

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 698 527**

51 Int. Cl.:

C12N 15/40 (2006.01)

C12N 15/85 (2006.01)

C12N 5/10 (2006.01)

C12N 5/074 (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.05.2013 PCT/US2013/041980**

87 Fecha y número de publicación internacional: **28.11.2013 WO13177133**

96 Fecha de presentación y número de la solicitud europea: **21.05.2013 E 13794196 (9)**

97 Fecha y número de publicación de la concesión europea: **22.08.2018 EP 2852671**

54 Título: **Generación de células iPS humanas mediante un ARN autorreplicante sintético**

30 Prioridad:

21.05.2012 US 201261649876 P
15.03.2013 US 201361798229 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.02.2019

73 Titular/es:

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (100.0%)
1111 Franklin Street, 5th Floor
Oakland, CA 94607-5200, US

72 Inventor/es:

DOWDY, STEVEN, F. y
YOSHIOKA, NAOHISA

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 698 527 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

Generación de células iPS humanas mediante un ARN autorreplicante sintético

Campo técnico

5 Se proporcionan métodos y composiciones útiles para producir y propagar células madre a partir de fibroblastos. La descripción se refiere a la producción de células madre pluripotentes inducidas (iPS, del inglés "induced Pluripotent Stem") y a métodos de uso de las mismas.

Antecedentes

Las células madre son una fuente potencial a partir de la cual se pueden regenerar órganos, se pueden reparar tejidos, se pueden preparar o administrar factores biológicos o se pueden tratar enfermedades o trastornos.

10 Hiroshi Ban et al (2011) PNAS; 108; 14234 describe vectores de virus Sendai sensibles a la temperatura que pueden usarse para generar células madre pluripotentes inducidas (IPSCs). El documento WO2009/067563 presenta la generación de IPSCs usando vectores retrovirales. Los documentos WO97/38087 y WO2005/116192 presentan replicones de ARN de alfavirus y su uso en la transformación de células madre. El documento WO2014/072061
15 para generar IPSCs.

Compendio

La generación de células madre pluripotentes inducidas (iPS) a partir de pacientes es importante para usar terapéuticamente células madre. La generación de células iPS requiere la expresión de varios factores de transcripción pluripotentes o de Factores de Reprogramación (RFs), que incluyen Oct4, Sox2, Klf4, cMyc, Glis1 (y potencialmente Nanog y Lin28). Sin embargo, debido a las preocupaciones con la integración de vectores de ADN (virus y ADN desnudo) en el genoma durante la generación de células iPS, se excluyen estas estrategias de ser usadas posteriormente en pacientes.

25 La descripción describe una estrategia para generar células madre pluripotentes inducidas (iPS) expresando ectópicamente RFs que usan un ARN autorreplicante sintético procedente de un alfavirus modificado (p.ej., el virus de Encefalitis Equina Venezolano (VEE)). El alfavirus se ha diseñado para expresar, en una realización, cuatro RFs que dieron como resultado las siguientes ventajas sobre las estrategias de transfección de ARNm: 1) se utiliza una única especie de ARN capaz de autorreplicarse por un número limitado de divisiones celulares, reduciendo de este modo el número de transfecciones; 2) es capaz de codificar en uno, dos, tres, cuatro o más marcos de lectura abiertos (ORFs) de RF; y 3) expresa consistentemente todos los genes de RF en niveles umbral elevados a lo largo de múltiples divisiones celulares. El uso de la cadena principal autorreplicante de un alfavirus (eliminando los genes estructurales) para expresar los RFs requiere solo de 3 a 4 transfecciones (e incluso de solo 1 o 2) en los fibroblastos humanos primarios para generar células iPS. La generación del transcrito RF-ARN de alfavirus utiliza un kit de transcripción SP6 (o T7) *in vitro* que no requiere de condiciones especiales y, por ello, simplifica adicionalmente la estrategia para un uso amplio. Expresando los cuatro RFs en niveles elevados consistentes en el tiempo en la misma célula, combinado
35 con la replicación del ARN de RF de alfavirus para un número limitado de generaciones de células múltiples, la estrategia de ARN de RF de alfavirus solventa los dos problemas de ineficiencia principales asociados al intento de generar células iPS mediante transfecciones diarias repetidas durante >14 días de cuatro mARNs de RF individuales. El ARN de RF de alfavirus es una estrategia ectópica que no utiliza un intermedio de ADN y, por tanto, no existe la posibilidad de que se pueda producir la mutación integrativa de las estrategias de células iPS basadas en vector de ADN. Además, la cadencia de pérdida de replicón de ARN por degradación se puede regular mediante la retirada de B18R del medio. Usando esta estrategia, se generaron >100 clones de células iPS a partir de protocolos de ARN de RF de alfavirus OCT4/KLF4/SOX2/c-MYC y OCT4/KLF4/SOX2/GLIS1 a partir de dos poblaciones de fibroblastos humanos parenterales independientes. Adicionalmente, la estrategia se puede modificar para expresar combinaciones de RF alternativas y/o la inserción de ORFs de RF adicionales en la cadena principal de ARN de RF para refinar la
45 generación de células iPS a partir de tipos celulares específicos o para uso en dirigir la transdiferenciación.

