

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 773 891**

51 Int. Cl.:

C12Q 1/68 (2008.01)

C12N 15/85 (2006.01)

C12N 9/06 (2006.01)

C12N 9/22 (2006.01)

C12N 9/00 (2006.01)

C12N 15/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.06.2012 E 18211367 (0)**

97 Fecha y número de publicación de la concesión europea: **25.12.2019 EP 3489366**

54 Título: **Procedimientos para la producción de líneas celulares de mamífero modificadas con transgenes amplificados**

30 Prioridad:

01.06.2011 US 201161492174 P

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

15.07.2020

73 Titular/es:

**PRECISION BIOSCIENCES, INC. (100.0%)
302 East Pettigrew Street, Dibrell Building, Suite
A-100
Durham, NC 27701, US**

72 Inventor/es:

**JANTZ, DEREK;
SMITH, JAMES, JEFFERSON y
NICHOLSON, MICHAEL, G.**

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 773 891 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

Procedimientos para la producción de líneas celulares de mamífero modificadas con transgenes amplificados

Campo de la invención

5 La invención se refiere al campo de la biología molecular y tecnología de ácido nucleico recombinante. En particular, la invención se refiere a procedimientos de inserción de genes en localizaciones determinadas del ADN cromosómico de líneas celulares de mamífero cultivadas que se someten a amplificación genética. La invención también se refiere a meganucleasas, vectores y líneas celulares modificadas necesarios para llevar a cabo los procedimientos, las líneas celulares resultantes de la aplicación de los procedimientos, y el uso de las líneas celulares para producir productos proteicos de interés.

Antecedentes de la invención

10 Las proteínas terapéuticas son el director primario del crecimiento en el mercado farmacéutico global (Kresse, Eur J Pharm Biopharm 72, 479 (2009)). En 2001, los biofarmacéuticos suponían 24,3 mil millones de \$ en ventas. En 2007, este número se había más que doblado a 54,4 mil millones de \$. Se estima actualmente que alcanza los 78 mil millones de \$ en 2012 (Pickering, Spectrum Pharmaceutical Industry Dynamics Report, Decision Resources, Inc., 5 (2008)).
15 Esto incluye las ventas de fármacos más "populares" tales como la eritropoyetina, activador del plasminógeno tisular, e interferón, así como numerosos fármacos "especializados" tales como terapias de reposición de enzimas para trastornos de almacenamiento lisosómico. El crecimiento del tamaño de mercado no es paralelo, sin embargo, está dirigido principalmente por la demanda meteórica de anticuerpos monoclonales completamente humanos y humanizados (Reichert, Curr Pharm Biotechnol 9, 423 (2008)). Debido a que tienen la capacidad de conferir un
20 espectro virtualmente ilimitado de actividades biológicas, los anticuerpos monoclonales se están convirtiendo rápidamente en la clase de agentes terapéuticos más potentes disponibles para los médicos. No es sorprendente que más del 25 % de las moléculas que actualmente se someten a ensayos clínicos en los Estados Unidos y Europa son anticuerpos monoclonales (Reichert, Curr Pharm Biotechnol 9, 423 (2008)).

25 A diferencia de los agentes farmacéuticos más tradicionales, las proteínas terapéuticas se producen en células vivas. Esto complica mucho el procedimiento de fabricación e introduce una heterogeneidad significativa en las formulaciones del producto (Field, Recombinant Human IgG Production from Myeloma and Chinese Hamster Ovary Cells, in Cell Culture and Upstream Processing, Butler, ed., (Taylor y Francis Group, New York, 2007)). Además, los fármacos proteicos se necesitan a dosis inusualmente altas, que necesitan procedimientos de fabricación altamente escalables y hace que los costes de fabricación sean un determinante del precio principal. Por estas razones, el tratamiento con un anticuerpo terapéutico típico (por ejemplo, el monoclonal anti-HER2-neu Herceptin® cuesta de 60.000\$- 80.000\$ para un curso completo de tratamiento (Fleck, Hastings Center Report 36, 12 (2006)). Para complicar más la economía de la producción biofarmacéutica está el hecho de que muchos de los biofarmacéuticos populares anteriores están fuera de patente (o estarán fuera de patente pronto) y los gobiernos de EE. UU. y EU esperan optimizar el procedimiento regulador de aprobación para agentes terapéuticos "biogénicos" y "biosimilares" (Kresse, Eur J Pharm Biopharm 72, 479 (2009)). Estos factores deberían dar lugar a un aumento significativo de la competición por las
35 ventas de muchos biofarmacéuticos destacados (Pickering, Spectrum Pharmaceutical Industry Dynamics Report, Decision Resources, Inc., 5 (2008)). Por lo tanto, existe un enorme interés en tecnologías que reduzcan los costes de fabricación de proteínas terapéuticas (Seth y col., Curr Opin Biotechnol 18, 557 (2007)).

40 Muchas de las proteínas farmacéuticas del mercado son glicoproteínas que no se pueden producir fácilmente en sistemas biológicos fáciles de manipular tales como bacterias o levaduras. Por esta razón, las proteínas terapéuticas recombinantes se producen casi exclusivamente en líneas celulares de mamífero, principalmente de ovario de hámster chino (por ejemplo, CHO-K1), mieloma de ratón (por ejemplo, NS0), riñón de hámster recién nacido (BHK), C127 murinas, riñón embrionario humano (por ejemplo, HEK 293), o células derivadas de la retina humana (por ejemplo, PER-C6) (Andersen y Krummen, Curr Opin Biotechnol 13, 117 (2002)). De estas, las células CHO son con mucho, la
45 plataforma más común para la bioproducción debido a que ofrecen la mejor combinación de altos niveles de expresión proteica, poco tiempo de doblaje, tolerancia a un amplio intervalo de condiciones de medio, protocolos de transfección amplificados establecidos, una incapacidad para propagar la mayoría de patógenos humanos, insuficiencia de bloquear la propiedad intelectual, y el mayor récord de recorrido de aprobación por la FDA (Field, Recombinant Human IgG Production from Myeloma and Chinese Hamster Ovary Cells, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)).

50 Los productos farmacéuticos para el gran mercado se producen normalmente en enormes depósitos biorreactores con agitado que contienen cientos de litros de células CHO que expresan establemente el producto proteico de interés (Chu y Robinson, Curr Opin Biotechnol 12, 180 (2001), Coco-Martin y Harmsen, Bioprocess International 6, 28 (2008)). En condiciones industriales optimizadas, dichos procedimientos de fabricación pueden dar lugar más de 5 g de proteína por litro de células por día (Coco-Martin y Harmsen, Bioprocess International 6, 28 (2008)). Debido al gran número de células implicadas (~ 50 mil millones de células por litro), el nivel de expresión proteico por célula tiene un efecto muy drástico sobre el resultado. Por esta razón, todas las células implicadas en la producción de un producto biofarmacéutico en particular se deben derivar de un único clon "de alta producción", cuya producción constituye una de las etapas más exigentes en tiempo y recursos del procedimiento de fabricación (Clarke y Compton, Bioprocess

International 6, 24 (2008)).

La primera etapa en la fabricación a gran escala de un producto biofarmacéutico es la transfección de células de mamífero con un ADN plasmídico que codifique el producto proteico de interés bajo el control de un promotor constitutivo fuerte. Los transfectantes estables se seleccionan utilizando como marcador un gen de selección también albergado en el plásmido. Más frecuentemente, este marcador es un gen de la dihidrofolato reductasa (DHFR), que cuando se transfecta en una línea celular deficiente en DHFR tal como DG44 permite la selección de transfectantes estables utilizando un medio deficiente en hipoxantina. La razón primaria para utilizar la DHFR como marcador genético es que hace posible un proceso denominado “amplificación genética”. Cultivando los transfectantes estables en concentraciones gradualmente crecientes de metotrexato (MTX), un inhibidor de la DHFR, es posible amplificar el número de copias del gen DHFR presente en el genoma. Debido a que el gen que codifica el producto proteico de interés está acoplado físicamente al gen DHFR, esto resulta en la amplificación de ambos genes con un aumento concomitante del nivel de expresión de la proteína terapéutica (Butler, Cell Line Development for Culture Strategies: Future Prospects to Improve Yields, in Cell Culture and Upstream Processing, Butler, ed., (Taylor y Francis Group, New York, 2007)). Los sistemas relacionados para la creación de líneas de bioproducción estables utiliza los genes de la glutamina sintetasa (GS) o la hipoxantina fosforribosil transferasa (HPRT) como marcadores genéticos y necesitan el uso de líneas celulares deficientes en GS o HPRT como huéspedes de la transfección (Clarke y Compton, Bioprocess International 6, 24 (2008)). En el caso del sistema HPRT, la amplificación genética se consigue cultivando las células en medio HAT, que contiene aminopterina, hipoxantina, y timidina (Kellems, ed. Gene amplification in mammalian cells: a comprehensive guide, Marcel Dekker, New York, 1993).

En todos estos sistemas el ADN plasmídico inicial comprende un casete de expresión del gen bioterapéutico y un marcador genético integrado en una localización aleatoria del genoma, dando como resultado una extrema variabilidad en la expresión de la proteína terapéutica de un transfectante estable a otro (Collingwood y Urnov, Targeted Gene Insertion to Enhance Protein Production from Cell Lines, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Por esta razón, es necesario explorar cientos o miles de transfectantes iniciales para identificar las células que expresan aceptablemente altos niveles de producto genético sea antes y después de la amplificación genética (Butler, Cell Line Development for Culture Strategies: Future Prospects to Improve Yields, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Una segunda consecuencia y más problemática de la integración genética aleatoria es el fenómeno del silenciamiento del transgén, en el que la expresión de la proteína recombinante se enlentece o cesa completamente con el tiempo (Collingwood y Urnov, Targeted Gene Insertion to Enhance Protein Production from Cell Lines, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Debido a que estos efectos a menudo no se manifiestan durante semanas o meses después de la transfección inicial y el proceso de exploración, generalmente es necesario llevar a cabo y expandir docenas de líneas clónicas independientes para identificar una que exprese la proteína de interés constantemente a lo largo del tiempo (Butler, Cell Line Development for Culture Strategies: Future Prospects to Improve Yields, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)).

Este gran número de etapas de exploración y expansión da como resultado un procedimiento muy largo y caro para generar simplemente la línea celular que, en último término, producirá el producto terapéutico de interés. Además, utilizando procedimientos convencionales, se necesita un mínimo de 10 meses (con una media de 18 meses) y un adelanto en investigación de decenas de millones de dólares en trabajo y material para producir un agrupamiento inicial de células que expresan la proteína adecuado para la fabricación industrial (Butler, Cell Line Development for Culture Strategies: Future Prospects to Improve Yields, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Si se tiene en cuenta la pérdida de tiempo en el mercado de una proteína terapéutica popular, las ineficacias en la producción de la línea celular pueden costar a los fabricantes biofarmacéuticos cientos de millones de dólares (Seth y col., Curr Opin Biotechnol 18, 557 (2007)).

Mucho del tiempo y gasto de la creación de la línea celular de bioproducción se puede atribuir a la integración genómica aleatoria del gen del bioproducto que da como resultado una variabilidad de clon a clon en el genotipo, por lo tanto, una variabilidad en la expresión genética. Una manera de superar esto es dirigir la integración genética en una localización determinada que se sepa que soporta un alto nivel de expresión genética. En este extremo, se han descrito varios sistemas que utilizan la Cree, F1p o ΦC31 recombinasas para dirigir la inserción de un gen de bioproducto (revisado en Collingwood y Urnov, Targeted Gene Insertion to Enhance Protein Production from Cell Lines, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Recientes realizaciones de estos sistemas, más notablemente el sistema F1p-in® comercializado por Invitrogen Corp. (Carlsbad, CA), acoplan la integración del gen del bioproducto con la reconstitución de un indicador genético de división de manera que se pueden seleccionar las células con los genes dirigidos correctamente. Como se esperaba, estos sistemas dan como resultado una heterogeneidad muy reducida de la expresión genética, y en algunos casos, los transfectantes estables individuales se pueden agrupar, obviando el tiempo y gasto asociado con la expansión de un único clon.

El principal inconveniente de los sistemas de direccionamiento genético basados en recombinasa es que los sitios de reconocimiento de recombinasa (sitios loxP, FRT, o attB/attP) no existente en los genomas de mamífero de origen natural. Por lo tanto, las células se tienen que modificar previamente para incorporar el sitio un sitio de reconocimiento para la recombinasa antes de que el sitio se pueda dirigir posteriormente la inserción genética. Debido a que el propio sitio de recombinasa se integra aleatoriamente en el genoma, sigue siendo necesario someterlo a una exploración y

evaluación extensas para identificar los clones que albergan el sitio en una localización que sea adecuada para la expresión genética con un nivel alto, y a largo plazo (Collingwood y Urnov, Targeted Gene Insertion to Enhance Protein Production from Cell Lines, in Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Además, la industria bioproductora tiene dudas para adoptar "nuevas" líneas celulares, tal como las que se han modificado para que alberguen un sitio de recombinasa, que no tienen un registro trazable de aprobación por la FDA. Por estas razones, los sistemas de modificación celular basados en recombinasa pueden no adoptarse fácilmente por la industria y es preferible una estrategia que permita a los biofabricantes utilizar sus líneas celulares existentes.

Sumario de la invención

La presente invención depende, en parte, del desarrollo de las líneas celulares de mamífero en las que se inserten las secuencias de interés (por ejemplo, transgenes transcritos activamente, exógenos) de manera proximal a un gen indicador endógeno en un locus que se puede amplificar, y el descubrimiento de que (a) la inserción de dichas secuencias exógenas de interés no inhibe la amplificación del gen de selección endógeno, (b) la secuencia exógena de interés se puede co-amplificar con el gen de selección endógeno, y (c) las líneas celulares resultantes, con una región amplificada que comprende múltiples copias del gen de selección endógeno y la secuencia exógena de interés, son estables durante periodos extensos incluso en ausencia del régimen de selección que se empleaba para inducir la amplificación. Por lo tanto, en un aspecto, la invención proporciona un procedimiento para la producción de líneas celulares que se pueden utilizar para la biofabricación de un producto proteico de interés dirigiendo específicamente la inserción de una secuencia exógena de interés capaz de expresar activamente el producto proteico de interés proximal a un gen de selección endógeno. En otro aspecto, la invención proporciona líneas celulares modificadas que se pueden utilizar para producir productos proteicos de interés (por ejemplo, proteínas terapéuticas tales como anticuerpos monoclonales) con altos niveles.

Se entiende que cualquiera de las realizaciones descritas posteriormente se puede combinar de cualquier manera deseada y se puede aplicar cualquier realización o combinación de realizaciones a cada uno de los aspectos descritos posteriormente, a menos de que el contexto dicte otra cosa.

En un aspecto, la invención proporciona una célula de mamífero recombinante que comprende un sitio diana modificado integrado establemente en un gen de selección con un locus que se puede amplificar, en el que el sitio diana modificado altera la función del gen de selección y en el que el sitio diana modificado comprende una secuencia de reconocimiento para una endonucleasa específica del sitio.

En algunas realizaciones, el gen de selección es de la glutamina sintetasa (GS) y el locus es de metionina sulfoximina (MSX) que se puede amplificar. En algunas realizaciones, el gen de selección es de la dihidrofolato reductasa (DHFR) y el locus es del metotrexato (MTX) que se puede amplificar.

En algunas realizaciones, el gen de selección se selecciona de entre el grupo que consiste en Dihidrofolato reductasa, glutamina sintetasa, hipoxantina fosforribosil transferasa, treonil ARNt sintetasa, Na-K-ATPasa, asparagina sintetasa, ornitina descarboxilasa, inosina-5'-fosfato deshidrogenasa, adenosina desaminasa, timidilato sintetasa, aspartato transcarbamilasa, metalotioneína, adenilato desaminasa (1,2), UMP-sintetasa y ribonucleótido reductasa.

En algunas realizaciones, el gen de selección se puede amplificar por selección con un agente de selección seleccionado de entre el grupo que consiste en metotrexato (MTX), metionina sulfoximina (MSX), aminopterina, hipoxantina, timidina, Borrelidina, ouabaina, albizina, beta-aspartil hidroxamato, alfa-difluorometilornitina (DFMO), ácido micofenólico, adenosina, alanosina, 2'-desoxicoformicina, fluorouracilo, pirazofurano, hidroxíurea, motexafin gadolinio, fludarabina, cladribina, gemcitabina, tezacitabina y triapina.

En algunas realizaciones, el sitio diana modificado se inserta en un exón del gen de selección. En algunas realizaciones, el sitio específico de endonucleasa es una meganucleasas, una nucleasa en dedos de zinc o una efector TAL nucleasa. En alguna realización, la célula recombinante comprende adicionalmente el sitio específico de endonucleasa.

En un aspecto, la invención proporciona una célula de mamífero recombinante que comprende un sitio diana modificado integrado establemente proximal a un gen de selección en un locus que se puede amplificar, en el que el sitio diana modificado comprende una secuencia de reconocimiento para un sitio específico de endonucleasa.

En algunas realizaciones, el sitio diana modificado está corriente abajo de la región reguladora 3' del gen de selección. En algunas realizaciones, el sitio diana modificado está de 0 a 100.000 pares de bases corriente abajo de la región reguladora 3' del gen de selección. En otras realizaciones, el sitio diana modificado está corriente arriba de la región reguladora 5' del gen de selección. En algunas realizaciones, el sitio diana modificado está de 0 a 100.000 pares de bases corriente arriba de la región reguladora 5' del gen de selección.

En otro aspecto, la invención proporciona un procedimiento para la inserción de una secuencia exógena en un locus que se puede amplificar de una célula de mamífero que comprende: (a) proporcionar una célula de mamífero que tiene un sitio diana endógena proximal a un gen de selección en un locus que se puede amplificar, en el que el sitio diana endógeno comprende : (i) una secuencia de reconocimiento para una meganucleasas modificada; (ii) una región

flanqueante 5' en 5' de la secuencia de reconocimiento; y
 (iii) una región 3' en 3' de la secuencia de reconocimiento; y (b) la introducción de una rotura de cadena doble entre las regiones flanqueantes 5' y 3' del sitio diana endógeno; (c) poner en contacto la célula con vector donante que comprende de 5' a 3': (i) una región 5' flanqueante donante homóloga a la región flanqueante 5' del sitio diana endógeno; (ii) una secuencia exógena; y (iii) una región flanqueante 3' donante homóloga de la región flanqueante 3' del sitio diana endógeno; de esta manera la región flanqueante 5' donante, la secuencia exógena y la región flanqueante 3' donante se insertan entre las regiones flanqueantes 5' y 3' del sitio diana endógeno mediante recombinación homóloga para proporcionar una célula modificada.

En algunas realizaciones, el procedimiento comprende adicionalmente el cultivo de la célula modificada en presencia de un compuesto que inhibe la función del gen de selección para amplificar el número de copias del gen de selección. En algunas realizaciones, la secuencia exógena comprende un gen de interés.

En algunas realizaciones, el sitio diana endógeno está corriente abajo de la región reguladora 3' del gen de selección. En algunas realizaciones, el sitio diana endógeno está de 0 a 100.000 pares de bases corriente abajo de la región reguladora 3' del gen de selección. En otras realizaciones, el sitio diana endógeno está corriente arriba de la región reguladora 5' del gen de selección. En algunas realizaciones, el sitio diana endógeno está de 0 a 100.000 pares de bases corriente arriba de la región reguladora 5' del gen de selección.

En un aspecto, la invención proporciona un procedimiento para la inserción de una secuencia exógena en un locus que se puede amplificar de una célula de mamífero que comprende: (a) proporcionar una célula de mamífero que tiene un sitio diana endógeno proximal a un gen de selección en el locus que se puede amplificar, en el que el sitio diana endógeno comprende: (i) una secuencia de reconocimiento para una meganucleasa modificada; (ii) una región flanqueante 5' en 5' de la secuencia de reconocimiento; y (iii) una región flanqueante 3' en 3' de la región de reconocimiento; y (b) introducir una rotura de doble cadena entre las regiones flanqueantes 5' y 3' del sitio diana endógeno; (c) poner en contacto la célula con un vector donante de un sitio diana modificado que comprende de 5' a 3': (i) una región 5' flanqueante donante homóloga con la región 5' flanqueante del sitio diana endógeno; (ii) una secuencia exógena que comprende un sitio diana modificado; y (iii) una región flanqueante 3' donante homóloga de la región flanqueante 3' del sitio diana endógeno por recombinación homóloga para proporcionar una célula de mamífero que comprende el sitio diana modificado; de manera que la región flanqueante 5' donante, la secuencia exógena y la región flanqueante 3' donante se insertan entre las regiones flanqueantes 5' y 3' del sitio diana endógeno mediante recombinación homóloga para proporcionar una célula de mamífero que comprende el sitio diana modificado; (d) introducir una rotura de doble cadena entre las regiones flanqueantes 3' y 5' del sitio diana modificado; (e) poner en contacto la célula que comprende el sitio diana modificado con un vector donante de la secuencia de interés que comprende de 5' a 3': (i) una región flanqueante donante 5' homóloga de la región flanqueante 5' del sitio diana modificado; (ii) una secuencia exógena que comprende una secuencia de interés; y (iii) una región flanqueante 3' donante homóloga de la región flanqueante 3' del sitio diana modificado; de manera que la región flanqueante 5', la secuencia exógena que comprende la secuencia de interés y la región flanqueante 3' donante se inserta entre las regiones flanqueantes 5' y 3' del sitio diana modificado por recombinación homóloga para proporcionar una célula de mamífero modificada que comprende la secuencia de interés.

En algunas realizaciones, el procedimiento comprende adicionalmente el cultivo de la célula de mamífero modificada en presencia de un compuesto que inhiba la función del gen indicador para amplificar el número de copias del gen de selección. En algunas realizaciones, la secuencia de interés comprende un gen.

En otro aspecto, la invención proporciona un procedimiento para insertar una secuencia exógena en un locus que se puede amplificar de una célula de mamífero que comprende: (a) proporcionar una célula de mamífero que tiene un sitio diana endógeno en un gen de selección en el locus que se puede amplificar, en el que el sitio diana endógeno comprende: (i) una secuencia de reconocimiento para una meganucleasa modificada; (ii) una región flanqueante 5' en 5' de la secuencia de reconocimiento; y (iii) una región flanqueante 3' en 3' de la secuencia de reconocimiento; y (b) introducir una rotura de cadena doble en las regiones flanqueantes 5' y 3' del sitio diana endógeno; (c) poner en contacto la célula con un vector donante del sitio diana modificado que comprende de 5' a 3': (i) una región flanqueante 5' donante homóloga con la región flanqueante 5' del sitio diana endógeno; (ii) una secuencia exógena que comprende un sitio diana modificado; y (iii) una región flanqueante 3' donante homóloga de la región flanqueante 3' del sitio diana endógeno; de manera que la región flanqueante 5' donante, la secuencia exógena y la región flanqueante 3' donante se insertan entre las regiones flanqueantes 5' y 3' del sitio diana endógeno por recombinación homóloga para proporcionar una célula de mamífero que comprende el sitio diana modificado; (d) introducir una rotura de doble cadena entre las regiones flanqueantes 5' y 3' del sitio diana modificado; (e) poner en contacto la célula que comprende el sitio diana modificado con un vector donante de la secuencia de interés que comprende de 5' a 3': (i) una región flanqueante 5' donante homóloga de la región flanqueante 5' del sitio diana modificado; (ii) una secuencia exógena que comprende una secuencia de interés; y (iii) una región flanqueante 3' donante homóloga de la región flanqueante 3' del sitio diana modificado; de manera que la región flanqueante 5', la secuencia exógena que comprende la secuencia de interés y la región flanqueante 3' donante se insertan entre las regiones flanqueantes 5' y 3' del sitio diana modificado mediante recombinación homóloga para proporcionar una célula de mamífero modificada que comprende la secuencia de interés.

En algunas realizaciones, el procedimiento comprende adicionalmente el cultivo de la célula de mamífero modificada

en presencia de un compuesto que inhibe la función del gen de selección para amplificar el número de copias del gen de selección.

En algunas realizaciones, la secuencia de interés comprende un gen.

5 En algunas realizaciones, el sitio diana endógeno está en un intrón del gen de selección. En otras realizaciones, el sitio diana endógeno está en un exón del gen de selección.

En un aspecto, la invención proporciona una meganucleasa recombinante que comprende un polipéptido que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 15.

En otro aspecto, la invención proporciona una meganucleasa recombinante que comprende la secuencia de aminoácidos de SEQ ID NO: 15.

10 En otro aspecto, la invención proporciona una meganucleasa recombinante que reconoce un escinde un sitio de reconocimiento que tiene un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 14. En una realización, la meganucleasa reconoce y escinde un sitio de reconocimiento de SEQ ID NO: 14.

15 En otro aspecto, la invención proporciona una meganucleasa recombinante que comprende un polipéptido que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 9. En una realización, la meganucleasa recombinante tiene la secuencia de la meganucleasa de SEQ ID NO: 9.

En otro aspecto, la invención proporciona una meganucleasa recombinante que reconoce un escinde un sitio de reconocimiento de al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 7. En una realización, la meganucleasa reconoce y escinde un sitio de reconocimiento de SEQ ID NO: 7.

20 En otro aspecto, la invención proporciona una meganucleasa recombinante que comprende un polipéptido que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 10. En una realización, la meganucleasa recombinante comprende el polipéptido de SEQ ID NO: 10.

25 En otro aspecto, la invención proporciona una meganucleasa recombinante que reconoce y escinde un sitio de reconocimiento que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 8. En una realización, la meganucleasa recombinante reconoce y escinde un sitio de reconocimiento de SEQ ID NO: 8.

En otro aspecto, la invención proporciona una meganucleasa recombinante que comprende un polipéptido que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 13. En una realización, la meganucleasa recombinante comprende el polinucleótido de SEQ ID NO: 13.

30 En otro aspecto, la invención proporciona una meganucleasa recombinante que reconoce y escinde un sitio de reconocimiento que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 12. En una realización, la meganucleasa reconoce y escinde un sitio de reconocimiento de SEQ ID NO: 12.

35 En otro aspecto, la invención proporciona una meganucleasa recombinante que comprende un polipéptido que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 29. En una realización, la meganucleasa recombinante comprende el polipéptido de SEQ ID NO: 29.

En otro aspecto, la invención proporciona una meganucleasa recombinante que reconoce y escinde un sitio de reconocimiento que tiene al menos un 75 %, 80 %, 85 %, 90 %, 95 %, 97 %, 98 % o 99 % de identidad de secuencia con la SEQ ID NO: 30. En una realización, la meganucleasa que y escinde un sitio de reconocimiento de SEQ ID NO: 30.

40 En otro aspecto la invención proporciona líneas celulares de mamífero recombinantes que continúan expresando un producto proteico de interés a partir de una secuencia exógena de interés presente en una región amplificada del genoma(es decir, presente en 2-1.000 copias, co-amplificado con un gen de selección en un locus que se puede amplificar) durante un periodo de al menos 8, 9, 10, 11, 12, 13 o 14 semanas después de retirar el agente de amplificación selección, y con una reducción de los niveles de expresión y/o número de copias de menos de 20, 25, 30, 35 o 40 %.

45 En otro aspecto, la invención proporciona procedimientos de producción de células recombinantes con regiones amplificadas que incluyen una secuencia de interés y un gen de selección sometiendo las células recombinantes descritas anteriormente a selección con un agente de selección que produce la co-amplificación de la secuencia de interés y el gen de selección.

50 En otro aspecto, la invención proporciona procedimientos de producción de un producto proteico de interés cultivando las células recombinantes descritas anteriormente, o las células recombinantes descritas con regiones amplificadas, y la obtención de un producto proteico de interés a partir del medio del cultivo o un lisado celular.

Breve descripción de las figuras

Figura 1. Una estrategia general para el direccionamiento de una secuencia de interés a un locus que se puede amplificar.

Figura 2. (A) Esquema del locus DHFR de CHO que muestra una región preferida para el direccionamiento de una secuencia de interés 5.000 a 60.000 pares de bases corriente abajo del gen DHFR. (B) Esquema del locus GS de CHO que muestra una región preferida para el direccionamiento de una secuencia de interés 5.000-55.000 pares de base corriente abajo del gen GS.

Figura 3. Estrategia para la inserción de una secuencia de interés en un locus que se puede amplificar en un procedimiento de dos etapas que implican una secuencia diana modificada integrada previamente.

Figura 4. Estrategia para la inserción de una secuencia diana modificada en un locus que se puede amplificar con la retirada concomitante de una parte del gen de selección, seguido por la inserción de una secuencia de interés y reconstitución del gen de selección.

Figura 5. Estrategia para la inserción de una secuencia diana modificada en un locus que se puede amplificar con alteración concomitante de la secuencia codificante de un gen de selección, seguido por inserción de una secuencia de interés y reconstitución del gen de selección.

Figura 6. Estrategia para la inserción de una secuencia diana modificada en un locus que se puede amplificar con la alteración concomitante del procesamiento de ARNm, seguido por la inserción de una secuencia de interés y reconstitución del gen de selección.

Figura 7. (A) Un ensayo de recombinación repetida directa para la actividad de endonucleasa específica del sitio. (B) resultados del ensayo en (A) aplicada a las meganucleasas CHO-23/24 y CHO-51/52. (C) Alineamiento de secuencias obtenidas de células CHO transfectadas con ARNm que codifica la meganucleasa CHO-23/24. (D) Alineamiento de secuencias obtenidas de células CHO transfectadas con ARNm que codifica la meganucleasa CHO-51/52.

Figura 8. (A) Estrategia para la inserción de una secuencia de ADN exógeno en el locus DHFR de CHO utilizando la meganucleasa CHO-51/52. (B) Productos de la PCR que demuestran la inserción de una secuencia diana modificada.

Figura 9. (A) Estrategia para la inserción de una secuencia diana modificada en el locus DHFR de CHO utilizando la meganucleasa CHO-23/24, seguido por la inserción mediada por Flp recombinasa de una secuencia de interés. (B) los productos PCR de clones resistentes a higromicina producidos en (A). (C) expresión de GFP por los 24 clones producidos en (B).

Figura 10. Resultados de los experimentos con una línea de CHO que expresa GFP producida integrando un casete de expresión del gen GFP en el locus DHFR utilizando una estrategia de secuencia diana como se muestra en la Figura 9.

Figura 11. (A) Un ensayo de recombinación repetida directa, como en la Figura 5A. (B) el ensayo de (A) aplicado a las meganucleasas CHO-13/14 y CGS-5/6. (C) Alineamiento de secuencias obtenidas a partir de células CHO transfectadas con ARNm que codifica la meganucleasa CGS-5/6.

Descripción detallada de la invención**1.1 Introducción**

La presente invención depende, en parte, del desarrollo de líneas celulares de mamífero en las que se habían insertado transgenes transcritos activamente exógenos proximales a un locus endógeno que se puede amplificar, y el descubrimiento de que (a) la inserción de dichos transgenes transcritos activamente exógenos no evitan o inhiben sustancialmente la amplificación del locus endógeno que se puede amplificar, (b) el transgén exógeno transcrito activamente puede co-amplificarse con el locus endógeno que se puede amplificar, y (c) la línea celular resultante, con una región amplificada que comprende múltiples copias del locus endógeno que se puede amplificar y el transgén exógeno transcrito activamente es estable durante periodos extensos incluso en ausencia del régimen de selección que se emplee para inducir la amplificación. Por lo tanto, en un aspecto, la invención proporciona un procedimiento para la producción de líneas celulares que se pueden utilizar para biofabricar un producto proteico de interés direccionando específicamente la inserción de un gen exógeno capaz de expresar activamente el producto proteico de interés proximal al locus endógeno que se puede amplificar. En otro aspecto, la invención proporciona líneas celulares modificadas que se pueden utilizar para producir productos proteicos de interés (por ejemplo, proteínas terapéuticas tales como anticuerpos monoclonales) a altos niveles.

1.2 Referencias y definiciones

La patente y bibliografía científica a la que se hace referencia en el presente documento establece un conocimiento que está disponible para los expertos en la técnica. Las divulgaciones completas de las patentes de Estados Unidos expedidas, las solicitudes pendientes, las solicitudes extranjeras publicadas y las referencias científicas y técnicas citadas en el presente documento, incluyendo las secuencias de bases de datos de proteínas y ácidos nucleicos, se incorporan por referencia en la misma medida que si cada una de ellas se indicara específica e individualmente como incorporada por referencia.

Como se utiliza en el presente documento "meganucleasas" se refiere a endonucleasas de asentamiento (a las que también se hace referencia como endonucleasas codificadas por el intrón del Grupo I) o endonucleasas no de origen

natural (por ejemplo, diseñadas o modificadas racionalmente) basadas en la secuencia de aminoácidos de una endonucleasa de asentamiento de origen natural. Ejemplos de meganucleasas de origen natural incluyen I-SceI, I-CreI, I-CeuI, I-DmI, I-MsoI, I-AniI, etc. Las meganucleasas diseñadas racionalmente se desvelan por ejemplo en los documentos WO 2007/047859 y WO 2009/059195, y se pueden alterar para tener una especificidad de unión al ADN modificada, actividad de escisión de ADN, afinidad de unión al ADN o propiedades de dimerización modificadas con respecto a la meganucleasas de origen natural. Una meganucleasa se puede unir a un ADN de doble cadena como un homodímero (por ejemplo, la I-CreI de tipo silvestre), o se puede unir al ADN como un heterodímero (por ejemplo, las meganucleasas modificadas desveladas en el documento WO 2007/047859). Una meganucleasas modificada también puede ser una “meganucleasas de cadena sencilla” en la que se une un par de dominios de unión al ADN derivados de una meganucleasas natural en un único polipéptido utilizando un engarce peptídico (por ejemplo, las meganucleasas de cadena sencilla desveladas en el documento WO 2009/059195).

Como se utiliza en el presente documento, la expresión “meganucleasas de cadena sencilla” se refiere a un polipéptido que comprende un par de subunidades de meganucleasas unidas por un engarce. Una meganucleasas de cadena sencilla tiene la organización: subunidad del extremo N – engarce – subunidad del extremo C. Las dos subunidades de meganucleasas en general no serán idénticas en la secuencia de aminoácidos y no reconocerán secuencias de ADN idénticas. Por lo tanto, las meganucleasas de cadena sencilla normalmente escindirán secuencias de reconocimiento pseudo palindrómicas o no palindrómicas. Los procedimientos de producción de meganucleasas de cadena sencilla se desvelan en el documento WO 2009/059195.

Como se utiliza en el presente documento, la expresión “endonucleasa específica del sitio” significa una meganucleasas, nucleasa en dedos de zinc o nucleasa efectora TAL.

Como se utiliza en el presente documento, con respecto a una proteína, el término “recombinante” significa que tiene una secuencia de aminoácidos alterada como resultado de la aplicación de técnicas de modificación genética a los ácidos nucleicos que codifican la proteína, y las células u organismos que expresan la proteína. Con respecto a un ácido nucleico, el término “recombinante” significa que tienen una secuencia de ácido nucleico alterada como resultado de la aplicación de técnicas de modificación genéticas. Las técnicas de modificación genético incluyen, pero no se limitan a, tecnologías de PCR y clonación de ADN; transfección, transformación y otras tecnologías de transferencia genética; recombinación genética, mutagénesis dirigida el sitio; y fusión genética. De acuerdo con esta definición, una proteína que tenga una secuencia de aminoácidos idéntica a una proteína de origen natural, pero producida por clonación y expresión en un huésped heterólogo, no se considera recombinante. Como se utiliza en el presente documento, el término “modificado” es sinónimo del término “recombinante”.

Como se utiliza en el presente documento, con respecto a una meganucleasas, la expresión “de tipo silvestre” se refiere a cualquier forma de meganucleasas de origen natural. La expresión “de tipo silvestre” no pretende significar la variante alélica más común de la enzima en la naturaleza sino más bien, cualquier variante alélica que se encuentra en la naturaleza. Las endonucleasas de asentamiento de tipo silvestre se distinguen de las meganucleasas recombinantes o de origen no natural.

Como se utiliza en el presente documento, la expresión “secuencia de reconocimiento” se refiere a una secuencia de ADN que se une y se escinde por una meganucleasas. Una secuencia de reconocimiento comprende un par de “medios sitios” de 9 pares de bases que están separadas por cuatro pares de bases. En el caso de las meganucleasas homo- o heterodiméricas, cada uno de los dos monómeros hace contactos específicos de base con un medio sitio. En el caso de una meganucleasa heterodimérica de cadena sencilla, el dominio del extremo N de la proteína se pone en contacto con un primer medio sitio y el dominio del extremo C de la proteína se pone en contacto con un segundo medio sitio. En el caso de I-CreI, por ejemplo, la secuencia de reconocimiento tiene 22 pares de bases y comprende un par de “medios sitios” de 9 pares de bases que están separados por cuatro pares de bases.

Como se utiliza en el presente documento, la expresión “sitio diana” se refiere a una región del ADN cromosómico de una célula que comprende una secuencia diana en la cual se puede insertar una secuencia de interés. Como se utiliza en el presente documento, la expresión “sitio diana modificado” se refiere a una secuencia exógena de ADN integrada en el ADN cromosómico de una célula que comprende una secuencia diana modificada en la que se puede insertar una secuencia de interés.

Como se utiliza en el presente documento, la expresión “secuencia diana” significa una secuencia de ADN en un sitio diana que incluye una o más secuencias de reconocimiento para una nucleasa, integrasa, transposasa y/o recombinasa. Por ejemplo, una secuencia diana puede incluir una secuencia de reconocimiento para una meganucleasa. Como se utiliza en el presente documento, una “secuencia diana modificada” significa una secuencia diana exógena que se introduce en un cromosoma que sirve como punto de inserción para otra secuencia.

Como se utiliza en el presente documento, la expresión “región flanqueante” o “secuencia flanqueante” se refiere a una secuencia de > 3 o preferentemente, > 50 o, más preferentemente, > 200 o, más preferentemente, > 400 pares de bases de ADN que está inmediatamente 5' o 3' de una secuencia de referencia (por ejemplo, una secuencia diana o una secuencia de interés).

Como se utiliza en el presente documento, las expresiones “locus que se puede amplificar” se refiere a una región del

ADN cromosómico de una célula que se puede amplificar mediante la selección con uno o más compuestos (por ejemplo, fármacos) en el medio de cultivo. Un locus que se puede amplificar normalmente comprenderá un gen que codifica una proteína que, en condiciones apropiadas, es necesario para la supervivencia de la célula. Inhibiendo la función de dicha proteína esencial, por ejemplo, con un fármaco de molécula pequeña, el locus que se pueden amplificar se duplica muchas veces como medio de aumento del número de copias del gen esencial. Un gen de interés, si se integra en un locus que se puede amplificar, también se duplicaría con el gen esencial. Ejemplos de loci que se pueden amplificar incluyen las regiones cromosómicas que comprenden genes de DHFR, GS y HPRT.

Como se utiliza en el presente documento, la expresión “locus amplificado” o “gen amplificado” o “secuencia amplificada” se refiere a un locus, gen o secuencia que está presente en 2-1.000 copias como resultado de la amplificación genética en respuesta a la selección de un indicador genético. Un gen o secuencia amplificados puede ser un gen o secuencia que se co-amplifica debido a la selección de un gen de selección en el mismo locus que se puede amplificar. En realizaciones preferidas, una secuencia de interés se amplifica hasta al menos 3, 4, 5, 6, 7, 8, 9 o 10 copias.

Como se utiliza en el presente documento, la expresión “gen de selección” se refiere a un gen endógeno que es esencial para la supervivencia celular en algunas condiciones de cultivo específicas (por ejemplo, la presencia o ausencia de un nutriente, toxina o fármaco). Los genes de selección son endógenos de la célula y se distinguen de los “marcadores genéticos” exógenos tal como genes de resistencia a antibióticos. Los genes indicadores existen en su contexto natural en el ADN cromosómico de la célula. Por ejemplo, el DHFR es un gen de selección que es necesario para la supervivencia celular en presencia de MTX en el medio del cultivo. El gen es esencial para el reconocimiento en ausencia de hipoxantina y timidina. Si se elimina el gen de selección de DHFR endógeno, las células son capaces de crecer en ausencia de hipoxantina y timidina si se les da una copia exógena del gen DHFR. Esta copia exógena del gen DHFR es un marcador genético, pero no es un gen de selección. Un locus que se puede amplificar comprende un gen de selección y un sitio diana. Un sitio diana se encuentra fuera de un gen de selección de manera que un gen de selección no comprende un sitio diana. Ejemplos de genes de selección se dan en la Tabla 1.

Como se utiliza en el presente documento, cuando se utiliza en conexión con la posición de un sitio diana, secuencia de reconocimiento o secuencia de interés insertada con respecto a la posición de un gen de selección, el término “proximal” significa que el sitio diana, secuencia de reconocimiento, o secuencia de interés insertada están en el mismo locus que se puede amplificar que el gen de selección, sea corriente arriba (5') o corriente abajo (3') del gen de selección, y preferentemente entre el gen de selección y el siguiente gen en la región (sea corriente arriba (5') o corriente abajo (3')). Normalmente, un sitio diana, secuencia de reconocimiento, o secuencia de interés insertada “proximales” estarán a < 100.000 pares de bases del gen de selección, según se mide desde el primer o último nucleótido del primer o último elemento regulador del gen de selección.

Como se utiliza en el presente documento, la expresión “recombinación homóloga” se refiere al proceso celular natural en el que una rotura de ADN de doble cadena se repara utilizando una secuencia de ADN homóloga como matriz de reparación (véase, por ejemplo, Cahill y col. (2006), *Front. Biosci.* 11:1958-1976). La secuencia de ADN homóloga puede ser una secuencia cromosómica endógena o un ácido nucleico exógeno que se suministra a la célula. Por lo tanto, para algunas aplicaciones de las meganucleasas modificadas, se suministra una nucleasa que se utiliza para escindir una secuencia de reconocimiento en una secuencia diana de un genoma y un ácido nucleico exógeno con homología o similitud de secuencia sustancial con la secuencia diana a la célula y se utiliza como una matriz para la reparación por recombinación homóloga. La secuencia de ADN del ácido nucleico exógeno, que puede ser significativamente diferente de la secuencia diana, se inserta o incorpora de esta manera en la secuencia cromosómica. El proceso de recombinación homóloga se produce primariamente en organismos eucariotas. El término “homología” se utiliza en el presente documento de manera equivalente a “similitud de secuencia” y no pretende que se necesite una identidad por relación descendente o filogenética.

Como se utiliza en el presente documento, la expresión “integrada establemente” significa que una secuencia de ADN exógeno o heteróloga se ha insertado covalentemente en un cromosoma (por ejemplo, por recombinación homóloga, unión de extremos no homóloga, transposición, etc.) y se mantiene en el cromosoma durante un periodo de al menos 8 semanas.

Como se utiliza en el presente documento, la expresión “unión de extremos no homóloga” o “NHEJ” se refiere al proceso celular natural en el que la rotura de ADN de doble cadena se repara por la unión directa de dos segmentos de ADN no homólogos (véase, por ejemplo, Cahill y col. (2006), *Front. Biosci.* 11:1958-1976). La reparación de ADN por unión de extremos no homólogos tiene tendencia a errores y frecuentemente da como resultado la adición o eliminación no de la matriz de secuencias de ADN en el sitio de reparación. Por lo tanto, para ciertas aplicaciones, se puede utilizar una meganucleasas modificada para producir una rotura de doble cadena en una secuencia de reconocimiento en un locus que se puede amplificar y se puede capturar una molécula de ácido nucleico exógena, tal como un producto de una PCR en el sitio de la rotura de EN mediante NHEJ (véase, por ejemplo, Salomon y col. (1998), *EMBO J.* 17:6086-6095). En tales casos, el ácido nucleico exógeno puede tener una homología o no, con la secuencia diana. El proceso de unión de extremos no homólogos se produce tanto en eucariotas como procariontes tales como bacterias.

Como se utiliza en el presente documento, la expresión “secuencia de interés” significa cualquier secuencia de ácido nucleico, si codifica una proteína, ARN o elemento regulador (por ejemplo, un amplificador, silenciador o secuencia promotora), que pueden insertarse en un genoma o utilizarse para reemplazar una secuencia de ADN genómico. Las secuencias de interés pueden tener secuencias de ADN heterólogo que permitan la marcación de una proteína o ARN que se expresa a partir de la secuencia de interés. Por ejemplo, se puede marcar una proteína con marcadores que incluyen, pero no se limitan a, un epítipo (por ejemplo, c-myc, FLAG) u otro ligando (por ejemplo, poli-His). Además, una secuencia de interés puede codificar una proteína de fusión de acuerdo con técnicas conocidas en la técnica (véase, por ejemplo, Ausubel y col., *Current Protocols in Molecular Biology*, Wiley 1999). En realizaciones preferidas, la secuencia de interés comprende un promotor unido operativamente a un gen que codifica una proteína de valor medicinal tal como un anticuerpo, un fragmento de anticuerpo, citocina, factor de crecimiento, hormona o enzima. Para algunas aplicaciones, la secuencia de interés está flanqueada por una secuencia de ADN que es reconocida por la meganucleasa modificada para la escisión. Por lo tanto, las secuencias flanqueantes se escinden permitiendo la inserción apropiada de la secuencia de interés en las secuencias genómicas de reconocimiento escindidas por una meganucleasa modificada. Para algunas aplicaciones, la secuencia de interés está flanqueada por secuencias de ADN con homología o similitud de secuencia sustancial con el sitio diana de manera que se inserta por recombinación homóloga la secuencia de interés en el genoma en el locus de la secuencia diana.

Como se utiliza en el presente documento, la expresión “ADN donante” se refiere a una molécula de ADN que comprende una secuencia de interés flanqueado por secuencias de ADN homóloga de un sitio diana. El ADN donante puede servir como almacén para la reparación de ADN por recombinación homóloga si se suministra a una célula con una nucleasa específica del sitio tal como una meganucleasa, nucleasa en dedos de zinc, o nucleasa efectora TAL. El resultado de dicha reparación de ADN es la inserción de la secuencia de interés en el ADN cromosómico de la célula. El ADN donante puede ser lineal, tal como un producto de la PCR, o circular, tal como un plásmido. En los casos en los que el ADN donante es un plásmido circular, se puede hacer referencia a este como un “plásmido donante”.

Como se utiliza en el presente documento, a menos de que se indique específicamente otra cosa, la palabra “o” se utiliza en sentido incluyente de “y/o” y no en el sentido excluyente de “o/o”.

2.1 Transgén direccionado a loci que se pueden amplificar

La presente invención proporciona procedimientos para la generación de líneas celulares de mamífero transgénicas que expresen un producto proteico deseado de interés, incluyendo líneas celulares de “alta producción”, dirigiendo la inserción de un gen que codifica el producto proteico de interés (por ejemplo, un casete de expresión genética de una proteína terapéutica) en regiones del genoma que se pueden amplificar. Dichas regiones de las células de mamífero incluyen los genes de DHFR, GS, y HPRT, así como otros que se muestran en la Tabla 1.

El mecanismo preciso de la amplificación genética no se conoce. Además, es muy probable que no haya un único mecanismo por el que se produce la amplificación genética, sino que se produce una variedad de diferentes aberraciones cromosómicas aleatorias, en combinación con una fuerte selección por amplificación, resulta en un número de copias genéticas creciente (revisado en Omasa (2002), *J. Biosci. Bioeng.* 94:600-605). Está claro que la localización cromosómica tiene un papel principal en la amplificación y el mantenimiento estable de genes amplificados (Brinton y Heintz (1995), *Chromosoma* 104:143-51). Se ha descubierto que los transgenes integrados en localizaciones cromosómicas adyacentes a los telómeros se amplifican más fácilmente y, una vez amplificados tienden a ser estables en un alto número de copias después de retirar el agente de selección (Yoshikawa y col. (2000), *Cytotechnology* 33:37-46; Yoshikawa y col. (2000), *Biotechnol Prog.* 16:710-715). Esto es significativo debido a que los agentes de selección tales como MTX y MSX son tóxicos y no se pueden incluir en el medio de cultivo en un procedimiento de biofabricación comercial. Por el contrario, los transgenes integrados en las regiones del genoma de CHO que no es adyacente a los telómeros se amplifican de manera ineficaz y pierden rápidamente el número de copias después de retirar los agentes de selección del medio. Por ejemplo, Yoshikawa y col. describieron que los transgenes integrados aleatoriamente unidos a un marcador genético DHFR se amplificada hasta más de 10 veces de número de copias cuando el sitio de integración estaba adyacente el telómero (Yoshikawa y col. (2000), *Biotechnol Prog.* 16:710-715). Estos investigadores también descubrieron que un transgén integrado en una región no telomérica perdería > 50 % de sus copias en solo 20 días después de la retirada de MTX del medio de cultivo. Ninguno de los genes de selección identificados en la Tabla 1 es adyacente al están en regiones no teloméricas en el genoma de ratón (www.ensembl.com) telómero y la similitud con la organización del genoma entre el ratón y CHO, hace que sea probable que estos genes estén también en la región no telomérica de las CHO (Xu y col. (2011), *Nat. Biotechnol.* 29:735-741). Por lo tanto, la técnica anterior instruye que los loci identificados en la Tabla 1, incluyendo los loci DHFR y GS, son localizaciones no preferidas para dirigir la inserción del transgén si el objetivo es la amplificación genética eficaz y estable.

Además, en el caso de la amplificación de un gen endógeno, está calor que las secuencias cromosómicas fuera de la secuencia del gen de selección tienen un papel importante facilitando la amplificación y definiendo la longitud de la secuencia de ADN que se co-amplifica con el gen seleccionado (Looney y Hamlin (1987), *Mol. and Cell. Biol.* 7:569-577). En particular, se ha demostrado que la secuencia y localización del origen de replicación del ADN en relación con el gen de selección tiene un papel principal en la amplificación. Por ejemplo, se ha demostrado que la amplificación del locus DHFR endógeno de CHO es dependiente al encontrar un par de orígenes de replicación descubiertos en la

región 5.000-60.000 pares de bases corriente abajo del gen que codifica la secuencia de DHFR (Anachkova y Hamlin (1989), *Mol. and Cell. Biol.* 9:532-540; Milbrandt y col. (1981), *Proc. Natl. Acad. Sci. USA* 78:6042-6047). Además, Brinton y Heintz en demostrado que esos mismos orígenes de replicación fallan en la promoción de la amplificación genética cuando se incorporan aleatoriamente en el genoma con una secuencia DHFR transgénica (Brinton y Heintz (1995), *Chromosoma.* 104:143-51). Esto demuestra claramente la importancia del mantenimiento de la secuencia y un contexto cromosómico apropiado de estos orígenes de replicación para promover la amplificación del gen de DHFR. Por lo tanto, la técnica instruye que la región corriente abajo del DFHR es crítica para la amplificación genética y no debería alterarse, por ejemplo, insertando un casete de expresión genética transgénica como se describe en la presente invención.

10 Sorprendentemente, los inventores han descubierto que las secuencias de ADN, incluyendo las secuencias exógenas transcripcionalmente activas, que se insertan proximales a (por ejemplo, en <100.000 pares de bases) los genes de selección en líneas celulares de mamífero (por ejemplo, CHO-K1) se co-amplificarán en presencia de los compuestos apropiados que se seleccionan para la amplificación. Por lo tanto, la presente invención proporciona procedimientos para la producción de líneas celulares isogénicas en las que se pueden amplificar transgenes que codifican productos proteicos de interés (por ejemplo, casetes de expresión genética bioterapéutica) pero en las que no es necesario explorar un gran número de líneas celulares generadas aleatoriamente para identificar las que expresan altos niveles del producto proteico de interés y son resistentes al silenciamiento genético.

Además, los inventores han descubierto sorprendentemente que las líneas celulares de mamífero de la invención, en las que se co-amplifica una secuencia de interés con un gen de selección en un locus que se puede amplificar son estables con respecto a la expresión de la secuencia de interés y/o número de copias de la secuencia de interés incluso en ausencia de una selección continua. Es decir, mientras que la técnica enseña que las secuencias amplificadas se reducirán a lo largo del tiempo si no se mantiene la selección (véase, por ejemplo, Yoshikawa y col. (2000), *Biotechnol Prog.* 16:710-715), los inventores han descubierto que las líneas celulares de acuerdo con los procedimientos de la invención continúan produciendo los productos proteicos de interés (codificados por las secuencias de interés) a niveles del 20 %-25 % de los niveles iniciales, incluso 14 semanas después de la retirada de la agente de selección. Esto es significativo, como se ha señalado anteriormente, debido a que los agentes de selección tales como MTX y MSX son tóxicos, y sería altamente deseable producir proteínas bioterapéuticas en líneas celulares que no necesitan continuar con la exposición a dichos agentes de selección. Por lo tanto, en algunas realizaciones, la invención proporciona líneas celulares de mamífero que continúan expresando un producto proteico de interés a partir de una secuencia exógena de interés presente en una región amplificada del genoma (es decir, presente en 2-1.000 copias, co-amplificada con un gen de selección en un locus que se puede amplificar) durante un periodo de al menos 8, 9, 10, 11, 12, 13, o 14 semanas después de la retirada del agente de selección amplificación, y con una reducción de los niveles de expresión y/o número de copias de menos del 20, 25, 30, 35, o 40 %.

