

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 887 776**

51 Int. Cl.:

C12N 15/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.09.2018 PCT/EP2018/075354**

87 Fecha y número de publicación internacional: **28.03.2019 WO19057774**

96 Fecha de presentación y número de la solicitud europea: **19.09.2018 E 18769378 (3)**

97 Fecha y número de publicación de la concesión europea: **18.08.2021 EP 3684925**

54 Título: **Vectores de ADN no integradores para la modificación genética de células**

30 Prioridad:

19.09.2017 EP 17191829

17.09.2018 WO PCT/US2018/051381

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.12.2021

73 Titular/es:

**DEUTSCHES KREBSFORSCHUNGSZENTRUM
(50.0%)**

Im Neuenheimer Feld 280

69120 Heidelberg, DE y

NATURE TECHNOLOGY CORP. (50.0%)

72 Inventor/es:

HARBOTTLE, RICHARD;

BOZZA, MATTHIAS y

WILLIAMS, JAMES, A.

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 887 776 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

Vectores de ADN no integradores para la modificación genética de células

La presente invención se refiere al campo de los vectores de expresión vertebrados episomales no integradores autorreplicantes, útiles para terapia génica, terapia celular ex vivo, terapia con células madre y, más particularmente, para mejorar la expresión de antígenos o genes terapéuticos codificados por el vector. Estas moléculas de ADN recombinante son útiles en biotecnología, organismos transgénicos, terapia génica, terapia con células madre, vacunación terapéutica, agricultura y vacunas de ADN. Más específicamente, se refiere a un polinucleótido que comprende al menos un promotor y un elemento S/MAR, en el que dicho elemento S/MAR está situado aguas abajo de dicho promotor y en el que la secuencia de ácido nucleico de dicho elemento S/MAR (secuencia S/MAR) comprende al menos 3 motivos de secuencia ATTA (SEQ ID NO:1) por cada 100 nucleótidos en un tramo de como máximo 200 nucleótidos; la presente invención se refiere además a una composición y a una célula huésped que comprende dicho polinucleótido, y al polinucleótido para su uso en medicina y para su uso en el tratamiento de enfermedades genéticas. La presente invención también se refiere a un kit y a un dispositivo que comprende dicho polinucleótido, y a procedimientos y usos relacionados con el polinucleótido.

La modificación genética de las células se utiliza de forma rutinaria en los cultivos celulares modernos con fines científicos. Sin embargo, el uso de las técnicas correspondientes en el tratamiento de enfermedades hereditarias causadas por mutaciones de genes, aunque es muy deseable, sigue viéndose obstaculizado por el problema de que los procedimientos disponibles normalmente sólo proporcionan una modificación transitoria, como los protocolos de transfección transitoria, mientras que los procedimientos que proporcionan una modificación estable de las células, como con vectores lentivirales virales o vectores de transposones no virales, suelen depender de la integración del transgén en el genoma de la célula huésped. Sin embargo, la integración de un transgén, incluso si está dirigido a un locus específico, conlleva el riesgo de inducir una mutación deletérea, que puede provocar, por ejemplo, un cáncer como efecto secundario del tratamiento.

Las regiones de fijación del andamio/matriz (S/MAR), que también se conocen como regiones de fijación del andamio (SAR) o regiones asociadas a la matriz (MAR) se conocen como secuencias en el genoma de los organismos eucariotas que median la fijación de la matriz nuclear. Los S/MARS son secuencias ricas en AT, y se comprobó que algunos motivos ricos en AT estaban más enriquecidos (Liebeich et al., (2002), NAR 30(15): 3433). Se ha propuesto una variedad de vectores para el mantenimiento estable en células basado en motivos S/MAR, por ejemplo en el documento US 6.410.314 B1 y en el documento Haase et al., (2010), BMC Biotechnology 10:20; además, se identificaron efectos epigenéticos que influyen en la replicación de dichos vectores (Haase et al., (013), PLOS One 8(11):e79262). No obstante, se necesitan vectores basados en S/MAR que sean lo suficientemente estables para su uso en terapia génica.

El nivel de expresión subóptimo, el silenciamiento del gen y la baja tasa de establecimiento representan las principales limitaciones de los vectores basados en S/MAR descritos en el arte.

Existe, por tanto, la necesidad de mejorar los medios y procedimientos para la transfección estable de células, en particular utilizando elementos S/MAR y evitando los riesgos que conlleva la integración del transgén en el genoma de la célula huésped. Este problema se resuelve con los medios y procedimientos aquí divulgados.

La presente invención se refiere a vectores útiles para la terapia génica episomal no integrativa y la terapia de células madre, y más particularmente, para mejorar la expresión de transgenes y la eficiencia de establecimiento del vector de un vector de expresión S/MAR episomal no integrador autorreplicante, y para eliminar la transferencia de genes marcadores de resistencia a los antibióticos mediante vectores no virales.

Se dan a conocer procedimientos y composiciones de vectores mejorados que mejoran la expresión y la eficiencia de establecimiento de un vector de expresión episomal no integrador S/MAR autorreplicante en una célula de vertebrado diana.

Un objeto de la invención es proporcionar una expresión mejorada de un vector de expresión S/MAR episomal autorreplicante no integrador en una célula de vertebrado diana.

Otro objeto de la invención es proporcionar una eficiencia de establecimiento mejorada de un vector de expresión S/MAR episomal no integrador autorreplicante en una célula de vertebrado diana.

En una realización, la presente tecnología proporciona un procedimiento para mejorar la eficiencia de expresión y establecimiento de un vector de expresión episomal S/MAR autorreplicante no integrador en una célula de vertebrado diana que comprende los siguientes pasos: a) proporcionar un vector de expresión episomal S/MAR que comprende i) una región de selección de replicación bacteriana que comprende un origen de replicación bacteriano y un marcador seleccionable; ii) una unidad de transcripción para la expresión de un transgén en una célula de vertebrado, que comprende un promotor, una UTR 5', un transgén y una UTR 3'; iii) un inserto S/MAR situado dentro de dicha UTR 3'; y b) modificar el vector de expresión S/MAR episomal de manera que el S/MAR esté flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' dentro de dicha UTR 3', con lo que el vector de expresión S/MAR episomal no integrador resultante ha mejorado la eficiencia de expresión y establecimiento tras la

transfección de una célula de vertebrado. En otra realización, dicho S/MAR contiene motivos internos de terminación de la transcripción AATAAA. En una realización adicional, dichos motivos de terminación de la transcripción AATAAA en dicha S/MAR se sustituyen por motivos AATATT. En otra realización, dicha S/MAR se selecciona del grupo que consiste en la S/MAR de Interferón beta humano, la S/MAR de M18, la S/MAR de ApoL1. En otra realización, dicho SMAR flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, y SEQ ID NO: 23. En otra realización, dicho origen bacteriano de replicación es un origen de replicación gamma R6K. En otra realización, dicho origen bacteriano es un origen de replicación gamma R6K con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, y SEQ ID NO: 4. En otra realización, dicho marcador seleccionable es una variante funcional de ARN-IN reguladora de ARN-OUT con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5, y SEQ ID NO: 7. En otra realización, dicho marcador seleccionable es un marcador seleccionable de ARN-OUT que codifica un ARN-IN regulador de ARN-OUT con al menos un 95% de identidad de secuencia con SEQ ID NO: 6. En otra realización, dicha región de replicación-selección bacteriana que comprende un origen de replicación bacteriano y un marcador seleccionable es una región de replicación-selección bacteriana de origen-ARN-ARN de salida R6K con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, y SEQ ID NO: 17. En otra realización, dicha UTR 5' codifica además un intrón. En una realización adicional, dicha unidad de transcripción codifica además un potenciador de expresión situado aguas arriba del promotor. En otra realización, dicho potenciador de expresión tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 27, y SEQ ID NO: 28. En otra realización, dicho sitio donante de empalme tiene al menos un 95% de identidad de secuencia con la SEQ ID NO: 25. En otra realización, dicho sitio aceptor de empalme tiene al menos un 95% de identidad de secuencia con SEQ ID NO: 26. En otra realización, dicho vector de expresión episomal no integrador de S/MAR se selecciona del grupo que consiste en un vector plasmídico, un vector nanoplásmido, un vector lentiviral de integración deficiente y vectores lentivirales no integradores.

En otra realización, la presente tecnología proporciona una molécula de ADN recombinante circular covalentemente cerrada libre de marcadores de antibióticos que comprende a) una unidad de transcripción libre de marcadores antibióticos para la expresión de un transgén en una célula de vertebrado, que comprende un promotor, una UTR de 5', un transgén y una UTR de 3'; b) un S/MAR situado dentro de dicha UTR de 3' en el que dicho S/MAR está flanqueado por un sitio donante de empalme de 5' y un sitio aceptor de empalme de 3'; c) un origen de replicación gamma R6K con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, y SEQ ID NO: 4; y d) un marcador seleccionable de ARN-OUT que comprende una variante funcional de ARN-IN reguladora de ARN-OUT con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5, y SEQ ID NO: 7. En otra realización, dicho origen de replicación gamma R6K y dicho marcador seleccionable de ARN-OUT comprenden una región de selección de replicación bacteriana de origen R6K-ARN-OUT con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, y SEQ ID NO: 17. En otra realización, dicha S/MAR se selecciona del grupo que consiste en la S/MAR de Interferón beta humano, la S/MAR de M18, la S/MAR de ApoL1. En otra realización, dicho S/MAR contiene motivos internos de terminación de la transcripción AATAAA. En una realización adicional, dichos motivos de terminación de la transcripción AATAAA en dicha S/MAR se sustituyen por motivos AATATT. En otra realización, dicha S/MAR se selecciona del grupo que consiste en la S/MAR de Interferón beta humano, la S/MAR de M18, la S/MAR de ApoL1. En otra realización, dicho SMAR flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, y SEQ ID NO: 23. En otra realización, dicha UTR 5' codifica además un intrón. En una realización adicional, dicha unidad de transcripción codifica además un potenciador de expresión situado aguas arriba del promotor. En otra realización, dicho potenciador de expresión tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 27, y SEQ ID NO: 28. En otra realización, dicho sitio donante de empalme tiene al menos un 95% de identidad de secuencia con la SEQ ID NO: 25. En otra realización, dicho sitio aceptor de empalme tiene al menos un 95% de identidad de secuencia con SEQ ID NO: 26.

En otra realización, la presente tecnología proporciona una molécula de ADN recombinante circular cerrada covalentemente que comprende a) una unidad de transcripción para la expresión de un transgén en una célula de vertebrado, que comprende un promotor, una UTR de 5', un transgén y una UTR de 3'; b) un S/MAR situado dentro de dicha UTR de 3', en el que dicho S/MAR está flanqueado por un sitio donante de empalme de 5' y un sitio aceptor de empalme de 3'; c) un origen de replicación gamma R6K con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, y SEQ ID NO: 4; y d) un marcador seleccionable de ARN-OUT que comprende una variante funcional de ARN-IN reguladora de ARN-OUT con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5, y SEQ ID NO: 7. En otra realización, dicho origen de replicación gamma R6K y dicho marcador seleccionable de ARN-OUT comprenden una región de selección de replicación bacteriana de origen R6K-ARN-OUT

con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, y SEQ ID NO: 17. En otra realización, dicha S/MAR se selecciona del grupo que consiste en la S/MAR de Interferón beta S/MAR humano, S/MAR de M18, S/MAR de ApoL1. En otra realización, dicho S/MAR contiene motivos internos de terminación de la transcripción AATAAA. En una realización adicional, dichos motivos de terminación de la transcripción AATAAA en dicha S/MAR se sustituyen por motivos AATATT. En otra realización, dicha S/MAR se selecciona del grupo que consiste en la S/MAR de Interferón beta humano, la S/MAR de M18, la S/MAR de ApoL1. En otra realización, dicho SMAR flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, y SEQ ID NO: 23. En otra realización, dicha UTR 5' codifica además un intrón. En una realización adicional, dicha unidad de transcripción codifica además un potenciador de expresión situado aguas arriba del promotor. En otra realización, dicho potenciador de expresión tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 27, y SEQ ID NO: 28. En otra realización, dicho sitio donante de empalme tiene al menos un 95% de identidad de secuencia con la SEQ ID NO: 25. En otra realización, dicho sitio aceptor de empalme tiene al menos un 95% de identidad de secuencia con SEQ ID NO: 26.

Los plásmidos resultantes con un S/MAR flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' dentro de la UTR 3' han mejorado sorprendentemente el establecimiento y la expresión del transgén que los plásmidos con un S/MAR dentro de la UTR 3' sin sitios donantes y aceptores de empalme flanqueantes.

Breve descripción de las figuras

FIG. 1 representa el intrón pCI, con el punto de ramificación del donante de empalme (SD) y las regiones del aceptor de empalme (SA). FIG. 2 representa el interferón beta S/MAR (arriba), y un derivado SD del interferón beta S/MAR SA (centro), así como un derivado SD del interferón beta S/MAR SA en el que se mutaron las señales de poliadenilación internas AATAAA (abajo).

FIG. 3 representa el derivado M18 del interferón beta S/MAR con sitios SD y SA flanqueantes.

FIG. 4 representa los 805 pb (arriba) o 525 pb (abajo) de apoB S/MAR con los sitios SD y SA flanqueantes.

FIG. 5 representa el vector pMAX-UCOE-coGFP P2A-PuroR-NP (pSMARt UCOE).

FIG. 6 representa los vectores NTC9385R-UCOE-CMV- coGFP P2A-PuroR -SMAR-SV40 pA (NP-UCOE) y NTC9385R-UCOE-CMV- coGFP P2A-PuroR - SD SMAR- SA SV40 pA (NP-UCOE-SP).

FIG. 7 representa los vectores NTC9385R-SP-UCOE-CMV-GFP SMARter (NP-SMARter-SP) y NTC9385R-SP-UCOE-CMV-GFP CMARter (NP-CMARter-SP).

FIG. 8 representa los vectores NTC9385R- UCOE EF1 -coGFP SD -SMAR SA SV40 pA (NP-UCOE-EF1-SP) y NTC9385R- UCOE EFI-coGFP-SD SMAR R6K-R-OUT-SA pA (UCOE-EF1-SP-NP).

FIG. 9 representa el vector NTC9385R-SP-ELE40-CMV-GFP CMARter (NP-Ele40-CMARter-SP).

FIG. 10 representa la mejora de la expresión de los vectores S/MAR establecidos con y sin sitios SD y SA flanqueantes. Panel izquierdo: MFI de células HEK293T establecidas con un vector S/MAR con y sin uniones de empalme. Los vectores contienen la región bacteriana NP, el aislante genómico UCOE, el casete de expresión GFP -2A-PuroR impulsado por el promotor CMV y el interferón beta S/MAR en la UTR 3' con (Nano-S/MAR-splice = NP-UCOE-SP; NTC9385R-UCOE-CMV- coGFP P2A- PuroR - SD SMAR- SA SV40 pA, Figura 6) o sin (NP-UCOE; NTC9385R-UCOE-CMV- coGFP P2A-PuroR -SMAR-SV40 pA, Figura 6) sitios S/MAR flanqueantes SD y SA. Panel derecho: la mejora de la expresión de la transcripción se confirma mediante el análisis de la PCR en tiempo real. La expresión del transgén GFP se normalizó con respecto al gen de mantenimiento GAPDH.

FIG. 11 representa la mejora de la expresión de los vectores S/MAR establecidos con y sin sitios SD y SA flanqueantes. MFI de células establecidas (HEK293T y fibroblastos embrionarios de ratón primarios) con vectores que albergan diferentes S/MARs flanqueados por uniones de empalme. Los nombres de los vectores son los de las figuras 5, 6 y 7.

FIG. 12 representa la mejora del establecimiento de los vectores S/MAR con y sin sitios SD y SA flanqueantes. Ensayo de formación de colonias realizado en HEK293T con vectores que albergan dos S/MAR diferentes con y sin sitios SD y SA flanqueantes. pEPI es un vector plasmídico promotor de CMV con un S/MAR de interferón beta 3' UTR.

FIG. 13: Eficiencia del establecimiento y análisis de la población de células modificadas genéticamente: A) Una placa de cultivo celular con colonias teñidas con Crystal Violet que se han formado después de 4 semanas de selección con Puomicina; la eficiencia del establecimiento del vector fue aproximadamente del

40%; b) Detección por FACS de la fluorescencia de GFP en las células seleccionadas con Puomicina; la fluorescencia es muy homogénea y el número de células no fluorescentes es extremadamente bajo.

FIG.14: Resultado del rescate de plásmidos de los vectores pS/MARt a partir de poblaciones celulares establecidas: El ADN de las colonias bacterianas (números 1 a 12) obtenidas en un experimento de rescate de plásmidos se digirió con BamHI y se resolvió mediante electroforesis en gel de agarosa; los carriles etiquetados con "p/SMART" contienen ADN de una colonia de bacterias que llevan el plásmido original tratado de la misma manera que el anterior.

FIG. 15: Southern blot de vectores pSMART mantenidos en células seleccionadas: se utilizaron oligonucleótidos que hibridaban con el gen GFP de pS/MARt como sondas para detectar el ADN del vector restringido por BamHI en extractos de células huésped (pS/MARt1 a 3); el vector no transfectado se utilizó como control ("pS/MARt(+)").

FIG. 16: Mapa vectorial de pS/MARt; ori: origen bacteriano de replicación, P2A: secuencia que codifica el péptido 2A de auto-escisión del teschovirus-1 porcino, apolipoB MAR: Secuencia S/MAR del gen de la apolipoproteína B.

Tabla 1: Vectores marcadores de selección R6K con sitio de clonación múltiple pNTC flanqueado por el Origen-ARN-OUT.

Tabla 2: Expresión transitoria de los vectores S/MAR tras la transfección en las líneas celulares A549 y HEK293.

Tabla 3: Expresión transitoria de los vectores S/MAR tras la transfección en las líneas celulares A549 y HEK293.

Tabla 4: Expresión transitoria de los vectores S/MAR tras la transfección en las líneas celulares A549 y HEK293.

SEQ ID NO:1: Origen de R6K gamma
 SEQ ID NO:2: origen de R6K gamma de 1 CpG
 SEQ ID NO:3: CpG libre R6K origen gamma
 SEQ ID NO:4: Origen del gamma R6K ampliado
 SEQ ID NO:5: Marcador seleccionable RNA-OUT
 SEQ ID NO:6: ARN-OUT ARN represor antisentido
 SEQ ID NO:7: 2 Marcador Seleccionable CpG RNA-OUT
 SEQ ID NO:8: Origen gamma R6K-RNA-OUT
 región bacteriana flanqueada por sitios de restricción NheI y KpnI
 SEQ ID NO:9: 1 CpG R6K origen gamma-2 CpG RNA-OUT región bacteriana flanqueada por sitios de restricción NheI y KpnI
 SEQ ID NO:10: casete de clonación polienlazador pNTC-NP1 polienlazador trpA R6K-RNA-OUT: EcoRI/HindIII
 SEQ ID NO:11: casete de clonación polienlazador pNTC-NP2 polienlazador trpA R6K-RNA-OUT: EcoRI/HindIII
 SEQ ID NO:12: casete de clonación polienlazador pNTC-NP3 polienlazador trpA R6K-RNA-OUT: EcoRI/HindIII
 SEQ ID NO:13: casete de clonación polienlazador pNTC-NP4 polienlazador trpA R6K-RNA-OUT: EcoRI/HindIII
 SEQ ID NO:14: casete de clonación polienlazador pNTC-NP5 polienlazador trpA R6K-RNA-OUT: KasI/HindIII
 SEQ ID NO:15: casete de clonación polienlazador pNTC-NP6 polienlazador trpA R6K-RNA-OUT: EcoRI/SacI
 SEQ ID NO:16: casete de clonación polienlazador pNTC-NP7 polienlazador trpA R6K-RNA-OUT: BssHII/BssHII
 SEQ ID NO:17: Casete de clonación del polienlazador pNTC-3xCpG NP1 polienlazador R6K-RNA-OUT: HindIII/EcoRI
 SEQ ID NO:18: Interferón beta S/MAR humano flanqueado por el sitio 5' BglII-XhoI y los sitios 3' de la enzima de restricción EcoRI
 SEQ ID NO:19: Donante de empalme - Interferón beta humano S/MAR - aceptor de empalme flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI
 SEQ ID NO:20: Donante de empalme - Interferón beta humano S/MAR (-AATAAA) - aceptor de empalme flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI
 SEQ ID NO:21: Donante de empalme-interferón beta humano M18 S/MAR-aceptante de empalme flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI
 SEQ ID NO:22: Donante de empalme-805 pb de apolipoproteína B humana S/MAR-aceptante de empalme flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI

SEQ ID NO:23: Donante de empalme-525 pb de apolipoproteína B humana S/MAR-aceptante de empalme
 flanqueado por el sitio 5' NsiI y los sitios de la enzima de restricción 3' BamHI
 SEQ ID NO:24: intrón pCI
 SEQ ID NO:25: Donante de empalme pCI
 SEQ ID NO:26: aceptor de empalme pCI (IgG murina)
 SEQ ID NO:27: Potenciador de expresión Ele40
 SEQ ID NO:28: Potenciador de la expresión A2UCOE
 SEQ ID NO:29: Secuencia de consenso del aceptor de empalme
 SEQ ID NO: 30: motivo de la secuencia
 SEQ ID NO: 31: motivo de la secuencia
 SEQ ID NO: 32: motivo de la secuencia
 SEQ ID NO: 33: motivo de la secuencia
 SEQ ID NO: 34: motivo de la secuencia
 SEQ ID NO: 35: motivo de la secuencia
 SEQ ID NO: 36: puromicina acetiltransferasa, construcción sintética
 SEQ ID NO: 37: secuencia codificadora de la puromicina acetiltransferasa, sintética
 SEQ ID NO: 38: elemento antirrepresivo40
 SEQ ID NO: 39: Promotor del CMV - Secuencia S/MAR
 SEQ ID NO: 40: Promotor CMV - Puromicina - Secuencia S/MAR
 SEQ ID NO: 41: Elemento40-GPF-P2A-Puromicina-S/MAR

Tal y como se utiliza en lo que sigue, los términos "tener", "comprender" o "incluir" o cualquier variación gramatical arbitraria de los mismos se utilizan de forma no exclusiva. Por lo tanto, estos términos pueden referirse tanto a una situación en la que, además de la característica introducida por estos términos, no hay otras características presentes en la entidad descrita en este contexto como a una situación en la que hay una o más características adicionales. A modo de ejemplo, las expresiones "A tiene B", "A comprende B" y "A incluye B" pueden referirse tanto a una situación en la que, además de B, no hay ningún otro elemento presente en A (es decir, una situación en la que A consiste única y exclusivamente en B) como a una situación en la que, además de B, hay uno o varios elementos más presentes en la entidad A, como el elemento C, los elementos C y D o incluso otros elementos.

Además, tal como se utiliza en lo que sigue, los términos "preferiblemente", "más preferiblemente", "más preferiblemente", "particularmente", "más particularmente", "específicamente", "más específicamente" o términos similares se utilizan junto con características opcionales, sin restringir otras posibilidades. Por lo tanto, las características introducidas por estos términos son características opcionales y no pretenden restringir el alcance de las reivindicaciones de ninguna manera. La invención puede, como reconocerá el experto, realizarse utilizando características alternativas. Del mismo modo, las características introducidas por "en una realización de la invención" o expresiones similares pretenden ser características opcionales, sin ninguna restricción en cuanto a otras realizaciones de la invención, sin ninguna restricción en cuanto al alcance de la invención y sin ninguna restricción en cuanto a la posibilidad de combinar las características introducidas de tal manera con otras características opcionales o no opcionales de la invención.

Además, si no se indica lo contrario, el término "aproximadamente" se refiere al valor indicado con la precisión técnica comúnmente aceptada en el campo correspondiente, preferiblemente se refiere al valor indicado $\pm 20\%$, más preferiblemente $\pm 10\%$, más preferiblemente $\pm 5\%$. Además, el término "esencialmente" indica que las desviaciones que influyen en el resultado o el uso indicados están ausentes, es decir, que las desviaciones potenciales no hacen que el resultado indicado se desvíe en más de $\pm 20\%$, más preferiblemente $\pm 10\%$, más preferiblemente $\pm 5\%$. Por lo tanto, "constituido esencialmente por" significa que incluye los componentes especificados pero excluye otros componentes, excepto los materiales presentes como impurezas, los materiales inevitables presentes como resultado de los procesos utilizados para proporcionar los componentes y los componentes añadidos con un propósito distinto al de lograr el efecto técnico de la invención. Por ejemplo, una composición definida utilizando la frase "que consiste esencialmente en" abarca cualquier aditivo, excipiente, diluyente, portador y similares conocidos y aceptables. Preferiblemente, una composición formada esencialmente por un conjunto de componentes comprenderá menos del 5% en peso, más preferiblemente menos del 3% en peso, incluso más preferiblemente menos del 1%, más preferiblemente menos del 0,1% en peso de componente(s) no especificado(s). En el contexto de las secuencias de ácidos nucleicos, el término "esencialmente idéntico" indica un valor de % de identidad de al menos el 80%, preferiblemente al menos el 90%, más preferiblemente al menos el 98%, más preferiblemente al menos el 99%. Como se entenderá, el término esencialmente idéntico incluye el 100% de identidad. Lo anterior se aplica al término "esencialmente complementario" mutatis mutandis.

El término "polinucleótido", tal como se utiliza aquí, se refiere a una molécula de ácido nucleico lineal o circular. El término abarca tanto los polinucleótidos de cadena simple como los de cadena doble parcial o total. Preferentemente, el polinucleótido es ARN o es ADN, incluido el ADNc. Además, también se incluyen polinucleótidos modificados químicamente, incluyendo polinucleótidos modificados de forma natural, como polinucleótidos glicosilados o metilados, o derivados modificados artificialmente, como polinucleótidos biotinilados. El polinucleótido de la presente invención se proporcionará, preferentemente, como polinucleótido aislado (es decir, aislado de su contexto natural) o en forma modificada genéticamente. El polinucleótido de la invención comprende al menos un

promotor activo en una célula huésped y un elemento S/MAR; además, el polinucleótido tiene la actividad biológica de replicarse episódicamente en una célula huésped, todo ello como se especifica a continuación. Preferentemente, el polinucleótido tiene una longitud de como máximo 1 Mb, más preferentemente de como máximo 500 kb, aún más preferentemente de como máximo 200 kb, más preferentemente de como máximo 100 kb. Preferiblemente, el polinucleótido es un polinucleótido no natural; así, preferiblemente, el nucleótido es un polinucleótido artificial. También preferentemente, el polinucleótido es un polinucleótido quimérico; más preferentemente, el polinucleótido comprende al menos una secuencia de ácido nucleico heteróloga a las restantes secuencias de ácido nucleico que comprende.

Tal como se utiliza aquí, el término polinucleótido, preferentemente, incluye variantes de los polinucleótidos específicamente indicados. Más preferentemente, el término polinucleótido se refiere a los polinucleótidos específicos indicados. El término "variante de polinucleótido", tal como se utiliza en el presente documento, se refiere a una variante de un polinucleótido relacionado con el presente documento que comprende una secuencia de ácido nucleico caracterizada porque la secuencia puede derivarse de la secuencia de ácido nucleico específica antes mencionada mediante al menos una sustitución, adición y/o supresión de nucleótidos, en la que la variante de polinucleótido tendrá la actividad o actividades biológicas especificadas para el polinucleótido específico. Por lo tanto, debe entenderse que una variante de polinucleótido, tal como se menciona de acuerdo con la presente invención, tendrá una secuencia de ácido nucleico que difiere debido a al menos una sustitución, supresión y/o adición de nucleótidos. Preferentemente, dicha variante polinucleotídica comprende un ortólogo, un paralog o otro homólogo del polinucleótido específico o de una subsecuencia funcional del mismo, por ejemplo, de un elemento S/MAR. También es preferible que dicha variante del polinucleótido comprenda un alelo natural del polinucleótido específico o de una subsecuencia funcional del mismo. Las variantes de polinucleótidos también abarcan los polinucleótidos que comprenden una secuencia de ácido nucleico que es capaz de hibridarse con los polinucleótidos específicos antes mencionados o con subsecuencias funcionales de los mismos, preferiblemente, en condiciones de hibridación estrictas. Un ejemplo preferido de condiciones de hibridación estrictas son las condiciones de hibridación en cloruro de sodio/citrato de sodio 6x (= SSC) a aproximadamente 45°C, seguidas de uno o más pasos de lavado en SSC 0,2x, SDS 0,1% a 50 a 65°C. El experto sabe que estas condiciones de hibridación difieren según el tipo de ácido nucleico y, por ejemplo, cuando hay disolventes orgánicos, en cuanto a la temperatura y la concentración del tampón. Por ejemplo, en "condiciones estándar de hibridación" la temperatura difiere según el tipo de ácido nucleico entre 42°C y 58°C en un tampón acuoso con una concentración de 0,1x a 5x SSC (pH 7,2). Si hay disolvente orgánico en el tampón mencionado, por ejemplo formamida al 50%, la temperatura en condiciones estándar es de aproximadamente 42°C. Las condiciones de hibridación para los híbridos de ADN:ADN son preferentemente, por ejemplo, 0,1x SSC y 20°C a 45°C, preferentemente entre 30°C y 45°C. Las condiciones de hibridación para los híbridos de ADN:ARN son preferentemente, por ejemplo, 0,1x SSC y 30°C a 55°C, preferentemente entre 45°C y 55°C. Las temperaturas de hibridación mencionadas se determinan, por ejemplo, para un ácido nucleico con una longitud de aproximadamente 100 pb (= pares de bases) y un contenido de G+C del 50% en ausencia de formamida; en consecuencia, el experto puede considerar más adecuadas otras condiciones más adecuadas para el ADN de bajo G+C, que en principio son conocidas por el experto. El trabajador experto sabe cómo determinar las condiciones de hibridación necesarias consultando los libros de texto estándar. Alternativamente, las variantes polinucleotídicas pueden obtenerse mediante técnicas basadas en la PCR, como la amplificación del ADN basada en cebadores oligonucleotídicos mixtos, es decir, utilizando cebadores degenerados contra dominios conservados de un polipéptido de la presente invención. Los dominios conservados de un polipéptido pueden identificarse mediante una comparación de la secuencia de ácido nucleico del polinucleótido o la secuencia de aminoácidos del polipéptido de la presente invención con secuencias de otros organismos. Como plantilla, puede utilizarse ADN o ADNc de bacterias, hongos, plantas o, preferentemente, de animales. Además, las variantes incluyen polinucleótidos que comprenden secuencias de ácido nucleico que son al menos el 70%, al menos el 75%, al menos el 80%, al menos el 85%, al menos el 90%, al menos el 95%, al menos el 98% o al menos el 99% idénticas a las secuencias de ácido nucleico específicamente indicadas o a sus subsecuencias funcionales. Además, también se incluyen los polinucleótidos que comprenden secuencias de ácido nucleico que codifican secuencias de aminoácidos que son al menos un 70%, al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 98% o al menos un 99% idénticas a las secuencias de aminoácidos específicamente indicadas. Los valores de identidad porcentual se calculan, preferentemente, sobre toda la región de la secuencia de aminoácidos o ácidos nucleicos. El trabajador especializado dispone de una serie de programas basados en diversos algoritmos para comparar diferentes secuencias. En este contexto, los algoritmos de Needleman y Wunsch o Smith y Waterman dan resultados especialmente fiables. Para llevar a cabo los alineamientos de secuencias, se utilizó el programa PileUp (J. Mol. Evolution, 25, 351-360, 1987, Higgins et al., CABIOS, 5 1989: 151-153) o los programas Gap y BestFit (Needleman y Wunsch (J. Mol. Biol. 48; 443-453 (1970)) y Smith y Waterman (Adv. Appl. Matemáticas. 2; 482-489 (1981)), se utilizan preferentemente. Preferiblemente, dichos programas se utilizan con sus parámetros estándar. Los valores de identidad de la secuencia mencionados anteriormente en porcentaje (%) deben determinarse, preferentemente, utilizando el programa GAP en toda la región de la secuencia con los siguientes ajustes: Peso de la brecha: 50, Peso de la longitud: 3, Partido Medio: 10.000 y el desajuste medio: 0.000, que, a menos que se especifique lo contrario, se utilizarán siempre como ajustes estándar para las alineaciones de secuencias.

Un polinucleótido que comprende un fragmento de cualquiera de las secuencias de ácido nucleico específicamente indicadas, dicho polinucleótido reteniendo la actividad o actividades indicadas, también se engloba como un

polinucleótido variante de la presente invención. Un fragmento, tal como se entiende en el presente documento, comprende preferentemente al menos 200, preferentemente al menos 300, más preferentemente al menos 400 nucleótidos consecutivos de cualquiera de las secuencias específicas de ácido nucleico; o codifica una secuencia de aminoácidos que comprende al menos 100, preferentemente al menos 200, más preferentemente al menos 300 aminoácidos consecutivos de cualquiera de las secuencias específicas de aminoácidos y que sigue teniendo la actividad indicada.