La descripción proporciona un ARN de replicón de alfavirus que comprende al menos un dominio de replicasa no estructural de un alfavirus y al menos una secuencia heteróloga no de alfavirus que codifica factores para inducir la generación de células madre pluripotentes cuando se expresa en una célula somática. En una realización, el replicón comprende secuencias obtenidas a partir de un alfavirus seleccionado del grupo que consiste en virus de Encefalitis Equina Oriental (EEE), virus de Encefalitis Equina Venezolana (VEE), virus de Everglades, virus Mucambo, virus Pixuna y virus de Encefalitis Equina Occidental (WEE). En otra realización, el replicón comprende secuencias obtenidas de un alfavirus seleccionado del grupo que consiste en virus de Sindbis, virus de Bosque de Semliki, virus Middelburg, virus Chikungunya, virus O'nyong-nyong, virus Ross River, virus del Bosque de Barnah, virus Getah, virus Sagiyama, virus Bebaru, virus Mayaro, virus Una, virus Aura, virus Whataroa, virus Babanki, virus Kyzylagach, virus J de las Highlands, virus Fort Morgan, virus Ndumu y virus Buggy Creek. En otra realización adicional, la al menos una secuencia heteróloga no de alfavirus comprende al menos 2, 3, 4 o 5 secuencias heterólogas no de alfavirus. En otra realización adicional, de cualquiera de las anteriores se selecciona la secuencia heteróloga no de alfavirus de un polinucleótido que codifica un polipéptido KLF, un polipéptido SOX-2, un polipéptido OCT-3/4, un polipéptido c-MYC o