La presente invención también proporciona los productos necesarios para practicar los procedimientos, y dirigir la inserción de secuencias de interés en loci que se pueden amplificar en líneas celulares de mamífero. Un procedimiento común para la inserción o modificación de una secuencia de ADN implica la introducción de una secuencia de ADN transgénica flanqueada por secuencias homólogas de la diana genómica y la selección o exploración de un evento de recombinación homóloga satisfactoria. La recombinación con el ADN transgénico se produce raramente, pero se puede estimular por una rotura de doble cadena en el ADN genómico en el sitio diana (Porteus y col. (2005), *Nat. Biotechnol.* 23: 967-73; Tzfira y col. (2005), *Trends Biotechnol.* 23: 567-9; McDaniel y col. (2005), *Curr. Opin. Biotechnol.* 16: 476-83). Se han empleado muchos procedimientos para crear roturas de ADN de doble cadena, incluyendo tratamientos con radiación y químicos. Aunque estos procedimientos estimulan eficazmente la recombinación, las roturas de doble cadena se dispersan aleatoriamente en el genoma, que pueden ser altamente mutagénicas y tóxicas. Hasta el presente, la incapacidad para dirigir modificaciones genéticas a sitios únicos en un escenario cromosómico es un impedimento principal de la modificación del genoma de rutina.

Una estrategia para conseguir este objetivo es estimular la recombinación homóloga en una rotura de cadena doble en un locus diana utilizando una nucleasa con especificidad por una secuencia que es lo suficientemente grande para estar presente en un único sitio en el genoma (véase, por ejemplo, Porteus y col. (2005), *Nat. Biotechnol.* 23: 967-73). La eficacia de esta estrategia se ha demostrado en una variedad de organismos utilizando ZFN (Porteus (2006), *Mol Ther* 13: 438-46; Wright y col. (2005), *Plant J.* 44: 693-705; Urnov y col. (2005), *Nature* 435: 646-51). Las endonucleasas de asentamiento son un grupo de nucleasas de origen natural que reconocen sitios de escisión de 15-40 pares de bases que se encuentran comúnmente en los genomas de plantas y hongos. Se asocian frecuentemente con elementos de ADN parásitos, tales como intrones e inteínas de auto corte y empalme del Grupo I. Promueven naturalmente la recombinación homóloga o inserción genética en localizaciones específicas del genoma huésped produciendo una doble rotura en el cromosoma, que recluta la maquinaria de reparación de ADN celular (Stoddard (2006), *Q. Rev. Biophys.* 38: 49-95). Las endonucleasas de asentamiento se agrupan comúnmente en cuatro familias: la familia LAGLIDADG, la familia GIY-YIG, la familia de la caja His-Cys y la familia HNH. Estas familias se caracterizan por motivos estructurales que afectan actividad catalítica y la secuencia de reconocimiento. Por ejemplo, los miembros de la familia LAGLIDADG se caracterizan por tener una o dos copias del motivo conservado LAGLIDADG (véase Chevalier y col. (2001), *Nucleic Acids Res.* 29(18): 3757-3774). Las endonucleasas de asentamiento LAGLIDADG con una única copia del motivo LAGLIDADG forma homodímeros, mientras que los miembros con dos copias del motivo LAGLIDADG se encuentran como monómeros.

Las endonucleasas de asentamiento, primariamente de la familia LAGLIDADG, se han utilizado para promover eficazmente la modificación del genoma específica del sitio en plantas, levaduras, *Drosophila*, células de mamífero y ratones, pero esta estrategia se ha limitado a la modificación de genes homólogos que conservan la secuencia de reconocimiento de endonucleasa (Monnat y col. (1999), *Biochem. Biophys. Res. Commun.* 255: 88-93) o en genomas modificados previamente en los que se ha introducido una secuencia de reconocimiento (Rouet y col. (1994), *Mol. Cell. Biol.* 14: 8096-106; Chilton y col. (2003), *Plant Physiol.* 133: 956-65; Puchta y col. (1996), *Proc. Natl. Acad. Sci. USA* 93: 5055-60; Rong y col. (2002), *Genes Dev.* 16: 1568-81; Gouble y col. (2006), *J. Gene Med.* 8(5):616-622).

La implementación sistemática de modificación genética estimulada por nucleasas necesita el uso de enzimas modificadas con especificidades a medida para dirigirse a las roturas de ADN en sitios existentes en un genoma, y, por lo tanto, ha habido gran interés en adaptar las endonucleasas de asentamiento para promover modificaciones genéticas en sitios médica o biotecnológicamente relevantes (Porteus y col. (2005), *Nat. Biotechnol.* 23: 967-73; Sussman y col. (2004), *J. Mol. Biol.* 342: 31-41; Epinat y col. (2003), *Nucleic Acids Res.* 31: 2952-62).

La I-Crel (SEQ ID NO: 1) es un miembro de la familia LAGLIDADG de endonucleasas de asentamiento que reconoce y corta una secuencia de reconocimiento de 22 pares de bases en el cromosoma de cloroplastos de las algas *Chlamydomonas reinhardtii*. Se han utilizado técnicas de selección genética para modificar la preferencia de sitio de escisión por la I-Crel (Sussman y col. (2004), *J. Mol. Biol.* 342: 31-41; Chames y col. (2005), *Nucleic Acids Res.* 33: e178; Seligman y col. (2002), *Nucleic Acids Res.* 30: 3870-9, Arnould y col. (2006), *J. Mol. Biol.* 355: 443-58). Más recientemente, se ha descrito un procedimiento para diseñar racionalmente endonucleasas de asentamiento mono-LAGLIDADG que sean capaces de rediseñar completamente la I-Crel y otras endonucleasas de asentamiento para dirigirse a sitios de ADN ampliamente divergentes, incluyendo sitios en genomas de mamíferos, levaduras, plantas, bacterianos y víricos (documento WO 2007/047859).

Por lo tanto, en una realización la invención proporciona meganucleasas modificadas derivadas de la secuencia de aminoácidos de I-Crel que reconoce y corta sitios de ADN en regiones que se pueden amplificar de genomas de mamífero. Estas meganucleasas modificadas se pueden utilizar de acuerdo con la invención para dirigir la inserción de casetes de expresión genética en localizaciones determinadas del ADN cromosómico de líneas celulares tales como de células CHO. La presente invención optimizará enormemente la producción de líneas celulares deseadas reduciendo el número de líneas que se deben explorar para identificar un clon de "alta producción" adecuados para la producción a escala comercial de una glicoproteína terapéutica.

La presente invención implica el direccionamiento las "secuencias de interés" de ADN transgénico en loci que se pueden amplificar. Los loci que se pueden amplificar son regiones del ADN cromosómico que contiene genes de selección que se convierten en amplificados en presencia de agentes de selección (por ejemplo, fármacos). Por ejemplo, el locus DHFR de la célula de ovario de hámster chino (CHO) se puede amplificar hasta ~1.000 copias cultivando las células en presencia de metotrexato (MTX), un inhibidor de la DHFR. La Tabla 1 enumera ejemplos adicionales de genes indicadores que se pueden amplificar utilizando fármacos de molécula pequeña (Kellems, ed. *Gene amplification in mammalian cells: a comprehensive guide*. Marcel Dekker, New York, 1993; Omasa (2002), *J. Biosci. Bioeng.* 94:6 600-605).

Tabla 1. Genes de selección

Nombre del gen de selección	Amplificado con
Dihidrofolato Reductasa	Metotrexato (MTX)
Glutamina Sintetasa	Metionina sulfoximina (MSX)
Hipoxantina fosforribosil transferasa	Aminopterina, hipoxantina, y timidina
Treonil ARNt Sintetasa	Borrelidina
Na, K-ATPasa	Ouabaína
Asparagina Sintetasa	Albizina o Beta-aspartil hidroxamato
Ornitina Descarboxilasa	alfa-difluorometilornitina (DFMO)
Inosina-5'-monofosfato deshidrogenasa	Ácido Micofenólico
Adenosina Desaminasa	Adenosina, Alanosina, 2'desoxicoformicina

(continuación)

Nombre del gen de selección	Amplificado con
Timidilato Sintetasa	Fluorouracilo
Aspartato Transcarbamilasa	N-Fosfonacetil-L-Aspartato (PALA)
Metalotioneína	Cadmio
Adenilato Desaminasa (1,2)	Adenina, Azaserina, Coformicina
UMP-Sintetasa	6-azauridina, pirazofurano
Ribonucleótido Reductasa	hidroxiurea, motexafin gadolinio, fludarabina, cladribina, gemcitabina, tezacitabina, triapina.

5 Deben tenerse en cuenta varias consideraciones cuando se selecciona un sitio diana específico para la inserción de una secuencia de interés en un locus que se puede amplificar. Primero, el sitio de inserción seleccionado debe co-
amplificarse con el gen bajo selección. En muchos casos, ya existen datos experimentales en la técnica que delimita
la cantidad de secuencia cromosómica flanqueante que se co-amplifica con un gen indicador de interés. Estos datos,
que definen con precisión la extensión del locus que se puede amplificar, existen para DHFR de CHO (Ma y col. (1988),
Mol Cell Biol. 8(6):2316-27), DHFR humano (Morales y col. (2009), Mol Cancer Ther. 8(2):424-432), y GS de CHO
10 (Sanders y col. (1987), Dev Biol Stand. 66:55-63). Cuando dichos datos no existen aún en la técnica, los inventores
predijeron que las secuencias de ADN <100.000 pares de bases corriente arriba o corriente abajo del gen de selección
que codifica la secuencia es probable que se co-amplifiquen. Por lo tanto, estas regiones podrían ser sitios adecuados
para el direccionamiento de la inserción de una secuencia de interés.

15 Segundo, los sitios diana se deberían seleccionar de manera que no tuvieran mucho impacto en la función del gen de
selección (por ejemplo, el gen endógeno de la DHFR, GS, o HPRT). Debido a que la amplificación necesita una copia
funcional del gen de selección, se deberían evitar los sitios de inserción en el promotor, exones, intrones, señales de
poliadenilación, u otras secuencias reguladoras que, si se alteran, impactaría enormemente en la transcripción o
traducción del gen de selección. Por ejemplo, el documento WO 2008/059317 desvela meganucleasas que escinden
20 los sitios de ADN diana en el gen HPRT. Hasta el punto en el que el documento WO 2008/059317 desvela la inserción
de genes en el locus del HPRT, enseña de que la secuencia codificante del gen de HPRT se debería alterar en el
proceso de inserción del transgén para facilitar la selección para el direccionamiento apropiado utilizando 6-tioguanina.
La 6-tioguanina es un análogo de nucleótido tóxico que destruye las células que tienen una actividad funcional de
HPRT. Debido a que las células producidas de acuerdo con el documento WO 2008/059317 no tienen actividad de
25 HPRT, no amplificarán un transgén insertado en respuesta al tratamiento con un inhibidor de HPRT, y, por lo tanto, no
se pueden utilizar en la presente invención. Para la presente invención, a menos de que ya se conozcan los límites
precisos de todas las secuencias reguladoras para un gen de selección particular, se deberían seleccionar sitios de
inserción >1.000 pares de bases, >2.000 pares de bases, >3.000 pares de bases, >4.000 pares de bases, o,
preferentemente, >5.000 pares de bases, corriente arriba o corriente abajo de la secuencia codificante del gen. Sin
embargo, si se conocen la localización de las secuencias reguladoras, la secuencia de interés se puede insertar
30 inmediatamente adyacente a la secuencia regulador más 5' o 3' (por ejemplo, inmediatamente 3' de la señal de
poliadenilación).

Por último, los sitios diana se deberían seleccionar para que no alteren otros genes cromosómicos que pueden ser
importantes para la fisiología celular normal. En general, los sitios de inserción genética deberían de estar a >1.000
pares de bases, >2.000 pares de bases, >3.000 pares de bases, >4.000 pares de bases, o, preferentemente, >5.000
pares de bases de cualquier secuencia genética codificante.

35 Se describen esquemáticamente distintos procedimientos de la invención en las figuras de la siguiente manera:

La Figura 1 representa una estrategia general para dirigir una secuencia de interés a un locus que se puede
amplificar. En la primera etapa, una endonucleasa específica del sitio introduce una rotura de doble cadena en el
ADN cromosómico de una célula en un sitio que está proximal a un gen de selección endógeno. El ADN
40 cromosómico escindido se somete entonces a recombinación homóloga con una molécula de ADN donante que
comprende una secuencia de interés flanqueada por secuencias de ADN homólogas de las secuencias que
flanquean la secuencia de reconocimiento de la endonucleasa del sitio diana. Como resultado, se inserta la
secuencia de interés en el ADN cromosómico de la célula adyacente al gen de selección endógeno. La célula
modificada se cultiva entonces en presencia de uno o más compuestos que inhiben la función del gen de selección
para inducir un aumento del número de copias (es decir, una amplificación) del gen de selección. La secuencia de
45 interés, que está unida genéticamente al gen de selección, se co-amplificará con el gen de selección. El resultado
es una línea celular transgénica estable que comprende múltiples copias de la secuencia de interés.

La Figura 2(A) representa un esquema del locus de DHFR de CHO que muestra una región preferida para el
direccionamiento de una secuencia de interés 5.000-60.000 pares de bases corriente abajo del gen de DHFR. La
Figura 2(B) representa un esquema del locus GS de CHO que muestra una región preferida para el
50 direccionamiento de una secuencia de interés 5.000-55.000 pares de bases corriente abajo del gen de GS. Los

promotores se muestran como flechas. Los exones se muestran como rectángulo, con los exones no codificantes en blanco y los exones codificantes de proteínas en gris.

La Figura 3 representa una estrategia para la inserción de una secuencia de interés en un locus que se puede amplificar en un procedimiento de dos etapas que implica una secuencia diana integrada previamente. En la primera etapa, el ADN cromosómico de una célula se escinde por una endonucleasa específica del sitio en un sitio que está proximal a un gen de selección. El ADN cromosómico escindido se somete entonces a recombinación homóloga con una molécula de ADN donante que comprende una secuencia diana exógena flanqueada por secuencias de ADN homólogas a las secuencias flanqueantes del sitio diana endógeno. Esto da como resultado la inserción de una secuencia diana modificada nueva en el ADN cromosómico de la célula proximal al gen de selección. Se puede direccionar posteriormente una secuencia de interés proximal al mismo gen de selección utilizando una nucleasa, integrasa, transposasa, o recombinasa que reconozca específicamente la secuencia diana modificada integrada previamente. La célula modificada se cultiva entonces en presencia de uno o más compuestos que co-amplifican el gen de selección y la secuencia de interés.

La Figura 4 representa una estrategia para la inserción de una secuencia diana modificada en un gen de selección (por ejemplo, DHFR) con la retirada concomitante de una parte del gen de selección. Se utilizó primero una endonucleasa específica del sitio para escindir el ADN cromosómico de la célula, proximal a o en la secuencia del gen de selección. Como se muestra en la figura, el sitio diana endógeno está entre los exones 2 y 3 del gen de DHFR de CHO (aunque el sitio diana podría estar en cualquier otro intrón o exón, y el gen de selección podría ser cualquier gen sometido a amplificación). El ADN cromosómico se somete entonces a recombinación homóloga con un primer ADN donante ("ADN donante nº 1") de manera que la secuencia del primer ADN donante se inserta en el ADN cromosómico de la célula. Como se muestra en la figura, esto da como resultado la sustitución del promotor y los dos primeros exones de DHFR por la secuencia diana modificada nueva (aunque el primer ADN donante podía remplazar más o menos del ADN cromosómico, tal como solo una parte del exón). Si dicha sustitución se hace en todos los alelos DHFR de una célula, la línea celular resultante es DHFR (-/-). Se puede dirigir posteriormente una secuencia de interés proximal a un gen de selección en la línea celular utilizando una endonucleasa, integrasa, transposasa, o recombinasa que reconoce la secuencia diana modificada. Como se muestra en la figura, el segundo ADN donante ("ADN donante nº 2") comprende una secuencia de interés, así como un promotor y los primeros dos exones de DHFR. El direccionamiento apropiado de esta segunda molécula de ADN donante da como resultado la inserción de la secuencia de interés en la secuencia diana modificada mientras que se reconstituye simultáneamente un gen DHFR funcional. Por lo tanto, las líneas celulares direccionadas apropiadamente serán DHFR+ y se pueden seleccionar utilizando medios deficientes en hipoxantina/timidina. Además, los que se representa en un diagrama aquí para la DHFR se puede aplicar a cualquier gen de selección en un locus que se pueda amplificar.

La Figura 5 representa una estrategia para la inserción de una secuencia diana modificada en un locus que se puede amplificar con una alteración concomitante de la secuencia codificante de un gen de selección. Primero se utiliza una endonucleasa específica del sitio para escindir el ADN cromosómico de la célula en la secuencia codificante del gen de selección. Como se muestran en la figura, el sitio diana endógeno está en el tercer exón del gen GS de CHO. El ADN cromosómico se somete entonces a recombinación genética con un primer ADN donante ("ADN donante nº 1") de manera que la secuencia del primer ADN donante se inserta en el ADN cromosómico de la célula. Esto da como resultado la inserción de una secuencia diana modificada nueva en la secuencia codificante de GS. Si se produce dicha inserción en ambos alelos del gen GS y da como resultado una mutación con cambio de fase u otra manera de alteración de la función del gen GS, la línea celular resultante será GS (-/-). Una secuencia de interés se puede direccionar posteriormente en proximal al locus que se puede amplificar de la línea celular utilizando una endonucleasa, integrasa, transposasa, o recombinasa que reconozca la secuencia diana modificada. Como se muestra en la figura, un segundo ADN donante ("ADN donante nº 2") comprende una secuencia de interés unida operativamente a un promotor así como la parte 3' de la secuencia codificante de GS que comprende los exones 3, 4, 5, y 6 (La figura muestra los exones 3, 4, 5 y 6 unidos en una única secuencia de nucleótidos (es decir, con intrones retirados), pero también se podría utilizar una secuencia que incluya o intrones de origen natural o uno o más intrones artificiales. El direccionamiento apropiado de la segunda molécula de ADN donante da como resultado la inserción de la secuencia de interés en la secuencia diana modificada a la vez que se reconstituye simultáneamente un gen de GS funcional. Por lo tanto, las líneas direccionadas apropiadamente serán GS+ y se pueden seleccionar utilizando medios deficientes en L-glutamina. Además, la secuencia de interés se puede co-amplificar con el gen GS utilizando una selección con MSX. La estrategia representada mediante diagramas aquí para GS se puede aplicar a cualquier gen de selección en un locus que se pueda amplificar.

La Figura 6 representa una estrategia para la inserción de una secuencia diana modificada en un locus que se puede amplificar con la alteración concomitante del procesamiento del ARNm de un gen de selección. Primero se utiliza una endonucleasa para escindir el ADN cromosómico de la célula en un intrón del gen de selección. Como se representa, el sitio diana endógeno está en el intrón entre el tercer y cuarto exones codificantes del gen de GS de CHO. El ADN cromosómico se somete entonces a recombinación homóloga con un ADN donante nº 1 de manera que la secuencia del ADN donante se inserta en el ADN cromosómico de la célula. Esto da como resultado la inserción de una nueva secuencia diana modificada en la secuencia codificante de GS con una secuencia adicional que produce que el ARNm de GS se procese incorrectamente. Como se representa, esta secuencia adicional comprende un fuerte receptor de corte y empalme. Si dicha inserción se produce en ambos alelos del gen GS, el receptor artificial de corte y empalme producirá que el ARNm de GS se corte y empalme incorrectamente, dando como resultado una pérdida de expresión de GS y la necesidad de crecimiento en cultivos que contengan L-glutamina. Se puede dirigir posteriormente una secuencia de interés a locus que se puede

amplificar en la línea celular utilizando una endonucleasa, integrasa, transposasa, o recombinasa que reconozca la secuencia diana modificada. Como se representa mediante diagramas, el ADN donante nº 2 comprende una secuencia de interés unida operativamente a un promotor, así como la parte 3' de la figura muestra los exones 4, 5, y 6 unido a una secuencia de nucleótidos única. (La figura muestra los exones 4, 5, y 6 unido en una secuencia de nucleótidos única (es decir, con los intrones retirados), pero también se podría emplear una secuencia que incluye intrones de origen natural o uno o más intrones artificiales). El direccionamiento apropiado de esta molécula de ADN donante nº 2 da como resultado la inserción de la secuencia de interés en la secuencia diana modificada mientras se reconstituye simultáneamente un gen GS funcional. Por lo tanto, las líneas celulares direccionadas apropiadamente serán GS+ y se pueden seleccionar utilizando medios deficientes en L-glutamina y la secuencia de interés se puede co-amplificar con el gen de GS utilizando una selección con MSX. La estrategia representada mediante diagramas aquí para GS se puede aplicar a cualquier gen de selección en un locus que se pueda amplificar.

La Figura 7(A) representa un ensayo de recombinación repetida directa de la actividad de endonucleasa específica del sitio. Se produce un plásmido indicador que comprende los dos tercios 5' del gen GFP ("GF"), seguido por una secuencia de reconocimiento por la endonucleasa, seguido por los dos tercios 3' del gen GFP ("FP"). Las células de mamífero se transfectaron con este plásmido indicador, así como el gen que codifica la endonucleasa. La escisión de la secuencia de reconocimiento por la endonucleasa estimula la recombinación homóloga entre repeticiones directas del gen de GFP para restaurar la función de GFP. Las células GFP+ se pueden contar y/o clasificar en un citómetro de flujo.

La Figura 7(B) representa los resultados del ensayo de la Figura 7(A) según se aplica a las meganucleasas CHO-23/24 y CHO-51/52. Las barras claras indican el porcentaje de células GFP+ cuando se transfectan las células con el plásmido indicador solo (-endonucleasa). Las barras oscuras indican el porcentaje de células GFP+ cuando se co-transfectan con el plásmido indicador y el gen de meganucleasas correspondiente (+endonucleasa). El ensayo se llevó a cabo por triplicado y se muestra la desviación típica.

La Figura 7(C) representa el alineamiento de secuencias obtenidas de las células CHO transfectadas con el ARNm que codifica la meganucleasa CHO-23/24. La secuencia superior es de una célula CHO de tipo silvestre (TS) con la secuencia de reconocimiento para la CHO-23/24 subrayada.

La Figura 7(D) representa el alineamiento de secuencias obtenidas de las células CHO transfectadas con el ARNm que codifica la meganucleasa CHO-51/52. La secuencia superior es de una célula CHO de tipo silvestre (TS) con la secuencia de reconocimiento para la CHO-51/52 subrayada.

La Figura 8(A) representa una estrategia para la inserción de una secuencia de ADN exógena en el locus de DHFR de CHO utilizando la meganucleasa CHO-51/52. Las células CHO se co-transfectaron con el ARNm que codifica la CHO-51/52 y un plásmido donante que comprendía un sitio EcoRI flanqueado por 543 pares de bases de la secuencia de ADN homóloga de la región corriente arriba del sitio de reconocimiento de CHO-51/52 y 461 pares de bases de la secuencia de ADN homóloga de la región corriente abajo del sitio de reconocimiento de CHO-51/52. 48 horas después de la transfección, el ADN genómico se aisló y se sometió a PCR utilizando cebadores específicos para la región corriente abajo del locus de DHFR (flechas discontinuas).

La Figura 8(B) representa los productos de la PCR que se clonaron en el pUC-19 y 48 clones plasmídicos individuales y se digirieron con EcoRI y se visualizaron en gel de agarosa. 10 plásmidos (calles numeradas) daban lugar a un fragmento de restricción de 647 bases, consistente con la escisión de un primer sitio EcoRI en el vector pUC-19 y un segundo sitio EcoRI en el fragmento de PCR clonado. Estos 10 plásmidos se secuenciaron para confirmar que albergaban un fragmento de la PCR que comprendían una parte corriente abajo del locus DHFR con un sitio de restricción EcoRI insertado en la secuencia de reconocimiento. Este patrón de restricción no se observó cuando se transfectaron las células con el plásmido donante solo.

La Figura 9(A) representa una estrategia para la inserción de una secuencia diana modificada en el locus de DHFR de CHO utilizando la meganucleasa CHO-23/24. Las células CHO se co-transfectaron con el ARNm que codifica la CHO-23/24 y un plásmido donante que comprende, en la orientación 5' a 3', un promotor SV40, un codón de inicio ATG, un sitio FRT, y un gen de resistencia a la Zeocina (Zeo). Las células resistentes a zeocina se clonaron por dilución limitante y se exploraron por PCR para identificar una línea celular clónica en la que se integra en la secuencia del plásmido donante en el sitio de reconocimiento para la CHO-23/24. Después de la expansión, esta línea celular se co-transfectó con un primer plásmido que codifica la Flp recombinasa unida operativamente a un promotor y un segundo plásmido (plásmido donante nº 2) que comprende un gen de la GFP bajo el control de un promotor CMV, un sitio FRT, y un gen de resistencia a la higromicina (Hig) que carece de un codón de inicio. Los sitios FRT entre la recombinación mediada por Flp resultaba en la integración de la secuencia del plásmido donante nº 2 en la secuencia diana modificada (es decir, el sitio FRT) de manera que se producía un casete de expresión genética Hig funcional. La Figura 9(B) representa los productos de PCR de clones resistentes a la higromicina producidos como en (A) que se clonaron por dilución limitante. Se extrajo el ADN genómico de 24 clones individuales y se amplificó por PCR utilizando un primer cebador en el locus DHFR y un segundo cebador en el gen Hig (líneas discontinuas). Los 24 clones daban lugar a un producto de la PCR consistente con la inserción del gen Hig en la secuencia diana modificada. La Figura 9(C) representa la expresión de los 24 clones producidos en (B) utilizando citometría de flujo. Se descubrió que todos los genes expresan altos niveles de GFP con una relativamente poca variabilidad clon a clon.

Figura 10. Se produjo una línea de CHO que expresa GFP integrando un casete de expresión del gen GFP en el locus DHFR utilizando una estrategia de secuencia diana modificada como se muestra en la Figura 9. Esta línea celular se cultivó entonces en MTX como se describe en el Ejemplo 2 para amplificar el gen GFP integrado. (A) gráficos de citometría de flujo que muestran la intensidad de GFP en el eje Y. En la línea celular pre-MTX, la

intensidad de GFP tenía un promedio de aproximadamente 2×10^3 mientras que en la línea celular cultivada en 250 nM de MTX, era visible una subpoblación distinta (en un círculo) en el que la intensidad de GFP se aproxima a 10^4 . (B) Las líneas celulares tratadas con MTX se clasificaron por FACS para identificar células individuales que expresan grandes cantidades de GFP. Cinco de dichas células de alta expresión se expandieron y se determinó la intensidad de GFP por citometría de flujo. Se descubrió que los cinco clones tenían un aumento significativo de la expresión de GFP con respecto a la línea celular pre-MTX. (C) El ADN genómico se aisló a partir de cinco líneas celulares clónicas producidas en (B) y se sometieron a PCR cuantitativa utilizando un par de cebadores específico para el gen de GFP. Se descubrió que los cinco clones de alta expresión tenían significativamente más copias del gen GFP que la línea celular pre-MTX. Estos resultados demuestran que el número de copias y el nivel de expresión de un transgén integrado corriente abajo del DHFR de CHO puede amplificarse en respuesta al tratamiento con MTX.

Figura 11. (A) Un ensayo de recombinación repetida directa, como en la Figura 5A. (B) El ensayo en (A) aplicado a las meganucleasas CHO-13/14 y CGS-5/6. Las barras claras indican el porcentaje de células GFP+ cuando se transfectan las células con el plásmido indicador solo (-endonucleasa). Las barras oscuras indican el porcentaje de células GFP+ cuando las células se co-transfectan con un plásmido indicador y el gen correspondiente de la meganucleasa (+endonucleasa). El ensayo se llevó a cabo por triplicado y se muestra la desviación típica. (C) Alineamiento de secuencias obtenidas de células CHO transfectadas con ARNm que codifica la meganucleasas CGS-5/6. La secuencia superior es de la célula CHO de tipo silvestre (TS) con la secuencia de reconocimiento para la CGS-5/6 subrayada. Los puntos indican bases eliminadas. Las bases que están en cursiva y negrita son mutaciones o inserciones puntuales con respecto a la secuencia de tipo silvestre. Nótese que se espera que las mutaciones observadas en al menos los clones 6d4, 6g5, 3b7, 3d11, 3e5, 6f10, 6hH8, 6d10, 6d7, 3g8 y 3a9 inactive la función del gen GS.

2.1.1 Direccionamiento genético al locus DHFR de CHO

El locus DHFR de CHO se representa en forma de diagrama en la Figura 2A. El locus comprende el gen de DHFR que codifica una secuencia en 6 exones que abarcan ~24.500 pares de bases. El gen Msh3 se localiza inmediatamente corriente arriba del DHFR y se transcribe divergentemente del mismo promotor de DHFR. Se puede encontrar un hipotético gen, 2BE2121, ~65.000 pares de bases corriente abajo de la secuencia codificante de DHFR. Por lo tanto, hay una región de ~65.000 pares de bases corriente abajo del gen DHFR que no alberga ningún gen conocido y es una localización adecuada para dirigir la inserción de secuencias de interés. Los sitios diana para la inserción de una secuencia de interés, en general no se debería seleccionar si son <1.000 pares de bases, y preferentemente no <5.000 pares de bases de los genes DHFR o 2BE2121. Esto limita la ventana de sitios diana preferidos en la región de 1.000-60.000 pares de bases corriente abajo de la secuencia codificante de la DHFR. La secuencia de esta región se proporciona como la SEQ ID NO: 2.

Los loci de DHFR humano y de ratón tienen una organización similar al locus de CHO. En ambos casos, el gen Msh3 está inmediatamente corriente arriba del DHFR, pero hay una gran área que carece de secuencias codificantes corriente abajo del DHFR. En los seres humanos el gen ANKRD34B tiene ~55.000 pares de bases corriente abajo del DHFR mientras que el gen ANKRD34B tiene ~37.000 pares de bases corriente abajo del DHFR de ratón. Por lo tanto, la región corriente abajo del DHFR es una localización apropiada para insertar genes de interés en líneas celulares y células CHO humanas y de ratón. Además, los casetes de expresión genética insertados en esta región se expresarán con un nivel alto, serán resistentes al silenciamiento genético y capaces de amplificarse mediante el tratamiento con MTX. Los procedimientos para la amplificación del locus DHFR en células CHO se conocen en la técnica (véase, por ejemplo, Kellems, ed., Gene amplification in mammalian cells: a comprehensive guide. Marcel Dekker, New York, 1993) y normalmente implican una concentración creciente gradualmente de la concentración de MTX en los medios de cultivo desde 0 hasta 0,8 mM durante un periodo de varias semanas.

2.1.2 Direccionamiento genético al locus de GS

Los loci en CHO humana y de ratón de la glutamina sintetasa (también conocida como "glutamato-amonio ligasa" o "GluL") comparten una organización común (Fig. 2B). El gen TEDDM1 está inmediatamente corriente arriba de GS en todas las especies (~5.000 pb corriente arriba en el caso de los seres humanos, ~7.000 pb corriente arriba en el caso del ratón y CHO). El gen corriente abajo más cercano está, sin embargo, ~46.000 pb más lejos en el caso del ser humano y ~117.000 pb más lejos en el caso del ratón y las CHO. Por lo tanto, los inventores predijeron que la región cromosómica de 1.000-41.000 pb o 5.000-41.000 pb corriente abajo del GS en células humanas y 1.000-100.000 pb, o 5.000-100.000 pb corriente abajo del GS en el ratón y células de CHO eran localizaciones apropiadas para dirigir la inserción de secuencias de interés. Debido a que los sitios de ADN distales a la secuencia codificante de GS es más probable que sean susceptibles al silenciamiento genético, se prefiere una región cromosómica de 5.000-60.000 pb corriente abajo de GS para dirigir la inserción de una secuencia de interés tanto en ratón o células CHO. La secuencia de esta región del genoma de CHO se proporciona como la SEQ ID NO: 3. Los casetes de expresión genética insertados en esta región se expresarán a un nivel alto, serán resistentes al silenciamiento genético y capaces de amplificarse mediante el tratamiento con MSX. Las regiones menos preferidas incluyen la región cromosómica entre los genes TEDDM1 y GS o la región <10.000 pb corriente abajo de TEDDM1 (véase la Figura 2B). Los procedimientos para la amplificación del locus de GS se conocen en la técnica (Bebbington y col. (1992), Biotechnology (N Y). 10(2):169-75).

2.2 Endonucleasas modificadas para el direccionamiento genético

Se puede insertar una secuencia de interés en un locus que se pueda amplificar utilizando una endonucleasa específica del sitio modificada. Los procedimientos para la generación de endonucleasas específicas del sitio que puedan dirigir roturas de ADN en loci predeterminados en un genoma se conocen en la técnica. Estas incluyen las nucleasas en dedos de zinc (Le Provost y col. (2010), Trends Biotechnol. 28(3):134-41), nucleasas efectoras TAL (Li y col. (2011), Nucleic Acids Res. 39(1):359-72), y meganucleasas modificadas (documentos WO 2007/047859; WO 2007/049156; WO 2009/059195). En una realización, la invención proporciona meganucleasas modificadas derivadas de I-Crel que se pueden utilizar para dirigir la inserción de un gen de interés en un locus que se pueda amplificar. Los procedimientos para producir dichas meganucleasas modificadas se conocen en la técnica (véase, por ejemplo, los documentos WO 2007/047859; WO 2007/049156; WO 2009/059195). En realizaciones preferidas, se utiliza una meganucleasa de “cadena sencilla” para dirigir la inserción genética en una región que se puede amplificar del genoma. Los procedimientos para la producción de dichas meganucleasas de “cadena sencilla” se conocen en la técnica (véase, por ejemplo, los documentos WO 2009/059195 y WO 2009/095742). En algunas realizaciones, la nucleasa modificada se fusiona a una señal de localización nuclear (NLS) para facilitar la captación nuclear. Ejemplos de señales de localización nuclear incluyen la NLS de SV40 (secuencia de aminoácidos MAPKKKRV) que se pueden fusionar en el extremo C o, preferentemente, en el extremo N de la proteína. Además, se puede marcar una nucleasa modificada con un epítipo peptídico (por ejemplo, un epítipo HA, FLAG o Myc) para monitorizar los niveles de expresión o localización o para facilitar la purificación.

2.3 Líneas celulares modificadas con secuencias de interés dirigidas a loci que se puedan amplificar

En algunas realizaciones, la invención proporciona procedimientos para la utilización de nucleasas modificadas en el direccionamiento de la inserción de transgenes en loci que se puedan amplificar en células de mamífero cultivadas. Este procedimiento tiene dos componentes primarios: (1) una nucleasa modificada; y (2) una molécula donante de ADN que comprende una secuencia de interés. El procedimiento comprende poner en contacto el ADN de la célula con la nucleasa modificada para crear una rotura de la doble cadena del ADN en una secuencia de reconocimiento endógena de un locus que se pueda amplificar seguido por la inserción de la molécula de ADN donante en el sitio de la rotura del ADN. Dicha inserción del ADN donante se facilita por la maquinaria de reparación de ADN y puede producirse por la ruta de unión de extremos no homólogos o por recombinación homóloga (Figura 1).

La nucleasa modificada se puede suministrar a la célula en forma de proteína o, preferentemente, como un ácido nucleico que codifique la nucleasa modificada. Dicho ácido nucleico puede ser un ADN (por ejemplo, un ADN plasmídico circular o que se ha hecho lineal o productos de PCR) o un ARN. Para las realizaciones en las que la secuencia codificante de nucleasa modificada se suministra en forma de ADN, debería estar unida operativamente a un promotor para facilitar la transcripción del gen de la nucleasa modificada. Los promotores de mamífero adecuados para la invención incluyen los promotores constitutivos tales como el promotor temprano de citomegalovirus (CMV) (Thomsen y col. (1984), Proc Natl Acad Sci USA. 81(3):659-63) o el promotor temprano del SV40 (Benoist y Chambon (1981), Nature. 290(5804):304-10), así como los promotores inducibles tales como el promotor inducible por tetraciclina (Dingermann y col. (1992), Mol Cell Biol. 12(9):4038-45).

En algunas realizaciones, se suministra a la célula el ARNm que codifica la nucleasa modificada debido a que esto reduce la probabilidad de que el gen que codifica la nucleasa modificada se integre en el genoma de la célula. Dicho ARNm que codifica la nucleasa modificada se puede producir utilizando procedimientos conocidos en la técnica tales como la transcripción *in vitro*. En algunas realizaciones, el ARN está protegido utilizando 7-metil-guanosina. En algunas realizaciones, el ARNm puede estar poliadenilado.

Las nucleasas proteicas modificadas purificadas se pueden suministrar a las células para escindir el ADN genómico, lo que permite la recombinación homóloga o la unión de extremos no homólogos en el sitio de escisión con una secuencia de interés, mediante una variedad de diferentes mecanismos conocidos en la técnica. Por ejemplo, la nucleasa proteica recombinante se puede introducir en la célula mediante técnicas que incluyen, pero no se limitan a, microinyección o transfecciones liposómica (véase, por ejemplo, Lipofectamine™, Invitrogen Corp., Carlsbad, CA). La formulación de liposomas se puede utilizar para facilitar la fusión de la bicapa lipídica con una célula diana, permitiendo de esta manera que los contenidos del liposoma o las proteínas asociadas con su superficie se introduzcan en la célula. De manera alternativa, la enzima se puede fusionar con un péptido de captación apropiado tal como el de la proteína TAT del VIH para dirigir la captación celular (véase, por ejemplo, Hudecz y col. (2005), Med. Res. Rev. 25: 679-736).

De manera alternativa, las secuencias genéticas que codifican la nucleasa proteica modificada se insertan en un vector y se transfectan en una célula eucariota utilizando técnicas conocidas en la técnica (véase, por ejemplo, Ausubel y col., Current Protocols in Molecular Biology, Wiley 1999). La secuencia de interés se puede introducir en el mismo vector, un vector diferente, o por otros medios conocidos en la técnica. Ejemplos no limitantes de vectores para la transfección de ADN incluyen vectores víricos, plásmidos, cósmidos y vectores YAC. La transfección de secuencias de ADN se puede conseguir mediante una variedad de procedimientos conocidos por los expertos en la técnica. Por ejemplo, se utilizan liposomas e inmunoliposomas para suministrar secuencias de ADN a las células (véase, por ejemplo, Lasic y col. (1995), Science 267: 1275-76). Además, se pueden utilizar virus para introducir vectores en las células (véase, por ejemplo, la Pat. de EE. UU. N.º 7.037.492). De manera alternativa, las estrategias de transfección

se pueden utilizar de manera que los vectores se introduzcan como ADN desnudo (véase, por ejemplo, Rui y col. (2002), *Life Sci.* 71(15): 1771-8).

Los procedimientos generales para el suministro de ácidos nucleicos en las células incluyen: (1) procedimientos químicos (Graham y col. (1973), *Virology* 54(2):536-539; Zatloukal y col. (1992), *Ann. N. Y. Acad. Sci.*, 660:136-153; (2) procedimientos físicos tales como la microinyección (Capecchi (1980), *Cell* 22(2):479-488, electroporación (Wong y col. (1982), *Biochim. Biophys. Res. Commun.* 107(2):584-587; Fromm y col. (1985), *Proc. Nat'l Acad. Sci. USA* 82(17):5824-5828; Pat. de EE. UU. N.º 5.384.253) e inyección balística (Johnston y col. (1994), *Methods Cell. Biol.* 43(A): 353-365; Fynan y col. (1993), *Proc. Nat'l Acad. Sci. USA* 90(24): 11478-11482); (3) vectores víricos (Clapp (1993), *Clin. Perinatol.* 20(1): 155-168; Lu y col. (1993), *J. Exp. Med.* 178(6):2089-2096; Eglitis y col. (1988), *Avd. Exp. Med. Biol.* 241:19-27; Eglitis y col. (1988), *Biotechniques* 6(7):608-614); y (4) mecanismos mediados por un receptor (Curiel y col. (1991), *Proc. Nat'l Acad. Sci. USA* 88(19):8850-8854; Curiel y col. (1992), *Hum. Gen. Ther.* 3(2):147-154; Wagner y col. (1992), *Proc. Nat'l Acad. Sci. USA* 89 (13):6099-6103). En algunas realizaciones preferidas, el ARNm protegido con 7-metil-guanosina que codifica la nucleasa modificada se suministra a las células utilizando electroporación.

La molécula donante de ADN comprende un gen de interés unido operativamente a un promotor. En muchos casos, una molécula donante puede comprender múltiples genes unidos operativamente al mismo o diferentes promotores. Por ejemplo, las moléculas donantes que comprenden casetes de expresión de anticuerpos monoclonales pueden comprender un gen que codifica la cadena pesada del anticuerpo y un segundo gen que codifica la cadena ligera del anticuerpo. Ambos genes pueden estar bajo el control de diferentes promotores o pueden estar bajo el control del mismo promotor utilizando, por ejemplo, un sitio interno de entrada en el ribosoma (IRES). Las moléculas donantes pueden comprender también un gen indicador genético unido operativamente a un promotor para facilitar la identificación de las células transgénicas. Dichos indicadores genéticos se conocen en la técnica e incluyen los genes de la neomicina fosfotransferasa (NEO), hipoxantina fosforribosiltransferasa (HPRT), glutamina sintetasa (GS), dihidrofolato reductasa (DHFR) y la higromicina fosfotransferasa (HIG).

En algunas realizaciones, las moléculas donantes de ADN comprenderán adicionalmente secuencias flanqueantes homólogas de las secuencias diana del ADN de la célula. Dichas secuencias flanqueantes homólogas comprenden > 3 o, preferentemente, < 50 o más preferentemente > 200 o más preferentemente > 400 pares de bases de ADN que son idénticas o casi idénticas en secuencia al locus cromosómico reconocido por la nucleasa modificada (Figura 1). Dichas secuencias de ADN homólogas facilitan la integración de la secuencia donante de ADN en el locus que se puede amplificar por recombinación homóloga.

La molécula "donante" de ADN puede ser circular (por ejemplo, un ADN plasmídico) o lineal (por ejemplo, un plásmido que se ha hecho lineal o productos de una PCR). Los procedimientos para el suministro de moléculas de ADN se conocen en la técnica como se ha expuesto anteriormente.

En algunas realizaciones, el gen de nucleasa modificada y el ADN donante se llevan a cabo en moléculas de ácido nucleico separadas que se co-transfectan en las células o líneas celulares. Por ejemplo, el gen de la nucleasa modificada unida operativamente a un promotor puede transfectarse en forma de plásmido simultáneamente con una molécula de ADN donante diferente en forma de plásmido o de producto de PCR. En una realización alternativa, la nucleasa modificada se puede suministrar en forma de ARNm con una molécula de ADN donante en forma de plásmido o de producto de PCR. En una tercera realización, el gen de la nucleasa modificada y el ADN donante se llevan en la misma molécula de ADN, tal como un plásmido. En una cuarta realización, las células se co-transfectan con una nucleasa proteica modificada y una molécula de ADN en forma de plásmido o producto de PCR.

Después de la transfección con la nucleasa modificada y el ADN donante, normalmente se permite a las células que se recuperen de la transfección (24-72 horas) antes de clonarse utilizando procedimientos conocidos en la técnica. Los procedimientos comunes para la clonación de una línea celular modificada genéticamente incluyen "dilución limitante" en la que las células transfectadas se transfieren a las placas de cultivo tisular (por ejemplo, placas de 48 pocillos, 96 pocillos, etc.) a una concentración de < 1 célula por pocillo y se expanden en poblaciones clónicas. Otras estrategias de clonación incluyen sistemas de aislamiento/identificación de clones robóticos tal como ClonePix™ (Genetix, Molecular Devices, Inc., Sunnyvale, CA). Las líneas celulares clónicas se pueden entonces explorar para identificar líneas celulares en las que se ha integrado la secuencia de interés en el sitio diana pretendido. Las líneas celulares pueden explorarse fácilmente utilizando análisis moleculares conocidos en la técnica tales como PCR o Transferencia de Southern. Por ejemplo, se puede aislar el ADN genómico de una línea celular clónica y someterse a amplificación por PCR utilizando un primer cebador (en sentido de la cadena) que se hibrida a una secuencia de ADN en la secuencia de interés y un segundo cebador (de cadena antisentido) que se hibrida a una secuencia del locus que se puede amplificar. Si la molécula de ADN donante comprende una secuencia de ADN homóloga del sitio diana, es importante que el segundo cebador se diseñe para hibridarse a una secuencia en el locus que se puede amplificar que está más allá de los límites de homología que lleva la molécula donante para evitar resultados falsos positivos. De manera alternativa, las líneas celulares se pueden explorar en cuanto a la expresión de la secuencia de interés. Por ejemplo, si la secuencia de interés codifica una proteína secretada tal como un anticuerpo, el medio de cultivo se tiene que muestrear de las líneas celulares clónicas aisladas y se ensaya en cuanto a la presencia del anticuerpo proteico utilizando procedimientos conocidos en la técnica tales como Transferencia de Western o ensayo de inmunoabsorción ligado a enzimas (ELISA). Este tipo de exploración funcional se puede utilizar para determinar cuál

de estas líneas celulares transgénicas llevan la secuencia de interés dirigida al locus que se puede amplificar de interés, como se ha descrito anteriormente.

El procedimiento de la invención se puede utilizar en cualquier tipo de célula que es pueda transfectar y cultivar tal como líneas celulares inmortalizadas y células madre. En realizaciones preferidas, el procedimiento de la invención se utiliza para modificar genéticamente líneas celulares inmortalizadas que se utilizan comúnmente para la biofabricación. Esto incluye:

1. Las líneas celulares de hámster tal como células de riñón recién nacido de hámster (BHK) y todas las variantes de células ováricas de hámster chino (CHO), por ejemplo, CHO-K1, CHO-S (Invitrogen Corp., Carlsbad, CA), DG44, o Potelligent™ (Lonza Group Ltd., Basel, Suiza). Debido a que las secuencias genómicas de las diferentes líneas celulares de hámster son casi idénticas, una meganucleasa modificada que se pueda utilizar para practicar la invención en un tipo celular de hámster (por ejemplo, células BHK) se puede utilizar en general para practicar la invención en otro tipo celular de hámster (por ejemplo, CHO-K1).

2. Las líneas celulares de ratón tal como las células de hibridoma de ratón o mieloma de ratón (por ejemplo, NS0). Debido a que las secuencias del genoma de las diferentes líneas celulares de ratón son casi idénticas, una meganucleasa modificada que se pueda utilizar para practicar la invención en un tipo celular de ratón (por ejemplo, células de hibridoma de ratón) se puede utilizar en general para practicar la invención en otro tipo celular de ratón (por ejemplo, NS0).

3. Las líneas celulares humanas tales como las células renales embrionarias humanas (por ejemplo, HEK-293 o 293S) y las células retinianas humanas (por ejemplo, PER.C6). Debido a que las secuencias genómicas de diferentes líneas celulares humanas son casi idénticas, una meganucleasa modificada que se pueda utilizar para practicar la invención en un tipo celular humano (por ejemplo, células HEK-293) se pueden utilizar en general para practicar la invención en otro tipo celular humano (por ejemplo, PER.C6).

2.6 Líneas celulares modificadas previamente con secuencias diana modificadas en loci que se puedan amplificar

En una realización, la invención proporciona líneas celulares que están modificadas previamente para que comprendan una "secuencia diana modificada" que se pueda direccionar para la inserción genética en un locus que se puede amplificar en una línea celular de mamífero (Figura 3). Una secuencia diana modificada comprende una secuencia de reconocimiento para una enzima que es útil para la inserción de ácidos nucleicos transgénicos en secuencias de ADN cromosómico. Dichas secuencias diana modificadas pueden incluir secuencias de reconocimiento para meganucleasas modificadas derivadas de I-Crel (por ejemplo, las SEQ ID NO: 37-87 del documento WO 2009/076292), secuencias de reconocimiento para nucleasas en dedos de zinc, secuencias de reconocimiento para nucleasas efectoras TAL (TALEN), el sitio LoxP (SEQ ID NO: 4) que es reconocido por la Cre recombinasa, el sitio FRT (SEQ ID NO: 5) que es reconocido por la recombinasa FLP, el sitio attB (SEQ ID NO: 6) que es reconocidos por la recombinasa lambda, o cualquier otra secuencia de ADN conocida en la técnica que es reconocida por una endonucleasa, recombinasa, integrasa, o transposasa específicas del sitio que sean útiles para el direccionamiento de la inserción de ácidos nucleicos en un genoma. Por lo tanto, la invención permite a un experto en la técnica utilizar una nucleasa modificada (por ejemplo, una meganucleasas, nucleasa en dedos de zinc, o nucleasa efectora TAL) para insertar una secuencia diana modificada en un locus que se puede amplificar en una línea celular de mamífero. La línea celular resultante comprende dicha secuencia diana modificada en un locus que se puede amplificar se puede poner en contacto con la enzima apropiada (por ejemplo, una segunda meganucleasas modificada, una segunda nucleasa en dedos de zinc, una segunda nucleasa efectora TAL, una recombinasa, una integrasa, o una transposasa) para dirigir la inserción de un gen de interés en el locus que se pueda amplificar en la secuencia diana modificada. Esta estrategia en dos etapas puede ser ventajosa debido a la eficacia de la inserción genética que se puede conseguir utilizando una meganucleasas óptima, nucleasa en dedos de zinc, recombinasa, integrasa, o transposasa puede ser mayor que la que se puede conseguir utilizando la endonucleasa inicial (por ejemplo, meganucleasas o nucleasa en dedos de zinc) que escinda el sitio diana endógeno para promover la inserción de la secuencia diana modificada.

En una realización alternativa, se produce una línea celular insertando una secuencia diana modificada en un locus que se puede amplificar con la retirada concomitante de todo o una parte del marcador genético endógeno adyacente (Figura 4). Por ejemplo, se puede utilizar una meganucleasa modificada, nucleasa en dedos de zinc o nucleasa efectora TAL para retirar los dos primeros exones en ambos alelos del gen DHFR de CHO y sustituirlos con una secuencia diana modificada por una meganucleasa modificada, ZFN, TALEN, recombinasa, integrasa, o transposasa diferentes. La línea celular resultante será deficiente en DHFR e incapaz de crecer en ausencia de hipoxantina/timidina. De manera alternativa, por ejemplo, se puede utilizar una meganucleasa modificada, ZFN o TALEN para retirar el primer exón de ambos alelos del gen GS de CHO y sustituirlo con una secuencia diana modificada para una meganucleasa modificada, ZFN, TALEN, recombinasa, integrasa, o transposasa diferentes (Figura 4). La línea celular resultante será deficiente en GS e incapaz de crecer en ausencia de L-glutamina. Dicha línea celular es útil debido a que se puede insertar un gen de interés en la secuencia diana modificada en la línea celular modificada previamente mientras que simultáneamente se reconstituye el gen de selección (por ejemplo, DHFR o GS). Por lo tanto, es posible seleccionar los transfectantes que albergan el gen de interés en el locus que se puede amplificar utilizando las condiciones del medio que seleccionan las células DHFR+ o GS+.

En una realización alternativa, se produce una línea celular en la que la secuencia diana modificada se inserta en un locus que se puede amplificar con alternación del gen de selección (Figuras 5, 6). Esto se puede conseguir, por

ejemplo, utilizando una meganucleasas que reconoce un sitio de ADN en la secuencia codificante del gen de selección. Dicha meganucleasas se puede utilizar para dirigir la inserción de una secuencia diana modificada en una secuencia codificante del gen de selección dando como resultado una alteración de la función del gen, por ejemplo, introduciendo un cambio de fase (Figura 5). De manera alternativa, por ejemplo, se puede insertar una secuencia diana modificada en un intrón de la secuencia del gen de selección con una secuencia adicional que promueve un procesamiento impropio de la transcripción del gen de selección (Figura 6). Dichas secuencias que promueven un procesamiento impropio incluyen, por ejemplo, receptores de corte y empalme artificiales o señales de poliadenilación. Las secuencias receptoras de corte y empalme se conocen en la técnica (Clancy (2008), "RNA Splicing: Introns, Exons and Spliceosome," Nature Education 1:1) y comprenden normalmente una secuencia rica en pirimidina de 20-50 pares de bases seguida por una secuencia (C/T)AG(A/G). La SEQ ID NO: 33 es un ejemplo de secuencia receptores de corte y empalme. De la misma manera, las señales de poliadenilación se conocen en la técnica e incluyen, por ejemplo, la señal de poliadenilación de SV40 (SEQ ID NO: 34) y la señal de poliadenilación BGH (SEQ ID NO: 35). En algunas realizaciones, la línea celular resultante que alberga la nueva secuencia diana modificada en todos los alelos del gen de selección será deficiente en la función del gen debido a la mala transcripción o mala traducción del gen de selección y será capaz de crecer solo en condiciones permisivas. Por ejemplo, una secuencia diana modificada se puede insertar en la secuencia del gen de GS utilizando una meganucleasas dando como resultado una línea celular que es GS-/- que solo puede crecer en presencia de L-glutamina en el medio de cultivo. En una etapa posterior, se puede insertar un gen de interés en la secuencia diana modificada mientras se reconstituye simultáneamente el gen indicador (por ejemplo, DHFR o GS). Por lo tanto, es posible seleccionar los transfectantes que albergan el gen de interés en el locus que se puede amplificar utilizando las condiciones del medio que seleccionan las células DHFR+ o GS+.

