAAAATACAGTTCACTCCC GAAGAAGATCGTTTTATTCTTGACTTTGTTAGGAGAAATCCTAAA CGAAGAAACACACATCAACTGTACACTGAGCTCGCTCAGCACATG AAAAACCATAACGAATCATTCTATCCGCCACAGATTTCTGTCGTAAT CTTTCCGCTCAACTTGATTGGGTTTATGATATCGATCCATTGACCA ACCAACCTCGAAAAGATGAAAACGGGAACTACATCAAGGTACAA GGCCTTCCA
66	Secuencia de la Región en la posición 5' usada para supresión genética de PpBMT2:	GGCCGAGCGGGCCTAGATTTTCACTACAAATTTCAAAACTACGCG GATTTATTGTCTCAGAGAGCAATTTGGCATTCTGAGCGTAGCAG GAGGCTTCATAAGATTGTATAGGACCGTACCAACAAATTGCCGAG GCACAACACGGTATGCTGTGCACTTATGTGGCTACTTCCCTACAA CGGAATGAAACCTTCCTCTTTCCGCTTAAACGAGAAAGTGTGTGCG CAATTGAATGCAGGTGCCTGTGCGCCTTGGTGTATTGTTTTGAG GGCCCAATTTATCAGGCGCCTTTTTTCTTGGTTGTTTTCCCTTAGC CTCAAGCAAGGTTGGTCTATTTTCATCTCCGCTTCTATACCGTGCCT GATACTGTTGGATGAGAACACGACTCAACTTCCTGCTGCTCTGTA TTGCCAGTGTGTTTGTCTGTGATTTGGATCGGAGTCTCCTTACTTG GAATGATAATAATCTTGGCGGAATCTCCCTAAACGGAGGCAAGG ATTCTGCCTATGATGATCTGCTATCATTGGGAAGCTTCAACGACA TGGAGGTCGACTCCTATGTCACCAACATCTACGACAATGCTCCAG TGCTAGGATGTACGGATTTGTCTTATCATGGATTGTTGAAAGTCA CCCCAAAGCATGACTTAGCTTGCGATTTGGAGTTCATAAGAGCTC AGATTTTGGACATTGACGTTTACTCCGCCATAAAAGACTTAGAAG ATAAAGCCTTGACTGTAAACAAAAGGTTGAAAAAACTGGTTTA CGTTTTATGGTAGTTCAGTCTTTCTGCCCGAACACGATGTGCATT ACCTGGTTAGACGAGTCATCTTTTCGGCTGAAGGAAAGGCGAACT CTCCAGTAACATC
67	Secuencia de la Región en la posición 3' usada para supresión genética de PpBMT2:	CCATATGATGGGTGTTTGCTCACTCGTATGGATCAAAATTCATG GTTTCTTCTGTACAACCTGTACACTTATTTGGACTTTTCTAACGGT TTTTCTGGTGATTTGAGAAGTCCTTATTTTGGTGTTCGCAGCTTAT CCGTGATTGAACCATCAGAAATACTGCAGCTCGTTATCTAGTTTC AGAATGTGTTGTAGAATACAATCAATTCTGAGTCTAGTTTGGGTG GGTCTTGGCGACGGGACCGTTATATGCATCTATGCAGTGTTAAGG TACATAGAATGAAAATGTAGGGGTAAATCGAAAGCATCGTTAATT TCAGTAGAACGTAGTTCTATTCCCTACCCAAATAATTTGCCAAGA ATGCTTCGTATCCACATACGCAGTGGACGTAGCAAATTTCACTTT GGACTGTGACCTCAAGTCGTTATCTTCTACTTGGACATTGATGGT CATTACGTAATCCACAAAGAATTGGATAGCCTCTCGTTTTATCTA GTGCACAGCCTAATAGCACTTAAGTAAGAGCAATGGACAAATTT GCATAGACATTGAGCTAGATACGTAACCTCAGATCTTGTTCACTCA TGGTGTACTCGAAGTACTGCTGGAACCGTTACCTCTTATCATTTC GCTACTGGCTCGTGAAACTACTGGATGAAAAAAGAGCT GAAAGCGAGATCATCCCATTTTGTATCATACAAATTCACGCTTG

		CAGTTTTGCTTCGTTAACAAGACAAGATGTCTTTATCAAAGACCC GTTTTTCTTCTTGAAGAATACTTCCCTGTTGAGCACATGCAAACC ATATTTATCTCAGATTTCACTCAACTTGGGTGCTTCCAAGAGAAG TAAAATTCTTCCCACTGCATCAACTTCCAAGAAACCCGTAGACCA GTTTCTCTTCAGCCAAAAGAAGTTGCTCGCCGATCACCGCGGTAA CAGAGGAGTCAGAAGGTTTCACACCCTTCCATCCCGATTTCAAAG TCAAAGTGCTGCGTTGAACCAAGGTTTTTCAGGTTGCCAAAGCCCA GTCTGCAAAAAGTAGTTCCAAATGGCCTATTAATTCCCATAAAAG TGTTGGCTACGTATGTATCGGTACCTCCATTCTGGTATTTGCTATT GTTGTCGTTGGTGGGTTGACTAGACTGACCGAATCCGGTCTTTCC ATAACGGAGTGGAACCTATCACTGGTTCGGTTCCTCCCACTGACT GAGGAAGACTGGAAGTTGGAATTTGAAAAATACAAACAAAGCCC TGAGTTTCAGGAACATAATTCTCACATAACATTGGAAGAGTTCAA GTTTATATTTTCCATGGAATGGGGACATAGATTGTTGGGAAGGGT CATCGGCCGTGTCGTTTGTCTTCCCACGTTTTACTTCATTGCCCGT CGAAAGTGTTCCAAAGATGTTGCATTGAAACTGCTTGCAATATGC TCTATGATAGGATTCCAAGGTTTCATCGGCTGGTGGATGGTGTAT TCCGGATTGGACAAACAGCAATTGGCTGAACGTAACCTCCAAACCA ACTGTGTCTCCATATCGCTTAACTACCCATCTTGGAAGTGCATTTG TTATTTACTGTTACATGATTTACACAGGGCTTCAAGTTTTGAAGA ACTATAAGATCATGAAACAGCCTGAAGCGTATGTTCAAATTTTCA AGCAAATTGCGTCTCCAAAATTGAAAACCTTCAAGAGACTCTCTT CAGTTCTATTAGGCCTGGTG
68	Secuencia de la Región en la posición 5' usada para supresión genética de BMT1	CATATGGTGAGAGCCGTTCTGCACAACTAGATGTTTTGAGCTTC GCATTGTTTCTGCGAGCTCGACTATTGAATTAAGATTTCCGGATA TCTCCAATCTCACAAAACTTATGTTGACCACGTGCTTTCCTGAG GCGAGGTGTTTTATATGCAAGCTGCCAAAAATGGAAAACGAATG GCCATTTTTTCGCCAGGCAAATTATTTCGATTACTGCTGTCATAAA GACAGTGTTGCAAGGCTCACATTTTTTTTTTAGGATCCGAGATAAA GTGAATACAGGACAGCTTATCTCTATATCTTGTAACCATTCGTGAA TCTTAAGAGTTTCGGTTAGGGGGACTCTAGTTGAGGGTTGGCACTC ACGTATGGCTGGGCGCAGAAATAAAATTCAGGCGCAGCAGCACT TATCGATG
69	Secuencia de la Región en la posición 3' usada para supresión genética de BMT1	GAATTCACAGTTATAAATAAAAAACAAAACTCAAAAAGTTTGGG CTCCACAAAATAACTTAATTTAAATTTTGTCTAATAAATGAATG TAATTCCAAGATTATGTGATGCAAGCACAGTATGCTTCAGCCCTA TGCAGCTACTAATGTCAATCTCGCCTGCGAGCGGGCCTAGATTTT CACTACAAATTTCAAAACTACGCGGATTTATTGTCTCAGAGAGCA ATTTGGCATTTCTGAGCGTAGCAGGAGGCTTCATAAGATTGTATA GGACCGTACCAACAAATTGCCGAGGCACAACACGGTATGCTGTG CACTTATGTGGCTACTTCCCTACAACGGAATGAAACCTTCCTCTTT CCGCTTAAACGAGAAAGTGTGTGCAATTGAATGCAGGTGCCTGT GCGCCTTGGTGTATTGTTTTTGAGGGCCCAATTTATCAGGCGCCT TTTTCTTGGTTGTTTTCCCTTAGCCTCAAGCAAGGTTGGTCTATT TCATCTCCGCTTCTATACCGTGCTGATACTGTTGGATGAGAACA CGACTCAACTTCTGCTGCTGTATTGCCAGTGTTTGTCTGTGA TTTGGATCGGAGTCCTCCTTACTTGGAATGATAATAATCTTGGCG GAATCTCCCTAAACGGAGGCAAGGATTCTGCCTATGATGATCTGC TATCATTGGGAAGCTT

70	Secuencia de la Región en la posición 5' usada para supresión genética de BMT3	GATATCTCCCTGGGGACAATATGTGTTGCAACTGTTTCGTTGTTGG TGCCCCAGTCCCCCAACCGGTACTAATCGGTCTATGTTCCCGTAA CTCATATTCGGTTAGAACTAGAACAAATAAGTGCATCATTGTTCAA CATTGTGGTTCAATTGTGCAACATTGCTGGTGCTTATATCTACAG GGAAGACGATAAGCCTTTGTACAAGAGAGGTAACAGACAGTTAA TTGGTATTTCTTTGGGAGTTCGTTGCCCTCTACGTTGTCTCCAAGAC ATACTACATTCTGAGAAACAGATGGAAGACTCAAAAATGGGAGA AGCTTAGTGAAGAAGAGAAAAGTTGCCTACTTGGACAGAGCTGAG AAGGAGAACCTGGGTTCTAAGAGGCTGGACTTTTTGTTTCGAGAGT TAACTGCATAAATTTTTCTAAGTAAATTCATAGTTATGAAATTT CTGCAGCTTAGTGTTTACTGCATCGTTTACTGCATACCCTGTAA ATAATGTGAGCTTTTTTCCTTCCATTGCTTGGTATCTTCCTTGCTG CTGTTT
71	Secuencia de la Región en la posición 3' usada para supresión genética de BMT3	ACAAAACAGTCATGTACAGAACTAACGCCTTTAAGATGCAGACCA CTGAAAAGAATTGGGTCCCATTTTTCTTGAAAGACGACCAGGAAT CTGTCCATTTTGTCTTACTCGTTCAATCCTCTGAGAGTACTCAACTG CAGTCTTGATAACGGTGCATGTGATGTTCTATTTGAGTTACCACA TGATTTTGGCATGTCTTCCGAGCTACGTGGTGCCACTCCTATGCT CAATCTTCCTCAGGCAATCCCGATGGCAGACGACAAAGAAATTTG GGTTTCATTCCCAAGAACGAGAATATCAGATTGCGGGTGTCTGA ACAATGTACAGGCCAATGTTAATGCTTTTTGTTAGAGAAGGAAC AACTTTTTTGCTGAGC
72	Secuencia de la Región en la posición 5' usada para supresión genética de BMT4	AAGCTTGTTACACCGTTGGGACTTTTCCGTGGACAATGTTGACTAC TCCAGGAGGGATTCCAGCTTTCTCTACTAGCTCAGCAATAATCAA TGCAGCCCCAGCGCGCCGTTCTGATGGCTTGATGACCGTTGTATT GCCTGTCACTATAGCCAGGGGTAGGGTCCATAAAGGAATCATAG CAGGGAAATTAAAAGGGCATATTGATGCAATCACTCCCAATGGCT CTCTTGCCATTGAAGTCTCCATATCAGCACTAACTTCCAAGAAGG ACCCCTTCAAGTCTGACGTGATAGAGCACGCTTGCTCTGCCACCT GTAGTCTCTCAAAACGTCACCTTGTGCATCAGCAAAGACTTTAC CTTGCTCCAATACTATGACGGAGGCAATTCTGTCAAAATTCTCTC TCAGCAATTCAACCAACTTGAAAGCAAATTGCTGTCTCTTGATGA TGGAGACTTTTTTCCAAGATTGAAATGCAATGTGGGACGACTCAA TTGCTTCTTCCAGCTCCTCTTCGGTTGATTGAGGAACTTTTGAAAC CACAAAATTGGTCGTTGGGTCAATGTACATCAAACCATTCTGTAGA TTTAGATTTCGACGAAAGCGTTGTTGATGAAGGAAAAGGTTGGAT ACGGTTTGTGCGGTCTCTTTGGTATGGCCGGTGGGGTATGCAATTG CAGTAGAAGATAATTGGACAGCCATTGTTGAAGGTAGAGAAAAG GTCAGGGAACTTGGGGGTATTATTAACCATTTTACCCACAAATA ACAACTGAAAAGTACCCATTCCATAGTGAGAGGTAACCGACGGA AAAAGACGGGCCCATGTTCTGGGACCAATAGAAGTGTGTAATCC ATTGGGACTAATCAACAGACGATTGGCAATATAATGAAATAGTTC GTTGAAAAGCCACGTCAGCTGTCTTTTCATTAACTTTGGTCGGAC ACAACATTTTCTACTGTTGTATCTGTCTACTTTGCTTATCATCTG CCACAGGGCAAGTGGATTTCCTTCTCGCGCGGCTGGGTGAAAAC GGTTAACGTGAA
73	Secuencia de la Región en la posición 3'	GCCTTGGGGGACTTCAAGTCTTTGCTAGAACTAGATGAGGTCAG GCCCTCTTATGGTTGTGTCCCAATTGGGCAATTTCACTCACCTAA AAAGCATGACAATTATTTAGCGAAATAGGTAGTATATTTTCCCTC

	usada para supresión genética de BMT4	ATCTCCCAAGCAGTTTCGTTTTTGCATCCATATCTCTCAAATGAGC AGCTACGACTCATTAGAACCAGAGTCAAGTAGGGGTGAGCTCAG TCATCAGCCTTCGTTTTCTAAAACGATTGAGTTCTTTTGTGCTACA GGAAGCGCCCTAGGGAACTTTCGCACCTTTGGAAATAGATTTTGAT GACCAAGAGCGGGAGTTGATATTAGAGAGGCTGTCCAAAGTACA TGGGATCAGGCCGGCCAAATTGATTGGTGTGACTAAACCATTGTG TACTTGGACACTCTATTACAAAAGCGAAGATGATTTGAAGTATTA CAAGTCCCGAAGTGTTAGAGGATTCTATCGAGCCCAGAATGAAAT CATCAACCGTTATCAGCAGATTGATAAACTCTTGGAAAGCGGTAT CCCATTTTCATTATTGAAGAACTACGATAATGAAGATGTGAGAGA CGGCGACCCTCTGAACGTAGACGAAGAAACAAATCTACTTTTGGG GTACAATAGAGAAAGTGAATCAAGGGAGGTATTTGTGGCCATAA TACTCAACTCTATCATTAATG
74	Secuencia de la Región en la posición 5' usada para supresión genética de PpPNO1 y PpMNN4:	TCATTCTATATGTTCAAGAAAAGGGTAGTGAAAGGAAAGAAAAG GCATATAGGCGAGGGAGAGTTAGCTAGCATACAAGATAATGAAG GATCAATAGCGGTAGTTAAAGTGCACAAGAAAAGAGCACCTGTT GAGGCTGATGATAAAGCTCCAATTACATTGCCACAGAGAAACAC AGTAACAGAAATAGGAGGGGATGCACCACGAGAAGAGCATTGAG TGAACAACCTTTGCCAAATTCATAACCCCAAGCGCTAATAAGCCAA TGTCAAAGTCGGCTACTAACATTAATAGTACAACAACATATCGATT TTCAACCAGATGTTTTGCAAGGACTACAAACAGACAGGTTACTGCG GATATGGTGACACTTGTAAGTTTTTGCACCTGAGGGATGATTTCA AACAGGGATGGAAATTAGATAGGGAGTGGGAAAATGTCCAAAAG AAGAAGCATAATACTCTCAAAGGGGTAAAGGAGATCCAAATGTT TAATGAAGATGAGCTCAAAGATATCCCGTTTAAATGCATTATATG CAAAGGAGATTACAAATCACCCGTGAAAACCTTCTTGCAATCATT TTTTTGCGAACAATGTTTCCTGCAACGGTCAAGAAGAAAACCAAA TTGTATTATATGTGGCAGAGACACTTTAGGAGTTGCTTTACCAGC AAAGAAGTTGTCCCAATTTCTGGCTAAGATACATAATAATGAAAG TAATAAAGTTTAGTAATTGCATTGCGTTGACTATTGATTGCATTG ATGTCGTGTGATACTTTCACCGAAAAAAAACACGAAGCGCAATA GGAGCGGTTGCATATTAGTCCCCAAAGCTATTTAATTGTGCCTGA AACTGTTTTTTAAGCTCATCAAGCATAATTGTATGCATTGCGACG TAACCAACGTTTAGGCGCAGTTTAATCATAGCCCCACTGCTAAGCC
75	Secuencia de la Región en la posición 3' usada para supresión genética de PpPNO1 y PpMNN4:	CGGAGGAATGCAAATAATAATCTCCTTAATTACCCACTGATAAGC TCAAGAGACGCGGTTTTGAAAACGATATAATGAATCATTTGGATT TATAATAAACCCCTGACAGTTTTTCCACTGTATTGTTTTAACTCA TTGGAAGCTGTATTGATTCTAAGAAGCTAGAAATCAATACGGCCA TACAAAAGATGACATTGAATAAGCACCGGCTTTTTTGATTAGCAT ATACCTTAAAGCATGCATTGCTGCTACATAGTTGTTAAAGGGCT TCTTCCATTATCAGTATAATGAATTACATAATCATGCACTTATATT TGCCCATCTCTGTTCTCTCACTCTTGCCCTGGGTATATTCTATGAAA TTGCGTATAGCGTGTCTCCAGTTGAACCCCAAGCTTGCGGAGTTT GAAGAGAATGCTAACCTTGCGTATTCCTTGCTTCAGGAAACATTC AAGGAGAAACAGGTCAAGAAGCCAAACATTTTGATCCTTCCCGA GTTAGCATTGACTGGCTACAATTTTCAAAGCCAGCAGCGGATAGA GCCTTTTTTGGAGGAAACAACCAAGGGAGCTAGTACCCAATGGG CTCAAAAAGTATCCAAGACGTGGGATTGCTTTACTTTAATAGGAT

		ACCCAGAAAAAAGTTTAGAGAGCCCTCCCCGTATTTACAACAGTG CGGTACTTGTATCGCCTCAGGGAAAAAGTAATGAACAACTACAGA AAGTCCTTCTTGTATGAAGCTGATGAACATTGGGGATGTTCCGAA TCTTCTGATGGGTTTCAAACAGTAGATTTATTAATTGAAGGAAAG ACTGTAAAGACATCATTGGAATTTGCATGGATTTGAATCCTTAT AAATTTGAAGCTCCATTACAGACTTCGAGTTCAGTGGCCATTGC TTGAAAACCGGTACAAGACTCATTTTGTGCCCAATGGCCTGGTTG TCCCCTCTATCGCCTTCCATTAATAAAGGATCTTAGTGATATAGAG AAAAGCAGACTTCAAAGTTCTACCTTGAAAAAATAGATACCCCG GAATTTGACGTTAATTACGAATTGAAAAAAGATGAAGTATTGCC ACCCGTATGAATGAAACGTTGGAAACAATTGACTTTGAGCCTTCA AAACCGGACTACTCTAATATAAATTATTGGATACTAAGGTTTTTT CCCTTTCTGACTCATGTCTATAAACGAGATGTGCTCAAAGAGAAT GCAGTTGCAGTCTTATGCAACCGAGTTGGCATTGAGAGTGATGTC TTGTACGGAGGATCAACCACGATTCTAAACTTCAATGGTAAGTTA GCATCGACACAAGAGGAGCTGGAGTTGTACGGGCAGACTAATAG TCTCAACCCCAAGTGTGGAAGTATTGGGGGCCCTTGGCATGGGTCA ACAGGGAATTCTAGTACGAGACATTGAATTAACATAATATACAAT ATACAATAAACACAAATAAAGAATACAAGCCTGACAAAAATTCA CAAATTATTGCCTAGACTTGTCTGTTATCAGCAGCGACCTTTTTCC AATGCTCAATTTACGATATGCCTTTTCTAGCTCTGCTTTAAGCTT CTCATTGGAATTGGCTAACTCGTTGACTGCTTGGTCAGTGATGAG TTTCTCCAAGGTCCATTCTCGATGTTGTTGTTTTCGTTTTCTTT AATCTCTTGATATAATCAACAGCCTTCTTTAATATCTGAGCCTTGT TCGAGTCCCCTGTTGGCAACAGAGCGGCCAGTTCCTTTATTCCGT GGTATATTTTTCTTCTACGCCTTTCTACTTCTTTGTGATTCTCT TTACGCATCTTATGCCATTCTTCAGAACCAGTGGCTGGCTTAACC GAATAGCCAGAGCCTGAAGAAGCCGCACTAGAAGAAGCAGTGGC ATTGTTGACTATGG
76	Secuencia de la Región en la posición 5' usada para supresión genética de PpMNN4L 1:	GATCTGGCCATTGTGAAACTTGACACTAAAGACAAAACCTTTAGA GTTCCAATCACTTAGGAGACGATGTTTCCTACAACGAGTACGAT CCCTCATTGATCATGAGCAATTTGTATGTGAAAAAAGTCATCGAC CTTGACACCTTGGATAAAAGGGCTGGAGGAGGTGGAACCACTG TGCAGGCGGTCTGAAAGTGTCAAGTACGGATCTACTACCAATA TACATCTGGTAACCTGAACGGCGTCAGGTTAGTATACTGGAACGA AGGAAAGTTGCAAAGCTCCAAATTTGTGGTTCGATCCTCTAATTA CTCTCAAAAGCTTGGAGGAAACAGCAACGCCGAATCAATTGACA ACAATGGTGTGGGTTTTGCCTCAGCTGGAGACTCAGGCGCATGG ATTCTTTCCAAGCTACAAGATGTTAGGGAGTACCAGTCATTCCT GAAAAGCTAGGTGAAGCTACGATGAGCATTTCGATTTCACGGT CTTAAACAGGAGACTTCTACTACAGGGCTTGGGGTAGTTGGTATG ATTCATTCTTACGACGGTGAGTTCAAACAGTTTGGTTTGTTCCT CCAATGACATCTATTCTACAAAGACTTCAACGAGTGACCAATGTA GAATGGTGTGTAGCGGGTTGCGAAGATGGGGATGTGGACACTGA AGGAGAACACGAATTGAGTGATTGGAACAACCTGCATATGCATA GTGATTCCGACTAGTCAGGCAAGAGAGAGCCCTCAAATTTACCTC TCTGCCCTCCTCACTCCTTTTGGTACGCATAATTGCAGTATAAA GAACTTGCTGCCAGCCAGTAATCTTATTTACATACGCAGTTCTATA TAGCACATAATCTTGCTTGTATGTATGAAATTTACCGCGTTTTAGT

		TGAAATTGTTTATGTTGTGTGCCTTGCATGAAATCTCTCGTTAGCC CTATCCTTACATTTAACTGGTCTCAAAACCTCTACCAATTCCATTG CTGTACAACAATATGAGGCGGCATTACTGTAGGGTTGGAAAAAA ATTGTCATTCCAGCTAGAGATCACACGACTTCATCACGCTTATTG CTCCTCATTGCTAAATCATTTACTCTTGACTTCGACCCAGAAAAG TTCGCC
77	Secuencia de la Región en la posición 3' usada para supresión genética de PpMNN4L 1:	GCATGTCAAACCTTGAACACAACGACTAGATAGTTGTTTTTCTAT ATAAAACGAAACGTTATCATCTTTAATAATCATTGAGGTTTACCC TTATAGTTCCGTATTTTCGTTTCCAAACTTAGTAATCTTTTGGAAA TATCATCAAAGCTGGTGCCAATCTTCTTGTTGAAGTTTCAAAC GCTCCACCAAGCTACTTAGAGACTGTTCTAGGTCTGAAGCAACTT CGAACACAGAGACAGCTGCCGCCGATTGTTCTTTTTTGTGTTTT CTTCTGGAAGAGGGGCATCATCTTGTATGTCCAATGCCCGTATCC TTTCTGAGTTGTCCGACACATTGTCTTTCGAAGAGTTTCCTGACA TTGGGCTTCTTCTATCCGTGTATTAATTTTGGGTTAAGTTCCTCGT TTGCATAGCAGTGGATACCTCGATTTTTTTGGCTCCTATTTACCTG ACATAATATTCTACTATAATCCAACCTTGGACGCGTCATCTATGAT AACTAGGCTCTCCTTTGTTCAAAGGGGACGCTTTCATAATCCACT GGCACGAAGTAAGTCTGCAACGAGGCGGCTTTTGAACAGAACG ATAGTGTGTTTTCGTACTTGGACTATGCTAAACAAAAGGATCTGT CAAACATTTCAACCGTGTTTCAAGGCACTCTTTACGAATTATCGA CCAAGACCTTCCTAGACGAACATTTCAACATATCCAGGCTACTGC TTCAAGGTGGTGCAAATGATAAAGGTATAGATATTAGATGTGTTT GGGACCTAAACAGTTCTTGCCTGAAGATTCCCTTGAGCAACAGG CTTCAATAGCCAAGTTAGAGAAGCAGTACCAAATCGGTAACAAA AGGGGGAAGCATATAAAACCTTTACTATTGCGACAAAATCCATCC TTGAAAGTAAAGCTGTTTGTTCATGTAAAGCATACGAAACGAAG GAGGTAGATCCTAAGATGGTTAGAGAACTTAACGGGACATACTC CAGCTGCATCCCATATTACGATCGCTGGAAGACTTTTTTCATGTA CGTATCGCCCACCAACCTTTCAAAGCAAGCTAGGTATGATTTTGA CAGTTCTCACAATCCATTGGTTTTTCATGCAACTTGAAAAAACCA ACTCAAACCTTCATGGGGATCCATACAATGTAAATCATTACGAGAG GGCGAGGTTGAAAAGTTTCCATTGCAATCACGTCGCATCATGGCT ACTGAAAGGCCTTAAC
78	Secuencia del locus de integración del gen PpTRP2:	TAATGGCCAAACGGTTTTCTCAATTACTATATACTACTAACCATTT ACCTGTAGCGTATTTCTTTTCCCTCTTCGCGAAAGCTCAAGGGCA TCTTCTTGACTCATGAAAAATATCTGGATTTCTTCTGACAGATCAT CACCTTGAGCCCAACTCTCTAGCCTATGAGTGTAAGTGATAGTC ATCTTGCAACAGATTATTTTGGAACGCAACTAACAAAGCAGATAC ACCCTTCAGCAGAATCCTTTCTGGATATTGTGAAGAATGATCGCC AAAGTCACAGTCCTGAGACAGTTCCTAATCTTTACCCCATTTACA AGTTCATCCAATCAGACTTCTTAACGCCTCATCTGGCTTATATCA AGCTTACCAACAGTTTCAGAACTCCAGTCCAAGTTTCTTGCTTG AAAGTGCGAAGAATGGTGACACCGTTGACAGGTACACCTTTATG GGACATTCCCCCAGAAAAATAATCAAGACTGGGCCTTTAGAGGG TGCTGAAGTTGACCCCTTGGTGCTTCTGGAAAAAGAACTGAAGG GCACCAGACAAGCGCAACTTCCTGGTATTCCTCGTCTAAGTGGTG GTGCCATAGGATACATCTCGTACGATTGTATTAAGTACTTTGAAC CAAAACTGAAAGAAAACTGAAAGATGTTTTGCAACTTCCGGAA

		GCAGCTTTGATGTTGTTTCGACACGATCGTGGCTTTTGACAATGTT TATCAAAGATTCCAGGTAATTGGAAACGTTTCTCTATCCGTTGAT GACTCGGACGAAGCTATTCTTGAGAAATATTATAAGACAAGAGA AGAAGTGGAAAAGATCAGTAAAGTGGTATTTGACAATAAACTG TTCCCTACTATGAACAGAAAGATATTATTCAAGGCCAAACGTTCA CCTCTAATATTGGTCAGGAAGGGTATGAAAACCATGTTTCGCAAGC TGAAAGAACATATTCTGAAAGGAGACATCTTCCAAGCTGTTCCCT CTCAAAGGGTAGCCAGGCCGACCTCATTGCACCCTTTCAACATCT ATCGTCATTTGAGAACTGTCAATCCTTCTCCATACATGTTCTATAT TGACTATCTAGACTTCCAAGTTGTTGGTGCTTCACCTGAATTACT AGTTAAATCCGACAACAACAACAAAATCATCACACATCCTATTGC TGGAACCTCTTCCCAGAGGTAAACTATCGAAGAGGACGACAATT ATGCTAAGCAATTGAAGTCGTTTGAAGACAGGGCCGAGCAC GTCATGCTGGTAGATTTGGCCAGAAATGATATTAACCGTGTGTGT GAGCCACCAGTACCACGGTTGATCGTTTATTGACTGTGGAGAGA TTTTCTCATGTGATGCATCTTGTGTCAGAAGTCAGTGGAACATTG AGACCAAACAAGACTCGCTTCGATGCTTTCAGATCCATTTTCCCA GCAGGAACCGTCTCCGGTGCTCCGAAGGTAAGAGCAATGCAACT CATAGGAGAATTGGAAGGAGAGAAAAGAGAGGTGTTTATGCGGGG CCGTAGGACACTGGTTCGTACGATGGAAAATCGATGGACACATGT ATTGCCTTAAGAACAATGGTCGTCAAGGACGGTGTGCTTACCTT CAAGCCGGAGGTGGAATTGTCTACGATTCTGACCCCTATGACGAG TACATCGAAACCATGAACAAAATGAGATCCAACAATAACACCATC TTGGAGGCTGAGAAAATCTGGACCGATAGGTTGGCCAGAGACGA GAATCAAAGTGAATCCGAAGAAAACGATCAATGAACGGAGGACG TAAGTAGGAATTTATGGTTTGGCCAT
79	Secuencia de la Región en la posición 5' usada para supresión genética de PpARG1:	GATCTGGCCTTCCCTGAATTTTACGTCCAGCTATACGATCCGTT GTGACTGTATTTCTGAAATGAAGTTTCAACCTAAAGTTTGGTT GACTTGTCCACCTACCACGGAAACTAATATCGAAACCAATGAA AAAGTAGAACTGGAATCGTCAATCGAAATTCGCAACCAAGTGGA ACCCAAAGACTTGAATCTTTCTAAAGTCTATTCTAGTGACACTAA TGGCAACAGAAGATTTGAGCTGACTTTTCAAATGAATCTCAATAA TGCAATATCAACATCAGACAATCAATGGGCTTTGTCTAGTGACAC AGGATCAATTATAGTAGTGTCTTCTGCAGGAAGAATAACTTCCCC GATCCTAGAAGTCGGGGCATCCGTCTGTGTCTTAAGATCGTACAA CGAACACCTTTTGGCAATAACTTGTGAAGGAACATGCTTTTCATG GAATTTAAAGAAGCAAGAATGTGTTCTAAACAGCATTTTCATTAGC ACCTATAGTCAATTCACACATGCTAGTTAAGAAAGTTGGAGATGC AAGGAACTATTCTATTGTATCTGCCGAAGGAGACAACAATCCGTT ACCCAGATTCTAGACTGCGAACTTTCCAAAAATGGCGCTCCAAT TGTGGCTCTTAGCACGAAAGACATCTACTTTATTCAAAGAAAAT GAAATGCTGGATCCATTTGATTGATTGAAATCTTTGAATTGTT GGGTGCTGACAATGCACTGTTTGAGTGTGTGGAAGCGCTAGAAG GTCCAATTGGAATGCTAATTCATAGATTGGTAGATGAGTTCTTCC ATGAAAACACTGCCGGTAAAAAACTCAAACCTTACAACAAGCGA GTAAGGAGGACCTTTCAAATTCAGTTGAAGAACTAGGTGAAAAT GCGTCTCAATTAAAGAGAGAACTTGACAACTCTATGGTGATGAG GTTGAGGCTTCTTGACCTCTTCTCTCTATCTGCGTTTCTTTTTTTT TTTTTTTTTTTTTTTTTTCAGTTGAGCCAGACCGCGCTAAACGCAT

		ACCAATTGCCAAATCAGGCAATTGTGAGACAGTGGTAAAAAAGA TGCTTGCAAAGTTAGATTACACAGTAAGAGAGATCCTACTCATA AATGAGGCGCTTATTTAGTAGCTAGTGATAGCCACTGCGGTTCTG CTTTATGCTATTTGTTGTATGCCTTACTATCTTTGTTTGGCTCCTT TTTCTTGACGTTTTCCGTTGGAGGGACTCCCTATTCTGAGTCATG AGCCGCACAGATTATCGCCCCAAAATTGACAAAATCTTCTGGCGAA AAAAGTATAAAAGGAGAAAAAAGCTCACCCTTTTCCAGCGTAGA AAGTATATATCAGTCATTGAAGAC
80	Secuencia de la Región en la posición 3' usada para supresión genética de PpARG1:	GGGACTTTAACTCAAGTAAAAGGATAGTTGTACAATTATATATAC GAAGAATAAATCATTACAAAAAGTATTCGTTTCTTTGATTCTTAA CAGGATTCATTTTCTGGGTGTCATCAGGTACAGCGCTGAATATCT TGAAGTTAACATCGAGCTCATCATCGACGTTTCATCACACTAGCCA CGTTTCCGCAACGGTAGCAATAATTAGGAGCGGACCACACAGTG ACGACATCTTTCTCTTTGAAATGGTATCTGAAGCCTTCCATGACC AATTGATGGGCTCTAGCGATGAGTTGCAAGTTATTAATGTGGTTG AACTCACGTGCTACTCGAGCACCGAATAACCAGCCAGCTCCACGA GGAGAAACAGCCCAACTGTCGACTTCATCTGGGTGAGACCAAAC CAAGTCACAAAATCCTCCTTCATGAGGGACCTCTTGCGCTCGGCT GAGAACTCTGATTTGATCTAACATGCGAATATCGGGAGAGAGAC CACCATGGATACATAATATTTTACCATCAATGATGGCACTAAGGG TAAAAAAGTCGAACACCTGGCAACAGTACTTCCAGACAGTGGTG GAACCATATTTATTGAGACATTCCTCATAAAATCCATAAACCTGA GTGATCTGTCTGGATTCATGATTTCCCCTTACCAATGTGATATGTT GAGGAAACTTAATTTTTAAATCATGAGTAACGTGAACGTCTCCA ACGAGAAATAGCCTCTATCCACATAGTCTCCTAGGAAGATATAGT TCTGTTTTATTCCATTAGAGGAGGATCCGGGAAACCCACCACTAA TCTTGAAAAGTTCCAGTAGATCGTGAAATTGGCCGTGAATATCTC CGCATACTGTCACTGGACTCTGCACTGGCTGTATATTGGATTCCCT CCATCAGCAAATCCTTCACCCGTTTCGCAAAGATGCTTCATATCAT TTTCACTTAAAGCCTTGACGCTTTTGACTTCTTCAAACCACTGATC TGGTCCTCTTTCTGGCATGATTAAGGTCTATAATATTTCTGAGCTG AGATGTAAAAAATAATAAAAAATGGGGAGTGAAAAAGTGTGT AGCTTTTAGGAGTTTGGGATTGATACCCCAAAATGATCTTTATGA GAATTAAGGTAGATACGCTTTTAATAAGAACACCTATCTATAG TACTTTGTGGTCTTGAGTAATTGAGATGTTTCAGCTTCTGAGGTTT GCCGTTATTCTGGGATAGTAGTGCGCGACCAAACAACCCGCCAG GCAAAGTGTGTTGTGCTCGAAGACGATTGCCAGAAGAGTAAGTC CGTCCTGCCTCAGATGTTACACACTTTCTTCCCTAGACAGTCGAT GCATCATCGGATTTAAACCTGAAACTTTGATGCCATGATACGCCT AGTCACGTCGACTGAGATTTTAGATAAGCCCCGATCCCTTTAGTA CATTCCTGTTATCCATGGATGGAATGCCCTGATA
81	Secuencia del marcador auxotrófico de PpARG1:	CAGTTGAGCCAGACCGCGCTAAACGCATACCAATTGCCAAATCAG GCAATTGTGAGACAGTGGTAAAAAAGATGCCTGCAAAGTTAGAT TCACACAGTAAGAGAGATCCTACTCATAAATGAGCGCTTATTTA GTAGCTAGTGATAGCCACTGCGGTTCTGCTTTATGCTATTTGTTG TATGCCTTACTATCTTTGTTTGGCTCCTTTTTCTTGACGTTTTCCG

		TTGGAGGGACTCCCTATTCTGAGTCATGAGCCGCACAGATTATCG CAAAAAATTGACAAAATCTTCTGGCGAAAAAAGTATAAAAGGAG AAAAAAGCTCACCTTTTCCAGCGTAGAAAGTATATATCAGTCAT TGAAGACTATTATTTAAATAACACAATGTCTAAAGGAAAAAGTTG TTTGGCCTACTCCGGTGGTTTGGATACCTCCATCATCCTAGCTTG GTTGTGGAGCAGGGATACGAAGTCGTTGCCTTTTGTAGCCAACAT TGGTCAAGAGGAAGACTTTGAGGCTGCTAGAGAGAAAGCTCTGA AGATCGGTGCTACCAAGTTTATCGTCAGTGACGTTAGGAAGGAAT TTGTTGAGGAAGTTTGTTCCTCCAGCAGTCCAAGTTAACGCTATCT ACGAGAACGTCTACTTACTGGGTACCTCTTTGGCCAGACCAGTCA TTGCCAAGGCCCAAATAGAGGTTGCTGAACAAGAAGGTTGTTTTG CTGTTGCCACGGTTGTACCGGAAAGGGTAACGATCAGGTTAGAT TTGAGCTTTCCTTTTATGCTCTGAAGCCTGACGTTGTCTGTATCGC CCCATGGAGAGACCCAGAATTCTTCGAAAGATTCTGCTGGTAGAA ATGACTTGCTGAATTACGCTGCTGAGAAGGATATTCCAGTTGCTC AGACTAAAGCCAAGCCATGGTCTACTGATGAGAACATGGCTCAC ATCTCCTTCGAGGCTGGTATTCTAGAAGATCCAAACACTACTCCT CCAAAGGACATGTGGAAGCTCACTGTTGACCCAGAAGATGCACC AGACAAGCCAGAGTTCTTTGACGTCCACTTTGAGAAGGGTAAGCC AGTTAAATTAGTTCTCGAGAACAAAAGTGAAGTCAACGATCCGGT TGAGATCTTTTGTACTGCTAACGCCATTGCTAGAAGAAACGGTGT TGGTAGAATTGACATTGTGAGAACAGATTATCGGAATCAAGTC CAGAGGTTGTTATGAAACTCCAGGTTTGTACTTACTGAGAACCAC TCACATCGACTTGGAAGGTCTTACCGTTGACCGTGAAGTTAGATC GATCAGAGACACTTTTGTACCCCAACCTACTCTAAGTTGTTATA CAACGGGTTGTACTTTACCCCAAGAGGTGAGTACGTCAGAACTAT GATTCAGCCTTCTCAAAACACCGTCAACGGTGTGTTAGAGCCAA GGCCTACAAAGGTAATGTGTATAACCTAGGAAGATACTCTGAAA CCGAGAAATTGTACGATGCTACCGAATCTTCCATGGATGAGTTGA CCGGATTCCACCCTCAAGAAGCTGGAGGATTTATCACAACACAAG CCATCAGAATCAAGAAGTACGGAGAAAGTGTGAGAGAGAAGGGA AAGTTTTTGGGACTTTAACTCAAGTAAAAGGATAGTTGTACAATT ATATATACGAAGAATAAATCATTACAAAAGTATTCTGTTTCTTTG ATTCTTAACAGGATTCATTTTCTGGGTGTCATCAGGTACAGCGCT GAATATCTTGAAGTTAACATCGAGCTCATCATCGACGTTTCATCAC ACTAGCCACGTTTCCGCAACGGTAGCAATAATTAGGAGCGGACC ACACAGTGACGACATC
82	Secuencia de la Región en la posición 5' que se usó para adición genética en el locus de PpADE1:	GAGTCGGCCAAGAGATGATAACTGTTACTAAGCTTCTCCGTAATT AGTGGTATTTTGTAACTTTTACCAATAATCGTTTATGAATACGGA TATTTTTCGACCTTATCCAGTGCCAAATCACGTAACCTAATCATG GTTTAAATACTCCACTTGAACGATTCAATTATCAGAAAAAAGTCA GGTITGGCAGAAACACTTGGGCGCTTTGAAGAGTATAAGAGTATT AAGCATTAAACATCTGAACCTTACCGCCCCAATATACTACTCTA GGAAACTCGAAAAATTCTTTCCATGTGTATCGCTTCCAACACA CTTTGCTGTATCCTTCCAAGTATGTCCATTGTGAACACTGATCTG GACGGAATCCTACCTTTAATCGCCAAAGGAAAGGTTAGAGACATT TATGCAGTCGATGAGAACAACCTTGCTGTTTCGTCGCAACTGACCGT ATCTCCGCTTACGATGTGATTATGACAAACGGTATTCTCTGATAAG GGAAAGATTTTGTACTCAGCTCTCAGTTTTCTGGTTTGATTTTTTG

		CACCCTACATAAAGAATCATTTGGTTGCTTCTAATGACAAGGAAG TCTTTGCTTTACTACCATCAAACTGTCTGAAGAAAAaTACAAATC TCAATTAGAGGGACGATCCTTGATAGTAAAAAAGCACAGACTGA TACCTTTGGAAGCCATTGTCAAGAGGTTACATCACTGGAAGTGCAT GGAAAGAGTACAAGAAGCTCAAAAAGCTGCCATGGAGTCAAGGTT GAAAACGAGAACCCTTCAAGAGAGCGACGCCTTTCCAACCTCCGATT TTCACACCTTCAACGAAAGCTGAACAGGGTGAACACGATGAAAA CATCTCTATTGAACAAGCTGCTGAGATTGTAGGTAAAGACATTG TGAGAAGGTGCTGTCAAGGCGGTGAGTTGTATTCTGCTGCAAA AAACCTCGCCCTTTTGAAGGGGATCATTATTGCTACTAGCAAAT CGAATTTGGACTGGACGAAAACAATGAATTGGTACTAGTAGTG AAGTTTTAACTCCAGATTCTTCTAGATTTTGGAATCAAAAGACTT ACCAAGTGGGTAAATCGCAAGAGAGTTACGATAAGCAGTTTCTC AGAGATTGGTTGACGGCCAACGGATTGAATGGCAAAGAGGGCGT AGCCATGGATGCAGAAATTGCTATCAAGAGTAAAGAAAAGTATA TTGAAGCTTATGAAGCAATTACTGGCAAGAAATGGGCTTGA
83	Secuencia de la Región en la posición 3' que se usó para adición genética en el locus de PpADE1:	ATGATTAGTACCCTCCTCGCCTTTTTTCAGACATCTGAAATTTCCCT TATTCTTCCAATTCATATAAAATCCTATTTAGGTAATTAGTAAAC AATGATCATAAAGTGAAATCATTCAAGTAACCATTCCGTTTATCG TTGATTTAAAATCAATAACGAATGAATGTGCGTCTGAGTAGTCAA TTTGTTCCTTGGAGCTCATTGGCAGGGGGTCTTTTGGCTCAGTA TGGAAGGTTGAAAGGAAAACAGATGGAAAGTGGTTTCGTCAGAAA AGAGGTATCCTACATGAAGATGAATGCCAAAGAGATATCTCAAG TGATAGCTGAGTTCAGAATTCCTTAGTGAGTTAAGCCATCCCAACA TTGTGAAGTACCTTCATCACGAACATATTTCTGAGAATAAACTG TCAATTTATACATGGAATACTGTGATGGTGGAGATCTCTCCAAGC TGATTCGAACACATAGAAGGAACAAAGAGTACATTTTCAAGAA AAAATATGGAGTATTTTACGCAGGTTTATTAGCATTGTATCGT TGTCATTATGGAACGTATTTACGGCTTCAAAGGAGTTTGAATCG CTCAATAAAGGTAATAGACGAACCCAGAATCCTTCGTGGGTAGA CTCGACAAGAGTTATTATTCACAGGGATATAAAACCCGACAACAT CTTCTGATGAACAATTCAAACCTTGTCAAACCTGGGAGATTTTGG ATTAGCAAAAATTCTGGACCAAGAAAACGATTTTGCCAAAACATA CGTCGGTACGCCGTATTACATGTCTCCTGAAGTGCTGTTGGACCA ACCCTACTCACCATTATGTGATATATGGTCTCTTGGGTGCGTCAT GTATGAGCTATGTGCATTGAGGCCTCCTT
84	MET16 5'	GGGTGGGCCTGGTAATGTTCACTCCTAGGAACTACTAGAAAACT GTGCTAAACGGATTACGTAATTATTATACAAATTCTCTATGGTCT ATGGTACATATGGGCTGGTTCAATAATGAATCTATGAAGAATTTG TGCCCATGGGGACCGTTTCTATAAACGTTCTCTTCTTTATGTTTT CACCTGCTCTTTGAGTTCCGGAAATTCGTTGACAATCTTTTGCCC AATGTCGATTGGGCGTATTTAAAGCCCAGCTGTTTTCTCTGAGA AATTGATTCAACTTCCTCACCACCTCCACAACTCACGCGTGTAT ATATCAGGGTTTCTACCGTCTTCGATATAATTGACTACGTCCACG GGGATGGGAATGTTCAAATCTGTGTTGTGGAGCTTTTGCAAGTGC TCTACAACCTTGTTAATGTTGTTGGAAAGACCCAATTGACTTTCC GCTGTACCGGCGTAATCGTGCACCTGAACACCCAATGGATGAG GGTTTCGATGAGTTGACTTAGTTCATTTCAACTTGATCTAATGTT

		GTCGCAGGTGCACTCATACTTGTTCATGGAGAATGAAAGTAAGTTG ATAGAGAGCAGACTTCGAGGATGGGATGAACTTGATTAGGTAAT CTTTGACAATGTCTTAGAGGTAGGCAGAGGATGCTGGAAAAAAA AAATTGAAAACGCCCAAGCTTCCAGCTTTGCAAGGAAAGAAGAA AAGGGAGTTGCCAGCACGAAATCGGCTTCCTCCGAAAGGTTTAC AATTGCAGAATTGTCACCATTCAAATGCCTTTACCCTTCATCTGTG GTACCTCAGGCTAAGAACGGGTACCGTGATATTTTCGACACTCATC GCCACAATATGTACTAGCAAGAACTTTTCAGATTTAGTAATCCGT TCGAAACGGG
85	MET16 3'	CTAGATTTGCACAATATTTGAAAGCTCAGCAAAACATATGAATAT AATTTTTTTTTTCTCTACACTATTTATCCTGTAAGTTTCTGTTCCC CATGTAGGATCTTTTTCTCCTTCTCTGTCTCCCATTTTTTTTGTTC CTGTAGTCTTGCCCTTGCCCTGAGATGCGAGCTCGTCCGCCCATCCA GTCGTGTGAAGGGCCTAGCTTTTCAAAAAGAAAATACCTCCCGCT AAAGGAGGCGTTGCCCTTCTATCAGTAGTGTGCTAACCAATTTT CACAAACAATAAAAAAAGGACACCAACAACGAAATCAACTATTT ACACACATCCAGATCCGTCCCCCTCCCATCCAAGAGTTAAAGAC AAATATGGCTGTTAATAATCCGTCTGAATTTAGAAAGAAGTTGGT CGTAGTAGGAGATGGTGCTTGCAGTAAACTTGTCTATTGATGGT GTTGCCGAGGGCGAGTTCCCTCCATCTTATGTTCCAAGTGTTCCT GAGAACTATGCCACCCAGTAGAGGTTGACAACAGAATAGTACA ACTCACTCTATGGGATACTGCCGGACAGGAAGATTATGATAGACT GAGACCTCTTTCCTATCCCGATGCCAATGTGGTCTTGATTTGTTT GCTATTGACATTCCTGACACCTTAGATAACGTTCAAGAGAAGTGG ATTAGTGAGGTGTTGCATTTCTGTCTGAGTCCCTATCATTTTA GTTGGTTGTAACTTGACTTGAGAAACGATCCAGAGGTTATCCGT GAATTACAAGCTGTTGGAAAGCAACCAGTCTCCACCAGTGAGGG TCAGGCCGTTGC
86	Secuencia del marcador auxotrófico de PpMET16:	CAACTTCCTCACCACCTCCACAACTCACGCGTGTATATATCAGG GTTCTACCGTCTTCGATATAATTGACTACGTCCACGGGGATGGG AATGTTCAAATCTGTGTTGTGGAGCTTTTGCAAGTGCTCTACAAC CTGTGTTAATGTTGTTGGAAAGACCCAATTGACTTTCGGCTGTACC GCGGTAATCGTGCACCTGAACACCCAAATGGATGAGGGTTTCGAT GAGTTGACTTAGTTTCATTTTCAACTTGATCTAATGTTGTGCGAGG TGCACCTCATACTTGTTCATGGAGAATGAAAGTAAGTTGATAGAGA GCAGACTTCGAGGATGGGATGAACTTGATTAGGTAATCTTTGACA ATGTCTTAGAGGTAGGCAGAGGATGCTGGAAAAAAAAAATTGAA AACGCCCAAGCTTCCAGCTTTGCAAGGAAAGAAGAAAAGGGAGT TGCCAGCACGAAATCGGCTTCCTCCGAAAGGTTACAATTGCAGA ATTGTCACCATTCAAATGCCTTTACCCTTCATCTGTGGTACCTCAG GCTAAGAACGGGTACCGTGATATTTTCGACACTCATCGCCACAATA TGTACTAGCAAGAACTTTTCAGATTTAGTAATCCGTTCGAAACGG GAAAAAATGTTTTTACCCTTCTATCAACTGCTAATCTTTCTAGGTT TATACTGCCAGCAGCCCGTTCCAGATACCAACATGCCATTCACTA TAGGCCAGTCAAAAACCAGTTTGAACCTCTCCAAGGTCCAAGTGG ACCACCTTAACCTTTCTCTTCAGAATCTCAGTCCAGAAGAAATCA TACAATGGTCTATCATTACCTTCCCACACCTGTATCAAACCTACGG CATTCGGATTGACTGGGTTGTGTATAACTGACATGGTTTCAAAAA TAACAGCCAAAAGAGGCAAAAAGCATGCTATTGACTTGATTTTCA

		TAGACACCTTACATCATTTTTCCACAGACTTTAGATCTCGTTGAAC GAGTCAAAGATAAATACCACTGCAATGTTTCATGTCTTCAAACCAC AGAATGCCACTACTGAGCTCGAGTTTGGGGCGCAATATGGCGAA AACTTATGGGAAACAGATGATAACAAGTATGACTACCTCGTAAA AGTTGAACCTCACAACGTGCCTACCATGCATTAGACGTCTGCGC CGTCTTCACAGGAAGAAGACGGTCTCAAGGTGGTAAAAGGGGAG AATTGCCCCGTGATTGAAATTGATGAAATTTCTCAGGTGGTCAAGA TTAATCCGTTAGCATCCTGGGGGTTTGAACAAGTTCAAAACTATA TCCAAGCTAATAGCGTTCATACAACGAATTGCTGGATTGGGAT ACAAGTCAGTTGGAGATTACCATTCCACACAACCCACTAAAAATG GTGAAGATGAAAGAGCAGGCAGGTGGAGAGTTAAACAAAAGAG TGAGTGTGGTATCCACGAAGCTTCTAGATTTGCACAATATTTGAA AGCTCAGCAAAACATATGAATATAATTTTTTTTTTCTCTACACTAT TTATCCTGTAAGTTTCTGTTTCCCCATGTAGGATCTTTTTCTCCTT CTCTGTCTCCCATTTTTTTTGTTCCTGTAGTCTTGCCTTGCCTGA GATGCGAGCTCGTCCGCCCATCCAGTCGTGTGAAGGGCCTAGCTT TTCAAAAAGAAAATACCTCCCGCTAAAGGAGGCGTTGCCCTTCT ATCAGTAGTGTGTAACCAATTTTCACAAACAATAAAAAAAGGAC ACCAACAACGAAATCAACTATTTACACACATCCAGATCCGTCCC
87	Secuencia de la Región en la posición 5' usada para supresión genética de PpHIS1:	TAACTGGCCCTTTGACGTTTCTGACAATAGTTCTAGAGGAGTCGT CCAAAACTCAACTCTGACTTGGGTGACACCACCGGGATCCGG TTCTTCCGAGGACCTTGATGACCTTGGCTAATGTAAGTGGAGTTT TAGTATCCATTTTAAAGATGTGTGTTTCTGTAGGTTCTGGGTTGGA AAAAAATTTTAGACACCAGAAGAGAGGAGTGAAGTGGTTTGCCT GGGTTTAGACTGTGTAAGGCACTACTCTGTGCAAGTTTATAGATAG GGGTTACCCGCTCCGATGCATGGGAAGCGATTAGCCCGGCTGTTG CCCGTTTGGTTTTTGAAGGGTAATTTTCAATATCTCTGTTTGAAGT ATCAATTTTCAATTTCAAAGATTCAAAAAACAAATCTGGTCCAAGG AGCGCATTTAGGATTATGGAGTTGGCGAATCACTTGAACGATAGA CTATTATTTGC
88	Secuencia de la Región en la posición 3' usada para supresión genética de PpHIS1:	GTGACATTCTTGTCTTTGAGATCAGTAATTGTAGAGCATAGATAG AATAATATTCAAGACCAACGGCTTCTCTTCGGAAGCTCCAAGTAG CTTATAGTGATGAGTACCGGCATATATTTATAGGCTTAAATTTTC GAGGGTTCATATATTCGTTTAGTGGGAAGAGTTTCTTTCACTCT TGTTATCTATATTGTCAGCGTGGACTGTTTATAACTGTACCAACTT AGTTTCTTTCAACTCCAGGTTAAGAGACATAAATGTCCTTTGATG CTGACAATAATCAGTGGGAATCAAGGAAGGACAATCCCGACCTC AATCTGTTTCAATTAATGAAGAGTTTCAATCGTCTTAAATCAAGCG CTAGACTCAATTGTCAATGAGAACCCTTTCTTTGACCAAGAACT ATAAATAGATCGAATGACAAAGTTGGAAATGAGTCCATTAGCTTA CATGATATTGAGCAGGCAGACCAAAATAAACCCTCTTTGAGAG CGATATTGATGGTTTCGGCGCCGTTGATAAGAGACGACAAATTGCC AAAGAAACAAAGCTGGGGGCTGAGCAATTTTTTTTCAAGAAGAA ATAGCATATGTTTACCACTACATGAAAATGATTCAAGTGTGTGTA AGACCGAAAGATCTATTGCAGTGGGAACACCCCATCTTCAATACT GCTTCAATGGAATCTCCAATGCCAAGTACAATGCATTTACCTTTT TCCCAGTCATCCTATACGAGCAATTCAAATTTTTTTTCAATTTATA

		CTTTACTTTAGTGGCTCTCTCTCAAGCGATACCGCAACTTCGCATT GGATATCTTTCTTCGTATGTCGTCCCACTTTTGTGTGACTCATAG TGACCATGTCAAAAGAGGCGATGGATGATATTCAACGCCGAAGA AGGGATAGAGAACAGAACAAATGAACCATATGAGGTTCTGTCCAG CCCATCACCAGTTTTGTCCAAAACTTAAATGTGGTCACTTGGT TCGATTGCATAAGGGAATGAGAGTGCCCGCAGATATGGTTCCTGT CCAGTCAAGCGAATCCACCGGAGAGTCATTTATCAAGACAGATCA GCTGGATGGTGAGACTGATTGGAAGCTTCGGATTGTTTCTCCAGT TACACAATCGTTACCAATGACTGAACTTCAAAATGTCGCCATCAC TGCAAGCGCACCCCTCAAAATCAATTCACTCCTTTCTTGGAAGATT GACCTACAATGGGCAATCATATGGTCTTACGATAGACAACACAAT GTGGTGTAATACTGTATTAGCTTCTGGTTCAGCAATTGCTGTAT AATTTACACAGGTAAAGATACTCGACAATCGATGAACACAACCTCA GCCCCAACTGAAAACGGGCTTGTTAGAAGTGGAAATCAATAGTTT GTCCAAGATCTTATGTGTTTGTGTGTTTGCATTATCTGTCTATCTTA GTGCTATTCCAAGGAATAGCTGATGATTGGTACGTCGATATCATG CGGTTTCTCATTCTATTCTCCACTATTATCCAGTGTCTCTGAGAG TTAACCTTGATCTTGGAAGTCAGTCCATGCTCATCAAATAGAAA CTGATAGCTCAATACCTGAAACCGTTGTTAGAAGTAGTACAATAC CGGAAGACCTGGGAAGAATTGAATACCTATTAAGTGACAAAACCT GGAACCTTACTCAAAATGATATGGAAATGAAAAAACTACACCTA GGAACAGTCTCTTATGCTGGTGATACCATGGATATTATTTCTGAT CATGTTAAAGGTCTTAATAACGCTAAAACATCGAGGAAAGATCTT GGTATGAGAATAAGAGATTGGTTACAACCTCTGGCCATCTG
89	Secuencia del marcador auxotrófico de PpHIS1:	CAAGTTGCGTCCGGTATACGTAACGTCTCACGATGATCAAAGATA ATACTTAATCTTCATGGTCTACTGAATAACTCATTTAAACAATTG ACTAATTGTACATTATATTGAACTTATGCATCCTATTAACGTAATC TTCTGGCTTCTCTCTCAGACTCCATCAGACACAGAATATCGTTCTC TCTAACTGGTCCTTTGACGTTTCTGACAAATAGTTCTAGAGGAGTC GTCCAAAACTCAACTCTGACTTGGGTGACACCACCACGGGATCC GGTCTTCCGAGGACCTTGATGACCTTGGCTAATGTAACCTGGAGT TTAGTATCCATTTTAAAGATGTGTGTTTCTGTAGGTTCTGGGTTGG AAAAAAATTTTAGACACCAGAAGAGAGGAGTGAACCTGGTTTGCG TGGGTTTAGACTGTGTAAGGCACTACTCTGTGCAAGTTTATAGATA GGGGTTACCCGCTCCGATGCATGGGAAGCGATTAGCCCGGCTGTT GCCCCGTTTGGTTTTTGAAGGGTAATTTTCAATATCTCTGTTTGAGT CATCAATTTCATATTCAAAGATTCAAAAACAAAATCTGGTCCAAG GAGCGCATTTAGGATTATGGAGTTGGCGAATCACTTGAACGATA GACTATTATTTGCTGTTCTTAAAGAGGGCAGATTGTATGAGAAAT GCGTTGAATTACTTAGGGGATCAGATATTCAGTTTCGAAGATCCA GTAGATTGGATATAGCTTTGTGCACTAACCTGCCCTGGCATTGG TTTTCTTCCAGCTGCTGACATTCCCACGTTTGTAGGAGAGGGTA AATGTGATTTGGGTATAACTGGTATTGACCAGGTTGAGGAAAGTG ACGTAGATGTCATACCTTTATTAGACTTGAATTTCCGTAAGTGCA AGTTGCAGATTCAAGTTCCCGAGAATGGTGACTTGAAAGAACCTA AACAGCTAATTGGTAAAGAAATTGTTTCTCTTACTAGCTTAA CCACCAGTACTTTGAACAACTGGAAGGAGTTAAGCCTGGTGAG CCACTAAAGACAAAAATCAAATATGTTGGAGGGTCTGTTGAGGC

		CTCTTGTGCCCTAGGAGTTGCCGATGCTATTGTGGATCTTGTGGA GAGTGGAGAAACCATGAAAGCGGCAGGGCTGATCGATATTGAAA CTGTTCTTTCTACTTCCGCTTACCTGATCTCTTCGAAGCATCCTCA ACACCCAGAACTGATGGATACTATCAAGGAGAGAATTGAAGGTG TACTGACTGCTCAGAAGTATGTCTTGTGTAATTACAACGCACCTA GAGGTAACCTTCTCAGCTGCTAAAACTGACTCCAGGCAAGAGA GCTGCTACCGTTTCTCCATTAGATGAAGAAGATTGGGTGGGAGTG TCCTCGATGGTAGAGAAGAAAGATGTTGGAAGAATCATGGACGA ATTAAAGAAACAAGGTGCCAGTGACATTCTTGTCTTTGAGATCAG TAATTGTAGAGCATAAGATAAGAATAATATTCAAGACCAACGGCTTC TCTTCGGAAGCTCCAAGTAGCTTATAGTGATGAGTACCGGCATAT ATTTATAGGCTTAAATTTTCGAGGGTTCACTATATTCTGTTTAGTG GGAAGAGTTCCTTTCACTCTTGTTATCTATATTGTCAGCGTGGAC TGTTTATAACTGTACCAACTTAGTTTCTTTCAACTCCAGGTTAAGA GACATAAATGTCTTTGATGC
90	Secuencia de la Región en la posición 5' que se usó para adición genética en el locus de PpPRO1:	GAAGGGCCATCGAATTGTCATCGTCTCCTCAGGTGCCATCGCTGT GGGCATGAAGAGAGTCAACATGAAGCGGAAACCAAAAAAGTTAC AGCAAGTGCAGGCATTGGCTGCTATAGGACAAGGCCGTTTGATA GGACTTTGGGACGACCTTTTCCGTCAGTTGAATCAGCCTATTGCG CAGATTTTACTGACTAGAACGGATTTGGTTCGATTACACCCAGTTT AAGAACGCTGAAAATACATTGGAACAGCTTATTAATAATGGGTATT ATTCCTATTGTCAATGAGAATGACACCCTATCCATTCAAGAAATC AAATTTGGTGACAATGACACCTTATCCGCCATAACAGCTGGTATG TGTCATGCAGACTACCTGTTTTTGGTGACTGATGTGGACTGTCTT TACACGGATAACCCCTCGTACGAATCCGGACGCTGAGCCAATCGTG TTAGTTAGAAATATGAGGAATCTAAACGTCAATACCGAAAGTGG AGGTTCCGCCGTAGGAACAGGAGGAATGACAACTAAATTGATCG CAGCTGATTTGGGTGTATCTGCAGGTGTTACAACGATTATTTGCA AAAGTGAACATCCCGAGCAGATTTTGGACATTGTAGAGTACAGTA TCCGTGCTGATAGAGTCGAAAATGAGGCTAAATATCTGGTCATCA ACGAAGAGGAACTGTGGAACAATTTCAAGAGATCAATCGGTCA GAACTGAGGGAGTTGAACAAGCTGGACATTCTTTGCATACACGT TTCGTTGGCCACAGTTTTAATGCTGTTAATAACAAAGAGTTTTGG TACTCCATGGACTAAAGGCCAACGGAGCCATTATCATTGATCCA GGTTGTTATAAGGCTATCACTAGAAAAAACAAAGCTGGTATTCTT CCAGCTGGAATTATTTCCGTAGAGGGTAATTTCCATGAATACGAG TGTGTTGATGTAAAGGTAGGACTAAGAGATCCAGATGACCCACAT TCACTAGACCCCAATGAAGAACTTTACGTCGTTGGCCGTGCCCGT TGTAATTACCCAGCAATCAAATCAACAAAATTAAGGGTCTACAA AGCTCGCAGATCGAGCAGGTTCTAGGTTACGCTGACGGTGAGTAT GTTGTTACAGGGACAACCTGGCTTTCCAGTATTTGCCGATCCA GAACTGTTGGATGTTGTTGAGAGTACCCTGTCTGAACAGGAGAG AGAATCCAAACCAATAAATAG
91	Secuencia de la Región en la posición 3' que se usó para adición genética en el	AATTTACATATGCTGCTTGATTATGTAATTATACCTTGCGTTCGA TGGCATCGATTTCTCTTCTGTCAATCGCGCATCGCATTAAAAAGT ATACTTTTTTTTTTTTCTTATAGTACTATTTCGCCTTATTATAAACTT TGCTAGTATGAGTTCTACCCCCAAGAAAGAGCCTGATTTGACTCC TAAGAAGAGTCAGCCTCCAAAGAATAGTCTCGGTGGGGGTAAAG GCTTTAGTGAGGAGGGTTTCTCCCAAGGGGACTTCAGCGCTAAGC

	locus de APPRO1:	ATATACTAAATCGTCGCCCTAACACCGAAGGCTCTTCTGTGGCTT CGAACGTCATCAGTTCGTTCATCATTGCAAAGGTTACCATCCTCTG GATCTGGAAGCGTTGCTGTGGGAAGTGTGTTGGGATCTTCGCCAT TAACTCTTTCTGGAGGGTTCCACGGGCTTGATCCAACCAAGAATA AAATAGACGTTCCAAAGTCGAAACAGTCAAGGAGACAAAGTGTT CTTTCTGACATGATTTCCACTTCTCATGCAGCTAGAAATGATCACT CAGAGCAGCAGTTACAACTGGACAACAATCAGAACAAAAAGAA GAAGATGGTAGTCGATCTTCTTTTTCTGTTTCTTCCCCCGCAAGA GATATCCGGCACCCAGATGTACTGAAAACGTGTCGAGAAACATCTT GCCAATGACAGCGAGATCGACTCATCTTTACAACCTCAAGGTGGA GATGTCCTAGAGGCATTTATCAATGGGTAAGTGGAGAAAGTAG TCAAAAAGATAACCCGCCTTTGAAACGAGCAAATAGTTTAAATGA TTTTTCTTCTGTGCATGGTGACGAGGTAGGCAAGGCAGATGCTGA CCACGATCGTGAAAGCGTATTCGACGAGGATGATATCTCCATTGA TGATATCAAAGTTCCGGGAGGGATGCGTCGAAGTTTATTATACA AAAGCATAGAGACCAACAACCTTTCTGGACTGAATAAAACGGCTC ACCAACCAAAAACAACCTTACTAAACCTAATTTCTTCACGAACAAC TTATAGAGTTTTTGGCATTGTATGGGCATTTTGCAGGTGAAGATT TGGAGGAAGACGAAGATGAAGATTTAGACAGTGGTTCGAATCA GTCGCAGTCAGTGATAGTGAGGGAGAATTCAGTGAGGCTGACAA CAATTTGTTGTATGATGAAGAGTCTCTCCTATTAGCACCTAGTAC CTCCAACCTATGCGAGATCAAGAATAGGAAGTATTCGTACTCTAC TTATGGATCTTTCAGTTCAAATGTTGGTTCTTCGTCTATTTCATCAG CAGTTAATGAAAAGTCAAATCCCGAAGCTGAAGAAACGTGGACA GCACAAGCATAAAACACAATCAAAAATACGCTCGAAGAAGCAAA CTACCACCGTAAAAGCAGTGTTGCTGCTATTAAA
92	ADN de hEPO truncado (optimizado por codón)	GCTCCACCAAGATTGATTTGTGACTCCAGAGTTTTGGAGAGATAC TTGTTGGAGGCTAAAGAGGCTGAGAACATCACTACTGGTTGTGCT GAACACTGTTCTTGAACGAGAACATCACAGTTCAGACACTAAG GTAACTTCTACGCTTGAAGAGAAATGGAAGTTGGACAACAGGC TGTTGAAGTTTGGCAAGGATTGGCTTTGTTGTCCGAGGCTGTTTT GAGAGGTCAAGCTTTGTTGGTTAACTCCTCCCAACCATGGGAACC ATTGCAATTGCACGTTGACAAGGCTGTTTCTGGATTGAGATCCTT GACTACTTTGTTGAGAGCTTTGGGTGCTCAGAAAGAGGCTATTTT TCCACCAGATGCTGCTTCAGCTGCTCCATTGAGAACTATCACTGC TGACACTTTCAGAAAGTTGTTTCAGAGTTTACTCCAACCTTCTTGAG AGGAAAGTTGAAGTTGTACACTGGTGAAGCTTGTAGAACTGGTG ACTAGTAA
93	Proteína hEPO truncada	APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDTKVNFYA WKRMEVGQQA VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLR ALGAQKEAIS PPDAASAAPL RTITADTFRK LFRVYSNFLR GKLLKLYTGEA CRTGD
94	ADN señal de lisosoma de pollo (CLSP)	ATGCTGGGTAAGAACGACCCAATGTGTCTTGTGTTTGGTCTTGTG GGATTGACTGCTTTGTTGGGTATCTGTCAAGGT
95	péptido señal de	MLGKNPDMCLVLVLLGLTALLGICQG
	lisosoma de pollo (CLSP)	