n-MYC o L-MYC, un polipéptido GLIS1, un polipéptido NANOG y cualquier combinación de los mismos. En una realización adicional, el polinucleótido que codifica el polipéptido KLF codifica un polipéptido KLF que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:8. En otra realización, el polinucleótido que codifica el polipéptido KLF codifica un polipéptido KLF que tiene una secuencia de la SEQ ID NO:8. En otra realización adicional, el polinucleótido que codifica el polipéptido KLF comprende una secuencia como la establecida en la SEQ ID NO:7, donde "T" es "U". En otra realización, el polinucleótido que codifica el polipéptido SOX-2 codifica un polipéptido SOX-2 que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:6. En otra realización, el polinucleótido que codifica el polipéptido SOX-2 que codifica un polipéptido SOX-2 que tiene una secuencia de la SEQ ID NO:6. En otra realización adicional, el polinucleótido que codifica el polipéptido SOX-2 comprende una secuencia como la establecida en la SEQ ID NO:5, donde "T" es "U". En otra realización, el polinucleótido que codifica el polipéptido OCT-4 codifica un polipéptido OCT-4 que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:4. En una realización adicional, el polinucleótido que codifica el polipéptido OCT-4 codifica un polipéptido OCT-4 que tiene una secuencia de la SEQ ID NO:4. En una realización adicional, el polinucleótido que codifica el polipéptido OCT-4 comprende una secuencia como la establecida en la SEQ ID NO:3, donde "T" es "U". En otra realización, el polinucleótido que codifica el polipéptido c-MYC codifica un polipéptido c-MYC que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:10. En una realización adicional, el polinucleótido que codifica el polipéptido c-MYC que codifica un polipéptido c-MYC que tiene una secuencia de la SEQ ID NO:10. En una realización adicional, el polinucleótido que codifica el polipéptido c-MYC comprende una secuencia como la establecida en la SEQ ID NO:9, donde "T" es "U". En otra realización, el polinucleótido que codifica el polipéptido GLIS1 codifica un polipéptido GLIS1 que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:34. En una realización adicional, el polinucleótido que codifica el polipéptido GLIS1, codifica un polipéptido GLIS1 que tiene una secuencia de la SEQ ID NO:34. En una realización adicional, el polinucleótido que codifica el polipéptido GLIS1 comprende una secuencia como la establecida en la SEQ ID NO:33, donde "T" es "U". En otra realización, el polinucleótido que codifica el polipéptido NANOG codifica un polipéptido NANOG que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:2. En una realización adicional, el polinucleótido que codifica el polipéptido NANOG codifica un polipéptido NANOG que tiene una secuencia de la SEQ ID NO:2. En otra realización adicional, el polinucleótido que codifica el polipéptido NANOG comprende una secuencia como la establecida en la SEQ ID NO:1, donde "T" es "U". En una realización de cualquiera de las anteriores, el replicón comprende de 5' a 3': (replicasa de ARN de VEE) – (promotor) – (RF₁) – (péptido de auto-ruptura) – (RF₂) – (péptido de auto-ruptura) – (RF₃) – (IRES o promotor central) – (RF₄) – (IRES o promotor opcional) – (marcador seleccionable opcional) – (3'UTR y cola poliA de VEE) – (marcador seleccionable opcional) – promotor; donde RF₁₋₄ son factores que inducen la des-diferenciación de una célula somática a células pluripotentes, donde RF₂₋₃ son opcionales, RF₃₋₄ son opcionales o RF₄ es opcional; donde RF₁₋₄ se seleccionan del grupo que consiste en Oct-4, Klf4, Sox-2, c-Myc, Nanog y Glis1. En otra realización, el replicón comprende una secuencia que es idéntica en un 90%, 95%, 98%, 99% o 100% con respecto a la SEQ ID NO:29, 30, 31 o 32 desde aproximadamente la posición 1 hasta aproximadamente la posición 7561, donde "T" de la secuencia es sustituido por "U", seguido de dos o más RFs, seguido de un 3'UTR y cola poliA, donde los dos o más RFs se seleccionan del grupo que consiste en Oct-3/4, Sox-2, Klf4, c-Myc, Nanog y Glis1; donde las más de una secuencias codificadoras de RF pueden estar separadas por un sitio de entrada de ribosoma interno (IRES) o un promotor pequeño. En una realización adicional, el replicón comprende una secuencia que es idéntica en al menos un 95%, 98%, 99% o 100% a una secuencia seleccionada del grupo que consiste en la SEQ ID NO:29, 30, 31 o 32, donde "T" es "U".

La descripción también proporciona una composición que comprende células humanas transformadas con un replicón como se ha descrito en cualquiera de las realizaciones anteriores y de realizaciones descritas adicionalmente en la presente memoria. En una realización, la composición comprende además medio condicionado con B18R. En otra realización, las células humanas son células somáticas. En una realización adicional, las células humanas son fibroblastos.

La descripción también proporciona un método para preparar células madre que comprende cultivar la composición descrita anteriormente y en otros sitios de la presente memoria, durante al menos 30 días en condiciones que expresen las secuencias codificadoras del replicón y aislar las células madre.

La descripción también proporciona un método para preparar células madre que comprende transformar células somáticas con un replicón de la descripción, cultivar las células somáticas en condiciones que promueven la expresión del replicón y aislar las células madre. El cultivo comprende cultivar las células en medio condicionado con B18R. En otra realización, el medio condicionado con B18R es producido mediante transfección de ARNm de B18R en fibroblastos humanos primarios.

