

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 746 529**

51 Int. Cl.:

**C12N 15/113** (2010.01)

**A61K 48/00** (2006.01)

**C12N 15/85** (2006.01)

**C12N 15/867** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.09.2015 PCT/US2015/048698**

87 Fecha y número de publicación internacional: **10.03.2016 WO16037138**

96 Fecha de presentación y número de la solicitud europea: **04.09.2015 E 15837740 (8)**

97 Fecha y número de publicación de la concesión europea: **07.08.2019 EP 3189143**

54 Título: **Terapia génica de globina para tratar hemoglobinopatías**

30 Prioridad:

**04.09.2014 US 201462045997 P**

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

**06.03.2020**

73 Titular/es:

**MEMORIAL SLOAN-KETTERING CANCER  
CENTER (50.0%)  
1275 York Avenue  
New York, NY 10065, US y  
UNIVERSITY OF WASHINGTON (50.0%)**

72 Inventor/es:

**SADELAIN, MICHEL;  
RIVIERE, ISABELLE;  
MANSILLA-SOTO, JORGE;  
WANG, XIUYAN;  
STAMATOYANNOPOULOS, GEORGE;  
STAMATOYANNOPOULOS, JOHN y  
LIU, MINGDONG**

74 Agente/Representante:

**ARIAS SANZ, Juan**

ES 2 746 529 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

Terapia génica de globina para tratar hemoglobinopatías

**Referencia cruzada a solicitudes relacionadas**

5 Esta solicitud reivindica la prioridad de la solicitud provisional estadounidense n.º 62/045.997, presentada el 4 de septiembre de 2014.

**Información sobre subvenciones**

Esta invención se realizó con apoyo del gobierno según la subvención n.º HL053750 del National Heart, Lung and Blood Institute. El gobierno tiene ciertos derechos en la invención.

**Introducción**

10 El contenido dado a conocer en el presente documento proporciona casetes de expresión y vectores que comprenden tales casetes de expresión que expresan una proteína globina, por ejemplo, una proteína  $\beta$ -globina humana. El contenido dado a conocer en el presente documento proporciona además casetes de expresión que comprenden un gen de globina o una parte funcional del mismo operativamente unido a una región de control de locus (LCR) de  $\beta$ -globina que comprende una pluralidad de sitios hipersensibles a ADNasa I. Los casetes de expresión del contenido dado a conocer en el presente documento comprenden uno o más aislantes que contrarrestan el efecto de elementos potenciadores. Los aislantes dados a conocer en el presente documento no tienen un impacto sustancialmente adverso sobre el título de un vector que comprende los casetes de expresión dados a conocer en el presente documento. Los casetes de expresión y vectores pueden usarse para tratar una hemoglobinopatía, por ejemplo,  $\beta$ -talasemia, y anemia de células falciformes.

**Antecedentes**

La  $\beta$ -talasemia y la anemia de células falciformes son anemias congénitas graves que están provocadas por una producción defectuosa de la cadena  $\beta$  de la hemoglobina. En la  $\beta$ -talasemia, el déficit de cadena  $\beta$  conduce a la precipitación intracelular de cadenas de  $\alpha$ -globina en exceso, provocando una eritropoyesis ineficaz y anemia hemolítica (Weatherall y Clegg (1981), Stamatoyannopoulos *et al.*, (1994), Weatherall (2001), Steinberg (2001)). En las formas más graves encontradas en homocigotos o heterocigotos compuestos, la anemia es letal dentro del plazo de los primeros años de vida en ausencia de cualquier tratamiento (Cooley y Lee (1925)). Se necesita terapia de transfusión durante toda la vida para corregir la anemia, suprimir la eritropoyesis ineficaz e inhibir la absorción de hierro gastrointestinal (Weatherall y Clegg (1981), Stamatoyannopoulos *et al.* (1994), Weatherall (2001), Steinberg (2001)). Sin embargo, la propia terapia de transfusión conduce a sobrecarga de hierro, la cual es letal si no se trata. La prevención y el tratamiento de la sobrecarga de hierro son los objetivos principales de la gestión de pacientes actual (Giardina (2001)). El único tratamiento curativo actual para curar la  $\beta$ -talasemia es proporcionar precursores eritroides que albergan genes de globina normales mediante trasplante de médula ósea (BMT) alogénico (Giardini y Lucarelli (1994), Boulad *et al.* (1998), Lucarelli *et al.* (1999), Tisdale y Sadelain (2001)).

En la anemia de células falciformes, la cadena  $\beta$  de hemoglobina está mutada en la posición de aminoácido 6 (Glu  $\rightarrow$  Val), conduciendo a la síntesis de  $\beta^S$  en lugar de la cadena  $\beta^A$  normal (Steinberg (2001), Pauling *et al.* (1949)). La hemoglobina resultante, HbS, provoca la destrucción acelerada de glóbulos rojos, hiperplasia eritroide y "crisis" vasoclusivas dolorosas (Steinberg (2001)). La vasoclusión puede dañar órganos, provocando eventualmente discapacidades a largo plazo (por ejemplo tras accidente cerebrovascular o necrosis ósea), y algunas veces la muerte repentina. Aunque es un trastorno muy grave, el transcurso de la enfermedad de células falciformes normalmente es impredecible (Steinberg (2001)). Aumentando la producción de hemoglobina fetal (Swank y Stamatoyannopoulos (1998)) y suprimiendo la hematopoyesis, la hidroxiurea puede producir un beneficio clínico medible (Platt *et al.* (1984)), Charache *et al.* (1992), Atweh y Loukopoulos (2001)). Dado que la hidroxiurea es un agente citotóxico, existe una gran necesidad de fármacos alternativos menos tóxicos para inducir la expresión del gen de  $\gamma$ -globina (Perrine *et al.* (2005), Stamatoyannopoulos (2005)). Al igual que para la  $\beta$ -talasemia, en la actualidad el trasplante de médula ósea (BMT) alogénico es la única terapia curativa para la enfermedad de células falciformes (Tisdale y Sadelain (2001), Vermilen *et al.* (1998), Luzzatto y Goodfellow (1989)).

Sin embargo, BMT no está disponible como opción terapéutica para la mayoría de los pacientes que padecen  $\beta$ -talasemia o enfermedad de células falciformes, debido a la falta de un donante de médula ósea de HLA coincidente para la mayoría de los individuos. Además, aunque es potencialmente curativo, el BMT alogénico no carece de complicaciones. El trasplante seguro requiere la identificación de un donante histocompatible para minimizar los riesgos de rechazo de injerto y enfermedad de injerto contra huésped (Tisdale y Sadelain (2001), Vermilen *et al.* (1998), Luzzatto y Goodfellow (1989)). Debido a los riesgos mayores asociados con trasplantes coincidentes no relacionados o no coincidentes, la mayoría de los pacientes tienen que conformarse con terapia de transfusión durante toda la vida, lo cual no corrige la eritropoyesis ineficaz y empeora la acumulación de hierro sistémico. Además, a pesar de la mejora considerable en la expectativa de vida en las últimas décadas (Borgna-Pignatti *et al.* (2004), Telfer *et al.* (2009), Ladis *et al.* (2011)), sigue existiendo el riesgo de que surjan algunas complicaciones

graves a largo plazo a partir de infecciones virales, toxicidad por hierro y cirrosis hepática (Mancuso *et al.* (2006)). Estos riesgos médicos, junto con el coste socioeconómico de la  $\beta$ -talasemia crónica, enfatizan la necesidad de terapias seguras, eficaces y curativas.

5 Los únicos medios para curar en vez de tratar la  $\beta$ -talasemia grave son proporcionar a los pacientes células madre hematopoyéticas (HSC) sanas. Las HSC dan normalmente lugar a todos los tipos de células sanguíneas, incluyendo 20 mil millones de RBC al día en adultos. Las HSC pueden recogerse a partir de un donante con genes de  $\beta$ -globina silvestre para proporcionar glóbulos rojos (RBC) de larga duración con un contenido normal en hemoglobina. Alternativamente, pueden corregirse genéticamente las HSC del propio paciente, lo cual resuelve la búsqueda de un 10 donante y a la vez elimina los riesgos de enfermedad de injerto contra huésped y rechazo de injerto asociados con BMT alogénico (Sadelain (1997), Sadelain *et al.* (2007)). La transferencia de genes de globina tiene el objetivo de restaurar la capacidad de las células madre formadoras de sangre del propio sujeto con  $\beta$ -talasemia para generar RBC con un contenido en hemoglobina suficiente (Sadelain *et al.* (2007), Persons y Tisdale (2004), Sadelain (2006)). El documento WO2014043131 describe vectores lentivirales para terapia génica de células madre de enfermedad de células falciformes.

15 El objetivo en pacientes con anemia de células falciformes es prevenir la falciformación, lo cual puede lograrse diluyendo la HbS endógena con una Hb que no produce falciformación que incorpora la cadena de globina codificada por vector. Las HSC del propio paciente son las células que tienen que modificarse genéticamente para garantizar beneficios terapéuticos de larga duración y lograr una terapia curativa basada en células madre.

20 La implementación de transferencia de genes de globina para el tratamiento de  $\beta$ -talasemia grave y anemia de células falciformes requiere la introducción eficaz de un gen de  $\beta$ -globina o de globina de tipo  $\beta$  humana regulado en HSC. El gen de  $\beta$ -globina (o variante de tipo  $\beta$ ) debe expresarse de manera específica de eritroides y a alto nivel, especialmente para el tratamiento de beta-cero-talasemias dependientes de transfusión.

25 Los vectores de globina desarrollados hasta la fecha presentan inconvenientes que pueden limitar o incluso evitar su uso seguro en pacientes con talasemia y células falciformes. Algunos de los componentes de región de control de locus (LCR) de  $\beta$ -globina contenidos en los vectores, en particular sitio hipersensible-2 (HS2) a ADNasa I, pueden tener actividad no eritroide, exponiendo a los pacientes al riesgo de oncogénesis por inserción tal como se observa con vectores de expresión no específicos. Además, el uso de grandes segmentos de LCR puede ser perjudicial para la producción de vectores de alto título y la transducción eficaz de HSC de pacientes. Por consiguiente, existe una 30 necesidad de casetes de expresión de globina novedosos que permitan la expresión terapéutica de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana) de manera específica de eritroides y específica de estadio de diferenciación con un riesgo mínimo de oncogénesis por inserción, y que permitan una transducción a alto nivel, mejorando por tanto su seguridad cuando se usan en el tratamiento de pacientes con talasemia y células falciformes.

### Sumario de la invención

35 La invención es tal como se define en las reivindicaciones.

El contenido dado a conocer en el presente documento proporciona de manera general aislantes de bloqueo de potenciador, y determinados aislantes presentan además actividad aislante de barrera. El contenido dado a conocer en el presente documento también proporciona casetes de expresión que comprenden uno o más aislantes y permiten la expresión de un gen de globina (por ejemplo, un gen de  $\beta$ -globina humana). También se proporcionan 40 vectores que comprenden tales casetes de expresión, células transducidas con tales casetes de expresión o tales vectores, y usos de tales casetes de expresión para tratar hemoglobinopatías (por ejemplo,  $\beta$ -talasemia y anemia de células falciformes).

En determinados aspectos de la divulgación, el contenido dado a conocer en el presente documento proporciona un 45 aislante que comprende la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, por ejemplo, pero sin limitarse a, un aislante que comprende SEQ ID NO: 24 o SEQ ID NO: 25, tal como un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1 (y véase más adelante). El contenido dado a conocer en el presente documento también proporciona casetes de expresión que comprenden al menos un aislante que comprende la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, por ejemplo, pero sin limitarse a, un aislante que comprende SEQ ID NO: 24 o SEQ ID NO: 25, tal como un aislante que tiene la secuencia de nucleótidos expuesta 50 en SEQ ID NO: 1. En una realización no limitativa, un casete de expresión comprende al menos un aislante que comprende la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, por ejemplo, pero sin limitarse a, un aislante que comprende SEQ ID NO: 24 o SEQ ID NO: 25, tal como un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1, y un gen de globina o una parte funcional del mismo operativamente unido a una región de control de locus (LCR) de  $\beta$ -globina. En determinadas realizaciones, la LCR de  $\beta$ -globina no comprende una región de sitio hipersensible 2 (HS2) a ADNasa I. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una secuencia central de HS2. En una realización no limitativa, la secuencia central de HS2 55 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 20. En una realización no limitativa, la secuencia central de HS2 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 21. En determinadas realizaciones, la LCR de  $\beta$ -

globina no comprende una región de HS2 que conserva la actividad potenciadora de HS2. En una realización no limitativa, la LCR de  $\beta$ -globina comprende una región de sitio hipersensible 1 (HS1) a ADNasa I, una región de sitio hipersensible 3 (HS3) a ADNasa I, y una región de sitio hipersensible 4 (HS4) a ADNasa I. En determinadas realizaciones, la región de HS3 está posicionada entre la región HS1 y la de HS4.

5 En determinadas realizaciones, la región de HS1 tiene aproximadamente 1,1 kb pb de longitud. En una realización no limitativa, la región de HS1 tiene entre aproximadamente 500 pb y aproximadamente 1000 pb de longitud. En una realización no limitativa, la región de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 2. En determinadas realizaciones, la región de HS1 tiene aproximadamente 600 pb de longitud. En una realización no limitativa, la región de HS1 tiene 602 pb de longitud. En determinadas realizaciones, la región de HS1 tiene entre  
10 aproximadamente 500 y aproximadamente 600 pb de longitud. En una realización no limitativa, la región de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 3. En determinadas realizaciones, la región de HS1 tiene aproximadamente 490 pb de longitud. En una realización no limitativa, la región de HS1 tiene 489 pb de longitud. En una realización no limitativa, la región de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 4. En una realización no limitativa, la LCR de  $\beta$ -globina comprende una región de HS1 que tiene una secuencia de nucleótidos  
15 expuesta en SEQ ID NO: 2, una región de HS3 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 6, y la región LCR de  $\beta$ -globina no comprende una región de HS2. En una realización no limitativa, la región LCR de  $\beta$ -globina comprende una región de HS1 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 3, una región de HS3 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene una secuencia de  
20 nucleótidos expuesta en SEQ ID NO: 8, y la LCR de  $\beta$ -globina no comprende una región de HS2. En una realización no limitativa, la LCR de  $\beta$ -globina comprende una región de HS1 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 4, una región de HS3 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 8, y la LCR de  $\beta$ -globina no comprende una región de HS2.

25 En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 y/o no comprende una región de HS2, y la LCR de  $\beta$ -globina no comprende una secuencia central de HS2. En determinadas realizaciones, la LCR de  $\beta$ -globina no comprende una secuencia central de HS1. En una realización no limitativa, la secuencia central de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 22. En una realización no limitativa, la secuencia central de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 23. En determinadas realizaciones, la LCR de  $\beta$ -globina no comprende una región de HS1 que conserva la función de HS1. En determinadas realizaciones, la LCR de  $\beta$ -globina comprende una región de HS3 y una región de HS4 y no  
30 comprende una secuencia central de HS1. En determinadas realizaciones, la región de HS3 está posicionada entre un gen de globina o parte funcional del mismo y la región de HS4. En determinadas realizaciones, la región de HS3 tiene entre aproximadamente 200 y aproximadamente 1400 pb de longitud, por ejemplo, entre aproximadamente 1300 y 1400 pb de longitud. En determinadas realizaciones, la región de HS3 tiene aproximadamente 1300 pb de longitud. En una realización no limitativa, la región de HS3 tiene 1301 pb de longitud. En una realización no limitativa, la región de HS3 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5. En determinadas realizaciones, la región de HS4 tiene entre aproximadamente 200 y aproximadamente 1200 pb de longitud, por ejemplo, entre aproximadamente 400 y 1100 pb de longitud. En determinadas realizaciones, la región de HS4 tiene  
35 aproximadamente 1,1 kb de longitud. En una realización no limitativa, la región de HS4 tiene 1065 pb de longitud. En una realización no limitativa, la región de HS4 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6. En una realización no limitativa, la región de HS4 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7. En determinadas realizaciones, la región de HS4 tiene aproximadamente 450 pb de longitud. En una realización no limitativa, la región de HS4 tiene 446 pb de longitud. En una realización no limitativa, la región de HS4 tiene la  
40 secuencia de nucleótidos expuesta en SEQ ID NO: 8. En una realización no limitativa, la región LCR de  $\beta$ -globina comprende una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5 y una región de HS4 que tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 6, y la región LCR de  $\beta$ -globina no comprende una región de HS1 o una región de HS2.

Alternativamente, la región LCR de  $\beta$ -globina puede comprender una región de HS2, una región de HS3, y una  
50 región de HS4. En determinadas realizaciones, la región de HS2 tiene entre aproximadamente 400 y aproximadamente 1000 pb de longitud, por ejemplo, entre aproximadamente 800 y 900 pb de longitud. En determinadas realizaciones, la región de HS2 tiene aproximadamente 860 pb de longitud. En una realización no limitativa, la región de HS2 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 9. En determinadas realizaciones, la región de HS3 tiene aproximadamente 1300 pb de longitud. En una realización no limitativa, la región de HS3 tiene 1301 pb de longitud. En una realización no limitativa, la región de HS3 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5. En determinadas realizaciones, la región de HS4 tiene aproximadamente 1,1 kb de longitud. En una realización no limitativa, la región de HS4 tiene 1065 pb de longitud. En una realización no limitativa, la región de HS4 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7. En una realización no limitativa, la región LCR de  $\beta$ -globina comprende una región de HS2 que tiene la secuencia de nucleótidos expuesta  
55 en SEQ ID NO: 9, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7. Adicionalmente, la región LCR de  $\beta$ -globina puede comprender además una región de HS1.

- 5 En determinadas realizaciones, el gen de globina se selecciona del grupo que consiste en gen de  $\beta$ -globina, gen de  $\gamma$ -globina y gen de  $\delta$ -globina. En una realización no limitativa, el gen de globina es gen de  $\beta$ -globina humana. En realizaciones no limitativas, el gen de  $\beta$ -globina humana se selecciona del grupo que consiste en un gen de  $\beta$ -globina humana silvestre, un gen de  $\beta$ -globina humana delecionado que comprende una o más deleciones de secuencias de intrones, y un gen de  $\beta$ -globina humana mutado que codifica para al menos un residuo de aminoácido anti-falciformación. En una realización no limitativa, el gen de  $\beta$ -globina humana es gen de  $\beta^A$ -globina humana que codifica para una mutación de treonina a glutamina en el codón 87 ( $\beta^{A-T87Q}$ ).
- 10 En determinadas realizaciones, el casete de expresión comprende un aislante que comprende la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, por ejemplo, pero sin limitarse a, un aislante que comprende SEQ ID NO: 24 o SEQ ID NO: 25, tal como un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1. En determinadas realizaciones, el casete de expresión comprende dos aislantes, comprendiendo cada uno la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, por ejemplo, pero sin limitarse a, en el que uno o ambos aislantes comprenden SEQ ID NO: 24 o SEQ ID NO: 25 y/o tienen la secuencia de nucleótidos expuesta en SEQ ID NO: 1.
- 15 En determinadas realizaciones, el casete de expresión comprende además un promotor de  $\beta$ -globina. En determinadas realizaciones, el promotor de  $\beta$ -globina está posicionado entre el gen de globina o parte funcional del mismo y la región LCR de  $\beta$ -globina. En determinadas realizaciones, el promotor de  $\beta$ -globina tiene entre aproximadamente 200 y aproximadamente 700 pb de longitud. En una realización no limitativa, el promotor de  $\beta$ -globina es un promotor de  $\beta$ -globina humana que tiene aproximadamente 613 pb de longitud. En una realización no limitativa, el promotor de  $\beta$ -globina humana tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 10. En otra realización no limitativa, el promotor de  $\beta$ -globina es un promotor de  $\beta$ -globina humana que tiene aproximadamente 265 pb de longitud. En una realización no limitativa, el promotor de  $\beta$ -globina humana  $\beta$  tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 11.
- 20 En determinadas realizaciones, el casete de expresión comprende además un potenciador en 3' de  $\beta$ -globina humana. En determinadas realizaciones, el potenciador en 3' de  $\beta$ -globina humana está posicionado aguas arriba del gen de globina o parte funcional del mismo. En determinadas realizaciones, el potenciador en 3' de  $\beta$ -globina tiene entre aproximadamente 700 y aproximadamente 900 pb de longitud, por ejemplo, entre aproximadamente 800 y 900 pb de longitud. En una realización no limitativa, el potenciador en 3' de  $\beta$ -globina humana tiene aproximadamente 879 pb de longitud. En una realización no limitativa, el potenciador en 3' de  $\beta$ -globina humana tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 12.
- 25 En determinadas realizaciones, el casete de expresión comprende además al menos un potenciador específico de eritroides. En determinadas realizaciones, el al menos un potenciador específico de eritroides está posicionado entre el gen de globina o parte funcional del mismo y la región LCR de  $\beta$ -globina. En determinadas realizaciones, el al menos un potenciador específico de eritroides tiene una secuencia de nucleótidos seleccionada del grupo que consiste en SEQ ID NOS: 13, 14, 15, 16 y 17. En determinadas realizaciones, el al menos un potenciador específico de eritroides tiene entre aproximadamente 100 y aproximadamente 200 pb de longitud. En determinadas realizaciones, el casete de expresión comprende uno, dos o tres potenciadores específicos de eritroides.
- 30 En determinadas realizaciones, el casete de expresión permite la expresión del gen de globina o parte funcional del mismo en un mamífero. En una realización no limitativa, el casete de expresión permite la expresión de un gen de  $\beta$ -globina humana. En determinadas realizaciones, la expresión del gen de globina o parte funcional del mismo se restringe a tejido eritroide.
- 35 El contenido dado a conocer en el presente documento también proporciona vectores recombinantes que comprenden los casetes de expresión descritos anteriormente. En determinadas realizaciones, el vector recombinante es un vector retroviral. En una realización no limitativa, el vector retroviral es un vector de lentivirus. En determinadas realizaciones, el casete de expresión comprendido en el vector recombinante comprende un aislante. En determinadas realizaciones, el vector recombinante comprende además un elemento post-regulador de la hepatitis de la marmota (WPRE) en la repetición terminal larga (LTR) en 3' del vector. En determinadas realizaciones, el vector recombinante comprende además una señal de poliadenilación de hormona de crecimiento bovina en la repetición terminal larga (LTR) en 3' del vector.
- 40 Además, el contenido dado a conocer en el presente documento proporciona sistemas de nucleasa modificados por ingeniería o que no se producen de manera natural que comprenden los casetes de expresión descritos anteriormente. En determinadas realizaciones, la nucleasa se selecciona del grupo que consiste en una nucleasa de dedos de zinc modificada por ingeniería o que no se produce de manera natural (ZFN), una meganucleasa modificada por ingeniería o que no se produce de manera natural, y una nucleasa efectora de tipo activador de la transcripción (TALEN) modificada por ingeniería o que no se produce de manera natural. En determinadas realizaciones, la nucleasa comprende un dominio de unión a ADN y un dominio de escisión de nucleasa. En determinadas realizaciones, la nucleasa se une a un sitio de locus seguro genómico. En determinadas realizaciones, la nucleasa genera una rotura de cadena doble (DSB) en el sitio de locus seguro genómico. En determinadas realizaciones, el casete de expresión comprendido en el sistema de nucleasa comprende dos del aislante que tiene
- 45
- 50
- 55

la secuencia de nucleótidos expuesta en SEQ ID NO: 1. En determinadas realizaciones, la nucleasa permite la administración dirigida del casete de expresión. El contenido dado a conocer en el presente documento también proporciona polinucleótidos que codifican para los sistemas de nucleasa descritos anteriormente, y vectores que comprenden los polinucleótidos. En una realización no limitativa, el vector es un vector lentiviral.

5 Además, el contenido dado a conocer en el presente documento proporciona sistemas de CRISPR-Cas modificados por ingeniería o que no se producen de manera natural que comprenden los casetes de expresión descritos anteriormente. En determinadas realizaciones, el sistema de CRISPR-Cas comprende una nucleasa de CRISPR-Cas y ARN guía individual. En determinadas realizaciones, el sistema de CRISPR-Cas se une a un sitio de locus seguro genómico. En determinadas realizaciones, el sistema de CRISPR-Cas genera una rotura de cadena doble (DSB) en el sitio de locus seguro genómico. En determinadas realizaciones, el casete de expresión comprendido en el sistema de CRISPR-Cas comprende dos del aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1. En determinadas realizaciones, el CRISPR-Cas permite la administración dirigida del casete de expresión. El contenido dado a conocer en el presente documento también proporciona polinucleótidos que codifican para los sistemas de CRISPR-Cas descritos anteriormente, y vectores que comprenden los polinucleótidos. En una realización no limitativa, el vector es un vector lentiviral.

En algunas realizaciones, el sitio de locus seguro genómico es un sitio de locus seguro genómico extragenómico. En determinadas realizaciones, el sitio de locus seguro genómico está ubicado en el cromosoma 1. En algunas realizaciones, el locus seguro genómico cumple la totalidad de los cinco criterios siguientes: (i) distancia de al menos 50 kb desde el extremo 5' de cualquier gen (por ejemplo, desde el extremo 5' del gen), (ii) distancia de al menos 300 kb desde cualquier gen relacionado con cáncer, (iii) dentro de una estructura de cromatina abierta/accesible (medido mediante escisión de ADN con nucleasas naturales o modificadas por ingeniería), (iv) ubicación fuera de una unidad de transcripción génica y (v) ubicación fuera de regiones ultraconservadas (UCR), microARN o ARN largo no codificante del genoma humano.

Adicionalmente, el contenido dado a conocer en el presente documento proporciona células transducidas con los casetes de expresión descritos anteriormente, células transducidas con los vectores recombinantes descritos anteriormente, células transducidas con las nucleasas descritas anteriormente, células transducidas con los sistemas de CRISPR-Cas descritos anteriormente. Además, el contenido dado a conocer en el presente documento proporciona células transducidas con los vectores descritos anteriormente. En determinadas realizaciones, la célula se selecciona del grupo que consiste en una célula madre hematopoyética, una célula madre embrionaria, una célula madre pluripotente inducida y una célula endotelial hemogénica. En una realización no limitativa, la célula madre hematopoyética es una célula madre hematopoyética CD34<sup>+</sup>. En determinadas realizaciones, la célula se transduce *ex vivo*.

También se proporcionan composiciones farmacéuticas que comprenden una cantidad eficaz de las células descritas anteriormente y un portador farmacéuticamente aceptable. El contenido dado a conocer en el presente documento también proporciona composiciones farmacéuticas para tratar una hemoglobinopatía que comprenden una cantidad eficaz de las células descritas anteriormente y un portador farmacéuticamente aceptable.

Además, el contenido dado a conocer en el presente documento proporciona kits para tratar una hemoglobinopatía que comprenden las células descritas anteriormente. En determinadas realizaciones, los kits comprenden además instrucciones escritas para usar la célula para tratar a un sujeto que tiene una hemoglobinopatía.

40 Además, el contenido dado a conocer en el presente documento proporciona métodos de tratamiento de una hemoglobinopatía en un sujeto, que comprenden administrar una cantidad eficaz de las células descritas anteriormente al sujeto, restaurando así la capacidad del sujeto para producir glóbulos rojos que contienen hemoglobina normal. La invención proporciona las células descritas anteriormente o composiciones para su uso en el tratamiento de una hemoglobinopatía en un sujeto, o para su uso en la restauración de la capacidad de un sujeto para producir glóbulos rojos que contienen hemoglobina normal.

En determinadas realizaciones, se produce un nivel terapéuticamente relevante de hemoglobina en el sujeto tras la administración de la célula al sujeto. En determinadas realizaciones, el método comprende administrar una cantidad eficaz de la célula transducida con el vector recombinante descrito anteriormente. En algunas realizaciones, el número de copias de vector del vector recombinante en la célula que proporciona el nivel terapéuticamente relevante de hemoglobina en el sujeto es de aproximadamente 0,5-2 números de copias de vector por célula. En determinadas realizaciones, el método corrige la eritropoyesis ineficaz en el sujeto. En determinadas realizaciones, el método no provoca el riesgo de enfermedad de injerto contra huésped en el sujeto. En determinadas realizaciones, el método no comprende administrar un agente inmunosupresor. En determinadas realizaciones, la célula se selecciona del grupo que consiste en una célula madre hematopoyética, una célula madre embrionaria, una célula madre pluripotente inducida, y una célula endotelial hemogénica. En una realización no limitativa, el sujeto es un ser humano. En determinadas realizaciones, la célula procede del sujeto. En una realización no limitativa, la célula procede de la médula ósea del sujeto.

Según el contenido dado a conocer en el presente documento, la hemoglobinopatía puede seleccionarse del grupo que consiste en enfermedad de hemoglobina C, enfermedad de células falciformes (SCD) de hemoglobina, anemia

de células falciformes, anemia hereditaria, talasemia,  $\beta$ -talasemia, talasemia mayor, talasemia intermedia,  $\alpha$ -talasemia y enfermedad de hemoglobina H. En una realización no limitativa, la hemoglobinopatía es  $\beta$ -talasemia. En otra realización no limitativa, la hemoglobinopatía es anemia de células falciformes.

**Breve descripción de las figuras**

5 La siguiente descripción detallada, facilitada a modo de ejemplo, pero que no se pretende que limite la invención a las realizaciones específicas descritas, puede entenderse junto con los dibujos adjuntos.

La figura 1 representa un vector recombinante que comprende un casete de expresión según una realización no limitativa del contenido dado a conocer en el presente documento.

10 La figura 2 representa un vector recombinante un casete de expresión según una realización no limitativa del contenido dado a conocer en el presente documento.

La figura 3 representa un vector recombinante un casete de expresión según una realización no limitativa del contenido dado a conocer en el presente documento.

La figura 4 representa un vector recombinante un casete de expresión según una realización no limitativa del contenido dado a conocer en el presente documento.

15 Las figuras 5A-C representan la genotoxicidad del aislante A1. (A) demuestra el ensayo de genotoxicidad de vector gamma-retroviral usado. (B) indica la supervivencia aumentada de ratones que reciben células 32D transducidas con vector gamma-retroviral aislado. También se indican los resultados obtenidos con cHS4 y con el control sin aislante. (C) muestra que el aislante A1 disminuyó el riesgo de genotoxicidad.

20 La figura 6 representa la expresión de cadena  $\beta$  normalizada en ratones Hbb<sup>th3/+</sup> con talasemia 8 y 44 semanas tras el tratamiento.

La figura 7 representa la evaluación de actividad potenciadora en células K562 no eritroides.

La figura 8 representa los potenciadores específicos de eritroides según determinadas realizaciones del contenido dado a conocer en el presente documento.

25 La figura 9 representa los potenciadores específicos de eritroides según determinadas realizaciones del contenido dado a conocer en el presente documento.

Las figuras 10A-B representan diversos vectores recombinantes que comprenden los casetes de expresión dados a conocer en el presente documento.

La figura 11 representa el título de los vectores recombinantes que comprenden los casetes de expresión dados a conocer en el presente documento.

30 La figura 12 representa el título de los vectores recombinantes que comprenden los casetes de expresión dados a conocer en el presente documento.

**Descripción detallada de la invención**

La invención es tal como se define en las reivindicaciones.

35 El contenido dado a conocer en el presente documento proporciona de manera general casetes de expresión que permiten la expresión de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana). En un ejemplo no limitativo, el casete de expresión comprende al menos un aislante que comprende la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, por ejemplo, pero sin limitarse a, un aislante que comprende SEQ ID NO: 24 o SEQ ID NO: 25, tal como un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1 y un gen de globina o una parte funcional del mismo operativamente unido a una región de región de control de locus (LCR) de  $\beta$ -globina.

40 La expresión del gen de globina inducida por los casetes de expresión dados a conocer en el presente documento es específica de eritroides, específica de estadio de diferenciación, de alto nivel y sostenida. El contenido dado a conocer en el presente documento también proporciona vectores recombinantes, nucleasas modificadas por ingeniería o que no se producen de manera natural, y sistemas de CRISPR-Cas modificados por ingeniería o que no se producen de manera natural que comprenden tales casetes de expresión, y células transducidas con tales casetes de expresión, vectores recombinantes, nucleasas y sistemas de CRISPR-Cas. Los casetes de expresión dados a conocer en el presente documento y vectores que comprenden los mismos proporcionan una terapia de transferencia génica segura dado que se logra la expresión transgénica terapéutica (por ejemplo, se produce un nivel terapéuticamente relevante de hemoglobina) con un número de copias de vector bajo por célula (por ejemplo, 0,5-2, 1-2 o incluso 0,5-1). Además, el contenido dado a conocer en el presente documento proporciona métodos de

45 uso de tales células transducidas para tratar una hemoglobinopatía (por ejemplo,  $\beta$ -talasemia y anemia de células falciformes).

## I. Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el significado entendido habitualmente por un experto en la técnica a la que pertenece esta invención. Las siguientes referencias proporcionan a un experto una definición general de muchos de los términos usados en esta invención:

5 Singleton *et al.*, Dictionary of Microbiology and Molecular Biology (2ª ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5ª ed., R. Rieger *et al.* (eds.), Springer Verlag (1991); y Hale & Marham, The Harper Collins Dictionary of Biology (1991). Tal como se usan en el presente documento, los siguientes términos tienen los significados que se les asignan a continuación, a menos que se especifique lo contrario.

10 Tal como se usa en el presente documento, el término “casete de expresión” se refiere a un constructo de ácido nucleico, generado de manera recombinante o sintética, con una serie de elementos de ácido nucleico especificados, que permiten la transcripción de un ácido nucleico particular en una célula diana. El casete de expresión puede incorporarse en un plásmido, cromosoma, ADN mitocondrial, ADN de plástido, virus o región o

15 ácido nucleico. La parte de casete de expresión puede incluir un gen que va a transcribirse y elementos que controlan la expresión del gen (por ejemplo, un promotor).

Tal como se usa en el presente documento, el término “región de región de control de locus (LCR) de  $\beta$ -globina” se refiere a un polinucleótido compuesto por una o más regiones de sitio hipersensible (HS) a ADNasa I, incluyendo una región de HS1, una región de HS2, una región de HS3 y una región de HS4. Se ha publicado la estructura de muchas LCR de los genes de  $\beta$ -globina, por ejemplo, de ser humano (Li *et al.*, J. Biol. Chem. (1985); 260: 14.901; Li

20 *et al.*, Proc. Natl. Acad. Sci. (1990) 87: 8207); de ratón (Shehee *et al.*, J. Mol. Biol. (1989); 205: 41); de conejo (Margot *et al.*, J. Mol. Biol. (1989); 205: 15); y de cabra (Li, Q., *et al.*, Genomics (1991); 9: 488), cada una de las cuales se incorpora como referencia en el presente documento. En determinadas realizaciones, la región LCR de  $\beta$ -globina comprende una región de HS2 (por ejemplo, una región LCR de  $\beta$ -globina que comprende una región de HS2, una región de HS3 y una región de HS4; y una región LCR de  $\beta$ -globina que comprende una región de HS1,

25 una región de HS2, una región de HS3 y una región de HS4). En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 (por ejemplo, una región LCR de  $\beta$ -globina que comprende una región de HS1, una región de HS3, una región de HS4). En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 o una región de HS1 (por ejemplo, una región LCR de  $\beta$ -globina que comprende una región de HS3 y una región de HS4).

30 Tal como se usa en el presente documento, el término “recombinante” incluye la referencia a una célula o vector que se ha modificado mediante la introducción de un ácido nucleico heterólogo o que la célula se deriva de una célula modificada de ese modo. Por tanto, por ejemplo, las células recombinantes expresan genes que no se encuentran de forma idéntica dentro de la forma nativa (no recombinante) de la célula o expresan genes nativos que por lo demás se expresan de manera anómala, se expresan de manera insuficiente o no se expresan como resultado de

35 intervención humana deliberada o pueden tener una expresión reducida o eliminada de un gen nativo.

Tal como se usa en el presente documento, el término “globina” se refiere a una familia de proteínas que contienen hemo que participan en la unión a, y el transporte de, oxígeno. El término globina incluye subunidades de hemoglobinas de vertebrados e invertebrados, mioglobinas de vertebrados e invertebrados o mutantes de las mismas.

40 Tal como se usa en el presente documento, el término “silvestre” se refiere al gen, virus u organismo normal encontrado en la naturaleza sin ninguna mutación o modificación.

Los términos “polinucleótido”, “nucleótido”, “secuencia de nucleótidos”, “ácido nucleico” y “oligonucleótido” se usan de manera intercambiable. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sean desoxirribonucleótidos o ribonucleótidos, o análogos de los mismos. Los polinucleótidos pueden tener cualquier

45 estructura tridimensional, y pueden realizar cualquier función, conocida o desconocida. Los siguientes son ejemplos no limitativos de polinucleótidos: regiones codificantes o no codificantes de un gen o región de gen, loci (locus) definidos a partir de análisis de unión, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ARN de interferencia corto (ARNic), ARN de horquilla corta (RNAhc), micro-ARN (miARN), ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier

50 secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. Un polinucleótido puede comprender uno o más nucleótidos modificados, tales como nucleótidos metilados y análogos de nucleótidos. En realizaciones particulares, el contenido dado a conocer en el presente documento proporciona polinucleótidos que codifican para uno o más genes de globina o partes funcionales de los mismos. Si están presentes, pueden conferirse modificaciones a la estructura de nucleótido antes o después del ensamblaje del polímero. La secuencia

55 de nucleótidos puede estar interrumpida por componentes distintos de nucleótidos. Un polinucleótido puede modificarse adicionalmente tras la polimerización, tal como mediante conjugación con un componente de marcaje. No se necesita que tales polinucleótidos sean idénticos al 100% con una secuencia de ácido nucleico endógena, pero normalmente mostrarán una identidad sustancial. Los polinucleótidos que tienen “identidad sustancial” con respecto a una secuencia endógena pueden normalmente hibridarse con al menos una cadena de una molécula de

60 ácido nucleico de cadena doble. Por “hibridarse” quiere decirse emparejar para formar una molécula de cadena

doble entre secuencias de polinucleótido complementarias (por ejemplo, un gen descrito en el presente documento), o partes de las mismas, en diversas condiciones de rigurosidad. (Véase, por ejemplo, Wahl, G. M. y S. L. Berger (1987) *Methods Enzymol.* 152: 399; Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507).

5 Por ejemplo, concentración de sales rigurosa será habitualmente de menos de aproximadamente NaCl 750 mM y citrato de trisodio 75 mM, preferiblemente menos de aproximadamente NaCl 500 mM y citrato de trisodio 50 mM, y más preferiblemente menos de aproximadamente NaCl 250 mM y citrato de trisodio 25 mM. Puede obtenerse hibridación de baja rigurosidad en ausencia de disolvente orgánico, por ejemplo, formamida, mientras que puede obtenerse hibridación de alta rigurosidad en presencia de formamida a al menos aproximadamente el 35%, y más preferiblemente formamida a al menos aproximadamente el 50%. Las condiciones de temperatura rigurosas incluirán habitualmente temperaturas de al menos aproximadamente 30°C, más preferiblemente de al menos aproximadamente 37°C, y lo más preferiblemente de al menos aproximadamente 42°C. Los expertos en la técnica conocen bien parámetros adicionales variables, tales como tiempo de hibridación, concentración de detergente, por ejemplo, dodecilsulfato de sodio (SDS), y la inclusión o exclusión de ADN portador. Se logran diversos niveles de rigurosidad combinando estas diversas condiciones según se necesite. En una realización preferida, se producirá hibridación a 30°C en NaCl 750 mM, citrato de trisodio 75 mM, y SDS al 1%. En una realización más preferida, se producirá hibridación a 37°C en NaCl 500 mM, citrato de trisodio 50 mM, SDS al 1%, formamida al 35%, y 100 µg/ml de ADN de esperma de salmón desnaturalizado (ADNes). En la realización más preferida, se producirá hibridación a 42°C en NaCl 250 mM, citrato de trisodio 25 mM, SDS al 1%, formamida al 50%, y 200 µg/ml de ADNes. Variaciones útiles de estas condiciones resultarán fácilmente evidentes para los expertos en la técnica.

20 Para la mayoría de las aplicaciones, las etapas de lavado que siguen a la hibridación también variarán en cuanto a la rigurosidad. Las condiciones de rigurosidad de lavado pueden definirse por la concentración de sal y por la temperatura. Como anteriormente, la rigurosidad de lavado puede aumentarse disminuyendo la concentración de sal o aumentando la temperatura. Por ejemplo, la concentración de sal rigurosa para las etapas de lavado será preferiblemente de menos de aproximadamente NaCl 30 mM y citrato de trisodio 3 mM, y lo más preferiblemente menos de aproximadamente NaCl 15 mM y citrato de trisodio 1,5 mM. Las condiciones de temperatura rigurosas para las etapas de lavado incluirán habitualmente una temperatura de al menos aproximadamente 25°C, más preferiblemente de al menos aproximadamente 42°C, e incluso más preferiblemente de al menos aproximadamente 68°C. En una realización preferida, se producirán etapas de lavado a 25°C en NaCl 30 mM, citrato de trisodio 3 mM, y SDS al 0,1%. En una realización más preferida, se producirán etapas de lavado a 42°C en NaCl 15 mM, citrato de trisodio 1,5 mM, y SDS al 0,1%. En una realización más preferida, se producirán etapas de lavado a 68°C en NaCl 15 mM, citrato de trisodio 1,5 mM, y SDS al 0,1%. Variaciones adicionales de estas condiciones resultarán fácilmente evidentes para los expertos en la técnica. Los expertos en la técnica conocen bien técnicas de hibridación y se describen, por ejemplo, en Benton y Davis (*Science* 196: 180, 1977); Grunstein y Rogness (*Proc. Natl. Acad. Sci., USA* 72: 3961, 1975); Ausubel *et al.* (*Current Protocols in Molecular Biology*, Wiley Interscience, Nueva York, 2001); Berger y Kimmel (*Guide to Molecular Cloning Techniques*, 1987, Academic Press, Nueva York); y Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Nueva York.

Tal como se usa en el presente documento, los términos “polipéptido” y “proteína” se usan de manera intercambiable para hacer referencia a un polímero de residuos de aminoácido y a variantes y análogos sintéticos del mismo. Por tanto, estos términos se aplican a polímeros de aminoácidos en los que uno o más residuos de aminoácido son aminoácidos sintéticos que no se producen de manera natural, tales como un análogo químico de un aminoácido que se produce de manera natural correspondiente, así como a polímeros de aminoácidos que se producen de manera natural. Realizaciones particulares del contenido dado a conocer en el presente documento también incluyen “variantes” de polipéptidos. “Variante” de polipéptido se refiere a polipéptidos que se distinguen de un polipéptido de referencia mediante la adición, delección, truncamiento y/o sustitución de al menos un residuo de aminoácido, y que conservan una actividad biológica. En determinadas realizaciones, una variante de polipéptido se distingue de un polipéptido de referencia mediante una o más sustituciones, que pueden ser conservativas o no conservativas, tal como se conoce en la técnica. En determinadas realizaciones, un polipéptido variante incluye una secuencia de aminoácidos que tiene al menos aproximadamente el 50%, el 55%, el 60%, el 65%, el 70%, el 75%, el 80%, el 85%, el 90%, el 91%, el 92%, el 93%, el 94%, el 95%, el 96%, el 97%, el 98%, el 99% o más de identidad o similitud de secuencia con respecto a una secuencia correspondiente de un polipéptido de referencia. En determinadas realizaciones, las adiciones o delecciones de aminoácidos se producen en el extremo C-terminal y/o el extremo N-terminal del polipéptido de referencia. En determinadas realizaciones, las delecciones de aminoácidos incluyen truncamientos C-terminales de aproximadamente 1, aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 15, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 35, aproximadamente 40, aproximadamente 45, 50, aproximadamente 55, aproximadamente 60, aproximadamente 65, aproximadamente 70, aproximadamente 75, aproximadamente 80, aproximadamente 85, aproximadamente 90, aproximadamente 95, aproximadamente 100, aproximadamente 105, aproximadamente 110, aproximadamente 115, aproximadamente 120, aproximadamente 125, aproximadamente 130, aproximadamente 135, aproximadamente 140, aproximadamente 145, aproximadamente 150, aproximadamente 155, aproximadamente 160, aproximadamente 165, aproximadamente 170, o aproximadamente 175 o más aminoácidos, incluyendo todos los números intermedios de aminoácidos, por ejemplo, 25, 26, 27, 29, 30 ... 100, 101, 102, 103, 104, 105 ... 170, 171, 172, 173, 174, etc.

Tal como se indicó anteriormente, polipéptidos del contenido dado a conocer en el presente documento pueden alterarse de diversas maneras incluyendo sustituciones, deleciones, truncamientos e inserciones de aminoácidos. Métodos para tales manipulaciones se conocen de manera general en la técnica. Por ejemplo, pueden prepararse variantes de secuencia de aminoácidos de un polipéptido de referencia mediante mutaciones en el ADN. En la técnica se conocen bien métodos para mutagénesis y alteraciones de secuencia de nucleótidos. Véase, por ejemplo, Kunkel (1985, Proc. Natl. Acad. Sci. USA. 82: 488-492), Kunkel *et al.*, (1987, Methods in Enzymol, 154: 367-382), patente estadounidense n.º 4.873.192, Watson, J. D. *et al.*, Molecular Biology of the Gene, cuarta edición, Benjamin/Cummings, Menlo Park, Calif, 1987) y las referencias citadas en los mismos. Puede encontrarse orientación sobre sustituciones de aminoácidos apropiadas que no afectan a la actividad biológica de la proteína de interés en el modelo de Dayhoff *et al.*, (1978) Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found., Washington, D.C.).

Tal como se usa en el presente documento, el término “sustancialmente idéntico” se refiere a un polipéptido o un polinucleótido que muestra una identidad de al menos el 50% con respecto a una secuencia de aminoácidos de referencia (por ejemplo, una cualquiera de las secuencias de aminoácidos descritas en el presente documento) o una secuencia de ácido nucleico (por ejemplo, una cualquiera de las secuencias de ácido nucleico descritas en el presente documento). Preferiblemente, una secuencia de este tipo es idéntica en al menos el 60%, más preferiblemente el 80% o el 85%, y más preferiblemente el 90%, el 95% o incluso el 99% a nivel de aminoácido o ácido nucleico con respecto a la secuencia usada para comparación.

La homología o identidad de secuencia se mide normalmente usando software de análisis de secuencias (por ejemplo, paquete de software de análisis de secuencias del Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, programas BLAST, BESTFIT, GAP o PILEUP/PRETTYBOX). Tal software hace coincidir secuencias idénticas o similares asignando grados de homología con respecto a diversas sustituciones, deleciones y/u otras modificaciones. En un enfoque a modo de ejemplo para determinar el grado de identidad u homología, puede usarse un programa BLAST, indicando una puntuación de probabilidad de entre  $e^{-3}$  y  $e^{-100}$  una secuencia estrechamente relacionada. El porcentaje de identidad entre dos secuencias también puede determinarse con programas tales como DNAMAN (Lynnon Biosoft, versión 3.2). Usando este programa pueden alinearse dos secuencias usando el algoritmo de alineación óptimo (Smith y Waterman, 1981). Tras la alineación de las dos secuencias puede calcularse la identidad en porcentaje dividiendo el número de nucleótidos idénticos entre las dos secuencias entre la longitud de las secuencias alineadas menos la longitud de todos los huecos.

Los términos que describen la orientación de polinucleótidos incluyen: 5' (normalmente el extremo del polinucleótido que tiene un grupo fosfato libre) y 3' (normalmente el extremo del polinucleótido que tiene un grupo hidroxilo (OH) libre). Las secuencias de polinucleótido pueden indicarse en la orientación de 5' a 3' o en la orientación de 3' a 5'.

Tal como se usa en el presente documento, un “ARN guía individual” o un “ARN guía sintético” se refiere a la secuencia de polinucleótido que comprende la secuencia de guía, la secuencia tracr y la secuencia pareja de tracr. El término “secuencia de guía” se refiere a la secuencia de aproximadamente 20 pb dentro del ARN guía que especifica el sitio diana y puede usarse de manera intercambiable con los términos “guía” o “espaciador”. El término “secuencia pareja de tracr” también puede usarse de manera intercambiable con el término “repetición/repeticiones directa(s)”.

Los términos “que no se produce de manera natural” o “modificado por ingeniería” se usan de manera intercambiable e indican la implicación del ser humano. Los términos, cuando hacen referencia a moléculas de ácido nucleico o polipéptidos, significan que la molécula de ácido nucleico o el polipéptido está al menos sustancialmente libre de al menos otro componente con el que están asociados de manera natural en la naturaleza y tal como se encuentran en la naturaleza.

Tal como se usa en el presente documento, el término “expresión” se refiere al procedimiento mediante el cual se transcribe un polinucleótido a partir de un molde de ADN (tal como para dar y ARNm u otro transcrito de ARN) y/o el procedimiento mediante el cual un ARNm transcrito se traduce posteriormente para dar péptidos, polipéptidos o proteínas. Los transcritos y polipéptidos codificados pueden denominarse de manera colectiva “producto génico”. Si el polinucleótido se deriva de ADN genómico, la expresión puede incluir el corte y empalme del ARNm en una célula eucariota.

Tal como se usa en el presente documento, el término “tratar” o “tratamiento” se refiere a la intervención clínica en un intento de alterar el transcurso de la enfermedad del individuo o la célula que está tratándose, y puede realizarse o bien para profilaxis o bien durante el transcurso de la patología clínica. Los efectos terapéuticos del tratamiento incluyen, sin limitación, prevención de la aparición o recidiva de la enfermedad, alivio de síntomas, disminución de cualquier consecuencia patológica directa o indirecta de la enfermedad, prevención de metástasis, disminución de la tasa de progresión de la enfermedad, mejora o alivio del estado patológico, y remisión o mejora del pronóstico. Al prevenir la progresión de una enfermedad o trastorno, un tratamiento puede prevenir el deterioro debido a un trastorno en un sujeto afectado o diagnosticado o un sujeto que se sospecha que tiene el trastorno, pero además un tratamiento puede prevenir la aparición del trastorno o un síntoma del trastorno en un sujeto en riesgo del trastorno o que se sospecha que tiene el trastorno.

Tal como se usa en el presente documento, el término “sujeto” se refiere a cualquier animal (por ejemplo, un mamífero), incluyendo, pero sin limitarse a, seres humanos, primates no humanos, roedores y similares (por ejemplo, que va a recibir un tratamiento particular, o del que se recogen células).

5 Tal como se usa en el presente documento, el término “célula aislada” se refiere a una célula que está separada de los componentes moleculares y/o celulares que acompañan de manera natural a la célula. Tal como se usa en el presente documento, el término “aislado” se refiere a material que está libre, sustancialmente libre, o esencialmente libre en diversos grados de componentes que lo acompañan normalmente tal como se encuentra en su estado nativo. “Aislar” indica un grado de separación a partir del entorno o la fuente original.