Los polinucleótidos de la presente invención consisten, consisten esencialmente en, o comprenden las secuencias de ácido nucleico antes mencionadas. Por lo tanto, también pueden contener otras secuencias de ácido nucleico. En concreto, los polinucleótidos de la presente invención pueden codificar, por ejemplo, proteínas de fusión o marcadores seleccionables. Dichas proteínas de fusión pueden comprender como parte adicional polipéptidos para controlar la expresión (por ejemplo, proteínas fluorescentes verdes, amarillas, azules o rojas, fosfatasa alcalina y similares) o las llamadas "etiquetas" que pueden servir como marcador detectable o como medida auxiliar para fines de purificación. Las etiquetas para los diferentes propósitos son bien conocidas en el arte y se describen en otra parte del presente documento.

También preferentemente, el polinucleótido comprende al menos una secuencia de carga. El término "secuencia de carga", tal y como se utiliza aquí, se refiere a una secuencia de ácido nucleico de interés para ser transferida y mantenida de forma estable en una célula huésped. Preferentemente, la secuencia de carga es una secuencia de ácido nucleico que codifica un polinucleótido, por ejemplo un ARN, y/o un polipéptido de interés. Preferentemente, el polipéptido de interés es un polipéptido terapéutico, más preferentemente un receptor de células T (TCR), más preferentemente un receptor de células T humano o quimérico, un receptor de antígeno quimérico (CAR), preferentemente MART1 TCR, o un polipéptido que falta en las células afectadas con una enfermedad genética como se especifica en otra parte del presente documento. Así, por ejemplo, preferentemente, el polinucleótido comprende al menos una secuencia de carga que codifica un polipéptido que proporciona actividad de fenilalanina-hidroxilasa (EC 1.14.16.1) para el tratamiento de la fenilcetonuria.

Preferiblemente, la secuencia que codifica un marcador seleccionable y la secuencia de carga están intervenidas por una secuencia que permite la expresión de dos (o más) polipéptidos en una célula eucariota a partir de un ARNm, por ejemplo, una secuencia de entrada ribosómica interna (IRES) o, más preferiblemente, una secuencia peptídica de auto-escisión, como, más preferiblemente, una secuencia de péptido 2A (P2A) del teschovirus-1 porcino. Las secuencias apropiadas son conocidas en la técnica, por ejemplo, de Kim et al. (2011) PLoS ONE 6(4): e18556.

Preferiblemente, el polinucleótido es un ADN. Preferentemente, el polinucleótido comprende otras secuencias de control de expresión que permiten la expresión de los genes en procariotas y/o eucariotas, preferentemente en células huésped eucariotas o en fracciones aisladas de las mismas. La expresión de dicho polinucleótido comprende la transcripción del polinucleótido, preferentemente en un ARNm traducible. Los elementos reguladores que garantizan la expresión en células eucariotas, preferentemente de mamíferos, son bien conocidos en la técnica. Preferentemente, comprenden secuencias reguladoras que aseguran el inicio de la transcripción y, opcionalmente, señales poli-A que aseguran la terminación de la transcripción y la estabilización del transcrito. Los elementos reguladores adicionales pueden incluir potenciadores transcripcionales y traslacionales. Ejemplos de elementos reguladores que permiten la expresión en células huésped eucariotas son el promotor AOX1 o GAL1 en la levadura o el promotor SMVP-, U6-, H1-, 7SK-, CMV- EFS-, SV40-, o RSV (virus del sarcoma de Rous), potenciador CMV, potenciador SV40 o un intrón de globina en células de mamíferos y otros animales. Además, las secuencias de control de expresión inducibles o específicas del tipo de célula pueden estar comprendidas en un polinucleótido de la presente invención. Las secuencias de control de la expresión inducible pueden comprender secuencias operadoras tet o lac o secuencias inducibles por choque térmico u otros factores ambientales. Las secuencias de control de expresión adecuadas son bien conocidas en la técnica. Además de los elementos responsables de la iniciación de la transcripción, dichos elementos reguladores también pueden comprender señales de terminación de la transcripción, como el sitio SV40-poly-A o el sitio tk-poly-A, aguas abajo del polinucleótido.

El término "célula huésped", tal como se utiliza aquí, se refiere a cualquier célula capaz de recibir y replicar de forma estable el polinucleótido. Preferiblemente, la célula huésped es una célula eucariota, preferiblemente una célula vegetal o de levadura, por ejemplo, una célula de una cepa de levadura de panadería, o es una célula animal. Más preferentemente, la célula huésped es una célula de insecto o una célula de mamífero, en particular una célula de ratón o de rata. Aún más preferentemente, la célula huésped es una célula de mamífero, más preferentemente es una célula humana. Preferiblemente, la célula huésped es una célula progenitora CD34+; un trombocito CD61+; un linfocito B CD19+; un monocito CD14+; un granulocito CD15+; un linfocito T citotóxico CD3+, preferiblemente también positivo para CD8 y CD45 un linfocito T ayudante CD3+, preferiblemente también positivo para CD4 y CD45; un linfocito T activado CD3+, preferiblemente también positivo para CD25 y CD45, un linfocito infiltrante de tumores o una célula asesina natural (NK). Como comprenderá el experto, el polinucleótido puede tener además secuencias que permitan la replicación en una célula bacteriana, en particular un origen de replicación bacteriano. Preferiblemente, la célula bacteriana es una célula de una cepa bacteriana de laboratorio, más preferiblemente una célula de *Escherichia coli*.

El término "promotor" es, en principio, conocido por el experto como un elemento genético que dirige, opcionalmente en concierto con otros elementos reguladores, el nivel de transcripción de un gen determinado. Un promotor puede

ser constitutivo, es decir, proporcionar un nivel constante de transcripción esencialmente independiente del estado de una célula huésped, o puede ser regulado, es decir, proporcionar niveles de transcripción en dependencia del estado de una célula huésped. Además, un promotor puede ser específico del tipo de célula y/o del tejido, es decir, proporcionar un nivel detectable de transcripción sólo en unos pocos o en un solo tipo de célula. Preferentemente, el promotor según la presente invención es activo en la célula huésped como se ha especificado anteriormente. Como comprenderá el experto, la selección del promotor puede depender del tipo de célula huésped a la que se dirige; se conocen en la técnica promotores adecuados para tipos celulares específicos, así como promotores constitutivos. Preferentemente, el promotor es un promotor eucariota, más preferentemente un promotor eucariota constitutivo, más aún un promotor eucariota fuerte. Preferiblemente, el promotor es un promotor EF1alfa (factor de elongación 1 alfa), un promotor UbiC (ubiquitina C), un promotor ROSA 26, un promotor PGK (fosfoglicerato quinasa), y/o un promotor CAG (alfa-actina de pollo), más preferiblemente es un promotor EF1alfa. También es preferible que el promotor sea un promotor eucariota específico de célula y/o tejido. En el presente documento, el término "promotor" se utiliza para el promotor especificado anteriormente, mientras que cualquier otro promotor potencialmente presente en el polinucleótido se denomina "promotor secundario". Así, preferentemente, el promotor es un promotor que dirige la transcripción hacia la secuencia S/MAR en una célula huésped; también preferentemente, un promotor que no dirige la transcripción hacia la secuencia S/MAR del polinucleótido, por ejemplo, por ser un promotor procariota, por estar aislado transcripcionalmente de la secuencia S/MAR, y/o por ser un promotor que dirige la transcripción lejos de la secuencia S/MAR, es un promotor secundario. Preferentemente, el promotor comprende menos de 1000, más preferentemente menos de 250, aún más preferentemente menos de 100, más preferentemente menos de 20 pares de bases contiguas correspondientes a un promotor de la Apolipoproteína B; así, preferentemente, el polinucleótido no comprende un promotor de la Apolipoproteína B humana, más preferentemente no comprende un promotor de la Apolipoproteína B.

Preferiblemente, la secuencia S/MAR está situada inmediatamente aguas abajo del promotor y, si está presente, del gen marcador seleccionable como se especifica a continuación. Preferiblemente, al estar situado "inmediatamente aguas abajo" carece de una señal de terminación de la transcripción intermedia, más preferiblemente carece de un gen intermedio. Así, preferentemente, los transcritos iniciados en el promotor y, si están codificados, incluyendo la secuencia marcadora detectable comprenden preferentemente una secuencia S/MAR transcrita, más preferentemente comprenden la secuencia S/MAR completa comprendida en el polinucleótido; como comprenderá el experto a la vista de la descripción en otra parte del presente documento, el polinucleótido puede incluir además sitios de empalme que medien la escisión de la secuencia S/MAR del transcrito primario; así, más preferentemente, al menos los transcritos primarios iniciados en el promotor y, si están codificados, incluyendo la secuencia marcadora detectable comprenden preferentemente una secuencia S/MAR transcrita, más preferentemente comprenden la secuencia S/MAR completa comprendida en el polinucleótido. También es preferible que el término "inmediatamente después" incluya un polinucleótido en el que el promotor y la secuencia S/MAR estén separados por secuencias de ácido nucleico alargadas, siempre que no intervenga una señal de terminación de la transcripción entre el promotor y la S/MAR.

Preferiblemente, la secuencia que interviene el promotor o, si está presente, el codón de parada del gen marcador seleccionable y la secuencia S/MAR tiene una longitud de como máximo 2 kb, más preferiblemente como máximo 0,5 kb, incluso más preferiblemente como máximo 0,2 kb, aún más preferiblemente como máximo 0,1 kb, más preferiblemente como máximo 50 pb.

El término "elemento S/MAR", también conocido bajo la designación de "región de unión de andamio/matriz", es, en principio, conocido por el experto para referirse a una secuencia de ADN que media la unión de la matriz nuclear de una célula eucariota a dicho ADN. Las secuencias S/MAR suelen proceder de secuencias del ADN de los cromosomas eucariotas. Existe una gran variedad de secuencias de S/MAR, y las secuencias están disponibles en bases de datos públicas, por ejemplo, como se describe en Liebich et al. (2002), Nucleic Acids Res. 30, 312-374. Según la presente invención, la secuencia de ácido nucleico de dicho elemento S/MAR (hasta ahora denominada secuencia S/MAR) comprende al menos 3 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de como máximo 200 nucleótidos. Así, el motivo comprendido en la secuencia S/MAR comprende una multitud del motivo de cuatro nucleótidos 5'-ATTA-3'. Preferentemente, la secuencia S/MAR tiene una longitud de al menos 200 nucleótidos, más preferentemente de al menos 300 nucleótidos, aún más preferentemente de al menos 400 nucleótidos, más preferentemente de al menos 500 nucleótidos. Preferentemente, la secuencia S/MAR tiene una longitud de como máximo 3 kb, más preferentemente de como máximo 2 kb, aún más preferentemente de como máximo 1,5 kb, aún más preferentemente de como máximo 1 kb, más preferentemente de como máximo 0,9 kb. Así, preferentemente, la secuencia S/MAR tiene una longitud de 0,2 kb a 3 kb, más preferentemente de 0,3 kb a 2 kb, aún más preferentemente de 0,4 kb a 1,5 kb, más preferentemente de 0,5 kb a 1 kb. Como se comprenderá, la indicación "comprende n motivos de secuencia por cada 100 nucleótidos" se refiere al número medio de dichos motivos de secuencia calculado por cada 100 pares de bases de secuencia y, en consecuencia, puede ser un número fraccionario. Por ejemplo, el número de motivos de la secuencia ATTA por cada 100 pares de bases en la SEQ ID NO:6 es de $34 / 525 \text{ pares de bases} \times 100 \text{ pares de bases} = 6,5$. Preferentemente, el número de motivos de secuencia por 100 pares de bases se determina a lo largo de toda la secuencia S/MAR; en caso de duda, por ejemplo, cuando no se puede determinar un límite de la secuencia S/MAR, el número de motivos de secuencia por 100 pares de bases de un polinucleótido, preferiblemente, es el número más alto determinable para cualquier ventana de 200 pb dentro de dicho polinucleótido, más preferiblemente es el número más alto determinable para

cualquier ventana de 500 pb dentro de dicho polinucleótido. Preferentemente, la secuencia S/MAR comprende al menos 4 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de como máximo 200 nucleótidos, más preferentemente al menos 5 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de como máximo 200 nucleótidos, y aún más preferentemente al menos 6 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de como máximo 200 nucleótidos. También es preferible que la secuencia S/MAR comprenda al menos 3 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 400 nucleótidos como máximo, más preferiblemente al menos 4 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 400 nucleótidos como máximo, incluso más preferiblemente al menos 5 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 400 nucleótidos como máximo, más preferiblemente al menos 6 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 400 nucleótidos como máximo. También es preferible que la secuencia S/MAR comprenda al menos 3 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 500 nucleótidos como máximo, más preferiblemente al menos 4 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 500 nucleótidos como máximo, aún más preferiblemente al menos 5 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 500 nucleótidos como máximo, más preferiblemente al menos 6 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de 500 nucleótidos como máximo. Así, preferentemente, la secuencia S/MAR comprende al menos 10 motivos de secuencia ATTA sobre una secuencia de 500 nucleótidos, más preferentemente al menos 20 motivos de secuencia ATTA sobre una secuencia de 500 nucleótidos, aún más preferentemente al menos 30 motivos de secuencia ATTA sobre una secuencia de 500 nucleótidos. Preferiblemente, al menos el 80%, más preferiblemente al menos el 90%, más preferiblemente al menos el 95% de los motivos ATTA en la secuencia S/MAR están separados por de 9 a 13, preferiblemente por 10 a 12, más preferiblemente por 11 pares de bases, respectivamente.

Preferentemente, el elemento S/MAR comprende motivos de secuencia adicionales, preferentemente dentro de la secuencia que comprende los motivos ATTA descritos anteriormente. Preferiblemente, el tramo de secuencia de dicho elemento S/MAR que comprende dichos motivos de secuencia ATTA comprende además al menos un motivo de secuencia ATTTA, preferiblemente al menos 2 motivos de secuencia ATTTA, más preferiblemente al menos 4 motivos de secuencia ATTTA, más preferiblemente al menos 8 motivos de secuencia ATTTA. También preferiblemente, el tramo de secuencia de dicho elemento S/MAR que comprende dichos motivos de secuencia ATTA y, opcionalmente, dicho(s) motivo(s) ATTTA, comprende además al menos uno, preferiblemente al menos dos, más preferiblemente al menos cuatro, más preferiblemente al menos seis motivos palindrómicos, preferiblemente motivos TAAATATTTTA (SEQ ID NO:30). Preferentemente, dichos motivos TAAATATTTTA son contiguos con al menos un motivo ATTA en el extremo 5' y/o en el extremo 3'. También preferiblemente, el tramo de secuencia del elemento S/MAR que comprende dichos motivos de secuencia ATTA comprende al menos uno, preferiblemente al menos dos, más preferiblemente al menos tres, aún más preferiblemente al menos cuatro, más preferiblemente al menos cinco motivos de secuencia ATTAAATATTTTAATTA (SEQ ID NO:31), más preferiblemente motivos de secuencia ATTTAAATATTTTAATTA (SEQ ID NO:32).

También es preferible que la secuencia S/MAR tenga un bajo contenido de G+C. El experto sabe cómo calcular el contenido de C+G de una secuencia conocida contando todas las bases de guanina y citidina de la secuencia y dividiendo el resultado acumulado por el número de nucleótidos de la secuencia. Preferentemente, el tramo de secuencia del elemento S/MAR que comprende dichos motivos de secuencia ATTA tiene un contenido de G+C de como máximo el 30%, más preferentemente de como máximo el 20%, aún más preferentemente de como máximo el 15%, aún más preferentemente de como máximo el 10%, más preferentemente de como máximo el 5%. Preferiblemente, en los casos en los que no se pueda determinar el límite de un elemento S/MAR, la secuencia utilizada para el cálculo del contenido de G+C es la misma que se utiliza para el cálculo del número de motivos ATTA por cada 100 pares de bases, tal y como se ha especificado anteriormente. También es preferible que la secuencia S/MAR tenga un bajo número de dinucleótidos CG. Preferentemente, el tramo de secuencia de dicho elemento S/MAR que comprende dichos motivos de secuencia comprende a lo sumo 6 motivos de secuencia CG, más preferentemente a lo sumo 4, aún más preferentemente a lo sumo 2, más preferentemente no comprende un motivo de secuencia CG.

Preferiblemente, la secuencia S/MAR comprende una secuencia S/MAR de un gen de apolipoproteína B, preferiblemente un gen de apolipoproteína B humana, más preferiblemente una secuencia S/MAR 3' de un gen de apolipoproteína B humana. Más preferentemente, la secuencia S/MAR comprende una variante de un gen de la apolipoproteína B humana, más preferentemente de una secuencia 3' S/MAR de un gen de la apolipoproteína B humana. Así, preferentemente, la secuencia S/MAR comprende una secuencia al menos 70% idéntica a la secuencia de la SEQ ID NO:33, preferentemente de la SEQ ID NO:34 o 35. Más preferentemente, la secuencia S/MAR comprende la secuencia de ácido nucleico de la SEQ ID NO:33, preferentemente de la SEQ ID NO:34, más preferentemente de la SEQ ID NO:35.

Preferentemente, el polinucleótido comprende una señal poli-A aguas abajo del elemento S/MAR: Más preferentemente, el polinucleótido comprende una señal de poli-A y una señal de terminación de la transcripción a continuación del elemento S/MAR. También es preferible que el elemento S/MAR esté flanqueado por un donante de empalme y un aceptor de empalme; así, preferiblemente, la secuencia S/MAR se empalma fuera del transcrito que codifica el marcador seleccionable después de la transcripción. También es preferible que el polinucleótido comprenda además un origen de replicación bacteriano (secundario) como se especifica en el presente documento

y/o un gen marcador seleccionable bacteriano. Preferiblemente, el origen de replicación bacteriano y el promotor que impulsa la expresión del gen marcador seleccionable bacteriano son procariotas específicos, es decir, más preferiblemente, no son funcionales en una célula huésped. También es preferible que el origen de replicación bacteriano y/o el gen marcador seleccionable bacteriano, preferiblemente todos los elementos activos en una célula procariota comprendidos en el polinucleótido, estén aislados de las secuencias residuales comprendidas en el polinucleótido por la presencia de al menos un elemento aislante, más preferiblemente por estar flanqueados por elementos aislantes. preferentemente, el origen de replicación bacteriano y/o el gen marcador seleccionable bacteriano, preferentemente todos los elementos activos en una célula procariota comprendidos en el polinucleótido, están aislados del promotor por la presencia de al menos un elemento aislante, más preferentemente por estar flanqueados por elementos aislantes. Preferentemente, dicho(s) elemento(s) aislante(s) es(son) un elemento antirrepresivo⁴⁰ (SEQ ID NO:38) o una variante del mismo y/o un elemento S/MAR.

Así, preferentemente, el polinucleótido comprende la secuencia de SEQ ID NO:34 o 35 o de una secuencia al menos 70% idéntica a la secuencia de SEQ ID NO:34 o 35; preferentemente de SEQ ID NO:39 o de una secuencia al menos 70% idéntica a la secuencia de SEQ ID NO:39, más preferentemente de la SEQ ID NO:40 o de una secuencia al menos 70% idéntica a la secuencia de la SEQ ID NO:40, más preferentemente de la SEQ ID NO:40 o de una secuencia al menos 70% idéntica a la secuencia de la SEQ ID NO:41. Preferentemente, el polinucleótido comprende la secuencia de SEQ ID NO:41 con la secuencia de ácido nucleico que codifica la GFP sustituida por una secuencia de ácido nucleico que codifica un polipéptido diferente, preferentemente un polipéptido terapéutico, más preferentemente el Receptor de Células T (TCR) humano, el Receptor de Antígeno Químico (CAR), preferentemente el TCR MART1.

Preferiblemente, el polinucleótido comprende además una secuencia codificadora que codifica un polipéptido marcador seleccionable, dicha secuencia marcadora seleccionable que interviene en el promotor del polinucleótido y en el elemento S/MAR del polinucleótido, preferiblemente en el que dicho promotor y dicha secuencia marcadora seleccionable constituyen conjuntamente un gen marcador seleccionable. Tal y como se utiliza aquí, el término "secuencia marcadora seleccionable" se utiliza como una abreviatura de la expresión "secuencia codificadora que codifica un polipéptido marcador seleccionable". El término "marcador seleccionable" es en principio entendido por el experto y se refiere a una secuencia de ácido nucleico que confiere, cuando se expresa en una célula huésped, resistencia a al menos una condición que media la presión selectiva a una célula huésped cuando se aplica a la misma. Los marcadores seleccionables son conocidos en la técnica para células procariotas y eucariotas. Preferiblemente, el marcador seleccionable es un marcador seleccionable de una célula eucariota. Preferentemente, el marcador seleccionable es un polipéptido marcador seleccionable, más preferentemente un polipéptido marcador seleccionable que tiene actividad transportadora y/o enzimática que elimina un compuesto selectivo de una célula caliente o que modifica dicho compuesto selectivo para hacerlo inactivo. Preferentemente, el gen marcador seleccionable codifica además al menos un intrón, preferentemente aguas arriba de la secuencia que codifica el polipéptido marcador seleccionable. Preferentemente, el marcador seleccionable es un marcador que media la resistencia a la puomicina, a la blasticidina, a la neomicina y/o a la zeocina, más preferentemente a la puomicina. Así, preferentemente, el promotor y el marcador seleccionable constituyen un gen de resistencia a la puomicina, un gen de resistencia a la blasticidina, un gen de resistencia a la neomicina o un gen de resistencia a la zeocina, más preferentemente un gen de resistencia a la puomicina. Preferentemente, el gen marcador seleccionable carece de una señal poli-A y de señal(es) de terminación de la transcripción.

Preferiblemente, el marcador seleccionable es la puomicina acetiltransferasa (Genbank Acc No. KX548903.1 (SEQ ID NO:36), codificada por los nucleótidos 535 a 1134 del Genbank Acc No. Así, el gen marcador seleccionable, preferentemente, comprende una secuencia de ácido nucleico que a) causa la expresión de un polipéptido de resistencia a la puomicina que comprende la secuencia de SEQ ID NO:36; b) causa la expresión de un polipéptido de resistencia a la puomicina que comprende una secuencia al menos 70% idéntica a la secuencia de SEQ ID NO:36; c) comprende la secuencia de SEQ ID NO:37; d) comprende una secuencia al menos 70% idéntica a la secuencia de SEQ ID NO:37, e) comprende una secuencia de ácido nucleico que codifica un polipéptido de resistencia a la puomicina que comprende, preferentemente, la secuencia de SEQ ID NO:36, y/o f) comprende una secuencia de ácido nucleico que codifica un polipéptido de resistencia a la puomicina que comprende, preferentemente, una secuencia al menos un 70% idéntica a la secuencia de SEQ ID NO:36.

Tal como se utiliza aquí, el término "replicante" se refiere a la actividad del polinucleótido para inducir la producción de al menos dos réplicas de dicho polinucleótido en una célula huésped durante un ciclo de replicación celular. Así, preferentemente, la replicación de un polinucleótido en una célula huésped se determina determinando la presencia del polinucleótido después de una serie de divisiones celulares, en las que se habría esperado que se diluyera un polinucleótido no replicante. Preferiblemente, la replicación es estable, es decir, es una replicación hasta tal punto que el polinucleótido sigue siendo detectable en una población de células huésped después de una media de 50 divisiones celulares, más preferiblemente después de una media de 100 divisiones celulares, más preferiblemente

después de una media de 250 divisiones celulares. Preferiblemente, la detección de un polinucleótido en una población de células huésped se realiza mediante PCR en condiciones estándar.

El término replicación "episomal" es, en principio, conocido por el experto para referirse a la replicación de un polinucleótido sin integrarse en el genoma celular, es decir, sin unirse covalentemente al genoma celular. Así, preferentemente, la replicación episomal de un polinucleótido es la replicación de dicho polinucleótido como una unidad de replicación autónoma. Preferentemente, la replicación episomal consiste en el mantenimiento del polinucleótido en la célula huésped en forma de una molécula de ADN de doble cadena cerrada circularmente. Como comprenderá el experto, la replicación real de dicho polinucleótido puede implicar otras formas, por ejemplo, en la replicación en círculo rodante. El mantenimiento episomal del ADN circular se verifica preferentemente mediante el procedimiento de rescate de plásmidos conocido por el experto; es decir, preferentemente, preparando un lisado de células huésped y transformando el ADN que contiene en células bacterianas apropiadas, por ejemplo, células *E. coli*; si un número adecuado de colonias bacterianas obtenidas por dicho procedimiento comprende el ADN circular como un plásmido que tiene el mismo patrón de restricción y/o secuencia que el ADN circular original, se asume preferentemente que el ADN circular se mantuvo episomalmente. Otro procedimiento para verificar el mantenimiento episomal, que también es conocido por el experto, es el ADN/blot de ADN (procedimiento "Southern Blot"); así, preferentemente, se prepara el ADN total de las células huésped y se digiere con una o más enzimas de restricción; si en un Southern Blot utilizando el plásmido original como sonda sólo son visibles las bandas correspondientes al ADN circular original, se concluye preferentemente que el plásmido se mantiene episomalmente. Más preferentemente, el mantenimiento episomal se verifica como se describe aquí en los Ejemplos.

De acuerdo, el término "replicar episomalmente", como se usa aquí, se refiere a la actividad de un polinucleótido para inducir la producción de al menos dos réplicas de dicho polinucleótido en una célula huésped durante un ciclo de replicación celular mientras dicho polinucleótido está presente en dicha célula como una entidad que se replica autónomamente; y la replicación episomal estable es la replicación episomal hasta tal punto que el polinucleótido es todavía detectable en la célula huésped después de al menos 50 divisiones celulares, preferiblemente después de al menos 100 divisiones celulares, más preferiblemente, después de al menos 250 divisiones celulares, más preferiblemente, después de al menos 500 divisiones celulares. Preferentemente, dicho número de divisiones celulares es el número medio de divisiones celulares para una población de células.

El polinucleótido de la presente invención preferentemente está desprovisto de un origen de replicación del virus simio 40 (SV40), un origen de replicación del papilomavirus bovino (BPV) y un origen de replicación del virus de Epstein-Barr (EBV), preferentemente está desprovisto de un origen de replicación de un poliomavirus, un origen de replicación de un papilomavirus y un origen de replicación de un herpesvirus; más preferentemente está desprovisto de un origen de replicación de un virus infectante de eucariotas. Más preferentemente, el vector carece de cualquier origen de replicación eucariota conocido. Sin embargo, preferentemente, el polinucleótido comprende además un origen de replicación procariota, preferentemente bacteriano, en particular un origen de replicación *E. coli*. Preferiblemente, el origen de replicación procariota es el único origen de replicación comprendido en el polinucleótido.

Ventajosamente, se comprobó en el trabajo subyacente a la presente invención que al combinar un elemento S/MAR como se especifica con un promotor que lee en dicho elemento S/MAR, se obtiene un polinucleótido que es altamente estable en forma episomal en las células huésped, incluso en ausencia de un origen de replicación dedicado. Además, se descubrió que la eficacia del establecimiento del polinucleótido podía mejorarse aún más utilizando un gen de resistencia a la puomicina, asegurando la transcripción en el elemento S/MAR a través del gen de resistencia, y aislando transcripcionalmente la combinación promotor - S/MAR de otros promotores potencialmente presentes en el polinucleótido.

Las definiciones anteriores se aplican, mutatis mutandis, a lo siguiente. Las definiciones y explicaciones adicionales que se hacen más adelante también se aplican a todas las realizaciones descritas en esta especificación mutatis mutandis.

La presente invención se refiere además a una composición que comprende un polinucleótido según la presente invención.

El término "composición", tal y como se utiliza aquí, se refiere a una composición de materia que comprende los compuestos tal y como se especifica y opcionalmente uno o más portadores aceptables. Preferiblemente, la composición es una composición farmacéuticamente aceptable; así, preferiblemente, el portador es un portador farmacéuticamente aceptable. Los compuestos de la presente invención pueden formularse como sales, preferiblemente farmacéuticamente aceptables. Las sales preferidas comprenden acetato, metiléster, HCl, sulfato, cloruro y similares.

El (los) portador(es) debe(n) ser aceptable(s) en el sentido de que sea(n) compatible(s) con los demás ingredientes de la formulación y no sea(n) perjudicial(es) para el receptor de la misma. El portador empleado puede ser, por ejemplo, un sólido, un gel o un líquido. Ejemplos de portadores farmacéuticos sólidos son la lactosa, la terra alba, la sacarosa, el talco, la gelatina, el agar, la pectina, la acacia, el estearato de magnesio, el ácido esteárico y similares. Ejemplos de portadores líquidos son la solución salina tamponada con fosfato, el jarabe, el aceite, como el aceite de

cacahuete y el aceite de oliva, el agua, las emulsiones, varios tipos de agentes humectantes, las soluciones estériles y similares. Del mismo modo, el portador o diluyente puede incluir material de retardo bien conocido en la técnica, como el monoestearato de glicerilo o el diestearato de glicerilo solos o con una cera. Dichos portadores adecuados comprenden los mencionados anteriormente y otros bien conocidos en la técnica, véase, por ejemplo Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pennsylvania. El o los diluyentes se seleccionan de manera que no afecten a la actividad biológica de los compuestos de la composición. Ejemplos de tales diluyentes son el agua destilada, la solución salina fisiológica, las soluciones de Ringer, la solución de dextrosa y la solución de Hank. Además, la composición o formulación farmacéutica también puede incluir otros portadores, adyuvantes o estabilizadores no tóxicos, no terapéuticos, no inmunogénicos y similares.

Preferiblemente, la composición media la entrada del polinucleótido en una célula huésped. Así, preferentemente la composición comprende al menos un agente de transfección. La selección de un agente de transfección adecuado puede depender de la célula huésped diana, así como de la aplicación específica prevista. Los agentes de transfección, las condiciones de transfección adecuadas y los criterios de selección de los mismos son bien conocidos en la técnica. También es preferible que la composición comprenda partículas similares a los virus. Así, preferentemente, el polinucleótido se empaqueta en partículas similares a un virus, es decir, preferentemente, el polinucleótido está comprendido en las partículas similares a un virus.

Las composiciones farmacéuticas se administran, preferentemente, por vía tópica o sistémica. Las vías de administración adecuadas utilizadas convencionalmente para la administración de fármacos son la oral, la intravenosa o la parenteral, así como la inhalación. Sin embargo, dependiendo de la naturaleza y el modo de acción de un compuesto, las composiciones farmacéuticas pueden administrarse también por otras vías. Por ejemplo, los compuestos polinucleotídicos pueden administrarse en un enfoque de terapia génica mediante el uso de vectores virales o virus o liposomas, como se especifica en el presente documento. Además, los compuestos pueden administrarse en combinación con otros fármacos, ya sea en una composición farmacéutica común o como composiciones farmacéuticas separadas, en las que dichas composiciones farmacéuticas separadas pueden suministrarse en forma de kit de piezas. Los compuestos se administran, preferentemente, en formas farmacéuticas convencionales preparadas mediante la combinación de los fármacos con portadores farmacéuticos estándar según procedimientos convencionales. Estos procedimientos pueden consistir en mezclar, granular y comprimir o disolver los ingredientes según convenga a la preparación deseada. Se apreciará que la forma y el carácter del portador o diluyente farmacéuticamente aceptable están dictados por la cantidad de ingrediente activo con el que se va a combinar, la vía de administración y otras variables conocidas.

Una dosis terapéuticamente eficaz de una composición farmacéutica se refiere a una cantidad de los compuestos que se utilizan en una composición farmacéutica de la presente invención que previene, mejora o trata los síntomas que acompañan a una enfermedad o afección a la que se hace referencia en esta especificación. La eficacia terapéutica y la toxicidad de dichos compuestos pueden determinarse mediante procedimientos farmacéuticos estándar en cultivos celulares o animales de experimentación, por ejemplo, la ED₅₀ (la dosis terapéuticamente eficaz en el 50% de la población) y la LD₅₀ (la dosis letal para el 50% de la población). La relación de dosis entre los efectos terapéuticos y los tóxicos es el índice terapéutico, y puede expresarse como la relación, LD₅₀/ED₅₀.