96	Secuencia del gen PpAde2 sin su promotor pero incluyendo sus secuencias de terminación	ATGGATTCTCAGGTAATAGGTATTCTAGGAGGAGGCCAGCTAGG CCGAATGATTGTTGAGGCCGCTAGCAGGCTCAATATCAAGACCGT GATTCTTGATGATGGTTTTTCACCTGCTAAGCACATTAATGCTGC GCAAGACCACATCGACGGATCATTCAAAGATGAGGAGGCTATCG CCAAGTTAGCTGCCAAATGTGATGTTCTCACTGTAGAGATTGAGC ATGTCAACACAGATGCTCTAAAGAGAGTTCAAGACAGAACTGGA ATCAAGATATATCCTTTACCAGAGACAATCGAACTAATCAAGGAT AAGTACTTGCAAAAGGAACATTTGATCAAGCACAACTTTCGGTG ACAAAGTCTCAGGGTATAGAATCTAATGAAAAGGCGCTGCTTTTG TTTGGAGAAGAGAATGGATTTCATATCTGTTGAAGTCCCGGACT ATGGCTTATGATGGAAGAGGCAATTTTGTAGTGGAGTCTAAAGA GGACATCAGTAAGGCATTAGAATTCTTGAAAGATCGTCCATTGTA TGCCGAGAAGTTTGCTCCTTTTGTTAAAGAATTAGCGGTAATGGT TGTGAGATCACTGGAAGGCGAAGTATTCTCTACCCAACCGTAGA AACTGTGCACAAGGACAATATCTGTCTATTTGTGTATGCTCCGGC CAGAGTTAATGACACCATCCAAAAGAAAGCTCAAATATTAGCTGA AAACACTGTGAAGACTTTCCAGGCGCTGGAATCTTCGGAGTTGA GATGTTCTATTGTCTGATGGAGAACTTCTTGTAATGAGATTGC TCCAAGGCCCCACAATTCTGGTCACTATACAATCGATGCATGTGT AACATCTCAGTTCGAAGCACATGTAAGAGCCATAACTGGTCTGCC AATGCCACTAGATTTACCAAATCTACTTCCAACACCAACGC TATTATGCTCAATGTTTTGGGTGCTGAAAAATCTCACGGGAATT AGAGTTTTGTAGAAGAGCCTTAGAAACACCCGGTGCTTCTGTATA TCTGTACGGAAAGACCACCCGATTGGCTCGTAAGATGGGTCTAT CAACATAATAGGATCTTCCATGTTGGAAGCAGAACAAAAGTTAG AGTACATTCTAGAAGAATCAACCCACTTACCATCCAGTACTGTAT CAGCTGACACTAAACCGTTGGTTGGAGTTATCATGGGTTCACT CTGATCTACCTGTGATTTGAAAGGTTGCGATATTTTAAACAGT TTGGTGTTCATTGGAAGTTACTATTGTCTCTGCTCATAGAACACC ACAGAGAATGACCAGATATGCCTTTGAAGCCGCTAGTAGAGGTA TCAAGGCTATCATTGCAGGTGCTGGTGGTGCTGCTCATCTTCCAG GAATGGTTGCTGCCATGACTCCGTTGCCAGTCATTGGTGTTCCTG TCAAGGGCTCTACGTTGGATGGTGTAGACTCGCTACACTCGATTG TCCAAATGCCTAGAGGTGTTCTGTGGCTACGGTTGCTATCAACA ACGCCACCAATGCCGCTCTGTTGGCCATCAGGATTTTAGGTACAA TTGACCACAAATGGCAAAGGAAATGTCCAAGTATATGAATGCA ATGGAGACCGAAGTGTGGGGAAGGCATCCAACCTGGAATCTGA AGGGTATGAATCCTATTTGAAGAATCGTCTTTGAATTTAGTATTGTT TTTTAATAGATGTATATATAATAGTACACGTAACCTTATCTATTCCATTCA TAATTTTATTTTAAAGGTTTCGGTAGAAATTTGTCTCCAAAAGTTGGT TAGAGCCTGGCAGTTTTGATAGGCATTATTATAGATTGGGTAATATTT ACCTGACCTGGAGGAACTTTGCAAAGAGCCTCATGTGC
97	PpADE2	MDSQVIGILGGGQLGRMIVEAASRLNIKTIVLDDGFSPAKHINAAQD HIDGSFKDEEAIKLAACDVLTVEIEHVNTDALKRVQDRGTGIKYP LPETIELIKDKYLQKEHLIKHNISVTKSQGIESNEKALLFGEENGFPY LLKSRTMAYDGRGNFVVESKEDISKALEFLKDRPLYAEKFAPFVKEL
		AVMVVRSLEGEVFSYPTVETVHKDNICHIVYAPARVNDTIQKKAQIL AENTVKTFPGAGIFGVEMFLSDGELLVNEIAPRPHNSGHYTIDACV TSQFEAHVRAITGLPMPLDFTKLSTSNNAIMNLVGAEKSHGELEF CRRALETPGASVYLYGKTTTLARKMGHINIIGSSMLEAEQKLEYILE ESTHLPSSSTVSADTKPLVGVIMGSDSDLPVISKGCDILKQFVPEFT IVSAHRTPQRMTRYAFEAAASRGIKAIAGAGGAHLPGMVAAMTLP VIGVPVKGSTLDGVDLSIVQMPRGVPVATVAINNATNAALLAIRI LGTIDHKWQKEMSKYMNAMETEVLGKASNLESEGYESYLKNRL

98	Pp TRP2: 5' y ORF	ACTGGGCCTTTAGAGGGTGCTGAAGTTGACCCCTTGGTGCTTCTG GAAAAAGAACTGAAGGGCACCAGACAAGCGCAACTTCCTGGTAT TCCTCGTCTAAGTGGTGGTGCCATAGGATACATCTCGTACGATTG TATTAAGTACTTTGAACCAAAAACTGAAAGAAAACTGAAAGATGT TTTGCAACTTCCGGAAGCAGCTTTGATGTTGTTGACACGATCGT GGCTTTTGACAATGTTTATCAAAGATTCCAGGTAATTGGAAACGT TTCTCTATCCGTTGATGACTCGGACGAAGCTATTCTTGAGAAATA TTATAAGACAAGAGAAGAAGTGGAAGATCAGTAAAGTGGTAT TTGACAATAAACTGTTCCCTACTATGAACAGAAAGATATTATTC AAGGCCAAACGTTACCTCTAATATTGGTCAGGAAGGGGATGAA AACCATGTTTCGCAAGCTGAAAGAACATATTCTGAAAGGACAT CTTCCAAGCTGTTCCCTCTCAAAGGGTAGCCAGGCCGACCTCATT GCACCCCTTCAACATCTATCGTCATTTGAGAACTGTCAATCCTTCT CCATACATGTTCTATATTGACTATCTAGACTTCCAAGTTGTTGGTG CTTCACCTGAATTACTAGTTAAATCCGACAACAACAACAAAATCA TCACACATCCTATTGCTGGAACCTTCCCAGAGGTAAAACTATCG AAGAGGACGACAATTATGCTAAGCAATTGAAGTCGTCTTTGAAA GACAGGGCCGAGCACGTCATGCTGGTAGATTTGCCAGAAATGA TATTAACCGTGTGTGTGAGCCCACCAGTACCACGGTTGATCGTTT ATTGACTGTGGAGAGATTTTCTCATGTGATGCATCTTGTGTCAGA AGTCAGTGGAACATTGAGACCAAACAAGACTCGCTTCGATGCTTT CAGATCCATTTTCCCAGCAGGTACCGTCTCCGGTGCTCCGAAGGT AAGAGCAATGCAACTCATAGGAGAATTGGAAGGAGAAAAGAGAG GTGTTTATGCGGGGGCCGTAGGACACTGGTCGTACGATGGAAAA TCGATGGACACATGTATTGCCTTAAGAACAATGGTCGTCAAGGAC GGTGTGCTTACCTTCAAGCCGGAGGTGGAATTGTCTACGATTCT GACCCCTATGACGAGTACATCGAAACCATGAACAAAATGAGATC CAACAATAACACCATCTTGGAGGCTGAGAAAATCTGGACCGATA GGTTGGCCAGAGACGAG AATCAAAGTGAATCCGAAGAAAACGATCAATGA
99	región en la posición 3' de PpTRP2	ACGGAGGACGTAAGTAGGAATTTATGTAATCATGCCAATACATCT TTAGATTTCTTCCTCTTCTTTTAAACGAAAGACCTCCAGTTTTCGA CTCTCGACTCTCTAGTATCTTCCCATTCTGTGTGCTGCAACCTCTT GCCTTCTGTTTCCTTCAATTGTTCTTCTTCTTCTGTTGCACTTGG CCTTCTTCCTCCATCTTTCGTTTTTTTTCAAGCCTTTTCAGCAGTTC TTCTTCCAAGAGCAGTTCTTTGATTTTCTCTCTCCAATCCACCAAA AACTGGATGAATTCAACCGGGCATCATCAATGTTCCACTTTCTT TCTCTTATCAATAATCTACGTGCTTCGGCATAACGAGGAATCCAGT TGCTCCCTAATCGAGTCATCCACAAGGTTAGCATGGGCCTTTTTC AGGGTGTCAAAGCATCTGGAGCTCGTTTATTCGGAGTCTTGTCT GGATGGATCAGCAAAGACTTTTTGCGGAAAGTCTTCTTATATCT
		TCCGGAGAACAACCTGGTTTCAAATCCAAGATGGCATAGCTGTCC AATTTGAAAGTGGAAGAATCCTGCCAATTTCTTCTCTCGTGTC AGCTCGTTCTCCTCCTTTTGCAACAGGTCCACTTCATCTGGCATTT TTCTTTATGTTAACTTTAATTATTATTAATTATAAAGTTGATTATC GTTATCAAAATAATCATATTCGAGAAATAATCCGTCCATGCAATA TATAAATAAGAATTCATAATAATGTAATGATAACAGTACCTCTGA TGACCTTTGATGAACCGCAATTTTCTTTCCAATGACAAGACATCC CTATAATACAATTATACAGTTTATATATCACAATAATCACCTTTT TATAAGAAAACCGTCCTCTCCGTAACAGAACTTATTATCCGCACG TTATGGTTAACACACTACTAATACCGATATAGTGTATGAAGTCGC TACGAGATAGCCATCCAGGAACTTACCAATTCATCAGCACTTTC ATGATCCGATTGTTGGCTTTATTCTTTGCGAGACAGATACTTGCC AATGAAATAACTGATCCCACAGATGAGAATCCGGTGCTCGT

100	región en la posición 5' de Pp ADE2	CTTAAAATCATCTGCCTCACCCACCGACCAATGGGAATTCTAGA AACAAATTTCAATTGCTCTTCTTCTCGTTACCATAAGAATCGGCTGTC ATGTTTGACTTAACGAACCCTGGAACAAGGGAATTCACGGTAATA CCTTTTGGAGCAAGTTCAACCGATAGAGCCTTCATTAATGAGTTG ATTGCACCTTTGGTGGTCGCATATACCGATTGATTCGGGTAGGTC ACTTCGAACTGTACAGGGAGGCAGTAAAGATGATCCTACCCTTA ATCTGGTTCTTAATAAAGTGTTTAGTGACTAGCTGTGTCAATCTA AATGGAAAATCGACATTTACCTTTTGGATAGCCGCGTAATCTTTC TCCGTAAAACCTTGTAACCTCAGATTTAATGGCAATGGCAGCGTTG TTGATTAAAATGTCAATCTTCCAGTGGAACCTTCTCCACCGCA GGACTCGTTACGGTCTCTTCCAGCTTTGCAAGATCGGCATCCACT AGATCCAACCTCAATTGTATGTATGGAGGCACCATCGGCATTTGAC ATTCTCACCTCTTCAATGAAAGCCGTTGGGTCTGTAGAAGGTCTA TGGATAAGAATAAGTTCTGCACCTGCTTCATAAAGTCCTCGAACT ATTCTTGGCCTAATCCGCTGGTACCACCGGTGATCAAGGCGACC TTACCATTCAAAGAAAAACAAATCAGCGGACATTAGCGCATTTGAAT AGGGAATGGGTTAGACAAATGAAAGCCGACGAGCCAGCACTTTA TAGTAAGTGCAGGTGAGTCAATAAGAATAAATGTATGGCTTGCTG TCCCTATCGCGTAAGAAGCTTACTAAGATCGCCTAAATTGAAAAG TTGAACAAATCAGTTCTAGCTGGCCTCCATCAGCATTTTCGTTCTC CTCTGATCATCTTTGCCAATCGCTAGCATGCCCTCAGCGTGCAAG GAAAAGCACGCTTCTTTCTTATCGACGTATTTTCAACTATGGCAG AGCCAGGTTAGCAAGTC
101	región en la posición 3' de Pp ADE2	ATTTAGTATTGTTTTTTAATAGATGTATATATAATAGTACACGTAA CTTATCTATTCCATTTCATAATTTTATTTTAAAGGTTTCGGTAGAAAT TTGTCCTCCAAAAAGTTGGTTAGAGCCTGGCAGTTTTGATAGGCA TTATTATAGATTGGGTAATATTTACCCTGCACCTGGAGGAACTTT GCAAAGAGCCTCATGTGCTCTAAAAGGATGTCAGAAATCCAACAT TTCAAAATTATATCTGCATGCGTCTGTAATACTGGAACGTGTTATTT TTCTGGTCAGGATTTACCCGCTCTTGTCGTCATGTTTCTCGTCGTC TGAAAGTAACTGACTTTCTCTTTCCATAAACACAAAAATCGAT TGCAACTTGGTTATTCTTGAGATTGAAATTTGCTGTGTCTTCAGT GCTTAGCTGAATATCAACAACTTACTTAGTACTAATAACGAAGC ACTATGGTAAGTGGCATAACATAGTGGTATTGAAGCGAACAGTG GATATTGAACCCAAGCATTGGCAACATCTGGCTCTGTTGATACTG ATCCGGATCGTTTGGCACCAATTCCTGAAACGGCGTAGTGCCACC

		AAGGTTTCGATTTGAGAACAGGTTTCATCATCAGAGTCAACCACCC CAATGTCAATGGCAGGCTCCAACGAAGTAGGTCCAACAACA GGAAGTATTTGACCTTGAAGATCTGTTCTTTATGATCCACCACA CCTTGCCCCAATTCCAATAACTTTACCAGTCCCGATGCAGACATG ATAACTGGTACTAATGATCTCCATTGATTTTCGTGCGGCACTACGT AAAGCCTCCAAAAATGAATTCAGAATATCTTCTGAAACTAGATTCT TGCTTCTGTGATTCAAGCATTGCTTTATGTAGACATCTCTTGAATA AAAGCAATTCTCCACATATTGGTGTGTGTAAGATAGATCTGGAAA GATGTATCTGGAATAGTCCAGTCAACGTTGTGCAATTGATTAGCA TTACCTTACTGTGAACATCTCTATCTACAACAACAGACTCAATTC GATAGACGTTCCGGGAAAGTTTTTCAAGCGCATTGATTGCTGT TGAACAAAGTGACTTTGCTTTCCAATGTGCAAATACCCCTGTATA TCAAGTCCATCACATCACTCAAGACCTTGGTGGAAAAGAATGAAA CAGCTGGAGCATAATTTTTCGAATGAATTAGGTAAGGTCACCTTCAT CCTTATCTGTTGTAATGCTATAATCAATAGCGGAACATAACATCTT CCCATGTAAACAGGTTTCTTGATCTCTGAATCTGAATCTTTATTTGA AAAAGAATTGAAAAAAGACTCATCACTCATTGGGAATTCGAAGGT CATTAGGGTATTCCATTGTTAGTTCTGGTCTAGGTTTAAAGGGAT CACCTTCGTTAAGACGATGGAAAATAGCTAATCTGTACAATAACC AGATACTTCTAACGAAGCTCTCTCTATCCATCAGTTGACGTGTTG AGGATATCTGAAGTCTCTTTCCACTGCGAATCAGGCATGCTCG TATAGCTGGCAAGCATGTTATTCAGCTTTACCAAGTTAGAAGCCC TTTGAAACCATCTATAGATTCCCGAAAAAACTTATACCCACTGA GGGTTTCACTGAGCATAGTCAGTGACATCAAAGAGCATTTCAAAT CCATCTCA
102	ORF de NATR	ATGGGTACCACTCTTGACGACACGGCTTACCGGTACCGCACCAGT GTCCCGGGGGACGCCGAGGCCATCGAGGCACTGGATGGGTCTT CACCACCGACACCGTCTTCCGCGTCACCGCCACCGGGGACGGCTT CACCCTGCGGGAGGTGCCGGTGGACCCGCCCTGACCAAGGTGT TCCCCGACGACGAATCGGACGACGAATCGGACGACGGGGAGGAC GGCGACCCGGACTCCCGGACGTTCTGTCGCGTACGGGGACGACGG CGACCTGGCGGGCTTCTGTTGTCGTCTCGTACTCCGGCTGGAACCG CCGGCTGACCGTTCGAGGACATCGAGGTGCGCCCGGAGCACCGGG GGCACGGGGTTCGGGCGCGCTTGATGGGGCTCGCGACGGAGTTT GCCCCGAGCGGGGGCGCCGGCACCTCTGGCTGGAGGTCACCAA CGTCAACGCACCGGCGATCCACGCGTACCGGCGGATGGGGTTCA CCCTCTGCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGG ACGGCGAGCAGGCGCTCTACATGAGCATGCCCTGCCCTAATCAG TACTG
103	ORF de HygR	ATGGGTAAAAAGCCTGAACTACCGCGACGTCTGTGAGAAAGTTT CTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCG GAGGGCGAAGAATCTCGTGCTTTTACGCTTCGATGTAGGAGGGCG TGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAA AGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGAT TCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCT ATTGCATCTCCCGCCGTGCACAGGGTGTACGTTGCAAGACCTGC CTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTTCGCGGAGGCC ATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTT CGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGC

		GTGATTTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGC AAACTGTGATGGACGACACCGTCAGTGCGTCCGTCCGCGCAGGCT CTCGATGAGCTGATGCTTTGGGCGGAGGACTGCCCCGAAGTCCG GCACCTCGTGACGCGGATTTTCGGCTCCAACAATGTCTTGACGGA CAATGGCCGCATAACAGCGGTCAATTGACTGGAGCGAGGCGATGT TCGGGGATTCCCAATACGAGGTGCGCAACATCTTCTTCTGGAGGC CGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGG AGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATAT GCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGG CAATTTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAA TCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCC CGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACT CGCCGATAGTGGAACCGACGCCCCAGCACTCGTCCGAGGGCAA AGGAATAG
104	<i>región de PpPEP4</i> (incluyendo fragmento de supresión genética cadena arriba, promotor, marco de lectura abierto y fragmento de supresión genética cadena abajo)	ATTTGAGTCACCTGCTTTAGGGCTGGAAGATATTTGGTTACTAGA TTTTAGTACAACTCTTGCTTTGTCAATGACATTAATAAGGCAA GAATCGCAAACTCAAATATTTTCATGGAGATGAGATATGCTTGT CAAAGATGCCAGAAAAAGAGCAACTCGTTTATAGGGTTCATAT TGATGATGGAACAGGCCTTTTCCAGGGAGGTGAAAGAACCCAAG CCAATTCTGATGACATTCTGGATATTGATGAGGTTGATGAAAAGT TAAGAGAACTATTGACAAGAGCCTCAAGGAAACGGCATATCACC CCTGCATTGGAACTCCTGATAAACGTGTAAAAAGAGCTTATTTG AACAGTATTACTGATAACTCTTGATGGACCTTAAAGATGTATAAT AGTAGACAGAATTCATAATGGTGAGATTAGGTAATCGTCCGGAA TAGGAATAGTGGTTTGGGGCGATTAATCGCACCTGCCTTATATGG TAAGTACCTTGACCGATAAGGTGGCAACTATTTAGAACAAAGCAA GCCACCTTTCTTTATCTGTAACCTCTGTCGAAGCAAGCATCTTTACT AGAGAACATCTAAACCATTTTACATTCTAGAGTTCCATTTCTCAA TTACTGATAATCAATTTAAAGATGATATTTGACGGTACTACGATG TCAATTGCCATTGGTTTGCTCTCTACTCTAGGTATTGGTGCTGAA GCCAAAGTTCATTCTGCTAAGATACACAAGCATCCAGTCTCAGAA ACTTTAAAAGAGGCCAATTTTGGGCAGTATGTCTCTGCTCTGGAA CATAAATATGTTTCTCTGTTCACGAACAAAATGCTTTGTCCAAG TCGAATTTTATGTCTCAGCAAGATGGTTTTGCCGTTGAAGCTTCG CATGATGCTCCACTTACAACTATCTTAACGCTCAGTATTTTACTG AGGTATCATTAGGTACCCCTCCACAATCGTTCAAGGTGATTCTTG ACACAGGATCCTCCAATTTATGGGTTTCTAGCAAAGATTGTGGAT CATTAGCTTGCTTCTTGATGCTAAGTATGACCATGATGAGTCTT CTACTTATAAGAAGAATGGTAGTAGCTTTGAAATTAGGTATGGAT CCGGTTCCATGGAAGGGTATGTTTCTCAGGATGTGTTGCAAATTG GGGATTTGACCATTTCCAAAGTTGATTTTGCTGAGGCCACATCGG AGCCGGGGTTGGCCTTCGCTTTTGGCAAATTTGACGGAATTTTGG GGCTTGCTTATGATTCAATATCAGTAAATAAGATTGTTCTCTCAA TTTACAAGGCTTTGGAATTAGATCTCCTTGACGAACCAAAATTTG CCTTCTACTTGGGGGATACGGACAAAGATGAATCCGATGGCGGTT TGGCCACATTTGGTGGTGTGGACAAATCTAAGTATGAAGGAAAG ATCACCTGGTGCCTGTCAGAAAGAAAGGCTTACTGGGAGGTCTCT TTTGATGGTGTAGGTTTGGGATCCGAATATGCTGAATTGCAAAAA ACTGGTGCAGCCATCGACACTGGAACCTCATTGATTGCTTTGCCC

		AGTGGCCTAGCTGAAATTCTCAATGCAGAAATTGGTGCTACCAAG GGTTGGTCTGGTCAATACGCTGTGGACTGTGACACTAGAGACTCT TTGCCAGACTTAACTTTAACCTTCGCCGGTTACAACCTTACCATT CTCCATATGACTATACTTTGGAGGTTTCTGGGTCATGTATTAGTG CTTTCACCCCCATGGACTTTCTGAAACCAATAGGTCCTTTGGCAA TCATTGGTGACTCGTTCTTGAGAAAATATTACTCAGTTTATGACC TAGGCAAAGATGCAGTAGGTTTAGCCAAGTCTATTTAGGCAAGA ATAAAAGTTGCTCAGCTGAACTTATTTGGTTACTTATCAGGTAGT GAAGATGTAGAGAATATATGTTTAGGTATTTTTTTTAGTTTTCT CCTATAACTCATCTTCAGTACGTGATTGCTTGTGAGCTACCTTGAC AGGGGCGCATAAGTGATATCGTGTACTGCTCAATCAAGATTTGCC TGCTCCATTGATAAGGGTATAAGAGACCCACCTGCTCCTCTTAA AATTCTCTCTTAACTGTTGTGAAAATCATCTTCGAAGCAAATTCG AGTTTAAATCTATGCGGTTGGTAACTAAAGGTATGTCATGGTGGT ATATAGTTTTTCATTTTACCTTTTACTAATCAGTTTTACAGAAGAG GAACGTCTTTCTCAAGATCGAAATAGGACTAAATACTGGAGACG ATGGGGTCCTTATTTGGGTGAAAGGCAGTGGGCTACAGTAAGGG AAGACTATTCCGATGATGGAGATGCTTGGTCTGCTTTTCTTTTG AGCAATCTCATTGAGAACTTATCGCTGGGGAGAGGATGGACTA GCTGGAGTCTCAGACAATCATCAACTAATTTGTTTCTCAATGGCA CTGTGGAATGAGAATGATGATATTTGAAGGAGCGATTATTTGGG GTCAGTGGAGAGGCTGCAAATCATGGAGAGGATGTTAAGGAGCT TTATTATTATCTTGATAATACACCTTCTCACTCTTATATGAAATAC CTTTACAAATATCCACAATCGAAATTTCTTACGAAGAATTGATT TCAGAGAACCGTAAACGTTCCAGATTAGAAAGAGAGTACGAGAT TACTGACTCTGAAGTACTGAAGGATAACAGATATTTTGATGTGAT CTTTGAAATGGCAAAGGACGATGAAGATGAGAATGAACTTTACTT TAGAATTACCGCTTACAACCGAGGTCCACCCCTGCCCTTTACA TGTCGCTCCACAGGTAACCTTTAGAAATACCTGGTCTGGGGTAT AGATGAGGAAAAGGATCACGACAAACCTATAGCTTGCAAGGAAT ACCAAGACAACAACCTATTCTATTTCGGTTAGATAGTT
105	Promotor de <i>TEF1</i> de <i>Ashbya gossypii</i>	GATCTGTTTAGCTTGCCTCGTCCCCGCCGGGTCAACCGGCCAGCG ACATGGAGGCCCAAGAATACCCTCCTTGACAGTCTTGACGTGCGCA GCTCAGGGGCATGATGTGACTGTGCGCCGTACATTTAGCCCATAC ATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCA CGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAG CAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCC GCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAG GTTCTTCTTTTATATACTTCTTTTAAATCTTGCTAGGATACAGT TCTCACATCACATCCGAACATAAACAACC
106	Secuencia de terminación de <i>TEF1</i> de <i>Ashbya gossypii</i>	TAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAAGTTGT CATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAA TGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATGTGCCA GATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCG TATGTGAATGCTGGTCGCTATACTGCTGTGATTGATACTAACG CCGCCATCCAGTGTGAAAAC

Listado de secuencias

5	<110> Bobrowicz, Piotr Gomathinayagam, Sujatha Hamilton, Stephen Li, Huijuan Setheraman, Natarajan Stadheim, Terrance A. Wildt, Stephan
10	<120> MÉTODO PARA PRODUCIR PROTEÍNAS EN <i>PICHA PASTORIS</i> QUE CARECEN DE ACTIVIDAD DE UNIÓN CRUZADA DETECTABLE A ANTICUERPOS FRENTE A ANTÍGENOS DE CÉLULA HOSPEDADORA
15	<130> GFI-MIS-00002
	<150> 61/252,312 <151> 16-10-2009
20	<160> 106
	<170> FastSEQ for Windows Versión 4.0
25	<210> 1 <211> 3029 <212> ADN <213> <i>Saccharomyces cerevisiae</i>
30	<220> <221> CDS <222> (909)...(2504) <223> Codifica invertasa
	<400> 1


```

aggcctcgca acaacctata attgagttaa gtgcctttcc aagctaaaaa gtttgagggtt 60
ataggggctt agcatccaca cgtcacaaatc tcgggtatcg agtatagtat gtagaattac 120
ggcaggaggt ttcccaatga acaaaggaca ggggcacggg gagctgtcga aggtatccat 180
tttatcatgt ttcgtttgta caagcacgac atactaagac atttaccgta tgggagttgt 240
tgtcctagcg tagttctcgc tccccagca aagctcaaaa aagtacgtca tttagaatag 300
tttgtgagca aattaccagt cggtagtcta cgttagaaaag gccacacagta ttcttctacc 360
aaaggcgtgc ctttggtgaa ctcgatccat tatgagggtt tccattattc ccgcattttt 420
tattactctg aacaggaata aaaagaaaaa acccagttta ggaaattatc cgggggcgaa 480
gaaatacgcg tagcgtaaat cgaccccacg tccagggttt ttccatggag gtttctggaa 540
aaactgacga ggaatgtgat tataaatccc tttatgtgat gtctaagact ttttaaggta 600
gcccgatggt tgcctattac catcatagag acgtttcctt tcgaggaatg cttaaacgac 660
tttgtttgac aaaaatgttg cctaagggtt ctatagtaaa ccatttgga gaaagatttg 720
acgacttttt ttttttggtt ttcgatccta taatccttcc tcttgaaaag aaacatataa 780
atagatatgt attattcttc aaaacattct cttgttcttg tgcttttttt ttaccatata 840
tcttactttt ttttttctct cagagaaaaca agcaaaaaca aaagcttttc ttttactaa 900
cgtatatg atg ctt ttg caa gct ttc ctt ttc ctt ttg gct ggt ttt gca 950
      Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala
        1             5             10

gcc aaa ata tct gca tca atg aca aac gaa act agc gat aga cct ttg 998
Ala Lys Ile Ser Ala Ser Met Thr Asn Glu Thr Ser Asp Arg Pro Leu
  15             20             25             30

gtc cac ttc aca ccc aac aag ggc tgg atg aat gac cca aat ggg ttg 1046
Val His Phe Thr Pro Asn Lys Gly Trp Met Asn Asp Pro Asn Gly Leu
          35             40             45

tgg tac gat gaa aaa gat gcc aaa tgg cat ctg tac ttt caa tac aac 1094
Trp Tyr Asp Glu Lys Asp Ala Lys Trp His Leu Tyr Phe Gln Tyr Asn

```

ES 2 589 655 T3

50								55					60						
cca	aat	gac	acc	gta	tgg	ggt	acg	cca	ttg	ttt	tgg	ggc	cat	gct	act		1142		
Pro	Asn	Asp	Thr	Val	Trp	Gly	Thr	Pro	Leu	Phe	Trp	Gly	His	Ala	Thr				
		65					70					75							
tcc	gat	gat	ttg	act	aat	tgg	gaa	gat	caa	ccc	att	gct	atc	gct	ccc		1190		
Ser	Asp	Asp	Leu	Thr	Asn	Trp	Glu	Asp	Gln	Pro	Ile	Ala	Ile	Ala	Pro				
	80					85					90								
aag	cgt	aac	gat	tca	ggt	gct	ttc	tct	ggc	tcc	atg	gtg	gtt	gat	tac		1238		
Lys	Arg	Asn	Asp	Ser	Gly	Ala	Phe	Ser	Gly	Ser	Met	Val	Val	Asp	Tyr				
	95				100					105					110				
aac	aac	acg	agt	ggg	ttt	ttc	aat	gat	act	att	gat	cca	aga	caa	aga		1286		
Asn	Asn	Thr	Ser	Gly	Phe	Phe	Asn	Asp	Thr	Ile	Asp	Pro	Arg	Gln	Arg				
				115					120					125					
tgc	gtt	gcg	att	tgg	act	tat	aac	act	cct	gaa	agt	gaa	gag	caa	tac		1334		
Cys	Val	Ala	Ile	Trp	Thr	Tyr	Asn	Thr	Pro	Glu	Ser	Glu	Glu	Gln	Tyr				
			130					135					140						
att	agc	tat	tct	ctt	gat	ggt	ggt	tac	act	ttt	act	gaa	tac	caa	aag		1382		
Ile	Ser	Tyr	Ser	Leu	Asp	Gly	Gly	Tyr	Thr	Phe	Thr	Glu	Tyr	Gln	Lys				
		145				150						155							
aac	cct	gtt	tta	gct	gcc	aac	tcc	act	caa	ttc	aga	gat	cca	aag	gtg		1430		
Asn	Pro	Val	Leu	Ala	Ala	Asn	Ser	Thr	Gln	Phe	Arg	Asp	Pro	Lys	Val				
		160				165					170								
ttc	tgg	tat	gaa	cct	tct	caa	aaa	tgg	att	atg	acg	gct	gcc	aaa	tca		1478		
Phe	Trp	Tyr	Glu	Pro	Ser	Gln	Lys	Trp	Ile	Met	Thr	Ala	Ala	Lys	Ser				
					180					185					190				
caa	gac	tac	aaa	att	gaa	att	tac	tcc	tct	gat	gac	ttg	aag	tcc	tgg		1526		
Gln	Asp	Tyr	Lys	Ile	Glu	Ile	Tyr	Ser	Ser	Asp	Asp	Leu	Lys	Ser	Trp				
				195					200					205					
aag	cta	gaa	tct	gca	ttt	gcc	aat	gaa	ggt	ttc	tta	ggc	tac	caa	tac		1574		
Lys	Leu	Glu	Ser	Ala	Phe	Ala	Asn	Glu	Gly	Phe	Leu	Gly	Tyr	Gln	Tyr				
			210					215				220							
gaa	tgt	cca	ggt	ttg	att	gaa	gtc	cca	act	gag	caa	gat	cct	tcc	aaa		1622		
Glu	Cys	Pro	Gly	Leu	Ile	Glu	Val	Pro	Thr	Glu	Gln	Asp	Pro	Ser	Lys				
		225				230						235							
tct	tat	tgg	gtc	atg	ttt	att	tct	atc	aac	cca	ggt	gca	cct	gct	ggc		1670		
Ser	Tyr	Trp	Val	Met	Phe	Ile	Ser	Ile	Asn	Pro	Gly	Ala	Pro	Ala	Gly				
		240				245						250							
ggt	tcc	ttc	aac	caa	tat	ttt	gtt	gga	tcc	ttc	aat	ggt	act	cat	ttt		1718		
Gly	Ser	Phe	Asn	Gln	Tyr	Phe	Val	Gly	Ser	Phe	Asn	Gly	Thr	His	Phe				
		255			260					265					270				
gaa	gcg	ttt	gac	aat	caa	tct	aga	gtg	gta	gat	ttt	ggt	aag	gac	tac		1766		
Glu	Ala	Phe	Asp	Asn	Gln	Ser	Arg	Val	Val	Asp	Phe	Gly	Lys	Asp	Tyr				
				275					280					285					
tat	gcc	ttg	caa	act	ttc	ttc	aac	act	gac	cca	acc	tac	ggt	tca	gca		1814		
Tyr	Ala	Leu	Gln	Thr	Phe	Phe	Asn	Thr	Asp	Pro	Thr	Tyr	Gly	Ser	Ala				
			290					295					300						
tta	ggt	att	gcc	tgg	gct	tca	aac	tgg	gag	tac	agt	gcc	ttt	gtc	cca		1862		
Leu	Gly	Ile	Ala	Trp	Ala	Ser	Asn	Trp	Glu	Tyr	Ser	Ala	Phe	Val	Pro				
		305					310					315							

ES 2 589 655 T3

```

act aac cca tgg aga tca tcc atg tct ttg gtc cgc aag ttt tct ttg 1910
Thr Asn Pro Trp Arg Ser Ser Met Ser Leu Val Arg Lys Phe Ser Leu
320 325 330

aac act gaa tat caa gct aat cca gag act gaa ttg atc aat ttg aaa 1958
Asn Thr Glu Tyr Gln Ala Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys
335 340 345 350

gcc gaa cca ata ttg aac att agt aat gct ggt ccc tgg tct cgt ttt 2006
Ala Glu Pro Ile Leu Asn Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe
355 360 365

gct act aac aca act cta act aag gcc aat tct tac aat gtc gat ttg 2054
Ala Thr Asn Thr Thr Leu Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu
370 375 380

agc aac tcg act ggt acc cta gag ttt gag ttg gtt tac gct gtt aac 2102
Ser Asn Ser Thr Gly Thr Leu Glu Phe Glu Leu Val Tyr Ala Val Asn
385 390 395

acc aca caa acc ata tcc aaa tcc gtc ttt gcc gac tta tca ctt tgg 2150
Thr Thr Gln Thr Ile Ser Lys Ser Val Phe Ala Asp Leu Ser Leu Trp
400 405 410

ttc aag ggt tta gaa gat cct gaa gaa tat ttg aga atg ggt ttt gaa 2198
Phe Lys Gly Leu Glu Asp Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu
415 420 425 430

gtc agt gct tct tcc ttc ttt ttg gac cgt ggt aac tct aag gtc aag 2246
Val Ser Ala Ser Ser Phe Phe Leu Asp Arg Gly Asn Ser Lys Val Lys
435 440 445

ttt gtc aag gag aac cca tat ttc aca aac aga atg tct gtc aac aac 2294
Phe Val Lys Glu Asn Pro Tyr Phe Thr Asn Arg Met Ser Val Asn Asn
450 455 460

caa cca ttc aag tct gag aac gac cta agt tac tat aaa gtg tac ggc 2342
Gln Pro Phe Lys Ser Glu Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly
465 470 475

cta ctg gat caa aac atc ttg gaa ttg tac ttc aac gat gga gat gtg 2390
Leu Leu Asp Gln Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val
480 485 490

gtt tct aca aat acc tac ttc atg acc acc ggt aac gct cta gga tct 2438
Val Ser Thr Asn Thr Tyr Phe Met Thr Thr Gly Asn Ala Leu Gly Ser
495 500 505 510

gtg aac atg acc act ggt gtc gat aat ttg ttc tac att gac aag ttc 2486
Val Asn Met Thr Thr Gly Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe
515 520 525

caa gta agg gaa gta aaa tagaggttat aaaacttatt gtctttttta 2534
Gln Val Arg Glu Val Lys
530

tttttttcaa aagccattct aaagggtttt agctaacgag tgacgaatgt aaaactttat 2594
gatttcaaag aatacctcca aaccattgaa aatgtatttt tattttttatt ttctcccgac 2654
cccagttacc tgggaatttgt tctttatgta ctttatataa gtataattct cttaaaaaatt 2714
tttactactt tgcaatagac atcatttttt cacgtaataa acccacaatc gtaatgtagt 2774
tgccttacac tactaggatg gacctttttg cctttatctg ttttgttact gacacaatga 2834
aacggggtaa agtattagtt atgtgaaaaa ttaaaagcat taagtagaag tataccatat 2894
tgtaaaaaaa aaaagcgttg tcttctacgt aaaagtgttc tcaaaaagaa gtagtgaggg 2954
aaatggatac caagctatct gtaacaggag ctaaaaaaatc tcagggaataa gcttctggtt 3014

```

tgggaaacgg tcgac

3029

5 <210> 2
<211> 532
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 2

Met	Leu	Leu	Gln	Ala	Phe	Leu	Phe	Leu	Leu	Ala	Gly	Phe	Ala	Ala	Lys
1				5					10					15	
Ile	Ser	Ala	Ser	Met	Thr	Asn	Glu	Thr	Ser	Asp	Arg	Pro	Leu	Val	His
			20					25					30		
Phe	Thr	Pro	Asn	Lys	Gly	Trp	Met	Asn	Asp	Pro	Asn	Gly	Leu	Trp	Tyr
		35					40					45			
Asp	Glu	Lys	Asp	Ala	Lys	Trp	His	Leu	Tyr	Phe	Gln	Tyr	Asn	Pro	Asn
	50					55					60				
Asp	Thr	Val	Trp	Gly	Thr	Pro	Leu	Phe	Trp	Gly	His	Ala	Thr	Ser	Asp
65					70					75					80
Asp	Leu	Thr	Asn	Trp	Glu	Asp	Gln	Pro	Ile	Ala	Ile	Ala	Pro	Lys	Arg
			85						90					95	
Asn	Asp	Ser	Gly	Ala	Phe	Ser	Gly	Ser	Met	Val	Val	Asp	Tyr	Asn	Asn
			100					105					110		
Thr	Ser	Gly	Phe	Phe	Asn	Asp	Thr	Ile	Asp	Pro	Arg	Gln	Arg	Cys	Val
		115					120					125			
Ala	Ile	Trp	Thr	Tyr	Asn	Thr	Pro	Glu	Ser	Glu	Glu	Gln	Tyr	Ile	Ser
	130					135					140				
Tyr	Ser	Leu	Asp	Gly	Gly	Tyr	Thr	Phe	Thr	Glu	Tyr	Gln	Lys	Asn	Pro
145					150					155					160
Val	Leu	Ala	Ala	Asn	Ser	Thr	Gln	Phe	Arg	Asp	Pro	Lys	Val	Phe	Trp
				165					170					175	
Tyr	Glu	Pro	Ser	Gln	Lys	Trp	Ile	Met	Thr	Ala	Ala	Lys	Ser	Gln	Asp
			180					185					190		
Tyr	Lys	Ile	Glu	Ile	Tyr	Ser	Ser	Asp	Asp	Leu	Lys	Ser	Trp	Lys	Leu
	195					200						205			
Glu	Ser	Ala	Phe	Ala	Asn	Glu	Gly	Phe	Leu	Gly	Tyr	Gln	Tyr	Glu	Cys
	210				215						220				
Pro	Gly	Leu	Ile	Glu	Val	Pro	Thr	Glu	Gln	Asp	Pro	Ser	Lys	Ser	Tyr
225					230					235					240
Trp	Val	Met	Phe	Ile	Ser	Ile	Asn	Pro	Gly	Ala	Pro	Ala	Gly	Gly	Ser
			245						250					255	
Phe	Asn	Gln	Tyr	Phe	Val	Gly	Ser	Phe	Asn	Gly	Thr	His	Phe	Glu	Ala
			260					265					270		
Phe	Asp	Asn	Gln	Ser	Arg	Val	Val	Asp	Phe	Gly	Lys	Asp	Tyr	Tyr	Ala
		275					280					285			
Leu	Gln	Thr	Phe	Phe	Asn	Thr	Asp	Pro	Thr	Tyr	Gly	Ser	Ala	Leu	Gly
	290				295						300				
Ile	Ala	Trp	Ala	Ser	Asn	Trp	Glu	Tyr	Ser	Ala	Phe	Val	Pro	Thr	Asn
305					310					315					320
Pro	Trp	Arg	Ser	Ser	Met	Ser	Leu	Val	Arg	Lys	Phe	Ser	Leu	Asn	Thr
			325						330					335	
Glu	Tyr	Gln	Ala	Asn	Pro	Glu	Thr	Glu	Leu	Ile	Asn	Leu	Lys	Ala	Glu
		340						345					350		
Pro	Ile	Leu	Asn	Ile	Ser	Asn	Ala	Gly	Pro	Trp	Ser	Arg	Phe	Ala	Thr
		355					360					365			
Asn	Thr	Thr	Leu	Thr	Lys	Ala	Asn	Ser	Tyr	Asn	Val	Asp	Leu	Ser	Asn
	370				375						380				
Ser	Thr	Gly	Thr	Leu	Glu	Phe	Glu	Leu	Val	Tyr	Ala	Val	Asn	Thr	Thr
385					390					395					400
Gln	Thr	Ile	Ser	Lys	Ser	Val	Phe	Ala	Asp	Leu	Ser	Leu	Trp	Phe	Lys
			405						410					415	
Gly	Leu	Glu	Asp	Pro	Glu	Glu	Tyr	Leu	Arg	Met	Gly	Phe	Glu	Val	Ser
			420					425					430		
Ala	Ser	Ser	Phe	Phe	Leu	Asp	Arg	Gly	Asn	Ser	Lys	Val	Lys	Phe	Val
		435					440					445			
Lys	Glu	Asn	Pro	Tyr	Phe	Thr	Asn	Arg	Met	Ser	Val	Asn	Asn	Gln	Pro

ES 2 589 655 T3

450		455		460											
Phe	Lys	Ser	Glu	Asn	Asp	Leu	Ser	Tyr	Tyr	Lys	Val	Tyr	Gly	Leu	Leu
465					470					475					480
Asp	Gln	Asn	Ile	Leu	Glu	Leu	Tyr	Phe	Asn	Asp	Gly	Asp	Val	Val	Ser
				485					490					495	
Thr	Asn	Thr	Tyr	Phe	Met	Thr	Thr	Gly	Asn	Ala	Leu	Gly	Ser	Val	Asn
			500					505					510		
Met	Thr	Thr	Gly	Val	Asp	Asn	Leu	Phe	Tyr	Ile	Asp	Lys	Phe	Gln	Val
		515				520						525			
Arg	Glu	Val	Lys												
530															

<210> 3
 <211> 2159
 <212> ADN
 <213> *K. lactis*

<220>
 <221> CDS
 <222> (1024)...(2007)
 <223> Codifica transportador de UDP-GlcNAc (KIMNN2-2)

<400> 3

```

aaacgtaacg cctggcactc tattttctca aacttctggg acggaagagc taaatattgt 60
gttgcttgaa caaacccaaa aaaacaaaaa aatgaacaaa ctaaaactac acctaaataa 120
accgtgtgta aaacgtagta ccatattact agaaaagatc acaagtgtat cacacatgtg 180
catctcatat tacatctttt atccaatcca ttctctctat cccgtctgtt cctgtcagat 240
tctttttcca taaaaagaag aagaccccgga atctcacccg tacaatgcaa aactgctgaa 300
aaaaaaagaa agttcactgg atacgggaac agtgccagta ggcttcacca catggacaaa 360
acaattgacg ataaaaataag caggtgagct tctttttcaa gtcacgatcc ctttatgtct 420
cagaaacaat atatacaagc taaacccttt tgaaccagtt ctctcttcat agttatgttc 480
acataaattg cggaacaag actccgctgg ctgtcaggta cacgttgtaa cgttttcgtc 540
cgcccaatta ttagcacaac attggcaaaa agaaaaactg ctcgttttct ctacaggtaa 600
attacaattt ttttcagtaa ttttcgctga aaaattttaa gggcaggaaa aaaagacgat 660
ctcgactttg catagatgca agaactgtgg tcaaaacttg aaatagtaat tttgctgtgc 720
gtgaactaat aaatatatat atatatatat atatatattt gtgtattttg tatatgtaat 780
tgtgcacgtc ttggctattg gatataagat tttcgcggtg tgatgacata gagcgtgtac 840
tactgtaata gttgtatatt caaaagctgc tgcgtggaga aagactaaaa tagataaaaa 900
gcacacattt tgacttcggt accgtcaact tagtgggaca gtcttttata tttggtgtaa 960
gctcatttct ggtactattc gaaacagaac agtgttttct gtattaccgt ccaatcgttt 1020
gtc atg agt ttt gta ttg att ttg tgc tta gtg ttc gga gga tgt tgt 1068
Met Ser Phe Val Leu Ile Leu Ser Leu Val Phe Gly Gly Cys Cys
    1             5             10             15

tcc aat gtg att agt ttc gag cac atg gtg caa ggc agc aat ata aat 1116
Ser Asn Val Ile Ser Phe Glu His Met Val Gln Gly Ser Asn Ile Asn
                20             25             30

ttg gga aat att gtt aca ttc act caa ttc gtg tct gtg acg cta att 1164
Leu Gly Asn Ile Val Thr Phe Thr Gln Phe Val Ser Val Thr Leu Ile
                35             40             45

cag ttg ccc aat gct ttg gac ttc tct cac ttt ccg ttt agg ttg cga 1212
Gln Leu Pro Asn Ala Leu Asp Phe Ser His Phe Pro Phe Arg Leu Arg
                50             55             60

cct aga cac att cct ctt aag atc cat atg tta gct gtg ttt ttg ttc 1260
Pro Arg His Ile Pro Leu Lys Ile His Met Leu Ala Val Phe Leu Phe
                65             70             75

ttt acc agt tca gtc gcc aat aac agt gtg ttt aaa ttt gac att tcc 1308
Phe Thr Ser Ser Val Ala Asn Asn Ser Val Phe Lys Phe Asp Ile Ser
    80             85             90             95

```

ES 2 589 655 T3

gtt ccg att cat att atc att aga ttt tca ggt acc act ttg acg atg	1356
Val Pro Ile His Ile Ile Ile Arg Phe Ser Gly Thr Thr Leu Thr Met	
100 105 110	
ata ata ggt tgg gct gtt tgt aat aag agg tac tcc aaa ctt cag gtg	1404
Ile Ile Gly Trp Ala Val Cys Asn Lys Arg Tyr Ser Lys Leu Gln Val	
115 120 125	
caa tct gcc atc att atg acg ctt ggt gcg att gtc gca tca tta tac	1452
Gln Ser Ala Ile Ile Met Thr Leu Gly Ala Ile Val Ala Ser Leu Tyr	
130 135 140	
cgt gac aaa gaa ttt tca atg gac agt tta aag ttg aat acg gat tca	1500
Arg Asp Lys Glu Phe Ser Met Asp Ser Leu Lys Leu Asn Thr Asp Ser	
145 150 155	
gtg ggt atg acc caa aaa tct atg ttt ggt atc ttt gtt gtg cta gtg	1548
Val Gly Met Thr Gln Lys Ser Met Phe Gly Ile Phe Val Val Leu Val	
160 165 170 175	
gcc act gcc ttg atg tca ttg ttg tcc ttg ctc aac gaa tgg acg tat	1596
Ala Thr Ala Leu Met Ser Leu Leu Ser Leu Leu Asn Glu Trp Thr Tyr	
180 185 190	
aac aag tac ggg aaa cat tgg aaa gaa act ttg ttc tat tcc cat ttc	1644
Asn Lys Tyr Gly Lys His Trp Lys Glu Thr Leu Phe Tyr Ser His Phe	
195 200 205	
ttg gct cta ccg ttg ttt atg ttg ggg tac aca agg ctc aga gac gaa	1692
Leu Ala Leu Pro Leu Phe Met Leu Gly Tyr Thr Arg Leu Arg Asp Glu	
210 215 220	
ttc aga gac ctc tta att tcc tca gac tca atg gat att cct att gtt	1740
Phe Arg Asp Leu Leu Ile Ser Ser Asp Ser Met Asp Ile Pro Ile Val	
225 230 235	
aaa tta cca att gct acg aaa ctt ttc atg cta ata gca aat aac gtg	1788
Lys Leu Pro Ile Ala Thr Lys Leu Phe Met Leu Ile Ala Asn Asn Val	
240 245 250 255	
acc cag ttc att tgt atc aaa ggt gtt aac atg cta gct agt aac acg	1836
Thr Gln Phe Ile Cys Ile Lys Gly Val Asn Met Leu Ala Ser Asn Thr	
260 265 270	
gat gct ttg aca ctt tct gtc gtg ctt cta gtg cgt aaa ttt gtt agt	1884
Asp Ala Leu Thr Leu Ser Val Val Leu Leu Val Arg Lys Phe Val Ser	
275 280 285	
ctt tta ctc agt gtc tac atc tac aag aac gtc cta tcc gtg act gca	1932
Leu Leu Leu Ser Val Tyr Ile Tyr Lys Asn Val Leu Ser Val Thr Ala	
290 295 300	
tac cta ggg acc atc acc gtg ttc ctg gga gct ggt ttg tat tca tat	1980
Tyr Leu Gly Thr Ile Thr Val Phe Leu Gly Ala Gly Leu Tyr Ser Tyr	
305 310 315	
ggt tcc gtc aaa act gca ctg cct cgc tgaacaatc cacgtctgta	2027
Gly Ser Val Lys Thr Ala Leu Pro Arg	
320 325	
tgatactcgt ttcagaattt ttttgatttt ctgccggata tggttttctca tctttacaat	2087
cgcattctta attataccag aacgtaattc aatgatccca gtgactcgta actcttatat	2147
gtcaatttaa gc	2159

<210> 4
 <211> 328
 <212> PRT
 <213> *K. lactis*

5

<400> 4

```

Met Ser Phe Val Leu Ile Leu Ser Leu Val Phe Gly Gly Cys Cys Ser
 1           5           10           15
Asn Val Ile Ser Phe Glu His Met Val Gln Gly Ser Asn Ile Asn Leu
      20           25           30
Gly Asn Ile Val Thr Phe Thr Gln Phe Val Ser Val Thr Leu Ile Gln
      35           40           45
Leu Pro Asn Ala Leu Asp Phe Ser His Phe Pro Phe Arg Leu Arg Pro
      50           55           60
Arg His Ile Pro Leu Lys Ile His Met Leu Ala Val Phe Leu Phe Phe
65           70           75           80
Thr Ser Ser Val Ala Asn Asn Ser Val Phe Lys Phe Asp Ile Ser Val
      85           90           95
Pro Ile His Ile Ile Ile Arg Phe Ser Gly Thr Thr Leu Thr Met Ile
      100          105          110
Ile Gly Trp Ala Val Cys Asn Lys Arg Tyr Ser Lys Leu Gln Val Gln
      115          120          125
Ser Ala Ile Ile Met Thr Leu Gly Ala Ile Val Ala Ser Leu Tyr Arg
      130          135          140
Asp Lys Glu Phe Ser Met Asp Ser Leu Lys Leu Asn Thr Asp Ser Val
145          150          155          160
Gly Met Thr Gln Lys Ser Met Phe Gly Ile Phe Val Val Leu Val Ala
      165          170          175
Thr Ala Leu Met Ser Leu Leu Ser Leu Leu Asn Glu Trp Thr Tyr Asn
      180          185          190
Lys Tyr Gly Lys His Trp Lys Glu Thr Leu Phe Tyr Ser His Phe Leu
      195          200          205
Ala Leu Pro Leu Phe Met Leu Gly Tyr Thr Arg Leu Arg Asp Glu Phe
      210          215          220
Arg Asp Leu Leu Ile Ser Ser Asp Ser Met Asp Ile Pro Ile Val Lys
225          230          235          240
Leu Pro Ile Ala Thr Lys Leu Phe Met Leu Ile Ala Asn Asn Val Thr
      245          250          255
Gln Phe Ile Cys Ile Lys Gly Val Asn Met Leu Ala Ser Asn Thr Asp
      260          265          270
Ala Leu Thr Leu Ser Val Val Leu Leu Val Arg Lys Phe Val Ser Leu
      275          280          285
Leu Leu Ser Val Tyr Ile Tyr Lys Asn Val Leu Ser Val Thr Ala Tyr
      290          295          300
Leu Gly Thr Ile Thr Val Phe Leu Gly Ala Gly Leu Tyr Ser Tyr Gly
305          310          315          320
Ser Val Lys Thr Ala Leu Pro Arg
      325

```

10 <210> 5
 <211> 108
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Codifica líder de Mnn2 (53)

<400> 5

atgctgctta ccaaaagggtt ttcaaagctg ttcaagctga cgttcatagt tttgatattg 60
tgcgggctgt tcgtcattac aaacaaatac atggatgaga acacgtcg 108

<210> 6
<211> 36
<212> PRT
<213> Secuencia artificial

<220>
<223> líder de Mnn2 (53)

<400> 6

Met	Leu	Leu	Thr	Lys	Arg	Phe	Ser	Lys	Leu	Phe	Lys	Leu	Thr	Phe	Ile
1				5				10						15	
Val	Leu	Ile	Leu	Cys	Gly	Leu	Phe	Val	Ile	Thr	Asn	Lys	Tyr	Met	Asp
			20					25					30		
Glu	Asn	Thr	Ser												
			35												

<210> 7
<211> 300
<212> ADN
<213> Secuencia artificial

<220>
<223> Codifica líder de Mnn2 (54). Los últimos 9 nucleótidos son el conector que contiene el sitio de restricción de Ascl)

<400> 7

atgctgctta ccaaaagggtt ttcaaagctg ttcaagctga cgttcatagt tttgatattg 60
tgcgggctgt tcgtcattac aaacaaatac atggatgaga acacgtcggt caaggagtac 120
aaggagtact tagacagata tgtccagagt tactccaata agtattcatc ttcttcagac 180
gccgccagcg ctgacgattc aacccattg agggacaatg atgaggcagg caatgaaaag 240
ttgaaaagct tctacaacaa cgtttttcaac tttctaattg ttgattcgcc cggggcgcgcc 300

<210> 8
<211> 100
<212> PRT
<213> Secuencia artificial

<220>
<223> líder de Mnn2 (54).