10 Tal como se usa en el presente documento, el término “población celular” se refiere a un grupo de al menos dos células que expresan fenotipos similares o diferentes. En ejemplos no limitativos, una población celular puede incluir al menos aproximadamente 10, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 400, al menos aproximadamente 500, al menos aproximadamente 600, al menos aproximadamente 700, al menos aproximadamente 800, al menos aproximadamente 900, al menos aproximadamente  $10^3$  células, al menos aproximadamente  $10^4$  células, al menos aproximadamente  $10^5$  células, al menos aproximadamente  $10^6$  células, al menos aproximadamente  $10^7$  células, o al menos aproximadamente  $10^8$  células que expresan fenotipos similares o diferentes.

15 Tal como se usa en el presente documento, el término “escisión” se refiere a la rotura de la estructura principal covalente de una molécula de ADN. La escisión puede iniciarse mediante una variedad de métodos incluyendo, pero sin limitarse a, hidrólisis enzimática o química de un enlace fosfodiéster. Son posibles tanto la escisión de cadena sencilla como la escisión de cadena doble, y la escisión de cadena doble puede producirse como resultado de dos acontecimientos de escisión de cadena sencilla diferenciados. La escisión de ADN puede dar como resultado la producción o bien de extremos romos o bien de extremos en bisel. En determinadas realizaciones, se usan polipéptidos de fusión para la escisión de ADN de cadena doble dirigida.

20 Tal como se usa en el presente documento, el término “semidominio de escisión” se refiere a una secuencia de polipéptido que, junto con un segundo polipéptido (o bien idéntico o bien diferente) forma un complejo que tiene actividad de escisión (preferiblemente actividad de escisión de cadena doble). Los términos “semidominios de escisión primero y segundo;” “semidominios de escisión + y -” y “semidominios de escisión derecho e izquierdo” se usan de manera intercambiable para hacer referencia a pares de semidominios de escisión que se dimerizan.

25 Tal como se usa en el presente documento, el término “cromosoma” se refiere a un complejo de cromatina que comprende la totalidad o una parte del genoma de una célula. El genoma de una célula se caracteriza con frecuencia por su cariotipo, que es la colección de todos los cromosomas que comprenden el genoma de la célula. El genoma de una célula puede comprender uno o más cromosomas.

30 Tal como se usa en el presente documento, el término “gen” incluye una región de ADN que codifica para un producto génico, así como todas las regiones de ADN que regulan la producción del producto génico, tanto si tales secuencias reguladoras son adyacentes a secuencias codificantes y/o transcritas como si no. Por consiguiente, un gen incluye, pero no se limita a, secuencias de promotor, terminadores, secuencias reguladoras de la traducción tales como sitios de unión a ribosoma y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aislantes, elementos de límite, orígenes de replicación, sitios de unión a matriz y regiones de control de locus.

35 Los términos “unión operativa” y “unido de manera operativa” (o “unido operativamente”) se usan de manera intercambiable con referencia a una yuxtaposición de dos o más componentes (tales como elementos de secuencia), en los que los componentes están dispuestos de tal manera que ambos componentes funcionan normalmente y permiten la posibilidad de que al menos uno de los componentes pueda mediar en una función que se ejerce sobre al menos uno de los otros componentes. A modo de ilustración, una secuencia reguladora de la transcripción, tal como un promotor, está unida de manera operativa a una secuencia codificante si la secuencia reguladora de la transcripción controla el nivel de transcripción de la secuencia codificante en respuesta a la presencia o ausencia de uno o más factores reguladores de la transcripción. Una secuencia reguladora de la transcripción está generalmente unida de manera operativa en cis con respecto a una secuencia codificante, pero no necesita estar directamente adyacente a la misma. Por ejemplo, un potenciador es una secuencia reguladora de la transcripción que está unida de manera operativa a una secuencia codificante, aunque no sean contiguas.

40 Una “región funcional” o “parte funcional” de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína, polipéptido o ácido nucleico de longitud completa, pero conserva la misma función que la proteína, polipéptido o ácido nucleico de longitud completa. Una región funcional puede presentar más, menos o el mismo número de residuos que la molécula nativa correspondiente, y/o puede contener una o más sustituciones de aminoácido o nucleótido. En la técnica se conocen bien métodos para determinar la función de un ácido nucleico (por ejemplo, función codificante, capacidad para hibridarse con otro ácido nucleico). De manera similar, se conocen bien métodos para determinar una función de proteína. Por ejemplo, la función de unión a ADN de un polipéptido puede determinarse, por ejemplo, mediante unión a filtro, desplazamiento de movilidad electroforética o ensayos de inmunoprecipitación. La escisión de ADN puede someterse a ensayo mediante electroforesis en gel. La capacidad de una proteína para interactuar con otra

proteína puede determinarse, por ejemplo, mediante coimmunoprecipitación, ensayos de dos híbridos o complementación, tanto genética como bioquímica.

5 Tal como se usa en el presente documento, el término “promotor” se refiere a un sitio de reconocimiento de un polinucleótido (ADN o ARN) al que se une una ARN polimerasa. El término “potenciador” se refiere a un segmento de ADN que contiene secuencias que pueden proporcionar transcripción potenciada y en algunos casos pueden funcionar de manera independiente de su orientación con respecto a otra secuencia de control. Un potenciador puede funcionar de manera cooperativa o aditiva con promotores y/u otros elementos potenciadores.

10 Tal como se usa en el presente documento, el término “vector” se refiere a cualquier elemento genético, tal como un plásmido, fago, transposón, cósmido, cromosoma, virus, virión, etc., que puede realizar replicación cuando se asocia con los elementos de control apropiados y que puede transferir secuencias génicas al interior de células. Por tanto, el término incluye vehículos de clonación y expresión, así como vectores virales y vectores de plásmido.

15 Tal como se usa en el presente documento, el término “modular” se refiere a alterar de manera positiva o negativa. Las modulaciones a modo de ejemplo incluyen un cambio de aproximadamente el 1%, aproximadamente el 2%, aproximadamente el 5%, aproximadamente el 10%, aproximadamente el 25%, aproximadamente el 50%, aproximadamente el 75%, o aproximadamente el 100%.

Tal como se usa en el presente documento, el término “aumentar” se refiere a alterar de manera positiva en al menos aproximadamente el 5%, incluyendo, pero sin limitarse a, alterar de manera positiva en aproximadamente el 5%, en aproximadamente el 10%, en aproximadamente el 25%, en aproximadamente el 30%, en aproximadamente el 50%, en aproximadamente el 75%, o en aproximadamente el 100%.

20 Tal como se usa en el presente documento, el término “reducir” se refiere a alterar de manera negativa en al menos aproximadamente el 5% incluyendo, pero sin limitarse a, alterar de manera negativa en aproximadamente el 5%, en aproximadamente el 10%, en aproximadamente el 25%, en aproximadamente el 30%, en aproximadamente el 50%, en aproximadamente el 75%, o en aproximadamente el 100%.

25 Tal como se usa en el presente documento, el término “alrededor de” o “aproximadamente” significa dentro de un intervalo de error aceptable para el valor particular tal como se determina por un experto habitual en la técnica, que dependerá en parte de cómo se mide o determina el valor, es decir, de las limitaciones del sistema de medición. Por ejemplo, “aproximadamente” puede significar dentro de 3 o más de 3 desviaciones estándar, según la práctica en la técnica. De manera alternativa, “aproximadamente” puede significar un intervalo de hasta el 20%, preferiblemente hasta el 10%, más preferiblemente hasta el 5%, y todavía más preferiblemente hasta el 1% de un valor dado. De manera alternativa, particularmente con respecto a procesos o sistemas biológicos, el término puede significar dentro de un orden de magnitud, preferiblemente dentro de 5 veces, y más preferiblemente dentro de 2 veces, de un valor.

## II. Aislantes

35 Se han notificado varios casos de transformación maligna relacionada con vector en entornos clínicos, asociados con la activación de oncogenes celulares mediante potenciadores codificados por vector (Baum *et al.* (2006), Nienhuis *et al.* (2006), Ramezani *et al.* (2006)) y se han realizado o propuesto diversas modificaciones de vector para reducir la genotoxicidad de vector (Baum *et al.* (2006), Nienhuis *et al.* (2006), Ramezani *et al.* (2006)). Se ha reconocido una clase de elementos de ADN conocidos como aislantes de cromatina como un enfoque para mejorar la seguridad y rendimiento de vectores (Emery (2011)).

40 Los aislantes son elementos de ADN que se producen de manera natural que ayudan desde los límites funcionales entre dominios de cromatina adyacentes. Los aislantes se unen a proteínas que modifican la cromatina y alteran la expresión génica regional. La colocación de aislantes en los vectores descritos en el presente documento ofrece diversos beneficios posibles incluyendo, pero sin limitarse a, 1) protección del vector frente a la variegación por efecto de posición de la expresión mediante cromosomas flanqueantes (es decir, actividad de barrera, que puede disminuir los efectos de posición y el silenciamiento de vector); y 2) protección de cromosomas flanqueantes frente a la trans-activación por inserción de la expresión de genes endógenos por el vector (bloqueo de potenciador). Hay dos clases básicas de aislantes de cromatina: (a) aislantes de barrera que bloquean la intrusión de heterocromatina de silenciamiento en regiones contiguas de cromatina abierta que son permisivas para la transcripción, y (b) aislantes de bloqueo de potenciador que impiden la activación de la transcripción mediada por potenciador de regiones contiguas. Las secuencias que median en estas actividades pueden separarse físicamente y son mecánicamente distintas (Recillas-Targa *et al.* (2002)). Los aislantes de cromatina no muestran por sí mismos actividades de represión o potenciación de la transcripción inherentes. Como tales, constituyen elementos ideales para reducir la interacción entre vectores de transferencia génica y el genoma de célula diana. Los aislantes pueden ayudar a conservar la función independiente de genes o unidades de transcripción incorporados en un genoma o contexto genético en el que su expresión puede verse influida de lo contrario por señales reguladoras dentro del genoma o el contexto genético (véase, por ejemplo, Burgess-Beusse *et al.* (2002) Proc. Nat'l Acad. Sci. USA, 99: 16433; y Zhan *et al.* (2001) Hum. Genet., 109: 471).

Los problemas creados por la mutagénesis por inserción de vectores virales se conocen ampliamente (Nienhuis

(2013), Baum *et al.* (2006), Nienhuis *et al.* (2006)) al igual que la evidencia de que los riesgos de genotoxicidad pueden reducirse mediante el uso de aislantes de cromatina (Arumugam *et al.* (2007), Emery (2011), Evans-Galea *et al.* (2007), Rivella *et al.* (2000), Emery *et al.* (2000), Emery *et al.* (2002), Yannaki *et al.* (2002), Hino *et al.* (2004), Ramezani *et al.* (2003), Ramezani *et al.* (2008)). El contenido dado a conocer en el presente documento proporciona  
 5 aislantes novedosos que son potentes aislantes de bloqueo de potenciador, y determinados aislantes presentan adicionalmente actividad aislante de barrera. En vertebrados, la función de aislantes de bloqueo de potenciador está mediada a través del factor de unión a ADN de dedos de zinc CTCF (Gaszner y Felsenfeld (2006), Wallace y Felsenfeld (2007)). En general, se piensa que estos elementos funcionan mediante estructuras de bucle físico, que se establecen mediante interacciones mediadas por CTCF entre elementos aislantes adyacentes o mediante anclaje  
 10 mediado por CTCF de la fibra de cromatina a elementos estructurales dentro del núcleo. El primer aislante de cromatina de vertebrados caracterizado está ubicado dentro de la región de control de locus de  $\beta$ -globina de pollo. Este elemento, que contiene un sitio hipersensible 4 (cHS4) a ADNasa I, parece constituir el límite en 5' del locus de  $\beta$ -globina de pollo (Prioleau *et al.* (1999) EMBO J. 18: 4035-4048). Una región de 1,2 kb que contiene el elemento cHS4 presenta actividades aislantes clásicas, incluyendo la capacidad de bloquear la interacción de promotores de  
 15 gen de globina y potenciadores en líneas celulares (Chung *et al.* (1993) Cell, 74: 505-514), y la capacidad para proteger casetes de expresión en *Drosophila* (*Id.*), líneas celulares transformadas (Pikaart *et al.* (1998) Genes Dev. 12: 2852-2862), y mamíferos transgénicos (Wang *et al.* (1997) Nat. Biotechnol., 15: 239-243; Taboit-Dameron *et al.* (1999) Transgenic Res., 8: 223-235) frente a efectos de posición. Gran parte de esta actividad está contenida en una región de 250 pb. Dentro de este tramo hay un elemento cHS4 de 49 pb (Chung *et al.* (1997) Proc. Natl. Acad. Sci.,  
 20 USA, 94: 575-580) que interacciona con la proteína de unión a ADN de dedos de zinc CTCF implicada en ensayos de bloqueo de potenciador (Bell *et al.* (1999) Cell, 98: 387-396).

Los aislantes, tales como cHS4, pueden bloquear la interacción entre potenciadores y promotores cuando se colocan entre estos elementos (Evans-Galea *et al.* (2007), Chung *et al.* (1997), Bell *et al.* (1999), Ryu *et al.* (2007), Ryu *et al.* (2008)). Varios estudios han demostrado la capacidad del aislante cHS4 para reducir el silenciamiento por  
 25 efectos de posición de vectores gamma-retrovirales (Evans-Galea *et al.* (2007), Rivella *et al.* (2000), Emery *et al.* (2000), Emery *et al.* (2002), Yannaki *et al.* (2002), Hino *et al.* (2004), Ramezani *et al.* (2006), Yao *et al.* (2003), Nishino *et al.* (2006), Aker *et al.* (2007), Li y Emery (2008)), y vectores lentivirales (Bank *et al.* (2005), Arumugam *et al.* (2007), Puthenveetil *et al.* (2004), Evans-Galea *et al.* (2007), Ramezani *et al.* (2003), Aker *et al.* (2007), Ma *et al.* (2003), Chang *et al.* (2005), Pluta *et al.* (2005)). Estos estudios diseñados de manera apropiada demostraron que la  
 30 inclusión de la versión de 1,2 kb del aislante cHS4 aumentaba la probabilidad y/o sistematicidad de la expresión transgénica de vector en al menos algunos entornos (Arumugam *et al.* (2007), Emery (2011), Evans-Galea *et al.* (2007), Emery *et al.* (2002), Yannaki *et al.* (2002), Hino *et al.* (2004), Ramezani *et al.* (2006), Aker *et al.* (2007), Li y Emery (2008), Pluta *et al.* (2005). Jakobsson *et al.* (2004)). No obstante, el grado de protección proporcionado por el  
 35 aislante cHS4 está lejos de ser completo. Además, la inclusión del cHS4 de 1,2 Kb puede afectar de manera adversa a los títulos de vector mientras que se ha demostrado que el núcleo de cHS4 más pequeño es ineficaz (Aker *et al.* (2007), Jakobsson *et al.* (2004)). En cambio, los aislantes del contenido dado a conocer en el presente documento no afectan de manera adversa a los títulos de vectores virales, y son más potentes y eficaces que el  
 aislante cHS4.

Los aislantes dados a conocer en el presente documento se identifican mediante enfoques genómicos, por ejemplo,  
 40 usando enfoques genómicos para identificar aislantes que son potentes bloqueadores de potenciador así como aislantes de barrera del genoma humano. Los aislantes dados a conocer en el presente documento potencian la seguridad de la terapia génica (por ejemplo, terapia génica de células madre, terapia génica de globina). Para la terapia génica de las hemoglobinopatías, se requieren potenciadores potentes para lograr niveles terapéuticos de  
 45 expresión de gen de globina. Por tanto, los aislantes potentes representan un medio para proteger el entorno genómico frente a los potenciadores potentes de los vectores de integración.

Los aislantes dados a conocer en el presente documento presentan una potente actividad de bloqueo de  
 50 potenciador. Por ejemplo, y no a modo de limitación, un aislante de la presente divulgación puede reducir la actividad de un elemento potenciador en al menos aproximadamente el 10%, al menos aproximadamente el 20%, al menos aproximadamente el 30%, al menos aproximadamente el 40%, al menos aproximadamente el 50%, al menos  
 55 aproximadamente el 55%, al menos aproximadamente el 60%, al menos aproximadamente el 65%, al menos aproximadamente el 70%, al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos  
 60 aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98% o al menos aproximadamente el 99%. En determinadas realizaciones, los aislantes presentan actividad de barrera además de actividad de bloqueo de potenciador. Los aislantes dados a conocer en el presente documento disminuyen sustancialmente los riesgos de mutagénesis por inserción y genotoxicidad asociados con vectores virales. Además, cuando se incorpora un aislante dado a conocer en el presente documento en un vector, el aislante no afecta de manera adversa a los títulos de vector del vector. En determinadas realizaciones, los aislantes (por ejemplo, aislante A1) aumentan la expresión *in vivo* del gen de globina o parte funcional del mismo.

Según la invención, el aislante comprende un sitio de unión a CTCF de represor de la transcripción, que tiene la

secuencia de nucleótidos expuesta en SEQ ID NO: 18, que se proporciona a continuación:

CACCAGGTGGCGCT [SEQ ID NO: 18].

5 En una realización no limitativa, el aislante tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1, que se proporciona a continuación, o una secuencia que es homóloga en al menos aproximadamente el 95 por ciento, o idéntica (homóloga) en al menos aproximadamente el 98 por ciento, a SEQ ID NO: 1. Este aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1 se denomina aislante A1.

TCCTTCCTTTCTAAATGACGAGAGAGACAGAAGAATTCTTCAAGGTTAGTGTGTCCAGCATG  
CAACCTTTCTTCCCTGGATGAGCATCCCTGGAGTAGGAGAGCCAGCCTGCCTCCTGCGCTGG  
CACAGAGCCCGGTTCCCTAGACAACCTGCCTCTCCAAATCTGATGTCCAGCGCCACCTGGTGT  
CCACATCAAGCAGACACAATTAATAGTCAACCTGTTTCAGGAAAACCTGTGAGGGGGAAAAAAA  
AGAAAGAGGATTTATGAAGGGAAAAGAAAGTTTAGAGGATATGCCACGATTGGCTAG [SEQ  
ID NO: 1]

10 En determinadas realizaciones, el aislante comprende una secuencia de nucleótidos tal como se expone en SEQ ID NO: 24, o una secuencia que es idéntica en al menos aproximadamente el 95 por ciento, o idéntica en al menos aproximadamente el 98 por ciento, a SEQ ID NO: 24.

CCAATC GTGGCATATC CTCTAAACTT TCTTTTCCCT TCATAAATCC TCTTTCTTTT  
TTTTCCCCT CACAGTTTTT CTGAACAGGT TGACTATTAA TTGTGTCTGC  
TTGATGTGGA CACCAGGTGG CGCTGGACAT CAGATTTGGA GAGGCAGTTG  
TCTAGGGAAC CGGGCTCTGT GCCAGCGCAG GAGGCAGGCT GGCTCTCCTA  
TTCCAGGGAT GCTCATCCAG GAAGGAAAGG TTGCATGCTG GACACACTAA  
CCTTGAAGAA TTCTTCTGTC TCTCTCGTCA TTAGAAAGG AAGGA [SEQ ID NO:  
24] .

En determinadas realizaciones, el aislante comprende una secuencia de nucleótidos tal como se expone en SEQ ID NO: 25 (que es el complemento inverso de SEQ ID NO: 1), o una secuencia que es idéntica en al menos aproximadamente el 95 por ciento, o idéntica en al menos aproximadamente el 98 por ciento, a SEQ ID NO: 25.

CTAGCCAATCGTGGCATATCCTCTAAACTTTCTTTTCCCTTCATAAATCCTCTTTCTTTTTT  
TTCCCCCTCACAGTTTTTCTGAACAGGTTGACTATTAAATTGTGTCTGCTTGATGTGGACACC  
AGGTGGCGCTGGACATCAGATTTGGAGAGGCAGTTGTCTAGGGAACCGGGCTCTGTGCCAGC  
GCAGGAGGCAGGCTGGCTCTCCTACTCCAGGGATGCTCATCCAGGAAGGAAAGGTTGCATGC  
TGGACACACTAACCTTGAAGAATTCTTCTGTCTCTCTCGTCATTTAGAAAGGAAGGA [SEQ  
15 ID NO: 25]

En determinadas realizaciones, el aislante comprende una secuencia de nucleótidos tal como se expone en las coordenadas de hg18 de 76229933 a 76230115 del cromosoma 1.

20 En determinadas realizaciones, el aislante comprende una secuencia de nucleótidos entre los residuos 68041 y 68160, o entre los residuos 68041 y 68210, o entre los residuos 68041 y 68280, o entre los residuos 68005 y 68305, del clon RP11-550H2 del cromosoma 1 de *Homo sapiens*, n.º de registro de GenBank AC092813.2, o una secuencia idéntica en al menos el 95 o el 98 por ciento a la misma.

### III. Casetes de expresión

25 El contenido dado a conocer en el presente documento proporciona casetes de expresión que comprenden uno o más los aislantes dados a conocer anteriormente (por ejemplo, aislante A1). En determinadas realizaciones, un casete de expresión comprende al menos un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1, y un gen de globina o una parte funcional del mismo operativamente unido a una región LCR de  $\beta$ -globina.

Región LCR de  $\beta$ -globina

La agrupación de genes de  $\beta$ -globina humana consiste en cinco genes incorporados dentro de una de muchas

matrices de genes de receptores olfativos (Bulger *et al.*, PNAS (1999); 96: 5129-5134). La agrupación abarca más de 80 kb en el cromosoma 11p15.4, e incluye los cinco genes de tipo  $\beta$  expresados y elementos reguladores que actúan en *cis* que dirigen su expresión específica de estadio durante la ontogenia (Forget (2001), Molecular Mechanism of Beta Thalassemia. Steinberg MH *et al.*, Eds. Disorders of Hemoglobin. Genetics, Pathophysiology and Clinical Management, Cambridge University Press, Cambridge). Los genes están dispuestos en el orden de su expresión de desarrollo (Stamatoyannopoulos *et al.*, (2001) Hemoglobin Switching. En: Stamatoyannopoulos G, *et al.*, Eds. Molecular Basis of Blood Disorders, W.B. Saunders, Filadelfia, PA), 5'- $\epsilon$ - $\zeta$ - $\gamma$ - $\psi$ - $\eta$ - $\delta$ - $\beta$ -3'. La agrupación de genes de globina de tipo  $\alpha$  (5'- $\xi$ - $\zeta$ - $\psi$ - $\xi$ -1- $\psi$ - $\alpha$ -2- $\psi$ - $\alpha$ -1- $\alpha$ -2- $\alpha$ -1- $\theta$ -3') está ubicada muy cerca del telómero del brazo corto del cromosoma 16 y abarca aproximadamente 40 kb. La expresión de genes codificados dentro de estas dos agrupaciones independientes está limitada a células eritroides y equilibrada de modo que la producción de las cadenas de tipo  $\beta$ -globina coincide con la de las cadenas  $\alpha$ . Este equilibrio finamente ajustado se regula a niveles de transcripción, tras la transcripción y tras la traducción.

La expresión específica de estadio de desarrollo se controla por varios elementos que actúan en *cis* proximales o distales y los factores de transcripción que se unen a los mismos. En el caso del gen de  $\beta$ -globina (*HBB*), los elementos reguladores proximales comprenden el promotor de  $\beta$ -globina y dos potenciadores aguas abajo, uno ubicado en el segundo intrón de  $\beta$ -globina y el otro aproximadamente 800 pb aguas abajo del gen (Antoniou *et al.*, EMBO J. (1988); 7: 377-384; Trudel *et al.*, Genes Dev. (1987); 1: 954-961; Trudel *et al.*, Mol. Cell. Biol. (1987); 7: 4024-4029). El elemento regulador distal más prominente es la LCR de  $\beta$ -globina, ubicada 50-60 kb aguas arriba del *HBB* y compuesta por varias subregiones con sensibilidad intensificada a ADNasa I en células eritroides (Forget (2001); Grosveld *et al.*, Cell (1987); 51: 975-985; Talbot *et al.*, Nature (1989); 338: 352). La propiedad más prominente de la LCR es su fuerte actividad potenciadora de la transcripción. Una secuencia de nucleótidos a modo de ejemplo de la región de  $\beta$ -globina humana en el cromosoma 11 se expone en SEQ ID NO: 19 (n.º de registro de GenBank: NG\_000007.3), que se proporciona a continuación:

ES 2 746 529 T3

ggatcctcacatgagttcagtatataaattgtaacagagaataaaaaatcaattatgtattcaagttgctagtgtctta  
agaggttcacatTTTTtactgattatcacaaaaacttcgagttacttttattataattcctgactacac  
atgaagagactgacacgtaggtgacctacttaggtaggttaagtaatttatccaaaaccacacaatgtagaaccta  
agctgatcggccatagaaacacaatattggtgataaatgagacagagggatttctctccttccctatgctgtcaga  
tgaatactgagatagaatatttagttcatctatcacacattaaacgggactttacatttctgtctgttgaagattt  
gggtgtgggataactcaaggtatcatatccaagggatggatgaaggcaggtgactctaacagaaaaggaaaggat  
gttggcaaggctatgttcatgaaagtatatgtaaaaatccacattaaagcttctttctgcatgcatggcaatgttta  
tgaataatgtgtatgtaaaagtgtgctgtatattcaaaagtgtttcatgtgacctaggggtgtcaaatactttgagt  
ttgtaagtataacttctctgtaattgtgtctgaatatctctatttacttgattctcaataagtaggtatcatagt  
aacatctgacaaaatglttgaggaacaatttagtgtttacctattcaccaaaatttattaaatgacctaatctgtatc  
agatatcaaatatctggcgaatctgtaattcctaatttaaacagctgtgtagcctaattagggataaaaggcatg  
caaacccataatTTgtgtaggtgaaatgagctatagaaaaatgcagtatatttatcagaagtcttttagggctatg  
aaaaggaatggtaactgacactgccagggactcatatgtaagagataactaatgtgaagtgactttaaaggagaa  
attagcagaagtttctttccatgtctcctcatcatgttacaataacggaagagattaaaacaacaaatcacattta  
gacagcaatgtttatcctgggttagatgttttaactctaaatctatcttggagtgttaaaatgcatttgcctcacctac  
tttaaaatataaatgaaggttaggaacctgtagatacaaaaagttggagaaaaaaagacaataaagatgacaaaaat  
ctattaatccttgatagaaaatgagaagagataaaaacactggttacataaaagaaaaataagatggatagatagcag  
atccttataaaaagtgataatttgagaaaaaaaatactccatattctgagtttcttcacataaaaataatacaaatct  
gctgtggttaagttacaaagagatagattttttatcattatataaaaagatattttaaacagagttatacaacaaagg  
aacagactatgtcatatattctcacttatcactataaacatctcagaaaaatctgcaaaatcatttcatagcattt  
taaatagtttaggaataatgtagaaaactgaaacagttctaaagtttccacaaacttagagctcaaatgttgcatt  
acctaaactaacctgcaaatattttatacaaatttgacacatgctactctagtcaaaaatataatgacattatgggta  
TTTTctgtgtgtaacttgggtctagttgcttctttcagaaatagcctctatttttgatttacctgataaaaatcaca  
ttcctctccaaagccttctaaatacttcagactaactacttttagtacatctaagaagaaaaggttttgtctc  
ttatccacctctgagtcaaaaagcagcatgtccatcaattgggtacatagttcccaagccccacttagctctggat  
tggagtctacttggcattgtttgcaactacatggacgtaaaatgcattggattctcttgaaaaaatgtttctgcca  
tgatgttctctgaaagagactaaccttccctcgttttgacagaaaagactcgtgtaatccttgacaatgtctctc  
atctatttatcccatgtctaccatattgtgaccttcatgtctttgctctaaagccccacatcctcaatctacaca  
ctaggatagatataaaagtaatagtaataatagtagtaatagtaataacaatacaatgattatggcttatactatac  
acaagacactgttgatataattattcatttagtatccacagtaactctgtgacctcaagtaactattgtaataccctt  
taagaggagaaaactgagggcacagggccctaaagtaataatccaagatgaagtggctactaactgacagagggcat  
aattcaactcatgatatttggctctagaatacatgtctgtaactattatacaataataattcatgaggaacattt  
tttaaagcctaagttatttggctctgaaataagacataaatttggggtagaaaagcttagattccatgaagattaca  
gcatttggtagctttttgactccaggtcttatttttactgcttaaacataataaacatattggtcagttatgcc  
tttgattttacaataatattcctgttatttttgggaagcacaggggtgtgggataatgctaattactagtgttagta  
ttgagaggtgacagcgtgctggcagctcctcacagccctcgtcgtcttggcgccctcctctgctgggctcccaca  
ttgggtggcacttgaggagcccttcagccggccgctgcaactgtgggagcccttttctgggctggccaaggccagagc  
cggctcctcagcttgccaggaggtgtggagggacagacggggcaggaaccgggctgtgcccgtgcttgagggga  
gttccgggtgggcatgggctccgaggaccccgcactcggagccggccagccggccccaccggccggcagtgagg  
ggcttagcactgggcccagcagctgctgtgctcaattcctcggccggccttagctgecttccctcggggcagggct  
cgggacctgcagcggccatgctgagcctccccaccttcatgggctcctgtgcccggccagcctcggccagcagc  
ggcggccctgctccagggcacccagttccatcgaccccaagggtgaagagtggggggcagcggcaggggact  
ggcagggcagctccccctgcagcccaggtgcccgatccactgggtgaagccggctaggctcctgagtttgcctgggga  
tgcgaagaaccccttatgtctagataagggattgtaaaatcaccaattggcactctgtatctagctcaaggtttgta  
aacacccaatcagcaccctgtgtctagctcaggggtttgtgaatgcaccaatcaacactctatctagctactctgg  
tggggccttggaacacctttatgtctagctcagggattgtaaaatcaccaatcggcagctctgtatctagctcaagg  
tttgtaaacacaccaatcagcaccctgtgtctagctcaggggtttgtgaatgcaccaatcaacactctgtatctagc  
tactctgggtggggacgtggagaacctttatgtctagctcagggattgtaaaatcacaccactcggcagctctgtatcta  
gotcaaggtttgtaaacacaccaatcagcaccctgtgtctagctcaggggtttgtgaatgcaccaatcaacactctg  
tatctagctactctgggtggggacttggagaacctttgtgtggacactctgtatctagctaatctgggtggggacgtg  
gagaacctttgtgtctagctcatggattgtaaatgcaccaatcagtgccctgtcaaacagaccactgggctctac  
caatcagcaggatgtgggtggggccagataagagaataaaaagcagggctgcccagccagcagtggaaccccgtctg  
ggccccctccacactgtggaagctttgttctttcgtctttgcaataaatcttgcctgctgctcactgtttgggtc  
tacactgectttatgagctgtaacgctcaccgggaaggtctgcagcttcaactctgaagccagcagaccacgaac

ccaccgggaggaacgaacaactccagaggcgccgcttaagagctggaacgttcaactgtgaaggctgcagcttca  
ctcctgagccagcgagaccacgaacccatcagaaggaagaaactccgaacacatccaaacatcagaacgaacaaac  
tcccacacgcagcctttaaagaactgtaaacactcaccacgagggtccccggcttcattctctgaagtcagtgaaacc  
aagaacccaccaatcccgacacagtatgtcagaacaatatgagtcactaaatcaatatactctccaacaatttc  
caacagcccttgcaatfaacttggccatgtgactggttgtgactaaaaatgtggagataaataatgtgttactcc  
ctaaggcagagtgcccttctatcattctcttcccttctctatgtggcagaaagttaaagatcttgaatgataa  
agtcaatcacaggaaggccctggactcctggccactgcttggaggagagcactcaggaccatgaacatctgact  
gtgacgtagcaataaagaacccacgcttcatatgaaactgcttaaaatfaatggcacaagtcagtttttgatgt  
tgcacatttgtctttatgtggcttgttttgcctccacatcaatccactcaaggcctacattctgctataatgca  
atltcaagttctttacaggccogagaaaaatgaatctgaaatccctgacctccaaaagtgatcaagatatttttagtt  
caggctccaaaattttctcattttcataggttttccctcgattgatcattttcatgatltgcaaggaatcattcaa  
tgtttctaaatctattactgcatcctgacacatatgacatttttaactatgttccagatttttgaatgaagagtg  
aaattttaaagttttccaccacaaaaataagtatgtgaagtggtggatttgttaattagccttatttaaccattt  
aataattgtacagtlacaccaaaagcatcagttgtaccoccatgaatacacacaattattttgtcaatttaaaatg  
aaaaataaaaaataacaaaggcattagcctctgcatgtccctttaccggctcactccacggctgactaacgcaaaaa  
acgttctattttcactccttacaacaacatccctatccttggatgacctcttggctctagatctctacccctcctgttct  
ctacgttattttatagggatcaccatcctggacaacatcaggacagatatccctcaccacagccaatgttctct  
ctctatgttggctcaaatgtccttgaacttccctttcaccaccctttccacagtcacaaaggatattgtagttta  
gocctcagagttcagcttttaagctcttgacaaaattatcttctctctttaggttctcctttatggaatctctgtac  
tgatggccatgtcctttaaactactatgtagatctctgctactacctgtattatgocctcactcttatttagcagag  
tatctgtactgttggcatgacaatcatttgttaatatgacttgccttctctttctgctattcttgatcaaatgg  
ctcctcttcttctcctctcatttctctctgoccttcaacttggacgttctcagtagtctgtgcttatgactggat  
taaaaattgatattgacttatcctaagtgtgtctgataaataatgggttttatggctcattattttctcctatgca  
ttgatctggagaaggcttcaatccttttactccttgggaaaaatctgtaaaacctctgggtcactctgctatag  
caattcagtttaggctagtaagcatgaggatgocctctctctgattttccacagctctgttggctcacagaata  
acctgagtgattactgatgaaagagtgagaatggttatgtatgtcacaatgacaaaaacaaacactacagtcac  
aatgtttctctttttattagtgatttatatttctgacctatctctggcaggactcttagagaggttagctgaagc  
tgctgttatgaccactagagggaagaagatacctgtggagctaatgggtccaagatgggtggagccccaagcaaggaa  
gttgttaaggagcccttttgattgaaggtgggtgccccaccttacagggacaggacatctggatactcctcccag  
ttctccagtttcccttttccataatatactcctgataaaaatgtctatactcacttcccatttctaaataaaa  
gcaaggctagttagtaagacatcccttgcattttgaaaatgoccatagacttcaaaattatttctacatcaggt  
ccttctttatttcaagagctcagaaaatggcaacattcactttgattcaatgtaatggaaagagctcttcaagga  
cagagaaaagaataaatttaatttcttcccccacacctctcctctgctcttaccctatcttctcctctcacc  
tcccatttctctctctcatttctcagaagtataattttgaaaggatccatagcagacagctaaaggctgggtttttc  
taagtgaagaagtgatattgagaaggtagggttgcagagcccttcagtttttttagtttatatacatctgtattg  
ttagaatgttttataataataaaaattatttctcagttatatactagctatgtaacctgtggatattccttaa  
gtattacaagctatacttaactcacttggaaaactcaataaatacctgcttcatagttatataaaggattaggt  
gagataatgcccataagattcctatttaataacagataaaatcacatacacacacacacacattgaaaggatcttact  
ttgtctaggaactalaalaagttcatttgatgcattatcattaaagttctaatttcaaacactagaaggcaggtat  
tatctaaaatttcaactggatccctccaaactcataaagataaataaattgccttttgtcataatatttcaaaa  
gggtaaaactcaactatggcttgtctaattttatatacacccctactgaacatgacctatttggatattttataa  
aattatctcaagttattatgaggatgttgaagacagagaggatgggggtgctatgccccaaatcagcctcacaat  
taagctaagcagctaaagctctgcagggtagtgtagggaccacaggggttaagggggcagtagaattatactcca  
ctttagtttctttcaacaatccatcacacacacagccctgagcacttacaatatactacgctctatacttttt  
gtttaaattgataaaaatagttggatgaaagaatagatagatagatagacagatagatgalagalagaataaattgctt  
gocctcatagctgtctcctcacttggttcaaaatgttctctgtccagaccaaagtaocttgccttcaactaaagta  
caattctcaggttatattctgagtgcaaggaagtcaaaagatgtgaaaaacaatttctgaccacacactcagct  
ttgtagatgactagatcaaaaaattcagccatcttcaacagtgagtgaacaggaaatctcctcttttccctaca  
cttgagatcccagcttctaaagaccttcaattctcactcttgatgcaacagaccttggaaagcatacaggagagctga  
acttggctcaacaaaggagaaaaagttgttggcctccaaaggcacagctcaaacctttcaagcctctctaaatctta  
aaggtaaacaagggtctcatttctttgagaacttcagggaaaaatagacaaggacttgcctggtgcttttggtaggg  
gagcttgcacttcccccttctggaggaaatatttatccccaggtagttccctttttgaccagtggttctttga  
agagacttccacctgggaacagttaaacagcaactacagggccttgaactgcaacacttccagtcgggtcctcacag  
ttgaaaagacctaaagcttgtgctgatttaagccttttggctcaaaaacattgaaattctaatctcctctcaacc  
ctacagctaccctttgggtatataaagatgtgttctactgtctagtatcctcaagttagtgcaggaataggt  
catttaaatagttctgcaagccaggagtggtggctcatgtctgtaattccagcacttgagaggtagaagtgaggga  
ctgcttgagctcaagagtttgatattatcctggacaacatagcaagacctcgtctctacttaaaaaaaaaaaaaaa  
attagccaggcatgtgatgtacacctgtagtcocagctactcaggaggccgaaatgggaggatcccttgagctcag  
gaggtcaaggctgagtgagacatgatcttgcactgcaactccagcctggacagcagagtgaaaccttgocctcacg  
aacagaatacaaaaaacaaacaaacaaaaaactgctccgcaatgcgcttcccttgatgctctaccacataggtctgg  
gtactttgtacacattatctcattgtgttcaataattgttagattaaattttgtaatatgtatatttcttagaaa  
gctgaggocccaagatgataacttttatttctggacttgaatagctttctcttgatccaccatgttgaacLL  
tcttagagtagtaacaatataaagttattgtgagtttttgcacacacagcaaacacacacaccccatatagacattg

atgtgaaatgtctatgtcaatttatgggaaacaagtatgtacttttctactaagccattgaaacaggaaataa  
cagaacaagattgaaagaatacattttccgaaattacttgagttatatacaaaagacaagcactggacctgggagg  
agggttatgtccatgactgggtgtgtggagacaaatgcaggttataatagatgggatggcatctagcgaatgac  
tttgccatcacttttagagagctcttggggaccccagtacacaagaggggacgcagggtatagttagacatctcat  
tcttttcttagtgtgagaataagaatagccatgacctgagtttatagacaatgagcccttttctctctcccactc  
agcagctatgagatggcttgcctgctctctactaggtgactcactccaaggcccagcaatgggagggctctg  
tcagggttttgatagcactatctgcagagccagggccgagaaggggtggactccagagactctccctcccattccc  
gagcagggttgtctatttatgcatttaaatgatataatttataaaagaaataaacaggagactgcccagcctg  
gctgtgacatggaaactatgtagaatattttgggttccatttttttctcttccagtttagaggaaaggggct  
cactgcacatacactagacagaaagtcaggagctttgaatccaagcctgatcatttccatgtcactgagaaagt  
ccccaccttctctgagcctcagttctcttttataagtaggagctggagtaaatgatltccaatggctctcat  
ttcaatcacaaaatttccgtttatataatgcatgagctctgttactccaagactgagaaggaaatgaaacctgaga  
ctcattgactggcaagatgtcccagaggtctctcattcagcaataaaaattctcaccctcaccagggcccactgagt  
gtcagatttgcattgactagttcaaggtgtgtaaaaaggaggatgcttcttctcttctgtattctcacatacctttag  
gaaacttagcaccctcccacacagccatcccataaactcatttcagtgactcaacccttgactttataaaaag  
tcttgggcagtatagagcagatataaggtacagatgctggagccagaccctgagtgatagtgaaagcagttt  
ctcttagtagttgtatgactcagttctctcatctgtaaaatggaggggttttataatagtttgtttttgagaaagg  
gtctcactctgtcaccocaaatgggagtgtagtggcaaaatctcggctcactgcaacttgacttcccaggtcaag  
cggctctcccacclcaacalctgagtagctggaaccacaggtacacaccaccatcctcgttaatttttctgatt  
tttggtagagatggggttccatggttacacaggatggtctcagactccggagctcaagcaatctgcccacctcag  
ccttccaaagtgtgggattataagcatgattacaggagtttaacaggctcataagattgttctgcagcccaggt  
gagttatacatgcaaagagtttaaagcagtgacttaataatgctaaactactctagaaatgttgcagtagtattt  
tgttaactgcaatcattctgtcaggtgaaaactagtggtctgtactttatgcccattcattttaaactgtaa  
taataaaaaaactgacattttatgaaggctatcagagactgtaattagtgctttgcataatcaatcatalttaat  
actcttggattcttccaggtagatactattatctcccatttactacagttaaaaaaactacctctcaacttgc  
lcaagcatacactctcacacacacaaacataaaactactagcaaatagtagaattgagatttggctcctaatatgtc  
tttgctcactatccaataaataattttatgacatgtacttcttggcagctctgtatgctggatgctgggatacaaaag  
atgtttaaatttaagctccagctctctgcttccaaagcctcccaggccaagttatccattcagaaagcatttttta  
ctcttgcattccactgtttttccaaagtactaaaaaattacactttatctgtctgtctcctgctctgggatgat  
agctgactttcctaacctgagcctaacatcctgacatcaggaaagactacaccatgtggagaaggggtgggtgggt  
tttgatgtctgtctctcagttagatgggttaactttgtgaagttgaaaactgtggctctctggttactgtttaga  
gttctggcacttctcactgctcctattatttaacaaatgcatgaatgcttcagaatctgggaatattatctctgg  
aatagggaaatcaagttatattatgtaaccaggattagaagattctctgtgtgtaagaatttcataaacatataag  
ctgtctagcaaaaagcaagggcttggaaaatctgtgagctcctcaccatatagaagcttttaaccatcatgaaat  
aaatccctataggggatttctaccctgagcaaaaggtggctcttgatgaatccccaaactcatatagctctgagaa  
agctatgctgttaacgtttctctgtctgtaccccacatatagcacaacaataaatgcaggcctaggcatgactg  
aaggctctctcataattcttgggtgcatgaatcagattatcaacagaaaatgttgagacaaactatggggaagcagg  
gtatgaaagagctctgaatgaaatggaaaccgcaatgcttctgcccattcagggtccagcatgtagaaaatctgg  
ggcttctggaagctggcttaaaatcagaagccccatgggataagagtaggggaagaacctagagcctacgctgag  
aggttctctcagtgacagggagcctctgccccagactccagggatcctctcttaagtgttctctggaat  
ctcctcacttctatctggaaatggttctccacagctccagcccctggctagttgaaagagttaccatgcagagggc  
cctcctagcatccagagactagtgcttagattcctactttcagcgttgacaacctggatccacttggccagtggt  
cttctctagttcctaccttgcacttgatcctcctttatcttctgaacctgctgagatgatctatgtggggaga  
atggctctttgagaaacatcttctctggttagtggctgcccctcattcccactttaatccagaatcactataa  
gaagaatataatagaggaataaactcttattataggttaagggaaaaatagaggcacaactgagtgaggatgagtaag  
agaggagagggaaagattatggagcataaaaactactactatttgttgagacctttatagttcaatcaattttg  
ctattgtttccatcctcagcctaactccataaaaaaacactattatctcttattttgcccatacaagactgag  
ctcagaagagtcagcatttgcctaaggtcggacatgtcagaggcagtgccagacctatgtgagactctgcagcta  
ctgctcatgggcccctgtgctgactgatgaggaggatcagatggatggggcaatgaagcaaaaggaatcattctgtg  
gataaaggagacagccatgaagaagttctatgactgtaaaatttgggagcaggagttcttaaggacttggatttcaag  
gaatttgactcagcaaacacaagaccctcacggtagctttgcagagctgggtgtccagatgtgtctatcagaggtt  
ccaggggaggggtggggtggggtcagggctggccaccagctatcagggccagatgggttataggctggcaggctcag  
atagggtggtaggtcaggttgggtgctgggtggagttccatgactcccaggagccaggagagatagaccatgagt  
agagggcagacatgggaaggtgggggagccagcagatagcagcattttcattctactactacatgggactgctc  
ccctatccccagctagggcaagtgccctgactcctatgttttcaggatcatcatctataaagtaagagtaata  
attgtgtctatctcatagggttattatgaggatcaaaaggagatgcacactctctggaccagtgccctaacagttca  
ggacagagctatgggttccctatgtatgggtcagtggtctcaatgtagcagggcaagttccagaagatagcatcaac  
cactgttagagatatactgcccagttctcagagcctgatgttaatttagcaatgggtgggacctcctccagtagaa  
ccttctaaccagctgctgcagtcaaaagtcgaatgcagctgggttagacttttttaatgaaagcttagcttctatta  
aagattaagctcctaagcagggccacagatgaaattgtctaacagcaactttgccatctaaaaaaatctgacttcc  
tggaaacatggaagcccagggtctgaacatgagaaattttaggaatctgcacaggagttgagagggaaacaaga  
tgggtgaagggactagaaaccacatgagagacacagaggaaatagtgtagatttaggtggaggttaaatgaaagagaa  
gtgggaattaatactactgaaactctctatattgtcaggtgcccattttatgatatttaataatctcatatcatat

ggtaattctgtgagatatgtattattgaacatactataattaactaataatgataagtaaacacctcttgagactta  
gtatatgctagaatcaaaatthaagtttatcatatgagggccgggacgggtggctcatatggtgatcacatgctgt  
aatcccagcactttgggaggccaaggcaatggatcaoctgaggtcaggagttccagaccagcctggccaacatgg  
tgaaccctctctactaaaaatacaaaaaatcagccaggtgtggtggcagcogtctataatcccagctactca  
ggaggctgagggcaggagaatcacttgaaccaggaggtggagggtgcagtgagctaagattgcaccactgcaactcc  
agcctaggggcagagtgagactccatctcaaaaaaaaaaagaagtttattatagaatlaacttagtttact  
cacaccaatactcagaagtagattattacotcatttattgatgaggagcccaatgacttgtagtgtagatcaact  
tattgaaagcacaagctaataagtagacaattagtaattagaagtcagatggctctgagctctcactactgtacat  
tacatgagctcttatttaactggggactcgaanaatcaagacatgaaataatttgtccaagcttacagaaccacca  
gtagtaaggctaggatgtagaccaggtctctgctacctctgaagacagtggtttttccacagcaaaacacaaactca  
gatattgtggatgagagaattagaagtagatattcctgcccgtggcccttgcctcttacttttactcttctgtcg  
attggaagttgtggtccaagccacagttgcagaccatacttctcaccataattgcatttcttcaggaaagtttg  
agggagaaaaaggtaaagaaaaattagaaaaacttcagaataaagagattttctctgggttacagagattgtc  
atagacaaaattataagcagacacttgagaaaactgaaggcccatgctgcccacaaatlaacctllgacctllggt  
caagctgcaactttggttaaaggaggtgttatgtgttatagtggttcaatttactcttctggtcacaaccttggct  
ccgtctcactcctgagtgacctcagtgccctcagaaaaacatacatatggttctgctagtttaagtttgggtgaaattc  
taactagcgtcaagaactgagggccctaaactatgctaggaatagtgctgtggtgctgtgataggtacacaagaaa  
tgagaagaaactgcagattctctgcatctcccttgcggggtctgacaacaaagtttccccaaaatltacaaatgc  
aagccatttctccatagctaaactactttaaaaatcatttgggggttcacattgtcttctcactctgtaaaaagaat  
ggaagaactcattcctacagaactccctatgtcttccctgatgggctagagttcctcttctcaaaaatlagccat  
lallglattlctcttaagccaaagctcagaggtcttgtatgcccagtgacatgcacactggctcaaaagtaggct  
aagtagaagggtactttcacaggaacagagagcaaaagaggtggglgaaalgagagggtaagtgagaaaagacaaat  
gagaagttacaacatgatggctgtgtgtcaaatctccttagggaaattattgtgagaggtctgaatagttgtgta  
aaataagctgaatctgctgccaacattaacagtcagaanaatccctccgaataactgtacctccaatttctttaa  
ggtagcatgcaactgtaatagttgcatgtatattttatcataataactgtaacagaaaaacacttactgaatatata  
ctgtgtccctagttctttacacaataaactaaactcctcactcataattctattagctaatacatattatcactcta  
tatttcagagacttcaagaagtaagcaacttgcctcaagatcatcctaagaagtaggggggtttctgggctcattt  
ggccctcctaatctctcatggcaacatggctgectaaagtggtgattgcctbaattcactcagggatgggctcata  
ctcactgcagaccttaactggcatcctcttcttctatgtgatctgctgacctagtagacttatgaaatttctga  
tgagaaggagagagagagaaggcagagctgactgtgatgagtgatgaagglgcttctcactctgggtaccagtg  
ggcctcacaagactaagctcactgtctcactgtctctagccagttccttacagcttgcctgatgggagatagag  
aatgggtatcctccaacaaaaataaatttctcattctcaaggtccaacttatgtttcttaatttttaaaaaaa  
tcttgaccattctccactctcaaaaataatccacagtgagagaaacattcttttccccatcccataaaactcct  
atlaaataatggaaaatctgggcatgggtgtctcacacctgtaatcccagcactttgggaggtgaggtgggtggact  
gcttggagctcaggagttcaagaccatcttgacaacatgggtgatccctgctctcaaaaaagtacaaaaattag  
cctggcatgggtgtgtgcacctgtaatcccagctattaggggtggctgagggcagagaaattgcttgaaccgggagg  
cggaggttgcagtgagctgagatcgtgocactgcactccagcctgggggacagagcaccattataatlaactgtat  
ttttacttggactcttgtggggaataagatacatgttttattctattttatgattcaagcactgaaaaatagttt  
tagcatccagcaggtgcttcaaaaccatttgcctgaatgataacttcttacaagctcagctccctctatccc  
ttccagcactcctctctgatataaataagcttcagtttctccttagttcctgttacatttctgtgtctccatt  
agtgacctcccatagtccaagcatgagcagttctggccaggccctgtcggggtcagtgccccacccccgctctt  
ggttctgtgtaaccttctcaagcaaaccttctggctcaagcacagcaatgctgagtcagtgatgagtcagctgagge  
ttaggggtgtgtgccagatgttctcagcctagagtgatgactcctatctgggtccccagcaggatgcttacagggc  
agatggcaaaaaaaggagaagctgaccacctgactaaaactccacctcaaacggcatcataaagaaaaatggatgc  
ctgagacagaatgtgacatattctagaatataattttcctgaatataatataatataacacataacgtatata  
tataatataatatttgtgtttatcaattggccatagaatgatagttattgtgaaatcaaatatttactctgca  
gggtgctctataccttagaagcggcagaaacaggctttatlaatacatgtgtatagatttttaggatctatacaca  
tgtattaatatgaaacaaggatagggaaggaaggcatgaaaacaggaaaagaaaacaaacttcttggcattt  
taaggcaccctggacagctaggtggcaaaaggcctgtgctgttagaggacacatgctcacatacggggctcagatc  
tgacttgggggtgctactgggaagctctcacttaaggatacatctcaggccagctctgggtgcattaggaagatgta  
ggcaactctgatcctgagaggaaagaaacattcctccaggagagctaaaagggttcacctgtgtgggttaactgtga  
aggactacaagaggatgaaaaaactgacagacagacataatgctgtgtgggagaaaaacaggaggtcaaggggat  
agagaaggcttccagaagaatggctllgaagctggctctgttaggagttcacagtgggcaaaagatgtttcagaaatg  
tgacatgacttaaggaaactatacaaaaaaggaacaaatlaaggagagggcagataaaatagttcaacagacatgcaa  
ggaaatttcagatgaatggtatgtctcactgagctcttgagggttagcagctgtgagggttttgcaggccagga  
ccattacaggacctcagctatacttgacactgtttttgtattcatttgtgaatgaatgacctcttctcagctcta  
ctcggttctcgtgtgaatgaatgatgtctgtcagcctacttgggttctgctaaagacacagagagaagatttagtg  
atgctatgtaaaaaacttcccttttgggttcaagtgatgtttgtgatagaaatgaagacaggtacatgatgcata  
ctaacataaacacaaacattaaagaaaggaaatcaacctgaagagttattatacagatacaaaaatacagagagtgga  
gttaaatgtgtaataactgtggcacaggctggaatatgagccatttaaatcaaaaatlaattagaaaaaaacagt  
ggggaaaaaatccatggatgggtctagaagagactagcattgttttaggttgagtgaggtgtttaaagggtgata  
tcagactaaactgaaatattgtggctaaataactagaatactcttatttttctgtatcatgaatagcagatag  
cttgatggccccatgcttgggttaacatcctgtctgctgacatgaaatccttaatttttgacaaaagggtctat