2.5 Líneas celulares transgénicas para la biofabricación

En algunas realizaciones, la invención proporciona líneas celulares transgénicas adecuadas para la producción de fármacos proteicos. Dichas líneas celulares transgénicas comprenden una población de células en la que se inserta un gen de interés, unido operativamente a un promotor en el genoma de la célula en un locus que se pueda amplificar en el que el gen de interés codifica una proteína terapéutica. Ejemplos de terapéuticos proteicos incluyen: anticuerpos monoclonales, fragmentos de anticuerpo, eritropoyetina, activador del plasminógeno de tipo tisular, Factor VIII, Factor IX, insulina, factores estimulantes de colonias, interferones (por ejemplo, interferón- α , interferón- β e interferón- γ), interleucinas (por ejemplo, interleucina-2), vacunas, factor de necrosis tumoral, y glucocerebrosidasa. También se hace referencia a los terapéuticos proteicos como "biológicos" o "biofarmacéuticos".

Para utilizarse para la biofabricación, una línea celular transgénica de la invención debería someterse a: (1) la adaptación de medio de cultivo libre de suero en suspensión; y (2) amplificación del gen de interés. En algunas realizaciones, la invención se practica en líneas celulares adherentes que se pueden adaptar al medio de cultivo en suspensión para facilitar su mantenimiento en biorreactores de matraces con agitado o contenedores con agitado como es típica de la biofabricación industrial. Los procedimientos para la adaptación de células adherentes a crecer en suspensión se conocen en la técnica (Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Es necesario en general adaptar adicionalmente las líneas celulares de biofabricación a medios definidos químicamente que carecen de componentes derivados de animales (es decir, medios "libres de suero"). Los procedimientos para la preparación de dichos medios y la adaptación de líneas celulares a este se conocen en la técnica (Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). Dichos medios también se pueden adquirir comercialmente (por ejemplo, medio CD-3 para el mantenimiento de células CHO, disponible en Sigma-Aldrich, St. Louis, MO) y las células se pueden adaptar a este siguiendo las instrucciones del fabricante. En algunas realizaciones, la línea celular se adapta al cultivo en suspensión y/o el medio libre de suero antes de transfectarse con la nucleasa modificada.

Finalmente, los procedimientos para la amplificación genética se conocen en la técnica (Cell Culture and Upstream Processing, Butler, ed. (Taylor y Francis Group, New York, 2007)). En general, el procedimiento implica la adición de un inhibidor de un producto genético de selección al medio de cultivo para seleccionar las células que expresan anormalmente grandes cantidades del producto genético debido a los eventos de duplicación genética. En general, la concentración de inhibidor añadido al medio de cultivo se aumenta lentamente durante un periodo de semanas hasta que se alcanza el nivel deseado de amplificación. El inhibidor se retira entonces en general del medio antes de iniciar la ejecución de la bioproducción para evitar la posibilidad de que el inhibidor contamine la formulación proteica terapéutica. Por ejemplo, se puede amplificar el locus DHFR de CHO aumentando lentamente la concentración de MTX en el medio de cultivo de 0 mM hasta 0,8 mM durante un periodo de varias semanas. El locus GS puede, de la misma manera, amplificarse en el medio de 0 μ M hasta 100 μ M durante un periodo de varias semanas. Los procedimientos para la evaluación de la amplificación genética se conocen en la técnica e incluyen la Transferencia de Southern y PCR en tiempo real cuantitativa (rtPCR). Además, o como alternativa, los niveles de expresión de la secuencia de interés, que se correlaciona en general con el número de copas del gen, se pueden evaluar determinando la concentración de proteína terapéutica en el medio de cultivo utilizando procedimientos convencionales tales como Transferencia de Western o ELISA. Después de la producción de la línea celular, adaptación, y amplificación, las proteínas terapéuticas se pueden producir y purificar utilizando procedimientos que son convencionales en la industria biofarmacéutica.

Ejemplos

La presente invención se ilustra adicionalmente por los siguientes ejemplos, que no deben interpretarse como limitantes. Los expertos en la materia reconocerán, o podrán determinar, utilizando no más que la experimentación de rutina, numerosos equivalentes de las sustancias y procedimientos específicos descritos en el presente documento. Se pretende que dichos equivalentes estén abarcados en el ámbito de las reivindicaciones que siguen a los siguientes ejemplos. El Ejemplo 1 se refiere a meganucleasas modificadas que se pueden utilizar para dirigir la inserción de un gen de interés corriente abajo del gen DHFR en células CHO. El Ejemplo 2 se refiere a meganucleasas modificadas que se pueden utilizar para dirigir la inserción de una secuencia diana modificada en el gen DHFR de CHO con la retirada concomitante de los exones 1 y 2 de DHFR. El Ejemplo 2 también se refiere a meganucleasas modificadas que se pueden utilizar para dirigir la inserción de una secuencia diana modificada en el gen GS de CHO. El Ejemplo 3 se refiere a meganucleasas que se pueden utilizar para dirigir la inserción de un gen de interés corriente abajo del gen GS en células CHO.

Ejemplo 1

Inserción genética dirigida en el locus DHFR de CHO utilizando meganucleasas modificadas

La secuencia de ADN genómico de CHO de 10.000-55.000 pares de bases corriente abajo del gen DHFR se exploró para identificar sitios de ADN susceptibles para el direccionamiento con meganucleasas modificadas. Se seleccionaron dos sitios (SEQ ID NO: 7 y SEQ ID NO: 8) que tienen, respectivamente, 35.699 y 15.898 pares de bases corriente abajo de la secuencia codificante de DHFR (Tabla 2).

Tabla 2. Ejemplo de sitios de reconocimiento para meganucleasas modificadas en el locus DHFR de CHO

SEQ ID NO:	Secuencias del sitio diana	Localización relativa a la secuencia codificante de DHFR de CHO
7	5'-TAAGGCCTCATATGAAAATATA-3'	35.699 pb corriente abajo
8	5'-ATAGATGTCTTGCATACTCTAG-3'	15.898 pb corriente abajo

1. Meganucleasas que reconocen la SEQ ID NO: 7 y SEQ ID NO: 8

Se produjo una meganucleasa modificada (SEQ ID NO: 9) que reconoce y escinde la SEQ ID NO: 7. Esta meganucleasa se denominó "CHO-23/24". Se produjo una segunda meganucleasa modificada (SEQ ID NO: 10) que reconoce y escinde la SEQ ID NO: 8. Esta nucleasa se llamó "CHO-51/52". Cada meganucleasa comprende una señal de localización de nucleasa en el extremo N derivada de SV40, una primera subunidad de meganucleasas, una secuencia de engarce, y una segunda subunidad de meganucleasas.

2. Escisión específica del sitio del ADN plasmídico por las meganucleasas CHO-23/24 y CHO-51/52

La CHO-23/24 y CHO-51/52 se evaluaron utilizando un ensayo de recombinación repetida directa como se ha descrito previamente (Gao y col. (2010), Plant J. 61(1):176-87, Figure 7A). Un casete indicador GFP deficiente se generó un casete clonando primero un fragmento 5' de 480 pb del gen GFP en pcDNA5/FRT digerido con NheI/HindIII (Invitrogen Corp., Carlsbad, CA) dando como resultado el plásmido pGF. A continuación, se clonó un fragmento 3' de 480 pb del gen GFP (incluyendo una secuencia de 240 pb duplicada en el fragmento 5' de 480 pb) en el pGF digerido por BamHI/XhoI. El plásmido resultante, pGFFP, consiste en dos tercios 5' del gen GFP seguido por dos tercios 3' del gen GFP, interrumpido por 24 pb del pcDNA/poliengarce FRT. Para insertar los sitios de reconocimiento de meganucleasas, se hibridaron los oligonucleótidos complementarios que comprenden la secuencia en sentido y antisentido de cada sitio de reconocimiento y se ligaron en el pGFFP digerido con HindIII/BamHI.

Las secuencias codificantes de las meganucleasas modificadas se insertaron en el vector de expresión de mamífero pCP bajo el control de un promotor constitutivo (CMV). Las células de ovario de hámster chino (CHO) con una confluencia aproximada del 90 % se transfectaron en placas de 96 pocillos con 150 ng de plásmido indicador pGFFP y 50 ng de vector de expresión de meganucleasas o para determinar el fondo, 50 ng de pCP vacío utilizando Lipofectamina 2000 de acuerdo con las instrucciones del fabricante (Invitrogen Corp., Carlsbad, CA). Para determinar la eficacia de la transfección, las células CHO se transfectaron con 200 ng de pCP GFP. Las células se lavaron en PBS 24 h después de la transfección, se tripsinizaron y se resuspendieron en PBS suplementado con un 3 % de suero bovino fetal. Las células se ensayaron en cuanto a la actividad de GFP utilizando un citómetro de flujo Cell Lab Quanta SC MPL y el software Cell Lab Quanta adjunto (Beckman Coulter, Brea, CA).

Los resultados se muestran en la Figura 7B. Se descubrió que ambas meganucleasas modificadas serán capaces de escindir sus sitios de reconocimiento pretendidos por encima del fondo en el contexto de un ensayo indicador basado en plásmido.

3. Escisión específica del sitio del locus DHFR de CHO por las meganucleasas CHO-23/24 y CHO-51/52

Para determinar si las CHO-23/24 y CHO-51/52 son capaces o no de escindir sus sitios diana pretendidos en el locus

DHFR de CHO, los inventores exploraron el ADN genómico de células CHO que expresan o CHO-23/24 o CHO-51/52 para identificar las pruebas de la escisión cromosómica en el sitio diana pretendido. Ese ensayo se basa en el hecho de que las roturas del ADN cromosómico se reparan frecuentemente por NHEJ de una manera que introduce mutaciones en el sitio de la rotura de ADN. Estas mutaciones, normalmente pequeñas eliminaciones o inserciones (conocidas colectivamente como "indel") dejan una cicatriz delatora que se puede detectar por secuenciación (Gao y col. (2010), Plant J. 61(1): 176-87).

Las células CHO se transfectaron con ARNm que codifican CHO-23/24 o CHO-51/52. El ARNm se preparó produciendo primero una matriz de PCR para una reacción de transcripción *in vitro* (SEQ ID NO: 29 y SEQ ID NO: 21). Cada producto de PCR incluía un promotor T7 y 609 pb de la secuencia de vector corriente abajo del gen de meganucleasas. El producto de PCR se purificó en gel para asegurar una única matriz. El ARN protegido (m7G) se generó utilizando el kit RiboMAX (Promega Corp., Fitchburg, WI) de acuerdo con las instrucciones del fabricante y el análogo Ribo cap m7G (Promega Corp., Fitchburg, WI) se incluyó en la reacción y 0,5 µg del producto de PCR meganucleasas purificadas servía como la matriz del ADN. El ARN protegido se purificó utilizando el Sistema de Aislamiento de ARN total SV (Promega Corp., Fitchburg, WI) de acuerdo con las instrucciones del fabricante.

Se nucleofectaron $1,5 \times 10^6$ células CHO-K1 con 3×10^{12} copias de ARNm de CHO-23/24 O CHO-51/52 (2×10^6 copias/célula) utilizando el dispositivo Amaxa Nucleofector II (Lonza Group Ltd., Basel, Suiza) y el programa U-23 de acuerdo con las instrucciones del fabricante. A las 48 horas después de la transfección, se aisló el ADN de las células utilizando un kit FlexiGene (Qiagen, Hilden, Alemania) de acuerdo con las instrucciones del fabricante. El ADN genómico se sometió entonces a una PCR para amplificar el sitio diana correspondiente. En el caso de las células transfectadas con el ARNm que codificaba la CHO-23/24, los cebadores de la PCR directo e inverso eran de las SEQ ID NO: 16 y SEQ ID NO: 17. En el caso de las células transfectadas con el ARNm que codificaba la CHO-51/52, los cebadores de la PCR directo e inverso eran de las SEQ ID NO: 18 y SEQ ID NO: 19. Los productos de la PCR se purificaron en gel y se clonaron en pUC-19. Se secuenciaron 40 plásmidos que albergaban los productos de la PCR derivados de las células transfectadas con el ARNm de la CHO-23/24, de los cuales se descubrió que 13 tenían mutaciones en el sitio diana de CHO-23/24 (Figura 7C). Se secuenciaron 44 plásmidos que albergaban los productos de la PCR derivados de las células transfectadas con el ARNm de la CHO-51/52, de los que se descubrió que 10 de los cuales tenían mutaciones en el sitio diana de la CHO-51/52 (Figura 7D). Estos resultados indican que CHO-23/24 y CHO-51/52 son capaces de cortar sus sitios diana pretendidos corriente abajo del gen de DHFR de CHO.

4. Integración específica del sitio en el locus DHFR de CHO utilizando una meganucleasas modificada.

Para evaluar la eficacia de la inserción de ADN en el locus DHFR de CHO utilizando una meganucleasas modificada, los inventores prepararon un plásmido donante (SEQ ID NO: 11) que comprende un sitio de la enzima de restricción EcoRI flanqueado por una secuencia homóloga de ADN al sitio de reconocimiento de la CHO-51/52 (Figura 8A). específicamente, el plásmido donante de SEQ ID NO: 11 comprende un vector pUC-19 que alberga un casete de recombinación homóloga insertado entre los sitios de restricción KpnI e HindIII. El casete de recombinación homóloga comprende, en el orden 5' a 3': (i) 543 pares de bases de un ADN idéntico a la secuencia inmediatamente corriente arriba del medio sitio de la secuencia de reconocimiento de la CHO-51/52 y la "secuencia central" de cuatro pares de bases que separa los dos medio sitios que comprenden la secuencia de reconocimiento de CHO-51/52; (ii) un sitio de enzima de restricción EcoRI (5'-GAATTC-3'); y (iii) 461 pares de bases de ADN idéntico a la secuencia inmediatamente corriente abajo del sitio de corte de la CHO-51/52, incluyendo el medio sitio corriente abajo de la secuencia de reconocimiento y la "secuencia central" de cuatro pares de bases que separan los dos medio sitios que comprenden la secuencia de reconocimiento de CHO-51/52. Nótese que esto da como resultado una duplicación de la "secuencia central" de cuatro pares de bases (5'-TTGC-3') para maximizar la probabilidad de la invasión de cadena por las tres protuberancias 3' generados por la escisión con CHO-51/52. Los inventores han descubierto que los plásmidos donantes que comprenden dicha duplicación de la secuencia central son sustratos óptimos para el direccionamiento genético por recombinación homóloga.

Se preparó el ARNm que codificaba la CHO-51/52 como se ha descrito anteriormente. Se nucleofectaron $1,5 \times 10^6$ células CHO-K1 con 3×10^{12} copias de ARNm de CHO-51/52 (2×10^6 copias/célula) y 1,5 µg de plásmido donante (SEQ ID NO: 11). La nucleofección se llevó a cabo utilizando el dispositivo Amaxa Nucleofector II (Lonza Group Ltd., Basel, Suiza) y el programa U-23 de acuerdo con las instrucciones del fabricante. A las 48 horas después de la transfección se aisló el ADN de las células utilizando un kit FlexiGene (Qiagen, Hilden, Alemania) de acuerdo con las instrucciones del fabricante. El ADN genómico se sometió entonces a una PCR utilizando los cebadores que flanquean el sitio de reconocimiento de la CHO-51/52 (SEQ ID NO: 18 y SEQ ID NO: 19). De manera importante, estos cebadores están más allá de los límites de la secuencia homóloga que lleva el plásmido donante, y, por lo tanto, se amplificará solo la secuencia de ADN cromosómica y no el plásmido donante. Los productos de la PCR se clonaron en un plásmido pUC-19 y se purificaron 48 clones y se digirieron con EcoRI (Figura 8B). 10 plásmidos daban lugar a un patrón de restricción consistente con la inserción de un sitio EcoRI en la secuencia de reconocimiento de CHO-51/52. Estos datos demuestran que es posible utilizar la CHO-51/52 para insertar un ADN corriente abajo del gen DHFR de CHO en la SEQ ID NO: 8.

5. Integración específica del sitio de una secuencia diana modificada en el locus DHFR de CHO

Se produjo un plásmido donante (SEQ ID NO: 25) que comprende una secuencia FRT (SEQ ID NO: 5) adyacente a

un gen de resistencia de zeocina bajo el control de un promotor temprano de SV40 (Figura 9A). Este casete estaba flanqueado por una secuencia de ADN homóloga al locus DHFR de CHO inmediatamente corriente arriba o corriente abajo de la secuencia de reconocimiento de la CHO-23/24. Las células CHO se co-transfectaron con este plásmido donante y el ARNm que codifica la CHO-23/24 como se ha descrito anteriormente. A las 72 horas después de la co-transfección, las células resistentes a la zeocina se clonaron por dilución limitante y se expandieron durante aproximadamente 3 semanas. Las poblaciones clónicas se exploraron entonces por PCR utilizando un primer cebador en el promotor SV40 (SEQ ID NO: 26) y un segundo cebador en el locus DHFR (SEQ ID NO: 16) para identificar las líneas celulares que llevaban la secuencia FRT/zeocina corriente abajo del gen DHFR. Una de dichas células que llevaba la inserción de la secuencia diana FRT se co-transfectó posteriormente con un segundo plásmido donante (SEQ ID NO: 27) y un plásmido que codifica Flp recombinasa. La SEQ ID NO: 27 comprende un gen GFP bajo el control de un promotor CMV, una secuencia FRT, y un gen de resistencia a higromicina no funcional que carece del codón de inicio ATG. La recombinación estimulada por Flp entre los sitios FRT en el genoma y el plásmido resultaba en la incorporación de la secuencia completa del plásmido en el genoma de CHO en el sitio de la secuencia diana modificada. Dicha recombinación restauraba la función del gen de resistencia a higromicina orientándolo corriente abajo de un codón de inicio ATG integrado como parte de la secuencia diana modificada. Como tal, las integraciones satisfactorias se podrían seleccionar utilizando higromicina.

Las células resistentes a la higromicina se clonaron por dilución limitante y se ensayaron 24 líneas clónicas individuales por PCR utilizando un primer cebador en el gen de resistencia a la higromicina (SEQ ID NO: 28). Los 24 clones daban lugar al producto de PCR esperado (Figura 9B), indicando que el casete de expresión del gen GFP se había insertado satisfactoriamente en la secuencia diana DHFR modificada en todos los casos. Entonces se evaluaron las 24 líneas celulares por citometría de flujo y se descubrió que expresaban niveles consistentes de GFP (Figura 9C).

6. Amplificación de transgenes

Se sembró una línea de CHO que expresaba GFP como se ha descrito anteriormente, con una densidad de 3×10^5 células/ml en 30 ml de medio que contenía 50 nM de MTX. Las células se cultivaron durante 14 días antes de sembrarse con la misma densidad en un medio que contenía 100 nM de MTX. Las células se cultivaron durante otros 14 días antes de sembrarse en un medio que contenía 250 nM de MTX. Después de los 14 días de cultivo, se evaluó la expresión de GFP en las células tratadas por citometría de flujo y se comparó con la expresión de GFP en la población celular parental (pre-MTX) (Figura 10A). Se descubrió que las células tratadas con MTX tenían una subpoblación distinta en las que la expresión de GFP estaba significativamente aumentada. Las células individuales con alta expresión de la población tratada con MTX se aislaron entonces utilizando un clasificador celular y se expandieron 5 clones durante 14 días en ausencia de MTX. La expresión de GFP en las 5 poblaciones celulares clónicas se evaluaron entonces por citometría de flujo y se compararon con la población celular parental (pre-MTX). Se descubrió que los clones tratados con MTX tenían aproximadamente 4-6 veces la intensidad de GFP que las células pre-MTX. La PCR cuantitativa se llevó a cabo entonces utilizando un conjunto de cebadores específicos del gen GFP y se descubrió que los clones tratados con MTX tenían todos aproximadamente 5-9 veces más copias del gen GFP que la población pre-MTX. Estos datos proporcionan una prueba concluyente de que un transgén insertado corriente abajo del gen DHFR de CHO se puede amplificar con el tratamiento de MTX.

7. Estabilidad de la amplificación genética

Las cinco líneas celulares clónicas que expresaban altos niveles de GFP que se produjeron en (6) anteriormente se pasaron entonces durante un periodo de 14 semanas con o sin 250 nM de MTX para evaluar la estabilidad de la amplificación genética. Se determinó en una base semanal y el ensayo de PCR cuantitativa usado para determinar el número de copias del gen GRP descrito anteriormente se repitió al final del periodo de evaluación de 14 semanas. Como se esperaba, los clones que se pasaron en el medio con MTX mantenían un alto nivel de expresión de GFP sin que ningún clon se desviara más del 20 % de la intensidad de GFP determinada la semana 1. La PCR cuantitativa reveló que el número de copias al igual se desviaba menos del 20 % para todos los clones. Sorprendentemente, la amplificación era igualmente estable en las líneas celulares cultivadas en medio que carecía de MTX. Al contrario de lo que se había predicho basándose en la técnica existente, la expresión del gen GFP no se reducía más del 18 % en cualquiera de las cinco líneas celulares durante el periodo de evaluación de 14 semanas. El número de copias del gen determinado por PCR cuantitativa también era estable con menos de un 24 % de desviación a lo largo del tiempo para todas las líneas celulares. Estos resultados indican que un transgén amplificado en el locus DHFR de CHO es estable durante un periodo de tiempo extendido, obviando la necesidad de cultivar las células en agentes de selección tóxicos que podría contaminar las formulaciones de bioproducto.

Ejemplo 2

Inserción de una secuencia diana modificada en las regiones codificantes del gen DHFR o GS de CHO

Como se representa mediante diagramas en la Figura 4, un procedimiento alternativo para direccionar una secuencia de interés a un locus que se puede amplificar implica la producción de una línea celular en la que una parte de un gen de selección se sustituye por una secuencia diana modificada. La ventaja de esta estrategia es que la posterior inserción de una secuencia de interés se puede acoplar con la reconstitución del gen de selección de manera que las líneas celulares que albergan la secuencia direccionada apropiadamente de interés se pueden seleccionar utilizando

las condiciones de medio adecuadas. Una línea celular que alberga dicha secuencia diana modificada puede producirse utilizando una recombinación homóloga inducida por nucleasa. En este caso, se prefiere una endonucleasa específica del sitio que corta una secuencia de reconocimiento cerca o en la secuencia del gen de selección.

1. Meganucleasas modificadas que cortan los genes DHFR o GS

5 Se produjo una meganucleasa llamada "CHO-13/14" (SEQ ID NO: 12) que corta una secuencia de reconocimiento en el gen DHFR de CHO (SEQ ID NO: 13). La secuencia de reconocimiento está en un intrón entre el Exón 2 y el Exón 3 del DHFR de CHO. Se produjo una meganucleasa llamada "CGS-5/6" (SEQ ID NO: 14) que corta una secuencia de reconocimiento en el gen GS de CHO (SEQ ID NO: 15). Cada meganucleasas comprende una señal de localización de nucleasa en el extremo N derivada de SV40, una primera subunidad de meganucleasa, una secuencia de engarce, 10 y una segunda subunidad de meganucleasas.

2. Escisión específica del sitio del ADN plasmídico por las meganucleasas CHO-13/14 y CGS-5/6

Se evaluaron las CHO-13/14 y CGS-5/6 utilizando un ensayo de recombinación de repetición directa como se describe en el Ejemplo 1 (Figura 7A). Se descubrió que ambas meganucleasas escindían eficazmente sus secuencias de reconocimiento pretendidas en el contexto de un ensayo indicador basado en plásmido (Figura 7B).

3. Escisión específica del sitio del gen GS de CHO por CGS-5/6

Las células CHO se transfectaron con ARNm que codifica CGS-5/6. El ARNm se preparó produciendo primero una matriz de PCR para una reacción de transcripción *in vitro* (SEQ ID NO: 22). Cada producto de PCR incluía un promotor T7 y 609 pb de la secuencia de vector corriente abajo del gen de meganucleasas. El producto de PCR se purificó en gel para asegurar una única matriz. El ARN protegido (m7G) se generó utilizando el kit RiboMAX (Promega Corp., Fitchburg, WI) de acuerdo con las instrucciones del fabricante y el análogo Ribo cap m7G (Promega Corp., Fitchburg, WI) se incluyó en la reacción y 0,5 µg del producto de PCR meganucleasas purificad servía como la matriz del ADN. El ARN protegido se purificó utilizando el Sistema de Aislamiento de ARN total SV (Promega Corp., Fitchburg, WI) de acuerdo con las instrucciones del fabricante.

25 Se nucleofectaron $1,5 \times 10^6$ células CHO-K1 con 3×10^{12} copias de ARNm de CGS-5/6 utilizando el dispositivo Amaxa Nucleofector II (Lonza Group Ltd., Basel, Suiza) y el programa U-23 de acuerdo con las instrucciones del fabricante. A las 48 horas después de la transfección, se aisló el ADN de las células utilizando un kit FlexiGene (Qiagen, Hilden, Alemania) de acuerdo con las instrucciones del fabricante. El ADN genómico se sometió entonces a una PCR para amplificar el sitio diana CGS-5/6 utilizando los cebadores de SEQ ID NO: 23 y SEQ ID NO: 24. Los productos de la PCR se purificaron en gel y se clonaron en pUC-19 y se digirieron 94 plásmidos que albergaban los productos de la PCR con la enzima de restricción BssSI, que reconocía y corta la secuencia 5'-CTCGTG-3' que se encuentra en la secuencia de reconocimiento de CGS-5/6. Se descubrió que 17 plásmidos eran resistentes a BssSI, sugiriendo que el sitio de reconocimiento de la CGS-5/6 estaba mutado. Estos 17 plásmidos se secuenciaron para confirmar la existencia de indels o mutaciones puntuales en la secuencia de reconocimiento de la CGS-5/6 (Figura 7C). Estos resultados indican que CGS-5/6 es capaz de cortar su sitio diana pretendido del gen de GS de CHO Debido a que la secuencia de reconocimiento de CGS-5/6 está en un exón de la secuencia codificante de GS, se espera que muchas de las mutaciones introducidas por CGS-5/6 cambien la fase del gen GS. Por lo tanto, CGS-5/6 es útil para inactivar la GS de CHO para producir líneas celulares GS (-/-). Dichas líneas celulares son útiles porque son susceptibles de selección por GS y de amplificación para la producción de líneas celulares.

Ejemplo 3

Meganucleasas para el direccionamiento de inserción genética en el locus GS de CHO

1. Meganucleasas modificadas que cortean corriente abajo del gen GS de CHO

Se produjo una meganucleasa modificada llamada "CHOX-45/46" (SEQ ID NO: 29) que reconoce una secuencia de ADN (SEQ ID NO: 30) de aproximadamente 7700 pares de bases corriente abajo de la secuencia codificante de GS de CHO. Las células CHO se transfectaron con ARNm que codifica la CHOX-45/46 como se describe en el Ejemplo 2. A las 72 horas después de la transfección, se extrajo el ADN genómico del agrupamiento de células transfectadas y se amplificó por PCR la región corriente abajo del gen GS de CHO utilizando un par de cebadores (SEQ ID NO: 31 y SEQ ID NO: 32) que flanqueaban la secuencia de reconocimiento de CHOX-45/46. Se clonaron entonces los productos de la PCR y se secuenciaron 24 productos clonados. Se descubrió que 14 de los 24 clones de los productos de PCR (un 58,3 %) tenían grandes mutaciones de la secuencia consistentes con la escisión del genoma inducida por meganucleasas seguido por la reparación mutagénica con unión de extremos no homólogos. A partir de estos datos, los inventores concluyeron que la meganucleasa CHOX-45/46 es capaz de escindir específicamente un sitio de ADN corriente abajo de la secuencia codificante del gen GS de CHO y era capaz igualmente de dirigir la inserción de transgenes en este locus que se puede amplificar del genoma.

LISTADO DE SECUENCIAS

55 <110> PRECISION BIOSCIENCES, INC.

<120> PROCEDIMIENTOS Y PRODUCTOS PARA LA PRODUCCIÓN DE LÍNEAS CELULARES DE MAMÍFERO MODIFICADAS CON TRANSGENES AMPLIFICADOS

- 5 <130> 111-077T1
- <140>
- <141> 01-06-2012
- <150> EP 12792301.9
- <151> 01-06-2012
- 10 <150> PCT/US2012/040599
- <151> 01-06-2012
- <150> 61/492.174
- <151> 01-06-2011
- <160> 64
- <170> PatentIn versión 3.5
- 15 <210> 1
- <211> 163
- <212> PRT
- <213> *Chlamydomonas reinhardtii*
- 20 <400> 1

```

Met Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe
1          5          10          15

Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser
          20          25          30

Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln Lys
          35          40          45

Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly Val
          50          55          60

Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu
65          70          75          80

Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys
          85          90          95

Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu Gln Leu
          100          105          110

Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys Thr Trp
          115          120          125
    
```

ES 2 773 891 T3

Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr
 130 135 140

Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys Lys Lys
 145 150 155 160

Ser Ser Pro

5 <210> 2
 <211> 50001
 <212> ADN
 <213> *Cricetulus griseus*

10 <220>
 <221> modified_base
 <222> (36646)..(36646)
 <223> a, c, t, g, desconocido u otro

<220>
 <221> modified_base
 <222> (38354)..(38354)
 <223> a, c, t, g, desconocido u otro

15 <400> 2

taaaactcaa gatgccagct ttgtagctag cttaggaaac aaagtagtaa aaaataataa 60
 tgggtgggtg aaggctctgaa gcatttacag agttctctca agacaaagca cagaggctgg 120
 tggccacata acttggcaac tgatttgggg gaacagaata caagaaagga aatttaataa 180
 ctgtttttct caatgttgaa ctatatgggc atagtcacag ctgcctaacc tatagagact 240
 ggaagctgga acctcggcta tctaagatag aataatcaag aaatgtcaat tatttgagaa 300
 aaacatcagg aataaatagc tgctaagtta caagttgggtg ctttagacat ttggagagga 360
 taggatgggg gctcccagac ctggggctcc ctaataaagc tgtgctggcc tacaagttcc 420
 agggatcctc cagtccatgc ctcccactgt tgggactgcg ggcgatggtt tctgacgtgg 480
 gtactgaggg cotgaactgt ccacacactt aagccacacg ccttttactg agtcatctcc 540
 tcatctcaga acattttcct ttaatctttc ttaatgaaa ggtcgcattt cttccgaggg 600
 ctagcctoct gttactctct atacatgtca cataaaacta catgaaaact ttgaaggcac 660
 tatatgtcca tactcagatg aaaagccatt agctgtgggtc atacaaaacc ccacagacca 720
 actgttggga aacatcagac ttttttcctg cagcgcctgc cctgatcttc cacagagaat 780
 tcagtctcac tttttccagg atgacttctg aactatcacc gtaagatgag aatttgaaac 840
 aaagatgtaa gtaatgaact tcatgtgttc tgaacacaca gcttagtgca ttgaaattac 900
 gtaacaccog cttccttata agccatttct caaaatgttc ccattacacc tgcacggggg 960
 atgggtccca gaatcttctt tttaaataaa cccccagag gattctgaag ctagaacacc 1020

ES 2 773 891 T3

aaggactgac agagagaagc atgcctgtgg gcgactccag acacctggga gctgcctgct 1080
ttcttgctac tgatttagaa ggcatttgcc cccgaatggg gctgggggac tgtcactatt 1140
tctcattctc gggactttga aaggaagcaa aacagaaaaac catgcaaagt ataagccacc 1200
atggaataat ggcagacgat ccggttgtgc agattagatt ttacatattg ctgattttga 1260
agctaaagac ctttcaactc ttaaataat aataaaatc atacaagagt attttgtgta 1320
ggtaactcag tcagatacaa ggtaagcaaa gtaaatgata ggtgcccctt aacaaaatgc 1380
attctcatag ttcatttatc aattatagaa atggtggact ggaggggaagg cttgaggtca 1440
ggagaatgtg ctgctcttcc agacagcccg ggttcttttc cccagcaatc tgggactcac 1500
gtctgcctgt agctccaggc ccaggggatc tggcaccttc ttctggcctc tgcaggcacc 1560
catacacaca tggcatacac acacatacac aaattctaaa attaaatagt aggttgtagg 1620
cctacacaaa aacatgcata cattaactaa ataattaata gttaataaat aaaaatcaac 1680
caaacacata cactgattaa gtaacatgac tctgtaaggt caaaggcggc tgaccagctg 1740
tggaaggggt taaataataa caatcacctt tgaaagaotg gacctggtga ttaaggatgt 1800
tccagctgtg tcgtggatga gaaatcaaat gcataattga atgagtgcc a ggaatagaac 1860
tgagacttt ctggtgagaa tgcttttact ggcagtagag tcctgttcta aacaggagag 1920
agacctgcag tagccctgtg gcggccctgc agtggccctg tgatggctct gcagttgtac 1980
tcttctgag ataggagaca cactagagag tgtttctaata gacagctcc tgtactttct 2040
gttccctgg agaccgcacg tgtttctccg ataatacatt gacatttctg ttaaaccatt 2100
ttcttcttg aacaaaaatg gagaacaaat cagattgggtg tgtggtcttt taaataactt 2160
ggtacttaat aacacaaaac aaaattatca gaggctggat tttaggtgct ctgagcatct 2220
gccaccctg agccatcagt caggtcttgg aggaacaatc tccaaggaga aaacagttct 2280
gtcctcagaa aagctggagg aatatgagat tttctacagc actcatagca aatcattta 2340
cggaagggat cctgagtaag atggcctctt cttcatcaca tgggtcatagt ctgcttcaat 2400
ggggagaata gttcaatcta gcatcgagaa atcgaagggt cccttttgac tggcaatgcc 2460
ccatagatag atagatatag attatgtata tattgtgtaa aacacacgta tgtatatata 2520
atacacatac atgtatgtgt atacatacat acatacatac atacatacat acatacatac 2580
atacatagat acgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt 2640
ttgagactga gtttctctac tatgtagctc tggctgtcct gaaagttgct aagtagacca 2700
gactggccag accagatcca cctcctctg cctcctaagt gctgagatta aaggcctgca 2760
cccaccccca cccagcccat cttatatttt gcttcatttc aaagtaagct ctatgcatca 2820
tttattcctg catattatta gccatggctc agtcttgttt gtgttttgga atatttactt 2880
aacaaaactt gaaaaacatt tttcaagatt tgtttgtttt taagatttat ttatttatta 2940

ES 2 773 891 T3

tgtataataa	taaataattat	tatgaaaaac	ggtgttctgc	ctgcagggca	gaagagggca	3000
ccagattgaa	ttacagatgg	ttgtgagcca	ccatgtgggt	gctgggactt	gaactcagga	3060
cctctggaag	agcagccagt	acttttaact	gctgagccat	ctccccaggc	ccaaaataca	3120
catcttaagt	gtattgccac	aagcatacat	cttcatggcc	caatcttctg	tccatcactt	3180
cagacagctc	tccttctttc	cctggccagt	cacaacaccc	tcagctatca	ggaaaggccc	3240
tatggggggt	gttttgtttt	cccactccag	ttoccttgcc	tgctctgacc	tcatgagtag	3300
actcatacag	gatgtgctca	cttcacttgg	gatgatttct	ttttcaccca	ttgttgcctc	3360
gcccagaatt	tgttcctttt	tattgtotta	gtgttaatca	actatcaaag	ccagcaacaa	3420
aaaatagtag	ggaaactttt	ttgatagggt	aaacctgatt	gattgcaggc	tttggttgcc	3480
ttgtttggtc	tatccccttg	agagtccctt	acaatgtgag	ttagttagtg	gctgctaact	3540
agttgaatct	caacttcctt	tttctttaat	gtgggtatth	gtaaggaata	gcccccttaa	3600
atctagattc	tgttctcaaa	tcaagcaagc	tcaaggctgt	aagcatggat	tcaccaactt	3660
tcctgctcaa	ggaatttaaa	tgtctggtct	ccatcatatt	actttaatag	taatagttta	3720
ttatacacat	gtgccagctg	tatatccctt	ttcttcttga	tggacctatg	aactctgttg	3780
aggtgagatt	tgaaccctt	agaaggtgct	agagaagagg	tacctgatgg	tcaaggcaag	3840
gctgataact	attcatgggt	cccacatctg	ctaagtgaag	caataacaga	taatatgctt	3900
tgtgtttaga	cccacagtgg	ttgcatgtac	actaagtatg	tatcatcatt	gtcttatcgt	3960
tcctttagaa	tacagctaata	aattatgacc	gctattctca	tagcatttat	attatatgag	4020
cattgtaaat	tattttgaaa	tgctttaaga	tatacttgag	aactatgcat	atcatgcgta	4080
tgttgttcta	ccagctggga	ccttgaaatg	agatcccttg	aggccagcat	aaagagaaaag	4140
ttttcatctc	aaacaaacaa	aagatacaact	tgataataga	tgagggataa	atgtcataact	4200
ttttatatag	tgattgagaa	tctacagatt	tgggtatcct	ggtcacttag	gagaccaagg	4260
gaggactatt	agctctagag	ctatgaactt	tatctccaga	ttccaaagcc	aatacaaaact	4320
ctagccaagt	tggggtgctg	ttacctgtat	ccctctgtca	aattccaagt	gttttcacca	4380
cctttactgt	atctttccaa	ctgttctctt	ttataaccac	acatagtcca	tggtctttcc	4440
ttctctcact	tgactgtgga	gtaacctaac	ttgctgtgtt	ccagttttcg	atctcttctc	4500
taaatctaca	ctagttaacc	acaaagaccc	tcttttctga	gctgtgtcta	ttctatcact	4560
gtcaccattc	cttaatgctc	tcccagatgc	agccaaactt	cactttgggc	ttgagagtct	4620
tctccaggty	acagtgacta	atgtctccag	attgagcacc	taccatctac	cctgtgtatt	4680
acacatgaat	agccttagct	tttcagcaat	agacagatag	atccatagtt	agccatgtca	4740
acacccttct	tcatgctggt	ctcacagtaa	taagtcctaa	ttcctgtttt	ctcccatcta	4800

ES 2 773 891 T3

aactcaaccc tgcctaaat accttactca aatcctaatt gtatctcttc cacaaacatt 4860
tcccccttct ctccattaca aggtggaaac tcagagatcc aggtgtcttg catgttggtg 4920
attctgtcct caacaaggaa ttccccaggt tcctgcacga aggaaagcat ggaggaccat 4980
acttgaggct actggtgtag tgggaagaca ggcccaaac atgtcacaga aaccatcac 5040
cagaaagttg ggggaggcag cccagttgtg gagcaggaga aggagaaaac aggcttgggg 5100
aactgctagc tatgctttgt cacagtcaca agaaaaaagg gccctagcct ggctacata 5160
ttctacaact tcctgaatct ttgctctgaa atgaagaggt ttggatggct gtctgggaat 5220
tcatcttgct tgcagtgaag ctcttgggg tatttgaaac caggaagttt gaaggagttg 5280
atgctaattg ttttctaaag tgtgtgagga gtactggcag agttcaggcc ttgtgaggaa 5340
agaatcctat atctagtctg cactcctggg cacatgagac attcagctat ctcccttata 5400
aagcatagaa agtactcttg tacttgacac agaaataatt tcagtatgta gagcattaaa 5460
aaaaagtatg aatgacttag agagatggct catcagttaa aagcacatac tgctcttcca 5520
gaggtcctga gttcaattcc caacaaccac aaaaactcac acatatgcat gtgattaaaa 5580
ataaaatctc tctctctctc tctctctgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgag 5640
tgtgtgtgtg tgtgtgtgag tgtgtgagtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 5700
tgtgtgtgtg tgtgtgtgtg tgtgtgatgg tgggcttgtg tttgcaagcc cagcactagg 5760
gagttaaggc ctactcaca gtgccaggcc agtctaggtt acagtgagtt ctagacagcc 5820
caagctacag agtaaggtag tgacaaagaa agaaagaaag aaaaaaagaa agaaagaaag 5880
aaagaaagaa agaaagaaag aaagaaagaa agaaagaaag aaagaaagga gagagggtgag 5940
agggagggaa ggaactggaa gggggaagga gggaaagaaa agaaaaagaa acaaccaaaag 6000
gaacaaacca ctgtatgcca ttatacatta gctttgggct ttacagggtta tacactctat 6060
attgtcatag ccaatgtctc aatattccat aagaggtgtc tagttgtggg tatgttcttt 6120
cttagtcctt ttatttagac tacatgacct gtttttgccct aataggccat tagtaatact 6180
gacttctcca catgctgccc tcaaaactta ctctggaag atctttattt aagctatgaa 6240
cgaaaatctt aaccctgtga cctgccaccc agaatgcctc tgggaacaac ctgaggcaac 6300
ctatcaagcc gcttttccaa catttggggc aacagggatt aaaattatga ttgttgtctg 6360
cctgctgagt tcaaaactcac agagggacca gaagctgact cactgatatc aagcagttct 6420
aaatcttcag tttaaaactc taattattaa acaggggatg tcctcagacc agcactcaag 6480
agaaggagat aggcagagct ctatgagttg agttatagcc cagcctgggtt ttcatagtga 6540
gttttagctc tccagagagt taccagcaag accctgtcac aaacaataa aaacaacaa 6600
acaattaggg gatatacata taactaaatg ataaagcctt acctagcaca ttcaagtccc 6660
caggttcaat tgctagccct ggggtgggat ttggacaaat ttaaaaagac cttttttgta 6720

ES 2 773 891 T3

tcacacataa atatgactgc actggttggt gttttccatg gaaacagaat caatgtggca 6780
 tgtattttac ggcattagct catatagttg tgcaggctgg caagtgtgga atgtataggg 6840
 caggccagga atcagaaatt gatacaaaat tcaggaaaga cctctgggtg caatggtgca 6900
 cacctttaat tcaagcactt gaaaggcaga ggcagggtgat ctttgtgagt tccaggccag 6960
 cctggtctac atagtgaatt cggggacagc cagggttca tagaaagaac ctgtctcaaa 7020
 acacacaaac aatcagaggg aagggttat tttgtttttg agacagggtc ttctatgtag 7080
 cccaggctgg cctcaaaactc atgctcttga tatgcccacc tcacaagtgc atgttaagat 7140
 tacaggtgcc tgacacacac cacttttgtg aagtgtgaa gagtaagccc agggcttcat 7200
 ggacgctggg caagcactgt gccagctgag ccacactccc cagtgtgcac gatactttgc 7260
 aaagatagat ccatatggat gctgtgcttc tatctaaaca gaatgacaac cacactctgg 7320
 caggttctgg ttcataactg agtcttattg gtcacctcct tctccatttt tcgctgggat 7380
 ttctcaagga gagaccacaa atgagaagtg aagcctaact tttaatgctg tctctcctat 7440
 gtcacctaaa ttctagctca aacagggttt ctggctctta ccttttctc gggtttctgg 7500
 atacttgaag tgtaaacggg catttctctt aaagaccaa tctggccaga ttcaaatggc 7560
 tggccttcaa ctcggaacac taggaacaat aatgtccgct gcatgtggct tgtagcactc 7620
 tgtttctatt catggacttg tgagtgattt ctgggaaaca cgaattataa gataagtctt 7680
 tttcagtgga cttcacaagt tcaccctcag gtagtatact gtcaggtaga aacgtctttc 7740
 agagaagcga gaggtgacaa gccctctggg ctggccattg tccctgctgg cattgaacag 7800
 cctgttcagc acatgaaagc atcgctgat gctcccaaag ctggagcact ggcagcccc 7860
 tgcagtcag tgtgtagggg gggttagcag ggggtgcttag gcggtttttg tagttacctt 7920
 ttcaacacaa atgcaaaagc cagagagaga gagagagaga gagagagaga gagagagaga 7980
 gagagggaga gagagagaga gagagagaga gagagagaga gagagagaga gcaggaaagc 8040
 atccaggctt tgaagcaagc cagccttcag ctctgtcctt gagccattct gagtggaatg 8100
 gagtaattgt ctgcttggag aactgaagaa tagcacatgg caaagaacaa tttgtacctg 8160
 gaatatattc attagcttgc atgtcaaaag gccacatgca gatagaaacc attatcttgg 8220
 cattctttaa aaccttgag ccttgagact tgaggtgcag aaaccacat gcccatgtga 8280
 ctgactacct gtgatctct ccagccctgc ctggctaaca gggacaatat aggggatgg 8340
 tgggagggga cagcttagac tcctgtggac ttggattgaa agaagaacag ggaagacagg 8400
 ggactgtgca aataagcact ctattaggac ctatttttgg tgtcttggga cctcctact 8460
 ggttagctt aaattgagag gggatttggg ttgcctcact agctgtttct tcccactcaa 8520
 ttcacaatta cagctttctt cattgtcatt aaaatacatt aaatgtgtac ttgttggggg 8580

ES 2 773 891 T3

aaggctttct gttgaaatct gcataaagac aatgtccaca gccccagtc agtggaaaga 8640
gcagtaggac cagaaggcat gtgtttccat cccgagtcta tattggaatg tttgttaaaa 8700
cctgcacttg taagagacaa aactagaac catcagcttg caggtctaca ggccagtgtt 8760
gccagtgcag ataatgccc aactggaacc taaagatgaa ggcctttggg agctgaggtg 8820
gaagagtcag ctgtgatctc ccagatgtcc tcctcatgcc ccattgccac tctagcctcc 8880
cacctccaag cacatttggg atccaactgc taaccctggg tgttcttttc ttagttgaaa 8940
ttctcagga ataacctaag agtctctgtc actcagtcta tggcatccta tgataacagc 9000
caaggctaaa tagccatcat tgttcttttt ccagatgtcc agcaatgagg atgcagaggt 9060
gaacaaaggt ggttcagggc tgccctgatg atgaatttga caagccagaa tctaacaaga 9120
tcagtcggta aacagaatcc tccttcctat ccagagatgt tggcttgttc tgtcactgga 9180
tgggcatcat ttactataag tcatacaggc accagacact cagagataaa taacatgaag 9240
ttccagctct tatgcagtc tgtctagtgt acttgccagt attctcaagg aagttccacc 9300
ccagcccctg gcatccatag accaaggact ctggaatgtt ctgggaaagc tccacctgag 9360
ctcctagcac ccatatatcc aaagagtctg gaaogttatg gtggaagccc cacctctctc 9420
tcccagacc tcgcccctc aaaaagtcca ccaaagactc cccaccccc acacaccccc 9480
agatgctcaa gaccacttcc atagagtatt taaactgcct ccagaaaaac agaattcatt 9540
ttttcagctc ctcttcccc tgtcctctca ggggtggggg caggggtatt agtattcaag 9600
cacctatact ggctgtctc tggggttctg acaagatatg acctcagcta cagccactaa 9660
gatcaccacc tgtgtatata cactatgtc ccttttaaaa gggccctgtc cacctcccat 9720
tctctctgtc tctctctctg tctctgtctc tgtgtgtgtg tgtctctgtc tctctctctc 9780
tttctctctc tctctgtctc tctctctctc tccttctctg cctgactctc cctccctccc 9840
ctgctctctt ctttctctgt gcttttctcc cttagaggcta gtctctctc tccccttccc 9900
ccttttcccc ttcaacttcc cccaataaaa aactctccac ccaagctcta tcacatggca 9960
tcattctctt gctccatgat tttaaaatca caatgaggag gggagcatgg aaaaattatc 10020
caggaagact ttatocatta aacctgggtg ctttttcttt cttccttctc tcctttcttt 10080
ccttctttct ttcttctttt ctttttctct ttcttctttt ctttttctct ttttctttt 10140
ctttttgttt tgttttgttt tgagacagcg tttctctgta gctttggaga ctgccctgaa 10200
actcaatctg tagagcaggc tggccttgag ctacacagaga tccacctgcc tctgcctccc 10260
atgtgcttga attaaaggtg tgcaccacca ctgctggct taaaactggg ctttttctaa 10320
gtcagtttga tttggattgc tgcattggca gagaggttta ttgggtgca gaaaccttct 10380
aaccagcttt tgagctaatg atagagagaa gctcaaggaa ttggagcaat gcttgactag 10440
ggatgtcaga gggaggctat ccagaggagc ttacaactga ggtaaaacta aaagttaggg 10500

ES 2 773 891 T3

agtttgtcaa cttcaaccca cagaatagag cagagccagg aggagctgag gcttctgagt 10560
 gttatgggtgg aagcatcacc ccaacccttg acatccatat gcctgaagag tctggaatgt 10620
 tatggtggaa gttccaccca agcctccctt cccggctgcc ctccaaacct tgctacatct 10680
 cagaaatccc accaaatgat gactccctcc cccagagata ttcaagacca ctcccacagg 10740
 gtatttaaac tgccccccaa cccccagaaa atagatgtgt ggttttccaa tctctctttc 10800
 ctatcacgtc tctggggagc tggcaggcca tttgggagca ttgtatccat taaacgactt 10860
 ctcagtggag actctgaaag ccagaagagc ctagacagat agatgtcttg catactctag 10920
 agactacaga tgccggccca gactattata tccagcaaaa gtttcaaca ccatacaaag 10980
 tcaaatttaa acagtatcta tctacaaatc caatattaca gaaggtgcta gtaggaaaac 11040
 tccaaactaa gattaactat acctgtgaag acacaggaaa taatctcaca ctggcaaaag 11100
 aagaaaaacc tctctctctc tctctctctc ctctctctct ctctctctct ctctctctct 11160
 ctctctctct ctctctcaca cacacacaca cacacacaca cacacacca caccaatacc 11220
 atgaacaaca aaataacagg aattaacaat aattgatgtg tgtgtatgtc cctgtgtgtg 11280
 tgtccttgtg tgtgtctgtt tgtgtgtctg tgtatatgtt tgtcacctga ggggtggctc 11340
 ttccttgggt tgtgaggttt ctaccaatc tataactccc ttttcttcat tcacttcctc 11400
 atgtccttac tagtctctat tgtggattaa ggaaactgtg tggagaacag ttttcttcta 11460
 gaaaagaaca ctagccatct catgtaatca aattggtgac tatcctaatt attatgagag 11520
 agcttccgtc cagtaagtgc tagaagtaga tgcagagatc cacagacaag cactgagcca 11580
 agctccagga gtccctgttga aaagagagag gaaggattgt aggagccaaa gagtcaagag 11640
 catgacaggg aaaccacag agacagctga cctgggcttg tgggtgggag ctcatggact 11700
 cttgaccaac aattagggaa cctgcatgag gccaacctag gaactctgca tgtgtgtgac 11760
 agttgtatag catggtctgt ttgtgaggct tctagcagtg ggatcagggc ctgtccttgg 11820
 cgcttgagct ggcttttggg aacctgttcc gcatgctgga ttaccacacc cagccttgat 11880
 gctgggggaa gcacttggtc ctgcctcaac ttgatgcgcc ttgcattggt ggattctcat 11940
 gggaggactg cccctttctg aaaaagaaca aggagaagtg aataggggag gggattggga 12000
 ggagaggaag gagaggaaac tgtgataggg atgtaaaata aattaaaaaa ttaattaatt 12060
 aaaaaagaac acttgtactg gtagattggc taaaatgaaa caagataaa agtacacagg 12120
 aaaaagagag gagaaacctg gggagggggg ctccaaagag aggtgagggg gggatgggaa 12180
 tggcagctta gtggaggaag gaagacatga cctacacgaa tcgagctgta gtttttatct 12240
 ggagcatagg gtaaagatgt ttgaggagaa ggaggaacac atgcttgtaa aacatggtct 12300
 tcagaaccag caacaatcat acagagtgtc cagggctccat gggcacatga aggacagacc 12360