<400> 8

ES 2 589 655 T3

```

Met Leu Leu Thr Lys Arg Phe Ser Lys Leu Phe Lys Leu Thr Phe Ile
 1          5          10          15
Val Leu Ile Leu Cys Gly Leu Phe Val Ile Thr Asn Lys Tyr Met Asp
 20          25          30
Glu Asn Thr Ser Val Lys Glu Tyr Lys Glu Tyr Leu Asp Arg Tyr Val
 35          40          45
Gln Ser Tyr Ser Asn Lys Tyr Ser Ser Ser Ser Asp Ala Ala Ser Ala
 50          55          60
Asp Asp Ser Thr Pro Leu Arg Asp Asn Asp Glu Ala Gly Asn Glu Lys
 65          70          75          80
Leu Lys Ser Phe Tyr Asn Asn Val Phe Asn Phe Leu Met Val Asp Ser
 85          90          95
Pro Gly Arg Ala
          100

```

<210> 9
 <211> 57
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Codifica secuencia pre señal del Factor de Acoplamiento a *S. cerevisiae*
 <400> 9
 atgagattcc catcatctt cactgctgtt ttgtcgctg cttctctgc ttggct 57

<210> 10
 <211> 19
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Secuencia pre señal del Factor de Acoplamiento a *S. cerevisiae*
 <400> 10

```

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1          5          10          15
Ala Leu Ala

```

<210> 11
 <211> 99
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Codifica Pp SEC12 (10)

<400> 11
 atgccagaa aaatatataa ctacttcatt ttgactgtat tcatggcaat tcttgctatt 60
 gttttacaat ggtctataga gaatggacat gggcgcgcc 99

<210> 12
 <211> 33
 <212> PRT
 <213> Secuencia artificial

<220>

ES 2 589 655 T3

<223> Pp SEC12 (10)

<400> 12

Met	Pro	Arg	Lys	Ile	Phe	Asn	Tyr	Phe	Ile	Leu	Thr	Val	Phe	Met	Ala
1				5				10					15		
Ile	Leu	Ala	Ile	Val	Leu	Gln	Trp	Ser	Ile	Glu	Asn	Gly	His	Gly	Arg
			20					25					30		
Ala															

5

<210> 13

<211> 183

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Codifica ScMnt1 (Kre2) (33)

15

<400> 13

atggccctct	ttctcagtaa	gagactgttg	agatttaccg	tcattgcagg	tgcggttatt	60
gttctcctcc	taacattgaa	ttccaacagt	agaactcagc	aatatattcc	gagttccatc	120
tccgctgcat	ttgattttac	ctcaggatct	atatcccctg	aacaacaagt	catcgggcgc	180
gcc						183

<210> 14

<211> 61

<212> PRT

<213> Secuencia artificial

20

<220>

<223> ScMnt1 (Kre2) (33)

25

<400> 14

Met	Ala	Leu	Phe	Leu	Ser	Lys	Arg	Leu	Leu	Arg	Phe	Thr	Val	Ile	Ala
1				5				10						15	
Gly	Ala	Val	Ile	Val	Leu	Leu	Leu	Thr	Leu	Asn	Ser	Asn	Ser	Arg	Thr
			20					25					30		
Gln	Gln	Tyr	Ile	Pro	Ser	Ser	Ile	Ser	Ala	Ala	Phe	Asp	Phe	Thr	Ser
		35					40					45			
Gly	Ser	Ile	Ser	Pro	Glu	Gln	Gln	Val	Ile	Gly	Arg	Ala			
	50					55					60				

<210> 15

<211> 318

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Codifica ScSEC12 (8)

35

<400> 15

ES 2 589 655 T3

```

atgaacacta tccacataat aaaattaccg cttaactacg ccaactacac ctcaatgaaa 60
caaaaaatct ctaaattttt caccaacttc atccttattg tgctgctttc ttacatttta 120
cagttctcct ataagcaciaa tttgcattcc atgcttttca attacgcgaa ggacaatttt 180
ctaacgaaaa gagacacccat ctcttcgccc tacgtagtgt atgaagactt acatcaaaca 240
actttgtttg gcaaccacgg tacaanaaaca tctgtaccta gcgtagattc cataaaagtg 300
catggcgtgg ggcgcgcc                                     318

```

<210> 16
 <211> 106
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> ScSEC12 (8)

<400> 16

Met	Asn	Thr	Ile	His	Ile	Ile	Lys	Leu	Pro	Leu	Asn	Tyr	Ala	Asn	Tyr
1				5					10					15	
Thr	Ser	Met	Lys	Gln	Lys	Ile	Ser	Lys	Phe	Phe	Thr	Asn	Phe	Ile	Leu
			20					25					30		
Ile	Val	Leu	Leu	Ser	Tyr	Ile	Leu	Gln	Phe	Ser	Tyr	Lys	His	Asn	Leu
		35					40					45			
His	Ser	Met	Leu	Phe	Asn	Tyr	Ala	Lys	Asp	Asn	Phe	Leu	Thr	Lys	Arg
	50					55					60				
Asp	Thr	Ile	Ser	Ser	Pro	Tyr	Val	Val	Asp	Glu	Asp	Leu	His	Gln	Thr
65					70					75					80
Thr	Leu	Phe	Gly	Asn	His	Gly	Thr	Lys	Thr	Ser	Val	Pro	Ser	Val	Asp
			85						90					95	
Ser	Ile	Lys	Val	His	Gly	Val	Gly	Arg	Ala						
			100					105							

<210> 17
 <211> 981
 <212> ADN
 <213> *Mus musculus*

<220>
 <221> CDS
 <222> (1)...(978)
 <223> codifica transportador de UDP-GlcNAc de MmSLC35A3

<400> 17

ES 2 589 655 T3

atg tct gcc aac cta aaa tat ctt tcc ttg gga att ttg gtg ttt cag	48
Met Ser Ala Asn Leu Lys Tyr Leu Ser Leu Gly Ile Leu Val Phe Gln	
1 5 10 15	
act acc agt ctg gtt cta acg atg cgg tat tct agg act tta aaa gag	96
Thr Thr Ser Leu Val Leu Thr Met Arg Tyr Ser Arg Thr Leu Lys Glu	
20 25 30	
gag ggg cct cgt tat ctg tct tct aca gca gtg gtt gtg gct gaa ttt	144
Glu Gly Pro Arg Tyr Leu Ser Ser Thr Ala Val Val Val Ala Glu Phe	
35 40 45	
ttg aag ata atg gcc tgc atc ttt tta gtc tac aaa gac agt aag tgt	192
Leu Lys Ile Met Ala Cys Ile Phe Leu Val Tyr Lys Asp Ser Lys Cys	
50 55 60	
agt gtg aga gca ctg aat aga gta ctg cat gat gaa att ctt aat aag	240
Ser Val Arg Ala Leu Asn Arg Val Leu His Asp Glu Ile Leu Asn Lys	
65 70 75 80	
ccc atg gaa acc ctg aag ctc gct atc ccg tca ggg ata tat act ctt	288
Pro Met Glu Thr Leu Lys Leu Ala Ile Pro Ser Gly Ile Tyr Thr Leu	
85 90 95	
cag aac aac tta ctc tat gtg gca ctg tca aac cta gat gca gcc act	336
Gln Asn Asn Leu Leu Tyr Val Ala Leu Ser Asn Leu Asp Ala Ala Thr	
100 105 110	
tac cag gtt aca tat cag ttg aaa ata ctt aca aca gca tta ttt tct	384
Tyr Gln Val Thr Tyr Gln Leu Lys Ile Leu Thr Thr Ala Leu Phe Ser	
115 120 125	
gtg tct atg ctt ggt aaa aaa tta ggt gtg tac cag tgg ctc tcc cta	432
Val Ser Met Leu Gly Lys Lys Leu Gly Val Tyr Gln Trp Leu Ser Leu	
130 135 140	
gta att ctg atg gca gga gtt gct ttt gta cag tgg cct tca gat tct	480
Val Ile Leu Met Ala Gly Val Ala Phe Val Gln Trp Pro Ser Asp Ser	
145 150 155 160	
caa gag ctg aac tct aag gac ctt tca aca ggc tca cag ttt gta ggc	528
Gln Glu Leu Asn Ser Lys Asp Leu Ser Thr Gly Ser Gln Phe Val Gly	
165 170 175	
ctc atg gca gtt ctc aca gcc tgt ttt tca agt ggc ttt gct gga gtt	576
Leu Met Ala Val Leu Thr Ala Cys Phe Ser Ser Gly Phe Ala Gly Val	
180 185 190	
tat ttt gag aaa atc tta aaa gaa aca aaa cag tca gta tgg ata agg	624
Tyr Phe Glu Lys Ile Leu Lys Glu Thr Lys Gln Ser Val Trp Ile Arg	
195 200 205	
aac att caa ctt ggt ttc ttt gga agt ata ttt gga tta atg ggt gta	672
Asn Ile Gln Leu Gly Phe Phe Gly Ser Ile Phe Gly Leu Met Gly Val	
210 215 220	
tac gtt tat gat gga gaa ttg gtc tca aag aat gga ttt ttt cag gga	720
Tyr Val Tyr Asp Gly Glu Leu Val Ser Lys Asn Gly Phe Phe Gln Gly	
225 230 235 240	
tat aat caa ctg acg tgg ata gtt gtt gct ctg cag gca ctt gga ggc	768
Tyr Asn Gln Leu Thr Trp Ile Val Val Ala Leu Gln Ala Leu Gly Gly	

ES 2 589 655 T3

	245	250	255	
ctt gta ata gct gct gtc atc aaa tat gca gat aac att tta aaa gga				816
Leu Val Ile Ala Ala Val Ile Lys Tyr Ala Asp Asn Ile Leu Lys Gly				
	260	265	270	
ttt gcg acc tcc tta tcc ata ata ttg tca aca ata ata tct tat ttt				864
Phe Ala Thr Ser Leu Ser Ile Ile Leu Ser Thr Ile Ile Ser Tyr Phe				
	275	280	285	
tgg ttg caa gat ttt gtg cca acc agt gtc ttt ttc ctt gga gcc atc				912
Trp Leu Gln Asp Phe Val Pro Thr Ser Val Phe Phe Leu Gly Ala Ile				
	290	295	300	
ctt gta ata gca gct act ttc ttg tat ggt tac gat ccc aaa cct gca				960
Leu Val Ile Ala Ala Thr Phe Leu Tyr Gly Tyr Asp Pro Lys Pro Ala				
	305	310	315	320
gga aat ccc act aaa gca tag				981
Gly Asn Pro Thr Lys Ala				
	325			

<210> 18

<211> 326

5 <212> PRT

<213> *Mus musculus*

<400> 18

Met	Ser	Ala	Asn	Leu	Lys	Tyr	Leu	Ser	Leu	Gly	Ile	Leu	Val	Phe	Gln
1				5					10					15	
Thr	Thr	Ser	Leu	Val	Leu	Thr	Met	Arg	Tyr	Ser	Arg	Thr	Leu	Lys	Glu
			20					25					30		
Glu	Gly	Pro	Arg	Tyr	Leu	Ser	Ser	Thr	Ala	Val	Val	Val	Ala	Glu	Phe
		35					40					45			
Leu	Lys	Ile	Met	Ala	Cys	Ile	Phe	Leu	Val	Tyr	Lys	Asp	Ser	Lys	Cys
	50					55					60				
Ser	Val	Arg	Ala	Leu	Asn	Arg	Val	Leu	His	Asp	Glu	Ile	Leu	Asn	Lys
65					70					75					80
Pro	Met	Glu	Thr	Leu	Lys	Leu	Ala	Ile	Pro	Ser	Gly	Ile	Tyr	Thr	Leu
				85					90					95	
Gln	Asn	Asn	Leu	Leu	Tyr	Val	Ala	Leu	Ser	Asn	Leu	Asp	Ala	Ala	Thr
			100					105					110		
Tyr	Gln	Val	Thr	Tyr	Gln	Leu	Lys	Ile	Leu	Thr	Thr	Ala	Leu	Phe	Ser
		115					120					125			
Val	Ser	Met	Leu	Gly	Lys	Lys	Leu	Gly	Val	Tyr	Gln	Trp	Leu	Ser	Leu
		130				135					140				
Val	Ile	Leu	Met	Ala	Gly	Val	Ala	Phe	Val	Gln	Trp	Pro	Ser	Asp	Ser
145					150					155					160
Gln	Glu	Leu	Asn	Ser	Lys	Asp	Leu	Ser	Thr	Gly	Ser	Gln	Phe	Val	Gly
			165						170					175	
Leu	Met	Ala	Val	Leu	Thr	Ala	Cys	Phe	Ser	Ser	Gly	Phe	Ala	Gly	Val
			180					185					190		
Tyr	Phe	Glu	Lys	Ile	Leu	Lys	Glu	Thr	Lys	Gln	Ser	Val	Trp	Ile	Arg
		195					200					205			
Asn	Ile	Gln	Leu	Gly	Phe	Phe	Gly	Ser	Ile	Phe	Gly	Leu	Met	Gly	Val
	210					215					220				
Tyr	Val	Tyr	Asp	Gly	Glu	Leu	Val	Ser	Lys	Asn	Gly	Phe	Phe	Gln	Gly
225					230					235					240
Tyr	Asn	Gln	Leu	Thr	Trp	Ile	Val	Val	Ala	Leu	Gln	Ala	Leu	Gly	Gly
			245						250					255	
Leu	Val	Ile	Ala	Ala	Val	Ile	Lys	Tyr	Ala	Asp	Asn	Ile	Leu	Lys	Gly
			260					265					270		
Phe	Ala	Thr	Ser	Leu	Ser	Ile	Ile	Leu	Ser	Thr	Ile	Ile	Ser	Tyr	Phe
		275					280					285			
Trp	Leu	Gln	Asp	Phe	Val	Pro	Thr	Ser	Val	Phe	Phe	Leu	Gly	Ala	Ile
	290					295					300				
Leu	Val	Ile	Ala	Ala	Thr	Phe	Leu	Tyr	Gly	Tyr	Asp	Pro	Lys	Pro	Ala
305					310					315					320
Gly	Asn	Pro	Thr	Lys	Ala										
				325											

<210> 19

<211> 1074

<212> ADN

5 <213> *Drosophila melanogaster*

<220>

<221> CDS

<222> (1)...(1071)

10 <223> Codifica DmUGT

<400> 19

ES 2 589 655 T3

atg aat agc ata cac atg aac gcc aat acg ctg aag tac atc agc ctg	48
Met Asn Ser Ile His Met Asn Ala Asn Thr Leu Lys Tyr Ile Ser Leu	
1 5 10 15	
ctg acg ctg acc ctg cag aat gcc atc ctg ggc ctc agc atg cgc tac	96
Leu Thr Leu Thr Leu Gln Asn Ala Ile Leu Gly Leu Ser Met Arg Tyr	
20 25 30	
gcc cgc acc cgg cca ggc gac atc ttc ctc agc tcc acg gcc gta ctc	144
Ala Arg Thr Arg Pro Gly Asp Ile Phe Leu Ser Ser Thr Ala Val Leu	
35 40 45	
atg gca gag ttc gcc aaa ctg atc acg tgc ctg ttc ctg gtc ttc aac	192
Met Ala Glu Phe Ala Lys Leu Ile Thr Cys Leu Phe Leu Val Phe Asn	
50 55 60	
gag gag ggc aag gat gcc cag aag ttt gta cgc tgc ctg cac aag acc	240
Glu Glu Gly Lys Asp Ala Gln Lys Phe Val Arg Ser Leu His Lys Thr	
65 70 75 80	
atc att gcg aat ccc atg gac acg ctg aag gtg tgc gtc ccc tgc ctg	288
Ile Ile Ala Asn Pro Met Asp Thr Leu Lys Val Cys Val Pro Ser Leu	
85 90 95	
gtc tat atc gtt caa aac aat ctg ctg tac gtc tct gcc tcc cat ttg	336
Val Tyr Ile Val Gln Asn Asn Leu Leu Tyr Val Ser Ala Ser His Leu	
100 105 110	
gat gcg gcc acc tac cag gtg acg tac cag ctg aag att ctc acc acg	384
Asp Ala Ala Thr Tyr Gln Val Thr Tyr Gln Leu Lys Ile Leu Thr Thr	
115 120 125	
gcc atg ttc gcg gtt gtc att ctg cgc cgc aag ctg ctg aac acg cag	432
Ala Met Phe Ala Val Val Ile Leu Arg Arg Lys Leu Leu Asn Thr Gln	
130 135 140	
tgg ggt gcg ctg ctg ctc ctg gtg atg ggc atc gtc ctg gtg cag ttg	480
Trp Gly Ala Leu Leu Leu Leu Val Met Gly Ile Val Leu Val Gln Leu	
145 150 155 160	
gcc caa acg gag ggt ccg acg agt ggc tca gcc ggt ggt gcc gca gct	528
Ala Gln Thr Glu Gly Pro Thr Ser Gly Ser Ala Gly Gly Ala Ala Ala	
165 170 175	
gca gcc acg gcc gcc tcc tct ggc ggt gct ccc gag cag aac agg atg	576
Ala Ala Thr Ala Ala Ser Ser Gly Gly Ala Pro Glu Gln Asn Arg Met	

ES 2 589 655 T3

180	185	190	
ctc gga ctg tgg gcc gca ctg ggc gcc tgc ttc ctc tcc gga ttc gcg Leu Gly Leu Trp Ala Ala Leu Gly Ala Cys Phe Leu Ser Gly Phe Ala 195 200 205			624
ggc atc tac ttt gag aag atc ctc aag ggt gcc gag atc tcc gtg tgg Gly Ile Tyr Phe Glu Lys Ile Leu Lys Gly Ala Glu Ile Ser Val Trp 210 215 220			672
atg cgg aat gtg cag ttg agt ctg ctc agc att ccc ttc ggc ctg ctc Met Arg Asn Val Gln Leu Ser Leu Leu Ser Ile Pro Phe Gly Leu Leu 225 230 235 240			720
acc tgt ttc gtt aac gac ggc agt agg atc ttc gac cag gga ttc ttc Thr Cys Phe Val Asn Asp Gly Ser Arg Ile Phe Asp Gln Gly Phe Phe 245 250 255			768
aag ggc tac gat ctg ttt gtc tgg tac ctg gtc ctg ctg cag gcc ggc Lys Gly Tyr Asp Leu Phe Val Trp Tyr Leu Val Leu Leu Gln Ala Gly 260 265 270			816
ggt gga ttg atc gtt gcc gtg gtg gtc aag tac gcg gat aac att ctc Gly Gly Leu Ile Val Ala Val Val Val Lys Tyr Ala Asp Asn Ile Leu 275 280 285			864
aag ggc ttc gcc acc tcg ctg gcc atc atc atc atc tcg tgc gtg gcc tcc Lys Gly Phe Ala Thr Ser Leu Ala Ile Ile Ile Ser Cys Val Ala Ser 290 295 300			912
ata tac atc ttc gac ttc aat ctc acg ctg cag ttc agc ttc gga gct Ile Tyr Ile Phe Asp Phe Asn Leu Thr Leu Gln Phe Ser Phe Gly Ala 305 310 315 320			960
ggc ctg gtc atc gcc tcc ata ttt ctc tac ggc tac gat ccg gcc agg Gly Leu Val Ile Ala Ser Ile Phe Leu Tyr Gly Tyr Asp Pro Ala Arg 325 330 335			1008
tcg gcg ccg aag cca act atg cat ggt cct ggc ggc gat gag gag aag Ser Ala Pro Lys Pro Thr Met His Gly Pro Gly Gly Asp Glu Glu Lys 340 345 350			1056
ctg ctg ccg cgc gtc tag Leu Leu Pro Arg Val 355			1074

<210> 20

<211> 357

<212> PRT

5 <213> *Drosophila melanogaster*

<400> 20

Met	Asn	Ser	Ile	His	Met	Asn	Ala	Asn	Thr	Leu	Lys	Tyr	Ile	Ser	Leu
1				5					10					15	
Leu	Thr	Leu	Thr	Leu	Gln	Asn	Ala	Ile	Leu	Gly	Leu	Ser	Met	Arg	Tyr
			20					25					30		
Ala	Arg	Thr	Arg	Pro	Gly	Asp	Ile	Phe	Leu	Ser	Ser	Thr	Ala	Val	Leu
		35					40					45			
Met	Ala	Glu	Phe	Ala	Lys	Leu	Ile	Thr	Cys	Leu	Phe	Leu	Val	Phe	Asn
	50					55					60				
Glu	Glu	Gly	Lys	Asp	Ala	Gln	Lys	Phe	Val	Arg	Ser	Leu	His	Lys	Thr
65					70					75					80
Ile	Ile	Ala	Asn	Pro	Met	Asp	Thr	Leu	Lys	Val	Cys	Val	Pro	Ser	Leu
				85					90					95	
Val	Tyr	Ile	Val	Gln	Asn	Asn	Leu	Leu	Tyr	Val	Ser	Ala	Ser	His	Leu
			100					105					110		
Asp	Ala	Ala	Thr	Tyr	Gln	Val	Thr	Tyr	Gln	Leu	Lys	Ile	Leu	Thr	Thr
		115					120					125			
Ala	Met	Phe	Ala	Val	Val	Ile	Leu	Arg	Arg	Lys	Leu	Leu	Asn	Thr	Gln
	130					135					140				
Trp	Gly	Ala	Leu	Leu	Leu	Leu	Val	Met	Gly	Ile	Val	Leu	Val	Gln	Leu
145					150					155					160
Ala	Gln	Thr	Glu	Gly	Pro	Thr	Ser	Gly	Ser	Ala	Gly	Gly	Ala	Ala	Ala
			165					170					175		
Ala	Ala	Thr	Ala	Ala	Ser	Ser	Gly	Gly	Ala	Pro	Glu	Gln	Asn	Arg	Met
		180						185					190		
Leu	Gly	Leu	Trp	Ala	Ala	Leu	Gly	Ala	Cys	Phe	Leu	Ser	Gly	Phe	Ala
	195					200					205				
Gly	Ile	Tyr	Phe	Glu	Lys	Ile	Leu	Lys	Gly	Ala	Glu	Ile	Ser	Val	Trp
	210					215				220					
Met	Arg	Asn	Val	Gln	Leu	Ser	Leu	Leu	Ser	Ile	Pro	Phe	Gly	Leu	Leu
225				230						235					240
Thr	Cys	Phe	Val	Asn	Asp	Gly	Ser	Arg	Ile	Phe	Asp	Gln	Gly	Phe	Phe
			245						250					255	
Lys	Gly	Tyr	Asp	Leu	Phe	Val	Trp	Tyr	Leu	Val	Leu	Leu	Gln	Ala	Gly
		260					265						270		
Gly	Gly	Leu	Ile	Val	Ala	Val	Val	Val	Lys	Tyr	Ala	Asp	Asn	Ile	Leu
	275						280					285			
Lys	Gly	Phe	Ala	Thr	Ser	Leu	Ala	Ile	Ile	Ile	Ser	Cys	Val	Ala	Ser
	290					295					300				
Ile	Tyr	Ile	Phe	Asp	Phe	Asn	Leu	Thr	Leu	Gln	Phe	Ser	Phe	Gly	Ala
305				310						315					320
Gly	Leu	Val	Ile	Ala	Ser	Ile	Phe	Leu	Tyr	Gly	Tyr	Asp	Pro	Ala	Arg
			325						330					335	
Ser	Ala	Pro	Lys	Pro	Thr	Met	His	Gly	Pro	Gly	Gly	Asp	Glu	Glu	Lys
		340						345					350		
Leu	Leu	Pro	Arg	Val											
		355													

<210> 21
 <211> 2100
 5 <212> ADN
 <213> *Saccharomyces cerevisiae*

<220>
 <221> CDS
 10 <222> (1)...(2097)
 <223> Codifica ScGAL10

<400> 21

ES 2 589 655 T3

```

---  ---
atg aca gct cag tta caa agt gaa agt act tct aaa att gtt ttg gtt 48
Met Thr Ala Gln Leu Gln Ser Glu Ser Thr Ser Lys Ile Val Leu Val
1      5      10      15

aca ggt ggt gct gga tac att ggt tca cac act gtg gta gag cta att 96
Thr Gly Gly Ala Gly Tyr Ile Gly Ser His Thr Val Val Glu Leu Ile
      20      25      30

gag aat gga tat gac tgt gtt gtt gct gat aac ctg tcg aat tca act 144
Glu Asn Gly Tyr Asp Cys Val Val Ala Asp Asn Leu Ser Asn Ser Thr
      35      40      45

tat gat tct gta gcc agg tta gag gtc ttg acc aag cat cac att ccc 192
Tyr Asp Ser Val Ala Arg Leu Glu Val Leu Thr Lys His His Ile Pro
      50      55      60

ttc tat gag gtt gat ttg tgt gac cga aaa ggt ctg gaa aag gtt ttc 240
Phe Tyr Glu Val Asp Leu Cys Asp Arg Lys Gly Leu Glu Lys Val Phe

```

ES 2 589 655 T3

65	70	75	80	
aaa gaa tat aaa att gat tgc gta att cac ttt gct ggt tta aag gct	Lys Glu Tyr Lys Ile Asp Ser Val Ile His Phe Ala Gly Leu Lys Ala	288		
	85	90	95	
gta ggt gaa tct aca caa atc ccg ctg aga tac tat cac aat aac att	Val Gly Glu Ser Thr Gln Ile Pro Leu Arg Tyr Tyr His Asn Asn Ile	336		
	100	105	110	
ttg gga act gtc gtt tta tta gag tta atg caa caa tac aac gtt tcc	Leu Gly Thr Val Val Leu Leu Glu Leu Met Gln Gln Tyr Asn Val Ser	384		
	115	120	125	
aaa ttt gtt ttt tca tct tct gct act gtc tat ggt gat gct acg aga	Lys Phe Val Phe Ser Ser Ser Ala Thr Val Tyr Gly Asp Ala Thr Arg	432		
	130	135	140	
ttc cca aat atg att cct atc cca gaa gaa tgt ccc tta ggg cct act	Phe Pro Asn Met Ile Pro Ile Pro Glu Glu Cys Pro Leu Gly Pro Thr	480		
	145	150	155	160
aat ccg tat ggt cat acg aaa tac gcc att gag aat atc ttg aat gat	Asn Pro Tyr Gly His Thr Lys Tyr Ala Ile Glu Asn Ile Leu Asn Asp	528		
	165	170	175	
ctt tac aat agc gac aaa aaa agt tgg aag ttt gct atc ttg cgt tat	Leu Tyr Asn Ser Asp Lys Lys Ser Trp Lys Phe Ala Ile Leu Arg Tyr	576		
	180	185	190	
ttt aac cca att ggc gca cat ccc tct gga tta atc gga gaa gat ccg	Phe Asn Pro Ile Gly Ala His Pro Ser Gly Leu Ile Gly Glu Asp Pro	624		
	195	200	205	
cta ggt ata cca aac aat ttg ttg cca tat atg gct caa gta gct gtt	Leu Gly Ile Pro Asn Asn Leu Leu Pro Tyr Met Ala Gln Val Ala Val	672		
	210	215	220	
ggt agg cgc gag aag ctt tac atc ttc gga gac gat tat gat tcc aga	Gly Arg Arg Glu Lys Leu Tyr Ile Phe Gly Asp Asp Tyr Asp Ser Arg	720		
	225	230	235	240
gat ggt acc ccg atc agg gat tat atc cac gta gtt gat cta gca aaa	Asp Gly Thr Pro Ile Arg Asp Tyr Ile His Val Val Asp Leu Ala Lys	768		
	245	250	255	
ggt cat att gca gcc ctg caa tac cta gag gcc tac aat gaa aat gaa	Gly His Ile Ala Ala Leu Gln Tyr Leu Glu Ala Tyr Asn Glu Asn Glu	816		
	260	265	270	
ggt ttg tgt cgt gag tgg aac ttg ggt tcc ggt aaa ggt tct aca gtt	Gly Leu Cys Arg Glu Trp Asn Leu Gly Ser Gly Lys Gly Ser Thr Val	864		
	275	280	285	
ttt gaa gtt tat cat gca ttc tgc aaa gct tct ggt att gat ott cca	Phe Glu Val Tyr His Ala Phe Cys Lys Ala Ser Gly Ile Asp Leu Pro	912		
	290	295	300	
tac aaa gtt acg ggc aga aga gca ggt gat gtt ttg aac ttg acg gct	Tyr Lys Val Thr Gly Arg Arg Ala Gly Asp Val Leu Asn Leu Thr Ala	960		
	305	310	315	320
aaa cca gat agg gcc aaa cgc gaa ctg aaa tgg cag acc gag ttg cag	Lys Pro Asp Arg Ala Lys Arg Glu Leu Lys Trp Gln Thr Glu Leu Gln	1008		
	325	330	335	

ES 2 589 655 T3

gtt gaa gac tcc tgc aag gat tta tgg aaa tgg act act gag aat cct	1056
Val Glu Asp Ser Cys Lys Asp Leu Trp Lys Trp Thr Thr Glu Asn Pro	
340 345 350	
ttt ggt tac cag tta agg ggt gtc gag gcc aga ttt tcc gct gaa gat	1104
Phe Gly Tyr Gln Leu Arg Gly Val Glu Ala Arg Phe Ser Ala Glu Asp	
355 360 365	
atg cgt tat gac gca aga ttt gtg act att ggt gcc gcc acc aga ttt	1152
Met Arg Tyr Asp Ala Arg Phe Val Thr Ile Gly Ala Gly Thr Arg Phe	
370 375 380	
caa gcc acg ttt gcc aat ttg ggc gcc agc att gtt gac ctg aaa gtg	1200
Gln Ala Thr Phe Ala Asn Leu Gly Ala Ser Ile Val Asp Leu Lys Val	
385 390 395 400	
aac gga caa tca gtt gtt ctt ggc tat gaa aat gag gaa ggg tat ttg	1248
Asn Gly Gln Ser Val Val Leu Gly Tyr Glu Asn Glu Glu Gly Tyr Leu	
405 410 415	
aat cct gat agt gct tat ata ggc gcc acg atc ggc agg tat gct aat	1296
Asn Pro Asp Ser Ala Tyr Ile Gly Ala Thr Ile Gly Arg Tyr Ala Asn	
420 425 430	
cgt att tgc aag ggt aag ttt agt tta tgc aac aaa gac tat cag tta	1344
Arg Ile Ser Lys Gly Lys Phe Ser Leu Cys Asn Lys Asp Tyr Gln Leu	
435 440 445	
acc gtt aat aac ggc gtt aat gcg aat cat agt agt atc ggt tct ttc	1392
Thr Val Asn Asn Gly Val Asn Ala Asn His Ser Ser Ile Gly Ser Phe	
450 455 460	
cac aga aaa aga ttt ttg gga ccc atc att caa aat cct tca aag gat	1440
His Arg Lys Arg Phe Leu Gly Pro Ile Ile Gln Asn Pro Ser Lys Asp	
465 470 475 480	
gtt ttt acc gcc gag tac atg ctg ata gat aat gag aag gac acc gaa	1488
Val Phe Thr Ala Glu Tyr Met Leu Ile Asp Asn Glu Lys Asp Thr Glu	
485 490 495	
ttt cca ggt gat cta ttg gta acc ata cag tat act gtg aac gtt gcc	1536
Phe Pro Gly Asp Leu Leu Val Thr Ile Gln Tyr Thr Val Asn Val Ala	
500 505 510	
caa aaa agt ttg gaa atg gta tat aaa ggt aaa ttg act gct ggt gaa	1584
Gln Lys Ser Leu Glu Met Val Tyr Lys Gly Lys Leu Thr Ala Gly Glu	
515 520 525	
gcg acg cca ata aat tta aca aat cat agt tat ttc aat ctg aac aag	1632
Ala Thr Pro Ile Asn Leu Thr Asn His Ser Tyr Phe Asn Leu Asn Lys	
530 535 540	
cca tat gga gac act att gag ggt acg gag att atg gtg cgt tca aaa	1680
Pro Tyr Gly Asp Thr Ile Glu Gly Thr Glu Ile Met Val Arg Ser Lys	
545 550 555 560	
aaa tct gtt gat gtc gac aaa aac atg att cct acg ggt aat atc gtc	1728
Lys Ser Val Asp Val Asp Lys Asn Met Ile Pro Thr Gly Asn Ile Val	
565 570 575	
gat aga gaa att gct acc ttt aac tct aca aag cca acg gtc tta ggc	1776
Asp Arg Glu Ile Ala Thr Phe Asn Ser Thr Lys Pro Thr Val Leu Gly	
580 585 590	

ES 2 589 655 T3

ccc	aaa	aat	ccc	cag	ttt	gat	tgt	tgt	ttt	gtg	gtg	gat	gaa	aat	gct	1824
Pro	Lys	Asn	Pro	Gln	Phe	Asp	Cys	Cys	Phe	Val	Val	Asp	Glu	Asn	Ala	
		595					600					605				
aag	cca	agt	caa	atc	aat	act	cta	aac	aat	gaa	ttg	acg	ctt	att	gtc	1872
Lys	Pro	Ser	Gln	Ile	Asn	Thr	Leu	Asn	Asn	Glu	Leu	Thr	Leu	Ile	Val	
	610					615				620						
aag	gct	ttt	cat	ccc	gat	tcc	aat	att	aca	tta	gaa	gtt	tta	agt	aca	1920
Lys	Ala	Phe	His	Pro	Asp	Ser	Asn	Ile	Thr	Leu	Glu	Val	Leu	Ser	Thr	
625					630					635					640	
gag	cca	act	tat	caa	ttt	tat	acc	ggg	gat	ttc	ttg	tct	gct	ggg	tac	1968
Glu	Pro	Thr	Tyr	Gln	Phe	Tyr	Thr	Gly	Asp	Phe	Leu	Ser	Ala	Gly	Tyr	
			645						650					655		
gaa	gca	aga	caa	ggg	ttt	gca	att	gag	cct	ggg	aga	tac	att	gat	gct	2016
Glu	Ala	Arg	Gln	Gly	Phe	Ala	Ile	Glu	Pro	Gly	Arg	Tyr	Ile	Asp	Ala	
			660					665					670			
atc	aat	caa	gag	aac	tg	aaa	gat	tgt	gta	acc	ttg	aaa	aac	ggg	gaa	2064
Ile	Asn	Gln	Glu	Asn	Trp	Lys	Asp	Cys	Val	Thr	Leu	Lys	Asn	Gly	Glu	
		675					680					685				
act	tac	ggg	tcc	aag	att	gtc	tac	aga	ttt	tcc	tga					2100
Thr	Tyr	Gly	Ser	Lys	Ile	Val	Tyr	Arg	Phe	Ser						
	690					695										

<210> 22

<211> 699

5 <212> PRT

<213> *Saccharomyces cerevisiae*

<400> 22

Met	Thr	Ala	Gln	Leu	Gln	Ser	Glu	Ser	Thr	Ser	Lys	Ile	Val	Leu	Val
1				5					10					15	
Thr	Gly	Gly	Ala	Gly	Tyr	Ile	Gly	Ser	His	Thr	Val	Val	Glu	Leu	Ile
			20					25					30		
Glu	Asn	Gly	Tyr	Asp	Cys	Val	Val	Ala	Asp	Asn	Leu	Ser	Asn	Ser	Thr
		35					40					45			
Tyr	Asp	Ser	Val	Ala	Arg	Leu	Glu	Val	Leu	Thr	Lys	His	His	Ile	Pro
	50					55					60				
Phe	Tyr	Glu	Val	Asp	Leu	Cys	Asp	Arg	Lys	Gly	Leu	Glu	Lys	Val	Phe
65					70					75					80
Lys	Glu	Tyr	Lys	Ile	Asp	Ser	Val	Ile	His	Phe	Ala	Gly	Leu	Lys	Ala
				85					90					95	
Val	Gly	Glu	Ser	Thr	Gln	Ile	Pro	Leu	Arg	Tyr	Tyr	His	Asn	Asn	Ile
			100					105					110		
Leu	Gly	Thr	Val	Val	Leu	Leu	Glu	Leu	Met	Gln	Gln	Tyr	Asn	Val	Ser
		115					120					125			
Lys	Phe	Val	Phe	Ser	Ser	Ser	Ala	Thr	Val	Tyr	Gly	Asp	Ala	Thr	Arg
	130					135					140				
Phe	Pro	Asn	Met	Ile	Pro	Ile	Pro	Glu	Glu	Cys	Pro	Leu	Gly	Pro	Thr
145					150					155					160
Asn	Pro	Tyr	Gly	His	Thr	Lys	Tyr	Ala	Ile	Glu	Asn	Ile	Leu	Asn	Asp
				165					170					175	
Leu	Tyr	Asn	Ser	Asp	Lys	Lys	Ser	Trp	Lys	Phe	Ala	Ile	Leu	Arg	Tyr
		180					185						190		
Phe	Asn	Pro	Ile	Gly	Ala	His	Pro	Ser	Gly	Leu	Ile	Gly	Glu	Asp	Pro
		195					200					205			
Leu	Gly	Ile	Pro	Asn	Asn	Leu	Leu	Pro	Tyr	Met	Ala	Gln	Val	Ala	Val
	210					215					220				
Gly	Arg	Arg	Glu	Lys	Leu	Tyr	Ile	Phe	Gly	Asp	Asp	Tyr	Asp	Ser	Arg
225					230					235					240

Asp	Gly	Thr	Pro	Ile	Arg	Asp	Tyr	Ile	His	Val	Val	Asp	Leu	Ala	Lys	
				245					250					255		
Gly	His	Ile	Ala	Ala	Leu	Gln	Tyr	Leu	Glu	Ala	Tyr	Asn	Glu	Asn	Glu	
			260					265					270			
Gly	Leu	Cys	Arg	Glu	Trp	Asn	Leu	Gly	Ser	Gly	Lys	Gly	Ser	Thr	Val	
		275					280					285				
Phe	Glu	Val	Tyr	His	Ala	Phe	Cys	Lys	Ala	Ser	Gly	Ile	Asp	Leu	Pro	
	290					295					300					
Tyr	Lys	Val	Thr	Gly	Arg	Arg	Ala	Gly	Asp	Val	Leu	Asn	Leu	Thr	Ala	
305				310						315					320	
Lys	Pro	Asp	Arg	Ala	Lys	Arg	Glu	Leu	Lys	Trp	Gln	Thr	Glu	Leu	Gln	
				325					330					335		
Val	Glu	Asp	Ser	Cys	Lys	Asp	Leu	Trp	Lys	Trp	Thr	Thr	Glu	Asn	Pro	
		340					345						350			
Phe	Gly	Tyr	Gln	Leu	Arg	Gly	Val	Glu	Ala	Arg	Phe	Ser	Ala	Glu	Asp	
		355					360					365				
Met	Arg	Tyr	Asp	Ala	Arg	Phe	Val	Thr	Ile	Gly	Ala	Gly	Thr	Arg	Phe	
	370					375					380					
Gln	Ala	Thr	Phe	Ala	Asn	Leu	Gly	Ala	Ser	Ile	Val	Asp	Leu	Lys	Val	
385				390						395					400	
Asn	Gly	Gln	Ser	Val	Val	Leu	Gly	Tyr	Glu	Asn	Glu	Glu	Gly	Tyr	Leu	
				405					410					415		
Asn	Pro	Asp	Ser	Ala	Tyr	Ile	Gly	Ala	Thr	Ile	Gly	Arg	Tyr	Ala	Asn	
			420					425					430			
Arg	Ile	Ser	Lys	Gly	Lys	Phe	Ser	Leu	Cys	Asn	Lys	Asp	Tyr	Gln	Leu	
		435					440					445				
Thr	Val	Asn	Asn	Gly	Val	Asn	Ala	Asn	His	Ser	Ser	Ile	Gly	Ser	Phe	
	450					455					460					
His	Arg	Lys	Arg	Phe	Leu	Gly	Pro	Ile	Ile	Gln	Asn	Pro	Ser	Lys	Asp	
465				470						475					480	
Val	Phe	Thr	Ala	Glu	Tyr	Met	Leu	Ile	Asp	Asn	Glu	Lys	Asp	Thr	Glu	
				485					490					495		
Phe	Pro	Gly	Asp	Leu	Leu	Val	Thr	Ile	Gln	Tyr	Thr	Val	Asn	Val	Ala	
			500					505					510			
Gln	Lys	Ser	Leu	Glu	Met	Val	Tyr	Lys	Gly	Lys	Leu	Thr	Ala	Gly	Glu	
		515					520					525				
Ala	Thr	Pro	Ile	Asn	Leu	Thr	Asn	His	Ser	Tyr	Phe	Asn	Leu	Asn	Lys	
	530					535					540					
Pro	Tyr	Gly	Asp	Thr	Ile	Glu	Gly	Thr	Glu	Ile	Met	Val	Arg	Ser	Lys	
545					550					555					560	
Lys	Ser	Val	Asp	Val	Asp	Lys	Asn	Met	Ile	Pro	Thr	Gly	Asn	Ile	Val	
				565					570					575		
Asp	Arg	Glu	Ile	Ala	Thr	Phe	Asn	Ser	Thr	Lys	Pro	Thr	Val	Leu	Gly	
			580					585					590			
Pro	Lys	Asn	Pro	Gln	Phe	Asp	Cys	Cys	Phe	Val	Val	Asp	Glu	Asn	Ala	
		595				600						605				
Lys	Pro	Ser	Gln	Ile	Asn	Thr	Leu	Asn	Asn	Glu	Leu	Thr	Leu	Ile	Val	
	610					615					620					
Lys	Ala	Phe	His	Pro	Asp	Ser	Asn	Ile	Thr	Leu	Glu	Val	Leu	Ser	Thr	
625					630					635					640	
Glu	Pro	Thr	Tyr	Gln	Phe	Tyr	Thr	Gly	Asp	Phe	Leu	Ser	Ala	Gly	Tyr	
				645					650					655		
Glu	Ala	Arg	Gln	Gly	Phe	Ala	Ile	Glu	Pro	Gly	Arg	Tyr	Ile	Asp	Ala	
			660					665					670			
Ile	Asn	Gln	Glu	Asn	Trp	Lys	Asp	Cys	Val	Thr	Leu	Lys	Asn	Gly	Glu	
		675					680					685				
Thr	Tyr	Gly	Ser	Lys	Ile	Val	Tyr	Arg	Phe	Ser						
	690					695										

<210> 23
 <211> 1068
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Codifica dominio catalítico de hGalT optimizado por codón (XB)

<400> 23

10

```

ggt aga gat ttg tct aga ttg cca cag ttg gtt ggt gtt tcc act cca 48
ttg caa gga ggt tct aac tct gct gct gct att ggt caa tct tcc ggt 96
gag ttg aga act ggt gga gct aga cca cct cca cca ttg gga gct tcc 144
tct caa cca aga cca ggt ggt gat tct tct cca gtt gtt gac tct ggt 192
cca ggt cca gct tct aac ttg act tcc gtt cca gtt cca cac act act 240
gct ttg tcc ttg cca gct tgt cca gaa gaa tcc cca ttg ttg gtt ggt 288
cca atg ttg atc gag ttc aac atg cca gtt gac ttg gag ttg gtt gct 336
aag cag aac cca aac gtt aag atg ggt ggt aga tac gct cca aga gac 384
tgt gtt tcc cca cac aaa gtt gct atc atc atc cca ttc aga aac aga 432
cag gag cac ttg aag tac tgg ttg tac tac ttg cac cca gtt ttg caa 480
aga cag cag ttg gac tac ggt atc tac gtt atc aac cag gct ggt gac 528
act att ttc aac aga gct aag ttg ttg aat gtt ggt ttc cag gag gct 576
ttg aag gat tac gac tac act tgt ttc gtt ttc tcc gac gtt gac ttg 624
att cca atg aac gac cac aac gct tac aga tgt ttc tcc cag cca aga 672
cac att tct gtt gct atg gac aag ttc ggt ttc tcc ttg cca tac gtt 720
caa tac ttc ggt ggt gtt tcc gct ttg tcc aag cag cag ttc ttg act 768
atc aac ggt ttc cca aac aat tac tgg gga tgg ggt ggt gaa gat gac 816
gac atc ttt aac aga ttg gtt ttc aga gga atg tcc atc tct aga cca 864
aac gct gtt gtt ggt aga tgt aga atg atc aga cac tcc aga gac aag 912
aag aac gag cca aac cca caa aga ttc gac aga atc gct cac act aag 960
gaa act atg ttg tcc gac gga ttg aac tcc ttg act tac cag gtt ttg 1008
gac gtt cag aga tac cca ttg tac act cag atc act gtt gac atc ggt 1056
act cca tcc tag 1068
    
```

<210> 24
 <211> 355
 <212> PRT
 <213> Secuencia artificial

15

<220>
 <223> dominio catalítico de hGalT (XB)

<400> 24

20

Gly	Arg	Asp	Leu	Ser	Arg	Leu	Pro	Gln	Leu	Val	Gly	Val	Ser	Thr	Pro
1				5					10					15	
Leu	Gln	Gly	Gly	Ser	Asn	Ser	Ala	Ala	Ala	Ile	Gly	Gln	Ser	Ser	Gly
		20					25						30		
Glu	Leu	Arg	Thr	Gly	Gly	Ala	Arg	Pro	Pro	Pro	Pro	Leu	Gly	Ala	Ser
		35					40					45			
Ser	Gln	Pro	Arg	Pro	Gly	Gly	Asp	Ser	Ser	Pro	Val	Val	Asp	Ser	Gly
	50					55					60				
Pro	Gly	Pro	Ala	Ser	Asn	Leu	Thr	Ser	Val	Pro	Val	Pro	His	Thr	Thr
65					70					75				80	
Ala	Leu	Ser	Leu	Pro	Ala	Cys	Pro	Glu	Glu	Ser	Pro	Leu	Leu	Val	Gly
				85						90				95	
Pro	Met	Leu	Ile	Glu	Phe	Asn	Met	Pro	Val	Asp	Leu	Glu	Leu	Val	Ala
			100					105					110		
Lys	Gln	Asn	Pro	Asn	Val	Lys	Met	Gly	Gly	Arg	Tyr	Ala	Pro	Arg	Asp
		115					120					125			
Cys	Val	Ser	Pro	His	Lys	Val	Ala	Ile	Ile	Ile	Pro	Phe	Arg	Asn	Arg
	130					135					140				
Gln	Glu	His	Leu	Lys	Tyr	Trp	Leu	Tyr	Tyr	Leu	His	Pro	Val	Leu	Gln
145					150					155					160
Arg	Gln	Gln	Leu	Asp	Tyr	Gly	Ile	Tyr	Val	Ile	Asn	Gln	Ala	Gly	Asp
				165					170					175	
Thr	Ile	Phe	Asn	Arg	Ala	Lys	Leu	Leu	Asn	Val	Gly	Phe	Gln	Glu	Ala
			180					185					190		
Leu	Lys	Asp	Tyr	Asp	Tyr	Thr	Cys	Phe	Val	Phe	Ser	Asp	Val	Asp	Leu
		195					200					205			
Ile	Pro	Met	Asn	Asp	His	Asn	Ala	Tyr	Arg	Cys	Phe	Ser	Gln	Pro	Arg

	210					215					220				
His	Ile	Ser	Val	Ala	Met	Asp	Lys	Phe	Gly	Phe	Ser	Leu	Pro	Tyr	Val
225					230					235				240	
Gln	Tyr	Phe	Gly	Gly	Val	Ser	Ala	Leu	Ser	Lys	Gln	Gln	Phe	Leu	Thr
			245						250					255	
Ile	Asn	Gly	Phe	Pro	Asn	Asn	Tyr	Trp	Gly	Trp	Gly	Gly	Glu	Asp	Asp
		260						265					270		
Asp	Ile	Phe	Asn	Arg	Leu	Val	Phe	Arg	Gly	Met	Ser	Ile	Ser	Arg	Pro
		275					280					285			
Asn	Ala	Val	Val	Gly	Arg	Cys	Arg	Met	Ile	Arg	His	Ser	Arg	Asp	Lys
	290					295					300				
Lys	Asn	Glu	Pro	Asn	Pro	Gln	Arg	Phe	Asp	Arg	Ile	Ala	His	Thr	Lys
305					310					315				320	
Glu	Thr	Met	Leu	Ser	Asp	Gly	Leu	Asn	Ser	Leu	Thr	Tyr	Gln	Val	Leu
				325					330					335	
Asp	Val	Gln	Arg	Tyr	Pro	Leu	Tyr	Thr	Gln	Ile	Thr	Val	Asp	Ile	Gly
			340					345					350		
Thr	Pro	Ser													
		355													

<210> 25

<211> 1224

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Codifica dominio catalítico de GnTI humano Optimizado por codón (NA)

10 <400> 25

ES 2 589 655 T3

```

tcagtcagtg ctcttgatgg tgaccagca agtttgacca gagaagtgat tagattggcc 60
caagacgcag aggtggagtt ggagagacaa cgtggactgc tgcagcaaatt cggagatgca 120
ttgtctagtc aaagaggttag ggtgcctacc gcagctcctc cagcacagcc tagagtgcatt 180
gtgacccctg caccagctgt gattcctatc ttggtcatcg cctgtgacag atctactgtt 240
agaagatgtc tggacaagct gttgcattac agaccatctg ctgagttggt ccctatcatc 300
gttagtcaag actgtggtca cgaggagact gcccaagcca tcgcctccta cggatctgct 360
gtcactcaca tcagacagcc tgacctgtca tctattgctg tgccaccaga ccacagaaag 420
ttccaaggtt actacaagat cgctagacac tacagatggg cattgggtca agtcttcaga 480
cagtttagat tccctgctgc tgtggtggtg gaggatgact tggaggtggc tcctgacttc 540
tttgagtact ttagagcaac ctatccattg ctgaaggcag acccatccct gtggtgtgtc 600
tctgcctgga atgacaacgg taaggagcaa atggtggacg cttctaggcc tgagctgttg 660
tacagaaccg acttctttcc tggctctgga tgggtgctgt tggctgagtt gtgggctgag 720
ttggagccta agtggccaaa ggcattctgg gacgactgga tgagaagacc tgagcaaaga 780
cagggttagg cctgtatcag acctgagatc tcaagaacca tgacctttgg tagaaaggga 840
gtgtctcacg gtcaattctt tgaccaaacac ttgaagtta tcaagctgaa ccagcaattt 900
gtgcacttca cccaactgga cctgtcttac ttgcagagag aggcctatga cagagatttc 960
ctagctagag tctacggagc tcctcaactg caagtggaga aagtgaggac caatgacaga 1020
aaggagttgg gagaggtgag agtgcagtac actggtaggg actcctttaa ggctttcgct 1080
aaggctctgg gtgtcatgga tgaccttaag tctggagttc ctgagctgg ttacagaggt 1140
attgtcacct ttcaattcag aggtagaaga gtccacttgg ctcctccacc tacttgggag 1200
ggttatgatc cttcttgga ttag 1224

```

<210> 26

<211> 407

5 <212> PRT

<213> Secuencia artificial

<220>

<223> dominio catalítico de GnTI humano (NA)

10

<400> 26

Ser Val Ser Ala Leu Asp Gly Asp Pro Ala Ser Leu Thr Arg Glu Val

1		5		10		15									
Ile	Arg	Leu	Ala	Gln	Asp	Ala	Glu	Val	Glu	Leu	Glu	Arg	Gln	Arg	Gly
		20		25		30									
Leu	Leu	Gln	Gln	Ile	Gly	Asp	Ala	Leu	Ser	Ser	Gln	Arg	Gly	Arg	Val
		35		40		45									
Pro	Thr	Ala	Ala	Pro	Pro	Ala	Gln	Pro	Arg	Val	His	Val	Thr	Pro	Ala
		50		55		60									
Pro	Ala	Val	Ile	Pro	Ile	Leu	Val	Ile	Ala	Cys	Asp	Arg	Ser	Thr	Val
		65		70		75									
Arg	Arg	Cys	Leu	Asp	Lys	Leu	Leu	His	Tyr	Arg	Pro	Ser	Ala	Glu	Leu
		85		90		95									
Phe	Pro	Ile	Ile	Val	Ser	Gln	Asp	Cys	Gly	His	Glu	Glu	Thr	Ala	Gln
		100		105		110									
Ala	Ile	Ala	Ser	Tyr	Gly	Ser	Ala	Val	Thr	His	Ile	Arg	Gln	Pro	Asp
		115		120		125									
Leu	Ser	Ser	Ile	Ala	Val	Pro	Pro	Asp	His	Arg	Lys	Phe	Gln	Gly	Tyr
		130		135		140									
Tyr	Lys	Ile	Ala	Arg	His	Tyr	Arg	Trp	Ala	Leu	Gly	Gln	Val	Phe	Arg
		145		150		155									
Gln	Phe	Arg	Phe	Pro	Ala	Ala	Val	Val	Val	Glu	Asp	Asp	Leu	Glu	Val
		165		170		175									
Ala	Pro	Asp	Phe	Phe	Glu	Tyr	Phe	Arg	Ala	Thr	Tyr	Pro	Leu	Leu	Lys
		180		185		190									
Ala	Asp	Pro	Ser	Leu	Trp	Cys	Val	Ser	Ala	Trp	Asn	Asp	Asn	Gly	Lys
		195		200		205									
Glu	Gln	Met	Val	Asp	Ala	Ser	Arg	Pro	Glu	Leu	Leu	Tyr	Arg	Thr	Asp
		210		215		220									
Phe	Phe	Pro	Gly	Leu	Gly	Trp	Leu	Leu	Leu	Ala	Glu	Leu	Trp	Ala	Glu
		225		230		235									
Leu	Glu	Pro	Lys	Trp	Pro	Lys	Ala	Phe	Trp	Asp	Asp	Trp	Met	Arg	Arg
		245		250		255									
Pro	Glu	Gln	Arg	Gly	Arg	Ala	Cys	Ile	Arg	Pro	Glu	Ile	Ser	Arg	
		260		265		270									
Thr	Met	Thr	Phe	Gly	Arg	Lys	Gly	Val	Ser	His	Gly	Gln	Phe	Phe	Asp
		275		280		285									
Gln	His	Leu	Lys	Phe	Ile	Lys	Leu	Asn	Gln	Gln	Phe	Val	His	Phe	Thr
		290		295		300									
Gln	Leu	Asp	Leu	Ser	Tyr	Leu	Gln	Arg	Glu	Ala	Tyr	Asp	Arg	Asp	Phe
		305		310		315									
Leu	Ala	Arg	Val	Tyr	Gly	Ala	Pro	Gln	Leu	Gln	Val	Glu	Lys	Val	Arg
		325		330		335									
Thr	Asn	Asp	Arg	Lys	Glu	Leu	Gly	Glu	Val	Arg	Val	Gln	Tyr	Thr	Gly
		340		345		350									
Arg	Asp	Ser	Phe	Lys	Ala	Phe	Ala	Lys	Ala	Leu	Gly	Val	Met	Asp	Asp
		355		360		365									
Leu	Lys	Ser	Gly	Val	Pro	Arg	Ala	Gly	Tyr	Arg	Gly	Ile	Val	Thr	Phe
		370		375		380									
Gln	Phe	Arg	Gly	Arg	Arg	Val	His	Leu	Ala	Pro	Pro	Pro	Thr	Trp	Glu
		385		390		395									
Gly	Tyr	Asp	Pro	Ser	Trp	Asn									
		405													

<210> 27

<211> 1407

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Codifica dominio catalítico de Mm ManI (FB)

10

<400> 27

```

gagcccgctg acgccaccat ccgtgagaag agggcaaaga tcaaagagat gatgacccat 60
gcttggaata attataaacg ctatgcgtgg ggcttgaacg aactgaaacc tatatcaaaa 120

gaaggccatt caagcagttt gtttggcaac atcaaaggag ctacaatagt agatgccctg 180
gatacccttt tcattatggg catgaagact gaatttcaag aagctaaatc gtggattaaa 240
aaatatttag attttaattg gaatgctgaa gtttctgttt ttgaagtcaa catacgcttc 300
gtcgggtggac tgctgtcagc ctactatttg tccggagagg agatatttcg aaagaaagca 360
gtggaacttg gggtaaaatt gctacctgca ttccatactc cctctggaat accttgggca 420
ttgctgaata tgaaaagtgg gatcggggcg aactggccct gggcctctgg aggcagcagt 480
atcctggcgg aatttggaac tctgcattta gagtttatgc acttgtcca cttatcagga 540
gaccagctct ttgccgaaaa ggttatgaaa attcgaacag tgttgaacaa actggacaaa 600
ccagaaggcc ttatctctaa ctatctgaac ccagtagtg gacagtgggg tcaacatcat 660
gtgtcggttg gaggacttgg agacagcttt tatgaatatt tgcttaaggc gtggttaattg 720
tctgacaaga cagatctcga agccaagaag atgtattttg atgctgttca ggccatcgag 780
actcacttga tccgcaagtc aagtggggga ctaacgtaca tcgcagagtg gaaggggggc 840
ctcctggaac acaagatggg ccacctgacg tgctttgcag gaggcattgt tgcacttggg 900
gcagatggag ctccggaagc ccgggcccac cactaccttg aactcggagc tgaaattgcc 960
cgcacttgtc atgaatctta taatcgtaca tatgtgaagt tgggaccgga agcgtttcga 1020
tttgatggcg gtgtggaagc tattgccacg aggcaaaatg aaaagtatta catcttacgg 1080
cccaggttca tcgagacata catgtacatg tggcgactga ctcacgaccc caagtacagg 1140
acctgggcct ggggaagcgt ggaggctcta gaaagtcact gcagagtga cggaggctac 1200
tcaggcttac gggatgttta cattgcccgt gagagtatat acgatgtcca gcaaagtctc 1260
ttcctggcag agacactgaa gtatttgtac ttgatatttt ccgatgatga ccttcttcca 1320
ctagaacact ggatcttcaa caccgaggct catcctttcc ctatactccg tgaacagaag 1380
aaggaaattg atggcaaaga gaaatga 1407

```

5

<210> 28
 <211> 468
 <212> PRT
 <213> Secuencia artificial

10

<220>
 <223> dominio catalítico de Mm ManI (FB)

15

<400> 28

Glu	Pro	Ala	Asp	Ala	Thr	Ile	Arg	Glu	Lys	Arg	Ala	Lys	Ile	Lys	Glu
1				5				10					15		
Met	Met	Thr	His	Ala	Trp	Asn	Asn	Tyr	Lys	Arg	Tyr	Ala	Trp	Gly	Leu
			20					25					30		
Asn	Glu	Leu	Lys	Pro	Ile	Ser	Lys	Glu	Gly	His	Ser	Ser	Ser	Leu	Phe
		35					40					45			
Gly	Asn	Ile	Lys	Gly	Ala	Thr	Ile	Val	Asp	Ala	Leu	Asp	Thr	Leu	Phe
	50					55					60				
Ile	Met	Gly	Met	Lys	Thr	Glu	Phe	Gln	Glu	Ala	Lys	Ser	Trp	Ile	Lys
65					70					75				80	
Lys	Tyr	Leu	Asp	Phe	Asn	Val	Asn	Ala	Glu	Val	Ser	Val	Phe	Glu	Val
				85					90					95	
Asn	Ile	Arg	Phe	Val	Gly	Gly	Leu	Leu	Ser	Ala	Tyr	Tyr	Leu	Ser	Gly
			100					105					110		
Glu	Glu	Ile	Phe	Arg	Lys	Lys	Ala	Val	Glu	Leu	Gly	Val	Lys	Leu	Leu
		115						120				125			
Pro	Ala	Phe	His	Thr	Pro	Ser	Gly	Ile	Pro	Trp	Ala	Leu	Leu	Asn	Met
	130					135					140				
Lys	Ser	Gly	Ile	Gly	Arg	Asn	Trp	Pro	Trp	Ala	Ser	Gly	Gly	Ser	Ser
145					150					155				160	
Ile	Leu	Ala	Glu	Phe	Gly	Thr	Leu	His	Leu	Glu	Phe	Met	His	Leu	Ser
				165					170					175	
His	Leu	Ser	Gly	Asp	Pro	Val	Phe	Ala	Glu	Lys	Val	Met	Lys	Ile	Arg
			180					185					190		
Thr	Val	Leu	Asn	Lys	Leu	Asp	Lys	Pro	Glu	Gly	Leu	Tyr	Pro	Asn	Tyr
		195					200					205			
Leu	Asn	Pro	Ser	Ser	Gly	Gln	Trp	Gly	Gln	His	His	Val	Ser	Val	Gly
	210					215					220				
Gly	Leu	Gly	Asp	Ser	Phe	Tyr	Glu	Tyr	Leu	Leu	Lys	Ala	Trp	Leu	Met
225					230					235				240	
Ser	Asp	Lys	Thr	Asp	Leu	Glu	Ala	Lys	Lys	Met	Tyr	Phe	Asp	Ala	Val
				245					250					255	
Gln	Ala	Ile	Glu	Thr	His	Leu	Ile	Arg	Lys	Ser	Ser	Gly	Gly	Leu	Thr
			260					265					270		
Tyr	Ile	Ala	Glu	Trp	Lys	Gly	Gly	Leu	Leu	Glu	His	Lys	Met	Gly	His
		275					280					285			
Leu	Thr	Cys	Phe	Ala	Gly	Gly	Met	Phe	Ala	Leu	Gly	Ala	Asp	Gly	Ala
	290					295					300				
Pro	Glu	Ala	Arg	Ala	Gln	His	Tyr	Leu	Glu	Leu	Gly	Ala	Glu	Ile	Ala
305					310					315				320	
Arg	Thr	Cys	His	Glu	Ser	Tyr	Asn	Arg	Thr	Tyr	Val	Lys	Leu	Gly	Pro
				325					330					335	
Glu	Ala	Phe	Arg	Phe	Asp	Gly	Gly	Val	Glu	Ala	Ile	Ala	Thr	Arg	Gln
			340					345					350		
Asn	Glu	Lys	Tyr	Tyr	Ile	Leu	Arg	Pro	Glu	Val	Ile	Glu	Thr	Tyr	Met
		355					360					365			
Tyr	Met	Trp	Arg	Leu	Thr	His	Asp	Pro	Lys	Tyr	Arg	Thr	Trp	Ala	Trp
	370					375					380				
Glu	Ala	Val	Glu	Ala	Leu	Glu	Ser	His	Cys	Arg	Val	Asn	Gly	Gly	Tyr
385					390					395				400	
Ser	Gly	Leu	Arg	Asp	Val	Tyr	Ile	Ala	Arg	Glu	Ser	Tyr	Asp	Asp	Val
				405					410					415	
Gln	Gln	Ser	Phe	Phe	Leu	Ala	Glu	Thr	Leu	Lys	Tyr	Leu	Tyr	Leu	Ile
			420					425					430		
Phe	Ser	Asp	Asp	Asp	Leu	Leu	Pro	Leu	Glu	His	Trp	Ile	Phe	Asn	Thr
		435					440					445			
Glu	Ala	His	Pro	Phe	Pro	Ile	Leu	Arg	Glu	Gln	Lys	Lys	Glu	Ile	Asp
	450					455					460				
Gly	Lys	Glu	Lys												
465															

<210> 29
 <211> 1494
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> Codifica dominio catalítico de Tr ManI

 10 <400> 29

 cgcgccggat ctcccaaccc tacgagggcg gcagcagtcaggccgcatt ccagacgtcg 60
 tggaaacgctt accaccattt tgccctttccc catgacgacc tccacccggg cagcaacagc 120
 tttgatgatg agagaaacgg ctgggggctcg tcggcaatcg atggcttgga caccgctatc 180
 ctcatggggg atgccgacat tgtgaacacg atccttcagt atgtaccgca gatcaacttc 240
 accacgactg cggttgcca ccaaggcacc tccgtgttcg agaccaacat tcggtacctc 300
 ggtggcctgc tttctgccta tgacctgttg cgaggtcctt tcagctcctt ggcgacaaaac 360
 cagaccctgg taaacagcct tctgaggcag gctcaaacac tggccaacgg cctcaagggt 420
 gcgttcacca ctcccagcgg tgtcccggac cctaccgtct tcttcaaccc tactgtccgg 480
 agaagtgggt catctagcaa caacgtcgct gaaattggaa gcctgggtgct cgagtggaca 540
 cggttgagcg acctgacggg aaacccgcag tatgcccagc ttgcgcagaa ggcgagtcg 600
 tatctcctga atccaaaggg aagcccgag gcattggcctg gcctgattgg aacgtttgtc 660
 agcacgagca acggtacctt tcaggatagc agcggcagct ggtccggcct catggacagc 720
 ttctacgagt acctgatcaa gatgtacctg tacgaccggg ttgcgtttgc aactacaag 780
 gatcgctggg tccttgctgc cgactcgacc attgogcacc tcgcctctca cccgtcgacg 840
 cgcaaggact tgaccttttt gtcttcgtac aacggacagt ctacgtcgcc aaactcagga 900
 catttgccca gttttgccgg tggcaacttc atcttgggag gcattctcct gaacgagcaa 960
 aagtacattg actttggaat caagcttgcc agctcgtaact ttgccacgta caaccagacg 1020
 gcttctggaa tcggccccga aggccttcgg tgggtggaca gcgtgacggg cgccggcggc 1080
 tcgcccgcct cgtcccagtc cgggtttctac tcgtcggcag gattctgggt gacggcaccg 1140
 tattacatcc tgccggccgga gacgctggag agcttgtact acgcataccg cgtcacgggc 1200
 gactccaagt ggcaggacct ggcgtgggaa gcgttcagtg ccattgagga cgcagccgc 1260
 gccggcagcg cgtactcgtc catcaacgac gtgacgcagg ccaacggcgg ggggtgcctc 1320
 gacgatatgg agagcttctg gtttgccgag gcgtcaagt atgcgtacct gatctttgcg 1380
 gaggagtcgg atgtgcaggt gcaggccaac ggcgggaaca aatttgtctt taacacggag 1440
 gcgcaccctt ttagcatccg ttcacatca cgacggggcg gccaccttgc ttaa 1494

 <210> 30
 <211> 497
 15 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> dominio catalítico de Tr ManI

 20 <400> 30

Arg	Ala	Gly	Ser	Pro	Asn	Pro	Thr	Arg	Ala	Ala	Ala	Val	Lys	Ala	Ala	
1				5					10					15		
Phe	Gln	Thr	Ser	Trp	Asn	Ala	Tyr	His	His	Phe	Ala	Phe	Pro	His	Asp	
			20					25					30			
Asp	Leu	His	Pro	Val	Ser	Asn	Ser	Phe	Asp	Asp	Glu	Arg	Asn	Gly	Trp	
		35					40					45				
Gly	Ser	Ser	Ala	Ile	Asp	Gly	Leu	Asp	Thr	Ala	Ile	Leu	Met	Gly	Asp	
	50					55					60					
Ala	Asp	Ile	Val	Asn	Thr	Ile	Leu	Gln	Tyr	Val	Pro	Gln	Ile	Asn	Phe	
65					70					75					80	
Thr	Thr	Thr	Ala	Val	Ala	Asn	Gln	Gly	Ile	Ser	Val	Phe	Glu	Thr	Asn	
			85						90					95		
Ile	Arg	Tyr	Leu	Gly	Gly	Leu	Leu	Ser	Ala	Tyr	Asp	Leu	Leu	Arg	Gly	
			100					105					110			
Pro	Phe	Ser	Ser	Leu	Ala	Thr	Asn	Gln	Thr	Leu	Val	Asn	Ser	Leu	Leu	
		115					120					125				
Arg	Gln	Ala	Gln	Thr	Leu	Ala	Asn	Gly	Leu	Lys	Val	Ala	Phe	Thr	Thr	
	130					135					140					
Pro	Ser	Gly	Val	Pro	Asp	Pro	Thr	Val	Phe	Phe	Asn	Pro	Thr	Val	Arg	
145					150					155					160	
Arg	Ser	Gly	Ala	Ser	Ser	Asn	Asn	Val	Ala	Glu	Ile	Gly	Ser	Leu	Val	
				165					170					175		
Leu	Glu	Trp	Thr	Arg	Leu	Ser	Asp	Leu	Thr	Gly	Asn	Pro	Gln	Tyr	Ala	
			180					185					190			
Gln	Leu	Ala	Gln	Lys	Gly	Glu	Ser	Tyr	Leu	Leu	Asn	Pro	Lys	Gly	Ser	
		195					200					205				
Pro	Glu	Ala	Trp	Pro	Gly	Leu	Ile	Gly	Thr	Phe	Val	Ser	Thr	Ser	Asn	
	210					215					220					
Gly	Thr	Phe	Gln	Asp	Ser	Ser	Gly	Ser	Trp	Ser	Gly	Leu	Met	Asp	Ser	
225					230					235				240		
Phe	Tyr	Glu	Tyr	Leu	Ile	Lys	Met	Tyr	Leu	Tyr	Asp	Pro	Val	Ala	Phe	
				245					250					255		
Ala	His	Tyr	Lys	Asp	Arg	Trp	Val	Leu	Ala	Ala	Asp	Ser	Thr	Ile	Ala	
			260					265					270			
His	Leu	Ala	Ser	His	Pro	Ser	Thr	Arg	Lys	Asp	Leu	Thr	Phe	Leu	Ser	
		275					280					285				
Ser	Tyr	Asn	Gly	Gln	Ser	Thr	Ser	Pro	Asn	Ser	Gly	His	Leu	Ala	Ser	
	290					295					300					
Phe	Ala	Gly	Gly	Asn	Phe	Ile	Leu	Gly	Gly	Ile	Leu	Leu	Asn	Glu	Gln	
305					310					315					320	
Lys	Tyr	Ile	Asp	Phe	Gly	Ile	Lys	Leu	Ala	Ser	Ser	Tyr	Phe	Ala	Thr	
				325					330					335		
Tyr	Asn	Gln	Thr	Ala	Ser	Gly	Ile	Gly	Pro	Glu	Gly	Phe	Ala	Trp	Val	
			340					345					350			
Asp	Ser	Val	Thr	Gly	Ala	Gly	Gly	Ser	Pro	Pro	Ser	Ser	Gln	Ser	Gly	
		355					360					365				
Phe	Tyr	Ser	Ser	Ala	Gly	Phe	Trp	Val	Thr	Ala	Pro	Tyr	Tyr	Ile	Leu	
	370					375					380					
Arg	Pro	Glu	Thr	Leu	Glu	Ser	Leu	Tyr	Tyr	Ala	Tyr	Arg	Val	Thr	Gly	
385					390					395					400	
Asp	Ser	Lys	Trp	Gln	Asp	Leu	Ala	Trp	Glu	Ala	Phe	Ser	Ala	Ile	Glu	
				405					410					415		
Asp	Ala	Cys	Arg	Ala	Gly	Ser	Ala	Tyr	Ser	Ser	Ile	Asn	Asp	Val	Thr	
			420					425					430			

ES 2 589 655 T3

Gln Ala Asn Gly Gly Gly Ala Ser Asp Asp Met Glu Ser Phe Trp Phe
 435 440 445
 Ala Glu Ala Leu Lys Tyr Ala Tyr Leu Ile Phe Ala Glu Glu Ser Asp
 450 455 460
 Val Gln Val Gln Ala Asn Gly Gly Asn Lys Phe Val Phe Asn Thr Glu
 465 470 475 480
 Ala His Pro Phe Ser Ile Arg Ser Ser Ser Arg Arg Gly Gly His Leu
 485 490 495
 Ala

<210> 31
 <211> 1068
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> codifica GnT II (TC) de Rata Optimizado por codón

<400> 31

```
tccttggttt accaattgaa cttcgaccag atgttgagaa acgttgacaa ggacgggtact 60
tggtctcctg gtgagttggt tttggttggt caggttcaca acagaccaga gtacttgaga 120
ttgttgatcg actccttgag aaaggctcaa ggtatcagag aggttttggt tatcttctcc 180
cacgatttct ggtctgctga gatcaactcc ttgatctcct ccgttgactt ctgtccagtt 240
ttgcagggtt tcttcccatt ctccatccaa ttgtacccat ctgagttccc aggttctgat 300
ccaagagact gtccaagaga cttgaagaag aacgctgctt tgaagttggg ttgtatcaac 360
gctgaatacc cagattcttt cggtcactac agagaggcta agttctccca aactaagcat 420
cattggtggt ggaagttgca ctttggttgg gagagagtta aggttttgca ggactacact 480
ggattgatct tgttcttgga ggaggatcat tacttggttc cagacttcta ccacgttttc 540
aagaagatgt ggaagttgaa gcaacaagag tgtccagggt gtgacgtttt gtccttgga 600
acttacacta ctatcagatc cttctacggt atcgctgaca aggttgacgt taagacttgg 660
aagtccactg aacacaacat gggattggct ttgactagag atgcttacca gaagttgatc 720
gagtgtactg acactttctg tacttacgac gactacaact gggactggac tttgcagtac 780
ttgactttgg cttgtttgcc aaaagtttgg aaggttttgg ttccacagga tccaagaatt 840
ttccacgctg gtgactgtgg aatgcaccac aagaaaactt gtagaccatc cactcagtcc 900
gctcaaattg agtccttggt gaacaacaac aagcagtact tgttcccaga gactttgggt 960
atcggagaga agtttccaat ggctgctatt tcccaccaa gaaagaatgg tggatgggg 1020
gatattagag accacgagtt gtgtaaatcc tacagaagat tgcagtag 1068
```

<210> 32
 <211> 355
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> GnT II (TC) de Rata Optimizado por codón

<400> 32

Ser	Leu	Val	Tyr	Gln	Leu	Asn	Phe	Asp	Gln	Met	Leu	Arg	Asn	Val	Asp
1				5					10					15	
Lys	Asp	Gly	Thr	Trp	Ser	Pro	Gly	Glu	Leu	Val	Leu	Val	Val	Gln	Val
			20					25					30		
His	Asn	Arg	Pro	Glu	Tyr	Leu	Arg	Leu	Leu	Ile	Asp	Ser	Leu	Arg	Lys
		35					40					45			
Ala	Gln	Gly	Ile	Arg	Glu	Val	Leu	Val	Ile	Phe	Ser	His	Asp	Phe	Trp
	50					55					60				
Ser	Ala	Glu	Ile	Asn	Ser	Leu	Ile	Ser	Ser	Val	Asp	Phe	Cys	Pro	Val
65					70					75					80
Leu	Gln	Val	Phe	Phe	Pro	Phe	Ser	Ile	Gln	Leu	Tyr	Pro	Ser	Glu	Phe
				85					90					95	
Pro	Gly	Ser	Asp	Pro	Arg	Asp	Cys	Pro	Arg	Asp	Leu	Lys	Lys	Asn	Ala
			100					105					110		
Ala	Leu	Lys	Leu	Gly	Cys	Ile	Asn	Ala	Glu	Tyr	Pro	Asp	Ser	Phe	Gly
		115					120					125			
His	Tyr	Arg	Glu	Ala	Lys	Phe	Ser	Gln	Thr	Lys	His	His	Trp	Trp	Trp
	130					135					140				
Lys	Leu	His	Phe	Val	Trp	Glu	Arg	Val	Lys	Val	Leu	Gln	Asp	Tyr	Thr
145					150					155					160
Gly	Leu	Ile	Leu	Phe	Leu	Glu	Glu	Asp	His	Tyr	Leu	Ala	Pro	Asp	Phe
				165					170					175	
Tyr	His	Val	Phe	Lys	Lys	Met	Trp	Lys	Leu	Lys	Gln	Gln	Glu	Cys	Pro
		180						185					190		
Gly	Cys	Asp	Val	Leu	Ser	Leu	Gly	Thr	Tyr	Thr	Thr	Ile	Arg	Ser	Phe
	195						200					205			
Tyr	Gly	Ile	Ala	Asp	Lys	Val	Asp	Val	Lys	Thr	Trp	Lys	Ser	Thr	Glu
	210					215					220				
His	Asn	Met	Gly	Leu	Ala	Leu	Thr	Arg	Asp	Ala	Tyr	Gln	Lys	Leu	Ile
225					230					235					240
Glu	Cys	Thr	Asp	Thr	Phe	Cys	Thr	Tyr	Asp	Asp	Tyr	Asn	Trp	Asp	Trp
			245						250					255	
Thr	Leu	Gln	Tyr	Leu	Thr	Leu	Ala	Cys	Leu	Pro	Lys	Val	Trp	Lys	Val
		260						265					270		
Leu	Val	Pro	Gln	Ala	Pro	Arg	Ile	Phe	His	Ala	Gly	Asp	Cys	Gly	Met
	275						280					285			
His	His	Lys	Lys	Thr	Cys	Arg	Pro	Ser	Thr	Gln	Ser	Ala	Gln	Ile	Glu
	290					295					300				
Ser	Leu	Leu	Asn	Asn	Asn	Lys	Gln	Tyr	Leu	Phe	Pro	Glu	Thr	Leu	Val
305					310					315					320
Ile	Gly	Glu	Lys	Phe	Pro	Met	Ala	Ala	Ile	Ser	Pro	Pro	Arg	Lys	Asn
				325					330					335	
Gly	Gly	Trp	Gly	Asp	Ile	Arg	Asp	His	Glu	Leu	Cys	Lys	Ser	Tyr	Arg
			340					345					350		
Arg	Leu	Gln													
		355													