El régimen de dosificación será determinado por el médico tratante y otros factores clínicos; preferentemente de acuerdo con cualquiera de los procedimientos descritos anteriormente. Como es bien sabido en las artes médicas, las dosis para cualquier paciente dependen de muchos factores, incluyendo el tamaño del paciente, la superficie corporal, la edad, el compuesto particular que se va a administrar, el sexo, el tiempo y la vía de administración, la salud general y otros medicamentos que se administran simultáneamente. Los progresos pueden ser controlados mediante evaluaciones periódicas. Una dosis típica puede estar, por ejemplo, en el rango de 1 a 1000 µg; sin embargo, se prevén dosis por debajo o por encima de este rango ejemplar, especialmente teniendo en cuenta los factores mencionados. Generalmente, el régimen de administración regular de la composición farmacéutica debe estar en el rango de 1 µg a 10 mg de unidades por día. Si el régimen es una infusión continua, también debe estar en el rango de 1 µg a 10 mg de unidades por kilogramo de peso corporal por minuto, respectivamente. Los progresos pueden ser controlados mediante evaluaciones periódicas. Sin embargo, dependiendo del sujeto y del modo de administración, la cantidad de administración de la sustancia puede variar en un amplio rango para proporcionar desde aproximadamente 0,01 mg por kg de masa corporal hasta aproximadamente 10 mg por kg de masa corporal. En caso de que se administre un vector viral, en particular un vector viral adeno-asociado, las dosis preferidas son de 5 x 10¹¹, a 2 x 10¹³ partículas virales o genomas virales / kg de peso corporal; como se entenderá, estas dosis ejemplares pueden modificarse dependiendo, además de los factores descritos anteriormente, de factores adicionales como el tipo de virus, el órgano diana, y similares.

Las composiciones y formulaciones farmacéuticas a las que se hace referencia en el presente documento se administran al menos una vez para tratar, mejorar o prevenir una enfermedad o afección mencionada en esta especificación. Sin embargo, dichas composiciones farmacéuticas pueden administrarse más de una vez, por ejemplo de una a cuatro veces al día hasta un número no limitado de días.

Las composiciones farmacéuticas específicas se preparan de una manera bien conocida en el arte farmacéutico y comprenden al menos un compuesto activo mencionado en el presente documento en mezcla o asociado de otro modo con un portador o diluyente farmacéuticamente aceptable. Para la elaboración de estas composiciones

farmacéuticas específicas, el o los compuestos activos se mezclarán normalmente con un portador o el diluyente, o se encerrarán o encapsularán en una cápsula, bolsita, caché, papel u otros contenedores o vehículos adecuados. Las formulaciones resultantes se adoptarán al modo de administración, es decir, en forma de comprimidos, cápsulas, supositorios, soluciones, suspensiones o similares. Las recomendaciones de dosificación se indicarán en las instrucciones de los prescriptores o de los usuarios con el fin de anticipar los ajustes de dosis en función del receptor considerado.

La presente invención también se refiere a un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para su uso en medicina. La presente invención se refiere además a un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para su uso en el tratamiento de enfermedades genéticas.

El término "enfermedad genética", tal como se utiliza aquí, se refiere a una enfermedad vinculada causalmente a una o más modificaciones, preferentemente mutaciones, en el genoma de un individuo. Así, preferentemente, la enfermedad genética está vinculada causalmente a uno o más cambios epigenéticos, más preferentemente está vinculada causalmente a una o más mutaciones genéticas. Como se comprenderá, los síntomas de una enfermedad genética suelen estar causados por la expresión de un gen mutado y/o la falta de expresión de un gen que proporciona la función normal del producto génico en uno o más tejidos y/o tipos de células específicos. Así, puede ser preferible tratar la enfermedad genética sólo en aquellas células en las que la mutación contribuye a la enfermedad. Preferiblemente, la enfermedad genética es una enfermedad monogénica, es decir, está causada por una alteración genética en un solo gen. Más preferentemente, la enfermedad genética es una enfermedad recesiva monogénica, es decir, está causada por alteraciones genéticas en ambos alelos de un gen; por lo tanto, preferentemente, se espera la mejora de los síntomas mediante la provisión de al menos una copia no alterada del gen afectado. Más preferentemente, la enfermedad genética es la fenilcetonuria, la alcaptonuria, la amaurosis congénita de Leber, la coroideremia o la enfermedad de Stargardt.

La presente invención también se refiere a un kit que comprende un polinucleótido según la presente invención y un compuesto que media la entrada en la célula.

El término "kit", tal y como se utiliza aquí, se refiere a una colección de los compuestos, medios o reactivos de la presente invención mencionados anteriormente, que pueden o no estar empaquetados juntos. Los componentes del kit pueden estar compuestos por viales separados (es decir, como un kit de piezas separadas) o suministrados en un único vial. Además, debe entenderse que el kit de la presente invención, preferentemente, debe utilizarse para practicar los procedimientos mencionados anteriormente. Preferiblemente, se prevé que todos los componentes se suministren listos para su uso para practicar los procedimientos mencionados anteriormente. Además, el kit, preferentemente, contiene instrucciones para llevar a cabo dichos procedimientos. Las instrucciones pueden ser proporcionadas por un manual de usuario en papel o en formato electrónico. Además, el manual puede incluir instrucciones para interpretar los resultados obtenidos al llevar a cabo los procedimientos mencionados utilizando el kit de la presente invención. Como se entenderá de lo anterior, la descripción del kit que comprende polinucleótidos, preferentemente, se refiere a un kit que comprende los vectores correspondientes mutatis mutandis.

Preferentemente, el kit comprende además al menos un compuesto mediador de la entrada celular para el polinucleótido que comprende, el término "compuesto mediador de la entrada celular" se refiere a cualquier medio adecuado para hacer que un polinucleótido del kit entre en el interior de una célula huésped, preferentemente una célula huésped. Los compuestos adecuados que median la entrada de células (medios de entrega) son conocidos en la técnica e incluyen en particular medios de transfección, composiciones de envasado y similares. Preferentemente, el polinucleótido de la presente invención está preenvasado en un medio de entrega, por ejemplo, en partículas virales, más preferentemente en partículas virales de replicación defectuosa, más preferentemente en partículas similares a virus (VLP). El experto conoce los medios de entrega que proporcionan diferentes especificidades para los receptores celulares, de manera que se pueden seleccionar los medios de entrega apropiados para una célula huésped determinada.

La presente invención se refiere además a un dispositivo que comprende un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención.

El término "dispositivo", tal como se utiliza aquí, se refiere a un sistema de medios que comprende al menos los medios vinculados operativamente entre sí para permitir la administración del compuesto o de la composición de la presente invención. Los medios preferidos para la administración de polinucleótidos, composiciones y células huésped son bien conocidos en la técnica. La forma de vincular los medios de forma operativa dependerá del tipo de medios incluidos en el dispositivo y del tipo de administración prevista. Preferiblemente, los medios están compuestos por un único dispositivo en tal caso. Dicho dispositivo puede incluir una unidad de entrega para la administración del compuesto o composición y una unidad de almacenamiento para guardar dicho compuesto o composición hasta su administración. Sin embargo, también se contempla que los medios de la presente invención pueden aparecer como dispositivos separados en dicha realización y son, preferentemente, empaquetados juntos como un kit. El experto en la materia se dará cuenta de cómo enlazar los medios sin más. Los dispositivos preferidos son los que pueden aplicarse sin los conocimientos particulares de un técnico especializado. En una realización

preferente, el dispositivo es una jeringa, más preferiblemente con una aguja, que comprende el compuesto o la composición de la invención. En otra realización preferente, el dispositivo es un equipo de infusión intravenosa (IV) que comprende el compuesto o la composición. En otra realización preferente, el dispositivo es un dispositivo endoscópico que comprende el compuesto o el medicamento para lavar un sitio de administración, o que comprende además una aguja para la aplicación tópica del compuesto o la composición, por ejemplo, a un tumor. En otra realización preferente, el dispositivo es un inhalador que comprende el compuesto de la presente invención, en el que, más preferentemente, dicho compuesto está formulado para su administración en forma de aerosol.

La presente solicitud también se refiere a un procedimiento para transfectar de forma estable una célula huésped, que comprende

- a) poner en contacto dicha célula huésped con un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, y,
- b) de este modo, transfectar de forma estable una célula huésped.

El procedimiento para transfectar de forma estable una célula huésped de la presente invención, preferentemente, es un procedimiento in vitro. Además, puede comprender otras etapas además de las mencionadas explícitamente. Por ejemplo, otros pasos pueden referirse, por ejemplo, a proporcionar una célula huésped o una muestra que comprenda la misma para el paso a), y/o aplicar presión selectiva a las células huésped contactadas. Además, una o varias de dichas etapas pueden ser realizadas por equipos automatizados.

El término transfectar de forma estable una célula huésped es entendido por el experto para referirse a la introducción de un polinucleótido, preferiblemente un polinucleótido heterólogo en una célula de tal manera que el polinucleótido es replicado de forma estable por la célula huésped como se especifica anteriormente. Preferentemente, la transfección estable comprende la replicación episomal estable del polinucleótido. Preferentemente, la transfección estable comprende, tras el contacto, la aplicación de presión selectiva a la célula huésped para seleccionar la presencia de un marcador seleccionable. La presión selectiva se aplica después del contacto, excluyendo opcionalmente un primer marco de tiempo que permita al polinucleótido establecerse dentro de la célula huésped; la duración de dicho primer marco de tiempo que permita al polinucleótido establecerse dentro de la célula huésped dependerá principalmente del tipo de célula huésped contactada y del tipo de marcador seleccionable utilizado; preferentemente, la duración de dicho primer marco de tiempo que permita al polinucleótido establecerse dentro de la célula huésped es de 1 h a 48 h, más preferentemente de 2 h a 24, más preferentemente de 3 h a 16 h. Sin embargo, la duración de dicho primer marco de tiempo que permite que el polinucleótido se establezca dentro de la célula huésped también puede ser cero, es decir, la presión selectiva puede aplicarse inmediatamente después del contacto o incluso durante el contacto. La presión selectiva puede aplicarse de forma continua, es decir, esencialmente en todos los puntos de tiempo después del primer plazo que permite que el polinucleótido se establezca dentro de la célula huésped, más preferiblemente para evitar que las células huésped que no comprenden el polinucleótido proliferen; o puede aplicarse de forma transitoria, más preferiblemente para eliminar las células que no han recibido el polinucleótido. Preferentemente, la aplicación transitoria de la presión selectiva se utiliza en los casos en que las células se transfieren de nuevo a un organismo después de dicho contacto. Sin embargo, también se prevé que no se aplique ninguna presión selectiva, en particular en los casos en los que se sabe que la eficacia de la transferencia del polinucleótido a las células huésped diana es suficientemente alta y/o cuando una población pura de células huésped transgénicas no es de gran importancia.

El término "contacto", tal como se utiliza en el contexto de los procedimientos de la presente invención, es entendido por el experto. Preferentemente, el término se refiere a poner al menos un polinucleótido, vector y/o célula huésped de la presente invención en contacto físico con una célula huésped, por ejemplo, permitiendo que la célula huésped y el o los compuestos interactúen. Preferentemente, el contacto incluye la entrega de al menos un polinucleótido de la presente invención en el interior de una célula huésped, preferentemente a través de un medio de entrega como el especificado anteriormente.

La presente invención también se refiere a un procedimiento para tratar enfermedades genéticas en un sujeto, que comprende

- a) poner en contacto a dicho sujeto con un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, y,
- b) por lo tanto, tratar la enfermedad genética en dicho sujeto.

El procedimiento para tratar la enfermedad genética de la presente invención, preferentemente, es un procedimiento in vivo. Además, puede comprender otras etapas además de las mencionadas explícitamente. Por ejemplo, los pasos adicionales pueden referirse, por ejemplo, a proporcionar una célula huésped o una muestra que comprenda la misma para el paso a), y/o volver a administrar dicha muestra o célula huésped al sujeto. Por lo tanto, el procedimiento para el tratamiento de enfermedades genéticas, comprende los pasos del procedimiento para transfectar de forma estable una célula huésped como se especifica anteriormente. Además, una o varias de dichas etapas pueden ser realizadas por equipos automatizados.

Además, la presente invención se refiere a un uso de un polinucleótido de la presente invención para modificar genéticamente de forma estable una célula huésped.

Asimismo, la presente invención se refiere a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la fabricación de un medicamento. Y a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la fabricación de un medicamento para tratar una enfermedad genética, preferentemente una enfermedad monogénica, más preferentemente una enfermedad monogénica recesiva, más preferentemente la fenilcetonuria, la alcaptonuria, la amaurosis congénita de Leber, la coroideremia o la enfermedad de Stargardt.

Asimismo, la presente invención se refiere a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la modificación genética de una célula primaria, preferentemente un fibroblasto dérmico primario, para la generación de una célula madre pluripotente inducida (IPSCs). Preferentemente, dicha célula primaria es una célula primaria de ratón o humana.

El término "célula primaria" es entendido por el experto como opuesto a una célula de una línea celular cultivada; así, preferiblemente, una célula primaria es una célula derivada de un organismo vivo y que ha sido cultivada durante un máximo de 20 pasajes, más preferiblemente un máximo de 15 pasajes, incluso más preferiblemente un máximo de 10 pasajes, y aún más preferiblemente un máximo de 5 pasajes. Lo más preferible es que las células primarias sean células derivadas directamente de un tejido de un ser vivo, preferiblemente un ratón o un humano.

El término "célula madre" también es entendido por el experto para referirse a una célula no diferenciada o poco diferenciada con el potencial de diferenciación en al menos dos tipos de células, preferiblemente al menos cinco tipos de células, más preferiblemente al menos un linaje celular completo. Preferiblemente, la célula madre es una célula madre totipotente, más preferiblemente una célula madre pluripotente. El término "célula madre pluripotente inducida" o "IPSC" se refiere a una célula madre pluripotente derivada de una célula diferenciada, preferiblemente una célula primaria diferenciada. Los procedimientos para generar IPSCs son conocidos en la técnica e incluyen, preferentemente, la expresión de cuatro factores de transcripción en la célula (por ejemplo, de Takahashi et al. (2006), Cell. 126 (4):663).

La presente invención también se refiere a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la modificación genética de células madre embrionarias. La presente invención también se refiere a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la fabricación de un medicamento para tratar una enfermedad genética, preferentemente una enfermedad monogénica, más preferentemente una enfermedad monogénica recesiva, más preferentemente fenilcetonuria, alcaptonuria, amaurosis congénita de Leber, coroideremia o enfermedad de Stargardt, en el que dicho medicamento comprende células huésped que comprenden un polinucleótido de la presente invención.

La presente invención también se refiere a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la modificación genética de células madre para generar un animal transgénico. La presente invención se refiere además a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la producción de un animal transgénico.

El término "animal transgénico", tal y como se utiliza aquí, se refiere a un animal que comprende al menos un polinucleótido heterólogo, preferiblemente introducido en dicho animal por procedimientos de ingeniería genética. Preferentemente, el animal transgénico comprende al menos una, más preferentemente al menos 10, aún más preferentemente al menos 1000, incluso más preferentemente al menos 10000 células que comprenden al menos un polinucleótido según la presente invención.

Asimismo, la presente invención se refiere a un uso de un polinucleótido según la presente invención, una composición según la presente invención, y/o una célula huésped según la presente invención, para la modificación genética de embriones unicelulares por inyección pronuclear.

Tal como lo entiende el experto, el término "inyección pronuclear" se refiere a la inyección de material genético, preferentemente un polinucleótido de la presente invención, en el núcleo de un ovocito fecundado, preferentemente para crear un animal transgénico.

Otras definiciones:

AF: Sin antibióticos.

amperio: Ampicilina.

ampliación: Gen de resistencia a la ampicilina.

Marcador seleccionable de antibióticos: Gen que confiere resistencia a un antibiótico, por ejemplo, gen de resistencia a la ampicilina, gen de resistencia a la kanamicina, gen de resistencia al cloranfenicol, gen de resistencia a la puromicina, gen de resistencia a la tetraciclina.

ApoB: Apolipoproteína B

5 Aproximadamente: Tal como se utiliza aquí, el término "aproximadamente" o "sobre", aplicado a uno o más valores de interés, se refiere a un valor que es igual o similar a un valor de referencia establecido.

Región bacteriana: Región de un vector plasmídico necesaria para la propagación y selección en el huésped bacteriano.

pb: pares de bases

10 ccc: Circular covalentemente cerrado.

cl: Represor Lambda.

clTs857: El represor Lambda incorpora además una mutación de C a T (Ala a Thr) que le confiere sensibilidad a la temperatura. clTs857 es un represor funcional a 28-30°C pero es mayormente inactivo a 37-42°C. También se denomina cl857.

15 CatR: Gen de resistencia al cloranfenicol.

cmv: Citomegalovirus.

E. coli: Escherichia coli, una bacteria gramnegativa.

EGFP: Proteína verde fluorescente mejorada.

ELE40: Elemento antirrepresor 40, STAR40 divulgado en Kwaks et al., 2003, Nat Biotechnol. 21:553

20 EP: Electroporación.

Eficiencia del establecimiento: El porcentaje de células en las que un vector de expresión S/MAR episomal no integrador se mantiene de forma estable como un episoma después de la transfección.

25 Vector de expresión eucariota: Un vector para la expresión de ARNm, antígenos proteicos, proteínas terapéuticas, ARNh, ARN o genes de microARN en una célula u organismo eucariótico diana utilizando promotores de ARN polimerasa I, II o III.

30 Región eucariótica: La región de un plásmido que codifica secuencias eucariotas y/o secuencias necesarias para la función del plásmido en el organismo diana. Esto incluye la región de un vector plasmídico necesaria para la expresión de uno o más transgenes en el organismo diana, incluidos los potenciadores del ARN Pol II, los promotores, los transgenes y las secuencias poliA. Esto también incluye la región de un vector de plásmido necesaria para la expresión de uno o más transgenes en el organismo diana utilizando promotores de ARN Pol I o ARN Pol III, transgenes o ARN expresados por ARN Pol I o ARN Pol III. La región eucariótica puede incluir opcionalmente otras secuencias funcionales, como los terminadores transcripcionales eucarióticos, las estructuras de dúplex de ADN desestabilizado por superenrollamiento (SIDD), los S/MAR, los elementos de frontera, etc.

35 Exon: Secuencia de nucleótidos codificada por un gen que se transcribe y está presente en un producto de ARNm maduro después de que se haya completado el empalme del ARN para eliminar los intrones. Potenciador de la expresión: Una secuencia de ADN que mejora la expresión de un promotor adyacente. Por ejemplo, Ele40, UCOE, elementos antirrepresores o elementos estabilizadores antirrepresores (STAR), como se revisa en Saunders et al., 2015 PloS One 10:e0120096

40 Vector de expresión: Un vector para la expresión de ARNm, antígenos proteicos, proteínas terapéuticas, ARNh, ARN o genes de microARN en un organismo diana.

g: Gramo, kg para kilogramo

45 gen de interés: gen que debe expresarse en el organismo diana. Incluye genes de ARNm que codifican antígenos proteicos o peptídicos, terapias proteicas o peptídicas, y ARNm, ARNh, ARN o microARN que codifican terapias de ARN, y ARNm, ARNh, ARN o microARN que codifican vacunas de ARN, etc.

GFP: Proteína verde fluorescente.

Hr(s): Hora(s).

ID: Intradérmica.

IM: Intramuscular.

respuesta inmunitaria: Respuestas celulares reactivas al antígeno (por ejemplo, células T reactivas al antígeno) o de anticuerpos (por ejemplo, IgG reactiva al antígeno).

5 Intron: Secuencia de nucleótidos codificada por un gen que se transcribe y posteriormente se elimina de un producto de ARNm maduro por empalme de ARN.

ITR: Repetición del terminal invertido.

kan: Kanamycin.

kanR: Gen de resistencia a la kanamicina.

10 Kd:Kilodalton.

secuencia kozak: Secuencia de ADN consenso optimizada gccRccATG (R = G o A) inmediatamente antes de un codón de inicio ATG que garantiza una iniciación de la tranlación eficiente.

MFI: Intensidad fluorescente media.

15 minicírculo: Derivados de plásmidos circulares cerrados covalentemente en los que la región bacteriana se ha eliminado del plásmido original por recombinación específica del sitio in vivo o in vitro o por digestión/ligación de restricción in vitro. Los vectores de minicírculo son incompetentes para la replicación en células bacterianas.

aARNm: ARN mensajero.

mSEAP: Fosfatasa alcalina secretada murina.

20 NA: No es aplicable.

Vector NanoplasmidTM: Vector con una región bacteriana que combina un marcador seleccionable de ARN con un origen de replicación relacionado con R6K, ColE2 o ColE2. Por ejemplo, los vectores NTC9385C, NTC9685C, NTC9385R, NTC9685R y las modificaciones descritas en Williams, 2014. Plásmidos de ADN con expresión mejorada. Solicitud de patente mundial WO2014035457 y que se incluye aquí por referencia.

25 Vector lentiviral no integrador: Un vector lentiviral con integrasa mutada y un S/MAR para el mantenimiento de círculos LTR episomales como los descritos en Verghese et al., 2014 Nucleic Acids Research 42:e53.

NP: Nanoplásmido.

30 NTC8385: Los plásmidos NTC8385, NTC8485 y NTC8685 son vectores de origen pUC sin antibióticos que contienen un marcador seleccionable de ARN corto (RNA-OUT) en lugar de un marcador de resistencia a los antibióticos como kanR. La creación y aplicación de estos vectores sin antibióticos basados en ARN-OUT se describen en Williams, JA 2008 Solicitud de patente mundial WO2008153733 y se incluye aquí por referencia.

35 NTC8485: El NTC8485 es un vector de origen pUC sin antibióticos que contiene un marcador seleccionable de ARN corto (RNA-OUT) en lugar de un marcador de resistencia a los antibióticos como kanR. La creación y aplicación del NTC8485 se describe en Williams, JA 2010 Solicitud de patente estadounidense 20100184158 y se incluye aquí por referencia.

NTC8685: El NTC8685 es un vector de origen pUC sin antibióticos que contiene un marcador seleccionable de ARN corto (RNA-OUT) en lugar de un marcador de resistencia a los antibióticos como kanR. La creación y aplicación del NTC8685 se describe en Williams, Supra, 2010 y se incluye aquí por referencia.

40 NTC9385R: El vector NTC9385R NanoplasmidTM descrito en Williams, Supra, 2014 incluido aquí por referencia tiene una región espaciadora codificada NheI- trpA terminador-R6K origen ARN-OUT -KpnI región bacteriana (SEQ ID NO:8) vinculada a través de los sitios NheI y KpnI flanqueantes a la región eucariótica.

OD600: densidad óptica a 600 nm.

45 PBS: Solución salina tamponada con fosfato.

PCR: Reacción en cadena de la polimerasa.

pDNA: ADN plasmídico.

pINT pR pL vector: El vector de expresión de integración pINT pR pL attHK022 se describe en Luke et al., 2011 Mol Biotechnol 47:43 y se incluye aquí por referencia. El gen diana a expresar se clona aguas abajo del promotor pL. El vector codifica el represor cl857 inducible por temperatura, lo que permite la expresión del gen diana inducible por calor.

Promotor de PL: El promotor de Lambda se fue. PL es un promotor fuerte que es reprimido por el represor cl que se une a los sitios de unión del represor OL1, OL2 y OL3. El represor cl857, sensible a la temperatura, permite controlar la expresión del gen por inducción de calor, ya que a 30 °C el represor cl857 es funcional y reprime la expresión del gen, pero a 37-42 °C el represor se inactiva, por lo que se produce la expresión del gen.

Promotor PL (OL1 G a T): El promotor de Lambda se fue. PL es un promotor fuerte que es reprimido por el represor cl que se une a los sitios de unión del represor OL1, OL2 y OL3. El represor cl857, sensible a la temperatura, permite controlar la expresión del gen por inducción de calor, ya que a 30 °C el represor cl857 es funcional y reprime la expresión del gen, pero a 37-42 °C el represor se inactiva, por lo que se produce la expresión del gen. La unión del represor cl a OL1 se reduce por la mutación G a T de OL1, lo que resulta en un aumento de la actividad del promotor a 30 °C y 37-42 °C, como se describe en Williams, Supra, 2014.

Plásmido: Una molécula extra de ADN cromosómico separada del ADN cromosómico que es capaz de replicarse independientemente del ADN cromosómico

Número de copias del plásmido: el número de copias de un plásmido por célula. El aumento del número de copias de plásmidos incrementa el rendimiento de la producción de plásmidos.

Pol: Polimerasa.

poliA: Señal o sitio de poliadenilación. La poliadenilación es la adición de una cola de poli(A) a una molécula de ARN. La señal de poliadenilación contiene el motivo de la secuencia reconocida por el complejo de corte del ARN. La mayoría de las señales de poliadenilación humanas contienen un motivo AAUAAA y secuencias conservadas 5' y 3' a él. Las señales de poliA comúnmente utilizadas se derivan de la β globina de conejo (RBG), de la hormona de crecimiento bovina (BGH), de las señales de poliA tempranas de SV40 o de las tardías de SV40.

origen pUC: origen de replicación derivado de pBR322, con transición de G a A que aumenta el número de copias a temperaturas elevadas y supresión del regulador negativo ROP.

pUC libre: Plásmido que no contiene el origen pUC. Pueden incluirse fragmentos no replicativos de origen pUC, por ejemplo el marcador seleccionable RNAI.

plásmido pUC: Plásmido que contiene el origen pUC.

PuroR: Gen de resistencia a la puromicina.

Plásmido R6K: NTC9385R, NTC9685R, NTC9385R2-O1, NTC9385R2-O2, NTC9385R2a-O1, NTC9385R2a-O2, NTC9385R2b-O1, NTC9385R2b-O2, NTC9385Ra-O1, NTC9385Ra-O2, NTC9385RaF, y NTC9385RbF, así como modificaciones y vectores alternativos que contienen un origen de replicación R6K que se describieron en Williams, Supra, 2014 y que se incluyen aquí por referencia. Los vectores alternativos R6K conocidos en la técnica incluyen, entre otros, los vectores pCOR (Gencell), los vectores pCpGfree (Invivogen) y los vectores libres de CpG de la Universidad de Oxford, incluido el pGM169.

Origen de replicación R6K: región que es reconocida específicamente por la proteína R6K Rep para iniciar la replicación del ADN. Incluye, pero no se limita a la secuencia del origen de replicación gamma R6K divulgada como SEQ ID NO:1, SEQ ID NO:2 y SEQ ID NO:4, y las versiones libres de CpG (por ejemplo, SEQ ID NO:3) como se describe en Drocourt et al., Patente de Estados Unidos 7244609 y que se incorpora al presente documento por referencia Origen de replicación R6K-Región bacteriana de ARN-OUT: Contiene un origen de replicación R6K para la propagación y el marcador seleccionable RNA-OUT (por ejemplo, SEQ ID NO:8; SEQ ID NO:9; SEQ ID NO:10; SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID NO:14; SEQ ID NO:15; SEQ ID NO:16; SEQ ID NO:17).

Rep: Replicación

Intermedios de replicación: Fragmentos lineales de ADN resultantes de la terminación prematura de la replicación del plásmido

Plásmido dependiente de la proteína Rep: Un plásmido en el que la replicación depende de una proteína de replicación (Rep) proporcionada en Trans. Por ejemplo, el origen de replicación R6K, el origen de

replicación ColE2-P9 y los plásmidos de origen de replicación relacionados con ColE2 en los que la proteína Rep se expresa a partir del genoma de la cepa huésped. Se conocen numerosos plásmidos adicionales dependientes de la proteína Rep en la técnica, muchos de los cuales se resumen en del Solar et al., Supra, 1998 que se incluye aquí por referencia

5 Vector retroviral: Vector viral integrador que puede infectar células en división. También se llama plásmido de transferencia. El plásmido codifica la unidad de expresión flanqueada por el LTR retroviral. El plásmido de transferencia se transfecta en las células de producción junto con los plásmidos de envoltura y empaquetamiento necesarios para fabricar partículas virales.

10 ARN-IN: La secuencia de inserción 10 (IS10) codifica el ARN-IN, un ARN complementario y antisentido a una porción del ARN RNA-OUT. Cuando el ARN-IN se clona en el líder no traducido de un ARNm, el recodo del ARN-IN con el ARN-OUT reduce la traducción del gen codificado aguas abajo del ARN-IN.

15 Marcador seleccionable regulado por ARN-IN: Un marcador selectivo regulado por ARN-IN expresado cromosómicamente. En presencia del ARN represor antisentido RNA-OUT (SEQ ID NO:6) portado por el plásmido, se reprime la expresión de una proteína codificada aguas abajo del ARN-IN. Un marcador seleccionable regulado por ARN-IN está configurado de tal manera que el ARN-IN regula 1) una proteína que es letal o tóxica para dicha célula per se o generando una sustancia tóxica (por ejemplo, SacB), o 2) una proteína represora que es letal o tóxica para dicha célula bacteriana reprimiendo la transcripción de un gen que es esencial para el crecimiento de dicha célula (por ejemplo, el gen esencial *murA* regulado por el gen represor *tetR* del ARN-IN). Por ejemplo, las líneas celulares de ARN-IN-SacB expresadas cromosómicamente para la selección/propagación de plásmidos de ARN-OUT se describen en Williams, Supra, 2008 y se incluyen aquí por referencia. Los marcadores de selección alternativos descritos en la técnica pueden sustituir a SacB.

25 ARN-OUT: La secuencia de inserción 10 (IS10) codifica el ARN-OUT, un ARN antisentido que se hibrida con el gen del transposón expresado aguas abajo del ARN-IN y reduce su traducción. La secuencia del ARN-OUT (SEQ ID NO:6) y las líneas celulares complementarias de ARN-IN SacB expresadas cromosómicamente pueden modificarse para incorporar pares funcionales alternativos de unión ARN-IN/ARN-OUT como los descritos en Mutalik et al., 2012 Nat Chem Biol 8:447, incluyendo, entre otros, el par ARN-OUT A08/ARN-IN S49, el par ARN-OUT A08/ARN-IN S08, y las modificaciones libres de CpG del ARN-OUT A08 que modifican el CG en la secuencia TTCGC del ARN-OUT 5' a una secuencia no CpG. En Williams 2015 se describió un ejemplo de marcador de selección de ARN-OUT libre de CpG, en el que se eliminan los dos motivos CpG en el ARN-OUT (uno de los cuales está presente en la región complementaria del ARN-IN). Vectores de minicirculo replicantes con expresión mejorada. Solicitud de patente de los Estados Unidos US 2015/0275221 y que se incluye aquí por referencia. Una multitud de sustituciones alternativas para eliminar los dos motivos CpG (mutando cada CpG a CpA, CpC, CpT, ApG, GpG o TpG) puede ser utilizada para hacer un ARN-OUT libre de CpG

35 RNA-OUT Marcador seleccionable: Un fragmento de ADN marcador seleccionable de ARN-OUT que incluye secuencias promotoras y terminadoras de la transcripción de *E. coli* que flanquean un ARN-OUT. Un marcador seleccionable de ARN-OUT, que utiliza las secuencias promotoras y terminadoras de ARN-OUT, que está flanqueado por sitios de enzimas de restricción *DraIII* y *KpnI*, y líneas celulares de ARN-IN-SacB expresadas cromosómicamente para la propagación del plásmido de ARN-OUT, se describen en Williams, Supra, 2008 y se incluyen aquí por referencia. Las secuencias promotoras y terminadoras del ARN-OUT de la SEQ ID NO: 5 que flanquean el ARN-OUT (SEQ ID NO:6) pueden ser sustituidos por secuencias promotoras y terminadoras heterólogas. Por ejemplo, el promotor RNA-OUT puede sustituirse por un promotor libre de CpG conocido en la técnica, por ejemplo el promotor I-EC2K o los promotores P5/6 5/6 o P5/6 6/6 descritos en Williams, Supra, 2008 e incluidos aquí por referencia. Un marcador seleccionable de ARN-OUT de 2 CpG en el que se eliminan los dos motivos CpG del promotor de ARN-OUT se da como SEQ ID NO: 7. Un ejemplo de una unidad de transcripción ARN-OUT libre de CpG, en la que se eliminan los dos motivos CpG en el ARN-OUT (uno de los cuales está presente en la región complementaria del ARN-IN) y los dos motivos CpG en el promotor del ARN-OUT se describió en Williams, Supra, 2015 y se incluye aquí por referencia. Los vectores que incorporan el marcador seleccionable RNA-OUT libre de CpG pueden seleccionarse para la resistencia a la sacarosa utilizando las líneas celulares RNA-IN-SacB para la propagación del plásmido RNA-OUT descritas en Williams, Supra, 2008. Alternativamente, la secuencia de ARN-IN en estas líneas celulares puede modificarse para incorporar el cambio de 1 pb necesario para que la región de ARN-OUT libre de CpG sea perfectamente complementaria al ARN-IN.