acagtttgaacacttgttatggctctattctctcattctttacaattacactagaaaaagccacaggcttctctgca  
aggcagccacagaatattatgacttgtgatalccaagtcattctctggataatgcaaaatctaacacaaaaatctagta  
gaatcatttgcctacatctatttttggctctgagaatatagatttagatacataatggaagcagaataattttaa  
ctggcctaaatctagaalccctaagcagctctttctctatcagtggtttacaagccttggcttatattttctctat  
aaaaataaaaaaaagtaagttatttgggtgtaaaagaatattcatttaaagttattttctcttagataataccatgaaa  
aacattcagtggaagtgaggggctacttacttaacaagaatctaatttatataatttttcactactaatagcatct  
aagaacagtaacaatattgactcttcagggttaaacatattgctataaattagccagaaagatttaagaaaaattgg  
atgcttctcttggtaaaatagggcatcttacagtttttagaatcctgcatagaacttaagaaattacaatgctaaa  
gcaaacccaaacaggcaggaattaatcttcatcgaatttgggtgtttcttctaaaagtcctttatacttaaagt  
cttaagacatacatagatttttactaattttaattatagacaataaagaatattcttactgattacttt  
ttctgactgtctaattcttctgactctctggatggccataaaccttatctctctgaactttgggcttttaatat  
aggaaagaaaagcaataatccattttctcatygtatctcatatgataaacaaaataaaatgcttaaaaatgagcaggt  
gaagcaatttatcttgaaccaacaagcatcgaagcaataatgagactgcccgcagcctacctgactctctgagtcag  
gatttataagccttgttactgagacacaaacctgggctttcaatgctataacctttcttgaagctctctctacc  
accttagccataaaggaaacctggaatgggtcagatccctggatgcaagccaggtctggaacctatggcagtaagg  
agagaagaaaaatgtggctctgcaactggctccgaggagcaggagaggatcaaccccatactctgaatctaagag  
aagactgggtgtccatactctgaatgggaagaatgatgggataccatagggtgttttagggagaaacctgttc  
tccaaactctggccttgagataacctggctcttattctctggactttggcaatgtctgacctcacattcaagttc  
tgaggaaagggccactgcttccatactgtggatctgtagcaaatccccctgaaaaccagagctgtatcttaattg  
gttaaaaaaattatattatctcaacgactgttctctctgagtagccaagctcagcttggltcaagctacaagca  
gctgagctgcttttgtctagtcattgttcttttatttcagtggtcaaaatcgttctttccaaacctaggatctt  
gtcttcttaggctatataatttgtcccaggaagtcttaactgggggtccacagaaacctagggggtgtgaagtt  
tatagaaaaaaatctgtatttttacttacatgtaactgaaatttagcattttcttctactttgaatgcaaggac  
aaactagaatgacatcatcagtaacctattgcatagttataaagagaaaccacagatattttcactacacctag  
gtattgcagatcttttgttttggttttgtttgagatggagtttctgctcttattgcccaggctggagtgagtg  
catgatttgggctcactgcaacctccccctctctgcatcaagcaattctctgcttggcctctctgagtagctggg  
gattacagggcactgccacctgcccagctcaattttggatttttagtagagatgggggttctgcccagctggccag  
gctggctctgaaactcctgacctcagatgatctgcccgccttggcctctctgaagtgtctgggattaggtgtgagcc  
accagcctggcccattgcagatatttttaattcacatttatctgcatcactactggatcttaaggtagctgtag  
accaatcttagatctaatgctttcataaagaagcaataataaataactataccacaaatgtaattgtttgatgt  
ctgataatgatatttcagtgtaattaaacttagcactctatgtaatttttagtgcaataaaaaactattttt  
tagcacttacagctctgcaaacctggcctgtgacacaaaaaagtttaggaattcctggtttgtctgttagcca  
atgggtagaatatagtctcagaagataccatgggttaatagctaaaagaaaaaggagtagaaaatcagtggcctg  
gaataataacaatttgggcagtcatttaagtcaggtgaagactcttggaaatcatgggagaaaaagcaagggagacatt  
cttacttgccacaagtggttt  
gggtatagaagagaaaaagcctcttagtaaaacttggaaataggaatccccaaagccacttgaactgggagacagga  
gccatactgctaagtgaaaaagacgaagaacctctagggctgacatacaggaaattgttaggaacagaaaatctct  
agatctgggtggggcaaggggagccataggagaaagaaatggtagaaatggatgggagcggagggcagaggtgggag  
atcatgaggtcaagagatcagagacctctggcaaacatggtagaaatccccctctactaaaaataaaaaattag  
ctgggcatgggtggcatgcccctgtagctccagctgctcgggaggtgaggcaggagaaatcgtttgaaccaggagg  
cgaaggtgagctgagctgagatagtgccattgcaactccagctctggcaacagagtgagactccgtctcaaaaaaa  
aaaaaaagaaagaaagaaaaagaaaaagaaaaagaaaaataaatggatgtagaacaagccagaaggaggaaactg  
ggctggggcaatgagattatgggtgatgtaagggactttatagaataaacaatgctggaaatttgggaaactctgct  
tctatttatcccccaatcattactctctgacattgatagttaaataatctctgtgaatttatctctgattctaa  
aatatgaggataatgacaatgggtattataagggcagattaaagtgatagcatgagcaatatcttccagggcacatg  
gatcgaatgaaatacactgtaaatccccacttccagttccagctctaccaagtaaaagagctagcaagtcataaaaa  
tgggyacatacagaaaaaaagggacactagaggaataatatacctgactctagcctgatttaataatctgatt  
cactttttctctgtttgatgacaaatctggcttttaataaattttaggattttaggcttctcagctccctcccc  
gtgagaagtataagcaggacagacagggcaagcaagaagagagccccaggcaatactcacaagtagccaatgtccc  
ctgtggtcatagagaaatgaaaagagagaggattctctggaagcactggatgtaattctttctgtctgctctct  
agggaaatcccccaaggtactgtactttgggattaaggttttagtcccactgtggactacttgcattctgttcag  
ttctagaaggaactatgtacggttttggctctccctagagaaactaaggtacagaagttttgtttacaatgcaactc  
cttaagagagctagaactgggtgagattctgttttaacagctttattttcttttcttggcctgtttttgtcact  
gtcccaccttttaagggcaatgttaaatgogctttggctgaaacttttttctctattttgagattttgctctcttat  
atgaggtcttctctggaaaggagaatgggagagatggatcatctttggaaagatgatgaagagggtaaaaaagggg  
acaaatggaaatttgggtgagatagatgaggagcaacaaaaaagagcctcaggatccagcacacattatcaca  
aacttagtgccatccatcactgctgacctctccggacctgactcccccctgagggacacaggtcagccttgac  
caatgacttttaagtaacctggagaacagggggccagaaactccggcagtaaaagataaaaagccagacagagaggc  
agcagcacatattctgcttccgacacagctgcaatcactagcaagctctcaggcctggcatcatgggtgcaatttact  
gctgaggagaaaggtgcccctcactagcctgtggagcaagatgaatgtggaagaggctggaggtgaagccttgggca  
ggtaagcatlggllctcaatgcatgggaatgaaggggtgaatattacctagcaagttgattgggaaagctctcaag  
atttttgcatctctaattttgtatctgabatgggtgcatttcctagactcctcgttgtttaccctggaccagag  
gatttttgacagcttggaaacctgctgctccctctgcccactcctgggcaaccccaaggtcaaggccatggcaa

gaaggtgctgacttcccttggagatgctattaaaaacatggacaacctcaagcccgcttttgctaaagctgagtgag  
ctgcactgtgacaagctgcatgtggatcctgagaacttcaaggtgagttcaggtgctgggtgatgtgatttttggc  
tttatattttgacatbaattgaagctcataatcttatggaaagaccaacaaagatctcagaatcatgggtcgag  
cttgatgttagaacagcagacttctagtgagcataaccaaaacttacctgatcagaactagtgacagttaaggac  
tactaacagcctgaattggcttaacttttcaggaaactcttgccagaacttgatgtgtttatcccagagaattgtat  
tatagaattgtagactgtgaaagaagaatgaaatttggcttttggtagatgaaagtccatttcaaggaaatagaa  
atgcctattttatgtgggtcatgataattgaggtttagaaagagattttgcaaaaaaaataaaagatttgcctca  
aagaaaaataagcacatcttctaaaaatgttaatttcccatcagttattgtgaccaagtgaggcttgtttccg  
aatttgttggggattttaaactcccgtgagaactctgagcactcacattctacatttcaaaaaattagacaat  
tgcttaaagaaaaacagggagagaggggaacccaataactggtaaaatgggggaaggggtgaggtgtaggtagg  
tagaatgttgaatgtagggctcatagaataaaattgaacctaaagctcatctgaatttttgggtgggcaacaaact  
tggaaacagtttggagtcaggggtgtctaggaatgtaggtataaagcgtttttgtttgtttgttttttctc  
aagttgttttggaaaacttctactcaacatgctgtgtgttattttgtcttttgcctaacagctcctgggtaactg  
gatgggtgatlatctggctactcactttggcaaggagttcacccctgaagtgacaggtgcctggcagaagctgggtg  
tctgtgtcgccattgcccggccataagtaacctgagttctctctccagtttgcaggtgttccgtgacccctga  
cacctcctctcgacatggggactgggcttggccttgagagaaaagccttctgtttaaataaagtaacatttctca  
gtaatacaaaaattgcaattttatctctccatcttttactctgtgtttaaaggaaaaagtttcatgggctgagg  
gatggagagaaacataggaagaaccaagagcttccctaaagaatgtatgggggcttgtaaaatlaatgtggatgtt  
atgggagaattccaggattccaaggaggatgatgatggagaaaaatcttctcgggggtgggaaaaatggttaatt  
aagtgagacagagactcctaggcagtttttactgcacoggggaaagaaggagctgttagtggtacctgagaaagcag  
atgtgtggtacatgtcactttcattaaaaaacaaaaacaaaaacaaaaacttcatagatatccaagataatagg  
ctagaattactattttaatttactcttatttaccatttgaagtagctagcttgtcacatgttttatgaaattgat  
tggagataagatgagtggtatcaacaatagcctgctcttccatgaaggatccattatctcatgggttagctga  
agctaagacacatgatatactattgtgcattatctctgatagaatgtaacatgcactaaaaataaagttagagtagg  
acctgagtgggaaagttttggagagtgatgaagacttccgtgggagatagaataactaataaaggcttaaatt  
ctaaaaaccagcaagctagggcttctgtgactgcatgaaactggctctctggaagtagaaggagagtaagacatac  
gtagaggactaggaagaccagatagtagcagggcctggctacaaaaatacaagcttttactatgctattgcaatac  
taaacgataagcattaggatgttaagtgactcaggaataagatttgggaaaaagtaactctgcttatgtgcacaa  
aatggatccaagttgcagataaaaaataaaataggtgatgattcaaggggacagatacaatggttcaaaccaag  
aggagcagtgagctgtggaatttgaaggatggacaaaggtgggggtgagaaagacatagttatcgactgactgtgg  
gagatgagaaggaagaggaggtgataaattgactgaaagctcccagactggtagataacaggaggaaacctgca  
actgacctggtgacttccatgtgtggaagggtagaggatataacagatttacttttaggaagtgctagatgggt  
cagggagtttgacctcaggtctgtgtcttccatcaaggaaacttgcatttccaagttagagtgccatata  
tttggcaataataactttatagtaattttatagtgctctcacattgatcagacttttctctgtgaattactttg  
aatttggctgtatataccagaatatgggagagagacaaataatlatgtagttgcaggctatcaacaactggt  
ctctctgagccttataaccttcaatatgcccataaacagagtaaacagggatattcatggcactaaatatttcc  
acctagtcagtcacaaaatgggagcaatgtgcatttttgatacataattttatataatttatggggtacatgtgat  
acttacatgcctagaacatgtgatgattlaagctagatatttaggatatccattgctttgagcatttatcattct  
atgtattgagaaaaattcaaatcctcatttctagaaatataataaataaactgtaataactcaactcaccac  
cctactcaaatatacaacattatggcttaactctctatccaactgtgtttgtacctattaaccaacatctcttaa  
atccccctcccatacacactcacacttttccagcctctgataactatcattctactctctaccaccatgagaccca  
cttttttagctcccacagatgaataaaaaacatgtgatattgactttctgtatctggcttattttatbatctatct  
ctttggcataccaagagtttgtttttgttctgcttcagggctttcaattaacataatgacctctgggtccatccat  
gttgctacaaaatgacaagatttcatcttttccatggcaaaatagtaactgtgcaaaaatacaattttttaactcgt  
tcactctgtgatagacacttaggtgatcccaaccttaactattgtgaatagtgcttcaataaacatgagtgtaa  
tgtgtccattggatatactgatttcccttcttttggataaataaaccactagtgagattgctggattgtatgatagt  
tctgttttagtttactgagaaactctcactgctttccataatgggtgtaactatttcaatcccacaaacagct  
gtgtaagaaagagttcccttttctccatctctcacaaggatctgttattttttgtcttttttgttaatagcgtt  
tlaactagaglaagtagatatactcattgtagttttgatttgcatttccctgatcattagtgatgttgaatttt  
tcataatgtttgttggctattgtatatacttttctgagaattgtctgttccatgtccttagcctacttttatggg  
attgtttgttattttcttgataactatttctgttccattttagagcctggatattattctttgtcagatgtatag  
attgtgaagattttctccactctgtgggttctgttttattctgcagactcttcccttttggcatgcaaaagctct  
ttagtttaatttagtccagatattttctttgtttttatgtatttgcatttgtgttcttgggtcatgaaatccttcc  
ctaagccaatgtgtagaaggggttttccogalgttatttctagaattggttacagtttccagggcttagatttaagt  
cttgatccatcttgagttgatttttgtataaggtgagagatgaagatccagttcattctctacatgtaactgctc  
cagctatcccgcaccatttgttgaataggggtgccccttcccactttatgtttttgtttgctttgtcaagatca  
ggtggatgtaagtatttgagtttatttctgggtctctatctctgttccattgggtgatgtgcttatttgtacacca  
gcatcagctgttttgggtgactatggccttattgtatagttgaaatgaggtaatgtaatgcttccagatttgttcc  
ttttttttagacttgcctgtttatttgggctcttttttgggtccataaagaattttaggatgttttttctagtctg  
tgaagactaatgggtggtattttgatgggaattgcaatgaaattttaggttgcctctggcattatggcattttccac  
aatattgatttcaaccatctatgagaatggcatgtgtttccatttgtttgtgtcttatatgattactttcagccgt  
gttttgtagttttcttctgtagatgcttctccctccttgggttaggtatataatccaaagttttgtttgtttgt  
tttgttttttgcagctattgtaaaaggggttaggttcttgattttatctcagcttgggtcatttgcctggtatgtaag

aaagcaactcattgggtgacgttaattttgtatccagaaactttgctgaattatcttatcagttctagggggttt  
ggaggagtcttttagagttttctacatacacaatcatalcatcagcaaacagtgacagtttgactttctcttaaca  
atctggatgtgcttactttgtttctctgtctgattgctcttgctaggacttccagtaatatgttaaagagaagt  
gtgagagtgggtatccttgtctcattccagttttcagacagaatgcttttaactttttccattcaatataatggt  
ggctgtgtgtttaccatagctggctttttatcattgaggtatgctcttgtaaaaccgattttgctgagtttagt  
cataaagtgatggtgaattttgttgaatgcagttctgtggctattgagataatcacatgattttgtttccaatt  
ctctttatggtgtatcacacttattgacttgcgtatgttaaaccaatccgtgcacccctogcatgaaacccactt  
gatcatgggttttgatagctgtcggatgctattagctagtattttgtcaaggatggtggcatctatgttcatcag  
ggatattgatctgtagtgtttttttttttgggtatgctcttccagttttgggtattaaggatgactggctca  
tagaatgatttagggaggattctctcttctctatcttgtagaatactgtcaataggattggtatcaattctctct  
tgaatgtctggtagaattcagctgtgaatctatctggtcctggactttttgttgttggtaaatttttatcat  
ttcagctctgctgcttattactggctgttcagggtatctaattctcctgacttaagctagagccctgtatcttt  
ccaggaatcgaacgtctccttaggtttctagtttatgcatgtaaagggttccatagtagccttgaataatctt  
ttgtatctctgtggtatcagtaaatagttatctctgtttgtttctaaattgagtttatttgcacttctctctctt  
tcttggtaaatcttgcataatggctatcagttttatttatcttttcaagaaccagctttttatttcattagctt  
ttgtattttttgcagttgttttaatttcatttagttctcctcttatcttagttattcccttctctttgctgggtt  
ttggttctgtttgtttttgtttctctagttcttctgtggtgtgaccttatattgtctgtctgctctcttccagactc  
tttgacatcgacatttagggctgtgaactttccttttagcaccatcttctgtatcctagagggtttgatagggtt  
gtgtcactattgtoggcagttcaagtaactttgttgttcttattactttaagttctgggatcacatgtgcagaa  
tgtgcagggtttgttacataggtatagatgtgccatggtgggttgcctgcacccatcaacctgtcatctacattagg  
atctctttaatgttatccctctcctaaccctcaccctccgacagggcctgggtgtgtgatgtccctccctgt  
gtccatgtgttctcattggtcaactcccacttatgagtgagaacgtgtgggtgtttgggtttctctgttctctgtt  
gtttgcacagaatgatggtttccactctcactatgctcctgcacaaagacatgaactcatctttttatggctgt  
agtattccatgggtgatattgtgcccactttctttatccattatctogctgatggcatttgggttgggtccaagt  
ctttgctattgtgaatagtgccacaataaacatacgtgtgcacgtgtctttatagtagaatgatttctaattctt  
gggtatatacccagtaattgggtatgctgggtcaaacagttttctggtcttagatccttgaggaatcggcactg  
tcttccacaatggttgaactaatttacacacccatcaacagtgtaaaattttctctatcttccacatctctcca  
gcacctttgtttcctgactttttaataatgcccattctaactggcatgagatgggtatctcattgtgggttttgatt  
tgcatttctcfaatgaccagtgatgatgagcttctttcatgtgtttcttggccacataaatgacttctttagaga  
agcactctgttccatctcttctcacttttgatgggtctgttaggtttttcttgtaaaattgttgaagtcttt  
gtagattttggatggttagccctttgtcagatggatagattgcaaaaattttctccattctgtaggttgctctgtc  
actctgatgattcttttgcctgtgcagaagctctttagtttaattagatcccatatgtcaattttggcctttgtt  
gtcatgtcttttgatgttttagtctgtgaattttgcccatagctatgctcctgaatgggtattgcccagggttatctt  
taggatttttatggttttaggttgcacatttaagctttaaaccacttgagtttaattttgtataagggtgaagg  
aaggggtacagtttcagttttatgcatattgctagccagttttccagcaccatttataaatagggaattctttc  
tccattgctttttgtgatgtttgtcaagaacagatggtcgttagatgtgtggcatatttctgaggctctgttctg  
ttccactgggtctatatactgttttgggtaccagttaccatgctgtttttgttactgtagccttgtagtagttga  
agtcaggtagcatcatgctccagctttgtctttttgttttaggattgtcttggctatatgggctctttttgatt  
ccatagacatttaaagtagtttttctaatctttgaaaaaagtcagtggttagcttgatggggatagcatgaat  
ctataaataactttgggcagtagttggccattttaaagatattgattctttctatctatgagcatggaatgttttcc  
atgtgtttgtgtctctcttatttctctgagcagtgagtggtttgtagctctccttgaagaggttcttccacatcc  
ttagaagttgtatttctaggtattttattttattctctttgcagcaattgtgaatgggagttcaccatgatttgg  
ctctctgctttgtctatatttgggtgtataggaacgcttgtgatttctgcacactgattttgtatcttgagactttgc  
tgaagctgtttatcagcttaagattttgggtcagatgacagggctcttcaaatatacaatcatgtcatctgcaaa  
cagagacaatttgacttctctctctctatttgaatatgctttatttcttctcttggcctgattgtcctggcgaga  
acttccaatactatgttgagtaagagtgccgagagggcatcctgtcttgtgcccgggttttcaagcaaatgatttt  
taaaattccactcttgatttctgttgacccaatgatcattccaggagcaggttatttaattccctgtattttgcat  
ggttttgaaggttcctttttagttgatttccaattttattctactgtggtctgagagagtgcttgatataatttc  
aatttttaaaaatttattgaggcttgttttgggcatatcatatggcctatcttggagaaagttccatgtgctgat  
gaatagaatgtgtattctgcagttgttgggtagaatgtcctgtaaaatctgttaagtccatttgttctttaaactc  
cattgtttctttgttagactgtcttgatgacctgcctagtgccagtcagtgaggatttgaagccccactattatta  
tgttctgtctagctagtagtaattgtttataaatttgggtatctccagtttagatgcatatataaagaatt  
gtaataattctccattggacaagggttttatcattatagatgtccctcttctgtctttttaaactgctgttctt  
taaagtttgtttgttgacataagaatagctgctttggctcgttttgggtgtccattttgtgtggaatgtcatttt  
ccaccctttaccttaagtttatgtgactcttattgtttaggtgagctcctgaaggcggcagataactgggtgg  
tgaattcttattcattctgcaattctgtatcttttaagtggagcatttagtccatttaccattcaacatcagtttg  
aggtgtgaggtactattccattctctgtggtattttgttgcctgtgtatcttttatctgtattttgttgtatag  
tctatgggatttatgctttaaagaggttctgtttttagatgtgcttccaggggtttatttcaagatttagagctcct  
ttatcagttctttagtgttggcttggtagtgcgaattctctcagcatttgttttctgaaaaacactgtgtatt  
ttcttcatttgtgaagcttagtttcaactggatataaaatcttggctgataattgttttgttttaagaaggctgaag  
atagggccatattcaactctagcttttaagggttctgctgagaaatctgctgttaactctgatagggtttcttctcat  
aggttaactggtagtttcaactcacagctcttaagattctctttgtctttagataactttggatactctgatgaca  
atgtacctaggcaatgatattttgcaatgaatttccagggttatttagacttcttgtattttggatattctagggt



aacttttcaaaaattggtacatgctttaactttaaactacaggcctcactggagctacagacaagaaggtgaaaa  
cggctgacaaaagaagtcctggatcttctatggtgggagaagaaaactagctaaagggagaataaataagagaa  
aaatgggaatgactgaatcggacaaggcaaggctataaaaaaaatagcagcagatctcttgggggcccct  
tccccacactatctcaatgcaaatatctgtctgaaacgggtcccctggctaaaactcccacccatgggttggccagcctt  
gccttgaccaatagccttgacaaggcaaaccttgaccaatagctcttagagatccagtgaggccaggggcccggggc  
tggctagggatgagaataaaaagggaagcacccttcagcagttccacacactcgttctggaaactctgaggttatc  
aataagctcctagtccagacggcatgggtcatttccagaggagacaaggctactatccaaagcctgtggggcaa  
gggtgaatgtggaagatgctggaggagaaaaccctgggaaggttagcctctggtgaccaggcaaggagggaagga  
gacctgtgctggcaaaaagtcagggtcgtctctcaggatctgtggcaccctctgactgtcaaaactgtctctgtca  
atctccaggctcctgggtgtctacccatggaccagaggtctcttgacagcttggcaacctgtcctctgctct  
gccatcatgggcaacccccaaagtcaaggcacatggcaagaaggtgctgactccttgggagatgccataaagcacc  
tggatgatctcaagggcacccttggccagctgagtgaaactgcactgtgacaagctgcatgtggatcctgagaactt  
caaggtagtccaggagatgttccagcactgttgcctttagtctcaggggcaacttagacaactgagtatgactctg  
agcacagcaggggtgtgagctgttgaagatactgggggtgggagtgagaactgcagaggactaacctgggctgag  
accagtggaactgtttagggcctaaggagtgccctgaaaaactagatggcaacttgcactttagaagaaagag  
aggtggaaatgaggaaaatgactttctttagatctcggtagaaagaacttccaccttcccctattttgtt  
attcgttttaaaacatctatctggaggcaggacaagatgggtcattaaaaagatgcaggcagaaggcatatattgg  
ctcagtcaaaagtggggaacttgggtggcaaacatacatattgctaaggctattcctatatcagctggacacataaa  
aatgtgctaatgcttcatcaaaacttatatcctttaattccagatgggggcaaagatgtccaggggtagggaa  
caattgaaacatttgggctggagtagatttggaaagtcagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt  
ttgtgtgtgtgtgagagcgtgtgttcttttaacgtttccagcctacagcatacaggggtcattgggtggcaaga  
taacaagatttaaatatggccagtgactagtgctgcaagaagaacaactaccctgcatttaattgggaagcaaaat  
ctcaggtttagaggaagttaacataggcttgattctgggtggaagcttgggtgtgtgtgtgtgtgtgtgtgtgt  
ggagctctcagctcactatgggttcatcttattgtctcctttcatctcaacagctcctgggaaatgtgctggtga  
cgttttggcaatccatttggcaagaattcaccctgaggtgcaggctcctggcagaagatgggtgactggaggt  
ggccagtgccctgtcctccagataccctgagctcactgcccatgatgcagagcttcaaggataggctttattct  
gcaagcaatcaaaataaataatctattctgtcaagagatcacacatgggtgtctcagttctttttttatgtcttt  
taaatatagagccacaaaagggttttatgttagggatgtgtttatgtgtatttatacatggctatgtgtgtttgt  
gtcatgtgcacactccacactttttgtttacgttagatgtgggttttgatgagcaaaataaagaactaggcaata  
aagaacttgtacatgggagttctgcaagtgaggagtaaaaggtgcaggagaaatctgggtggaagaagacctcta  
taggcagagctcctcagaaaacagatgtttggaagagatggggaaaggttcagtgaagggggtgaacccccctc  
cctggatgacagcagcagcaggaagggtcacaagaaagaaaggtgtccaagcttaggaagtcagggttta  
ggcagggatagccattctattttataggggcaatactatttccacggcatctggcttttctcagccctgtgag  
gctctacagggaggttaggtgttagagatcagagcaggaacaggtttttcttccacggtaactacaatgaagt  
gatccttactttactaaggaacttttcttttaagtgttgacgcacgctaaagaggtgaaatcaatccatacc  
ttaagtctacagactggctcagcagcatttcaaggaggagacctcattgtaagcttctagggaaggtggggacttaggt  
gaaggaaatgagccagcagaagctcacaagtcagcactcagcgtgtcatgtctcagcagcagaacagcagcaggtcaga  
tgaataatagtgtagaagaatttgtataacatttaattgagaagcagattcactggagttcttatataattgaaa  
ttaatgcagtttaatacagcagcaggtttagtttaagtgttgactgaggtgtatgaaactaacgcttgtgtctccagaaat  
tcacatgctgaaatccccaaactccaattggctccatttggggggaggcttggaaaagtaatcaggttttagagga  
gctcatgagagcagatccccatcatagaattattttcctcatcagaagcagagagattagccatttctctcctc  
tggctagggacacagtggggaagtcagccactgcaaccagggaagagagcctgaccaggaaccagcagaaaagtga  
gaaaaaatcctgtgttgaagtcacccagctctatgctattttgttatagcaccctgcaactaagtaagcagatgaa  
gaaagagaaaaaataagctcgggtgtcagtggttagaaaccatgtttatctcaggtttacaaatctccacttg  
tctctgtgtttcagaataaaaataccaactctactactctcatctgtaagatgcaaatagtaagcctgagccctc  
tgtetaactttgaattctattttttcttcaacgtactttaggcttgtaattgtgtttatatacagtgaaatgtcaag  
ttctttctttatatttctttctttctttttctcctcagctcagagttttccacatgccccttctacttccagga  
acttctttctccaaagctctctgctggctccatcaaatcataaaggaccacttcaaatgocatcactcactac  
catttccaaattcgcactttctttctttgtcttttttttttagtaaaacaagtttataaaaaattgaaggaata  
aatgaatggctacttcataggcagagtagacgcaagggtcactgggtgcccattttttattgttatttttcaatagt  
atgctaaacaagggtagatttttatgctgccatttttagaccataaaagataacttctctgatgttgccatggc  
attttttctcttttaattttatcttctatttttaatttogaaggtacatgtgcaggatgtgcaggcttgttaca  
tgggtaaatgtgtgtctttctggccttttagccatctgtatcaatgagcagatataagctttacacaggatcatga  
aggatgaaagaatttccaatatataataaattcaactcaacctgatagcttaggggataaactaatttgaagat  
acagcttgcctcagataagcagaattccagagcttctggcattataacttagcaaggttagagatcatggatcac  
tttcagagaaaaacaaaaacaaactaaccaaaagcaaacagaaacaaaaaccacataaatacttctcaccctg  
ttaatggttcaaatatgtcagaaacagcactgtgttagaataaagctgtctaaagtacactaatattcaggttata  
atagtggtggactattagtcataaaaaacaccccttgcctctttagagttgttttccatgtacacgcacatctta  
tgtcttagagtaagatccctgagaagtgaaactagcatttatacaagataaataattctaatccacagtaacctgc  
caaagaacattctaccatcatcttactgagcatagaagagctacgcaaaaccctgggtcatcagccagcacaca  
cacttatccagtggaatacacatcatctgggtgatacacatacactgaatatggaatcaaatatttttctaag  
atgaaacagtcagatttatctcaaataggtacggataagtagatattgaggtlaagcattaggtcttatattatgt  
aacactaatctatctactgcgctgaaactgtggctttatagaaattgttttctactgcactatgagaaatgaagaga

taatggcaaaagtcacaaagagtataattcaaaaagaagtagcaacttttctcttagaaaccactgctaactgaaa  
gagactaagatttgtcccgctcaaaaatcctggacctaagcctaaaacacatttcacaatccctgaacttttcaaaa  
atttggtacatgcttttagcttttaactacagccctcactggagctagagacaagaaggtaaaaaacggcctgacaaaa  
gaagtccttggatcctctatgatgggagaaggaaactagctaaagggaagaataaattagagaaaaactggaatga  
ctgaatcgggaacaaggcaaggctataaaaaaaattagcagtagctctcttggggggcccttccccacactatctca  
atgcaaatatctgtctgaaacggctccctggctaaactccacccatgggttggccagccttgccttgaccaatagcc  
ttgacaaggcaacttgaccaatagctcttagagtagctccagtgaggccaggggcccggcggctggctagggatgaaga  
ataaaagggaagcacccttcagcagttccacacactcgcctcttggaaactctgaggttatcaataagctcctagctc  
agacgccatgggtcatctcacagaggaggacaaggctactatcacaaagcctgtggggcaaggatggtggaagat  
gctggaggagaaacccctgggaaggtaggctctggtagaccaggacaaggagggaaggacccctgtgctggca  
aaagtcacaggtcgtctctcaggatttggggcaccctctgactgtcaactgtctcttglcaatctccacaggtcctg  
gttcttaccatggaccagaggttctttagcagcttggcaaccctgctctgctctgcatcatgggcaacc  
ccaaagtcaaggcacaaggcaagaaggctgacttccctgggagatgccacaaggcaccctggatgatctcaaggg  
cacccttggccagctgagtgaaactgcaactgtgacaagctgcatgtggatcctgagaacttcaaggtagtccagga  
gatgttccagccctgttgcctttagtctcagaggcaacttagacaacggagtagatctgagcagcagcaggggtg  
agctgttgaagatcactgggttgggggtgaagaaactgcagaggactaactgggctgagaccagtggtaatgtt  
ttagggcctaaggagtgcctctaaaaatctagatggacaattttagctttagaanaagagaggtggaatgaggaa  
aatgacttttcttattagattccagtagaagaacttcatcttccctcatttttggttgttttaaaacatctat  
ctggagcaggaacaagtagtggctgtaaaaagatgcagggcagaaggcatatattggctcagtcagggtgggaact  
ttgggtggccaaacatcattgctaaaggctattcctataatcagctggacacatataaaatgctgctaatgctcatt  
acaaacttatatcctttaattccagatgggggcaagtagtccaggggtgaggaacaattgaaacatttgggctg  
gagtagatttgaagtcagctctgt  
gttctttaaactcctcagcctacaacatcacagggttcatggggcaagaagatagcaagatttaaatatagggcc  
agctactagtgcttgaagggaacaactaccctgcatttaatgggaaggcaaaactcaggctttaggggaagttaa  
cataggttgattctgggtggaagcttgggt  
ttcatcttattgtctccttctatctcaacagctcctgggaaatgtgctggtagccgtttggcaatccatttggg  
caagaattccaccctgaggtgcaggttctcctggcagaagatggtagctgcagtgggcagtgccctgtcctccaga  
taccactgagctcactgccatgattcagagcttccaaggataggtcttattctgcaagcaatacaaaataaataat  
ctattctgctgagagatcacacatgattttctcagctctttttttacatcttttaaatatagagccacaag  
ggtttatattgaggaagtggtatgtgtatctctgcatgctgttgggttgggtgtgtgtgtgtgtgtgtgtgtgtgt  
tattttatagatgtgcattttgatgagcaataaaagcagtaagacacttgtaacagggaggttctgcaagt  
gggagttaaatgggttaggagaaatccgggtgggaagaaagacctctataggacaggaacttccagaaacagatgtt  
tggaaagagatgggaaaagggtcagtgaaagcctggggctggattgattgcagctgagtagcaaggatggtctta  
aggaagggaaggtgtccaagctttaggaattcaagggttagtcaggtgtagcaattctattttatagagggaat  
actatttctaattggcacttagctttccacagccctgtggatgctcaagaaagtgaatttaatcccatgcccctcaa  
gtgtgcagatgggtcacagcattccaaggagagacctcattgtaagactctgggggagggggacttaggtgta  
agaaatgaatcagcagaggctcacaagtcagcatgagcatgttatgtctgagaaacagaccagcactgtgagatca  
aaatgtagtgggaagaatttgtacaacatttaattggaaggcttacttaattggaatttttgtatagttggatgttag  
tgcatctataaagaaagatttaatatgattgggttccagcctaaatgttgggtctcctcaaaaatcacatgct  
gaaatcccaactcccaactgaccttatctgtgggggaccttttgaanaagtaattaggttagatgagctcataag  
agcagatcccatcataaaattatcttcttatcagaagcagagagacaagccatttctcttctcctccgggtgagg  
acacagtgagaagtcggccatctgcaatccaggaagagaacctgaccacagctcagcctcagaaatgtgagaaa  
aaactctgttggtagaccaccagctcttggatatttggatagcacttgactgagtaaggcagatgaagaag  
gagaaaaaaataagcttgggttttgagtggactacagaccatgttatctcaggttggcaagctccctcctgccc  
ctatgttccagtagtaaaatcctactctactctctcatctataagaccaaaataaagcctggcctctctctc  
taactttgatttctcctatttttacttcaacatgcttactctagccttgaatgtctttacatacagtgaaatgt  
aaagttctttattcttttttcttcttcttcttcttctcctcagcctcagaatttggcactgccccttctcttct  
caggaacttctccaaactctctgctggctccatcatalcataaaggctccacttcaaatgcagtcactaccgtt  
cagaatagcacttcttcttcttttggtttttggttttttaaagtaaaagcaaatttcttgagagagtaaaagaaa  
taaacgaatgactactgcataggcagagcagcccgaggccgctggttgttcttttatggttatttcttgatga  
tatgttaacaagtttggattatattatgcttctcttttaggcatatagggttaacttctgacattgcccag  
cattttcttttaatttaactgttaccttaattcaggggtacacgtacaggatagcaggttggttttata  
ggtaaaagtgtgcatgggttttaattgggttttttttcttggtaaggttggtttaagtttcttggttactctggata  
ttaggcctttgtcagaagaatagattggaanaacttttccattctgtagatgtctttcgtctctgaggtagtt  
cttttctgagcagaggtctttagtttaattagattccattgggtcaaattttgccttggcttggctgcaatttcc  
acgttctcatcatgaaatctgtgcccgtgttlatatcatgaatagtattgccttgattttttctaggctttlat  
agtttgggggttttctatttaagctcttaatccatctggagtttaatttggataaggtataaggaaggagtcagtt  
tcattttcagcatalggctagccagttctccccatcatttataaattgaaaatccttccccattgcttgcct  
ttgtcaggtttctaaaagaccagatggttgtaggtacaatatgcagtttcttcaagctcataataaccatctgaaa  
tctcttataaattcatttcttttagtatgtatgctggctcctctgctcactatagtgaggggcaccattagccaga  
gaaatctgtctgtctagttcatgtaagattctcagaatgaagaaaaatggatggcatatgaaatgaaacttcatggat  
gacataggaatctaataatglatttgttgaatttaatgcataagatgcaacagagagaagttgacaactgcaatgat  
aacctggatttgatgataaagagctctatagatcacagtagaagcaataatcatggaaaacaattggaaatgggga

acagccacaaacaagaagaatcaatacttccaggaaagtgcactgacgggtcacttttctggaggggtgagagaa  
aagtggaggttagcagtaactgctgaattcctgggtggctgatggaaagatggggcagctgttccactggtagcag  
ggttttagatgtatgtacctaaagatgatgaggtatggcaatgaacagaaattctttgggaatgagtttagggcc  
atataaggacatgacctgaagtttccctctgaggccagtccccacaactcaatataaaatggtttccctgcataatg  
caaagttgcccactcttttcttcatatcatcgatctctgctcttaagataatcttgggtttggcctcaaaactggt  
tgtcactacaaaactttcccactgttccctaaagtaaaacaggttaactgctctcaactatatacaagtagactaaaata  
ttgtgtctctaataatcagaaaatcagctttaaataatattgggtttaactctttgaaatttagagttctccctgaaata  
cacatgggggtgatctcctaaactttatttcttgtaaggatttatctcaggggtaacacacaaaaccagctcctga  
acctctaagatgaggacagtaagccttaagaatataaaaataaactggttctctctctgcccgtggaagtgtgccc  
tgtctattcctgaaattgcttgtttgagacgcatgagacgtgcagcacatgagacacgtgcagcagcctgtggaat  
attgtcagtgagaatgtcttggcctgattagatataaagacaagttaaacacagcatttagactatagatcaagcc  
tgtgcccagacacaaatgacctaatgcccagcacggggccacggaatctcctatcctcttgcctgaaacagagcagc  
acttctcccccaacactattagatgttctggcataattttgtagatagtaggatttgacatggactattggtcaaa  
tgattcagaggaatctcctttggttcagataagtaactgactactaaatggattaaaaaacacagtaataaaacc  
cagttttccccttacttcccactagtttggtttcttcttctgcttctcctcctcctcctcctcctcctcctcctcct  
tctattctaaaaagtgatgaaatggccgggcccgttggctcacactgtaatcccagcactttgggaggtgagg  
tgggcccagatcacaggtcaggagatcaagaccatcctggctaacaatggtgaaaccccactctcactaaaaatacaa  
aaaattagccagagacagtggggggtgctgtagtcccagctactgggaggctgaggcaggagaatggcgtgaac  
ctgggaggcagagcttggcgtgagcagagatcgccactgcacactccagcctgggtgacaaagcagactccat  
ctcaaaaaaaaaaaaaaaaaaagaaaaagaaagaaagaaagaaaaaaaactgatgaaattgtgtattcaatgta  
gtctcaagagaattgaaaaccaagaaaggctgtggcttctccacataaagcctggatgaaataacaggataaacacg  
ttgttacattgtcacaactcctgatccaggaattgatggctaagatattcgtaattcttctccttttctcagttgta  
cttattcctatttgcagcatccaggttattagcggctgctggcgaagtcttgagaaataaactgcacactggct  
gggtgggggtagtgtaggaaaatggaggggaaggaagttaagtttcaaattaagcctgaacagcaaaagtctccctga  
gaaggccactggattctatcagaaaactcgaatgtccatcttgcaaaaactccttggccaaaccccaccctggag  
tcacaaccccacttgaccaatagattcatttactgagggaggcaagggtggtcaatagattcatttccactgg  
gagaggcaaaaggctggggggccagagaggagaagttaaaagccacacatgaagcagcaatgcaggcatgcttctgg  
ctcatctgtgatcaccaggaactcccagatctgacactgtagtgcatttctcctgctgacaagaaggctgctgcca  
ccagcctgtgaagcaagggttaaggtgagaaggctggaggtgagattctgggaggttaggtactggaagccgggaca  
agggtcagaaaaggcagaaatgcttctgaaagaggatagccctgtgtcttccatagctgactttgcacctgct  
ctgtgatattgactatcccactgctcctgggttctcctaccatggacctagaggtactttgaaagttttggatct  
tgggctctgactgtgcaataatgggcaaccccaagtaaacggcacatggcaagaagggtgctgatctcctcggaaa  
agctgttatgctcagggatgacctcaaaggcactttgctacactgagtgacctgcaactgttaacaagctgcacgtg  
gacctgagaacttctgggtgagtagtaagtacactcacgcttcttcttacccttagatatttgcactatgggt  
acttttgaagcagaggtggcttctcttgtgttatgagtcagctatgggatatgatattcagcagtgaggatttt  
gagagttatggtgctgtaaatcaataactaaaatttggtagagcaaggactatgaaataaggaggccacttacc  
atltgatagctctgaaaaacacatcttataaaaaattctggccaaaatcaaacagaggtttttggatgagggaac  
agaagttgagatagagaaaaataacatcttcttqctcagcgaattttctataaaaaatcaatagtcactttctc  
gcalagctctggaggttagaaaaagatcaactgaacaaagttagtgggaagctgttaaaaagaggattglttcccctc  
cgaatgatgatggtatacttttgtacgcatggtagcaggttcttggttatgagtggttgggaaaaattgtatgtatg  
tatgtatgtatgtatgtatgactggggacttatectatccattactgttcttgaagtaactattatcctactttt  
taaaaggacgaagtctcaaaaaaaaaaatgaaacaatcacaatattgtgggtagtgagttggcatagcaagtaag  
agaaggataggacacaatgggaggtgcagggtgcccagtcataatgaagctgatcttagccataatggtgagag  
ttgctcaaacctctggtgaaaaaggatgtaagtggtatctatcttactgcaagtcagcctgaggcctctctatca  
ctatgtaccatttcttttttcttctcactcccctcccagctcttaggcaacgtgatattgattggtttggcaacc  
cactcagcagaggttttaccctacagatcacaggtctcttggcagtaactaacaatgctgtggttaatgctgtag  
cccacaagaccactgagttccctgtccactatggttggtaacctatggttggtaacctatgtcccaaaaat  
ctcatctccttttagatgggggaggttggggagaagagcagtatcctgectgctgatccagttcctgcatgataaaa  
atagaataaagaaatagctctcctaagaaatcattgtactcttttctgtctttatatttaccctgatccagc  
caaaaggacgcactatttctgatggaaatgagaatgtggagaatgggagtttaaggacagagaagatacttctt  
gcaatcctgcaagaaaagagagaactcgtgggtggatttagtgggtagttactccttaggaaggggaaatcgtctc  
tagaataagacaatgttttacagaaaaggaggtcaatggaggtactcttggaggtgtaagaggattgttggtag  
tg-gtagaggtatgttaggactcaaatagaagttctgtataggctatttattgtatgaaactcaggatatagctc  
atltgtgactgcagttcactctacttatttaacaacatatttttatttataatgaagtggggtagggg  
cttctagagaccaatcagggccaacttgaacttctcttaacgtctcctaatgggtatcaatagaaattatct  
ctaaggcatgtgaactggctgtcttgggtttctctgtacttcatctgctacctctgtgacctgaaacatatttat  
aattccatgaagctgtgcataatgatagatttatcatalgtattttctttaaaggatttttgaagaactaatgaa  
ttgatacctgtaaaagtctttatcacactacccaataaataaataatctcttggctcagctctctgtttctataaat  
atgtacaagttttatggttttagtggttagtattttatctcttctatataatcacacacacatgtgtgcattc  
ataaatatatacaatttttatgaataaaaaattattagcaatcaatattgaaaaccactgatttttggttatgtga  
gcaaacagcagattaaaaggctgagatttaggaaacagcaggttaagtcaggtgatagaggagaataggacatt  
taaaagggcaggtgatataaattagggaaactggatgcagagaccagatgaagtaagaaaaatagctatcgtt  
ttgagcaaaaatcactgaagtttcttgcataatgagagtgacataataaatagggaacgtagaaaaattgatccaca

tgtatataatataatagaactgattagacaaaagtctaaacttgggtatagtcagaggagcttgctgtaattatattg  
aggtgatggataaagaactgaagttgatggaacaatgaagtttaagaaaaaaatcgagtaagagaccatttgggc  
agtgatgacacagaactggaaaacattgtgaaacagagagtcagagatgacagctaaaatccctgtctgtgaatga  
aaagaaggaaatatttgacagaacagcaaatgcctacaagccccctgtttggatctggcaatgaacgtagccatt  
ctgtggcaatcacttcaaactcctgtacccaagacccttaggaagtatgtagcaccctcaaacctaaaactcaaa  
gaaagaggtttttagaagatataatacccttctctctcoagtttccattaatcccaaacctcttctcaaaagtatt  
cctctatgtgtccaccccaagagctcacctcaccatatactcttgagtgaggacacatagataggcggtgctacca  
tctaacagcttctgaaattcctttgtcatatttttgagtccccactaataaccacaaaagcagaataaaataccagt  
tgctcatgtacaaataactcaactgctgtcttgtagcatacattaattaagcacattcttgaataattactgt  
gtccaaaacaatcacactttaaaatctcacacttgtgctatcccttgccttctgaaatgtcactctgtattttaa  
gaagagatgaggggtgaatttctctgtgttacttattgttcaatttctcgatgaggagtttcaacattcaactttagt  
ggaaaacacataagtagacacattacaggaataatataccaaactgacatgtagcatgaatgcttgtgcatgtagt  
catataaaatctttagcaatgtaaacattctctgatatacacatacagatgtgtctatagtctacacaattct  
tatgtccatgaacaaacattccatgacacataagaacacacactgttacagatgcatacttgagtgcattgaca  
aaattaccocagtcacatagagaatttggatttctgcaatttgactctgttagcttggtagactgttctcaattac  
tctgggtgatgtcttccctcattttgccttgtctatcttgaactcactttaaagtcctaaacttatatgttatct  
caactaagaagctattttttttaaatttaactgggetttaaagccctgtctataaaactctgtctacaattatgggct  
cttctctataaattatgtgttttctactaatgtaacttaactctgtctcattgtatattctaccactaaaattta  
acctctttatggtagagacattgctctgtaaactcttatttccctagatttggagatgaaaaaaagattaaat  
tatccaaaatttagatctctcttctacatfatgagttatcacactatccatagagaagtttgtttgagacctaaac  
tgaggaaaccttgggtctaaaatgactatgtgatctcttagtatttataggctcatgaggttctctctctgctct  
gctatagtttgatttagtcaacaagcatgtgtcatgcatatttattcacatcagaatttcatacactaataagacatag  
tatcagagctcagttatttagttatcagttagggtccatcaaggaaaggacaaaccattatcagttactcaacc  
tagaataaatacagctcttaataagttaatatccttgtatttggaaagagctaaaatatacaataaaggacagtgca  
gaaatctagatgttagtaacatcagaaaacctctccgccattaggcctagaagggcagaaggagaaaaatgtttat  
accaccagagtcagaaaccagagccataaccagaggtccactggattcagtgagctagtgggtgctccttggaga  
gagccagaactgtctaatgggggcatcaaaagtatcagccataaaaaaacataaaaaagactgtctgtctgtaggaga  
tccgttcagagagagagagagaccagaaataatcttgccttatgcttccctcagccagtgcttaccattgcagaat  
gtacatgcgactgaaagggtagggaaacctgggaaatgtcagttcctcaaatcacagagaacactgaggggaaggatg  
agaaataaagtgtgaaagcagacatgaatggttaattgacagaaggaaactaggatgtgtccagtaaatgaataatta  
cagtgctgagctgatttatgcaabgattaatgtattgataagataaatatgaaaaacagaaatcaaacacagtgaa  
ctgagatagaatttgggagagcactggcatttaagaatgtcacactagaatgtgtctctaggcattgttcttctg  
cacatacatctcaatattcatfatctgaaaaatfatgaattaggtaaaaagctcaaaataatttatttttcaggtt  
agcaagaactttttttttttttctgagatagagcattgctatggttggccaggtggagtgcaatggcatgac  
caggtcactgcaacatctgctcctcaggttcaagcgattctcctgctcagcctcccaagtagctggcactacag  
gcatgtgccaccacatgctggcctaaatttctatatttagtagataggggggttccaccatgttggtaggctgat  
ctcgaactcctaacatcaggtgatccacctcctcggcctctgaaagtgtgggatcacaggcgtgagccaccaca  
cccagccaagaatgtgaattttagtagaaggatataaccatatttctctgacctagagctcttagtataccctcc  
ataccatctctattaatcaccactcttactgcccaggtcaggtcctcatgtttcctgaacaagagtagatgctatt  
ctgtgtcaaaatctagaaccttatctcctacctgctctgaaaccaacagcaagttgacttccattctaaaccacat  
tggcattacactaataaaatcgatactgagttctaaaaatcctggggattttggggactatgttacttcaac  
ttccttgagatttcaacttaaatgttgggttctcaataaaggctcctcatttaactttgtattcctcacactctgg  
attcacagttatatactaaactctaaatacagcctgtataatcccaattcccaactctgatttcaacctctgacc  
tccaaactcagtgccaaaccatatacaacaatgtactgggcttatattatagatgtcctataggcaccctcag  
actcagcatgggtatllcacttgttatactaaaactgtttctctccaglgtttccatttttagtcattagatagc  
tacttgccatttccaaggtcacagataaaaatcatttccctacctctaatcaacagttcagattctgtctcaatt  
tgtcctatctattaatcaccactcttactgcccaggtcaggtcctcatgtttcctgaacaagagtagatgctatt  
ctttccacttttagaccttatctggtggatgcggtggctcaggttgtaaaaccagcacttggggaggccaagg  
caggcagatcacttgaggtcaggagttcaagaccagcctgaccaacatgggtgaaaaccctctctactaaaaatc  
aaaatcagccgggctgtgtggtgcatgctgcagtcacagctattcaggtggctgaggcaggagaattgcttgaacc  
caggagggcagaggttggggtgagcctagattgcaccattgcactctagcttgggcaatagggatgaaactccatct  
cagaagagaaaaagaaaaaagaccttatctgttatatacaaatcctctcaatgcaatccatatagaataaacatgta  
accagatctcccaatgtgtaaaatcatttcaggtagaacagaattaaagtgaaaagccaagctcttgggaattaca  
gcaaaagatcaaaataacagctcctcatggccttaagaatttaccctaacatttttttagaatcaatttcttatata  
tgaattggaaacataaattcctcctcacaacacattctaaagattttaaaggagatattgatgaagtacatcatctg  
tcatttttaacaggtagtggttagtgattcacacagcacattatgatctgttcttgtatgtctgttccattctgta  
ttcttgacctgggtgtattctttctgagctccagatccacatatactaaagtacatctttttgcattttacaagagtg  
catacaatacaatgtatccaagactgtatttctgattttatcgtaccactaaactcaaaaatgtggccctattctt  
gtgtcagcactgacatcaccgtcatgggtccaagctctgataatagaaatggcatgtgcacttctctccctactgca  
acagaagcccagctatttgtctcccatttctctactcttaaaatcacattcttccactaagtgagaataatctttt  
aaagacacaaatcaaacatgcccacccttcttgaattattcaatactcttctgttggctccagggttacagaaa  
aataacttgaacaaagtttlaaaggtcattcaggtcctctctaccctattttataacatttcccttgtgatca  
gaatctcaggcacatccatctttctatatacaaaataaagtcataatagtttgaactcactctgggtactttta