<210> 33
 <211> 3105
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Codifica ManII de *Drosophila melanogaster* optimizado por codón (KD)

<400> 33

agagacgatac	caattagacc	tccattgaag	gttgctagat	ccccaagacc	aggtcaatgt	60
caagatggtg	ttcaggacgt	cccaaacgtt	gatgtccaga	tggtggagtt	gtacgataga	120
atgtccttca	aggacattga	tggtggtgtt	tggaagcagg	gttggaaacat	taagtacgat	180
ccattgaagt	acaacgctca	tcacaagttg	aaggtcttcg	ttgtcccaca	ctcccacaac	240
gacccctggt	ggattcagac	cttcgaggaa	tactaccagc	acgacaccaa	gcacatcttg	300
tccaacgctt	tgagacattt	gcacgacaac	ccagagatga	agttcatctg	ggctgaaatc	360
tcctacttcg	ctagattcta	ccacgatttg	ggtgagaaca	agaagttgca	gatgaagtcc	420
atcgtcaaga	acggtcagtt	ggaattcgtc	actggtggat	gggtcatgcc	agacgaggct	480
aactcccact	ggagaaaacgt	tttggtgcag	ttgaccgaag	gtcaaacttg	gttgaagcaa	540
ttcatgaacg	tactccaac	tgcttcctgg	gctatcgatc	cattcggaca	ctctccaact	600
atgccatata	ttttgcagaa	gtctggtttc	aagaatatgt	tgatccagag	aaccctactac	660
tccgttaaga	aggagttggc	tcaacagaga	cagttggagt	tcttgtggag	acagatctgg	720
gacaacaaaag	gtgacactgc	tttggtcacc	cacatgatgc	cattctactc	ttacgacatt	780
cctcatacct	gtggtccaga	tccaaaggtt	tggtgtcagt	tcgatttcaa	agaatgggt	840
tccttcgggt	tgtcttgcct	atggaaggtt	ccacctagaa	ctatctctga	tcaaatggtt	900
gctgctagat	cagatttggt	gggtgatcag	tggaagaaga	aggctgagtt	gtacagaacc	960
aacgtcttgt	tgaattccatt	gggtgacgac	ttcagattca	agcagaacac	cgagtggggt	1020
gttcagagag	tcaactacga	aagattgttc	gaacacatca	actctcaggc	tcacttcaat	1080
gtccaggctc	agttcgggtac	tttgcaggaa	tacttcgatg	ctgttcacca	ggctgaaaga	1140
gctggacaag	ctgagttccc	aaccttgtct	ggtgacttct	tcacttacgc	tgatagatct	1200

gataactact	ggtctggtta	ctacacttcc	agaccatacc	ataagagaat	ggacagagtc	1260
ttgatgcact	acgttagagc	tgctgaaatg	ttgtccgctt	ggcactcctg	ggacgggatg	1320
gctagaatcg	aggaaagatt	ggagcaggct	agaagagagt	tgctccttgt	ccagcaccac	1380
gacggtatta	ctggtactgc	taaaactcac	gttgtcgtcg	actacgagca	agaatgcag	1440
gaagctttga	aagcttgtca	aatgggtcatg	caacagttctg	tctacagatt	gttgactaag	1500
ccatccatct	actctccaga	cttctccttc	tcctacttca	ctttggacga	ctccagatgg	1560
ccagggttctg	gtggtgagga	ctctagaact	accatcatct	tgggtgagga	tatcttgcca	1620
tccaagcatg	ttgtcatgca	caacaccttg	ccacactgga	gagagcagtt	ggttgacttc	1680
tacgtctcct	ctccattcgt	ttctgttacc	gacttggtca	acaatccagt	tgaggctcag	1740
gtttctccag	tttggtcttg	gcaccacgac	actttgacta	agactatcca	cccacaaggt	1800
tccaccacca	agtacagaat	catcttcaag	gctagagttc	caccaatggg	tttggctacc	1860
tacgttttga	ccatctccga	ttccaagcca	gagcacacct	cctacgcttc	caatttggtg	1920
cttagaaaaga	acccaacttc	cttgccattg	ggtcaatacc	cagaggatgt	caagttcggg	1980
gatccaagag	agatctcctt	gagagttggt	aacgggtccaa	ccttggcttt	ctctgagcag	2040
ggtttgttga	agtcattcca	gttgactcag	gattctccac	atgttccagt	tcacttcaag	2100
ttcttgaagt	acggtgttag	atctcatggt	gatagatctg	gtgcttactt	gttcttgcca	2160
aatgggtccag	cttctccagt	cgagttgggt	cagccagttg	tcttgggtcac	taagggtaaa	2220
ttggagtctt	cgttttctgt	tgggttgcca	tctgtcgttc	accagaccat	catgagaggt	2280
ggtgctccag	agattagaaa	tttggtcgat	attggttctt	tggacaacac	tgagatcgtc	2340
atgagattgg	agactcatat	cgactctggt	gatatcttct	acactgattt	gaatggattg	2400
caattcatca	agaggagaag	attggacaag	ttgccattgc	aggctaacta	ctaccaattt	2460
ccatctggta	tggttcattga	ggatgctaag	accagattga	ctttgttgac	cgggtcaacca	2520
ttgggtggat	cttctttggc	ttctggtgag	ttggagatta	tgcaagatag	aagattgggt	2580
tctgatgatg	aaagaggttt	gggtcagggg	gttttggaca	acaagccagt	tttgcatatt	2640
tacagatttg	tcttgagaga	ggttaacaac	tgtgtcagac	catctaagtt	gcacccagct	2700
ggttacttga	cttctgctgc	tcacaaagct	tctcagttct	tggtggatcc	attggacaag	2760
ttcatcttcg	ctgaaaatga	gtggatcggt	gctcaggggc	aattcgggtg	tgatcatcca	2820
tctgctagag	aggatttgga	tgtctctgtc	atgagaagat	tgaccaagtc	ttctgctaaa	2880
accagagagc	ttggttacgt	tttgcacaga	accaatttga	tgcaatgtgg	tactccagag	2940
gagcatactc	agaagttgga	tgtctgtcac	ttgttgccaa	atgttgctag	atgtgagaga	3000
actaccttga	ctttcttgca	gaatttggtg	cacttggtatg	gtatggttgc	tccagaagtt	3060
tgtccaatgg	aaaccgctgc	ttacgtctct	tctcactctt	cttga		3105

<210> 34
 <211> 1034
 <212> PRT
 <213> Secuencia artificial

<220>

<223> ManII de *Drosophila melanogaster* optimizado por codón (KD)

<400> 34

5

```

Arg Asp Asp Pro Ile Arg Pro Pro Leu Lys Val Ala Arg Ser Pro Arg
 1      5      10      15
Pro Gly Gln Cys Gln Asp Val Val Gln Asp Val Pro Asn Val Asp Val
 20      25      30
Gln Met Leu Glu Leu Tyr Asp Arg Met Ser Phe Lys Asp Ile Asp Gly
 35      40      45
Gly Val Trp Lys Gln Gly Trp Asn Ile Lys Tyr Asp Pro Leu Lys Tyr
 50      55      60
Asn Ala His His Lys Leu Lys Val Phe Val Val Pro His Ser His Asn
 65      70      75      80
Asp Pro Gly Trp Ile Gln Thr Phe Glu Glu Tyr Tyr Gln His Asp Thr
 85      90      95
Lys His Ile Leu Ser Asn Ala Leu Arg His Leu His Asp Asn Pro Glu
 100     105     110
Met Lys Phe Ile Trp Ala Glu Ile Ser Tyr Phe Ala Arg Phe Tyr His
 115     120     125
Asp Leu Gly Glu Asn Lys Lys Leu Gln Met Lys Ser Ile Val Lys Asn
 130     135     140
Gly Gln Leu Glu Phe Val Thr Gly Gly Trp Val Met Pro Asp Glu Ala
 145     150     155     160
Asn Ser His Trp Arg Asn Val Leu Leu Gln Leu Thr Glu Gly Gln Thr
 165     170     175

```

Trp	Leu	Lys	Gln	Phe	Met	Asn	Val	Thr	Pro	Thr	Ala	Ser	Trp	Ala	Ile
			180					185					190		
Asp	Pro	Phe	Gly	His	Ser	Pro	Thr	Met	Pro	Tyr	Ile	Leu	Gln	Lys	Ser
		195					200					205			
Gly	Phe	Lys	Asn	Met	Leu	Ile	Gln	Arg	Thr	His	Tyr	Ser	Val	Lys	Lys
	210					215					220				
Glu	Leu	Ala	Gln	Gln	Arg	Gln	Leu	Glu	Phe	Leu	Trp	Arg	Gln	Ile	Trp
225					230					235					240
Asp	Asn	Lys	Gly	Asp	Thr	Ala	Leu	Phe	Thr	His	Met	Met	Pro	Phe	Tyr
			245						250					255	
Ser	Tyr	Asp	Ile	Pro	His	Thr	Cys	Gly	Pro	Asp	Pro	Lys	Val	Cys	Cys
		260						265					270		
Gln	Phe	Asp	Phe	Lys	Arg	Met	Gly	Ser	Phe	Gly	Leu	Ser	Cys	Pro	Trp
	275						280						285		
Lys	Val	Pro	Pro	Arg	Thr	Ile	Ser	Asp	Gln	Asn	Val	Ala	Ala	Arg	Ser
	290					295					300				
Asp	Leu	Leu	Val	Asp	Gln	Trp	Lys	Lys	Lys	Ala	Glu	Leu	Tyr	Arg	Thr
305					310					315					320
Asn	Val	Leu	Leu	Ile	Pro	Leu	Gly	Asp	Asp	Phe	Arg	Phe	Lys	Gln	Asn
			325						330					335	
Thr	Glu	Trp	Asp	Val	Gln	Arg	Val	Asn	Tyr	Glu	Arg	Leu	Phe	Glu	His
			340					345						350	
Ile	Asn	Ser	Gln	Ala	His	Phe	Asn	Val	Gln	Ala	Gln	Phe	Gly	Thr	Leu
	355						360						365		
Gln	Glu	Tyr	Phe	Asp	Ala	Val	His	Gln	Ala	Glu	Arg	Ala	Gly	Gln	Ala
	370				375						380				
Glu	Phe	Pro	Thr	Leu	Ser	Gly	Asp	Phe	Phe	Thr	Tyr	Ala	Asp	Arg	Ser
385					390					395					400
Asp	Asn	Tyr	Trp	Ser	Gly	Tyr	Tyr	Thr	Ser	Arg	Pro	Tyr	His	Lys	Arg
			405						410					415	
Met	Asp	Arg	Val	Leu	Met	His	Tyr	Val	Arg	Ala	Ala	Glu	Met	Leu	Ser
			420					425					430		
Ala	Trp	His	Ser	Trp	Asp	Gly	Met	Ala	Arg	Ile	Glu	Glu	Arg	Leu	Glu
	435					440						445			
Gln	Ala	Arg	Arg	Glu	Leu	Ser	Leu	Phe	Gln	His	His	Asp	Gly	Ile	Thr
	450					455					460				
Gly	Thr	Ala	Lys	Thr	His	Val	Val	Val	Asp	Tyr	Glu	Gln	Arg	Met	Gln
465					470					475					480
Glu	Ala	Leu	Lys	Ala	Cys	Gln	Met	Val	Met	Gln	Gln	Ser	Val	Tyr	Arg
			485						490					495	
Leu	Leu	Thr	Lys	Pro	Ser	Ile	Tyr	Ser	Pro	Asp	Phe	Ser	Phe	Ser	Tyr
		500						505					510		
Phe	Thr	Leu	Asp	Asp	Ser	Arg	Trp	Pro	Gly	Ser	Gly	Val	Glu	Asp	Ser
	515						520						525		
Arg	Thr	Thr	Ile	Ile	Leu	Gly	Glu	Asp	Ile	Leu	Pro	Ser	Lys	His	Val
	530					535					540				
Val	Met	His	Asn	Thr	Leu	Pro	His	Trp	Arg	Glu	Gln	Leu	Val	Asp	Phe
545					550					555					560
Tyr	Val	Ser	Ser	Pro	Phe	Val	Ser	Val	Thr	Asp	Leu	Ala	Asn	Asn	Pro
			565						570					575	
Val	Glu	Ala	Gln	Val	Ser	Pro	Val	Trp	Ser	Trp	His	His	Asp	Thr	Leu
		580						585					590		
Thr	Lys	Thr	Ile	His	Pro	Gln	Gly	Ser	Thr	Thr	Lys	Tyr	Arg	Ile	Ile
	595						600						605		
Phe	Lys	Ala	Arg	Val	Pro	Pro	Met	Gly	Leu	Ala	Thr	Tyr	Val	Leu	Thr
	610					615					620				
Ile	Ser	Asp	Ser	Lys	Pro	Glu	His	Thr	Ser	Tyr	Ala	Ser	Asn	Leu	Leu
625					630					635					640
Leu	Arg	Lys	Asn	Pro	Thr	Ser	Leu	Pro	Leu	Gly	Gln	Tyr	Pro	Glu	Asp
			645						650					655	
Val	Lys	Phe	Gly	Asp	Pro	Arg	Glu	Ile	Ser	Leu	Arg	Val	Gly	Asn	Gly
		660						665					670		
Pro	Thr	Leu	Ala	Phe	Ser	Glu	Gln	Gly	Leu	Leu	Lys	Ser	Ile	Gln	Leu
	675						680						685		
Thr	Gln	Asp	Ser	Pro	His	Val	Pro	Val	His	Phe	Lys	Phe	Leu	Lys	Tyr

<210> 35
 <211> 1014
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Codifica transportador de CMP-ácido siálico de Ratón (MmCST) Optimizado por codón

<400> 35

10

690		695		700
Gly Val Arg Ser His	Gly Asp Arg Ser Gly Ala Tyr Leu Phe Leu Pro			
705	710	715		720
Asn Gly Pro Ala Ser	Pro Val Glu Leu Gly Gln Pro Val Val Leu Val			
	725	730		735
Thr Lys Gly Lys Leu Glu Ser Ser Val Ser Val Gly Leu Pro Ser Val				
	740	745		750
Val His Gln Thr Ile Met Arg Gly Gly Ala Pro Glu Ile Arg Asn Leu				
	755	760		765
Val Asp Ile Gly Ser Leu Asp Asn Thr Glu Ile Val Met Arg Leu Glu				
	770	775		780
Thr His Ile Asp Ser Gly Asp Ile Phe Tyr Thr Asp Leu Asn Gly Leu				
	785	790		795
Gln Phe Ile Lys Arg Arg Arg Leu Asp Lys Leu Pro Leu Gln Ala Asn				
	805	810		815
Tyr Tyr Pro Ile Pro Ser Gly Met Phe Ile Glu Asp Ala Asn Thr Arg				
	820	825		830
Leu Thr Leu Leu Thr Gly Gln Pro Leu Gly Gly Ser Ser Leu Ala Ser				
	835	840		845
Gly Glu Leu Glu Ile Met Gln Asp Arg Arg Leu Ala Ser Asp Asp Glu				
	850	855		860
Arg Gly Leu Gly Gln Gly Val Leu Asp Asn Lys Pro Val Leu His Ile				
	865	870		875
Tyr Arg Leu Val Leu Glu Lys Val Asn Asn Cys Val Arg Pro Ser Lys				
	885	890		895
Leu His Pro Ala Gly Tyr Leu Thr Ser Ala Ala His Lys Ala Ser Gln				
	900	905		910
Ser Leu Leu Asp Pro Leu Asp Lys Phe Ile Phe Ala Glu Asn Glu Trp				
	915	920		925
Ile Gly Ala Gln Gly Gln Phe Gly Gly Asp His Pro Ser Ala Arg Glu				
	930	935		940
Asp Leu Asp Val Ser Val Met Arg Arg Leu Thr Lys Ser Ser Ala Lys				
	945	950		955
Thr Gln Arg Val Gly Tyr Val Leu His Arg Thr Asn Leu Met Gln Cys				
	965	970		975
Gly Thr Pro Glu Glu His Thr Gln Lys Leu Asp Val Cys His Leu Leu				
	980	985		990
Pro Asn Val Ala Arg Cys Glu Arg Thr Thr Leu Thr Phe Leu Gln Asn				
	995	1000		1005
Leu Glu His Leu Asp Gly Met Val Ala Pro Glu Val Cys Pro Met Glu				
	1010	1015		1020
Thr Ala Ala Tyr Val Ser Ser His Ser Ser				
1025	1030			

<210> 36
 <211> 336
 <212> PRT
 <213> Secuencia artificial

15

<220>

<223> transportador de CMP-ácido siálico de Ratón (MmCST) Optimizado por codón

<400> 36

5

```

atggctccag ctagagaaaa cgtttccttg ttcttcaagt tgtactgttt ggctgttatg 60
acttttggtt ctgctgctta cactgttgct ttgagatata ctagaactac tgctgaggag 120
ttgtacttct ccactactgc tgtttgtatc actgaggtta tcaagttgtt gatctccgtt 180
ggtttggttg ctaaggagac tggttctttg ggaagattca aggccttcctt gtccgaaaac 240
gttttggttg ccccaaagga gttggctaag ttgtctgttc catccttggg ttacgctgtt 300
cagaacaaca tggctttctt ggctttgtct aacttggacg ctgctgttta ccaagttact 360
taccagttga agatcccatg tactgctttg tgtactgttt tgatgttgaa cagaacattg 420
tccaagttgc agtggatctc cgttttcatg ttgtgtggtg gtgttacttt ggttcagtgg 480

aagccagctc aagcttccaa agttgttggt gctcagaacc cattgttggg tttegggtgct 540
attgctatcg ctgtttttgt ttccggtttc gctgggtgtt acttcgagaa ggttttgaag 600
tcctccgaca cttctttgtg ggtagaaac atccagatgt acttgtccgg tatcgttggt 660
actttggctg gtacttactt gtctgacggg gctgagattc aagagaaggg attcttctac 720
ggttacactt actatgtttg gttcgttatc ttcttggctt ccgttggttg tttgtacact 780
tccgttggtg ttaagtacac tgacaacatc atgaagggat tctctgctgc tgctgctatt 840
gttttgctca ctatcgcttc cgttttggtt ttccgattgc agatcacatt gtcctttgct 900
ttgggagctt tgttggtttg tgtttccatc tacttgtacg gattgccaag acaagacact 960
acttccattc agcaagaggc tacttccaag gagagaatca tcggtgttta gtag 1014

```


Met	Ala	Pro	Ala	Arg	Glu	Asn	Val	Ser	Leu	Phe	Phe	Lys	Leu	Tyr	Cys
1				5					10					15	
Leu	Ala	Val	Met	Thr	Leu	Val	Ala	Ala	Ala	Tyr	Thr	Val	Ala	Leu	Arg
			20					25					30		
Tyr	Thr	Arg	Thr	Thr	Ala	Glu	Glu	Leu	Tyr	Phe	Ser	Thr	Thr	Ala	Val
		35				40						45			
Cys	Ile	Thr	Glu	Val	Ile	Lys	Leu	Leu	Ile	Ser	Val	Gly	Leu	Leu	Ala
	50					55					60				
Lys	Glu	Thr	Gly	Ser	Leu	Gly	Arg	Phe	Lys	Ala	Ser	Leu	Ser	Glu	Asn
65				70						75					80
Val	Leu	Gly	Ser	Pro	Lys	Glu	Leu	Ala	Lys	Leu	Ser	Val	Pro	Ser	Leu
			85						90					95	
Val	Tyr	Ala	Val	Gln	Asn	Asn	Met	Ala	Phe	Leu	Ala	Leu	Ser	Asn	Leu
			100					105						110	
Asp	Ala	Ala	Val	Tyr	Gln	Val	Thr	Gln	Leu	Lys	Ile	Pro	Cys	Thr	
		115				120					125				
Ala	Leu	Cys	Thr	Val	Leu	Met	Leu	Asn	Arg	Thr	Leu	Ser	Lys	Leu	Gln
	130					135					140				
Trp	Ile	Ser	Val	Phe	Met	Leu	Cys	Gly	Gly	Val	Thr	Leu	Val	Gln	Trp
145					150					155					160
Lys	Pro	Ala	Gln	Ala	Ser	Lys	Val	Val	Val	Ala	Gln	Asn	Pro	Leu	Leu
				165					170					175	
Gly	Phe	Gly	Ala	Ile	Ala	Ile	Ala	Val	Leu	Cys	Ser	Gly	Phe	Ala	Gly
			180					185					190		
Val	Tyr	Phe	Glu	Lys	Val	Leu	Lys	Ser	Ser	Asp	Thr	Ser	Leu	Trp	Val
		195					200					205			
Arg	Asn	Ile	Gln	Met	Tyr	Leu	Ser	Gly	Ile	Val	Val	Thr	Leu	Ala	Gly
	210					215					220				
Thr	Tyr	Leu	Ser	Asp	Gly	Ala	Glu	Ile	Gln	Glu	Lys	Gly	Phe	Phe	Tyr
225				230						235					240
Gly	Tyr	Thr	Tyr	Tyr	Val	Trp	Phe	Val	Ile	Phe	Leu	Ala	Ser	Val	Gly
				245					250					255	
Gly	Leu	Tyr	Thr	Ser	Val	Val	Val	Lys	Tyr	Thr	Asp	Asn	Ile	Met	Lys
			260					265					270		
Gly	Phe	Ser	Ala	Ala	Ala	Ala	Ile	Val	Leu	Ser	Thr	Ile	Ala	Ser	Val
		275					280					285			
Leu	Leu	Phe	Gly	Leu	Gln	Ile	Thr	Leu	Ser	Phe	Ala	Leu	Gly	Ala	Leu
	290					295					300				
Leu	Val	Cys	Val	Ser	Ile	Tyr	Leu	Tyr	Gly	Leu	Pro	Arg	Gln	Asp	Thr
305					310					315					320
Thr	Ser	Ile	Gln	Gln	Glu	Ala	Thr	Ser	Lys	Glu	Arg	Ile	Ile	Gly	Val
				325					330					335	

<210> 37

<211> 2172

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> Codifica UDP-GlcNAc 2-epimerasa/N-acetilmanosamina quinasa Humana (HsGNE) optimizada por codón

<400> 37

```

atggaaaaga acggttaacaa cagaaagttg agagtttgtg ttgctacttg taacagagct 60
gactactcca agttgggtcc aatcatgttc ggtatcaaga ctgagccaga gttcttcgag 120
ttggacgttg ttgttttggg ttcccacttg attgatgact acggtaacac ttacagaatg 180
atcgagcagg acgacttcga catcaacact agattgcaca ctattgttag aggagaggac 240
gaagctgcta tggttgaatc tgttggattg gcttttggtta agttgccaga cgttttgaac 300
agattgaagc cagacatcat gattgttcac ggtgacagat tcgatgcttt ggctttgggt 360
acttccgctg ctttcatgaa cattagaatc ttgcacatcg aggggtggtga agtttctggt 420
actatcgacg actccatcag acacgctatc actaagttgg ctactacca tgtttgttgt 480
actagatccg ctgagcaaca cttgatttcc atgtgtgagg accacgacag aattttgttg 540
gctggttgtc catcttacga caagttgttg tccgctaaga acaaggacta catgtccatc 600
atcagaatgt ggttgggtga cgacgttaag tctaaggact acatcgttgc tttgcagcac 660
ccagttacta ctgacatcaa gcactccatc aagatgttcg agttgacttt ggacgctttg 720
atctccttca acaagagaac tttggttttg ttcccaaaca ttgacgctgg ttccaaagag 780
atggttagag ttatgagaaa gaagggtatc gaacaccacc caaacttcag agctgttaag 840
cacgttccat tcgaccaatt catccagttg gttgctcatg ctggttgtat gatcggtaac 900
tcctcctgtg gtgttagaga agttggtgct ttcgggtactc cagttatcaa cttgggtact 960
agacagatcg gtagagagac tggagaaaac gttttgcatg ttagagatgc tgacactcag 1020
gacaagattt tgcaggcttt gcacttgcaa ttcggaagc agtaccatg ttccaaaatc 1080
tacggtgacg gtaacgctgt tccaagaatc ttgaagtttt tgaagtccat cgacttgcaa 1140
gagccattgc agaagaagtt ctgtttccca ccagttaagg agaactctc ccaggacatt 1200
gaccacatct tggagacatt gtccgctttg gctgttgatt tgggtggaac taacttgaga 1260
gttgcctatc tttccatgaa gggagagatc gtttaagaagt aactcagtt caacccaaag 1320
acttacgagg agagaatcaa cttgatcttg cagatgtgtg ttgaagctgc tgctgaggct 1380
gttaagttga actgtagaat cttgggtgtt ggtatctcta ctggtggtag agttaatcca 1440
agagagggta tcgttttgca ctccactaag ttgattcagg agtggaactc cgttgatttg 1500
agaactccat tgtccgacac attgcacttg ccagtttggg ttgacaacga cggtaattgt 1560
gctgcttttg ctgagagaaa gttcgggtcaa ggaaagggat tggagaactt cgttactttg 1620
atcactggta ctggtattgg tgggtggtatc attcaccagc acgagttgat tcacggttct 1680
tccttctgtg ctgctgaatt gggacacttg gttgtttctt tggacggtcc agactgttct 1740
tgtggttccc acggttgtat tgaagcttac gcatcaggaa tggcattgca gagagaggct 1800
aagaagttgc acgacgagga cttgttggtg gttgagggaa tgtctgttcc aaaggacgag 1860
gctgttggtg ctttgcatth gatccaggct gctaagttgg gtaatgctaa ggctcagttc 1920
atcttgagaa ctgctggtac tgctttggga ttgggtgttg ttaatatctt gcacactatg 1980
aaccatcctt tggttatctt gtccggtgtt ttggcttctc actacatcca catcgtttaag 2040
gacgttatca gacagcaagc tttgtcctcc gttcaagacg ttgatgttgt tgtttccgac 2100
ttggttgacc cagctttgtt ggggtgctgt tccatggttt tggactacac tactagaaga 2160
atctactaat ag 2172

```

<210> 38

<211> 722

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> UDP-GlcNAc 2-epimerasa/N-acetilmanosamina quinasa Humana (HsGNE) optimizada por codón

<400> 38

Met	Glu	Lys	Asn	Gly	Asn	Asn	Arg	Lys	Leu	Arg	Val	Cys	Val	Ala	Thr
1				5					10					15	
Cys	Asn	Arg	Ala	Asp	Tyr	Ser	Lys	Leu	Ala	Pro	Ile	Met	Phe	Gly	Ile
			20					25					30		
Lys	Thr	Glu	Pro	Glu	Phe	Phe	Glu	Leu	Asp	Val	Val	Val	Leu	Gly	Ser
		35					40					45			
His	Leu	Ile	Asp	Asp	Tyr	Gly	Asn	Thr	Tyr	Arg	Met	Ile	Glu	Gln	Asp
	50					55					60				
Asp	Phe	Asp	Ile	Asn	Thr	Arg	Leu	His	Thr	Ile	Val	Arg	Gly	Glu	Asp
65					70					75					80
Glu	Ala	Ala	Met	Val	Glu	Ser	Val	Gly	Leu	Ala	Leu	Val	Lys	Leu	Pro
			85						90					95	
Asp	Val	Leu	Asn	Arg	Leu	Lys	Pro	Asp	Ile	Met	Ile	Val	His	Gly	Asp
			100					105					110		
Arg	Phe	Asp	Ala	Leu	Ala	Leu	Ala	Thr	Ser	Ala	Ala	Leu	Met	Asn	Ile
		115						120				125			
Arg	Ile	Leu	His	Ile	Glu	Gly	Gly	Glu	Val	Ser	Gly	Thr	Ile	Asp	Asp
	130					135					140				
Ser	Ile	Arg	His	Ala	Ile	Thr	Lys	Leu	Ala	His	Tyr	His	Val	Cys	Cys
145					150					155					160
Thr	Arg	Ser	Ala	Glu	Gln	His	Leu	Ile	Ser	Met	Cys	Glu	Asp	His	Asp
			165						170					175	
Arg	Ile	Leu	Leu	Ala	Gly	Cys	Pro	Ser	Tyr	Asp	Lys	Leu	Leu	Ser	Ala
		180						185					190		
Lys	Asn	Lys	Asp	Tyr	Met	Ser	Ile	Ile	Arg	Met	Trp	Leu	Gly	Asp	Asp
		195					200					205			
Val	Lys	Ser	Lys	Asp	Tyr	Ile	Val	Ala	Leu	Gln	His	Pro	Val	Thr	Thr
	210					215					220				
Asp	Ile	Lys	His	Ser	Ile	Lys	Met	Phe	Glu	Leu	Thr	Leu	Asp	Ala	Leu
225					230					235					240
Ile	Ser	Phe	Asn	Lys	Arg	Thr	Leu	Val	Leu	Phe	Pro	Asn	Ile	Asp	Ala
			245						250					255	
Gly	Ser	Lys	Glu	Met	Val	Arg	Val	Met	Arg	Lys	Lys	Gly	Ile	Glu	His
			260					265					270		
His	Pro	Asn	Phe	Arg	Ala	Val	Lys	His	Val	Pro	Phe	Asp	Gln	Phe	Ile
		275					280					285			
Gln	Leu	Val	Ala	His	Ala	Gly	Cys	Met	Ile	Gly	Asn	Ser	Ser	Cys	Gly
	290					295					300				
Val	Arg	Glu	Val	Gly	Ala	Phe	Gly	Thr	Pro	Val	Ile	Asn	Leu	Gly	Thr
305					310					315					320
Arg	Gln	Ile	Gly	Arg	Glu	Thr	Gly	Glu	Asn	Val	Leu	His	Val	Arg	Asp
			325						330					335	
Ala	Asp	Thr	Gln	Asp	Lys	Ile	Leu	Gln	Ala	Leu	His	Leu	Gln	Phe	Gly
			340					345					350		
Lys	Gln	Tyr	Pro	Cys	Ser	Lys	Ile	Tyr	Gly	Asp	Gly	Asn	Ala	Val	Pro
		355					360					365			
Arg	Ile	Leu	Lys	Phe	Leu	Lys	Ser	Ile	Asp	Leu	Gln	Glu	Pro	Leu	Gln
	370					375					380				
Lys	Lys	Phe	Cys	Phe	Pro	Pro	Val	Lys	Glu	Asn	Ile	Ser	Gln	Asp	Ile
385					390					395					400
Asp	His	Ile	Leu	Glu	Thr	Leu	Ser	Ala	Leu	Ala	Val	Asp	Leu	Gly	Gly
			405						410					415	
Thr	Asn	Leu	Arg	Val	Ala	Ile	Val	Ser	Met	Lys	Gly	Glu	Ile	Val	Lys
			420					425					430		
Lys	Tyr	Thr	Gln	Phe	Asn	Pro	Lys	Thr	Tyr	Glu	Glu	Arg	Ile	Asn	Leu
		435					440					445			
Ile	Leu	Gln	Met	Cys	Val	Glu	Ala	Ala	Ala	Glu	Ala	Val	Lys	Leu	Asn
	450					455					460				
Cys	Arg	Ile	Leu	Gly	Val	Gly	Ile	Ser	Thr	Gly	Gly	Arg	Val	Asn	Pro
465					470					475					480
Arg	Glu	Gly	Ile	Val	Leu	His	Ser	Thr	Lys	Leu	Ile	Gln	Glu	Trp	Asn
			485						490					495	
Ser	Val	Asp	Leu	Arg	Thr	Pro	Leu	Ser	Asp	Thr	Leu	His	Leu	Pro	Val
			500					505				510			
Trp	Val	Asp	Asn	Asp	Gly	Asn	Cys	Ala	Ala	Leu	Ala	Glu	Arg	Lys	Phe

[illegible]

<210> 39

<211> 1308

<212> ADN

<213> Secuencia artificial

$\langle 220 \rangle$

<223> Codifica CMP-ácido siálico sintasa Humana (HsCSS) Optimizada por codón

<400> 39

atggactctg	ttgaaaagg	tgctgctact	tctgtttcca	acccaagagg	tagaccatcc	60
agaggtagac	ctcctaagtt	gcagagaaac	tccagagggtg	gtcaaggtag	aggtgttgaa	120
aagccaccac	acttggctgc	tttgatcttg	gctagaggag	gttctaagg	tatcccattg	180
aagaacatca	agcacttggc	tggtgttcca	ttgattggat	gggttttgag	agctgctttg	240
gactctgggtg	ctttccaatc	tgtttgggtt	tccactgacc	acgacgagat	tgagaacggt	300
gctaagcaat	tccgtgctca	ggttcacaga	agatcctctg	aggtttccaa	ggactcttct	360
acttccttgg	acgctatcat	cgagttcttg	aactaccaca	acgaggttga	catcgttggt	420
aacatccaag	ctacttcccc	atgtttgcac	ccaactgact	tgcaaaaagt	tgctgagatg	480
atcagagaag	agggttacga	ctccgttttc	tccgttgtta	gaaggcacca	gttcagatgg	540
tccgagattc	agcaagggtgt	tagagagggt	acagagccat	tgaacttgaa	cccagctaaa	600
agaccaagaa	ggcaggattg	ggacggtgaa	ttgtacgaaa	acggttcctt	ctacttcgct	660
aagagacact	tgatcgagat	gggatacttg	caaggtggaa	agatggctta	ctacgagatg	720
agagctgaac	actccgttga	catcgacgtt	gatatcgact	ggccaattgc	tgagcagaga	780
gttttgagat	acggttactt	cggaaggag	aagttgaagg	agatcaagtt	gttggtttgt	840
aacatcgacg	gttgtttgac	taacggtcac	atctacgttt	ctggtgacca	gaaggagatt	900
atctcctacg	acgttaagga	cgctattggt	atctccttgt	tgaagaagtc	cggtatcgaa	960
gttagattga	tctccgagag	agcttgttcc	aagcaaakat	tgctcctctt	gaagttggac	1020
tgtaagatgg	aggtttccgt	ttctgacaag	ttggctgttg	ttgacgaatg	gagaaaggag	1080
atgggtttgt	gttggaagga	agttgcttac	ttgggttaacg	aagtttctga	cgaggagtgt	1140
ttgaagagag	tctgttttgc	tggtgttcca	cgtagtgctt	gttccactgc	tcaaaaggct	1200
gttggttaca	tctgtaagtg	taacgggtgt	agaggtgcta	ttagagagtt	cgctgagcac	1260
atctgtttgt	tgatggagaa	agttaataac	tcctgtcaga	agtagtag		1308

<210> 40

<211> 434

<212> PRT

<213> Secuencia artificial

5

<220>

<223> CMP-ácido siálico sintasa Humana (HsCSS) Optimizada por codón

<400> 40

10

Met	Asp	Ser	Val	Glu	Lys	Gly	Ala	Ala	Thr	Ser	Val	Ser	Asn	Pro	Arg
1				5					10					15	
Gly	Arg	Pro	Ser	Arg	Gly	Arg	Pro	Pro	Lys	Leu	Gln	Arg	Asn	Ser	Arg
			20					25					30		
Gly	Gly	Gln	Gly	Arg	Gly	Val	Glu	Lys	Pro	Pro	His	Leu	Ala	Ala	Leu
		35					40					45			
Ile	Leu	Ala	Arg	Gly	Gly	Ser	Lys	Gly	Ile	Pro	Leu	Lys	Asn	Ile	Lys
	50					55					60				
His	Leu	Ala	Gly	Val	Pro	Leu	Ile	Gly	Trp	Val	Leu	Arg	Ala	Ala	Leu
65				70					75						80
Asp	Ser	Gly	Ala	Phe	Gln	Ser	Val	Trp	Val	Ser	Thr	Asp	His	Asp	Glu
			85					90					95		
Ile	Glu	Asn	Val	Ala	Lys	Gln	Phe	Gly	Ala	Gln	Val	His	Arg	Arg	Ser
			100					105					110		
Ser	Glu	Val	Ser	Lys	Asp	Ser	Ser	Thr	Ser	Leu	Asp	Ala	Ile	Ile	Glu
		115					120					125			
Phe	Leu	Asn	Tyr	His	Asn	Glu	Val	Asp	Ile	Val	Gly	Asn	Ile	Gln	Ala
	130					135					140				
Thr	Ser	Pro	Cys	Leu	His	Pro	Thr	Asp	Leu	Gln	Lys	Val	Ala	Glu	Met
145				150						155					160
Ile	Arg	Glu	Glu	Gly	Tyr	Asp	Ser	Val	Phe	Ser	Val	Val	Arg	Arg	His
				165					170					175	
Gln	Phe	Arg	Trp	Ser	Glu	Ile	Gln	Lys	Gly	Val	Arg	Glu	Val	Thr	Glu
			180					185					190		
Pro	Leu	Asn	Leu	Asn	Pro	Ala	Lys	Arg	Pro	Arg	Arg	Gln	Asp	Trp	Asp
		195					200					205			
Gly	Glu	Leu	Tyr	Glu	Asn	Gly	Ser	Phe	Tyr	Phe	Ala	Lys	Arg	His	Leu
	210					215					220				
Ile	Glu	Met	Gly	Tyr	Leu	Gln	Gly	Gly	Lys	Met	Ala	Tyr	Tyr	Glu	Met
225					230					235					240
Arg	Ala	Glu	His	Ser	Val	Asp	Ile	Asp	Val	Asp	Ile	Asp	Trp	Pro	Ile
				245					250					255	
Ala	Glu	Gln	Arg	Val	Leu	Arg	Tyr	Gly	Tyr	Phe	Gly	Lys	Glu	Lys	Leu
			260					265					270		
Lys	Glu	Ile	Lys	Leu	Leu	Val	Cys	Asn	Ile	Asp	Gly	Cys	Leu	Thr	Asn
		275					280					285			
Gly	His	Ile	Tyr	Val	Ser	Gly	Asp	Gln	Lys	Glu	Ile	Ile	Ser	Tyr	Asp
	290					295					300				
Val	Lys	Asp	Ala	Ile	Gly	Ile	Ser	Leu	Leu	Lys	Lys	Ser	Gly	Ile	Glu
					310					315					320
Val	Arg	Leu	Ile	Ser	Glu	Arg	Ala	Cys	Ser	Lys	Gln	Thr	Leu	Ser	Ser
				325					330					335	
Leu	Lys	Leu	Asp	Cys	Lys	Met	Glu	Val	Ser	Val	Ser	Asp	Lys	Leu	Ala
			340					345					350		
Val	Val	Asp	Glu	Trp	Arg	Lys	Glu	Met	Gly	Leu	Cys	Trp	Lys	Glu	Val
		355					360					365			
Ala	Tyr	Leu	Gly	Asn	Glu	Val	Ser	Asp	Glu	Glu	Cys	Leu	Lys	Arg	Val
	370					375						380			
Gly	Leu	Ser	Gly	Ala	Pro	Ala	Asp	Ala	Cys	Ser	Thr	Ala	Gln	Lys	Ala
					390					395					400
Val	Gly	Tyr	Ile	Cys	Lys	Cys	Asn	Gly	Gly	Arg	Gly	Ala	Ile	Arg	Glu
				405					410					415	
Phe	Ala	Glu	His	Ile	Cys	Leu	Leu	Met	Glu	Lys	Val	Asn	Asn	Ser	Cys
			420					425					430		

Gln Lys

<210> 41

<211> 1080

<212> ADN

<213> Secuencia artificial

<220>

5 <223> Codifica N-acetilneuraminato-9-fosfato sintasa Humana (HsSPS) Optimizada por codón

<400> 41

```

atgccattgg aattggagtt gtgtcctggt agatggggtg gtggtcaaca cccatgtttc 60
atcatcgctg agatcgggtca aaaccaccaa ggagacttgg acgttgctaa gagaatgata 120
agaatggcta aggaatgttg tgctgactgt gctaagttcc agaagtcga gttggagttc 180
aagttcaaca gaaaggcttt ggaaagacca tacacttcca agcactcttg gggaaagact 240
tacggagAAC acaagagaca cttggagttc tctcacgacc aatacagaga gttgcagaga 300
tacgctgagg aagttggtat cttcttctact gcttctggaa tggacgaaat ggctgttgag 360
ttcttgcacg agttgaacgt tccattcttc aaagttgggt ccggtgacac taacaacttc 420
ccatacttgg aaaagactgc taagaaaggt agaccaatgg ttatctcctc tggaatgcag 480
tctatggaca ctatgaagca ggtttaccag atcgtaagc cattgaaccc aaacttttgt 540
ttcttgcagt gtacttccgc ttaccattg caaccagagg acgttaattt gagagttatc 600
tccgagtacc agaagttgtt cccagacatc ccaattgggt actctggtca cgagactggt 660
attgctatatt ccgttgctgc tgttgctttg ggtgctaagg ttttgagag acacatcact 720
ttggacaaga cttggaaggg ttctgatcac tctgcttctt tggaacctgg tgagttggct 780
gaacttgtaa gatcagttag attggttgag agagctttgg gttccccaac taagcaattg 840
ttgccatgtg agatggcttg taacgagaag ttgggaaagt ccgttggtgc taaggtaag 900
atcccagagg gtactatctt gactatggac atgttgactg ttaaagttgg agagccaaag 960
ggttaccac cagaggacat ctttaacttg gttggtaaaa aggttttggg tactgttgag 1020
gaggacgaca ctattatgga ggagttggtt gacaaccacg gaaagaagat caagtcctag 1080

```

10

<210> 42

<211> 359

<212> PRT

<213> Secuencia artificial

15

<220>

<223> N-acetilneuraminato-9-fosfato sintasa Humana (HsSPS) Optimizada por codón

<400> 42

20

```

Met Pro Leu Glu Leu Glu Leu Cys Pro Gly Arg Trp Val Gly Gly Gln
1          5          10          15
His Pro Cys Phe Ile Ile Ala Glu Ile Gly Gln Asn His Gln Gly Asp
20          25          30
Leu Asp Val Ala Lys Arg Met Ile Arg Met Ala Lys Glu Cys Gly Ala
35          40          45
Asp Cys Ala Lys Phe Gln Lys Ser Glu Leu Glu Phe Lys Phe Asn Arg
50          55          60
Lys Ala Leu Glu Arg Pro Tyr Thr Ser Lys His Ser Trp Gly Lys Thr
65          70          75          80
Tyr Gly Glu His Lys Arg His Leu Glu Phe Ser His Asp Gln Tyr Arg
85          90          95
Glu Leu Gln Arg Tyr Ala Glu Glu Val Gly Ile Phe Phe Thr Ala Ser
100         105         110
Gly Met Asp Glu Met Ala Val Glu Phe Leu His Glu Leu Asn Val Pro
115         120         125
Phe Phe Lys Val Gly Ser Gly Asp Thr Asn Asn Phe Pro Tyr Leu Glu

```

130						135					140				
Lys	Thr	Ala	Lys	Lys	Gly	Arg	Pro	Met	Val	Ile	Ser	Ser	Gly	Met	Gln
145					150					155					160
Ser	Met	Asp	Thr	Met	Lys	Gln	Val	Tyr	Gln	Ile	Val	Lys	Pro	Leu	Asn
				165					170					175	
Pro	Asn	Phe	Cys	Phe	Leu	Gln	Cys	Thr	Ser	Ala	Tyr	Pro	Leu	Gln	Pro
			180					185					190		
Glu	Asp	Val	Asn	Leu	Arg	Val	Ile	Ser	Glu	Tyr	Gln	Lys	Leu	Phe	Pro
		195					200				205				
Asp	Ile	Pro	Ile	Gly	Tyr	Ser	Gly	His	Glu	Thr	Gly	Ile	Ala	Ile	Ser
210						215					220				
Val	Ala	Ala	Val	Ala	Leu	Gly	Ala	Lys	Val	Leu	Glu	Arg	His	Ile	Thr
225					230					235					240
Leu	Asp	Lys	Thr	Trp	Lys	Gly	Ser	Asp	His	Ser	Ala	Ser	Leu	Glu	Pro
				245					250					255	
Gly	Glu	Leu	Ala	Glu	Leu	Val	Arg	Ser	Val	Arg	Leu	Val	Glu	Arg	Ala
			260					265					270		
Leu	Gly	Ser	Pro	Thr	Lys	Gln	Leu	Leu	Pro	Cys	Glu	Met	Ala	Cys	Asn
		275					280					285			
Glu	Lys	Leu	Gly	Lys	Ser	Val	Val	Ala	Lys	Val	Lys	Ile	Pro	Glu	Gly
		290				295					300				
Thr	Ile	Leu	Thr	Met	Asp	Met	Leu	Thr	Val	Lys	Val	Gly	Glu	Pro	Lys
305					310					315					320
Gly	Tyr	Pro	Pro	Glu	Asp	Ile	Phe	Asn	Leu	Val	Gly	Lys	Lys	Val	Leu
				325					330					335	
Val	Thr	Val	Glu	Glu	Asp	Asp	Thr	Ile	Met	Glu	Glu	Leu	Val	Asp	Asn
			340					345					350		
His	Gly	Lys	Lys	Ile	Lys	Ser									
			355												

<210> 43

<211> 1092

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> Codifica dominio catalítico de alfa-2,6-sialil transferasa de Ratón (MmmST6) Optimizado por
codón

<400> 43

ES 2 589 655 T3

```

gtttttcaaa tgccaaagtc ccaggagaaa gttgctgttg gtccagctcc acaagctggt 60
ttctccaact ccaagcaaga tccaaaggag ggtgttcaaa tcttgtccta cccaagagtt 120
actgctaagg ttaagccaca accatccttg caagtttggg acaaggactc cacttactcc 180
aagttgaacc caagattggt gaagatttgg agaaactact tgaacatgaa caagtacaag 240
gtttcctaca aggggtccagg tccagggtgtt aagttctccg ttgaggcctt gagatgtcac 300
ttgagagacc acgttaacgt ttccatgatc gaggtactg acttcccatt caacactact 360
gaatgggagg gatacttgcc aaaggagAAC ttcagaacta aggcctggtcc atggcataag 420
tgtgctgttg tttcttctgc tggttccttg aagaactccc agttgggtag agaaattgac 480
aaccacgacg ctgttttgag attcaacggt gctccaactg acaacttcca gcaggatggt 540
ggtactaaga ctactatcag attggttaac tccaatttgg ttactactga gaagagattc 600
ttgaaggact ccttgtacac tgagggaatc ttgattttgt gggacccatc tgtttaccac 660
gctgacattc cacaatggta tcagaagcca gactacaact tcttcgagac ttacaagtcc 720
tacagaagat tgcacccatc ccagccattc tacatcttga agccacaaat gccatgggaa 780
ttgtgggaca tcatccagga aatttcccca gacttgatcc aaccaaaacc accatcttct 840
ggaatgttgg gtatcatcat catgatgact ttgtgtgacc aggttgacat ctacgagttc 900
ttgccatcca agagaaagac tgatgtttgt tactaccacc agaagttctt cgactccgct 960
tgtactatgg gagcttacca cccattgttg ttcgagaaga acatgggttaa gcacttgaac 1020
gaaggtactg acgaggacat ctacttggtc ggaaaggcta ctttgtccgg tttcagaaac 1080
aacagatggt ag 1092

```

<210> 44

<211> 363

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Dominio catalítico de alfa-2,6-sialil transferasa de Ratón (MmmST6) Optimizado por codón

<400> 44

Val	Phe	Gln	Met	Pro	Lys	Ser	Gln	Glu	Lys	Val	Ala	Val	Gly	Pro	Ala
1				5					10					15	
Pro	Gln	Ala	Val	Phe	Ser	Asn	Ser	Lys	Gln	Asp	Pro	Lys	Glu	Gly	Val
			20					25					30		
Gln	Ile	Leu	Ser	Tyr	Pro	Arg	Val	Thr	Ala	Lys	Val	Lys	Pro	Gln	Pro
		35					40					45			
Ser	Leu	Gln	Val	Trp	Asp	Lys	Asp	Ser	Thr	Tyr	Ser	Lys	Leu	Asn	Pro
	50					55					60				
Arg	Leu	Leu	Lys	Ile	Trp	Arg	Asn	Tyr	Leu	Asn	Met	Asn	Lys	Tyr	Lys
65					70					75					80
Val	Ser	Tyr	Lys	Gly	Pro	Gly	Pro	Gly	Val	Lys	Phe	Ser	Val	Glu	Ala
				85					90					95	
Leu	Arg	Cys	His	Leu	Arg	Asp	His	Val	Asn	Val	Ser	Met	Ile	Glu	Ala
			100					105					110		
Thr	Asp	Phe	Pro	Phe	Asn	Thr	Thr	Glu	Trp	Glu	Gly	Tyr	Leu	Pro	Lys
		115					120					125			
Glu	Asn	Phe	Arg	Thr	Lys	Ala	Gly	Pro	Trp	His	Lys	Cys	Ala	Val	Val
	130					135					140				
Ser	Ser	Ala	Gly	Ser	Leu	Lys	Asn	Ser	Gln	Leu	Gly	Arg	Glu	Ile	Asp
145					150					155					160
Asn	His	Asp	Ala	Val	Leu	Arg	Phe	Asn	Gly	Ala	Pro	Thr	Asp	Asn	Phe
				165					170					175	
Gln	Gln	Asp	Val	Gly	Thr	Lys	Thr	Thr	Ile	Arg	Leu	Val	Asn	Ser	Gln
			180					185					190		
Leu	Val	Thr	Thr	Glu	Lys	Arg	Phe	Leu	Lys	Asp	Ser	Leu	Tyr	Thr	Glu
		195					200					205			
Gly	Ile	Leu	Ile	Leu	Trp	Asp	Pro	Ser	Val	Tyr	His	Ala	Asp	Ile	Pro
	210					215					220				
Gln	Trp	Tyr	Gln	Lys	Pro	Asp	Tyr	Asn	Phe	Phe	Glu	Thr	Tyr	Lys	Ser
225					230					235					240
Tyr	Arg	Arg	Leu	His	Pro	Ser	Gln	Pro	Phe	Tyr	Ile	Leu	Lys	Pro	Gln
				245					250					255	
Met	Pro	Trp	Glu	Leu	Trp	Asp	Ile	Ile	Gln	Glu	Ile	Ser	Pro	Asp	Leu
			260					265					270		
Ile	Gln	Pro	Asn	Pro	Pro	Ser	Ser	Gly	Met	Leu	Gly	Ile	Ile	Ile	Met
		275					280					285			
Met	Thr	Leu	Cys	Asp	Gln	Val	Asp	Ile	Tyr	Glu	Phe	Leu	Pro	Ser	Lys
	290					295					300				
Arg	Lys	Thr	Asp	Val	Cys	Tyr	Tyr	His	Gln	Lys	Phe	Phe	Asp	Ser	Ala
305					310					315					320
Cys	Thr	Met	Gly	Ala	Tyr	His	Pro	Leu	Leu	Phe	Glu	Lys	Asn	Met	Val
				325					330					335	
Lys	His	Leu	Asn	Glu	Gly	Thr	Asp	Glu	Asp	Ile	Tyr	Leu	Phe	Gly	Lys
			340					345					350		
Ala	Thr	Leu	Ser	Gly	Phe	Arg	Asn	Asn	Arg	Cys					
		355					360								

<210> 45

<211> 1037

5 <212> ADN

<213> Secuencia artificial

<220>

<223> promotor de PpPMA1

10

<400> 45

ES 2 589 655 T3

```

aaatgcgtac ctcttctacg agattcaagc gaatgagaat aatgtaatat gcaagatcag 60
aaagaatgaa aggagttgaa aaaaaaaacc gttgcgtttt gaccttgaat ggggtggagg 120
tttccattca aagtaaagcc tgtgtcttgg tattttcggc ggcacaagaa atcgtaat 180
tcattcttcta aacgatgaag atcgcagccc aacctgtatg tagttaaccg gtcggaatta 240
taagaaagat tttcgatcaa caaaccttag caaatagaaa gcagggttac aactttaaac 300
cgaagtcaca aacgataaac cactcagctc ccacccaaat tcattcccac tagcagaaag 360
gaattattta atccctcagg aaacctcgat gattctcccg ttcttccatg ggcgggtatc 420
gcaaaatgag gaatttttca aattttctcta ttgtcaagac tgtttattat ctaagaaata 480
gcccaatccg aagctcagtt ttgaaaaaat cacttccgag tttctttttt acagcccgat 540
gaatatccaa atttggaata tggattactc tatcgggact gcagataata tgacaacaac 600
gcagattaca ttttaggtta ggcataaaca ccagccagaa atgaaacgcc cactagccat 660
ggtcgaatag tccaatgaat tcagatagct atggtctaaa agctgatgtt ttttattggg 720
taatggcgaa gagtccagta cgaactccag cagagctgag atggccattt ttgggggtat 780
tagtaacttt ttgagctctt ttcacttcga tgaagtgtcc cattcgggat ataatcggat 840
cgcgctggtt tctcgaaaat acagcttagc gtcgtccgct tgttgtaaaa gcagcaccac 900
attcctaate tcttatataa acaaaacaac ccaaattatc agtgctgttt tcccaccaga 960
tataagtttc ttttctcttc cgttttttga ttttttatct ctttccctta aaaacttctt 1020
taccttaaag ggcggcc 1037

```

<210> 46
 <211> 512
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> terminador de PpPMA1

<400> 46

```

taagcttcac gatttgtgtt ccagttttat ccccttttat ataccgttaa ccttttccct 60
gttgagctga ctgttgttgt attaccgcaa tttttccaag tttgccatgc ttttcgtgtt 120
atttgaccga tgtctttttt cccaaatcaa actatatattg ttaccattta aaccaagtta 180
tcttttgtat taagagtcta agtttgttcc caggcttcat gtgagagtga taaccatcca 240
gactatgatt ctgtgtttttt attgggtttg tttgtgtgat acatctgagt tgtgattcgt 300
aaagtatgtc agtctatcta gatttttaaat agttaattgg taatcaatga cttgttttgt 360
ttaactttta aattgtgggt cgtatccacg cgttttagtat agctgttcat ggctgttaga 420
ggagggcgat gtttatatac agaggacaag aatgaggagg cggcgtgtat ttttaaaatg 480
gagacgcgac tcctgtacac cttatcggtt gg 512

```

<210> 47
 <211> 798
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> promotor de PpOCH1

<400> 47

ES 2 589 655 T3

```

tggacacagg agactcagaa acagacacag agcgttctga gtcctggtgc tcctgacgta 60
ggcctagaac aggaattatt ggctttatct gtttgtccat ttcataaggct tggggtaata 120
gatagatgac agagaaatag agaagaccta atattttttg ttcataggcaa atcgcggtt 180
cgcggtcggg tcacacacgg agaagtaatg agaagagctg gtaatctggg gtaaaagggt 240
tcaaaagaag gtcgcctggt agggatgcaa tacaagggtg tcttgaggtt tacattgacc 300
agatgatttg gctttttctc tgttcaattc acatttttca gcgagaatcg gattgacgga 360
gaaatggcgg ggtgtggggg ggatagatgg cagaaatgct cgcaatcacc gcgaaagaaa 420
gactttatgg aatagaacta ctgggtggtg taaggattac atagctagtc caatggagtc 480
cgttggaaag gtaagaagaa gctaaaaccg gctaagtaac tagggaagaa tgatcagact 540
ttgatttgat gaggtctgaa aatactctgc tgccttttca gttgcttttt ccctgcaacc 600
tatcattttc cttttcataa gcctgccttt tctgttttca cttatatgag ttccgcccag 660
acttcccaa attctctcct ggaacattct ctatcgctct ccttccaagt tgcgccccct 720

```

```

ggcactgcct agtaatatta ccacgcgact tatattcagt tccacaattt ccagtgttcg 780
tagcaaatat catcagcc 798

```

<210> 48
 <211> 302
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> terminador de PpALG12
 <400> 48

```

aatatatacc tcatttgttc aatttgggtg aaagagtgtg gcggatagac ttcttgtaaa 60
tcaggaaagc tacaattcca attgctgcaa aaaataccaa tgcccataaa ccagtatgag 120
cggtgccttc gacggattgc ttactttcgc accctttgtc gtttgattct tctgcctttg 180
gtgagtcagt ttgtttcgac tttatatctg actcatcaac ttcccttaac gttgcgtttt 240
taatcataat tttagccggt ggcttattat cccttgagtt ggtaggaggt ttgatgatgc 300
tg 302

```

<210> 49
 <211> 435
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> promotor de PpSEC4
 <400> 49

```

gaagtaaagt tggcgaaact ttgggaacct ttggttaaaa ctttgtaatt tttgtcgcta 60
cccattaggc agaactctgca tcttgggagg gggatgtggt ggcgttctga gatgtacgcg 120
aagaatgaag agccagtggg aacaacaggc ctagagagat acgggcataa tgggtataac 180
ctacaagtta agaatgtagc agccctggaa accagattga aacgaaaaac gaaatcattt 240
aaactgtagg atgttttggc tcattgtctg gaaggctggc tgtttattgc cctgttcttt 300
gcatgggaat aagctattat atccctcaca taatcccaga aaatagattg aagcaacgcg 360
aaatccttac gtatcgaagt agccttctta cacattcacg ttgtacggat aagaaaacta 420
ctcaaacgaa caatc 435

```

<210> 50
 <211> 404
 <212> ADN
 <213> Secuencia artificial

<220>
<223> terminador de PpOCH1

<400> 50

5

```

aatagatata gcgagattag agaatgaata ccttcttcta agcgatcgtc cgtcatcata 60
gaatatcatg gactgtatag tttttttttt gtacatatata tgattaaacg gtcattccaac 120
atctcgttga cagatctctc agtacgcgaa atccctgact atcaaagcaa gaaccgatga 180
agaaaaaac aacagtaacc caaacaccac aacaaacact ttatcttctc ccccccaaca 240
ccaatcatca aagagatgtc ggaacacaaa caccaagaag caaaaactaa ccccatataa 300
aaacatcctg gtagataatg ctggtaaccc gctctccttc catattctgg gctacttcac 360
gaagtctgac cggctctcagt tgatcaacat gatcctcgaa atgg 404

```

<210> 51
<211> 600
<212> ADN
<213> Secuencia artificial

10

<220>
<223> promotor de PpTEF1

15

<400> 51

```

ttaagggttg gaacaacact aaactacctt gcggtactac cattgacact acacatcctt 60
aattccaatc ctgtctggcc tccttcacct ttttaaccatc ttgcccatc caactcgtgt 120
cagattgcgt atcaagtga aaaaaaaaaa ttttaaatct ttaacccaat caggtaataa 180
ctgtcgctc ttttatctgc cgcactgcat gaggtgtccc cttagtggga aagagtactg 240
agccaaccct ggaggacagc aagggaacaa tacctacaac ttgcttcata atggctgtaa 300
aaacaatcct tgctcgatat aagtgttgta gactgtccct tatcctctgc gatgttcttc 360
ctctcaaagt ttgcgatttc tctctatcag aattgccatc aagagactca ggactaattt 420
cgcagtccca cagcactcgc tacatgattg gctgaaattt ccctaaagaa tttctttttc 480
acgaaaattt tttttttaca caagattttc agcagatata aaatggagag caggacctcc 540
gctgtgactc ttcttttttt tctttttattc tcaactacata catttttagtt attogccaac 600

```

20

<210> 52
<211> 301
<212> ADN
<213> Secuencia artificial

25

<220>
<223> terminador de PpTEF1

<400> 52

```

attgcttgaa gctttaattt attttattaa cataataata atacaagcat gatataatttg 60
tattttgttc gttaacattg atgttttctt cattttactgt tattgtttgt aactttgatc 120
gattttatctt ttctacttta ctgtaatatg gctggcgggt gagccttgaa ctccctgtat 180
tactttacct tgctattact taatctattg actagcagcg acctcttcaa ccgaagggca 240
agtacacagc aagttcatgt ctccgtaagt gtcattcaacc ctggaaacag tgggccatgt 300
c 301

```

30

<210> 53
<211> 486
<212> ADN
<213> Secuencia artificial

35

ES 2 589 655 T3

<220>

<223> promotor de PpGAPDH

<400> 53

5

```

tttttgtaga aatgtcttgg tgtcctcgtc caatcaggta gccatctctg aaatatctgg 60
ctccgttgca actccgaacg acctgctggc aacgtaaaat tctccgggggt aaaacttaaa 120
tgtggagtaa tggaaccaga aacgtctctt cccttctctc tccttccacc gcccgttacc 180
gtccctagga aattttactc tgctggagag cttcttctac ggcccccttg cagcaatgct 240
cttcccagca ttacgttgcg ggtaaaacgg aggtcgtgta cccgacctag cagcccaggg 300
atggaaaagt cccggccgtc gctggcaata atagcgggcg gacgcagtgc atgagattat 360
tggaaccac cagaatcgaa tataaaaggc gaacaccttt cccaattttg gtttctcctg 420
acccaaagac tttaaattta atttatttgt ccctatttca atcaattgaa caactatcaa 480
aacaca                                           486

```

<210> 54

<211> 376

10

<212> ADN

<213> Secuencia artificial

<220>

<223> terminador de PpALG3

15

<400> 54

```

atttacaatt agtaatatta aggtggtaaa aacattcgta gaattgaaat gaattaatat 60
agtatgacaa tggttcatgt ctataaatct ccggcttcgg taccttctcc ccaattgaat 120
acattgtcaa aatgaatggt tgaactatta ggttcgccag tttcgttatt aagaaaactg 180
ttaaaatcaa attccatata atcggttcca gtgggaggac cagttccatc gccaaaatcc 240
tgtaagaatc cattgtcaga acctgtaaag tcagtttgag atgaaatfff tccggtcttt 300
gttgacttgg aagcttcggt aaggttaggt gaaacagttt gatcaaccag cggctcccgt 360
tttcgtcgct tagtag                                           376

```

20

<210> 55

<211> 934

<212> ADN

<213> Secuencia artificial

25

<220>

<223> promotor de PpAOX1 y locus de integración

<400> 55

ES 2 589 655 T3

```

aacatccaaa gacgaaaggt tgaatgaaac ctttttgcca tccgacatcc acaggtccat 60
tctcacacat aagtgcctaa cgcaacagga ggggatacac tagcagcaga ccgttgcaaa 120
cgcaggacct ccactcctct tctcctcaac acccactttt gccatcgaaa aaccagccca 180
gttattgggc ttgattggag ctgcgtcatt ccaattcctt ctattaggct actaacacca 240
tgactttatt agcctgtcta tcttgcccc cctggcgagg ttcatgtttg tttatttccg 300
aatgcaacaa gctccgcatt acaccggaac atcactccag atgagggttt tctgagtgtg 360
gggtcaaata gtttcatgtt ccccaaatgg cccaaaactg acagttttaa cgctgtcttg 420
gaacctaaata tgacaaaagc gtgatctcat ccaagatgaa ctaagtttg ttcgttgaaa 480
tgctaacggc cagtttgtca aaaagaaact tccaaaagtc ggcataccgt ttgtcttgtt 540
tggtattgat tgacgaatgc tcaaaaataa tctcattaat gcttagcgca gtctctctat 600
cgcttctgaa ccccggtgca cctgtgccga aacgcaaag gggaaacacc cgcttttttg 660
atgattatgc attgtctcca cattgtatgc ttccaagatt ctggtgggaa tactgtctgat 720
agcctaacgt tcatgatcaa aatttaactg ttctaaccct tacttgacag caatatataa 780
acagaaggaa gctgcctgtt cttaaaccct tttttttatc atcattatta gcttactttc 840
ataattgcga ctggttccaa ttgacaagct tttgatttta acgactttta acgacaactt 900
gagaagatca aaaaacaact aattattcga aacg                                     934

```

<210> 56
 <211> 293
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> terminador de ScCYC1
 <400> 56

```

acaggccctt tttcctttgt cgatatcatg taattagtta tgtcacgctt acattcacgc 60
cctcctccca catccgctct aaccgaaaag gaaggagtta gacaacctga agtctaggtc 120
cctattttatt ttttttaata gttatgttag tattaagaac gttatttata tttcaaattt 180
ttcttttttt tctgtacaaa cgcgtgtacg catgtaacat tatactgaaa accttgcttg 240
agaagggttt gggacgctcg aaggctttaa tttgcaagct gccggctctt aag                                     293

```

<210> 57
 <211> 427
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> promotor de ScTEF1
 <400> 57

```

gatcccccac acaccatagc ttcaaaatgt ttctactcct tttttactct tccagatttt 60
ctcggactcc gcgcacgccc gtaccacttc aaaacaccca agcacagcat actaaatttc 120
ccctctttct tctctaggg tgtcggtaat taccggtact aaagggtttg aaaagaaaaa 180
agagaccgcc tcgtttcttt ttcttcgtcg aaaaaggcaa taaaaatttt tatcacgttt 240
ctttttcttg aaaatttttt tttttgattt ttttctcttt cgatgacctc ccattgatat 300
ttaagttaat aaacgggtct caatttctca agtttcagtt tcatttttct tgttctatta 360
caactttttt tacttcttgc tcattagaaa gaaagcatag caatctaate taagttttta 420
ttacaaa                                     427

```

<210> 58
 <211> 375
 <212> ADN
 <213> Secuencia artificial

ES 2 589 655 T3

<220>

<223> Codifica ORF de Sh ble (marcador de resistencia a Zeocina):