ES 2 773 891 T3

aacacatatt taacagtaaa gtgtccatat ttggtatgaa agtgatgggt aaattgtcct 12420
gggactgtaa tttagttgta aaggacttgt ctggcatgtg ggtattcttg ggttccctcc 12480
ttagcactga aaaaaaaaaa aacacacac acacacacac atatattcta gtgttttgta 12540
gaaaaggatt caaagaaagc catgatttct cttttgataa atccagaata atgtaataag 12600
aacacacagt ggtgtgattt cagcaatcaa gtacaggttg cttgtctggt tgttgtatgg 12660
gatggttggg tggttgttt cttggtttgt aagatgggtg ggtgggttgg tgggtgggtg 12720
cttggttggg tagttgggtg ggtgattggg tgggtgggta tttggttggg tgggtgggtg 12780
gttggttggt cgtttgggtg ggtgggttgg gttttgtttt gagacagga tttactctat 12840
atctcagttt gtctcaaaact cactatgtgc acatgagtat gtgatgagat tatctaagac 12900
catagtgtct gtgttcatgg aatgtctctc tagcttagag aatttaaaaa atggccatgt 12960
agggaaacc ctcagaaaag gagtttctat ggccccaag aataagaatg gatcctccta 13020
gctcggagtc agcaaggaac tgaagccctt aattttatag acacaaagga atccattgtg 13080
tggctccttc ccagccaagt ctcagatgag tcacagacct gcatggcacc ttatgcagtc 13140
ttttgaggtc ccaagaatag gatgcagata agccatgcca gaatcccaac acacaaagcc 13200
ttagtgatat agtaaatatg tattgtgtct aggctgctgc atttctgggt atgctactgt 13260
gcagtaatac acaactaata cagatgtgat ggtaaatatt atgtgacaac ttgagtgggg 13320
cacagaggta cagacacttg gtaaaccatt ctgggtgcac gtaaggatag ttttgatga 13380
cataaacatt tagattagta tgctgggtaa aatacattgt ccatcccaat gggcatgggc 13440
tttgtccaac tagatgacag ctggaataga aaagtctgcc tctctcatag ttctcaggcc 13500
tttgagctca gactagacag aactcacagg ttctctgagc tttccagctt gatgaatgtc 13560
catggcagtc ttcacactta acacctgaca gacttaatga tcatatgaac caattcaaat 13620
ctgaccatca ctgggtcat tcttttgatt ctgtcacttt ggagaactaa taccgaggac 13680
ataaaatgcc atcacatcgt tattttcttc ctgtctgtga atatttttct ttttttctt 13740
ggtttttttt tttttttttt tttttttttt tttgttttct tctgtgtagc tttggagcct 13800
atcctggcac ttgctctgga gaccaggctg accttgaact ctcagagatc cgcctgcctc 13860
tgcctcccga gtgctgggat taaaggcgtg taccaccaac gctcggcctg tctgtgaata 13920
tttaaaatga aaactttgga aatgttctga aaccagctgg tgcagatag tcagagaact 13980
ttcgtaaggt aggtgtgggt tatagcataa tcccacacaa gaggtgaag caggaggatt 14040
ttgtgtttga gggcagctag agccacatgg tgagtccctg cctcaaaaca caaaagcaag 14100
acaaaaacaa gctccaaata agattcactg ggcccttctt ttccttctt ctcagtgagt 14160
ccacttgctt taaaatcagg tcttaaagac gcactagatg ctgaacttaa cagtaataat 14220
aaatatcttc tcttacagta cagattatgc tctataaaca ctgcactgat aaagttcagc 14280

ES 2 773 891 T3

cttaaccttt gttctgtaaa tgtttcctag tttttctact gccgtattat aagacaaatg 14340
 tcagcatgaa ggcaggtttt tcagaaaaca cagcagctcc acagatggcc tctaateccat 14400
 aatcattaaa gacaagactg caactttttc aactggaaat cattcaagat gtttttctga 14460
 agtccctacc aggacacaag ccaccctggt tgctgtgtga catcagttag gtagactctg 14520
 aactggcttc ccaagaaatt atacaaaagc aagggtgtcac ctagtattag cataacttct 14580
 gataactact gtcttagctg gggtttctat tgctgtgaag agacaccatg accacagaaa 14640
 ctcttataaa ggaagcaat tattgggtcc agcttacagt tcagaggttt aatccattgt 14700
 catgattgca ggaagtatgg tggcccacag gcagacatgg tgctggagaa gtagatgaga 14760
 gttctatatac agattgacac acttcttcca acaaggccac acctccactc actctgagcc 14820
 tatggggcca ttttcattca aaccacaaa gctacaaggt agcttatacc ccagcttgct 14880
 atttctgatg agacttagta aatagtctta aaagccata aatgactca aaactagttt 14940
 ttttattatt attattagtt caaattagga agaagcttgc tttacatgtc aatcccttct 15000
 ccctctccct catcaaaact agttttttgt ttttttaggt ttttttcaag acagggtttc 15060
 tctgtgtagc tttggagcct atcctggcac togtctgga gaccaggctg gcctcgaact 15120
 cacagagatc tgcctgcctt tgcctcccga gtgctgggat taaaggcatg caccaccaac 15180
 acctggccaa aattagtttt aagtccagtt ctaggagctc caatgcctc ttttggttc 15240
 catgggaacc aggaacacta tatatatata tatatatata tatatatata tatatatata 15300
 tatatatattca ggcaaatatt tatgcatata aaaataaaat aaatcttttt tccttttttt 15360
 tttaaagaag tgacattgtc ttggaatttt tgtggctgct ctgcccttat gtgtaactgg 15420
 aactaccag catctaaaca ctggcctgaa accagccaaa gaaaacctt gtgccaggtc 15480
 ctgtgtcaaa gtattatggt ccttttagga tctctatat cctaaaggat ttattttact 15540
 gatagcatct taacttctt tgaaaggttg gtcttctcaa gcagtcctcg tggagctggc 15600
 tcctcagcta atgccagggg acaataatga tcccctcca aaaccaaca gaaaacctg 15660
 gcaactctgg tttccttggg cagcacctgc ttttaagaatg agcaaatgac caatcagctc 15720
 atgaaactaa atactctatt attactaaaa ttttttttg agacagggca tggaaatcat 15780
 cacatagttc aggttggcct tgaactcaga gagactcact tacctttgcc tcccacgtgc 15840
 tggaaataaa ggcattgaacc accacaccaa acataaact tgaattttgg aagagtcctt 15900
 cttccaatag atttgaggtt ttgaaaatgt ggcacagaaa atatgaattc aatataatg 15960
 aaaacaagag ataactttca actaagtttc tataggttct tgctaggaat cctaagcttg 16020
 tctgaaactc tagagcttct gtttctagct tctgagtgtt agtattgtag gtatgtgccc 16080
 tgcctcagtg tgatgttttt gataatctta aagaaatcaa agaaatttta taaaagacta 16140

ES 2 773 891 T3

gactgtgcta cacaaaaaga atattcagat gccaaagaaag agttcttaga aattaagaaa 16200
tatgctacta gtataaatcc tttataaagt ggaatgacaa atctgatgaa atcttactaa 16260
aagtagaaaa acataaacat caaagacatg aataataaga aaatcatatt gtgcatatga 16320
ttaacctaaa acattaactt gcaaaaatag aatagtccca aaaagtaaac aaaataaata 16380
aatcaccaag aacatgatac aaggacaatt cctaggatga taaaacaaga atattcatta 16440
taaaaggccc tatcactaaa gcacaacaga aacagactca aaagataaat cttcattgtc 16500
actggagaga agtccatact atcatagcac tcagaaggaa ataaaaatca aaatgtcaaa 16560
aaggacctca gcctctgaaa cacaaatata aaatatgtcc cgccttcttg acacgcatta 16620
ctcttcaatt aacattttaa gaaaactata aactgttaaa gagagcttag tattttaaga 16680
aatctgtagc tatttctttt ataagcatga caactaagtt tccttgattt aaacagacct 16740
aaaaaacggg tgaagtgagt ggagaaaggg gatacgaaga cagcatccca catgactgct 16800
cccagtaaag gcaaggtctt catccatttt atcctgaact ctgggaaatt tataaagaac 16860
agaaatgtat ttctctcagt tctggagcct cagtccagga cactaagtct aggtactaca 16920
ctctcacatg gtggaaagta gaaagcaagc tcacttgtca ctactacct gatgcctctt 16980
tcatcaatcc cattgataag gaagagacct ggcatctcag tttcctaagg actcagctct 17040
tactaacatt agctgtcatt tctgggtcac tgtaacagaa agcctgacag aagcaaccca 17100
ggggaagaag gatgtatfff ggctcactgt ctctgaggat ttcaacttat cccagcaata 17160
aagggataaa ggcattgcag caggaatatg tgtggcagaa gctgtttatg tcacaataaa 17220
caataaaca cacgctagcg cgcgcgcaca cacacacaca cacacacaca cacacacaca 17280
cacagagaga gagagagaga gagagagaga gagagagaga gagagggggg ggggcagaca 17340
gacagacaga gagggagaga ggcagagagg gagagagaga gagagagaga gagagagaga 17400
gagagagaga gagagagaga gagagaaatc aaaggccac ctccatcaga ctggtcccat 17460
atcccaaat tctagaacct cctaaaacaa caccatcaac tgaggggagac atttttggat 17520
tgaaagcata atgccattac ccaggcagaa tctgcctgtc tgggggagtc acatttaagc 17580
catggtatca attgacctca tgtaatttca gaatactaca taaaactatc agatattttt 17640
catgatgaat ttctaaagct tgaaattccc tttgaataaa ggaccaacta cagaattttg 17700
ctgagtctac aattacatac atgaaaatgt aactacgaag tggccagcca caatgaaaat 17760
taaagtgttt ggggtggtctg tctctattga tgctcttctt tgccctgttt ttttttaata 17820
ttgttgatgg tttgtttttc ttttaagata cttggcccca agaaaaaaaa tgacagcctt 17880
aattaatfff gtttactctc ctgacatttt aaaagacaaa tttatgaaga cctgactggt 17940
ccatgtagta ttagaaagat gtaaaattaa ggggttgctta agctgtgtag aattgaagag 18000
cacagcattt gagtgcaggg gtacaattag agatcatcag ggatgtggca caaagtgtac 18060

ES 2 773 891 T3

tcaacctcac cttttcctgc ttagcagaga acaggggtgcc tccggtgagat aggaaattaa 18120
tcaaatagaa gaagaaatag taattttaga aggatcaaat tttcctgggtt agaatgatca 18180
aaactacaag acttgtaact aaaatatagt caaacccatt tcaactggaa tctgtgctat 18240
tcatgtatag attaactaga atctaatttt taaattttca tcttacttcc aaaaatattt 18300
gtccaaatac tctgtgaatg cattagtttc ttatgggaaa acatcatatc ttttgtacaa 18360
tgtgtttott agcttgaggt tctctccaaa caggaccaag acgaggccag gaccatgtga 18420
tacaacccat agtcctcaag aatagttgt cattttctta ttccaattgc atccaaggt 18480
ctcatctcat tttgcgtgtg cctttgacac cccataccca cataaactaa ggtggtgtta 18540
ttttttgagg ccctgaaggt atcttcagga atccataagt gagccttaag ctgcatctgg 18600
atataggaat ctgaaaggt cccttctctg catgatctct tctttcagtt tttcaagtca 18660
gtgtgccaca ggaatcagga acgataaatg gagaggggaa gtgcagttgc ttggtataga 18720
caccocagag ggctatttgc atcctgtcct tcaaaatctc tctgagcctt cctgcctaag 18780
ctgttttgag ttgggtttgt ggtaccagaa ccctgcccc cgccccattc tgactaatga 18840
gagagagaga gagagagaga gagagagaga gagagagaga gcagcagagc atagaatgaa 18900
agtaggttag aagggcaggt aaaagcactt tagacaagag caggtataag ggccttgac 18960
tcctcccca gaacacacac atgaaggtaa acgatggtta aaggatacag ataggatgtc 19020
gaagctggac gatcacttgc ttttgtgtgc ttgaagtgac aggctgtggc tttcgggttc 19080
atggggctctg ttgttgagtt cacagctca ccatgtagc aagcatgtca ctattaagct 19140
ctatccccgc ccccttttt tgagacatgg tcttgctaac ataccagac cggcctagga 19200
agcactttgc agtctcagct cccctgagtg ctatgatcac tccgtgtgagc tacagtacc 19260
aaaccagaat atgtgtgttg ggtgttatga gagtttacac attgctgcct tgaatgctgc 19320
tctgcttgag ttcctgtagg aagctgagct gggaaacctaa gcttcctcct cccagatagc 19380
agtaaccctg cagagacctc ccaccaagac tagctaacct ctccttcttg tgetgtactt 19440
agcaagaacc ccaaggttct gggccttgt gctacagttc cagaagagta tgaacaatct 19500
tagctttttot gtatatgtgt ctgtgtctgt cctgtcagat caagtcccag cctcactgta 19560
tgcaacatga aaggctgtga aaactgtgca ttttgagaat gaacatcatt agtctccagt 19620
aagttcaaaa acaaatgaag gcagccactc ataagggctot ttaatgaggc aagggggcaa 19680
aaggggtggt tctgtttgtt caaagaagcc tgtcatacat tttcagaaaa tttagaaca 19740
cgtatcatgt catttcacgt tagtatgaag tccttataat tcatttcata ttaaatgatt 19800
tcctttggtt agaagcaaaa ttatgcataa aatgtgttcc tttgtgtttg gagcaaaatt 19860
acaagttaca ttattagtta atattctagt tcttattttt cccaatctcc aagaagcaaa 19920

ES 2 773 891 T3

atattcccct aaaccctaaa gcatcaaatt atcctatcac acagtgacca gtcacgtaa 19980
cctaaatatt aaagcatcag attatcctgt ctatggtgac cagtcattgt aacctaata 20040
ttattgtaat gtggattaga gttaactata ccttttcato aactataat gtaaactc 20100
tccaaatcct tcaaagtctt gaaaacacaa tttataaata ctgtgtctg tttgtttga 20160
gacctgatcc ggttaggaat ttcaggctgt cctcaaactc atcatcttc tgcctcactc 20220
aggcctaag tgctgagatt aaaggctctat gctaccacag ccatacgaat gccatgtctc 20280
catcagctta tcacttctta acttttttct tttcttcttc tacatactgc tgagtaggag 20340
catgatgac ctgagcctag taggaatggg tcccatgtga acccttaatc tgttaggaaga 20400
tgctggactt cttcattaa gactgatctc cattedgaact tgacttctct ctctctgtg 20460
tgagctacc atcccatata taatcttctg gttataaac agattgcttt accctcaaga 20520
tcctttgcta gcgcagcaat gtaagtttta atacaaacag taaggctctc gattggagtg 20580
tcatggtttg gttaagtgcc ctttccaagg gcccatatag ttaaggctc aaccaccaag 20640
tgatgcttgt ggataggag cagggcctag tggacagtct ttaggtcatg gagctatgct 20700
gttgaggggg actgtggggg cctggctctt tcccactcc ttttaggtc ctagctatga 20760
ggtgagtgg tttgtcctat caagcacctc tgtcctgcca tgggtgaatt gattataact 20820
acaacctctg aaactaagcc agtataacct atttatctca agatgtaact tacaggtaat 20880
ggtaagataa agctaacaaa agacaaattg ttataatcca ggcaagcctg gccccatccc 20940
ttgggggcat ggcacagagt gtgtcaccca tctgtgcatg gcaagcagta ccctgactct 21000
gtatgctgat tcaaaggctc cttaaagcaa actcctocca ctctctctct tttctgcca 21060
ttctctgag gagggaggcc actgtctctc tgtctctctc tgtgtctctt tttctatctt 21120
cctctccctc tcttccctt ccccaataaa ctttccacat taagtttctg ctgaaggat 21180
ctgtttgtct ctccccgcc ttttaggcc cacctaccat gggatctgcc aaaggctctca 21240
cctcgagctg tattcataac acaaatgaca gacaaagatc aaccctgaag actagtagga 21300
tgtagaaggc ctggagctga cctgaagaac actgctgact tcaacattgc ccatccgtca 21360
gttatgtagc attaaagtta tagtggttcc tcagaaagca gtctcctttg aaaacttctc 21420
gttttgtgtc taaatggaat taaatacctt gttccogaat aattgtttta gttctcttga 21480
aagatcccgt atacttacta ttaagatgta tataaacctc aagctgaaag aatgacttcc 21540
cctatggcca gatcacaaga ctctccactg atgtgcccgt tgcaacctga ttagaggaag 21600
agggtcaaag ttccccaaaga ttoagctgag ttcattgcaag ttttagaaaa aaaacaagat 21660
gttctccac agttagaaag gagtggggct ggagggatga ctactgaga aaggttattg 21720
tcgtacaagc atgaagacct gagctcgaag cctggcacc atgtaaaaag aaaccatgca 21780
tggtagtgtg catcttcaat cccagcattg gggagacaga gaaagagaaa gggacatccc 21840

ES 2 773 891 T3

tagagcttcc tggtcagcca gccttggcaa gccagtgaac tccaggttca gtgagagacc 21900
tgtctgggga ggaaaaaggg agggagggag ggagagagag agagacacac acacacacac 21960
acacacacag agagagagag agagagagag agagagagag agagagagag agagattgag 22020
gaagatacct gatatcaacc tcacacactc atgtacccat gtatgtaggt accttcacac 22080
acacacacac acacacacac acacacacac acacacacac acacacacac acacacacac 22140
acggatggtg ttgaattcta aggetottat ccacacatat atggagacaa atagaagaat 22200
tacagtcgtc cctgcctttg acgetactct gtttctccaa ccttgcctcc cagatatttt 22260
tcaacatcta ctacagcctg agtggttgca ctctgacccc aggacctctt tctgtgactt 22320
ccttggcctc ctgttttgtt tttctgatgc taaaaactga atctggggcc tcatgcacac 22380
aggaagatgc tataccaatg agctacaatt ttgttgccct ttttaatttt tgagatggtc 22440
tcactaaatt gttcaggatg gcccaactgt aattctcctg ccttagcttc ccaagtagct 22500
gggcttttat acagatctgt gcttccacac ctggctgagc agacactcat gatttcattt 22560
ctgctaataca ggtagttttc ttgccctcgt ctgccatttc ctacctgctt ttccttgcca 22620
actaaactgg tccccacaag cgacaggcta tcatttctca gctcttccac aggttagctg 22680
tgcaatttgg tatgaatcat ttagcaagcc cagttctcct ctttgtaaaa cagatgattt 22740
agatgaaatt tttcacaagt tctctttgaa ttaaaactat cactgccttg cttgctctct 22800
gactcttggg gaccatggcc tatccctgat tagtccttgg tccacagaag gatgggtggc 22860
attggatgtg ctgaacaatc aggtactttc atgtcacttg gagtcttaca gtaactgcat 22920
gtttcaaatg aatcctttct ggctctatta gtttcttttt tgtcactgtg aaaaaaacac 22980
ctgaaagaaa caaggcacgg tttgttctga ctctcggttc agaggatata gttcaccatg 23040
gaggcaggag cttctcacag ctgtaacagc catggagtca ggtggctagt tacagtcagc 23100
tggccttagc agtcagagag ccaagagagc tcagttgagg agagtccagc caggctgtag 23160
cccttaggac ctgctcccca gagatccact ttctacagta tcttctaaac agtgtcacta 23220
gatggtgacc aggtagtcaa gcacatgagc ctgagggata atatcattca aaccatagga 23280
ttagtctaga actgaaccag atcaagaacc aggttttctt ctacacataat agataccaca 23340
catcatgttc tcatatagag tgtgatctag gtattgtttc tccaaatgga gaagccaaca 23400
ctggatgact tacatagaaa gaaagagagg gaggaacaa gcaagggagg ggaagagtg 23460
agaattattg gaacagtacc agtgcctcaa aatccttggg ggactagaga attagcctca 23520
ggaagaagcg actaggcttc ttacagcata gacatacagt tcttaccaga ggcacagcca 23580
tcatgggtgc catggggagc atgaagttca gctccatcca gccattccta gcgatttctg 23640
gcaacctctg tcctttgaga cacttctgga agatataaga gtccaggag agacatctga 23700

ES 2 773 891 T3

ttgctttgat cccaggatct tgggatggaa ttggtgttgt ctctgctcca gctccagggt 23760
 caggaagggtg aaactggaaa cacaagctag cttttcttac ttagcaaaaa cccacagggtg 23820
 acataaaaga cagattgaca cgagaacagc atggcagatt tatttagtca aagttttacc 23880
 agacacaagc accttcagaa aggtaaagtc agagacctta ggggaatttt cttgccagaa 23940
 tttttccaga agaatcaaca gccgtgtaac aataggacta gataaacaag taagactgga 24000
 cctgcagcac aatgtgaca ataggagttg gaatccccag gactcacata aagccatggg 24060
 agccgaatgt aatggctact tgtagtttca gcctcagatg ggggtgggga ttctccagaa 24120
 taagcaggct agcaagacta gccatgttgc caagctctgg gttatattga gacactctgc 24180
 ctcaatgagt aagtggaaga atgatggagg ccaacttcaa ccttggactt ccacatgaac 24240
 acacatacac aatgcaacca tgcattccaca gtgtatgtac acacacacac acacacacac 24300
 acacacacac acacacacac acgcaaagtg acaaagaaag aggtaaaacc tacaaggaat 24360
 caactgaaca gaagccaact ggtctgcctg ttcagatcct ttttggcctc tctgtgtgct 24420
 tccctttctc ctgggcatgg ggcaggcagg atctgtatgg ggtgagggtc ttcagagaag 24480
 cgaacagcct tcctaggttt tatggctcag tttgggtggag aggggatcta gtttctctta 24540
 atcatctttt taaaaattta ttaatttatt ttttatattc caatcccagt tttccctccc 24600
 tcctctcttc cctccccca cctcccatct gttccttaga gagggtaaga cctcctctag 24660
 gaagtctact aagtctgccc catcatctca ttgagggcagg accaaggcac ctctccaccc 24720
 ctacactctg gtgtctaggc agaacaaggt atctctccat atagaatggg ctccactaag 24780
 tcagtttggtg cattagtggt agatcttggg cccacttcca gtggcctcat atattgtccc 24840
 agtcacatcg ttgtcaccta tattaaggga gtctagttcg gtcttatgca ggttccccat 24900
 ttgtcagact ggagtcagtg atctctcact agctctggtc agctgattct gtggtttccc 24960
 catcatgate ttgactcctt tgttcatatt gtcactcttg cctcacttca attgtactcc 25020
 aggagcttgc ccattgggta gttgtggatt tctgcatctg cttccatcta tttctggaag 25080
 agggttctat cttctctggg gttgtgaatt gtagactggg tatcttttgc tttatgtctg 25140
 gtatatgctt atgagtgagt acatacaaca tttgtccttc tgggtctggg ttaccccact 25200
 caggatgttt tttttctagt tctgtccatt tgctgcaaa ttttagaatg tcattgtttc 25260
 ttactgctga gtagtactgc attgtgtaaa tgtaccacat tttctttatc cattcttcag 25320
 ttgaggggca tctaggttgt ttccaagttc tggttattac aaataatggt cctatgaata 25380
 tagttgagca aatgtccttg tggatgaat gtgcctcctt tgggtatatg cacaaaagtg 25440
 atatttcagg gtcttgaggt aggttgattc ctaattttct gagaaatoga cataactaatt 25500
 tccatggagg ctgtacaagt ttgcactccc accagcaatg gaggagtgtt ctctttactc 25560
 cacatcctct ccaccataag ctgtcatcag tgtttttgat cttagccttt ctgatcagct 25620

ES 2 773 891 T3

taaaatggta tctcaggggt gttttgtaa tcatcttgag aaaaaggaat tctattttct 25680
 gtgactggct ctgagagaga gagaagaggg aaaggtggga ggaatgtgtg ctttcaagac 25740
 cttgtgttct cccttagctc aaagtactca ccatgaaaaa ccaccagcct ttggaggagc 25800
 atgctcttgc agaggcaaga tcctggcttc ctcccatctt gaatttgcca aaatagcaaa 25860
 gatgtttggg tgctggacag ccaaaaatga cagctgctca cttcacagct tcctcacgta 25920
 tgattacaac tccactcacc atcaagcttt aattacatca tgagcaggct tatggctgag 25980
 ccgttatcct cgcaccctt cgtctcatca ctgattcaca caaatcacta ggtgctccgg 26040
 ttaatgaaaa catattcacc agtacagtga ctaattcacc aggccaacat ttacatggct 26100
 cctctgcatg acaaaaatga atgtttagaa tgaataatga gtcaccagag gtgggggaca 26160
 tcttctgagc acaggttgcc cttgtcttcc ctggactca atcccggctg aagagctgaa 26220
 caaagctgag gttattttcc ccatgacagt gcatttggtt ttagagatct gtaagcggct 26280
 tatcttgatt ggcagtttga ttggttctgg gatgactaa gagacgtgcc tcattggcat 26340
 ttccagaaaag aattaactga gggggaagct cctogccccg agaatgggta ggagcatctg 26400
 gtgggttaca gatgtaaagt ggtccaaggg agaagccgca tggcctgcct gccttcaacc 26460
 cttgctgctg agtgtgttta tcccatctat ccogttgttg cttctgttgc agttgcaacc 26520
 ctgcttctcc aggccccagc gtagactgaa cagtggctgc ccagaaatcc ccaattgaag 26580
 cagccgaatg gtggactgag cacctctcag tcttcagtct ctctagtttg taggcaacca 26640
 ttgttgacc caactcttag tagtaagcca atctactaaa tacagaaagg ccagtgagat 26700
 ggctcagtat aggtgcttac caccaagctt ggtgaccgca gttcaatccc caagactcat 26760
 aaggaaagaa ctaactaccg agagttgttc tctgagctcc acacatgctg aaacatgggc 26820
 ctccacatgt catgaacatg ttcacacaat acatatttat ctctatata tcatttctta 26880
 taattttag aaaatttcat tttatgtata tgagtgtttt atctgtttgt atgtctgtgt 26940
 accacatgca tgcctggtgc ctgaagaagt cataagaacg tatcagattc cctctaactg 27000
 gagctaaaag aagattgaga ggtacctacc atctgagtgcc taggaaccaa acctgtgtct 27060
 tctggaagat cagtaagcat gcttaaccac tgagccatca tgccacttat ttgtaacaca 27120
 tatccatcct attggttaca gtcctgactc atacagttag atagctgagg aacctagaat 27180
 tcttctgctt ttttattaca aaacaaagaa ttttatctga cttacagttc tggccttagt 27240
 cagggagctg cattgggaga tggcttctct actgtcagag tccagagggt gccgtaaagt 27300
 atcatatgac atgaggcaga aagtctaact tacttgagag ttaacttggg aatgtccaaa 27360
 gagacagggg gctaagtccc tcttattgaa gagaccttcc atagaagtta gcctgacaga 27420
 tggccttgcc tgaactgcat tgacagtctt acttgggaagg cctgttttgg ttccctaagaa 27480

ES 2 773 891 T3

attcaaggat ccaccagaga agtgtgcagc cagcaagctg gactccctat cccaagcccc 27540
 agctcctcct cagggacctc agcagtcctg tgtctagctt acctcagcga tggggggaaa 27600
 gatgctgttt tcctgctaag agcacactat tttatattat tgttgacaca ggttggactg 27660
 catgtaacag actctccaac aacacagtga agatacaagt gtgttttgct gcatttaaata 27720
 gtctcccat ctgtccctgc taagacacct actgtccttc acatgtcact gaaaactcca 27780
 cccottatga gaagtcttc ctgatgcat ctagacaagc taagagtgct ctgctctgca 27840
 ctgagcagct tctcaactct ggggttatca ttgctctgca tcacaattag cacacgtggt 27900
 agtggctgtg tttgtgtttt tccacacat gagtccagac agcatccctc tcaccagcac 27960
 gccataggca caagtgtca agagtagcag gacttgaaca tgtgtggttt atcatacaga 28020
 cagctgctgc tcagagacca gatcaaattc aaagcaaat agagagatga tggttcctgc 28080
 catgagcgtg ctgaacaagg acaaacatca ccatcataag gaactcagct gacagggagc 28140
 ggtcaccaaa cttttttttc tgtaaagtga caaaaatagt taagtatttt gccctagaca 28200
 tagtgggtgg tacacatgta atctcagcat ttgtcagagt gaggcagaga gttgaatgct 28260
 gggctacgta gatagtctca aaaaataaat aaataagtaa ataaataaat aaataaataa 28320
 aaggaagaaa taaaaaaaaag aatttgttac tcaactctgc acaatggtgc aaaagaaaca 28380
 ataagcatta tgtaacctag tgggtattgg ctgtttcact ttactaacag gcattgaaat 28440
 ttcaattttg caaaattttc atgttccata ttacccttat ttttattctc ccctataaat 28500
 ggtgactcac caatacgaac ctggataaga ttagggtatt tttattaggg aatatgcctt 28560
 acttacagag cacctaacca gccagcagga aacatagtaa agtagcgcac gccgatgaaa 28620
 caaggaaaaa gaagaactac catgtgtgac ccctaaccct taaaacctct cccacatcac 28680
 cctgaccatg cccattaggg gtggtcacct agccagcccc taggaggcat ggttacgggtg 28740
 tccccctaca ctccccaat catttaaaga tgcaaatgca tgcttgggtg tgggctaacc 28800
 ttggctcatg ggctaactct ggctcatggg ctaaccttag ctcatgggct aataatcaag 28860
 gtttactaat ctctgtcaga cagccatfff tttttgagc agaagaatcc ccatcttgg 28920
 atcattttatt tattcctttt gtatatttga tgcaatttat aaccacaaga acctactatg 28980
 tgactgcact gtgccagatg gcagagaaag ctaagccccg attcttgtgg catggactca 29040
 cacaactcca gtacaggact gttagtgaca atctccttaa ggcataagca tactgcagtg 29100
 gcagcctctg ggttaggaga caaggataca gtttatgaca cctggtatct ggaaggcatg 29160
 aaacatgtca aatgctggct acacctaaga atcagcaaca tctagtctgg ccatagccta 29220
 ggatgaatgt cacagggctc taggccagaa atgtatggcc gagctgtagc agggtcctct 29280
 ctagggccag aattaattcc agtgtgatgg acagccaaga ccacaggat aacaaatgag 29340
 cagtgccaat gacacgtgct tctccttatt attgctgcac agtgtttggt acacatagca 29400

ES 2 773 891 T3

ttttcgcaca gtaatataat gtgcttgggt catcttgctt catatcccat cactccctcc 29460
 atctccctag tgccctcccct gttacctttg cttctcagtt ttgtttctgc tttgatgtca 29520
 acagcacata caagatttta tgcaatacat cacttcctga atggctctat ttggaaatca 29580
 ctaaaaggta atttatggaa catttgggggt ctttttgatt ttctaattta ccaaaaaatc 29640
 cacctgggga aagacaatgg agttcaagga cttctaagag gggaatgtac catggtatgc 29700
 tccagccagg ggaaccagtg cttcccagga gctatggctt acaaagtggg ttatcacatg 29760
 aaagcaagac taaaataatc atctcaaata ttcattagat gtgggactcc taaccatctc 29820
 acaatgcctc cctcgggtcta cattaataaa gaaacctcca ttttgtgctt tgcgagaaaa 29880
 tgactgaaga ttatacattt ggccttgaag tggaagtatt tttgaaaatc atgaatagga 29940
 aaataataaa tctctcattt caacataaaa tataagggac aaggacatct actcatgctc 30000
 caaggacgga cactgaattt tccatcaggt agttgcagaa cgctgtgtcg ctcaatcaaa 30060
 aattcaggat gcattgctca gagtgcatta tattaanaaga tagcatcttg gaacacagga 30120
 tgctcaggaa atgggagggg cattaatctg catgcagtga tcatctcctg caaagcgggc 30180
 atgagagcct gatgggagac aagccatcca gatgccata cccaggggag ctgtactggg 30240
 ctgcagccct gggccattca gccatgcacc aggtactcc ctctcttcc agctttctcc 30300
 ttctgatggc cataggatta gaagataagg gactctagtg caggtcaact gctgaccagt 30360
 gtgaaaatgc acagactaca tgctgtaga tcagcacttc aaactactgt tcaccatcat 30420
 ctctggaata agcactacat ttacagggtt caaacctcaa tgaatataaa caaacaaaac 30480
 acacctccct tccttactg tctcccattt ctttggttcc catctccaca tagaatttat 30540
 aattaaaatt tctaagtatc tttccagaaa tacttcacac atgttataag caaatgtgct 30600
 tttaaagata ctattttaaa ttatgaaaat ggttatatta gttgagataa aagaatagaa 30660
 tgggaagtcc cagaatttaa ggcctcatat gaaaatataa agcgctttct cttttaagtc 30720
 taggtaggt gtactagatc agcgctcagc tccataccat gaagccatcc aggagtcaga 30780
 cctctctgac agccctgcc a ttgtcacaga gaagtttctg tcaccagtgc tcatgctgtc 30840
 agaggagcga aggagaaaag atgtgagacc tcccagtc aagtcatcta tggataaaaac 30900
 cttagttgca tggcacacca gtgttaggga gtcggggaaa cacagccata gccagcttc 30960
 ctctctgttc ttgctcttat taccaccaga aagaggttgc ttagacaacc caaaccaaga 31020
 cacagggtc tgtgggaggg aatcagtcctc aggtctctgg cacatgctat gtcaccggaa 31080
 agccccagcc ctactccgaa tccccacaag tacagcaaat atcagattat agcattttaa 31140
 ggggcaactct tgccaaagag aagcaccatt ggaatagcca tgcttgagaa ctggtcctac 31200
 ttactgcaga accatggata caggctccct tttgtagatg ggcttaataa atacttctat 31260

ES 2 773 891 T3

aagtgatact ctgctttgtg aaaatgacct cgtcaatatt caaagtaatc ctctggttta 31320
ggactactat gaacctgtgg ggttcattgt tcatgtggtt aaacagcaaa gagtagttag 31380
acagttgtcc tacgtcacag agggggacat atgctatgct tggttaaata gctgtcctgg 31440
tcagagggga ggcatgctat tctgcccttt ctgacagacc ctgattgcat agacatttca 31500
gtgagataaa ggaaggaagg gaagaaggag gaaagacaac attttttgc tctgttaagg 31560
tagagactat ctgtgatcca gttcagcaca gtgcctgtga gtagaagcta caggtcagggc 31620
aggagccaag gaaatgtatt gcttttctaa ttgaacaaag gacacacagc tgccatttat 31680
tttcttcatt ttgaccttc agccctgcac tgtggatag acatcaagaa actaagcagc 31740
cattttgtga aatgagatc taagttagta aatgtggctg aaaaagaagc cagctgcatc 31800
ctccctggat ttacgagggg gaaatgtagg catactaaat taaaacacta aaattgacct 31860
aaagctattt tgactgatat taaatatag attctgctcc tggacattcc agagttcata 31920
ggacagttgc ttctgttcag aggattcctc ttcggggttg cctctcctc cttaggcctg 31980
cttgtcctgc ccaaagctgc ccaagtgcac caggcccaa accaacttct ccatcctgac 32040
gcacagcaga ctaaatatgc aactttgtgt ctcttcatcc caggacaaaa ctttcaccca 32100
gcccctgaca totgagactc tactacaggt tatctattaa atcttttata aagaccaaga 32160
aacaagtgt tggcatcaa actttggtaa atcatagcct ttaataaag tcaaatggac 32220
caatgtactc taacaaaaaa atatgggtct ctcatctctg aatggcagat ttcaagccct 32280
aagaaccaca atgctcacct actgggcaac actgagttac agagaccag cccccacc 32340
cctcaccaag ccagagaaac actctatctg aacaatcctt ggtccatgga gcaagaatta 32400
gacatagaat ttgtatctca ttgtttttta ggaaaacccc aaaggctatt atgaagtcat 32460
ttttctggg caccttttct tcccatgac aacgagttgt gggcagtctc agcagaatac 32520
tgaagctgtg gcttggggag acagagcata tactggattg gagttcatgg gtgggtgcat 32580
ggaatcaatg cogggcatgg gattcaagac cttatgcatg tgggtagatg ctttgttact 32640
gggataaatc cccacctgg gatctgactt caagcacaat ctttgggaag cggcattggc 32700
tctctgctaa ttttctagc acttttatto cacttatttt ctgcttgttt gctttgggag 32760
ttttgttctg tataagacag tcttgctgtg tatcctaggc tgatcacaaa cctgtggcag 32820
tccttttgtc agcaggccaa aattcccact ttatctctga agacagaaag tagattgagg 32880
aatatatgat aaagacactc atcaaagcca ggcatctatc tttacttttc ttaaagcatg 32940
ttttgaaatg gcataaaacc atgtagacaa ggagtcttat gttgtacatg gtctactttt 33000
gtcacttaca atataggata ctttcaataa gcttggtagc ccttgccta ttctacttat 33060
tctgttctct ctccctcggg tcttggggag ccttcttacc aggtggggtg gcataaaggg 33120
aaaagtcaca aagctcttcc tattcctggt tcccctcta agtgtacctt gctggtggcc 33180

ES 2 773 891 T3

ttgctagcaa atgtagtata acatctgact tatctcctct cagatatggt tgttgactt 33240
 agataaattt aatctagaaa ctcaagctgt atgtctttgg ggaccagcat tacagagctc 33300
 ttcccttctt gtccttacct caccttggct actgtagtaa gttaatcctg atgattcctc 33360
 catgagtcct gaaactgatt agttccaaga gctggaggat gagaagggat atagcctggg 33420
 gcagggacac tttccaatga ccacaagacc ttgcacaagg tacacatgga atgtgttaga 33480
 ctgtctcctt tctgtcccta gcctcagttg ccccagtggt tatcaatggt tattaacatt 33540
 gccctagcaa aaatactaca gactaggaag cttgggtaca attgaaaaga gcttctcagg 33600
 gttctggata ccgggaagtg caaaggttca gcatctggac agggctgcta ttgtagtttc 33660
 aatggttctt gctgcaacac ccctttgaga gaatgaacac tgcttttcac atggtggaga 33720
 gtgcacagac accaacccea ctccctgaagg ccctttctcg agggctctaa tccatcatga 33780
 gggccatact ctcaggactc attacctccc caacatcccc tctctaaata gtaccacact 33840
 gcatttgcat ttcaatatat cactggagat atataaatct ccagaccaca gcataccata 33900
 aatcagataa ggcaggcctg ccttctatag cctttcactc agcaaagggtg tttctagccc 33960
 aaagcagtct ggactctcac tctgaaacct cttgggagtg gtggccagaa atgacttccc 34020
 atcatccctc tctcctgacc tggccagca ccaggtcacc aggaaatcct ccaagtttca 34080
 ttatccccac cccaattgt ctcttgtctc tagcaaacct cttccaatac ttccttctt 34140
 ggtgggtgta gcaagccaga tgatagcctg ccaaagaagt tcacagcctc atttctggag 34200
 cctatgaata tgttacattg tgtggtaaaa ggaactttgt aggtgtgatt aaattatgaa 34260
 tcttgaagtg ggcagattat ccaagtgagt ccagtgaaat tgcaaaggta catcaccaac 34320
 agtgaggcag gaaggccaga gggggagaag gaagcagaga ggcagagggg ggaagagaca 34380
 agccagggga ggggagtggg gggaaagaaa ggagagagag agagagagag agagagagag 34440
 agagagagag agagagagag aaatatcaca cacacacaca cacacacaca cacacacaca 34500
 cacacacaca cacacctgaa cctgattgtg gaggaagaaa ccactaacca aggcattcga 34560
 ggcagccttt gaaagtcaca agagacaggg aaaacagatt ctctccctcg gcccttcaga 34620
 atcaacacag ccccaact gctgatttta gtcattgtaa agccaagttg gacttctgac 34680
 tgccaaaact ttagacgagc aaataaatct gcactatfff aagataccaa tgtgatttgt 34740
 tcatgaaaac aatcaataag gaactaataa agtagaagtg aaaattggat cacttctgaa 34800
 gtttgtaat atccacagaa actggacaca tgctgacttt gtgagccata gctccacacc 34860
 caggtatgcc ccctacagaa atgtgtatat aggtgggcag gagatgtcac ctgctgtgtt 34920
 catagtgcga cctttagact ttcccaagcc tgagaatagc ccaaacacct accaggagca 34980
 aaataaattg agatatacag acgcagtggt atactacact tctaaaagaa tgagaaaacc 35040

ES 2 773 891 T3

acgctataca ctgtatatcg toggaacagt aacacagggg tgacaatcag gcaataggac 35100
 atattctcta tggctttaga aaacataaaa atagcataac agttctgtta gtggcaatgt 35160
 gttctgtttt gtgatctgta tgatgcttcg gtttgtgcaa aagctctgga cttacctttt 35220
 aaatgtatgg tggctctatac cttttaaatg tatgctagat atacatgagt aaaaatgatt 35280
 aaaagagatg gaggggagga gactcatgcc ttcataaaaag tttgttctgt cctttctggc 35340
 actgtccaag tgaatgtgtg taaacaaaga gtgaccacc ccaggtagtc caccttctta 35400
 gaacctactt ctgctacaac atgtcctgtg aatgtgcacc aaatgtttac taagggatca 35460
 tgccacaggg tttgttttaa ataaagtatg tctacctagg ggtatattga ttgtctttcc 35520
 ttttgagggg gggctctcaa actacaaact agtttgtttt gagacaagta tgtagcccag 35580
 gatggccttg aactcacacc ttctgtcctg cctctttccc agcactagga tggcaggtga 35640
 gactatcagc ctggcccag gaaactatct ttgattgaca ttatctggtc agaaaagatc 35700
 taccttttcc tccaccaggt cctccaaata catgaagagc tgaaacagtt ctgtctaccg 35760
 aatttcoctt tttcttgatg tttctgtgga atttaataca taaattttaa tttgcatttt 35820
 tagcttttct attaagcctt aattagagta taatgaagtt atgaatttat aaaaataaaa 35880
 acaaaaocgg tgcctcccaca atcactcagt cttgaagtga ggttctgact ttacctgaag 35940
 tgggggaaga gagtgaggaa agggacctgc ggaagctgaa tctcagacc acaagatgga 36000
 tctgagatcc atccaagcga acgtggacgc agaccgggag tagggacatc caggggtcat 36060
 cttcatctgt cctcgctgtg cttctgcccc tttgctcctc taccagtctc agctgtcaaa 36120
 gctcagtggc ctggagggga gatggggcgg ggcttaggat cgaaggcggg gcctcggaga 36180
 gcatcttctg gccccgggg cctggactgg cccgcgcgcc ccacctgcag cgcggcggag 36240
 cgcgggcgcg tcaactccag cgggaagcgc agcctcgcgt ctggcgaggt gcgcgcttcg 36300
 cggctcccgc tccagagctt cgtggcccgc ctgtgtctgc agagcagggg cgggggcccg 36360
 gcggcaccga ctgggactg agatccaagt agccactgaa tcgtagacag tcacctcagct 36420
 cggacagcgc gtcggggcgg gagcagatcg ggaaggtgaa ggaccactgc ggatccgaca 36480
 gcgcgtocca ggtcagtcct cccgctgcac ttggggaaac tttgggatgc ggtgacggct 36540
 gcgagatgag gacactgagg gtcgcgaggc cgcgtggccc ctgtgaacce cgcgaacctg 36600
 tacctgccgc gcacctgaca ccgcagctgc cagggcgggg accganacct tgctgcccg 36660
 gaccactgcg ggcaccaag ggctagcggg cttcaggggc ctctcgggag cctccggett 36720
 gcccgcgccc agccgcgcgc ctccggtcct cgcgggtccc cagctccttt tggcggctcg 36780
 cgcgggacc ccgcgggct gggattccg ccgtcttcgg gcctcgtggc gctggaggag 36840
 cggcccgggg gcccatggct gcaggggtggc ggccccggg cgggagcggc gcgtgctcgg 36900
 ccggtggagc gcgcgggtcg cggggttcgg ctggagcgcg tggccgcag tgccctgtggc 36960

ES 2 773 891 T3

cgctgggcag cggaggtgag agcgcgggct ggggacgcgg agcggattgc aacctctggc 37020
 tgcaggaacc agggtogetg ggtgagcagt cctgtccccg cgcttccgg gcgtgcacat 37080
 ccctggcacc cggcatccag accccatcag ctggaggcgg gctgcagagc ggcgcctgcc 37140
 cgggccgagg accagtgcct cctgctctga cacgccatct caccaacgag ggcggggtgc 37200
 tagattggcg ggctgcgcgg ggaccactgg ccagggcctt ctggcacaag ccctttctgt 37260
 ggacagctgc ctgctctggc ttggagtgga ggagacgaaa tgagtacccc gccccatca 37320
 gcgccccaac actgtcgccc cagtcacctt cctttgccct tctccgacag caccttggac 37380
 ttgctccctc ccgaattggg gaaaatctga ggaaaccagg cagggacctt ggagataccg 37440
 cagcctgcat actcaacagc ctggaaatcc agtcaccttg gtacctogct gcttcccaga 37500
 cactttggag gagcaggttt gccatttcta ccccacatcc gtaccccatc ccccgtccgt 37560
 ctctgctgag gaagggactc ttatgagaga agttgggatc taggtacccc ttaaggtagc 37620
 cccagagtct gtggttaacta ggctcatagg taactaaaag gcacctagc tctgtagctt 37680
 tgtgagggaa acaaacctta ccaactaatt ccttcccttt ctgaatattt cttagaagac 37740
 tggagaccaa cggaaaccga ctgttctggc cagtctttgc accctttgct tggctctgac 37800
 tctccttccct aggcagagaa acattttgct tatgacctct ggctggcctc cttccaatcg 37860
 ctgcctggcc ttggactgcc catcaggact gtgatttttt ttttttttta agacctgatt 37920
 aggaaaggct gcaagcctcc ggttctagaa ggctcaaact caggggtata ctcttctctg 37980
 ataccatgt gctccctaact tccactgtgg caacacctct gcccttcaact cccacaagaa 38040
 aattggttgt caaacctctt ggggaagatg atggaggcat ccctgtggga gcagatgcag 38100
 gatttgaag caaccaggaa acaaccagga gtgaggaatc ttttttaaag gctcacatga 38160
 ttctggaact aagaaaagat ggagatgcca ccagtgtatg aagcttggcc tctcctcggc 38220
 ccatcccacc caactcaggg aactggcata tgcaggacct gtattgggtg atgcatattt 38280
 ggaacctagt acttattgaa ttcctaagca gtaaacacat tccgaatttg aaattcctca 38340
 caatcatcta ctgnaatgta gatattaaac ccccaactta tgaatgatag ccccaaaatt 38400
 gttaacattg agagagccca ggttccctgc cacctcttcc acaacaggac aggaactagg 38460
 acaatgaata ggaccatttg agctttaggg tcatgtgcc actttacagc tccatagcca 38520
 gacaactggt ttataagaga gggcacaaaag gaaaatcact gtcctgtcca aatgaataga 38580
 aagctgggga tgggtggcagg acaaaggcaa caggaaaaat catctccaac aaggctttcc 38640
 aagcatatca gtcttatact actgccatgt tgggtaccac acaaatcagg tatctcaaac 38700
 tggacgctgc ctagggagggt ctgtcatcta aaaaggcagg gagatattga gataaaatca 38760
 acagaagcta gtatttaact ccaggctggc agataatagg aatgaccttg ggagggtgtg 38820

ES 2 773 891 T3

cttaccttcc cttctctctt gaacaaaatg tggactggac cagatgagca ccaaggctcc 38880
 accaactcta acagaccttg tgtggtgggc ttgacctgaa acagacttga gctaggttgc 38940
 tgtgctgagg atccattcca gactcattta caaactcgta gtcagtgaaa tgtgataaac 39000
 cgaacactgt agggatttct aaacaaggaa ttaaaaaact cgactccaaa tgggagagat 39060
 gcaggcaaca aatcgacagt gtttatgtgc ctctgaatag ctttgatttc cttcggtagg 39120
 agctgacagc tggctgacag aaagctcacc caggagagaga agagagaaaa atcaagtatg 39180
 agattaggaa taatgttttc aggtaacttt ctattcccat tcggagtggg tgtctggaag 39240
 ggcgagtgta gttatggctt gaattgctcc atttatccac agatattttc ttccaaggg 39300
 ctctgatttc taagatgctg ggctttgctt ctgtctccta gtttctggt agcagggtag 39360
 agagctgggg gtcccagcat tcagcctgca tattcttcct ctatctcac tatctgctgc 39420
 ctccattatt tgtggtcttt tggatctatt tggctcagaga gtcagtcttt ggtttcttgc 39480
 cctggaaact gcttggttgc acttggtgtg ggggcagcat ttggaagtcc aggtgctctg 39540
 cccacaaact ttcaaccat ctttgtttt tcatccctt ctcatgcca ctttgtgtgg 39600
 tgcttgggac ttctgggacc tatagttcaa gggctatata taccaatggc tcacatgaca 39660
 gcactgatca ctctgccagc tctctctctt ttgcaaaact tatttcagat ttttcatttg 39720
 acaatacctt tctccagtt gtctttattc ttggcagcat atgccttgta accttaaaa 39780
 aggaaggtaa ataatttgag aaaaaatgta ccaagtctc agtgatacat tcttactaaa 39840
 gactcccagt ttaacaagg agttgggctg gagccatggc tcaacagtta agagcactac 39900
 ctgctcttcc aaaggacaca aattccattc ccagaacca catggcccct tccaaacatt 39960
 gataactctc gttccagggc acctcatgcc ctttctggtc atctgagaga accagcataa 40020
 acatacatgc aggtgaacat tcatacacat aaaatgaaca ttaaaaaaga aatgaaatag 40080
 agaaagggtt tacataacta ttaataact aagactgcct aataatgtag ggaccataa 40140
 agaaaatcta gtaagttttt acaagattcc actcaatcag accaaacatt actgttactg 40200
 acagagtaaa aagtcacttc caatagtcca agaacaactt tgtttcattt ctcaggcact 40260
 gtctgttttg tggcatatgt gcatggtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 40320
 tgtgtacagg tgaatgctgc tcgtgtatga gcacatgcag gtgtgtgttt gcatggtgtg 40380
 tagacagagt ttctgacctg cctggctcca cagctgtttg gccacaaata aacatacaga 40440
 ggcttatatt aattagaaac tgtttggcct atggcttagg cttctcactg gctatctctg 40500
 tcttaattat taaccataa ctactaatct atgtatttct acgtggcggt atcttaccgg 40560
 agaatacttg gtgtcctatc ttctcagcaa ctacatggcg tcttctctct gcgtcttctc 40620
 cccagaattc tctctgctg gttgccccgc ctatacttcc tacctggcta ctggccaatc 40680
 agtgttttat tcatcagcca ataagagaaa catatgtgaa gaaggacatt tcctatcaa 40740

ES 2 773 891 T3

tgggtgtgtgt gtgtgtgttt gtgtgtgtgt gtgtgtgtgt gtgtgtatgt gtgtacatgg 40800
 gtatgtgagc acatgtgggt atatgggtgc atgtgcacct gtgtgtgtgc atgggtggcta 40860
 gagttgaggt tagatgtcct ccttggtgc tctccacctt tttttattg aagctctcac 40920
 tgaacttaga gctcactgat tcagctagtc tagctacccg gcctgctctg ggggtcccct 40980
 gccttcactt tccatgtggc taccatatct actttacatt tatgtgggta atggggatct 41040
 gaactatggg gtcctcatgc ttgcatggca agtgctttat ggactaagac atctttctag 41100
 cctttacctt tttttttttt gaaagagttt ttttttgcta actgggaact caacaccaga 41160
 tagctagtct actggtcact gaggcccaggt gatctactat ttctgcttct cttcccaagt 41220
 gctgggacta cagactgtac caccatatcc atatttcttt tagcatgagc tctggaagtc 41280
 aaactcaggt cctcacgctc acaaagtaag tgttttatct accaagccat cttcccatct 41340
 ctgttgtttt aaaaggcttt gaatatggga tgtgatgaag ggaggtgaaa ttctgagata 41400
 aatttcttga aaagaagaat gaatcaagta ggagaacctc ctctgtgtgc tgtctttcag 41460
 ttccatgtcc acacagcata aacattatga ttatcattcc acagattgta attagtcttt 41520
 ctctgttttg ccagtctgct cccaaaaaat gacacagaga gacttcttat taatgatgaa 41580
 agctttgcct tagcttaggc ttgtttctaa ctaactcttg taacttaaat taaccattt 41640
 ctattcatct acctgctgcc acgtgattca tgacttttac ctctctctca ttctgcatat 41700
 cctgcttctc ctgcttctgg ctcatgatcc cgcttttctt cctctccgag tgctctgtcc 41760
 ccagaagtcc cgcctaacct ctctctgctt agcaattgcc catttggtctc ttactaaac 41820
 caatcacagt gacacatctt cacgcagtgt aaaggagtat tctgcaaca caggtgatga 41880
 agccaacatt ccaagaggcc agggcttgc tagggcacat agctaactta agaaaattag 41940
 gatcgcatc tacatctgtc tgactctgaa ttggatctga actgtgactt gcatggaaga 42000
 cccaaagacc ctgagaaagt acaatgacaa aggggctgac tctgtccaca tgggtgtagc 42060
 ccaggtttcc cacaggagga aaacctatcc taggcaagag aagtggctct catcaaacac 42120
 tctatgaaaa gcaaatcaga ctcaaatgtc aggatttgtg ctttacagat cgatccggtg 42180
 agatgaaaga acttctgaa agtgtgtgaa ggccctaaagt cagggtctgt catggaaggc 42240
 actgactaca gaatgaggtg ccagaagcct agtcagagcc tctagggaat aaagtgtcag 42300
 atgatcttct aaaaaagttg aagtttcacc agtaacagaa tggccccaact attaaaatgt 42360
 gagcaaacctc agaagtcatt gtagcatata gaagcacaga cctatggatt gctggatgga 42420
 gccaggtat tcactccatc ctgaatagcc agctggggag ctagctcagt cagttaagta 42480
 tttgctatgc aaatctgagg accagacttt ggtctcctgc atccacagaa atgggtgcaca 42540
 cttgtaatct cagcactggg gaagcagtca gccagatcca acagctgctt agccagcgga 42600