gcaagatgttggcctaaaaatccaaataccagagaattcatgagaacatcacctggatgggacatgtgccgagca  
cacacaattactatagcttaggcatttgctatcttcatattgaagatgaggaggtcaagagatgaaaaagacttgg  
caccttgttgtatattaaaaatatttggtagagtagagcttttgaagagctaggaggtgggagctaaatgat  
gatcacacalggacacaaaaatagatcaacagacaccagggcctacttgaggggttgagggg:gggagagggagacg  
atgaaaaagaacctatgggtatlaagttcatcactgagtgatgaaataatctgtacatcaagaccagtgatag  
caatttacctatataacttgtacatgtacccccaaatttaaaatgaaagttaaaacaaagtaggaatggaatta  
attcctcaagatltggctttaattttatttgataatttatcaaatgcttggttttcttttctcactatggcgttgc  
tttataaactatgttcagtatgtctgaatgaaaggggtgtgtgtgtgtgaaagagagggagagaggaagggaaga  
gaggacgtataatgtgaatttgagttcatgaaaatttttcaataaaaataatttaatgtcaggagaatlaagccta  
atagttcctaaatcatccatctcttgagcttcagagcagtcctctgaattaatgocctacatgtttgtaagggtg  
ttcagactgaagcoagatltcactcetaagagatgcaatctcaaatttatctgaagactgtacctctgctctcc  
ataaattgacaccatggcccacttaatgaggttaaaaaaagctaattctgaatgaaaactgagcccagtgaggg  
aaatattaatgaaacaggtgcagactgaaatataaattttctgttaataattatgcatatactttagcaaggttct  
gtctatgttgactttattgctttttggtaagaaatacaactttttaaagtgaactaaactatcctatttccaaact  
atlttgtgtgtgtggcgtttgtttctatgggttctgggtttcttggagcatttttattttcatttttaatttaatt  
ctgagagctgctgagttgtgtttactgagagattgtgtatctgcgagagaagtctgtagcaagtagctagactgt  
gctgacctaggaacatatacagtagattgctaaaatgtctcacttggggaatttttagactaaaacagtagagcatg  
tataaaaactctagtcaggtgctgctttgaaacaaatgataaaaaccacactcccatagatgagtgctatgatt  
ttcatggagggaagttaatatcactcctctaagtataccagacttagggcattctgatataaaacattaggaacta  
agaaagatlaatagactggagtaaggaaatggacctctgctctctcctgctgtctcttttttgaggacttgtgtgt  
gt  
tggcatccattcacagagtagaagcaagctcacaatagtgaaagatgtcagtaagcttgaatagtttttcaggaact  
ttgaatgctgatttagatttgaactgaggtctgaccataaccaaatttgactatttatgtctcttgaactt  
atltgctggatggcctgggcttttgatggcttttagtalagcttgcagccttgcctgcaaggtattatgggtaa  
tagaaagaaaagctcggttacactctagtcacactaagtaactaccattggaaaagcaaccctgccttgaagcc  
aggatgatgggtatctgcagcagttgccacacaagagaaggatccatagttcatcatttaaaaaagaaaacaaat  
agaaaagggaaaactatcttgagcataagaagttgtagggttaagcttttaagaaggtgacaattcttgccaatca  
ggatttcaaaagctcttgccttgacaattttgggtcttcagaatactataaaatataacctatattataatttcaata  
agtctgtgcattttcttgaccaggatatttgcaaaagacatattcaaaactccgcagaacactttatttccat  
atcacatgctcttatatcagggaagtgaacaggggtctgaaaactgtctaaatctaaaacaaatgctaatgcaggt  
ttaaatttaataaaaataaaaactcaaaaactcaacagccaagctcaaatctgcatgttttaacatttaaaatattttaa  
agagctcttttccaggatccaacatgtgaaatctttctcaggatcacagctgtgctcctagatcctcattgcttta  
gtttttacagaggaatgaatataaaaagaaaacttaaattttatccctcttaacctctataatcatacatagggc  
ataattttttaacctaggctccagatagccatagaagaacaaacactttctgctgtgtgagaataatcagagtg  
agattttttcacaagtagctgatggggttgagacaggtagaanaagtgagagatctctatttattagcaataat  
agagaaaagcatttaagagaataaagcaatggaaataagaaaatttgaatattcctctctgataactagaaatagagg  
atccagtttcttttggtaacctaaattttattttcattttattgtttttattttattttattttattttttgtg  
laatcgtagttcagagtgtagagctgaaaggaagaagtaggagaacatgcaaaagtaaaagtataaacactttcc  
ttactaaaccgacatgggtttccaggtagggcaggatccaggatgactgacagggccttagggcaacctgagac  
cctacgctgacctcataaatgcttgcactttgtgttttaattacatcttttaatagcaggaagcagaactctg  
cacttcaaaaagtlttctcactcagggagtlaatttagtacaaggggaaaaagtagcagggggatgggagaaaaggc  
gatcagctggggaagctatagagaaagaagagtaaattttagttaaaggaggtttaaacaacaaaataaaagaga  
aataggaacttgaatcaaggaatgatlttaaaacgcagttatcttagtgactagaggaaaaaataaactcagc  
caagtagaagacctttccctcctaccctactttctaagtcacagaggtttttgttccccagacactcttgc  
agattagtcaggcagaaacagttagatgtcccagttaacctcctatttgacaccactgatcaccattgatag  
tcacactttgggttgaagtgactttttattttattttgtatttttgactgcattaaagaggtctctagttttttatct  
cttgtttcccaaaacctaaatgaactaactgaacagagcacttgattttgtatttatctatttttagacataat  
ttattagcatgcatgagcaaatlaagaaaaacaacaaatgaatgcatatataatgtatgtatgtgtgtat  
atacacacatatatatatatttttcttttcttaccagaaggttttaatccaaataaggagaagatagcttag  
aacggaggtagagtttccatccattctgctcgtgaaglattttgcatattctggagacgcaggaagagatccatct  
acatataccaaagctgaattatggtagacaaaactctccacttttagtgcatcaacttcttattttgtgtaataag  
aaaattgggaaaacgatcttcaatagcttaccagctgtgattccaaatattacgtaaatcacactgcaaaaggag  
gatgttttagtagcaatttgaactgatgggtatggggccaagagatatcttagagggagggctgaggggttga  
gtccaactcctaagccagtgcoagaagagccaaggacaggtcaggtctcacttagacctcaccctgtggagc  
cacacctaggttggcctactcctcaggagcagggcagggcagggcagggcagggcagggcagggcagggcagggc  
gagccatctattgcttacatttgcctctgacacaactgtgtcacttagcaacctcaaacagacacacalgggtgcatc  
tgactcctgaggagaagctgtccgttactgocctgtggggcaaggtgaaactggatgaagtgggtgggtgagggcct  
gggcaggttgggtatcaaggttacaagacaggtttaaaggagaccaatagaactgggcagtgaggagacagagaagac  
tcttgggtttctgatagggcactgactctctctgctatgtgtctattttccaccttaggctgctgggtgctac  
ccttggaccagaggttctttgagtcctttggggatctgtccactcctgatgctgttatgggcaacctaaaggtga  
aggctcatggcaagaaagtgctcgggtgcttttagtgatggcctggctcactggacaacctcaagggcactttgc  
cacactgagtgagctgcactgtgacaagctgcacgtggatcctgagaactlcaggggtgagctatgggacgcttga  
tgttttcttcccttcttttctatgggttaagttcatgtcataggaaggggataagtaacaggggtacagtttagaa

tgggaaacagacgaatgattgcatcagtggtggaagctcaggatcgttttagtctcttttatttggctgttcataaac  
aattggtttcttttggtttaactcttgccttctctttttttctctccgcaattttactattatactbaatgcctt  
aacattgtgtataacaaaaggaaatatctctgagatacattaagtaacttaaaaaaaaaactttacacagctcgcct  
agtacattactatttgggaatatatgtgtgcttatttgcataatcotaatctccctactttatctttttatctttt  
aattgatacataatcattatacatatcttgggttaaaagtgaatgttttaatatgtgtacacataattgaccaa  
cagggtaattttgcatttgtaatttttaaaaaatgcttctctcttaataatacttttttggtttatcttattttctaa  
tactttccctaactctcttcttccagggaataatgatataatglatcatgctctttgcaccattctaaagaata  
acagtgataaattctgggttaaggcaatagcaatatctctgcataaaaatttctgcatataaattgtaactgat  
gtaagaggtttcatttgctaatagcagctacaatccagctaccattctgcttttattttatgggtgggataaggc  
tggattattctgagtcacagctaggcccttttgcataatcatggttcatacctcttatcttctcccacagctcctgg  
gcaacgtgctggctctgtgtgctggccctcactttggcaagaattcaccaccagtgagcctgcctatcagaa  
agtgtggctggctgtggttaaatgcccctggcccacaagatcactaaagctcgccttcttggctgtccaattctatta  
aaggttcccttggttccctaagtcccaactactaaactgggggatattatgaaggcccttgagcatctggattctgpc  
taataaaaaacatttttcatgcaatgatgtatttaattttctgaaatatttactaaaaaggggaatggtg  
gaggtcagtgcatttaaaaacataaagaatgaagagctagtccaaccttgggaaaatacactatacttaaacctc  
catgaaaagaaggtgaggtgcaaacagctaatgcacatlygcaacagcccctgatgcatacgcttattcatccct  
cagaaaaggattcaagtagaggttgatttggaggttaaaagtttgcatagctgtattttacattacttattgttt  
tagctgtcctcatgaaatgtctttcactaccatttgccttatctgcatctctcagccttgactccactcagttct  
cttgccttagagataccacccttcccctgaagtggttcttccatgtttacggcgagatgggtttctcctcgcctggc  
cactcagccttagttgtctctgtgtcttatagaggtctacttgaagaaggaaaaacaggggtcatggtttgactg  
tccctgtgagccctcttcccctgcctccccactcacagtgaccgggaatctgcagtgctagtctcccggaaactatc  
actcttccacagtgcttgggaaggactgggcttagtatgaaaagttaggactgagaagaatttgaaggcggct  
ttttgtagcttgatctcactactgtcttatcctctgcataaggcccaccccaaatggaagtccattcttctctc  
aggatgtttaagattagcattcagggaagatcagaggtctgctggctcccttatcatgtcccttatggtgctctc  
ggctctgcagttattagcatagtggtaccatcaaccacttaacttcatttttcttattcaataccttaggtaggta  
gatgctagattctggaataaaaatagagctcaagtggtccttgcctctctcccagtcacaaatctgaaatctagt  
tggcaagattctgaaatcaaggcatataatcagtaataagtgatgataagaaggatataagaagaattttattata  
tgagaggggtgaaacccctcaaaatgaaatgaaatcagacccttgccttacaccataaacaaaaataaatttgaatgg  
gttaagaattaaactaagacctaaaaccataaaaaattttaaagaatcaaaagaagaaaaattctaataatcacg  
ttgcaagccgtttttgaaatttgatagagaagcaaggcaacaaaaggaaaaataaagaagtgaggctacatcaaa  
ctaaaaaattccacacaaaaaacaacaaatgaacaaatgaaaggtgaacctgaaatggcatatttggcaaacca  
aatatttcttaaatatttgggttaaatccaaaaatataaagaacacagatgattcaatacaaaaacaaaaaatta  
aaaataggaaaaataaaaaaattaaaaagaagaaaatcctgccatttatggcagaatgatgaacctggaggatgta  
aaactaagaaaaataagcctgacacaaaaagacaaatactacacacaccttgcctcatatgtgaaacataaaaaagtc  
actctcatggaaacagacagtagaggtatgggttccaggggttgggggtgggagaatcaggaaactattactcaaa  
gggtataaaaattcagttatgtgggatgaaataattctagatatactaatgtacagcatcgtgactgtagttaattg  
tactgtaaagtataattaaaatttgcaaaagagagtagatttttttttttttagatggagttttgctcttgttgc  
caggtcggagtgcaatggcaagatcttggctcactgcacacctccgctcctgggttcaagcaaatctcctgcctca  
gctcccgagtagctgggaltcacaggctgcgacacctagcccagctaaatttggatttttagtagagacgggggtt  
tctccatgttggctcaggctgatccgctcctcggccaccacaaagggttgggattacaggcgtgagccacgggct  
tggccgagagtagatctaaaagcatttaccacaagaaaaaggtaactatgtgagataatgggtatgttaattagc  
ttgatgtggtaatcatttccacaggtatacatatataaaaacatcatggttgcacacctaaatataacaatttt  
tatttgtgaatgatcctcaataaagttgaagaataataaaaaagaaatagacatcacatgaattaaaaaactaaaa  
aataaaaaaatgcatcttgatgattagaattgcattcttgattttcagatacaaaatctcatttgactgtttact  
cttttccaaaacaatacaataaalltagcactttatctcattttcccctcccacatctataallatataatata  
atattttagatatttgtatagtttactccctagatttcttagtgttattatlaaatagtgagaagaatgtttaca  
cttatgtacaaaatgttttgcattgcttttctcatttctaacattctctcctaaagtttattctattttttctgatt  
atccttaataattatctcttctggtgaaatacattgttacttttgggtttatctaaaaatggcttcatlttcttca  
ttctaaaaatcatgttaaatataaccactcatgtgtaagtaagatagtggaataaaatagaatccaaaaactaaat  
ctcactaaaaatataaataatgtgataataaaaaatagcttttaaatttagcttggaaataaaaaacaaacagtaa  
ttgaacaaactatactttttgaaaagagtaagtgaaatgcttaactgcataatccacaatcgattacacaattagg  
tgtgaaggtaaaattcagtcacgaaaaaactagaataaaaaatagggagacatgtatataatcttagagataaca  
ctgttatttaattatcaaccacaaagtagaaactatcaaggggagaataaaattcagtcacaataaaaagcatttaag  
aagttattctaggctgggaggggtggctcacaacctgcaattgcagcactttgggagggctagacagggcggatcacg  
acgtcaaggagttcaagatcagcctggccacatagtgaaacctcatcactactaaaaataaaaaaacttagcctgg  
cgtgggtggcaggcatgtgtaatcccagcaatttgggagggctgaggcaggagaatcgtctgatcctgggagggcagag  
ggtgcagtgagccaaagatttgcactgcattccagcccaggtgacagcatgagactcctgcacaaaaaaaaga  
aaaaaaaagggggggggggaggggtggagccaagatgaccgaataggaacagctccagctctatagctcccactcgtg  
agtgcagcagaagacgggtgatcttctgcatttccactgaggtaccaggttcatctcacaggggaagtgccagggcag  
tgggtgcaggacagtaggtgcagtgcaactgtgcataagcgaagcagggcagggcatcacctcaccgggaagc  
aaggggtcagggaaatcccttctcagtcacaaagaaagggtgacagatggcacctggaaaaatcgggtcactccgc  
cctaactcgcctcttccacaagcttaacaaatggcacaccaggagattatcccctgctggctcagagggyl  
cctacgcccattggagcctcgcctcattgctagcacagcagctcagggtaaaactgcaaggtggcagtgaggtcggg

gaggggtgccaccatgtccagccttgagcaggtaaaacaaagccgctggaagctcgaactgggtggagcccacc  
 acagctcaaggaggcctgctgctctgtaggctccacctctaggggcaggggcacagacaaacaaaagacaacaag  
 aacctctgcagacttaaaatgtccctgtctgacagctttagagagagtagtgggtctctccagcacatagcttccagat  
 ctgagaacagycagactgectcctcaagtgggtccctgacccccagtagcctaaactgggaggcatccccagtag  
 gggcagactgacacctccatggctggtactcctctaagacaaaacttccagaggaaatgatcaggcagcagcattt  
 ggggttccaataatccactgttctgcagocaccgctgttgatacccaggaaaaacagcttctggagtggacctcca  
 gtaaacctccaacagacctgcagctgagggtcctgactgttagaaggaaaactaaacaaacagaaaggacatccacac  
 caaaaacccatctgtacatcgcctcatcaaaagacaaaggtagataaaacataaagatggggaaaaagcagagc  
 agaaaaactggacactctaaaaatgagagtgcctctcctcctccaaagtaacgcagctcctcaccagcaatggaac  
 aaagctgggcagagaaatgactttgacgagttgagagaggaggcttcagaagatcaaactactccaagctaaaggga  
 ggaagtlogaacaaaacggcaaaagagtaaaaaactttgaaaaaaaattagatgaatggataactagaataaccat  
 gcacagaagtcccttaaaggacctgatggagctgaaaaacaaaggcaggagaactacgtgacaaatacacaagcctca  
 gtaaccgatgagatcaactggaagaaagggtatcaatgacgaaagatgaaatgaatgaaatgaaagcatgaaagaa  
 gtttagagaaaaaagaataaaaaaaacgaaacaaagcctccaagaaatatgggactatgtgaaaagaccaaatcta  
 catctaattgggtgtagctgaaagtgtggggagaatggaaccaagtggaaaaactctgcaggatattatccagg  
 agaacttcccctcctagcaagcaagcccaaatccacttcaggaaatcacagagaacgcacaaagatactccta  
 gagaaaaacactccaagacacataactgtcagattcaccaaagttagaatgaaggaaaaaatgttaagggcagcc  
 agagagaaaggctcgggttaccacaaaagggaagcccatcagactaacagctgatctatcggcagaaactctacaag  
 ccagaagaaagtggggcccaatattcaacattggttaaagaaaagaattttcaaccogaatttcataatccagccaa  
 actaagcttcataagtgaaggagaaataaaatcctttacagacaagcaaatgtgagagatttgtcaccaccagg  
 cctgcttacaagagctcctgaaaggaagcactaaacatggaaaggaacaaactagatcagccactgcaaaaaatg  
 ccaaatgttaaagaccatcaaggctaggaagaaactgcataacagagcaaaaataaccagctaacatcataatgaca  
 ggtcaaatccatacatacaataactcacttaaatgtaaataggctaaatgctccaattaaagacacagactgg  
 caaattggataaaggagtcaggaccctctgtgttclgtattcaggaaacccatctcaagcagagacacacatag  
 gctcgaataaaaaggatggaggaatatctaccaagcaaatggaaaaacaaaaaaggcagggttgcactcctagtc  
 tctgataaaaacagattttaaaccaacaaagatcaaaagagacaaagaggccattacataatggcaagggatcta  
 ttcaagaagaagaactaactataactaaatataatgaccccaatacaggagcaccocagattcataaaacaagtct  
 gagtgccttacaagagacttagatgcccacacaataataatgggagactttaaaccacccactgtcaacattagac  
 agatcaacgagacagaaagttaacaaggatataccaggaaatggactcagctctgcaccaagcagacctaatagaca  
 tctacagaactctccaccccaatcaacagaaatatacattctttcagcaccacacccacactattccaaaactga  
 ccacatagttggaagttaaagctcctcctcagcaaatgtaaaagacagaaactatacaaaactgtctctcagaccac  
 agtgcactcaaaactagaactcaggattaaagaaactcaaaacccactcagctacatggaaactgaacagcctgc  
 tctgaatgactactgggtacataacaaaatgaaggcagaaataaagatgttctttgaaaccaaagagacaaaaga  
 cacacacaccagaatctctgagacacattcaagcagtggttagagggaaatttatagcactaaatgcccacaag  
 ggaaagcaggaaagatctaaaatgacaccctaacatcacaatlaaaaaactagagaagcaggagcaaacacattc  
 aaaagctaacagaagacaagaataaactaaagatcagagcagaagtgaaggacatagagacacaaaaaaccttca  
 aaaaaatcaatgaatccagaagctgttttttgaaaagatcaacaaaaatlgatagactgctagcaagactlaataaa  
 gaagaaaagagagaagaatcaaatagacgcaataaaaaatgacacggggtatcaccactgatcccacagaaataca  
 aactacgctcagagaaactataaacacctctacgcaaataaaaactagaaaatctagaagaatggataaaatcctc  
 gacacatacactctgccaagctaaaccaggaagagtgtatctctgaatagaccaataacaggctctgaaattg  
 aggcaataatataagcttatcaaccaaaaaaagtccgggaccagtaggattcatagccgaattctaccagaggta  
 caaggaggagctggtaccattcctctgaaactattccaatcaatagaaaaagagggaatcctccttaactcattt  
 tatgaggccagcatcatcctgataccaagcctgacagagacacacaaaaaaagagaatgttacaccaatcctc  
 tgatgaacattgatgcaaaaaatcctcaataaaaatactggcaactgatccaccatgatcaagtgggcttcatcct  
 gccatgcaaggctggttcaacatacgaaaaatcaataaacataatccagcatataaacagaaacaaagacacaaaac  
 atatgattatctcaatagatgcagaaaaaggcctttgacaaaaatcaacaacgcttcatgctaaaaactctcaataa  
 attaggatttgatgggacatactcaaaaataaagagctatctatgacaaaacccacagccaatctcactgagt  
 ggacaaaaactggaagcatcctcttgaaaactggcacaaggcagggatgcctctctcaccactcctattcaaca  
 tagtgttgaagtctggccagggcaatcaggcaggagaaggaaataaagggcattcaattaggaaaagagggaagt  
 gaaattgtccctgtttgcagatgacatgattgtatctagaaaaacccattgtctcagcccaaatctccttaag  
 ctgataagcaactcagcaaaagtctcaggatataaaatcagtggtgcaaaaatcacaagtatcctatgcaccaata  
 acagacaaaacagagagccaaatcatgagtgaactcccatcccaattgcttcaagagaaataaaataccataggaat  
 ccaacttacaagggatgtgaaggacctctcaaggagaactacaaaacactgctcaatgaaataaaagaggataca  
 acaaaatggaagaacatccatgctcatgggttaggaagaatcaatctgtagaaatggctcactgcccagggttaa  
 tttatagattcaatgocctcccacagctaccaatgactttctcagagaactggaaaaaactactttaaagtt  
 catatggaacaaaaaagagccacatcaccaaaggcaatcctaagccaaaagaacaaagctggaggcatcacgcta  
 cctgactcaaaactataactacaatgctacggtaacccaaaacagcatggtactggtacccaaaacagagatctagacc  
 aatggaacagaaacagagccctcagaataatgcccataatctacaactatctgatctttgacaaacctgagagaaa  
 caagcaatggggaaggatcctctattttaataaatggtgctgggaaaactggctagccatagtagaaagctgaaa  
 ctggatcctctccttacaccttatacaaaaatcaatcaagatggattaaagacttcatggttagacctaaaacca  
 taaaaacccctagaaaaaaactaggcaataaccattcaggacataggcatgggcaaggacttcatgcttaaaacacc  
 aaaagcaatggcaacaaaagacaaaatggcaaacgggatctaataaactaaagagcttctgcacagctaaagaa  
 actaccatcagagtgaaacaggcaacctacaaaatgggagaaaattttgcaatctactcatctgacaaaaggctaa

tatccagaatctacaatgaactcaaacaattttacaagaaaaacaacaaccccatcaaaaagtgggcaaaggat  
 atgaacagacacttcgcaaaagaagacatttatgtaatcaaaaaacacatgaaaaatgctcatcactagcca  
 tcagagaaatgcaaatcaaaaccacaatgagataccatctcacaccagttagaatggcgatcattaaaaagtcagg  
 aaacaacaggtgctggagaggatgtggagaaacaggaacaacttttacactgttggtgggactgtaaaactagtcca  
 accattgcggaagtgcagtggtgcaattcctcaggaatctagaactagaataccatttgaccagccatccatta  
 ctgggtacatacccaaaggattataaatcatgctgctataaagacacatgcacacgtatgtttattgcagcactat  
 tcacaatagcaaagacttggaaaccaacccaaatgtccaacaacgatagactggattaagaaaatgtggcacatata  
 caccatggaatactatgcagccataaaaaatgatgagttcatgtcctttgtagggacatggatgaagctggaaact  
 atcattctcagcaaaactatcacaaggagaataaaaccaaacaccgcagttctcactcataggtgggaattgaaaca  
 tgagaacacatggacacatgaagaggaaacatcacactctggggactggtatggggtggggggcaggggcagggata  
 gcactagagatatacctaatgctaataatgacgagtttaattgggtgcagcacaccaacatggcacatgtatacatata  
 taacaaaactgcagttgtgacatgtaccctaaaacttgaagtataataaaaaaaagtattcctattaaaaac  
 tgatctcacacatccgtagaccattatcaagtctttctctttgaaatagacagaaattagtggtttctcagtc  
 gttaac [SEQ ID NO:19]

Se han identificado cinco sitios de sitio hipersensible (HS) en 5' (HS1-HS5) y un HS en 3' en la LCR de  $\beta$ -globina humana (Stamatoyannopoulos *et al.*, (2001)). Los HS 1-4 en 5' son sitios hipersensibles a ADNasa I. Los elementos HS2 y HS3 son los elementos individuales más potentes dentro de la LCR (Ellis *et al.*, EMBO J. (1996), 15: 562-568; Collis *et al.*, EMBO J. (1990) 9: 233-240), tal como se corrobora por muchos grupos. Deleción HS2 en el contexto de  $\beta$ YAC en ratones transgénicos afecta gravemente a la formación de sitios HS así como a la expresión de todos los genes de  $\beta$ -globina humana en cualquier estadio de desarrollo (Bungert *et al.*, Mol. Cell Biol. (1999); 19: 3062-3072). Se notificó que la deleción de HS2 reducía de manera mínima la expresión de los genes de globina  $\epsilon$ y y  $\beta$ hi embrionarios en eritrocitos derivados de saco vitelino (Ley *et al.*, Ann. N.Y. Acad. Sci. (1998); 850: 45-53; Hug *et al.*, Mol. Cell Biol. (1996); 26: 2906-2912). HS2 funciona principalmente como potenciador.

En determinadas realizaciones, la región LCR de  $\beta$ -globina comprende una región de HS2. En ejemplo no limitativo, la región LCR de  $\beta$ -globina comprende una región de HS2, una región de HS3, y una región de HS4. En determinadas realizaciones, la región de HS2, región de HS3 y región de HS4 dentro de la región LCR de  $\beta$ -globina son contiguas. En una realización no limitativa, la región LCR de  $\beta$ -globina consiste esencialmente en una región de HS2, una región de HS3 y una región de HS4. En otra realización, la región LCR de  $\beta$ -globina comprende dos sitios de unión a GATA-1 introducidos en la unión entre la región de HS3 y la región de HS4. La región de HS3 puede encontrarse entre la región de HS2 y la región de HS4. La longitud y la secuencia de la región de HS2 pueden variar. La región de HS2 puede tener una longitud de desde aproximadamente 400 pb hasta aproximadamente 1000 pb, por ejemplo, desde aproximadamente 400 pb hasta aproximadamente 500 pb, desde aproximadamente 500 pb hasta aproximadamente 600 pb, desde aproximadamente 600 pb hasta aproximadamente 700 pb, desde aproximadamente 700 pb hasta aproximadamente 800 pb, desde aproximadamente 800 pb hasta aproximadamente 900 pb, o desde aproximadamente 900 pb hasta aproximadamente 1000 pb. En una realización no limitativa, la región de HS2 tiene una longitud de 860 pb. En un ejemplo no limitativo, la región de HS2 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 9, que se proporciona a continuación:

GTATATGTGTATATATATATATATATATATATTCAGGAAATAATATATTTCTAGAATATGTCA  
 CATTCTGTCTCAGGCATCCATTTTTCTTTATGATGCCGTTTGAGGTGGAGTTTTAGTCAGG  
 TGGTCAGCTTCTCCtttttttttGCCATCTGCCCTGTAAGCATCCTGCTGGGGACCCAGAT  
 AGGAGTCACTACTCTAGGCTGAGAACATCTGGGCACACACCCTAAGCCTCAGCATGACTC  
 ATCATGACTCAGCAATTGCTGTGCTTGAGCCAGAAGTTTTGCTTAGAAGGTTACACAGAAC  
 CAGAAGGCGGGGGTGGGGCACTGACCCCGACAGGGGCCGGCCAGAAGTCTCATGCTTG  
 GACTATGGGAGGTCACTAATGGAGACACACAGAAATGTAACAGGAACTAAGGAAAACTG  
 AAGCTTATTTAATCAGAGATGAGATGCTGGAAGGGATAGAGGGAGCTGAGCTTGTA  
 AGTATAGTAATCATTTCAGCAAATGGTTTTGAAGCACCTGCTGGATGCTAAACACTATTTT  
 CAGTGCTTGAATCATAAATAAGAATAAAACATGTATCTTATTTCCCAAGAGTCCAAGT  
 AAAAAATAACAGTTAATTATAATGTGCTCTGTCCCCAGGCTGGAGTGCAGTGGCAGCAT  
 CTCAGCTCACTGCAACCTCCGCCCTCCCGGTTCAAGCAATTTCTCCTGCCTCAGCCACCCT  
 AATAGCTGGGATTACAGGTGCACACCACCATGCCAGGCTAATTTTTGTACTTTTTGTAGA  
 GGCAGGGTATCACCATGTTGTCCAAGATGGTCTTGAACCTCTGAGCTCCAAGCAGTCCAC  
 CCACCTCAGCCTCCCAAAGTGCT [SEQ ID NO: 9]

En determinadas realizaciones, la región de HS2 tiene una longitud de aproximadamente 840 pb. En determinadas realizaciones, la región de HS2 tiene una longitud de aproximadamente 650 pb (por ejemplo, 646 pb). En determinadas realizaciones, la región de HS2 tiene una longitud de aproximadamente 420 pb (por ejemplo, 423 pb).

La longitud y la secuencia de la región de HS3 pueden variar. La región de HS3 puede tener una longitud de desde aproximadamente 200 pb hasta aproximadamente 1400 pb, por ejemplo, desde aproximadamente 200 pb hasta aproximadamente 300 pb, desde aproximadamente 300 pb hasta aproximadamente 400 pb, desde aproximadamente 400 pb hasta aproximadamente 500 pb, desde aproximadamente 500 pb hasta aproximadamente



TGAGCCCCTTTTCTCTAACTGAAAGAAGGAAAAAAAAAATGGAACCCAAAATATTCTACATA  
 GTTTCCATGTACAGCCAGGGCTGGGCAGTCTCCTGTTATTTCTTTAAAATAAATATATCAT  
 TTAAATGCATAAATAAGCAAACCCCTGCTCGGGAATGGGAGGGAGAGTCTCTGGAGTCCACCCC  
 TTCTCGGCCCTGGCTCTGCAGATAGTGCTATCAAAGCCCTGACAGAGCCCTGCCCATTTGCTGG  
 GCCTTGGAGTGAGTCAGCCTAGTAGAGAGGCAGGGCAAGCCATCTCATAGCTGCTGAGTGGGA  
 GAGAGAAAAGGGCTCATTGTCTATAAACTCAGGTCATGGCTATTCTTATTCTCACACTAAGAA  
 AAAGAATGAGATGTCTACATATAACCTGCGTCCCCCTCTTGTGTACTGGGGCCCCCAAGAGCTC  
 TCTAAAAGTGATGGCAAAGTCATTGCGCTAGATGCCATCCCATCTATTATAAACCTGCATTTG  
 TCTCCACACACCAGTCATGGACAATAACCCCTCCAGGTCACAGTGGCTTGTCTTTGTATAA  
 TACTCAAGTAATTTTCGAAAAATGTATTCTTTCAATCTTGTTCCTGTTATTCCTGTTTCAATGGC  
 TTAGTAGAAAAAGTACATACTTGTTCCTCCATAAAATGACAAATAGACAATTTACATCAATGT  
 CTATATGGGTCGTTGTGTTTGTCTGTGTTGCAAAAACCTCACAATAACTTTATATTGTTACTAC  
 TCTAAGAAAAGTTACAACATGGTGAATACAAGAGAAAGCTATTACAAGTCCAGAAAATAAAAGT  
 TATCATCTTGAGGCCCTCAGCTTTCTAGGAATAATATCAATATTACAATAATTTAATCTAACAAT  
 TATGAACAGCAATGAGATAATATGTACAAAGTACCCAGACCTATGTGGTAGAGCATCAAGGAA  
 GCGCATTGCGGAGCAGTTTTTTTGTGTTTGTGTTTGTATTCTGTTTCGTGAGGCAAGGTTTC  
 ACTCTGCTGTCCAGGCTGGAGTGCAGTGGCAAGATCATGTCTCACTGCAGCCTTGAC [SEQ  
 ID NO:6]

En un ejemplo no limitativo, la región de HS4 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7, que se proporciona a continuación:

TGAGCCCCTTTTCTCTAACTGAAAGAAGGAAAAAAAAAATGGAACCCAAAATATTCTAC  
 ATAGTTTCCATGTACAGCCAGGGCTGGGCAGTCTCCTGTTATTTCTTTAAAATAAATA  
 TATCATTTAAATGCATAAATAAGCAAACCCCTGCTCGGGAATGGGAGGGAGAGTCTCTGGA  
 GTCCACCCCTTCTCGGCCCTGGCTCTGCAGATAGTGCTATCAAAGCCCTGACAGAGCCCT  
 GCCCATTTGCTGGGCCTTGGAGTGAGTCAGCCTAGTAGAGAGGCAGGGCAAGCCATCTCAT  
 AGCTGCTGAGTGGGAGAGAGAAAAGGGCTCATTGTCTATAAACTCAGGTCATGGCTATTCT  
 TTATTCTCACACTAAGAAAAAGAATGAGATGTCTACATATAACCTGCGTCCCCCTCTTGTG  
 TACTGGGCCCCCCAAGAGCTCTCTAAAAGTGTATGGCAAAGTCATTGCGCTAGATGCCATC  
 CCATCTATTATAAACCTGCATTTGTCTCCACACACCAGTCATGGACAATAACCCCTCCTCC  
 CAGGTCCACCTGCTTGTCTTTGTATAATACTCAAGTAATTTTCGAAAAATGTATTCTTTCA  
 ATCTTGTTCGTTATTCTCTGTTTCAATGGCTTAGTAGAAAAAGTACATACTTGTTTTCCC  
 ATAAATGACAATAGACAATTTACATCAATGTCTATATGGGTCGTTGTGTTTGTGTTG  
 TTGCAAAAACCTCACAATAACTTTATATTGTTACTACTCTAAGAAAGTTACAACATGGTGA  
 ATACAAGAGAAAAGCTATTACAAGTCCAGAAAATAAAAGTTATCATCTTGAGGCCTCAGCT  
 TTCTAGGATAAATATCAATATTACAATAATTAATCTAACAATTTATGAACAGCAATGAGAT  
 AATATGTACAAAGTACCCAGACCTATGTGGTAGAGCATCAAGGAAGCGCATTGCGGAGCA  
 GTTTTTGTGTTGTTTGTGTTTGTATTCTGTTTCGTGAGGCAAGGTTTCACTCTGCTGTCC  
 5 AGGCTGGAGTGCAGTGGCAAGATCATGTCTCACTGCAGCCTTGACAC [SEQ ID NO:7]

En determinadas realizaciones, la región de HS4 tiene una longitud de menos de aproximadamente 1,0 kb, por ejemplo, menos de aproximadamente 900 pb, menos de aproximadamente 700 pb, menos de aproximadamente 600 pb, o menos de aproximadamente 500 pb. En determinadas realizaciones, la región de HS4 tiene una longitud de menos de aproximadamente 500 pb. En determinadas realizaciones, la región de HS4 tiene una longitud de aproximadamente 450 pb. En una realización no limitativa, la región de HS4 tiene una longitud de aproximadamente 446 pb. En un ejemplo no limitativo, la región de HS4 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 8, que se proporciona a continuación:

TGGAACCCAAAATATTCTACATAGTTTCCATGTACAGCCAGGGCTGGGCAGTCTCCTGTTAT  
 TTCTTTTAAAATAAATATATCATTTAAATGCATAAATAAGCAAACCCCTGCTCGGGAATGGGAG  
 GGAGAGTCTCTGGAGTCCACCCCTTCTCGGCCCTGGCTCTGCAGATAGTGCTATCAAAGCCCT  
 GACAGAGCCCTGCCCATTTGCTGGGCCCTTGGAGTGAGTCAGCCTAGTAGAGAGGCAGGGCAAGC  
 CATCTCATAGCTGCTGAGTGGGAGAGAGAAAAGGGCTCATTGTCTATAAACTCAGGTCATGGC  
 TATTCTTATTCTCACACTAAGAAAAAGAATGAGATGTCTACATATAACCTGCGTCCCCCTCTTG  
 TGTACTGGGGTCCCCAAGAGCTCTCTAAAAGTGTATGGCAAAGTCATTGCGCTAGATGCCATCC  
 CATCT [SEQ ID NO:8]

En determinadas realizaciones, la región de HS4 tiene una longitud de aproximadamente 280 pb (por ejemplo, 283 pb). En determinadas realizaciones, la región de HS4 tiene una longitud de aproximadamente 240 pb (por ejemplo, 243 pb).



GGCATCTGTGAAGGAAAGAAACATCTCCTCTAAACCACTATGCTGCTAGAGCCTCTTTTCTGTAICTCAA  
 GCCTCATTCAGACACTAGTGTCCAGCTCTCCTCATATACCTATTGTATTTTCTTCTTCTTGCTGGTTT  
 AGTCATGTTTTCTGGGAGCTTAGGGGCTTATTTTATTTTGTGTTTTGTTTTCTAATCAACAGAGATGGGCA  
 AACCCATTATTTTTTCTTTAGACTTGGGATGGTGATAGCTGGGCAGCGTCAGAAACTGTGTGTGGATA  
 TAGATAAGAGCTCGGACTATGCTGAGCTGTGATGAGGGAGGGACCTAGCCAAAGGCAGTGAGAGTCAGA  
 ATGCTCCTGCTATTGCCTTCTCAGTCCCCACGCTTGGTTTTCTACACAAGTAGATACATAGAAAAGGCTA  
 TAGGTTAGTGTGTTGAGAGTCTGCATGAGTTAGTTGCTCAGAAATGCCCGATAAATATGTTATGTGTGT  
 TTATGTATATATATGTTTTATATATATATATATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  
 TGTGATTATCATCAAAACGTGAGGGCTAAAGTGACCAGATAACTTGCAGG [SEQ ID NO:3]

5 En determinadas realizaciones, la región de HS1 tiene una longitud de menos de aproximadamente 500 pb. En determinadas realizaciones, la región de HS1 tiene una longitud de aproximadamente 490 pb. En una realización no limitativa, la región de HS1 tiene una longitud de 489 pb. En un ejemplo no limitativo, la región de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 4, que se proporciona a continuación:

GGCATCTGTGAAGGAAAGAAACATCTCCTCTAAACCACTATGCTGCTAGAGCCTCTTTTCTGTAICTCAA  
 GCCTCATTCAGACACTAGTGTCCAGCTCTCCTCATATACCTATTGTATTTTCTTCTTCTTGCTGGTTT  
 AGTCATGTTTTCTGGGAGCTTAGGGGCTTATTTTATTTTGTGTTTTGTTTTCTAATCAACAGAGATGGGCA  
 AACCCATTATTTTTTCTTTAGACTTGGGATGGTGATAGCTGGGCAGCGTCAGAAACTGTGTGTGGATA  
 TAGATAAGAGCTCGGACTATGCTGAGCTGTGATGAGGGAGGGACCTAGCCAAAGGCAGTGAGAGTCAGA  
 ATGCTCCTGCTATTGCCTTCTCAGTCCCCACGCTTGGTTTTCTACACAAGTAGATACATAGAAAAGGCTA  
 TAGGTTAGTGTGTTGAGAGTCTGCATGAGTTAGTTGCTCAGAAATGCCCGATAAATATGTTATGTGTGT  
 TTATGT [SEQ ID NO:4]

10 Recientes estudios han mostrado que HS2 no es específica de eritroides, sino que se expresa en otros linajes y líneas celulares (véanse el ejemplo 3 y la figura 7) y también está presente en células madre embrionarias humanas no diferenciadas (Chang *et al.*, Stem cell reviews (2013); 9: 397-407). Debido a la actividad no eritroide de HS2, los vectores de globina que contienen HS2 pueden suponer un riesgo para su uso seguro en el tratamiento clínico, por ejemplo, para tratar a pacientes con talasemia y células falciformes. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una secuencia central de HS2. Una secuencia central de HS2 proporciona una expresión de alto nivel independiente de la posición. Además, una secuencia central de HS2 conserva la actividad potenciadora de HS2.

15 Por ejemplo, la secuencia central de HS2 potencia la transcripción de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana). Adicionalmente, una secuencia central de HS2 comprende uno o más sitios de unión o motivos de unión para proteínas (por ejemplo, factores de transcripción) ubicuas así como específicas de tejido (por ejemplo, específicas de eritroides), incluyendo, pero sin limitarse a, miembros de la familia de proteínas AP1 (por ejemplo, NF-E2), GATA-1 (también conocido como "NF-E1" o "NFE1"), proteínas de dedos de Zn de tipo Krüppel (por ejemplo, proteínas ubicuas Spl e YY1, y factor restringido a eritroides factor de tipo Krüppel eritroide (EKLF)), y

20 proteínas de hélice-bucle-hélice básicas (bHLH) (cajas E) (por ejemplo, USF y TAL1). Se requieren sitios de unión a AP1 para la potenciación e inducción (Moi y Kan (1990); Ney *et al.*, (1990); Talbot y Grosfeld (1991)). Además, la unión de NF-E2 puede provocar la alteración de la cromatina reconstituida *in vitro* en HS2 (Armstrong y Emerson (1996)). Mutaciones en los sitios de unión a GATA-1 pueden provocar una reducción de la actividad potenciadora de HS2 en ratones transgénicos (Caterina *et al.*, (1994)). Aunque tanto sitios de unión a AP1 (por ejemplo, AP1/NF-E2) como a GATA1 son importantes para la función principal, los ratones que carecen de estos factores no muestran una expresión de gen de globina alterada (Weiss *et al.*, 1994).

25 En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende la longitud completa de una secuencia central de HS2. En determinadas realizaciones, la secuencia central de una región de HS2 es una secuencia central de HS2 humano. En una realización no limitativa, la secuencia central de HS2 humano comprende un par en tándem de sitios de unión para miembros de la familia de proteínas AP1 (por ejemplo, NF-E2) (denominados sitios de unión a "AP1/NF-E2") (por ejemplo, GCTGAGTCA, y GATGAGTCA), un sitio de unión para proteínas de dedos de Zn de tipo Kruppel (por ejemplo, AGGGTGTGT), un sitio de unión a GATA-1 (por ejemplo, CTATCT), y tres cajas E (CANNTG, por ejemplo, CAGATG, y CACCTG). En una realización no limitativa, la región LCR de  $\beta$ -globina no

30 comprende la longitud completa de una secuencia central de 388 pb de HS2 humano, que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 20 proporcionada a continuación:

TAAGCTTCAGTTTTTTCCTTAGTTCCTGTTACATTTCTGTGTGTCTCCATTAGTGACCTCCCATAGTCCA  
 AGCATGAGCAGTTCTGGCCAGGCCCTGTCCGGGTGAGTGCCCCACCCCGCCTTCTGGTTCTGTGTAA  
 CTTCTAAGCAAACCTTCTGGCTCAAGCACAGCAATGCTGAGTCATGATGAGTCATGCTGAGGCTTAGG  
 GTGTGTGCCAGATGTTCTCAGCCTAGAGTGATGACTCCTATCTGGGTCCCCAGCAGGATGCTTACAGG  
 GCAGATGGCAAAAAAAGGAGAAGCTGACCACCTGACTAAAACCTCCACCTCAAACGGCATCATAAAGAA  
 AATGGATGCCTGAGACAGAATGTGACATATTCTAGAATATATT [SEQ ID NO:20]

La secuencia de nucleótidos expuesta en SEQ ID NO: 20 corresponde a nucleótidos de la posición 16671 a la posición 17058 de SEQ ID NO: 19 (n.º de registro de GenBank: NG\_000007.3). En SEQ ID NO: 20, un sitio de unión a AP1/NF-E2 que tiene la secuencia de nucleótidos de GCTGAGTCA está ubicado de la posición 175 a la posición 183, un sitio de unión a AP1/NF-E2 que tiene la secuencia de nucleótidos de GATGAGTCA está ubicado de la posición 185 a la posición 193, un sitio de unión para proteínas de dedos de Zn de tipo Krüppel que tienen la secuencia de nucleótidos de AGGGTGTGT está ubicado de la posición 205 a la posición 213, dos cajas E, cada una de las cuales tiene la secuencia de nucleótidos de CAGATG, están ubicadas de la posición 217 a la posición 222, y de la posición 278 a la posición 283, un sitio de unión a GATA-1 que tiene la secuencia de nucleótidos de CTATCT está ubicado de la posición 246 a la posición 251, una caja E que tiene la secuencia de nucleótidos de CACCTG está ubicada de la posición 306 a la posición 311.

En una realización no limitativa, la región LCR de  $\beta$ -globina no comprende la longitud completa de una secuencia central de 387 pb de HS2 humano, que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 21 proporcionada a continuación:

TAAGCTTCAGTTTTTTCCTTAGTTCCTGTTACATTTCTGTGTGTCTCCATTAGTGACCTCCCATAGTCCA  
 AGCATGAGCAGTTCTGGCCAGGCCCTGTCCGGGTGAGTGCCCCACCCCGCCTTCTGGTTCTGTGTAA  
 CTTCTAAGCAAACCTTCTGGCTCAAGCACAGCAATGCTGAGTCATGATGAGTCATGCTGAGGCTAGGG  
 TGTGTGCCAGATGTTCTCAGCCTAGAGTGATGACTCCTATCTGGGTCCCCAGCAGGATGCTTACAGGG  
 CAGATGGCAAAAAAAGGAGAAGCTGACCACCTGACTAAAACCTCCACCTCAAACGGCATCATAAAGAAA  
 ATGGATGCCTGAGACAGAATGTGACATATTCTAGAATATATT [SEQ ID NO:21]

En SEQ ID NO: 21, un sitio de unión a AP1/NF-E2 que tiene la secuencia de nucleótidos de GCTGAGTCA está ubicado de la posición 175 a la posición 183, un sitio de unión a AP1/NF-E2 que tiene la secuencia de nucleótidos de GATGAGTCA está ubicado de la posición 185 a la posición 193, un sitio de unión para proteínas de dedos de Zn de tipo Krüppel que tienen la secuencia de nucleótidos de AGGGTGTGT está ubicado de la posición 204 a la posición 212, dos cajas E, cada una de las cuales tiene la secuencia de nucleótidos de CAGATG, están ubicadas de la posición 216 a la posición 221, y de la posición 277 a la posición 282, un sitio de unión a GATA-1 que tiene la secuencia de nucleótidos de CTATCT está ubicado de la posición 245 a la posición 250, una caja E que tiene la secuencia de nucleótidos de CACCTG está ubicada de la posición 305 a la posición 310.