<400> 58

5

```
atggccaagt tgaccagtgc cgttcgggtg ctcaccgcgc gcgacgtcgc cggagcggtc 60
gagttctgga ccgaccgggt cgggtttctcc cgggacttcg tggaggacga cttcgccgggt 120
gtggtccggg acgacgtgac cctgttcctc agcgcgggtcc aggaccaggt ggtgccggac 180
aacaccctgg cctgggtgtg ggtgcgcggc ctggacgagc tgtacgccga gtggtcggag 240
gtcgtgtcca cgaacttcgc ggacgcctcc gggccggcca tgaccgagat cggcgagcag 300
ccgtgggggc gggagttcgc cctgcgcgac ccggccggca actgcgtgca cttcgtggcc 360
gaggagcagg actga 375
```

<210> 59

<211> 898

10

<212> ADN

<213> Secuencia artificial

<220>

<223> Región en la posición 5' de PpURA5

15

<400> 59

```
atcggccttt gttgatgcaa gttttacgtg gatcatggac taaggagttt tatttggacc 60
aagttcatcg tcctagacat tacggaaagg gttctgctcc tctttttgga aacttttttg 120
aacctctgag tatgacagct tgggtggattg taccatgggt atggcttcct gtgaatttct 180
attttttcta cattggattc accaatcaaa acaaattagt cgccatggct ttttggcttt 240
tgggtctatt tgtttggacc ttcttggaaat atgctttgca tagatttttg ttccacttgg 300
actactatct tccagagaat caaattgcat ttaccattca tttcttattg catgggatac 360
accactattt accaatggat aaatacagat tgggtgatgcc acctacactt ttcattgtac 420
tttgctaccc aatcaagacg ctgcgtcttt ctgttctacc atattacatg gcttgttctg 480
gatttgcagg tggattcctg ggctatatca tgtatgatgt cactcattac gttctgcac 540
actccaagct gcctcgttat ttccaagagt tgaagaaata tcatttggaa catcactaca 600
agaattacga gttaggcttt ggtgtcactt ccaaattctg ggacaaagtc tttgggactt 660
atctgggtcc agacgatgtg tatcaaaaga caaattagag tatttataaa gttatgtaag 720
caaatagggg ctaataggga aagaaaaatt ttggttcttt atcagagctg gctcgcgcgc 780
agtgtttttc gtgctccttt gtaatagtca tttttgacta ctgttcagat tgaaatcaca 840
ttgaagatgt cactcgaggg gtaccaaaaa aggttttttg atgctgcagt ggcttcgc 898
```

20

<210> 60

<211> 1060

<212> ADN

<213> Secuencia artificial

25

<220>

<223> Región en la posición 3' de PpURA5

<400> 60


```

ggctcttttca acaaagctcc attagtgagt cagctggctg aatcttatgc acaggccatc 60
attaacagca acctggagat agacgttgta tttggaccag cttataaagg tattcctttg 120
gctgctatta ccgtgttgaa gttgtacgag ctccggcgca aaaaatacga aaatgtcggg 180
tatgcgttca atagaaaaga aaagaaagac cacggagaag gtggaagcat cgttggagaa 240
agtctaaaga ataaaagagt actgattatc gatgatgtga tgactgcagg tactgctatc 300
aacgaagcat ttgctataat tggagctgaa ggtgggagag ttgaaggtag tattattgcc 360
ctagatagaa tggagactac aggagatgac tcaaatacca gtgctaccca ggctgttagt 420
cagagatatg gtaccctgt cttgagtata gtgacattgg accatattgt ggcccatattg 480
ggcgaaactt tcacagcaga cgagaaatct caaatggaaa cgtatagaaa aaagtatttg 540
cccaaataag tatgaatctg cttcgaatga atgaattaat ccaattatct tctcaccatt 600
atcttcttct gtttcggagc tttgggcacg gcgggcggtg gtgcgggctc aggttccctt 660
tcataaacag atttagtact tggatgctta atagtgaatg gcgaatgcaa aggaacaatt 720
tcgttcatct ttaacccttt cactcggggt acacgttctg gaatgtacct gccctgttgc 780
aactcaggtg gaccgggcaa ttcttgaact ttctgtaacg ttgttggatg ttcaaccaga 840
aattgtccta ccaactgtat tagtttcctt ttggtcttat attgttcacg gagatacttc 900
ccactctcct tgatagccac tctcactcct cctggattac caaatcttg aggatgagtc 960
ttttcaggct ccaggatgca aggtatatcc aagtaacctg aagcatctaa tattgtcttt 1020
gccagggggt tctccacacc atactccttt tggcgcatgc 1060

```

<210> 61
 <211> 957
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Codifica marcador auxotrófico de PpURA5

<400> 61

```

tctagaggga cttatctggg tccagacgat gtgtatcaaa agacaaatta gattatattat 60
aaagttatgt aagcaaatag gggctaatag ggaaagaaaa attttgggtc tttatcagag 120
ctggctcgcg cgcagtgttt ttctgtctcc tttgtaatag tcatttttga ctactgttca 180
gattgaaatc acattgaaga tgtcactgga ggggtaccaa aaaagggttt tggatgctgc 240
agtggcttcg caggccttga agtttggaaac tttcaccttg aaaagtggaa gacagtctcc 300
atacttcttt aacatgggtc ttttcaacaa agctccatta gtgagtcagc tggctgaatc 360
ttatgctcag gccatcatta acagcaacct ggagatagac gttgtatttg gaccagctta 420
taaaggtatt cctttggctg ctattaccgt gttgaagttg tacgagctgg gcggcaaaaa 480
atcagaaaat gtcggatatg cgttcaatag aaaagaaaag aaagaccacg gagaagggtg 540
aagcatcggt ggagaaagtc taaagaataa aagagtactg attatcgatg atgtgatgac 600
tgcaggtact gctatcaacg aagcatttgc tataattgga gctgaagggtg ggagagttga 660
aggttgtatt attgccctag atagaatgga gactacagga gatgactcaa ataccagtgc 720
taccagggct gttagtcaga gatatggtac ccctgtcttg agtatagtga cattggacca 780
tattgtggcc catttgggcg aaactttcac agcagacgag aaatctcaa tggaaacgta 840
tagaaaaaag tatttggcca aataagtatg aatctgcttc gaatgaatga attaatacaa 900
ttatcttctc accattatatt tcttctgttt cggagctttg ggcacggcgg cggtatcc 957

```

<210> 62
 <211> 709
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Parte del gen Ec lacZ

<400> 62

```

cctgcactgg atggtggcgc tggatggttaa gccgctggca agcgggtgaag tgcctctgga 60
tgtcgctcca caaggtaaac agttgattga actgcctgaa ctaccgcagc cggagagcgc 120
cgggcaactc tggctcacag tacgcgtagt gcaaccgaac gcgaccgcac ggtcagaagc 180
cgggcacatc agcgctggc agcagtggtg tctggcggaa aacctcagtg tgacgctccc 240
cgccgcgtcc cagcccatcc cgcactctgac caccagcgaa atggattttt gcatcgagct 300
gggtaataag cgttggcaat ttaaccgcca gtcaggcttt ctttcacaga tgtggattgg 360
cgataaaaaa caactgctga cgccgctgcg cgatcagttc acccgtgcac cgctggataa 420
cgacattggc gtaagtgaag cgaccgcac tgaccctaac gcctgggtcg aacgctggaa 480
ggcggcgggc cattaccagg ccgaagcagc gttgttgtag tgcacggcag atacacttgc 540
tgatgcggtg ctgattacga ccgctcacgc gtggcagcat caggggaaaa ccttatttat 600
cagccggaaa acctaccgga ttgatggtag tgggtcaaag gcgattaccg ttgatgttga 660
agtggcgagc gataaccgc atccggcgcg gattggcctg aactgccag 709

```

<210> 63

<211> 222

5 <212> PRT

<213> *Pichia pastoris*

<400> 63

```

Met Ser Leu Glu Gly Tyr Gln Lys Arg Phe Leu Asp Ala Ala Val Ala
 1          5          10          15
Ser Gln Ala Leu Lys Phe Gly Thr Phe Thr Leu Lys Ser Gly Arg Gln
 20          25          30
Ser Pro Tyr Phe Phe Asn Met Gly Leu Phe Asn Lys Ala Pro Leu Val
 35          40          45
Ser Gln Leu Ala Glu Ser Tyr Ala Gln Ala Ile Ile Asn Ser Asn Leu
 50          55          60
Glu Ile Asp Val Val Phe Gly Pro Ala Tyr Lys Gly Ile Pro Leu Ala
 65          70          75          80
Ala Ile Thr Val Leu Lys Leu Tyr Glu Leu Gly Gly Lys Lys Tyr Glu
 85          90          95

```

```

Asn Val Gly Tyr Ala Phe Asn Arg Lys Glu Lys Lys Asp His Gly Glu
 100          105          110
Gly Gly Ser Ile Val Gly Glu Ser Leu Lys Asn Lys Arg Val Leu Ile
 115          120          125
Ile Asp Asp Val Met Thr Ala Gly Thr Ala Ile Asn Glu Ala Phe Ala
 130          135          140
Ile Ile Gly Ala Glu Gly Gly Arg Val Glu Gly Cys Ile Ile Ala Leu
 145          150          155          160
Asp Arg Met Glu Thr Thr Gly Asp Asp Ser Asn Thr Ser Ala Thr Gln
 165          170          175
Ala Val Ser Gln Arg Tyr Gly Thr Pro Val Leu Ser Ile Val Thr Leu
 180          185          190
Asp His Ile Val Ala His Leu Gly Glu Thr Phe Thr Ala Asp Glu Lys
 195          200          205
Ser Gln Met Glu Thr Tyr Arg Lys Lys Tyr Leu Pro Lys Glx
 210          215          220

```

10

<210> 64

<211> 2875

<212> ADN

15 <213> Secuencia artificial

<220>

<223> Región en la posición 5' de PpOCH1

<400> 64

aaaacctttt	ttcctattca	aacacaaggc	attgcttcaa	cacgtgtgcg	tatccttaac	60
acagatactc	catacttcta	ataatgtgat	agacgaatac	aaagatgttc	actctgtgtt	120
gtgtctacaa	gcattttctta	ttctgatttg	ggatattcta	gttacagcac	taaacaactg	180
gcgatacaaa	cttaaattaa	ataatccgaa	tctagaaaat	gaacttttgg	atgggtccgcc	240
tgttggttgg	ataaatcaat	accgattaaa	tggattctat	tccaatgaga	gagtaatcca	300
agacactctg	atgtcaataa	tcattttgctt	gcaacaacaa	acccgtcatc	taatcaaagg	360
gtttgatgag	gcttaccttc	aattgcagat	aaactcattg	ctgtccactg	ctgtattatg	420
tgagaatatg	ggtgatgaat	ctggctcttct	ccactcagct	aacatggctg	tttgggcaaa	480
ggtggtacaa	ttatacggag	atcaggcaat	agtgaattg	ttgaatatgg	ctactggacg	540
atgcttcaag	gatgtacgtc	tagtaggagc	cgtgggaaga	ttgctggcag	aaccagttgg	600
cacgtcgcaa	caatcccaa	gaaatgaaat	aagtgaatac	gtaacgtcaa	agacagcaat	660
ggagtcaata	ttgataaac	cactggcaga	gcggttcgta	cgtcgttttg	gagccgatat	720
gaggctcagc	gtgctaacag	cacgattgac	aagaagactc	tcgagtgcga	gtaggttgag	780
taaagtattc	gcttagattc	ccaaccttcg	ttttattctt	tcgtagacaa	agaagctgca	840
tgcgaacata	gggacaactt	ttataaatcc	aattgtcaaa	ccaacgtaaa	accctctggc	900
accattttca	acatatattt	gtgaagcagt	acgcaatctc	gataaatact	caccgttggt	960
tgtaacagcc	ccaacttgca	tacgccttct	aatgacctca	aatggataag	ccgcagcttg	1020
tgctaacata	ccagcagcac	cgcccgcggt	cagctgcgcc	cacacatata	aaggcaatct	1080
acgatcatgg	gaggaattag	ttttgaccgt	caggctctca	agagttttga	actcttcttc	1140
ttgaactgtg	taacctttta	aatgacggga	tctaataacg	tcattggatga	gatcatgtgt	1200
gtaaaaactg	actccagcat	atggaatcat	tccaaagatt	gtaggagcga	accacagata	1260
aaagtttccc	aaccttgcca	aagtgtctaa	tgctgtgact	tgaaatctgg	gttctctcgt	1320
gaagaccctg	cgtactatgc	ccaaaaactt	tcctccacga	gccctattaa	cttctctatg	1380
agtttcaaat	gccaaacgga	cacggattag	gtccaatggg	taagtgaata	acacagagca	1440
aaccccgagc	aatgagccgg	ccagtaaccg	tcttgaggct	gtttcataag	agtcattagg	1500
gatcaataac	gttctaattc	gttcataaca	tacaaatttt	atggctgcat	agggaaaaat	1560
tctcaacagg	gtagccgaat	gacctgata	tagacctgcg	acaccatcat	acccatagat	1620
ctgctgaca	gccttaaga	gcccgcataa	agaccgggaa	aaccgagaga	actctggatt	1680
agcagtctga	aaaagaatct	tactctgtgc	tagtgaggca	attaatgtct	tagcggcact	1740
tccgtctact	ccgccagcta	ctcctgaata	gatcacatac	tgcaaagact	gcttgtcgat	1800
gaccttgggg	ttattttagct	tcaagggcaa	tttttgggac	atthttggaca	caggagactc	1860
agaaacagac	acagagcggt	ctgagtcctg	gtgctcctga	cgtaggccta	gaacaggaat	1920
tattggcttt	atthgtttgt	ccattttcata	ggcttggggg	aatagataga	tgacagagaa	1980
atagagaaga	cctaataattt	tttgttcatt	gcaaatcgcg	gggttcgagg	cgggtcacac	2040
acggagaagt	aatgagaaga	gctggtaatc	tggggtaaaa	gggttcaaaa	gaaggtcgcc	2100
tggtagggat	gcaatacaag	gttgtcttgg	agtttacatt	gaccagatga	tttggctttt	2160
tctctgttca	attcacattt	ttcagcgaga	atcggattga	cggagaaatg	gcgggggtgt	2220
gggtggatag	atggcgagaa	tgctcgcaat	caccgcgaaa	gaaagacttt	atggaataga	2280
actactgggt	ggtgtaagga	ttacatagct	agtccaatgg	agtccgttgg	aaaggtaaga	2340
agaagctaaa	accggctaag	taactagggg	agaatgatca	gactttgatt	tgatgaggtc	2400
tgaaaatact	ctgctgcttt	ttcagttgct	ttttccctgc	aacctatcat	tttctttttc	2460
ataagcctgc	cttttctgtt	ttcacttata	tgagttccgc	cgagacttcc	ccaaattctc	2520
tcctggaaca	ttctctatcg	ctctccttcc	aagttgcgcc	ccctggcact	gcctagtaat	2580
attaccacgc	gacttatatt	cagttccaca	atttccagtg	ttcgtagcaa	atatcatcag	2640
ccatggcgaa	ggcagatggc	agtttgctct	actataatcc	tcacaatcca	cccagaaggt	2700
attacttcta	catggctata	ttcgccggtt	ctgtcatttg	cgttttgtac	ggacctcac	2760
aacaattatc	atctccaaaa	atagactatg	atccattgac	gctccgatca	cttgatttga	2820
agactttgga	agctccttca	cagttgagtc	caggcaccgt	agaagataat	cttcg	2875

5

<210> 65

<211> 997

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Región en la posición 3' de PpOCH1

<400> 65

```

aaagctagag taaaatagat atagcgagat tagagaatga ataccttctt ctaagcgatc 60
gtccgtcatc atagaatatc atggactgta tagttttttt tttgtacata taatgattaa 120
acgggtcatcc aacatctcgt tgacagatct ctccgtacgc gaaatccctg actatcaaag 180
caagaaccga tgaagaaaaa aacaacagta acccaaacac cacaacaaac actttatctt 240
ctcccccca acaccaatca tcaaagagat gtcggaacca aacaccaaga agcaaaaaact 300
aaccatcatat aaaaacatcc tggtagataa tgctggtaac ccgctctcct tccatattct 360
gggctacttc acgaagtctg accggtctca gttgatcaac atgatcctcg aaatgggtgg 420
caagatcggt ccagacctgc ctccctcgtt agatggagtg ttgtttttga caggggatta 480
caagtctatt gatgaagata ccctaaagca actgggggac gttccaatat acagagactc 540
cttcatctac cagtgttttg tgcacaagac atctcttccc attgacactt tccgaattga 600
caagaacgtc gacttggttc aagatttgat caatagggcc cttcaagagt ctgtggatca 660
tgtcacttct gccagcacag ctgcagctgc tgctgttggt gtcgctacca acggcctgtc 720
ttctaaacca gacgctcgtc ctagcaaaat acagttcact cccgaagaag atcgttttat 780
tcttgacttt gttaggagaa atcctaaacg aagaaacaca catcaactgt aactgagct 840
cgctcagcac atgaaaaacc atacgaatca ttctatccgc cacagatttc gtcgtaattc 900
ttccgctcaa cttgattggg tttatgatat cgatccattg accaaccaac ctcgaaaaa 960
tgaaaacggg aactacatca aggtacaagg ccttcca 997

```

5 <210> 66
 <211> 870
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Región en la posición 5' de PpBMT2
 <400> 66

```

ggccgagcgg gcctagattt tcactacaaa tttcaaaact acgcggattt attgtctcag 60
agagcaattt ggcatttctg agcgtagcag gaggcttcat aagattgtat aggaccgtac 120
caacaaattg ccgaggcaca acacgggatg ctgtgcactt atgtggctac ttccctacaa 180
cggaatgaaa ccttcctctt tccgcttaaa cgagaaagtg tgtcgcaatt gaatgcagg 240
gcctgtgcgc cttggtgtat tggttttgag ggccaattt atcaggcgcc ttttttcttg 300
gttggtttcc cttagcctca agcaagggtg gtctatttca tctccgcttc tataccgtgc 360
ctgatactgt tggatgagaa cacgactcaa ctctctgctg ctctgtattg ccagtgtttt 420
gtctgtgatt tggatcgag tcctccttac ttggaatgat aataatcttg gcggaatctc 480
cctaaacgga ggcaaggatt ctgcctatga tgatctgcta tcattgggaa gcttcaacga 540
catggagggtc gactcctatg tcaccaacat ctacgacaat gctccagtgc taggatgtac 600
ggatttgtct tatcatggat tgttgaaagt caccctaaag catgacttag cttgcgattt 660
ggagttcata agagctcaga ttttgacat tgacgtttac tccgccataa aagacttaga 720
agataaagcc ttgactgtaa aacaaaagg tgaacaaac tggtttacgt tttatggtag 780
ttcagtcctt ctgcccgaac acgatgtgca ttacctggtt agacgagtca tcttttcggc 840
tgaaggaaag gcgaactctc cagtaacatc 870

```

15 <210> 67
 <211> 1733
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Región en la posición 3' de PpBMT2
 25 <400> 67


```

ccatatgatg ggtgtttgct cactcgtatg gatcaaaatt ccatggtttc ttctgtacaa 60
cttgtacact tatttggact tttctaacgg tttttctggg gatttgagaa gtccttattt 120
tgggtgttcg agcttatccg tgattgaacc atcagaaata ctgcagctcg ttatctagtt 180
tcagaatgtg ttgtagaata caatcaattc tgagtctagt ttgggtgggt cttggcgacg 240
ggaccgttat atgcatctat gcagtgttaa ggtacataga atgaaaatgt aggggttaat 300
cgaaagcatc gttaatttca gtagaacgta gttctattcc ctacccaaat aatttgccaa 360
gaatgcttcg tatccacata cgcagtggac gtagcaaatt tcactttgga ctgtgacctc 420
aagtcgttat cttctacttg gacattgatg gtcattacgt aatccacaaa gaattggata 480
gcctctcgtt ttatctagtg cacagcctaa tagcacttaa gtaagagcaa tggacaaatt 540
tgcatagaca ttgagctaga tacgtaactc agatcttggt cactcatggt gtactcgaag 600
tactgctgga accgttacct cttatcattt cgctactggc tcgtgaaact actggatgaa 660
aaaaaaaaa gagctgaaag cgagatcatc ccattttgtc atcatacaaa ttcacgcttg 720
cagttttgct tcgttaacaa gacaagatgt ctttatcaaa gaccggtttt ttcttcttga 780
agaatacttc cctgttgagc acatgcaaac catattttatc tcagatttca ctcaacttgg 840
gtgcttccaa gagaagtaaa attcttccca ctgcatcaac ttccaagaaa cccgtagacc 900
agtttctctt cagccaaaag aagttgctcg ccgatacccg cggtaacaga ggagtcagaa 960
ggtttcacac ccttccatcc cgatttcaaa gtcaaagtgc tgcgttgaa caaggttttc 1020
aggttgccaa agcccagtct gcaaaaacta gttccaaatg gcctattaat tcccataaaa 1080
gtgttggtc cgtatgtatc ggtacctcca ttctggtatt tgctattggt gtcgttggtg 1140
ggttgactag actgaccgaa tccggtcttt ccataacgga gtggaaacct atcactggtt 1200
cgggtccccc actgactgag gaagactgga agttggaatt tgaaaaatac aaacaaagcc 1260
ctgagtttca ggaactaaat tctcacataa cattggaaga gttcaagttt atattttcca 1320
tggaatgggg acatagattg ttgggaaggg tcatcggcct gtcgtttggt cttcccacgt 1380
tttacttcat tgcccgtcga aagtgttcca aagatgttgc attgaaactg cttgcaatat 1440
gctctatgat aggattccaa ggtttcatcg gctggtggat ggtgtattcc ggattggaca 1500
aacagcaatt ggctgaacgt aactccaaac caactgtgtc tccatatcgc ttaactacct 1560
atcttggaa tgcatttgtt atttactggt acatgattta cacagggtt caagttttga 1620
agaactataa gatcatgaaa cagcctgaag cgtatgttca aattttcaag caaattgcgt 1680
ctccaaaatt gaaaactttc aagagactct cttcagttct attaggcctg gtg 1733

```

<210> 68
 <211> 411
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 5' de PpBMT1
 <400> 68

```

catatgggtga gagccgttct gcacaactag atgttttcga gcttcgcatt gtttcctgca 60
gctcgactat tgaattaaga tttccggata tctccaatct cacaaaaact tatgttgacc 120
acgtgctttc ctgaggcgag gtgttttata tgcaagctgc caaaaatgga aaacgaatgg 180
ccatttttcg cccaggcaaa ttattcgatt actgctgtca taaagacagt gttgcaaggc 240
tcacattttt ttttaggatc cgagataaag tgaatacagg acagcttata tctatatctt 300
gtaccattcg tgaatcttaa gagttcgggt agggggactc tagttgaggg ttggcactca 360
cgtatggctg ggccgagaaa taaaattcag gcgcagcagc acttatcgat g 411

```

<210> 69
 <211> 692
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 3' de PpBMT1
 <400> 69

gaattcacag ttataaataa aaacaaaaac tcaaaaagtt tgggctccac aaaataactt 60
aatttaaatt tttgtctaataaatgaatgt aattccaaga ttatgtgatg caagcacagt 120

atgcttcagc cctatgcagc tactaatgtc aatctcgcct gcgagcgggc ctagattttc 180
actacaaatt tcaaaactac gcggatttat tgtctcagag agcaatttgg catttctgag 240
cgtagcagga ggcttcataa gattgtatag gaccgtacca acaaattgcc gaggcacaac 300
acggtatgct gtgcacttat gtggctactt ccctacaacg gaatgaaacc ttctcttttc 360
cgcttaaacg agaaaagtgtg tcgcaattga atgcaggtgc ctgtgcgcct tgggtgtattg 420
tttttgaggg cccaatttat caggcgcctt ttttcttggg tgttttccct tagcctcaag 480
caaggttggg ctatttcctc tccgcttcta taccgtgcct gatactgttg gatgagaaca 540
cgactcaact tcctgctgct ctgtattgcc agtgttttgg ctgtgatttg gatcgagagc 600
ctccttactt ggaatgataa taatcttggc ggaatctccc taaacggagg caaggattct 660
gcctatgatg atctgctatc attggaagc tt 692

<210> 70
<211> 546
<212> ADN
<213> Secuencia artificial

<220>
<223> Región en la posición 5' de PpBMT3

<400> 70

gatattctccc tggggacaat atgtgttgca actgttcggt gttggtgccc cagtccccca 60
accggtacta atcgggtctat gttcccgtaa ctcatattcg gttagaacta gaacaataag 120
tgcattcattt ttcaacattg tgggttcaatt gtccgaacatt gctggtgctt atatctacag 180
ggaagacgat aagcctttgt acaagagagg taacagacag ttaattggta tttctttggg 240
agtcgttgcc ctctacgttg tctccaagac atactacatt ctgagaaaca gatggaagac 300
tcaaaaatgg gagaagctta gtgaagaaga gaaagttgcc tacttgagca gagctgagaa 360
ggagaacctg ggttctaaga ggctggactt tttgttcgag agttaaaactg cataattttt 420
tctaagttaa tttcatagtt atgaaatttc tgcagcttag tgtttactgc atcgtttact 480
gcatcaccct gtaaataatg tgagcttttt tccctccatt gcttggtatc ttcttgctg 540
ctgttt 546

<210> 71
<211> 378
<212> ADN
<213> Secuencia artificial

<220>
<223> Región en la posición 3' de PpBMT3

<400> 71

acaaaacagt catgtacaga actaacgcct ttaagatgca gaccactgaa aagaattggg 60
tcccattttt cttgaaagac gaccaggaat ctgtccattt tgtttactcg ttcaatcctc 120
tgagagtact caactgcagt cttgataacg gtgcatgtga tgtttctattt gagttaccac 180
atgatttttg catgtcttcc gagctacgtg gtgccactcc tatgctcaat ctctctcagg 240
caatcccgat ggcagacgac aaagaaattt gggtttcatt cccaagaacg agaatatcag 300
attgcgggtg ttctgaaaca atgtacaggc caatgttaat gctttttgtt agagaaggaa 360
caaacttttt tgctgagc 378

<210> 72
<211> 1043

<212> ADN

<213> Secuencia artificial

<220>

5 <223> Región en la posición 5' de PpBMT4

<400> 72

```

aagcttgttc accgttggga cttttccgtg gacaatgttg actactccag gagggattcc 60
agctttctct actagctcag caataatcaa tgcagcccca ggcgcccgtt ctgatggcctt 120
gatgaccgtt gtattgcctg tcactatagc caggggtagg gtccataaag gaatcatagc 180
agggaaatta aaagggcata ttgatgcaat cactcccaat ggctctcttg ccattgaagt 240
ctccatatca gcactaactt ccaagaagga ccccttcaag tctgacgtga tagagcacgc 300
ttgctctgcc acctgtagtc ctctcaaaac gtcaccttgt gcatcagcaa agactttacc 360
ttgctccaat actatgacgg aggcaattct gtcaaaattc tctctcagca attcaaccaa 420
cttgaagaca aattgctgtc tcttgatgat ggagactttt ttccaagatt gaaatgcaat 480

gtgggacgac tcaattgctt cttccagctc ctcttcgggt gattgaggaa cttttgaaac 540
cacaaaattg gtcgttgggt catgtacatc aaaccattct gtagatttag attcgacgaa 600
agcgttgttg atgaaggaaa aggttggata cggtttgtcg gtctcttttg tatggccggt 660
ggggtatgca attgcagtag aagataattg gacagccatt gttgaaggta gagaaaagg 720
cagggaaactt gggggttatt tataccattt taccacacaa ataacaactg aaaagtacc 780
attccatagt gagaggtaac cgacggaaaa agacgggccc atgttctggg accaatagaa 840
ctgtgtaatc cattgggact aatcaacaga cgattggcaa tataatgaaa tagttcgttg 900
aaaagccacg tcagctgtct ttccattaac tttggtcgga cacaacattt tctactgttg 960
tatctgtcct actttgctta tcatctgcca cagggcaagt ggatttcctt ctgcgcggc 1020
tgggtgaaaa cggttaacgt gaa                                     1043

```

10

<210> 73

<211> 695

<212> ADN

15 <213> Secuencia artificial

<220>

<223> Región en la posición 3' de PpBMT4

<400> 73

20

```

gccttggggg acttcaagtc tttgctagaa actagatgag gtcaggccct cttatggttg 60
tgtcccaatt gggcaatttc actcacctaa aaagcatgac aattatttag cgaaatagg 120
agtatatttt cctcatctc ccaagcagtt tcgtttttgc atccatatct ctcaaagtag 180
cagctacgac tcattagaac cagagtcaag taggggtgag ctgagtcac agccttcgtt 240
tctaaaacga ttgagttctt ttgttgctac aggaagcgcc ctagggaact ttcgcacttt 300
ggaaatagat tttgatgacc aagagcggga gttgatatta gagaggctgt ccaaagtaca 360
tgggatcagg cgggccaaat tgattgggtg gactaaacca ttgtgtactt ggacactcta 420
ttacaaaagc gaagatgatt tgaagtatta caagtcccga agtgtagtag gattctatcg 480
agcccagaat gaaatcatca accgttatca gcagattgat aaactcttg aaagcgggat 540
cccattttca ttattgaaga actacgataa tgaagatgtg agagacggcg accctctgaa 600
cgtagacgaa gaaacaaatc tacttttggg gtacaataga gaaagtgaat caaggagagt 660
atttgtggcc ataatactca actctatcat taatg                                     695

```

<210> 74

25 <211> 937

<212> ADN

<213> Secuencia artificial

<220>

30 <223> Región en la posición 5' de PpPNO1 y PpMNN4

<400> 74

```

tcattctata tgttcaagaa aagggtagtg aaaggaaaga aaaggcatat aggcgagggga 60
gagttagcta gcatacaaga taatgaagga tcaatagcgg tagttaaagt gcacaagaaa 120
agagcacctg ttgaggctga tgataaagct ccaattacat tgccacagag aaacacagta 180
acagaaatag gaggggatgc accacgagaa gagcattcag tgaacaactt tgccaaattc 240
ataaccccaa gcgctaataa gccaatgtca aagtcggcta ctaacattaa tagtacaaca 300
actatcgatt ttcaaccaga tgtttgcaag gactacaaac agacagggtta ctgcggatat 360
ggtgacactt gtaagttttt gcacctgagg gatgatttca aacagggatg gaaattagat 420
agggagtggtg aaaatgtcca aaagaagaag cataatactc tcaaaggggt taaggagatc 480
caaatgttta atgaagatga gctcaaagat atcccgttta aatgcattat atgcaaagga 540
gattacaaat caccctgtaa aacttcttgc aatcattatt tttgcgaaca atgtttcctg 600
caacggtcaa gaagaaaacc aaattgtatt atatgtggca gagacacttt aggagttgct 660
ttaccagcaa agaagttgtc ccaatttctg gctaagatac ataataatga aagtaataaa 720
gttttagtaat tgcattgcgt tgactattga ttgcattgat gtcgtgtgat actttcaccg 780
aaaaaaaaaca cgaagcgcaa taggagcggg tgcattattag tcccaaagc tatttaattg 840
tgctgaaac tgttttttaa gtcctcaag cataattgta tgcattgcga cgtaaccaac 900
gtttaggcgc agtttaatca tagccactg ctaagcc 937

```

5 <210> 75
 <211> 1906
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Región en la posición 3' de PpPNO1 y PpMNN4

<400> 75


```

cggaggaatg caaataataa tctccttaat taccactga taagctcaag agacgcgggt 60
tgaaaacgat ataatgaatc atttggattt tataataaac cctgacagtt tttccactgt 120
attgttttaa cactcattgg aagctgtatt gattctaaga agctagaaat caatacggcc 180
atacaaaaga tgacattgaa taagcaccgg cttttttgat tagcatatac cttaaagcat 240
gcattcatgg ctacatagtt gttaaagggc ttcttccatt atcagtataa tgaattacat 300
aatcatgcac ttatatttgc ccatctctgt tctctcactc ttgcctgggt atattctatg 360
aaattgcgta tagcgtgtct ccagttgaac cccaagcttg gcgagtttga agagaatgct 420
aaccttgctt attccttgct tcaggaaaca ttcaaggaga aacaggtcaa gaagccaaac 480
atcttgatcc ttcccgagtt agcattgact ggctacaatt ttcaaagcca gcagcggata 540
gagccttttt tggaggaaac aaccaagggg gctagtaccc aatgggctca aaaagtatcc 600
aagacgtggg attgctttac tttaatagga taccagaaaa aaagtttaga gagccctccc 660
cgtatttaca acagtgcggg acttgatcg cctcagggaa aagtaatgaa caactacaga 720
aagtccttct tgtatgaagc tgatgaacat tggggatgtt cggaatcttc tgatgggttt 780
caaacagtag atttattaat tgaaggaaag actgtaaaaga catcatttgg aatttgcag 840
gatttgaatc cttataaatt tgaagctcca ttcacagact tcgagttcag tggccattgc 900
ttgaaaaccg gtacaagact cattttgtgc ccaatggcct ggttgtcccc tctatcgct 960
tccattaaaa aggatcttag tgatatagag aaaagcagac ttcaaaagt ctaccttgaa 1020
aaaatagata ccccggaatt tgacgttaat tacgaattga aaaaagatga agtattgccc 1080
accgctatga atgaaacgtt ggaaacaatt gactttgagc cttcaaaacc ggactactct 1140
aatataaatt attggatact aaggtttttt cctttctga ctcattgtct taaacgagat 1200
gtgctcaaag agaatgcagt tgcagctcta tgcaaccgag ttggcattga gagtgtgtc 1260
ttgtacggag gatcaaccac gattctaaac ttcaatggta agttagcatc gacacaagag 1320
gagctggagt tgtacgggca gactaatagt ctcaacccca gtgtggaagt attggggggc 1380
cttggcatgg gtcaacaggg aattctagta cgagacattg aattaacata atatacaata 1440
tacaataaac acaataaaag aatacaagcc tgacaaaaat tcacaaatta ttgcctagac 1500
ttgtcgttat cagcagcgac ctttttccaa tgctcaattt cacgatatgc cttttctagc 1560
tctgctttta gcttctcatt ggaattggct aactcgttga ctgcttggtc agtgatgagt 1620
ttctccaagg tccatttctc gatgttggtt ttttcgtttt cctttaatct cttgatataa 1680
tcaacagcct tctttaatat ctgagccttg ttcgagtcct ctgttggtgaa cagagcggcc 1740
agttccttta ttccgtggtt tatattttct cttctacgcc tttctacttc tttgtgattc 1800
tctttacgca tcttatgcca ttcttcagaa ccagtggtcg gcttaaccga atagccagag 1860
cctgaagaag ccgcactaga agaagcagtg gcattgttga ctatgg 1906

```

<210> 76
 <211> 1128
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 5' de PpMNN4L1

<400> 76

```

gatctggcca ttgtgaaact tgacactaaa gacaaaactc ttagagtttc caatcactta 60
ggagacgatg tttcctacaa cgagtacgat ccctcattga tcatgagcaa tttgtatgtg 120
aaaaaagtc tgcaccttga caccttggat aaaagggctg gaggaggtgg aaccacctgt 180
gcaggcggtc tgaaagtgtt caagtacgga tctactacca aatatacatc tggtaacctg 240
aacggcggtc ggtaggtata ctggaacgaa ggaaagtgtc aaagctccaa atttgtggtt 300
cgatcctcta attactctca aaagcttggg ggaaacagca acgccgaatc aattgacaac 360
aatggtgtgg gttttgcctc agctggagac tcaggcgcat ggattctttc caagctacaa 420
gatgttaggg agtaccagtc attcactgaa aagctaggtg aagctacgat gaggattttc 480
gatttccacg gtcttaaaaca ggagacttct actacagggc ttggggtagt tggtagtatt 540
cattcttacg acggtgagtt caaacagttt gggttgttca ctccaatgac atctattcta 600
caaagacttc aacgagtgc caatgtagaa tgggtgttag cgggttgcca agatggggat 660
gtggacactg aaggagaaca cgaattgagt gatttggaa aactgcatat gcatagttag 720
tccgactagt caggcaagag agagccctca aatttacctc tctgccctc ctcactcctt 780
ttggtacgca taattgcagt ataaagaact tgctgccagc cagtaatctt atttcatacg 840
cagttctata tagcacataa tcttgcttgt atgtatgaaa tttaccgctg tttagttgaa 900
attgtttatg ttgtgtgcct tgcataaaat ctctcgtag ccctatcctt acatttaact 960
ggtctcaaaa cctctaccaa ttccattgct gtacaacaat atgaggcggc attactgtag 1020
ggttggaaaa aaattgtcat tccagctaga gatcacacga cttcatcacg cttattgtct 1080
ctcattgcta aatcatttac tcttgacttc gaccagaaa agttcgcc 1128

```

<210> 77
 <211> 1231
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 3' de PpMNN4L1
 <400> 77

```

gcatgtcaaaa cttgaacaca acgactagat agttgttttt tctatataaaa acgaaacggt 60
atcatcttta ataatcattg aggtttaccc ttatagttcc gtattttcgt ttccaaactt 120
agtaatcttt tggaatatc atcaaagctg gtgccaatct tcttgtttga agtttcaaac 180
tgctccacca agctacttag agactgttct aggtctgaag caacttcgaa cacagagaca 240
gctgccgccc attgttcttt tttgtgtttt tcttctggaa gaggggcatc atcttgtatg 300
tccaatgccc gtatcctttc tgagttgtcc gacacattgt ccttcgaaga gtttcctgac 360
attgggcttc ttctatccgt gtattaattt tgggttaagt tcctcgtttg catagcagtg 420
gatacctcga tttttttggc tcctatttac ctgacataat attctactat aatccaactt 480
ggacgcgtca tctatgataa ctaggctctc ctttgttcaa aggggacgtc ttcataatcc 540
actggcacga agtaagtctg caacgaggcg gcttttgcaa cagaacgata gtgtcgtttc 600
gtacttggac tatgctaaac aaaaggatct gtcaaacatt tcaaccgtgt ttcaaggcac 660
tctttacgaa ttatcgacca agaccttcct agacgaacat ttcaacatat ccaggctact 720
gcttcaaggt ggtgcaaatg ataaaggtat agatattaga tgtgtttggg acctaaaaca 780
gttcttgctt gaagattccc ttgagcaaca ggcttcaata gccaggttag agaagcagta 840
ccaaatcggg aacaaaaggg ggaagcatat aaaaccttta ctattgcgac aaaatccatc 900
cttgaaagta aagctgtttg ttcaatgtaa agcatacgaa acgaaggagg tagatcctaa 960
gatggttaga gaacttaacg ggacatactc cagctgcac ccatattacg atcgctggaa 1020
gacttttttc atgtacgtat cgcccaccaa ctttcaaag caagctaggt atgattttga 1080
cagttctcac aatccattgg ttttcatgca acttgaaaaa acccaactca aacttcatgg 1140
ggatccatac aatgtaaatc attacgagag ggcgaggttg aaaagtttcc attgcaatca 1200
cgctcgcatca tggctactga aaggccttaa c 1231

```

<210> 78
 <211> 1815
 <212> ADN
 <213> Secuencia artificial

<220>

<223> locus de integración del gen PpTRP2

<400> 78

```

taatggccaa acggtttctc aattactata tactactaac catttacctg tagcgtattt 60
cttttccctc ttgcgaaaag ctcaaggcca tcttcttgac tcatgaaaaa tatctggatt 120
tcttctgaca gatcatcacc cttgagccca actctctagc ctatgagtgt aagtgatagt 180
catcttgcaa cagattattht tggaaacgca ctaacaaagc agatacaccc ttcagcagaa 240
tcctttctgg atattgtgaa gaatgatcgc caaagtcaca gtcctgagac agttcctaatt 300
ctttacccca tttaacaagt catccaatca gacttcttaa cgcctcatct ggcttatatc 360
aagcttacca acagttcaga aactcccagt ccaagtttct tgcttgaaaag tgcgaagaat 420
ggtgacaccg ttgacaggta cacttttatg ggacattccc ccagaaaaat aatcaagact 480
gggccttttag aggggtgctga agttgacccc ttggtgcttc tggaaaaaga actgaagggc 540
accagacaag cgcaacttcc tggatttcct cgtctaagtg gtggtgccat aggatacatc 600
tcgtacgatt gtattaagta ctttgaacca aaaactgaaa gaaaactgaa agatgttttg 660
caacttccgg aagcagcttt gatgttggtc gacacgatcg tggcttttga caatgtttat 720
caaagattcc aggttaattgg aaacgtttct ctatccgttg atgactcggg cgaagctatt 780
cttgagaaat attataagac aagagaagaa gtggaaaaga tcagtaaagt ggtatttgac 840
aataaaactg ttccctacta tgaacagaaa gatattatc aaggccaaac gttcacctct 900
aatattgggtc aggaagggtg tgaaaacccat gttcgcaagc tgaaagaaca tattctgaaa 960
ggagacatct tccaagctgt tccctctcaa agggtagcca ggccgacctc attgcacctc 1020
ttcaacatct atcgtcattt gagaactgtc aatccttctc catacatgtt ctatattgac 1080
tatctagact tccaagttgt tgggtgcttc cctgaattac tagttaaatc cgacaacaac 1140
aacaaaatca tcacacatcc tattgctgga actcttccca gaggtaaaac tatcgaagag 1200
gacgacaatt atgctaagca attgaagtcg tctttgaaag acagggccga gcacgtcatg 1260
ctggtagatt tggccagaaa tgatattaac cgtgtgtgtg agcccaccag taccacgggt 1320
gatcgtttat tgactgtgga gagattttct catgtgatgc atcttgtgtc agaagtcatg 1380
ggaacattga gaccaaacaa gactcgtctc gatgctttca gatccatttt cccagcagga 1440
accgtctccg gtgctccgaa ggtaagagca atgcaactca taggagaatt ggaaggagaa 1500

```

5

```

aagagaggtg tttatgcggg ggccgtagga cactgggtcgt acgatggaaa atcgatggac 1560
acatgtattg ccttaagaac aatgggtcgtc aaggacgggtg tcgcttacct tcaagccgga 1620
ggtggaattg tctacgattc tgacctctat gacgagtaca tcgaaacccat gaacaaaatg 1680
agatccaaca ataacaccat cttggaggct gagaaaatct ggaccgatag gttggccaga 1740
gacgagaatc aaagtgaatc cgaagaaaac gatcaatgaa cggaggacgt aagtaggaat 1800
ttatggtttg gccat                                     1815

```

<210> 79

<211> 1373

<212> ADN

<213> Secuencia artificial

<220>

<223> Región en la posición 5' de PpARG1

<400> 79

ES 2 589 655 T3

```

gatctggcct tccctgaatt tttacgtcca gctatacgat ccgttgtgac tgtatttcct 60
gaaatgaagt ttcaacctaa agttttgggt gtacttgctc cacctaccac ggaaactaat 120
atcgaaacca atgaaaaagt agaactggaa tcgtcaatcg aaattcgcaa ccaagtggaa 180
cccaaagact tgaatctttc taaagtctat tctagtgaac ctaatggcaa cagaagattt 240
gagctgactt ttcaaatgaa tctcaataat gcaatatcaa catcagacaa tcaatgggct 300
ttgtctagtg acacaggatc aattatagta gtgtcttctg caggaagaat aacttccccg 360
atcctagaag tcggggcatc cgtctgtgtc ttaagatcgt acaacgaaca ccttttggca 420
ataacttgtg aaggaacatg cttttcatgg aatttaaaga agcaagaatg tgttctaaac 480
agcatttcat tagcacctat agtcaattca cacatgctag ttaagaaagt tggagatgca 540
aggaactatt ctattgtatc tgccgaagga gacaacaatc cgttacccca gattctagac 600
tgcgaaactt ccaaaaatgg cgctccaatt gtggctctta gcacgaaaga catctactct 660
tattcaaaga aaatgaaatg ctggatccat ttgattgatt cgaaatactt tgaattgttg 720
ggtgctgaca atgcaactgt tgagtgtgtg gaagcgctag aaggtccaat tggaatgcta 780
attcatagat tggtagatga gttcttccat gaaaacactg ccggtaaaaa actcaaactt 840
tacaacaagc gagtactgga ggacctttca aattcacttg aagaactagg tgaaaatgca 900
tctcaattaa gagagaaact tgacaaactc tatggtgatg aggttgaggc ttcttgacct 960
cttctctcta tctgcgtttc tttttttttt tttttttttt tttttttcag ttgagccaga 1020
ccgcgctaaa cgcataccaa ttgccaaatc aggcaattgt gagacagtgg taaaaaagat 1080
gcctgcaaag ttagattcac acagtaagag agatcctact cataaatgag gcgcttattt 1140
agtagctagt gatagccact gcggttctgc tttatgctat ttgttgatg ccttactatc 1200
tttgtttggc tcttttttct tgacgttttc cgttggaggg actccctatt ctgagtcatt 1260
agccgcacag attatcgccc aaaattgaca aaatcttctg gcgaaaaaag tataaaagga 1320
gaaaaaagct cacccttttc cagcgtagaa agtatatatc agtcattgaa gac 1373

```

<210> 80

<211> 1470

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Región en la posición 3' de PpARG1

10

<400> 80

ES 2 589 655 T3

```

gggacttttaa ctcaagtaaa aggatagttg tacaattata tatacgaaga ataaatcatt 60
acaaaaagta ttctgtttctt tgattcttaa caggattcat tttctgggtg tcatcaggta 120
cagcgctgaa tatcttgaag ttaacatcga gctcatcatc gacgttcac acactagcca 180
cgtttccgca acggtagcaa taattaggag cggaccacac agtgacgaca tctttctctt 240
tgaaatggta tctgaagcct tccatgacca attgatgggc tctagcgatg agttgcaagt 300
tattaatgtg gttgaactca cgtgctactc gagcaccgaa taaccagcca gctccacgag 360
gagaaacagc ccaactgtcg acttcatctg ggtcagacca aaccaagtca caaaatcctc 420
cttcatgagg gacctcttgc gctcggctga gaactctgat ttgatctaac atgcgaatat 480
cgggagagag accaccatgg atacataata ttttaccatc aatgatggca ctaaggggta 540
aaaagtcgaa cacctggcaa cagtacttcc agacagtggg ggaaccatat ttattgagac 600
attcctcata aaatccataa acctgagtga tctgtctgga ttcatgattt ccccttacca 660
atgtgatatg ttgaggaaac ttaattttta aaatcatgag taacgtgaac gtctccaacg 720
agaaatagcc tctatccaca tagtctccta ggaagatata gttctgtttt attccattag 780
aggaggatcc gggaaaccca ccactaatct tgaaaagtgc cagtagatcg tgaaattggc 840
cgtgaatata tccgcatact gtcactggac tctgcaactg ctgtatattg gattcctcca 900
tcagcaaatc cttcaccctg tcgcaaagat gcttcatatc attttcaactt aaagccttgc 960

```

```

agcttttgac ttcttcaaac cactgatctg gtcctctttc tggcatgatt aaggtctata 1020
atatttctga gctgagatgt aaaaaaaaaa aataaaaaatg gggagtgaaa aagtgtgtag 1080
cttttaggag tttgggattg ataccccaaa atgatcttta tgagaattaa aaggtagata 1140
cgcttttaat aagaacacct atctatagta ctttgtgggc ttgagtaatt gagatgttca 1200
gcttctgagg tttgccgtta ttctgggata gtagtgcgcg accaaacaac ccgccaggca 1260
aagtgtgttg tgctcgaaga cgattgccag aagagtaagt ccgtcctgcc tcagatgtta 1320
cacactttct tccctagaca gtcgatgcat catcggattt aaacctgaaa ctttgatgcc 1380
atgatacgcc tagtcacgtc gactgagatt ttagataaag cccgatccct ttagtacatt 1440
cctgttatcc atggatggaa tggcctgata 1470

```

<210> 81
 <211> 1854
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> marcador auxotrófico de PpARG1

<400> 81

```

cagttgagcc agaccgcgct aaacgcatac caattgccaa atcaggcaat tgtgagacag 60
tggtaaaaaa gatgcctgca aagttagatt cacacagtaa gagagatcct actcataaat 120
gaggcgctta tttagtagct agtgatagcc actgcgggtc tgctttatgc tatttggtgt 180
atgccttact atctttgttt ggctcctttt tcttgacgtt ttccggttga gggactccct 240
attctgagtc atgagccgca cagattatcg cccaaaattg acaaaatcct ctggcgaaaa 300
aagtataaaa ggagaaaaaa gctcacccct ttccagcgta gaaagtatat atcagtcatt 360
gaagactatt atttaaataa cacaatgtct aaaggaaaag tttgtttggc ctactccggt 420
ggtttgata cctccatcat cctagccttg ttgttgagc agggatacga agtcgttgcc 480
tttttagcca acattggtca agaggaagac tttgaggctg ctagagagaa agctctgaag 540
atcggtgcta ccaagtttat cgtcagtgac gttaggaagg aatttggtga ggaagttttg 600
ttcccagcag tccaagttaa cgctatctac gagaacgtct acttactggg tacctctttg 660
gccagaccag tcattgccaa ggcccaaata gaggttgctg aacaagaagg ttgttttgct 720
gttgcccacg gttgtaccgg aaagggtaac gatcagggtta gatttgagct ttccttttat 780
gctctgaagc ctgacgttgt ctgtatcgcc ccatggagag acccagaatt ctgcgaaaga 840
ttcgctggta gaaatgactt gctgaattac gctgctgaga aggatattcc agttgctcag 900
actaaagcca agccatggtc tactgatgag aacatggctc acatctcctt cgaggctggt 960
attctagaag atccaaacac tactcctcca aaggacatgt ggaagctcac tgttgaccca 1020
gaagatgcac cagacaagcc agagttcttt gacgtccact ttgagaaggg taagccagtt 1080
aaattagttc tcgagaacaa aactgaggtc accgatccgg ttgagatctt tttgactgct 1140
aacgccattg ctagaagaaa cgggtgttgt agaattgaca ttgtcgagaa cagattcatc 1200
ggaatcaagt ccagaggttg ttatgaaact ccaggtttga ctctactgag aaccactcac 1260
atcgacttgg aaggtcttac cgttgaccgt gaagttagat cgatcagaga cacttttggt 1320
accccaacct actctaagtt gttatacaac ggggttgact ttaccccaga aggtgagtag 1380
gtcagaacta tgattcagcc ttctcaaaac accgtcaacg gtgttggttag agccaaggcc 1440
tacaaaggta atgtgtataa cctaggaaga tactctgaaa ccgagaaatt gtacgatgct 1500
accgaatcct ccatggatga gttgaccgga ttccaccctc aagaagctgg aggatattatc 1560
acaacacaag ccatcagaat caagaagtag ggagaaagtg tcagagagaa gggaaagttt 1620
ttgggacttt aactcaagta aaaggatagt tgtacaatta tatatacgaa gaataaatca 1680
ttacaaaaag tattcgtttc tttgattctt aacaggattc attttctggg tgtcatcagg 1740
tacagcgctg aatatcttga agttaacatc gagctcatca tcgacgttca tcacactagc 1800
cacgtttccg caacggttagc aataattagg agcggaccac acagtgacga catc 1854

```

<210> 82

<211> 1250

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Región en la posición 5' de PpADE1

10

<400> 82

```

gagtcggcca agagatgata actgttacta agcttctccg taattagtggt tattttgtaa 60
cttttaccaa taatcgttta tgaatacggg tatttttcga ccttatccag tgccaaatca 120
cgtaacttaa tcatggttta aatactccac ttgaacgatt cattattcag aaaaaagtca 180
ggttggcaga aacacttggg cgctttgaag agtataagag tattaagcat taaacatctg 240
aactttcacc gccccaatat actactctag gaaactcgaa aaattccttt ccatgtgtca 300

```


ES 2 589 655 T3

```

tcgcttccaa cacactttgc tgtatccttc caagtatgtc cattgtgaac actgatctgg 360
acggaatcct acctttaatc gccaaaaggaa aggttagaga catttatgca gtcgatgaga 420
acaacttgct gttcgtcgca actgaccgta tctccgctta cgatgtgatt atgacaaacg 480
gtattcctga taaggggaaag attttgaactc agctctcagt tttctggttt gatttttttg 540
caccctacat aaagaatcat ttggttgctt ctaatgacaa ggaagtcttt gctttactac 600
catcaaaact gtctgaagaa aaatacaaat ctcaattaga gggacgatcc ttgatagtaa 660
aaaagcacag actgatacct ttggaagcca ttgtcagagg ttacatcact ggaagtgcac 720
ggaaagagta caagaactca aaaactgtcc atggagtcaa ggttgaaaac gagaaccttc 780
aagagagcga cgcctttcca actccgattt tcacaccttc aacgaaagct gaacaggggtg 840
aacacgatga aaacatctct attgaacaag ctgctgagat tgtaggtaaa gacattttgtg 900
agaaggtcgc tgtcaaggcg gtcgagttgt attctgctgc aaaaaacctc gcccttttga 960
aggggatcat tattgctgat acgaaattcg aatttgact ggacgaaaac aatgaattgg 1020
tactagtaga tgaagtttta actccagatt cttctagatt ttggaatcaa aagacttacc 1080
aagtgggtaa atcgcaagag agttacgata agcagtttct cagagattgg ttgacggcca 1140
acggattgaa tggcaaagag ggcgtagcca tggatgcaga aattgctatc aagagtaaag 1200
aaaagtatat tgaagcttat gaagcaatta ctggcaagaa atgggcttga 1250

```

<210> 83
 <211> 882
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 3' de PpADE1
 <400> 83

```

atgattagta ccttcctcgc ctttttcaga catctgaaat ttcccttatt cttccaattc 60
catataaaat cctatttagg taattagtaa acaatgatca taaagtgaat tcattcaagt 120
aaccattccg tttatcgttg atttaaaatc aataacgaat gaatgtcggc ctgagtagtc 180
aatttggtgc cttggagctc attggcaggg ggtcttttgg ctcagtatgg aagggtgaaa 240
ggaaaacaga tggaaagtgg ttcgtcagaa aagaggtatc ctacatgaag atgaatgcca 300
aagagatata tcaagtgata gctgagttca gaattcttag tgagttaagc catcccaaca 360
ttgtgaagta ccttcacac gaacatattt ctgagaataa aactgtcaat ttatacatgg 420
aatactgtga tgggtggagat ctctccaagc tgattcgaac acatagaagg aacaaagagt 480
acatttcaga agaaaaata tggagtattt ttacgcaggc tttattagca ttgtatcgtt 540
gtcattatgg aactgatttc acggcttcaa aggagtttga atcgctcaat aaaggttaata 600
gacgaaccca gaatccttcg tgggtagact cgacaagagt tattattcac agggatataa 660
aaccgacaa catctttctg atgaacaatt caaaccttgt caaactggga gatttttgat 720
tagcaaaaat tctggacca gaaaacgatt ttgccaaaac atacgtcggc acgccgtatt 780
acatgtctcc tgaagtgtcg ttggaccaac cctactcacc attatgtgat atatggtctc 840
ttgggtgcgt catgtatgag ctatgtgcac tgaggcctcc tt 882

```

<210> 84
 <211> 909
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 5' de MET16
 <400> 84

ES 2 589 655 T3

```

gggtgggcct ggtaatgttc actcctagga actactagaa aaactgtgct aaacggatta 60
cgtaattatt atacaaattc tctatggctc atgggtacata tgggctgggt caataatgaa 120
tctatgaaga atttgtgccc atggggaccg tttctataaa cgttctcttc tttatgtttt 180
ccacctgctc tttgagttcc ggaaattcgt tgacaatctt ttgtcccaat gtcgattggg 240
cgtattttaa gccagctgt tttcctctga gaaattgatt caacttcctc accacctcca 300
caaaactcac cgtgtatata tcagggtttc taccgtcttc gatataattg actacgtcca 360
cggggatggg aatgttcaaa tctgtgttgt ggagcttttg caagtgtctc acaaccttgt 420
taatgttgtt ggaaagaccc aattgacttt ccgctgtacc ggcgtaatcg tgcacctgaa 480
cacccaaatt gatgagggtt tcgatgagtt gacttagttc attttcaact tgatctaatt 540
ttgtcgcagg tgcactcata cttgtcatgg agaataaaag taagttgata gagagcagac 600
ttcgaggatg ggatgaactt gattaggtta tctttgacaa tgtcttagag gtaggcagag 660
gatgctggaa aaaaaaatt gaaaacgcc aagcttcag ctttgcaagg aaagaagaaa 720
agggagttgc cagcacgaaa tcggcttcct ccgaaagggt cacaattgca gaattgtcac 780
cattcaaatt cctttacctc tcactctgtg tacctcaggc taagaacggg tcacgtgata 840

tttcgacact catcgccaca atatgtacta gcaagaactt ttcagattta gtaatccgtt 900
cgaaacggg 909

```

<210> 85
 <211> 825
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 3' de MET16

<400> 85

```

ctagatttgc acaatatttg aaagctcagc aaaacatatg aatataattt tttttttctc 60
tacactattt atcctgtaag tttctgtttc cccatgtagg atctttttct ctttctctgt 120
ctcccatttt ttttgttccc tgtagtcttg ccttgctga gatgcgagct cgtccgcca 180
tcagtcgtg tgaagggcct agcttttcaa aaagaaaata cctcccgcta aaggaggcgt 240
tgcccccttc atcagtagtg tcgtaaccaa ttttcacaaa caataaaaaa aggacaccaa 300
caacgaaatc aactatttac acacatccag atccgtcccc ctccccatcc aagagttaaa 360
gacaaatatg gctgttaata atccgtctga atttagaaag aagttggctg tagtaggaga 420
tggtgcttgc ggtaaaactt gtctattgat ggtgtttgcc gagggcgagt tccctccatc 480
ttatgttcca actgtttttg agaactatgc caccacagta gaggttgaca acagaatagt 540
acaactcact ctatgggata ctgccggaca ggaagattat gatagactga gacctcttc 600
ctatcccgat gccaatgtgg tcttgatttg ttttgctatt gacattcctg acaccttaga 660
taacgttcaa gagaagtgga ttagtgaggt gttgcatttc tgcctggag tccctatcat 720
tttagttggt tgtaaaactt acttgagaaa cgatccagag gttatccgtg aattacaagc 780
tgttggaaaag caaccagtct ccaccagtga ggtcaggcc gttgc 825

```

<210> 86
 <211> 1796
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> marcador auxotrófico de PpMET16

<400> 86


```

caacttcctc accacctcca caaactcacg cgtgtatata tcagggtttc taccgtcttc 60
gatataattg actacgtcca cggggatggg aatgttcaaa tctgtgttgt ggagcttttg 120
caagtgtctt acaaccttgt taatgttgtt ggaaagaccc aattgacttt ccgctgtacc 180
ggcgtaatcg tgcacctgaa caccctaatg gatgaggggt tcgatgagtt gacttagttc 240
attttcaact tgatctaatt ttgtcgcagg tgcactcata cttgtcatgg agaataaaag 300
taagttgata gagagcagac ttcgaggatg ggttgaaact gattaggtaa tctttgacaa 360
tgtcttagag gtaggcagag gatgctggaa aaaaaaaatt gaaaacgccc aagcttccag 420
ctttgcaagg aaagaagaaa agggagttgc cagcacgaaa tcggcttcct ccgaaagggt 480
cacaattgca gaattgtcac cattcaaagt cctttaccct tcatctgttg tacctcaggc 540
taagaacggg tcacgtgata tttcgacact catcgccaca atatgtacta gcaagaactt 600
ttcagattta gtaatccgtt cgaaacggga aaaaatgttt ttacccttct atcaactgct 660
aatctttcta ggtttatact gccagcagcc cgttccagat accaacatgc cattcactat 720
aggccagtca aaaaccagtt tgaacctctc caaggtccaa gtggaccacc ttaacctttc 780
tcttcagaat ctcatgtccag aagaaatcat acaatggtct atcattacct tcccacacct 840
gtatcaaaact acggcattcg gattgactgg gttgtgtata actgacatgg ttcacaaaat 900
aacagccaaa agaggcaaaa agcatgctat tgacttgatt ttcatagaca ccttacatca 960
ttttccacag acttttagatc tcgttgaacg agtcaaagat aaataccact gcaatgttca 1020
tgtttcaaaa ccacagaatg ccactactga gctcgagttt ggggcgcaat atggcgaaaa 1080
cttatgggaa acagatgata acaagtatga ctacctcgta aaagttgaac cctcacaaacg 1140
tgccatccat gcattagacg tctgcgccgt cttcacagga agaagacggg ctcaagggtg 1200
taaaaggggg gaattgcccg tgattgaaat tgatgaaatt tctcagggtg tcaagattaa 1260
tccgttagca tcctgggggt ttgaacaagt tcaaaactat atccaagcta atagcgttcc 1320
atacaacgaa ttgctggatt tgggatacaa gtcagttgga gattaccatt ccacacaacc 1380
cactaaaaat ggtgaagatg aaagagcagg cagggtggaga ggtaaacaaa agagtgaagt 1440
tggtatccac gaagcttcta gatgtgcaca atatttgaaa gctcagcaaa acatataaat 1500
ataatttttt ttttctctac actattttat ctgtaagttt ctgtttcccc atgtaggatt 1560
tttttctcct tctctgtctc ccattttttt tgttccctgt agtcttgcc tgcctgagat 1620
gcgagctcgt ccgcccattc agtcgtgtga agggcctagc ttttcaaaaa gaaaataacct 1680
cccgttaaag gaggcgttgc cccttctatc agtagtgtcg taaccaattt tcacaaacaa 1740

```

```

taaaaaaagg acaccaacaa cgaaatcaac tatttacaca catccagatc cgtccc 1796

```

<210> 87
 <211> 461
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 5' de PpHIS1
 <400> 87

```

taactggccc tttgacgttt ctgacaatag ttctagagga gtctgtccaaa aactcaactc 60
tgacttgggt gacaccacca cgggatccgg ttcttccgag gacottgatg accttggcta 120
atgtaactgg agtttttagta tccattttta gatgtgtgtt tctgtaggtt ctgggttggg 180
aaaaaatttt agacaccaga agagaggagt gaactgggtt gcgtgggttt agactgtgta 240
aggcactact ctgtcgaagt tttagatagg ggttaccgcg tccgatgcac gggaagcgat 300
tagcccggtt gttgcccggt tggtttttga agggtaattt tcaatatctc tgtttgagtc 360
atcaattttc tattcaaaga ttcaaaaaca aaatctggtc caaggagcgc atttaggatt 420
atggagttgg cgaatcactt gaacgataga ctattatttg c 461

```

<210> 88
 <211> 1841
 <212> ADN
 <213> Secuencia artificial

<220>

<223> Región en la posición 3' de PpHIS1

<400> 88

```

gtgacattct tgtctttgag atcagtaatt gtagagcata gatagaataa tattcaagac 60
caacggcttc tcttcggaag ctccaagtag cttatagtga tgagtaccgg catatattta 120
taggcttaaa atttcgaggg ttcaactatat tctgttagtg ggaagagttc ctttcaactct 180
tggtatctat attgtcagcg tggactgttt ataactgtac caacttagtt tctttcaact 240
ccagggttaag agacataaat gtcctttgat gctgacaata atcagtggaa ttcaaggaag 300
gacaatcccg acctcaatct gtccattaat gaagagttcg aatcgtcctt aaatcaagcg 360
ctagactcaa ttgtcaatga gaaccctttc tttagaccaag aaactataaa tagatcgaat 420
gacaaagttg gaaatgagtc cattagctta catgatattg agcaggcaga ccaaaataaa 480
ccgtcctttg agagcgatat tgatgggtcg ggcgcgttga taagagacga caaattgcc 540
aagaaacaaa gctgggggct gagcaatttt ttttcaagaa gaaatagcat atgtttacca 600
ctacatgaaa atgattcaag tgttggttaag accgaaagat ctattgcagt gggaacaccc 660
catcttcaat actgcttcaa tggaaatctcc aatgccaaagt acaatgcatt tacctttttc 720
ccagtcatacc tatacagca attcaaattt ttttcaatt tatactttac ttttagtggt 780
ctctctcaag cgataccgca acttcgcatt ggatatcttt ctctgtatgt cgtccactt 840
ttgtttgtac tcatagtga catgtcaaaa gaggcgatgg atgatattca acgcccgaaga 900
agggatagag aacagaacaa tgaaccatat gaggttctgt ccagcccatc accagttttg 960
tccaaaaact taaaatgtgg tcaattgggt cgattgcata agggaaatgag agtgcccgc 1020
gatatgggtc ttgtccagtc aagcgaatcc accggagagt catttatcaa gacagatcag 1080
ctggatgggt agactgattg gaagcttcgg attgtttctc cagttacaca atcgttacca 1140
atgactgaac ttcaaaatgt cgccatcact gcaagcgcac cctcaaaatc aattcactcc 1200
tttcttgga gattgacct caatgggcaa tcatatggtc ttacgataga caacacaatg 1260
tggtgtaata ctgtattagc ttctgggtca gcaattgggt gtataattta cacaggtaaa 1320
gatactcgac aatcgatgaa cacaactcag cccaaactga aaacgggctt gttagaactg 1380
gaaatcaata gtttgtccaa gatcttatgt gtttgtgtgt ttgcattatc tgtcatctta 1440
gtgctattcc aaggaatagc tgatgattgg tacgtcgata tcatgcggtt tctcattcta 1500
ttctccacta ttatcccagt gtctctgaga gttaaccttg atcttggaaa gtcagtccat 1560
gctcatcaaa tagaaactga tagctcaata cctgaaaccg ttgttagaac tagtacaata 1620
ccggaagacc tgggaagaat tgaataccta ttaagtgaac aaactggaac tcttactcaa 1680
aatgatatgg aaatgaaaaa actacaccta ggaacagctc cttatgctgg tgataccatg 1740
gatattattt ctgatcatgt taaaggtctt aataacgcta aaacatcgag gaaagatctt 1800
ggtatgagaa taagagattt gggtacaact ctggccatct g 1841