La descripción también proporciona células madre aisladas obtenidas a partir de los métodos descritos en la presente memoria, donde las células madre están libres de ADN o ARN retroviral.

La descripción también proporciona un método que comprende poner en contacto una célula somática humana *in vitro* con un replicón de ARN autorreplicante ectópico que comprende polinucleótidos que codifican al menos dos factores de des-diferenciación seleccionados del grupo que consiste en un (i) KLF4, (ii) OCT4, (iii) SOX2, (iv) c-MYC o n-MYC o L-MYC, (v) GLIS1 y (vi) NANOG; cultivar la célula somática para expresar el factor de des-diferenciación; seleccionar células que presentan una morfología de célula madre y/o marcadores de célula madre; y subcultivar las células para

obtener una población de células madre inducidas. En una realización, las células se seleccionan detectando la expresión de un antígeno de rechazo de tumor 1-60 y/o 1-81.

La descripción también proporciona un sistema de vector para producir células madre humanas, que comprende al menos un replicón de ARN autorreplicante que comprende uno o más polinucleótidos heterólogos no de alfavirus que codifica al menos dos factores de des-diferenciación seleccionados del grupo que consiste en un KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, GLIS1 y NANOG. En una realización, el replicón comprende (a) Oct4, Sox2, Klf4 y c-Myc, o (b) Oct4, Sox2, Klf4 y Glis1. El al menos un replicón de ARN autorreplicante es derivado de un alfavirus. En una realización adicional, el alfavirus es VEE.

La descripción también proporciona una célula somática humana aislada que comprende un replicón de ARN ectópico que comprende dos o más secuencias de polinucleótido de des-diferenciación que codifican factores de des-diferenciación seleccionados del grupo que consiste en Klf4, Oct4, Sox2, c-Myc o n-Myc o L-Myc, NANOG y GLIS1. En una realización adicional, donde en las condiciones de cultivo para expresar los polinucleótidos de des-diferenciación en el replicón de ARN ectópico, la célula somática se des-diferencia.

La descripción también proporciona una población celular que comprende la célula somática humana que contiene un replicón de ARN ectópico que comprende una o más secuencias de polinucleótido de des-diferenciación.

La descripción también proporciona una población celular obtenida poniendo en contacto una célula somática humana con un replicón de ARN autorreplicante ectópico que comprende polinucleótidos que codifican al menos cuatro factores de des-diferenciación seleccionados del grupo que consiste en un (i) KLF4, (ii) OCT4, (iii) SOX2, (iv) c-MYC o n-MYC o L-MYC, (v) GLIS1 y (vi) NANOG; cultivar la célula somática para expresar el factor de des-diferenciación; seleccionar células que presentan una morfología de célula madre y/o marcadores de célula madre; y subcultivar las células para obtener una población de células madre inducidas. Las células pueden ser seleccionadas detectando la expresión de un antígeno de rechazo de tumor 1-60 y/o 1-81.

La descripción también proporciona células de fibroblasto humanas recombinantes que contienen una molécula de ARN ectópica que codifica B18R. El ARN que codifica B18R puede comprender la SEQ ID NO:39, donde "T" es reemplazada con "U". O el ARN puede codificar un polipéptido que comprende la secuencia establecida en la SEQ ID NO:40.

La descripción también proporciona un método para preparar medio acondicionado con B18R que comprende cultivar una célula de fibroblasto humano transformada con ARN que codifica B18R en las condiciones que permiten la expresión de B18R, y aislar el medio del cultivo.

Descripción de las figuras

Figura 1A-E: muestra la construcción y la persistencia de replicones de ARN de RF de VEE sintéticos en fibroblastos humanos primarios. (A) Esquema de replicón de ARN de RF de VEE. Extremo 5' de nsP1-4: proteínas no estructurales 1-4; extremo 3' C, E2, E1: proteínas estructurales. Localizaciones de promotor interno 26S, péptido 2A de cambio de ribosoma, secuencia IRES, gen de resistencia de puromicina (Puro) y las regiones correspondientes a detecciones de PCR de replicón según se indica. (B) La co-transfección de ARNm de B18R con replicón de ARN de VEE permite expresar GFP de VEE en el día 1. (C) El medio acondicionado con B18R (B18R-CM) y la selección de puromicina son requeridos para la persistencia de ARN de GFP de VEE a lo largo de 7 días. (D) B18R-CM y puromicina son requeridos para la retención de ARN de GFP de VEE. Fotografías de expresión de GFP en el día 7 según se indica. Barra, 200 μ m. (E) Análisis de inmunotinción de factores de reprogramación expresados por ARN de VEE expresados en células HFFs en el día 1 frente a expresión de retrovirus (RV-4Fs).