60 Promotor de la ARN polimerasa II: Promotor que recluta a la ARN polimerasa II para sintetizar ARNm, la mayoría de los ARN nucleares pequeños y los microARN. Por ejemplo, promotores constitutivos como el promotor del CMV humano o murino, el promotor del factor de elongación 1 (EF1), el promotor de la β -actina del pollo, el promotor de la β -actina de otras especies, el promotor del factor de elongación-1 α (EF1 α), el promotor de la fosfogliceroquinasa (PGK) el promotor del virus del sarcoma de Rous (RSV), el promotor de la seroalbúmina humana (SA), el promotor del virus formador de focos del bazo (SFFV), el

- promotor de la α -1 antitripsina (AAT), el promotor de la globulina fijadora de tiroxina (TBG), el promotor del citocromo P450 2E1 (CYP2E1), etc. Los vectores también pueden utilizar promotores combinados, como el promotor de la β -actina de pollo/el potenciador del CMV (CAG), los elementos potenciadores derivados del CMV humano o murino combinados con los promotores del factor de elongación 1 α (EF1 α), versiones libres de CpG de los elementos potenciadores derivados del CMV humano o murino combinados con los promotores del factor de elongación 1 α (EF1 α), el promotor de la albúmina combinado con un potenciador de la α -fetoproteína MERII, etc, o la diversidad de promotores específicos de tejido o inducibles conocidos en la técnica, como los promotores específicos de músculo creatina quinasa (MCK), y C5-12 o los promotores específicos de hígado ApoE-hAAT, apolipoproteína A-I (ApoAI), etc.
- Promotor de la ARN polimerasa III: Promotor que recluta a la ARN polimerasa III para sintetizar ARNt, ARN ribosomal 5S y otros ARN pequeños. Por ejemplo, promotores de clase I como el promotor del ARNr 5s, promotores de clase II como los promotores del ARNt, promotores de clase III como el promotor del ARN nuclear pequeño U6 o el promotor de la RNasa P nuclear H1, etc.
- Marcador seleccionable de ARN: Un marcador de ARN seleccionable es un ARN no traducido expresado por un plásmido que regula un gen diana expresado por el cromosoma para permitir la selección. Puede tratarse de un ARNt supresor del sinsentido portado por un plásmido que regula una diana cromosómica seleccionable supresora del sinsentido, como se describe en Crouzet J y Soubrier F 2005 Patente estadounidense 6,977,174 incluida aquí por referencia. También puede tratarse de un ARN represor antisentido portado por un plásmido; una lista no limitativa incluida aquí por referencia incluye el ARN-OUT que reprime las dianas reguladas por el ARN-IN (Williams, Supra, 2008), el ARNI codificado por el origen del plásmido pMB1 que reprime las dianas reguladas por el ARNII (Grabherr R, Pfaffenzeller I. 2006 solicitud de patente estadounidense US20060063232; Cranenburgh RM. 2009 Patente estadounidense 7.611.883), el plásmido de origen IncB pMU720 codifica el ARNI que reprime las dianas reguladas por el ARN II (Wilson IW, Siemering KR, Praszkie J, Pittard AJ. 1997. J Bacteriol 179:742-53), el locus ParB Sok del plásmido R1 que reprime las dianas reguladas por Hok, el locus FlmB del plásmido F que reprime las dianas reguladas por flmA (Morsey MA, 1999 Patente estadounidense US5922583). Un marcador seleccionable de ARN puede ser otro ARN represor antisentido natural conocido en la técnica, como los descritos en Wagner EGH, Altuvia S, Romby P. 2002. Adv Genet 46:361-98 y Franch T, y Gerdes K. 2000. Current Opin Microbiol 3:159-64. Un marcador seleccionable de ARN también puede ser un ARN represor diseñado, como los pequeños ARN sintéticos expresados en los andamiajes SgrS, MicC o MicF, como se describe en Na D, Yoo SM, Chung H, Park H, Park JH, Lee SY. 2013. Nat Biotechnol 31:170-4. Un marcador seleccionable de ARN también puede ser un ARN represor diseñado como parte de un marcador seleccionable que reprime un ARN diana fusionado a un gen diana que debe regularse, como SacB, como se describe en Williams, Supra, 2015
- ROP: Represor de la imprimación.
- RSM: Marcador seleccionable de ARN.
- SA: Aceptador de empalme, la secuencia consenso YYYYYYAGRW se presenta como SEQ ID NO:29. Para crear un sitio SA empalmado de forma eficiente, se incluye un punto de bifurcación de empalme (secuencia consenso YTNAY) aguas arriba del sitio aceptor de empalme (véase la Figura 1).
- SacB: Gen estructural que codifica la levansucrasa de *Bacillus subtilis*. La expresión de SacB en bacterias gramnegativas es tóxica en presencia de sacarosa.
- SD: Donante de empalme, secuencia de consenso AGGTRAGT.
- SEAP: Fosfatasa alcalina secretada.
- Marcador seleccionable: Un marcador seleccionable, por ejemplo un gen de resistencia a la kanamicina o un marcador seleccionable de ARN.
- Marcador de selección: Un marcador seleccionable, por ejemplo un gen de resistencia a la kanamicina o un marcador seleccionable de ARN.
- SIDD: estructuras desestabilizadas de dúplex de ADN inducidas por superenrollamiento (SIDD). Estos sitios, cuando se incorporan a un vector, pueden alterar la susceptibilidad de otras secuencias dentro del vector a ser desestabilizadas. Esto puede alterar la función. Por ejemplo, la adición de un sitio SIDD a un vector de expresión puede reducir la desestabilización helicoidal de un promotor. Esto puede aumentar o disminuir la actividad del promotor, dependiendo del promotor, ya que algunos promotores tienen una mayor expresión con la desestabilización helicoidal del promotor, mientras que otros tendrán una expresión reducida con la desestabilización helicoidal del promotor.
- aRNhc: ARN de horquilla corta.

S/MAR: Región de fijación del andamio/matriz como se especifica en este documento. Secuencias eucariotas que median la fijación del ADN a la matriz nuclear.

Región del espaciador: Tal y como se utiliza aquí, la región espaciadora es la región que une los extremos 5' y 3' de las secuencias de la región eucariota. Los extremos 5' y 3' de la región eucariota suelen estar separados por el origen de replicación bacteriano y el marcador seleccionable bacteriano en los vectores plasmídicos.

SR: Región del espaciador.

Origen del SV40: ADN genómico del Virus Simio 40 que contiene el origen de replicación.

Potenciador del SV40: ADN genómico del Virus Simio 40 que contiene las repeticiones del potenciador de 72 pb y, opcionalmente, de 21 pb.

antígeno diana: Epítipo proteico o peptídico inmunogénico, o combinación de proteínas y epítopos, contra el que se puede montar una respuesta inmunitaria. Los antígenos diana pueden derivarse de un patógeno para aplicaciones de enfermedades infecciosas o alergias o derivarse de un organismo huésped para aplicaciones como el cáncer, la alergia o las enfermedades autoinmunes. Los antígenos diana están bien definidos en la técnica. Algunos ejemplos se describen en Williams, Supra, 2008 y se incluyen aquí por referencia.

Tampón TE: Una solución que contenga aproximadamente 10mM de Tris pH 8 y 1 mM de EDTA.

TetR: Gen de resistencia a la tetraciclina.

Terminador de la transcripción: Bacteriana: Secuencia de ADN que marca el final de un gen u operón para la transcripción. Puede tratarse de un terminador intrínseco de la transcripción o de un terminador transcripcional dependiente de Rho. Para un terminador intrínseco, como el terminador trpA, se forma una estructura de horquilla dentro del transcrito que interrumpe el complejo ternario ARNm-ARN polimerasa. Alternativamente, los terminadores transcripcionales dependientes de Rho requieren el factor Rho, un complejo proteico de ARN helicasa, para interrumpir el complejo ternario ARNm-ARN polimerasa naciente. Eucariota: Las señales de PolyA no son "terminadores", sino que la escisión interna en los sitios de PolyA deja un extremo 5'sin tapar en el ARN 3'UTR para la digestión de la nucleasa. La nucleasa alcanza al ARN Pol II y provoca la terminación. La terminación puede promoverse dentro de una región corta del sitio poli A mediante la introducción de sitios de pausa del ARN Pol II (terminador de transcripción eucariótico). La pausa de la ARN Pol II permite que la nucleasa introducida en el ARNm 3' UTR después de la escisión de PolyA alcance a la ARN Pol II en el lugar de la pausa. Una lista no limitativa de terminadores de transcripción eucariotas conocidos en la técnica incluye el C2x4 y el terminador de gastrina. Los terminadores de transcripción eucariotas pueden elevar los niveles de ARNm al mejorar el procesamiento adecuado del extremo 3 del ARNm.

transfección: Procedimiento para introducir ácidos nucleicos en las células [por ejemplo poli(lactida-co-glicolida) (PLGA), ISCOMs, liposomas, niosomas, virosomas, quitosano y otros polímeros biodegradables, micropartículas, microesferas, nanopartículas, nanocápsulas, electroporación, nucleofección, permeabilización piezoeléctrica, sonoporación, iontoforesis, ultrasonido, disrupción de la membrana mediada por deformación celular de alta velocidad SQZ, plasma corona, entrega facilitada por plasma, plasma tolerable por el tejido, microporación láser, energía de ondas de choque, campos magnéticos, magnetopermeabilización sin contacto, pistola de genes, microagujas, microdermoabrasión, entrega hidrodinámica, inyección de alta presión en la vena de la cola, etc.] como se conoce en la técnica y se incluye aquí por referencia.

Transgénico: Gen de interés que se clona en un vector para su expresión en un organismo diana. ts: Sensible a la temperatura

µg: Microgramo

µl: Microlitro

UCOE: Elemento de apertura de la cromatina ubicuo, como el A2UCOE o derivados mínimos como se divulga en Muller-Kuller et al., 2015, Nucleic Acids Research 43:1577.

UTR: Región no traducida de un ARNm (5' o 3' de la región codificante).

Vector: Un vehículo de administración de genes, incluyendo vectores virales (por ejemplo, alfavirus, poxvirus, lentivirus, retrovirus, adenovirus, virus relacionados con el adenovirus, vectores lentivirales deficientes en integración, etc.) y no virales (por ejemplo, plásmidos, nanoplásmidos, MIDGE, fragmentos

de PCR transcripcionalmente activos, minicírculos, bacteriófagos, etc.). Estos son bien conocidos en la técnica y se incluyen aquí por referencia.

Vector de fondo: Región eucariota y región bacteriana de un vector, sin la región codificante del transgén o del antígeno diana.

- 5 Vector de expresión de vertebrados: Un vector para la expresión de ARNm, antígenos proteicos, proteínas terapéuticas, ARNhc, ARN o genes de microARN en una célula u organismo vertebrado diana utilizando promotores de ARN polimerasa I, II o III.

Descripción detallada de las realizaciones preferentes

- 10 La tecnología actual se refiere en general a procedimientos y composiciones de vectores de expresión vertebrados episomales no integradores autorreplicantes que mejoran la replicación episomal y la expresión de transgenes. La tecnología actual puede practicarse para mejorar la expresión y la replicación episomal de vectores tales como los vectores no virales y los vectores virales (por ejemplo, vector lentivirus episomal deficiente para la integración, vectores lentivirales no integradores, vector retroviral episomal, etc.).

- 15 La replicación episomal mejorada se define en el presente documento como el establecimiento y/o mantenimiento mejorado del vector episomal no integrador in vitro o in vivo en comparación con un plásmido que no incorpora la tecnología actual. La expresión mejorada del plásmido se define en el presente documento como la mejora del nivel de expresión del transgén y/o la duración de la expresión in vitro o in vivo en comparación con un plásmido que codifica el transgén que no incorpora la tecnología actual. Debe entenderse que todas las referencias citadas aquí se incorporan por referencia en su totalidad.

- 20 Los procedimientos de modificación de plásmidos de la presente tecnología se han encontrado sorprendentemente para proporcionar una solución para proporcionar vectores episomales no integradores autorreplicantes con un establecimiento eficiente.

- 25 Los procedimientos y composiciones de vectores aquí divulgados son composiciones SD-SMAR-SA 3' UTR con expresión mejorada y/o establecimiento episomal (rendimiento mejorado) en comparación con las versiones no SD-SA. La mejora del rendimiento no es específica de la S/MAR, ya que se observa una mejora del rendimiento con varias S/MAR. La mejora del rendimiento tampoco es específica de la unidad de transcripción del vector, ya que la mejora del rendimiento se observa con SD-SMAR-SA vinculado a varios promotores, 5' UTRs, transgenes y señales polyA. Se observa un mejor rendimiento con o sin intrones aguas arriba. Por lo tanto, los vectores SD-SMAR-SA 3' UTR de la divulgación son ampliamente aplicables para mejorar el rendimiento de los vectores de expresión de vertebrados episomales no integradores autorreplicantes.

- 35 La mejora del rendimiento de la UTR 3' SD-SMAR-SA revelada en comparación con las versiones no SD-SA es sorprendente a la luz del estado de la técnica. Por ejemplo, Le Hir et al., 2003 Trends in Biochemical Sciences 28:215 enseña 'Matsumoto et al. [51] comprobó que estos efectos traslacionales son altamente dependientes de la posición del intrón. En su estudio, un intrón situado en la UTR 5' era altamente estimulante, mientras que el mismo intrón situado en la UTR 3' reprimía la traducción por debajo del nivel del ARNm correspondiente sin intrón' 'No obstante, para los investigadores interesados en optimizar la expresión de los transgenes, es importante tener en cuenta que la posición del intrón es una variable importante. Además de inhibir potencialmente la traducción, los intrones en la UTR 3' pueden desencadenar la decadencia mediada por el sinsentido (NMD) del ARNm, como se describe a continuación, lo que da lugar a una expresión proteica aún más baja' Barrett et al., 2012 Cell. Mol. Life Sci. 69:3613 enseña "En contraste con las 5'UTRs, se comprobó que las 3'UTRs tienen relativamente pocos intrones (5 %) [21]. Un estudio que analizaba casos raros de adquisición de intrones en genes de mamíferos retropuestos descubrió que la presencia de un intrón en el 3'UTR de estos genes daba lugar a una regulación a la baja de la expresión génica por desintegración mediada por el sinsentido [52]. Este efecto negativo sobre la expresión ofrece una explicación para la baja prevalencia de los intrones 3'UTR" Aunque no limita la aplicación de la invención, la adición de sitios de empalme donantes y aceptores de empalme flanqueantes puede tener un beneficio inesperado en la invención divulgada en la que la UTR 3' codifica una secuencia S/MAR.

- 50 Tal y como se utiliza aquí, el término "identidad de secuencia" se refiere al grado de identidad entre cualquier secuencia de consulta dada, por ejemplo, SEQ ID NO: 2, y una secuencia de temas. Una secuencia objeto de estudio puede, por ejemplo, tener al menos un 90 por ciento, un 95 por ciento o un 99 por ciento de identidad de secuencia con una determinada secuencia de consulta. Para determinar el porcentaje de identidad de la secuencia, una secuencia de consulta (por ejemplo, una secuencia de ácido nucleico) se alinea con una o más secuencias del sujeto utilizando cualquier programa adecuado de alineación de secuencias que sea bien conocido en la técnica, por ejemplo, el programa informático ClustalW (versión 1.83, parámetros por defecto), que permite realizar alineaciones de secuencias de ácido nucleico en toda su longitud (alineación global). Chema et al., 2003 Nucleic Acids Res., 31:3497-500. En un procedimiento preferido, el programa de alineación de secuencias (por ejemplo, ClustalW) calcula la mejor coincidencia entre una consulta y una o más secuencias del sujeto, y las alinea para poder determinar identidades, similitudes y diferencias. Se pueden insertar huecos de uno o más nucleótidos en una secuencia de consulta, en una secuencia de sujeto, o en ambas, para maximizar las alineaciones de secuencias.

Para los alineamientos rápidos por pares de secuencias de ácidos nucleicos, se pueden seleccionar los parámetros por defecto adecuados para el programa de alineamiento en particular. El resultado es un alineamiento de secuencias que refleja la relación entre las mismas. Para determinar además el porcentaje de identidad de una secuencia de ácido nucleico del sujeto con una secuencia de consulta, las secuencias se alinean utilizando el programa de alineación, el número de coincidencias idénticas en la alineación se divide por la longitud de la secuencia de consulta, y el resultado se multiplica por 100. Cabe señalar que el valor de la identidad porcentual puede redondearse a la décima más cercana. Por ejemplo, 78.11, 78.12, 78.13 y 78.14 se redondean a 78.1, mientras que 78.15, 78.16, 78.17, 78.18 y 78.19 se redondean a 78.2.

Volviendo a los dibujos, la FIG. 1. muestra mapas anotados del intrón pCI (arriba), de la región donante de empalme (SD) (centro) y del punto de ramificación y de la región aceptora de empalme (SA) (abajo). FIG. 2 muestra mapas anotados del interferón beta S/MAR (arriba), y un derivado SD del interferón beta S/MAR SA (centro), así como un derivado SD del interferón beta S/MAR SA en el que se mutaron las señales de poliadenilación internas AATAAA (abajo). FIG. 3 muestra mapas anotados del derivado M18 del interferón beta S/MAR con sitios SD y SA flanqueantes. FIG. 4 muestra mapas anotados de los 805 pb (arriba) o 525 pb (abajo) de apoB S/MAR con sitios SD y SA flanqueantes. FIG. 5 muestra un mapa anotado del vector pMAX-UCOE-coGFP P2A-PuroR-NP (pSMART UCOE) FIG. 6 muestra mapas anotados de los vectores NTC9385R-UCOE-CMV-coGFP P2A-PuroR -SMAR-SV40 pA (NP-UCOE) y NTC9385R-UCOE-CMV- coGFP P2A-PuroR - SD SMAR- SA SV40 pA (NP-UCOE-SP). FIG. 7 muestra mapas anotados de los vectores NTC9385R-SP-UCOE-CMV-GFP SMARter (NP-SMARter-SP) y NTC9385R-SP-UCOE-CMV-GFP CMARter (NP-CMARter-SP). FIG. 8 muestra mapas anotados de los vectores NTC9385R- UCOE EF1 -coGFP SD -SMAR SA SV40 pA (NP-UCOE-EF1-SP) y NTC9385R- UCOE EF1-coGFP-SD SMAR R6K-R-OUT-SA pA (UCOE-EF1-SP-NP). FIG. 9 muestra mapas anotados del vector NTC9385R-SP-ELE40-CMV-GFP CMARter (NP-Ele40-CMARter-SP). FIG. 10 muestra la mejora de la expresión de los vectores S/MAR establecidos con y sin sitios SD y SA flanqueantes. Panel izquierdo: MFI de células HEK293T establecidas con un vector S/MAR con y sin uniones de empalme. Los vectores contienen la región bacteriana NP, el aislante genómico UCOE, el casete de expresión GFP -2A-PuroR impulsado por el promotor CMV y el interferón beta S/MAR en la UTR 3' con (Nano-S/MAR-splice = NP-UCOE-SP; NTC9385R-UCOE-CMV- coGFP P2A-PuroR - SD SMAR- SA SV40 pA, Figura 6) o sin (NP-UCOE; NTC9385R-UCOE-CMV- coGFP P2A-PuroR -SMAR-SV40 pA, Figura 6) sitios S/MAR flanqueantes SD y SA. Panel derecho: la mejora de la expresión de la transcripción se confirma mediante el análisis de la PCR en tiempo real. La expresión del transgén GFP se normalizó con respecto al gen de mantenimiento GAPDH. FIG. 11 muestra la mejora de la expresión de los vectores S/MAR establecidos con y sin sitios SD y SA flanqueantes. MFI de células establecidas (HEK293T y fibroblastos embrionarios de ratón primarios) con vectores que albergan diferentes S/MARs flanqueados por uniones de empalme. Los nombres de los vectores son los de las figuras 5, 6 y 7. FIG. 12 muestra la mejora del establecimiento de los vectores S/MAR con y sin los sitios SD y SA flanqueantes. Ensayo de formación de colonias realizado en HEK293T con vectores que albergan dos S/MAR diferentes (interferón beta S/MAR; ApoB S/MAR, 805 pb) con y sin sitios SD y SA flanqueantes. pEPI es un vector plasmídico promotor del CMV con un interferón beta S/MAR 3' UTR.

Ejemplos

Los procedimientos de la tecnología actual se ilustran además con los siguientes ejemplos. Se proporcionan a modo de ilustración y no pretenden en modo alguno limitar el alcance de la divulgación.

Ejemplo 1: producción de plásmidos de origen de replicación pUC y R6K

Fondo del marcador seleccionable sin antibióticos RNA-OUT: La selección sin antibióticos se realiza en cepas de *E. coli* que contienen el sitio de unión del fago lambda integrado cromosómicamente en pCAH63-CAT RNA-IN-SacB (P5/6 6/6) como se describe en Williams, Supra, 2008. SacB (*Bacillus subtilis* levansucrase) es un marcador contraselectivo que es letal para las células de *E. coli* en presencia de sacarosa. La traducción de SacB a partir del transcrito RNA-IN-SacB es inhibida por el RNA-OUT codificado por el plásmido. Esto facilita la selección de plásmidos en presencia de sacarosa, mediante la inhibición de la letalidad mediada por SacB.

Antecedentes de replicación y producción del vector de origen R6K: El origen de replicación del plásmido R6K gamma requiere una única proteína de replicación del plásmido π que se une como monómero iniciador de la replicación a múltiples "iterones" repetidos. El uso de un origen de replicación condicional como el R6K gamma, que requiere una línea celular especializada para su propagación, añade un margen de seguridad, ya que el vector no se replicará si se transfiere a la flora endógena del paciente.

Un origen de replicación derivado de la gamma R6K altamente minimizado (SEQ ID NO:1) que contiene secuencias centrales necesarias para la replicación se describió en el documento Williams, Supra, 2014 y se incluye aquí por referencia. La columna vertebral del nanoplásmido TM NTC9385R, que incluye este origen R6K minimizado y el marcador seleccionable RNA-OUT AF en la región espaciadora, se describió en el documento Williams, Supra, 2014 y se incluye aquí por referencia. Williams, Supra, 2014 describe cepas huésped que expresan el sitio de fijación del fago HK022 integrado en el promotor pL inducible por calor π P42L, P106L y F107S proteína de replicación mutante de alta copia (Rep) para la selección y propagación de vectores NanoplasmidTM de origen R6K. Este es un factor de

seguridad adicional de Nanoplasmid™ ya que los vectores de origen R6K sólo pueden replicarse dentro de la cepa huésped de *E. coli* que expresa la proteína Rep.

Producción en matraz agitado: la producción del plásmido de origen pUC se realizó en la cepa DH5α de *E. coli* [F-Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-, mK+) phoA supE44 λ- thi-1 gyrA96 relA1] (Invitrogen, Carlsbad CA). La selección por sucrosa de los vectores Nanoplasmid™ de origen R6K se realizó en las cepas huésped NTC940211 DH5α attλ::P5/6 6/6-RNA-IN- SacB, catR; attHK022::pL (OL1-G a T) P42L-P106I-F107S o NTC1050811 DH5α attλ::P5/6 6/6-RNA-IN- SacB, catR; attHK022::pL (OL1-G a T) P42L-P106I-F107S P113S (P3-), SpecR StrepR; attφ80::pARA-CI857ts, tetR. La producción en frascos agitados se realizó utilizando el medio de cultivo agitado patentado Plasmid+. Los cultivos de semillas se iniciaron a partir de reservas o colonias de glicerol y se extendieron en placas de agar LB con 50 µg/mL de antibiótico (para plásmidos de selección ampR o kanR) o 6% de sacarosa (para plásmidos de selección RNA-OUT). Las placas se cultivaron a 30-32°C; las células se resuspendieron en medios y se utilizaron para proporcionar inóculos de aproximadamente 2,5 OD600 para los matraces de agitación Plasmid+ de 500 mL que contenían 50 µg/mL de antibiótico para los plásmidos de selección ampR o kanR o 0,5% de sacarosa para seleccionar los plásmidos RNA-OUT. Los frascos se cultivaron con agitación hasta la saturación.

Ejemplo 2: Construcción del vector S/MAR

Los vectores pNTC-NP1, pNTC-NP2, pNTC-NP3, pNTC-NP4, pNTC-NP5, pNTC-NP6, pNTC-NP7, codifican la región de selección de replicación bacteriana del origen gamma R6K-ARN-OUT (SEQ ID NO:8) clonada en la región polienlazador de un vector basado en pUC57. El vector pNTC-3xCpG NP1 codifica la región de selección de replicación bacteriana 1 CpG R6K origen gamma- 2 CpG RNA-OUT (SEQ ID NO:9) clonada en la región polienlazador de un vector basado en pUC57. Cada vector tiene diferentes sitios de restricción flanqueantes que pueden utilizarse para adaptar un vector diana a la selección R6K de replicación-ARN-OUT. Las secuencias de polienlazador 5' y 3' que flanquean el inserto R6K-RNA-OUT en los vectores pNTC-NP 1-7 y pNTC-3xCpG NP1 se muestran en la Tabla 1.

Tabla 1: Vectores marcadores de selección R6K flanqueados por sitios de clonación múltiple pNTC Origin-RNA-OUT

Vector	R6K 5' flanqueando sitios de restricción	término trpA	Origen del R6K	Sitio de enlace	ARN OUT Marcador de selección	Sitio de restricción de flanqueo de ARN-OUT 3'
pNTC-NP1 (SEQ ID NO:10)	EcoRI, SacI, KpnI, NruI, NsiI, XmaIII, NotI, <i>NheI</i>	Sí	SEQ ID NO: 1	DrallI ^a	SEQ ID NO: 5	<i>NheI</i> BamHI, XmaI, ApaI, Sall, HincII, PstI, StuI, AatI, SphI, <i>HindIII</i> (en R6K)
pNTC-NP2 (SEQ ID NO:11)	EcoRI, SacI, KpnI, NruI, NsiI, XmaIII, NotI, <i>NheI</i>	Sí	SEQ ID NO: 1	DrallI ^a	SEQ ID NO: 5	<i>SpeI</i> , <i>XmaI</i> , <i>SspI</i> BamHI, <i>XmaI</i> , ApaI, Sail, HincII, PstI, StuI, AatI, SphI, <i>HindIII</i> (en R6K)
pNTC-NP3 (SEQ ID NO:12)	EcoRI, <i>SacI</i> , <i>KpnI</i> , NruI, NsiI, XmaIII, NotI, <i>NheI</i>	Sí	SEQ ID NO: 1	DrallI ^a	SEQ ID NO: 5	<i>KpnI</i> , <i>SacI</i> BamHI, XmaI, ApaI, Sail, HincII, PstI, StuI, AatI, SphI, <i>HindIII</i> (en R6K)
pNTC-NP4 (SEQ ID NO:13)	<i>NheI</i> , XmaIII, NotI, NsiI, NruI, <i>KpnI</i> , <i>SacI</i> BamHI, XmaI, ApaI, Sail, HincII, SfcI, PstI, StuI, AatI, SphI, <i>HindIII</i> (en R6K)	Sí	SEQ ID NO: 1	DrallI ^a	SEQ ID NO: 5	EcoRI, <i>SacI</i> , <i>KpnI</i>
pNTC-NP5 (SEQ ID NO:14)	KasI, <i>NheI</i>	Sí	SEQ ID NO: 1	DrallI ^a	SEQ ID NO: 5	<i>KpnI</i> AfIII PstI, AatI, SphI, <i>HindIII</i> (en R6K)
pNTC-NP6 (SEQ ID NO:15)	EcoRI, PstI, EcoRV, BstXI, NotI, <i>NheI</i>	Sí	SEQ ID NO:	DrallI ^a	SEQ ID NO: 5	<i>KpnI</i> , ApaI, PvuI, Sall, <i>SacI</i>

Vector	R6K 5' flanqueando sitios de restricción	término trpA	Origen del R6K	Sitio de enlace	ARN OUT Marcador de selección	Sitio de restricción de flaqueo de ARN-OUT 3'
pNTC-NP7 (SEQ ID NO:16)	BssHII PacI NheI	Sí	SEQ ID NO:	Dralll ^a	SEQ ID NO: 5	KpnI PacI BssHII
pNTC-3x CpG NP1 (SEQ ID NO:17)	XhoI, XbaI, ApaI, Sall, HincII, PstI, StuI, AatI, SphI,	No	SEQ ID NO: 2	BsrGI	SEQ ID NO: 7	EcoRI, SacI, <i>KpnI</i> , NruI, NsiI, XmaII, NotI, NheI, <i>KpnI</i>

a un sitio Dralll de extremo pegajoso NNN no palindrómico único de 3 pb (CACNNNGTG) que separa la R6K y el ARN-OUT de la secuencia CACGTTGTG puede utilizarse para ensamblar la R6K y el ARN-OUT a partir de vectores bpNTC separados en reacciones de ligación direccional de fragmentos múltiples. El vector S/MAR pUC origen-selección antibiótica retroalimentación bacteriana a la R6K-ARN-OUT (es decir, Los vectores nanoplásmidos, NP, fueron realizados por:

- 1) Selección de los sitios de restricción que flanquean el origen pUC y la región del marcador de selección de antibióticos en el vector S/MAR diana;
- 2) Identificación de un casete polienlazador -R6K-RNA-OUT polienlazador compatible con pNTC-NP (ya sea pNTC-NP1, 2, 3, 4, 5, 6 o 7; Tabla 1);
- 3) Extirpación de la región del marcador de selección de antibióticos del origen pUC y reemplazarla por la región seleccionada del origen R6K RNA-OUT utilizando el enfoque de digestión de restricción seleccionado y la clonación mediada por ligasa estándar.

En algunos casos, el origen R6K y las unidades de ARN-OUT se ensamblaron en ligaciones multifragmentarias a partir de fragmentos de restricción separados utilizando el sitio enlazador Dralll no palindrómico (véase la Tabla 1).

- 5 Ejemplos de mapas vectoriales y características del vector marcador de selección original pUC origen-antibiótico (por ejemplo, pSMART UCOE; Figura 5) y el vector marcador de selección libre de antibióticos R6K adaptado origen-ARN-OUT (por ejemplo, NP-UCOE; Figura 6).

- 10 Los SD -S/MAR-SA 3' UTRs se hicieron como genes sintéticos de la siguiente manera. Un sitio donante de empalme (SEQ ID NO: 25) con sitios de clonación 5' BglII y NsiI y un sitio de clonación 3' XhoI (Figura 1) fue incorporado 5' a la S/MAR, mientras que un sitio aceptor de empalme (SEQ ID NO: 26) con sitios de clonación 5' EcoRI y 3' BamHI (Figura 1) se incorporó 3' a la S/MAR. Los genes se sintetizaron en Genscript (Piscataway, NJ) y se clonaron en lugar de los S/MAR en los vectores SMAR-NP existentes utilizando la clonación estándar mediada por ligadura de fragmentos de restricción. Por ejemplo, el interferón beta S/MAR (SEQ ID NO: 18) (por ejemplo, el vector NP-UCOE, Figura 6) fue sustituido por el aceptor de empalme donante-interferón beta S/MAR (SEQ ID NO:19) (por ejemplo, el vector NP-UCOE-SP NP-UCOE-SP vector, Figura 6; NP-UCOE-EF1-SP, Figura 8) o aceptor de empalme donante-interferón beta S/MAR (-AATAAA) (SEQ ID NO:20) o aceptor de empalme donante-interferón beta M18 S/MAR (SEQ ID NO:21). Se diseñó el empalme donante-interferón beta S/MAR (-AATAAA) para eliminar las señales de terminación de la transcripción AATAAA(N) codificadas en S/MAR con un motivo AATATT(T) MAR (Figura 2). La ApoB S/MAR de 805 pb fue sustituida por la versión donante de empalme-805 pb ApoB S/MAR-aceptante de empalme (SEQ ID NO: 22) (por ejemplo, NP-SMARter-SP, Figura 7) mientras que la ApoB S/MAR de 525 pb fue sustituida por la versión donante de empalme-525 pb de ApoB S/MAR-aceptante de empalme (SEQ ID NO: 23) (por ejemplo, NP-CMARter-SP, figura 7; NP-Ele40-CMARter-SP, figura 9). Se realizaron construcciones NP adicionales con transgenes alternativos, promotores, intrones 5' UTR o elementos ELE40 o UCOE mediante clonación mediada por ligadura de fragmentos de restricción estándar. Todas las construcciones se verificaron correctamente mediante digestión de restricción y secuenciación.

25 Ejemplo 3: Expresión del vector S/MAR tras la transfección transitoria

Las líneas celulares adherentes HEK293 (riñón embrionario humano) y A549 (carcinoma de pulmón humano) se obtuvieron de la American Type Culture Collection (Manassas, VA, USA). Las líneas celulares se propagaron en medio Eagle modificado de Dulbecco/F12 con un 10% de suero bovino fetal

- 30 y dividido (0,25% tripsina-EDTA) utilizando reactivos de Invitrogen (Carlsbad, CA, USA) y metodologías convencionales. Para las transfecciones, las células se colocaron en placas de cultivo de tejidos de 24 pocillos. Los plásmidos se transfectaron en las líneas celulares utilizando Lipofectamine 2000 siguiendo las instrucciones del fabricante (Invitrogen).

Los lisados celulares totales para la determinación de la EGFP se prepararon resuspendiendo las células en tampón de lisis celular (CellLytic M, Sigma, St Louis, MO, EE.UU.), lisando las células mediante incubación durante 30 minutos a 37°C, seguido de un ciclo de congelación y descongelación a -80 °C. Las células lisadas se clarificaron por centrifugación y los sobrenadantes se analizaron para EGFP mediante un lector de fluorescencia de microplacas FLX800 (Bio-Tek, Winooski, VT, EE.UU.). Los resultados se resumen en las tablas 2-4.