```

5

<210> 89

<211> 1729

<212> ADN

10 <213> Secuencia artificial

<220>

<223> marcador auxotrófico de PpHIS1

15 <400> 89

```

caagttgcgt cccgtatcacg taacgtctca cgatgatcaa agataatact taatcttcat 60
ggtctactga ataactcatt taaacaattg actaattgta cattatattg aacttatgca 120
tcctattaac gtaatcttct ggcttctctc tcagactcca tcagacacag aatatcggtc 180
tctctaactg gtcctttgac gtttctgaca atagttctag aggagtcgtc caaaaaactca 240
actctgactt ggggtgacacc accacgggat ccggttcttc cgaggacctt gatgaccttg 300
gctaattgtaa ctggagtttt agtatccatt ttaagatgtg tgtttctgta ggttctgggt 360
tggaaaaaaa ttttagacac cagaagagag gagtgaactg gtttgcgtgg gtttagactg 420
tgtaaggcac tactctgtcg aagttttaga taggggttac ccgctccgat gcatgggaag 480
cgattagccc ggctgttgcc cgtttggttt ttgaagggtta attttcaata tctctgtttg 540
agtcattcaat ttcatattca aagattcaaa aacaaaatct ggtccaagga ggcatttag 600
gattatggag ttggcgaatc acttgaacga tagactatta tttgctgttc cttaaagaggg 660
cagattgtat gagaaatgcg ttgaattact taggggatca gatattcagt ttcgaagatc 720
cagtagattg gatatagctt tgtgcactaa cctgcccctg gcattgggtt tccttccagc 780
tgctgacatt ccacggtttg taggagaggg taaatgtgat ttgggtataa ctggtattga 840
ccaggttcag gaaagtgcg tagatgtcat acccttatta gacttgaatt tcggttaagt 900
caagttgcag attcaagttc ccgagaatgg tgacttgaaa gaacctaaac agctaattgg 960
taaagaaatt gtttctctct ttactagctt aaccaccagg tactttgaac aactggaagg 1020
agttaagcct ggtgagccac taaagacaaa aatcaaatat gttggagggg ctggttgaggc 1080
ctcttgtgcc ctaggagttg ccgatgctat tgtggatctt gttgagagtg gagaaacat 1140
gaaagcggca gggctgatcg atattgaaac tgttctttct acttccgctt acctgatctc 1200
ttcgaagcat cctcaacacc cagaactgat ggatactatc aaggagagaa ttgaaggtgt 1260
actgactgct cagaagtatg tcttgtgtaa ttacaacgca cctagaggta accttctca 1320
gctgctaaaa ctgactccag gcaagagagc tgctaccgtt tctccattag atgaagaaga 1380
ttgggtggga gtgtcctcga tggtagagaa gaaagatgtt ggaagaatca tggacgaatt 1440
aaagaaacaa ggtgccagtg acattcttct ctttgagatc agtaattgta gagcatagat 1500
agaataatat tcaagaccaa cggcttctct tcggaagctc caagtagctt atagtgtga 1560
gtaccggcat atatttatag gcttaaaatt tcgaggggtc actatattcg tttagtggga 1620
agagttcctt tcaactctgt tatctatatt gtcagcgtgg actgtttata actgtaccaa 1680
cttagtttct ttcaactcca ggttaagaga cataaatgtc ctttgatgc 1729

```

<210> 90
 <211> 1231
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 5' de PpPRO1

<400> 90

```

gaagggccat cgaattgtca tcgtctcctc aggtgccatc gctgtgggca tgaagagagt 60
caacatgaag cggaaaccaa aaaagttaca gcaagtgcag gcattggctg ctataggaca 120
aggccgtttg ataggacttt gggacgacct tttccgtcag ttgaatcagc ctattgcgca 180
gattttactg actagaacgg atttggtcga ttacaccagc ttttaagaacg ctgaaaatac 240
attggaacag cttattaaaa tgggtattat tcctattgtc aatgagaatg acaccctatc 300
cattcaagaa atcaaatttg gtgacaatga caccttatcc gccataacag ctggtatgtg 360
tcattgcagac tacctgtttt tgggtactga tgtggactgt ctttacacgg ataaccctcg 420
tacgaatccg gacgctgagc caatcgtgtt agttagaaat atgaggaatc taaacgtcaa 480
taccgaaagt ggaggttccg ccgtaggaac aggaggaatg acaactaaat tgatcgagc 540
tgatttgggt gtatctgcag gtgttacaac gattatttgc aaaagtgaac atcccgagca 600
gattttggac attgtagagt acagtatccg tgctgataga gtcgaaaatg aggctaaata 660
tctggctcgc aacgaagagg aaactgtgga acaatttcaa gagatcaatc ggtcagaact 720
gaggaggttg aacaagctgg acattccttt gcatacacgt ttcgttggcc acagttttaa 780
tgctgttaat aacaaagagt tttggttact ccatggacta aaggccaacg gagccattat 840
cattgatcca ggttgttata aggctatcac tagaaaaaac aaagctggta ttcttccagc 900
tggaattatt tccgtagagg gtaatttcca tgaatacagag tgtgttgatg ttaaggtagg 960
actaagagat ccagatgacc cacattcact agacccaat gaagaacttt acgtcgttgg 1020
ccgtgcccgt tgtaattacc ccagcaatca aatcaacaaa attaaaggtc tacaaagctc 1080
gcagatcgag caggttctag gttacgctga cggtagatg gttgttcaca gggacaactt 1140
ggctttccca gtatttgccg atccagaact gttggatgtt gttgagagta ccctgtctga 1200
acaggagaga gaatccaaac caaataaata g
1231

```

<210> 91
 <211> 1425
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Región en la posición 3' de PpPRO1
 <400> 91

```

aatttcacat atgctgcttg attatgtaat tataccttgc gttcgatggc atcgatttcc 60
tcttctgtca atcgcgcatc gcattaaaag tatacttttt tttttttcct atagtactat 120
tcgccttatt ataaactttg ctagtatgag ttctaccccc aagaaagagc ctgatttgac 180
tcctaagaag agtcagcctc caaagaatag tctcgggtggg ggtaaaggct ttagtgagga 240
gggtttctcc caaggggact tcagcgctaa gcataacta aatcgctcgc ctaacaccga 300
aggctcttct gtggcttcga acgtcatcag ttcgtcatca ttgcaaaggc taccatcctc 360
tggtatctga agcgttgctg tgggaagtgt gttgggatct tcgccattaa ctctttctgg 420
agggttccac gggcttgatc caaccaagaa taaaatagac gttccaaagt cgaaacagtc 480
aaggagacaa agtgttcttt ctgacatgat ttccacttct catgcagcta gaaatgatca 540
ctcagagcag cagttacaaa ctggacaaca atcagaacaa aaagaagaag atggtagtcg 600
atcttctttt tctgtttctt ccccgcgaag agatatccgg caccagatg tactgaaaac 660
tgtcgagaaa catcttgcca atgacagcga gatcgactca tctttacaac ttcaagggtg 720
agatgtcact agaggcattt atcaatgggt aactggagaa agtagtcaaa aagataaccc 780
gcctttgaaa cgagcaaata gttttaatga tttttcttct gtgcatggtg acgaggtagg 840
caaggcagat gctgaccacg atcgtgaaag cgtattcgac gaggatgata tctccattga 900
tgatatcaaa gttccgggag ggatgcgtcg aagtttttta ttacaaaagc atagagacca 960
acaactttct ggactgaata aaacggctca ccaacccaaa caacttacta aacctaat 1020
cttcacgaac aactttatag agtttttggc attgtatggg cattttgcag gtgaagattt 1080
ggaggaagac gaagatgaag atttagacag tggttccgaa tcagtcgcag tcagtgatag 1140
tgagggagaa ttcagtgagg ctgacaacaa tttgttgtat gatgaagagt ctctcctatt 1200
agcacctagt acctccaact atgcgagatc aagaatagga agtattcgta ctctactta 1260
tggatctttc agttcaaatg ttggttcttc gtctattcat cagcagttaa tgaaaagtc 1320
aatcccgaag ctgaagaaac gtggacagca caagcataaa acacaatcaa aaatacgctc 1380
gaagaagcaa actaccaccg taaaagcagt gttgctgcta ttaaa
1425

```

<210> 92
 <211> 501

<212> ADN

<213> Secuencia artificial

<220>

5 <223> Codifica ADN de hEPO Truncada (Optimizado por codón)

<400> 92

```
gctccacca gattgatttg tgactccaga gttttggaga gatacttggt ggaggctaaa 60
gaggctgaga acatcactac tgggtgtgct gaacactggt ccttgaacga gaacatcaca 120
gttccagaca ctaagggttaa cttctacgct tggaagagaa tggaagttgg acaacaggct 180
gttgaagttt ggcaaggatt ggctttggtg tccgaggctg ttttgagagg tcaagctttg 240
ttggttaact cctcccaacc atgggaacca ttgcaattgc acgttgacaa ggctgtttct 300
ggattgagat ccttgactac tttgttgaga gctttgggtg ctcagaaaga ggctattttct 360
ccaccagatg ctgcttcagc tgctccattg agaactatca ctgctgacac tttcagaaaag 420
ttgttcagag tttactccaa cttcttgaga ggaaagttga agttgtacac tgggtgaagct 480
tgtagaactg gtgactagta a                                     501
```

<210> 93

<211> 165

<212> PRT

<213> Secuencia artificial

<220>

<223> hEPO Truncada

<400> 93

```
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
 1          5          10          15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
 20          25          30
Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
 35          40          45
Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
 50          55          60
Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
 65          70          75          80
Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
 85          90          95
Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
100         105         110
Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
115         120         125
Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
130         135         140
Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145         150         155         160
Cys Arg Thr Gly Asp
165
```

<210> 94

<211> 78

<212> ADN

<213> Secuencia artificial

<220>

<223> codifica péptido señal de lisozima de pollo (CLSP)

<400> 94

5

```
atgctgggta agaacgaccc aatgtgtctt gttttggctt tggtgggatt gactgctttg 60
ttgggtatct gtcaaggt                                     78
```

<210> 95

<211> 26

10

<212> PRT

<213> Secuencia artificial

<220>

<223> péptido señal de lisozima de pollo (CLSP)

15

<400> 95

```
Met Leu Gly Lys Asn Asp Pro Met Cys Leu Val Leu Val Leu Leu Gly
 1      5      10      15
Leu Thr Ala Leu Leu Gly Ile Cys Gln Gly
      20      25
```

20

<210> 96

<211> 1892

<212> ADN

<213> Secuencia artificial

25

<220>

<223> Codifica PpAde2

<400> 96

30

```
atggattctc aggtaatagg tattctagga ggaggccagc taggccgaat gattgttgag 60
gccgctagca ggctcaatat caagaccgtg attcttgatg atggtttttc acctgctaag 120
cacattaatg ctgcgcaaga ccacatcgac ggatcattca aagatgagga ggctatcgcc 180
aagtttagctg ccaaattgtga tgttctcact gtagagattg agcatgtcaa cacagatgct 240
ctaaagagag ttcaagacag aactggaatc aagatatatc ctttaccaga gacaatcgaa 300
ctaatacagg ataagtactt gcaaaaggaa catttgatca agcacaacat ttcggtgaca 360
aagtctcagg gtatagaatc taatgaaaag gcgctgcttt tgtttgagga agagaatgga 420
```



```

tttccatatac tgttgaagtc cgggactatg gcttatgatg gaagaggcaa tttttagtagt 480
gagtctaaag aggacatcag taaggcatta gaattcttga aagatcgtcc attgtatgcc 540
gagaagtttg ctccctttgt taaagaatta ggggtaattg ttgtgagatc actggaaggc 600
gaagtattct cctacccaac cgtagaaact gtgcacaagg acaatatctg tcatattgtg 660
tatgctccgg ccagagttaa tgacaccatc caaaagaaag ctcaaattt agctgaaaac 720
actgtgaaga ctttcccagg cgctggaatc ttgggagttg agatgttct attgtctgat 780
ggagaacttc ttgtaaatga gattgctcca aggcccccaca attctgggtca ctatacaatc 840
gatgcatgtg taacatctca gttcgaagca catgtaagag ccataactgg tctgccaatg 900
ccactagatt tcaccaaact atctacttcc aacaccaacg ctattatgct caatgttttg 960
gggtgctgaaa aatctcacgg ggaattagag tttttagtaga gagccttaga aacacccggg 1020
gcttctgtat atctgtacgg aaagaccacc cgattggctc gtaagatggg tcatatcaac 1080
ataataggat cttccatgtt ggaagcagaa caaaagttag agtacattct agaagaatca 1140
accacattac catccagtac tgtatcagct gacactaaac cgttgggttg agttatcatg 1200
ggttcagact ctgatctacc tgtgatttgc aaagggttgc atatttttaa acagtttggt 1260
gttccattcg aagttactat tgtctctgct catagaacac cacagagaat gaccagatat 1320
gcctttgaag ccgctagtag aggtatcaag gctatcattg cagggtgctg tgggtgctgct 1380
catcttccag gaatggttgc tgccatgact ccgttgccag tcattgggtg tcctgtcaag 1440
ggctctacgt tggatggtgt agactcgcta cactcgattg tccaaatgcc tagaggtgtt 1500
cctgtggcta cggttgctat caacaacgcc accaatgccg ctctgttggc catcaggatt 1560
ttaggtacaa ttgaccacaa atggcaaaaag gaaatgtcca agtatatgaa tgcaatggag 1620
accgaagtgt tggggaaggc atccaacttg gaatctgaag ggtatgaatc ctatttgaag 1680
aatcgtcttt gaatttagta ttgtttttta atagatgtat atataatagt acacgtaact 1740
tatctattcc attcataatt ttatttttaa ggttcggtag aaatttgtcc tccaaaaagt 1800
tggtagagc ctggcagttt tgataggcat tattatagat tgggtaatat ttaccctgca 1860
cctggaggaa ctttgcaaag agcctcatgt gc 1892

```

<210> 97

<211> 563

5 <212> PRT

<213> *Pichia pastoris*

<400> 97

Met	Asp	Ser	Gln	Val	Ile	Gly	Ile	Leu	Gly	Gly	Gly	Gln	Leu	Gly	Arg
1				5					10					15	
Met	Ile	Val	Glu	Ala	Ala	Ser	Arg	Leu	Asn	Ile	Lys	Thr	Val	Ile	Leu
			20					25					30		
Asp	Asp	Gly	Phe	Ser	Pro	Ala	Lys	His	Ile	Asn	Ala	Ala	Gln	Asp	His
		35					40					45			
Ile	Asp	Gly	Ser	Phe	Lys	Asp	Glu	Glu	Ala	Ile	Ala	Lys	Leu	Ala	Ala
	50					55					60				
Lys	Cys	Asp	Val	Leu	Thr	Val	Glu	Ile	Glu	His	Val	Asn	Thr	Asp	Ala
65					70					75				80	
Leu	Lys	Arg	Val	Gln	Asp	Arg	Thr	Gly	Ile	Lys	Ile	Tyr	Pro	Leu	Pro
				85					90					95	
Glu	Thr	Ile	Glu	Leu	Ile	Lys	Asp	Lys	Tyr	Leu	Gln	Lys	Glu	His	Leu
			100					105					110		
Ile	Lys	His	Asn	Ile	Ser	Val	Thr	Lys	Ser	Gln	Gly	Ile	Glu	Ser	Asn
		115					120					125			
Glu	Lys	Ala	Leu	Leu	Leu	Phe	Gly	Glu	Glu	Asn	Gly	Phe	Pro	Tyr	Leu
	130					135					140				
Leu	Lys	Ser	Arg	Thr	Met	Ala	Tyr	Asp	Gly	Arg	Gly	Asn	Phe	Val	Val
145					150					155				160	
Glu	Ser	Lys	Glu	Asp	Ile	Ser	Lys	Ala	Leu	Glu	Phe	Leu	Lys	Asp	Arg
				165					170					175	
Pro	Leu	Tyr	Ala	Glu	Lys	Phe	Ala	Pro	Phe	Val	Lys	Glu	Leu	Ala	Val
			180					185					190		
Met	Val	Val	Arg	Ser	Leu	Glu	Gly	Glu	Val	Phe	Ser	Tyr	Pro	Thr	Val
		195					200					205			
Glu	Thr	Val	His	Lys	Asp	Asn	Ile	Cys	His	Ile	Val	Tyr	Ala	Pro	Ala
	210					215					220				
Arg	Val	Asn	Asp	Thr	Ile	Gln	Lys	Lys	Ala	Gln	Ile	Leu	Ala	Glu	Asn
225					230					235				240	
Thr	Val	Lys	Thr	Phe	Pro	Gly	Ala	Gly	Ile	Phe	Gly	Val	Glu	Met	Phe
				245					250					255	
Leu	Leu	Ser	Asp	Gly	Glu	Leu	Leu	Val	Asn	Glu	Ile	Ala	Pro	Arg	Pro

[illegible]

<210> 98
<211> 1302
<212> ADN
<213> Secuencia artificial

<220>
<223> Pp TRP2: 5' y ORF
<400> 98

ES 2 589 655 T3

```

actgggcctt tagaggggtgc tgaagttgac cccttggtgc ttctggaaaa agaactgaag 60
ggcaccagac aagcgcaact tcctgggtatt cctcgtctaa gtggtgggtgc cataggatac 120
atctcgtacg attgtattaa gtactttgaa ccaaaaaactg aaagaaaact gaaagatgtt 180
ttgcaacttc cggaagcagc tttgatgttg ttcgacacga tcgtggcctt tgacaatgtt 240
tatcaaagat tccaggtaat tggaaacgtt tctctatccg ttgatgactc ggacgaagct 300
attccttgaga aatattataa gacaagagaa gaagtggaaa agatcagtaa agtgggtattt 360
gacaataaaa ctgttcctta ctatgaacag aaagatatta ttcaaggcca aacgttcacc 420
tctaattattg gtcaggaagg gtatgaaaac catgttcgca agctgaaaga acatattctg 480
aaaggagaca tcttccaagc tgttcctctc caaagggtag ccaggccgac ctcattgcac 540
cctttcaaca tctatcgtca tttgagaact gtcaatcctt ctccatacat gttctatatt 600
gactatctag acttccaagt tgttgggtgt tcacctgaat tactagttaa atccgacaac 660
aacaacaaaa tcatcacaca tcctattgct ggaactcttc ccagaggtaa aactatcgaa 720
gaggacgaca attatgctaa gcaattgaag tcgtctttga aagacagggc cgagcacgtc 780
atgctggttag atttggccag aaatgatatt aaccgtgtgt gtgagcccac cagtaccacg 840
gttgatcgtt tattgactgt ggagagattt tctcatgtga tgcactctgt gtcagaagtc 900

```

```

agtggaaacat tgagacccaaa caagactcgc ttcgatgctt tcagatccat tttcccagca 960
ggtaccgtct ccggtgctcc gaaggtaaga gcaatgcaac tcataggaga attggaagga 1020
gaaaagagag gtgtttatgc gggggccgta ggacactggc cgtacgatgg aaaatcgatg 1080
gacacatgta ttgccttaag aacaatggct gtcaaggacg gtgtcgctta ccttcaagcc 1140
ggaggtggaa ttgtctacga ttctgacccc tatgacgagt acatcgaaac catgaacaaa 1200
atgagatcca acaataacac catcttggag gctgagaaaa tctggaccga taggttggcc 1260
agagacgaga atcaaagtga atccgaagaa aacgatcaat ga 1302

```

<210> 99
 <211> 1085
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Pp TRP2: 3'
 <400> 99

```

acggaggacg taagtaggaa tttatgtaat catgccaaata catctttaga tttcttcttc 60
ttcttttttaa cgaaagacct ccagttttgc actctcgact ctctagtatc ttccattttc 120
tgttgctgca acctcttgcc ttctgtttcc ttcaattgtt cttctttctt ctgttgcaact 180
tggccttctt cctccatctt tcgttttttt tcaagccttt tcagcagttc ttcttccaag 240
agcagttctt tgattttctc tctccaatcc accaaaaaac tggatgaatt caaccgggca 300
tcatcaatgt tccactttct ttctcttatc aataatctac gtgcttcggc atacgaggaa 360
tccagttgct cctaatacga gtcattccaca aggttagcat gggccttttt caggggtgtca 420
aaagcatctg gagctcgttt attcggagtc ttgtctggat ggatcagcaa agactttttg 480
cggaagatct ttcttatatc ttccggagaa caacctgggt tcaaataccaa gatggcatag 540
ctgtccaatt tgaaagtggg aagaatcctg ccaatttctt tctctcgtgt cagctcgttc 600
tcctcctttt gcaacaggtc cacttcatct ggcatttttc tttatgttaa ctttaattat 660
tattaattat aaagttgatt atcgttatca aaataatcat attcgagaaa taatccgtcc 720
atgcaatata taaataagaa ttcataataa tgtaatgata acagtacctc tgatgacctt 780
tgatgaaccg caattttctt tccaatgaca agacatccct ataatacaat tatacagttt 840
atatatcaca aataatcacc tttttataag aaaaccgtcc tctccgtaac agaacttatt 900
atccgcacgt tatgggttaac aactactaa taccgatata gtgtatgaag tcgctacgag 960
atagccatcc aggaaactta ccaattcatc agcactttca tgatccgatt gttggcttta 1020
ttcttttgca gacagatact tgccaatgaa ataactgatc ccacagatga gaatccggtg 1080
ctcgt 1085

```

<210> 100
 <211> 1007
 <212> ADN
 <213> Secuencia artificial

<220>

<223> región en la posición 5' de Pp ADE2

5 <400> 100

```

cttaaaatca tctgcctcac cccaccgacc aatgggaatt ctagaaacaa tttcattgct 60
cttcttctcg ttaccataag aatcgggtgt catgtttgac ttaacgaacc ctggaacaag 120
ggaattcacg gtaatacctt ttggagcaag ttcaaccgat agagccttca ttaatgagtt 180
gattgcacct ttggtggtcg catataccga ttgattcggg taggtcactt cgaaactgta 240
cagggaggca gtaaagatga tccctaccctt aatctggttc ttaataaagt gtttagtgac 300
tagctgtgtc aatctaaatg gaaaatcgac atttaccttt tggatagccg cgtaatcttt 360
ctccgtaaaa cttgtaaact cagatttaat ggcaatggca gcgttggtga ttaaaatgtc 420
aatctttcca gtggaactct tctccaccgc aggactcgtt acggtctctt ccagctttgc 480
aagatcggca tccactagat ccaactcaat tgtatgtatg gaggcaccat cggcatttga 540
cattctcacc tcttcaatga aagccgttgg gtctgtagaa ggtctatgga taagaataag 600
ttctgcacct gcttcataaa gtctcgaac tattccttgg cctaaccgc tggtagcacc 660
ggtgatcaag gcgaccttac cattcaaaga aaacaaatca gcggacatta gcgacttgaa 720
tagggaatgg gttagacaaa tgaaagccga cgagccagca ctttatagta agtgcagggtg 780
agtcaataag aataaatgta tggcttgctg tccctatcgc gtaagaagct tactaagatc 840
gcctaaattg aaaagttgaa caaatcagtt ctatctggcc tccatcagca tttcgtttctc 900
ctctgatcat ctttgccaat cgctagcatg ccctcagcgt gcaaggaaaa gcacgcttct 960
ttcttatcga cgtattttca actatggcag agccagggtta gcaagtc 1007

```

<210> 101

<211> 1676

<212> ADN

<213> Secuencia artificial

10

<220>

<223> región en la posición 3' de Pp ADE2 3'

15

<400> 101

```

atttagtatt gttttttaat agatgtatat ataatagtac acgtaactta tctattccat 60
tcataatttt attttaaagg ttccgtagaa atttgcctc caaaaagttg gttagagcct 120
ggcagttttg ataggcatta ttatagattg ggtaatat tt accctgcacc tggaggaact 180
ttgcaaagag cctcatgtgc tctaaaagga tgtcagaatt ccaacatttc aaaattatat 240
ctgcatgcgt ctgtaatact ggaactgtta tttttctggt caggatttca ccgctcttgt 300
cgtcatgttt ctgctcgtct gaaagtaaac tgactttcct cttccataa acacaaaaat 360
cgattgcaac ttggttattc ttgagattga aatttgctgt gtcttcagt cttagctgaa 420
tatcaacaaa cttacttagt actaataacg aagcactatg gtaagtggca taacatagt 480
gtattgaagc gaacagtggg tattgaaccc aagcattggc aacatctggc tctgttgata 540
ctgatccgga tcgtttggca ccaattcctg aaacggcgta gtgccacca ggtttcgatt 600
tgagaacagg ttcatcatca gagtcaacca cccaatgtc aatggcaggc tccaacgaag 660
taggtccaac aacaacagga agtatttgac cttgaagatc tgttccttta tgatccacca 720
caccttgccc caattccaat aactttacca gtcccgatgc agacatgata actggtacta 780
atgatctcca ttgattttcg tcggcactac gtaaagcctc caaaaatgaa ttcagaatat 840
cttctgaaac tagattctgc ttctgtgatt caagcattgc tttatgtaga catctcttga 900
ataaaagcaa ttctccacat attggtgtgt gtaagataga tctggaaaga tgtatctgga 960
atagtcagat caacgttgtg caattgatta gcattacctt actgtgaaca tctctatcta 1020
caacaacaga ctcaattcga tagacgttcc gggaaagttt ttcaagcgca ttcagtttgc 1080
tgttgaacaa agtgactttg ctttccaatg tgcaaatacc cctgtatata aagtccatca 1140
catcactcaa gaccttggtg gaaaagaatg aaacagctgg agcataattt tcgaatgaat 1200
taggtaaggt cacttcatcc ttatctgttg taatgctata atcaatagcg gaactaacat 1260
cttcccattg aacaggtttc ttgatctctg aatctgaatc tttatttgaa aaagaattga 1320
aaaaagactc atcactcatt gggaattcaa ggtcattagg gtattccatt gttagttctg 1380
gtctaggttt aaagggatca cttcgttaa gacgatggaa aatagctaata ctgtacaata 1440
accagatact tctaacgaag ctctctctat ccatcagttg acgtgttgag gatatctgaa 1500
ctagctcttt ccactgcgaa tcaggcatgc tcgtatagct ggcaagcatg ttattcagct 1560
ttaccaagtt agaagccctt tggaaacat ctatagattc ccgaaaaaac ttatacccac 1620
tgagggtttc actgagcata gtcagtgaca tcaaagagca tttcaaatac atctca 1676

```

<210> 102
 <211> 582
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> ORF de NATR

<400> 102

```

atgggtacca ctcttgacga cacggcttac cggtagcgca ccagtgtccc gggggacgcc 60
gaggccatcg aggcactgga tgggtccttc accaccgaca ccgtcttccg cgtcaccgcc 120
accggggacg gcttcaccct gcgggaggtg ccggtggacc cgcacctgac caaggtgttc 180
cccagcgacg aatcggacga cgaatcggac gacggggagg acggcgaccc ggactcccgg 240
acgttcgtcg cgtacgggga cgacggcgac ctggcgggct tcgtggtcgt ctcgtaactcc 300
ggctggaacc gccggctgac cgtcggggac atcgaggtcg ccccgagca ccgggggcac 360
ggggtcgggc gcgcgttgat ggggctcgcg acggagttcg cccgcgagcg gggcgccggg 420
cacctctggc tggaggtcac caacgtcaac gcaccggcga tccacgcgta ccggcggtatg 480
gggttcaccc tctgcggcct ggacaccgcc ctgtacgacg gcaccgcctc ggacggcgag 540
caggcgctct acatgagcat gccctgcccc taatcagtag tg 582

```

<210> 103
 <211> 1029
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> ORF de HygR

ES 2 589 655 T3

<400> 103

```

atgggtaaaa agcctgaact caccgagacg tctgtcgaga agtttctgat cgaaaagttc 60
gacagcgtct ccgacctgat gcagctctcg gagggcgaaag aatctcgtgc tttcagcttc 120
gatgtaggag ggcgtggata tgtcctgcgg gtaaatagct gcgccgatgg tttctacaaa 180
gatcggttatg tttatcggca ctttgcacog gccgcgctcc cgattccgga agtgcttgac 240
attggggaat tcagcgagag cctgacctat tgcactctcc gccgtgcaca ggggtgcacg 300
ttgcaagacc tgcctgaaac cgaactgccc gctgttctgc agccggctgc ggaggccatg 360
gatgcatcg ctgcggccga tcttagccag acgagcgggt tcggcccatt cggaccgcaa 420
ggaatcggtc aatacactac atggcgtgat ttcataatgc cgattgctga tccccatgtg 480
tatcactggc aaactgtgat ggacgacacc gtcagtgcgt ccgtcgcgca ggctctcgat 540
gagctgatgc tttgggcccga ggactgccc gaagtccggc acctcgtgca cgcggatttc 600
ggctccaaca atgtcctgac ggacaatggc cgcataacag cggtcattga ctggagcgag 660
gcgatgttcg gggattccca atacgaggtc gccaacatct tcttctggag gccgtggttg 720
gcttgtatgg agcagcagac gcgctacttc gagcggaggg atccggagct tgcaggatcg 780
ccgcggctcc gggcgatatat gctccgcatt ggtcttgacc aactctatca gagcttggtt 840
gacggcaatt tcgatgatgc agcttgggcg cagggtcgat gcgacgcaat cgtccgatcc 900
ggagccggga ctgtcgggcg tacacaaatc gcccgagaa gcgcggccgt ctggaccgat 960
ggctgtgtag aagtactcgc cgatagtgga aaccgacgcc ccagcactcg tccgagggca 1020
aaggaatag                                     1029

```

5 <210> 104
 <211> 2957
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> región de PpPEP4
 <400> 104

```

atttgagtc cctgcttttag ggctggaaga tatttggtta ctagatttta gtacaaactc 60
ttgctttgtc aatgacatta aaataggcaa gaatcgcaaa actcaaatat ttcattggaga 120
tgagatatgc ttgttcaaag atgcccagaa aaaagagcaa ctctgtttata gggttcatat 180
tgatgatgga acaggccttt tccaggaggg tgaaagaacc caagccaatt ctgatgacat 240
tctggatatt gatgagggtt atgaaaagtt aagagaacta ttgacaagag cctcaaggaa 300
acggcatatc acccctgcat tggaaactcc tgataaacgt gtaaaaagag cttattttgaa 360
cagtattact gataactctt gatggacctt aaagatgtat aatagtagac agaattcata 420
atgggtgagat taggtaatcg tccggaatag gaatagtggg ttggggcgat taatcgacc 480
tgccttatat ggtaagtacc ttgaccgata aggtggcaac tatttagaac aaagcaagcc 540
acctttcttt atctgtaact ctgtcgagc aagcatcttt actagagAAC atctaaacca 600
ttttacattc tagagttcca tttctcaatt actgataatc aatttaaaga tgatatttga 660
cgggtactacg atgtcaattg ccattgggtt gctctctact ctaggtattg gtgctgaagc 720
caaagttcat tctgctaaga tacacaagca tccagtctca gaaactttaa agaggccaa 780
ttttgggcag tatgtctctg ctctggaaca taaatatgtt tctctgttca acgaacaaaa 840
tgctttgtcc aagtgcgaatt ttatgtctca gcaagatggg tttgccgttg aagcttcgca 900
tgatgctcca cttacaaact atcttaacgc tcagtatttt actgaggtat cattaggtac 960
ccctccacaa tctttcaagg tgattcttga cacaggatcc tccaatttat gggttcctag 1020
caaagattgt ggatcattag cttgcttctt gcatgctaag tatgaccatg atgagtcttc 1080
tacttataag aagaatggta gtagctttga aattagggtat ggatccggtt ccatggaagg 1140
gtatgtttct caggatgtgt tgcaaattgg ggatttgacc attcccaaag ttgattttgc 1200
tgaggccaca tccgagccgg ggttgccctt cgcttttgcc aaatttgacg gaattttggg 1260
gcttgcttat gattcaatat cagtaaataa gattgttctt ccaatttaca aggctttgga 1320
attagatctc cttgacgaac caaaatttgc cttctacttg ggggatacgg acaaagatga 1380
atccgatggc ggtttggcca catttggtgg tgtggacaaa tctaagtatg aaggaaagat 1440
cacctgggtg cctgtcagaa gaaaggctta ctgggaggtc tcttttgatg gtgtagggtt 1500
gggatccgaa tatgctgaat tgcaaaaaac tgggtgcagc atcgacactg gaacctcatt 1560
gattgctttg cccagtggcc tagctgaaat tctcaatgca gaaattgggtg ctaccaaggg 1620
ttgggtctgg caatacgtg tggactgtga cactagagac tctttgccag acttaacttt 1680
aaccttcgcc ggttacaact ttaccattac tccatatgac tatacttttg aggtttcttg 1740
gtcatgtatt agtgctttca ccccatgga ctttcttgaa ccaatagggtc ctttggaact 1800
cattgggtgac tctgtcttga gaaaatatta ctcagtttat gacctaggca aagatgcagt 1860
agggttagcc aagtctatct aggcagaagt aaaagttgct cagctgaact tatttggtta 1920
cttatcaggt agtgaagatg tagagaatat atgtttagggt attttttttt agttttctc 1980
ctataactca tcttcagtac gtgattgctt gtcagctacc ttgacagggg cgcataagtg 2040
atatcgtgta ctgctcaatc aagatttgcc tgcctcattg ataagggtat aagagaccca 2100
cctgctctc tttaaaattc tctcttaact gttgtgaaaa tcatcttcga agcaaatcgc 2160

agtttaaatc tatgcggttg gtaactaaag gtatgtcatg gtggtatata gtttttctatt 2220
ttacctttta ctaatcagtt ttacagaaga ggaacgtctt tctcaagatc gaaataggac 2280
taaatactgg agacgatggg gtccttattt ggggtgaaagg cagtgggcta cagtaaggga 2340
agactattcc gatgatggag atgcttggtc tgcttttctt tttagagcaat ctcatattgag 2400
aacttatcgc tggggagagg atggactagc tggagtctca gacaatcatc aactaatttg 2460
tttctcaatg gcactgtgga atgagaatga tgatattttg aaggagcgat tatttggggt 2520
cactggagag gctgcaaatc atggagagga tgttaaggag ctttattatt atcttgataa 2580
tacaccttct cactcttata tgaaataact ttacaaatat ccacaatcga aatttcctta 2640
cgaagaattg atttcagaga accgtaaacg ttccagatta gaaagagagt acgagattac 2700
tgactctgaa gtactgaagg ataacagata ttttgatgtg atctttgaaa tggcaaggga 2760
cgatgaagat gagaatgaac tttacttttag aattaccgct tacaaccgag gtccacccc 2820
tgccctttta catgtcgtc cacaggtaac ctttagaat acctggctct ggggtataga 2880
tgaggaaaag gatcacgaca aacctatagc ttgcaaggaa taccaagaca acaactattc 2940
tattcggtta gatagtt 2957

```

<210> 105
 <211> 388
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> no

<400> 105

ES 2 589 655 T3

```

gatctgttta gcttgcctcg tccccgccgg gtcacccggc cagcgacatg gagggccaga 60
ataccctcct tgacagtctt gacgtgcgca gtcagggggc atgatgtgac tgtcgcccg 120
acatttagcc catacatccc catgtataat catttgcac catacatttt gatggccgca 180
cggcgcgaag caaaaattac ggctcctcgc tgcagacctg cgagcagggg aacgctcccc 240
tcacagacgc gttgaattgt cccacgcgcg cgccctgta gagaaatata aaagggttagg 300
atttgccact gaggttcttc ttcatatac ttcttttaa aatcttgcta ggatacagtt 360
ctcacatcac atccgaacat aaacaacc 388

```

<210> 106

<211> 247

5 <212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de terminación de TEF1 de *Ashbya gossypii*

10

<400> 106

```

taatcagtac tgacaataaa aagattcttg ttttcaagaa cttgtcattt gtatagtttt 60
tttatattgt agttgttcta ttttaataca atgttagcgt gatattatatt ttttttcgcc 120
tcgacatcat ctgcccagat gcgaagttaa gtgcgcagaa agtaatatca tgcgtcaatc 180
gtatgtgaat gctggtcgct atactgctgt cgattcgata ctaacgccgc catccagtgt 240
cgaaaac 247

```

REIVINDICACIONES

1. Un método para producir una glicoproteína recombinante en *Pichia pastoris* que carece de actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora, que comprende:

(a) proporcionar una célula hospedadora recombinante de *Pichia pastoris* que no presente ninguna actividad de β -manosiltransferasa 1, 2, 3 y 4 con respecto a un *N*-glicano u *O*-glicano, en donde los genes de β -manosiltransferasa 1, 2 y 3 han experimentado delección o interrupción, incluyendo la célula hospedadora una molécula de ácido nucleico que codifica la glicoproteína recombinante

(b) cultivar la célula hospedadora en un medio en condiciones eficaces para expresar la glicoproteína recombinante; y

(c) recuperar la glicoproteína recombinante del medio para producir la glicoproteína recombinante que carece de actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora.

2. El método de la reivindicación 1, en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en un ELISA de tipo sándwich.

3. El método de la reivindicación 1, en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en una transferencia de Western.

4. El método de la reivindicación 1, en el que la glicoproteína recombinante es una glicoproteína terapéutica.

5. El método de la reivindicación 4, en el que la glicoproteína terapéutica se selecciona entre el grupo que consiste en eritropoyetina (EPO); citoquinas tales como interferón α , interferón β , interferón γ e interferón ω ; y factor estimulante de colonias de granulocitos (G-CSF); GM-CSF; factores de coagulación tales como factor VIII, factor IX y proteína C humana; antitrombina III; trombina; cadena α de receptor de IgE soluble; inmunoglobulinas tales como IgG, fragmentos de IgG, fusiones de IgG e IgM; inmunoadhesiones y otras proteínas de fusión de Fc tales como proteínas de fusión a receptor de TNF-Fc solubles; proteínas de fusión de RAGE-Fc; interleuquinas; uroquinas; quimasa; e inhibidor de urea y tripsina; proteína de unión a IGF; factor de crecimiento epidérmico; factor de liberación de hormona de crecimiento; proteína de fusión de anexina V; angiostatina; factor 2 de crecimiento endotelial vascular; factor 1 inhibitorio de precursor mieloide; osteoprotegerina; α -1-antitripsina; α -feto proteínas; DNasa II; kringle 3 de plasminógeno humano; glucocerebrosidasa; proteína 1 de unión del TNF; hormona foliculoestimulante; antígeno 4 - Ig asociado a linfocitos T citotóxicos; activador transmembrana y modulador de calcio y ligando de ciclofilina; proteína 1 de tipo glucagón; y agonista de receptores de IL-2.

6. El método de la reivindicación 1, en el que la célula hospedadora está modificada por ingeniería genética para producir glicoproteínas que tengan *N*-glicanos similares a los humanos.

7. El método de la reivindicación 1, en el que la célula hospedadora está modificada por ingeniería genética para producir glicoproteínas que tengan predominantemente un *N*-glicano seleccionado entre $\text{Man}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_5\text{GlcNAc}_2$, $\text{GalGlcNAcMan}_5\text{GlcNAc}_2$, $\text{NANA}\text{GalGlcNAcMan}_5\text{GlcNAc}_2$, $\text{GlcNAcMan}_3\text{GlcNAc}_2$, $\text{GlcNAc}(1-4)\text{Man}_3\text{GlcNAc}_2$, $\text{Gal}(1-4)\text{GlcNAc}(1-4)\text{Man}_3\text{GlcNAc}_2$ y $\text{NANA}(1-4)\text{Gal}(1-4)\text{GlcNAc}(1-4)\text{Man}_3\text{GlcNAc}_2$.

8. El método de la reivindicación 1 para producir una eritropoyetina humana madura en *Pichia pastoris* que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad detectable de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora, que comprende:

(a) proporcionar una célula hospedadora recombinante de *Pichia pastoris* modificada por ingeniería genética para producir *N*-glicanos biantenarios terminados en ácido siálico y que no presenta ninguna actividad de β -manosiltransferasa 1, 2, 3 y 4 con respecto a un *N*-glicano u *O*-glicano, en donde los genes de β -manosiltransferasa 1, 2 y 3 han experimentado delección o interrupción, incluyendo la célula hospedadora dos o más moléculas de ácido nucleico, cada una codificando una proteína de fusión que comprende una eritropoyetina humana madura fusionada con un péptido señal que se dirige al RE y que se retira cuando la proteína de fusión está en el RE;

(b) cultivar la célula hospedadora en un medio en condiciones eficaces para expresar y procesar las proteínas primera y segunda de fusión; y

(c) recuperar la eritropoyetina humana madura del medio para producir la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada con anticuerpos preparados frente a antígenos de célula hospedadora.

9. El método de la reivindicación 8, en el que el péptido señal es un péptido señal α MATpre de *S. cerevisiae* o un péptido señal de lisozima de pollo.

10. El método de la reivindicación 8, en el que al menos una molécula de ácido nucleico codifica una proteína de fusión en donde la eritropoyetina está fusionada con el péptido señal α MATpre de *S. cerevisiae* y al menos una molécula de ácido nucleico codifica una proteína de fusión, estando la eritropoyetina fusionada con el péptido señal

α MATpre de *S. cerevisiae* o un péptido señal de lisozima de pollo.

11. El método de la reivindicación 8, en el que los codones de la secuencia de ácidos nucleicos de la molécula de ácido nucleico que codifica la eritropoyetina están optimizados para su expresión en *Pichia pastoris*.
12. El método de la reivindicación 8, en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en un ELISA de tipo sándwich.
13. El método de la reivindicación 8, en el que la actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora se determina en una transferencia de Western.
14. El método de la reivindicación 8, en el que la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora del medio, incluye una etapa de cromatografía de intercambio catiónico.
15. El método de la reivindicación 8, en el que la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora del medio, incluye una etapa de cromatografía de hidroxipatita.
16. El método de la reivindicación 8, en el que la recuperación de la eritropoyetina humana madura que comprende predominantemente *N*-glicanos biantenarios terminados en ácido siálico y que no tiene actividad de unión cruzada detectable con anticuerpos preparados frente a antígenos de célula hospedadora del medio incluye una etapa de cromatografía de intercambio aniónico.

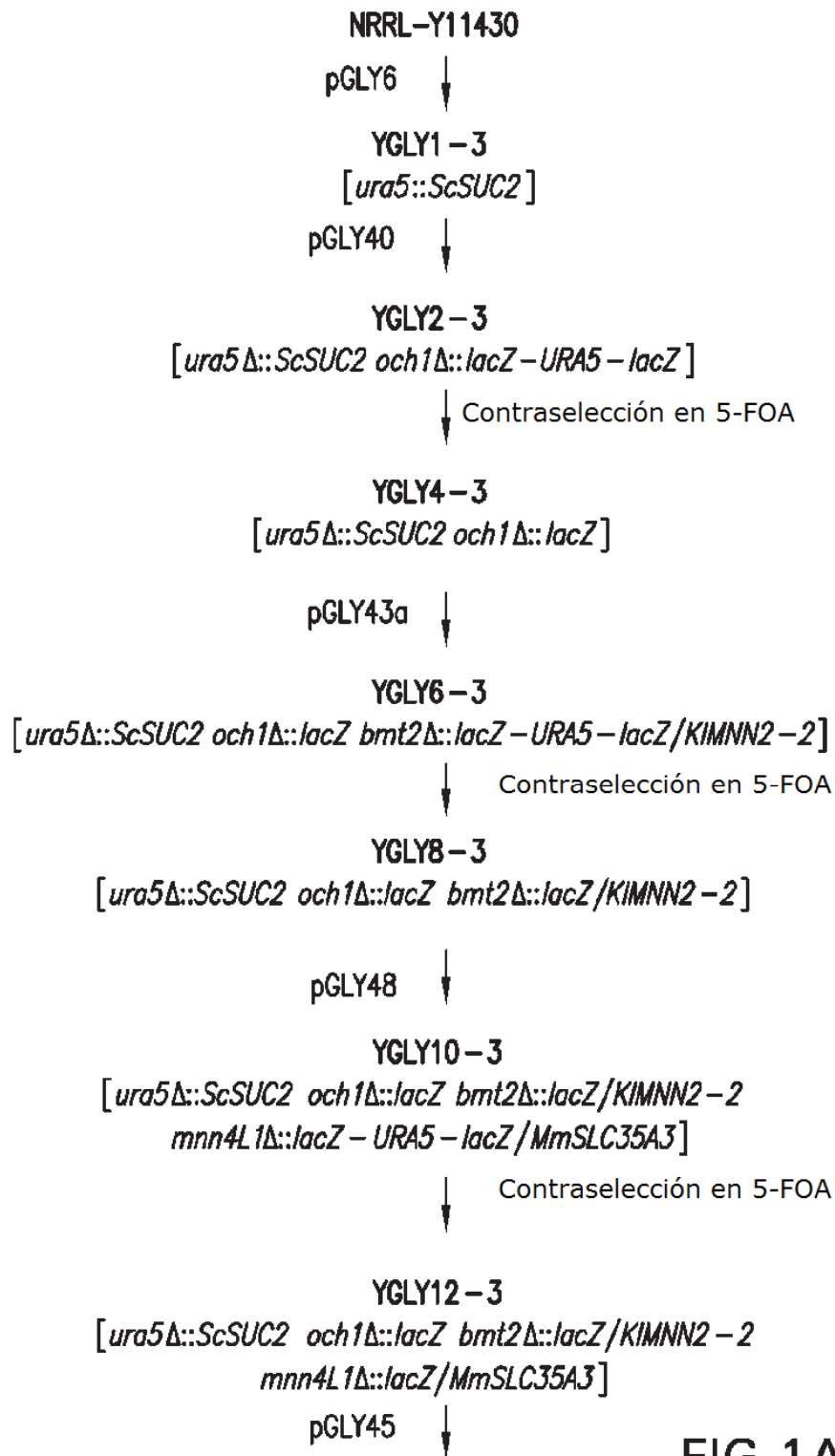


FIG.1A

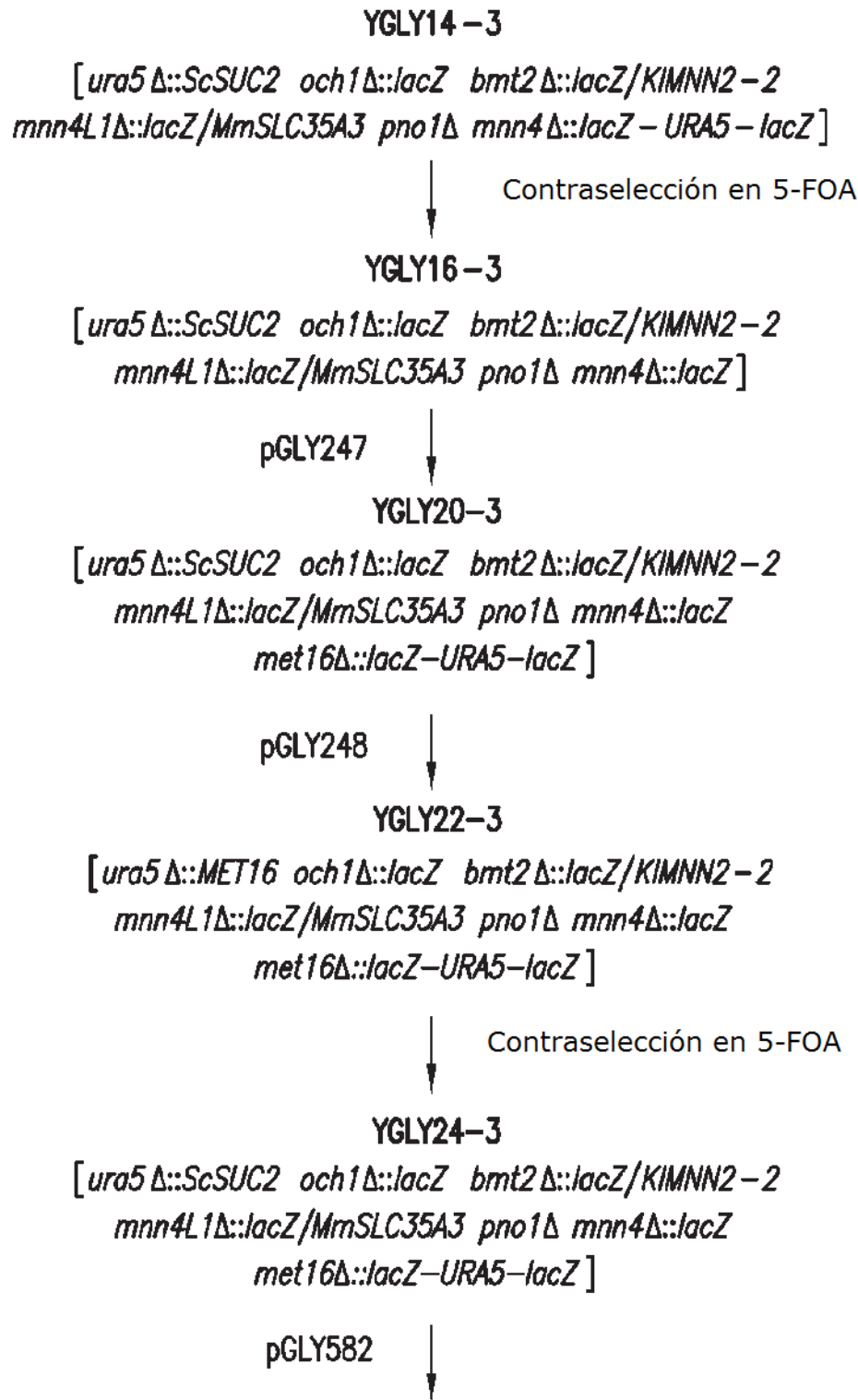


FIG. 1 B

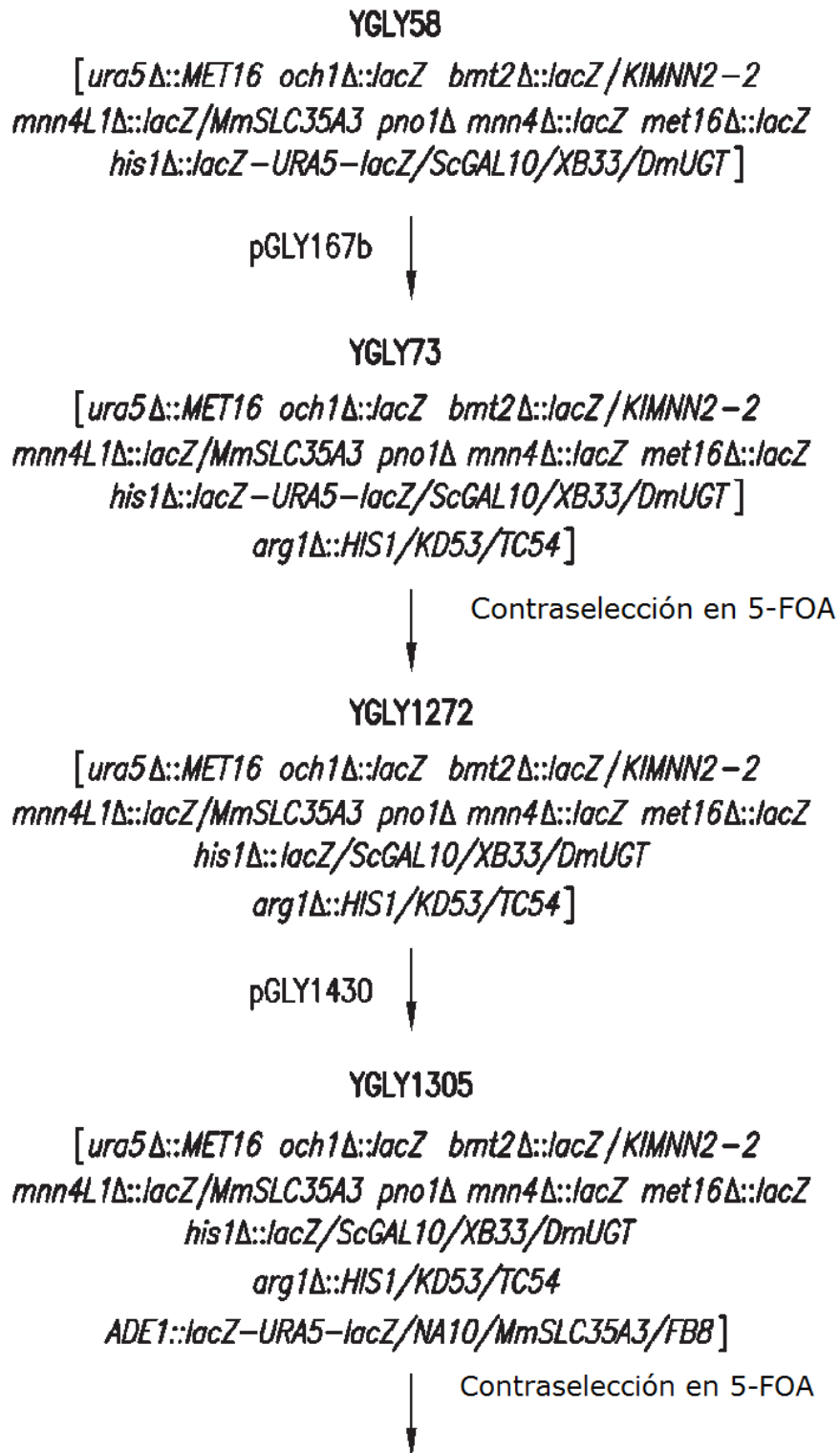


FIG. 1 C

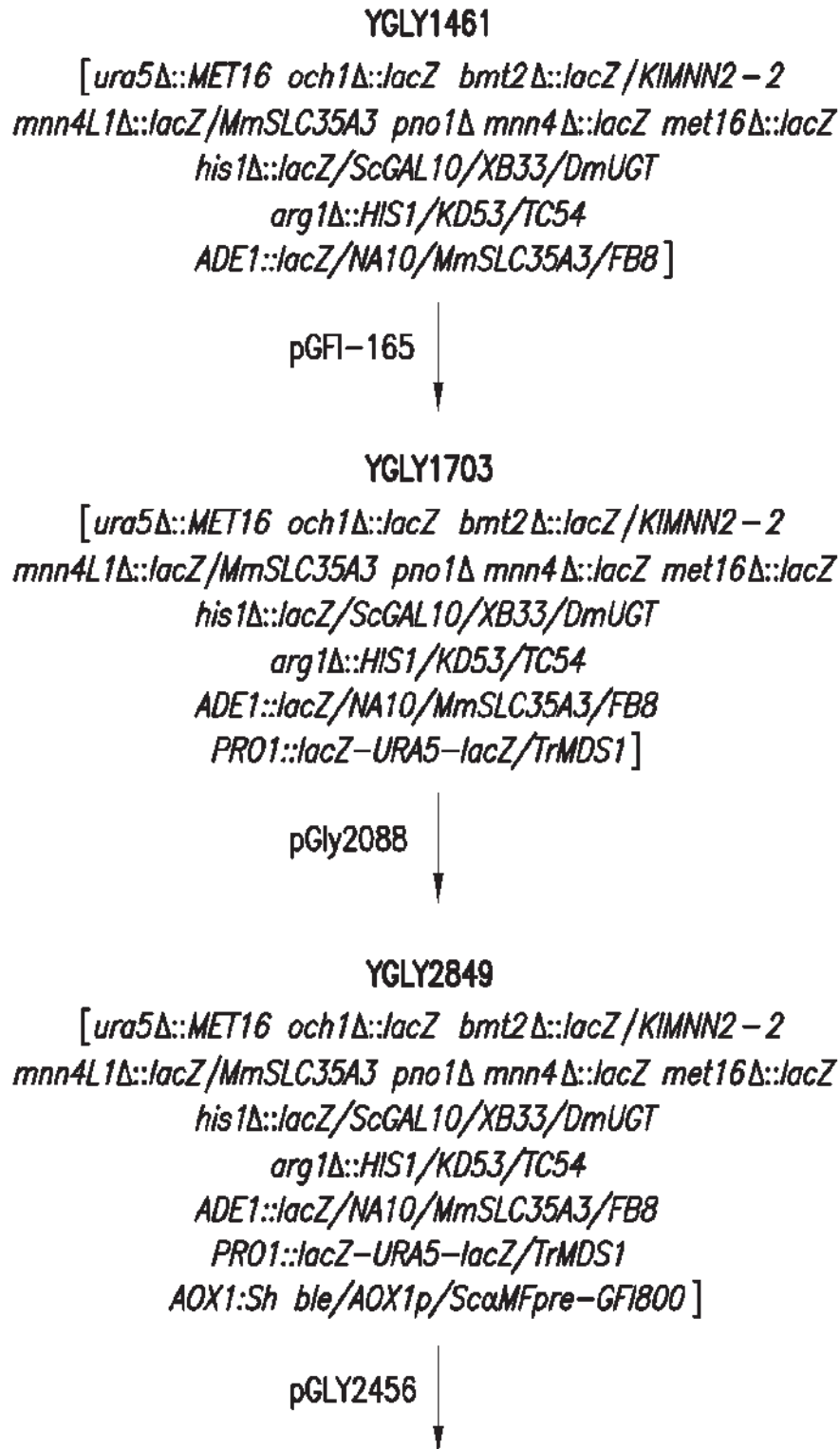


FIG. 1 D

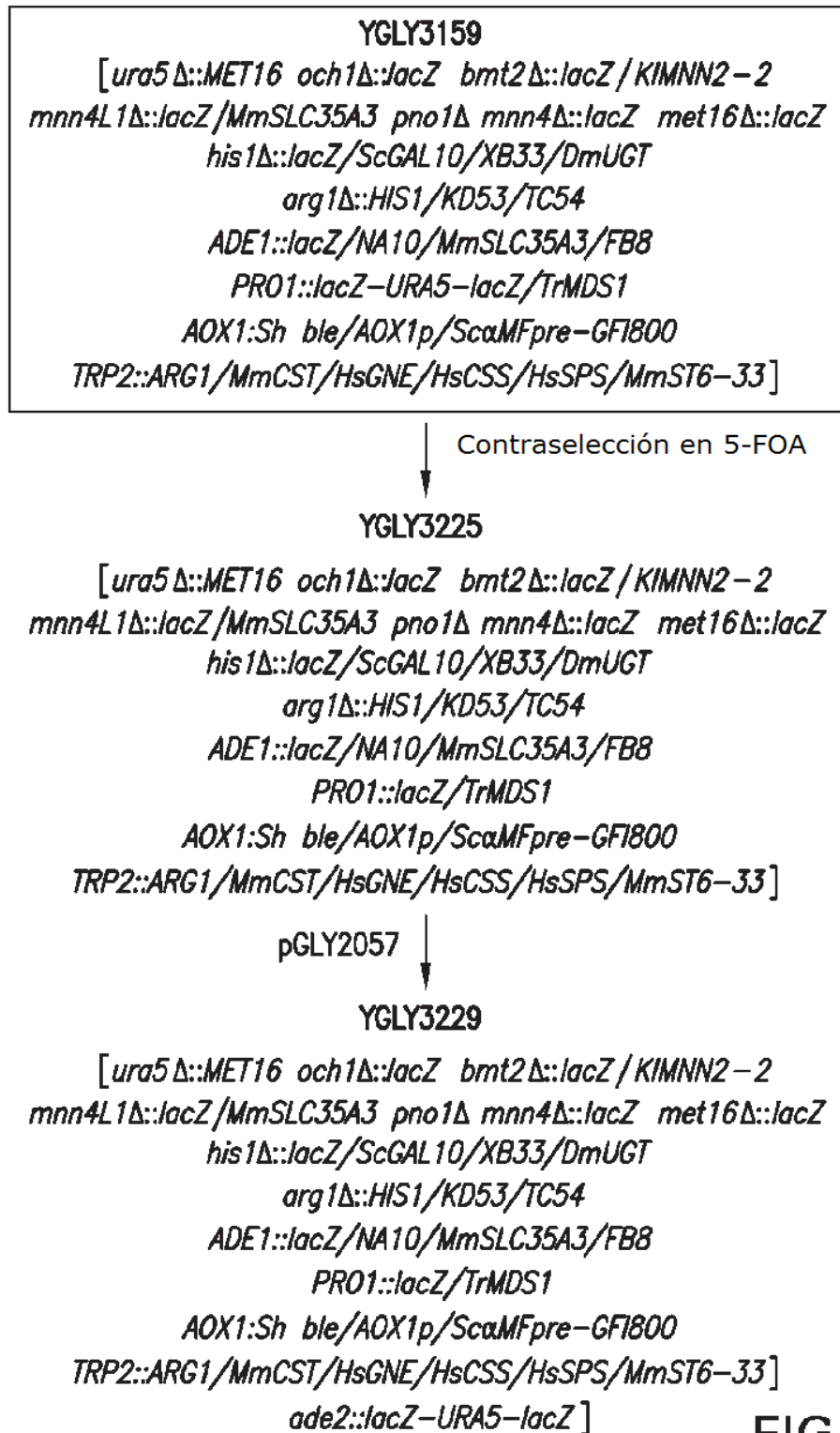


FIG.1E

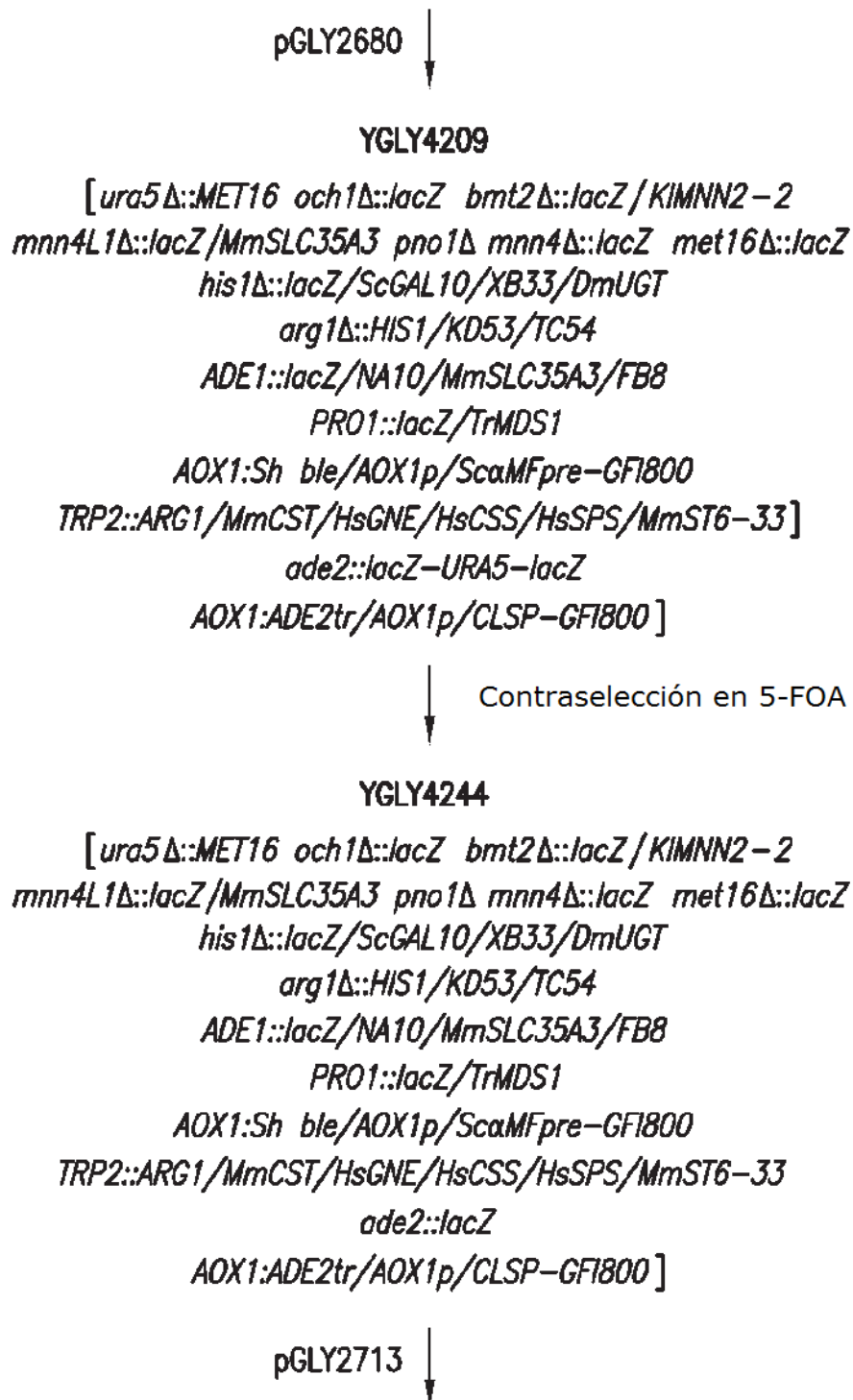


FIG.1F

YGLY5053

[*ura5Δ::MET16 och1Δ::lacZ bmt2Δ::lacZ/KIMNN2-2*
mnn4L1Δ::lacZ/MmSLC35A3 pno1Δ mnn4Δ::lacZ met16Δ::lacZ
his1Δ::lacZ/ScGAL10/XB33/DmUGT
arg1Δ::HIS1/KD53/TC54
ADE1::lacZ/NA10/MmSLC35A3/FB8
PRO1::lacZ/TrMDS1
AOX1:Sh ble/AOX1p/ScαMFpre-GFI800
TRP2::ARG1/MmCST/HsGNE/HsCSS/HsSPS/MmST6-33]
ade2::lacZ
AOX1:ADE2tr/AOX1p/CLSP-GFI800
pep4::lacZ-URA5-lacZ/PpPNO1]