Figura 2A-E: muestra la generación de células iPS mediante ARN de RF de VEE. (A) Esquema del protocolo de generación de células iPS con ARN de RF de VEE epigenético. Se llevaron a placa fibroblastos humanos en el día 0 (d0) y se co-transfectaron (Tfx) con replicón de ARN de RF de VEE más ARNm de B18R (ratio 3:1) en el día 1 (confluente, $\sim 4 \times 10^5$ células) y se trataron con puromicina hasta el día 7 (o 10) según se indica. Las células fueron cultivadas en B18R-CM hasta que las colonias de células iPS fueron aisladas en el día 25-30. (B) Se generaron colonias de células iPS teñidas con fosfatasa alcalina con ARN VEE-OKS-iM, pero no con ARN VEE-OMKS. (C) Tinción con fosfatasa alcalina de colonias de células iPS generadas a partir de BJ o HFFs del protocolo de transfección en los días 1, 4, 7, 10, según se indica. (D) Imágenes típicas de colonias de células iPS en el día 26 para ARN de VEE-OKS-iM y en el día 22 para ARN de VEE-OKS-iG a partir de fibroblastos BJ o HFFs según se indica. Barra, 100 μ m. (E) Tinción de inmunohistoquímica de genes marcadores de ES pluripotentes en clones de células iPS aisladas generados como se indica. Se obtuvieron resultados similares para 26 clones de células iPS adicionales (30 clones en total). Barra, 100 μ m; cuadro interior, ampliación 10x.

Figura 3A-E: muestra la caracterización de clones de células iPS de ARN de RF de VEE. (A) Expresión de genes marcadores de ES mediante análisis qRT-PCR de clones de iPS de VEE-OKS-iM de BJ y HFF como se indica. (B) Análisis de metilación de ADN de regiones promotoras NANOG y OCT4. Círculo relleno, metilado; Círculo hueco, desmetilado. Los números de la parte superior indican el número de CpG relativo al sitio de inicio de la transcripción. (C) Análisis de gráfico de puntos del perfil de secuencia de ARNm a escala de genoma de BJ-OKS-iM n°2 y BJ-OKS-iG n°5 en comparación con fibroblastos BJ humanos originales y células madre embrionarias HUES9 humanas con

pluripotencia NANOG, OCT4, SOX2 según se indica. (D) Dendrograma jerárquico no supervisado del análisis de secuencias de ARN a escala de genoma que muestra la agrupación de cuatro clones de células iPS independientes con HUES9 en comparación con los fibroblastos BJ. (E) Formación de teratoma del clon BJ-OKS-iM n°21 en ratones nude. AE1/AE3 (citoqueratina), NF-1 (células neuronales) y GFAP (células neuronales) usados para marcadores de ectoderma; Desmin (células musculares) usadas para marcador de mesoderma; y AFP (endoderma primitivo y definitivo) usado para marcador de endoderma. Barra, 100 µm.

Figura 4: muestra el análisis RT-PCR para comprobar la existencia de replicón de ARN. Medida de la sensibilidad de PCR con el plásmido de replicón OKS-iM-RNA. Se realizaron PCRs para las regiones nsP2, nsP4 y OCT4-T2A-KLF4 (OK) con 100, 10 y 1 fg de plásmido (Panel superior). RT-PCR de clones de iPSCs HFF-OKS-iM. +; control positivo, el ARN total se preparó desde un día después de la transfección de replicón OKS-iM-RNA. -; control negativo, el ARN total se preparó a partir de HFFs transfectadas de imitación. Los ARNs totales de clones de células iPS se prepararon a partir del pasaje 8 (Panel inferior).