5

Tabla 2: Expresión transitoria de los vectores S/MAR tras la transfección en las líneas celulares A549 y HEK293

Plásmido	Promotor	Intron	3-UTR	A549 GFP ^a	HEK GFP ^a
NTC9385R- UCOE EF1 - coGFP	UCOE EF1	Ninguno	hIFNB SMAR- SV40	552 ± 95	5168 ± 202
NTC9385R- UCOE EF1 - coGFP SD -SMAR SA SV40 pA	UCOE EF1	Ninguno	SD hIFNB SMAR SA - SV40 pA	1139 ± 181	13909 ± 1068
NTC9385R- UCOE EF1-coGFP- SMAR R6K-ROUT- pA	UCOE EF1	Ninguno	hIFNB SMAR-R6K-R-OUT RBG	607 ± 217	7552 ± 1754
NTC9385R- UCOE EF1-coGFP- SD SMAR R6K-R-OUT-SA pA (UCOE-EF1-SP-NP-Figura 8)	UCOE EF1	Ninguno	SD hIFNB SMAR-R6K-R-OUT SA RBG pA	961 ± 83	12956 ± 848
NTC9385R- UCOE EF1-coGFP-SD M18 SMAR R6K-R-OUT-	UCOE EF1	Ninguno	SD M18 SMAR-R6K-R-OUT SA RBG pA	2088 ± 449	16761 ± 954
NTC9385R- UCOE EF1 - coGFP SD -M18 SMAR SA SV40 pA	UCOE EF1	Ninguno	SD M18 SMAR SA-SV40 pA	3190 ± 386	22640 ± 1129
a Los resultados presentados son unidades fluorescentes medias ± desviación estándar a los 2 días después de la transfección					

Tabla 3: Expresión transitoria de los vectores S/MAR tras la transfección en las líneas celulares A549 y HEK293

Plásmido	Promotor	Intron	3-UTR	A549 GFP	HEK GFP
NTC9385R- EF1-coGFP - SMAR SV40 pA	EF1	Ninguno	hIFNB SMAR-SV40 pA	1221 ±	2038 ±
NTC9385R- UCOE EF1 - coGFP -SMAR SV40 pA	UCOE EF1	Ninguno	hIFNB SMAR-SV40 pA	2251 ±	7339 ±
NTC9385R- UCOE EF1 - coGFP SD -SMAR SA SV40 pA (NP-UCOE-EF1-SP-Figura 8)	UCOE EF1	Ninguno	SD hIFNB SMAR SA - SV40 pA	6205 ± 420	24507 ± 2501
NTC9385R- UCOE EF1 - coGFP SD -SMAR(-AATAAA) SA SV40 pA	UCOE EF1	Ninguno	SD hIFNB SMAR-AATAAA SA - SV40 pA	4708 ± 359	18910 ± 1278
NTC9385R- EF1-coGFP-SMAR R6K-R-OUT- pA	EF1	Ninguno	hIFNB SMAR-R6K-R-OUT RBG pA	1240 ± 164	1896 ± 189
NTC9385R- UCOE EF1-coGFP-SMAR R6K-R-OUT-pA	UCOE EF1	Ninguno	hIFNB SMAR-R6K-R-OUT RBG pA	1540 ± 180	4996 ± 322
NTC9385R- UCOE EF1-coGFP-SD SMAR R6K-R-OUT-SA pA (UCOE-EF1-SP-NP-Figura 8)	UCOE EF1	Ninguno	SD hIFNB SMAR-R6K-R-OUT SA RBG pA	4843 ± 604	19247 ± 1693

Plásmido	Promotor	Intron	3-UTR	A549 GFP	HEK GFP
NTC9385R- UCOE EF1-coGFP-SD M18 SMAR R6K-R-OUT-SA pA	UCOE EF1	Ninguno	SD M18 SMAR-R6K-R-OUT SA RBG pA	10021 ± 753	27981 ± 1121
NTC9385R- UCOE EF1-coGFP SD -M18 SMAR SA SV40 pA	UCOE EF1	Ninguno	SD M18 SMAR SA - SV40 pA	9751 ± 821	29019 ± 2744
NTC9385R-UCOE-CMV-coGFP P2A-PuroR - SMAR-SV40 pA (NP-UCOE -Figura 6)	UCOE CMV	pCI	hIFNB SMAR-SV40 pA	2104 ± 74	8478 ± 320
NTC9385R-UCOE-CMV-coGFP P2A-PuroR - SD SMAR- SA SV40 pA (NP-UCOE-SP -Figura 6)	UCOE CMV	pCI	SD hIFNB SMAR SA - SV40 pA	3526 ± 102	14278 ± 2664
NTC9385R-UCOE-CMV-coGFP P2A-PuroR - SD SMAR(-AATAAA)- SA SV40 pA	UCOE CMV	pCI	SD hIFNB SMAR-AATAAA SA - SV40 pA	2876 ± 376	13425 ± 1331
a Los resultados presentados son unidades fluorescentes medias ± desviación estándar a los 2 días después de la transfección					

Los resultados presentados en las Tablas 2 y 3 demuestran que con un promotor UCOE-EF1 sin intrón coGFP la unidad de transcripción del IFNB SMAR humano flanqueado por SD/SA mejora la expresión en las líneas celulares HEK293 y A549 en comparación con el IFNB SMAR humano sin sitios SD/SA. Se observó una mejora de la expresión en 2 configuraciones SD/SA (flanqueo de SMAR, o flanqueo de SMAR+R6K-ARN-OUT región bacteriana). El SMAR M18 (derivado del SMAR IFNB humano) flanqueado por SD/SA tiene mayor expresión que el SMAR IFNB humano equivalente flanqueado por SD/SA.

Además, los resultados de la Tabla 3 muestran una expresión mejorada en la unidad de transcripción del transgén UCOE-CMV pCI intrón coGFP (es decir, una expresión mejorada con dos promotores diferentes, con o sin un intrón codificado en 5' UTR). También se observa una mejor expresión con diferentes señales de poliadenilación (derivadas de SV40 o RBG).

Tabla 4: Expresión transitoria de los vectores S/MAR tras la transfección en las líneas celulares A549 y HEK293

Plásmido	Promotor	5' UTR Intron	3' UTR	T=2 días A549 GFP	T=2 días HEK GFP
NTC9385R- EF1-coGFP - SMAR SV40 pA	EF1	Ninguno	hIFNB SMAR-SV40	525 ± 37	1377 ± 111
NTC9385R- UCOE EF1 - coGFP -SMAR SV40 pA	UCOE EF1	Ninguno	hIFNB SMAR-SV40	1848 ± 163	12980 ± 1005
NTC9385R- UCOE EF1-coGFP SD - SMAR SA SV40 pA	UCOE EF1	Ninguno	SD hIFNB SMARSA -SV40 pA	3091 ± 169	22354 ± 1686
NTC9385R- UCOE EF1-coGFP SD - SMAR(-AATAAA) SA SV40 pA	UCOE EF1	Ninguno	SD hIFNB SMAR-AATAAA SA -	2311 ± 413	14768 ± 1628
NTC9385R- UCOE EF1 -coGFP SD-	UCOE EF1	Ninguno	SD M18 SMAR SA-	4833 ± 462	21254 ± 6296
NTC9385R-SP-UCOE-EF1-GFP SMARter =	UCOE EF1	Ninguno	SD SMARter SA - SV40	2878 ± 233	13688 ± 1873
NC9385R-SP-Ele40-EF1-GFP SMARter = coGFP	ELE40 EF1	Ninguno	SD SMARter SA - SV40 pA	990 ± 175	3349 ± 341

Plásmido	Promotor	5' UTR Intron	3' UTR	T=2 días A549 GFP	T=2 días HEK GFP
MAX-UCOE-coGFP P2A-PuroR-NP	UCOE CMV	pCI	SMAR-SV40 pA	933 ± 117	6193 ± 533
NTC9385R-UCOE-CMV-coGFP P2A-PuroR -SMAR-SV40 pA	UCOE CMV	pCI	SMAR-SV40 pA	1081 ± 85	8216 ± 211
NTC9385R-UCOE-CMV-coGFP P2A-PuroR - SD SMAR- SA SV40 pA	UCOE CMV	pCI	SD SMAR SA-SV40 pA	1857 ± 207	12596 ± 1531
NTC9385R-UCOE-CMV-coGFP P2A-PuroR - SD SMAR(-AATAAA)- SA SV40 pA	UCOE CMV	pCI	SD SMAR-AATAAA SA-SV40 pA	2204 ± 70	13901 ± 1024
NTC9385R-SP-UCOE-CMV-GFP SMARter = coGFP P2A-PuroR (NP-	UCOE CMV	pCI	SD SMARter SA - SV40 pA	917 ± 113	8091 ± 449
NTC9385R-SP-UCOE-CMV-GFP CMARter = coGFP P2A-PuroR (NP-	UCOE CMV	pCI	SD CMARter SA-SV40 pA	3875 ± 230	12020 ± 624
NTC9385R-SP-Ele40-CMV-GFP SMARter = coGFP P2A-PuroR	ELE40 CMV	pCI	SD SMARter SA - SV40 pA	1524 ± 59	5483 ± 393
a Los resultados presentados son unidades fluorescentes medias ± desviación estándar a los 2 días después de la transfección					

Los resultados presentados en la Tabla 4 demuestran además que el IFNB SMAR humano flanqueado por SD/SA mejora la expresión tanto en las líneas celulares HEK293 como en las A549 en comparación con el IFNB SMAR humano sin sitio SD/SA con la unidad de transcripción del transgén UCOE-EF1 sin intrón coGFP y con la unidad de transcripción del transgén UCOE-CMV pCI intrón coGFP (es decir, expresión mejorada con dos promotores diferentes, con o sin un intrón codificado en 5' UTR). Además, el SMAR de CMARter flanqueado por SD/SA tiene mayor expresión que el SMAR de IFNB humano flanqueado por SD/SA. Además, la sustitución de las señales de terminación de la transcripción S/MAR AATAAA(N) por un motivo MAR AATATT(T) dio lugar a un S/MAR funcional, demostrando que este enfoque puede utilizarse para eliminar las señales de terminación de la transcripción de los elementos S/MAR descritos en el arte. Los motivos alternativos pueden ser sustituidos por AATATT(T), por ejemplo, los motivos ricos en AT enriquecidos en S/MARs como se describe en Liebeich et al., Supra, 2002. Este procedimiento de sustitución del motivo AATAAA permite la adaptación de los S/MARs en el arte para ser utilizados en las UTRs 3' de la invención, sin reducir la expresión a través de la terminación prematura de la transcripción mediada por el motivo AATAAA.

En conjunto, los resultados demuestran que los vectores de la presente invención resuelven la limitación del nivel de expresión subóptimo de los vectores basados en S/MAR descritos en el arte.

Ejemplo 4: Expresión del vector S/MAR tras el establecimiento del episoma

Se determinó la expresión de NP-UCOE (Figura 6) y NP-UCOE-SP (Figura 6) tras el establecimiento episomal en la línea celular HEK293. Las células se establecieron con los protocolos estándar que requerían la aplicación de puomicina (0,5 µg/ml) durante una semana antes de la expansión durante al menos 30 días (Wong y Harbottle, 2013 Mol Ther Nucleic Acids 2:e115). Las poblaciones establecidas se analizaron para la expresión del gen reportero GFP mediante FACS y los niveles de ARN de GFP se evaluaron mediante qPCR. Los resultados (Figura 10) demuestran que el SMAR de IFNB humano flanqueado por SD/SA mejora la transcripción del ARNm y la expresión del transgén GFP en comparación con el SMAR de IFNB humano sin el sitio SD/SA tras el establecimiento episomal en la línea celular HEK293. Un segundo experimento demostró que la expresión del transgén GFP de los vectores SD-S/MAR-SA NP-UCOE-SP (Figura 6), NP- SMARter-SP (Figura 7) y NP-CMARter-SP (Figura 7) mejoró en comparación con el vector no SD-SA NP-UCOE (Figura 6) tras el establecimiento episomal en la línea celular HEK293 y en células primarias de fibroblastos embrionarios de ratón.

Estos resultados con líneas celulares establecidas demuestran que los vectores de la presente invención resuelven la limitación de silenciamiento de genes de los vectores basados en S/MAR descritos en el arte.

Ejemplo 5: Expresión del vector S/MAR tras el establecimiento del episoma

La eficacia en el establecimiento de células también se probó en HEK293T mediante el ensayo de formación de colonias (Wong y Harbottle, Supra, 2013) con vectores que albergan dos S/MAR diferentes (interferón beta S/MAR; ApoB S/MAR, 805 pb) con y sin sitios SD y SA flanqueantes. Los resultados demostraron (Figura 12) que tanto con el interferón beta S/MAR como con el ApoB S/MAR flanqueando los sitios SD y SA se mejoraba drásticamente la eficacia en la generación de células establecidas (es decir, produciendo el mayor número de colonias). Estos resultados demuestran que los vectores de la presente invención resuelven la limitación de la baja tasa de establecimiento de los vectores basados en S/MAR descritos en la técnica.

Ejemplo 6: Eficiencia del establecimiento y análisis de la población de células modificadas genéticamente (Fig. 1)

La eficacia en la generación de células con expresión estable se evaluó en un ensayo de formación de colonias utilizando pS/MARt (Fig. 4, SEQ ID NO:41). Tras la entrega del ADN, se aislaron las células positivas para la expresión del transgén GFP mediante clasificación FACS (FACS Aria II) y se colocaron 100 células en una placa de cultivo celular de 6 cm. A continuación, se cultivaron durante 4 semanas en presencia de 0,5 µg/ml de puromicina. Después de 4 semanas, las células se fijaron con PFA y las colonias se tiñeron con Crystal Violet. El número de colonias se considera como la eficiencia del establecimiento del vector, es decir, el número de colonias que se forman por el número de células clasificadas por FACS que se colocan. La generación de líneas celulares estables es muy eficaz, ya que más del 40% de las células transfectadas se establecen (Fig. 1A)). El número de células que expresan el transgén (GFP) se estimó mediante citometría de flujo. Como se muestra en la Fig. 1 b), pS/MARt genera poblaciones modificadas en las que la expresión del transgén es homogénea, sin un número significativo de células negativas.

Ejemplo 7: Rescate de plásmidos de vectores pS/MARt a partir de poblaciones celulares establecidas (Fig. 2)

Un procedimiento eficaz para determinar si los vectores de ADN se mantienen episomalmente con integridad dentro de las células modificadas es verificar que pueden ser rescatados intactos en bacterias ingenuas. Para ello, las líneas celulares establecidas de forma persistente y modificadas con el plásmido pS/MARt se cultivaron en presencia del antibiótico Puromicina (0,5 µg/ml) durante 1 semana y se expandieron durante al menos 30 días sin antibiótico para evaluar la integridad del vector. El ADN total se preparó a partir de las células utilizando el kit Blood&Tissue DNAeasy (Qiagen) y se transformó en células E. coli DH10B. Las bacterias se cultivaron en placas LB-Agar con Kanamicina (50µg/ml). Se cultivaron 12 colonias en medio LB líquido con Kanamicina (50 µg/ml) durante una noche y se extrajo el ADN plasmídico con el MiniprepKit (Qiagen). Para el análisis, las minipreparaciones de ADN se digirieron con la enzima de restricción BamHI (Thermo Fisher) durante 10 minutos a 37 °C y el patrón de restricción se abordó en un gel de agarosa al 1%. Como control, el ADN utilizado para la transfección de las células al principio del procedimiento de establecimiento se digirió con la misma enzima y se utilizó como referencia. Estos geles ilustran que se pudo aislar ADN pS/MARt intacto de líneas celulares modificadas estables y que en todos los casos el ADN era idéntico al de los vectores transfectados originalmente.

Ejemplo 8: los vectores pS/MARt se mantienen episódicamente en células modificadas (Fig. 3)

Para demostrar aún más que los vectores pS/MARt modificaban las células de mamífero como un episoma, se determinó físicamente la estructura mediante el análisis de Southern Blot. Se analizaron poblaciones de células Hek293T cultivadas durante al menos 30 días después de la transfección de ADN. El ADN genómico se extrajo con el kit Blood&Tissue DNAeasy (Qiagen) y se digirió durante la noche a 37°C con la enzima de restricción BamHI (NEB). A continuación, el ADN celular total se separó en un gel de agarosa al 1% y se transfirió a una membrana de nailon. Se utilizaron oligonucleótidos correspondientes al gen GFP del vector para generar la sonda radiactiva utilizada para detectar el ADN pS/MARt dentro del ADN celular. La presencia en las muestras de una sola banda que tiene el mismo tamaño del vector de control demuestra el estado episomal de pS/MARt en las poblaciones de células de mamíferos establecidas. La ausencia de manchas y/o bandas alternativas demuestra que los vectores no se reorganizaron ni se integraron en el genoma celular.

Los procedimientos y composiciones de vectores aquí divulgados y las evaluaciones presentadas anteriormente demuestran que las composiciones SD-SMAR-SA de 3' UTR mejoran la expresión y/o el establecimiento episomal en comparación con las versiones no SD-SA. La mejora del rendimiento no es específica de la S/MAR, ya que se observa una mejora del rendimiento con varias S/MAR. La mejora del rendimiento tampoco es específica de la unidad de transcripción del vector, ya que la mejora del rendimiento se observa con SD-SMAR-SA vinculado a varios promotores, 5' UTRs, transgenes y señales polyA. Se observa un mejor rendimiento con o sin intrones aguas arriba. Por lo tanto, los vectores SD-SMAR-SA 3' UTR de la divulgación son ampliamente aplicables para mejorar el rendimiento de los vectores de expresión de vertebrados episomales no integradores autorreplicantes.

Los vectores de la tecnología actual cualquier sitio donante de empalme intrónico descrito en el arte podría ser sustituido por el donante de empalme derivado del intrón pCI. Asimismo, cualquier sitio aceptor de empalme intrónico descrito en la técnica podría ser sustituido por el aceptor de empalme derivado del intrón pCI. Por ejemplo, los donantes y aceptores de empalme pueden derivarse del intrón híbrido HTLV- IR-Conejo β globina, intrón híbrido HTLV- IR CMV, intrón CMV, intrón libre de CpG I 140, intrón quimérico de β globina humana-IgG murina, Intrón quimérico de IgG murina con líder de adenovirus, intrón de globina β de conejo, intrón truncado de CMV, intrón CAG

(globina β de pollo y conejo), intrón híbrido de globina β de conejo y CMV divulgado en Williams, Supra, 2014 u otros intrones descritos en la técnica.

Los diversos S/MAR alternativos descritos en el arte también podrían utilizarse en los vectores de la tecnología actual. Si es necesario, los sitios internos de poliA pueden eliminarse mediante la sustitución de motivos, como se describe en el presente documento.

Los vectores pueden codificar una diversidad de transgenes diferentes de los ejemplos proporcionados en el presente documento, por ejemplo, genes de antígenos para una variedad de patógenos, o genes terapéuticos como el factor inducible por hipoxia, el factor de crecimiento de queratinocitos, el factor IX, el factor VIII, la proteína del grupo A de complementación de la anemia de Fanconi, la homogentisato dioxigenasa, etc. o poliproteínas como una poliproteína del factor de reprogramación.

Asimismo, los vectores pueden utilizar una diversidad de promotores de ARN Pol II diferentes de los ejemplos de promotores de CMV y del factor de elongación 1 (EF1) proporcionados en el presente documento, por ejemplo, promotores constitutivos como el promotor de β -actina de pollo, el promotor de β -actina de otras especies el promotor de la fosfogliceroquinasa (PGK), el promotor del virus formador de focos del bazo (SFFV), el promotor del virus del sarcoma de Rous (RSV), el promotor de la seroalbúmina humana (SA), el promotor de la globulina fijadora de tiroxina (TBG), el promotor del citocromo P450 2E1 (CYP2E1), etc. Los vectores también pueden utilizar promotores combinados, como el promotor de la β -actina del pollo/el potenciador del CMV (CAG), los elementos potenciadores derivados del CMV humano o murino combinados con los promotores del factor de elongación 1 α (EF1 α), las versiones sin CpG de los elementos potenciadores derivados del CMV humano o murino combinados con los promotores del factor de elongación 1 α (EF1 α), el promotor de la albúmina combinado con un potenciador de la α -fetoproteína MERII, etc, o la diversidad de promotores específicos de tejido o inducibles conocidos en la técnica como los promotores específicos de músculo creatina quinasa (MCK) y C5-12 o los promotores específicos del hígado apolipoproteína A-I (ApoAI), promotor de la α -1 antitripsina (AAT), promotor de la AAT-TTR, promotor de la SERP-TTR y ApoE-hAAT, o promotores de células T como hTCR81, CD4 y WASp. Además, en la espina dorsal del nanoplásmido, se pueden utilizar varias orientaciones del origen de replicación R6K y el marcador seleccionable de ARN. Por ejemplo, puede utilizarse cualquiera de las ocho orientaciones del origen de replicación R6K, y el marcador seleccionable de ARN en los vectores de la tecnología actual (es decir, $\leftarrow$ Origen de replicación de Pol III RSM $\rightarrow$; origen de replicación de Pol III $\leftarrow$ RSM; origen de replicación de Pol III $\rightarrow$ RSM $\rightarrow$; origen de replicación de Pol III $\rightarrow$ $\leftarrow$ RSM; $\leftarrow$ RSM $\leftarrow$ origen de replicación de Pol III; RSM $\rightarrow$ origen de replicación de Pol III $\rightarrow$; RSM $\rightarrow$ $\leftarrow$ Origen de replicación de Pol III). Además, una variedad de marcadores seleccionables de ARN conocidos en el arte pueden ser sustituidos por ARN-OUT. Así, el lector verá que los vectores de expresión vertebrados episomales no integradores mejorados de la tecnología actual proporcionan un enfoque para mejorar la expresión de transgenes codificados por plásmidos de replicación episomal no integrativa.

Por consiguiente, el alcance de la divulgación debe determinarse no por las realizaciones ilustradas, sino por las reivindicaciones adjuntas.

LISTA DE SECUENCIAS

<110> Deutsches Krebsforschungszentrum
Nature Technology Coropration
Bozza, Matthias
Harbottle, Richard
Williams, James

<120> VECTORES DE ADN NO INTEGRADORES PARA MODIFICACIÓN GENÉTICA DE CÉLULAS

<130> DK14791PC1

<160> 41

<170> BiSSAP 1.3.6

<210> 1

<211> 281

<212> ADN

<213> Secuencia artificial

<220>

<223> R6K origen gamma

<400>1

ES 2 887 776 T3

	ggcttgttgt ccacaaccgt taaaccttaa aagctttaaa agccttatat attctttttt	60
	ttcttataaa acttaaaacc ttagaggcta tttaagttgc tgatttatat taattttatt	120
	gttcaaacat gagagcttag tacgtgaaac atgagagctt agtacgtag ccatgagagc	180
	ttagtacgtt agccatgagg gtttagttcg ttaaacaatga gagcttagta cgttaaacaat	240
	gagagcttag tacgtactat caacaggttg aactgctgat c	281
	<210> 2	
	<211> 281	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> 1 CpG R6K origen gamma	
	<400>2	
	ggcttgttgt ccacaaccat taaaccttaa aagctttaaa agccttatat attctttttt	60
	ttcttataaa acttaaaacc ttagaggcta tttaagttgc tgatttatat taattttatt	120
	gttcaaacat gagagcttag tacgtgaaac atgagagctt agtacattag ccatgagagc	180
	ttagtacatt agccatgagg gtttagttca ttaaacaatga gagcttagta cattaacaat	240
	gagagcttag tacatactat caacaggttg aactgctgat c	281
10	<210> 3	
	<211> 260	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> CpG libre R6K origen gamma	
	<400>3	
	aaaccttaaa accttttaaaa gccttatata ttcttttttt tcttataaaa cttaaaacct	60
	tagaggctat ttaagttgct gatttatatt aattttattg ttcaaacaatg agagcttagt	120
	acatgaaaca tgagagctta gtacattagc catgagagct tagtacatta gccatgaggg	180
	tttagttcat taaacaatgag agcttagtac attaaacaatg agagcttagt acatactatc	240
	aacaggttga actgctgatc	260
	<210> 4	
	<211> 389	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Origen gamma R6K ampliado	
	<400>4	

ES 2 887 776 T3

	tgtcagccgt taagtgttcc tgtgtcactg aaaattgctt tgagaggctc taagggcttc	60
	tcagtgcgtt acatccctgg cttgttgtcc acaaccgtta aaccttaaaa gctttaaaag	120
	ccttatatat tctttttttt cttataaaac ttaaaacctt agaggctatt taagttgctg	180
	atttatatta attttattgt tcaaaccatga gagcttagta cgtgaaacat gagagcttag	240
	tacgttagcc atgagagctt agtacgttag ccatgagggt ttagttcgtt aaacatgaga	300
	gcttagtacg ttaaacatga gagcttagta cgtgaaacat gagagcttag tacgtactat	360
	caacagggtg aactgctgat cttcagatc	389
5	<210> 5 <211> 139 <212> ADN <213> Secuencia artificial	
	<220> <223> Marcador seleccionable RNA-OUT	
	<400>5	
	gtagaattgg taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta	60
	ttgatTTTTg gcgaaaccat ttgatcatat gacaagatgt gtatctacct taacttaatg	120
	atTTtgataa aaatcatta	139
10	<210> 6 <211> 69 <212> ADN <213> Secuencia artificial	
	<220> <223> ARN-OUT antisentido represor	
15	<400>6	
	tcgcacatct tgttgtctga ttattgattt ttggcgaaac catttgatca tatgacaaga	60
	tgtgtatct	69
20	<210> 7 <211> 139 <212> ADN <213> Secuencia artificial	
	<220> <223> 2 Marcador seleccionable CpG RNA-OUT	
	<400>7	
	gtagaattgg taaagagagt tgtgtaaaat attgagttcg cacatcttgt tgtctgatta	60
	ttgatTTTTg gcgaaaccat ttgatcatat gacaagatgt gtatctacct taacttaatg	120
25	atTTtgataa aaatcatta	139
	<210> 8 <211> 466 <212> ADN <213> Secuencia artificial	
30	<220> <223> Origen gamma R6K -RNA-OUT región de replicación-selección bacteriana flanqueada por sitios de restricción NheI y KpnI	
	<400>8	

gctagcccg ctaatgagcg ggcttttttt tggtttgttg tccacaaccg ttaaaccctta 60
aaagctttta aagccttata tattcttttt tttcttataa aacttaaaac cttagaggct 120
atttaagttg ctgatttata ttaattttat tgttcaaaca tgagagctta gtacgtgaaa 180
catgagagct tagtacgtta gccatgagag cttagtagct tagccatgag ggttttagttc 240
gttaaacatg agagcttagt acgttaaaca tgagagctta gtacgtacta tcaacagggtt 300
gaactgctga tccacgttgt ggtagaattg gtaaagagag tcgtgtaaaa tatcgagttc 360
gcacatcttg ttgtctgatt attgattttt ggcgaaacca tttgatcata tgacaagatg 420
tgtatctacc ttaacttaat gattttgata aaaatcatta ggtacc 466

<210> 9

<211> 439

<212> ADN

<213> Secuencia artificial

<220>

<223> 1 origen gamma CpG R6K - 2 región de selección de replicación bacteriana CpG RNA-OUT
flanqueada por sitios de restricción NheI y KpnI

<400>9

gctagctggc ttgttgtcca caaccattaa accttaaaag ctttaaaagc cttatatatt 60
cttttttttc ttataaaact taaaaccta gaggtatatt aagttgctga tttatattaa 120
ttttattggt caaacatgag agcttagtac gtgaaacatg agagcttagt acattagcca 180
tgagagctta gtacattagc catgagggtt tagttcatta aacatgagag cttagtagat 240
taaacatgag agcttagtac atactatcaa caggttgaac tgctgatctg tacagtagaa 300
ttggtaaaga gagttgtgta aaatattgag ttcgcacatc ttgttgtctg attattgatt 360
tttggcgaaa ccatttgatc atatgacaag atgtgtatct accttaactt aatgattttg 420
ataaaaaatca ttaggtacc 439

<210> 10

<211> 565

<212> ADN

<213> Secuencia artificial

<220>

<223> Casete de clonación del polienlazador pNTC-NP1 polienlazador trpA R6K-RNA-OUT: EcoRI-HindIII

<400>10

ES 2 887 776 T3

```

gaattcgagc tcggtacctc gcgaatgcat ctaggggacg gccgctagcc cgcctaataga      60
gcgggcgtttt ttttggttg ttgtccacaa ccgttaaacc ttaaaagctt taaaagcctt      120
atatattctt ttttttctta taaaacttaa aaccttagag gctatttaag ttgctgattt      180
atattaattt tattgttcaa acatgagagc ttagtacgtg aaacatgaga gcttagtacg      240
ttagccatga gagcttagta cgtagccat gagggtttag ttcgttaaac atgagagctt      300
agtacgttaa acatgagagc ttagtacgta ctatcaacag gttgaactgc tgatccacgt      360
tgtggtagaa ttggtaaaga gagtcgtgta aaatatcgag ttgcacatc ttgttgtctg      420
attattgatt tttggcgaaa ccatttgatc atatgacaag atgtgtatct accttaactt      480
aatgattttg ataaaaatca ttaggagcta gcattgggtc atcggatccc gggcccgtcg      540
actgcagagg cctgcatgca agctt                                           565

```

<210> 11

<211> 584

<212> ADN

<213> Secuencia artificial

<220>

<223> casete de clonación polienlizador pNTC-NP2 polienlizador trpA R6K-RNA-OUT: EcoRI-HindIII

<400>11

```

gaattcgagc tcggtacctc gcgaatgcat ctaggggacg gccgctagcc cgcctaataga      60
gcgggcgtttt ttttggttg ttgtccacaa ccgttaaacc ttaaaagctt taaaagcctt      120
atatattctt ttttttctta taaaacttaa aaccttagag gctatttaag ttgctgattt      180
atattaattt tattgttcaa acatgagagc ttagtacgtg aaacatgaga gcttagtacg      240
ttagccatga gagcttagta cgtagccat gagggtttag ttcgttaaac atgagagctt      300
agtacgttaa acatgagagc ttagtacgta ctatcaacag gttgaactgc tgatccacgt      360
tgtggtagaa ttggtaaaga gagtcgtgta aaatatcgag ttgcacatc ttgttgtctg      420
attattgatt tttggcgaaa ccatttgatc atatgacaag atgtgtatct accttaactt      480
aatgattttg ataaaaatca ttaggactag tcccgggcgc tagttattaa tattgggtca      540
tcggatcccc ggcccgtcga ctgcagaggc ctgcatgcaa gctt                                           584

```

<210> 12

<211> 557

<212> ADN

<213> Secuencia artificial

<220>

<223> casete de clonación polienlizador pNTC-NP3 polienlizador trpA R6K-RNA-OUT: EcoRI-HindIII

<400>12

ES 2 887 776 T3

```

gaattcgagc tccgtacctc gcgaatgcat ctaggggacg gccgctagcc cgcctaataga      60
gcggggctttt ttttggttg ttgtccacaa ccgttaaacc ttaaaagctt taaaagcctt      120
atatattctt ttttttctta taaaacttaa aaccttagag gctatttaag ttgctgattt      180
atattaattt tattgttcaa acatgagagc ttagtacgtg aaacatgaga gcttagtacg      240
ttagccatga gagcttagta cgttagccat gaggggttag ttcgttaaac atgagagctt      300
agtacgttaa acatgagagc ttagtacgtg ctatcaacag gttgaactgc tgatccacgt      360
tgtggtagaa ttggtaaaga gagtcgtgta aaatatcgag ttgcacatc ttgttgtctg      420
attattgatt tttggcgaaa ccatttgatc atatgacaag atgtgtatct accttaactt      480
aatgattttg ataaaaatca ttaggtaccg agctcggatc cggggcccg cgaactgcaga      540
ggcctgcatg caagctt

```

<210> 13

<211> 557

<212> ADN

<213> Secuencia artificial

<220>

<223> casete de clonación polienlizador pNTC-NP4 polienlizador trpA R6K-RNA-OUT: HindIII-EcoRI

<400>13

```

aagcttgcat gcaggcctct gcagtcgacg ggcccgggat ccgagctcgg tacctcgca      60
atgcatctag gggacggccg ctagcccgcc taatgagcgg gctttttttt ggcttggtgt      120
ccacaaccgt taaaccttaa aagctttaa agccttataat attctttttt ttcttataaa      180
acttaaaacc ttagaggcta tttaagttgc tgatttatat taattttatt gttcaaaccat      240
gagagcttag tacgtgaaac atgagagctt agtacgttag ccatgagagc ttagtacgtt      300
agccatgagg gtttagttcg ttaaaccatga gagcttagta cgttaaaccat gagagcttag      360
tacgtactat caacagggtg aactgctgat ccacgttgtg gtagaattgg taaagagagt      420
cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta ttgatttttg gcgaaaccat      480
ttgatcatat gacaagatgt gtatctacct taacttaatg attttgataa aaatcattag      540
gtaccgagct cgaattc