ES 2 773 891 T3

aacagcctta tcagaaactc atgggtcctg gtgaaagata ttatctcaa taacaagggtg 42660
ggaagctcct gaaggacact ggaggttaac ttctggataa acataggctc gcccaccac 42720
cagtgagcat gtgcctaaat cegtacataa caatgatgta aagatggaat tcattccagt 42780
gaaaagtaag cctcctggac tctttttttt tttttgttgc tagatattct cgagacctca 42840
ggagagaagg tttgccatca tctatataac atgggtactca acttccctgt agtccacaac 42900
attcctatth ctatatgatg gagaagaggc cactgcccct cccagacatc tcagtctcaa 42960
atthgttacc agttccctct cctaataagt gcttaggggt agtgtttag agaagggctt 43020
taccatgaagt gtgtgtgtgt gtgtgtgggt tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 43080
tgtgtgtgtg tgtgtgtaac ctaaaggctt tccatgtttc cactgaaa ggttcttaag 43140
actgagaaca accagataag agtccaaatt ctagaaacca tgggaaagt taatattgaa 43200
agtcaaca aggcattgtg gtgctcacct tgaaccac cactggggc agaggcagtc 43260
agatctctgt gagttcaagg cccagcctgg tctacagact gtacatagt agttccagg 43320
ccagaactac atagtgat cttgtctggc caaaaatata taagtaaata aaataaatca 43380
gtacatggta acttgttctt atthcagtt ctgtttctca agcatgactt tggcttaag 43440
atthttcca acttgtttt gtgattgcca ctgtatcatt tctttgtgtg aagttactaa 43500
gtggtttctg tatttgatat tatgttctga cctagtttct tttcatatta aaccatttg 43560
tatatgaaa ctgcaaagaa gtgggtttt tgttttttg gttttttttt gttgtttgt 43620
ttgtttgtt tttcttgggt ttctcatgtg acctttcca tgtttgttc cagaatagac 43680
ctgcaagttg ggatccacac tgccatctga agtctgcac cccaagttc aggtatgtt 43740
tgatggcaga atagctttc tagactgtga caatagggc ataaagccac aaagcattcg 43800
ctttcctaca ggttatgcac cactctctg agtgattggc tgtgcatcat gaatattac 43860
aaaatggagg cagttcagtt tggagtgtg tcttttatgc gcttattcat ggcaatgcca 43920
atggaacatt cggcaacata tactactaat catgcatggt aactgaactg tgttgtgcaa 43980
ggaagacctc atatgacctc cctttgcata tgcctgacct ttctgtgaca gactcctata 44040
atactgagag tggactgtg tggagagtg tgtgaaaatg tattgtttaa ataacagaca 44100
gatgcctcta aatacaacac ccaagcagag aaatggagca tcaactggcac tttggaggcc 44160
tctgggtaac cttccagat cactgttt tcttctctc accaataacc actttcctt 44220
tggatgctac tcatagttaa catcttact tttgttgtt tccactgat gctaagaaa 44280
ataacttcaa ctagcaagca caacactaga tgaattaaga gtgatattga ctgtgtgtgg 44340
tgagtctcag aagactagct gcctcaggat tcatgaatgc ttacaggaac cctttagcaa 44400
ggtcaggaat gagtcttag atccatgtg ctcatagtct ccagcctgga catggagtag 44460
cacagtgtct gagtgcccc aaggaatggg cttgttcagg ctccctccc cgtcccagt 44520

ES 2 773 891 T3

tccaacaggt ctcagatcca ggacatcaga gctgagtгаа gagcagagct aaaaggagca 44580
 ccatcggagc cctagaagca gaataggggg ggacacagca cacagagaca agaactgagg 44640
 ccaggctgct gtgtgctttg ggcctaagtt gacagatгаа acatggtagg gtgaccacat 44700
 ggaggatgtc tgtgcacatc catcaaactg gcaggtcccc ccagcatttt ctgggagctt 44760
 ggggtcctct tttccatgat cttcagcttc tgtattctat gtgcgctggt accatttcat 44820
 cttggtagag tctatccttc tgttatttct tgagagtatg tcccaattct tgctggagg 44880
 tttggctaaa tatagaattc taagcagagg gtcatttctc cttcagatat ttaaagacac 44940
 tttctgtatt gtgcctcatt gccattggtg atatacctga atctaaattg atcccttggg 45000
 gcgtgactta tccccacagc caagggcccc ttccctctcg gtctgtgctc tggaagtctg 45060
 caggcacatg gtatgggtag ccactgtttc attcatagtt caatgctcog ataggccctt 45120
 ttgatttgat aactctatcc ctttccccca ttccogtga tgatttcttc ttttgttccc 45180
 cttttgatat agtttcttg ctgatgctgt gctaaaatat tctacccaaa aacaacctgg 45240
 ggaggagagg cttcatttgg cttacaattc cagctcacag tcattgaggg aagtcagggc 45300
 aggaactcaa ggcagggagc atggaggaat tgctgctggt cttcctctct gacttactca 45360
 caggttcttg taggctagct ttctgataac atctcaggac cacctgctta gcaatagtgt 45420
 ggtccacagc aggtttgaac cttctgcac agttaactaat caagacattt gcccaaagac 45480
 atgccacag gccagattga tgtaggcagt tcttaaatca agtctttttt gtcaagtgac 45540
 tctagactgt caagtcgaca gttgatgcta actaggacac tattctacca cttttcttgg 45600
 tagaaatatt attcggatat tggagttctt ggactagttt ttctggttct cttttcttt 45660
 cttttctgt tatttatatt tgttttatga gataggtct ctctgtgaag ttgtcctaga 45720
 ctttctggcc ctctgctta taattcctaa gaactgatat tacaggcagg tgccatgagc 45780
 ccaacgtttt ttcttttctt ttcactgcac tctgtttgag agtctcatcg tcacagtcac 45840
 tcacatcttc tattgtcttg tttttctttt taaatgtgca ttggtgtttt gcctgtatgt 45900
 atgtctgtgt gagggtgtca gatcttggaa ttacagttcc aaataatatt tctaccaaga 45960
 aaaagtggta gttgtatcct agttggcacc aatgtcacc ttgacagcct tgagtcacct 46020
 gagaagaaag acttgattta ggagctacca tgtggttgct ggtaattgaa cccaggacct 46080
 ctggaagagc acccagtgtc ctttaactgct gagccatctc tctggcttcc ttctattgac 46140
 ttttgcaggc ttctttcttg ttcttttgca atttcatggt ctctgactgt tcttcacaga 46200
 ctcttacctc atgcttaaga tgtctcttac tcttcaagg atactgagtt tttgaagttt 46260
 taattctcct gactactgtc ttttccctcc tgtttgtcat tctctgtttg ccctggcctc 46320
 tgtctttcat gcaggaagac ttttcatttg cttttaggtt tttattttaa ctattgggtc 46380

ES 2 773 891 T3

atgactaaag ggctagatga aaaggccagt gagaaggctg gagcatatgg gtgatacttg 46440
tcaaccggga gcctcaactgt ggaatgcttc agtggcatgt gaaatcctgt ggtatttgct 46500
caggcaagtg cagctgttga atgcagacca gagcagcttc cttcgaagga gtcagatggt 46560
gctgactgtc tttctgcagc tggtcaggaa ggtgggatag acttcagctc ttttcaaaca 46620
gtggtcacca aacaaccact tgcccagaga ctttgtgctt taccattctc agagaacaga 46680
cctctggatg gccccatggt ggaagcagcg cacctgtcta tcacaggtgc tctgaaggag 46740
ttggaagaac taccattgt ccacatttcc cacattttca catgccagct tcaactctggg 46800
atctgggtga cagtgggct gacataatgg caggggttgc agtttcagac tcagagtatg 46860
tggtaggaat gctgctgtct gaggggaagac tcatctgagc agtggaggct ttgcctgttc 46920
cctggcatca tttgacctgc ccctccttag aactgggaac ccagttcta aagctccctg 46980
ctttaaagat tctgtgttg ggtaagttct tagctttctc aggctaggtc ctctgctctt 47040
gggtttccac ggcaactgtt tttccctct ggctttgtga gtggttgtct tttgaaaaac 47100
tagttagttt ggaaaatttt gggaggagat caaataagat gtatgcattt tgccatgtaa 47160
gtcctaacca agccatctgc tgtggtattt tcoctgagttt ggttctgccc ctataggcag 47220
agtctgtcat cacagataat tgcattttga acttgagcat ctcccttct tctttgtctg 47280
cctgaaaaag tctctttata aaaaaatgta atgttaattt aaaaagtatt cattattctt 47340
gtgttgtgat acatgagtat atatatgcta tgatgcataat gtgcaggttg gaggacaact 47400
ttctgtagtt ggttctctct ttctccctc atgtaggttc tggggatcga acccaagtca 47460
tcaagcttgc acaacagcac ctttaccttc taagccttct catcagccct ttttttattg 47520
attgattggt tgattgattg attgattgat gctagggata gagcctaggg tcttttacat 47580
gctaagaaaa tgctctacca ctgaactgca ctccctagccc aacctgctaa attcttacac 47640
tgtcttcaaa aagaagctct gatgctggat tctgcaaagt ccatttttat ccctaaattc 47700
ctaaaagctgt ttaaactctg tgagtcttac tgtacagacc agctctgtgc accatcttcc 47760
acaatctoca tgacctctc aggatgggct ggtatctctg cagctctgcc cagtgcctac 47820
caggaactta caggtgtcac caatgaattt attggtgcat gctcaactca tcttgtccct 47880
atccactttc tgctttgact ccttctggta agagacaagt gtgttaacta cttgtgctat 47940
caccacacag aatccatat ccataatct tagtcctttt tatttactta tttttgagac 48000
agggtcacac tctgtagctc ccacactggc cttaaacact gacctogaac tcatgggtgat 48060
tctcctgect aaacttctca aataccatga ttacaagagt gacacacat gctgggagtc 48120
ataatcttaa gtttaaaagt gagggactgg tcagtttact gtgctagggt gacattgtat 48180
agaaatgaac agccatggtg gtctggaaat gttcctagtt ttcatttgta caaggatatg 48240
cagtgtgtga aatagggaga gtcttaccta tgtggggttg atcacagcaa ttaataaaat 48300

ES 2 773 891 T3

atgctctaaa taatgaaaa agccagtaac tagtagtggt tctgaatcct cactaaagct 48360
 ttaatacatc ataaataata taccactgca gattatgtct acatgttata catatcacat 48420
 ttatagtaca atctgatcct tgtcacctac tgtaagcaca actgaaaaac aaatcttctc 48480
 atagctcaat attaagtcac tattatcccc ataataagta attattatcc ccataatgaa 48540
 actatctatt gagggagtca gaatctgaga tagttaaata aatttaagca tgtattttta 48600
 gtgtcaatgg taaaaattaa atgttcataa agcctgtatg actcctttta aagtagtttt 48660
 aatcttatgt gtatacatat atgcatgttt tgccttcttg tatgtctgag taccacttgt 48720
 atgtctgggt cctgaggagg ccagaacgta tcagatcccc tgaaactggt attacagttt 48780
 tgagctacta tgtggctggt ggggaattgaa cctggatgct ctgaaagagc agccagtgct 48840
 cttaatgact aggccatctc tccattttct taaaaaaaaa tttaaacaat ttactctaag 48900
 atttactttt atgtaggtgc gtgtgtgaat gtgtatgggt tatgcattgg ggtggggagg 48960
 atggattagc acagtcacag aagactagag gagggctctc actattgctt tctgtcttct 49020
 acccttgaga cagggctctc cactaaacct gaaactcacc tttgcagctg gggtagctgg 49080
 tcagaaagat cctggaatct gtctttctcc ctggccctaa tgcttgagtt acaggcccat 49140
 gtgaccatac ctgtcgtttt actgggggtc tacagagtca aaccaagtc ctcacgcttg 49200
 catagccagc gattttaccg actgagacat ttatctgccc caattcataa ttcttctctg 49260
 cttccattaa taatcccac tctgtcccct tcatacatat ttctgaaata gacaaaatga 49320
 atacaagtta gacatcgagt ctgattaatc ttcaacttct ttgataacca ggtattgatt 49380
 tctgactttt gaagatggat gaaggcacag aagtctccac tgatggaaat tccctgatca 49440
 aagctgtcca tcagagccgg cttcgcctca caagactttt gctcgaaggt ggtgcttaca 49500
 tcaacgagag caatgaccgt ggcgaaacac ctttaatgat tgcttgtaag accaaacaca 49560
 ttgaccagca gagcgttggt agagccaaga tggttaaata ccttctagag aacagtgctg 49620
 accccaacat ccaggacaaa tctgggaaaa gcgctctgat gcacgcatgc ttggaaagag 49680
 cgggcccgga agtggtttcc ttgctgctca agagtggggc tgacctcagc ttgcaggacc 49740
 attctggcta ctcagctctg gtgtatgcta taaatgcaga agacagagat accctcaaag 49800
 tcctccttag tgcttgccag gcgaaaggaa aagaggatcat tatcataacc acagcaaagt 49860
 caccctctgg gaggcatacc acccagcagt acctcaacat gcctcccgca gacatggatg 49920
 agagccatcc gccagccacg ccttcagaaa ttgacatcaa gacagcctcc ttgccactct 49980
 catgttcttc agagacggac c 50001

5
 <210> 3
 <211> 50000
 <212> ADN
 <213> *Cricetulus griseus*
 <220>

<221> modified_base
 <222> (13508)..(13531)
 <223> a, c, t, g, desconocido u otro

5 <220>
 <221> modified_base
 <222> (21446)..(21479)
 <223> a, c, t, g, desconocido u otro

<400> 3

```

gggctcaggc atttatcgtt cagagattga ctgagctgta aagatggaaa gacaaacttt      60
tttttttttt gattgagtcg gggtttctct atgtaacagc cctggctgtc caggaactca      120
ctctgtagac caggctggcc ttgaactcac agagatctgc ctgcccctgc ctgtcgaatg      180
ttgggattaa aggtgtgagc caccaccgcc ccgctgacaa actagacttt tagaatgtat      240
tatgagataa ggttttgta tgttgcccag gctggactca gatctgtagc aatctatctg      300
ctccagactc ctgagtgctg ggatatacag acctgagtta cctgtacagc tttctaataca      360
tcccccgctc ccccagagac agggtttctc tttattgttt tggagcctgt cctggcactg      420
gcactcactc tgtagaccag gttggcctcg aactcacaga gatccacctg tctctgcctc      480
ctgagtgccg agattaaag tgtgcaccac caacacccta ctttctaatt cttaaagcaa      540
ggctcccaac tcctccctg tgtgtaatca acaaggttct tagaccctgt ctgcagtgtg      600
gattccact aataagacag tggcggcaca gtgctgtgtg gcagagcaag cgtccatcta      660
gttcctattg tcattctatg atttgctctt ctgggagcct tgtcattcag caagttcctg      720
ggcttgtctt gggattgcaa tgtgcctcag cttggctagt tcctctgagg cagaagcagt      780
gtttgaactc agtgggact cagtcactac atctaacttg tttgagggct ctctgcattt      840
gctttccaat taaggtttag gatgactcct ccctgtgact cttatcatcc tgcctattaa      900
tgctaaatta gagaggcatt caagataact gccgaagatc taataaataa atggggtggg      960
tgggtaggac tataaaccag tttatagcat gcaagaaagc tctgagcacc acattcaaaa     1020
ataaagtgct gtgagcctgg tgggtggggc tcacaccctg atcccagaac tcaagaagta     1080
gacagaaggc tcagattcaa gattcaagtt cttccactat acagccaatt tgaagtcagc     1140
ccagactaca tgagaccctg tctcaactaa gcaaatgaaa gcaaaactggg gtccaaatag     1200
gcactattcg atgttttgat gcaagtttgt gactgaggag tggaggtggc aatgaagac     1260
ttttttcttc ctctctctct tctcctggg tcccgttttt tttaggggtg tcttaggata     1320
tgtatgtctc attggcacta ctaagaagtg tggggctctag ggaacttctt gttatgtata     1380
caagctaadc ttcaacaat tgtgtgggct gttttggtaa ctactcaaat aatgctatag     1440
aaaattgtac aatatattgg ggaaggaagg gagttttaca caggagtcaa catgactctt     1500
    
```

ES 2 773 891 T3

gtctctggaa agcaacttgt gatccaatga ggagctaaat ttagagacac aattcaggaa 1560
 gagaatccaa tcagagcttc ctgttaaaac aactcacctt cacaaacaag ttcattccta 1620
 atcgaattta aggtctagaa actgccaaacc tattaatggt tctataaata cacttgggggt 1680
 caactacgta gccaaaggaaa tctttaataa attgaacaca aattgtcagg ggaaggttat 1740
 tgctgggact cctggaagca tgtataagca gggtaggggt gacatagggg tggggggcag 1800
 ttaactcaca gatattagtc tcagatatta atggcttgtg tgtgagctgt ctgccacact 1860
 taatgtcagt cacottgccg ggaactattt ttctctctga ttccaaatgt agctattggg 1920
 ctattaaatg attaacttcc acagaaactg ataatatcct tatggaatct gactgtggta 1980
 agcctgtaca cccccgccc aatttccttc tagatttaga attccattcc atgagccatc 2040
 acaccacgc tgaaaaaaga aaacctgttg aatcaaattt gtgttttga gggaagagc 2100
 cacccttcca atttataagg ctgtctattt ctttgggggg ggggaaatga accagtatct 2160
 tctattagta aaaggagtgt ttgagcatgg gcactacaac ccacttcttt caggagatt 2220
 ctttttctc tgagaactca gcctctctgt gctggtgcca caggaattct taaactcttt 2280
 caactctcca attaaccaga gagcaaaccc agcactttcc atctatgaga aatctacacc 2340
 actcatggaa tcattgtgtg ccctctctca ctgcctaaca ggggtaccct tgccaaagaa 2400
 aagcaacta atgccaaaaa ggtgcatcac ctggcactgc ttccgaggat gggcaatgtg 2460
 caagcacttt gttcagtgcc tctgccttgg ggtctcttga ggggaggcag gttacctggg 2520
 gtgggggggc acactctctg aagggtgggt gcgttcagtt tcctgcttca ggggctcctt 2580
 catagtaccg cccoctgatg agtttctgct cagactggaa ggtgtcaggt cccaaagaaa 2640
 cctgggacaa ggctcactca gtacctgtcg cttctcccag cacgtctcac cccacccta 2700
 ccctaaactt ctctagccca gaggtgggc tcccccttc tctttctac ataaccctgc 2760
 cattttagct gtgagctctc tccgtcttta gctcctctac tgttctttta tcctctcttt 2820
 tctctctcct cttcttctct cacccccacc cccaccccga tctctcccc catggtctgg 2880
 ttcagtctgg accctttcag atgcctctgt ctgaactctc cctcatatct caataaaacc 2940
 cttctcttca gccacgcctt ggagaggtca taggctcatt ttcgttcaga aggcctatca 3000
 aagaatctgt gggcttatct ttacattcac aataggcagc ttggccctga gaccacagtc 3060
 caggttaaag tgttaacctg gaaagaaagt cttttattca aggtgtctgg tttctttct 3120
 tgtttttgtt tttgttttg gagacaggggt ttctctgtat tattttggag gctgtcctgg 3180
 aactcgctct gtagaccagg ctggccttga actcacagag atccgcctgc ctctacctcc 3240
 tgagtctctg gattaaaggc gtgagccacc aacgcccggc tcaagtgtct ggtttctttt 3300
 gatgtcttta gtttctttaa tcccataatt cctttaatta taccctcttg tctgtcggag 3360

ES 2 773 891 T3

aatgacatca aggatatcca gttcaaggtt tcctatgtag ttcagtcata gagtgccttgc 3420
ccagctgcca gactctgtca gatgcccagc accacacaca tacaagcat tccagctct 3480
gtgtctgtgt caattactcc tgtctgcttc tccatcccca gacaccagga gggcccacaa 3540
gaagcttgga gcaggaaga ataaagagc aatatccata gacacacaaa acctccaaag 3600
tacttatgca ttgaggaatt acagcttaca aatccagtca cagtatctat attcatgtta 3660
gcctgatttc aatccccag ctacatattc tccatgagc tagctccttt cctattcaag 3720
actcccttga taatagttgt tatcagactt taccctatt aaaatatttg gaccgtttga 3780
gagcaatagc tcacctctat aatctagaac ccaggaagtt aaaacaagat gtttgctgca 3840
agtttgatgc cagcctgggc tacatagcaa tttccagaac atcctgagct acagggcaaa 3900
attctatctt aaaaaacaaa aagtagacag atcaggtggt tcaccttgtt tcaaaaaatg 3960
caaaaaatat tttttaattg tagaaatata tacgctaatt cctttggtac cctaggccaa 4020
gtgactagat gggtagtct tccttctggt cctcacagaa gaaagttaag ttctcagcag 4080
gaataataaa aatatataa aaaaaaaca agctgcaaaa ttctgttggt gttctgccaa 4140
agtgttctca ggagtgaggg catactggga tttagtcaag cagatatttc tgtttgaata 4200
actaggatct gggagccatg ggacaccacc cccaccata agggctactg aaaaccacc 4260
ctggaaatct gtaaatattg ctaaggctct acccttttgc tcagagaaca accaccaca 4320
aggatagggg ataagttagt tctgtagtag agtgcttgc tagcacacag aaagtcttc 4380
tctctctgtc tttctctctg tctctctctc tgtctctctc tctctctctc tctcacacac 4440
acacacacac acaaacacac acatgagtgc acaagaaact tctaggtgct actaaactaa 4500
tgtaaaatca tgcaaagttc atagagaatt caacagctag tgacaggatg acccgaacac 4560
aagattctgc cctagtcctt gtattctgta gtcccagtt tctctttact gccacagtct 4620
cctatctctg acagcctccc tctttgcaga tctggcagtt tctggcctg gaactgcttt 4680
ggtagaatgt ctgtacagca tgcactaggc actgggtttg atccccagca ctgcataaat 4740
caactttgat gtcacaccta taattcagc acttggcagg gatcgaagca ggaggatcag 4800
aggtgaatca aggccagcct gggctacttg aaaccctggg gagagggata gaagaagggg 4860
gaggggggag ggagaagaaa ggaaggaggg ggagggaga ggagaggaag agaggagggg 4920
gagggagggg aacagggagg gaggaagaga aggagggaga gagggaggag ggaggagag 4980
actagtgtaa gcagaacctg taagttctct cctcagcctc aacacacccc agctccctgc 5040
tgtctcccgg tccagggctt cagggcctgg caggacaggc agcaggttgt tttgctctca 5100
taaagccatg ttacataact aactaatggt ttgagcagtg gagctgagcc aatctaggtc 5160
acatcaagag ggaatgggga aagaggatga tcacggaagt ggtgagagga agggaaacaa 5220
gaagggagga ataaaaaaaa gaggcgagag tggaaatggg gtgcgattat ttaatatctg 5280

ES 2 773 891 T3

ctgcctgttc atagttcctg gtccttaggg acagcatata ttatcctgaa aagtcctctc 5340
 tctatTTTTat ctaggcattc tgtcatccta tagccccac tctggatggc tgaactctgt 5400
 gccagcagcc tgcaggtatc accccttatt ggagtgaggt ctattcctta ttggaagcag 5460
 tggcaggctg gtaggaaaca aacaggcctg gtgttgtgga atgctgtcct cccagcatga 5520
 ccatcattag accttatgga agcagagcga ggggggcatt gtcctcctcc ccaggctcct 5580
 gcaagcctac tcagctcaac tggttccccg ggccagactt aggtgcaaga gttgctttgg 5640
 tttgttattg gtggcctgtg tagctgagta gacacatgct cacctacatg atatatgatg 5700
 gcttgcaacc ttctaaaagt tcagtttcag gagatccaga accctctttt gccctccaag 5760
 gacaccagac acctatgtgg taccatacag tacatgcggg caaaacactt gtgcatataa 5820
 aataaaaaga gatggctccg tggctaagaa tgctccctac ctccagctca cccacatctt 5880
 cacaactgac tgtgaatcca tccatggttc tctctgacc tcggagggca cctgtgcccc 5940
 tggggcatac acatacacat acacaaaaca agtatgtaa taaataaata tttaaaattg 6000
 gggctggaga tggcttagtg gttgagagca ctggctgatc ctccagaggt ccagagttca 6060
 attcccagca cctacatggt ggctcccaat cacctaaagt gggacctgat gtcctcttct 6120
 gacataaggt catacatgca gatagaggac tcaaatgcat aaaataaata aataaatctt 6180
 tagaaaataa gtacataata aataaatatt taaaatgacc caaattaaga aaaaaatgaa 6240
 gccaggcagt ggtggtacac tcagaaggca gaggcaggca gatctctgag tttgagacca 6300
 gcagttccag gacagccaga gttacacaga gaaactctgt ctcaaaaaaa aaaaagaaaa 6360
 aaaaacagag aaagaagaga ggagaaaaac aagaacaaaa aataacaaaa caaaaacatg 6420
 gctttccctt catggcatct gcttcatctg cctatttggg aatgatcagg gcactacaca 6480
 cccagtgctt cataccctgg ccattgttct gttcttggtg tcaccaccaa gtttactaaa 6540
 gatggttcca gaggacatt agcagcccca caccctaatt gcagctagca gttgaggaga 6600
 tttctggctt tttgtetaag aggaaggttc tttggctagg agatatactg agaaggacta 6660
 ggaaaagggg tgtctaagaa acttgagagag cacatttttc aagtcagaaa gaacatagac 6720
 atattctggg ggtgggggta gtaagataat ggaccctcct aagggaagga ttgtgggggt 6780
 tgctgaagg ggctgaagca gaccactgag caggccagac caccagcagc ttttgagagg 6840
 tgggaacact gcagctgaag tcacttgtca ccttcccagg tagttottac ttccagctct 6900
 ggcagggcta gatagcctag gaactcccag ataggagttc tagttcttct tctcccagc 6960
 tgacagaacg tgagctcaga gtctagggac actccaggtt aaggacgggg ccattcttga 7020
 ttgtcagcac agatagattt taattagaga gcaatgacat gacagataaa cagcccctta 7080
 tctaaagggg tacatcccaa gaccctggag gactcttgaa aaccagata ggagccagcc 7140

ES 2 773 891 T3

acggaagcat atacctttaa tctaagatt tgggaggctg aggtaggagg atctctgtga 7200
 gtttgaggcc agtcttgtct acaaagtga ttttgggaca gctacacaga gaaaccctgt 7260
 aagaaaaaaa aaaaaaagaa agaaaggaag gaaggaagga aggaaggaag gaaggaaggg 7320
 aaaggaagaa aaagataaag gaagaaaatc caaataggaa agaatoecat atataccata 7380
 tttttcttaa acatacatag gtttattcat tctctctgtg tctgtgtgtc tgtgtgtctg 7440
 tgtgtctgtg tgtctgtgtc tgtctgtctg tctgtctgtc tgtctctctc tctctctctc 7500
 tctttctctc cctctctctc tctttcttgt ctcataaatc tcaaacactca gggaccocaga 7560
 agatatocca gtggtaaga atacacactg ctcttgaga cctaaactca gttccttgtc 7620
 cctacttggg gcagctcaca accacacctg taagtctagc tccagggaat ccacaccttc 7680
 tggcctgtgc aggcacctgt gtgaaggagc acatatacctt ccccataatt aaaaaacaat 7740
 cattgaaaaa taaaactcaa ccccctcccc cgggactcaa accagaggta gtctccctgc 7800
 cgtagggcgt caaaaactgg actttcaggt gtgagcctct aggccaggct gcttttctta 7860
 actggctacc gtgctcttgc ctgaaacttc cagcttgaga cctcatagta aaaagaacat 7920
 acacgtcttc tgtctgtact attttacaga cgctgacat gttcatacca cgtattttag 7980
 caatttcagc acttggata ttttctgtca ttctcaaata actttcacct tgccacttag 8040
 ggcagtccaa ggctcctctt agatataatc aaattatcag ccaccaacttc tgcctttact 8100
 aagtaagaca gggacttaa catggagtac ttaacacaag cactgtgatc tgaaggtgga 8160
 gactgcttgc tactcagtca cagcttagca ttgctagaac aaatcctgaa caaagggtaa 8220
 ttcattgacc aggcagggca gaggcggatg gctgttcttg ctctcagaa acccctgtgt 8280
 ataatttcaa gcttaggagt tgtttgtctt tggatggaga gggctcagacc tagggcttca 8340
 ctcacactag gcaagcaccg caggtctacc ttogaagaga agaattttca cttagcgttt 8400
 tcagatatag gtcaacctca gctggctgaa actttgacta agtgagcaac tgtgaggggtg 8460
 gggaacacat gcatgcattt ctcatgtta taacatctat ttatacataa acatatacata 8520
 taaatatatt ctattgcata taaatataca taaatgcaca ctcatgtata gatatacaatc 8580
 acataattta tgcttttatt catagattat ctctgggagg tgtacaatta ctgacaatac 8640
 ctgcacatga tagtacagt tgttctagtt aggtttcttt tgctgtgaca aacaccacaa 8700
 ccaaaagcaa cttgcagagg gaagggttta tttcagctta cagttgtatt cattatgaag 8760
 agttgggaag tcaggacagg aacctggagg caggaactga agcagaaacc atggaataat 8820
 gctgcttact ggtttaccba ccatgactca acctgctttc ttatatacacc aggactgctt 8880
 gccagggat agaaccacac atggggactg tacctcccac aacaatcatt gatcaagaaa 8940
 tgccctagag tcagggatgg tggcaaatgc ttttaatccc agcactcggg aggcagaacc 9000
 aggccttgac tgtgaggtca aggccaggct ggtctacaga ttgagttcca ggacagccag 9060

ES 2 773 891 T3

ggctactcag agaaacccatg tctcatggaa aagaaaagga ggaggaggag aaaggagaag 9120
 gaaaaagagg aggaggagga ggaggaggag gagggaggag aggaggaaag aagaagaaga 9180
 agaagaagaa gaagaagaag aagaagaaga agaagtagaa gaagaagtgt ccaactggaca 9240
 atctgatggt ggcgtttccc aattgaagtt ccccttccaa gataactcca ggatgtgtca 9300
 agcagacaaa aacaagaacc aagacacatg tttataatcc caaactggg gaagtggaat 9360
 aagaggttg gcagtttaag gccattttca gctacatagg gagtccaga ctatcctggc 9420
 tacatgagac cctgtctcaa aacacccaaa tgcaagggaa aaacaaaaag caaataatg 9480
 agtacaata gcagtgacat tctggggaga cagcctggag ggggggattg cttattatct 9540
 ctccctaccg tttggagttt ttaaatcat gaatctaacc ccagaaaaa aagcattgag 9600
 attctgggac actcgggtg tagagaagat catctgatcc tgtcacctt cgggtacgtc 9660
 actttattaa tctctctgag attcagttc atcacctctg aagtggttg tgtcgacgta 9720
 cagtcctcag gactaagtaa ggccacttg tggtgtgcc aaagcactgt gtcagggaca 9780
 cggcagatgt ctgacacatc ttgtagatt ccttttctgt cctccgctcc cctacccag 9840
 aggtgggtac agcccatg caccctcatct ttaatggctt gggtttctt tctccagcca 9900
 ggaaagtgt cgctttggtg acagctattt taagtcaact gaccttct gcaaatgatc 9960
 cagatgcctc tatcttaggc tggtagtac gaagatggcc tatgacggg ttctggggg 10020
 tgtgtggga ggtggggcag gggggggcc cggcatttgt cagaccata tgatctctg 10080
 gctcccgggc tctgcagatt tctcctgctg gagatgccta cctgccagca atcttgaga 10140
 agacagaaat agcagcttg ggttccaggt cccctcctcc cttggcca atgtagctag 10200
 agctttggtt tcctgctgct gtcttgggtc ctggagccct ctctggatg tcatggagtc 10260
 ttgtcagaga agcaacttg ggtggcaga cagtcattcc agaagacatg atctggaaaa 10320
 actgctcat cgtttcttc agaggcactg tcccagccc atttcttgt ctggttctg 10380
 aatctcagg gatgccatca gaagaagtg ttcttgtgt tactttggac atggtttct 10440
 gtagtgcaga ctgccctta actctacgta gctgaaaatg accttggctt ccagacctt 10500
 tgatctgtca gcatccctgg gaaatccagg gttctgtaat cctcccctt caccttgact 10560
 tactgtacca gcatcaaca tccaaaca atccagtgt tagccaaata cagcgggtca 10620
 tgtctgtaat cccagccacc tgggaagccg aggcagaagg attaaggag ctggaggcca 10680
 gtctgtgcaa tttagcagga ctgtctcaa acaaaattta atggttaggg gtgggcatgt 10740
 ctttatttg actcttatca catgaacaca cctgtaatct catcacgaaa cgacaaggca 10800
 ggaaaatcaa aagttcaaag tcatcttgg ctacatagca agttctaacc tgacctaggg 10860
 tatgtaagac cttgtctcaa aagcaaaaa acaaacccca aataacaaca acaacaaaac 10920

ES 2 773 891 T3

aaaaagcaaa caaggagagg gtgtgcagct agggatataa ttcaatgggt gagggcttac 10980
ctcacatgca cgaggccttg gtttcaactt ccagttgaaa tgaagtttag tggttagagtt 11040
ctgtgcaagg ctgtagtttc agctctccat actgcaaact ggaaagaaca acagtgacaa 11100
acagaaacaa aaaacccccca caaacaatgt gctttctcac tcaataaaac cacctcttta 11160
catacaacta caactgctaa gaaagttctt cagtgtteta gagcctgagc acctcaaatg 11220
gtttccataa agctgtatgc aaacactgat aagccacgag aagcaactgt acaaagcacc 11280
ctttgattht catagthtat ctacacaagg attctaggaa agtgtgctag gaaaatttta 11340
tgtatcagcc ttgctgggtt gtccaatagt tttagattht gccagtgaag atthtcctth 11400
ctthatttht tacatgggaa ggaagthtaa ttgggggaaag ggacgggagt gggctthatt 11460
ththatttht aatgagacta gcatttgcat tgggtggacat tgaaggaaac agthtccct 11520
ccctaatgtg tgtgggcctc acctaaactca ttgaaagtct tagataaaac taagctgagt 11580
gagtgagttg gccatacct gtagatggaa ggaaaagggt cttgagtht gthttatcct 11640
agagagaact tgatcccca aacaccaaac thtcaaacca aacccagcc tctcagtgt 11700
gaagggatgc tgttacatga ccacctatgg actcagacaa cctctctcc ctgagtctgc 11760
tggcttactc atcagagtct gggctcacga agccgccaca catatatgag cctcgttctc 11820
cccactcttc tcttgtggca ctgaggttca aaccaaggac ctgcacatg atagcaaata 11880
ctgtactgaa ccatagagcc agcccttgtc agthtcttaa cacaaacata tagatgtata 11940
tgtatatgaa tthtccatg ctaccaattc cthtthtctca gagaaccaa gaatacacca 12000
agtagtcaca cttgaaattc tgttctgaga ttgaataaaa cctgatcaaa tgtgaattcg 12060
gtcccttctc cccatccct gacgccacca cgttgcata cagaccaggc acaaactctt 12120
ctccttgtga atgtgtgtaa cacatgttac cactgtgctt ggctthtga gttagaagg 12180
tggttgatat ttaaaaaaa actthaatat ttagtcatta cthtthtagta aagatttgcc 12240
ttgctthtat thtattcatg tgcattgtg tgtatctgtg tgagtgtatg ccactgtgt 12300
ttgggtgcct ctggagattg gaaaagaatg tcaaatccc aggacctgga gttccaggca 12360
gttgtaaact tcccaatgtg ggtaattata atgaacttg atcctctaaa agagcagaac 12420
tcactcttaa ctgatgagtt atcctctac ccccaaattt atthttht gthtthttht 12480
tthtthttht agagggtctc actgtgtagc tctgacagta ttagaattta ctatgtagac 12540
cagacttgat aatgtctaa ccctagaaaa aatagtht gthtthgatt tatgtctgtg 12600
ccatccactc cttgaacata tthtthgtat ctgtgaagcc agtgaaggct gthtthtccc 12660
ttaggactgg agttacagat ggctctgagc taccatgtgc atgctgggaa acaaactcag 12720
gtcctthtga agagcaaaaa atgtccttht atggtggtgg thtgaatgag aattgccta 12780
tcgagcataa aaacttgga gctthtggca catgthtctg gattaagagt caagaaggat 12840

ES 2 773 891 T3

acaagaaagc ggttgtggaa tcatccccc a tggttaagga aaaccaccaa agccaggctt 12900
 gtggcagggg agttcctgca tggaggccaa gagaagccac tatgtcaagc tgtgaagggtg 12960
 aagcctggat tgtgttggag acccaagcta ctggagatgt aagagatgtg agataatgcc 13020
 caggagagct gcagacaggg catggaatca ggccaagcga gagaagtgtg ttgcagtcag 13080
 cagaactggg aggggaagagt catctaagtc ctttgtcatc agacatagag atacaggatc 13140
 tgaaatthgc tctgctgggt tttggtcttg atttggccca gtacttccca actatgtccc 13200
 cttttctccc ttttagaata ctaatttata ttctgtgcca ttgccgggtg atcaggatgg 13260
 ttctcagata ctgttttagt tccatgcctg tctacttccc gtcatgacag tcatgcaacta 13320
 acactctaaa actgtaagca agctcccaat gaaatgtttt catttataga ggtgccttga 13380
 tcatgctgtc tcttcacagc aatacaacag tgattaagtc agctgctgag caatctctct 13440
 ggccccagaa gtatgcatgt gtgcaattgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt 13500
 gtgtgtgnnn nnnnnnnnnn nnnnnnnnnn naggaaatgt cattctgtaa atatgtttat 13560
 cttattgggt gatgaataaa aactgttgg ccaatagggc aacaaaatag gtggggccag 13620
 gatataagga ggattttggg aagtgtaggc agaggggaat tgtcatatga tcccaggaag 13680
 agacatagat gggcagaaac tgcctctagc taacataga ggtctggagg tctgtacaga 13740
 caggcaggaa gtgatgtagc tggagaatc agaataaag caggaacaaa caggaaatcg 13800
 agctcttctt ctctctccac ttccagagatg ctgaacagtt gagatgcagg atgccagaag 13860
 agtaagaggt ccctggacct ttctccagta agataagacc atgtggaaat agattgatag 13920
 aaatgggtta gagattaagt cagagctagc caataagaag ccgtagatat tggccaaccg 13980
 tttcataatt aatatagcat ctgtgtatth atttggggga cctggtagac cagaaaactc 14040
 gtgttagaga catcttatca aagttgaaaa aagaaaaaat gtgataaagt taggaaaaaa 14100
 tatagtaaat gttaaaagct aaattctaaa actacaactt atttatcatt tccataaatgt 14160
 ttaaaaatat tattttataa tgaagatact taaaattcat ttctctgtct tttgagacag 14220
 ggtctcagtg tccctggaact cattatatac agcaggctgg cttggaactc acagagatcc 14280
 acctgcctct gtctcctaaa tgctgggatt aaaggtgtgt gccaccaagc ctcaattaaa 14340
 atgcgtttct ttttcttctt ttcttctgt ctttcatttt tttgtttgt tagatttttt 14400
 ttttagaca gggtttctct gttagcatta gttgtactgg aactcactct gtagaccagg 14460
 ctggccatga actgagagat ctgcctgcct ctgcctctct agtgctagga ttaaaggcat 14520
 gcaccaccac tgccaggctt aaaatgtatt tcttttttta atttagaaat ttattctgtt 14580
 taatccacac gctttatata gctttagtta agaaataaaa taaaatgaaa cagtgaacc 14640
 aagagactat gtccaagtc aggtcctccc agcctgccaa tgccaagagc tctttagttc 14700

ES 2 773 891 T3

tgtgtaccaa ttggaagagt aagaaaaaaaa tatggatggg aaccacacag tttcataaaa 14760
 cagatttatg gaactgaagg gtccttgctg agtctagcaa attgccttta caaaagagaa 14820
 agaaaaaagg gggaggtaga aaaacaaaac aatcaaccc aaagaggaca aatcccaga 14880
 gttctaaatt gaacttagaa cctgtcacac tgggacagaa gcttcagcat ccatgagctg 14940
 tgcctcccct gctctctaga gctgggatct cgagggtgtca gcagagaccc cacaggtaac 15000
 aggagcaaaa aactcactc agacctttgt ggtacttcaa cagtggctc acttctgggc 15060
 aagcttacia acctatacaa agttgaagg gtactttaca tgagtgctaa acttcaagag 15120
 gaaggaagaa aaaaagggag gtggagggga cagagagaga gagaaaaaaaa caaaacaaaa 15180
 caaaaacaac cacctcagga gaggcaagg catttaaagg aaccacaaga atgccaacga 15240
 tattaanaatg tatttcttaa tagtaaattt tatgggaaaa gagagtctcc tcttctcca 15300
 agtaggctag gtaagtacct tgccactgag ctctatctat acccttcaa gtggacaaaa 15360
 tgacaaagat agttcatctc ccccaaaggc cctggtgggg tgctgattgt cacatctggt 15420
 gagatttctg tttttgtttt ttttcaaga cagggcctct ctacatagat agtctggct 15480
 gccctggaac tcaactctga gaccaggtg gcctggaact catagacca cttgcttctg 15540
 tctcccaagt gctggtgcta aagggtgtca ctgccactct ttttaagtaa ctatgagttt 15600
 caaaacaaat taaagagcac tgtaaaagt gcttgtttgt taagcctagc ttcaagtcaa 15660
 aggcccagag cccccctacc aaccagctgc tatcacctag aactgtctg tagatcttgc 15720
 actgactcaa aactgtggcc taaggctaaa ataatggtct tcctggattc tgatgtgagt 15780
 gagattgtgt aggagggctg gccgctggcc tggcttgagt cactctcagc tggtttcatc 15840
 ccattcctgc aactctgtgt aagaggtgga tgatccttgc ttaactgatg aagaaaccaa 15900
 agctgtagaa aggatcattt gcttaactct tcacagatgg caagaggcag agtcaggatt 15960
 ggcagagtca cttctgcaa cttcacctc ctgctaactc caccctctg ctaactccac 16020
 cctcttgett atacttgaca gtggaggaaa agccactgag ggaattaaaa gttgttactg 16080
 gtaatggtca gaaaaaagc tgaacaaagg agattagatt cagggatctt tttctgaaa 16140
 gaaagaaaga aagggggact atagtctaga aatgctgaga taaaaggggtg gattatcata 16200
 tctactctca aactaaagaa gcaactacta gtctcaaata ctttatattg gtatggattt 16260
 ttgtgtattg gtacaaattt aaggttattt ttgttatact gtatatatgt ttttctttct 16320
 tgtttaaggt attgtacctg tatagcttat ttaaaaatgc aatgtaaca tatagtcctt 16380
 gaaaactatt taagataata aagaaataca ggttaatagt catctatagc aatcaaactt 16440
 atagtcatgt taggtatggt ttcaaggga tacagaaata aatttgagat agataggtea 16500
 tcttcaaaaca ctccagagat ctacagaaaa tggcatttat aaaatgtttt aatgacataa 16560
 gatTTTTcat gatagtgaga aatgtctact cttggcagca ccaatttact tcaaaaatgg 16620

ES 2 773 891 T3

acaatgggca ttgaagaaac tccatgtgga ttttgctttc tttgtggcaa aaatctagct 16680
 atctgggcaa gaaacttccc ttaccttgac tgctgtccta actggacaag caggacataa 16740
 aagaaattga ctgctgaact ttgccaagat agtatacatt agtctttcaa aaatccctgc 16800
 tttacaaaaa agtctatcag atattctaag cttctaggcc aaagatggat gcttcaatgt 16860
 taacagagga atcttctgtg actgatgttt ctgtcatttc tatagttttg aaaattgctt 16920
 gctctgttct tccctgtttg ctcaggtagt attatttctt tcttgagtgt ctaatggagt 16980
 taaagactag atagttatag ctacagtttt ccttghtaacc aaattcagaa aagaaactcc 17040
 caaaagaggt gtaaaagtat gaggctgaga aatataaaaa cttaaattta tctaagaaaa 17100
 tgttttgtta tctaaaaaaa aataattttg ggtagtaat acaagttagg atagaaaatg 17160
 aattaggtac aaaacttttg actcatcaag aaaaaataga taatggagta ttttctctga 17220
 atttgccaaa tacaaataga ctgggtattg taaatgtaat tcttacttga taattgttct 17280
 tattgtttat agtttattat gttagagtca aaacctttct tttttattta gacaaaaagg 17340
 gggaatgtag aatatttctt tacactgtgt gaagatgtat cactgtgatt ggtttaataa 17400
 agagctgaat agccaatagt taggcaggaa gaggttaggt gagacttctg ggaacagaag 17460
 tctcagggaa ggaaacaggc taggtcacca gctaaatgaa gaggaaatag gacactcagg 17520
 aggagaggta acagccacaa gccaaagtggg ggaatataga tgaatggaaa tgggttaatt 17580
 taagtcatag gagctagtta gaaacaagcc tgagctaaag ctgagctgtc ataactaaaa 17640
 gtggagcttt cataattagt aagtctctgt gtcatgattt gggggctgac ggcccaaaaa 17700
 agcctgctac ccaagttctt ttcaattttc aagttctagg attctggcct tttattggaa 17760
 aacctgtca agtttctata gaggtctgac tccacagtgt tgctctgtca atgaaattta 17820
 ttaattttat tccgaggcct tgtgcactct ggataatcac tgtaccactt aatctatatt 17880
 cccatccttc attataattt aaaatggtct tattaatctg gtcacttggc tttttttttt 17940
 ttttttttct gagacaggat ttctctgtgt agccttggcc atcctagaac ttgctctgta 18000
 gaccagcctg gcctggaact cacagagatc cacctgcctc ccctccagag ttctgggatt 18060
 aaaggcgtgt gccaccacct cccagtgagt ttatgtcttt gcaaattata catggtttca 18120
 gttttttttt ctgtttgtaa gtcactttat ttcaaagtga aagtttaaaa caagaagcaa 18180
 attactatga atttttgtta acagtcattt tccttaacta ataagtttta aattttcatt 18240
 aatatgtttt gatcatattt tttccatgcc ccaacacctc caaaatctcc ccaactcattc 18300
 agttctttct ctatctcaa aaatgaaaaa tccaagcaaa caaccattag acaaaaaata 18360
 acaaaacaaa acaaagcaaa gcaaaaataaa agcacacggg ctggagagat ggctcagagg 18420
 ttaagagcac cgactgctct tccagaggtc ctgagttcaa ttcccagcaa ccacatgggtg 18480

ES 2 773 891 T3

gctcacaacc atctgtaatg agatctggtg ccctcttctg gtgtacagat atacatggaa 18540
 gcagaatggt gtatacataa taaataaata aaatctaaaa aaaaaaaga aaaaagcaca 18600
 caaaaaaacc agagagtgtg tattgagttg gttaacccct actcctctgg agtgtgattg 18660
 atacagccag tgccgctatt ggagaacact gattgtccct gtccttacag gtatcaattg 18720
 tgtgtagctc cttggttagg aatggggcct tgtgtgcaact tcccctttca gctttgtaaa 18780
 ggggtgccga ttgaagtctg tatcttctgg gagagcataa aatcaaaaa agataaatgg 18840
 actccagtga aaaaggagca agcggcacct atctttaagg tagagaggca gaggagtgtg 18900
 gtgtggcctg tcacaaacac ccaattccca atcagotggc gtctaccagg ctgctttcac 18960
 ttagatgaac cctgacctcc atgtctcctt aacattgcc a ttgtttaact gttagtgagt 19020
 ctgccctctg ttcactgaaa gactttcaga aggtgggtgc gcctgccttt aatcctagca 19080
 ctcgggagtc agaagcaggt agatagagct ctgtgagttt gaggccaggc tggctctgcag 19140
 agttccagga caggctacag agtgaaacc agtctcacia acaccgctc caccacaaaa 19200
 aaaaaaggaa acaagataga gtgaacaaac ccagctacct agacatctat ctggtaaact 19260
 gactcatccc aatcctccct gccctcccaa agagcttggc tggctcactt ccccaaatgc 19320
 tcttcccctt taacatttaa ctagttcttg tctcttgat ggtttccttt taactgtatc 19380
 caccaccctt acctgactt ttgtcctggt tggtttttaa ttgtaaactt gacacacaaa 19440
 gtcacctggg aaaagggaa ctttaattgaa gaattgtctt agattggcct gtgggtgtat 19500
 ttatagggca ttgtcttgat tgccaattga ttcgggggtg ggagtgggag ggtaggggtg 19560
 ggggtgggagc agcccaactat gggactcact tcccttaggc agatggctat attagaaagg 19620
 tagctgagcc taagccagcg ggtgagccga gccagcaagt agcattctc tatggtttct 19680
 ttctttcttt ttctttttct ttttctttt ctttttcttt ttctttttct ctttctttc 19740
 ttttcttttt tttttttct tcccagaca gggtttcttt gtgtagcttt ggagcctatc 19800
 ctggcactcg ctctggagac caggctggcc tcaaactcac agagatctc ctgcctctgc 19860
 ctcccagtg ctgggattaa aggcattgct caccaacgcc cagctcttct gtggtttctg 19920
 cttcagattt ctgctttgag ttctgtctg acttccctca ataattggtt gtaacctagg 19980
 agtgaagac aatgaacc tttcatcccc aagtagctat ggatttagag tggtttatca 20040
 cagccacaga gtgaaaccag aacaactttc tagtagcctc ttgttctact ccagctgctc 20100
 ctctgactat tcctaaaagg tagttgggct cagggaaacca catcccgaga gattcagccc 20160
 atatgaaaat agctccattg tgttgaagaa atgtgaccct ccaggatttc aggcatcagg 20220
 attccatggt gaaaatgaaa acaattattt tcctctctct caagattcct ttagtcacct 20280
 tcccttacc cagttcctg ctttcttct aaacaaatgt tcagggaggt tcaacaaac 20340
 agctgtgaag agcagcatcc cataccccca ccttcogacc caacacttgc cagtgtata 20400

ES 2 773 891 T3

agtagactgg gatcatccct ggacactgtg ttaaattacc catgaccaac cttctagcaa 20460
 gctctccttt tcaggatttt gttgtttgtt tgggtttgtt tgtttgtagac ttgatctcat 20520
 gtaagctgac ctggaatttg cttaatagcc aaggatagac ttacaacctg tgatgctcca 20580
 gcctctgact cctgagtacc agggattaca catgtgtggc atcacaatga aagattttag 20640
 tttgctgaga gaaaaagttt ttaaagattt tagttcacag agagaataag tttcccacag 20700
 gccttggtcc aggacaagga agttggtccc aacccgaggg cagacaaaca atcctttttg 20760
 ggtcacacct ggctggccaa cagacaataa aggacttctc agggtagatt ctatggttga 20820
 ccactctaac atgagatcat actttgtaat caatcacttt gtgccccttg cctgtatgct 20880
 gatctgcggt tttttacagg ctctatata aggagtctgt aacccttgct ggggtgtgca 20940
 gcttccccga tattgctgac accogaatga gcattcgctc aataaacctt cttgcttttg 21000
 cagctcttgg tctggtttct gagtcttggg gcctccttgg gatcctgaga cccttaaggg 21060
 tctgggggtc tttcaacct taactttcct gtttttaagt aggaagatct gaaatcccag 21120
 attcctgact ccattgcaca ttttctgtat tagaggctgt agctctgtat agtgggttgt 21180
 gtggcttaca catgctctga gctggagatt ctagggacac ttagggtaaa gtggagtgtc 21240
 agcccccttc cctgctagac tgaggccttt ctgttctttc ctaactggga ggctgtatag 21300
 cacccaatgt gttcattaaa ctccatattg tagcactgca tggaatctga cacacacaca 21360
 cacacacaca caccctctac caccaccatc atcagcacca cccccatcag caccaccctc 21420
 atccccccac cccccacct gccccnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnc 21480
 aactggaggg tagcattagc acccagatgc cattaatgtg ccaaattttt gcttgcttgc 21540
 ttgcttgttt gttccagcat ccttagtgaa tgctcctgcc ctcttggtta aagatggctt 21600
 tggcatctct tggcatcttt cttgtattct aggcctgaaa tagggatgaa tggatgaagg 21660
 caaggagctc aagtgtcact taccacctgc acttgtccct ttaaggggtt tccttagaag 21720
 cagtctacat ttcattagcc agagctttgt cacctggcta cttgtgaagg aggtggtgaa 21780
 gaagccttac ctttgactct gccacttggg gccaaagttag gattctctcc ctggaaagga 21840
 aatggaagat taataccttg ttggttgtaa gacctagccc attatgcgcc atgaggaaag 21900
 agagacaaca gtgggtcact gattgatcag ggttacagga caaggagcct tgtttctcct 21960
 aacagctctg agcggagaca gaagtggagt atataggcat aaaattcaca aacatttgc 22020
 gccacgttac aggtacattt tttcaccagt cagaaatcaa agattagggg ctttgcttgt 22080
 gtgttccatc actgtcaact gacatacacg gcaagccttt tagtccaacc aatcagaatc 22140
 atttgctcct tctgttgtaa ggagcagcca taatgattct aaagaactaa caatgcataa 22200
 tgactatfff tgtagtttag ggatgaggtg tgtcagccat tggacagttc tcagctcccc 22260