Figura 5: muestra el análisis del cariotipo de clones de células iPS. El cariotipado de banda-G de los clones HFF-OKS-iM-1, BJ-OKS-iM-2, BJ-OKS-iM-21 y BJ-OKS-iG-5 se llevó a cabo en veinte células de metafase de banda-G de cada clon y se evaluó como cariotipo humano masculino normal en todos los clones (línea celular GENETICS).

Figura 6A-B: muestra clones de células iPS cultivados con células cebadoras STO. Las células fueron recolectadas y después inyectadas intramuscularmente o subcutáneamente en los músculos de la pata trasera o en el flanco dorsal de ratones nude. Después de 5 a 8 semanas de la inyección, los tumores fueron diseccionados y fijados con paraformaldehído al 4% (A) Análisis de teratoma de clon HFF-OKS-iM n°1 en ratones nude. AE1/AE3 (citoqueratina), NF-1 (células neuronales) usados para marcadores de ectoderma; Desmin (células musculares) usadas para marcador de mesoderma; y AFP (endoderma primitivo y definitivo) usado para marcador de endoderma. Barra, 100 µm. (B) Tinción H&E de teratomas de clones 3 y 5 de BJ-OKS-iG. Barra, 100 µm.

Figura 7A-D: muestra que (A) el medio acondicionado con B18R es útil para la existencia persistente de replicón de ARN de VEE. Arriba; % de células positivas de GFP, Abajo: valor medio de fluorescencia GFP en la población positiva de GFP. (B) Fotografías de células. Barra, 200 µm. (C) Expresión de proteínas de RFs en el día 10 como se indica. (D) El B18R-CM es requerido para la generación de células iPS en cultivo de cebador. Se co-transfectaron HFFs con ARN de OKS-iM y ARNm de B18R como se ha indicado, y a continuación las células fueron cultivadas en presencia de B18R-CM y puromicina. Las células fueron pasadas a células cebadoras STO en el día 10 (transfecciones en los días 1, 3, 8) o en el día 11 (transfecciones en los días 1, 4, 7, 10), y se cultivaron en presencia o en ausencia de B18R-CM más/menos puromicina.

Descripción detallada

Tal como se usa en la presente memoria y en las reivindicaciones anexas, las formas singulares “un”, “una” y “el/la” incluyen los referentes plurales a menos que el contexto claramente indique lo contrario. Por tanto, por ejemplo, la referencia a “una célula” incluye una pluralidad de dichas células y la referencia al “agente” incluye referencia a uno o más agentes conocidos por los especialistas en la técnica, etcétera.

Asimismo, el uso de “o” significa “y/o” a menos que se especifique lo contrario. De forma similar, “comprende”, “comprende”, “que comprende”, “incluye” y “que incluye” son intercambiables y no pretenden ser limitativos.

Debe entenderse asimismo que cuando las descripciones de las diversas realizaciones usan el término “que comprende”, los especialistas en la técnica entenderá que en algunos casos específicos una realización puede ser descrita alternativamente usando el lenguaje “que consiste esencialmente en” o “que consiste en”.

Aunque se pueden usar métodos y materiales similares o equivalentes a los descritos en la presente memoria al llevar a la práctica los métodos y composiciones descritos, los ejemplos de métodos, dispositivos y materiales se describen en la presente memoria.

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en la presente memoria tienen el mismo significado empleado habitualmente por los especialistas en el campo de la técnica a la cual pertenece la descripción. Por tanto, tal como se usan a lo largo de la presente solicitud, los siguientes términos tendrán los siguientes significados.

Aunque las células madre pluripotentes inducidas (células iPS) son virtualmente idénticas a las células ES a un nivel molecular y funcional, existen impedimentos críticos para la traducción de sus potenciales terapéuticos en aplicaciones médicas. Uno de los problemas es que debido a que los protocolos estándar actuales para la reprogramación y propagación de células iPS incluyen materiales derivados de animales que no son adecuados para fines clínicos potenciales, es necesario desarrollar un método completamente definido para generar y expandir células hiPS.