En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que comprende una secuencia central de HS2. Una región de HS2 que comprende una secuencia central de HS2 puede variar en cuanto a la longitud y secuencia. En ejemplos no limitativos, una región de HS2 que comprende una secuencia central de HS2 tiene desde aproximadamente 400 pb hasta aproximadamente 1000 pb, por ejemplo, desde aproximadamente 400 pb hasta aproximadamente 500 pb, desde aproximadamente 500 pb hasta aproximadamente 600 pb, desde aproximadamente 600 pb hasta aproximadamente 700 pb, desde aproximadamente 700 pb hasta aproximadamente 800 pb, desde aproximadamente 800 pb hasta aproximadamente 900 pb, o desde aproximadamente 900 pb hasta aproximadamente 1000 pb, de longitud. En una realización no limitativa, la región LCR de  $\beta$ -globina no comprende una región de HS2 de 840 pb (por ejemplo, la región de HS2 comprendida en el vector de globina TNS9 dado a conocer en el documento U.S. 7.541.179). En una realización no limitativa, la región LCR de  $\beta$ -globina no comprende una región de HS2 de 860 pb. En una realización no limitativa, la región LCR de  $\beta$ -globina no comprende una región de HS2 de aproximadamente 650 pb. En un ejemplo no limitativo, la región LCR de  $\beta$ -globina no comprende una región de HS2 de 646 pb (por ejemplo, la región de HS2 comprendida en el vector de globina LentiGlobin™, también conocido como "β87"). En una realización no limitativa, la región LCR de  $\beta$ -globina no comprende una región de HS2 de aproximadamente 420 pb. En un ejemplo no limitativo, la región LCR de  $\beta$ -globina no comprende una región de HS2 de 423 pb (por ejemplo, la región de HS2 comprendida en el vector de globina dado a conocer en Sadelain *et al.*, Proc. Nat'l Acad. Sci. (USA) (1995); 92: 6728-6732).

En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que conserva la actividad potenciadora de HS2. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que puede potenciar la transcripción de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana).

En ejemplos no limitativos, la región LCR de  $\beta$ -globina no comprende una región de HS2 cuya capacidad para potenciar la transcripción de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana) es de no menos del 60%, no menos del 70%, no menos del 80%, no menos del 90%, o no menos del 95% en comparación con una región de HS2 nativa.

5 En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que comprende uno, dos, tres, cuatro, cinco, seis o siete de los siguientes sitios de unión: dos (un par en tándem de) sitios de unión a AP1/NF-E2 (por ejemplo, GCTGAGTCA, y GATGAGTCA), un sitio de unión para proteínas de dedos de Zn de tipo Kruppel (por ejemplo, AGGGTGTGT), un sitio de unión a GATA-1 (por ejemplo, CTATCT), y tres cajas E (CANNTG, por ejemplo, CAGATG, y CACCTG). En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una  
 10 región de HS2 que comprende seis de los sitios de unión descritos anteriormente. Por ejemplo, en determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que comprende dos sitios de unión a AP1/NFE2, un sitio de unión para proteínas de dedos de Zn de tipo Kruppel, un sitio de unión a GATA-1, y dos, no tres, cajas E. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que  
 15 comprende uno, no dos, sitios de unión a AP1/NFE2, un sitio de unión para proteínas de dedos de Zn de tipo Kruppel, un sitio de unión a GATA-1, y tres cajas E. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que comprende dos sitios de unión a AP1/NF-E2, un sitio de unión a GATA-1, y tres cajas E y no comprende un sitio de unión para proteínas de dedos de Zn de tipo Kruppel. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS2 que comprende dos sitios de unión a AP1/NF-E2, un sitio de unión para proteínas de dedos de Zn de tipo Kruppel, y tres cajas E, y no comprende un sitio  
 20 de unión a GATA-1.

En determinadas realizaciones, la región LCR de  $\beta$ -globina comprende una región de HS1, una región de HS3, y una región de HS4, y no comprende una región de HS2. En determinadas realizaciones, la región de HS1, región de HS3 y región de HS4 dentro de la región LCR de  $\beta$ -globina son contiguas. En una realización no limitativa, la región LCR de  $\beta$ -globina consiste esencialmente en una región de HS1, una región de HS3 y una región de HS4. En otra  
 25 realización, la región LCR de  $\beta$ -globina comprende dos sitios de unión a GATA-1 introducidos en la unión entre la región de HS3 y la región de HS4. La región de HS3 puede encontrarse entre la región de HS1 y la región de HS4.

En determinadas realizaciones no limitativas, la región LCR de  $\beta$ -globina comprende una región de HS1 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 22 o SEQ ID NO: 23, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que  
 30 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8, y la región LCR de  $\beta$ -globina no comprende una región de HS2.

En una realización no limitativa, la región LCR de  $\beta$ -globina comprende una región de HS1 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 2, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6, y la región LCR de  
 35  $\beta$ -globina no comprende una región de HS2, tal como se muestra en la figura 2.

En una realización no limitativa, la región LCR de  $\beta$ -globina comprende una región de HS1 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 3, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 8, y la región LCR de  $\beta$ -globina no comprende una región de HS2, tal como se muestra en la figura 3.

40 En una realización no limitativa, la región LCR de  $\beta$ -globina comprende una región de HS1 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 4, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 8, y la región LCR de  $\beta$ -globina no comprende una región de HS2.

En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 o una región de HS2.  
 45 En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una secuencia central de HS1. Una secuencia central de HS1 conserva la actividad de HS1, por ejemplo, actividad potenciadora, o que funciona como elemento facilitador o regulador para anclar la actividad potenciadora de otras regiones de HS, por ejemplo, HS2-4. Además, una secuencia central de HS1 comprende uno o más sitios de unión o motivos de unión para proteínas (por ejemplo, factores de transcripción) ubicuas así como específicas de tejido (por ejemplo, específicas de eritroides),  
 50 incluyendo, pero sin limitarse a, GATA-1, y proteínas de dedos de Zn de tipo Kruppel (por ejemplo, factor restringido a eritroides EKLF).

En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende la longitud completa de una secuencia central de HS1. En determinadas realizaciones, la secuencia central de una región de HS1 es una secuencia central de HS1 humano. En una realización no limitativa, la secuencia central de HS1 humano comprende dos sitios de  
 55 unión a GATA-1 (por ejemplo, TTATCT, y CTATCA), y un sitio de unión para EKLF (por ejemplo, CCACACACA). En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende la longitud completa de una secuencia central de 286 pb de HS1 humano. En una realización no limitativa, la secuencia central de 286 pb de HS1 humano tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 22 proporcionada a continuación:

CTGAGCAACTAACTCATGCAGGACTCTCAAACACTAACCTATAGCCTTTTCTATGTATCTACTTGTGTA  
 GAAACCAAGCGTGGGGACTGAGAAGGCAATAGCAGGAGCATTCTGACTCTCACTGCCTTTGGCTAGGTC  
 CCTCCCTCATCACAGCTCAGCATAGTCCGAGCTCTTATCTATATCCACACACAGTTTCTGACGCTGCC  
 AGCTATCACCATCCCAAGTCTAAAGAAAAAATAATGGGTTTGCCCATCTCTGTTGATTAGAAAACAAA  
 ACAAATAAAA [SEQ ID NO:22]

5 En SEQ ID NO: 22, un sitio de unión a GATA-1 que tiene la secuencia de nucleótidos de TTATCT está ubicado de la posición 173 a la posición 178, un sitio de unión a GATA-1 que tiene la secuencia de nucleótidos de CTATCA está ubicado de la posición 210 a la posición 215, y un sitio de unión para EKLF que tiene la secuencia de nucleótidos de CCACACACA está ubicado de la posición 183 a la posición 191.

En otra realización no limitativa, la secuencia central de 286 pb de HS1 humano tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 23 proporcionada a continuación:

CTGAGCAACTAATCATGCAGGACTCTCAAACACTAACCTATAGCCTTTTCTATGTATCTACTTGTGTAG  
 AAACCAAGCGTGGGGACTGAGAAGGCAATAGCAGGAGCATTCTGACTCTCACTGCCTTTAGCTAGGCC  
 CTCCCTCATCACAGCTCAGCATAGTCCGAGCTCTTATCTATATCCACACACAGTTTCTGACGCTGCC  
 AGCTATCACCATCCCAAGTCTAAAGAAAAAATAATGGGTTTGCCCATCTCTGTTGATTAGAAAACAAA  
 ACAAATAAAA [SEQ ID NO:23]

10 La secuencia de nucleótidos expuesta en SEQ ID NO: 23 corresponde a los nucleótidos de la posición 21481 a la posición 21766 de SEQ ID NO: 19 (n.º de registro de GenBank: NG\_000007.3). En SEQ ID NO: 23, un sitio de unión a GATA-1 que tiene la secuencia de nucleótidos de TTATCT está ubicado de la posición 173 a la posición 178, un sitio de unión a GATA-1 que tiene la secuencia de nucleótidos de CTATCA está ubicado de la posición 210 a la posición 215, y un sitio de unión para EKLF que tiene la secuencia de nucleótidos de CCACACACA está ubicado de la posición 183 a la posición 191.

15 En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 que comprende una secuencia central de HS1. Una región de HS1 que comprende una secuencia central de HS1 puede variar en cuanto a la longitud y secuencia. En ejemplos no limitativos, una región de HS1 que comprende una secuencia central de HS1 tiene desde aproximadamente 300 pb hasta aproximadamente 1200 pb, por ejemplo, desde aproximadamente 300 pb hasta aproximadamente 400 pb, desde aproximadamente 400 pb hasta aproximadamente 500 pb, desde aproximadamente 500 pb hasta aproximadamente 600 pb, desde aproximadamente 600 pb hasta aproximadamente 700 pb, desde aproximadamente 700 pb hasta aproximadamente 800 pb, desde aproximadamente 800 pb hasta aproximadamente 900 pb, desde aproximadamente 900 pb hasta aproximadamente 1000 pb, desde aproximadamente 1000 pb hasta aproximadamente 1100 pb, o desde aproximadamente 1100 pb hasta aproximadamente 1200 pb, de longitud. En una realización no limitativa, la región LCR de  $\beta$ -globina no comprende una región de HS1 de aproximadamente 1,0 kb pb. En una realización no limitativa, la región LCR de  $\beta$ -globina no comprende una región de HS1 de aproximadamente 1,1 kb.

20 En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 que conserva la actividad de HS1, por ejemplo, actividad potenciadora, o que funciona como elemento facilitador o regulador para anclar la actividad potenciadora de otras regiones de HS, por ejemplo, HS2-4. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 que puede potenciar la transcripción de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana). En ejemplos no limitativos, la región LCR de  $\beta$ -globina no comprende una región de HS1 cuya capacidad para potenciar la transcripción de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana) es de no menos del 60%, no menos del 70%, no menos del 80%, no menos del 90%, o no menos del 95% en comparación con una región de HS1 nativa. En ejemplos no limitativos, la región LCR de  $\beta$ -globina no comprende una región de HS1 cuya capacidad para anclar la actividad potenciadora de uno o más de HS2-HS4 es de no menos del 60%, no menos del 70%, no menos del 80%, no menos del 90%, o no menos del 95% en comparación con una región de HS1 nativa.

35 En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 que comprende uno, dos o tres de los siguientes sitios de unión: dos sitios de unión a GATA-1 (por ejemplo, TTATCT, y CTATCA), y un sitio de unión para EKLF (por ejemplo, CCACACACA). En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 que comprende dos de los sitios de unión descritos anteriormente. Por ejemplo, en determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 que comprende dos sitios de unión a GATA-1 y no comprende un sitio de unión para EKLF. En determinadas realizaciones, la región LCR de  $\beta$ -globina no comprende una región de HS1 que comprende uno, no dos, sitios de unión a AP1/NFE2 y un sitio de unión para EKLF.

En determinadas realizaciones, la región LCR de  $\beta$ -globina comprende una región de HS3 y una región de HS4, y la región LCR de  $\beta$ -globina no comprende una región de HS1 o una región de HS2. En determinadas realizaciones, la región de HS3 y región de HS4 dentro de la región LCR de  $\beta$ -globina son contiguas. En una realización no limitativa, la región LCR de  $\beta$ -globina consiste esencialmente en una región de HS3 y una región de HS4. En otra realización, la región LCR de  $\beta$ -globina comprende dos sitios de unión a GATA-1 introducidos en la unión entre la región de HS3 y la región de HS4. La región de HS3 puede encontrarse entre el gen de globina o parte funcional del mismo y la región de HS4.

En determinadas realizaciones, la región LCR de  $\beta$ -globina comprende una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5 y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8, y la región LCR de  $\beta$ -globina no comprende una región de HS1 o una región de HS2.

En una realización no limitativa, la región LCR de  $\beta$ -globina comprende una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5 y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6, y la región LCR de  $\beta$ -globina no comprende una región de HS1 o una región de HS2, tal como se muestra en la figura 4.

#### *Gen de globina*

Según el contenido dado a conocer en el presente documento, el casete de expresión comprende un gen de globina o una parte funcional del mismo. El gen de globina puede ser un gen de  $\beta$ -globina, un gen de  $\gamma$ -globina, o un gen de  $\delta$ -globina. En determinadas realizaciones, el casete de expresión comprende un gen de  $\beta$ -globina humana. Según el contenido dado a conocer en el presente documento, el gen de  $\beta$ -globina humana puede ser un gen de  $\beta$ -globina humana silvestre, un gen de  $\beta$ -globina humana delecionado que comprende una o más deleciones de secuencias de intrones, o un gen de  $\beta$ -globina humana mutado que codifica para al menos un residuo de aminoácido anti-falciformación. En una realización no limitativa, un casete de expresión dado a conocer en el presente documento comprende un gen de  $\beta$ -globina humana silvestre. En otra realización, el un casete de expresión dado a conocer en el presente documento comprende un gen de  $\beta^A$ -globina humana que codifica para una mutación de treonina a glutamina en el codón 87 ( $\beta^{A-T87Q}$ ). El residuo de glutamina en la posición 87 en la cadena de gamma-globina aumenta la actividad anti-falciformación de la cadena gamma con respecto a la cadena beta, al tiempo que conserva características de unión a oxígeno de adulto de la cadena beta (Nagel *et al.*, Proc. Natl. Acad. Sci. U.S.A. (1979); 76: 670-672). En determinadas realizaciones, una parte funcional de un gen de globina tiene una identidad de al menos el 80%, al menos el 90%, al menos el 95%, o al menos el 99% con respecto a una secuencia de polinucleótido de referencia silvestre correspondiente.

#### *Promotores y potenciadores*

Según el contenido dado a conocer en el presente documento, el casete de expresión puede comprender además un promotor de  $\beta$ -globina. En determinadas realizaciones, el promotor de  $\beta$ -globina está posicionado entre el gen de globina o parte funcional del mismo y la región LCR de  $\beta$ -globina. La longitud y la secuencia del promotor de  $\beta$ -globina pueden variar. En determinadas realizaciones, el promotor de  $\beta$ -globina tiene desde aproximadamente 100 pb hasta aproximadamente 1600 pb de longitud, por ejemplo, desde aproximadamente 200 pb hasta aproximadamente 700 pb, desde aproximadamente 100 pb hasta aproximadamente 200 pb, desde aproximadamente 200 pb hasta aproximadamente 300 pb, desde aproximadamente 300 pb hasta aproximadamente 400 pb, desde aproximadamente 400 pb hasta aproximadamente 500 pb, desde aproximadamente 500 pb hasta aproximadamente 600 pb, desde aproximadamente 600 pb hasta aproximadamente 700 pb, desde aproximadamente 700 pb hasta aproximadamente 800 pb, desde aproximadamente 800 pb hasta aproximadamente 900 pb, desde aproximadamente 900 pb hasta aproximadamente 1000 pb, desde aproximadamente 1000 pb hasta aproximadamente 1100 pb, desde aproximadamente 1100 pb hasta aproximadamente 1200 pb, desde aproximadamente 1200 pb hasta aproximadamente 1300 pb, desde aproximadamente 1300 pb hasta aproximadamente 1400 pb, desde aproximadamente 1400 pb hasta aproximadamente 1500 pb, o desde aproximadamente 1500 pb hasta aproximadamente 1600 pb de longitud. En determinadas realizaciones, el promotor de  $\beta$ -globina un promotor de  $\beta$ -globina humana que tiene aproximadamente 130 pb, aproximadamente 613 pb, aproximadamente 265 pb, o aproximadamente 1555 pb, de longitud. En una realización, el promotor de  $\beta$ -globina es un promotor de  $\beta$ -globina humana que tiene aproximadamente 613 pb de longitud. En un ejemplo no limitativo, el promotor de  $\beta$ -globina humana tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 10, que se proporciona a continuación:

AAGCAATAGATGGCTCTGCCCTGACTTTTTATGCCAGCCCTGGCTCCTGCCCTCCCTGCTCCT  
 GGGAGTAGATTGGCCAACCCCTAGGGTGTGGCTCCACAGGGTGAGGTCTAAGTGATGACAGCCG  
 TACCTGTCCCTGGCTCTTCTGGCACTGGCTTAGGAGTTGGACTTCAAACCCCTCAGCCCTCCCT  
 CTAAGATATATCTCTTGGCCCCATACCATCAGTACAAATTGCTACTAAAAACATCCTCCTTTG  
 CAAGTGATTTACGTAATATTTGGAATCACAGCTTGGTAAGCATATTGAAGATCGTTTTCCCA  
 ATTTTCTTATTACACAAAATAAGAAATTGATGCACATAAAGTGGAAAGAGTTTTGTCTACCATAA  
 TTCAGCTTTGGGATATGTAGATGGATCTCTTCCCTGCGTCTCCAGAATATGCAAATACTTACA  
 GGACAGAATGGATGAAAACCTCTACCTCAGTTCCTAAGCATATCTTCTCCTTATTTGGATTAAAA  
 CCTTCTGGTAAGAAAAGAAAAAATATATATATATATGTGTATATATACACACATACATATA  
 CATATATATGCATTTCATTTGTTGTTGTTTTCTTAATTTGCTCATG [SEQ ID NO:10]

En una realización, el promotor de  $\beta$ -globina es un promotor de  $\beta$ -globina humana que tiene aproximadamente 265 pb de longitud. En un ejemplo no limitativo, el promotor de  $\beta$ -globina humana tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 11.

5 AAGCAATAGATGGCTCTGCCCTGACTTTTTATGCCAGCCCTGGCTCCTGCCCTCCCTGCTCCT  
 GGGAGTAGATTGGCCAACCCCTAGGGTGTGGCTCCACAGGGTGAGGTCTAAGTGATGACAGCCG  
 TACCTGTCCCTGGCTCTTCTGGCACTGGCTTAGGAGTTGGACTTCAAACCCCTCAGCCCTCCCT  
 CTAAGATATATCTCTTGGCCCCATACCATCAGTACAAATTGCTACTAAAAACATCCTCCTTTG  
 CAAGTGATTTAC [SEQ ID NO:11]

Adicional o alternativamente, un casete de expresión dado a conocer en el presente documento puede comprender además un potenciador en 3' de  $\beta$ -globina humana. En determinadas realizaciones, el potenciador en 3' de  $\beta$ -globina humana está posicionado aguas arriba del gen de globina o parte funcional del mismo. En determinadas realizaciones, el potenciador en 3' de  $\beta$ -globina tiene desde aproximadamente 500 pb hasta aproximadamente 1000 pb de longitud, por ejemplo, desde aproximadamente 500 pb hasta aproximadamente 600 pb, desde aproximadamente 600 pb hasta aproximadamente 700 pb, desde aproximadamente 700 pb hasta aproximadamente 800 pb, o desde aproximadamente 800 pb hasta aproximadamente 900 pb de longitud. En una realización, el potenciador en 3' de  $\beta$ -globina humana tiene aproximadamente 879 pb de longitud. En un ejemplo, el potenciador en 3' de  $\beta$ -globina humana tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 12.

10 TAGGTATTGAATAAGAAAAATGAAGTTAAGGTGGTTGATGGTAACACTATGCTAATAAC  
 TGCAGAGCCAGAAGCACCATAAAGGACATGATAAGGGAGCCAGCAGACCTCTGATCTCTT  
 CCTGAATGCTAATCTTAAACATCCTGAGGAAGAATGGGACTTCCATTTGGGGTGGGCCTA  
 TGATAGGTAAATAAGACAGTAGTGAATATCAAGCTACAAAAGCCCTTTCAAATTCTT  
 CTCAGTCCTAACCTTTTCACTAAGCCAGTCTTCCAAAGCAGACTGTGAAAGAGTGAT  
 AGTTCCGGGAGACTAGCACTGCAGATTCGGGTCACTGTGAGTGGGGGAGGCAGGGAAGA  
 AGGGCTCACAGGACAGTCAAACCATGCCCTGTTTTTCTTCAAGTAGACCTCTAT  
 AAGACAACAGAGACAACCTAAGGCTGAGTGGCCAGGCGAGGAGAAACCATCTCGCCGTAAA  
 ACATGGAAGGAACACTTCAGGGGAAAGGTGGTATCTCTAAGCAAGAGAACTGAGTGGAGT  
 CAAGGCTGAGAGATGCAGGATAAGCAAATGGGTAGTAAAAGACATTCATGAGGACAGCT  
 AAAACAATAAGTAATGTAATAACAGCATAGCAAACCTTTAACCTCCAAATCAAGCCTCT  
 ACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTG  
 CATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAAGATATAGTGTATTTTCCCA  
 AGGTTTGAACCTAGCTCTTCAATTTCTTTATGTTTTAAATGCACCTGCCACATTCCCT  
 TTTTAGTAAAATATTCAGAAAATAATTTAAATACATCATTG [SEQ ID NO:12]

Además, un casete de expresión dado a conocer en el presente documento puede comprender además al menos un potenciador específico de eritroides. El casete de expresión dado a conocer en el presente documento permite la expresión de un gen de globina (por ejemplo, gen de  $\beta$ -globina humana) de una manera específica de eritroides. El potenciador específico de eritroides puede potenciar la expresión del gen de globina de una manera específica de eritroides. Por ejemplo, el potenciador específico de eritroides carece de actividad potenciadora en tejidos no eritroides. Particularmente, para la región LCR de  $\beta$ -globina que carece de una región de HS2, que funciona principalmente como potenciador de la expresión, la adición de uno o más potenciadores específicos de eritroides puede compensar la actividad de potenciación de una región de HS2. Además, los potenciadores específicos de eritroides dados a conocer en el presente documento no disminuyen o reducen el título de un vector que comprende el casete de expresión. La longitud del potenciador específico de eritroides puede variar, por ejemplo, desde aproximadamente 100 pb hasta aproximadamente 200 pb, desde aproximadamente 100 pb hasta aproximadamente 120 pb, desde aproximadamente 120 pb hasta aproximadamente 140 pb, desde aproximadamente 140 pb hasta aproximadamente 200 (por ejemplo, desde aproximadamente 140 pb hasta aproximadamente 150 pb, desde aproximadamente 150 pb hasta aproximadamente 160 pb, desde aproximadamente 160 pb hasta aproximadamente 170 pb, desde aproximadamente 170 pb hasta aproximadamente 180 pb, desde aproximadamente 180 pb hasta aproximadamente 190 pb, o desde aproximadamente 190 pb hasta aproximadamente 200 pb). En determinadas realizaciones, el potenciador específico de eritroides tiene una longitud de desde aproximadamente 140 pb hasta

aproximadamente 200 pb. En una realización no limitativa, el potenciador específico de eritroides tiene una longitud de 152 pb, que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 13, que se proporciona a continuación:

TCTCCCACGCCCTGGTCTCAGCTTGGGGAGTGGTCAGACCCCCAATGGCGATAAACTCTGGCAA  
CTTTATCTGTGcaCTGCAGGCTCAGCCCCAAcaGCTTTAGCTTTTCAAGCAGGCAGGGGAAG  
GGAAACACATATCTCCAGATATGAGG [SEQ ID NO:13]

5 En una realización no limitativa, el potenciador específico de eritroides tiene una longitud de 157 pb, que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 14, que se proporciona a continuación:

CTAAACCCCTCCCCACCCTAGCCCCAAGCTTCATCTTAGCTCCACTCCTGACCCTATCCAGC  
TAAAGGTCCCCACCCAGCTCCTGCCTATCTAGTCATTCATATGGCAAGACTTGAAAGTCCTA  
TCTCAAAGCAGCAGAATTATCAGCTACGACT [SEQ ID NO:14]

En una realización no limitativa, el potenciador específico de eritroides tiene una longitud de 141 pb, que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 15, que se proporciona a continuación:

CCATCCCCCAGCACTCCCTGCCCCACAGCCAGACTTGACCAACTCCCAGCTccGCCTGGGA  
CTTCCAGATATGGGGCCCCACCCTGCAGGCCCTTGGGGACGCTGAAGATATTGACTATCTGCC  
TGCCggAAAAGGGTG [SEQ ID NO:15]

10 En una realización no limitativa, el potenciador específico de eritroides tiene una longitud de 171 pb, que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 16, que se proporciona a continuación:

AAAGGCTGGGGGTGGGAGTAGCGGATTTGAAGCACTTGTGGCCTACAGAGGTGTGGCAAGCA  
GAGCACCTCAGAACTCAGGCGTACTGCCCGCCGCCGAGCCCTGCGAGGGCCGATAGCGAGGG  
TGTGGCCCTTATCTGCACCCAGCAGAGCGCCGGCGGGGTACGGTC [SEQ ID NO:16]

En una realización no limitativa, el potenciador específico de eritroides tiene una longitud de 195 pb, que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 17, que se proporciona a continuación:

CAGTTGCCCTCAGCTGAGTATGTCTTCTAAAGATAATGTCGATTGTGTATGGCTGATGGGATTC  
TAGGACCAAGCAAGAGGTTTTTTTTTTCCCCACATACTTAACGTTTCTATATTTCTATTTG  
AATTCGACTGGACAGTTCCATTTGAATTAATTTCTCTCTCTCTCTCTCTGACACATTTTATC  
15 TTGCCA [SEQ ID NO:17]

Pueden identificarse y determinarse potenciadores específicos de eritroides mediante cualquier método adecuado conocido en la técnica. Los potenciadores específicos de eritroides pueden estar posicionados en la LTR en 3' (aguas abajo) o la LTR en 5' (aguas arriba) de la región LCR de  $\beta$ -globina. En una realización, el al menos un potenciador específico de eritroides está posicionado en la LTR en 5' de la región LCR de  $\beta$ -globina, por ejemplo, aguas arriba de la región de HS3. El casete de expresión puede comprender uno, dos, tres, cuatro o cinco potenciadores específicos de eritroides. En una realización, el casete de expresión comprende un potenciador específico de eritroides. En otra realización, el casete de expresión comprende dos potenciadores específicos de eritroides. En aún otra realización, el casete de expresión comprende tres potenciadores específicos de eritroides. En determinadas realizaciones, el casete de expresión comprende cuatro potenciadores específicos de eritroides. En una realización no limitativa, el casete de expresión comprende cinco potenciadores específicos de eritroides.

*Aislantes*

Según el contenido dado a conocer en el presente documento, el casete de expresión comprende al menos uno de los aislantes descritos anteriormente. Según la invención, un casete de expresión dado a conocer en el presente documento comprende al menos un aislante que comprende la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, por ejemplo, pero sin limitarse a, un aislante que comprende SEQ ID NO: 24 o SEQ ID NO: 25, tal como un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1 (es decir, aislante A1). En diversas realizaciones no limitativas, el aislante puede incorporarse o insertarse en una o ambas LTR o en otra parte en la región de un casete de expresión dado a conocer en el presente documento que se integra en el genoma celular. En una realización, el aislante está posicionado en el extremo 3' del casete de expresión. En una realización, el aislante está posicionado en el extremo 5' del casete de expresión. En una realización, el casete de expresión comprende dos del aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1, en el que un aislante está posicionado en el extremo 3' y el otro aislante está posicionado en el extremo 5' del casete de expresión.

Los aislantes dados a conocer en el presente documento presentan una potente actividad de bloqueo de potenciador. En determinadas realizaciones, los aislantes presentan actividad de barrera además de actividad de bloqueo de potenciador. Los aislantes dados a conocer en el presente documento disminuyen sustancialmente los riesgos de mutagénesis por inserción y genotoxicidad asociados con vectores virales. Además, cuando se incorpora un aislante dado a conocer en el presente documento en un vector, el aislante no afecta de manera adversa a los títulos de vector del vector. En determinadas realizaciones, los aislantes (por ejemplo, aislante A1) aumentan la

expresión *in vivo* del gen de globina o parte funcional del mismo.

Con fines de ilustración y no de limitación, las figuras 1-4 muestran vectores recombinantes que comprenden casetes de expresión a modo de ejemplo según determinadas realizaciones del contenido dado a conocer en el presente documento. La figura 1 muestra un vector recombinante que comprende un casete de expresión dado a conocer en el presente documento que comprende un gen de  $\beta^{A-T87Q}$ -globina humana, que está operativamente unido a una región LCR de  $\beta$ -globina que comprende una región de HS2 de 860 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 9), una región de HS3 de 1301 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5), y una región de HS4 de 1065 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7).

La figura 2 muestra un vector recombinante a modo de ejemplo que comprende un casete de expresión según una realización del contenido dado a conocer en el presente documento. La figura 2 muestra un vector recombinante que comprende un casete de expresión dado a conocer en el presente documento que comprende un gen de  $\beta^{A-T87Q}$ -globina humana, que está operativamente unido a una región LCR de  $\beta$ -globina que comprende una región de HS1 de 1,1 kb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 2), una región de HS3 de 1301 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5), y una región de HS4 de 1065 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6).

La figura 3 muestra un vector recombinante a modo de ejemplo que comprende un casete de expresión según una realización del contenido dado a conocer en el presente documento. La figura 3 muestra un vector recombinante que comprende un casete de expresión dado a conocer en el presente documento que comprende un gen de  $\beta^{A-T87Q}$ -globina humana, que está operativamente unido a una región LCR de  $\beta$ -globina que comprende una región de HS1 de 602 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 3), una región de HS3 de 1301 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5), y una región de HS4 de 446 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 8).

La figura 4 muestra un vector recombinante a modo de ejemplo que comprende un casete de expresión según una realización del contenido dado a conocer en el presente documento. La figura 4 muestra un vector recombinante que comprende un casete de expresión dado a conocer en el presente documento que comprende un gen de  $\beta^{A-T87Q}$ -globina humana, que está operativamente unido a una región LCR de  $\beta$ -globina que comprende una región de HS3 de 1301 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5), y una región de HS4 de 1065 pb (por ejemplo, una que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6). El casete de expresión mostrado en la figura 4 también comprende los cinco potenciadores específicos de eritroides siguientes (mostrados como "EE5" en la figura 4): un potenciador específico de eritroides que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 13, un potenciador específico de eritroides que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 14, un potenciador específico de eritroides que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 15, un potenciador específico de eritroides que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 16, y un potenciador específico de eritroides que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 17.

Tal como se muestra en las figuras 1-4, cada uno de los casetes de expresión comprende un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1 (es decir, aislante A1). Además, tal como se muestra en las figuras 1-4, cada uno de los casetes de expresión comprende un potenciador en 3' de  $\beta$ -globina humana de 879 pb, que está posicionado aguas arriba del gen de  $\beta$ -globina humana. Además, tal como se muestra en las figuras 1-4, cada uno de los vectores recombinantes comprende un elemento post-regulador de la hepatitis de la marmota (WPRE) y una señal de poliadenilación de hormona de crecimiento bovina en la repetición terminal larga (LTR) en 3' del vector (por ejemplo, 3' con respecto a la región R en la LTR en 3').

### III. Vectores, nucleasas y sistemas de CRISPR-Cas

El contenido dado a conocer en el presente documento proporciona vectores y sistemas de administración (por ejemplo, unas nucleasas modificadas por ingeniería o que no se producen de manera natural o un sistema de CRISPR-Cas) que comprenden los casetes de expresión descritos anteriormente. Los vectores y sistemas de administración son vehículos de administración adecuados para la introducción estable de gen de globina (por ejemplo,  $\beta$ -globina humana) en el genoma de una amplia gama de células diana para aumentar la expresión de la proteína globina (proteína  $\beta$ -globina humana) en la célula.

En determinadas realizaciones, el vector es un vector retroviral (por ejemplo, gamma-retroviral o lentiviral) que se emplea para la introducción o transducción del casete de expresión descrito anteriormente en el genoma de una célula huésped (por ejemplo, una célula madre hematopoyética, una célula madre embrionaria, una célula madre pluripotente inducida, o una célula endotelial hemogénica). En determinadas realizaciones, el vector retroviral comprende un casete de expresión que comprende uno de los aislantes descritos anteriormente, por ejemplo, aislante A1. El aislante puede estar posicionado en el extremo 3' o 5' del casete de expresión. En una realización, el aislante está posicionado en el extremo 3' del casete de expresión. Durante la transcripción inversa e integración de vector, el aislante posicionado en el extremo 3' se copia en el extremo 5' del casete de expresión. La topología resultante coloca copias del aislante entre las regiones genómicas ubicadas en la LTR en 5' y la LTR en 3' del virus integrado y actividad potenciadora de la LTR en 5' y el promotor de paquete interno, pero no contiene el potenciador

en la LTR en 3'. Esta topología puede disminuir la genotoxicidad, dando por tanto como resultado una formación de tumor reducida y supervivencia aumentada de los animales.

En determinadas realizaciones, el vector recombinante comprende además un elemento post-regulador de la hepatitis de la marmota (WPRE) en la repetición terminal larga (LTR) en 3' del vector (por ejemplo, 3' con respecto a la región R en la LTR en 3' del vector). En determinadas realizaciones, el vector recombinante comprende además una señal de poliadenilación de hormona de crecimiento bovina además del WPRE en la repetición terminal larga (LTR) en 3' del vector (por ejemplo, 3' con respecto a la región R en la LTR en 3' del vector). Una característica esencial de vectores de globina terapéuticos es lograr un alto título, suficiente para una transducción eficaz de células del paciente. Debido a su gran carga útil, que comprende un gen, promotor, potenciadores y/o elementos de LCR, los vectores lentivirales de globina tienen inherentemente un título bajo, complicando su fabricación y limitando su uso clínico. Este problema se complica adicionalmente por la incorporación de elementos genómicos adicionales tales como un aislante, que aumentan adicionalmente el tamaño del vector. El WPRE puede aumentar el título del vector recombinante. La adición de una señal de poliadenilación de hormona de crecimiento bovina al WPRE puede aumentar adicionalmente el título del vector recombinante. En determinadas realizaciones, el WPRE y la señal de poliadenilación de hormona de crecimiento bovina no están comprendidos dentro del casete de expresión, y por tanto, no se transfieren a las células transducidas con el vector recombinante. La incorporación de estos elementos para potenciar la producción de vectores lentivirales de globina es esencial para producir títulos superiores y por tanto para la utilidad clínica de los vectores descritos en esta solicitud.

En un ejemplo no limitativo, un casete de expresión dado a conocer en el presente documento puede clonarse en el interior de un vector retroviral y puede impulsarse la expresión a partir de su promotor endógeno, a partir de la repetición terminal larga retroviral, o a partir de un promotor interno alternativo. También son adecuadas combinaciones de vector retroviral y una línea de empaquetamiento apropiada, en las que las proteínas de cápside serán funcionales para infectar células humanas. Se conocen diversas líneas celulares productoras de virus anfótropo, incluyendo, pero sin limitarse a, PA12 (Miller, *et al.* (1985) *Mol. Cell. Biol.* 5: 431-437); PA317 (Miller, *et al.* (1986) *Mol. Cell. Biol.* 6: 2895-2902); y CRIP (Danos, *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85: 6460-6464). También son adecuadas partículas no anfótropas, por ejemplo, partículas pseudotipadas con envuelta de VSVG, RD114 o GALV y cualquier otra conocida en la técnica.

Los métodos de transducción adecuados también incluyen el cultivo conjunto directo de las células con células productoras, por ejemplo, mediante el método de Bregni, *et al.* (1992) *Blood* 80: 1418-1422, o cultivando con sobrenadante viral solo o disoluciones madre de vector concentradas con o sin factores de crecimiento apropiados y policaciones, por ejemplo, mediante el método de Xu, *et al.* (1994) *Exp. Hemat.* 22: 223-230; y Hughes, *et al.* (1992) *J. Clin. Invest.* 89: 1817.

Pueden usarse vectores virales de transducción para expresar un gen de globina (por ejemplo, un gen de  $\beta$ -globina humana) en una célula huésped (por ejemplo, células madre hematopoyéticas, una célula madre embrionaria, o una célula madre pluripotente inducida). Preferiblemente, el vector elegido muestra una alta eficacia de infección e integración y expresión estables (véase, por ejemplo, Cayouette *et al.*, *Human Gene Therapy* (1997); 8: 423-430; Kido *et al.*, *Current Eye Research* (1996); 15: 833-844; Bloomer *et al.*, *Journal of Virology* (1997); 71: 6641-6649; Naldini *et al.*, *Science* (1996); 272: 263-267; y Miyoshi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 94: 10319, 1997). Otros vectores virales que pueden usarse incluyen, por ejemplo, vectores adenovirales, lentivirales y virales adenoasociados, virus vaccinia, un virus de papiloma bovino, o un virus del herpes, tales como virus de Epstein-Barr (véase también, por ejemplo, los vectores de Miller, *Human Gene Therapy* (1990); 15-14; Friedman, *Science* (1989); 244: 1275-1281; Eglitis *et al.*, *Bio-Techniques* 6: 608-614, 1988; Tolstoshev *et al.*, *Current Opinion in Biotechnology* (1990); 1: 55-61; Sharp, *The Lancet* (1991); 337: 1277-1278; Cornetta *et al.*, *Nucleic Acid Research and Molecular Biology* (1987)36: 311-322; Anderson, *Science* (1984); 226: 401-409; Moen, *Blood Cells* (1991); 17: 407-416; Miller *et al.*, *Biotechnology* (1989); 7: 980-990; Le Gal La Salle *et al.*, *Science* (1993); 259: 988-990; y Johnson, *Chest* (1995); 107: 77S-83S). Los vectores retrovirales están particularmente bien desarrollados y se han usado en entornos clínicos (Rosenberg *et al.*, *N. Engl. J. Med* (1990); 323: 370; Anderson *et al.*, patente estadounidense n.º 5.399.346).

El requisito de administración e integración eficaces hacen que los vectores retrovirales sean adecuados para transducir un casete de expresión dado a conocer en el presente documento. Pueden derivarse vectores retrovirales partir de tres géneros de retrovirus: los  $\gamma$ -retrovirus (también conocidos como oncoretrovirus o retrovirus murinos de tipo C), los lentivirus, y los spumavirus (también conocidos como virus espumosos). Están disponibles varias revisiones que detallan enfoques moleculares para la generación de partículas de retrovirus defectuosos para la replicación (Cornetta *et al.* (2005); Cockrell & Kafri (2007)). El propio vector, que codifica para el ADNc o transgén terapéutico, conserva las secuencias virales mínimas necesarias para permitir el empaquetamiento en partículas virales en una línea celular de empaquetamiento, la transcripción inversa y la integración. La célula de empaquetamiento expresa las proteínas estructurales y enzimas necesarias que se requieren para ensamblar una partícula recombinante infecciosa que contiene la secuencia de vector y la maquinaria necesaria para su transcripción inversa e integración en la célula transducida.

Aunque los aspectos de fabricación de todos los tipos de vectores retrovirales siguen los mismos principios generales, los vectores  $\gamma$ -retrovirales, lentivirales y spumavirales difieren en algunas de sus propiedades biológicas

intrínsecas. Los gamma-retrovirus, incluyendo los virus de leucemia murina prototípico (VLM), infectan eficazmente muchos tipos de células pero no pueden integrarse en células que no proceden a la fase S poco después de su infección. En cambio, los lentivirus y sus derivados de vector pueden transducir células que no se dividen (Follenzi & Naldini, 2002; Salmon & Trono, 2002) debido a su capacidad para translocarse al núcleo e integrarse en ausencia de división celular (Lewis & Emerman, 1994; Goff, 2001). Otro atributo fundamental de vectores lentivirales es su estabilidad genómica relativa, tal como se establece para vectores lentivirales de globina (May *et al.*, 2000), lo cual contrasta con la inestabilidad genómica de vectores de globina basados en VLM (Leboulch *et al.*, 1994; Sadelain *et al.*, 1995). Los vectores lentivirales y espumosos proporcionan además una mayor capacidad de empaquetamiento (Kumar *et al.*, 2001; Retwilm, 2007). Los tres tipos de vector se han usado satisfactoriamente para la transducción de HSC activadas por citocina (Miyoshi *et al.*, 1999; Josephson *et al.*, 2002; Leurs *et al.*, 2003).

Estos tres sistemas de vector difieren en cuanto a sus patrones de integración. El patrón de integración de retrovirus es semialeatorio y está desviado hacia genes y sus inmediaciones en aproximadamente dos tercios de todos los acontecimientos de integración (Schroder *et al.*, 2002; Wu *et al.*, 2003; Mitchell *et al.*, 2004; De Palma *et al.*, 2005; Trobridge *et al.*, 2006). Sin embargo, hay diferencias sutiles y posiblemente significativas en cuanto a su distribución exacta. Los gamma-retrovirus tienen propensión para integrarse aguas arriba de genes transcritos, mientras que los lentivirus y vectores lentivirales seleccionan como diana toda la secuencia génica transcrita. Los vectores espumosos parecen ser menos propensos a integración intragénica (Trobridge *et al.*, 2006). En una realización, el vector que comprende el casete de expresión es un vector de lentivirus. Los vectores pueden derivarse de virus de inmunodeficiencia humana 1 (VIH-1), virus de inmunodeficiencia humana 2 (VIH-2), virus de inmunodeficiencia del simio (VIS), virus de inmunodeficiencia felina (VIF), virus de inmunodeficiencia bovina (VIB), virus de enfermedad de Jembrana (VEJ), virus de anemia infecciosa equina (VAIE), virus de artritis y encefalitis caprina (VAEC) y similares. En una realización no limitativa, el vector lentiviral es un vector de VIH. Los constructos basados en VIH son los más eficaces en la transducción de células humanas.

El patrón semialeatorio de la integración de vector expone a los pacientes al riesgo de oncogénesis por inserción cuando el vector trans-activa un oncogén cercano. Esto puede dar como resultado la expansión clonal (Ott *et al.*, 2006; Cavazzana-Calvo *et al.*, 2010), mielodisplasia (Stein *et al.*, 2010) o leucemia (Hacein-Bey-Abina *et al.*, 2003, 2008; Howe *et al.*, 2008). Las estrategias de administración génica dirigida, que usan una nucleasa modificada por ingeniería o que no se produce de manera natural (incluyendo, pero sin limitarse a, nucleasas de dedos de zinc (ZNF), meganucleasa, nucleasa efectora de tipo activador de la transcripción (TALEN)), o un sistema de CRISPR-Cas, pueden reducir o incluso eliminar la preocupación por oncogénesis por inserción que es inherente al uso de vectores retrovirales.

Las células eucariotas usan dos mecanismos de reparación del ADN diferenciados en respuesta a roturas de cadena doble (DSB) de ADN: recombinación homóloga (HR) y unión de extremos no homóloga (NHEJ). La activación de la maquinaria de reparación de HR depende del estado del ciclo celular, y está restringida a las fases S y G2; en cambio, la ruta de NHEJ es activa a lo largo de todo el ciclo celular. Desde el punto de vista mecánico, HR es un mecanismo de reparación del ADN libre de errores, porque requiere un molde homólogo para reparar la cadena de ADN dañada. Por otro lado, NHEJ es un mecanismo de reparación independiente de molde que no es preciso, debido al procesamiento de extremos de ADN durante la reparación que conduce a inserciones o deleciones en el sitio de rotura del ADN (Moynahan y Jasin, 2010). Debido a su mecanismo basado en homología, se ha usado HR como herramienta para modificar por ingeniería de manera específica del sitio el genoma de diferentes especies. Desde un punto de vista terapéutico, se ha usado HR de manera satisfactoria para reparar genes mutados, ofreciendo por tanto un enfoque prometedor al tratamiento mediado por células de enfermedades monogénicas (Porteus *et al.*, 2006).

El direccionamiento génico mediante HR requiere el uso de dos brazos de homología que flanquean al transgén/sitio diana de interés. Generalmente, se han usado ADN de plásmido convencionales para suministrar brazos de homología de 5-10 kb junto con transgenes para la selección positiva y negativa. Este método se usa habitualmente para desactivar/insertar genes en células madre embrionarias de ratón (mES) (Capecchi, 2005; figura 2B). En células humanas, el uso de este enfoque ha permitido el direccionamiento génico con eficacias del orden de  $10^{-6}$ , que son menores que en células mES y no resulta terapéuticamente práctico. La eficacia de HR puede aumentarse mediante la introducción de roturas de cadena doble (DSB) de ADN en el sitio diana usando endonucleasas de corte poco frecuente específicas, dando como resultado un aumento de más de 1.000 veces en el direccionamiento génico correcto (Jasin, 1996). El descubrimiento de este fenómeno motivó el desarrollo de métodos para crear DSB específicas de sitio en el genoma de diferentes especies. Se han diseñado diversas enzimas químicas para este fin a lo largo de la última década, concretamente nucleasas de dedos de zinc (ZFN), meganucleasas y nucleasas efectoras de tipo activadoras de la transcripción (TALEN).

Las ZFN son proteínas químicas modulares que contienen un dominio de unión a ADN (DBD) basado en ZF y un dominio de nucleasa FokI (Porteus y Carroll, 2005). El DBD está compuesto habitualmente por tres dominios de ZF, cada uno con una especificidad de 3 pares de bases; el dominio de nucleasa FokI proporciona una actividad de mellado de ADN, que se dirige mediante dos ZFN flanqueantes. Debido a la naturaleza modular del DBD, en principio puede seleccionarse como diana cualquier sitio en un genoma. Sin embargo, dado que una única ZFN puede unirse a, y mellar, ADN, existe posibilidad de un alto número de efectos fuera de la diana, dando como resultado la activación de la ruta de NHEJ que puede o bien introducir inserciones/deleciones o bien integrar el

vector de direccionamiento de una manera no específica. Recientemente se notificaron dominios de FokI obligados que pueden mellar su cadena de ADN respectiva únicamente cuando forman un heterodímero (Doyon *et al*, 2011). El uso de tales ZFN obligadas puede reducir los efectos genotóxicos de este enfoque.

5 Las meganucleasas (MN)/endonucleasas dirigidas (HE) son nucleasas de ADNcd que reconocen y escinden sitios de ADN grandes (14-40 pb) con bajas frecuencias de escisión en genomas eucariotas (Paques y Duchateau, 2007). Aunque esto limita los posibles sitios diana, se han usado estructuras de MN-ADN como guía para modificar de manera específica residuos de interacción con ADN con el fin de cambiar la especificidad de MN (Marcaida *et al*, 2010). Se ha modificado satisfactoriamente por ingeniería I-Crel para generar meganucleasas quiméricas que seleccionan como diana los genes XPC y RAG1 humanos, y se ha mostrado que estimulan la actividad de HR en  
10 células de mamífero sin genotoxicidad evidente (Redondo *et al*, 2008; Grizot *et al*, 2009). Se necesitará comparar la genotoxicidad de este enfoque con la de ZFN y nucleasas TALE.

Las TALEN son similares a ZFN excepto porque el DBD se deriva de efectores de tipo activadores de la transcripción (TALE), que son factores virulentos usados por bacterias fitopatógenas (Herbers, 1992). El DBD de TALE es modular, y está compuesto por repeticiones de 34 residuos, y su especificidad de ADN está determinada por el número y orden de repeticiones (Herbers, 1992). Cada repetición se une a un único nucleótido en la secuencia diana a través de tan sólo dos residuos (Boch, 2011). La ventaja con respecto a la tecnología de ZFN es la rápida construcción de DBD.

Varios estudios han usado estas enzimas quiméricas para estimular HR o bien para la adición de genes o bien para la reparación de genes en su sitio diana (Paques y Duchateau, 2007; Urnov *et al*, 2010). Porteus diseñó una ZFN para una secuencia de medio sitio a partir del HBB humano que rodea al nucleótido de mutación de células falciformes (Porteus, 2006). Esta ZFN selecciona como diana la secuencia y estimula la HR en una diana de ADN quimérico cuando se combina con una ZFN que selecciona como diana el sitio de unión a Zif268. Ha habido avances recientes en la selección como diana de genes en células CD34<sup>+</sup> de sangre de cordón umbilical. El uso de lentivirus no integrantes para administrar ZFN y el ADN donante en estas células para seleccionar como diana el gen de CCR5 se notificó en Lombardo *et al*, 2007. Lombardo *et al*, 2007 mostraron la adición de genes en este locus con un direccionamiento correcto en el 80% de las células seleccionadas de manera positiva.

El contenido dado a conocer en el presente documento proporciona un sistema de nucleasa modificado por ingeniería o que no se produce de manera natural que comprende un casete de expresión dado a conocer en el presente documento, tal como se describió anteriormente. Las nucleasas adecuadas incluyen, pero no se limitan a, ZFN, meganucleasas y TALEN. Una nucleasa dada a conocer en el presente documento comprende un dominio de unión a ADN y un dominio de escisión de nucleasa. El dominio de unión a ADN de la nucleasa puede modificarse por ingeniería para unirse a una secuencia de elección, por ejemplo, un sitio predeterminado. Un dominio de unión a ADN modificado por ingeniería puede tener una especificidad de unión diferenciada, en comparación con una nucleasa que se produce de manera natural. Los métodos de modificación por ingeniería incluyen, pero no se limitan a, diseño racional y diversos tipos de selección. Cualquier dominio de escisión adecuado puede unirse de manera operativa a un dominio de unión a ADN para formar una nucleasa. Por ejemplo, pueden fusionarse dominios de unión a ADN de proteínas de dedos de zinc (ZFP) a dominios de escisión de nucleasa para crear ZFN, una entidad funcional que puede reconocer su diana de ácido nucleico prevista a través de su dominio de unión a ADN de ZFP modificado por ingeniería y hacer que se corte el ADN cerca del sitio de unión a ZFP mediante la actividad nucleasa. Véase, por ejemplo, Kim *et al*. Proc Nat'l Acad Sci USA (1996); 93(3): 1156-1160. Asimismo, pueden fusionarse dominios de unión a ADN de TALE a dominios de escisión de nucleasa para crear TALEN. Véase, por ejemplo, la publicación estadounidense n.º 20110301073.

El dominio de escisión puede ser heterólogo con respecto al dominio de unión a ADN, por ejemplo, un dominio de unión a ADN de meganucleasa y dominio de escisión de una nucleasa diferente. Los dominios de escisión heterólogos pueden obtenerse a partir de cualquier endonucleasa o exonucleasa. Las endonucleasas a modo de ejemplo a partir de las cuales puede derivarse un dominio de escisión incluyen, pero no se limitan a, endonucleasas de restricción y endonucleasas dirigidas. Véase, por ejemplo, 2002-2003 Catalog, New England Biolabs, Beverly, Mass.; y Belfort *et al*. (1997) Nucleic Acids Res. 25: 3379-3388. Se conocen enzimas adicionales que escinden el ADN (por ejemplo, nucleasa S1; nucleasa de judía mungo; ADNasa I pancreática; nucleasa de micrococos; endonucleasa HO de levadura; véase también Linn *et al*. (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993). Pueden usarse una o más de estas enzimas (o regiones funcionales de las mismas) como fuente de dominios de escisión y semidominios de escisión.