ES 2 773 891 T3

tagggcttgg gaacttgaac tttatttcat cctgcatgta atggagtctg aagtcaaaat 22320
ggcagtactt aggtcaaggt gctcgtgcct gctgccttca aggtggtttc ccattcccac 22380
cataccagag acttcctact gcatctccag tcaaggacac aaacactttt aagtctgac 22440
tggtgattca atctatatag ttaccagcat agaggctaag agtcacactg gcttgcaggg 22500
gacttctcta gcatatgtga agccccgttt gaatcctaaa cacaagagtc taagctttgg 22560
agtcagagac aagcatgttc aaatctgtac gtcaccaccc tatagacata gacaagtccc 22620
ttgggctcag ttttttcaact acagagagta attggtattt cagattccta gggttgtggt 22680
aattaaatag ttgaaagata tagcccatgg aacataaaaa aaactcaaaa ccaggcacag 22740
tggcacatgt cttaatttc agcactcaag agacagagggc aagtggatct ctgtgagttt 22800
gaggccagggc tggctctatat agagagttcc aggtctacac agagaaacag gctcaaaacc 22860
aaagcaaaag caaaacctca actaatgttc ataaaattat gaaattgctg gtaccagtga 22920
catgactcat tggtaaagac acttgctagc aagtttaatg atctgagttt tatctccggg 22980
atctacaatg tagaagaaga aaaacaactc tcaagagttg tcctctgatt tccacttatg 23040
caaaaatagca tgggaacaca cttaagcagg taggtaggta ggtagataga tagatagata 23100
gatagataga tagatagata atagacataa ttaagaacgt tcagttgcag cacagttcat 23160
actgaactgc atttggacac ctctgtgaaa agtcaggagc tctcctgtcc tcttgggtgac 23220
atttaaacat tgaaggcaac tattttaact gtcagttata taaaaatcca ctggccttgt 23280
aaaattttaa aacataacag agggaggctaa agtcctgttt aacaaccctc tccttttacc 23340
atcccaggaa gccaaaattg ttcacaattt gttctcttcc ctcaaggcctt ccatatttca 23400
aataccacat aaaacaccta tggaaaaaca tgaggтата aaaatgtcac ttggaaatcc 23460
ttcttcaaac aagcttgttc tttctttttt cttttatgta cagtgaatgg aatccaggac 23520
ctttgcagat gctaggcgag tcctttacct cattcctctt tcgatttaaa actttttctt 23580
gttttgtgga gacagggttt ctctgtgtag ccatagatgt cctagaacta gctctgtaga 23640
ctaggctggt ctcaaatca gaagccagtc tgcctctgcc tcgggagcgc taggattaaa 23700
ggtgtgggca gagtgttagg atgaaaggta tgcacaccac cactcctggt tgattttaaa 23760
aagatgcttt ttaaaaaaaaa tgatgtgtag gtagtggggg gagagacggt ttcatgccta 23820
agagcactga cagctcttct agaggactca ggttcaattc ccagcaccca catggcagct 23880
cataaccatc tgtaaccccg gtcccaggga atccaacacc ctcttctggt ctctgtgaat 23940
gacagatatg catgggatat acaaacatat acgcagacaa aacactgtat acattaaata 24000
agtacaaatt taaaatatgt gtaggcatgt atgtctgcat gtgggtatgt gtacactgaa 24060
tgcaagttca cttggaggcc agagatatat agatccoctg gagttgcagt tacagatact 24120
tgcgagctgc tgtgagtgtg ctgggaacca aatcctctgg aacagcagca agtgctctca 24180

ES 2 773 891 T3

cctgctgagc catttcttca cccgcttctt tctacttttt attttgagac aaggtcttac 24240
taagttatat attcacttgg ggcttgaatt cattttgtca gcaggcagac cataaacttg 24300
ccttcctctt gcctcgggct cctgagtagc tgagacttca ccatgaggtc tggctttgat 24360
tacatTTTTtC tttgttttct ttttgggggt ggggctgatc atgaactcta aatagccaag 24420
gattgatagt gaagtcCaga ttccccacc tatcaccggg tgggaattaca ggtgtgcact 24480
accacaccca atttggtttg atTTTTTTTT tTTTTtCag gacaagctct ccttttatag 24540
ctctgactgg gttggaattt actatgtaga ctaggctagt gtcaaaatca cagagatctt 24600
cctgtccctg cttcctgagt actgggatta aaggcatgta ccaccacacc ttcgggtgtg 24660
gtgatgcaca gctttaatcc cagcaactcag gcaggcgaat ctctctgagt ttgaggctag 24720
cctagtcttc agagtgagtt ccagaacagc caaggctaca cagagacact ttgtttcgaa 24780
aaacaaaca aaacaaaaga ggctagcctg aaactcctga ttotaccagc acctccaag 24840
ggctgggatg acaggttgtg gccccatgct ctctgccggg gcctctctt tctttcttct 24900
gtttgaggta gaggcttact aggttggctg ggtgagttgt gaactcactc tgcagcccac 24960
acaggaactg atcttgtgat cctcctgcct cagtctccct agcagctagg attgcaggcc 25020
tgcaccatca ggcccatcgt acaactgtttt ctgagtttga aaattgcctc tgttgttgac 25080
talaaggcal gctctcctcc taacattgtc ctlggtgcct ctgccaccct ttgggaclag 25140
agagaacaga tcttattcct atttcacatg ctgtgccaac ccagtaacaa actcagatcc 25200
ctgcttccgc cccaccacc cccatctaatt tgttcagtgt ttctgtgaag ataaacacga 25260
tcactcttgt gaaagccact taagttcctt tcaaggttgg gatataagtt agagtgatag 25320
cttgttccca ggggtgggag agcatgtgaa ttcccctctc gctcaagtag gctatactaa 25380
tttcatTTta gatatttctg aggcaaagtc tcatgtctgc catccacctg ccttagcttc 25440
tcaagtgctt ggattacagg catgagctac aatatctggc ttagtttcaa ggttgtgaaa 25500
attatactgt gttctgatga cctgagttca attccctgga cctgggtgat ggacggagag 25560
gacagacccc tgcagattgt cctttgacct ccctgtcact atgtgaacac tctgttacac 25620
acacacacac acacacacac acacacacta aatgaatgta ataaaatata aaaaggtgtt 25680
cactagttaa taagacatga gagaaaaagc ttaccatccc taatcaatgg ggaagcattg 25740
aatataagtg actgtggtca tggaaagcag tatagaggtt cctcaataaa ctggaatata 25800
gcagcatata cttgtaagcc tcccacaaca ggagaaaggt aaagaggggc ggccactctg 25860
gaatattatt aatatcctgt ttcataaaca agtaaataga acaaacctc caacaacaag 25920
aaccgggtgtg ctggcacaca cctgcaatcc cagcatttgg gacttggagg cagcacaatt 25980
gaagttcgtt cttggtcatc ctCagctatg tatgaaatct gaagcctgcc tggcctacag 26040

ES 2 773 891 T3

gagaccctgt ctcaaaaaaa taaactaaat agattaaaaat gaaaattaga agcaggtagt 26100
gtggaagttg aataagaata gccgccatgg gctcatgtat ttgaatgttt agtggcacia 26160
cttgagttag ttaggaggtg tggcctggtg gagttgtgtg tcaactgggag tgagctttgg 26220
gattttagaa gcccaagcca ggcccagggg cttgctctct tctgctgatc tgaggaactg 26280
gatgtagaac gcttagctac ttcttcagca ccatgtctgc ctgcatgctg ccatgttccc 26340
tgtcaaaatg ataatggact gaccctctga aacttggctc cttttggctg aggagttagc 26400
aaggaagag gtggctgtgg cttgctcttg tttctctctc tctgatcttt catcattttc 26460
tcccgatctt ggctgtgggt ttttattatt aagagtaatt agaactcatg ttacagtggg 26520
acatgcatgc cacagacca gtgtggatgc cagaggacia catgtgtaaa tttttctttt 26580
ccttgtagtg gcgtccaggc tagtttcaga cttgtgggct tctgcttcag cctcccaaag 26640
gtggggacca caggcttata tacctacact cacctcttta ttcccagtgg atgtgtgtgt 26700
gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtttgtgt gttttacaca gacctgtacc 26760
acattcattt ggtaactttt ttttctgca ttttgttttt aggtagggtc tcaactatgta 26820
accctgactg tcttgaaca tgctatttag attagactga cctgctggtc cctaccttcc 26880
gagtgtggg attaaagggt tgtactacca tacctgggtg ttagtttgtc ttttgagact 26940
gggtctcttg tagcccaggc tggctctgaa ctctgggttt tccagactct accttccaaa 27000
tattgatatt gcaggtggtc actaccatgt gtggaattta tttttgagca gtgttctgtg 27060
ggtggatgat aaggctcatg ctatggtaaa attgtttcta ataagatga atagcttcat 27120
gtgtgtatgc atctatcagg tttgttcaac ctgaagtga ggcctaatat ttggatttat 27180
ttagccagtg atagctatga attgagcca gaaaaaatca taaacttgac taaaacatct 27240
taagaatttt gtaacttctt ttgtaactca actgtattgt ttctgagcat gaatgttga 27300
aatgacaatg tcagctgcca tgtcaaaagg ttgaacatta cttggcagtg gtggcacaca 27360
cctttaactc caacactcag gaggcagagg caggcagatc tctgagttag aggccagcct 27420
ggtccacata gggagttcca caccagctaa ggtgacagag tgagaccttg tctaattttt 27480
ttttaagggt ggacatgtat aattccagag aataattttt cactaatcgg aaaagaggca 27540
gtttcaactt ggagttcaca agatttaate tttctttgaa gatttattta tttttagtta 27600
tgtgtgtgta tatatgtatg tatgtatgta tgtattgggtg tgttaaaccc ctggggctgg 27660
aattacaggt ggttgtgaac ctgatgttgt aataagctcc cagaccgtag cacaaatgac 27720
tctatgaaga aagtaccatt caggctgtaa aatccacata gacagcacca cctggaaaaa 27780
ctaaaacaaa aatccaatcc atcaactcc acagatctgg gaaagtatct aatgcacta 27840
accttgattt ttggcttctg tagttctgct tctggctaac tattcttgtt aactgaagta 27900
tgtgaaccca caacatgggt tttgtgctta aaagttctct gttctacaga atgaattcca 27960

ES 2 773 891 T3

ggacagccag agctgcatgg agaaaatctg cctcaaaaaca aaacaaacaa ataaaaacct 28020
 tgagaaaggc tcagggctat actggtatcc catacactca gtgtagtcgc caactgtcaa 28080
 agactttttg ttgacttaaa cccatttcta agcagtattc tcttatggat accccttaca 28140
 agtgggtgct gggacttgaa ctcaggtcct ctggaaaagc agaggatttc tcacctgctg 28200
 agcacctctc caggcccata agatctatct taagacaaga cctgagcagc cttatggaga 28260
 tggcagtctg ggaaccact ggtgcgctt ttcttctgct ggtcacaac tgctgtggga 28320
 atttccatct gaagttoctg cctcttctca cattccatga tatgagaaag ctatcaatgt 28380
 tctaaatctg tttgctttct gctttgcaag acctttctct ttcttaggtc accctccaag 28440
 agttcttgac ctcagccccg actggtgtct tgggatgggt gactgggttc tgggggcttc 28500
 cctgtgcctt ggaatatggt aaaagagcat ctcaggtatt cactcagtag atgctagtag 28560
 cactccctcc ctccatttct gtctacagat gttgctagct ggcccctatg aggtagtctt 28620
 tgcccctttg ttattgctgc agactcagaa aaaagaggaa atatagaact cctcgtggtc 28680
 ttctactcaa tatccaagca agggggaaca actgagcacc catacactgc tgttttggct 28740
 tctcaattgc ttgcttgtag atcaccaaga agctttcatt ggtcagtgtg aacaagatct 28800
 gggagttgat ggtagagcag ttggatgagt gactctgtct ttcacctttg ttgagtcatt 28860
 tgggtgtgtc acattgtggg tccctgctc gcttccatt aaatgtcaag gtgaacttta 28920
 tgaggttgaa acttttatat gtagtgcaac tgtactcctt cctctctatc tcttcttca 28980
 ttttcttcc ttcaccttct cttcctttaa aaaaagaaaa actttaaaaa atgtgaatct 29040
 gatgtatccc aggatggcct caaactggtt gctttctcag aagatgacct tgaactttca 29100
 atcctcctgc ctccaacctc caaatgctgg gcttacagga attcatcacc atgcctgggt 29160
 ttctctctc ctggtgagtg aatccagggc ttcattgctt ccaggcaagt gttctgctga 29220
 ctgagttaca tgcttagcct gtatccacat cttgactgag taatttctgc accaaaactt 29280
 taggtttcat ctcagtgact ctgccaatgt gtttccattt tagagtgacg actggcctta 29340
 gaggagagtg taagagaaat agagtctctt tccttggctt gctttttaa ttttaatttc 29400
 ttttagaca tcttatattt attcatgcat gtgtgtgtat aactagcaga actcagctgt 29460
 ctctttctac cactcaggtc accaggcttg gtggcagga ctcttacctg ccttcagaca 29520
 ggctctgcc tccttttggg gaaactggtt tgcagaagga agagacagca cagctcagaa 29580
 gacagccgtg ctttcagatg cctgagaatc ctgccaagga cactgctgca ttctcctatt 29640
 cttttgtaag ggtcccatct ctgctgagct aaactgggct ttctcagccc ttctcctctg 29700
 acagtatttt aaaacctac ctaaaggggg atggagagat ggctcagcaa ttaggagcat 29760
 atcctactct tccggagacc cctacttctg ttcccagcac caatgctggt caatttacia 29820

ES 2 773 891 T3

ctgtaactct gctccaggtc atcggatgct gctatcctcc tcaggcaact tcaactcatgt 29880
 gcacatacac atacttaaaa acaaaaataag tctttaaaaa tcacctaaga aatataaagg 29940
 cacatatcat aattcagcct gctgtgacgt atagctatag tcccagaatt ctgaaggcag 30000
 aggcaagagg atcacctcaa gcttggggcc agcgtgggtct acagtgagac cctggagact 30060
 ttaatctcaa aatatgtaac aaaacaaata tgtaaataga catatatcac aatttatatt 30120
 taagtaaaat ggggggcatt ggagagatag ctttgtgggtt aagagcatgt actgttcttg 30180
 tcaaggaccc aagtttgatt cccagtgtct aactggttg gtctccaacc caattccaag 30240
 agatctgctg ccttctctc ctctctactg gaactgcatt catgtgcaa tgtccatatg 30300
 cacacacata cccacatgca tacacacaaa cacatacata ctcatcttgc ctgacatcgt 30360
 ggtaaagtgg gaagacttgt tgccctatta ctgtgtcttc atttgctat gagcaccatg 30420
 ttggcatgaa ctcatcatt aatatctttc ctgtacaact cccaataac caagatgaca 30480
 cttggcacac attaattgct aagtataatg aaaatttagt ttaaattagc taaataattt 30540
 aaagttcccc ctcaagcctc atgcctgatt taaagtagta ctattaatg ctgggcctgg 30600
 tggcatacat ttctaattct aacacttagg aggctgaggc aggaggatgg ccaattcaag 30660
 gccagcttag ccagcttagt aagaccttgt ctccaagcaa attacagcaa agtctgagat 30720
 atagttcagt aattaggggtg tttgtctacc atgtgtgaag acctgagttc agtttctaac 30780
 aacaaaacaa aactaaacaa accagaacct agaggttacc atttattttt ttatttttat 30840
 ttttttttgg agtttatgcc tttggattat ccattctatg tccagacatc agtactgcca 30900
 tgttacagtc aataaaagtc ttccttcac acccttaac ttatcaccac taaagtctct 30960
 acttgacaga catgccatac ataattatag ctgttacctt ctatcataaa gtagacattt 31020
 tattttattt gtgtattcat tttcatttat tttgttgggtg ttgttgtttt atgagacaga 31080
 gtttctctgt gcagccctgg ttatcctgga actcactctg cagaccaggc tggcttcaaa 31140
 cacacagaga tccacctgcc tctgcctcct gagtgctaag attaaaggag tgtgctgcca 31200
 tcttcccagc aacattctaa attatttttt gtttatgttt tgaaatggtc taatgtagct 31260
 gaggtgggccc tcaagcttgt tatatagctg gggaaccttg aacttgtgtt ctccctacct 31320
 ctagaactct ggagtgctgg aattacaggt atgaaccatc acattccagt tttaatcaaa 31380
 tccagacttc atgggtacta ggaaagcact ctacaaatta aacttcaccc ctagtccata 31440
 tatatatatg tgtgtgtgtg tccatgtatg tatgcctaca tgattttatg tgtgccacat 31500
 gtgtgcagggt gctcttggag gtcagagggt gtcaaatccc ctggcacctg agttataggt 31560
 ggttgtgagc cacctgatgt ggattctggg aactgaactt tggctctctg caggagaagt 31620
 cactgttccct ctgagtgaac gtttctactt tttaatatac tcccattcg aattagaaag 31680
 tagaagctct cggagggtga gaccttacct aaagtcaccc aactagtaag aaaactaaaa 31740

ES 2 773 891 T3

tatcaacttg gttttctgag ttttaaatat tttttcccaa tgtgtaatta cacaggagaa 31800
 ttaatgggga cacttcaagg taaaacagaa gcttttagaca tagcaaggca tggtaggacaca 31860
 catcccattg agaggcagga ggatcaggag gccagctttg gctgcatact taagaggcat 31920
 ccagggctac atgaggcgct acctaaaaaa attaaattag gcagggcggt ggtggcgcac 31980
 gcctttaatc ccagcactcg ggaggcagag gcaggcggat ctctgtgagt tcaaggctag 32040
 cctggtcttc agagcgagtg ccaggatagg ctccaaagct acacagagaa accctgtctt 32100
 gaaaaaccaa aaaagcactg gtcattgtca ttttctttcc taacagggca ctggaacctt 32160
 gatgttggtt ggctcctaga tttcttctcc acagcagaga gttcttgccc tgtagagacc 32220
 agaaggatgc tctggagagt cagtatatag caaagcaggg tcatctggag tagtaaaaac 32280
 cctctggcac agtcagacct catttctct tgtcctgtgc tcgtggctct agcattatgc 32340
 aaggagaggc gcaaacagca aacaatttg aaggcctagc acttgagcaa ctctttgtag 32400
 ctctctcttc tctactcttt tgcctctggc ttctactgga acaggtgaet ttccattgca 32460
 ttgcattctc caaactcaga tgattttgag aatgtggcac tactaaaagt cacatggaca 32520
 tacaaggtac aactagaact atccccggaa acagtgatac acgatctagt ttgaggcctt 32580
 gagccatagc ttgtcagaag ctcagaaatg attgagtctc tgggagccct cacctcagca 32640
 tcctgcttg caaaaggctt cttgaagtag taaaaactgc tgggaccttg tctaggctgg 32700
 gtaaccttgc ataattactc aaccttactg agctcagtcc cctcctctat aaaataagtg 32760
 caacagtatt taccttagtg gcccacctga aaacatcaca gctgccatag ctagctcttg 32820
 gctttgttc tatctctctc tccccctact ttctcttccc tcctccctc cctccctcat 32880
 ttttctttat tcctttcttt gtattttttt cttttttctt cctcacacct ctctttatc 32940
 cccacctcc tctctctctc tccttccca ctctcttctc ttctatggca ggatatcatg 33000
 tatcctagct atacttgaat tcactatata gctgaagagg agcttcagc ccttttgctt 33060
 ctgcctcca agtgctgaga ttataggtgt ccacctccac gtctacttat gctttgctaa 33120
 ggatcaaacc agggctttgt atgtgcatgc taggcaagag ccaactacat cgccagacct 33180
 atataatacc cctttctcag cgaaactggg gttgctgatg gctggtgttg ggggaaggca 33240
 ctaaataatt agcagaagta taggaaaact ctagaagtct agagatcctc aaagtaagtt 33300
 tggagagcct tggccttttc ttagttgaaa gtcatgggtc ctactcaett tgactgctca 33360
 aggaatatcc attcaccacc tggaaataag aaaggaggga gaaccagcta gggatgtgac 33420
 ttagtagtag agcacttgtc tagcatgagc gtggctctgg gttcaagctc cagtacaaag 33480
 gctgggtggg ggggtggaga aaggcttctt tccatggcg ttctagagat ggcggggaga 33540
 aaccaccaat ccacatctat ctacaacagt tcaagtagaa ctaatcttgg tggtaggctt 33600

ES 2 773 891 T3

atagtagtcc taatcccac tcagggatgc ttctctttgc aattgatata aaacacatta 33660
 cagaaaacca cagtgaatca aaatgcagag ttgtgggtgcc tagttccaat ggatgcatct 33720
 acagtacaac tcccatgcct aaggctcagg gatcattgtg gaagacaaaag atcctcccag 33780
 gagatcaggg agtttgctgt ctccatagaa tttcagaaaa tacatctgta aaggctcacc 33840
 aacgtgaatt cctaaacatg agctgaacaa ggatgacaat agacatgcta acaaggatgg 33900
 gaaaaagccc ttgaagcctc agacctacac aaagagccgc agttgattaa ggaatgctga 33960
 ttgtggggaga aaccatcttc ccaaattggt atctaatacc acatagtcag ccctgaaaac 34020
 acacatgcaa ataagattat acaaaacaag ggggttgtag atatgtattt aggaatatat 34080
 atatatatat atatatatat atatatatat gtaacaataa ttaatagaaa aagagaccat 34140
 gaatttgaaa aagaacaagg aggggtacat ggaagggttt aggatgcttt gaccctttaa 34200
 tatagtttct tgtgttggg tgaccccaat cataaaatta tttttgttgc tagttcacia 34260
 ctgtaatttt gctgctgta tgaattgtaa agtaaatacc tatggttttt gatgatctta 34320
 ggcaatccct gttaaactgt cattcagtc ccaaaggggt caagaccac aggttgagaa 34380
 ctgctgattt agagagagga aaggaaggg ggggtgaaat gctgtaatta taattccaaa 34440
 aaaaaatttt taaaatttc ttaaaggaac tgaagaaaag agctgaacat tctaagctta 34500
 aggggggaaa ggttctggaa tgttacattt ttctggtttc cttagtctca gcaacaggct 34560
 cccagccttc tgtttgaca gtggtttaca ggcatgtgag ctccaggaac actcttccaa 34620
 gtgaatcaga cttcaggaga agacattcag ttcaggccc tggggaaagt aaggacagaa 34680
 ctccattcct gagaattacc aggtttgctc agaagataaa actggtgagc ccaatggctg 34740
 tgtgcacaac cctgacctca gtgtctagga tagctggact ctactgctga gaagatagtc 34800
 agagggccat ccttccctg aggctaactc gtgaatcaag taaactacag tcaggaaggg 34860
 agctggagat gggggcccag caaacaggtc ccccttaaag cccagcacat aggtggggaa 34920
 cccaacctcc cattttgtct tcaacctacc accaggcctt taccaaggcc cgaggttgcc 34980
 actatthtca gcttgccagg ctctttgcag ttttaggggg atgaggagga gatgctctga 35040
 ggtgctggga ggcacatggc ggggtctatt tatggcttgg gctgaactcc gatgtcctag 35100
 aaagagtgtt tctgacactt tctgccttct gggaatcagg agactcatga caaacactgc 35160
 ctggcagtggt ttctttcttg ttcacagcaa gaagtgtgca gtccatggca cgaaagaggc 35220
 ctgagcaggg caagatggac acgatgacat cactgaagga gcttcccagg ggctgtcttg 35280
 actgcttcat taactcattc atgcagttta ttcagcagct atgcctgtca gacccattc 35340
 tgtctgcaca agacacatgg caacaaagga gacttactat tcccatcttc atgggtttta 35400
 tgttctggca agaggaagat agtaataatt tttaaaagt aaccagtctt gagagcatga 35460
 taaatatggt tgataacaat atgctatatt ttaaagttg tgagatagta tactttaagt 35520

ES 2 773 891 T3

gttctcaaaa caaaatgatg aatatgggtg atataacatg ttaattggtt taatttagcc 35580
 atgcctttgt gaacatactg tatcgtgat cataattgtg catgacttta tttatgagct 35640
 aaataaatga atggaaaaaa aagtaaccag tcttgatgct tacctgccat cctggaagga 35700
 aatggaaata ggatctgccg ccgcagcatt gccotatgct cttatttctt ctcttgaaga 35760
 ggtaggggtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgttactaga gactgagcta 35820
 ccggcctcac acattctagg caaatgctct actttatatt aaacacttta taaaacatta 35880
 agcctttcag ggtcagcaag gtagctcaga gagtccgggc atttgctacc aagcctgaca 35940
 acctgagttc gattgatgat cccccagact cacgtgatag gaggaagctg acacctgtgg 36000
 gttgtcctct gactataggc atgcacacac atcccatgaa taaatattta tacattttca 36060
 aatcatactt attttacaat gatttttatt tgtttgcctg tctttctgtc tgtgtagaga 36120
 caaggtttca tgcagctcag gttggcctca aactcactct gtggcaagga tgccttaact 36180
 tcaggtcttc caggtccagg taacaaaatg ttcaggagga acctggtacc tcatcataac 36240
 cggttctaga tggcttccc agggctgctg taagaaagtg ctacacgacg agttatttca 36300
 aacattctca cagttctggg gattagaagt ttgaaactaa ggtgctgaag agattagttc 36360
 cttctggaag ctcagaagag ccatctggtc catacttttc tccaggtttc tcttagtttt 36420
 tggcaatcct tggaatccct tggttttag atgcagcttc caaagctcaa gatctctctc 36480
 caatgctgtg tggcatttcc ccgtgtttat gagtgtctaa atggctttta aaaacatttt 36540
 tgagatgtga aattctggct gaccagaat atataaacca ggctgacctt tgtctcccag 36600
 agatctccct gcctctgctt cccaaaacctt ttgattaaag gtgtgtgtca agtgcccaga 36660
 ccaaagccc ttcttgaag gacaacggtc atattggatt tagtgtctaa gtgagtcccc 36720
 tatgaactca tctcgaactc agtttgcata gaacactgta ccatgcaaaa taaatgacac 36780
 agagactgat attggggttc acacttcaag ctgaaggtca gaaaagcaaa gcattgggcc 36840
 actagctctt accactacct caggctgaac gggctgatcc tgetgectct cctcagcatg 36900
 gctggagaat atcttcatat cctcattgtg gctggaaaat gaatgcctga tatggagaac 36960
 ttgtctctgt tttatataac tccctaatac tgggattaaa gatgtgtgat ccaggtgct 37020
 gagatcatct ttgtgtgagc tgtttctctt taggactgga tcaattttgt gtagatctgg 37080
 atggctttgg gctcactgag atctatctac ctcttaatcc ctggctctag gattaaaggt 37140
 atgtaccacc acatcctagc ttctggctgc tgggattaaa ggtgtatgcc tggcttcgat 37200
 ggcttgtggc tgactttgct ttctgaatcc gcaggcaagc ttaaaaaaat cataaataat 37260
 atatcaccat agaccacact tccaaatagg cttccattta gaggcgccag tgggtgataa 37320
 tgtaggcggg tttactcagt tttgtgcaga tggctggcgt cctgtctggt gagttcagat 37380

ES 2 773 891 T3

tttttttttt ttttttttta agttcagaat cttaccacagc tcagcttttc aggctgcatt 37440
 cagtgtccgg cttttttctc accgtcttga ctteectgtcc tgcaccccat ttctcagcct 37500
 ggaccctgcc agtctatcag atagataaca taaacaaaat tgtactggat taatgggagc 37560
 tgtttggaca tttcctactt ttgccttttc accaatgatt tgcatactta agcctgcaac 37620
 tacagccccg atgcagtaag ctcaagtctct ggcaagcaaa ggtctctctg gggctctgtt 37680
 taagaaccag ctcaaggctgc tggctctgtt ggcagtggag gtatttccta taatgggatg 37740
 atgggatggg ttattcacac acatctcagt tactgggcta catggatcca aatcagccac 37800
 ccaaggggtt gcagtcacat gtgagtcact tagcacagag aaagaagcct ggaggaggag 37860
 gggctctccc agcttcagga gggttttcca ggatataggc ttctagtctc gttttggatc 37920
 aatttatcag ttttgattg ggtctaataa ctctttcctg agcctggact gggctcaaag 37980
 gcatgagtat gtgaggggaa tttactagaa ttcacctgta gtttctgtat cattcctaga 38040
 gaaggggaag tagagacact ggtgatggga aataaaaaaca aaacaaaacc taaatattgg 38100
 gagcacagag gtccctgttc cacagctctt gatagaagtc aggaatgtta tgtatgtaca 38160
 attgcccttg aaaaggaaag gatgtatgac ctgtttttct gtcccgaagg ctgggaactg 38220
 gggatgatta acagcctggt gatctgcatt atctgaaggg ctaggccata tcaagctccc 38280
 acagctagca ctgaaggaga atagggcctt acaaagggaa ttccctcttt ggatcgaacc 38340
 taggaacatc ttctgtttta ccgctctctc cttgtttcat ctgcaaaggg aggagcttgg 38400
 tagtgatgtt gaggcaggca ccacttgtat ttttctaagc cacagagact gtttccctac 38460
 cttacaaaca tccctgtgca tcaactgcagc tctgtctctt atggcagtgt ctcaagttagg 38520
 gcttctattg ctgcgactaa acaccatgac caaaaaagct cacacttcca tactcctggt 38580
 cattattgaa gaatgtcagg actggagcgc aaacagggca gggctctgga ggcaggagct 38640
 gatgcagagg tcatggagga aggctgctta ctggcttgcct ctccatggct tgctcagcct 38700
 gctttcttat agaaccagg accacctgcc cagggatgac accacctaca atgggctggg 38760
 cgctaatatg agggatcaaa gagatggagt tgtgggaggg acagaggggg agagcaatga 38820
 aagagataat cttgatagag ggagccgtta tggggtttagg gagaaacctg gtgctagaga 38880
 aattcccagg aatccacaag gaagacccca gctaagactc ctagcaataa tgaagaggat 38940
 gtctgaacgg gtcttcccct ttaatcagat tagtgactac cctaattgtc atcacagaac 39000
 ctacatccag taactgatgg aagcagatgc agtgatccac agccaagcac tgggctgagc 39060
 ttcgggagtt cagttgaaga gagaagggat catgtgagca aggggggtggg ggaagtcaag 39120
 atcatgatgg ggaaaaccac agagacagct gaccogagct agtgggagct catggactat 39180
 gaaacgccag acgttgtaga ctccctaagg aaggccttac cccctctgaa gagtggatgg 39240
 ggggtgggaa gtggggacgc tgggggacag gagaaagga gggaggggga actgggttgg 39300

ES 2 773 891 T3

tttgtaaaat gaaaaaatag atttttttta aataaaaaaa gaaagtgcct tacatctgga 39360
 tttcatggag gcattttctt aactgaagct ccttcctctc tggcgactct agtttgtgtc 39420
 aagttaacac agaaccagcc agtacaggca gcagaaatac cttgcagaaa tatcttagtt 39480
 caggagtcca cgggtggtctc agtcacttcc tcatgtgcc a cctgagtta acattcccca 39540
 aaacttgga cacaggccac cacatcatgg agccctggct taaagctcaa gttttatggt 39600
 attttctttt atcactgtct ataattccta aacatgctac aatggtgtga gcctcaccg 39660
 tctcctaggt ccatagtgac ttcttggcat taatagactg tgccccaaga gctctatggc 39720
 cagcaccacc acctgccatt cccctcccc tccatggtcc cagcctcact tcttcacttc 39780
 ctggtccttc cgagcccaat gtgcaaacc acagaatctg tctgcttatg taagtctcct 39840
 ggtcactgag tgggggtgact cagcaccaag gtggtgccct gcgatttccc agccccaggc 39900
 agcagaacaa ctgaaatgga aaacaagtcc cgtaaatagg gtccagctga gagcctccct 39960
 ttctcagga gtctggcaaa tctactctc ggggaactgc cctgggcagt ggaattctcc 40020
 agctccctgc tcatttcta gttcctctc cctcttctca cctttggctg aggatcagaa 40080
 aggttcccac tgaggctctgc tttgccctgg gcttctctt ttcagagtcc ctttttggga 40140
 atgaattttt tttgtctcct actttcaagt tcacatattg aagccattat tgccaagtg 40200
 atggtatcag aaggaggac ctttgggaga tgaatggatg gattccaaga ggttatgtgg 40260
 gcagagcacc catgatgggg ttggtgcctt cataggaaga agacacagta gaagggaaaag 40320
 agatgccgac tgaaaaacag gaagtctcct ggagtaggcc actcagccta tgacacgcca 40380
 gcactcagat ctgggacttc ccatctccca aatggtgata aacaaatgct gttgtccagg 40440
 ctgcacagtc tacggcattt tgttgcaagg gcttgacca accaggctca ggcaggaagt 40500
 gaatctagtg tgggaggatg tacagactgc cactcagtct ggacacaaac tgtcctcagg 40560
 gatcacctga gccacatcta cctaagaatg gctattcttt ccatttgta acatcaaatg 40620
 ccaagcccct actgtatgta ggctcttgct agcagtggat atgatgctat gtgagatggg 40680
 agcaatctc tctgcacaga actatacata gaactatgca tagaagacca acagggagac 40740
 atcagataac tattaactgt gatagctctg tgggagacaa acagaatgag ggaatggaca 40800
 atgactttga ggaaaaacta tgattgaaaa tactctatct ggctgggcgg tgggtggcga 40860
 tgcctttaat cccagcactt gggaggcaga ggcaggtaga tctctgtgag ttcgagacca 40920
 gcctggtcta taagagctag ttccaggaca gcctccaaag ccacagagaa accctgtctc 40980
 aaaaaaaca aaacaaacac acaaaaaaga aaatattctg tgaggtaaac aagcatctgg 41040
 aagggttggg agataatgca ggcaaaaatg cattagacag cacacagtac aacacagcaa 41100
 tcaaacttaa tataaacaca gcaaatgtca tctttggct ttgccccatt tctgatctg 41160

ES 2 773 891 T3

accataacag cctagtgtct ggaaagcaca ctaaagccat ttacgtcaca caggagttca 41220
atggtgagtt cagagggagg gggtagggg cagattagcg aggtacaagt tctggtcctt 41280
ttgatgaagt gttgatgtac ccatcgacac cacacaaata taccatcatg ctccatgtta 41340
gggtcagtga aggattgcat atgtgacggg ggcccactgg gctgagaaaag ccctattgct 41400
tagtgacatc tgtgataatg acatgcgagc cctattgctt agtgacatca ctcttctcat 41460
agtgtgggat ccaatgtggt tcttgtagac ttgtgataat gacatgcaaa caagtctatt 41520
gtgcgccag tcacacaaaa aatatattat gtgcagtcag gaacagtcca tagtacttga 41580
ttgggacagc acaagtctgt gttgctgggt cacacattaa tcattaccac tgttttagtg 41640
tgctcctata tatatatatt taaaaattac tataaaatga tacaccgtgc tgagcaatag 41700
cacctcttat accttgtggt tactggatgt actcaagcta ttttctcttg tgcttgattt 41760
atgtgtattt gtatttttga gagaacctca tctagtccat gctggcttca aacttgttat 41820
aaagctgagg atggcttga actcctgac cccagcctc tgcctcccaa atgatgagat 41880
tacaggcata tgctaccaa catgactttt atttattttt attacttagg tggatgggt 41940
ggtttgaatg agactgtccc ctttggctta tatatttga ggtggacctt tggaaaggtt 42000
taacaggtat gaccatagtg gaggcagtgt gtcagtaggg gaggtctttg gggaacccaa 42060
tactcaatca attccaagtt agggctgtct gtctgtctgt cccctgattg tgtcacaagg 42120
cagaaactct cagctactgc tctagttcta tgcctacca cctggttcca tggtccttgc 42180
catgatggtc atgtacttca accctttgga taggtggccc ccaaattaa tggtttcttt 42240
tataagttgc cttggtcatg gtgttttgtc atggcgataa gaaagtgact gagacaggtt 42300
tgttgctggt gttacaaggt ttagtccagg catctggcac cacctctggc ctgtgcttga 42360
ttcaatcatg ttaccttag aatagcagg ctaaaggaca tatacctgtg tacgtatatg 42420
tgtacgtata tattagctgt atagtctaag tgtgcacctg actctaatat ctaggtttgt 42480
gtaagtagac tccaccaagc tcactaagca atggtatcac agttttcaga tagtgttcag 42540
ogatgcttgg ctgagtgtta gttctttttt taatatttta tttatttatt atgtatacaa 42600
cattctgctt ccatgtatct ctgcacacca gaagaggaca ccaaactca taacggatgg 42660
ttttgagcca ccatgtggtt gctgggaatt gaactcagga cctctggaag agcagtcggt 42720
gctcttaacc tctgagccat ctctccagcc cctgagtgtt tttaaataca ggaaaaaagc 42780
ctgaggggaag ggagctcagg ctgaagggga ggagtcaaga cagtctgacc ccaaggcatt 42840
gtgggacgta aagagttctg ggacaagact gaggtctctt ccttctcaga gactgtgggc 42900
ttcagtttcc ttggtagccg gaagcaaagc taatccatgg cttaaaatat aatactcagt 42960
gtaaccttgt gttgtagaag tgacttgctt gtcttcttcc ataattctaa aacatcttta 43020
agagcaggat ccaggaaggg aaaaggagag attctcatct tcttcaaaag gcagctttcc 43080

ES 2 773 891 T3

ctaaagcatt ttctgatgaa atttaagttc taaaaccagc agtgggtataa tcccatcatg 43140
 aatggggatc tctgagtta aggccagcct ggtctacaga gcaagttcca ggacagccac 43200
 ggttacacaa agaaatcctg tcttaaaaca aaacaaaacc caaaacaaac ataaacaaaa 43260
 actatccaaa accaaccaac ccccccaact cagaaagaaa gaaagaaaga aatcaagaaa 43320
 gaactgccc a cgggtgttg gtggtgcaag cctttaatcc cagcactcgg gaggcagagg 43380
 caggcagatc tctgtgagtt tgaggccaac ctggtctcca gaaagagtgc caggataggc 43440
 tccaaagcta cacagagaaa ccctgtcttg aaaaaagaaa agaaagaact acccatgacc 43500
 aacagttcc atggccaggt agagaatgag gacgctgaaa gtcacacctt ctcagagtct 43560
 caaactgcac atctggcctc aaagtccaga aatgagtgca agaccattaa tgacagtctt 43620
 tgaaacaaa ccagacaaa gaacatttg ctctgatac atattctgag ggtcacatag 43680
 aaagaaagat ctgcctttgg ccacctcctt ttgaagtggg gaattttatt ttctctgca 43740
 tgaaacttc atgtaggtat ttgagaatac atacagacat gcaggtgcac atgcacggac 43800
 atgaacacac acatacacc cgggtaggca ggcaagaaag tgtgtggaat aacacttgaa 43860
 cttcccttc agaacagaag ccctctgaag tgtgacattc atgctggctg catggggctc 43920
 gatcagtact agtgagtgga ggtggagggg taggaaacat ggggatgata ataggttgtc 43980
 aggaaagtgg tgcccaggt agcacagagt agaaatttgt cccccaaat ccttttgaac 44040
 ccagttgatt tgaatgccgt gccctgcca ccaggcttc agagctaagt gacttatgtc 44100
 ttcaggtcag tgatgattac cacggttgca gtgctaacac agatgcttta tctaccagga 44160
 cagaaacaag aaagatgctc cttcccaggc cccttagcac tctctgggtg gggaggattg 44220
 ccccacctc caaaaataga atactgtttt ggtaaacagc cactttgagc ccatgaggat 44280
 atcttcatta gctatggaga caggttttag taagaaagca agatgagagg ctaaaaaacc 44340
 cttggggagc aggaactggg aagactgtgg taccttgttc ccagatccac cagaaacctt 44400
 gccaccagac gatgtgtcca ggccccacat atttcacaaa aagttggatc tgataacaat 44460
 gaggatggaa tcccgtctt aaggtgggtt tggggtggga agaggcggga taatgggtga 44520
 gaggtcggg ggggacaggt gagatgggt atggtgggga gaggtggaat ggggtgggt 44580
 gggttgagat ggagtatgt acagcgggga gggatagaat tgtctttcc ctgtaccaca 44640
 gagaagttg actgctacc ttggcaatta atcaattata gaaaatgcaa ctttgctttt 44700
 aaaatgtgtc tatttccaaa ggcttcttc cctcccctac ctaggagaa ggaaagaatg 44760
 gataatgcta ctgtagagga gggtagcatc actatagagg cctcagtatc tgccccaggg 44820
 agctgggaga gagttctatc acacaaacac agcccagatc acatactcaa caaaccac 44880
 aaaacaaaac aacaataatg aagatacaaa atctcattat gtagccagg ctagtctag 44940

ES 2 773 891 T3

atttctgttt tctttttttg tttttcgaga cagggtttct ctgtgtagct ttggagccta 45000
 tcctggcact tgctctgaag cccaggctgc cctcactcac agagatccgc ctgcctctgt 45060
 ctccagagtg ctgggattaa aggcgtgcac cactaatgcc tggctagtcc tagatttttt 45120
 tatcctcctg cctcaggctc ccaactggtg ggtttacttt tgggagtcca tttcttcca 45180
 gcatggattc tttgaattga aattcagatt atcaggtttc tgtagcaatc ccaccagccc 45240
 atttttttgt ctgacactgc ttgttttgag acacagtctc ccaactgctgt agcccaggct 45300
 gccctagatt ttctatgtag cccaggctgg ccttgaactc ccaggagtcc tctggcctct 45360
 cccttttgat tactggaact agaagaagtc actatgcttg acttgggaact aatattagaa 45420
 caaaatatat ttttcattga gattcaactt tgaaatcctg atgctcctgc ctcaactcagg 45480
 tcatcagggt tggcagcaag agcctttatc cactgagtca tattgggccc tgacctgctt 45540
 ttaaattttg cctttagggc tggagatgta gctcggctgg ttcagtgctt gcctggtacc 45600
 cacgaagccc tgggtttgat ctacaacaca gtataagcca ggctgatgg cgtatacatg 45660
 taatcctaac acttggggag caagaggag gccaaagcca tcctctgcta cttggtgagc 45720
 ttgaggccag cctgggatcc ttgagaccct gtttcaaac aataacaaca aacacagact 45780
 actaaaaaaaa attaataagg gccagactgg gtggtgtatt cctttaatcc aagcaatgag 45840
 gaggcagagg caggcaaagtg tctgtgagtc tggggacagc ctggtctact gagcagcagg 45900
 ccaactaagg ctacatagtg agactatctc aaaaaagca aaataacaat aacagacca 45960
 gttccccatc tcctattttg cctttacctc ctattccctg ctcaagcagg tattttttgt 46020
 tcctgcatct tggttcaactg atctgtaaac ttgtctgaat aagtaggtac agggttgttt 46080
 taaaattaga taatatattc aatgagaagg gctaccaagt gctcaaccaa tgtatgcata 46140
 tgtatgatg tatgatgta tgtatgtatt tatttttgtt ttgtttttca agataaggtt 46200
 tctctgtgta gttttagagt ctgtcctaga acttgctctg aagagcaggc tggctctgaa 46260
 ctcaaaaaga tccacctgtc tctgcctccc aagtctggg attaaaggca tgtgacacca 46320
 ccccaaaagc caatgttctt ataggcatct ttgatttttt ttctctttct ttgagtggag 46380
 tctgactaag tagccatac tagctctgca ttacaatct gaacacatgg ataagagtgg 46440
 tgaaaattat caagatcatg ttatgctatg cctcctgagt caccatgcc tgcttcagac 46500
 ttctttgtat taaagaactg tgtaaaaaa aaaaaaaga catttgaagg cacataatca 46560
 gaggaatttg tcagtgattt ttcacatact gtcttattttg tggccaagg t aagcctagag 46620
 agtatttctt aaaattaaaa atagtgggca gattttggag gcgatctgat atgaaaatcc 46680
 cttcccacc caggtagtca tgggctgact atcaaggata cattctgaga catatatcct 46740
 caagcagttt ctgccttacg caaatatcat aggtcatagc acaactgagac tatgtggcag 46800
 tctatgtgtc tatatacaca tgggtgtggc tattgttccc atggtcacia agaacaaaac 46860

ES 2 773 891 T3

aactttttca caaggcttta cccctagagg aagagctaca ggcaatcaat ggttgctgag 46920
aggagtatca gtcttctcca gggacttagc caatcccaag aggtcagcca cgcataggaa 46980
cgcttagcca cgcttgtata gaacatctca aacaacaacc acctcagtgt aaagcaagca 47040
cacaaggaac tgatgcaact aagagacaaa gggcccgggtg tgtgtggccc gtagctgtca 47100
tcccagcact tgagactaag gaaggaaggt tgagaatttg aggccagcat ggactccaca 47160
gaaagaccgt tttctttctc agaaaaaaga agcaaaaacc aagaacaagg tgtatgggaa 47220
tgctactgtc ttggcatatt gtttatagaa aactttttta tatataaaag gaatgcacta 47280
caaaaattat aaactactgt aatattaact gcatagatct ataacatggt catttattat 47340
tgagtatgat tatctatcta cccacgctgc aggtttagac agttgcaacta cagtagatct 47400
gtttgcagta gcatcattat tagacatttt ggacaaagcc aagtggtaat ggcacatgcc 47460
tttaatocca gcacttggga agcagaggta ggcggatctc tgtgagtacag agaccagcct 47520
ggtctacaaa gaactagttc caggagagtc tccaaggcca cagagaaacc ctgtctcgaa 47580
aaaccaaaag aaaaaaagaa aacaaaaaac taaaaataa ataaatttgg ggcaatatct 47640
tgtcctatga tgttactggg taatgggatt tcctcctctt gtattatttt ttctttgggg 47700
gttttactta ttatttactt gagacagagt ctcatttatg acaggctggc ctcaaacagg 47760
aatgaagcc aaggaagacc ttgaagacct aatccttctg tttcttcctc ctatatggtg 47820
agttaaaggc atacagtacc atgcccagtc tattcaactgc ccagggett c atgcatgcta 47880
gcaaagcacc aactgagctg catccccacc cctcctcctg gcttccatct ccttatgtag 47940
ctagaaatga gcctgtctgt ctcaaatact gggattatgg gtgtgtgcca ccacacctgg 48000
cttcctatta tagccttggtg ggatcactgt tgtttactga agcattgtga cacactgcag 48060
attgctggaa cagcgtctgc catcatcatg acacaacttc agagaaagag agagttccca 48120
accagccaca cacttaactc aatgcctgta gcccttatto tgtaagacg atttctgcc 48180
atcttactca aagaccctct ttaactcggg aggaacatct gttacactga aagtcctgcc 48240
tgttgctcca ctgacctcct tcacaaatta ttatatattg gagccaatc tgaaccagg 48300
ttttctgagt gacacatttt agtatttttt ttttctttct attttcttc atggaaagtc 48360
tcttgttact gttcacatga ccaaggatca ctgcatcatc ttccaaggcc aattttggat 48420
gtttcagcaa gggagactga agatcctgag tctcagtgtt gatctccttt agaatgtcct 48480
ctggagaagg tagtgacaac actgcaagga taataggtga ataaaggga gccagagtgt 48540
cctctgggat gtgcggcact tacatgaagg attcatttat aaattttaag ttatggagta 48600
taataataag actaaatatg tagtgtcgtat attttataac tatacatatg tatatagtaa 48660
atataaattt atatgtaatg tatttatagt aagtgtacat agaattgaac atatgttaca 48720

ES 2 773 891 T3

taaatggcag aaaggaatga ttctcaattg ctttttttct aattataatt tctattgctc 48780
 tttgtggatt tcacaccatg cattctgatc ccacttatct ccttgtctcc ttgcatttgc 48840
 cctctgccct tgcaacctca cccccaaatc aaagccaaat ttaaaaaaaaa aaccaaaatc 48900
 caaacaaaac agagacaaaa caaaaataaa agcaacaaca aaaaaaggag aatcttgtca 48960
 tggtagctgt agtgtggcct gttgaatcac acagtatacc ctttagtcca ttcatctttt 49020
 ctccaagtg ttcatgata caagtcacgg tctggctcga ggattctggt ttctgctata 49080
 ttactaataa tgggctctca ctggggctcc ccttggatat cctattgtcc tgtgttatgg 49140
 agagcctgct gttttggata tgtaggtttg tccccttcac atgctataac aattcataaa 49200
 ttcagtgaat gttgggtgg gccaaactcat agccctgggt ctgggcttgg gtggtattat 49260
 taaaccact gatggagaat aagaccacta ccataattta aaagccaaat tgaagcaagt 49320
 ttaattcaa tactgcccag gtggacaggc tctggctagg tccatctctg agtttccagg 49380
 aggtggccct gactcacggt ttacagtggc ttgagtattt tccataaggc ccaatcaggg 49440
 gcaagcatac atcctgatgt acctccagtc tatatccaat cgggggcaag tgtacatctt 49500
 gatgtatttc ctgcctgtga acctactgcc cacatgtgat caagcacatc cgggtgcagtt 49560
 ggggtcaaaca gacttgttta gggcaatgaa aaacacatgg ctttttatct cccataaaca 49620
 atagcctcca gcggttcagg gactatttgt ccttgggcaa ggaatttaca gatcctatag 49680
 gtgagtcagg gtcagcatcc tgctctcatg ccctcagggc tggtcactt gttacctccc 49740
 cgaccctctc tcaacagggt cagctctgag gtgctgocca ggtgggggtgc agggcctact 49800
 cttccgcatg ttgcagctgg tcagggttag ttctctcata tgccacaggt ggcaatgggt 49860
 gaagggggag ggcatgttc cctcatcaac gccattacat ggggggatgg ggtcagctct 49920
 catgccctta gggttggctc acctgcatcc ttgaccatag ggtcagctct agtatgctgc 49980
 tcaagtgagg cgcacaccta 50000

5 <210> 4
 <211> 34
 <212> ADN
 <213> Fago de enterobacterias P1

<400> 4
 ataacttctg atagcataca ttatcgaag ttat 34

10 <210> 5
 <211> 34
 <212> ADN
 <213> *Saccharomyces cerevisiae*

15 <400> 5
 gaagttccta ttctctagaa agtataggaa cttc 34

<210> 6
 <211> 23
 <212> ADN

ES 2 773 891 T3

<213> *Escherichia coli*

<400> 6
cctgcctttt tatactaact tga 23

5 <210> 7
<211> 22
<212> ADN
<213> *Cricetulus griseus*

10 <400> 7
taaggcctca tatgaaaata ta 22

15 <210> 8
<211> 22
<212> ADN
<213> *Cricetulus griseus*

<400> 8
atagatgtct tgcatactct ag 22

20 <210> 9
<211> 364
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: polipéptido sintético

25 <400> 9

Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys Tyr Asn
1 5 10 15

Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser
20 25 30

Ile Lys Ala Gln Ile Phe Pro Asn Gln Cys Tyr Lys Phe Lys His Gln
35 40 45

Leu Arg Leu Arg Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
50 55 60

Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Thr Asp Arg
65 70 75 80

Gly Ser Val Ser Asp Tyr Met Leu Ser Gln Ile Lys Pro Leu His Asn
85 90 95

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
100 105 110

ES 2 773 891 T3

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
 115 120 125

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala
 130 135 140

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
 145 150 155 160

Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro Ser Gln
 165 170 175

Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Pro Gly Ser Gly Ile
 180 185 190

Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr Asn Lys
 195 200 205

Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile
 210 215 220

Ile Ala Gln Ile Lys Pro Gly Gln Ser Tyr Lys Phe Lys His Thr Leu
 225 230 235 240

Gln Leu Val Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu
 245 250 255

Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Ile Asp Arg Gly
 260 265 270

Ser Ala Ser Asp Tyr Arg Leu Ser Glu Ile Lys Pro Leu His Asn Phe
 275 280 285

Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn
 290 295 300

Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro
 305 310 315 320

Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu
 325 330 335

Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val
 340 345 350

Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro

355

360

5
 <210> 10
 <211> 364
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: polipéptido sintético
 <400> 10

```

Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys Tyr Asn
 1           5                10                15

Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser
          20                25                30

Ile Ile Ala Gln Ile Pro Pro Asn Gln Ser Cys Lys Phe Lys His Gln
          35                40                45

Leu Arg Leu Thr Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
 50                55                60

Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg
65                70                75                80

Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His Asn
          85                90                95

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
          100                105                110

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
          115                120                125

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala
130                135                140

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
145                150                155                160

Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro Ser Gln
          165                170                175

Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Ser Pro Gly Ser Gly Ile
          180                185                190

Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr Asn Lys
195                200                205
    
```

ES 2 773 891 T3

Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile
 210 215 220
 Tyr Ala Gly Ile Ala Pro Asn Gln Ser Cys Lys Phe Lys His Gln Leu
 225 230 235 240
 Arg Leu Trp Phe Val Val Ser Gln Lys Thr Gln Arg Arg Trp Phe Leu
 245 250 255
 Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Ile Asp Asn Gly
 260 265 270
 Ser Val Ser His Tyr Arg Leu Ser Glu Ile Lys Pro Leu His Asn Phe
 275 280 285
 Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn
 290 295 300
 Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro
 305 310 315 320
 Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu
 325 330 335
 Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val
 340 345 350
 Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
 355 360