```

<210> 14

<211> 503

<212> ADN

<213> Secuencia artificial

<220>

<223> casete de clonación polienlizador pNTC-NP5 polienlizador trpA R6K-RNA-OUT: KasI-HindIII

<400>14

ES 2 887 776 T3

```

ggcgccgcta gccgcgctaa tgagcgggct tttttttggc ttgttgcca caaccgttaa      60
accttaaaag ctttaaaagc cttatatatt cttttttttc ttataaaaact taaaacctta      120
gaggctattht aagttgctga tttatattaa ttttattgtt caaacatgag agcttagtac      180
gtgaaacatg agagcttagt acgttagcca tgagagctta gtacgttagc catgagggtt      240
tagttcgtta aacatgagag ctttagtacgt taaacatgag agcttagtac gtactatcaa      300
caggttgaac tgctgatcca cgttgtggta gaattggtaa agagagtcgt gtaaaatata      360
gagttcgcac atcttggtgt ctgattattg atttttggcg aaaccatttg atcatatgac      420
aagatgtgta tctaccttaa cttaatgatt ttgataaaaa tcattaggta ccacatgtcc      480
tgcagaggcc tgcattgcaag ctt                                     503

```

<210> 15
 <211> 539
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Casete de clonación del polienlizador pNTC-NP6 polienlizador trpA R6K-RNA-OUT: EcoRI-SacI
 <400>15

```

gaattctgca gatattcatc aacttgccgg ccgctagccc gcctaattgag cgggcttttt      60
tttggtttgt tgtccacaac cgttaaacct taaaagcttt aaaagcctta tatattcttt      120
tttttcttat aaaacttaaa accttagagg ctatttaagt tgctgattta tattaatttt      180
attgttcaaa catgagagct tagtacgtga aacatgagag ctttagtacgt tagccatgag      240
agcttagtac gttagccatg aggggtttagt tgcgttaaaca tgagagctta gtacgttaaa      300
catgagagct tagtacgtac tatcaacagg ttgaactgct gatccacgtt gtggtagaat      360
tggtaaagag agtcgtgtaa aatattcgagt tcgcacatct tgttgtctga ttattgattt      420
ttggcgaaac catttgatca tatgacaaga tgtgtatcta ccttaactta atgattttga      480
taaaaatcat taggtaccgg gccccccctc gatcgaggtc gacggtatcg ggggagctc      539

```

<210> 16
 <211> 500
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Casete de clonación polienlizador pNTC-NP7 polienlizador trpA R6K-RNA-OUT: BssHII-BssHII
 <400>16

ES 2 887 776 T3

	gcgcgcagcc ttaattaagc tagcccgcc aatgagcggg cttttttttg gcttggtgtc	60
	cacaaccggt aaaccttaaa agcttttaaa gccttatata ttcttttttt tcttataaaa	120
	cttaaaacct tagaggctat ttaagttgct gatttatatt aattttattg ttcaaaccatg	180
	agagcttagt acgtgaaaca tgagagctta gtacgttagc catgagagct tagtacgtta	240
	gccatgaggg tttagttcgt taaacatgag agcttagtac gttaaaccatg agagcttagt	300
	acgtactatc aacagggtga actgctgac cacgttggtg tagaattggt aaagagagtc	360
	gtgtaaaata tcgagttcgc acatcttggt gtctgattat tgatttttgg cgaaaccatt	420
	tgatcatatg acaagatgtg tatctacctt aacttaatga ttttgataaa aatcattagg	480
	taccttaatt aactgcgcgc	500
5	<210> 17 <211> 530 <212> ADN <213> Secuencia artificial	
	<220> <223> Casete de clonación del polienlazador pNTC-3xCpG NP1 polienlazador R6K-RNA-OUT: HindIII-EcoRI	
	<400>17	
10	aagcttgcac gcaggcctct gcagtcgacg ggccctctag actcgagctg gcttggtgtc	60
	cacaaccatt aaaccttaaa agcttttaaa gccttatata ttcttttttt tcttataaaa	120
	cttaaaacct tagaggctat ttaagttgct gatttatatt aattttattg ttcaaaccatg	180
	agagcttagt acgtgaaaca tgagagctta gtacattagc catgagagct tagtacatta	240
	gccatgaggg tttagttcat taaacatgag agcttagtac attaaaccatg agagcttagt	300
	acatactatc aacagggtga actgctgac tgtacagtag aattggtaaa gagagttgtg	360
	taaaatattg agttcgaca tcttggtgtc tgattattga tttttggcga aaccatttga	420
	tcatatgaca agatgtgtat ctaccttaac ttaatgattt tgataaaaat cattaggtac	480
	cgctagcggc cgtcccttag atgcattcgc gaggtaccga gctcgaattc	530
15	<210> 18 <211> 1973 <212> ADN <213> Secuencia artificial	
	<220> <223> Interferón beta S/MAR humano flanqueado por el sitio 5' BglII-XhoI y los sitios de la enzima de restricción 3' EcoRI	
20	<400>18	

ES 2 887 776 T3

```

agatctcgag ctcgatgata atgaatgtct aagttaatgc agaaacggag agacatacta      60
tattcatgaa ctaaaagact taatattgtg aaggatatact ttctttccac ataaatttgt      120
agtcaatatg ttcaccccaa aaaagctgtt tgtaacttg ccaacctcat tctaaaatgt      180
atatagaagc ccaaaagaca ataacaaaaa tattcttgta gaacaaaatg ggaagaatg      240
ttccactaaa tatcaagatt tagagcaaag catgagatgt gtgggtagatg acagtgaggc      300
tgataaaata gagtagagct cagaaacaga cccattgata tatgtaagt acctatgaaa      360
aaaatatggc attttacaat gggaaaatga tgatcttttt cttttttaga aaaacaggga      420
aatatattta tatgtaaaaa ataaaaggga acccatatgt cataccatac acacaaaaaa      480
attccagtga attataagtc taaatggaga aggcaaaact ttaaactctt tagaaaaata      540
tatagaagca tgccatcatg acttcagtgt agagaaaaat ttcttatgac tcaaagtcct      600
aaccacaaag aaaagattgt taattagatt gcatgaatat taagacttat ttttaaatt      660
aaaaaaccat taagaaaagt caggccatag aatgacagaa aatatttgca acaccccgat      720
aaagagaatt gtaatatgca gattataaaa agaagtctta caaatcagta aaaaaataaa      780
ctagacaaaa atttgaacag atgaaagaga aactctaaat aatcattaca catgagaaac      840
tcaatctcag aaatcagaga actatcattg catatacact aaattagaga aatattaaaa      900
ggctaagtaa catctgtggc aatattgatg gtatataacc ttgatatgat gtgatgagaa      960
cagtacttta ccccatgggc ttctcccca aacccttacc ccagtataaa tcatgacaaa     1020
tatactttta aaaccattac cctatatcta accagtactc ctcaaaactg tcaaggtcat     1080
caaaaataag aaaagtctga ggaactgtca aaactaagag gaaccaagg agacatgaga     1140
attatatgta atgtggcatt ctgaatgaga tcccagaaca gaaaaagAAC agtagctaaa     1200
aaactaatga aatataaata aagtttgaac tttagttttt tttaaaaaag agtagcatta     1260
acacggcaaa gccattttca tatttttctt gaacattaag tacaagtcta taattaaaaa     1320
ttttttaaat gtagtctgga acattgccag aaacagaagt acagcagcta tctgtgctgt     1380
cgctaacta tccatagctg attggtctaa aatgagatac atcaacgctc ctccatgttt     1440
tttgttttct ttttaaatga aaaactttat tttttaagag gagtttcagg ttcatagcaa     1500
aattgagagg aaggtacatt caagctgagg aagttttcct ctattcctag tttactgaga     1560
gattgcatca tgaatgggtg ttaaattttg tcaaagtctt tttctgtgtc tatcaatatg     1620
accatgtgat tttcttcttt aacctgttga tgggacaaat tacgttaatt gattttcaaa     1680
cgttgaacca cccttacata tctggaataa attctacttg gttgtggtgt atattttttg     1740
atacattctt ggattctttt tgctaataatt ttgttgaaaa tgtttgtatc tttgttcatt     1800
agagatatgg gtctgttgtt ttcttttctt gtaatgtcat tttctagttc cggtatataag     1860
gtaatgctgg cctagttgaa tgatttagga agtattccct ctgcttctgt cttctgaaag     1920
agattgtaga aagttgatac aatttttttt tcttttaata tttgatagaa ttc           1973

```

<210> 19

<211> 2085

<212> ADN

<213> Secuencia artificial

<220>

<223> Donante de empalme-interferón beta humano S/MAR-aceptante de empalme flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI

<400>19

ES 2 887 776 T3

```

agatctatgc atgcagaagt tggtcgtgag gcaactgggca ggtaagtatc aaggttacaa      60
gacaggtctc gagctcgatg ataatgaatg tctaagttaa tgcagaaacg gagagacata      120
ctatatcatc gaactaaaag acttaatatg gtgaagggtat actttctttc cacataaatt      180
tgtagtcaat atgttcaccc caaaaaagct gtttggttaac ttgccaacct cattctaaaa      240
tgtatataga agcccaaaag acaataacaa aaatattctt gtagaacaaa atgggaaaga      300
atgttccact aaatatcaag atttagagca aagcatgaga tgtgtgggga tagacagtga      360
ggctgataaa atagagtaga gctcagaaac agaccattg atatatgtaa gtgacctatg      420
aaaaaaatat ggcattttac aatgggaaaa tgatgatctt tttctttttt agaaaaacag      480
ggaaatatat ttatatgtaa aaaataaaag ggaaccata tgtcatacca tacacacaaa      540
aaaattccag tgaattataa gtctaaatgg agaaggcaaa actttaaatc ttttagaaaa      600
taatatagaa gcatgccatc atgacttcag tgtagagaaa aatttcttat gactcaaagt      660
cctaaccaca aagaaaagat tgtaatttag attgcatgaa tattaagact tattttttaa      720
attaaaaaac cattaagaaa agtcaggcca tagaatgaca gaaaatattt gcaacacccc      780
agtaaagaga attgtaatat gcagattata aaaagaagtc ttacaaatca gtaaaaaata      840
aaactagaca aaaatttgaa cagatgaaag agaaactcta aataatcatt acacatgaga      900
aactcaatct cagaaatcag agaactatca ttgcatatac actaaattag agaaatatta      960
aaaggctaag taacatctgt ggcaatattg atggtatata accttgatat gatgtgatga     1020
gaacagtact ttaccccatg ggcttcctcc ccaaaccctt accccagtat aaatcatgac     1080
aaatatactt taaaaaccat taccctatat ctaaccagta ctctcaaaa ctgtcaaggt     1140
catcaaaaat aagaaaagtc tgaggaactg tcaaaaactaa gaggaacca aggagacatg     1200
agaattatat gtaatgtggc attctgaatg agatcccaga acagaaaaag aacagtagct     1260
aaaaaactaa tgaaatataa ataaagtttg aacttttagtt ttttttataa aagagtagca     1320
ttaacacggc aaagccattt tcatattttt ctggaacatt aagtacaagt ctataattaa     1380
aaatttttta aatgtagtct ggaacattgc cagaacaga agtacagcag ctatctgtgc     1440
tgtcgcttaa ctatccatag ctgattggtc taaaatgaga tacatcaacg ctctccatg     1500
ttttttgttt tcttttttaa tgaaaaactt tattttttta gaggagtttc aggttcatag     1560
caaaattgag aggaaggtag attcaagctg aggaagtttt cctctattcc tagtttactg     1620
agagattgca tcatgaatgg gtgttaaatt ttgtcaaatg ctttttctgt gtctatcaat     1680
atgacctatg gattttcttc tttaacctgt tgatgggaca aattacgtta attgattttc     1740
aaacgttgaa ccacccttac atatctggaa taaattctac ttgggtgtgg tgtatatttt     1800
ttgatacatt ctggtattct ttttgctaatt attttggtga aaatgtttgt atctttgttc     1860
atgagagata ttggtctggt gttttctttt cttgtaatgt cattttctag ttccggtatt     1920
aaggtaatgc tggcctagtt gaatgattta ggaagtattc cctctgcttc tgtcttctga     1980
aagagattgt agaaagtga tacaattttt ttttctttta atatttgata gaattcttac     2040
tgacatccac tttgcctttc tctccacagg tgtccactcg gatcc      2085

```

<210> 20

<211> 2085
<212> ADN
<213> Secuencia artificial

5 <220>
<223> Donante de empalme-interferón humano beta S/MAR (-AATAAA)-aceptante de empalme flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI
<400>20

agatctatgc atgcagaagt tggctcgtgag gcactgggca ggtaagtatc aaggttacaa 60
 gacaggtctc gagctcgatg ataatgaatg tctaagttaa tgcagaaacg gagagacata 120
 ctatatcat gaactaaaag acttaatat gtgaaggat actttcttcc cacataaatt 180
 tgtagtcaat atgttcaccc caaaaaagct gtttggttaac ttgccaacct cattctaaaa 240
 tgtatataga agcccaaaag acaataacaa aaatatctct gtagaacaaa atgggaaaga 300
 atgttccact aaatatcaag atttagagca aagcatgaga tgtgtgggga tagacagtga 360
 ggctgataaa atagagtaga gctcagaaac agaccattg atatatgtaa gtgacctatg 420
 aaaaaatat ggcattttac aatgggaaaa tgatgatctt tttctttttt agaaaaacag 480
 ggaaatatat ttatatgtaa aaaatatgtt ggaaccata tgtcatacca tacacacaaa 540
 aaaattccag tgaattataa gtctaaatgg agaaggcaaa actttaaatc ttttagaaaa 600
 taatatagaa gcatgccatc atgacttcag tgtagagaaa aatttcttat gactcaaagt 660
 cctaaccaca aagaaaagat tgttaattag attgcatgaa tattaagact tattttttaa 720
 attaaaaaac cattaagaaa agtcaggcca tagaatgaca gaaaatattt gcaacacccc 780
 agtaaagaga attgtaatat gcagattata aaaagaagtc ttacaaatca gtaaaaaata 840
 tttctagaca aaaatttgaa cagatgaaag agaaactcta aataatcatt acacatgaga 900
 aactcaatct cagaaatcag agaactatca ttgcatatac actaaattag agaaatatta 960
 aaaggctaag taacatctgt ggcaatattg atggtatata accttgatat gatgtgatga 1020
 gaacagtact ttaccccatg ggcttcctcc ccaaaccctt acccagtat aaatcatgac 1080
 aaatatactt taaaaaacat taccctatat ctaaccagta ctctcaaaa ctgtcaaggt 1140
 catcaaaaat aagaaaagtc tgaggaactg tcaaaactaa gaggaaccca aggagacatg 1200
 agaattatat gtaatgtggc attctgaatg agatccaga acagaaaaag aacagtagct 1260
 aaaaaactaa tgaaatataa atattttttg aacttttagt tttttttaa aagagtagca 1320
 ttaacacggc aaagccattt tcatattttt cttgaacatt aagtacaagt ctataattaa 1380
 aaatttttta aatgtagtct ggaacattgc cagaaacaga agtacagcag ctatctgtgc 1440
 tgtcgctaa ctatccatag ctgattgggc taaaatgaga tacatcaacg ctctcccatg 1500
 ttttttggtt tcttttttaa tgaaaaactt tattttttta gaggagtttc aggttcatag 1560
 caaaattgag aggaaggtac attcaagctg aggaagtttt cctctattcc tagtttactg 1620
 agagattgca tcatgaatgg gtgttaaat ttgtcaaatg ctttttctgt gtctatcaat 1680
 atgacctgt gattttcttc tttaacctgt tgatgggaca aattacgtta attgattttc 1740
 aaacgttgaa ccacccttac atatctggaa tattttctac ttggttgtgg tgtatatatt 1800
 ttgatacatt cttggattct ttttgctaat attttgttga aaatgtttgt atctttgttc 1860
 atgagagata ttggtctgtt gttttctttt cttgtaatgt cattttctag ttccgggtatt 1920
 aaggtaatgc tggcctagtt gaatgattta ggaagtattc cctctgcttc tgtcttctga 1980
 aagagattgt agaaagttga tacaattttt ttttctttta atatttgata gaattcttac 2040
 tgacatccac tttgccttcc tctccacagg tgtccactcg gatcc 2085

<210> 21
 <211> 868
 <212> ADN
 <213> Secuencia artificial

<220>

<223> Donante de empalme-interferón beta humano M18 S/MAR-aceptante de empalme flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI

<400>21

agatctatgc atgcagaagt tggctcgtgag gcactgggca ggtaagtatc aaggttacia	60
gacaggtctc gagaagcttg atctaaataa acttataaat tgtgagagaa attaatgaat	120
gtctaagtta atgcagaaac ggagagacat actatattca tgaactaaaa gacttaatat	180
tgtgaaggta tactttcttt ccacataaat ttgtagtcaa tatgttcacc ccaaaaaagc	240
tgtttgttaa ctgccaacc tcattctaaa atgtatatag aagcccaaaa gacaataaca	300
aaaatattct tgtagaacaa aatgggaaag aatgttcac taaatatcaa gatttagagc	360
aaagcatgag atgtgtgggg atagacagtg aggtcgataa aatagagtag agctcagaaa	420
cagacccatt gatatatgta agtgacctat gaaaaaata tggcatttta caatgggaaa	480
atgatgatct ttttcttttt tagaaaaaca gggaaatata tttatatgta aaaaataaaa	540
gggaacccat atgtcatacc atacacacaa aaaaattcca gtgaattata agtctaaatg	600
gagaaggcaa aacttttaat cttttagaaa ataatataga agcatgccat catgacttca	660
gtgtagagaa aaatttctta tgactcaaag tcctaaccac aaagaaaaga ttgttaatta	720
tgaatgattt aggaagtatt ccctctgctt ctgtcttctg aaagagattg tagaaagttg	780
atacaatttt tttttcttta aatatttgat agagaattct tactgacatc cactttgcct	840
ttctctccac aggtgtccac tcggatcc	868

5

<210> 22

<211> 927

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Donante de empalme-805 pb aceptor de apolipoproteína B S/MAR humana flanqueado por el sitio 5' BglII y los sitios de la enzima de restricción 3' BamHI

<400>22

agatctatgc atgcagaagt tggctcgtgag gcaactgggca ggtaagtatc aaggttacia 60
gacaggtctc gaggcctacc ggagagcctt gccttgcaaa ggcagacagt cagtgaggaa 120
gactatgtgg cacatgaaga caccagaggt gttcctcagg atcaaagtat gtacaagcct 180
ttgtgaatat tttttccttc tcacttggca aatacaattc ctgagatcaa taacctcgtc 240
tttttaattt tttcctcgtc tttttaacta tttataaaat attgaattat aaaatatgta 300
attataaata ctttaattat aaaatatgta attataaata ctttaattat aaaatatgta 360
attataaata ctttataaaa tatgtaatta taaaatatgt aattataaac attttaatta 420
taaaatatgt aattataaac attttaatta taaaatatgt aattataaac attttaatta 480
taaaatatgt aattataaac attttaatta taaaatatgt aattataaac attttaatta 540
taaaatatgt aattataaac attttaatta taaaatatgt aattataaac attttaatta 600
taaaatatgt aattataaac attttaatta taaaatatgt aattataaac attttaatta 660
taaaatatgt aattataaac attttaatta taaaatatgt aattataaac attttaatta 720
taaaatatgt aattataaac attttaatta taaaatatgt aattataaac attttaatta 780
taaaatatgt aattataaac attttaatta taaaatatgt aattataaac attttaatta 840
ctccgtttcc atttttattc tgttacttaa atgaattctt actgacatcc actttgcctt 900
tctctccaca ggtgtccact cggatcc 927

5

<210> 23
<211> 647
<212> ADN
<213> Secuencia artificial

10

<220>
<223> Donante de empalme-525 pb de apolipoproteína B humana S/MAR-aceptante de empalme
flanqueado por el sitio 5' Nsil y los sitios de la enzima de restricción 3' BamHI

<400>23

atgcatgcag aagttggtcg tgaggcactg ggcaggtaag tatcaaggtt acaagacagg 60
tctcgagatt tataaaatat tgaattataa aatatgtaat tataaatact ttaattataa 120
aatatgtaat tataaatact ttaattataa aatatgtaat tataaatact ttataaaata 180
tgtaattata aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat 240
tttaattata aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat 300
tttaattata aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat 360
tttaattata aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat 420
tttaattata aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat 480
tttaattata aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat 540
tttaattata aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat 600
actgacatcc actttgcctt tctctccaca ggtgtccact cggatcc 647

15

<210> 24
<211> 197
<212> ADN
<213> Secuencia artificial

<220>

ES 2 887 776 T3

<223> intrón pCI

<400>24

gcagaagttg gtcgtgaggc actgggcagg taagtatcaa ggttacaaga caggtttaag	60
gagaccaata gaaactgggc ttgtcgagac agagaagact cttgcgtttc tgataggcac	120
ctattggtct tactgacatc cactttgcct ttctctccac aggtgtccac tcccagttca	180
attacagctc ttaaggc	197

5 <210> 25
 <211> 42
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> donante de empalme pCI

10 <400> 25 gcagaagttg gtcggagg actgggcagg taagtatcaa gg 42

<210> 26
 <211> 44
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> pCI Aceptor de empalme (IgG murina)

<400> 26 ctactgaca tccacttgc cttctctcc acaggtcc actc 44

<210> 27
 <211> 1025
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Potenciador de la expresión de Ele40

<400>27

```

gatcaagaaa gcactccggg ctccagaagg agccttccag gccagctttg agcataagct      60
gctgatgagc agtgagtgtc ttgagtagtg ttcagggcag catgttacca ttcattgcttg      120
acttctagcc agtgatgaga gaggctggag tcaggtctct agagagttga gcagctccag      180
ccttagatct cccagtctta tgcggtgtgc ccattogctt tgtgtctgca gtcccctggc      240
cacacccagt aacagttctg ggatctatgg gagtagcttc cttagtgage tttcccttca      300
aatactttgc aaccaggtag agaagtttgg agtgaagggt ttgttctttg tttcttcaca      360
atatggatat gcatcttctt ttgaaaatgt taaagtaaag tacctctctt ttcagatact      420
gtcttcatgc gaacttggtg tcctgtttcc atcccagcct tctataaccc agtaacatct      480
tttttgaaac cagtgggtga gaaagacacc tggtcaggaa cgcgggaccac aggacaactc      540
aggctcacc cccagcatcag actaaaggca aacaaggact ctgtataaag taccggtggc      600
atgtgtatta gtggagatgc agcctgtgct ctgcagacag ggagtcacac agacactttt      660
ctataatttc ttaagtgcct tgaatgttca agtagaaagt ctaacattaa atttgattga      720
acaattgtat attcatggaa tatttttgaa cggaatacca aaaaatggca atagtgggtc      780
tttctggatg gaagacaaac ttttcttttt aaattttatc ttatatattt gaggttgacc      840
acatgacctt aaggatacat atagacagta aactgggttac tacagtgaag caaattaaca      900
tatctaccat cttacatagt tacatttttt tgtgtgacag gaacagctaa aatctacgta      960
tttaacaaaa atcctaaaga caatacatct ttatttaacta tagccctcat gatgtacatt     1020
agatc                                     1025

```

5

<210> 28
 <211> 1503
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Potenciador de la expresión A2UCOE

10

<400>28

```

--- --
gcctacagct caagccacat ccgaaggggg agggagccgg gagctgcgcg cggggccgcc 60
ggggggaggg gtggcaccgc ccacgccggg cggccacgaa gggcggggca gcgggcgcg 120
gcgcggcggg gggagggggc ggccgcgcgc ccgctgggaa ttggggccct agggggaggg 180
cggagggcgcc gacgaccgcg gcacttacgg ttcgcggcgt ggcccccggg ggtccccaag 240
gggaggggaa ggggagggcg ggcgaggaca gtgaccggag tctcctcagc ggtggctttt 300
ctgcttgga gcctcagcgg ctggcgccaa aaccggactc cggccacttc ctgcgccgcc 360
ggtgcgaggg tgtggaatcc tccagacgct gggggagggg gagttgggag cttaaaaact 420
agtaccctt tgggaccact ttcagcagcg aactctcctg tacaccaggg gtcagttcca 480
cagacgcggg ccagggggtg gtcattgcgg cgtgaacaat aatttgacta gaagttagtt 540
cgggtgtttc cggaaggggc cgagtcaatc cggcgagttg gggcacggaa aacaaaaagg 600
gaaggctact aagatttttc tggcgggggt tatcattggc gtaactgcag ggaccacctc 660
ccgggttag ggggttgat ctccaggtg cggattaagc cctcccgtc ggcgttaatt 720
tcaaactgcg cgaccgtttc tcacctgcct tgcgccagg cagggggcg gaccctattc 780
caagaggtag taactagcag gactctagcc ttccgcaatt cattgagcgc atttacggaa 840
gtaacgtcgg gtactgtctc tggccgcaag ggtgggagga gtacgcattt ggcgtaaggt 900
ggggcgtaga gccttccgcg cattggcggc ggatagggcg tttacgcgac ggctgacgt 960
agcgggaagc gcgttagtg gggggaaggt tctagaaaag cggcggcagc ggctctagcg 1020
gcagtagcag cagcgccggg tcccgtcgg aggtgctcct cgcagagttg tttctcgagc 1080
agcggcagtt ctactacag cgccaggacg agtccgggtc gtgttcgtcc gcggagatct 1140
ctctcatctc gctcggctgc gggaaatcgg gctgaagcga ctgagtcgc gatggaggt 1200
acgggtttga aatcaatgag ttattgaaaa gggcatggcg aggcgttg cgcctcagtg 1260
gaagtcggcc agccgcctcc gtgggagaga ggcaggaaat cggaccaatt cagtagcagt 1320
ggggcttaag gtttatgaac ggggtcttga gcggaggcct gagcgtacaa acagcttccc 1380
caccctcagc ctcccggcgc catttccctt cactgggggt gggggatggg gagctttcac 1440
atggcggacg ctgccccgct ggggtgaaag tggggcgcgg aggcgggaat tcttattccc 1500
ttt 1503

```

5

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

<220>
 <223> Secuencia de consenso del aceptor de empalme

10

<400> 29 yyyyyyyyyy yagrw 15

<210> 30
 <211> 11
 <212> ADN
 <213> Homo sapiens

15

<400> 30 taaatatttt a 11

<210> 31
 <211> 19

	<212> ADN	
	<213> Homo sapiens	
	<400> 31 attataaata ttttaatta 19	
5	<210> 32	
	<211> 24	
	<212> ADN	
	<213> Homo sapiens	
	<400> 32 atttaattat aaatatattta atta 24	
10	<210> 33	
	<211> 525	
	<212> ADN	
	<213> Homo sapiens	
	<400>33	
	atttataaaa tattgaatta taaaatatgt aattataaat actttaatta taaaatatgt	60
	aattataaat actttaatta taaaatatgt aattataaat actttataaaa atatgtaatt	120
	ataaaatatg taattataaaa cattttaatt ataaaatatg taattataaaa cattttaatt	180
	ataaaatatg taattataaaa cattttaatt ataaaatatg taattataaaa cattttaatt	240
	ataaaatatg taattataaaa cattttaatt ataaaatatg taattataaaa cattttaatt	300
	ataaaatatg taattataaaa cattttaatt ataaaatatg taattataaaa cattttaatt	360
	ataaaatatg taattataaaa cattttaatt ataaaatatg taattataaaa cattttaatt	420
	ataaaatatg taattataaaa cattttaatt ataaaatatg taattataaaa cattttaatt	480
	ataaaatatg taattataaaa cattttaatt ataaaatatg taattataaaa cattttaatt	525
15	<210> 34	
	<211> 825	
	<212> ADN	
	<213> Homo sapiens	
	<400>34	

gcaggctgag tgaataaag gacttggtat ttcattctga ggctaccgg agagccttgc 60
 cttgcaaagg cagacagtca gtgaggaaga ctatgtggca catgaagaca ccagagggtgt 120
 tcctcaggat caaagtatgt acaagccttt gtgaatattt tttccttctc acttggcaaa 180
 tacaattcct gagatcaata acctcgtctt ttttaattttt tcctcgtctt ttttaactatt 240
 tataaaatat tgaattataa aatatgtaat tataaatact ttaattataa aatatgtaat 300
 tataaatact ttaattataa aatatgtaat tataaatact ttataaaata tgtaattata 360
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 420
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 480
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 540
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 600
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 660
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 720
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 780
 aaatatgtaa ttataaacat ttttaattata aaatatgtaa ttataaacat ttttaattata 825

<210> 35
 <211> 854
 <212> ADN
 <213> Homo sapiens

<400>35
 cttctccact cctggcaggc tgagtgaat aaaggacttg ttatttcatc tcgaggccta 60
 ccggagagcc ttgccttgca aaggcagaca gtcagtggag aagactatgt ggcacatgaa 120
 gacaccagag gtgttcctca ggatcaaagt atgtacaagc ctttgtgaat attttttctc 180
 tctcacttgg caaatataat tcttgagatc aataacctcg tctttttaat tttttcctcg 240
 tctttttaac tatttataaa atattgaatt ataaaatatg taattataaa tacttttaatt 300
 ataaaatatg taattataaa tacttttaatt ataaaatatg taattataaa tactttataa 360
 aatatgtaat tataaaatat gtaattataa acattttaat tataaaatat gtaattataa 420
 acattttaat tataaaatat gtaattataa acattttaat tataaaatat gtaattataa 480
 acattttaat tataaaatat gtaattataa acattttaat tataaaatat ttaattataa 540
 acattttaat tataaaatat ttaattataa atattttaat tataaaatat ttaattataa 600
 atattttaat tataaaatat ttaattataa atattttaat tataaaatat ttaattataa 660
 atactttaat tataaaatat ttaattataa atattttaat tataaaatat ttaattataa 720
 atattttaat tataaaatat ttaattataa atattttaat tataaaatat ttaattataa 780
 atctttttaa atatttttgc aaaatatttc cctccataat ttctccgttt ccatttttat 840
 tctgttactt aaat 854

<210> 36
 <211> 199
 <212> PRT
 <213> Secuencia artificial
 <220>

<223> puromicina acetiltransferasa, construcción sintética

<400>36

```

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val
1      5      10      15
Pro Arg Ala Val Arg Thr Leu Ala Ala Phe Ala Asp Tyr Pro Ala
20      25      30
Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu

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

```

5

<210> 37

<211> 600

<212> ADN

<213> Secuencia artificial

10

<220>

<223> secuencia codificadora de la puromicina acetiltransferasa, sintética

<400>37

```

atgaccgagt acaagcccac ggtgcgctc gccacccgcg acgacgtccc cggggccgta      60
cgccacctcg ccgccgcgtt cgccgactac cccgccacgc gccacaccgt cgaccgggac      120
cgccacatcg agcgggtcac cgagctgcaa gaactcttcc tcacgcgcgt cgggctcgac      180
atcggaagg tgtgggtcgc ggacgacggc gccgcggtgg cggctctggac cacgccggag      240
agcgtcgaag cggggggcgt gttcgccgag atcgcccgcc gcatggccga gttgagcgg      300
tcccggctgg ccgcgcagca acagatggaa ggcctcctgg cgccgcaccg gccaaggag      360
cccgctggt tcctggccac cgctggcgct tcgcccgaacc accagggcaa gggctctggc      420
agcgccgtcg tgctccccgg agtggaggcg gccgagcgcg cgggggtgcc cgccttctg      480
gagacctccg cgccccgcaa cctccccttc tacgagcggc tcggcttcac cgtcaccgcc      540
gacgtcgagg tgcccgaagg accgcgcacc tggatcatga cccgcaagcc cggtgcccta      600