Contraselección en 5-FOA

YGLY5597

[*ura5Δ::MET16 och1Δ::lacZ bmt2Δ::lacZ/KIMNN2-2*
mnn4L1Δ::lacZ/MmSLC35A3 pno1Δ mnn4Δ::lacZ met16Δ::lacZ
his1Δ::lacZ/ScGAL10/XB33/DmUGT
arg1Δ::HIS1/KD53/TC54
ADE1::lacZ/NA10/MmSLC35A3/FB8
PRO1::lacZ/TrMDS1
AOX1:Sh ble/AOX1p/ScαMFpre-GFI800
TRP2::ARG1/MmCST/HsGNE/HsCSS/HsSPS/MmST6-33]
ade2::lacZ
AOX1:ADE2tr/AOX1p/CLSP-GFI800
pep4::lacZ/PpPNO1]

pGLY3411



FIG. 1 G

YGLY5618

[*ura5Δ::MET16 och1Δ::lacZ bmt2Δ::lacZ / KIMNN2-2*
mnn4L1Δ::lacZ / MmSLC35A3 pno1Δ mnn4Δ::lacZ met16Δ::lacZ
his1Δ::lacZ / ScGAL10 / XB33 / DmUGT
arg1Δ::HIS1 / KD53 / TC54
ADE1::lacZ / NA10 / MmSLC35A3 / FB8
PRO1::lacZ / TrMDS1
AOX1:Sh ble / AOX1p / ScaMFpre-GFI800
TRP2::ARG1 / MmCST / HsGNE / HsCSS / HsSPS / MmST6-33]
ade2::lacZ
AOX1:ADE2tr / AOX1p / CLSP-GFI800
pep4::lacZ / PpPNO1
bmt4::lacZ-URA5-lacZ]

pGLY3430 ↓

YGLY7110

[*ura5Δ::MET16 och1Δ::lacZ bmt2Δ::lacZ / KIMNN2-2*
mnn4L1Δ::lacZ / MmSLC35A3 pno1Δ mnn4Δ::lacZ met16Δ::lacZ
his1Δ::lacZ / ScGAL10 / XB33 / DmUGT
arg1Δ::HIS1 / KD53 / TC54
ADE1::lacZ / NA10 / MmSLC35A3 / FB8
PRO1::lacZ / TrMDS1
AOX1:Sh ble / AOX1p / ScaMFpre-GFI800
TRP2::ARG1 / MmCST / HsGNE / HsCSS / HsSPS / MmST6-33]
ade2::lacZ
AOX1:ADE2tr / AOX1p / CLSP-GFI800
pep4::lacZ / PpPNO1
bmt4::lacZ-URA5-lacZ bmt1::Nat^R]

pGLY4472 ↓

FIG. 1H

YGLY7113-7122

[*ura5Δ::MET16 och1Δ::lacZ bmt2Δ::lacZ/KIMNN2-2*
mnn4L1Δ::lacZ/MmSLC35A3 pno1Δ mnn4Δ::lacZ met16Δ::lacZ
his1Δ::lacZ/ScGAL10/XB33/DmUGT
arg1Δ::HIS1/KD53/TC54
ADE1::lacZ/NA10/MmSLC35A3/FB8
PRO1::lacZ/TrMDS1
AOX1:Sh ble/AOX1p/ScαMFpre-GFI800
TRP2::ARG1/MmCST/HsGNE/HsCSS/HsSPS/MmST6-33]
ade2::lacZ
AOX1:ADE2tr/AOX1p/CLSP-GFI800
pep4::lacZ/PpPNO1
bmt4::lacZ-URA5-lacZ bmt1::Nat^R bmt3::Hyg^R]

FIG. 1I

Glosario

ScSUC2	Invertasa de <i>S. cerevisiae</i>
OCH1	Alfa-1,6-manosiltransferasa
MMNN2-2	Transportador de UDP-GlcNAc de <i>K. lactis</i>
BMT1	Transferencia de beta-manosa (eliminación de beta-manosa)
BMT2	Transferencia de beta-manosa (eliminación de beta-manosa)
BMT3	Transferencia de beta-manosa (eliminación de beta-manosa)
BMT4	Transferencia de beta-manosa (eliminación de beta-manosa)
MNN4L1	1 similar a MNN4 (eliminación de carga)
MmSLC35A3	Homólogo de ratón del transportador de UDP-GlcNAc
PN01	Fosfomanosilación de <i>N</i>-glicanos (eliminación de carga)
MNN4	Manosiltransferasa (eliminación de carga)
ScGALW	UDP-glucosa 4-epimerasa
XB33	HsGalTI truncado fusionado con el líder <i>ScKRE2</i>
DmUGT	Transportador de UDP-Galactosa
KD53	DmMNSII truncado fusionado con el líder <i>ScMNN2</i>
TC54	RnGNTII truncado fusionado con el líder <i>ScMNN2</i>
NA10	HsGNTI truncado fusionado con el líder <i>PpSEC12</i>
FB8	MmMNSIA truncado fusionado con el líder <i>ScSEC12</i>
TrMDS1	<i>MNS1</i> de <i>T. reseei</i> secretada
Sh ble	Marcador de resistencia a zeocina
ScαMFpre-GFI800	Fusión de presecuencia del factor de acoplamiento de Sc alfa con HsEPO
MmCST	Transportador de CMP-ácido siálico de ratón
HsGNE	UDP-GlcNAc 2-epimerasa humana /
	<i>N</i>-acetilmanosamina quinasa
HsCSS	CMP-ácido siálico sintasa humana
HsSPS	<i>N</i>-acetilneuraminato-9-fosfato sintasa humana
MmST6-33	α-2,6-Sialil transferasa de ratón truncada
	Fusionada con el líder <i>ScKRE2</i>
ADE2tr	Marcador <i>ADE2</i> truncado
CLSP	Péptido señal de Lisozima de Pollo
PEP4	Proteinasa A

Fig 1J

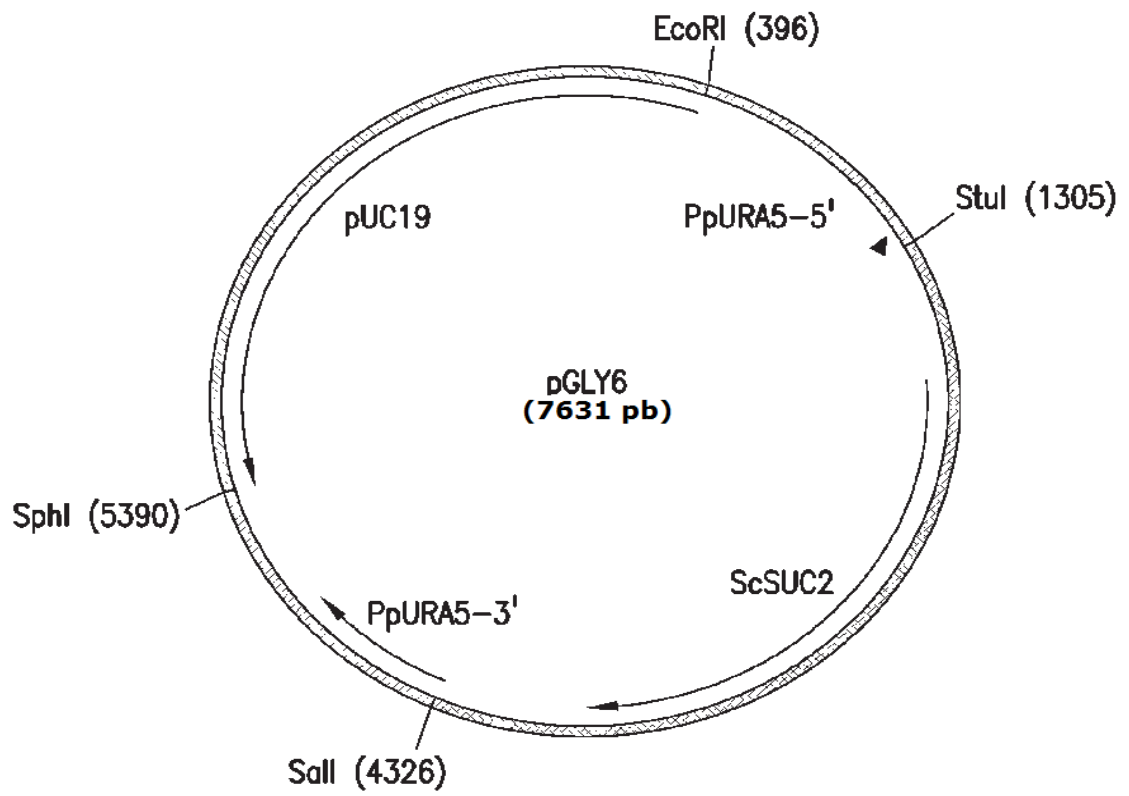


FIG.2

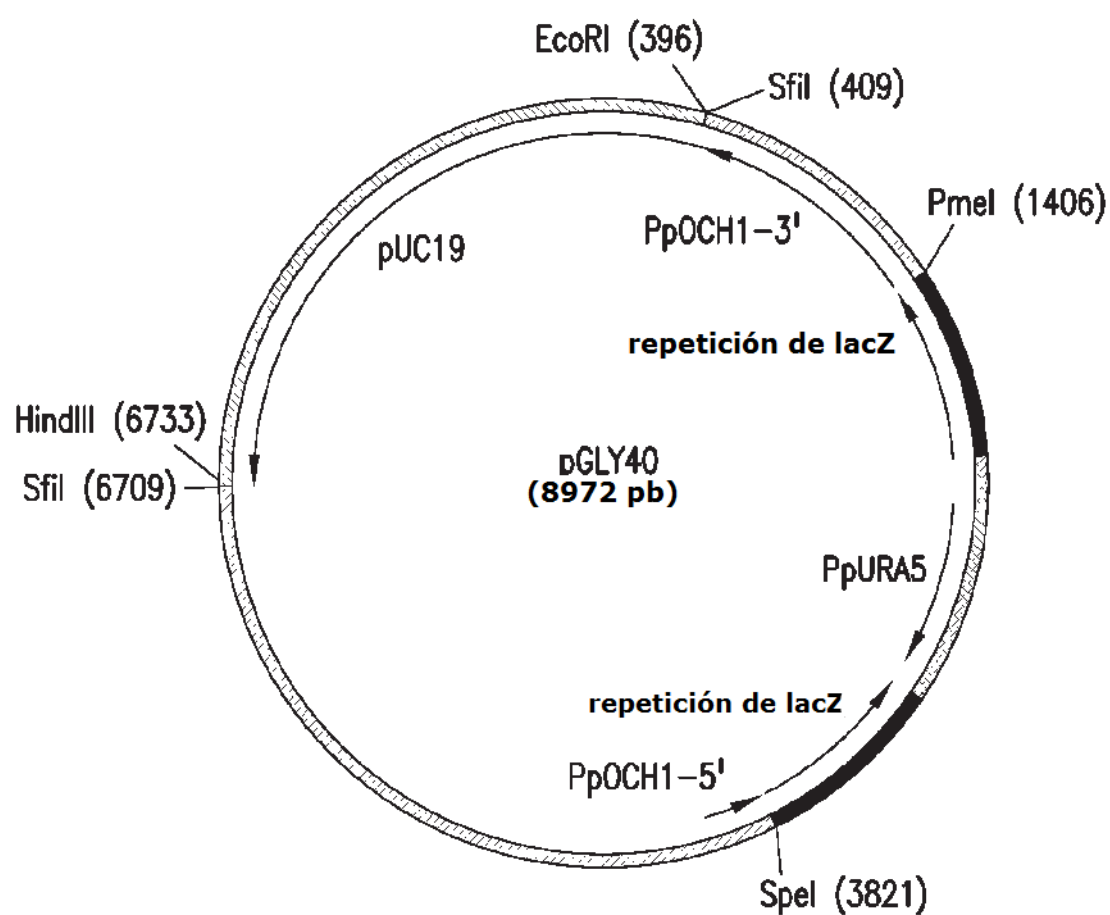


FIG.3

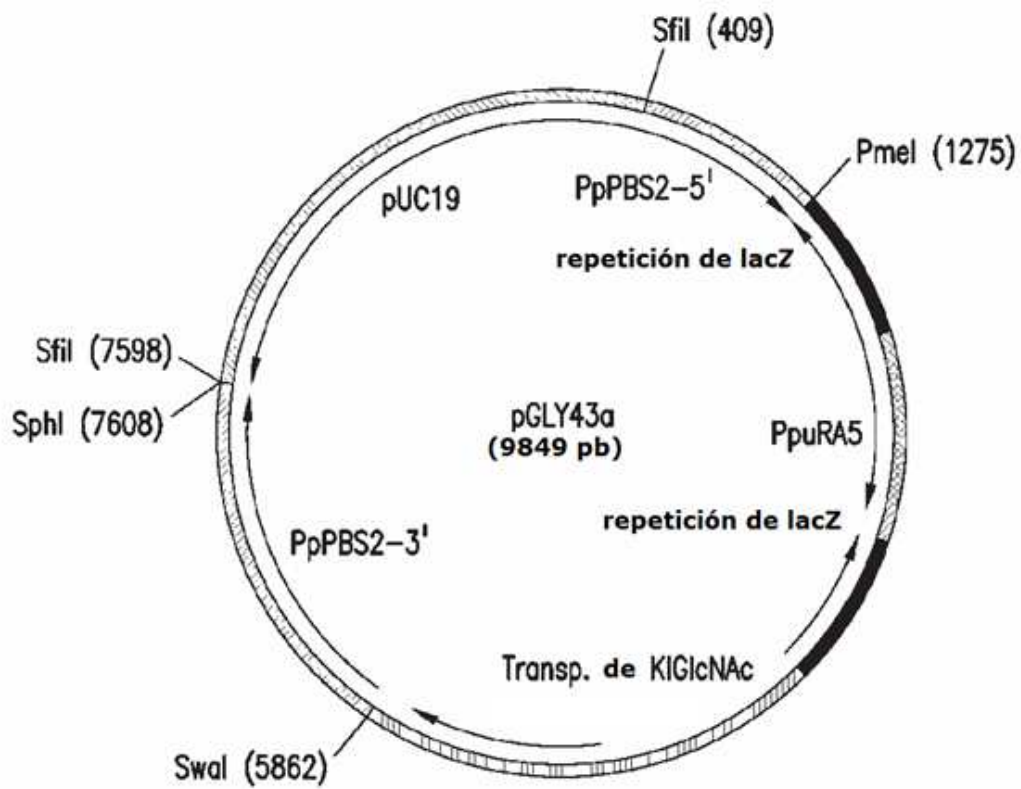


FIG.4

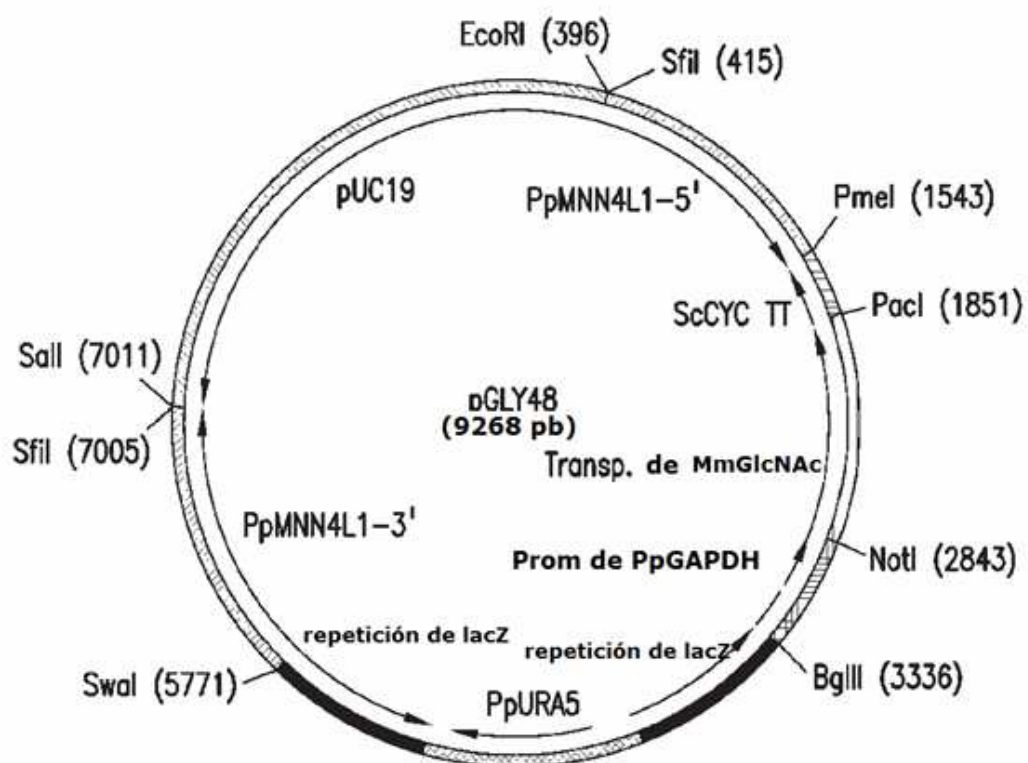


FIG.5

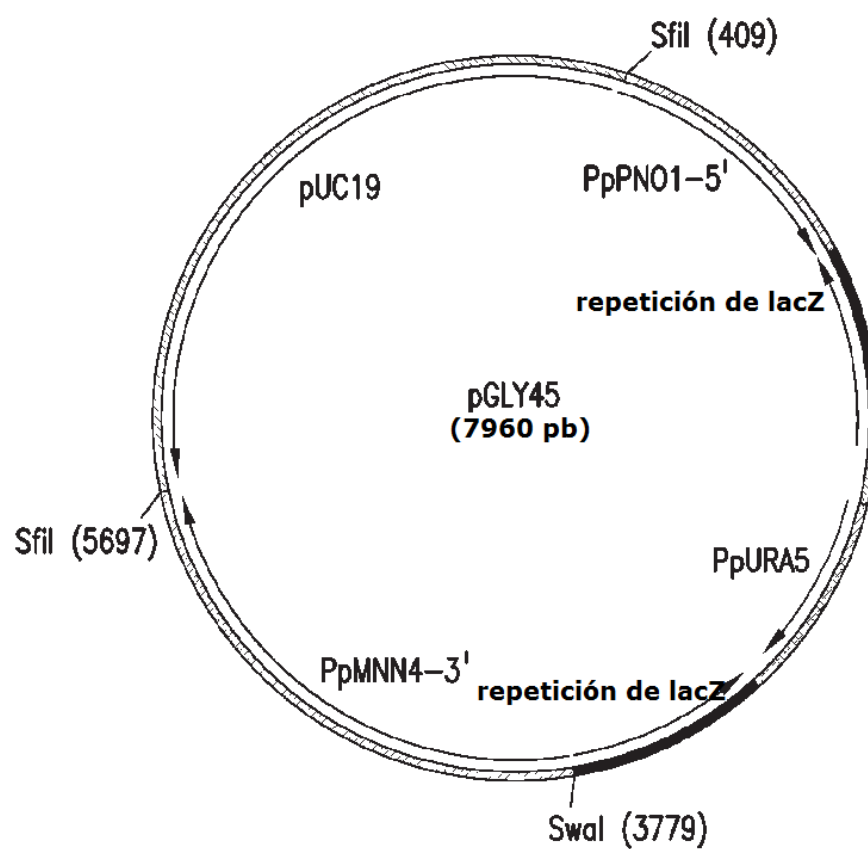


FIG.6

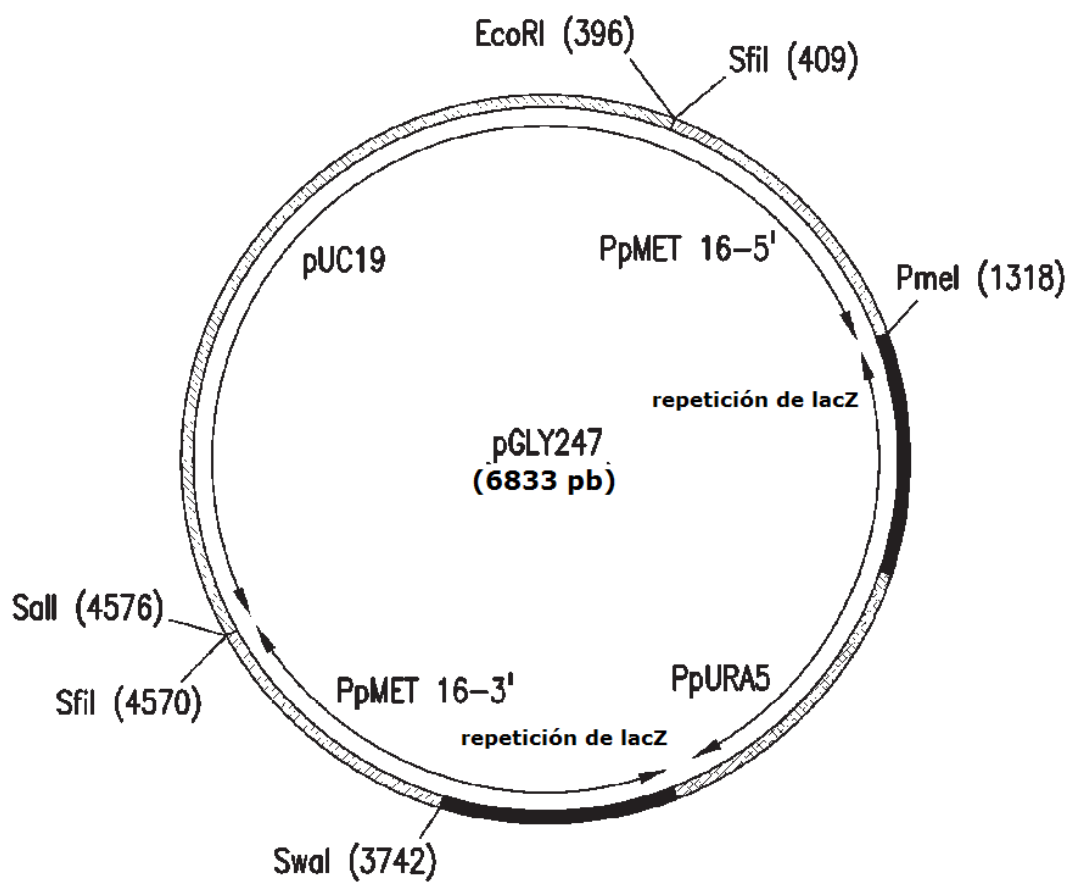


FIG.7

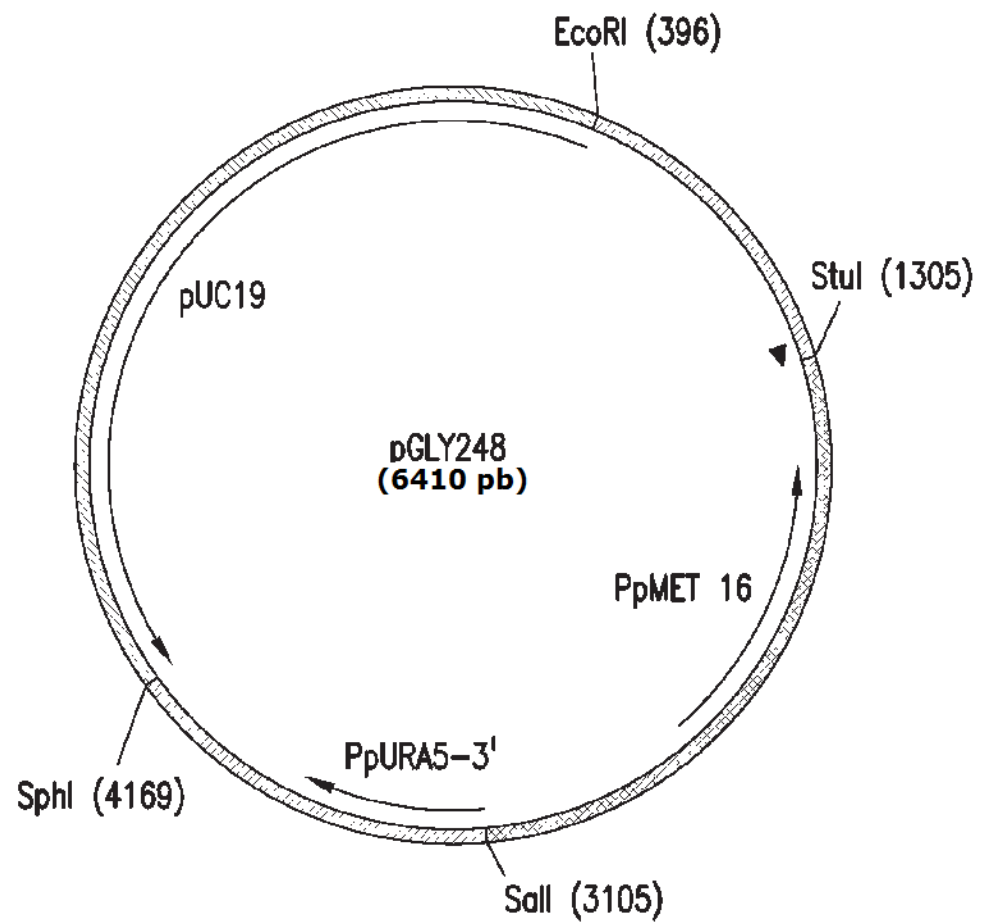


FIG.8

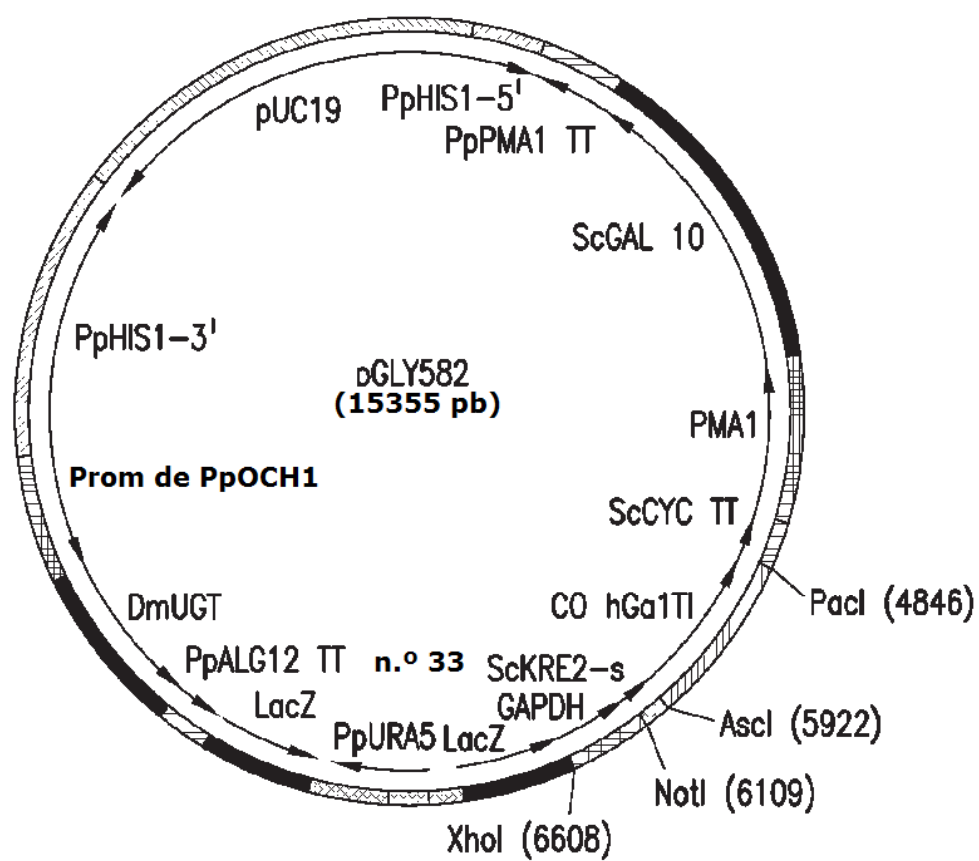


FIG.9

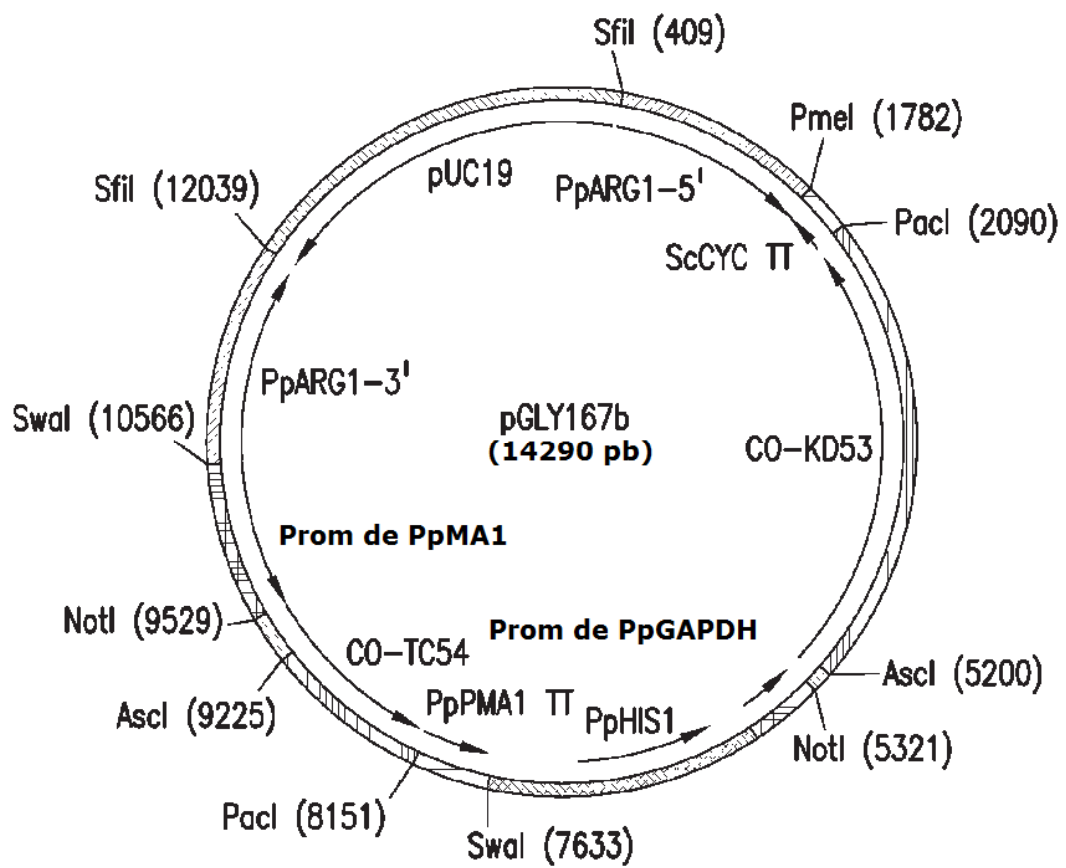


FIG.10

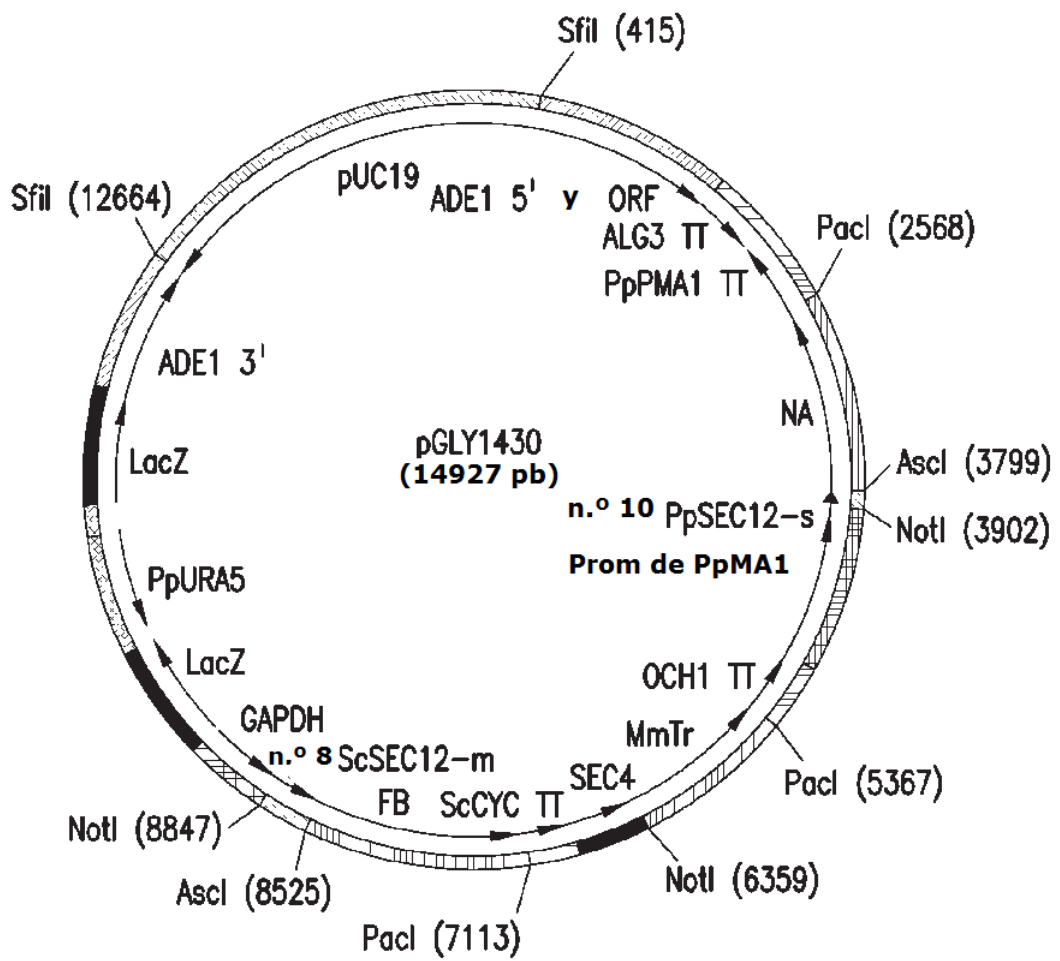


FIG.11

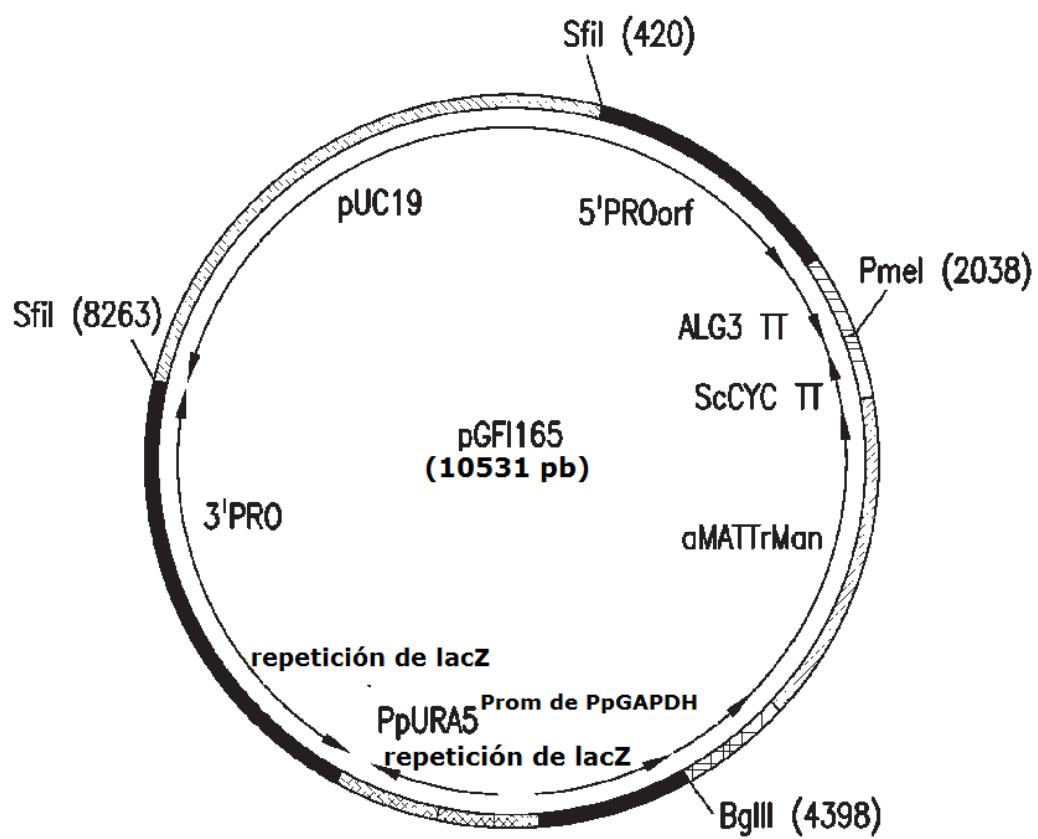


FIG.12

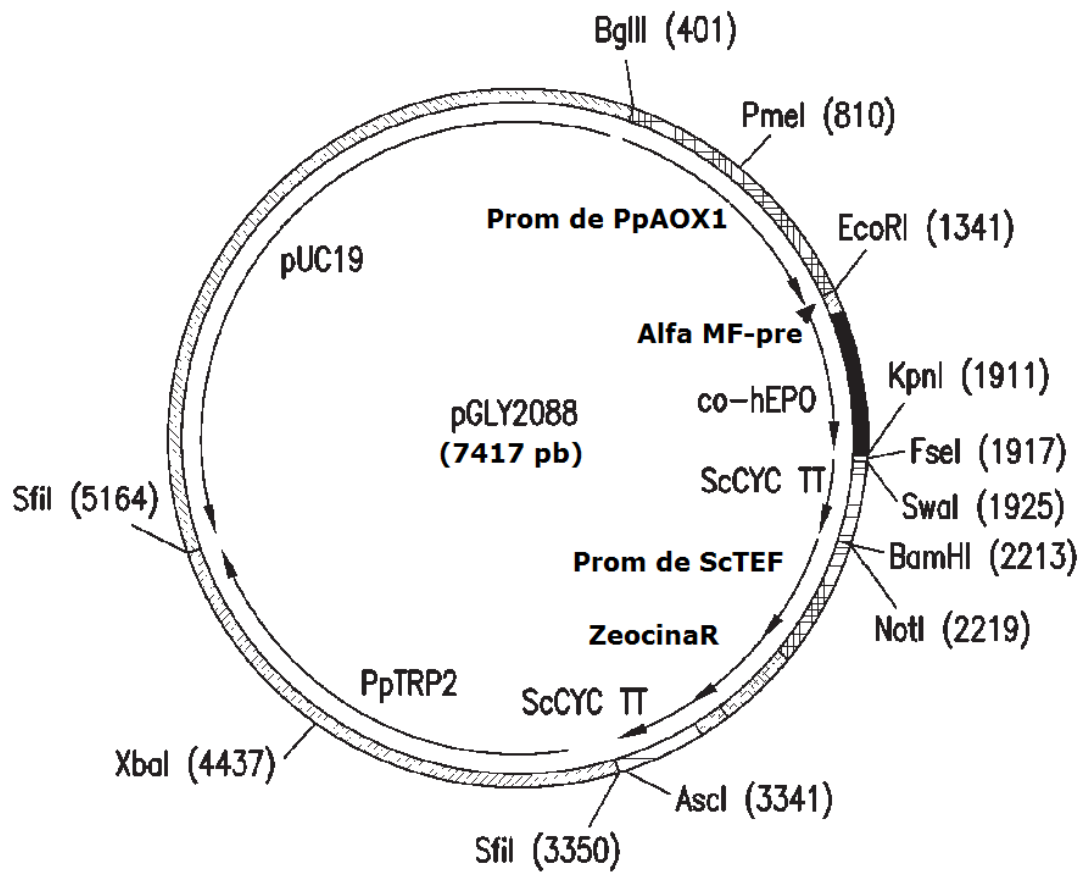


FIG.13

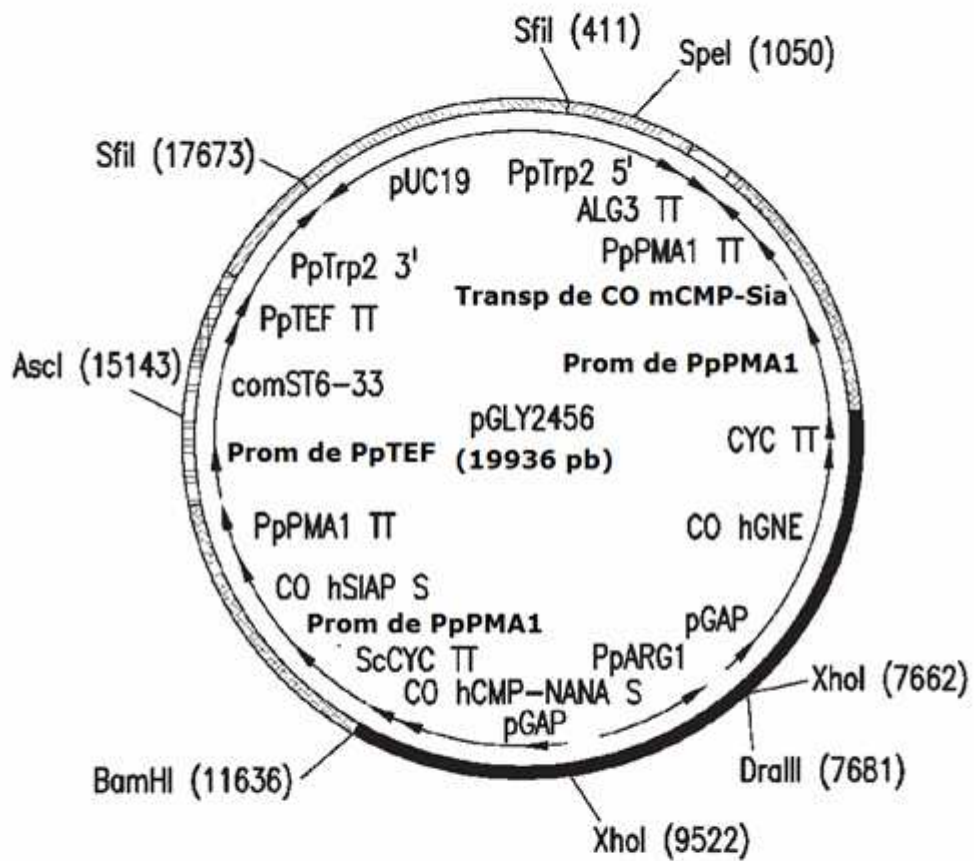


FIG.14

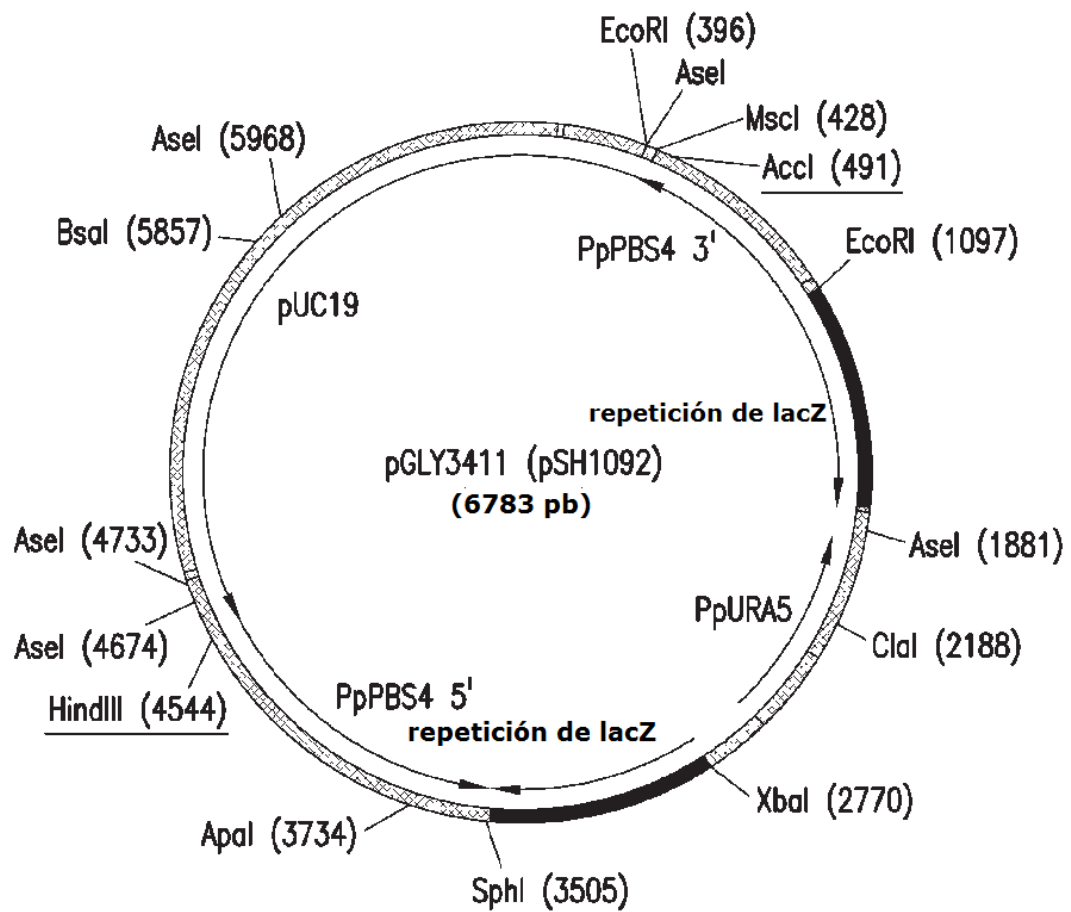


FIG.15

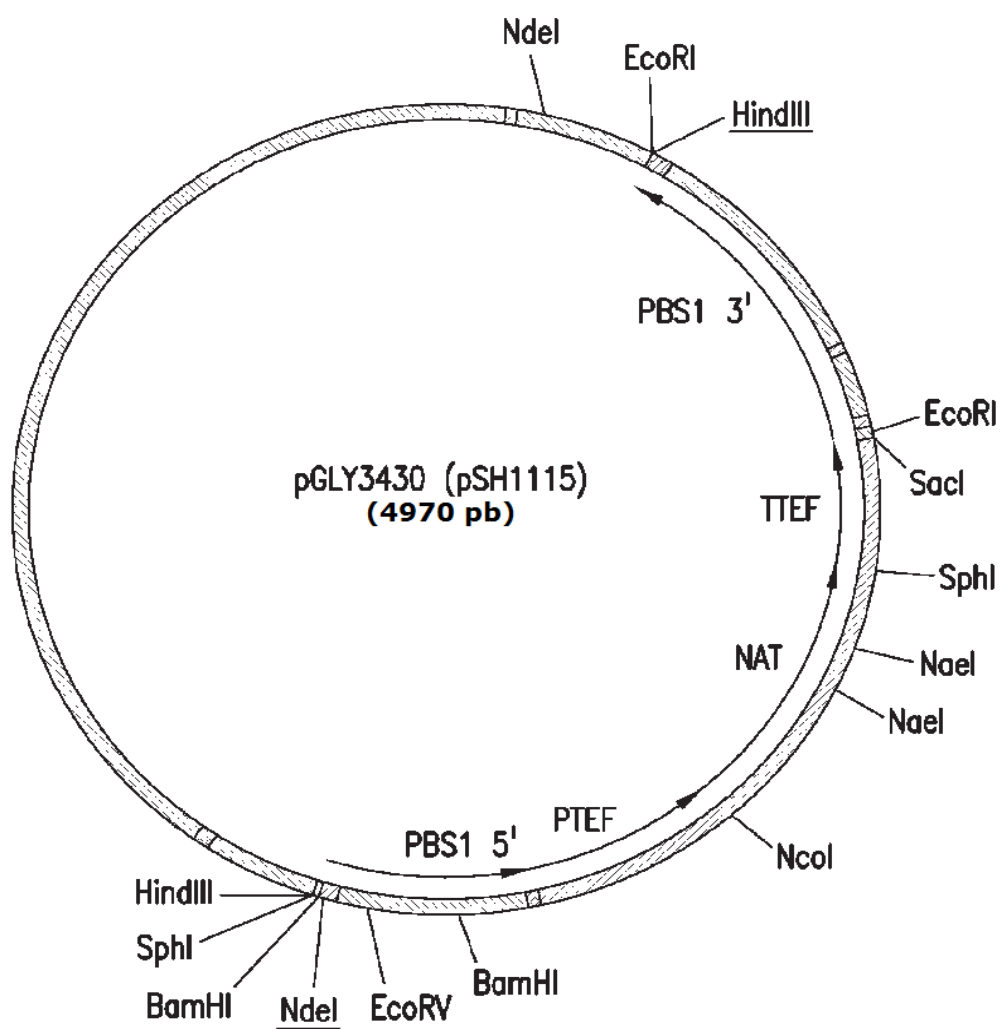


FIG.16

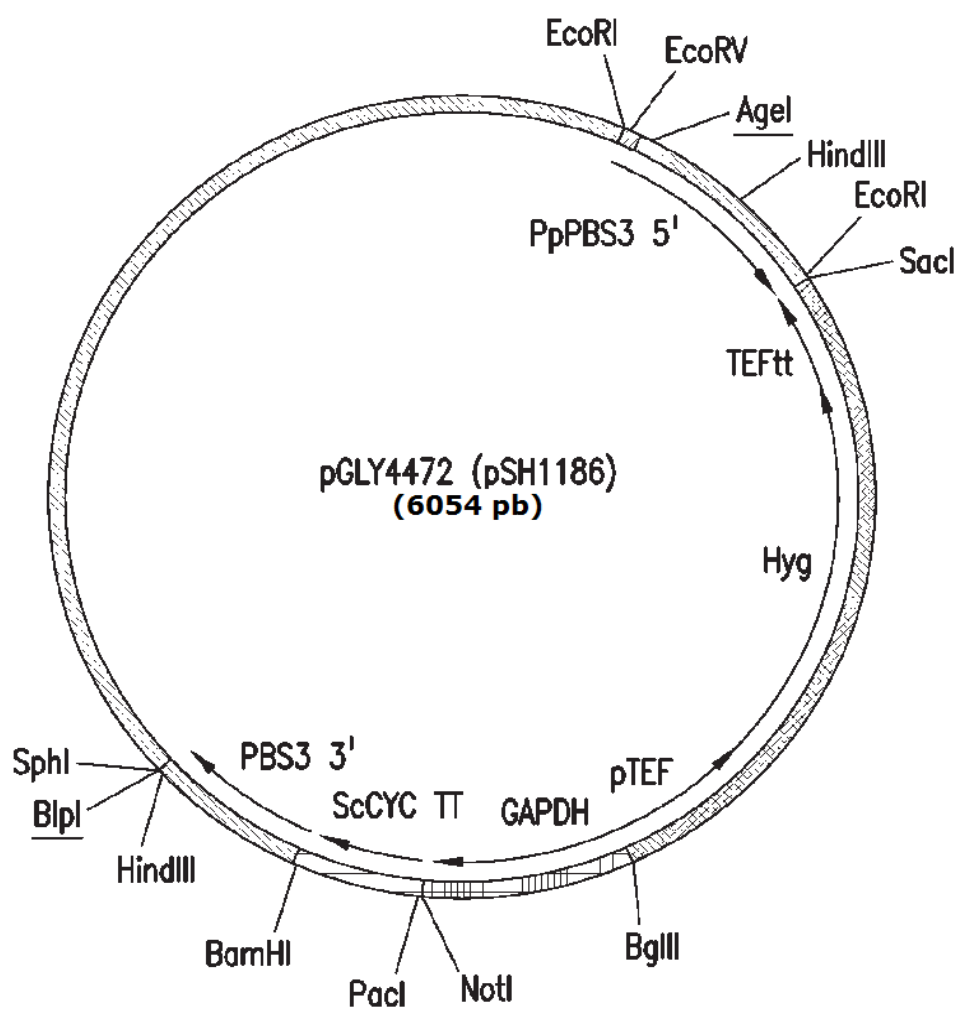


FIG.17

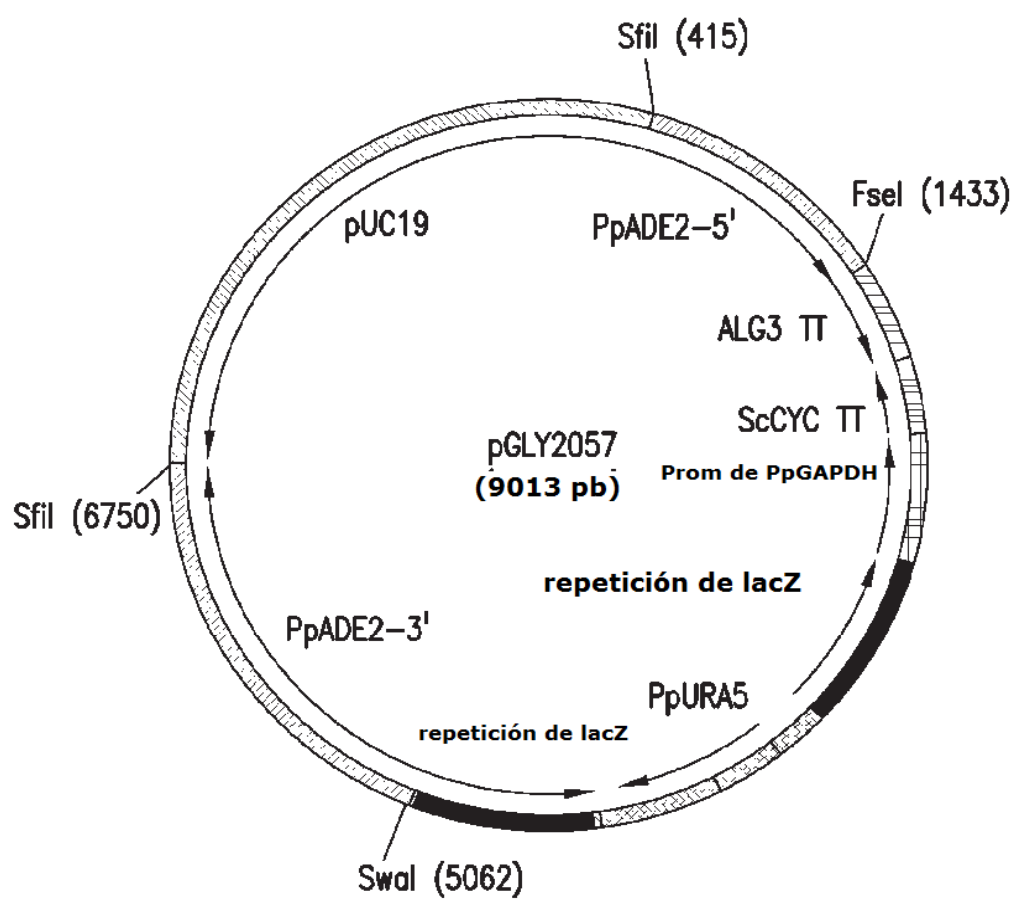


FIG.18

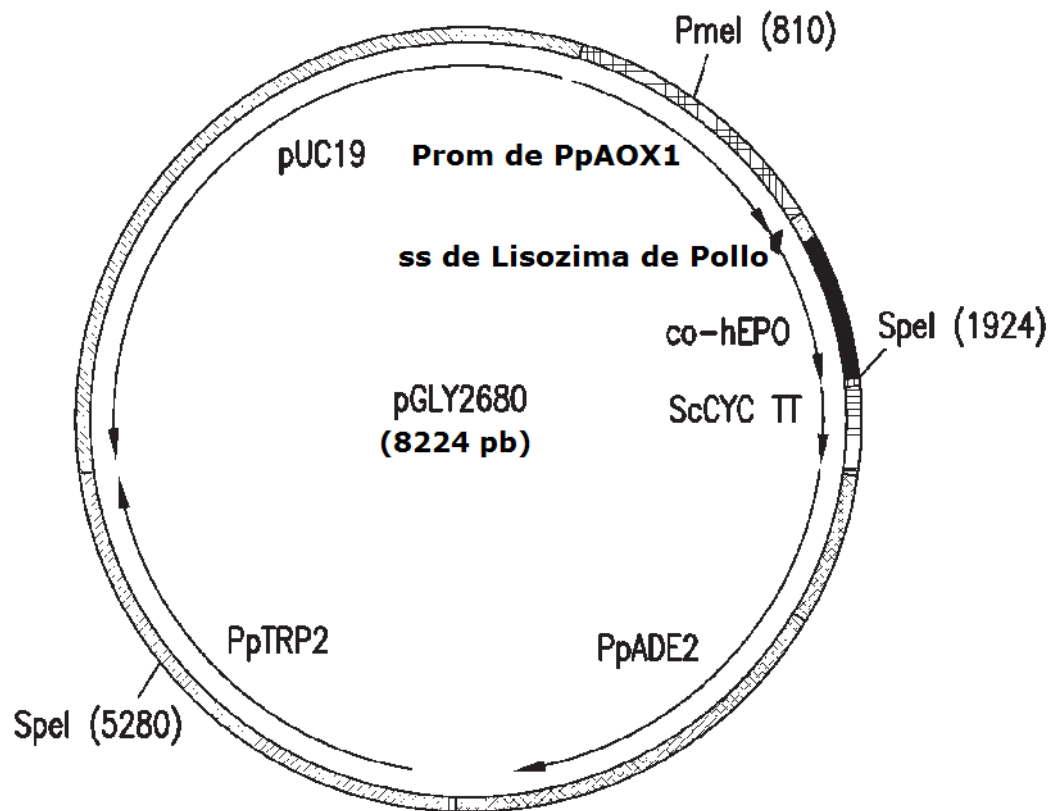


FIG.19

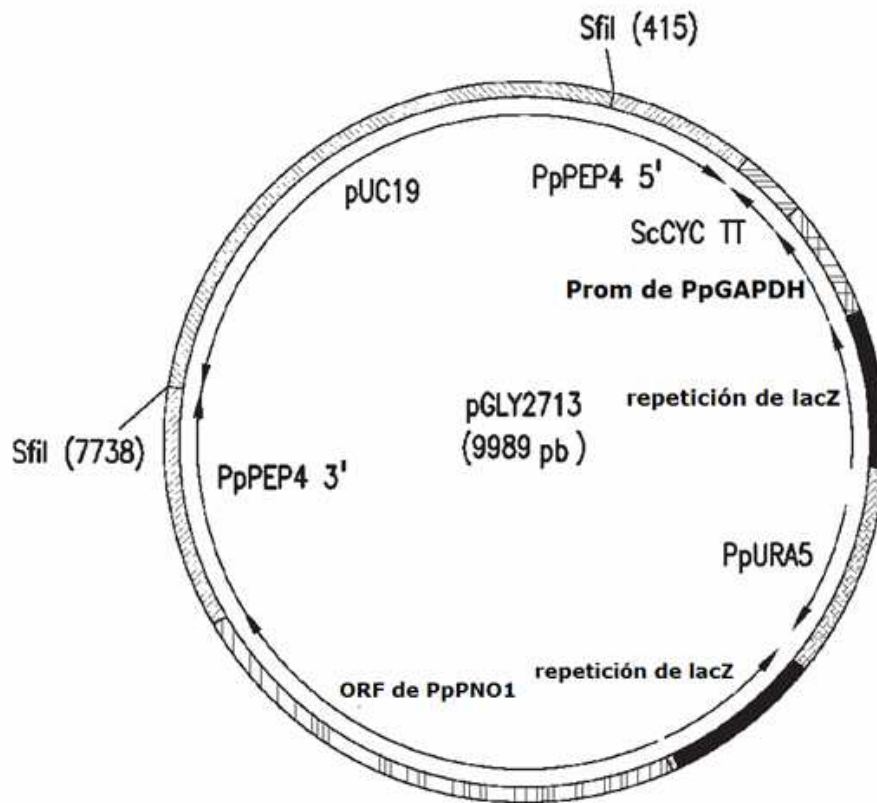


FIG.20

Flujo del Proceso de Fermentación

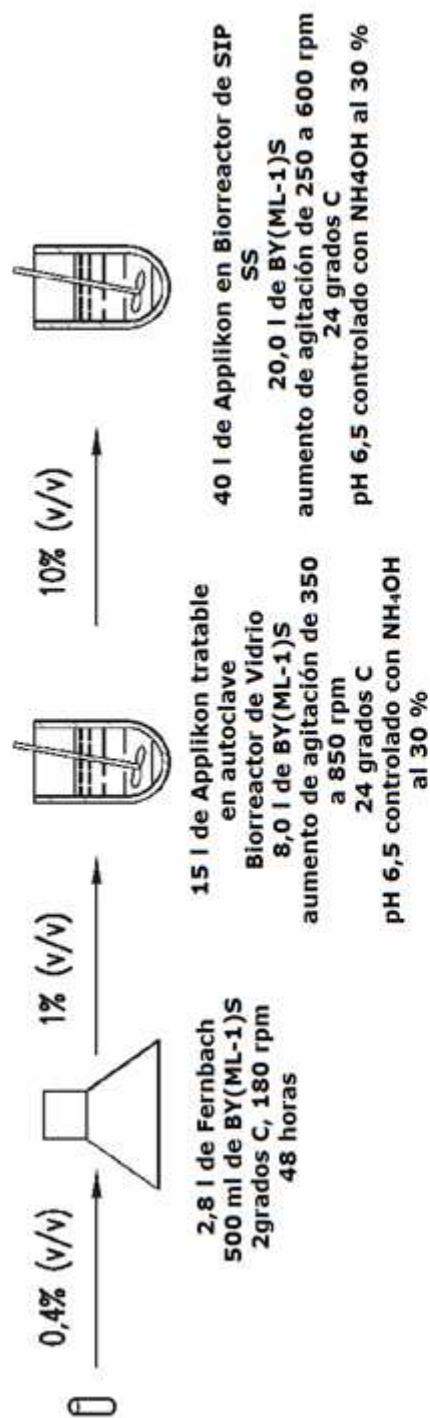


FIG.21

Transferencia de Western del anticuerpo HCA

Detección de una proteína de aproximadamente el tamaño de rhEPO

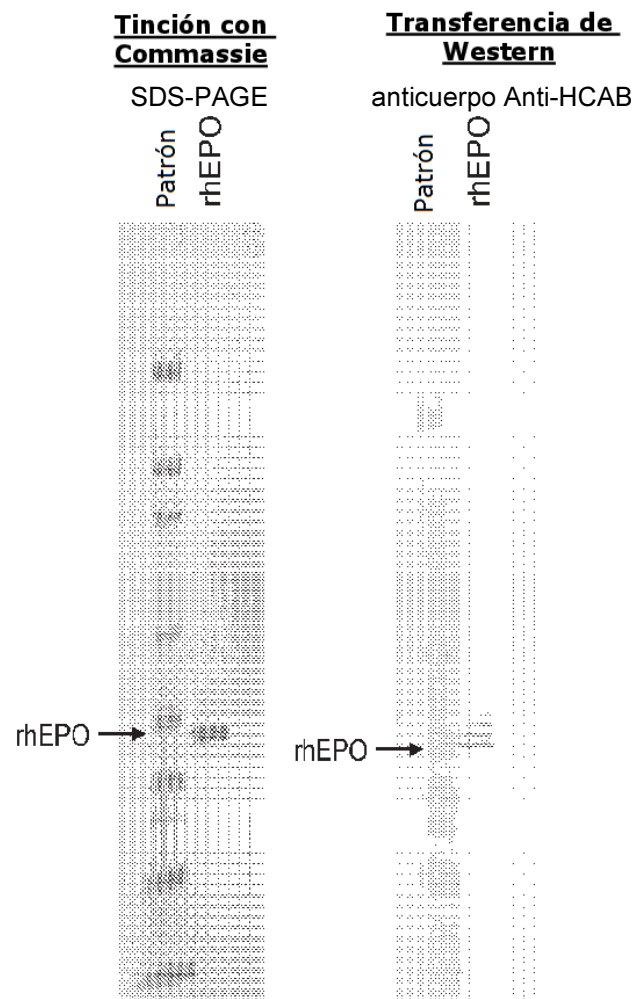


FIG.22

Transferencia de Western del anticuerpo Anti-HCA
Sin detección de la forma desglucosilada de esta proteína

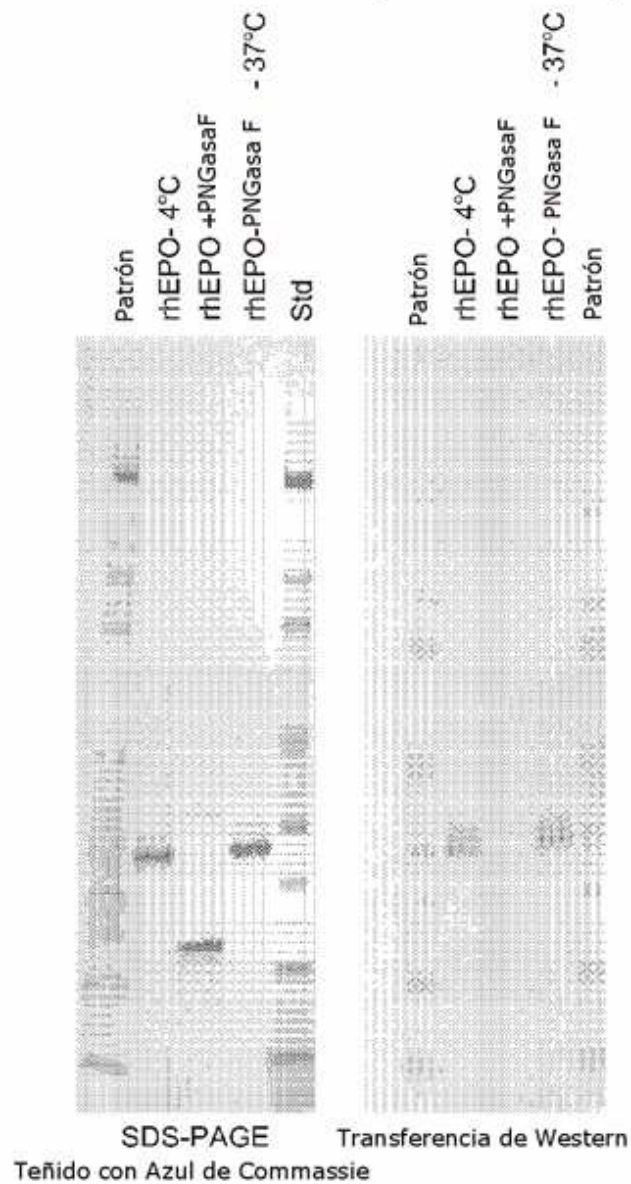


FIG.23

Reactividad de anti-HCP con rhlgG1 producido
en *Pichia pastoris*

- Calle 1: rhlgG1 producido en GS2.0 que contiene *N*-glicanos de Man₉GlcNAc₂ detectables que eran resistentes a la α 1,2-manosidasa
Calle 2: Marcador
Calle 3: rhlgG1 producido en WT
Calle 4: rhlgG1 producido en WT+PNGasaF
Calle 5: Marcador