De manera similar, puede derivarse un semidominio de escisión a partir de la nucleasa descrita anteriormente que requiere la dimerización para la actividad de escisión. En general, se requieren dos proteínas de fusión para la escisión si las proteínas de fusión comprenden semidominios de escisión. Alternativamente, puede usarse una única proteína que comprende dos semidominios de escisión. Los dos semidominios de escisión pueden derivarse a partir de la misma endonucleasa (o partes funcionales de la misma), o cada semidominio de escisión puede derivarse a partir de una endonucleasa diferente (o partes funcionales de la misma).

En determinadas realizaciones, el sistema de nucleasa comprende un casete de expresión que comprende dos de los aislantes descritos anteriormente, por ejemplo, dos del aislante que tiene la secuencia de nucleótidos expuesta

en SEQ ID NO: 1. Uno de los dos aislantes está posicionado en el extremo 3' del casete de expresión, y el otro aislante está posicionado en el extremo 5' del casete de expresión.

El contenido dado a conocer en el presente documento también proporciona un sistema de CRISPR-Cas modificado por ingeniería o que no se produce de manera natural que comprende el casete de expresión descrito anteriormente.

5 El sistema de CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas)-Cas (asociado a CRISPR) es un sistema de nucleasa modificado por ingeniería basado en un sistema bacteriano que puede usarse para la modificación por ingeniería del genoma. Se basa en parte de la respuesta inmunitaria adaptativa de muchas bacterias y arqueas. Cuando un virus o plásmido invade una bacteria, segmentos del ADN del invasor se convierten en ARN de CRISPR (ARNcr) mediante la respuesta "inmunitaria". Después el ARNcr se asocia, a través de una

10 región de complementariedad parcial, con otro tipo de ARN denominado ARNtracr para guiar una nucleasa de CRISPR-Cas a una región homóloga al ARNcr en el ADN diana denominada "protoespaciador". La nucleasa de CRISPR-Cas escinde el ADN para generar extremos romos en el DSB en sitios especificados por una secuencia guía de 20 nucleótidos contenida dentro del transcrito de ARNcr. La nucleasa de CRISPR-Cas requiere tanto el ARNcr como el ARNtracr para el reconocimiento y la escisión de ADN específicos del sitio. El sistema se ha

15 modificado por ingeniería de tal manera que el ARNcr y el ARNtracr pueden combinarse para dar una molécula (el "ARN guía individual"); y la parte equivalente a ARNcr del ARN guía individual puede modificarse por ingeniería para guiar la nucleasa de CRISPR-Cas para seleccionar como diana cualquier secuencia deseada (véase Jinek *et al.*, Science (2012); 337: 816-821). Por tanto, el sistema de CRISPR-Cas puede modificarse por ingeniería para crear una DSB en una diana deseada en un genoma. En determinadas realizaciones, el sistema de CRISPR-Cas

20 comprende una nucleasa de CRISPR-Cas y un ARN guía individual. Los ejemplos adecuados de nucleasas de CRISPR-Cas incluyen, pero no se limitan a, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocida como Csn1 y Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homólogos de las mismas, o versiones modificadas de las

25 mismas. Estas nucleasas de CRISPR-Cas se conocen; por ejemplo, la secuencia de aminoácidos de proteína Cas9 de *S. pyogenes* puede encontrarse en la base de datos SwissProt con el número de registro Q99ZW2. En algunas realizaciones, la nucleasa de CRISPR-Cas tiene actividad de escisión de ADN, por ejemplo, Cas9. En determinadas realizaciones, la nucleasa de CRISPR-Cas es Cas9. La nucleasa de CRISPR-Cas puede dirigir la escisión de una o ambas cadenas en la ubicación de una secuencia diana (por ejemplo, un sitio de locus seguro genómico).

30 Adicionalmente, la nucleasa de CRISPR-Cas puede dirigir la escisión de una o ambas cadenas dentro de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500, o más pares de bases desde el primer o el último nucleótido de una secuencia diana.

Las nucleasas y el sistema de CRISPR-Cas dados a conocer en el presente documento permiten la administración dirigida del casete de expresión. En determinadas realizaciones, un sistema de CRISPR-Cas dado a conocer en el

35 presente documento o el dominio de unión a ADN de una nucleasa dada a conocer en el presente documento se une a un sitio de locus seguro genómico. Una nucleasa o el sistema de CRISPR-Cas genera una rotura de cadena doble en el sitio de locus seguro genómico. Los sitios de locus seguros genómicos son regiones intragénicas o extragénicas del genoma humano que pueden adaptarse a la expresión predecible de ADN integrado sin efectos adversos sobre el organismo o la célula huésped. Un locus seguro útil debe permitir suficiente expresión

40 transgénica como para proporcionar niveles deseados de la proteína codificada por vector o ARN no codificante. Además, un sitio de locus seguro genómico no debe predisponer a las células a transformación maligna ni alterar las funciones celulares. Se describen métodos para identificar sitios de locus seguros genómicos en Sadelain *et al.*, "Safe Harbours for the integration of new DNA in the human genome", Nature Reviews (2012); 12: 51-58; Papapetrou *et al.*, "Genomic safe harbors permit high  $\beta$ -globin transgene expression in thalassemia induced pluripotent stem cells" Nat Biotechnol. (2011) enero; 29(1): 73-8, que se incorporan como referencia en su totalidad.

45 Un sitio de locus seguro genómico dado a conocer en el presente documento cumple uno o más (uno, dos, tres, cuatro, o cinco) de los cinco criterios siguientes: (i) distancia de al menos 50 kb desde el extremo 5' de cualquier gen (por ejemplo, desde el extremo 5' del gen), (ii) distancia de al menos 300 kb desde cualquier gen relacionado con cáncer, (iii) dentro de una estructura de cromatina abierta/accesible (medido mediante escisión de ADN con

50 nucleasas naturales o modificadas por ingeniería), (iv) ubicación fuera de una unidad de transcripción génica y (v) ubicación fuera de regiones ultraconservadas (UCR), microARN o ARN largo no codificante del genoma humano. Dado que el acontecimiento de oncogénesis por inserción más común es la transactivación de genes promotores de tumor cercanos, los dos primeros criterios excluyen la parte del genoma humano ubicada cerca de promotores de genes, en particular, genes relacionados con cáncer, que son genes funcionalmente implicados en cánceres humanos o los homólogos humanos de genes implicados en cáncer en organismos de modelo. La proximidad a genes de miARN es un criterio de exclusión porque los miARN están implicados en la regulación de muchos procesos celulares, incluyendo proliferación y diferenciación celular. Dado que la integración de vector dentro de una

55 unidad de transcripción puede alterar la función génica mediante la pérdida de función de un gen supresor de tumor o la generación de un producto génico sometido a corte y empalme de manera aberrante, el cuarto (iv) criterio excluye todos los sitios ubicados dentro de genes transcritos. También se excluyen UCR, que son regiones que están altamente conservadas a lo largo de múltiples vertebrados y se sabe que están enriquecidas para potenciadores y exones, y ARN largos no codificantes. En determinadas realizaciones, el sitio de locus seguro genómico es un sitio de locus seguro genómico extragénico. En determinadas realizaciones, el sitio de locus seguro genómico está ubicado en el cromosoma 1.

60

El contenido dado a conocer en el presente documento también proporciona polinucleótidos que codifican para los sistemas de nucleasa descritos anteriormente, vectores que comprenden los polinucleótidos que codifican para los sistemas de nucleasa descritos anteriormente, polinucleótidos que codifican para el sistema de CRISPR-Cas descrito anteriormente, y vectores que comprenden los polinucleótidos que codifican para el sistema de CRISPR-Cas descrito anteriormente.

Las nucleasas y polinucleótidos que codifican para estas nucleasas, y el sistema de CRISPR-Cas y polinucleótidos que codifican para el sistema de CRISPR-Cas, pueden administrarse *in vivo* o *ex vivo* mediante cualquier medio adecuado. Por ejemplo, nucleasas y sistema de CRISPR-Cas tal como se describen en el presente documento pueden administrarse a una célula (por ejemplo, una célula madre hematopoyética, una célula madre embrionaria, una célula madre pluripotente inducida, o una célula endotelial hemogénica) mediante un vector que comprende polinucleótidos que codifican para la nucleasa o el sistema de CRISPR-Cas. Puede usarse cualquier vector incluyendo, pero sin limitarse a, vectores de plásmido, vectores retrovirales (por ejemplo, vectores  $\gamma$ -retrovirales, vectores lentivirales y vectores de virus espumosos), vectores de adenovirus, vectores de poxvirus; vectores de virus del herpes y vectores virales adenoasociados, etc. En una realización, el vector que comprende un polinucleótido que codifica para una nucleasa descrita anteriormente o un sistema de CRISPR-Cas descrito anteriormente es un vector lentiviral. En una realización particular, el vector lentiviral es un vector lentiviral no integrante. Se describen ejemplos de vector lentiviral no integrante en Ory *et al.* (1996) Proc. Natl. Acad. Sci. USA 93: 11382-11388; Dull *et al.*, (1998) J. Viral. 72: 8463-8471; Zuffery *et al.* (1998) J. Viral. 72: 9873-9880; Follenzi *et al.*, (2000) Nature Genetics 25: 217-222; publicación de patente estadounidense n.º 2009/054985.

Adicionalmente, también pueden emplearse enfoques no virales para la expresión de un gen de globina en células. Por ejemplo, puede introducirse una molécula de ácido nucleico en una célula administrando el ácido nucleico en presencia de lipofección (Feigner *et al.*, Proc. Natl. Acad. Sci. U.S.A. 84: 7413, 1987; Ono *et al.*, Neuroscience Letters 17: 259, 1990; Brigham *et al.*, Am. J. Med. Sci. 298: 278, 1989; Staubinger *et al.*, Methods in Enzymology 101: 512, 1983), conjugación de asialoorosomucoide-polilisina (Wu *et al.*, Journal of Biological Chemistry 263: 14621, 1988; Wu *et al.*, Journal of Biological Chemistry 264: 16985, 1989), o mediante microinyección en condiciones quirúrgicas (Wolff *et al.*, Science 247: 1465, 1990). Otros medios no virales para la transferencia génica incluyen transfección *in vitro* usando fosfato de calcio, DEAE-dextrano, electroporación y fusión de protoplastos. Los liposomas también pueden ser posiblemente beneficiosos para la administración de ADN al interior de una célula. El trasplante de genes normales al interior de tejidos afectados de un sujeto también puede lograrse transfiriendo un ácido nucleico normal al interior de un tipo de célula cultivable *ex vivo* (por ejemplo, una célula primaria autóloga o heteróloga o progenie de la misma), tras lo cual la célula (o sus descendientes) se inyectan en el interior de un tejido seleccionado como diana o se inyectan de manera sistémica. También pueden derivarse receptores recombinantes u obtenerse usando transposasas. Puede obtenerse expresión transitoria mediante electroporación de ARN.

#### IV. Células

La modificación genética de células (por ejemplo, células madre hematopoyéticas, células madre embrionarias, células madre pluripotentes inducidas y células endoteliales hemogénicas) puede lograrse transduciendo una composición de células sustancialmente homogénea con una un constructo de ADN o ARN recombinante (por ejemplo, un vector o un sistema de administración que comprende el casete de expresión descrito anteriormente). El contenido dado a conocer en el presente documento proporciona células transducidas con los casetes de expresión descritos anteriormente, células transducidas con los vectores descritos anteriormente, y células transducidas con las nucleasas descritas anteriormente o con vectores que comprenden polinucleótidos que codifican para las nucleasas, y célula transducida con el sistema de CARISPR-Cas descrito anteriormente o con vectores que comprenden polinucleótidos que codifican para el sistema de CARISPR-Cas, que se denominan de manera colectiva "células transducidas". Tal como se describió anteriormente, los vectores, nucleasas y sistema de CRISPR-Cas se emplean para la transducción del casete de expresión en las células para expresar un gen de globina (por ejemplo, un gen de  $\beta$ -globina humana). En determinadas realizaciones, las células transducidas se administran a un sujeto para tratar y/o prevenir una enfermedad, trastorno o estado hematopoyético. Los aislantes dados a conocer en el presente documento pueden potenciar la eficacia de la transducción del casete de expresión en células.

Las células transducidas adecuadas incluyen, pero no se limitan a, células madre, células progenitoras y células diferenciadas. Tal como se usa en el presente documento, el término "progenitor" o "células progenitoras" se refiere a células que tienen la capacidad para autorrenovarse y diferenciarse para dar células más maduras. Las células progenitoras tienen una potencia reducida en comparación con células madre pluripotentes y multipotentes. Muchas células progenitoras se diferencian a lo largo de un único linaje, pero también pueden tener una capacidad proliferativa bastante extensa.

En determinadas realizaciones, las células transducidas son células madre. Las células madre tienen la capacidad para diferenciarse para dar los tipos de células apropiados cuando se administran a un nicho biológico particular, *in vivo*. Una célula madre es una célula no diferenciada que puede realizar (1) autorrenovación a largo plazo, o la capacidad para generar al menos una copia idéntica de la célula original, (2) diferenciación a nivel de célula individual para dar múltiples, y en algunos casos un único, tipos de células especializadas y (3) regeneración funcional *in vivo* de tejidos. Las células madre se subclasifican según su potencial de desarrollo como totipotentes, pluripotentes, multipotentes y oligo/unipotentes. Tal como se usa en el presente documento, el término "pluripotente"

significa la capacidad de una célula para formar todos los linajes del cuerpo o soma (es decir, el embrión propiamente dicho). Por ejemplo, las células madre embrionarias son un tipo de células madre pluripotentes que pueden formar células de cada una de las tres capas germinales, el ectodermo, el mesodermo y el endodermo. Tal como se usa en el presente documento, el término "multipotente" se refiere a la capacidad de una célula madre adulta para formar múltiples tipos de células de un linaje. Por ejemplo, las células madre hematopoyéticas pueden formar todas las células del linaje de células sanguíneas, por ejemplo, células linfoides y mieloides.

En determinadas realizaciones, las células transducidas son células madre embrionarias, células madre de médula ósea, células madre de cordón umbilical, células madre de placenta, células madre mesenquimatosas, células madre neuronales, células madre hepáticas, células madre pancreáticas, células madre cardíacas, células madre renales, y/o células madre hematopoyéticas. En una realización, las células transducidas son células madre hematopoyéticas (HSC). Las HSC dan lugar a células progenitoras hematopoyéticas (HPC) comprometidas que pueden generar todo el repertorio de células sanguíneas maduras a lo largo de toda la vida de un organismo. El término "célula madre hematopoyética" o "HSC" se refiere a células madre multipotentes que dan lugar a todos los tipos de células sanguíneas de un organismo, incluyendo linajes mieloides (por ejemplo, monocitos y macrófagos, neutrófilos, basófilos, eosinófilos, eritrocitos, megacariocitos/plaquetas, células dendríticas) y linfoides (por ejemplo, células T, células B, células NK). Cuando se trasplantan al interior de seres humanos o animales que reciben radiación letal, las células progenitoras y madre hematopoyéticas pueden volver a poblar la reserva de células hematopoyéticas eritroides, neutrófilos-macrófagos, megacariocitos y linfoides.

Las HSC pueden aislarse o recogerse de médula ósea, sangre de cordón umbilical o sangre periférica. Las HSC pueden identificarse según determinados marcadores fenotípicos o genotípicos. Por ejemplo, las HSC pueden identificarse por su pequeño tamaño, ausencia de marcadores de linaje (lin), baja tinción (población secundaria) con colorantes vital tales como rodamina 123 (rodamina APAGADAS, también denominadas rolo) o Hoechst 33342, y presencia de diversos marcadores antigénicos sobre su superficie, muchos de los cuales pertenecen a la agrupación de serie de diferenciación (por ejemplo, CD34, CD38, CD90, CD133, CD105, CD45, Ter119, y c-kit, el receptor para factor de células madre). En una realización, la célula transducida es una HSC CD34<sup>+</sup>.

En una realización, la célula transducida es una célula madre embrionaria. En otra realización, la célula transducida es una célula madre pluripotente inducida. En aún otra realización, la célula transducida es una célula endotelial hemogénica.

Aunque las HSC son el vehículo natural para restaurar la hematopoyesis a largo plazo, su uso tiene algunas limitaciones importantes. La primera es su relativa escasez, lo cual puede impedir eventualmente la terapia con HSC autólogas cuando el producto celular recogido es demasiado pequeño. La segunda es la dificultad para realizar pruebas de bioseguridad tales como análisis de sitio de integración y por consiguiente para seleccionar células con sitios de integración elegidos, dado que las HSC adultas no pueden replicarse *in vitro*. La tercera limitación es que la recombinación homóloga usando las tecnologías actuales es prácticamente imposible, poniendo por tanto en peligro la aparición de corrección génica. Todas estas limitaciones se deben en última instancia al hecho de que las HSC adultas no pueden expandirse *in vitro* sin perder su potencia de células madre. Estas limitaciones explican la importancia crítica de vectores virales tales como vectores gamma-retrovirales y lentivirales, que son notablemente rápidos y eficaces en lograr una transferencia génica estable. Esto es esencial cuando se trata con HSC que sólo están disponibles en cantidades limitadas.

El uso de células ES y madre pluripotentes inducidas (iPS) para la terapia génica de globina se da a conocer en Moi *et al.*, *Haematol*, 1 de marzo, 2008; 93(3): 325-330. Las células madre embrionarias (ES) pueden someterse a direccionamiento y corrección génicos, lo cual requiere una división celular *in vitro* no limitada sin perder la multipotencia. Chang *et al.*, *Proc Natl Acad Sci USA* 2006; 103: 1036-40 proporcionaron prueba de concepto de la viabilidad del enfoque de recombinación homóloga en ratones con anemia de células falciformes. Takahashi *et al.* *Cell* 2006; 126: 663-76 notificaron la reprogramación satisfactoria de fibroblastos para dar un estado de tipo madre embrionario. Se produjeron células obtenidas mediante este proceso de diferenciación inversa, denominadas células madre pluripotentes inducidas (iPS), exponiendo cultivos de fibroblastos en masa embrionarios o jóvenes adultos a vectores gamma-retrovirales que codificaban para 4 factores de transcripción, que son fisiológicamente activos en las células madre embrionarias, pero se desactivan generalmente cuando progresa la diferenciación. Las células cultivadas formaron colonias similares a colonias de células ES. Estos hallazgos se han confirmado y extendido por terceros a fibroblastos tanto de ratón como humanos (Meissner *et al.*, *Nat Biotechnol* 2007; 25: 1177-81; Nakagawa *et al.*, *Nat Biotechnol* 2007; 26: 101-6; Okita *et al.*, *Nature* 2007; 448: 313-7; Park *et al.*, *Nature* 2007; 451: 141-6; Takahashi *et al.*, *Nat Protoc* 2007; 2: 3081-9; Takahashi K *et al.*, *Cell* 2007; 131: 861-72; Wernig *et al.*, *Nature* 2007; 448: 318-24; Yu J *et al.*, *Science* 2007; 318: 1917-20). Rudolf Jaenisch y colaboradores lograron una terapia génica satisfactoria en un modelo de ratón de enfermedad de células falciformes, usando recombinación homóloga en células iPS de tipo ES (Hanna *et al.*, *Science* 2007; 318: 1920-3). Hasta ahora se ha aplicado el procedimiento principalmente a fibroblastos recogidos a partir de una biopsia de piel, que después se induce que se conviertan en iPS mediante transducción con vectores retrovirales que codifican para cuatro factores de transcripción de células madre. Las iPS pueden someterse a la corrección de la mutación de SC mediante técnicas de recombinación homóloga convencionales y después pueden diferenciarse *in vitro* para dar cantidades ilimitadas de células madre hematopoyéticas. El procedimiento completo termina con el trasplante autólogo de las HSC corregidas al interior del ratón donante original, cuya enfermedad de SC se habrá curado ahora. Esta técnica no sólo es útil para la

recombinación homóloga, sino que también puede potenciar la transferencia de gen de globina mediada por lentivirus para el tratamiento de  $\beta$ -talasemia proporcionando un medio para realizar análisis de sitio de integración detallado y expansión celular *in vitro* adecuada antes de administrar células por infusión al interior del receptor.

- 5 La célula del contenido dado a conocer en el presente documento puede ser autóloga (“propia”) o no autóloga (“no propia”, por ejemplo, alogénica, singénica o xenogénica). Tal como se usa en el presente documento, “autólogo” se refiere a células del mismo sujeto. Tal como se usa en el presente documento, “alogénico” se refiere a células de la misma especie que difieren genéticamente de la célula en comparación. Tal como se usa en el presente documento, “singénico” se refiere a células de un sujeto diferente que son genéticamente idénticas a la célula en comparación. Tal como se usa en el presente documento, “xenogénico” se refiere a células de una especie diferente de la célula en comparación. En determinadas realizaciones, la célula es autóloga, por ejemplo, se administra una célula transducida con el casete de expresión dado a conocer en el presente documento a un sujeto a partir del cual se recoge la célula, por ejemplo, la célula se recoge a partir de médula ósea, sangre de cordón umbilical, sangre periférica y/o tejido adiposo del sujeto. En determinadas realizaciones, la célula se obtiene o se recoge a partir de la médula ósea de un sujeto.
- 10
- 15 En determinadas realizaciones, antes de la transducción con el casete de expresión, se estimula previamente la célula, por ejemplo, en presencia de una o más citocinas (por ejemplo, IL-3, IL-1 $\alpha$ , IL-6, ligando de Kit (también conocido como “factor de células madre (SCF)”), y ligando de Flt-3), y/o una o más glicoproteínas (por ejemplo, trombopoyetina y fibronectina). En un ejemplo no limitativo, la célula se estimula previamente en presencia de ligando de Flt-3, SCF, trombopoyetina, interleucina 3 y fibronectina. La célula puede estimularse previamente durante aproximadamente 24 horas o más, por ejemplo, aproximadamente 48 horas, o aproximadamente 36 horas. Posteriormente, se transduce la célula con un casete de expresión dado a conocer en el presente documento, o un vector u otro sistema de administración que comprende tal casete de expresión. La transducción puede realizarse en una célula nueva o en una célula congelada. Se aísla ADN genómico de la célula para determinar el número de copias de vector y analizar el sitio de integración o estructura de vector integrado, por ejemplo, mediante análisis por transferencia de tipo South y/o mediante PCR cuantitativa. Para la cuantificación de ARNm de globina (por ejemplo, análisis de transgenes de  $\beta$ -globina humana), se extrae ARN total a partir de la célula. Puede usarse un ensayo de extensión de cebadores cuantitativo para la cuantificación de ARNm de globina.
- 20
- 25

#### V. Composiciones y formulaciones

- 30 El contenido dado a conocer en el presente documento proporciona composiciones farmacéuticas que comprenden una célula transducida dada a conocer en el presente documento tal como se describió anteriormente y un portador farmacéuticamente aceptable. Tal como se usa en el presente documento, “portador farmacéuticamente aceptable” incluye todos y cada uno de los disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardantes de la absorción, y similares que son fisiológicamente compatibles, incluyendo medios de cultivo celular farmacéuticamente aceptables. El portador farmacéuticamente aceptable puede ser adecuado para administración parenteral (por ejemplo, intravenosa, intramuscular, subcutánea o intraperitoneal), espinal o epidérmica (por ejemplo, mediante inyección, infusión o implante). Dependiendo de la vía de administración, el compuesto activo, por ejemplo, la célula transducida, puede recubrirse en un material para proteger el compuesto frente a la acción de ácidos y otras condiciones naturales que pueden inactivar el compuesto.
- 35

- 40 Los portadores farmacéuticamente aceptables incluyen dispersiones o disoluciones acuosas estériles y polvos estériles para la preparación extemporánea de dispersión o disoluciones inyectables estériles. El uso de tales medios y agentes para principios farmacéuticamente activos se conoce bien en la técnica. Excepto en la medida en que cualquier medio o agente convencional sea incompatible con las células transducidas, se contempla el uso de los mismos en las composiciones farmacéuticas de la invención.

- 45 Las composiciones farmacéuticas del contenido dado a conocer en el presente documento pueden comprender además uno o más polipéptidos, polinucleótidos, vectores que comprenden los mismos, células transducidas, etc., tal como se describe en el presente documento, formulados en disoluciones farmacéuticamente aceptables o fisiológicamente aceptables para su administración a una célula o un animal, ya sean solos, o en combinación con una o más de otras modalidades de terapia. Si se desea, las composiciones farmacéuticas del contenido dado a conocer en el presente documento pueden administrarse en combinación con otros agentes, incluyendo, pero sin limitarse a, citocinas, factores de crecimiento, hormonas, moléculas pequeñas o diversos agentes farmacéuticamente activos. Cualquier agente adicional que no afecte de manera adversa a la capacidad de la composición para administrar la terapia génica prevista puede incluirse en las composiciones.
- 50

- 55 En las composiciones farmacéuticas del contenido dado a conocer en el presente documento, los expertos habituales en la técnica conocen bien la formulación de disoluciones de portador y excipientes farmacéuticamente aceptables, al igual que el desarrollo de regímenes de tratamiento y dosificación adecuados para usar las composiciones particulares descritas en el presente documento en una variedad de regímenes de tratamiento, incluyendo, por ejemplo, formulación y administración oral, parenteral, intravenosa, intranasal e intramuscular.

Las composiciones farmacéuticas del contenido dado a conocer en el presente documento pueden administrarse por vía parenteral (por ejemplo, por vía intravenosa, intramuscular o intraperitoneal) tal como se describe, por ejemplo,

en la patente estadounidense n.º 5.543.158; patente estadounidense n.º 5.641.515 y patente estadounidense n.º 5.399.363. Pueden prepararse disoluciones de los compuestos activos como base libre o sales farmacológicamente aceptables en agua mezcladas de manera adecuada con un tensioactivo, tal como hidroxipropilcelulosa. También pueden prepararse dispersiones en glicerol, polietilenglicoles líquidos, y mezclas de los mismos o en aceites. En condiciones habituales de almacenamiento y uso, estas preparaciones contienen un conservante para prevenir el crecimiento de microorganismos.

Los portadores farmacéuticamente aceptables incluyen disoluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de dispersión o disoluciones inyectables estériles. El uso de tales medios y agentes para principios farmacéuticamente activos se conoce en la técnica. Excepto en la medida en que cualquier agente o medio convencional sea incompatible con el compuesto activo, se contempla el uso de los mismos en las composiciones farmacéuticas de la invención. También pueden incorporarse compuestos activos complementarios en las composiciones.

Las composiciones terapéuticas deben ser normalmente estériles y estables en las condiciones de fabricación y almacenamiento. La composición puede formularse como una disolución, microemulsión, liposoma u otra estructura ordenada adecuada para una alta concentración de fármaco. El portador farmacéuticamente aceptable puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares), y mezclas adecuadas de los mismos. La fluidez apropiada puede mantenerse, por ejemplo, mediante el uso de un recubrimiento tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de dispersión y mediante el uso de tensioactivos. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol, o cloruro de sodio en la composición. Puede proporcionarse absorción prolongada de las composiciones inyectables incluyendo en la composición un agente que retarda la absorción, por ejemplo, sales de monoestearato y gelatina.

Las composiciones farmacéuticas del contenido dado a conocer en el presente documento pueden proporcionarse de manera conveniente como preparaciones líquidas estériles, por ejemplo, disoluciones acuosas isotónicas, suspensiones, emulsiones, dispersiones o composiciones viscosas, que pueden tamponarse a un pH seleccionado. Las preparaciones líquidas son normalmente más fáciles de preparar que los geles, otras composiciones viscosas y composiciones sólidas. Adicionalmente, las composiciones líquidas son algo más convenientes de administrar, especialmente mediante inyección. Por otro lado, pueden formularse composiciones viscosas dentro del intervalo de viscosidad apropiado para proporcionar periodos de contacto más prolongados con tejidos específicos. Las composiciones líquidas o viscosas pueden comprender portadores, que pueden ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, solución salina, solución salina tamponada con fosfato, poliol (por ejemplo, glicerol, propilenglicol, polietilenglicol líquido, y similares) y mezclas adecuadas de los mismos.

Pueden prepararse disoluciones inyectables estériles incorporando las composiciones del contenido dado a conocer en el presente documento en la cantidad requerida del disolvente apropiado con diversas cantidades de los demás componentes, según se desee. Tales composiciones pueden estar en mezcla con un portador, diluyente o excipiente adecuado tal como agua estéril, solución salina fisiológica, glucosa, dextrosa, o similares. Las composiciones también pueden liofilizarse. Las composiciones pueden contener sustancias auxiliares tales como agentes humectantes, dispersantes o emulsionantes (por ejemplo, metilcelulosa), agentes tamponantes del pH, aditivos gelificantes o de potenciación de la viscosidad, conservantes, agentes saborizantes, colores y similares, dependiendo de la vía de administración y la preparación deseada. Pueden consultarse textos convencionales, tales como "REMINGTON'S PHARMACEUTICAL SCIENCE", 17ª edición, 1985, incorporado en el presente documento como referencia, para preparar preparaciones adecuadas, sin experimentación excesiva.

Pueden añadirse diversos aditivos que potencian la estabilidad y esterilidad de las composiciones, incluyendo conservantes antimicrobianos, antioxidantes, agentes quelantes y tampones. La prevención de la acción de microorganismos puede garantizarse mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico y similares. Puede provocarse la absorción prolongada de la forma farmacéutica inyectable mediante el uso de agentes que retrasan la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Las composiciones pueden ser isotónicas, es decir, pueden tener la misma presión osmótica que la sangre y el líquido lagrimal. La isotonicidad deseada de las composiciones del contenido dado a conocer en el presente documento puede lograrse usando cloruro de sodio u otros agentes farmacéuticamente aceptables tales como dextrosa, ácido bórico, tartrato de sodio, propilenglicol u otros solutos orgánicos o inorgánicos. Se prefiere particularmente cloruro de sodio para tampones que contienen iones de sodio.

Para la administración parenteral en una disolución acuosa, por ejemplo, la disolución debe tamponarse de manera adecuada si es necesario y hacer en primer lugar que el diluyente líquido sea isotónico con glucosa o solución salina suficiente. Pueden prepararse disoluciones inyectables estériles incorporando el compuesto activo en la cantidad requerida en un disolvente apropiado con uno o una combinación de componentes indicados anteriormente, según se requiera, seguido por esterilización por microfiltración. Generalmente, se preparan dispersiones incorporando el compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los demás componentes requeridos de los indicados anteriormente. En el caso de polvos estériles para la preparación de disoluciones

inyectables estériles, los métodos de preparación preferidos son secado a vacío y secado por congelación (liofilización) que proporcionan un polvo del principio activo más cualquier componente deseado adicional a partir de una disolución previamente esterilizada por filtración de los mismos.

5 En determinadas realizaciones, las composiciones pueden administrarse mediante pulverizaciones intranasales, inhalación y/u otros vehículos de administración de aerosol. Se describen métodos para administrar genes, polinucleótidos y composiciones de péptidos directamente a los pulmones mediante pulverizaciones de aerosoles nasales, por ejemplo, en la patente estadounidense n.º 5.756.353 y la patente estadounidense n.º 5.804.212. Se describen métodos de administración de fármacos usando compuestos de lisofosfatidil-glicerol, por ejemplo, en la patente estadounidense n.º 5.725.871. Se describe la administración de fármacos transmucosa en forma de una matriz de soporte de politetrafluoroetileno, por ejemplo, en la patente estadounidense n.º 5.780.045. Las composiciones del contenido dado a conocer en el presente documento pueden formularse para su administración encapsuladas en una partícula lipídica, un liposoma, una vesícula, una nanoesfera, una nanopartícula o similares. La formulación y el uso de tales vehículos de administración pueden llevarse a cabo usando técnicas conocidas y convencionales. Las formulaciones y composiciones del contenido dado a conocer en el presente documento pueden comprender uno o más represores y/o activadores que comprenden una combinación de cualquier número de polipéptidos, polinucleótidos y moléculas pequeñas, tal como se describen en el presente documento, formulados en disoluciones farmacéuticamente aceptables o fisiológicamente aceptables (por ejemplo, medio de cultivo) para su administración a una célula o un animal, o bien solos o bien en combinación con una o más de otras modalidades de terapia.

20 En determinados aspectos, el contenido dado a conocer en el presente documento proporciona formulaciones o composiciones adecuadas para la administración de sistemas de vector viral (es decir, transducción mediada por virus) incluyendo, pero sin limitarse a, vectores retrovirales (por ejemplo, lentivirales). Las formulaciones a modo de ejemplo para administración *ex vivo* también pueden incluir el uso de diversos agentes de transfección conocidos en la técnica, tales como fosfato de calcio, electoporación, choque térmico y diversas formulaciones de liposomas (es decir, transfección mediada por lípidos). Los liposomas son bicapas lipídicas que atrapan una fracción de líquido acuoso. El ADN se asocia de manera espontánea a la superficie externa de liposomas catiónicos (gracias a su carga) y estos liposomas interaccionarán con la membrana celular.

30 El experto en la técnica puede determinar fácilmente la cantidad de células y aditivos, vehículos y/o portador opcionales en composiciones y que van a administrarse en métodos del contenido dado a conocer en el presente documento. Normalmente, cualquier aditivo (además de la(s) célula(s) transducida(s) y/o agente(s)) está presente en una cantidad de desde aproximadamente el 0,001% hasta aproximadamente el 50% en peso) disolución en solución salina tamponada con fosfato, y el principio activo está presente en el orden de microgramos a miligramos, tal como desde aproximadamente el 0,0001% en peso hasta aproximadamente el 5% en peso, desde aproximadamente el 0,0001% en peso hasta aproximadamente el 1% en peso, desde aproximadamente el 0,0001% en peso hasta aproximadamente el 0,05% en peso, desde aproximadamente el 0,001% en peso hasta aproximadamente el 20% en peso, desde aproximadamente el 0,01% en peso hasta aproximadamente el 10% en peso, o desde aproximadamente el 0,05% en peso hasta aproximadamente el 5% en peso. Para cualquier composición que va a administrarse a un animal o ser humano, y para cualquier método de administración particular, debe determinarse la toxicidad, tal como determinando la dosis letal (DL) y DL50 en un modelo de animal adecuado, por ejemplo, roedor tal como ratón; y, la dosificación de la(s) composición/composiciones, concentración de componentes en la misma y momento de administración de la(s) composición/composiciones, que provocan una respuesta adecuada. Tales determinaciones no requieren experimentación excesiva a partir del conocimiento del experto en la técnica, esta divulgación y los documentos citados en el presente documento. Y el momento para administraciones secuenciales puede determinarse sin experimentación excesiva.

#### 45 VI. Usos y métodos

Los vectores y otros sistemas de administración (nucleasas y sistemas de CRISPR-Cas) que comprenden el casete de expresión dado a conocer en el presente documento proporcionan métodos de terapia génica mejorados. Tal como se usa en el presente documento, el término "terapia génica" se refiere a la introducción de un polinucleótido en el interior del genoma de una célula que restaura, corrige o modifica el gen y/o la expresión del gen. En diversas realizaciones no limitativas, un vector dado a conocer en el presente documento u otro sistema de administración (por ejemplo, una nucleasa o un sistema de CRISPR-Cas) comprende un casete de expresión que comprende un gen de globina o una parte funcional del mismo que codifica para una proteína globina (por ejemplo, proteína  $\beta$ -globina humana), que proporciona beneficios curativos, preventivos o de mejora para un sujeto al que se le ha diagnosticado o que se sospecha que tiene una enfermedad, trastorno o estado del sistema hematopoyético. El vector u otros sistemas de administración (por ejemplo, una nucleasa y el sistema de CRISPR-Cas) puede infectar y transducir la célula *in vivo*, *ex vivo* o *in vitro*. En realizaciones *ex vivo* e *in vitro*, las células transducidas pueden administrarse entonces a un sujeto que necesita terapia. El contenido dado a conocer en el presente documento contempla que los vectores y otros sistemas de administración (por ejemplo, nucleasas o sistemas de CRISPR-Cas), partículas virales y células transducidas del contenido dado a conocer en el presente documento van a usarse para tratar, prevenir y/o mejorar una enfermedad, trastorno o estado del sistema hematopoyético en un sujeto, por ejemplo, una hemoglobinopatía.

Tal como se usa en el presente documento, el término “hemoglobinopatía” o “estado hemoglobinopático” incluye cualquier trastorno que implica la presencia de una molécula de hemoglobina anómala en la sangre. Los ejemplos de hemoglobinopatías incluyen, pero no se limitan a, enfermedad de hemoglobina C, enfermedad de células falciformes (SCD) de hemoglobina, anemia de células falciformes y talasemias. También se incluyen hemoglobinopatías en las que está presente una combinación de hemoglobinas anómalas en la sangre (por ejemplo, enfermedad de células falciformes/Hb-C).

Tal como se usa en el presente documento, “talasemia” se refiere a un trastorno hereditario caracterizado por producción defectuosa de hemoglobina. Los ejemplos de talasemias incluyen  $\alpha$  y  $\beta$ -talasemia. Las  $\beta$ -talasemias están provocadas por una mutación en la cadena de beta-globina, y puede producirse en una forma mayor o menor. En la forma mayor de  $\beta$ -talasemia, los niños son normales en el nacimiento, pero desarrollan anemia durante el primer año de vida. La forma leve de  $\beta$ -talasemia produce glóbulos rojos pequeños y las talasemias están provocadas por la delección de un gen o genes de la cadena de globina. La  $\alpha$ -talasemia resulta normalmente de delecciones que implican a los genes HBA1 y HBA2. Ambos de estos genes codifican para  $\alpha$ -globina, que es un componente (subunidad) de hemoglobina. Hay dos copias del gen HBA1 y dos copias del gen HBA2 en cada genoma celular. Como resultado, hay cuatro alelos que producen  $\alpha$ -globina. Los diferentes tipos de una talasemia resultan de la pérdida de alguno o todos de estos alelos. El síndrome de Hb de Bart, la forma más grave de talasemia, resulta de la pérdida de los cuatro alelos de  $\alpha$ -globina. La enfermedad de HbH está provocada por una pérdida de tres de los cuatro alelos de [alfa]-globina. En estos dos estados, un acortamiento de [alfa]-globina impide que las células produzcan hemoglobina normal. En vez de eso, las células producen formas anómalas de hemoglobina denominadas hemoglobina de Bart (Hb Bart) o hemoglobina H (HbH). Estas moléculas de hemoglobina anómala no pueden transportar oxígeno de manera eficaz a los tejidos del cuerpo. La sustitución de hemoglobina normal por Hb de Bart o HbH provoca anemia y los demás problemas de salud graves asociados con una talasemia.

Tal como se usa en el presente documento, el término “enfermedad de células falciformes” se refiere a un grupo de trastornos de la sangre genéticos recesivos y autosómicos, que resultan de mutaciones en un gen de globina y que se caracterizan por glóbulos rojos que adoptan una forma falciforme, rígida, anómala. Se definen por la presencia de gen  $\beta^S$  que codifica para una variante de cadena de  $\beta$ -globina en la que se sustituye ácido glutámico por valina en la posición de aminoácido 6 del péptido, y un segundo gen  $\beta$  que tiene una mutación que permite la cristalización de HbS conduciendo a un fenotipo clínico. Tal como se usa en el presente documento, el término “anemia de células falciformes” se refiere a una forma específica de enfermedad de células falciformes en pacientes que son homocigóticos para la mutación que provoca HbS. Otras formas comunes de enfermedad de células falciformes incluyen HbS/ $\beta$ -talasemia, HbS/HbC y HbS/HbD.

En determinadas realizaciones, el contenido dado a conocer en el presente documento es para su uso en métodos de terapia génica para tratar, prevenir o mejorar una hemoglobinopatía que se selecciona del grupo que consiste en: enfermedad de hemoglobina C, enfermedad de células falciformes (SCD) de hemoglobina, anemia de células falciformes, anemia hereditaria, talasemia,  $\beta$ -talasemia, talasemia mayor, talasemia intermedia,  $\alpha$ -talasemia y enfermedad de hemoglobina H. En una realización no limitativa, la hemoglobinopatía es  $\beta$ -talasemia. En otra realización no limitativa, la hemoglobinopatía es anemia de células falciformes.

En diversas realizaciones no limitativas, vectores u otros sistemas de administración (por ejemplo, nucleasas o sistemas de CRISPR-Cas) que comprenden un casete de expresión dado a conocer en el presente documento se administran mediante inyección directa a una célula, tejido u órgano de un sujeto que necesita terapia génica, *in vivo*. En diversas otras realizaciones, se transducen células *in vitro* o *ex vivo* con vectores u otros sistemas de administración (por ejemplo, nucleasas o sistemas de CRISPR-Cas) del contenido dado a conocer en el presente documento, y opcionalmente se expanden *ex vivo*. Después se administran las células transducidas a un sujeto que necesita terapia génica, por ejemplo, dentro de una formulación farmacéutica dada a conocer en el presente documento.

El contenido dado a conocer en el presente documento proporciona un método de proporcionar una célula transducida a un sujeto. En diversas realizaciones no limitativas, el método comprende administrar (por ejemplo, por vía parenteral) una o más células (una población de células) transducidas con un casete de expresión dado a conocer en el presente documento o un vector u otro sistema de administración (por ejemplo, una nucleasa o sistema de CRISPR-Cas) que comprende tal casete de expresión al sujeto.

El contenido dado a conocer en el presente documento proporciona un método de tratamiento de una hemoglobinopatía en un sujeto. En diversas realizaciones no limitativas, el método comprende administrar una cantidad eficaz de una célula transducida dada a conocer en el presente documento o una población de las células transducidas dadas a conocer en el presente documento (por ejemplo, HSC, células madre embrionarias, o iPSC) al sujeto.

Para el tratamiento, la cantidad administrada es una cantidad eficaz en la producción del efecto deseado. Una cantidad eficaz puede proporcionarse en una o en una serie de administraciones. Una cantidad eficaz puede proporcionarse en un bolo o mediante perfusión continua. Una “cantidad eficaz” (o “cantidad terapéuticamente eficaz”) es una cantidad suficiente para producir un resultado clínico beneficioso o deseado tras el tratamiento. Una

cantidad eficaz puede administrarse a un sujeto en una o más dosis. En términos de tratamiento, una cantidad eficaz es una cantidad que es suficiente para aliviar, mejorar, estabilizar, revertir o ralentizar la progresión de la enfermedad, o reducir de otro modo las consecuencias patológicas de la enfermedad. La cantidad eficaz se determina generalmente por el médico según cada caso y está dentro de las habilidades de un experto en la técnica.

5 Normalmente se tienen en cuenta varios factores cuando se determina una dosificación apropiada para lograr una cantidad eficaz. Estos factores incluyen edad, sexo y peso del sujeto, el estado que está tratándose, la gravedad del estado y la forma y concentración eficaz de las células inmunosensibles administradas.

En un ejemplo no limitativo, tras la administración de una o más de las células transducidas dadas a conocer en el presente documento, se extrae sangre periférica del sujeto y se miden los niveles de hemoglobina. Un nivel terapéuticamente relevante de hemoglobina se produce tras la administración de una o más de las células transducidas dadas a conocer en el presente documento. Un nivel terapéuticamente relevante de hemoglobina es un nivel de hemoglobina que es suficiente (1) para mejorar o corregir la anemia, (2) para restaurar la capacidad del sujeto para producir glóbulos rojos que contienen hemoglobina normal, (3) para corregir la eritropoyesis ineficaz en el sujeto, (4) para corregir la hematopoyesis extramedular (por ejemplo, hematopoyesis extramedular esplénica y hepática), y/o (5) para reducir la acumulación de hierro, por ejemplo, en tejidos y órganos periféricos. El nivel terapéuticamente relevante de hemoglobina puede ser de al menos aproximadamente 7 g/dl de Hb, al menos aproximadamente 7,5 g/dl de Hb, al menos aproximadamente 8 g/dl de Hb, al menos aproximadamente 8,5 g/dl de Hb, al menos aproximadamente 9 g/dl de Hb, al menos aproximadamente 9,5 g/dl de Hb, al menos aproximadamente 10 g/dl de Hb, al menos aproximadamente 10,5 g/dl de Hb, al menos aproximadamente 11 g/dl de Hb, al menos aproximadamente 11,5 g/dl de Hb, al menos aproximadamente 12 g/dl de Hb, al menos aproximadamente 12,5 g/dl de Hb, al menos aproximadamente 13 g/dl de Hb, al menos aproximadamente 13,5 g/dl de Hb, al menos aproximadamente 14 g/dl de Hb, al menos aproximadamente 14,5 g/dl de Hb, o al menos aproximadamente 15 g/dl de Hb. Adicional o alternativamente, el nivel terapéuticamente relevante de hemoglobina puede ser de desde aproximadamente 7 g/dl de Hb hasta aproximadamente 7,5 g/dl de Hb, desde aproximadamente 7,5 g/dl de Hb hasta aproximadamente 8 g/dl de Hb, desde aproximadamente 8 g/dl de Hb hasta aproximadamente 8,5 g/dl de Hb, desde aproximadamente 8,5 g/dl de Hb hasta aproximadamente 9 g/dl de Hb, desde aproximadamente 9 g/dl de Hb hasta aproximadamente 9,5 g/dl de Hb, desde aproximadamente 9,5 g/dl de Hb hasta aproximadamente 10 g/dl de Hb, desde aproximadamente 10 g/dl de Hb hasta aproximadamente 10,5 g/dl de Hb, desde aproximadamente 10,5 g/dl de Hb hasta aproximadamente 11 g/dl de Hb, desde aproximadamente 11 g/dl de Hb hasta aproximadamente 11,5 g/dl de Hb, desde aproximadamente 11,5 g/dl de Hb hasta aproximadamente 12 g/dl de Hb, desde aproximadamente 12 g/dl de Hb hasta aproximadamente 12,5 g/dl de Hb, desde aproximadamente 12,5 g/dl de Hb hasta aproximadamente 13 g/dl de Hb, desde aproximadamente 13 g/dl de Hb hasta aproximadamente 13,5 g/dl de Hb, desde aproximadamente 13,5 g/dl de Hb hasta aproximadamente 14 g/dl de Hb, desde aproximadamente 14 g/dl de Hb hasta aproximadamente 14,5 g/dl de Hb, desde aproximadamente 14,5 g/dl de Hb hasta aproximadamente 15 g/dl de Hb, desde aproximadamente 7 g/dl de Hb hasta aproximadamente 8 g/dl de Hb, desde aproximadamente 8 g/dl de Hb hasta aproximadamente 9 g/dl de Hb, desde aproximadamente 9 g/dl de Hb hasta aproximadamente 10 g/dl de Hb, desde aproximadamente 10 g/dl de Hb hasta aproximadamente 11 g/dl de Hb, desde aproximadamente 11 g/dl de Hb hasta aproximadamente 12 g/dl de Hb, desde aproximadamente 12 g/dl de Hb hasta aproximadamente 13 g/dl de Hb, desde aproximadamente 13 g/dl de Hb hasta aproximadamente 14 g/dl de Hb, desde aproximadamente 14 g/dl de Hb hasta aproximadamente 14,5 g/dl de Hb, desde aproximadamente 14,5 g/dl de Hb hasta aproximadamente 15 g/dl de Hb, desde aproximadamente 7 g/dl de Hb hasta aproximadamente 8 g/dl de Hb, desde aproximadamente 8 g/dl de Hb hasta aproximadamente 9 g/dl de Hb, desde aproximadamente 9 g/dl de Hb hasta aproximadamente 10 g/dl de Hb, desde aproximadamente 10 g/dl de Hb hasta aproximadamente 11 g/dl de Hb, desde aproximadamente 11 g/dl de Hb hasta aproximadamente 12 g/dl de Hb, desde aproximadamente 12 g/dl de Hb hasta aproximadamente 13 g/dl de Hb, desde aproximadamente 13 g/dl de Hb hasta aproximadamente 14 g/dl de Hb, desde aproximadamente 14 g/dl de Hb hasta aproximadamente 15 g/dl de Hb. En determinadas realizaciones, el nivel terapéuticamente relevante de hemoglobina se mantiene en el sujeto durante al menos aproximadamente 6 meses, durante al menos aproximadamente 12 meses (o 1 año), durante al menos aproximadamente 24 meses (o 2 años). En determinadas realizaciones, el nivel terapéuticamente relevante de hemoglobina se mantiene en el sujeto durante hasta aproximadamente 6 meses, durante hasta aproximadamente 12 meses (o 1 año), durante hasta aproximadamente 24 meses (o 2 años). En determinadas realizaciones, el nivel terapéuticamente relevante de hemoglobina se mantiene en el sujeto durante aproximadamente 6 meses, durante aproximadamente 12 meses (o 1 año), durante aproximadamente 24 meses (o 2 años). En determinadas realizaciones, el nivel terapéuticamente relevante de hemoglobina se mantiene en el sujeto durante desde aproximadamente 6 meses hasta aproximadamente 12 meses (por ejemplo, desde aproximadamente 6 meses hasta aproximadamente 8 meses, desde aproximadamente 8 meses hasta aproximadamente 10 meses, desde aproximadamente 10 meses hasta aproximadamente 12 meses), desde aproximadamente 12 meses hasta aproximadamente 18 meses (por ejemplo, desde aproximadamente 12 meses hasta aproximadamente 14 meses, desde aproximadamente 14 meses hasta aproximadamente 16 meses, o desde aproximadamente 16 meses hasta aproximadamente 18 meses), o desde aproximadamente 18 meses hasta aproximadamente 24 meses (por ejemplo, desde aproximadamente 18 meses hasta aproximadamente 20 meses, desde aproximadamente 20 meses hasta aproximadamente 22 meses, o desde aproximadamente 22 meses hasta aproximadamente 24 meses).

60 En determinadas realizaciones, el método comprende administrar una o más células transducidas con un vector recombinante que comprende un casete de expresión dado a conocer en el presente documento tal como se describió anteriormente. El número de copias de vector del vector recombinante en las células que proporcionan el nivel terapéuticamente relevante de hemoglobina (por ejemplo, 9-10 g/dl) en el sujeto es de desde aproximadamente 0,5 hasta aproximadamente 2, desde aproximadamente 0,5 hasta aproximadamente 1, o desde aproximadamente 1

hasta aproximadamente 2 números de copias de vector por célula. En determinadas realizaciones, el número de copias de vector del vector dado a conocer en el presente documento es de aproximadamente 0,5, aproximadamente 0,6, aproximadamente 0,7, aproximadamente 0,8, aproximadamente 0,9, aproximadamente 1,0, aproximadamente 1,1, aproximadamente 1,2, aproximadamente 1,3, aproximadamente 1,4, aproximadamente 1,5, aproximadamente 1,6, aproximadamente 1,7, aproximadamente 1,8, aproximadamente 1,9, o aproximadamente 2,0 número de copias de vector por célula.