```

15

<210> 38

<211> 1031

<212> ADN

<213> Secuencia artificial

<220>

<223> elemento antirrepresivo40

20

<400>38

ES 2 887 776 T3

```

gatcaagaaa gcactccggg ctccagaagg agccttccag gccagctttg agcataagct      60
gctgatgagc agtgagtgtc ttgagtagtg ttcagggcag catgttacca ttcattgcttg      120
acttctagcc agtgtgacga gaggtggag tcagggtctct agagagttga gcagctccag      180
ccttagatct cccagtctta tgcggtgtgc ccattcgctt tgtgtctgca gtcccctggc      240
cacaccagc aacagttctg ggatctatgg gagtagcttc cttagtgagc tttcccttca      300
aatactttgc aaccaggtag agaagtttgg agtgaagggtt ttgttcttgc tttcttcaca      360
atatggatat gcatcttctt ttgaaaatgt taaagtaa atacctctt ttcagatact      420
gtcttcatgc gaacttggtg tcctgtttcc atccagcct tctataaccc agtaacatct      480
tttttgaaac cagtgggtga gaaagacacc tggtcaggaa cgcgaccac aggacaactc      540
aggctcacc acggcatcag actaaaggca aacaaggact ctgtataaag taccggtggc      600
atgtgtatta gtggagatgc agcctgtgct ctgcagacag ggagtcacac agacactttt      660
ctataatttc ttaagtgtt tgaatgttca agtagaaagt ctaacattaa atttgattga      720
acaattgtat attcatggaa tatttttgga cggaatacca aaaaatggca atagtgggtc      780
tttctggatg gaagacaaac ttttcttctt taaaataaat tttattttat atatttgagg      840
ttgaccacat gaccttaagg atacatatag acagtaaact gggtactaca gtgaagcaaa      900
ttaacatatc taccatcgta catagttaca tttttttgtg tgacaggaac agctaaaatc      960
tacgtattta acaaaactcc taaagacaat acattttttat taactatagc cctcatgatg     1020
tacattagat c                                                                1031

```

<210> 39

<211> 1438

<212> ADN

<213> Secuencia artificial

<220>

<223> Promotor del CMV - Secuencia S/MAR

<400>39

```

ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc      60
catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgccca      120
acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggg      180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc      240
aagtgtatca tatgccaaagt cggcccccta ttgacgtcaa tgacggtaaa tggcccgccct      300
ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat      360
tagtcatcgc tattaccatg gtgatgcggt tttggcagta caccaatggg cgtggatagc      420
ggtttgactc acgggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt      480
ggcaccaaaa tcaacgggac tttccaaaat gtcgtaataa ccccgccccg ttgacgcaaa      540
tgggcggtag gcgtgtacgg tgggaggtct atataagcag aggtcttctc cactcctggc      600
aggctgagtg aaataaagga cttgttattt catctcgagg cctaccggag agccttgccct      660
tgcaaaaggca gacagtcagt gaggaagact atgtggcaca tgaagacacc agagggtgtc      720
ctcaggatca aagtatgtac aagcctttgt gaatatTTTT tccttctcac ttggcaaata      780
caattcctga gatcaataac ctctgtcttt taattttttc ctctgtcttt taactattta      840
taaaatattg aattataaaa tatgtaatta taaatacttt aattataaaa tatgtaatta      900
taaatacttt aattataaaa tatgtaatta taaatacttt ataaaatatg taattataaa      960
atatgtaatt ataaacattt taattataaa atatgtaatt ataaacattt taattataaa     1020
atatgtaatt ataaacattt taattataaa atatgtaatt ataaacattt taattataaa     1080
atatgtaatt ataaacattt taattataaa atatttaatt ataaacattt taattataaa     1140
atatttaatt ataaatattt taattataaa atatttaatt ataaatattt taattataaa     1200
atatttaatt ataaatattt taattataaa atatttaatt ataaatactt taattataaa     1260
atatttaatt ataaatattt taattataaa atatttaatt ataaatattt taattataaa     1320
tattttaatt ataaaatatt taattataaa aacacaatta cctcatcttt ttaaataatt     1380
ttgcaaaata tttccctcca taatttctcc gtttccattt ttattctgtt acttaaat     1438

```

<210> 40

<211> 2038

<212> ADN

<213> Secuencia artificial

<220>

<223> Promotor del CMV - Puromicina - Secuencia S/MAR

<400>40

```

ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc      60
catatatgga gttccgcgtt acataactta cggtaaattg cccgcctggc tgaccgccc      120
acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggg      180
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc      240
aagtgtatca tatgccaaagt cggcccccta ttgacgtcaa tgacggtaaa tggcccgcc      300
ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat      360
tagtcatcgc tattaccatg gtgatgcggt tttggcagta caccaatggg cgtggatagc      420
ggtttgactc acgggggattt ccaagtctcc accccattga cgtcaatggg agtttgttt      480
ggcacaaaaa tcaacgggac tttccaaaat gtcgtaataa ccccgccccg ttgacgcaaa      540
tgggcggtag gcgtgtacgg tgggaggtct atataagcag aggtatgacc gagtacaagc      600
ccacggtgcg cctcgccacc cgcgacgacg tccccggggc cgtacgcacc ctgcgcgcg      660
cgttcgcca ctaccccgcc acgcgccaca ccgtcgaccc ggaccgccac atcgagcggg      720
tcaccgagct gcaagaactc ttcctcacgc gcgtcgggct cgacatcggc aaggtgtggg      780
tcgcggaacg cggcgccgcg gtggcggtct ggaccacgcc ggagagcgtc gaagcggggg      840
cgggtgtcgc cgagatcggc ccgcgcattg ccgagttgag cggttcccg      900
agcaacagat ggaaggcctc ctggcgccgc accggcccaa ggagcccgcg tggttcctgg      960
ccaccgtcgg cgtctcgccc gaccaccagg gcaagggtct gggcagcgcc gtcgtgctcc      1020
ccggagtggg ggcggccgag cgcgcggggg tgcccgccct cctggagacc tccgcgcccc      1080
gcaacctccc cttctacgag cggctcggct tcaccgtcac cgccgacgtc gaggtgcccg      1140
aaggaccgcg cacctggtgc atgaccgca agcccggtgc ctgacttctc cactcctggc      1200
aggctgagtg aaataaagga cttgttattt catctcgagg cctaccggag agccttgct      1260
tgcaaaggca gacagtcagt gaggaagact atgtggcaca tgaagacacc agaggtgttc      1320
ctcaggatca aagtatgtac aagcctttgt gaatattttt tccttctcac ttggcaaata      1380
caattcctga gatcaataac ctcgctcttt taattttttc ctcgctcttt taactattta      1440
taaaatattg aattataaaa tatgtaatta taaatacttt aattataaaa tatgtaatta      1500
taaatacttt aattataaaa tatgtaatta taaatacttt ataaaatatg taattataaa      1560
atatgtaatt ataaacattt taattataaa atatgtaatt ataaacattt taattataaa      1620
atatgtaatt ataaacattt taattataaa atatgtaatt ataaacattt taattataaa      1680
atatgtaatt ataaacattt taattataaa atatttaatt ataaacattt taattataaa      1740
atatttaatt ataaatattt taattataaa atatttaatt ataaatattt taattataaa      1800
atatttaatt ataaatattt taattataaa atatttaatt ataaatattt taattataaa      1860
atatttaatt ataaatattt taattataaa atatttaatt ataaatattt taattataaa      1920
tattttaatt ataaaatatt taattataaa aacacaatta cctcatcttt ttaaataatt      1980
ttgcaaaata tttccctcca taatttctcc gtttccattt ttattctgtt acttaaat      2038

```

<210> 41
 <211> 6189
 <212> ADN

<213> Secuencia artificial

<220>

<223> Elemento40-GPF-P2A-Puromicina-S/MAR

<400>41

	ggcattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc	60
5	catatatgga gttccgcgtt acataactta cggtaaattgg cccgcctggc tgaccgcca	120
	acgacccccg cccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggg	180
	ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgccacttg gcagtacatc	240
	aagtgtatca tatgccaagt cggcccccta ttgacgtcaa tgacggtaaa tggccgcct	300
	ggcattatgc ccagtacatg accttacggg actttcctac ttggcagtac atctacgtat	360
	tagtcatcgc tattaccatg gtgatgoggt tttggcagta caccaatggg cgtggatagc	420
	ggtttgactc acggggattt ccaagtctcc acccattga cgtcaatggg agtttgtttt	480
	ggcaccaaaa tcaacgggac tttccaaaat gtcgtaataa ccccgccccg ttgacgcaaa	540
	tgggcggtag gcgtgtacgg tgggaggtct atataagcag aggtcgttta gtgaaccgtc	600
	agatcactag tagcttttatt gcggtagttt atcacagtta aattgctaac gcagtcagtg	660
	ctcgactgat cacaggtaag tatcaaggtt acaagacagg tttaaggagg ccaatagaaa	720
	ctgggcttgt cgagacagag aagattcttg cgtttctgat aggcacctat tggcttact	780
	gacatccact ttgcctttct ctccacaggg gtaccgaagc cgctagcgct accggtcgcc	840
	accatgcccg ccatgaagat cgagtgcgc atcaccggca ccctgaacgg cgtggagttc	900
	gagctggtgg gcggcgga gggcaccccc gagcagggcc gcatgaccaa caagatgaag	960
	agcaccaaag gcgcccgtgac cttcagcccc tacctgctga gccacgtgat gggctacggc	1020
	ttctaccact tcggcaccta ccccgagggc tacgagaacc cttcctgca cgccatcaac	1080
	aacggcggt acaccaacac ccgcatcgag aagtacgagg acggcgcggt gctgcacgtg	1140
	agcttcagct accgtacga ggccggccgc gtgatcggg acttcaaggt ggtgggcacc	1200
	ggcttccccg aggacagct gatcttcacc gacaagatca tccgcagcaa cgccaccgtg	1260
	gagcacctgc accccatggg cgataacgtg ctggtgggca gcttcgcccc caccttcagc	1320
	ctgcgcgacg gcggctacta cagcttcgtg gtggacagcc acatgcaact caagagcgcc	1380
	atccacccca gcacctgca gaacgggggc cccatgttcg cttccgccc cgtggaggag	1440
	ctgcacagca acaccgagct gggcatcgtg gattaccagc acgccttcaa gacccccatc	1500
	gccttcgcca gatctggcag cggcgccacc aacttcagcc tgctgaagca ggccggcgac	1560
	gtggaggaaa accctgggccc catgaccgag tacaagccca cgggtgcgct cgccaccgc	1620
	gacgacgtcc cccgggccc acgcaccctc gccgcccgt tcgcccacta ccccgccacg	1680
	cgccacaccg tcgaccggga ccgccacatc gagcgggtca ccgagctgca agaactcttc	1740
	ctcacgcgcg tcgggctcga catcggaag gtgtgggtcg cggacgacgg cgccgcggtg	1800
	gcggtctgga ccacgcggga gagcgtcgaa gcggggggcg tgctcgccga gatcgggccc	1860
	cgcattggccg agttgagcgg tccccggtg gccgcgcagc aacagatgga aggcctcctg	1920
	gcgcgcgacc ggcccaagga gcccgctgg ttcctggcca ccgtcgcgct ctgcgcccac	1980
	caccagggca agggctctggg cagcgccgtc gtgctcccc gagtggaggc ggccgagcgc	2040

gccgggggtgc ccgccttcct ggagacctcc gcgccccgca acctccccctt ctacgagcgg	2100
ctcggettca ccgtcaccgc cgacgtcgag gtgcccgaag gaccgcgcac ctggtgcatg	2160
acccgcaagc ccggtgcctg aagatctatg catgcagaag ttggtcgtga ggcactgggc	2220
aggtaagtat caaggttaca agacaggtcg acttctccac tcctggcagg ctgagtgaaa	2280
taaaggactt gttatttcat ctcgaggcct accggagagc cttgccttgc aaaggcagac	2340
agtcagttag gaagactatg tggcacatga agacaccaga ggtgttcctc aggatcaaag	2400
tatgtacaag cctttgtgaa tattttttcc ttctcacttg gcaaatataa ttcttgagat	2460
caataacctc gtctttttta ttttttcctc gtctttttta ctattttata aatattgaat	2520
tataaaatat gtaattataa atactttaat tataaaatat gtaattataa atactttaat	2580
tataaaatat gtaattataa atactttata aaatatgtaa ttataaaata tgtaattata	2640
aacattttta ttataaaata tgtaattata aacattttta ttataaaata tgtaattata	2700
aacattttta ttataaaata tgtaattata aacattttta ttataaaata tgtaattata	2760
aacattttta ttataaaata tttaattata aacattttta ttataaaata tttaattata	2820
aatattttta ttataaaata tttaattata aatattttta ttataaaata tttaattata	2880
aatattttta ttataaaata tttaattata aatactttta ttataaaata tttaattata	2940
aatattttta ttataaaata tttaattata aatattttta ttataaatat tttaattata	3000
aaatatttta ttataaaaac acaattacct catcttttta aatatttttg caaaatattt	3060
ccctccataa tttctcgtt tccattttta ttctgttact taaataattc tgcagtcgac	3120
ggttactgac atccactttg ctttctctc cacagggtgc cactctaccg cgggcccggg	3180
atccaccgga tctagataac tgatcataat cagccatacc acattttag aggttttact	3240
tgcttttaaa aacctccac acctccccct gaacctgaaa cataaaatga atgcaattgt	3300
tggtgttaac ttgtttattg cagcttataa tggttacaaa taaagcaata gcatcacaaa	3360
tttcacaaat aaagcatttt tttcactgca ttctagtgtt ggtttgtcca aactcatcaa	3420
tgtatcttac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt	3480
gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag	3540
tcagaggtag cgaaaccgga caggactata aagataccag gcgtttcccc ctggaagctc	3600
cctcgtgcgc tctcctgttc cgacctgcc gcttacggga tacctgtccg cctttctccc	3660
ttcggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggttaggt	3720
cgttcgtcc aaagctgggt gtgtgcacga accccccgtt cagcccgacc gctgcgcctt	3780
atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc	3840
agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa	3900

gtggtggcct aactacggct acactagaag aacagtattht ggtatctgcg ctctgctgaa	3960
gccagttacc ttccgaaaaa gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg	4020
tagcgggtggt ttttttgtht gcaagcagca gattacgcgc agaaaaaaag gatctcaaga	4080
agatcccttg atcttttcta cggggtctga cgtcagtggt aacgaaaaact cacgttaagg	4140
gatttttggtc atgccgtctc agaagaactc gtcaagaagg cgatagaagg cgatgcgctg	4200
cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg tcagcccatc cgccgccaag	4260
ctcttcagca atatcacggg tagccaacgc tatgtcctga tagcgggtccg ccacaccag	4320
ccggccacag tcgatgaatc cagaaaagcg gccatthtcc accatgatat tcggcaagca	4380
ggcatcgcca tgggtcacga cgagatctc gccgtcgggc atgctcgctt tgagcctggc	4440
gaacagttcg gctggcgcca gcccctgatg ctcttcgtcc agatcatcct gatcgacaag	4500
accggcttcc atccgagtac gtgctctctc gatgcgatgt ttcgcttggt ggtcgaatgg	4560
gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca tcagccatga tggatacttht	4620
ctcggcagga gcaaggtgag atgacaggag atcctgcccc ggcacttcgc ccaatagcag	4680
ccagtccctt cccgcttcag tgacaacgtc gagtacagct gcgcaaggaa cgcctcgt	4740
ggccagccac gatagccgcg ctgcctcgtc ttgcagttca ttcagggcac cggacaggtc	4800
ggtcttgaca aaaagaaccg ggcgccctg cgtgacagc cggaacacgg cggcatcaga	4860
gcagccgatt gtctgttgtg cccagtcata gccgaatagc ctctccaccc aagcggccgg	4920
agaacctgcg tgcaatccat cttgttcaat cataatatta ttgaagcatt tatcagggtt	4980
cgtctcgtcc cgggtctctc ccatgcatgt caatatggc cattagccat attattcatt	5040
ggttatatag cataaatcaa tattggctat tggccattgc atacgttgta tctatatcat	5100
aatatgtaca aagcttgatc aagaaagcac tccgggctcc agaaggagcc ttccaggcca	5160
gctttgagca taagctgctg atgagcagtg agtgtcttga gtagtgttca gggcagcatg	5220
ttaccattca tgcttgactt ctagccagtg tgacgagagg ctggagtcag gtctctagag	5280
agttgagcag ctccagcctt agatctccca gtcttatgcg gtgtgcccac tcgctttgtg	5340
tctgcagtcc cctggccaca cccagtaaca gttctgggat ctatgggagt agcttcctta	5400
gtgagcttht ccttcaaata ctttgcaacc aggtagagaa gtttgagtg aaggthttht	5460
tcttcgtht ttcaaatat ggatatgcat cttctthtga aaatgttaa gtaaattacc	5520
tctctthtca gatactgtct tcatgcgaac ttggtatcct gthtccatcc cagccttcta	5580
taaccacgta acatcttht tgaaaccagt gggtgagaaa gacacctggt caggaaacgcg	5640
gaccacagga caactcaggc tcaccacgg catcagacta aaggcaaaca aggactctgt	5700
ataaagtacc ggtggcatgt gtattagtgg agatgcagcc tgtgctctgc agacaggag	5760
tcacacagac actthtctat aathtctta gtgctthtga tgttcaagta gaaagtctaa	5820

ES 2 887 776 T3

cattaaattt gattgaacaa ttgtatatc atggaatatt ttggaacgga ataccaaaaa	5880
atggcaatag tggttctttc tggatggaag acaaactttt cttctttaaa ataaatttta	5940
ttttatatat ttgaggttga ccacatgacc ttaaggatac atatagacag taaactgggt	6000
actacagtga agcaaattaa catatctacc atcgtacata gttacatttt tttgtgtgac	6060
aggaacagct aaaatctacg tatttaacaa aactcctaaa gacaatacat tttattaac	6120
tatagccctc atgatgtaca ttagatctgt acatttatat tggetcatgt ccaatatgac	6180
cgccatggt	6189

REIVINDICACIONES

1. Un procedimiento *in vitro* para mejorar la expresión y la eficiencia de establecimiento de un vector de expresión episomal no integrador S/MAR autorreplicante en una célula de vertebrado diana que comprende los siguientes pasos:
 - 5 a. proporcionar un vector de expresión S/MAR episomal que comprende:
 - i. una región de selección-replicación bacteriana que comprende un origen de replicación bacteriano y un marcador seleccionable;
 - ii. una unidad de transcripción para la expresión de un transgén en una célula de vertebrado, que comprende un promotor, una UTR 5', un transgén y una UTR 3';
 - 10 iii. un inserto S/MAR situado dentro de dicha UTR 3'; y
 - b. modificar el vector de expresión de S/MAR episomal de manera que el S/MAR esté flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' dentro de dicha UTR 3', por lo que el vector de expresión de S/MAR episomal no integrador autorreplicante resultante ha mejorado la expresión y la eficiencia de establecimiento tras la transfección de una célula de vertebrado.
- 15 2. El procedimiento de la reivindicación 1, en el que dicho inserto S/MAR contiene motivos internos de terminación de la transcripción AATAAA.
3. El procedimiento de la reivindicación 2, en el que dichos motivos de terminación de la transcripción AATAAA en dicho S/MAR se sustituyen por motivos AATATT.
- 20 4. El procedimiento de la reivindicación 1, en el que dicho S/MAR se selecciona del grupo que consiste en Interferón beta humano S/MAR, M18 S/MAR, Apolipoproteína B S/MAR.
5. El procedimiento de la reivindicación 1, en el que dicho SMAR flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, y SEQ ID NO: 23.
- 25 6. El procedimiento de la reivindicación 1, en el que dicho origen de replicación bacteriano es un origen de replicación gamma R6K.
7. El procedimiento de la reivindicación 1, en el que dicho marcador seleccionable es una variante funcional de ARN-IN que regula el ARN-OUT con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5, y SEQ ID NO: 7.
- 30 8. El procedimiento de la reivindicación 1, en el que dicho vector de expresión episomal no integrador de S/MAR autorreplicante se selecciona del grupo que consiste en el vector plasmídico, el vector nanoplásmido, el vector lentivirus de integración deficiente y los vectores lentivirales no integradores.
- 35 9. Un polinucleótido que comprende al menos un promotor y un elemento S/MAR, en el que dicho elemento S/MAR está situado aguas abajo de dicho promotor y en el que la secuencia de ácido nucleico de dicho elemento S/MAR (secuencia S/MAR) comprende al menos 3 motivos de secuencia ATTA por cada 100 nucleótidos en un tramo de como máximo 200 nucleótidos, en el que dicho elemento S/MAR está flanqueado por un donante de empalme y un aceptor de empalme, y en el que dicho polinucleótido comprende además un origen de replicación gamma R6K con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO: 3 y SEQ ID NO: 4.
- 40 10. El polinucleótido de la reivindicación 9, en el que dicho promotor está comprendido en una unidad de transcripción para la expresión de un polipéptido de carga y/o un marcador seleccionable en una célula huésped.
11. El polinucleótido según la reivindicación 10, en el que dicha unidad de transcripción comprende un promotor, un UTR 5', un transgén y un UTR 3'.
- 45 12. El polinucleótido de la reivindicación 11, en el que dicho S/MAR está localizado dentro de dicho UTR 3'.
13. Una molécula de ADN recombinante circular cerrada covalentemente que comprende:
 - a. una unidad de transcripción para la expresión de un transgén en una célula de vertebrado, que comprende un promotor, una UTR 5', un transgén y una UTR 3';
 - 50 b. un S/MAR situado dentro de dicha UTR 3', en el que dicho S/MAR está flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3';

- c. un origen de replicación gamma R6K con una identidad de secuencia de al menos el 95% con una secuencia seleccionada del grupo formado por SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, y SEQ ID NO: 4; y
 - d. un marcador de ARN seleccionable de ARN-OUT que comprende una variante funcional de ARN-IN reguladora de ARN-OUT con al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5, y SEQ ID NO: 7.
- 5
- 14. El polinucleótido de una cualquiera de las reivindicaciones 9 a 12 o la molécula de ADN recombinante de la reivindicación 13, en los que dicho S/MAR se selecciona del grupo que consiste en el Interferón beta humano S/MAR, M18 S/MAR, Apolipoproteína B S/MAR.
- 10
- 15. El polinucleótido de una cualquiera de las reivindicaciones 9 a 12 o la molécula de ADN recombinante de la reivindicación 13, en los que dicho SMAR flanqueado por un sitio donante de empalme 5' y un sitio aceptor de empalme 3' tiene al menos un 95% de identidad de secuencia con una secuencia seleccionada del grupo que consiste en SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, y SEQ ID NO: 23.