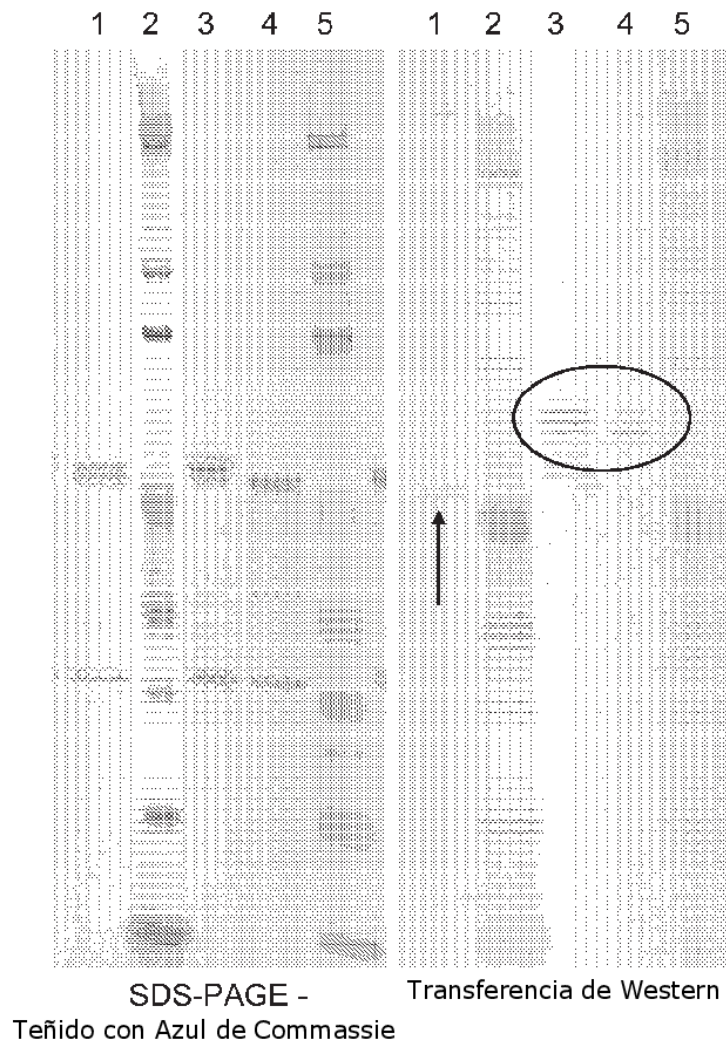


FIG.24

Detección de HCA en rhEPO y Otras Preparaciones de Proteína Producidas en *Pichia pastoris*

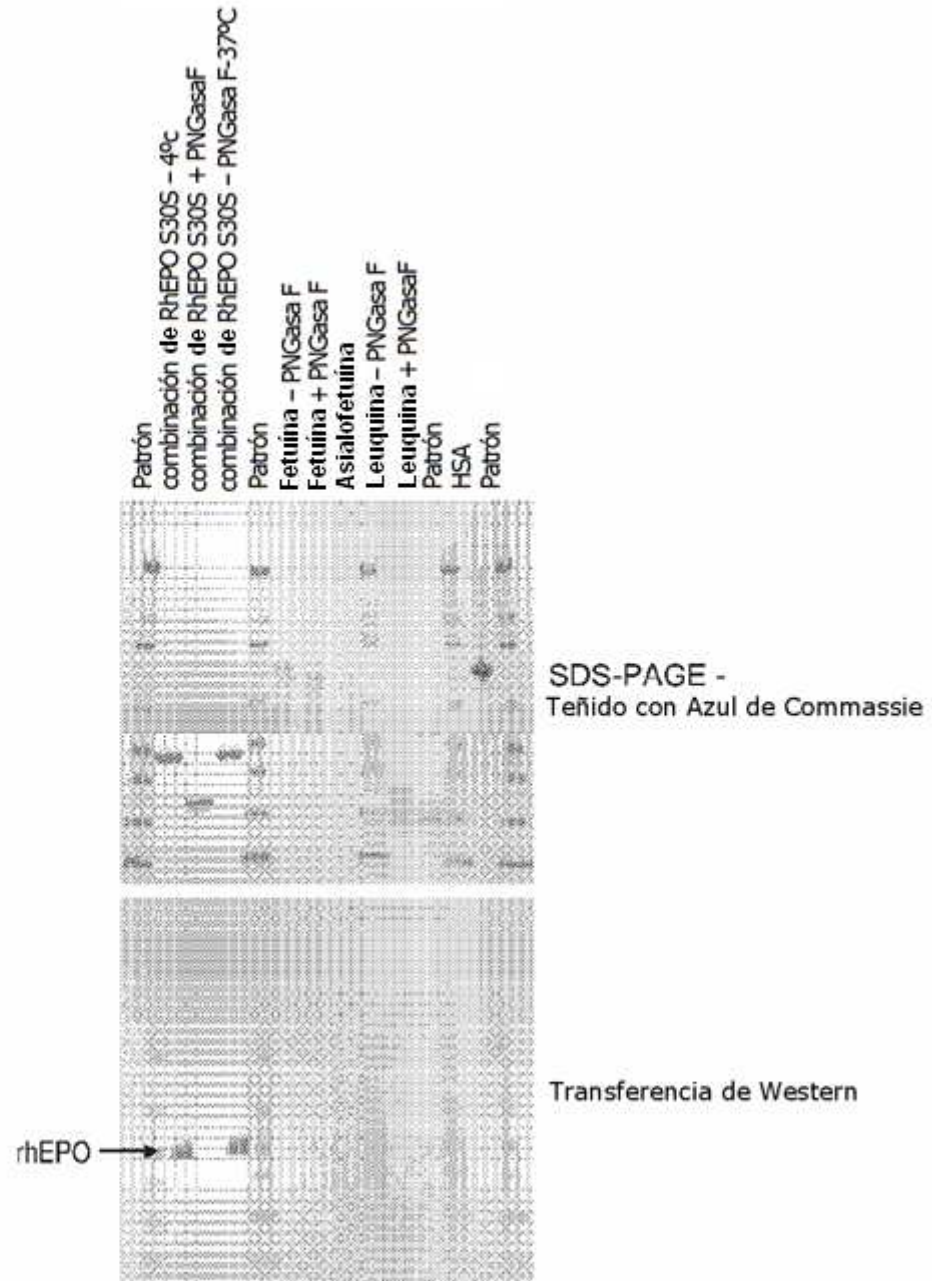
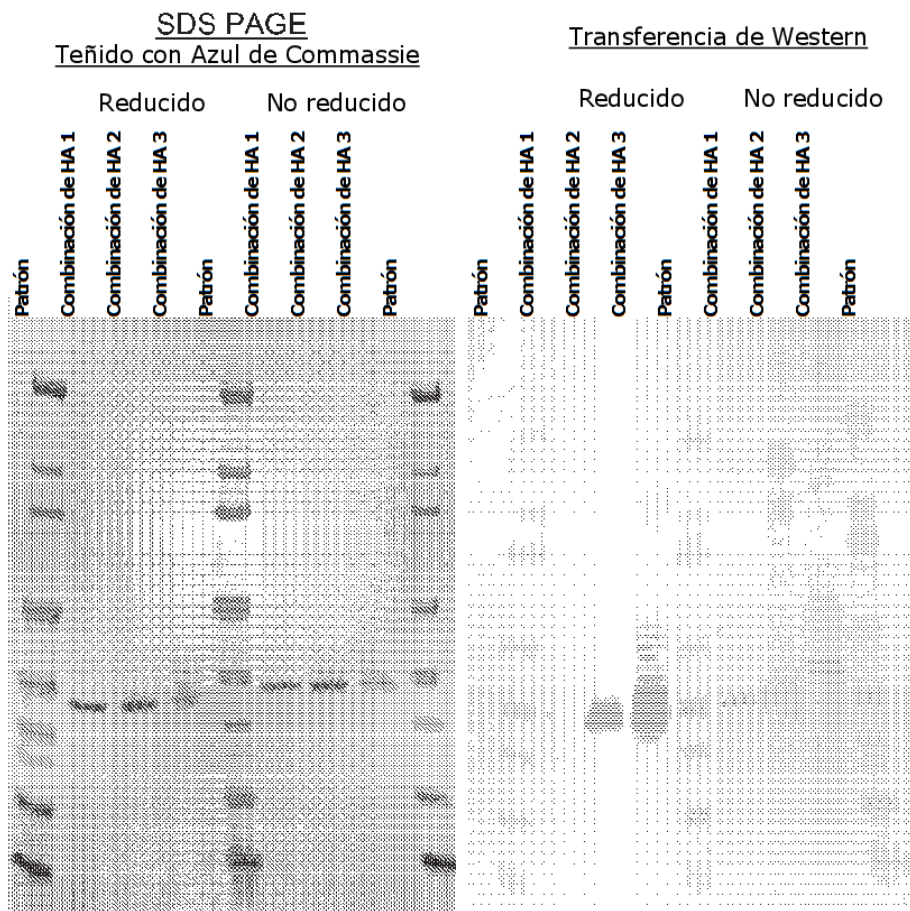


FIG.25

Detección de HCA- rhEPO**Purificación a 40 µm de Mezcla de F071401 - Hydroxi apatita de Tipo I****HPLC para Análisis de N-Glicanos**

combinaciones	% neutro	% mono sialilado	% híbrido sialilado	% bi sialilado
combinación de HA 1	3,49%	4,90%	1,21%	90,40%
combinación de HA 2	13,51%	15,76%	5,86%	64,87%
combinación de HA 3	46,79%	11,50%	10,33%	31,38%

FIG.26

Cromatograma de purificación de Q sepharose FF

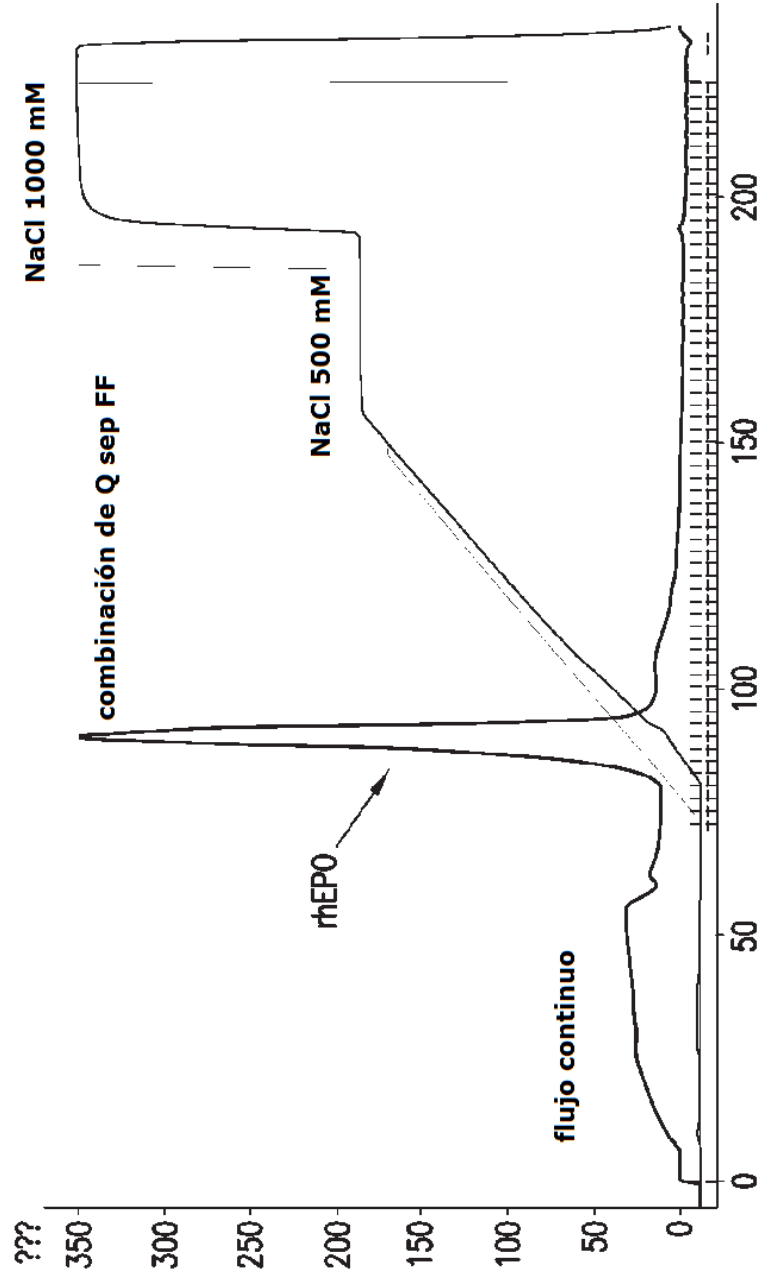
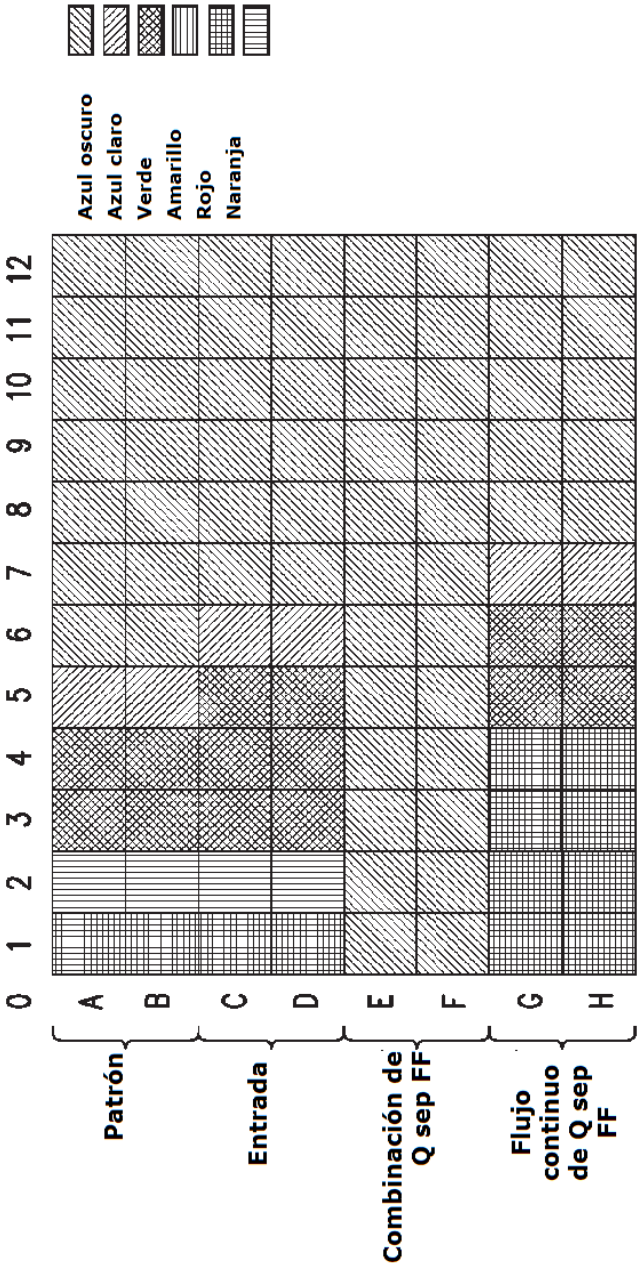


FIG.27A

ELISA de Sándwich para actividad de CBA asociada con rhEPO



Diluciones en placa anti-sueros

	1	2	3	4	5	6	7	8	9	10	11	12
A	1:800	1:1600	1:3200	1:6400	1:12800	1:25600	1:51200	1:102400	1:204800	1:409600	1:819200	Control NEG.
B	1:800	1:1600	1:3200	1:6400	1:12800	1:25600	1:51200	1:102400	1:204800	1:409600	1:819200	Control NEG.

FIG. 27B

Semana 44 - rhEPO - Sepharose 6 FF Azul
Combinación de Azul Azul +/- PNGasa F

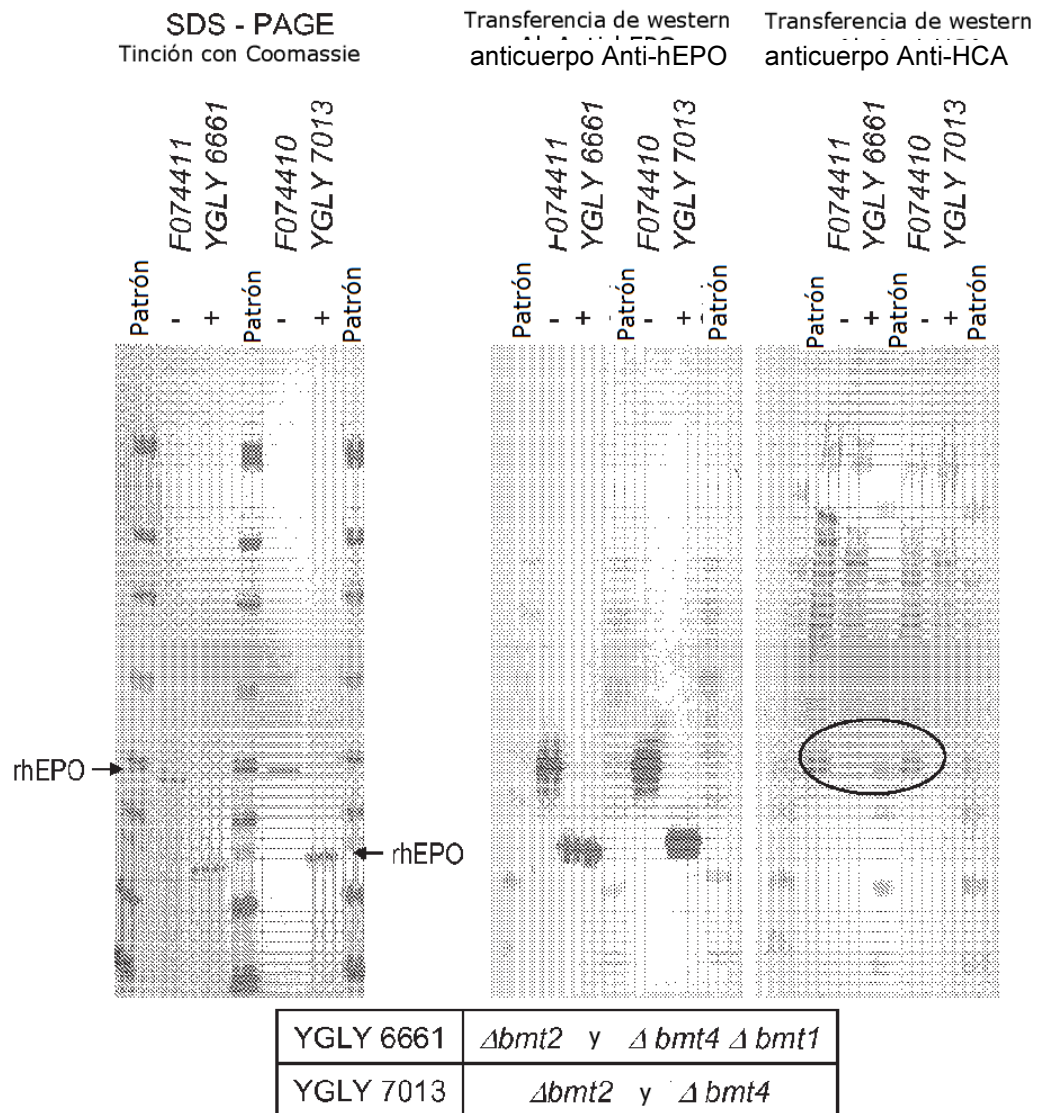
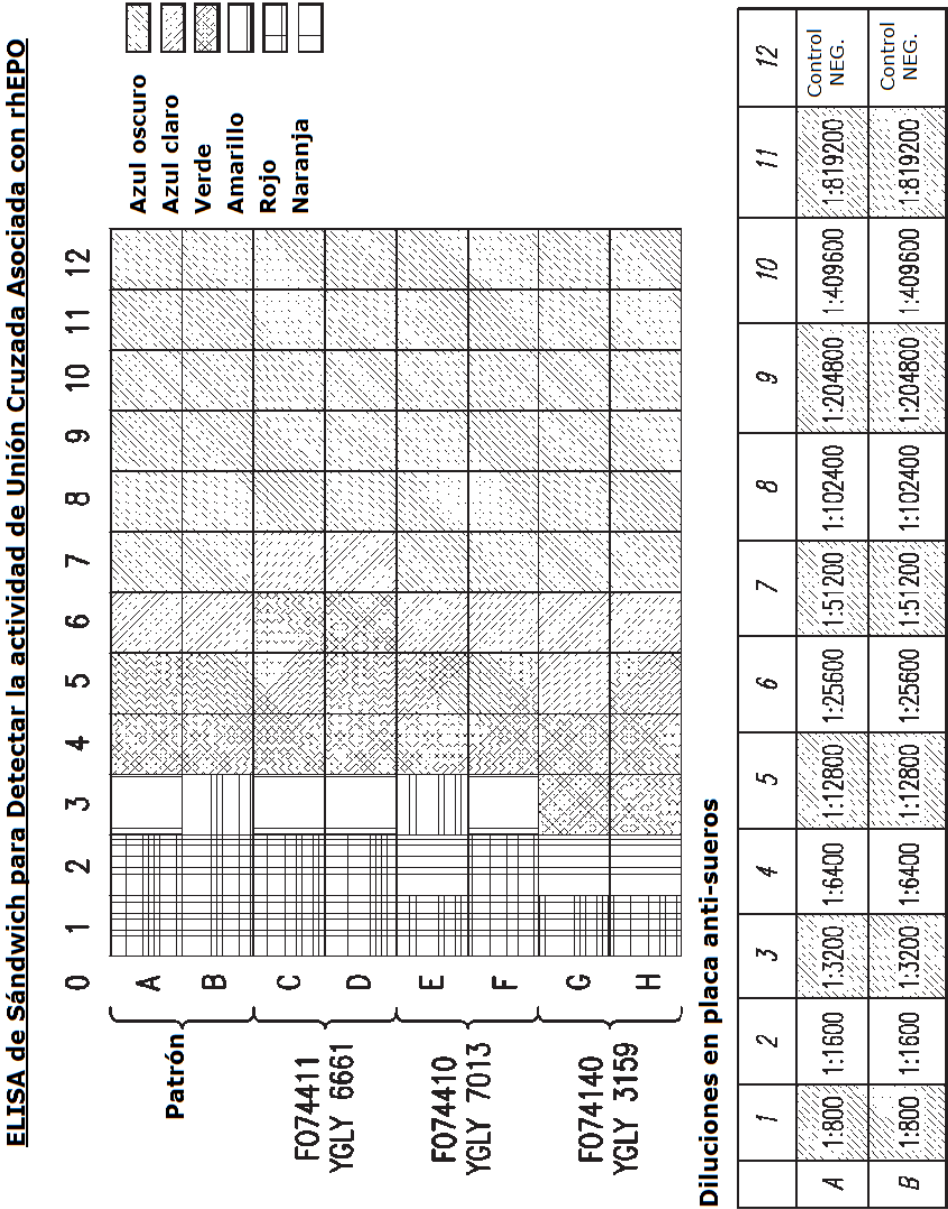


FIG.28



▷YGLY 6661 y 7013 presentaban actividad de Unión Cruzada asociada con rhEPO

FIG.29

Semana 45 - rhEPO - Sepharose 6 FF Azul
Combinación de Azul +/- PNGasa F

SDS - PAGE
Tinción con Coomassie

Transferencia de Western
anticuerpo Anti-hEPO

Transferencia de Western
anticuerpo Anti-HCA

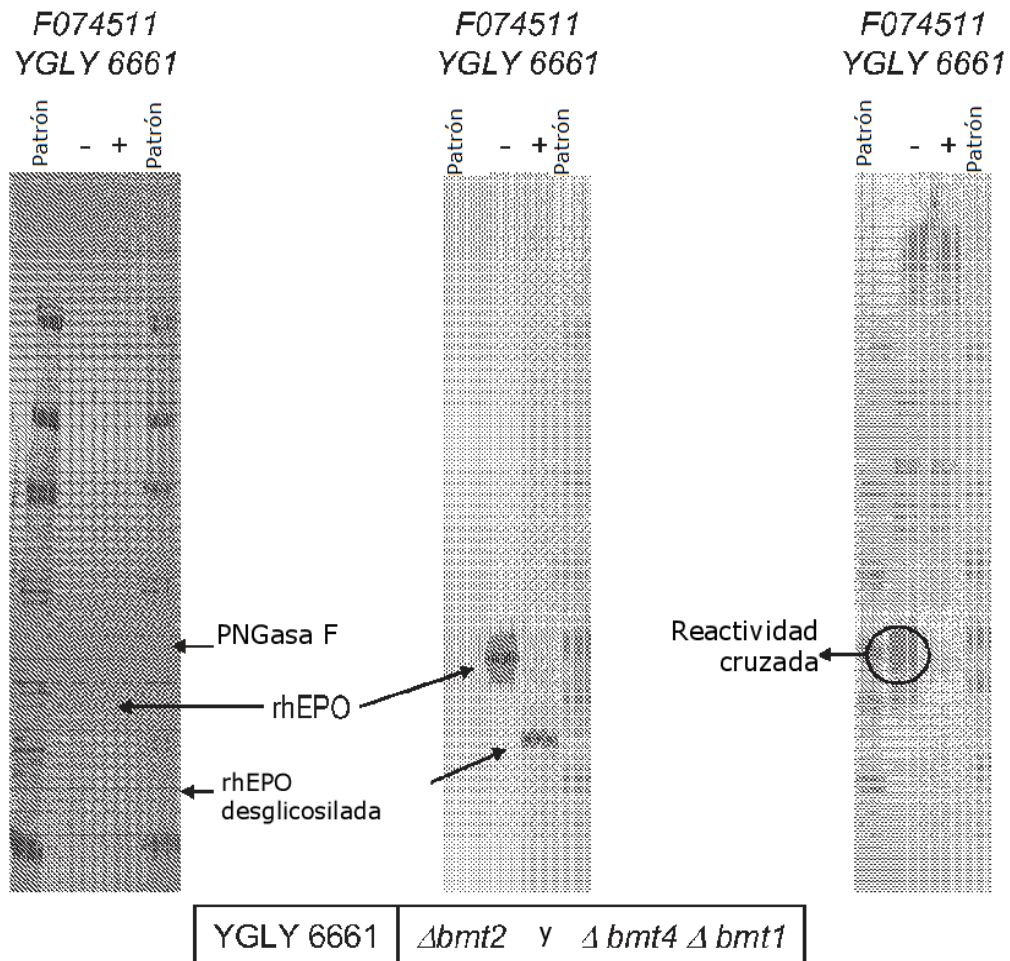


FIG.31

**Semana 46 a - Combinaciones de SixFors -
Supresión genética de *BMT* Cuádruple - Azul**

Tratamiento con +/- PNGasa F

teñido con Azul de Coomassie por SDS-PAGE

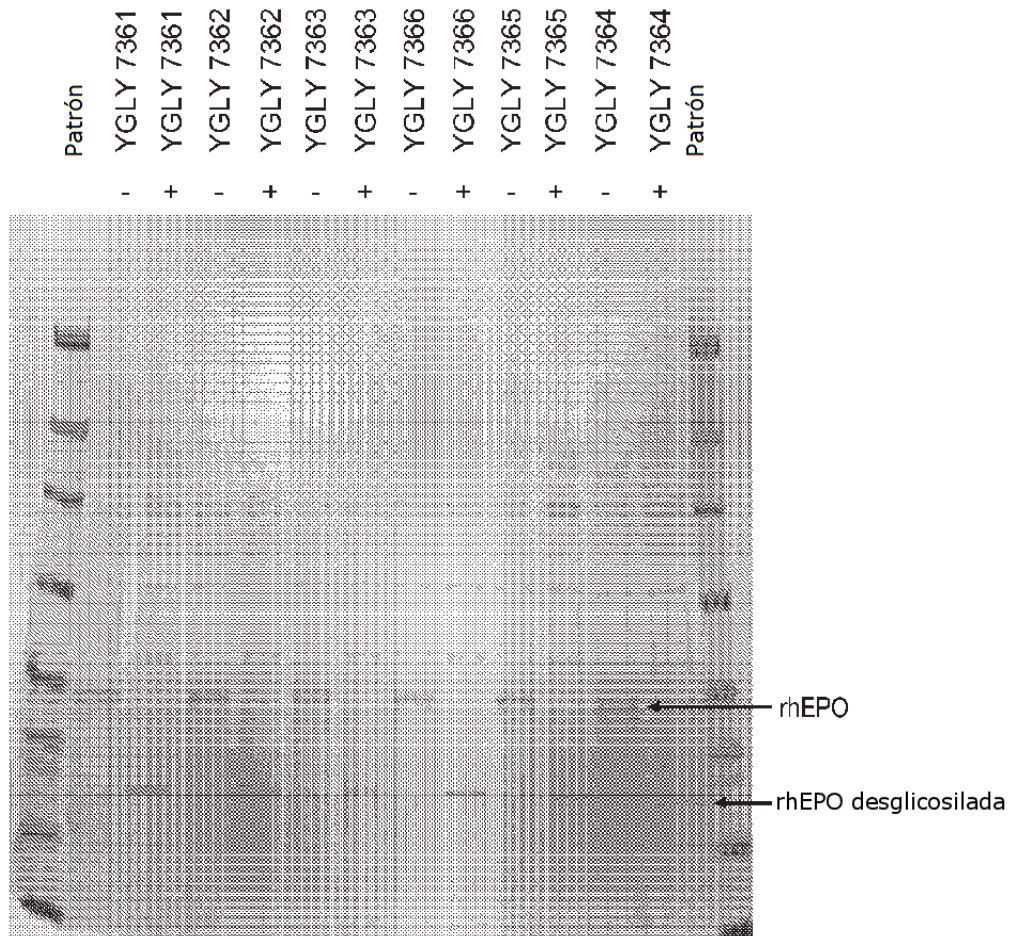


FIG.32A

**Semana 46 b - Combinaciones de SixFors -
Supresión genética de *BMT* Cuádruple - Azul**

Tratamiento con +/- PNGasa F

teñido con Azul de Coomassie por SDS-PAGE

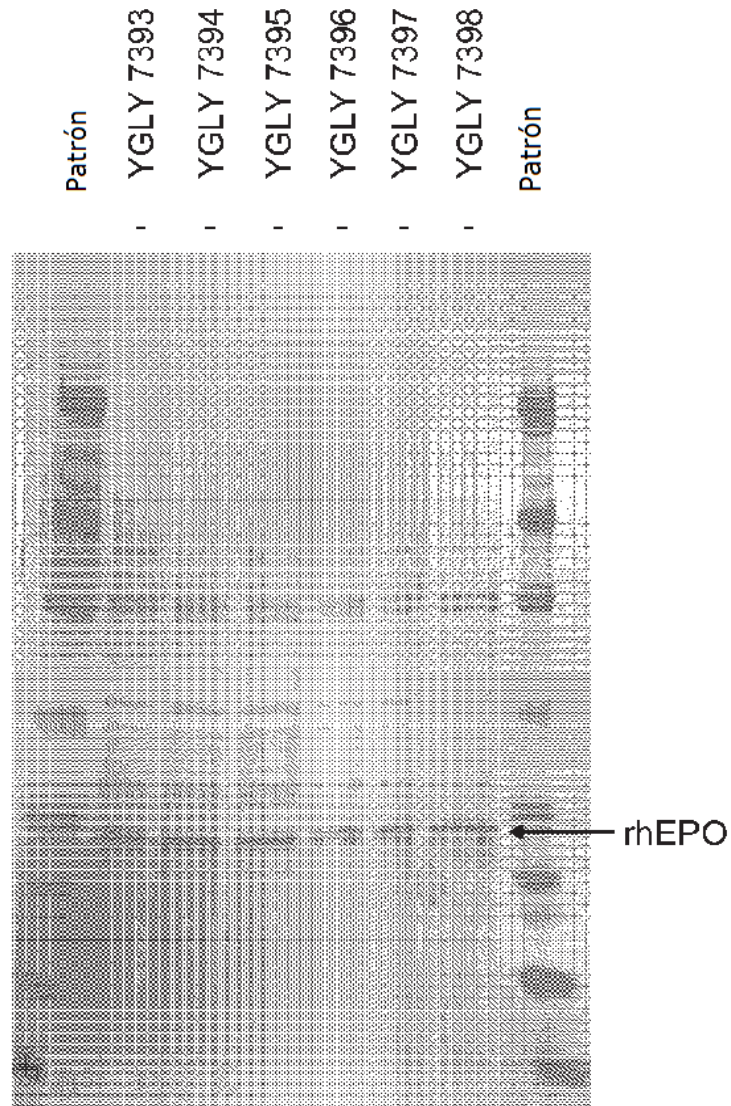
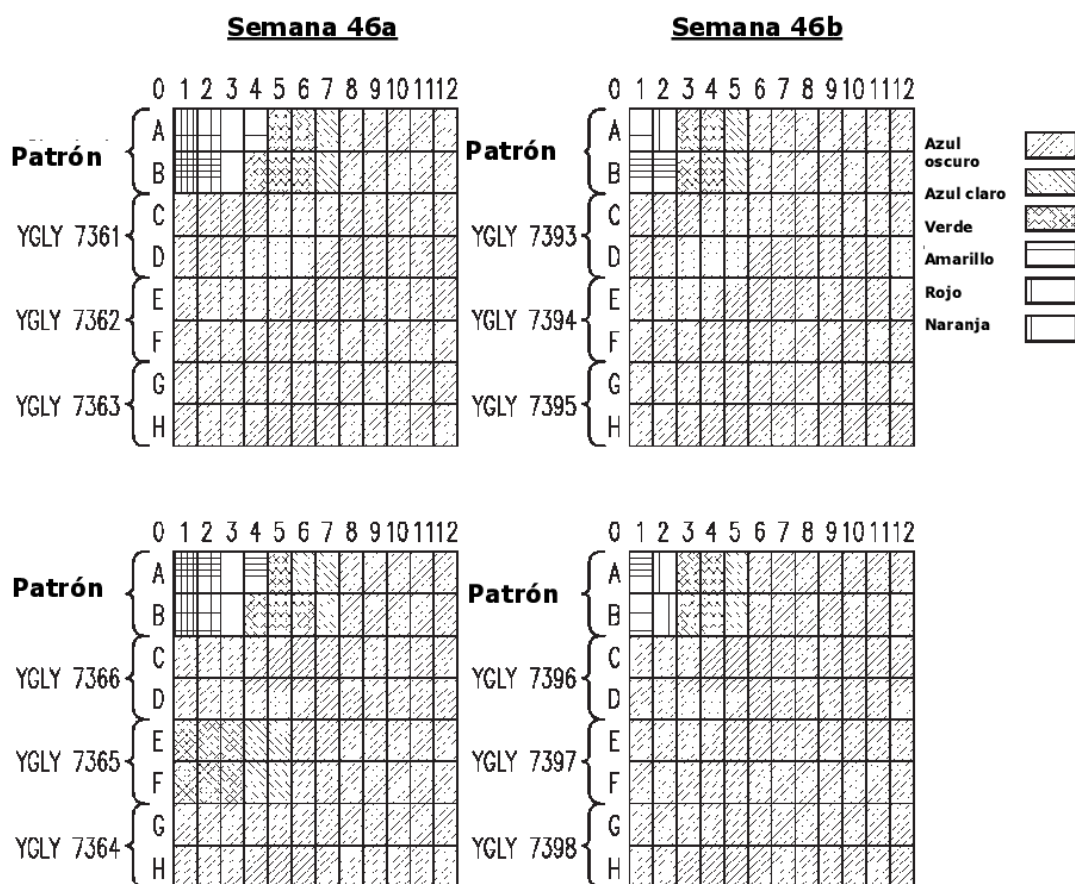


FIG.32B

ELISA de Sándwich para Detectar la actividad de Unión Cruzada Asociada con rhEPO**Semana 46 - Combinaciones de SixFors - Supresión genética de *BMT* Cuádruple - Azul****Diluciones en placa anti-sueros**

	1	2	3	4	5	6	7	8	9	10	11	12
A	1:800	1:1600	1:3200	1:6400	1:12800	1:25600	1:51200	1:102400	1:204800	1:409600	1:819200	Control NEG.
B	1:800	1:1600	1:3200	1:6400	1:12800	1:25600	1:51200	1:102400	1:204800	1:409600	1:819200	Control NEG.

FIG.33

Resultados de HPLC de N-Glicanos

Semanas 46 a y b - Combinaciones de SixFors - Supresión genética de *BMT*
Cuádruple - Azul

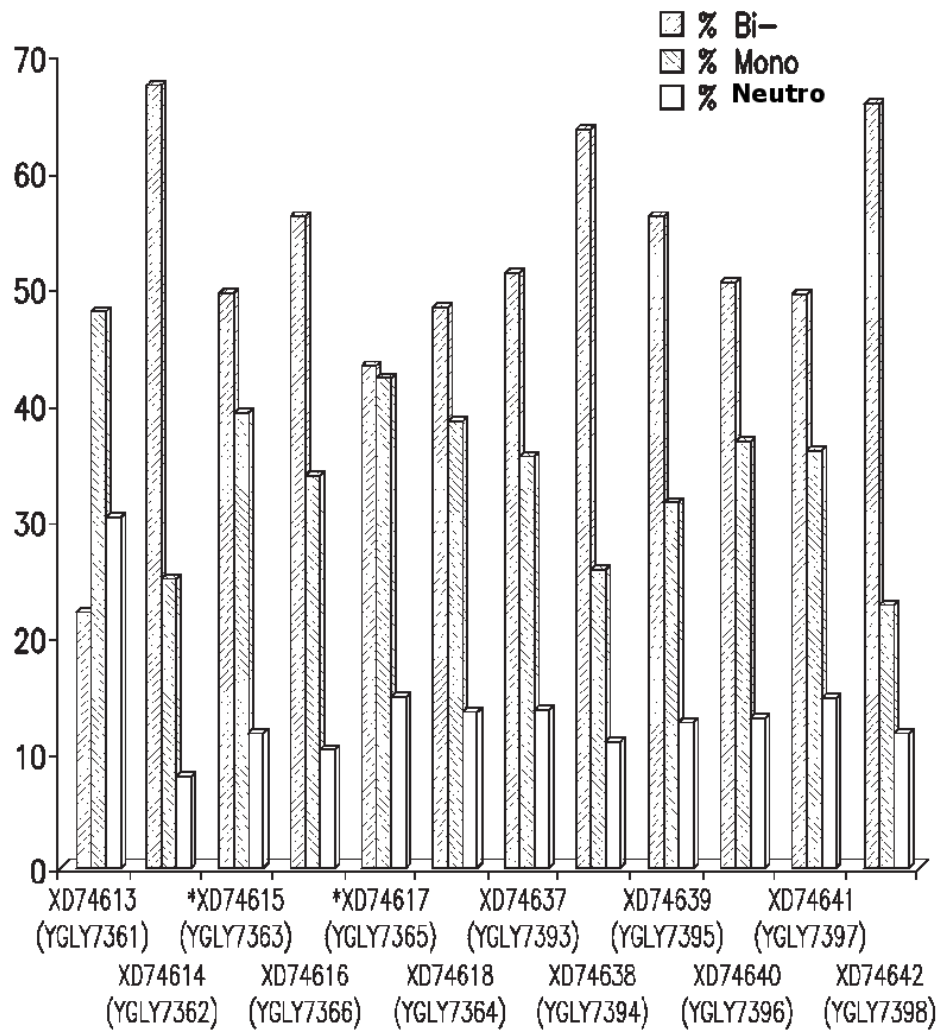


FIG.34

Semana 50 - identificación sistemática de cepas de supresión genética de BMT cuádruple para combinaciones de rhEPO azul y la combinación HA 1s

Tinción con Coomassie por SDS-PAGE

Patrón	YGLY 7362	YGLY 7366	YGLY 7396	YGLY 3159	YGLY 7398	YGLY 7398	Patrón	YGLY 7362	YGLY 7366	YGLY 7396	YGLY 3159	YGLY 7398	YGLY 7398	Patrón
	Combinaciones de Azul							Combinaciones de HA						

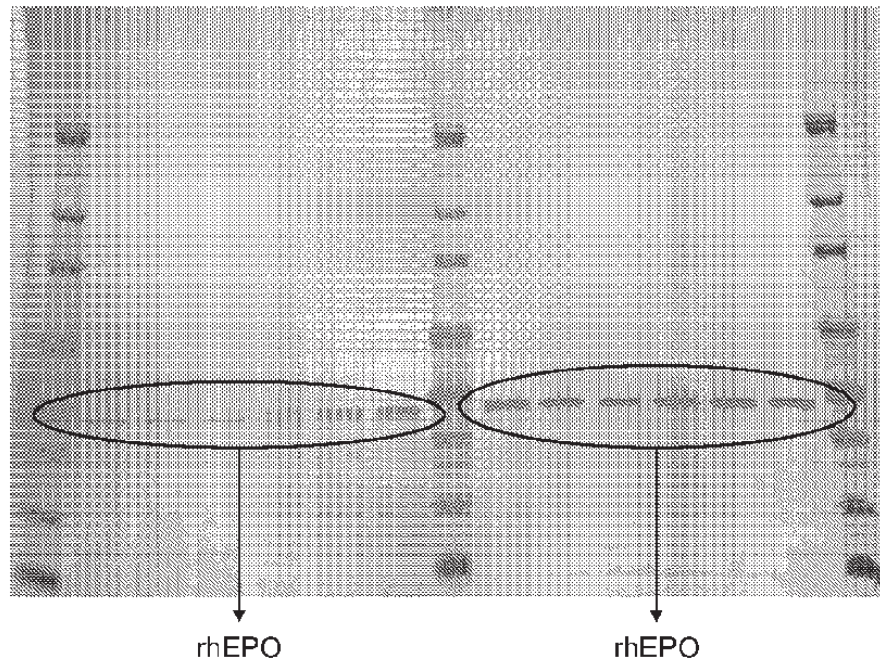


FIG.35A

**Semana 50 – identificación sistemática de cepas de
supresión genética de BMT cuádruple para combinaciones
de rhEPO azul y la combinación HA 1s**

Transferencia de Western - Ab anti-HCA

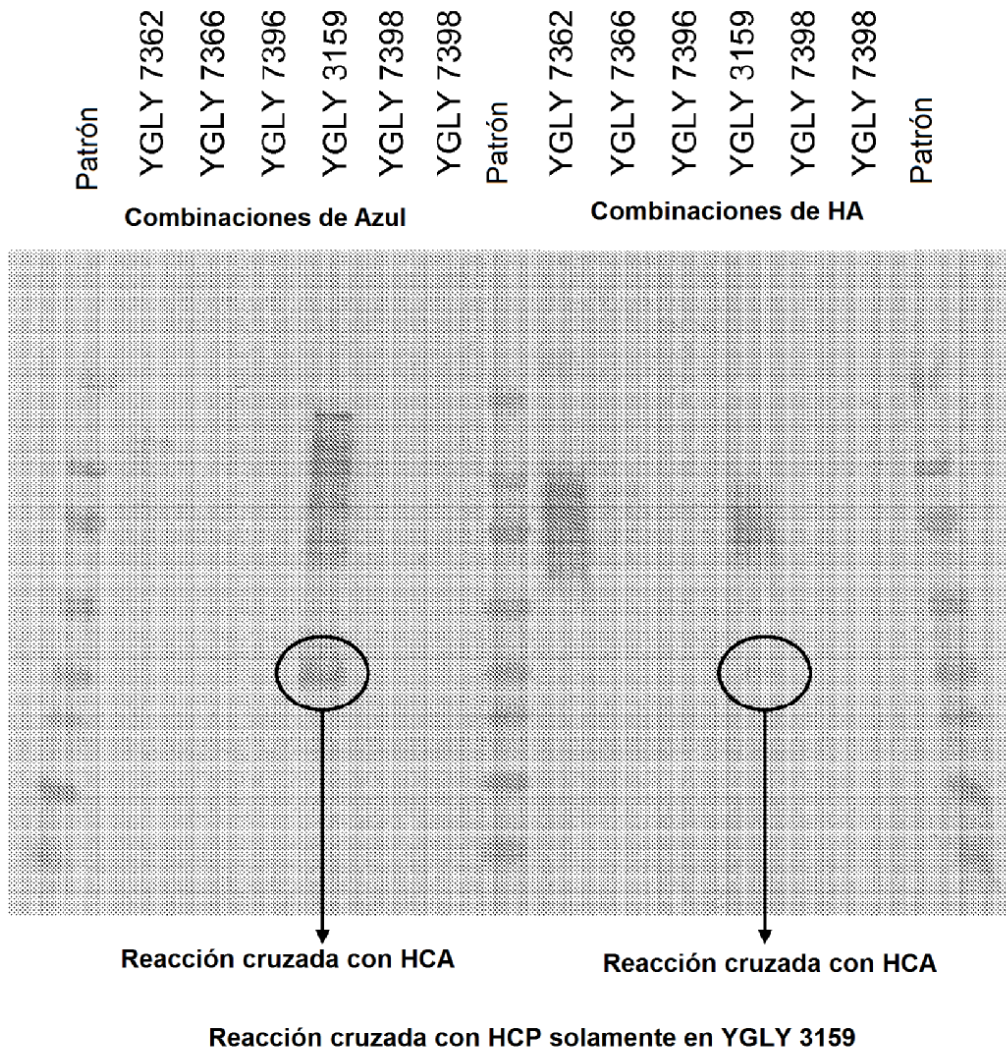


FIG.35B

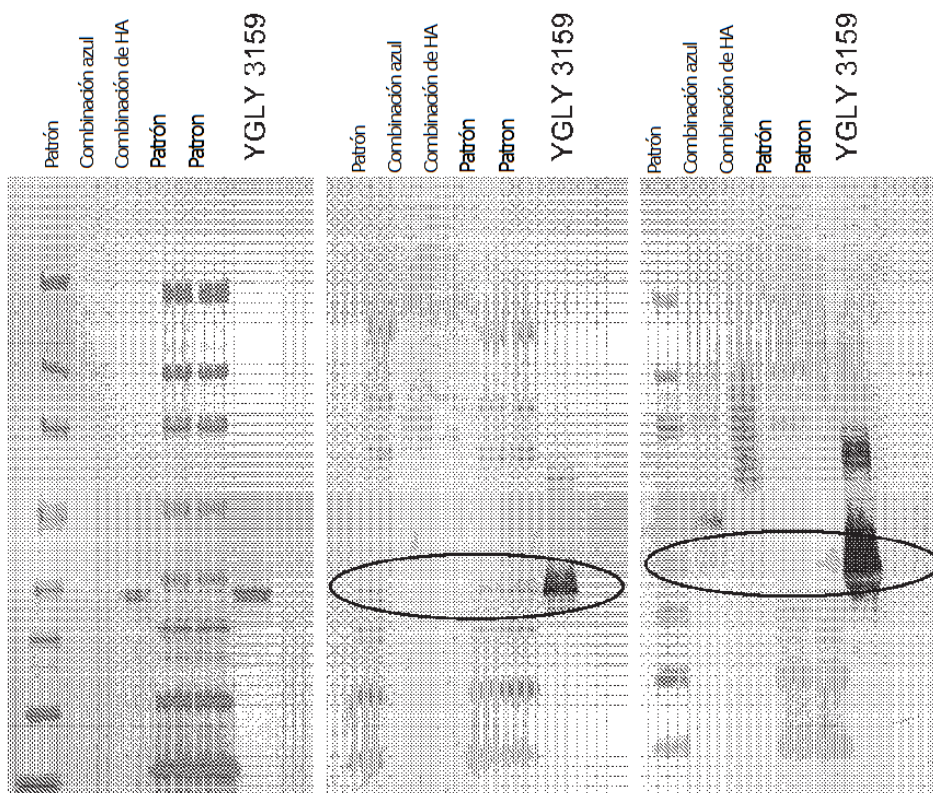
**Identificación sistemática de cepas de supresión genética de *BMT* cuádruple
teñidas con Azul de Coomassie en Gel por SDS - PAGE**

F08Q240 - YGLY 7398 - combinación de Azul/combinación de HA 1

SDS - PAGE
Tinción con Coomassie

Transferencia de Western
anticuerpo Anti-HCA
1:3000

Transferencia de Western
anticuerpo Anti-HCA
1:2000

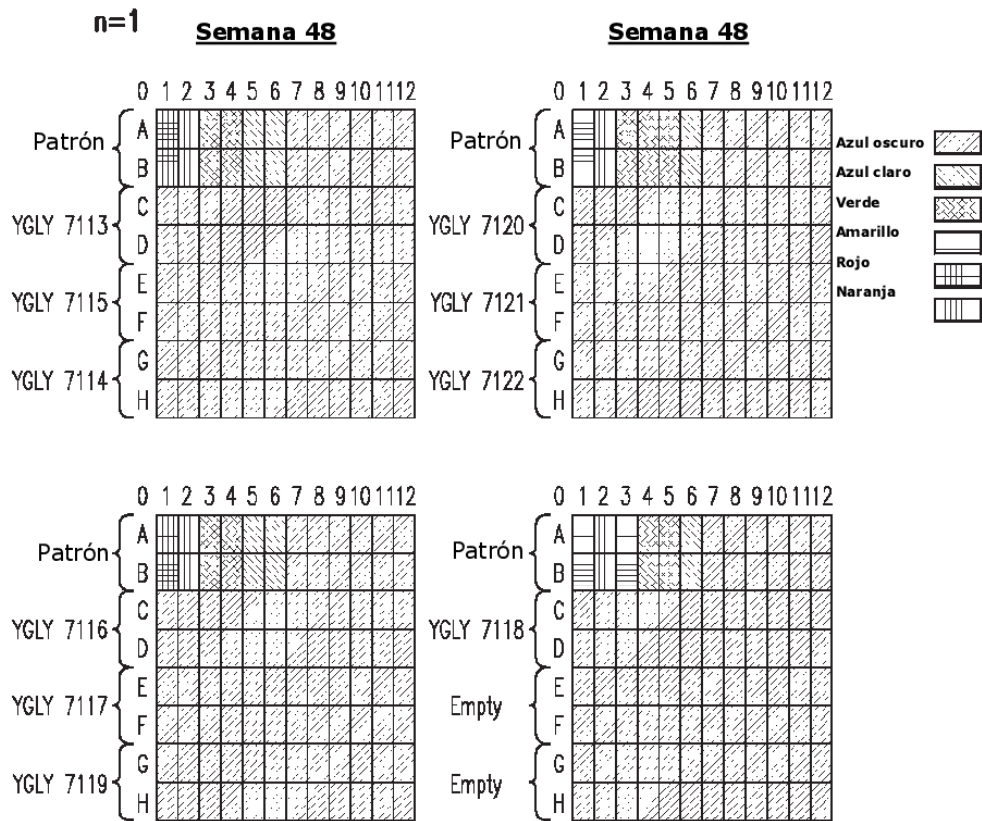


*GiF2 conejo policlonal::6316

FIG.36

ELISA de Sándwich para Detectar la actividad de Unión Cruzada Asociada con rhEPO

Combinaciones de SixFors - Supresión genética de *BMT* Cuádruple - Azul



Diluciones en placa anti-sueros:

	1	2	3	4	5	6	7	8	9	10	11	12
A	1:800	1:1600	1:3200	1:6400	1:12800	1:25600	1:51200	1:102400	1:204800	1:409600	1:819200	Control NEG.
B	1:800	1:1600	1:3200	1:6400	1:12800	1:25600	1:51200	1:102400	1:204800	1:409600	1:819200	Control NEG.

FIG.37

Resultados de HPLC de N-Glicanos

Semana 48 - Combinaciones de SixFors - Supresión genética de *BMT* Cuádruple - Azul

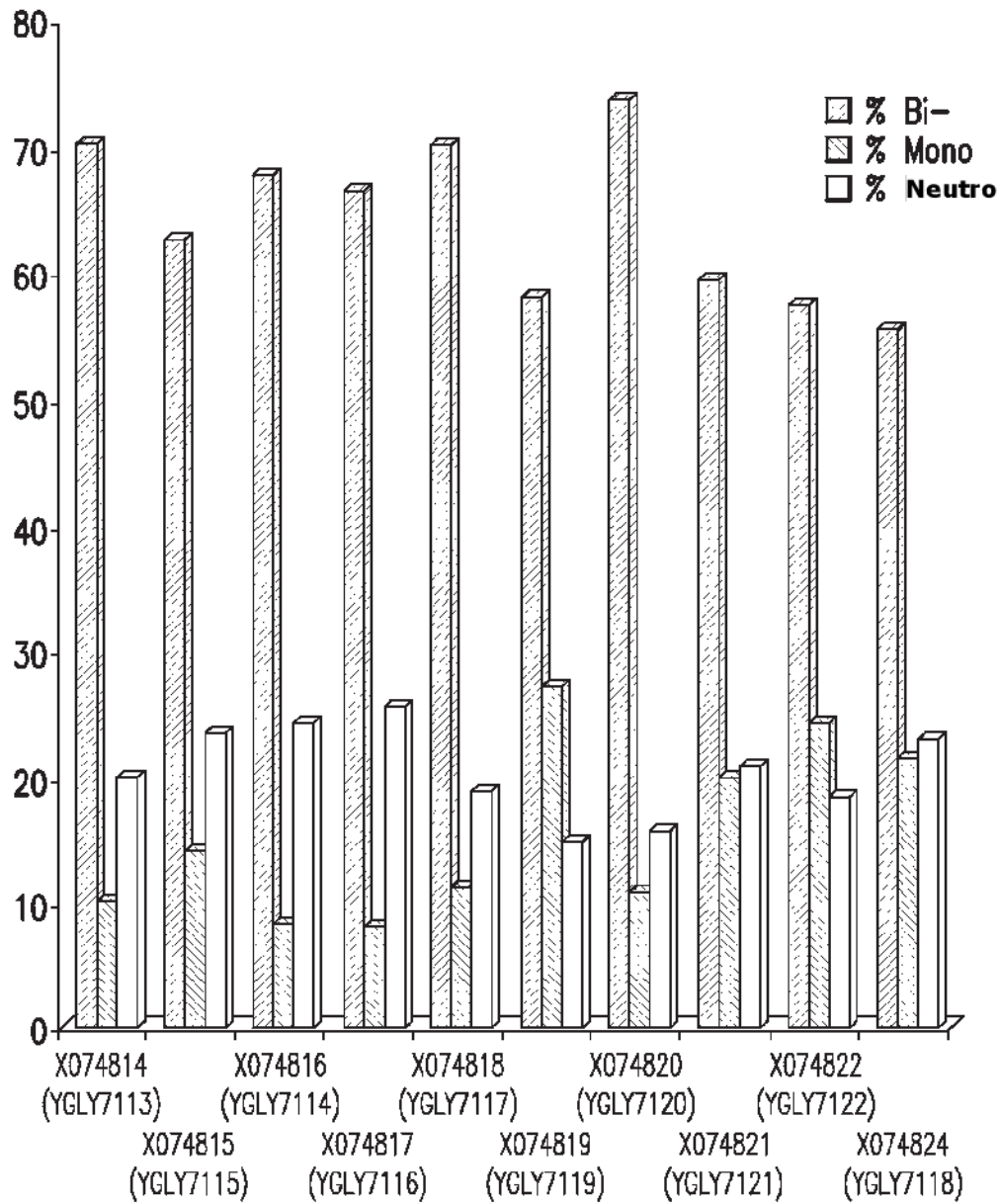


FIG.38

Semana 51 – identificación sistemática de cepas de supresión genética de BMT cuádruple para combinaciones de rhEPO azul y la combinación HA 1s

Tinción con Coomassie por SDS-PAGE

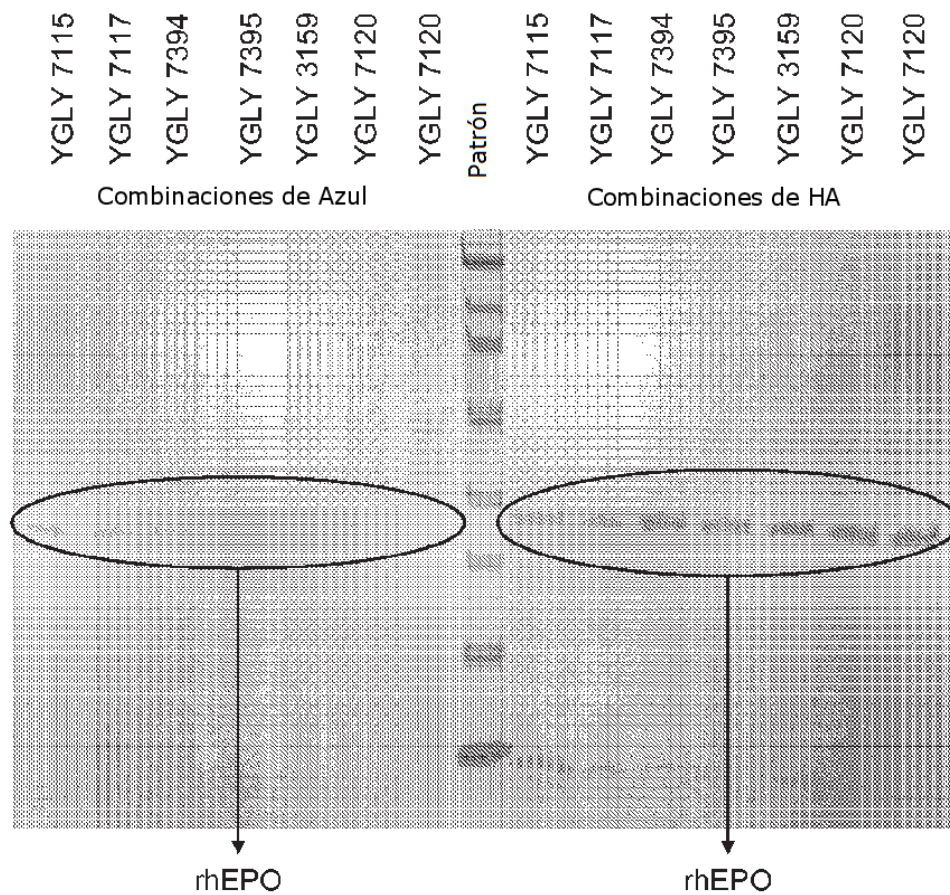


FIG. 39A

Semana 51 - identificación sistemática de cepas de supresión genética de *BMT* cuádruple para combinaciones de rhEPO azul y la combinación HA 1s

Transferencia de Western -anticuerpo Anti-HCA

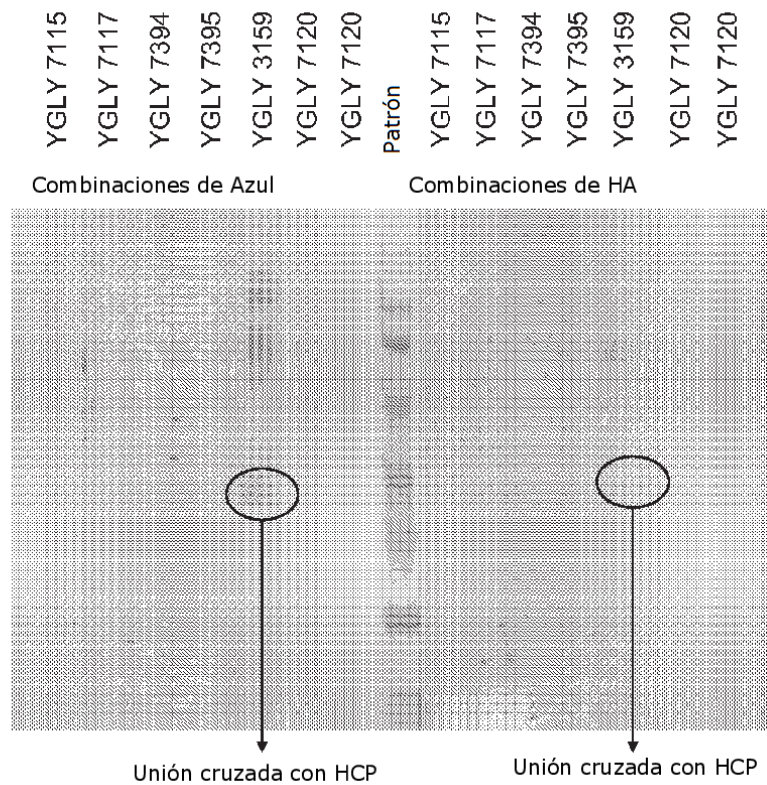


FIG.39B

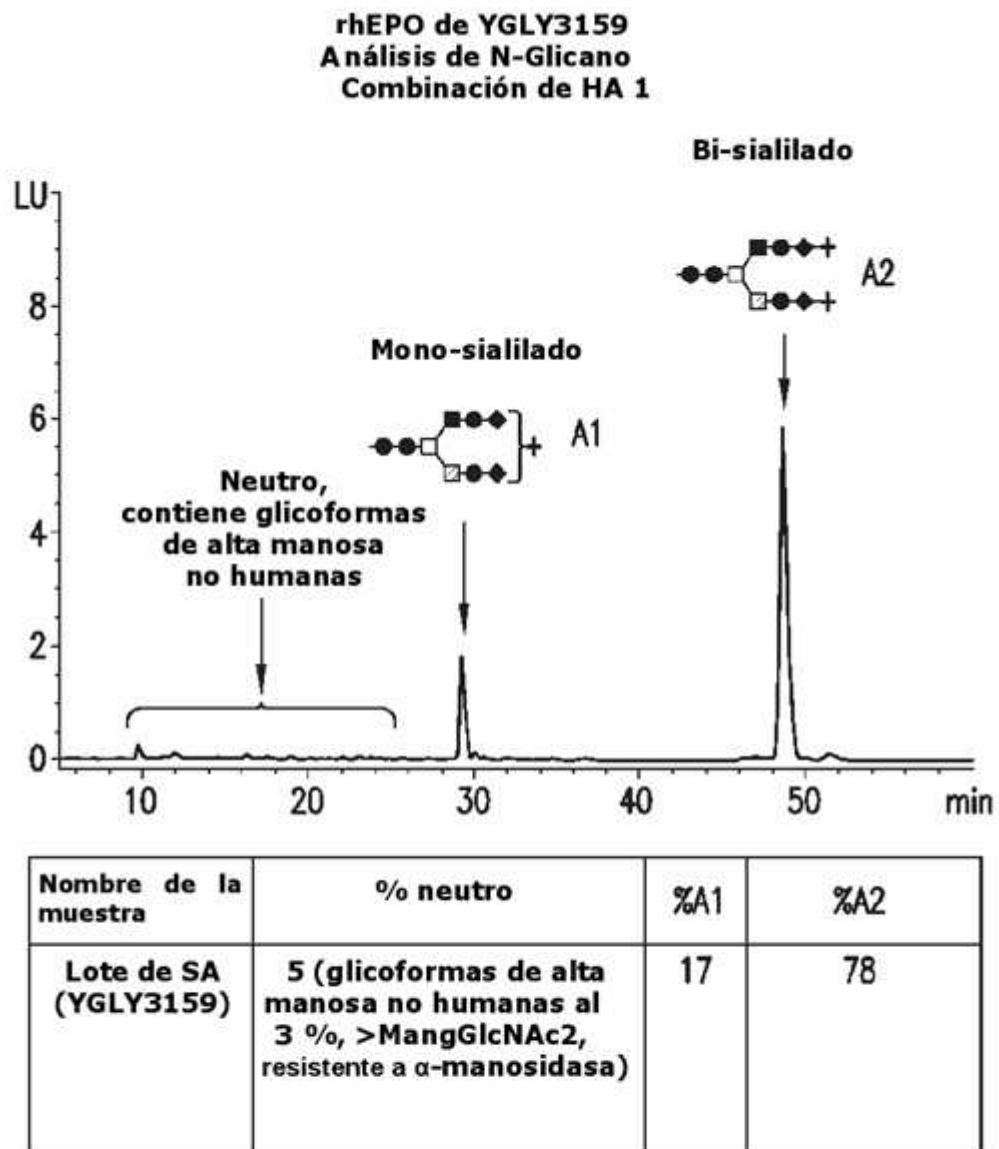


FIG.40A

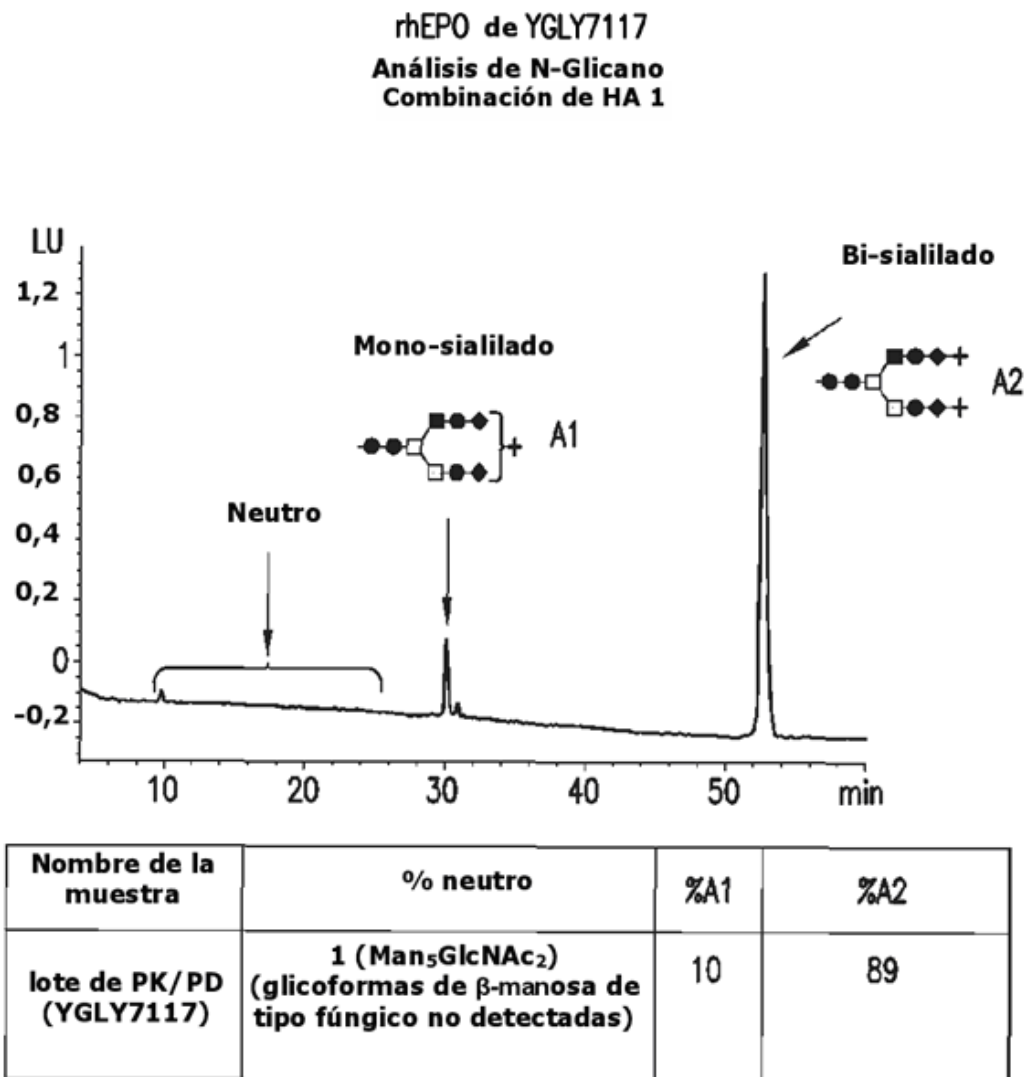


FIG.40B