<210> 11

<211> 3663

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de secuencia artificial: polinucleótido sintético

<400> 11

5

10

ES 2 773 891 T3

tgcgcggtt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggta	60
cagcttgtct	gtaagcggat	gcogggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcagggcgc	240
attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	ggtgcgggcc	tcttcgctat	300
tacgccagct	ggcgaaaggg	ggatgtgctg	caagggcatt	aagttgggta	acgccagggt	360

ES 2 773 891 T3

tttcccagtc acgacgttgt aaaacgacgg ccagtgatt cgagctcggg acccagaaac 420
 ctttcaacca gcttttgagc taatgataga gagaagctca aggaattgga gcaatgcttg 480
 actagggatg tcagagggag gctatccaga ggagcttaca actgaggtaa acttaaaagt 540
 tagggagttt gtcaacttca acccacagaa tagagcagag ccaggaggag ctgaggcttc 600
 tgagtgttat ggtggaagca tcaccccaac ccttgacatc catatgcctg aagagtctgg 660
 aatgttatgg tggaaagtcc acccaagcct cccttcccgg tcgccctcca aaccctgcta 720
 catctcagaa atcccaccaa atgatgactc cctccccag agatattcaa gaccactccc 780
 acagggattt taaactgccc cccaaccccc agaaaataga tgtgtggttt tccaatctct 840
 ctttcctatc acgtctctgg ggagctggca ggccatttgg gagcattgta tccattaaac 900
 gacttctcag tggagactct gaaagccaga agagcctaga cagatagatg tcttgccaat 960
 tcttgcatat tctagagact acagatgccg gccagacta ttatatccag caaaagtctc 1020
 aaacaccata caaagtcaaa tttaaacagt atctatctac aaatccaata ttacagaagg 1080
 tgctagtagg aaaactccaa actaagatta actatacctg tgaagacaca ggaaataatc 1140
 tcacactggc aaaagaagaa aaacctctct ctctctctcc tctctctctc tctctctctc 1200
 tctctctctc tctctctctc tctctctctc tcacacacac acacacacac acacacacac 1260
 accaacacca ataccatgaa caacaaaata acaggaatta acaataattg atgtgtgtgt 1320
 atgtccctgt gtgtgtgtcc ttgtgtgtgt ctgtttgtgt gtctgtgtat atgtttgtca 1380
 cctgaggggt ggctcttctt tggtttgtga ggtttctacc caaaagcttg gcgtaatcat 1440
 ggtcatagct gtttctctgt tgaattgtt atccgctcac aattccacac aacatacagag 1500
 ccggaagcat aaagtgtaaa gcctgggggt cctaagagt gagctaactc acattaattg 1560
 cgttgccgct actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 1620
 tcggccaacg cgcggggaga ggcggtttgc gtattgggag ctcttccgct tctcgcctca 1680
 ctgactcgtc gcgctcggtc gttcggctgc ggcgagcggg atcagctcac tcaaaggcgg 1740
 taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaaggcc 1800
 agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggtccgcc 1860
 cccctgacga gcatcacaaa aatcgacgct caagtccagag gtggcgaaac ccgacaggac 1920
 tataaagata ccaggcgttt cccctggaa gctccctcgt gcgctctct gttccgacct 1980
 tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcata 2040
 gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc 2100
 acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 2160
 acccgtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 2220

ES 2 773 891 T3

cgaggtatgt aggcggtgct acagagttct tgaagtgggtg gcctaactac ggctacacta 2280
 gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 2340
 gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 2400
 agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 2460
 ctgacgctca gtggaacgaa aactcacggt aagggatttt ggtcatgaga ttatcaaaaa 2520
 ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 2580
 atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 2640
 tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 2700
 gggagggcct accatctggc cccagtgcctg caatgatacc gcgagacca cgctcaccgg 2760
 ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg 2820
 caactttatc cgctccatc cagtctatta attggtgccg ggaagctaga gtaagtagtt 2880
 cgccagttaa tagtttgccg aacgttggtg ccattgctac aggcatcgtg gtgtcacgct 2940
 cgtcgtttgg tatggcttca ttcagctccg gtccccaacg atcaaggcga gttacatgat 3000
 cccccatggt gtgcaaaaaa gcggttagct ccttcgggtcc tccgatcgtt gtcagaagta 3060
 agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca 3120
 tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 3180
 agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac 3240
 atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 3300
 ggatcttacc gctggtgaga tccagttcga tgtaaccac tccgtgcaccc aactgatctt 3360
 cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 3420
 caaaaaaggg aataagggcg acacggaat gttgaatact catactcttc ctttttcaat 3480
 attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 3540
 agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtct 3600
 aagaaacat tattatcatg acattaacct ataaaaatag gcgtatcacg aggcccttc 3660
 gtc 3663

5 <210> 12
 <211> 22
 <212> ADN
 <213> *Cricetulus griseus*

<400> 12
 tacatgatg tacaaaaat at 22

10 <210> 13
 <211> 364
 <212> PRT
 <213> Secuencia artificial

ES 2 773 891 T3

<220>

<223> Descripción de secuencia artificial: polipéptido sintético

<400> 13

```

Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys Tyr Asn
1           5           10           15

Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser
           20           25           30

Ile Phe Ala Ser Ile Thr Pro Arg Gln Cys Tyr Lys Phe Lys His Glu
           35           40           45

Leu Gln Leu Thr Phe Val Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
50           55           60

Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Ile Asp Gln
65           70           75           80

Gly Ser Val Ser His Tyr Arg Leu Ser Glu Ile Lys Pro Leu His Asn
           85           90           95

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
100          105          110

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
115          120          125

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala
130          135          140

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
145          150          155          160

Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro Ser Gln
165          170          175

Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Ser Pro Gly Ser Gly Ile
180          185          190

Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr Asn Lys
195          200          205

Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile
210          215          220

```

ES 2 773 891 T3

Ile Ala Gln Ile Lys Pro Asn Gln Ser Cys Lys Phe Lys His Gln Leu
225 230 235 240

Met Leu Thr Phe Thr Val Ala Gln Lys Thr Gln Arg Arg Trp Phe Leu
245 250 255

Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Ile Asp Ile Gly
260 265 270

Ser Val Ser Glu Tyr Arg Leu Ser Gln Ile Lys Pro Leu His Asn Phe
275 280 285

Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn
290 295 300

Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro
305 310 315 320

Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu
325 330 335

Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val
340 345 350

Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
355 360

5 <210> 14
<211> 22
<212> ADN
<213> *Cricetulus griseus*

<400> 14
aaggcactcg tgtaaacgga ta 22

10 <210> 15
<211> 364
<212> PRT
<213> Secuencia artificial

15 <220>
<223> Descripción de secuencia artificial: polipéptido sintético

<400> 15

Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys Tyr Asn
1 5 10 15

Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser
20 25 30

ES 2 773 891 T3

Ile Lys Ala Ile Ile Arg Pro Glu Gln Ser Tyr Lys Phe Lys His Arg
 35 40 45

Leu Arg Leu Val Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
 50 55 60

Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Tyr Asp Arg
 65 70 75 80

Gly Ser Val Ser Asp Tyr Tyr Leu Ser Glu Ile Lys Pro Leu His Asn
 85 90 95

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
 100 105 110

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
 115 120 125

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala
 130 135 140

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
 145 150 155 160

Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro Ser Gln
 165 170 175

Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Ser Pro Gly Ser Gly Ile
 180 185 190

Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr Asn Lys
 195 200 205

Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile
 210 215 220

Trp Ala Arg Ile Lys Pro Gly Gln Ser Tyr Lys Phe Lys His Thr Leu
 225 230 235 240

Glu Leu Val Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Ile Leu
 245 250 255

Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Thr Asp Ala Gly
 260 265 270

Ser Ala Ser Val Tyr Arg Leu Ser Glu Ile Lys Pro Leu His Asn Phe
 275 280 285

ES 2 773 891 T3

Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn
 290 295 300

Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro
 305 310 315 320

Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu
 325 330 335

Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val
 340 345 350

Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
 355 360

5 <210> 16
 <211> 29
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: cebador sintético

10 <400> 16
 ggagggacat taactgcat gcagtgatc 29

15 <210> 17
 <211> 29
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: cebador sintético

20 <400> 17
 gtctgggtt ggggtgtcta agcaacctc 29

<210> 18
 <211> 39
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Descripción de secuencia artificial: cebador sintético

<400> 18
 cacaggtgc cactcccagt tcaattacag ctcttaagg 39

30 <210> 19
 <211> 27
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Descripción de secuencia artificial: cebador sintético

<400> 19
 cgatggccca ctacgtgaac catcacc 27

ES 2 773 891 T3

5 <210> 20
 <211> 1821
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: polinucleótido sintético

<400> 20

cacaggtgtc cactcccagt tcaattacag ctcttaaggc tagagtactt aatcagactc	60
actataggct agcctcgagc cgccaccatg gcaccgaaga agaagcgcaa ggtgcatatg	120
gcaccgaaga agaagcgcaa ggtgcatatg aacaccaagt acaacaagga gttcctgctc	180
tacctggcgg gcttcgctga cggggacggc tccatcaagg cccagatctt tccgaaccag	240
tgctacaagt tcaagcatca gctgaggctc cgtttcaggc tcaccagaa gacacagcgc	300
cgttggttcc tcgacaagct ggtggacgag atcggggtgg gctacgtgac tgaccgcggc	360
agcgtctccg actacatgct gagccagatc aagcctctgc acaacttctt gaccagctc	420
cagcccttcc tgaagctcaa gcagaagcag gccaacctcg tgctgaagat catcgagcag	480
ctgccctccg ccaaggaatc cccggacaag ttctggaggc tgtgcacgtg ggtggaccag	540
atcgcgccc tcaacgacag caagaccgac aagacgacct cggagacggt gcgggcggtc	600
ctggactccc tcccaggatc cgtgggaggt ctatcgccat ctcaggcatc cagcgcgca	660
tcctcggett cctcaagccc gggttcaggc atctccgaag cactcagagc tggagcaggt	720
tccggcactg gatacaaaa ggaattcctg ctctacctgg cgggcttctt ggacggggac	780
ggctccatca tcgcccagat caagccgggt cagtcctaca agttcaagca tacctgcag	840
ctcgttttcc aggtcagca gaagacacag cgccgttggc tcctcgaaa gctggtggac	900
gagatcgggg tgggctatgt gatcgaccgc ggcagcgctt ccgactaccg cctgagcgag	960
atcaagcctc tgcacaactt cctgaccagc ctccagccct tcctgaagct caagcagaag	1020
caggccaacc tcgtgctgaa gatcatcgag cagctgcctt ccgccaagga atccccggac	1080
aagttcctgg aggtgtgcac ctgggtggac cagatcgccg ctctgaacga ctccaagacc	1140
cgcaagacca cttccgagac cgtccgagcc gttctagaca gtctctccga gaagaagaag	1200
tcgtccccct agacagtctc tccgagaaga agaagtcgtc cccctagcgg ccgcttcgag	1260
cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa	1320

10

ES 2 773 891 T3

aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 1380
 ataaacaagt taacaacaac aattgcattc attttatggt tcaggttcag ggggagatgt 1440
 gggaggtttt ttaaagcaag taaaacctct acaaagtgtg taaaatcgat aagatcttga 1500
 tccgggctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgctgag 1560
 cctgaatggc gaatggacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt 1620
 acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc 1680
 ccttcctttc tcgccacggt cgccggcttt ccccgtaag ctctaaatcg ggggctccct 1740
 ttagggttcc gatttagtgc tttacggcac ctcgaccca aaaaacttga ttaggggtgat 1800
 ggttcacgta gtgggccatc g 1821

5 <210> 21
 <211> 1821
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: polinucleótido sintético
 10 <400> 21

cacaggtgtc cactcccagt tcaattacag ctcttaaggc tagagtactt aatacgactc 60
 actataggct agcctcgagc cgccaccatg gcaccgaaga agaagcgcaa ggtgcatatg 120
 gcaccgaaga agaagcgcaa ggtgcatatg aacaccaagt acaacaagga gttcctgctc 180
 tacctggcgg gcttcgtgga cggggacggc tccatcatcg ccagatccc gccgaaccag 240
 tcctgcaagt tcaagcatca gctgcgctc accttccagg tcacgcagaa gacacagcgc 300
 cgttggttcc tcgacaagct ggtggacgag atcgggggtg gctacgtgcg cgaccgcggc 360
 agcgtctccg actacatcct gagcgagatc aagcctctgc acaacttctt gaccagctc 420
 cagcccttcc tgaagctcaa gcagaagcag gccaacctcg tgctgaagat catcgagcag 480
 ctgccctccg ccaaggaatc cccggacaag ttctctggagg tgtgcacctg ggtggaccag 540
 atcgccgctc tgaacgactc caagaccgc aagaccactt ccgagactgt ccgcgccgtt 600
 ctagacagtc tcccaggatc cgtgggaggc ctatcgccat ctgagcatc cagcgccgca 660
 tcctcggctt cctcaagccc gggttcaggg atctccgaag cactcagagc tggagcaggt 720
 tccggcactg gatacaaaa ggaattcctg ctctacctgg cgggcttctg ggacggggac 780
 ggctccatct acgccgggat cgcgccgaac cagtccctgca agttcaagca tcagctgcgc 840
 ctctggttcg tggtcagcca gaagacacag cgccggttgg tcctcgaaa gctggtggac 900
 gagatcgggg tgggctacgt gattgacaat ggcagcgtct ccattaccg cctgagcag 960
 atcaagcctc tgcacaactt cctgaccag ctccagcctt tcctgaagct caagcagaag 1020

ES 2 773 891 T3

caggccaacc tcgtgctgaa gatcatcgag cagctgccct ccgccaagga atccccggac 1080
 aagttcctgg aggtgtgcac ctgggtggac cagatcgccg ctttgaacga ctccaagacc 1140
 cgcaagacca cttccgagac tgtccgcgcc gttctagaca gtctctccga gaagaagaag 1200
 tcgtccccct agacagtctc tccgagaaga agaagtcgtc ccctagcgg ccgcttcgag 1260
 cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 1320
 aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 1380
 ataaacaagt taacaacaac aattgcattc attttatggt tcaggttcag ggggagatgt 1440
 gggagggttt ttaaagcaag taaaacctct acaaagtgg taaaatcgat aagatcttga 1500
 tccgggctgg cgtaatagcg aagaggcccg caccgatcgc ccttccaac agttgcgcag 1560
 cctgaatggc gaatggacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt 1620
 acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgetttcttc 1680
 ccttcctttc tcgccacggt cgccggcttt ccccgtaag ctctaaatcg ggggctccct 1740
 ttagggttcc gatttagtgc tttacggcac ctcgaccca aaaaacttga ttagggatgat 1800
 ggttcacgta gtgggccatc g 1821

5 <210> 22
 <211> 1821
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: polinucleótido sintético
 10 <400> 22

ES 2 773 891 T3

cacaggtgtc cactcccagt tcaattacag ctcttaaggc tagagtactt aatacgactc 60
actataggct agcctcgagc cgccaccatg gcaccgaaga agaagcgcaa ggtgcatatg 120
gcaccgaaga agaagcgcaa ggtgcatatg aacaccaagt acaacaagga gttcctgtctc 180
tacctggcgg gcttcgtcga cggggacggc tccatcaagg ccattatccg gccagagcag 240
tcctacaagt tcaagcatcg cctgcggctc gttttccagg tcacgcagaa gacacagcgc 300
cgttggttcc tcgacaagct ggtggacgag atcgggggtgg gctacgtgta cgaccgcggc 360
agcgtctccg actactatct gagcgagatc aagcctctgc acaacttctt gaccagctc 420
cagcccttcc tgaagctcaa gcagaagcag gccaacctcg tgctgaagat catcgagcag 480
ctgccctccg ccaaggaatc cccggacaag ttcttgaggg tgtgcacgtg ggtggaccag 540
atcggggccc tcaacgcagc caagaccgc aagacgacct cggagacggt gcgagcggtc 600
ctggactccc tcccaggatc cgtgggaggt ctatcgccat ctcaaggcatc cagcgcgccca 660
tcctcggctt cctcaagccc gggttcaggg atctccgaag cactcagagc tggagcaggt 720
tccggcactg gatacaacaa ggaattcctg ctctacctgg cgggcttctt ggacggggac 780
ggctccatct gggcccggat caagccgggg cagtcctaca agttcaagca taccctggag 840
ctcgtgttcc aggtcaccca gaagacacag cgccgttggg tcctcgacaa gctggtggac 900
gagatcgggg tgggctacgt gaccgacgcc ggcagcgcct ccgtctaccg cctgagcag 960
atcaagcctc tgcacaactt cctgaccag ctccagccct tcctgaagct caagcagaag 1020
caggccaacc tcgtgctgaa gatcatcgag cagctgccct ccgccaagga atccccggac 1080
aagttcctgg aggtgtgcac ctgggtggac cagatcgccg ctctgaacga ctccaagacc 1140
cgcaagacca ctcccgagac cgtccgcgcc gttctagaca gtctctccga gaagaagaag 1200
tcgtccccct agacagtctc tccgagaaga agaagtcgtc cccctagcgg ccgcttcgag 1260
cagacatgat aagatacatt gatgagttt gacaaaccac aactagaatg cagtgaaaaa 1320
aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 1380
ataaacaagt taacaacaac aattgcattc attttatggt tcaggttcag ggggagatgt 1440
gggaggtttt ttaaagcaag taaaacctct acaaatgtgg taaaatcgat aagatcttga 1500
tccgggctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag 1560
cctgaatggc gaatggacgc gcctgtagc ggcgcattaa gcgcggcggg tgtggtggtt 1620
acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc 1680
ccttcctttc tcgccacggt cggcggcttt ccccgtaag ctctaaatcg ggggctccct 1740
ttagggttcc gatttagtgc ttacggcac ctcgaccca aaaaacttga ttagggatgat 1800
ggttcacgta gtgggccatc g 1821

<210> 23
<211> 34

<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: cebador sintético

5 <400> 23
tgacagctct ggccttaagt gcctacgaaa ctag 34

10 <210> 24
<211> 39
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: cebador sintético

15 <400> 24
gtcttcctc ttgctgtag ccttgtaga actactgcc 39

20 <210> 25
<211> 5653
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: polinucleótido sintético

<400> 25

ES 2 773 891 T3

togcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	60
cagcttgtct	gtaagcggat	gocgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggtg	cttaactatg	cggcacacaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcaggcgcc	240
attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	ggtgcgggce	tcttcgctat	300
tacgccagct	ggcgaaaggg	ggatgtgctg	caaggcgatt	aagttgggta	acgccagggt	360
tttoccagtc	acgacgttgt	aaaacgacgg	ccagtgaatt	ccataccag	gggagctgta	420
ctgggctgca	gcctgcgcc	attcagccat	gcaccaggct	actccctcct	cttccagctt	480
totocttctg	atggocatag	gattagaaga	taagggactc	tagtgcaggt	caactgctga	540
ccagtgtgaa	aatgcacaga	ctacatgctg	gtagatcagc	acttcaaact	actgttcacc	600
atcatctctg	gaataagcac	tacatttaca	gggttcaaac	ctcaatgaat	ataaacaaac	660
aaaacacacc	tcccttcctt	cactgtctcc	catttctttg	gttcccatct	ccacatagaa	720
tttataatta	aaatttctaa	gtatctttcc	agaaatactt	cacacatggt	ataagcaaat	780
gtgcttttaa	agatactatt	ttaaattatg	aaaatggtta	tattagttga	gataaaaagaa	840
tagaatggga	agttccagaa	tttaaggcct	catatggatc	ccagctgtgg	aatgtgtgtc	900
agttaggggtg	tggaaagtcc	ccaggctccc	cagcaggcag	aagtatgcaa	agcatgcac	960
tcaattagtc	agcaaccagg	tgtggaaagt	cccaggctc	cccagcaggc	agaagtatgc	1020
aaagcatgca	tctcaattag	tcagcaacca	tagtcccgcc	cctaactccg	cccatcccgc	1080
ccctaactcc	gccagttcc	gcccattctc	cgcccatg	ctgactaatt	ttttttat	1140
atgcagaggc	cgaggccgcc	toggcctctg	agctattcca	gaagttagtga	ggaggctttt	1200
ttggaggcta	ccatggagaa	gttactatc	ogaagttcct	attctctaga	aagtatagga	1260
acttcaagct	tggcactggg	taccgccaag	ttgaccagt	ccgttccgg	gctcaccgcg	1320
cgcgacgtcg	ccggagcgg	cgagttctg	accgaccggc	tcgggttctc	ccgggacttc	1380
gtggaggacg	acttcgccc	tgtggtccg	gacgacgtga	ccctgttcat	cagcgcggtc	1440
caggaccagg	tgggtccgga	caacaccctg	gcctgggtgt	gggtgcgcgg	cctggacgag	1500

ES 2 773 891 T3

ctgtacgccg agtggtcgga ggtcgtgtcc acgaacttcc gggacgcctc cgggcccggc 1560
 atgaccgaga tcggcgagca gccgtggggg cgggagttcg ccctgcgcga cccggccggc 1620
 aactgcgtgc acttcgtggc cgaggagcag gactgacacc cgagcgaaaa cggctctgcgc 1680
 tgcgggacgc gcgaattgaa ttatggccca caccagtggc gcggcgactt ccagttcaac 1740
 atcagccgct acagtcaaca gcaactgatg gaaaccagcc atcgccatct gctgcacgcg 1800
 gaagaaggca catggctgaa tatcgacggg ttccatatgg ggattggtgg cgacgactcc 1860
 tggagcccgt cagtatcggc ggaattccag ctgagcgccg gtogctacca ttaccagttg 1920
 gtctggtgtc aaaaataata ataaccgggc aggggggatc tgcattggatc tttgtgaagg 1980
 aaccttactt ctgtggtgtg acataattgg acaaactacc tacagagatt taaagctcta 2040
 aggtaaatat aaaattttta agtgtataat gtgttaaaact actgattcta attgtttgtg 2100
 tattttagat tccaacctat ggaactgatg aatgggagca gtggtggaat gcctttaatg 2160
 aggaaaaact gttttgctca gaagaaatgc catctagtga tgatgaggct actgctgact 2220
 ctcaacattc tactcctcca aaaaagaaga gaaaggtaga agacccaag gactttcctt 2280
 cagaattgct aagttttttg agtcatgctg tgtttagtaa tagaactctt gcttgctttg 2340
 ctatttacac cacaaaggaa aaagctgcac tgctatacaa gaaaattatg gaaaaatatt 2400
 ctgtaacctt tataagtagg cataacagt ataatcataa catactgttt tttcttactc 2460
 cacacaggca tagagtgtct gctattaata actatgctca aaaattgtgt acctttagct 2520
 ttttaatttg taaaggggtt aataaggaat atttgatgta tagtgccttg actagagatc 2580
 ataatcagcc ataccacatt tgtagaggtt ttacttgctt taaaaaacct cccacacctc 2640
 cccctgaacc tgaaacataa aatgaatgca attggtgttg ttaacttggt tattgcagct 2700
 tataatgggt acaaataaag caatagcatc acaaatttca caaataaagc atttttttca 2760
 ctgcattcta gttgtggttt gtccaaactc atcaatgat cttatcatgt ctggatcccc 2820
 aggaagctcc tctgtgtcct cataaacctt aacctcctct acttgagagg acattccaat 2880
 cataggctgc ccatccacc tactagtata tgaaaatata aagcgctttc tcttttaagt 2940
 ctagggtagg tgtactagat cagcgctcag ctccatacca tgaagccatc caggagtcag 3000
 acctctctga cagccctgcc attgtcacag agaagtttct gtcaccagtg ctcatgctgt 3060
 cagaggagcg aaggagaaaa gatgtgagac ctccaagtc aaagtcactc atggataaaa 3120
 ccttagttgc atggcacacc agtgttaggg agtcggggaa acacagccat agcccagctt 3180
 cctctctggt cttgctetta ttaccaccag aaagaggttg cttagacaac ccaaaccaag 3240
 acacagggct ctgtgggagg gaatcagtcc caggcttctg gcacatgcta tgtcaccgga 3300
 aagccccagc cctactccga atccccaca gtacagcaaa taccagatta tagcatttaa 3360
 aggggcactc ttgcaaaga gaagcaccat tggaatagcc atgcttgaga actaagcttg 3420

ES 2 773 891 T3

gcgtaatcat ggtcatagct gttcctgtg tgaattggt atccgctcac aattccacac 3480
aacatacagag ccggaagcat aaagtgtaaa gcctgggggtg cctaatgagt gagctaactc 3540
acattaattg cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg 3600
cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattgggag ctcttccgct 3660
tcctcgcctc ctgactcgtt gcgctcggtc gttcggctgc ggcgagcggg atcagctcac 3720
tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga 3780
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 3840
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 3900
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 3960
gttccgacct tgcgccttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 4020
ctttctcata gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg 4080
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 4140
cttgagtcca acccggttaag acacgactta tcgccactgg cagcagccac tggtaacagg 4200
attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 4260
ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 4320
aaaagagttg gtagctcttg atccggcaaa caaacaccgg ctggtagcgg tggttttttt 4380
gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt 4440
tctacggggt ctgacgctca gtggaacgaa aactcacggt aagggatttt ggtcatgaga 4500
ttatcaaaaa ggatcttcac ctagatcctt ttaaattaa aatgaagttt taaatcaatc 4560
taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 4620
atctcagcga tctgtctatt tcggtcatcc atagttgcct gactccccgt cgtgtagata 4680
actacgatac gggagggctt accatctggc occagtgctg caatgatacc gcgagaccca 4740
cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga 4800
agtgtcctg caactttatc cgcctccatc cagtctatta attggtgccc ggaagctaga 4860
gtaagtagtt cgcaggttaa tagtttgcgc aacgttggtt ccattgctac aggcatcgtg 4920
gtgtcacgct cgtcgttttg tatggcttca ttcagctccg gttcccaacg atcaaggcga 4980
gttacatgat ccccatggt gtgcaaaaaa gcggttagct ccttcgggtcc tccgatcgtt 5040
gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct 5100
cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca 5160
ttctgagaat agtgtatgcg gcgaccgagt tgctcttgc cggcgtcaat acgggataat 5220
accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga 5280

ES 2 773 891 T3

aaactctcaa ggatottacc gctggtgaga tccagttoga tgtaaccac tcgtgcaccc 5340
 aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg 5400
 caaaatgccg caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc 5460
 ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt 5520
 gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcc 5580
 cctgacgtct aagaaacat tattatcatg acattaacct ataaaaatag gcgtatcacg 5640
 aggccctttc gtc 5653

5 <210> 26
 <211> 29
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: cebador sintético
 <400> 26
 10 agatgcatgc ttgcatact tctgcctgc 29

15 <210> 27
 <211> 5785
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: polinucleótido sintético
 <400> 27

gaaggatcgg gagatctccc gatcccctat ggtgcaactct cagtacaatc tgctctgatg 60
 ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggtcgct gagtagtgcg 120
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
 ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
 gattattgac tagttattaa tagtaatcaa ttacggggtc attagtcat agcccatata 300
 tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
 cccgccatt gacgtcaata atgacgtatg ttcccattag aacgccaata gggactttcc 420
 attgacgtca atgggtggag tatttacggg aaactgccc cttggcagta catcaagtgt 480
 atcatatgcc aagtacgcc octattgacg tcaatgacgg taaatggccc gcctggcatt 540
 atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
 tcgctattac catggtgatg oggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
 actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc 720
 20 aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg 780

ES 2 773 891 T3

gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
 ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggetagc 900
 gtttaaactt aagcttagcc accatggtga gcaagggcga ggagctgttc accggggtgg 960
 tgcccatcct ggtcgagctg gacggcgacg taaacggcca caagttcagc gtgtccggcg 1020
 agggcgaggg cgatgccacc tacggcaagc tgaccctgaa gttcatctgc accaccggca 1080
 agctgcccggt gccctggccc accctcgtga ccaccctgac ctacggagtg cagtgettca 1140
 gccgctaccg cgaccacatg aagcagcacg acttcttcaa gtccgccatg cccgaaggct 1200
 acgtccagga gcgcaccatc ttcttcaagg acgacggcaa ctacaagacc cgcgccgagg 1260
 tgaagttcga gggcgacacc ctggtgaacc gcatcgagct gaagggcatc gacttcaagg 1320
 aggacggcaa catcctgggg cacaagctgg agtacaacta caacagccac aacgtctata 1380
 tcatggccga caagcagaag aacggcatca aggtgaactt caagatccgc cacaacatcg 1440
 aggacggcag cgtgcagctc gccgaccact accagcagaa caccoccatc ggcgacggcc 1500
 ccgtgctgct gcccgacaac cactacctga gcaccagtc cgcctgagc aaagacccca 1560
 acgagaagcg cgatcacatg gtccctgctgg agttcgtgac cgcgcgccggg atcaactctcg 1620
 gcatggacga gctgtacaag taaggatcca ctagtccagt gtggtggaat tctgcagata 1680
 tccagcacag tggcggcgcg tcgagcttag agggcccgtt taaacccgct gatcagcctc 1740
 gactgtgcct tctagttgcc agccatctgt tgtttgccc tccccgtgc ctccctgac 1800
 cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcatg 1860
 tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga 1920
 ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcgga 1980
 aagaaccagc tggggctcta gggggtatcc ccacgcgcc tgtagcggcg cattaagcgc 2040
 ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgcgc tagcgcgcgc 2100
 tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct 2160
 aaatcggggg ctcccttag ggttcgatt tagtgcttta cggcacctcg accccaaaaa 2220
 acttgattag ggtgatggtt cacgtacctg gaagttccta ttccgaagtt cctattctct 2280
 agaaagtata ggaacttctt tggccaaaaa gcctgaactc accgcgacgt ctgtcgagaa 2340
 gtttctgac gaaaagttcg acagcgtctc cgacctgatg cagctctcgg agggcgaaga 2400
 atctcgtgct ttcagcttcg atgtaggagg gcgtggatat gtccctgcggg taaatagctg 2460
 cgccgatggt ttctacaaag atcgttatgt ttatcggcac tttgcatcgg ccgcgctccc 2520
 gattccggaa gtgcttgaca ttggggaatt cagcagagac ctgacctatt gcctctcccg 2580
 ccgtgcacag ggtgtcacgt tgcaagacct gcctgaaacc gaactgcccg ctgttctgca 2640

ES 2 773 891 T3

gcccgtcgcg gagcccatgg atgcatcgc tgcggccgat cttagccaga cgagcgggtt 2700
cgccccattc ggaccgcaag gaatcggcca atacactaca tggcgtgatt tcatatgcgc 2760
gattgctgat ccccatgtgt atcaactggca aactgtgatg gacgacaccg tcagtgcgtc 2820
cgtcgcgcag gctctcgatg agctgatgct ttgggcccag gactgccccg aagtccggca 2880
cctcgtgcac gcggatttcg gctccaacaa tgtcctgacg gacaatggcc gcataacagc 2940
ggtcattgac tggagcgagg cgatgttcgg ggattcccaa tacgaggtcg ccaacatctt 3000
cttctggagg ccgtggttgg cttgtatgga gcagcagacg cgctacttcg agcggaggca 3060
tccggagctt gcaggatcgc cgcggctccg ggcgtatatg ctccgcattg gtcttgacca 3120
actctatcag agcttggttg acggcaattt cgatgatgca gcttgggccc agggtcgatg 3180
cgacgcaatc gtcgatccg gagccgggac tgtcggggcg acacaaatcg cccgcagaag 3240
cgccggccgtc tggaccgatg gctgtgtaga agtactcgc gatagtggaa accgacgccc 3300
cagcactcgt ccgagggcaa aggaatagca cgtactacga gatttcgatt ccaccgccc 3360
cttctatgaa aggttgggct tcggaatcgt tttccgggac gccggctgga tgatctcca 3420
gcgcggggat ctcatgctgg agttcttcgc ccacccaac ttgtttattg cagcttataa 3480
tggttacaaa taaagcaata gcatcacaaa tttcacaaat aaagcatttt tttcactgca 3540
ttctagttgt ggtttgtcca aactcatcaa tgtatcttat catgtctgta taccgctgac 3600
ctctagctag agcttgccgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc 3660
gctcacaatt ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta 3720
atgagtgagc taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa 3780
cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgat 3840
tgggcgctct tccgcttccg cgtcactga ctcgctgcgc tcggtcgctt ggctgcggcg 3900
agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc 3960
aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt 4020
gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag 4080
tcagaggtgg cgaaaccoga caggactata aagataccag gcgtttcccc ctggaagctc 4140
cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc 4200
ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtagt 4260
cgttcgtctc aagctgggct gtgtgcacga accccccgtt cagcccagacc gctgcgcctt 4320
atccggtaac tctcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc 4380
agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttctttaa 4440
gtgggtggcct aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa 4500
gccagttacc ttcggaaaaa gagttgtag ctcttgatcc ggcaaaaaa ccaccgctgg 4560

ES 2 773 891 T3

tagcgggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaag gatctcaaga 4620
 agatoccttg atcttttcta cggggtctga cgctcagtg aacgaaaact cacgtaag 4680
 gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg 4740
 aagttttaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt 4800
 aatcagtgag gcacctatct cagcgatctg tctatttctg tcatccatag ttgcctgact 4860
 ccccgctgtg tagataacta cgatacggga gggttacca tctggcccca gtgctgcaat 4920
 gataaccgga gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg 4980
 aagggccgag cgcagaagtg gtccctgcaac tttatccgcc tccatccagt ctattaattg 5040
 ttgcccggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat 5100
 tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc 5160
 ccaacgatca aggcgagtta catgatcccc catgttctgc aaaaaagcgg ttagctcctt 5220
 cggctcctcg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc 5280
 agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga 5340
 gtactcaacc aagtcattct gagaatagtg tatgcccga cagagttgct cttgcccgcc 5400
 gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgctca tcattgaaa 5460
 acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta 5520
 acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg 5580
 agcaaaaaaca ggaaggcaaa atgccgcaaa aaaggaata agggcgacac ggaaatggtg 5640
 aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat 5700
 gagcggatac atatttgaat gtatttagaa aaataaaca ataggggttc cgcgcacatt 5760
 tccccgaaa gtgccacctg acgtc 5785

5 <210> 28
 <211> 30
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Descripción de secuencia artificial: cebador sintético
 10 <400> 28
 cagaaacttc tcgacagacg tcgcggtgag 30

15 <210> 29
 <211> 364
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> Descripción de secuencia artificial: polipéptido sintético
 <400> 29

ES 2 773 891 T3

Met Ala Pro Lys Lys Lys Arg Lys Val His Met Asn Thr Lys Tyr Asn
1 5 10 15

Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser
20 25 30

Ile Cys Ala Ser Ile Arg Pro Glu Gln Glu Arg Lys Phe Lys His Arg
35 40 45

Leu Val Leu Arg Phe Glu Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
50 55 60

Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Tyr Asp Ser
65 70 75 80

Gly Ser Val Ser Arg Tyr Tyr Leu Ser Gln Ile Lys Pro Leu His Asn
85 90 95

Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
100 105 110

Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser
115 120 125

Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala
130 135 140

Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala
145 150 155 160

Val Leu Asp Ser Leu Pro Gly Ser Val Gly Gly Leu Ser Pro Ser Gln
165 170 175

Ala Ser Ser Ala Ala Ser Ser Ala Ser Ser Ser Pro Gly Ser Gly Ile
180 185 190

Ser Glu Ala Leu Arg Ala Gly Ala Gly Ser Gly Thr Gly Tyr Asn Lys
195 200 205

Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile
210 215 220

Phe Ala Thr Ile Cys Pro Arg Gln Gln Tyr Lys Phe Lys His Gln Leu
225 230 235 240

ES 2 773 891 T3

Arg Leu Arg Phe Glu Val Asp Gln Lys Thr Gln Arg Arg Trp Phe Leu
245 250 255

Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Tyr Asp Leu Gly
260 265 270

Ser Val Ser Arg Tyr Gly Leu Ser Glu Ile Lys Pro Leu His Asn Phe
275 280 285

Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn
290 295 300

Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro
305 310 315 320

Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu
325 330 335

Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val
340 345 350

Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
355 360

- 5 <210> 30
- <211> 22
- <212> ADN
- <213> Secuencia artificial
- <220>
- <223> Descripción de secuencia artificial: oligonucleótido sintético
- 10 <400> 30
- cagcagctct caccccaccc ct 22
- <210> 31
- <211> 28
- <212> ADN
- 15 <213> Secuencia artificial
- <220>
- <223> Descripción de secuencia artificial: cebador sintético
- <400> 31
- ggaatctgac tgtgtaagc ctgtacac 28
- 20 <210> 32
- <211> 24
- <212> ADN
- <213> Secuencia artificial
- 25 <220>
- <223> Descripción de secuencia artificial: cebador sintético

ES 2 773 891 T3

<400> 32
cagcactcag gaggtagagg cagg 24

5 <210> 33
<211> 36
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: oligonucleótido sintético

10 <400> 33
tcttactgac atccacttg ccttctctc cacagg 36

15 <210> 34
<211> 130
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: polinucleótido sintético

20 <400> 34

acttgtttat tgcagcttat aatggttaca aataaagcaa tagcatcaca aatttcacaa	60
ataaagcatt ttttctactg cattctagtt gtggtttgtc caaactcatc aatgtatctt	120
atcatgtctg	130

25 <210> 35
<211> 225
<212> ADN
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: polinucleótido sintético

30 <400> 35

ctgtgccttc tagttgccag ccatctggtg ttgcccctc ccccgctgct tccttgacct	60
tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca tgcattgtc	120
tgagtaggtg tcattctatt ctggggggtg ggggtgggca ggacagcaag ggggaggatt	180
gggaagacaa tagcaggcat gctggggatg cgggtgggctc tatgg	225

35 <210> 36
<211> 9
<212> PRT
<213> Secuencia artificial

<220>
<223> Descripción de secuencia artificial: péptido sintético

40 <400> 36

Met Ala Pro Lys Lys Lys Arg Lys Val
1 5

ES 2 773 891 T3

5 <210> 37
 <211> 53
 <212> ADN
 <213> *Cricetulus griseus*

 <400> 37
 gaatgggaag ttccagaatt taaggcctca tatgaaaata taaagcgctt tct 53

10 <210> 38
 <211> 36
 <212> ADN
 <213> *Cricetulus griseus*

 <400> 38
 gaatgggaag ttccagaatt taataaagcg ctttct 36

15 <210> 39
 <211> 28
 <212> ADN
 <213> *Cricetulus griseus*

20 <400> 39
 gaatgggaag ttccagaaag cgctttct 28

25 <210> 40
 <211> 27
 <212> ADN
 <213> *Cricetulus griseus*

 <400> 40
 gaatgggaag ttccagaagc gctttct 27

30 <210> 41
 <211> 69
 <212> ADN
 <213> *Cricetulus griseus*

35 <400> 41

tctgaaagcc agaagagcct agacagatag atgtcttgca tactctagag actacagatg 60

ccggcccag 69

40 <210> 42
 <211> 58
 <212> ADN
 <213> *Cricetulus griseus*

 <400> 42
 tctgaaagcc agaagagcct agacatgcat actctagaga ctacagatgc cggcccag 58

45 <210> 43
 <211> 49
 <212> ADN
 <213> *Cricetulus griseus*

50 <400> 43
 tctgaaagcc agaagagcct agacagatag atgtcagatg cggcccag 49

ES 2 773 891 T3

<210> 44
 <211> 36
 <212> ADN
 <213> *Cricetulus griseus*
 5 <400> 44
 tctgaaagcc agaagagcct acagatgccg gccag 36

 <210> 45
 <211> 51
 <212> ADN
 <213> *Cricetulus griseus*
 10 <400> 45
 tctgaaagcc agaagagcct agacagatag atgtcttgca tgccggccca g 51

 <210> 46
 <211> 65
 <212> ADN
 <213> *Cricetulus griseus*
 15
 <400> 46
 20 tctgaaagcc agaagagcct agacagatag atgtcttgct ctagagacta cagatgccgg 60
 cccag 65

 <210> 47
 <211> 67
 <212> ADN
 <213> *Cricetulus griseus*
 25 <400> 47
 tctgaaagcc agaagagcct agacagatag atgtcttata ctctagagac tacagatgcc 60
 ggcccag 67

 <210> 48
 <211> 38
 <212> ADN
 <213> *Cricetulus griseus*
 30
 <400> 48
 35 tctgaaagcc agaagagcct agacagatgc cggcccag 38

 <210> 49
 <211> 54
 <212> ADN
 <213> *Cricetulus griseus*
 40 <400> 49
 tctgaaagcc agaagagcct agacagatag atgtctttac agatgccgac ccag 54

 <210> 50
 <211> 65
 <212> ADN
 <213> *Cricetulus griseus*
 45 <400> 50

ES 2 773 891 T3

		tctgaaagcc agaagagcct agacagatag atgtcttgct ctagagacta cagatgccgg	60
		cccag	65
5		<210> 51 <211> 69 <212> ADN <213> <i>Cricetulus griseus</i> <400> 51	
		tctgaaagcc agaagagcct agacagatag atgtctttgc atactctaga gactacagat	60
10		gccggccca	69
15		<210> 52 <211> 75 <212> ADN <213> <i>Cricetulus griseus</i> <400> 52	
		gatgctttat tcctagagac caatttaagg cactcgtgta aacggataat ggacatggtg	60
		agcaaccagc acccc	75
20		<210> 53 <211> 69 <212> ADN <213> <i>Cricetulus griseus</i> <400> 53	
25		gatgctttat tcctagagac caatttaagg cactcgtgtg taatggacat ggtgagcaac	60
		cagcaccccc	69
30		<210> 54 <211> 74 <212> ADN <213> <i>Cricetulus griseus</i> <400> 54	
		gatgctttat tcctagagac caatttaagg cactcgtgta accgataatg gacatggtga	60
		gcaaccagca cccc	74
35		<210> 55 <211> 71 <212> ADN <213> <i>Cricetulus griseus</i> <400> 55	
40			

ES 2 773 891 T3

	gatgctttat tcctagagac caatttaagg cactcaaacg gataatggac atggtgagca	60
	accagcacc c	71
5	<210> 56 <211> 73 <212> ADN <213> <i>Cricetulus griseus</i> <400> 56	
	gatgctttat tcctagagac caatttaagg cactcgtaaa cggataatgg acatggtgag	60
	caaccagcac ccc	73
10	<210> 57 <211> 74 <212> ADN <213> <i>Cricetulus griseus</i>	
15	<400> 57	
	gatgctttat tcctagagac caatttaagg cgctgtgtaa acggataatg gacatggtga	60
	gcaaccagca cccc	74
20	<210> 58 <211> 75 <212> ADN <213> <i>Cricetulus griseus</i> <400> 58	
	gatgctttat tcctagagac caatttaagg cattcgtgta aacggataat ggacatggtg	60
25	agcaaccagc acccc	75
30	<210> 59 <211> 75 <212> ADN <213> <i>Cricetulus griseus</i> <400> 59	
	gatgctttat tcctagagac caatttaagg cactcatgta aacggataat ggacatggtg	60
	agcaaccagc acccc	75
35	<210> 60 <211> 75 <212> ADN <213> <i>Cricetulus griseus</i>	
40	<400> 60	

ES 2 773 891 T3

	gatgctttat tcctagagac caatttaagg cgctcgtgta aacggataat ggacatggtg	60
	agcaaccagc acccc	75
5	<210> 61 <211> 75 <212> ADN <213> <i>Cricetulus griseus</i> <400> 61	
	gatgctttat tcctagagac caatttaagg cactcgcgta aacggataat ggacatggtg	60
	agcaaccagc acccc	75
10	<210> 62 <211> 75 <212> ADN <213> <i>Cricetulus griseus</i>	
15	<400> 62	
	gatgctttat tcctagagac caatttaagg cacacgtgta aacggataat ggacatggtg	60
	agcaaccagc acccc	75
20	<210> 63 <211> 74 <212> ADN <213> <i>Cricetulus griseus</i>	
	<400> 63	
	gatgctttat tcctagagac caatttaagg cactcgtgtg taaacggata atggacatgg	60
25	tgagcaacca gcac	74
30	<210> 64 <211> 57 <212> ADN <213> <i>Cricetulus griseus</i> <400> 64	
	gatgctttat tcctagagac caatttaagg catggtgagt aaccgagcaa ccagcac	57
35		

REIVINDICACIONES

1. Un procedimiento para la inserción de una secuencia exógena en un locus que se pueda amplificar de una célula de mamífero que comprende:

5 (a) proporcionar una célula de mamífero que tiene un sitio diana endógeno proximal a un gen de selección endógeno en el locus que se puede amplificar, en el que el sitio diana endógeno está de 0 a 100.000 pares de bases corriente abajo de la región reguladora 3' del gen de selección endógeno o de 0 a 100.000 pares de bases corriente arriba de la región reguladora 5' del gen de selección endógeno, en el que el sitio diana endógeno comprende:

10 (i) una secuencia de reconocimiento para una endonucleasa específica del sitio;
 (ii) una región flanqueante 5' en 5' de la secuencia de reconocimiento; y
 (iii) una región flanqueante 3' en 3' de la secuencia de reconocimiento; y

(b) introducir una rotura de doble cadena entre las regiones flanqueantes 5' y 3' del sitio diana endógeno;
 (c) poner en contacto la célula con un vector donante que comprende de 5' a 3':

15 (i) una región flanqueante donante 5' homóloga de la región flanqueante 5' del sitio diana endógeno;
 (ii) una secuencia exógena; y
 (iii) una región flanqueante donante 3' homóloga de la región flanqueante 3' del sitio diana endógeno;

de manera que la región flanqueante donante 5', la secuencia exógena y la región flanqueante donante 3' se insertan entre las regiones flanqueantes 5' y 3' del sitio diana endógeno mediante recombinación homóloga para proporcionar una célula modificada.

20 2. El procedimiento de la reivindicación 1, que comprende adicionalmente el cultivo de la célula modificada en presencia de un compuesto que inhibe la función del gen de selección endógeno para co-amplificar el número de copias del gen de selección endógeno y el número de copias de la secuencia exógena, y opcionalmente/o en el que la secuencia exógena comprende un gen de interés.

3. El procedimiento de una cualquiera de las reivindicaciones 1-2, en el que:

25 (a) el gen de selección endógeno es de la glutamina sintetasa (GS) y el locus es de metionina sulfoximina (MSX) que se puede amplificar;
 (b) el gen de selección endógeno es de la dihidrofolato reductasa (DHFR) y el locus es de metotrexato (MTX) que se puede amplificar; o
 30 (c) el gen de selección endógeno se selecciona de entre el grupo que consiste en dihidrofolato reductasa, glutamina sintetasa, hipoxantina fosforribosil transferasa, treonil ARNt sintetasa, Na,K-ATPasa, asparagina sintetasa, ornitina descarboxilasa, inosina-5' - monofosfato deshidrogenasa, adenosina desaminasa, timidilato sintetasa, aspartato transcarbamilasa, metalotioneína, adenilato desaminasa (1,2), UMP-sintetasa y ribonucleótido reductasa, preferentemente en el que el gen de selección se puede amplificar por selección con un agente de selección
 35 seleccionado de entre el grupo que consiste en metotrexato (MTX), metionina sulfoximina (MSX), aminopterina, hipoxantina, timidina, borrelidina, ouabaína, albizina, beta-aspartil hidroxamato, alfa-difluorometilornitina (DFMO), ácido micofenólico, adenosina, alanosina, 2'desoxicoformicina, fluorouracilo, N-fosfonacetil-L-aspartato (PALA), cadmio, adenina, azaserina, coformicina, 6-azauridina, pirazofurano, hidroxiourea, motexafin gadolinio, fludarabina, cladribina, gemcitabina, tezacitabina y triapina.

40 4. Un procedimiento para la inserción de una secuencia exógena en un locus que se pueda amplificar de una célula de mamífero que comprende:

(a) proporcionar una célula de mamífero que tiene un sitio diana endógeno proximal a un gen de selección endógeno en el locus que se puede amplificar, en el que el sitio diana endógeno está de 0 a 100.000 pares de bases corriente abajo de la región reguladora 3' del gen de selección endógeno o de 0 a 100.000 pares de bases corriente arriba de la región reguladora 5' del gen de selección endógeno, en el que el sitio diana endógeno
 45 comprende:

(i) una secuencia de reconocimiento para una endonucleasa específica del sitio;
 (ii) una región flanqueante 5' en 5' de la secuencia de reconocimiento; y
 (iii) una región flanqueante 3' en 3' de la secuencia de reconocimiento; y

50 (b) introducir una rotura de doble cadena entre las regiones flanqueantes 5' y 3' del sitio diana endógeno;
 (c) poner en contacto la célula con un vector donante del sitio diana modificado que comprende de 5' a 3':

(i) una región flanqueante donante 5' homóloga de la región flanqueante 5' del sitio diana endógeno;
 (ii) una secuencia exógena que comprende un sitio diana modificado; y
 (iii) una secuencia flanqueante donante 3' homóloga de la región flanqueante 3' del sitio diana endógeno;

de manera que la región flanqueante donante 5', la secuencia exógena y la región flanqueante donante 3' se

insertan entre las regiones flanqueantes 5' y 3' del sitio diana endógeno mediante recombinación homóloga para proporcionar una célula de mamífero que comprende el sitio diana modificado;

(d) introducir una rotura de doble cadena entre las regiones flanqueantes 5' y 3' del sitio diana modificado;

5 (e) poner en contacto la célula que comprende el sitio diana modificado con un vector donante de la secuencia de interés que comprende de 5' a 3':

(i) una región flanqueante donante 5' homóloga de la región flanqueante 5' del sitio diana modificado;

(ii) una secuencia exógena que comprende una secuencia de interés; y

(iii) una región flanqueante donante 3' homóloga de la región flanqueante 3' del sitio diana modificado;

10 de manera que la región flanqueante donante 5', la secuencia exógena que comprende la secuencia de interés y la región flanqueante donante 3' se insertan entre las regiones flanqueantes 5' y 3' del sitio diana modificado mediante recombinación homóloga para proporcionar una célula de mamífero modificada que comprende la secuencia de interés.

15 5. El procedimiento de la reivindicación 4, que comprende adicionalmente el cultivo de la célula de mamífero modificada en presencia de un compuesto que inhiba la función del gen de selección endógeno para co-amplificar el número de copias del gen de selección endógeno y el número de copias de la secuencia de interés, y opcionalmente/o en el que la secuencia de interés comprende un gen.

6. El procedimiento de una cualquiera de las reivindicaciones 4-5, en el que:

20 (a) el gen de selección endógeno es de la glutamina sintetasa (GS) y el locus es de metionina sulfoximina (MSX) que se puede amplificar;

(b) el gen de selección endógeno es de la dihidrofolato reductasa (DHFR) y el locus es de metotrexato (MTX) que se puede amplificar; o

25 (c) el gen de selección endógeno se selecciona de entre el grupo que consiste en dihidrofolato reductasa, glutamina sintetasa, hipoxantina fosforribosil transferasa, treonil ARNt sintetasa, Na,K-ATPasa, asparagina sintetasa, ornitina descarboxilasa, inosina-5' - monofosfato deshidrogenasa, adenosina desaminasa, timidilato sintetasa, aspartato transcarbamilasa, metalotioneína, adenilato desaminasa (1,2), UMP-sintetasa y ribonucleótido reductasa, preferentemente en el que el locus es metotrexato (MTX), metionina sulfoximina (MSX), aminopterina, hipoxantina, timidina, borrelidina, ouabaina, albizina, beta-aspartil hidroxamato, alfa-difluorometilornitina (DFMO), ácido micofenólico, adenosina, alanosina, 2'desoxicoformicina, fluorouracilo, N-Fosfonacetil-L-Aspartato (PALA),
30 cadmio, adenina, azaserina, coformicina, 6-azauridina, pirazofurano, hidroxurea, motexafin gadolinio, fludarabina, cladribina, gemcitabina, tezacicabina y/o triapina que se puedan amplificar.