FIG 1

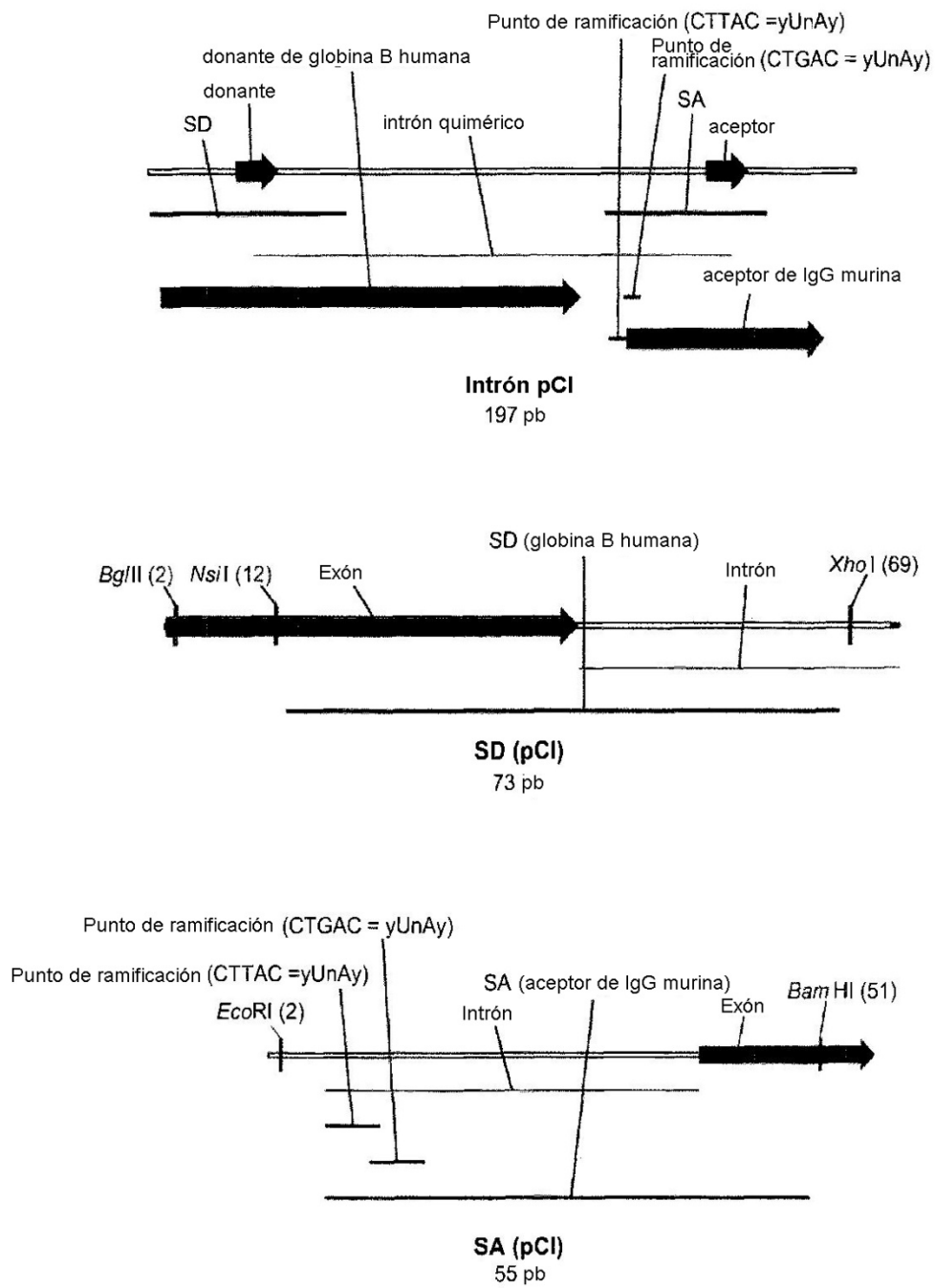


FIG 2

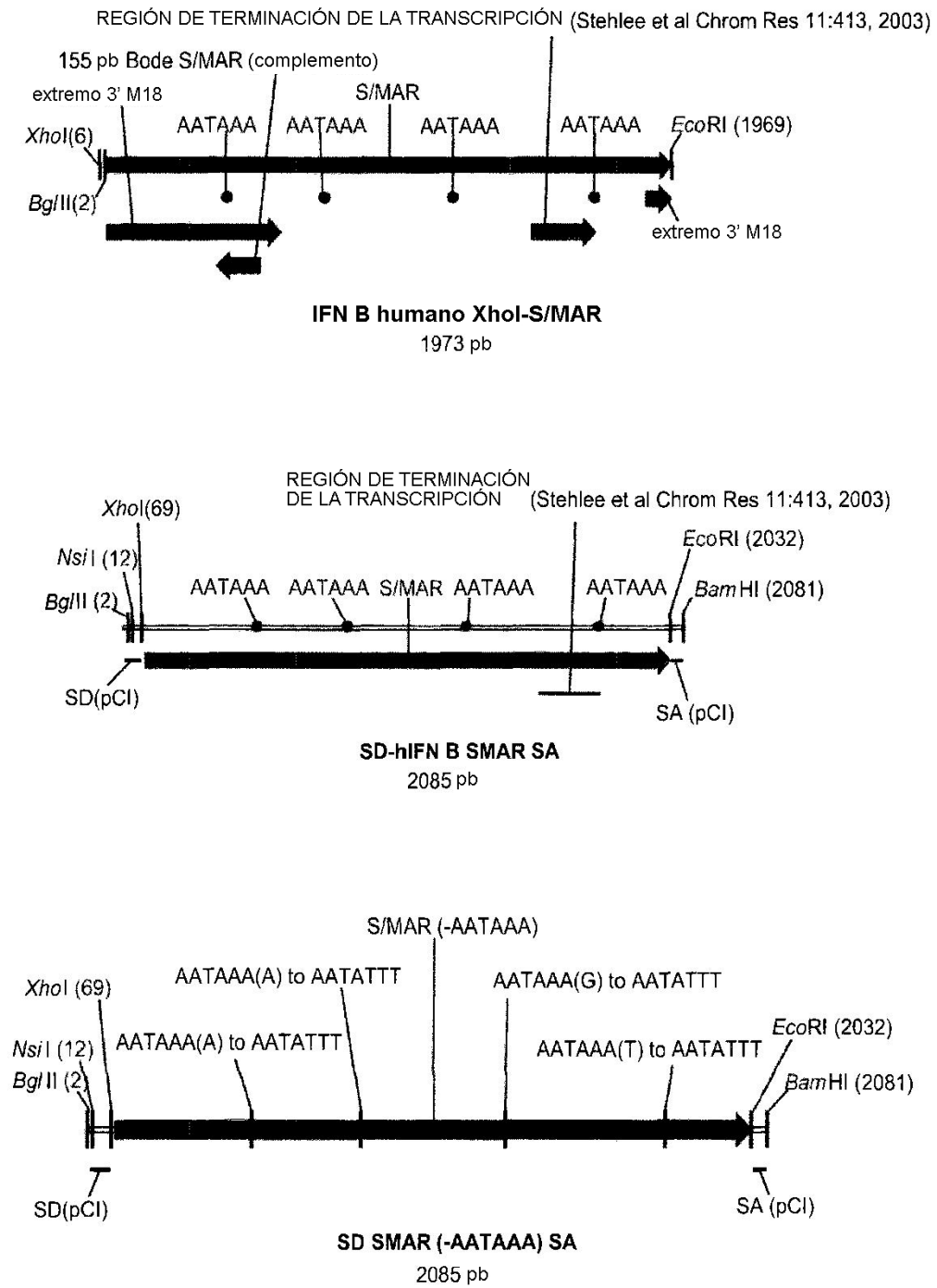


FIG 3

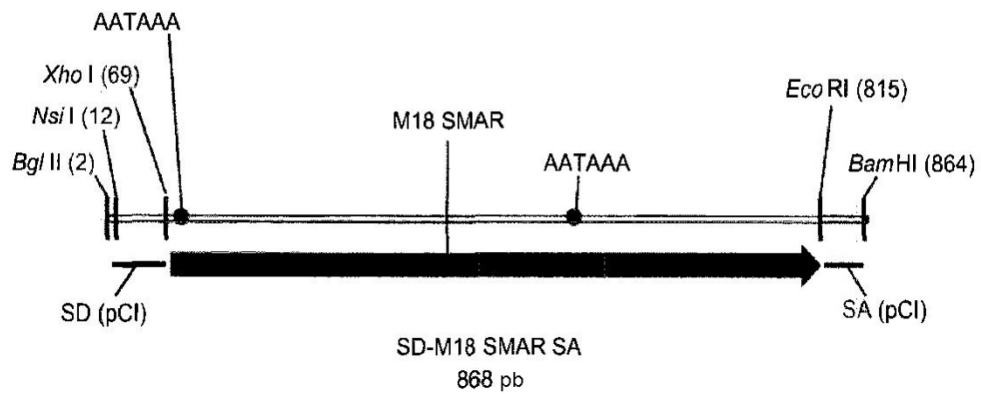


FIG 4

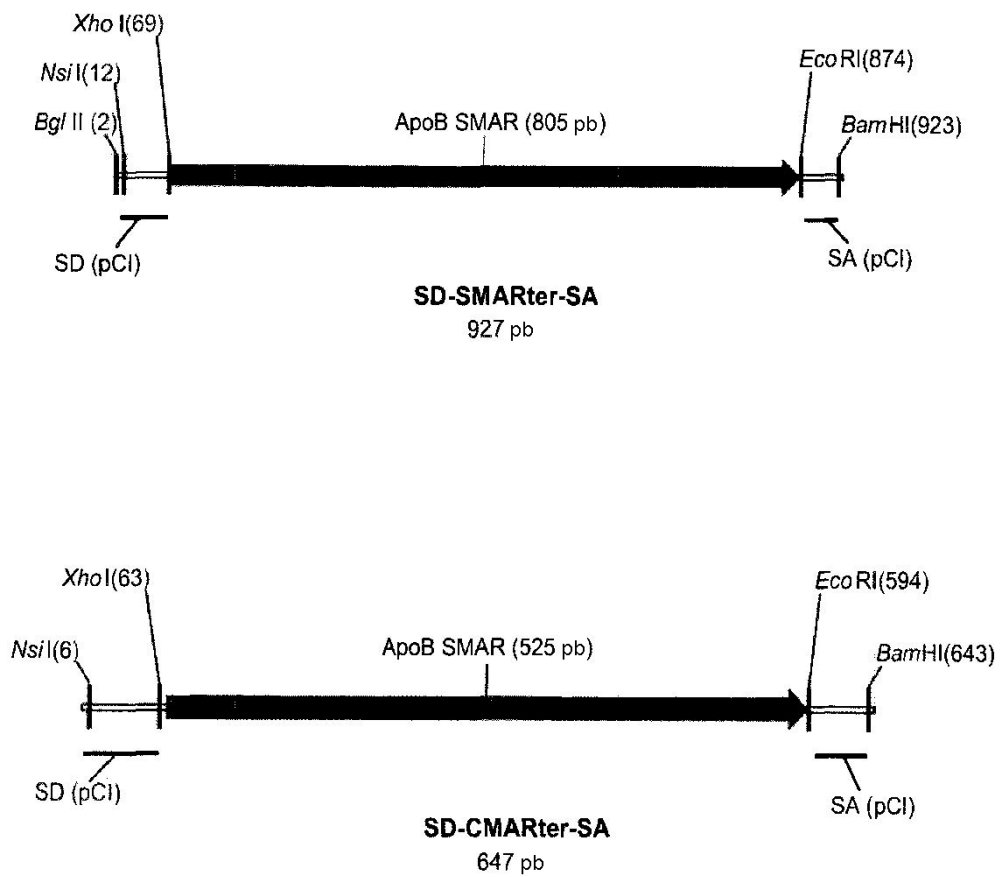


FIG 5

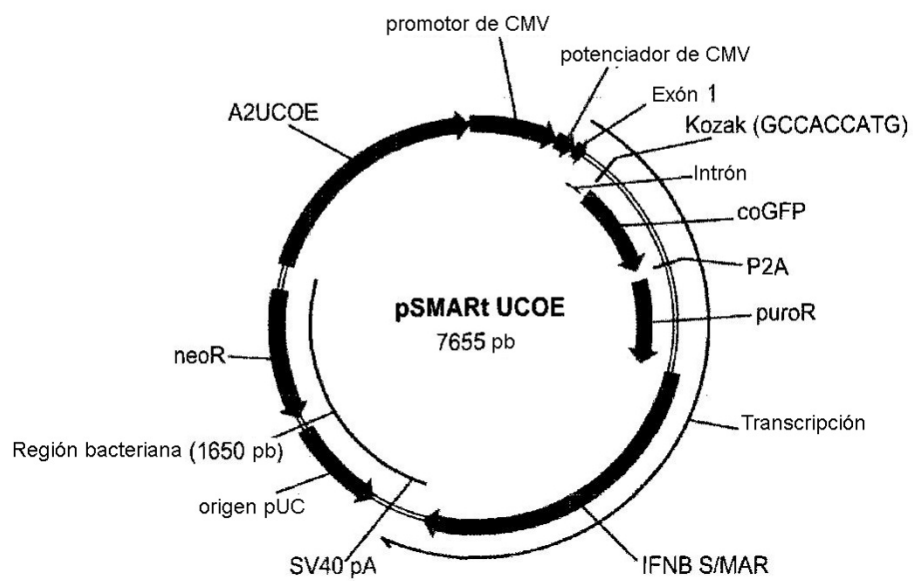


FIG 6

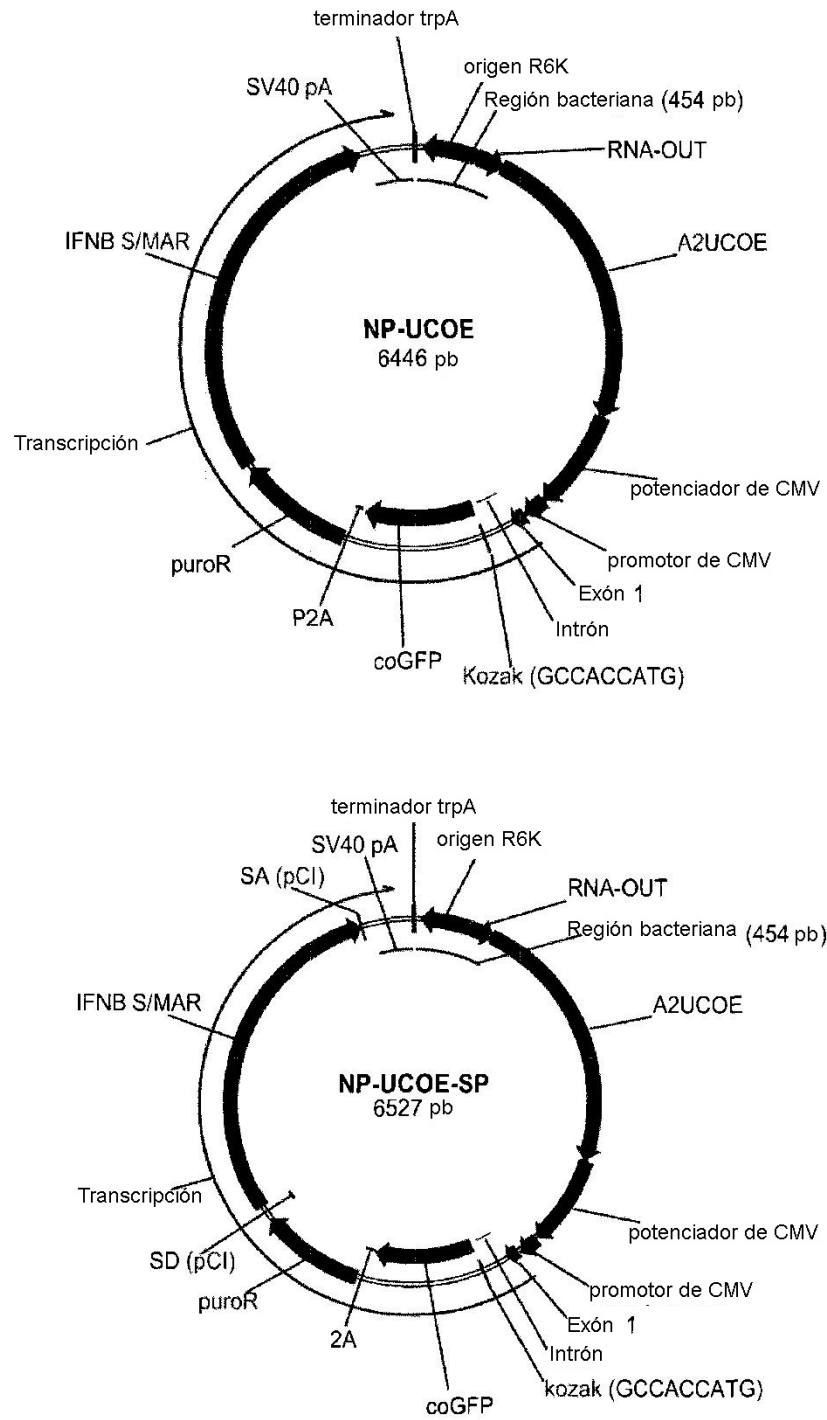


FIG 7

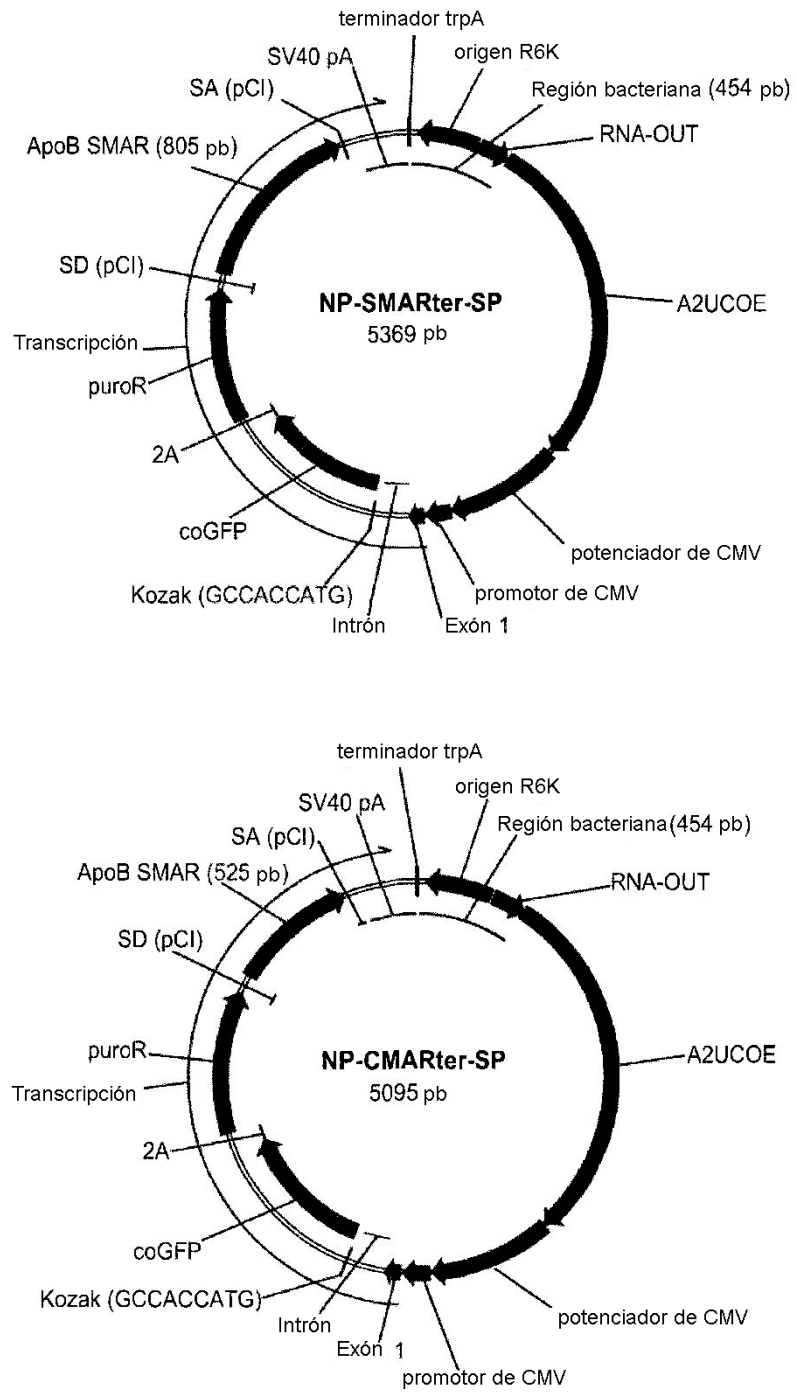


FIG 8

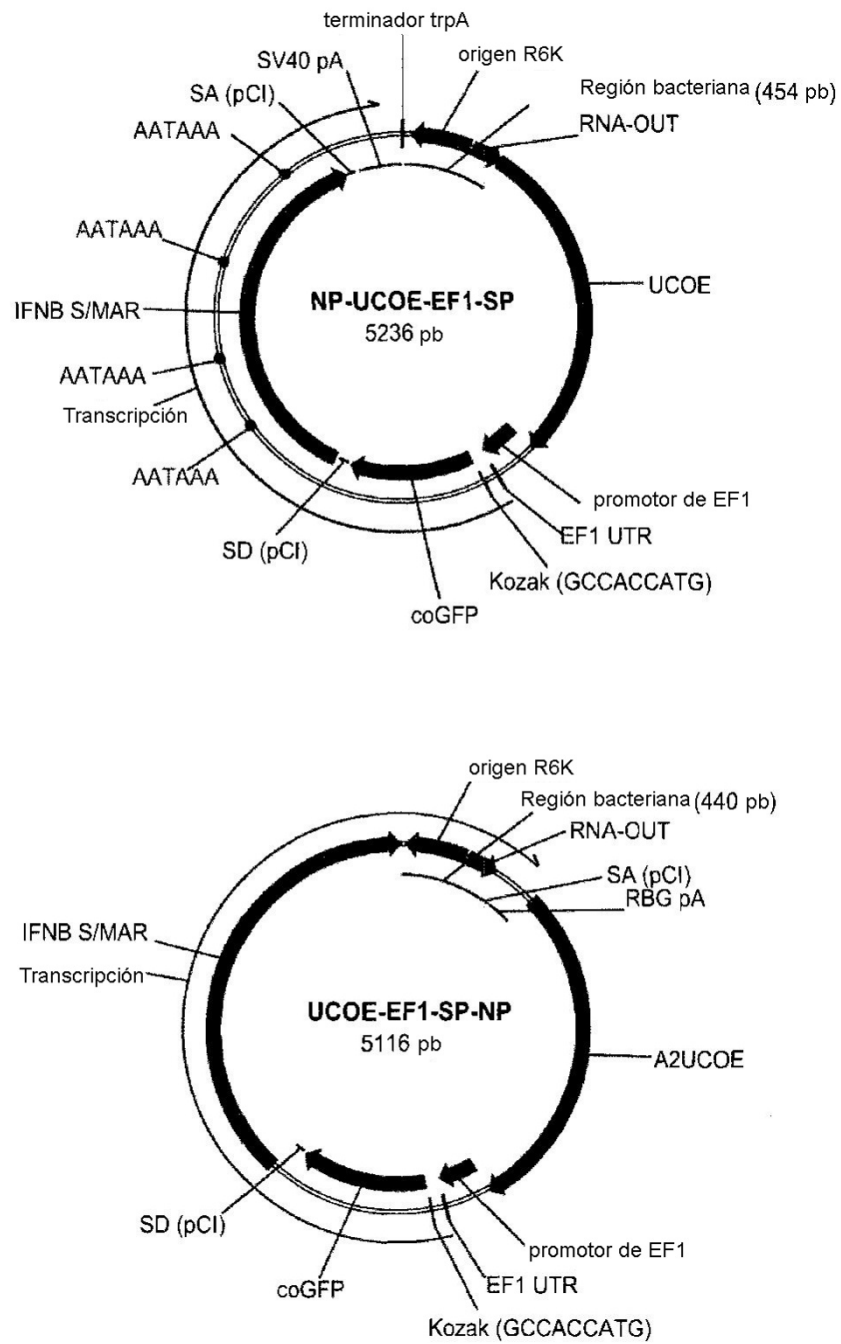


FIG 9

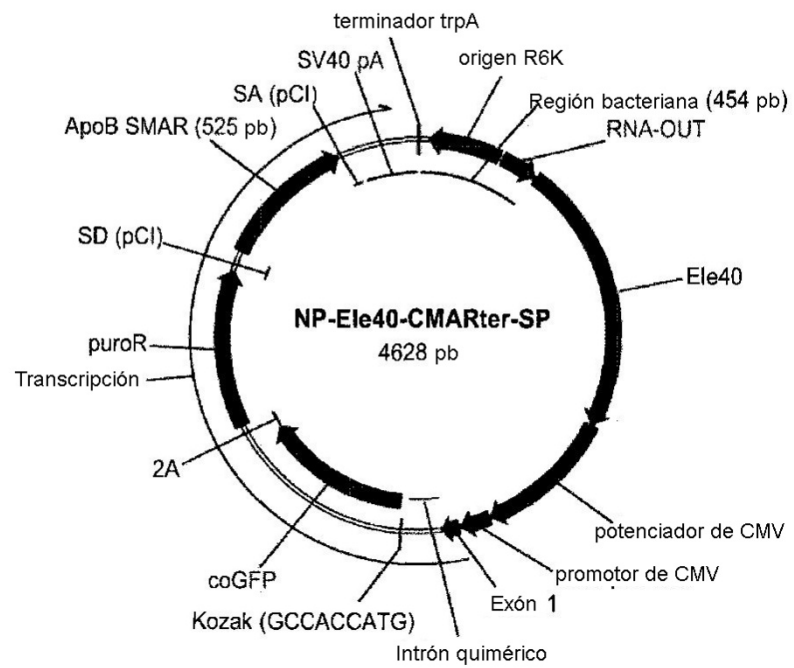


FIG 10

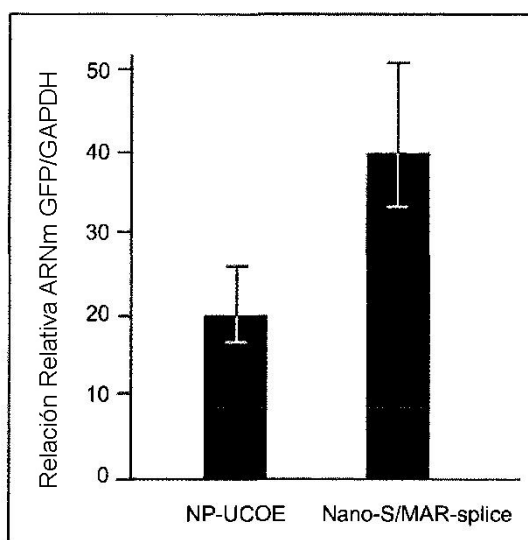
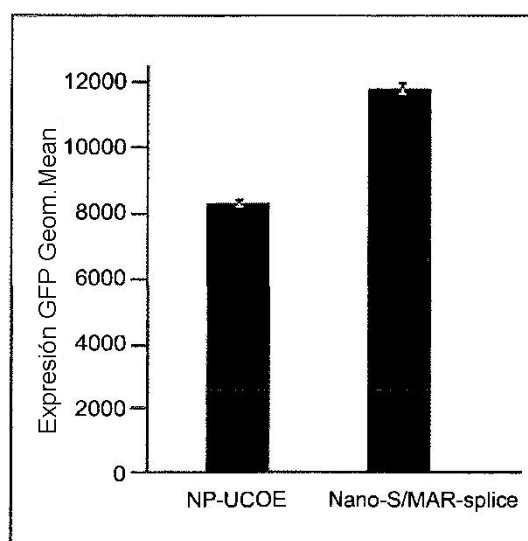


FIG 11

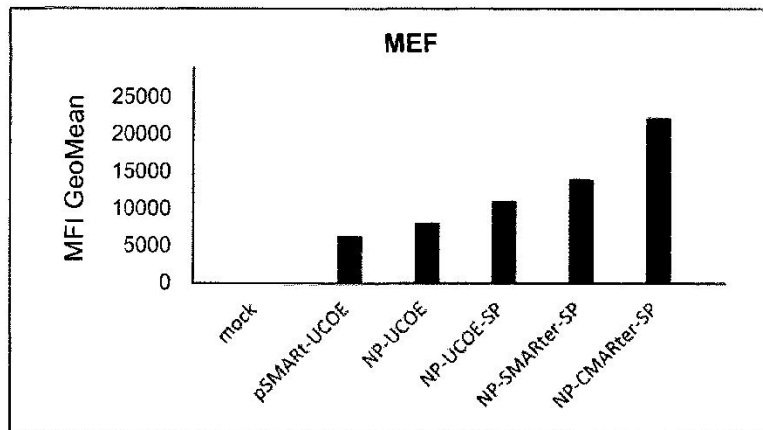
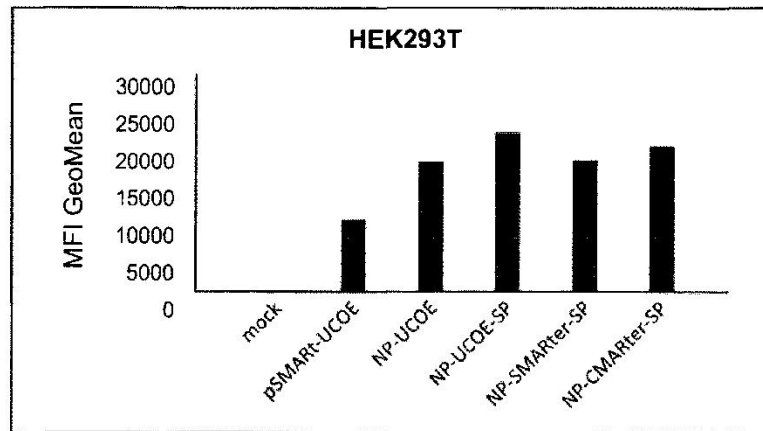
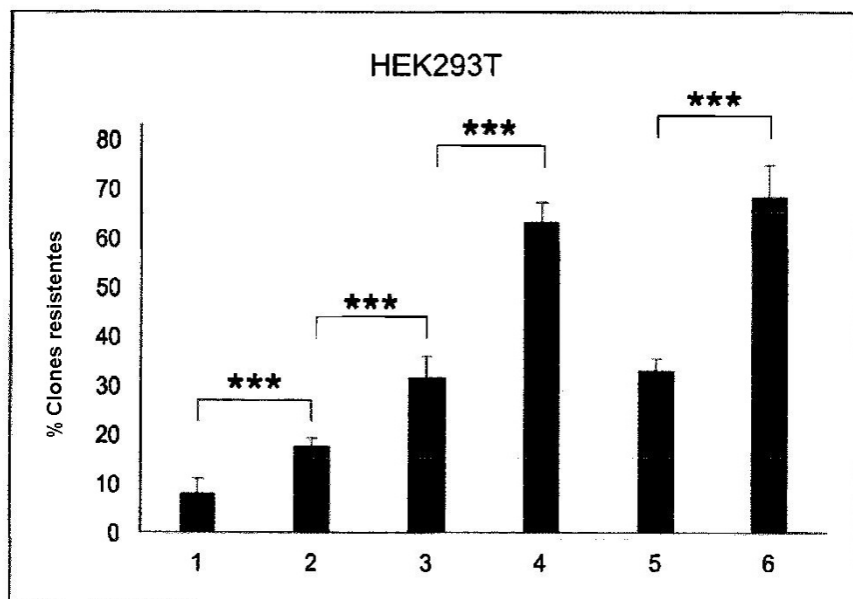


FIG 12



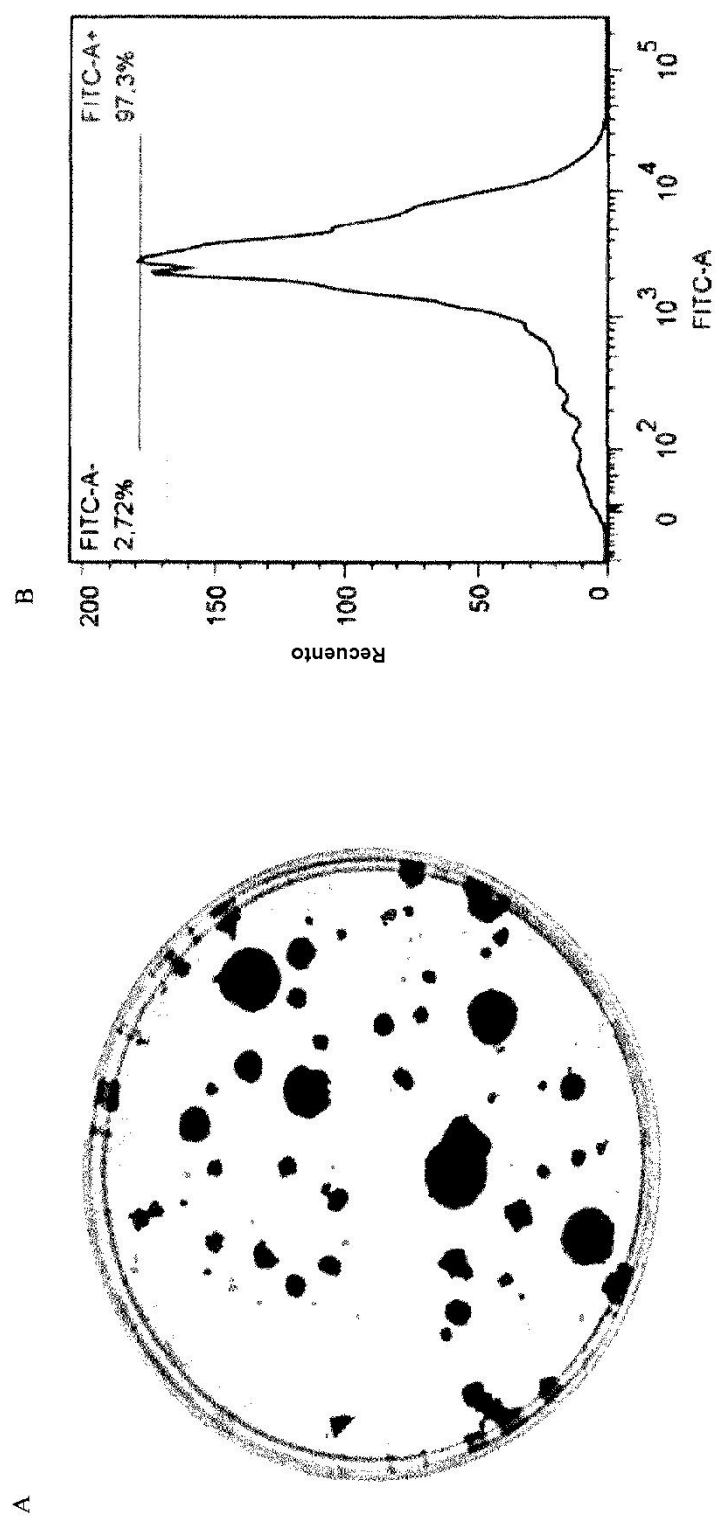


Fig. 13

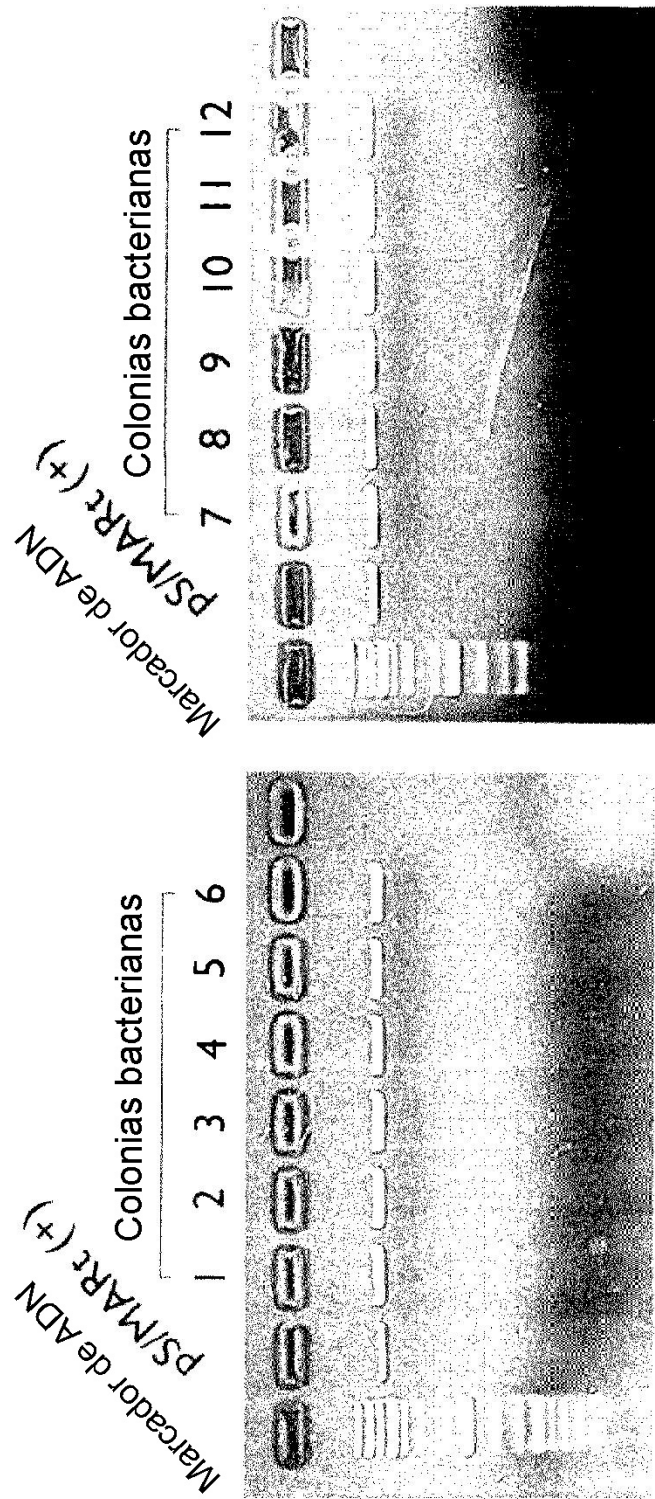


Fig.14

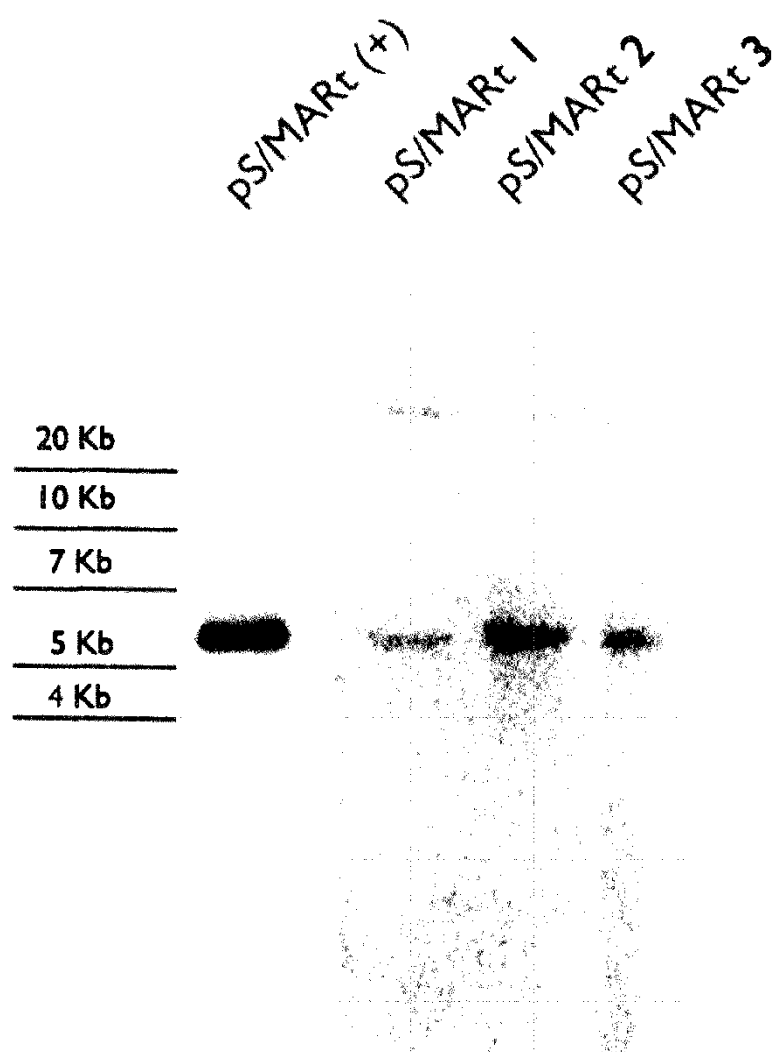


Fig. 15

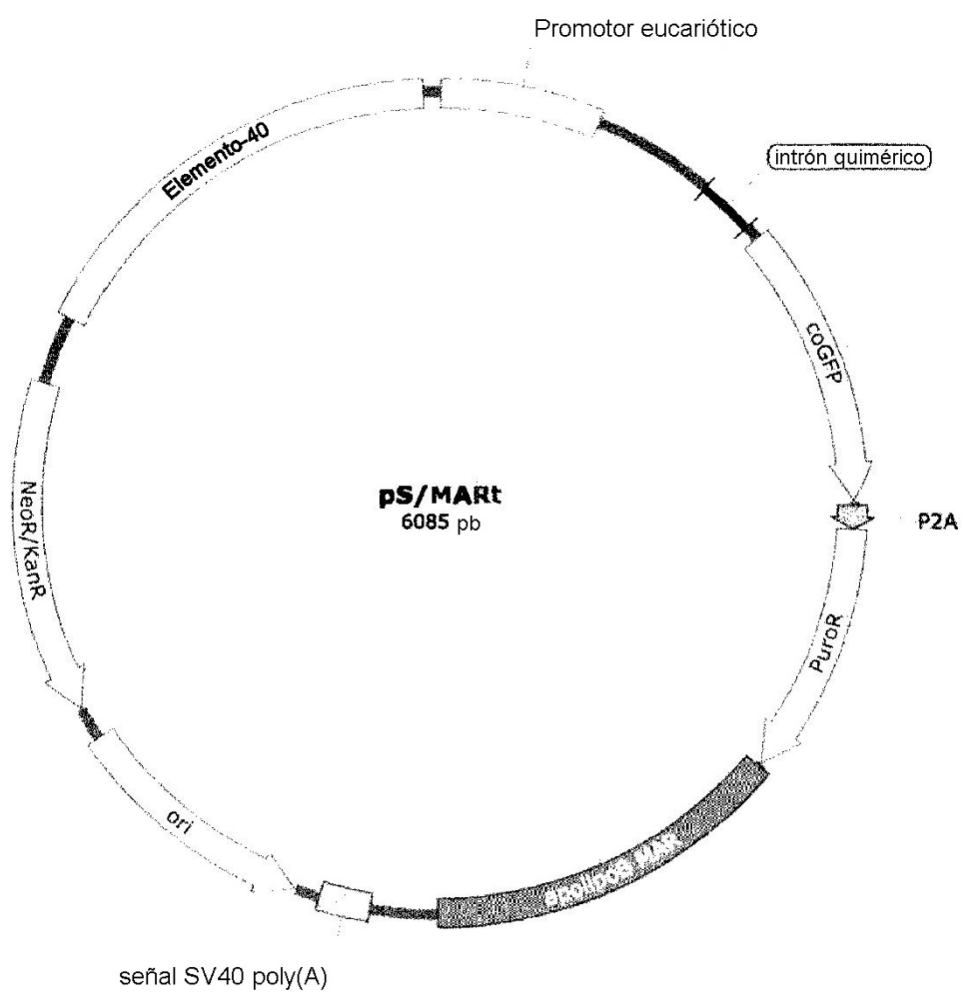


Fig. 16