En determinadas realizaciones, el sujeto carece de donante coincidente de antígeno leucocitario humano (HLA). En determinadas realizaciones, la célula transducida procede del mismo sujeto. En una realización, la célula transducida procede de la médula ósea del mismo sujeto. Por tanto, la administración de las células transducidas no produce el riesgo de enfermedad de injerto contra huésped en el sujeto. El método no requiere inmunosupresión para prevenir el rechazo de injerto, por ejemplo, el método no comprende administrar un agente inmunosupresor al sujeto.

El contenido dado a conocer en el presente documento también proporciona un método de aumento de la proporción de glóbulos rojos o eritrocitos en comparación con glóbulos blancos o leucocitos en un sujeto. En diversas realizaciones no limitativas, el método comprende administrar una cantidad eficaz de una célula transducida dada a conocer en el presente documento o una población de las células transducidas dadas a conocer en el presente documento (por ejemplo, HSC, células madre embrionarias o iPSC) al sujeto, en el que la proporción de células de progenie de glóbulos rojos de las células madre hematopoyéticas se aumenta en comparación con las células de progenie de glóbulos blancos de las células madre hematopoyéticas en el sujeto.

Sin desear limitarse a ninguna teoría particular, una ventaja importante proporcionada por el casete de expresión, vectores y otros sistemas de administración (por ejemplo, nucleasas y sistemas de CRISPR-Cas), composiciones, y métodos del contenido dado a conocer en el presente documento es la alta eficacia de la terapia génica de globina que puede lograrse administrando poblaciones de células que comprenden porcentajes inferiores de células transducidas en comparación con los métodos existentes. Esto proporciona ventajas de seguridad importantes asociadas con probabilidades reducidas de mutación, transformación o activación de oncogenes perjudiciales de genes celulares en células transducidas. Las células transducidas pueden administrarse como parte de un trasplante de médula ósea o de sangre de cordón umbilical en un individuo que se ha sometido o no a terapia de ablación de médula ósea.

Una consideración con respecto al uso terapéutico de las células transducidas dadas a conocer en el presente documento con el casete de expresión descrito en el presente documento ("células transducidas") es la cantidad de células necesarias para lograr un efecto óptimo. La cantidad de células transducidas que van a administrarse variará para el sujeto que esté tratándose. En una realización, se administran desde aproximadamente  $1 \times 10^4$  hasta aproximadamente  $1 \times 10^5$  células/kg, desde aproximadamente  $1 \times 10^5$  hasta aproximadamente  $1 \times 10^6$  células/kg, desde aproximadamente  $1 \times 10^6$  hasta aproximadamente  $1 \times 10^7$  células/kg, desde aproximadamente  $1 \times 10^7$  hasta aproximadamente  $1 \times 10^8$  células/kg, desde aproximadamente  $1 \times 10^8$  hasta aproximadamente  $1 \times 10^9$  células/kg, o desde aproximadamente  $1 \times 10^9$  hasta aproximadamente  $1 \times 10^{10}$  células/kg de las células transducidas dadas a conocer en el presente documento a un sujeto. Pueden administrarse células más eficaces en números incluso menores. En algunas realizaciones, se administran al menos aproximadamente  $1 \times 10^8$  células/kg, al menos aproximadamente  $2 \times 10^8$  células/kg, al menos aproximadamente  $3 \times 10^8$  células/kg, al menos aproximadamente  $4 \times 10^8$  células/kg, o al menos aproximadamente  $5 \times 10^8$  células/kg de las células transducidas dadas a conocer en el presente documento a un sujeto. La determinación precisa de lo que se considerará una dosis eficaz puede basarse en factores individuales para cada sujeto, incluyendo su tamaño, edad, sexo, peso y estado del sujeto particular. Los expertos en la técnica pueden determinar fácilmente las dosificaciones a partir de esta divulgación y del conocimiento en la técnica.

En diversas realizaciones, los casetes de expresión, vectores y otros sistemas de administración (nucleasas y sistemas de CRISPR-Cas), composiciones y métodos del contenido dado a conocer en el presente documento ofrecen métodos de terapia génica mejorados usando terapia génica *ex vivo* y trasplante autólogo. El trasplante de células transducidas con el casete de expresión o al interior de sujetos que tienen una hemoglobinopatía da como resultado la corrección a largo plazo de la enfermedad.

Pueden administrarse una o más células transducidas dadas a conocer en el presente documento mediante cualquier método conocido en la técnica, incluyendo, pero sin limitarse a, administración parenteral (por ejemplo, administración intramuscular, administración intravenosa, administración subcutánea o administración intraperitoneal), administración espinal, y administración epidérmica. En una realización no limitativa, se administran una o más células transducidas a un sujeto por vía intravenosa. Pueden administrarse una o más células transducidas dadas a conocer en el presente documento mediante inyección, infusión o implante. En una realización no limitativa, se administran una o más células transducidas mediante inyección. En otra realización no limitativa, se administran una o más células transducidas mediante inyección intravenosa.

Los sujetos pueden tener una forma avanzada de enfermedad, en cuyo caso el objetivo de tratamiento puede incluir mitigación o inversión de la progresión de la enfermedad, y/o mejora de los efectos secundarios. Los sujetos pueden tener una historia del estado, para el que ya se han tratado, en cuyo caso el objetivo terapéutico incluirá

normalmente una disminución o retardo del riesgo de recidiva.

#### VII. Kits

El contenido dado a conocer en el presente documento proporciona kits para el tratamiento o la prevención de una hemoglobinopatía. En una realización, el kit comprende una composición terapéutica o profiláctica que contiene una cantidad eficaz de una célula transducida con el casete de expresión dado a conocer en el presente documento en forma de dosificación unitaria. En una realización no limitativa, el kit comprende uno o más casetes de expresión dados a conocer en el presente documento. En determinadas realizaciones, el kit comprende uno o más vectores que comprenden un casete de expresión dado a conocer en el presente documento. En algunas realizaciones, el kit comprende un recipiente estéril, que puede ser una caja, una ampolla, una botella, un vial, un tubo, una bolsa, un saco, un envase de tipo blíster, u otras formas de recipiente adecuadas conocidas en la técnica. Tales recipientes pueden fabricarse de plástico, vidrio, papel laminado, lámina de metal u otros materiales adecuados para contener medicamentos.

Si se desea, la célula transducida se proporciona junto con instrucciones para administrar la célula a un sujeto que tiene o corre el riesgo de desarrollar una hemoglobinopatía. Las instrucciones incluirán generalmente información sobre el uso de la composición para el tratamiento o la prevención de una hemoglobinopatía. En otras realizaciones, las instrucciones incluyen al menos uno de los siguientes: descripción del agente terapéutico; calendario de dosificación y administración para el tratamiento o la prevención de una hemoglobinopatía o síntomas de la misma; precauciones; advertencias; indicaciones; contraindicaciones; información de sobredosis; reacciones adversas; farmacología en animales; estudios clínicos; y/o bibliografía. Alternativa o adicionalmente, el kit puede incluir instrucciones para transducir una célula con el uno o más casetes de expresión y/o vectores que comprenden tales casetes de expresión. Las instrucciones pueden imprimirse directamente en el recipiente (cuando está presente), o como etiqueta aplicada al recipiente, o como hoja, folleto, tarjeta o carpeta independiente suministrada en o con el recipiente.

#### Ejemplos

La práctica del contenido dado a conocer en el presente documento emplea, a menos que se indique lo contrario, técnicas convencionales de biología molecular (incluyendo técnicas recombinantes), microbiología, biología celular, bioquímica e inmunología, que están dentro del alcance del experto en la técnica. Tales técnicas se explican completamente en la bibliografía, tal como, "Molecular Cloning: A Laboratory Manual", segunda edición (Sambrook, 1989); "Oligonucleotide Synthesis" (Gait, 1984); "Animal Cell Culture" (Freshney, 1987); "Methods in Enzymology" "Handbook of Experimental Immunology" (Weir, 1996); "Gene Transfer Vectors for Mammalian Cells" (Miller y Calos, 1987); "Current Protocols in Molecular Biology" (Ausubel, 1987); "PCR: The Polymerase Chain Reaction", (Mullis, 1994); "Current Protocols in Immunology" (Coligan, 1991). Estas técnicas son aplicables a la producción de los polinucleótidos y polipéptidos del contenido dado a conocer en el presente documento, y, como tales, pueden tenerse en cuenta en la realización y práctica del contenido dado a conocer en el presente documento. En las siguientes secciones se comentarán técnicas particularmente útiles para realizaciones particulares.

Los siguientes ejemplos se exponen para proporcionar a los expertos habituales en la técnica una divulgación y descripción completas de cómo realizar y usar los casetes de expresión, vectores, sistemas de administración y métodos terapéuticos del contenido dado a conocer en el presente documento, y no se pretende que limiten el alcance de lo que los inventores consideran como su invención.

#### 40 *Ejemplo 1: Descubrimiento de aislantes novedosos*

Los problemas creados por la mutagénesis por inserción de vectores virales se conocen ampliamente (Nienhuis (2013), Baum *et al.* (2006), Nienhuis *et al.* (2006)) al igual que la evidencia de que los riesgos de genotoxicidad pueden reducirse mediante el uso de aislantes de cromatina (Arumugam *et al.* (2007), Emery (2011), Evans-Galea *et al.* (2007), Rivella *et al.* (2000), Emery *et al.* (2000), Emery *et al.* (2002), Yannaki *et al.* (2002), Hino *et al.* (2004), Ramezani *et al.* (2003), Ramezani *et al.* (2008)). Se han desarrollado enfoques que permiten la identificación eficaz de aislantes de bloqueo de potenciador en el genoma humano. Estos nuevos aislantes son cortos, en promedio de 150 pb, y no afectan de manera adversa a los títulos de vectores virales y son varias veces más potentes que el aislante cHS4. Se usan enfoques genómicos para descubrir los aislantes de barrera y de bloqueo de potenciador más potentes del genoma humano. Para terapia génica de las hemoglobinopatías, se requieren potenciadores potentes para lograr niveles terapéuticos de expresión de gen de globina. Por tanto, los aislantes potentes pueden proporcionar unos medios para proteger el entorno genómico de los potenciadores potentes de los vectores de integración.

Varios estudios han demostrado la capacidad del aislante cHS4 para reducir el silenciamiento por efecto de posición de vectores gamma-retrovirales (Evans-Galea *et al.* (2007), Rivella *et al.* (2000), Emery *et al.* (2000), Emery *et al.* (2002), Yannaki *et al.* (2002), Hino *et al.* (2004), Ramezani *et al.* (2006), Yao *et al.* (2003), Nishino *et al.* (2006), Aker *et al.* (2007), Li y Emery (2008)), y vectores lentivirales (Evans-Galea *et al.* (2007), Ramezani *et al.* (2003), Puthenveetil *et al.* (2004), Arumugam *et al.* (2007), Bank *et al.* (2005), Aker *et al.* (2007), Ma *et al.* (2003), Chang *et al.* (2005), Pluta *et al.* (2005)). Los estudios que estaban diseñados de manera apropiada demostraron que la

inclusión de la versión de 1,2 kb del aislante cHS4 aumentó la probabilidad y/o sistematicidad de expresión de transgén de vector en al menos algunos entornos (Arumugam *et al.* (2007), Evans-Galea *et al.* (2007), Emery *et al.* (2002), Yannaki *et al.* (2002), Hino *et al.* (2004), Ramezani *et al.* (2006), Aker *et al.* (2007), Li y Emery (2008), Pluta *et al.* (2005), Jakobsson *et al.* (2004)). No obstante, el grado de protección proporcionado por el aislante cHS4 está lejos de ser completo. Además, la inclusión del cHS4 de 1,2 Kb puede afectar de manera adversa a los títulos de vector mientras que se ha demostrado que el núcleo de cHS4 más pequeño es ineficaz (Aker *et al.* (2007), Jakobsson *et al.* (2004)).

Se sometieron a prueba los efectos sobre la genotoxicidad usando un ensayo *in vivo* basado en la cuantificación de la formación de tumor en ratones. Los vectores aislados mediante aislante A1 disminuyeron la formación de tumor inducida por integración de vector aleatoria en quimeras hematopoyéticas en comparación con ratones que recibieron controles sin aislar o aislados con cHS4.

Para evaluar los efectos sobre los títulos de vector, se introdujo aislante A1 en la región de copia doble de un vector lentiviral de tercera generación que expresaba GFP a partir de un promotor de paquete constitutivo, y se midieron los títulos virales y la expresión de GFP. El aislante A1 no afectó de manera adversa a la expresión de GFP de vector.

En el ensayo de genotoxicidad *in vivo*, una línea celular transducida con vectores gamma-retrovirales produjo tumores tras el trasplante en ratones y permitió la cuantificación de efectos genotóxicos midiendo las tasas de supervivencia libre de tumor. Se cuantificaron los efectos de un aislante sobre la genotoxicidad mediante el número de tumores formados en los ratones y las tasas de supervivencia libre de tumor. Se insertó aislante A1 en la parte proximal de la LTR en 3', a partir de la cual se copia en la LTR en 5' durante la transcripción inversa e integración de vector. La topología resultante pone copias del aislante entre las regiones genómicas ubicadas en 5' y 3' del provirus integrado y la actividad potenciadora de la LTR viral en 5' y el promotor Pgk interno, pero no contiene el potenciador en la LTR en 3'. Esto puede disminuir la genotoxicidad dando por tanto como resultado una disminución de la formación de tumor y un aumento de la supervivencia de los animales. Se usaron vectores indicadores gamma-retrovirales flanqueados con aislante A1 o regiones de control para transducir la línea celular dependiente de factor de crecimiento 32D, y se trasplantaron 10 subcombinaciones independientes para cada vector al interior de ratones C3H/HeJ singénicos. Los 10 ratones en los que se trasplantaron células transducidas simuladas permanecieron libres de tumores derivados de células 32D, mientras que casi todos los ratones en los que se trasplantaron células 32D transducidas con vectores que no contenían insertos o contenían un espaciador neutro de 790 pb desarrollaron tumores dentro de una mediana de 16 semanas (figura 5B). Flanquear este vector con el aislante cHS4 retrasó la aparición de la formación de tumor en varias semanas, y redujo la frecuencia de animales que desarrollaron tumores hasta 6 de 10. En cambio, sólo dos de 10 animales desarrollaron tumores tras el trasplante con células 32D transducidas con el vector flanqueado con aislante A1 (figura 5B). La frecuencia de animales con tumores y el número de acontecimientos de transducción de vector en las subcombinaciones originales sugirieron que flanquear el vector con aislante A1 redujo la tasa global de formación de tumor 12 veces, desde 46,9 tumores por  $10^5$  provirus hasta 3,9 tumores por  $10^5$  provirus (figura 5C). En comparación, el aislante cHS4 redujo la tasa global de formación de tumor 2,8 veces (hasta 16,9 tumores por  $10^5$  provirus), mientras que el espaciador neutro no tuvo ningún efecto estadísticamente apreciable sobre la tasa de formación de tumor. Estos resultados indican que los aislantes de bloqueo de potenciador descubiertos pueden reducir sustancialmente los riesgos de mutagénesis por inserción y genotoxicidad.

#### *Ejemplo 2: Caracterización de vectores de globina que comprenden al menos un aislante.*

Se generó un casete de expresión dado a conocer en el presente documento (denominado "casete de expresión 1"; tal como se muestra en la figura 1), que comprende aislante A1, y un gen de  $\beta^A$ -globina humana que codifica para una mutación de treonina a glutamina en el codón 87 ( $\beta^{A-T87Q}$ ) operativamente unido a una región LCR de  $\beta$ -globina que comprende una región de HS2 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 9, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7. El fundamento para usar la cadena  $\beta$  variante ( $\beta^A$ ) es facilitar la detección del gen de  $\beta$ -globina codificado por vector, distinguiéndolo de cadenas beta endógenas o transfundidas. El residuo de glutamina (GLN) en la posición 87 en la cadena de  $\gamma$ -globina aumenta la actividad anti-falciformación de la cadena gamma con respecto a la cadena  $\beta$ , mientras que conserva las características de unión a oxígeno de adulto de la cadena  $\beta$  (Nagel *et al.* (1979)). En el vector 1, una mutación puntual que altera el codón 87 ( $\beta^{A-T87Q}$ , o  $\beta 87$ ) sustituye la treonina normal por glutamina y aumenta la actividad anti-falciformación de la cadena  $\beta$  codificada por vector. Esta cadena  $\beta 87$  se ha usado de manera segura en un paciente con HbE-talasemia (Cavazzana-Calvo *et al.* (2010)).

Se incorporó el casete de expresión 1 o se introdujo en un vector de lentivirus (denominado "vector 1"). Se introdujo el vector 1 en células de médula ósea de ratones C57BL/6-Hbb th3/+ y se trasplantaron a pacientes singénicos sometidos a radiación letal tal como se describió anteriormente (May *et al.* (2000), May *et al.* (2002), Lisowski *et al.* (2007)). El título de vector de V1 era comparable al de un vector de lentivirus que comprende un casete de expresión que carece de aislante A1. Se comparó la expresión de  $\beta$ -globina del vector 1 con la de un vector de lentivirus (denominado "vector 2") que comprende un casete de expresión que carece de un aislante y comprende un gen de

β-globina humana silvestre operativamente unido a una región LCR de β-globina que comprende una región de HS2 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 9, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6. En comparación con el vector 2, la expresión de β-globina del vector 1 normalizada con respecto a las copias de vector era equivalente o estaba ligeramente aumentada, sugiriendo un beneficio añadido para la expresión *in vivo* proporcionado por los elementos de barrera flanqueantes, tal como se muestra en la figura 6.

*Ejemplo 3: Evaluación de la actividad potenciadora en células K562 no eritroides*

Se evaluó la actividad potenciadora de HS2 en células K562 no eritroides. Tal como se muestra en la figura 7, la expresión de GFP en células K562 transducidas con vectores impulsada por un promotor mínimo no unido a ningún potenciador ("vacío", HS2, HS3-4, HS2-3-4 o el potenciador *runx1* usado como control positivo ("RUNX1"). La expresión de fondo era del orden o el 0,01% ("vacío"), pero aumentó más de 10 veces con HS2-3-4 ("Lcr9", 0,17%). Esta potenciación se debía principalmente a HS2 (0,15%) pero no a HS3-4 (0,05%). Todas las líneas celulares se transdujeron de manera comparable (número de copias de vector medio de 2,5). Los resultados respaldan que HS2, pero no HS3-HS4, puede suponer un riesgo oncogénico en células progenitoras y madre hematopoyéticas no eritroides.

*Ejemplo 4: Potenciadores específicos de eritroides novedosos*

Tal como se muestra en las figuras 8 y 9, cinco potenciadores específicos de eritroides sustituyeron a HS2: intrón 1 de ALAS, intrón 8 de ALAS, BLVRB, PPOX y espectrina alfa. Los inventores han mostrado que la totalidad de estos cinco potenciadores son potenciadores potentes, y carecen de actividad potenciadora en tejidos no eritroides, y no reducen el título de vector.

*Ejemplo 5: Aumento de la producción de vector lentiviral de globina mediante modificaciones de LTR en 3'*

Una característica esencial de los vectores de globina terapéuticos es lograr un título alto, suficiente para la transducción eficaz de células de pacientes. Debido a su gran carga útil, que comprende un gen, promotor, potenciadores y/o elementos de LCR, los vectores lentivirales de globina tienen de manera inherente un título bajo, lo cual complica su fabricación y limita su uso clínico. Este problema se complica adicionalmente por la incorporación de elementos genómicos adicionales tales como un aislante, que aumenta adicionalmente el tamaño del vector.

Los inventores exploraron diferentes modificaciones de la repetición terminal larga (LTR) en 3' de vectores de globina para aumentar el título de vectores de globina. Se evaluaron más de 62 variaciones, numeradas de 1 a 62, modeladas en un vector de lentivirus que comprende un gen de β-globina humana operativamente unido a una región LCR de β-globina que comprende una región de HS2 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 9, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7. Dicho de otro modo, todos los vectores n.º 1 a vector 62 comprenden una región LCR de β-globina que comprende una región de HS2 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 9, una región de HS3 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y una región de HS4 que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 7. El vector n.º 18 sirvió como referencia, que comprende una delección U3 convencional en la LTR en 3'. El vector n.º 1 (no representado) comprendía una LTR completa, es decir, silvestre, que no podía usarse de manera clínica. En las figuras 10A y 10B se representan modificaciones de la LTR en 3', y sus títulos se muestran en las figuras 11 y 12 (el eje Y muestra el número de copias de vector de disoluciones madre de vector producidas y sometidas a prueba en condiciones estrictamente idénticas). Se midieron las titulaciones en repeticiones triples, realizadas en paralelo por dos operarios, y se repitieron en múltiples experimentos.

Tal como se muestra en las figuras 11 y 12, el vector n.º 55 mostró de manera repetida un título superior. Este vector comprende un elemento post-regulador de la hepatitis de la marmota (WPRE) y una señal de poliadenilación de hormona de crecimiento bovina en 3' con respecto a la región R en la LTR en 3'. Por tanto, el elemento WPRE no se transfiere a las células transducidas.

La incorporación de estos elementos para potenciar la producción de vectores lentivirales de globina es esencial para proporcionar títulos superiores y por tanto para la utilidad clínica de los vectores descritos en esta solicitud.

**Bibliografía**

1. Weatherall, D.J. y Clegg, J.B. The Thalassemia Syndrome. Blackwell Scientific Oxford (1981).
2. Stamatoyannopoulos, G., Nienhuis, A.W., Majerus, P. y Varmus, H. The Molecular Basis of Blood Diseases. WB Saunders, Filadelfia (1994).
3. Weatherall, D.J. Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. Nat Rev Genet 2, 245-255. (2001).
4. Steinberg, M.H., Forget, B.G., Higgs, D.R. y Nagel, R.L. Molecular Mechanism of β Thalassemia; Bernard G.

- Forget, (Cambridge University Press, Cambridge, R.U., 2001).
5. Cooley, T.B. y Lee, P. A series of cases of splenomegaly in children with anemia and peculiar bone changes. *Trans. Am. Pediatr. Soc.* 37, 29 (1925).
- 5 6. Giardina, P.J. y Grady, R.W. Chelation therapy in beta-thalassemia: an optimistic update. *Semin Hematol* 38, 360-366. (2001).
7. Giardini, C. y Lucarelli, G. Bone marrow transplantation in the treatment of thalassemia. *Current opinion in hematology* 1, 170-176. (1994).
8. Boulad, F., Giardina, P., Gillio, A., Keman, N., Small, T., Brochstein, J., Van Syckle, K., George, D., Szabolcs, P. y O'Reilly, R.J. Bone marrow transplantation for homozygous beta-thalassemia. The Memorial Sloan-Kettering Cancer Center experience. *Ann N Y Acad Sci* 850, 498-502. (1998).
- 10 9. Lucarelli, G., Clift, R.A., Galimberti, M., Angelucci, E., Giardini, C., Baronciani, D., Polchi, P., Andreani, M., Gaziev, D., Erer, B., Ciaroni, A., D'Adamo, F., Albertini, F. y Muretto, P. Bone marrow transplantation in adult thalassemic patients. *Blood* 93, 1164-1167. (1999).
- 15 10. Tisdale, J. y Sadelain, M. Toward gene therapy for disorders of globin synthesis. *Semin Hematol* 38, 382-392 (2001).
11. Pauling, L., Itano, H.A., Singer, S.J. y Wells, I.C. Sickle cell anemia, a molecular disease. *Science* 110, 543-546 (1949).
12. Swank, R.A. y Stamatoyannopoulos, G. Fetal gene reactivation. *Curr Opin Genet Dev* 8, 366-370 (1998).
- 20 13. Platt, O.S., Orkin, S.H., Dover, G., Beardsley, G.P., Miller, B. y Nathan, D.G. Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. *J Clin Invest* 74, 652-656. (1984).
14. Charache, S., Dover, G.J., Moore, R.D., Eckert, S., Ballas, S.K., Koshy, M., Milner, P.F., Orringer, E.P., Phillips, G., Jr., Platt, O.S. y *et al.* Hydroxyurea: effects on hemoglobin F production in patients with sickle cell anemia. *Blood* 79, 2555-2565. (1992).
- 25 15. Atweh, G.F. y Loukopoulos, D. Pharmacological induction of fetal hemoglobin in sickle cell disease and betathalassemia. *Semin Hematol* 38, 367-373. (2001).
16. Perrine, S.P., Castaneda, S.A., Boosalis, M.S., White, G.L., Jones, B.M. y Bohacek, R. Induction of fetal globin in beta-thalassemia: Cellular obstacles and molecular progress. *Ann N Y Acad Sci* 1054, 257-265 (2005).
17. Stamatoyannopoulos, G. Prospects for developing a molecular cure for thalassemia. *Hematology* 10 Sup. 1, 255-257 (2005).
- 30 18. Vermylen, C., Cornu, G., Ferster, A., Brichard, B., Ninane, J., Ferrant, A., Zenebergh, A., Maes, P., Dhooge, C., Benoit, Y., Beguin, Y., Dresse, M.F. y Sariban, E. Haematopoietic stem cell transplantation for sickle cell anaemia: the first 50 patients transplanted in Belgium. *Bone Marrow Transplant* 22, 1-6 (1998).
19. Luzzatto, L. y Goodfellow, P. Sickle cell anaemia. A simple disease with no cure. *Nature* 337, 17-18 (1989).
- 35 20. Sadelain, M. Genetic treatment of the haemoglobinopathies: recombinations and new combinations. *Br J Haematol* 98, 247-253 (1997).
21. Sadelain, M., Boulad, F., Galanello, R., Giardina, P., Locatelli, F., Maggio, A., Rivella, S., Riviere, I. y Tisdale, J. Therapeutic options for patients with severe beta-thalassemia: the need for globin gene therapy. *Hum Gene Ther* 18, 1-9 (2007).
- 40 22. Borgna-Pignatti, C., Rugolotto, S., De Stefano, P., Zhao, H., Cappellini, M.D., Del Vecchio, G.C., Romeo, M.A., Forni, G.L., Gamberini, M.R., Ghilardi, R., Piga, A. y Cnaan, A. Survival and complications in patients with thalassemia major treated with transfusion and deferoxamine. *Haematologica* 89, 1187-1193 (2004).
23. Telfer, P.T., Warburton, F., Christou, S., Hadjigavriel, M., Sitarou, M., Kolnagou, A. y Angastiniotis, M. Improved survival in thalassemia major patients on switching from desferrioxamine to combined chelation therapy with desferrioxamine and deferiprone. *Haematologica* 94, 1777-1778 (2009).
- 45 24. Ladis, V., Chouliaras, G., Berdoukas, V., Chatziliami, A., Fragodimitri, C., Karabatsos, F., Youssef, J., Kattamis, A. y Karagiorga-Lagana, M. Survival in a large cohort of Greek patients with transfusion-dependent beta thalassaemia and mortality ratios compared to the general population. *European journal of haematology* 86, 332-338 (2011).
25. Mancuso, A., Sciarrino, E., Renda, M.C. y Maggio, A. A prospective study of hepatocellular carcinoma incidence

- in thalassemia. Hemoglobin 30, 119-124 (2006).
26. Persons, D.A. y Tisdale, J.F. Gene therapy for the hemoglobin disorders. *Semin Hematol* 41, 279-286 (2004).
27. Sadelain, M. Recent advances in globin gene transfer for the treatment of beta-thalassemia and sickle cell anemia. *Current opinion in hematology* 13, 142-148 (2006).
- 5 28. May, C., Rivella, S., Callegari, J., Heller, G., Gaensler, K.M., Luzzatto, L. y Sadelain, M. Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. *Nature* 406, 82-86 (2000).
29. May, C., Rivella, S., Chadburn, A. y Sadelain, M. Successful treatment of murine beta-thalassemia intermedia by transfer of the human beta-globin gene. *Blood* 99, 1902-1908 (2002).
- 10 30. Rivella, S., May, C., Chadburn, A., Riviere, I. y Sadelain, M. A novel murine model of Cooley anemia and its rescue by lentiviral-mediated human beta-globin gene transfer. *Blood* 101, 2932-2939 (2003).
31. Sadelain, M., Boulad, F., Lisowski, L., Moi, P. y Riviere, I. Stem cell engineering for the treatment of severe hemoglobinopathies. *Curr Mol Med* 8, 690-697 (2008).
32. Bank, A., Dorazio, R. y Leboulch, P. A phase I/II clinical trial of beta-globin gene therapy for beta-thalassemia. *Ann N Y Acad Sci* 1054, 308-316 (2005).
- 15 33. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., Cavalleco, R., Gillet-Legrand, B., Caccavelli, L., Sgarra, R., Maouche-Chretien, L., Bernaudin, F., Girot, R., Dorazio, R., Mulder, G.J., Polack, A., Bank, A., Soulier, J., Larghero, J., Kabbara, N., Dalle, B., Gourmel, B., Socie, G., Chretien, S., Cartier, N., Aubourg, P., Fischer, A., Cornetta, K., Galacteros, F., Beuzard, Y., Gluckman, E., Bushman, F., Hacein-Bey-Abina, S. y Leboulch, P. Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature* 467, 318-322 (2010).
- 20 34. Braun, C.J., Boztug, K., Paruzynski, A., Witzel, M., Schwarzer, A., Rothe, M., Modlich, U., Beier, R., Gohring, G., Steinemann, D., Fronza, R., Ball, C.R., Haemmerle, R., Naundorf, S., Kuhlcke, K., Rose, M., Fraser, C., Mathias, L., Ferrari, R., Abboud, M.R., Al-Herz, W., Kondratenko, I., Marodi, L., Glimm, H., Schlegelberger, B., Schambach, A., Albert, M.H., Schmidt, M., von Kalle, C. y Klein, C. Gene therapy for Wiskott-Aldrich syndrome—long-term efficacy and genotoxicity. *Sci Transl Med* 6, 227ra233 (2014).
- 25 35. Chang, A.H. y Sadelain, M. The genetic engineering of hematopoietic stem cells: the rise of lentiviral vectors, the conundrum of the Itr, and the promise of lineage-restricted vectors. *Mol Ther* 15, 445-456 (2007).
- 30 36. Pawliuk, R., Westerman, K.A., Fabry, M.E., Payen, E., Tighe, R., Bouhassira, E.E., Acharya, S.A., Ellis, J., London, I.M., Eaves, C.J., Humphries, R.K., Beuzard, Y., Nagel, R.L. y Leboulch, P. Correction of sickle cell disease in transgenic mouse models by gene therapy. *Science* 294, 2368-2371 (2001).
37. Emery, D.W., Chen, H., Li, Q. y Stamatoyannopoulos, G. Development of a condensed locus control region cassette and testing in retrovirus vectors for A gamma-globin. *Blood Cells Mol Dis* 24, 322-339 (1998).
38. Miccio, A., Cesari, R., Lotti, F., Rossi, C., Sanvito, F., Ponzoni, M., Routledge, S.J., Chow, C.M., Antoniou, M.N. y Ferrari, G. In vivo selection of genetically modified erythroblastic progenitors leads to long-term correction of betathalassemia. *Proc Natl Acad Sci U S A* 105, 10547-10552 (2008).
- 35 39. Sadelain, M., Wang, C.H., Antoniou, M., Grosveld, F. y Mulligan, R.C. Generation of a high-titer retroviral vector capable of expressing high levels of the human beta-globin gene. *Proc Natl Acad Sci U S A* 92, 6728-6732 (1995).
- 40 40. Samakoglu, S., Lisowski, L., Budak-Alpdogan, T., Usachenko, Y., Acuto, S., Di Marzo, R., Maggio, A., Zhu, P., Tisdale, J.F., Riviere, I. y Sadelain, M. A genetic strategy to treat sickle cell anemia by coregulating globin transgene expression and RNA interference. *Nat Biotechnol* 24, 89-94 (2006).
41. Pestina, T.I., Hargrove, P.W., Jay, D., Gray, J.T., Boyd, K.M. y Persons, D.A. Correction of murine sickle cell disease using gamma-globin lentiviral vectors to mediate high-level expression of fetal hemoglobin. *Mol Ther* 17, 245-252 (2009).
- 45 42. Hanawa, H., Yamamoto, M., Zhao, H., Shimada, T. y Persons, D.A. Optimized lentiviral vector design improves titer and transgene expression of vectors containing the chicken beta-globin locus HS4 insulator element. *Mol Ther* 17, 667-674 (2009).
43. Arumugam, P.I., Scholes, J., Perelman, N., Xia, P., Yee, J.K. y Malik, P. Improved human beta-globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (cHS4) insulator element. *Mol Ther* 15, 1863-1871 (2007).
- 50 44. Fraser, P., Pruzina, S., Antoniou, M. y Grosveld, F. Each hypersensitive site of the human beta-globin locus

- control region confers a different developmental pattern of expression on the globin genes. *Genes & development* 7, 106-113 (1993).
- 5 45. Navas, P.A., Peterson, K.R., Li, Q., Skarpidi, E., Rohde, A., Shaw, S.E., Clegg, C.H., Asano, H. y Stamatoyannopoulos, G. Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion. *Molecular and cellular biology* 18, 4188-4196 (1998).
46. Li, Q. y Stamatoyannopoulos, G. Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. *Blood* 84, 1399-1401 (1994).
- 10 47. Li, Q., Zhang, M., Han, H., Rohde, A. y Stamatoyannopoulos, G. Evidence that DNase I hypersensitive site 5 of the human beta-globin locus control region functions as a chromosomal insulator in transgenic mice. *Nucleic Acids Res* 30, 2484-2491 (2002).
48. Puthenveetil, G., Scholes, J., Carbonell, D., Qureshi, N., Xia, P., Zeng, L., Li, S., Yu, Y., Hiti, A.L., Yee, J.K. y Malik, P. Successful correction of the human beta-thalassemia major phenotype using a lentiviral vector. *Blood* 104, 3445-3453 (2004).
- 15 49. Wilber, A., Nienhuis, A.W. y Persons, D.A. Transcriptional regulation of fetal to adult hemoglobin switching: new therapeutic opportunities. *Blood* 117, 3945-3953 (2011).
50. Arumugam, P.I., Higashimoto, T., Urbinati, F., Modlich, U., Nestheide, S., Xia, P., Fox, C., Corsinotti, A., Baum, C. y Malik, P. Genotoxic potential of lineage-specific lentivirus vectors carrying the beta-globin locus control region. *Mol Ther* 17, 1929-1937 (2009).
- 20 51. Chang, K.H., Fang, X., Wang, H., Huang, A., Cao, H., Yang, Y., Bonig, H., Stamatoyannopoulos, J.A. y Papayannopoulou, T. Epigenetic modifications and chromosome conformations of the beta globin locus throughout development. *Stem cell reviews* 9, 397-407 (2013).
52. Papayannopoulou, T., Priestley, G.V., Rohde, A., Peterson, K.R. y Nakamoto, B. Hemopoietic lineage commitment decisions: in vivo evidence from a transgenic mouse model harboring micro LCR-beta-pro-LacZ as a transgene. *Blood* 95, 1274-1282 (2000).
- 25 53. Nienhuis, A.W. Development of gene therapy for blood disorders: an update. *Blood* 122, 1556-1564 (2013).
54. Baum, C., Kustikova, O., Modlich, U., Li, Z. y Fehse, B. Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. *Hum Gene Ther* 17, 253-263 (2006).
55. Nienhuis, A.W., Dunbar, C.E. y Sorrentino, B.P. Genotoxicity of retroviral integration in hematopoietic cells. *Mol Ther* 13, 1031-1049 (2006).
- 30 56. Emery, D.W. The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. *Hum Gene Ther* 22, 761-774 (2011).
57. Evans-Galea, M.V., Wielgosz, M.M., Hanawa, H., Srivastava, D.K. y Nienhuis, A.W. Suppression of clonal dominance in cultured human lymphoid cells by addition of the cHS4 insulator to a lentiviral vector. *Mol Ther* 15, 801-809 (2007).
- 35 58. Rivella, S., Callegari, J.A., May, C., Tan, C.W. y Sadelain, M. The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. *J Virol* 74, 4679-4687 (2000).
59. Emery, D.W., Yannaki, E., Tubb, J. y Stamatoyannopoulos, G. A chromatin insulator protects retrovirus vectors from chromosomal position effects. *Proc Natl Acad Sci U S A* 97, 9150-9155 (2000).
- 40 60. Emery, D.W., Yannaki, E., Tubb, J., Nishino, T., Li, Q. y Stamatoyannopoulos, G. Development of virus vectors for gene therapy of beta chain hemoglobinopathies: flanking with a chromatin insulator reduces gamma-globin gene silencing in vivo. *Blood* 100, 2012-2019 (2002).
61. Yannaki, E., Tubb, J., Aker, M., Stamatoyannopoulos, G. y Emery, D.W. Topological constraints governing the use of the chicken HS4 chromatin insulator in oncoretrovirus vectors. *Mol Ther* 5, 589-598 (2002).
- 45 62. Hino, S., Fan, J., Taguwa, S., Akasaka, K. y Matsuoka, M. Sea urchin insulator protects lentiviral vector from silencing by maintaining active chromatin structure. *Gene Ther* 11, 819-828 (2004).
63. Ramezani, A., Hawley, T.S. y Hawley, R.G. Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator. *Blood* 101, 4717-4724 (2003).
- 50 64. Ramezani, A., Hawley, T.S. y Hawley, R.G. Combinatorial incorporation of enhancer-blocking components of the

- chicken beta-globin 5'HS4 and human T-cell receptor alpha/delta BEAD-1 insulators in self-inactivating retroviral vectors reduces their genotoxic potential. *Stem Cells* 26, 3257-3266 (2008).
65. Yannaki, E., Emery, D.W. y Stamatoyannopoulos, G. Gene therapy for beta-thalassaemia: the continuing challenge. *Expert reviews in molecular medicine* 12, e31 (2010).
- 5 66. Persons, D.A. The challenge of obtaining therapeutic levels of genetically modified hematopoietic stem cells in beta-thalassemia patients. *Ann N Y Acad Sci* 1202, 69-74 (2010).
67. Perumbeti, A. y Malik, P. Therapy for beta-globinopathies: a brief review and determinants for successful and safe correction. *Ann N Y Acad Sci* 1202, 36-44 (2010).
- 10 68. Johnson, K.D., Grass, J.A., Park, C., Im, H., Choi, K. y Bresnick, E.H. Highly restricted localization of RNA polymerase II within a locus control region of a tissue-specific chromatin domain. *Molecular and cellular biology* 23, 6484-6493 (2003).
69. Vieira, K.F., Levings, P.P., Hill, M.A., Crusselle, V.J., Kang, S.H., Engel, J.D. y Bungert, J. Recruitment of transcription complexes to the beta-globin gene locus in vivo and in vitro. *J Biol Chem* 279, 50350-50357 (2004).
- 15 70. Levings, P.P., Zhou, Z., Vieira, K.F., Crusselle-Davis, V.J. y Bungert, J. Recruitment of transcription complexes to the beta-globin locus control region and transcription of hypersensitive site 3 prior to erythroid differentiation of murine embryonic stem cells. *The FEBS journal* 273, 746-755 (2006).
71. Felsenfeld, G. y Groudine, M. Controlling the double helix. *Nature* 421, 448-453 (2003).
72. Felsenfeld, G. Chromatin as an essential part of the transcriptional mechanism. *Nature* 355, 219-224 (1992).
- 20 73. Brownell, J.E. y Allis, C.D. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev* 6, 176-184 (1996).
74. Kingston, R.E. y Narlikar, G.J. ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. *Genes & development* 13, 2339-2352 (1999).
75. Tsukiyama, T. y Wu, C. Chromatin remodeling and transcription. *Curr Opin Genet Dev* 7, 182-191 (1997).
- 25 76. Wolffe, A.P., Wong, J. y Pruss, D. Activators and repressors: making use of chromatin to regulate transcription. *Genes to cells : devoted to molecular & cellular mechanisms* 2, 291-302 (1997).
77. Kadonaga, J.T. Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* 92, 307-313 (1998).
78. Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes & development* 12, 599-606 (1998).
- 30 79. Gross, D.S. y Garrard, W.T. Nuclease hypersensitive sites in chromatin. *Annual review of biochemistry* 57, 159-197 (1988).
80. Elgin, S.C. Anatomy of hypersensitive sites. *Nature* 309, 213-214 (1984).
81. Wu, C. The 5' ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I. *Nature* 286, 854-860 (1980).
- 35 82. Felsenfeld, G., Boyes, J., Chung, J., Clark, D. y Studitsky, V. Chromatin structure and gene expression. *Proc Natl Acad Sci U S A* 93, 9384-9388 (1996).
83. Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A. y Felsenfeld, G. The insulation of genes from external enhancers and silencing chromatin. *Proc Natl Acad Sci U S A* 99 Suppl 4, 16433-16437 (2002).
- 40 84. Elgin, S.C. DNAase I-hypersensitive sites of chromatin. *Cell* 27, 413-415 (1981).
85. McGhee, J.D., Wood, W.I., Dolan, M., Engel, J.D. y Felsenfeld, G. A 200 base pair region at the 5' end of the chicken adult beta-globin gene is accessible to nuclease digestion. *Cell* 27, 45-55 (1981).
86. Lowrey, C.H., Bodine, D.M. y Nienhuis, A.W. Mechanism of DNase I hypersensitive site formation within the human globin locus control region. *Proc Natl Acad Sci U S A* 89, 1143-1147 (1992).
- 45 87. Adams, C.C. y Workman, J.L. Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Molecular and cellular biology* 15, 1405-1421 (1995).

88. McArthur, M., Gerum, S. y Stamatoyannopoulos, G. Quantification of DNaseI-sensitivity by real-time PCR: quantitative analysis of DNaseI-hypersensitivity of the mouse beta-globin LCR. *J Mol Biol* 313, 27-34 (2001).
- 5 89. Dorschner, M.O., Hawrylycz, M., Humbert, R., Wallace, J.C., Shafer, A., Kawamoto, J., Mack, J., Hall, R., Goldy, J., Sabo, P.J., Kohli, A., Li, Q., McArthur, M. y Stamatoyannopoulos, J.A. High-throughput localization of functional elements by quantitative chromatin profiling. *Nat Methods* 1, 219-225 (2004).
90. Sabo, P.J., Kuehn, M.S., Thurman, R., Johnson, B.E., Johnson, E.M., Cao, H., Yu, M., Rosenzweig, E., Goldy, J., Haydock, A., Weaver, M., Shafer, A., Lee, K., Neri, F., Humbert, R., Singer, M.A., Richmond, T.A., Dorschner, M.O., McArthur, M., Hawrylycz, M., Green, R.D., Navas, P.A., Noble, W.S. y Stamatoyannopoulos, J.A. Genomescale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. *Nat Methods* 3, 511-518 (2006).
- 10 91. Sabo, P.J., Hawrylycz, M., Wallace, J.C., Humbert, R., Yu, M., Shafer, A., Kawamoto, J., Hall, R., Mack, J., Dorschner, M.O., McArthur, M. y Stamatoyannopoulos, J.A. Discovery of functional noncoding elements by digital analysis of chromatin structure. *Proc Natl Acad Sci U S A* 101, 16837-16842 (2004).
- 15 92. Sabo, P.J., Humbert, R., Hawrylycz, M., Wallace, J.C., Dorschner, M.O., McArthur, M. y Stamatoyannopoulos, J.A. Genome-wide identification of DNaseI hypersensitive sites using active chromatin sequence libraries. *Proc Natl Acad Sci U S A* 101, 4537-4542 (2004).
- 20 93. Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., Garg, K., John, S., Sandstrom, R., Bates, D., Boatman, L., Canfield, T.K., Diegel, M., Dunn, D., Ebersol, A.K., Frum, T., Giste, E., Johnson, A.K., Johnson, E.M., Kutayavin, T., Lajoie, B., Lee, B.K., Lee, K., London, D., Lotakis, D., Neph, S., Neri, F., Nguyen, E.D., Qu, H., Reynolds, A.P., Roach, V., Safi, A., Sanchez, M.E., Sanyal, A., Shafer, A., Simon, J.M., Song, L., Vong, S., Weaver, M., Yan, Y., Zhang, Z., Zhang, Z., Lenhard, B., Tewari, M., Dorschner, M.O., Hansen, R.S., Navas, P.A., Stamatoyannopoulos, G., Iyer, V.R., Lieb, J.D., Sunyaev, S.R., Akey, J.M., Sabo, P.J., Kaul, R., Furey, T.S., Dekker, J., Crawford, G.E. y Stamatoyannopoulos, J.A. The accessible chromatin landscape of the human genome. *Nature* 489, 75-82 (2012).
- 25 94. Stergachis, A.B., Neph, S., Reynolds, A., Humbert, R., Miller, B., Paige, S.L., Vernot, B., Cheng, J.B., Thurman, R.E., Sandstrom, R., Haugen, E., Heimfeld, S., Murry, C.E., Akey, J.M. y Stamatoyannopoulos, J.A. Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. *Cell* 154, 888-903 (2013).
95. Neph, S., Stergachis, A.B., Reynolds, A., Sandstrom, R., Borenstein, E. y Stamatoyannopoulos, J.A. Circuitry and dynamics of human transcription factor regulatory networks. *Cell* 150, 1274-1286 (2012).
- 30 96. Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutayavin, T., Stehling-Sun, S., Johnson, A.K., Canfield, T.K., Giste, E., Diegel, M., Bates, D., Hansen, R.S., Neph, S., Sabo, P.J., Heimfeld, S., Raubitschek, A., Ziegler, S., Cotsapas, C., Sotoodehnia, N., Glass, I., Sunyaev, S.R., Kaul, R. y Stamatoyannopoulos, J.A. Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190-1195 (2012).
- 35 97. Stergachis, A.B., Haugen, E., Shafer, A., Fu, W., Vernot, B., Reynolds, A., Raubitschek, A., Ziegler, S., LeProust, E.M., Akey, J.M. y Stamatoyannopoulos, J.A. Exonic transcription factor binding directs codon choice and affects protein evolution. *Science* 342, 1367-1372 (2013).
- 40 98. Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., Maurano, M.T., Humbert, R., Rynes, E., Wang, H., Vong, S., Lee, K., Bates, D., Diegel, M., Roach, V., Dunn, D., Neri, J., Schafer, A., Hansen, R.S., Kutayavin, T., Giste, E., Weaver, M., Canfield, T., Sabo, P., Zhang, M., Balasundaram, G., Byron, R., MacCoss, M.J., Akey, J.M., Bender, M.A., Groudine, M., Kaul, R. y Stamatoyannopoulos, J.A. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489, 83-90 (2012).
99. Ramezani, A., Hawley, T.S. y Hawley, R.G. Stable gammaretroviral vector expression during embryonic stem cell-derived in vitro hematopoietic development. *Mol Ther* 14, 245-254 (2006).
- 45 100. Recillas-Targa, F., Pikaart, M.J., Burgess-Beusse, B., Bell, A.C., Litt, M.D., West, A.G., Gaszner, M. y Felsenfeld, G. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A* 99, 6883-6888 (2002).
101. Gaszner, M. y Felsenfeld, G. Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet* 7, 703-713 (2006).
- 50 102. Wallace, J.A. y Felsenfeld, G. We gather together: insulators and genome organization. *Curr Opin Genet Dev* 17, 400-407 (2007).
103. Chung, J.H., Bell, A.C. y Felsenfeld, G. Characterization of the chicken beta-globin insulator. *Proc Natl Acad Sci U S A* 94, 575-580 (1997).

104. Bell, A.C., West, A.G. y Felsenfeld, G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98, 387-396 (1999).
105. Ryu, B.Y., Persons, D.A., Evans-Galea, M.V., Gray, J.T. y Nienhuis, A.W. A chromatin insulator blocks interactions between globin regulatory elements and cellular promoters in erythroid cells. *Blood Cells Mol Dis* 39, 221-228 (2007).
106. Ryu, B.Y., Evans-Galea, M.V., Gray, J.T., Bodine, D.M., Persons, D.A. y Nienhuis, A.W. An experimental system for the evaluation of retroviral vector design to diminish the risk for proto-oncogene activation. *Blood* 111, 1866-1875 (2008).
107. Yao, S., Osborne, C.S., Bharadwaj, R.R., Pasceri, P., Sukonnik, T., Pannell, D., Recillas-Targa, F., West, A.G. y Ellis, J. Retrovirus silencer blocking by the cHS4 insulator is CTCF independent. *Nucleic Acids Res* 31, 5317-5323 (2003).
108. Nishino, T., Tubb, J. y Emery, D.W. Partial correction of murine beta-thalassemia with a gammaretrovirus vector for human gamma-globin. *Blood Cells Mol Dis* 37, 1-7 (2006).
109. Aker, M., Tubb, J., Groth, A.C., Bukovsky, A.A., Bell, A.C., Felsenfeld, G., Kiem, H.P., Stamatoyannopoulos, G. y Emery, D.W. Extended core sequences from the cHS4 insulator are necessary for protecting retroviral vectors from silencing position effects. *Hum Gene Ther* 18, 333-343 (2007).
110. Li, C.L. y Emery, D.W. The cHS4 chromatin insulator reduces gammaretroviral vector silencing by epigenetic modifications of integrated provirus. *Gene Ther* 15, 49-53 (2008).
111. Ma, Y., Ramezani, A., Lewis, R., Hawley, R.G. y Thomson, J.A. High-level sustained transgene expression in human embryonic stem cells using lentiviral vectors. *Stem Cells* 21, 111-117 (2003).
112. Chang, L.J., Liu, X. y He, J. Lentiviral siRNAs targeting multiple highly conserved RNA sequences of human immunodeficiency virus type 1. *Gene Ther* 12, 1133-1144 (2005).
113. Pluta, K., Luce, M.J., Bao, L., Agha-Mohammadi, S. y Reiser, J. Tight control of transgene expression by lentivirus vectors containing second-generation tetracycline-responsive promoters. *J Gene Med* 7, 803-817 (2005).
114. Jakobsson, J., Rosenqvist, N., Thompson, L., Barraud, P. y Lundberg, C. Dynamics of transgene expression in a neural stem cell line transduced with lentiviral vectors incorporating the cHS4 insulator. *Experimental cell research* 298, 611-623 (2004).
115. Leboulch, P., Huang, G.M., Humphries, R.K., Oh, Y.H., Eaves, C.J., Tuan, D.Y. y London, I.M. Mutagenesis of retroviral vectors transducing human beta-globin gene and beta-globin locus control region derivatives results in stable transmission of an active transcriptional structure. *EMBO J* 13, 3065-3076 (1994).
116. Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenko, V.V. y Ren, B. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128, 1231-1245 (2007).
117. Yusufzai, T.M. y Felsenfeld, G. The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci U S A* 101, 8620-8624 (2004).
118. Phillips, J.E. y Corces, V.G. CTCF: master weaver of the genome. *Cell* 137, 1194-1211 (2009).
119. Giles, K.E., Gowher, H., Ghirlando, R., Jin, C. y Felsenfeld, G. Chromatin boundaries, insulators, and long-range interactions in the nucleus. *Cold Spring Harbor symposia on quantitative biology* 75, 79-85 (2010).
120. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. y Zhao, K. High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837 (2007).
121. Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., Thurman, R.E., Kaul, R., Myers, R.M. y Stamatoyannopoulos, J.A. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome research* 22, 1680-1688 (2012).
122. Schmidt, D., Schwalie, P.C., Wilson, M.D., Ballester, B., Goncalves, A., Kutter, C., Brown, G.D., Marshall, A., Flicek, P. y Odom, D.T. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* 148, 335-348 (2012).
123. Renda, M., Baglivo, I., Burgess-Beusse, B., Esposito, S., Fattorusso, R., Felsenfeld, G. y Pedone, P.V. Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. *J Biol Chem* 282, 33336-33345 (2007).
124. Dickson, J., Gowher, H., Strogantsev, R., Gaszner, M., Hair, A., Felsenfeld, G. y West, A.G. VEZF1 elements

mediate protection from DNA methylation. PLoS Genet 6, e1000804 (2010).