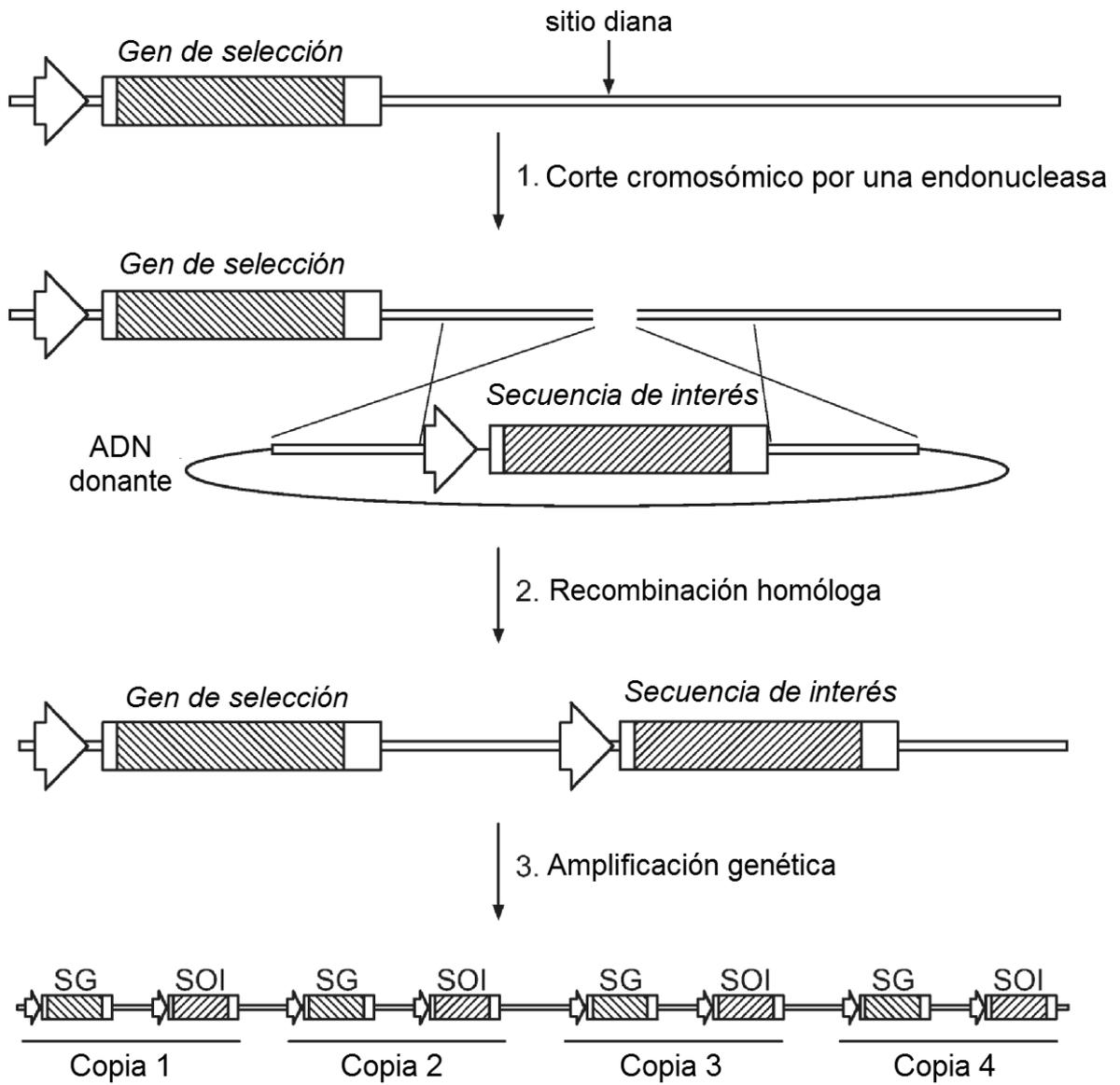


FIG. 1

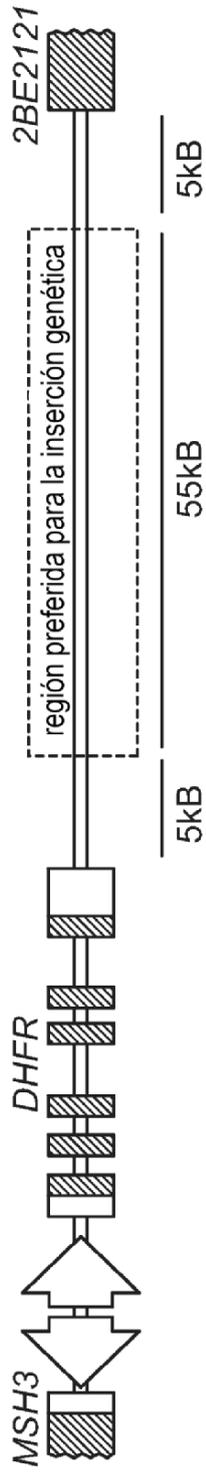


FIG. 2A

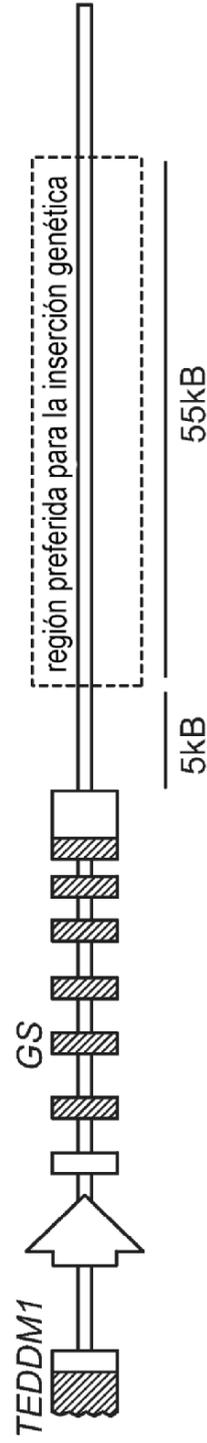


FIG. 2B

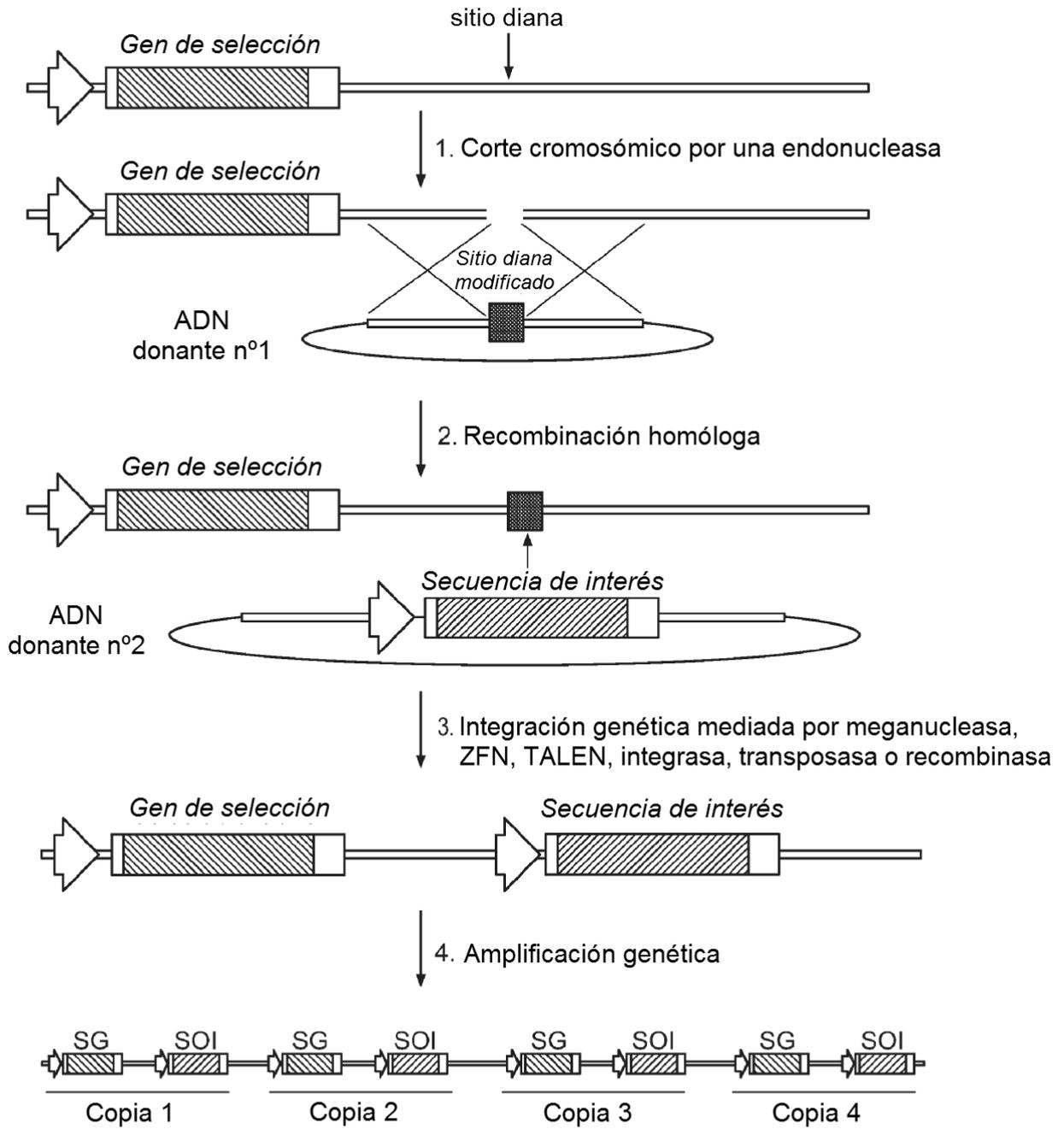


FIG. 3

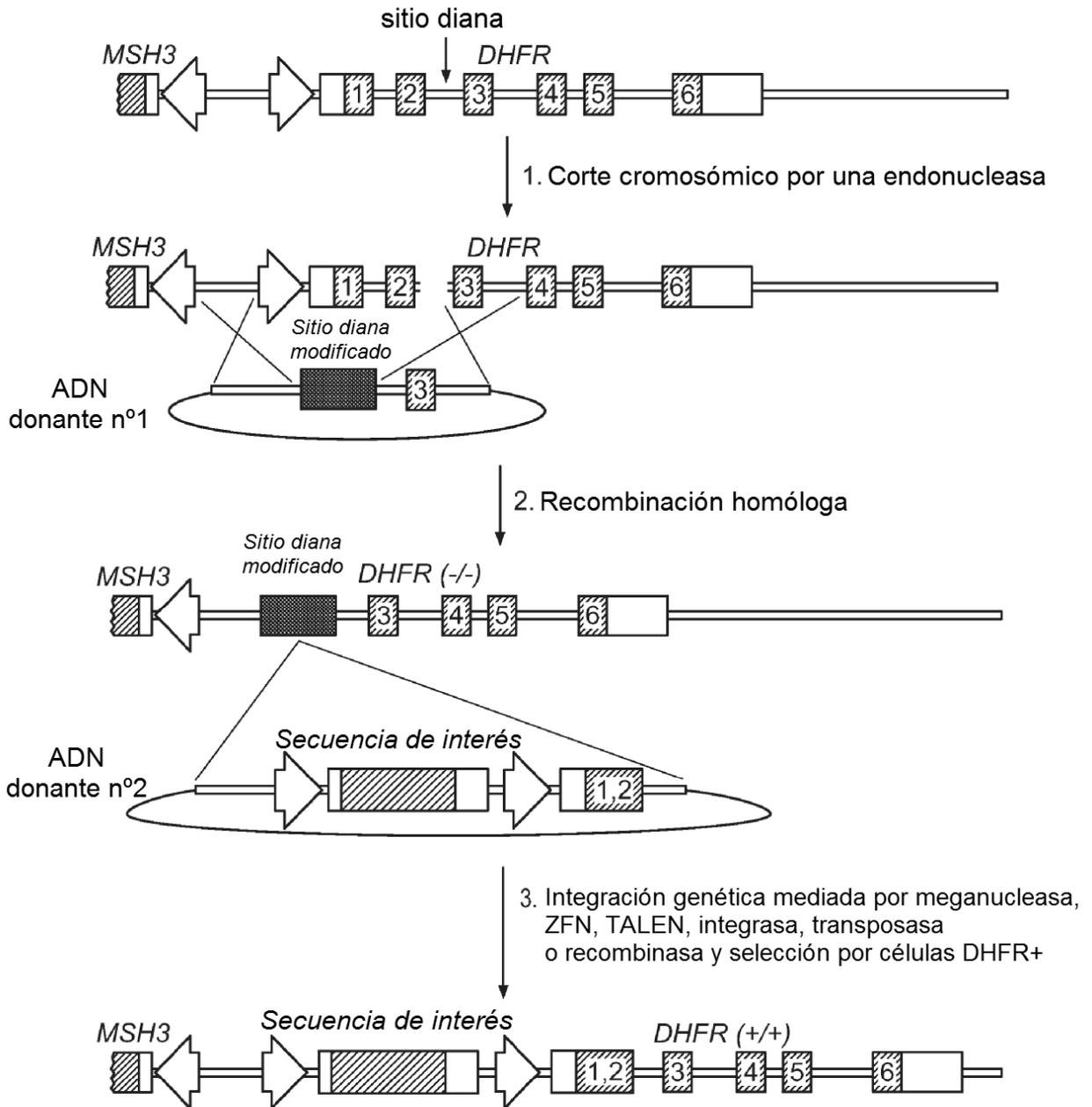


FIG. 4

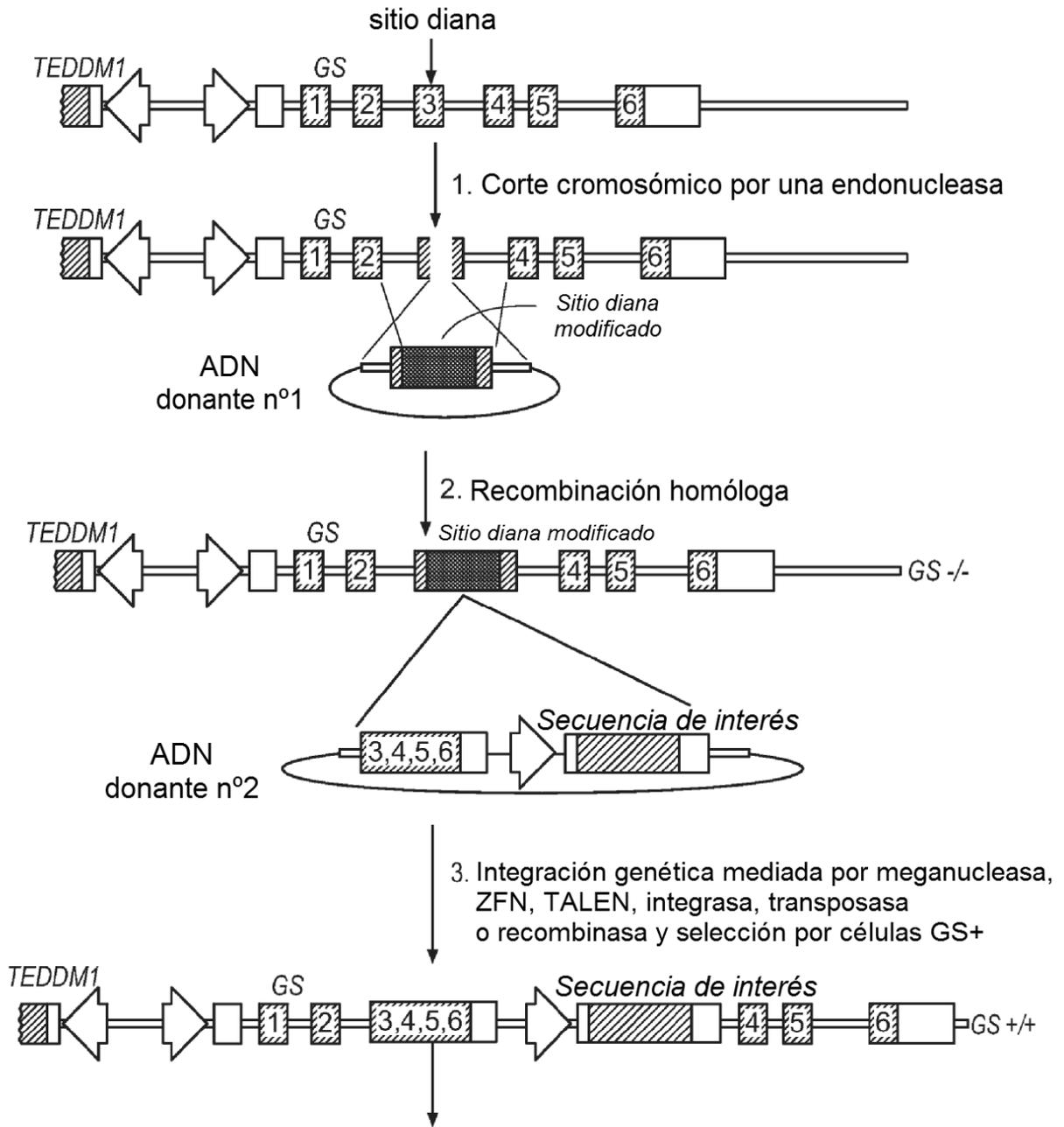


FIG. 5

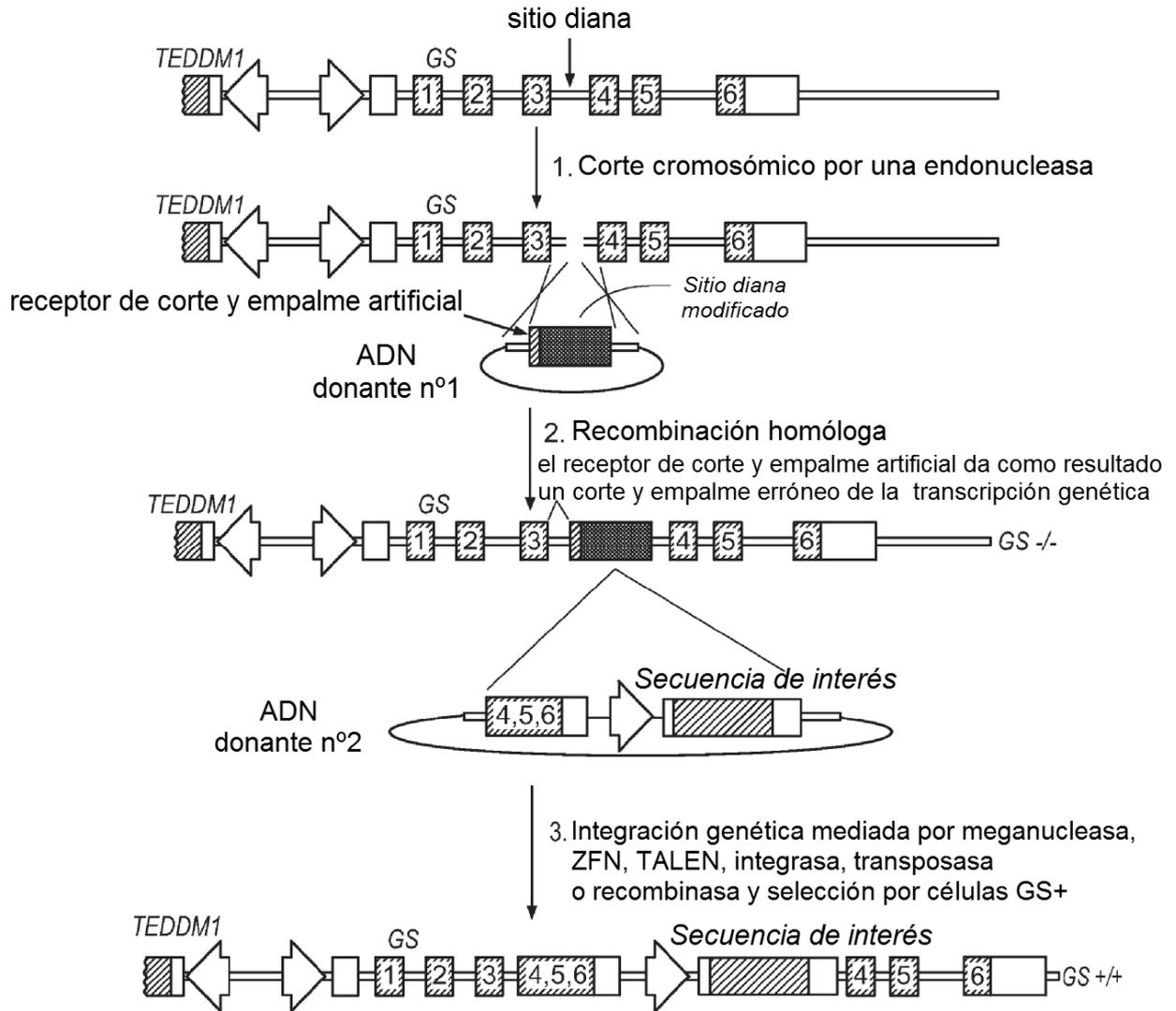


FIG. 6

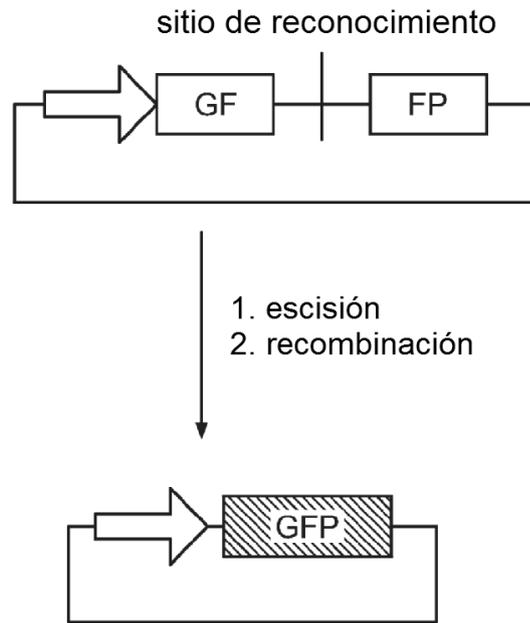


FIG. 7A

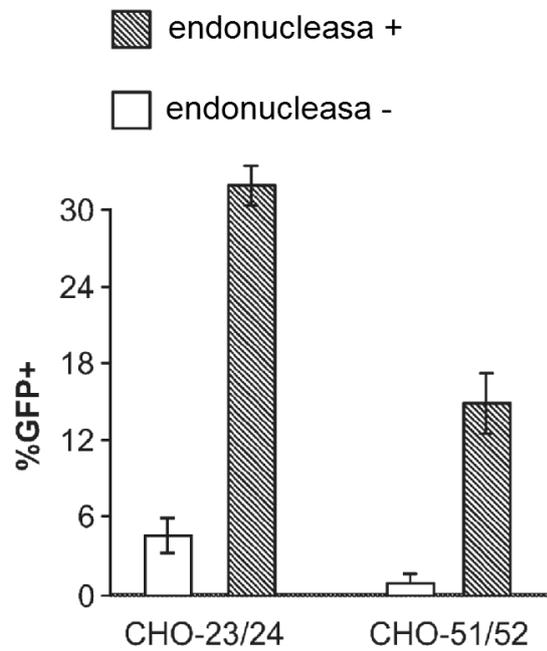


FIG. 7B

sitio CHO-23/24

TS: GAATGGGAAGTTCAGAAATTTAAGGCCTCATATGAAAATAATAAGCGCTTTCT
 F8: GAATGGGAAGTTCAGAAATTTAA-----TAAAGCGCTTTCT
 F10: GAATGGGAAGTTCAGAA-----AGCGCTTTCT
 G2: GAATGGGAAGTTCAGAAATTTAA-----TAAAGCGCTTTCT
 H4: GAATGGGAAGTTCAGAA-----AGCGCTTTCT
 H5: GAATGGGAAGTTCAGAAATTTAA-----TAAAGCGCTTTCT
 G1: GAATGGGAAGTTCAGAAATTTAA-----TAAAGCGCTTTCT
 G3: -----eliminación de 561 pb-----
 G8: -----eliminación de 348 pb-----
 F1: -----eliminación de 562 pb-----
 G8: -----eliminación de 348 pb-----
 F1: -----eliminación de 562 pb-----
 E8: -----eliminación de 568 pb-----
 G12: -----eliminación de 587 pb-----

FIG. 7C

sitio CHO-51/52

TS TCTGAAAGCCAGAAAGAGCCTTAGACAGATAGATGCTTGCATACTCTAGAGACTACAGATGCCGGCCCCAG
F3 TCTGAAAGCCAGAAAGAGCCTAGACA-----TGCATACTCTAGAGACTACAGATGCCGGCCCCAG
G10 TCTGAAAGCCAGAAAGAGCCTAGACAGATAGATGTC-----AGATGCCGGCCCCAG
F1 TCTGAAAGCCAGAAAGAGCCTA-----CAGATGCCGGCCCCAG
E3 TCTGAAAGCCAGAAAGAGCCTAGACAGATAGATGCTTGC-----ATGCCGGCCCCAG
E9 TCTGAAAGCCAGAAAGAGCCTAGACAGATAGATGCTTGC-----TCTAGAGACTACAGATGCCGGCCCCAG
F11 TCTGAAAGCCAGAAAGAGCCTAGACAGATAGATGCTT--ATACTCTAGAGACTACAGATGCCGGCCCCAG
F6 TCTGAAAGCCAGAAAGAGCCTAGACAGAT-----GCCGGCCCCAG
G5 TCTGAAAGCCAGAAAGAGCCTAGACAGATAGATGCTT-----TACAGATGCCGACCCCAG
H10 TCTGAAAGCCAGAAAGAGCCTAGACAGATAGATGCTTGC-----TCTAGAGACTACAGATGCCGGCCCCAG
F11 TCTGAAAGCCAGAAAGAGCCTAGACAGATAGATGCTTTGGCATACTCTAGAGACTACAGATGCCGGCCCCA

FIG. 7D

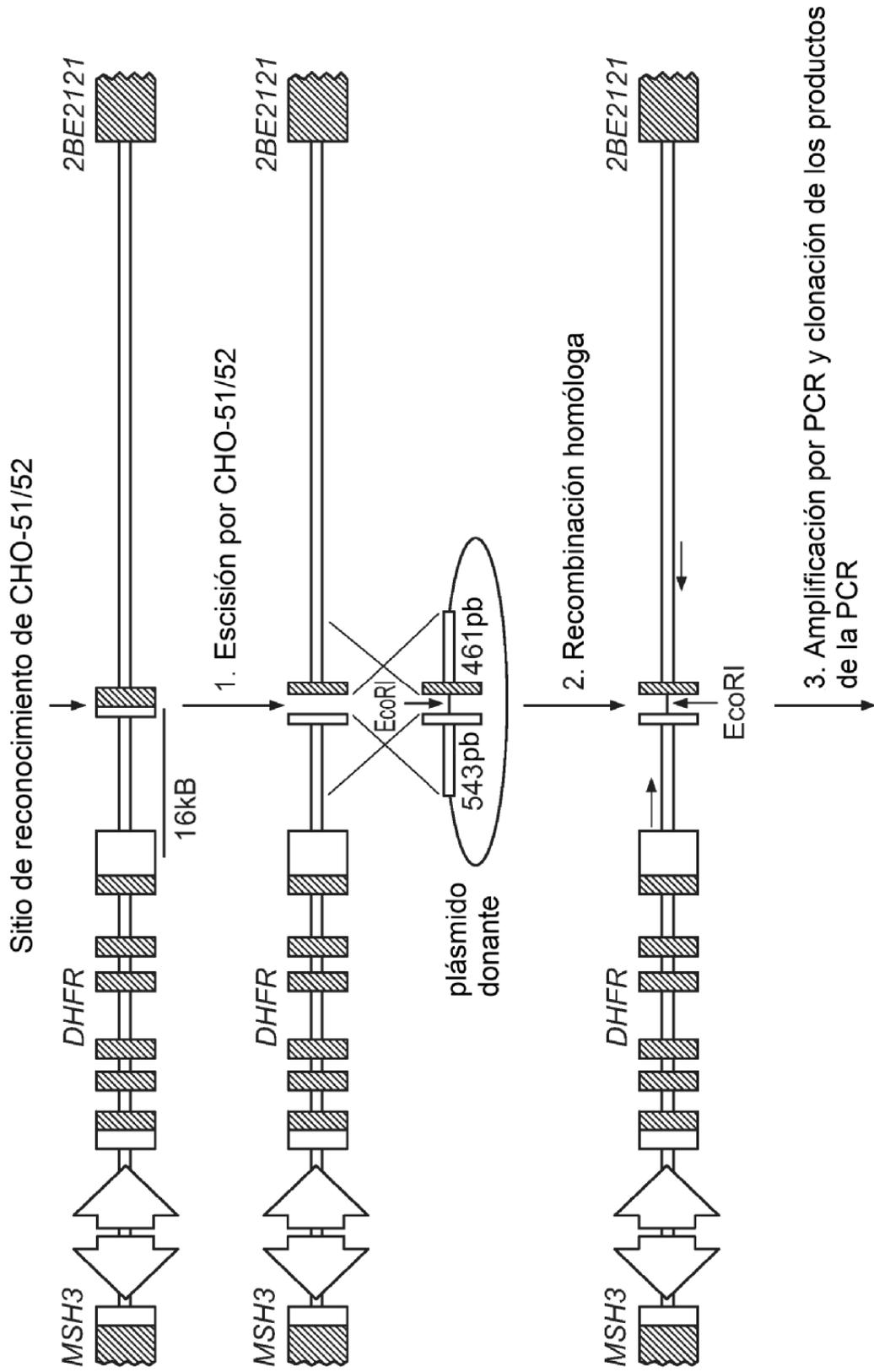


FIG. 8A

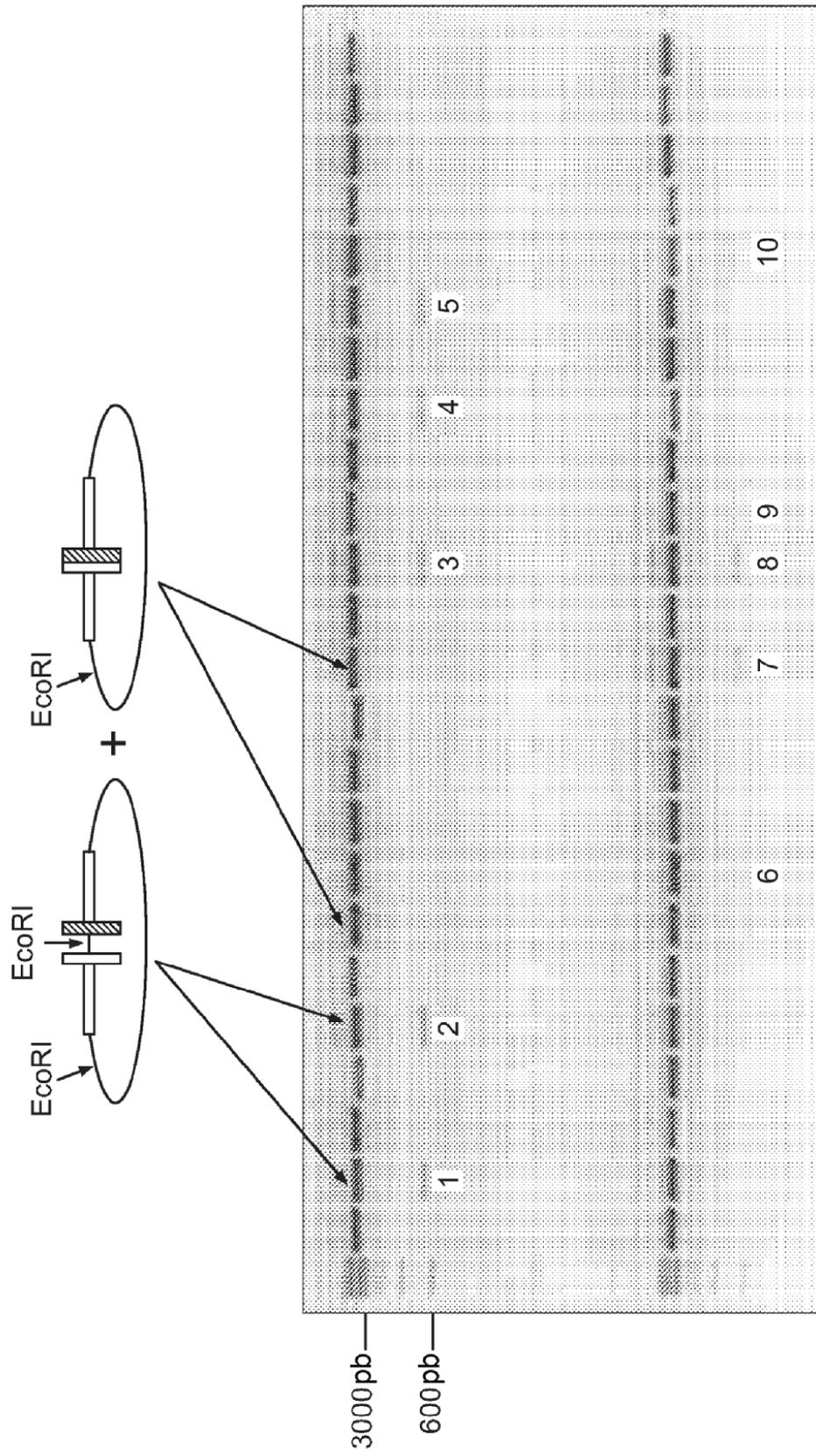


FIG. 8B

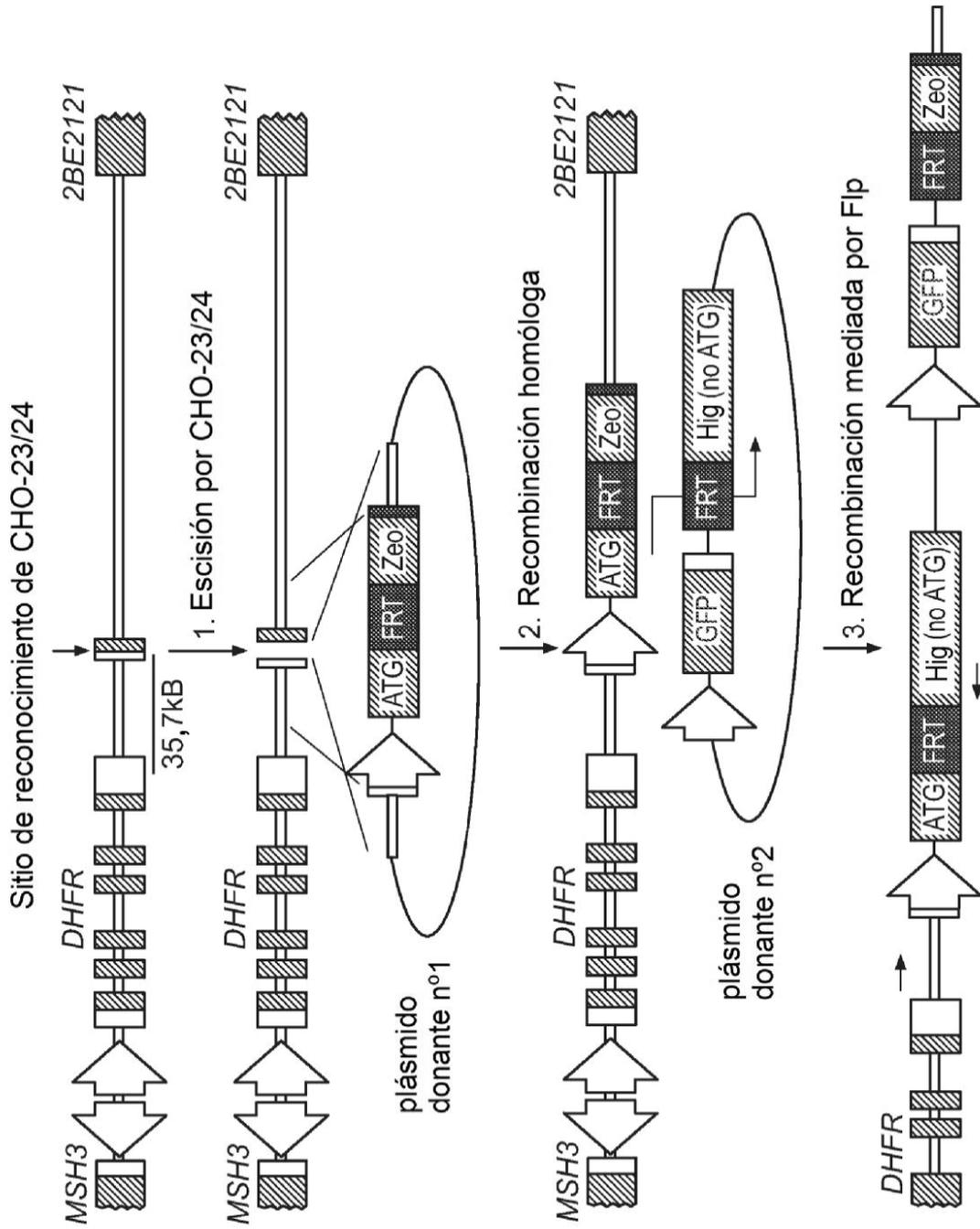


FIG. 9A

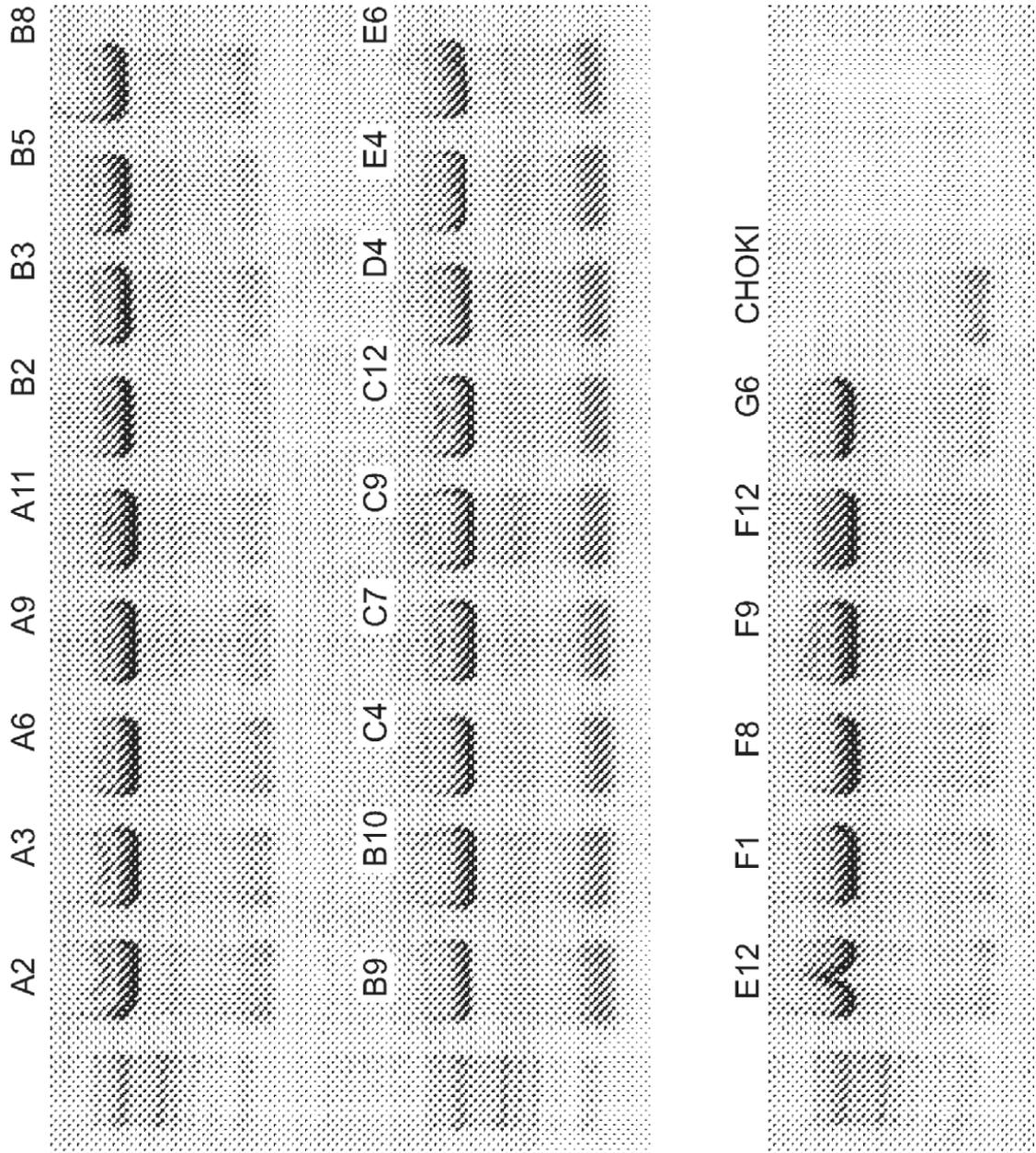


FIG. 9B

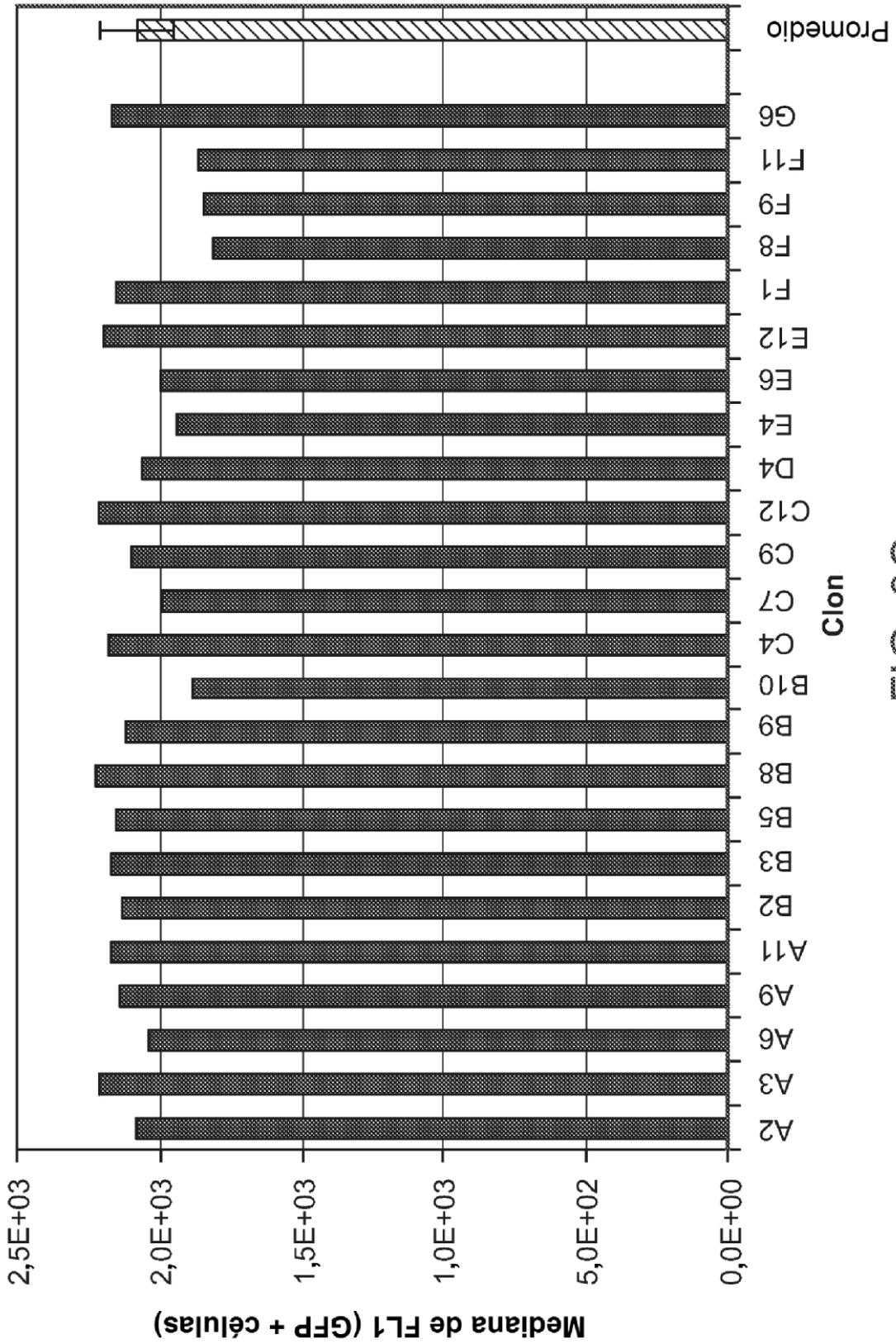


FIG. 9C

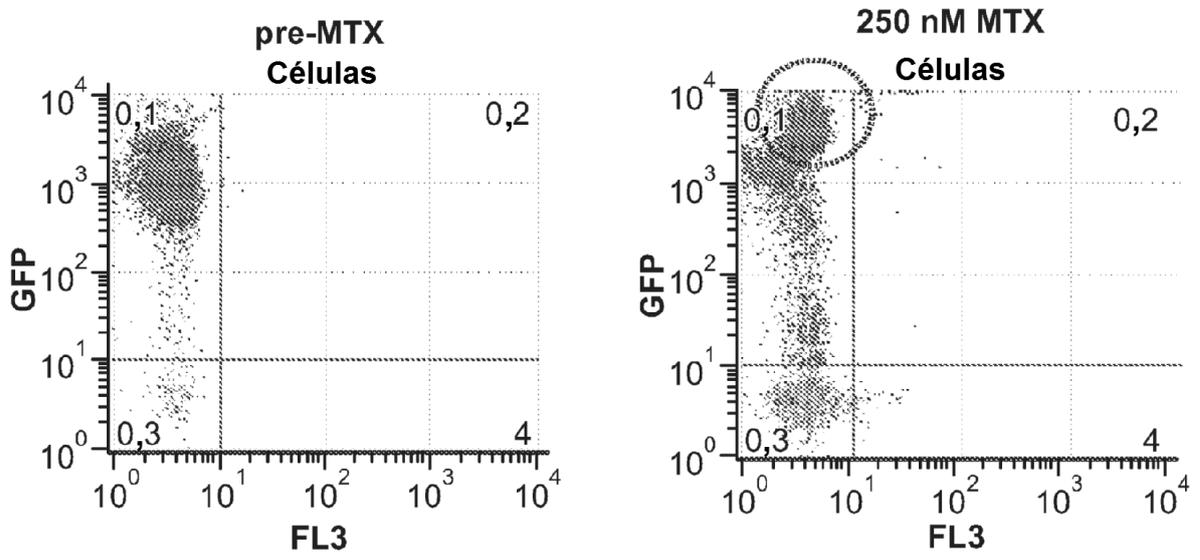


FIG. 10A

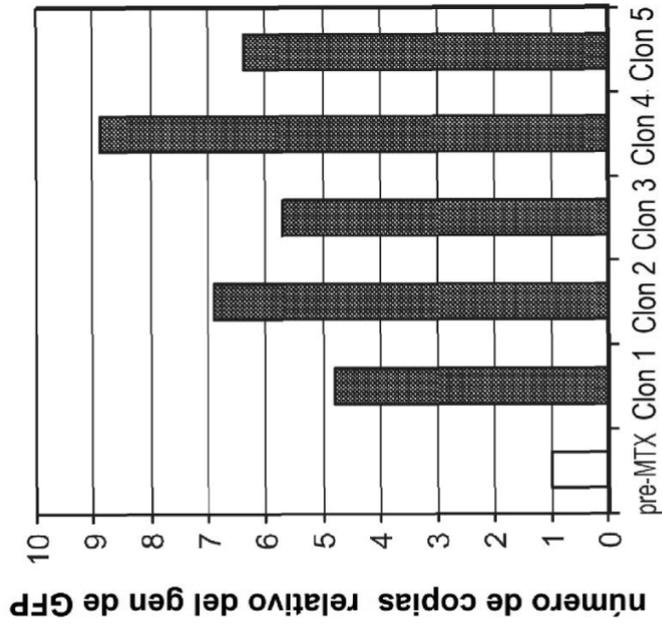


FIG. 10C

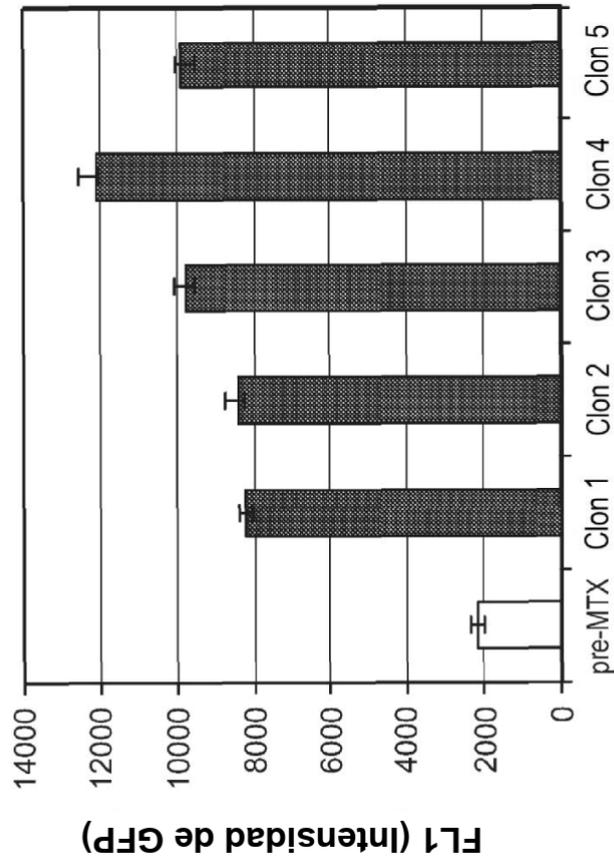


FIG. 10B

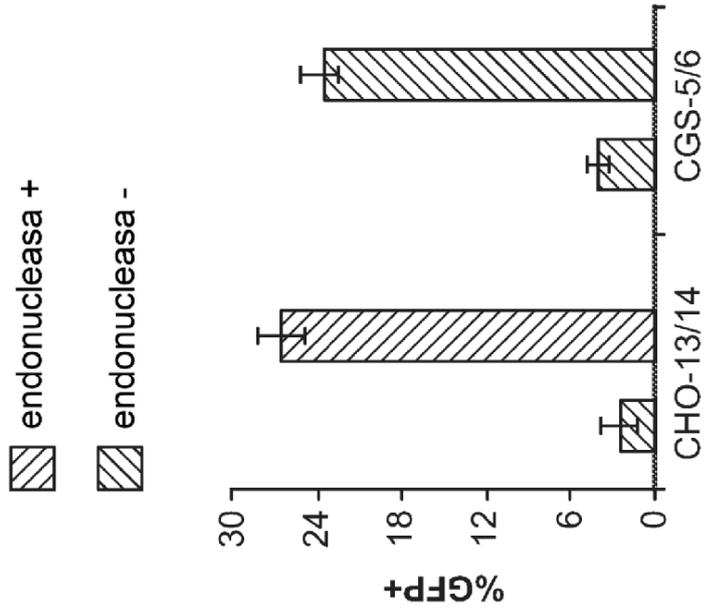


FIG. 11B

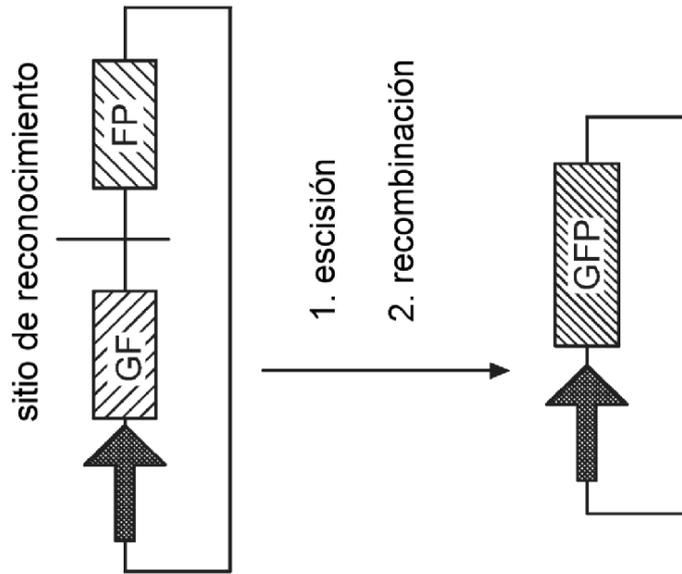


FIG. 11A

sitio de CGS-5/6

TS GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGTGTAAACGGATAAATGGACATGGTGAGCAACCAGCACCCC
 3d2 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGTGT----G-TAATGGACATGGTGAGCAACCAGCACCCC
 6d4 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGTGTAA-CGGATAATGGACATGGTGAGCAACCAGCACCCC
 6g5 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTC---AAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 3b7 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGT--AAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 3d11 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGT--AAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 3e5 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGT--AAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 6f10 GATGCTTTATTCCTAGAGACCAAATTAAGGCCT-GTGTAAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 6a2 GATGCTTTATTCCTAGAGACCAAATTAAGGCATTCGTGTAAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 6b3 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCATGTAAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 6c9 GATGCTTTATTCCTAGAGACCAAATTAAGGCCTCGTGTAAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 6f12 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGCGTAAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 6h7 GATGCTTTATTCCTAGAGACCAAATTAAGGCACACCGTGTAAACGGATAATGGACATGGTGAGCAACCAGCACCCC
 6H8 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGTGTGTAAACGGATAATGGACATGGTGAGCAACCAGCAC
 6d10 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGTGTGTAAACGGATAATGGACATGGTGAGCAACCAGCAC
 6d7 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGTGTGTAAACGGATAATGGACATGGTGAGCAACCAGCAC
 3g8 GATGCTTTATTCCTAGAGACCAAATTAAGGCACCTCGTGTGTAAACGGATAATGGACATGGTGAGCAACCAGCAC
 3a9 GATGCTTTATTCCTAGAGACCAAATTAAGGCA----TGGTGAGTAACC-----GAGCAACCAGCAC

FIG. 11C