125. Li, C.L., Xiong, D., Stamatoyannopoulos, G. y Emery, D.W. Genomic and functional assays demonstrate reduced gammaretroviral vector genotoxicity associated with use of the cHS4 chromatin insulator. *Mol Ther* 17, 716-724 (2009).
- 5 126. Lisowski, L. y Sadelain, M. Locus control region elements HS1 and HS4 enhance the therapeutic efficacy of globin gene transfer in beta-thalassemic mice. *Blood* 110, 4175-4178 (2007).
127. Nagel, R.L., Bookchin, R.M., Johnson, J., Labie, D., Wajcman, H., Isaac-Sodeye, W.A., Honig, G.R., Schiliro, G., Crookston, J.H. y Matsutomo, K. Structural bases of the inhibitory effects of hemoglobin F and hemoglobin A2 on the polymerization of hemoglobin S. *Proc Natl Acad Sci U S A* 76, 670-672 (1979).
- 10 128. Sadelain *et al.*, *Proc. Nat'l Acad Sci. (USA)* (1995);92:6728-6732.
129. Armstrong, J.A., Emerson, B.M., 1996. NFE2 disrupts chromatin structure at human fl-globin locus control region hypersensitive site 2 in vitro. *Mol. Cell. Biol.* 16, 5634-5644.
130. Caterina, J.J., Ciavatta, D.J., Donze, D., Behringer, R.R., Townes, T.M., 1994. Multiple elements in human flglobin locus control region 5' HS2 are involved in enhancer activity and position-independent transgene expression. *Nucleic Acids Res.* 22, 1006 1011.
- 15 131. Moi, P., Kan, Y.W., 1990. Synergistic enhancement of globin gene expression by activator protein-I-like proteins. *Proc. Natl. Acad. Sci. USA* 87, 9000-9004.
132. Ney, P., Sorrentino, B., McDonagh, K., Nienhuis, A., 1990. Tandem AP-1-binding sites within the human /j-globin dominant control region function as an inducible enhancer in erythroid cells. *Genes Dev.* 4, 993 1006.
- 20 133. Shivdasani, R.A., Rosenblatt, M.F., Zucker-Franklin, D., Jackson, C.W., Hunt, P., Saris, C.J.M., Orkin, S.H., 1995. Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. *Cell* 81,695-704.
134. Talbot, D., Grosveld, F., 1991. The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. *EMBOJ.* 10, 1391-1398.
- 25 135. Hardison *et al.*, *Gene* (1997);205:73-94.
136. Elnitski *et al.*, *The Journal of Biological Chemistry* (1997);272(1):369-378; Horak *et al.*, *PNAS* (2002);99(5):2924-2929.
137. Shimotsuma *et al.*, *Journal of Biological Chemistry* (2010);285(19):14495-14503.
- 30 A partir de la descripción anterior, resultará evidente que pueden realizarse variaciones y modificaciones en el contenido dado a conocer en el presente documento descrito en el presente documento para adaptarlo a diversos usos y condiciones. La invención es tal como se define en las siguientes reivindicaciones.

**REIVINDICACIONES**

1. Casete de expresión que comprende un aislante y un gen de globina o una parte funcional del mismo operativamente unido a una región de región de control de locus (LCR) de  $\beta$ -globina, en el que el aislante comprende:
  - 5 (a) la secuencia de sitio de unión a CTCF expuesta en SEQ ID NO: 18, opcionalmente el aislante comprende la secuencia de ácido nucleico expuesta en SEQ ID NO: 24 o SEQ ID NO: 25; o
  - (b) la secuencia de nucleótidos expuesta en SEQ ID NO: 1.
2. Casete de expresión según la reivindicación 1, en el que el aislante tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1.
- 10 3. Casete de expresión según la reivindicación 1 ó 2, en el que la región LCR de  $\beta$ -globina
  - a. no comprende una región de sitio hipersensible 2 (HS2) a ADNasa I, y opcionalmente
    - i. la región LCR de  $\beta$ -globina no comprende una secuencia central de HS2; y/o
    - ii. la región LCR de  $\beta$ -globina no comprende una región de HS2 que conserva la actividad potenciadora de HS2;
  - 15 b. comprende una región de sitio hipersensible 1 (HS1) a ADNasa I, una región de sitio hipersensible 3 (HS3) a ADNasa I, y una región de sitio hipersensible 4 (HS4) a ADNasa I, y opcionalmente la región de HS3 está posicionada entre la región de HS1 y la región de HS4;
  - c. no comprende una región de HS1, y opcionalmente
    - i. la región LCR de  $\beta$ -globina no comprende una secuencia central de HS1;
    - 20 ii. la región LCR de  $\beta$ -globina no comprende una región de HS1 que conserva la función de HS1; y/o
    - iii. la región LCR de  $\beta$ -globina comprende una región de HS3 y una región de HS4, y opcionalmente la región de HS3 está posicionada entre el gen de globina o parte funcional del mismo y la región de HS4; y/o
  - 25 d. comprende una región de HS2, una región de HS3, y una región de HS4, y opcionalmente que comprende además una región de HS1.
4. Casete de expresión según la reivindicación 3, en el que la secuencia central de HS2 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 20 o SEQ ID NO: 21.
5. Casete de expresión según la reivindicación 3 ó 4, en el que la región de HS1 tiene aproximadamente 1,1 kb de longitud, aproximadamente 600 pb de longitud, o aproximadamente 490 pb de longitud.
- 30 6. Casete de expresión según la reivindicación 5, en el que la región de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 2, SEQ ID NO: 3 o SEQ ID NO: 4.
7. Casete de expresión según la reivindicación 3, en el que la secuencia central de HS1 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 22 o SEQ ID NO: 23.
- 35 8. Casete de expresión según una cualquiera de las reivindicaciones 3-7, en el que la región de HS3 tiene aproximadamente 1300 pb de longitud.
9. Casete de expresión según la reivindicación 8, en el que la región de HS3 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 5.
10. Casete de expresión según una cualquiera de las reivindicaciones 3-9, en el que la región de HS4 tiene aproximadamente 1,1 kb de longitud o aproximadamente 450 pb de longitud.
- 40 11. Casete de expresión según la reivindicación 10, en el que la región de HS4 tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8.
12. Casete de expresión según la reivindicación 3, en el que
  - e. la región LCR de  $\beta$ -globina comprende una región de HS1, una región de HS3, y una región de HS4 y no comprende una región de HS2, en el que
    - 45 i. la región de HS1 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 2, la región de HS3

- comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y la región de HS4 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 6;
- 5 ii. la región de HS1 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 3, la región de HS3 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y la región de HS4 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 8; o
- iii. la región de HS1 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 4, la región de HS3 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y la región de HS4 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 8; o
- 10 f. la región LCR de  $\beta$ -globina comprende una región de HS3 y una región de HS4 y no comprende una región de HS1 o una región de HS2, en el que la región de HS3 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 5 y la región de HS4 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 6; o
- 15 g. la región LCR de  $\beta$ -globina comprende una región de HS2, una región de HS3, y una región de HS4, y la región de HS2 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 9, la región de HS3 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y la región de HS4 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 7.
13. Casete de expresión según la reivindicación 3, en el que la región LCR de  $\beta$ -globina comprende una región de HS2, una región de HS3, y una región de HS4 y la región de HS2 tiene 860 pb de longitud, y/o la región de HS3 tiene aproximadamente 1300 pb de longitud, y/o la región de HS4 tiene aproximadamente 1,1 kb de longitud.
- 20 14. Casete de expresión según la reivindicación 13, en el que la región de HS2 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 9, y/o la región de HS3 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 5, y/o la región de HS4 comprende la secuencia de nucleótidos expuesta en SEQ ID NO: 7.
- 25 15. Casete de expresión según una cualquiera de las reivindicaciones 1-14, en el que el gen de globina se selecciona del grupo que consiste en gen de  $\beta$ -globina, gen de  $\gamma$ -globina y gen de  $\delta$ -globina.
16. Casete de expresión según la reivindicación 15, en el que el gen de globina es gen de  $\beta$ -globina humana, y opcionalmente el gen de  $\beta$ -globina humana se selecciona del grupo que consiste en un gen de  $\beta$ -globina humana silvestre, un gen de  $\beta$ -globina humana deleciónado que comprende una o más deleciones de secuencias de intrones, y un gen de  $\beta$ -globina humana mutado que codifica para al menos un residuo de aminoácido anti-falci-formación.
- 30 17. Casete de expresión según la reivindicación 16, en el que el gen de  $\beta$ -globina humana es gen de  $\beta^A$ -globina humana que codifica para una mutación de treonina a glutamina en el codón 87 ( $\beta^{A-T87Q}$ ).
18. Casete de expresión según una cualquiera de las reivindicaciones 1-17, que comprende un aislante que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 1, o que comprende dos aislantes que tienen cada uno la secuencia de nucleótidos expuesta en SEQ ID NO: 1.
- 35 19. Casete de expresión según una cualquiera de las reivindicaciones 1-18, que comprende además:
- h. un promotor de  $\beta$ -globina, y opcionalmente el promotor de  $\beta$ -globina está posicionado entre el gen de globina o parte funcional del mismo y la región LCR de  $\beta$ -globina;
- 40 i. un potenciador en 3' de  $\beta$ -globina humana, y opcionalmente el potenciador en 3' de  $\beta$ -globina humana está posicionado aguas arriba del gen de globina o parte funcional del mismo; y
- j. al menos un (por ejemplo, uno, dos o tres) potenciador específico de eritroides, y opcionalmente el al menos un potenciador específico de eritroides está posicionado entre el gen de globina o parte funcional del mismo y la región LCR de  $\beta$ -globina.
- 45 20. Casete de expresión según la reivindicación 19, en el que
- k. el promotor de  $\beta$ -globina es un promotor de  $\beta$ -globina humana que tiene aproximadamente 610 pb de longitud y opcionalmente tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 10, o tiene aproximadamente 265 pb de longitud, y opcionalmente tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 11;
- 50 l. el potenciador en 3' de  $\beta$ -globina humana tiene aproximadamente 880 pb de longitud, y opcionalmente el potenciador en 3' de  $\beta$ -globina humana tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 12; y/o

m. el al menos un potenciador específico de eritroides tiene una secuencia de nucleótidos expuesta en SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17.

- 5 21. Casete de expresión según una cualquiera de las reivindicaciones 1-20, en el que el casete de expresión permite la expresión del gen de globina o una parte funcional del mismo en un mamífero, opcionalmente la expresión se restringe a tejido eritroide.
22. Vector recombinante que comprende el casete de expresión según una cualquiera de las reivindicaciones 1-21, y opcionalmente en el que el vector recombinante es un vector retroviral, y opcionalmente en el que el vector recombinante es un vector de lentivirus.
- 10 23. Vector recombinante según la reivindicación 22, que comprende además: un elemento post-regulador de la hepatitis de la marmota (WPRE) en la repetición terminal larga (LTR) en 3' del vector, y/o una señal de poliadenilación de hormona de crecimiento bovina en la repetición terminal larga (LTR) en 3' del vector.
24. Sistema de nucleasa modificado por ingeniería o que no se produce de manera natural que comprende el casete de expresión según una cualquiera de las reivindicaciones 1-21.
- 15 25. Sistema de nucleasa según la reivindicación 24, en el que la nucleasa se selecciona del grupo que consiste en una nucleasa de dedos de zinc modificada por ingeniería o que no se produce de manera natural (ZFN), una meganucleasa modificada por ingeniería o que no se produce de manera natural, y una nucleasa efectora de tipo activador de la transcripción (TALEN) modificada por ingeniería o que no se produce de manera natural, y/o la nucleasa comprende un dominio de unión a ADN y un dominio de escisión de nucleasa, y/o la nucleasa se une a un sitio de locus seguro genómico, y opcionalmente
- 20 n. la nucleasa genera una rotura de cadena doble (DSB) en el sitio de locus seguro genómico,  
o. el sitio de locus seguro genómico es un sitio de locus seguro genómico extragenómico,  
p. el sitio de locus seguro genómico está ubicado en el cromosoma 1, y/o
- 25 q. el sitio de locus seguro genómico cumple la totalidad de los cinco criterios siguientes: (1) distancia de al menos 50 kb desde el extremo 5' de cualquier gen (por ejemplo, desde el extremo 5' del gen), (ii) distancia de al menos 300 kb desde cualquier gen relacionado con cáncer, (iii) dentro de una estructura de cromatina abierta/accesible (medido mediante escisión de ADN con nucleasas naturales o modificadas por ingeniería), (iv) ubicación fuera de una unidad de transcripción génica y (v) ubicación fuera de regiones ultraconservadas (UCR), microARN o ARN largo no codificante del genoma humano.
- 30 26. Sistema de CRISPR-Cas modificados por ingeniería o que no se producen de manera natural que comprende el casete de expresión según una cualquiera de las reivindicaciones 1-21.
27. Sistema según la reivindicación 26, en el que el sistema de CRISPR-Cas comprende una nucleasa de CRISPR-Cas y un ARN guía individual, y opcionalmente el sistema de CRISPR-Cas se une a un sitio de locus seguro genómico.
- 35 28. Polinucleótido que codifica para el sistema de nucleasa según la reivindicación 24 ó 25 o el sistema de CRISPR-Cas según la reivindicación 26 ó 27.
29. Vector que comprende el polinucleótido según la reivindicación 28, y opcionalmente el vector es un vector lentiviral.
- 40 30. Célula transducida con el casete de expresión según una cualquiera de las reivindicaciones 1-21, el vector recombinante según la reivindicación 22 ó 23, el sistema de nucleasa según la reivindicación 24 ó 25, el sistema de CRISPR-Cas según la reivindicación 26 ó 27, o el vector según la reivindicación 29.
31. Célula según la reivindicación 30, en la que la célula se selecciona del grupo que consiste en una célula madre hematopoyética, una célula madre embrionaria, una célula madre pluripotente inducida, y una célula endotelial hemogénica, y opcionalmente la célula se transduce *ex vivo*, y opcionalmente la célula madre hematopoyética es una célula madre hematopoyética CD34<sup>+</sup>.
- 45 32. Composición farmacéutica que comprende una cantidad eficaz de la célula según la reivindicación 30 ó 31 y un portador farmacéuticamente aceptable.
- 50 33. Composición farmacéutica según la reivindicación 32 para su uso en el tratamiento de una hemoglobinopatía, opcionalmente en la que la hemoglobinopatía se selecciona del grupo que consiste en enfermedad de hemoglobina C, enfermedad de células falciformes (SCD) de hemoglobina, anemia de células falciformes, anemia hereditaria, talasemia,  $\beta$ -talasemia, talasemia mayor, talasemia intermedia,  $\alpha$ -talasemia y enfermedad de hemoglobina H.

34. Kit para tratar una hemoglobinopatía que comprende la célula según la reivindicación 30 ó 31, y opcionalmente que comprende además instrucciones escritas para usar la célula para tratar a un sujeto que tiene una hemoglobinopatía, y/o en el que la hemoglobinopatía se selecciona del grupo que consiste en enfermedad de hemoglobina C, enfermedad de células falciformes (SCD) de hemoglobina, anemia de células falciformes, anemia hereditaria, talasemia,  $\beta$ -talasemia, talasemia mayor, talasemia intermedia,  $\alpha$ -talasemia y enfermedad de hemoglobina H.
- 5
35. Célula según la reivindicación 30 ó 31 para su uso en el tratamiento de una hemoglobinopatía en un sujeto, o para su uso en la restauración de la capacidad de un sujeto para producir glóbulos rojos que contienen hemoglobina normal.

10

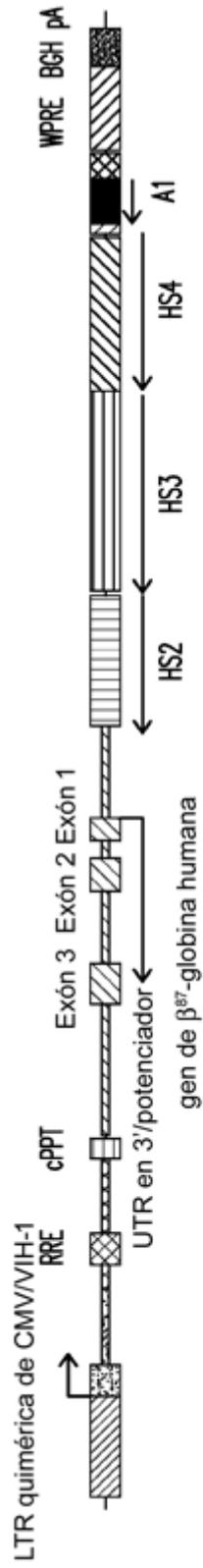


FIG. 1

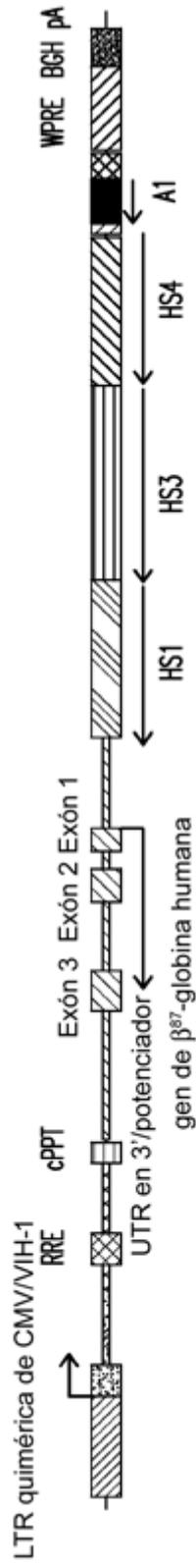


FIG. 2

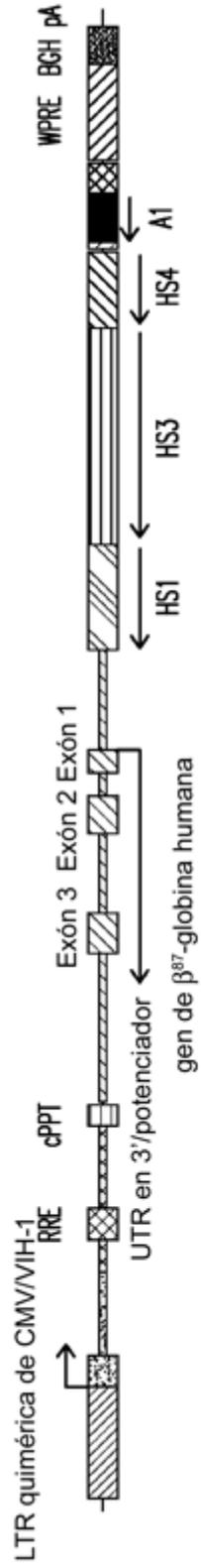


FIG. 3

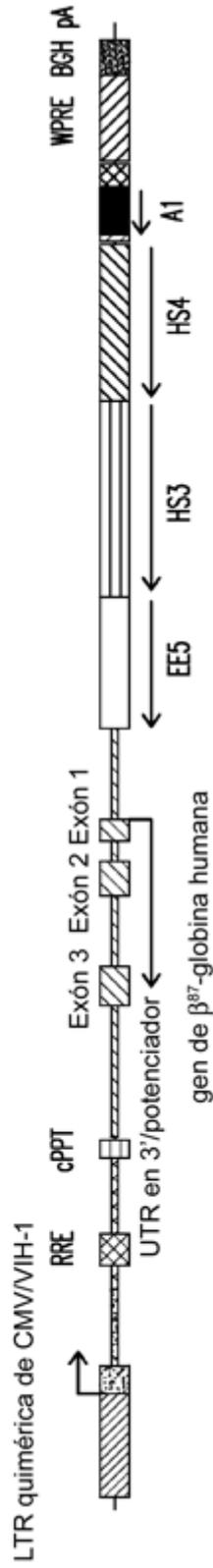
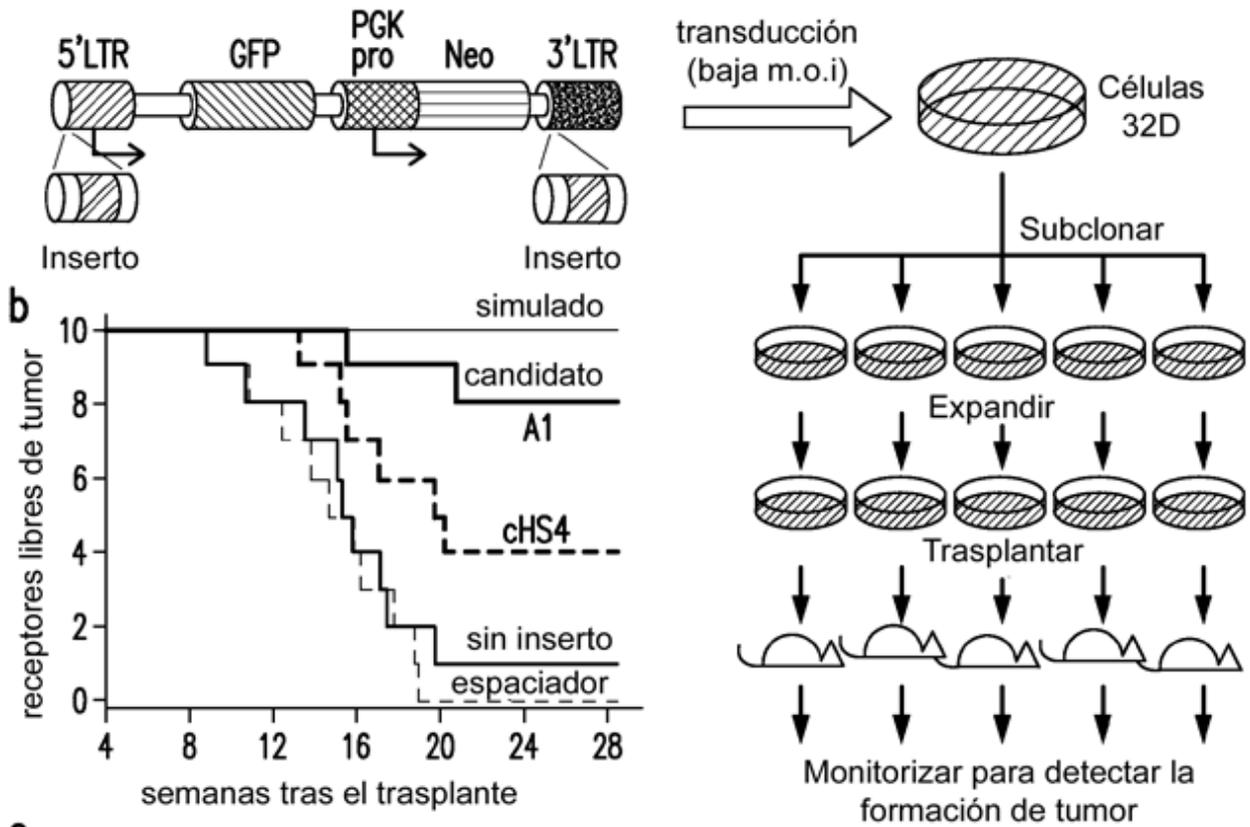


FIG. 4

**a** Ensayo de genotoxicidad de vector gamma-retroviral



Vector	(a) N.º de tumores estimado	(b) N.º de provirus estimado	Tumores por 10 <sup>5</sup> provirus	(c) Probabilidad frente a sin inserto	(c) Probabilidad frente a cHS4
simulado	0	0	0	<< 0,001	<< 0,001
sin inserto	23,0	0,49 x 10 <sup>5</sup>	46,9	--	<0,001
espaciador	≥30	0,67 x 10 <sup>5</sup>	≥44,7	n.e.	<0,001
cHS4	9,15	0,54 x 10 <sup>5</sup>	16,9	<0,001	--
A1	2,23	0,64 x 10 <sup>5</sup>	3,9	<0,001	<0,05

(a) Basándose en la distribución de posición para la fracción de receptores sin tumores;  
 (b) Basándose en el número de células inicial y las tasas de transducción iniciales;  
 (c) Basándose en la prueba de la Z para dos proporciones.

**FIG. 5**

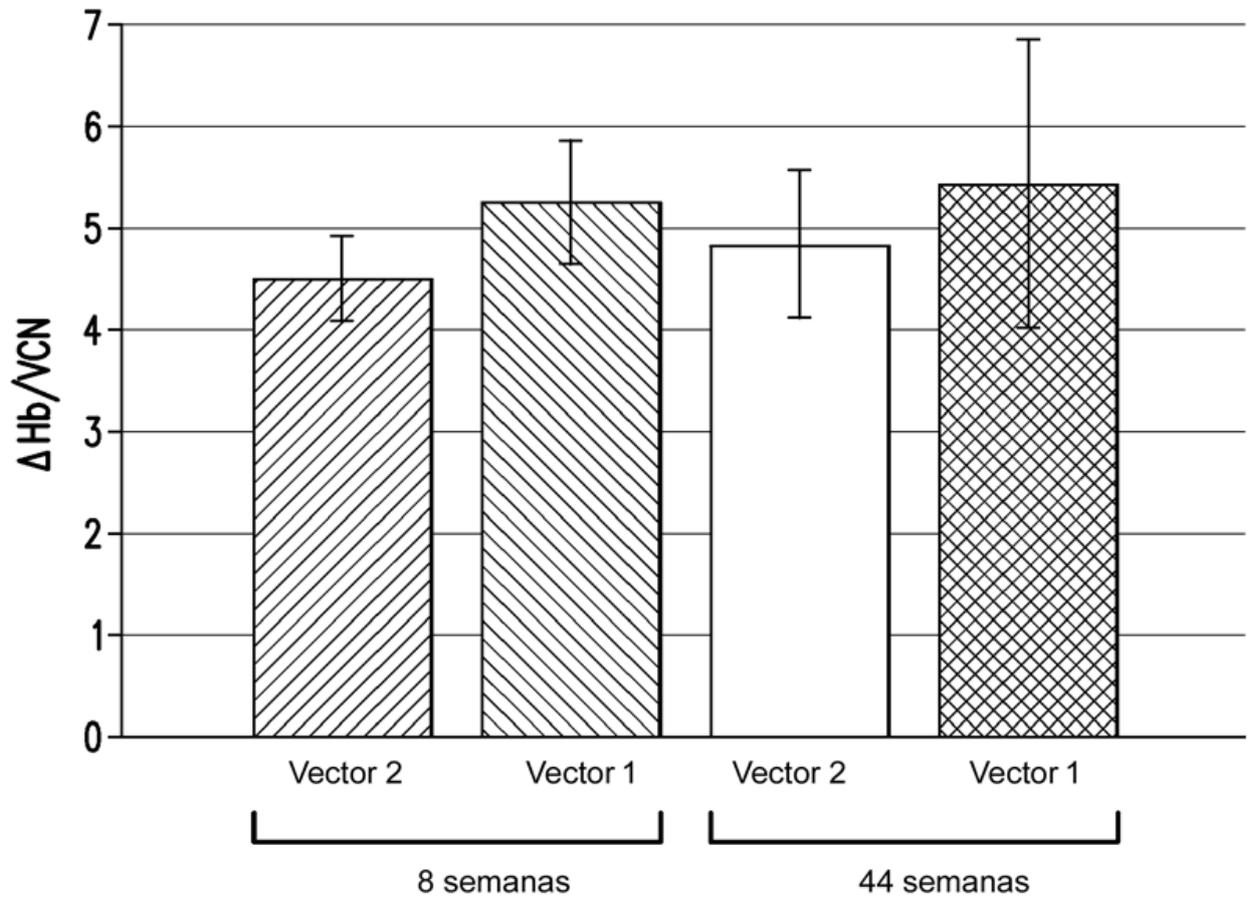


FIG. 6

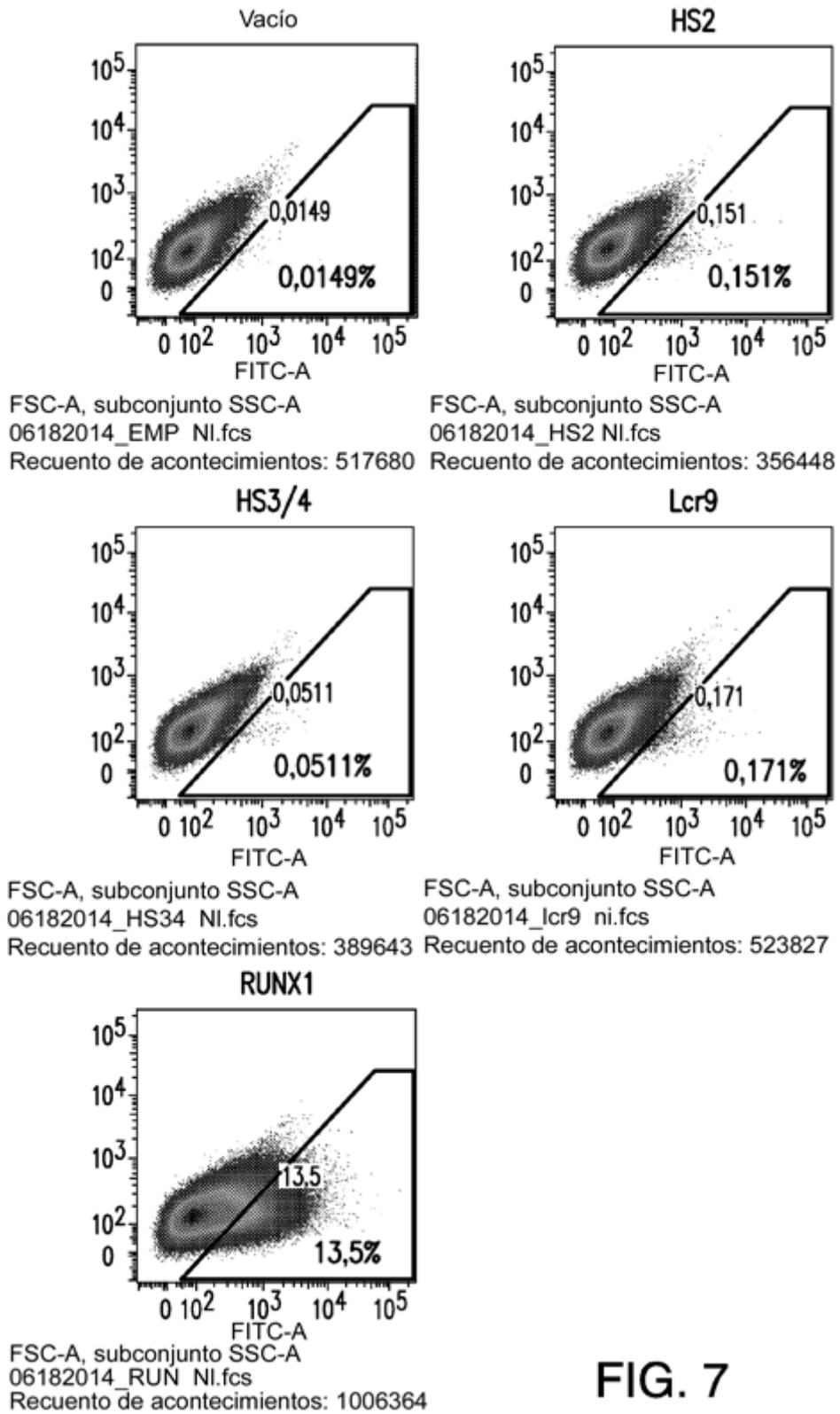
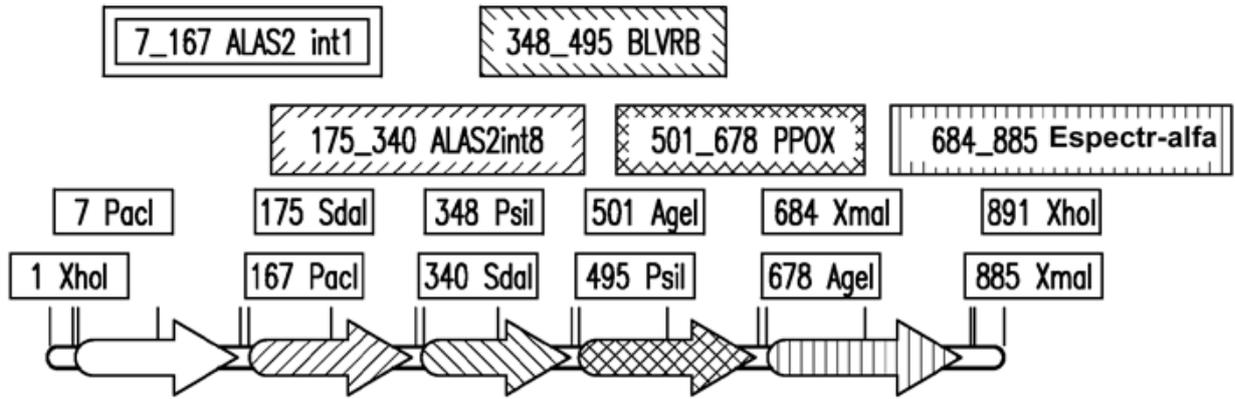


FIG. 7





nuevo potenciador- 896 nt

· Nuevo POTENCIADOR

• CTCGAGttaattaaTCTCCCACGCCCTGGTCTCAGCTTGGGGAGTGGTCAGACCCCAATGGCGATAA  
 ACTCTGGCAACTTTATCTGTGcaCTGCAGGCTCAGCCCCAAcaGCTTTAGCTTTCACAAGCAGGCAGGGG  
 AAGGAAACACATATCTCCAGATATGAGGttaattaaacctgcaggCTAAACCCCTCCCCACCCTAGCCC  
 CAAGCTTCATCTTAGCTCCACTCCTGACCCTATCCAGCTAAAGGTCCCACCCAGCTCCTGCCTATCTAG  
 TCATTGCATATGGCAAGACTTGAAAGTCCTATCTCAAAGCAGCAGAATTATCAGCTACGACTcctgcagg  
 ttataaCCATCCCCAGCACTCCCTGCCCCACAGCCCAGACTTGACCAACTCCCAGTccGCCTGGGAC  
 TTCCAGATATGGGGCCCCACCCTTGAGGCCTTGGGGACGCTGAAGATATTGACTATCTGCGTGCCggAA  
 AAGGGTgtataaaccggtAAAGGCTGGGGGTGGGAGTAGCGGATTTGAAGCACTTGTGGCCTACAGAG  
 GTGTGGCAAGCAGAGCACCTCAGAACTCAGGCGTACTGCCCCCGCCCGAGCCCTGCGAGGGCCGATAGC  
 GAGGGTGTGGCCCTTATCTGCACCCAGCAGAGCGCCGGGGGTACGGTCaccggtcccgggCAGTTGCC  
 TCAGCTGAGTATGTCTTCTAAAGATAATGTCGATTGTGTATGGCTGATGGATTCTAGGACCAAGCAAGA  
 GGTTTTTTTTTTCCCCACATACTTAACGTTTCTATATTTCTATTTGAATTCGACTGGACAGTTCCATT  
 TGAATTATTTCTCTCTCTCTCTCTCTGACACATTTTATCTTGCCAcccgggCTCGAG

FIG. 9

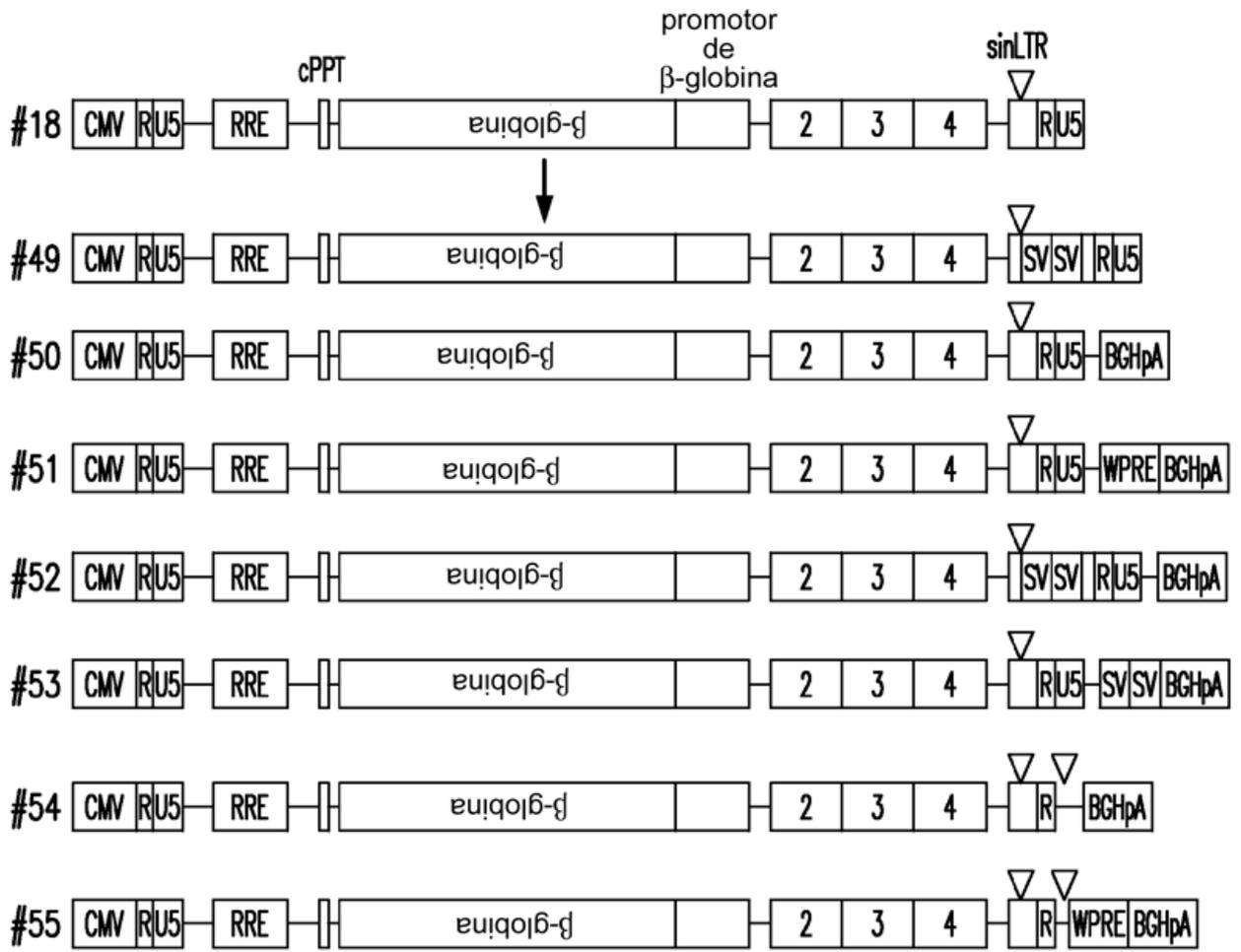


FIG. 10A

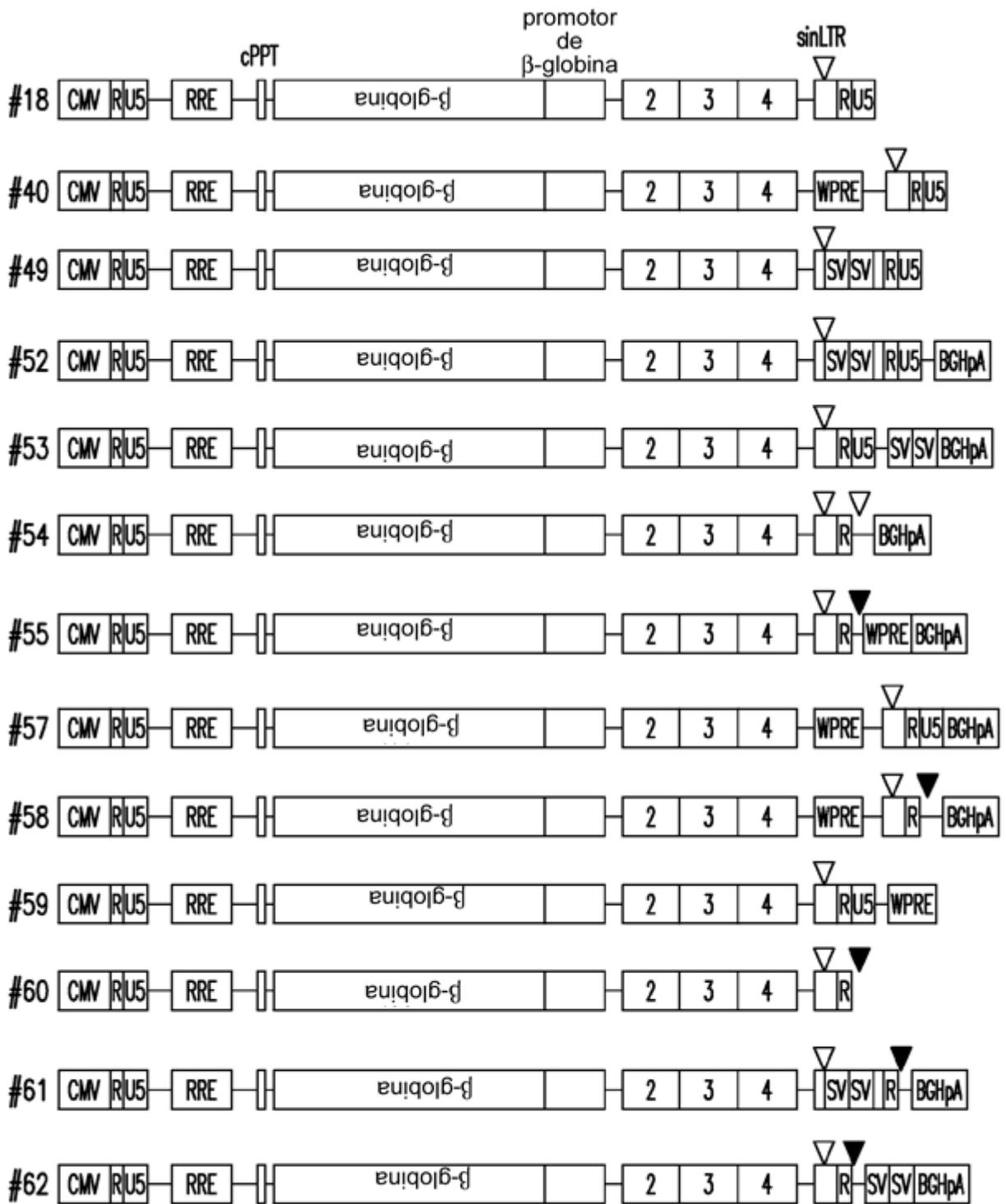


FIG. 10B

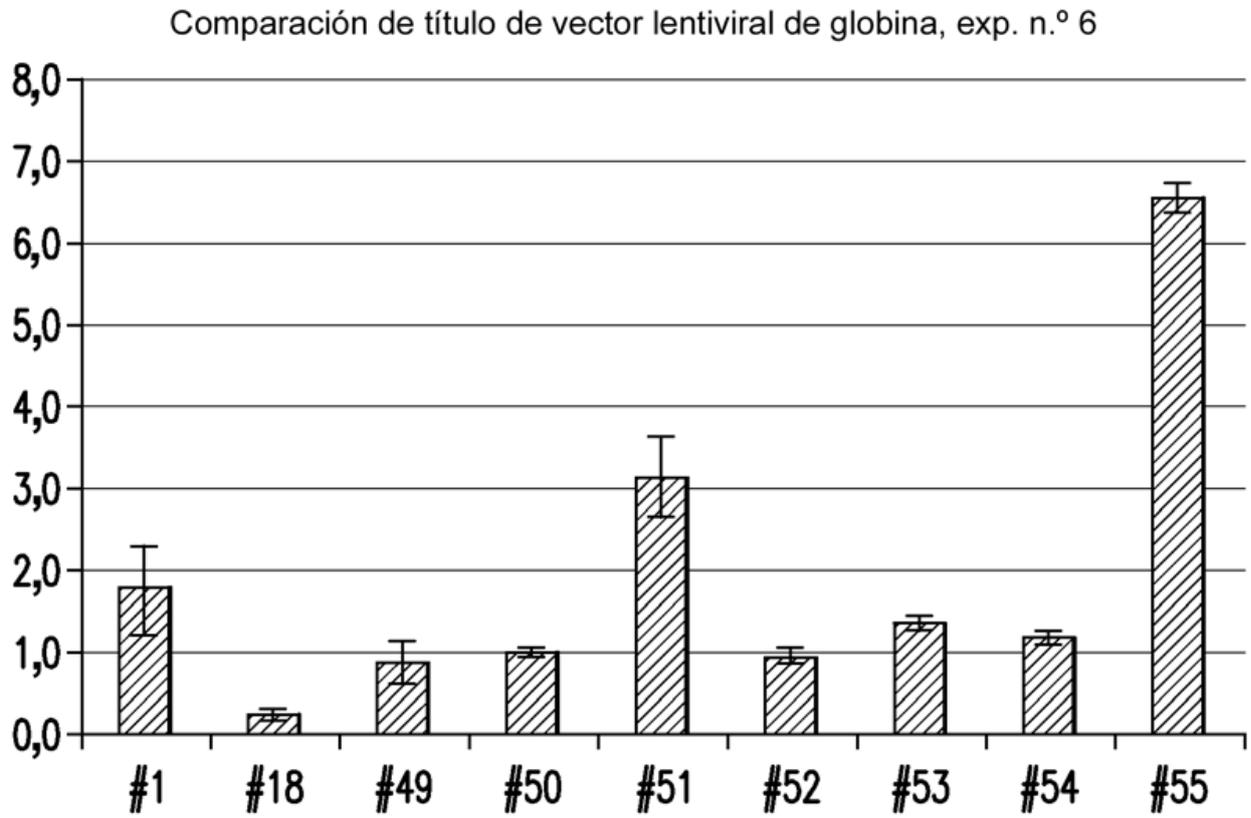


FIG. 11

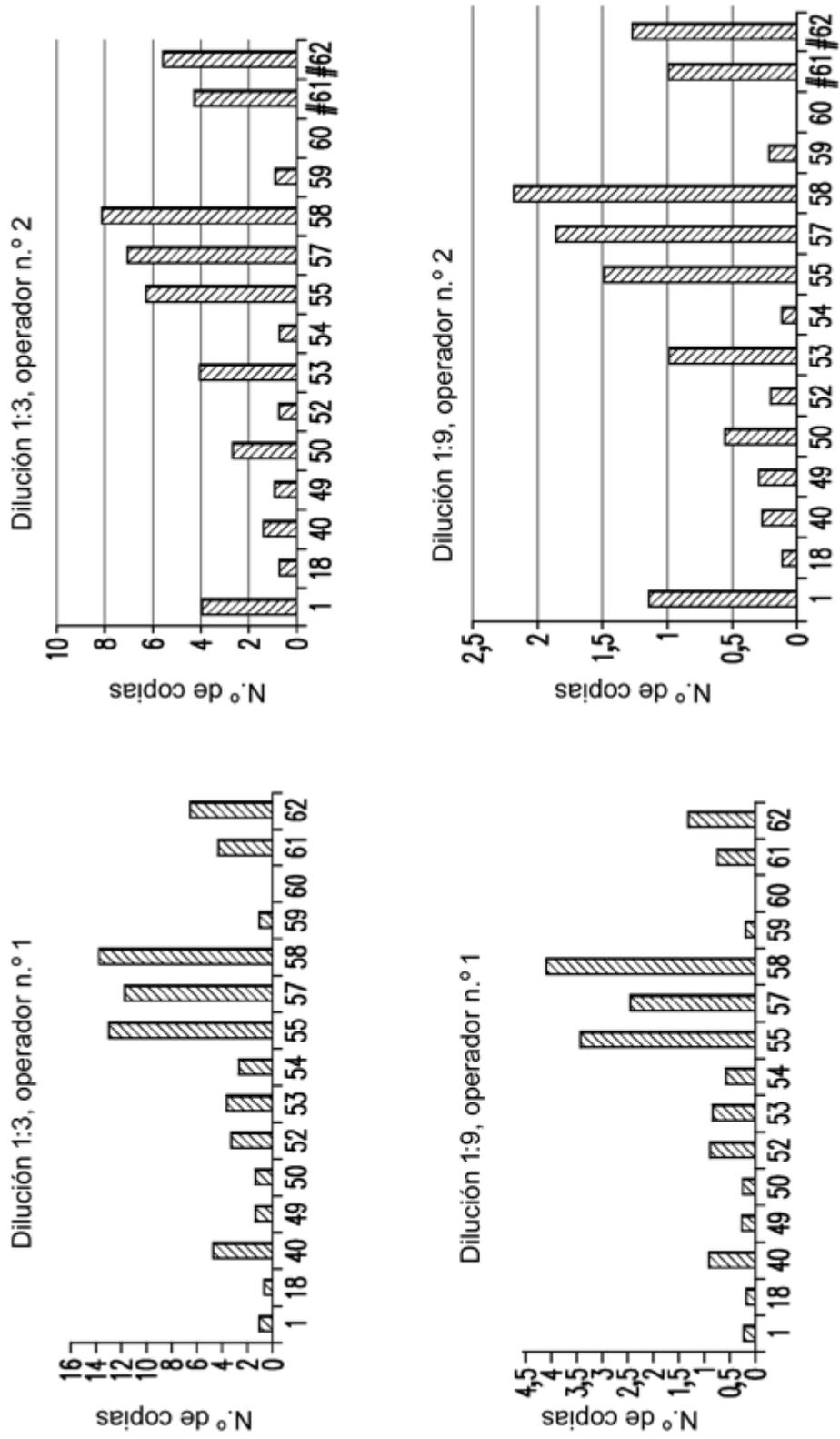


FIG. 12