Las células madre pluripotentes inducidas (iPS) son descritas por el equipo de Shinya Yamanaka de la Universidad de Tokio, Japón. Yamanaka identificó genes que son particularmente activos en células madre embrionarias, y usó retrovirus para transfectar fibroblastos de ratón con una selección de dichos genes. Eventualmente, se aislaron cuatro genes clave de pluripotencia esenciales para la producción de células madre pluripotentes; Oct-3/4, SOX2, c-Myc y

Klf4. Las células fueron aisladas mediante selección de antibióticos para células Fbx15⁺. El mismo grupo publicó un estudio junto a otros dos grupos de investigación independientes de Harvard, MIT, y la Universidad de California, Los Angeles, que demostró con éxito la reprogramación de fibroblastos de ratón en iPS e incluso la producción de una quimera viable.

La generación de células iPS humanas mediante expresión retroviral de cuatro factores de reprogramación (RFs; también denominados factores de des-diferenciación) abrió el potencial de terapias de medicina regenerativa basadas en células madre personalizadas, específicas de paciente. Sin embargo, el potencial mutágeno insercional de retrovirus combinado con el potencial de la activación de gen RF latente, especialmente *c-MYC*, descarta completamente las estrategias basadas en ADN integrativo para uso en terapias de medicina regenerativa. Se han desarrollado otras estrategias de iPS basadas en ADN empleando vectores episomales, adenovirus, transposón *piggyBac* integrado y escindido o lentivirus floxeado; sin embargo, estas estrategias o bien adolecen de una baja eficiencia de generación de células iPS o bien requieren estrategias de escisión genómica que dejan atrás una etiqueta de elemento de ADN insertado. Las estrategias de células iPS basadas en ARN con virus Sendai o transfección de ARNm evitan los potenciales problemas de integración asociados a las estrategias basadas en ADN y son métodos inherentemente más seguros para aplicaciones clínicas. Aunque el virus Sendai ofrece una estrategia de iPS razonablemente eficiente, los problemas asociados a la replicación persistente de virus Sendai en clones de células iPS requiere una etapa de selección negativa seguida de varias etapas de reclonación desde el nivel de célula individual hasta aislar iPS libres de virus aislado, dichos procesos resultan en una división y pasaje de células iPS excesivos. Una de las estrategias no basadas en ADN más prometedoras implica la transfección diaria de cuatro ARNms de RF individuales (más ARNm de GFP) a lo largo de 16 días. Desafortunadamente, esta estrategia sigue siendo problemática. Por ejemplo, se han llevado a cabo experimentos para reemplazar los retrovirus *KLF4* y *c-MYC* con los correspondientes ARNms transfectados y los resultados fueron validados; sin embargo los retrovirus *OCT4* y *SOX2* no pudieron ser reemplazados con ARNms transfectados. El problema parece tener su origen en la rápida degradación de los ARNms de RF combinada con el nivel inconsistente de expresión de umbral célula-a-célula con el tiempo, lo que deriva del intento de transfectar cuatro ARNms independientes en la misma célula en una base diaria durante >14 días durante la reprogramación. Consecuentemente, sigue existiendo una necesidad significativa de una estrategia no basada en ADN sencilla y altamente reproducible para generar células iPS humanas.

La descripción proporciona métodos y composiciones para generar células iPS a partir de células somáticas (p.ej., células de fibroblastos). Las composiciones y el método comprenden el uso de replicones derivados de alfavirus. Los replicones comprenden una secuencia de ARN que codifica proteínas de alfavirus no estructurales necesarias para la replicación y 1, 2, 3, 4 o más secuencias codificadoras heterólogas al alfavirus y que inducen la des-diferenciación de células somáticas en fenotipos de célula madre.

Tal como se usa en la presente memoria, el término "alfavirus" tiene su significado convencional en la técnica, e incluye diversas especies tales como virus de encefalitis equina venezolana (VEE), virus de encefalitis equina oriental (EEE), virus de Everglades (EVE), virus Mucambo (MUC), virus Pixuna (PIX) y virus de Encefalitis Equina Occidental, todos los cuales son miembros del grupo VEE/EEE de alfavirus. Otros alfavirus incluyen, p.ej., el virus de Bosque de Semliki (SFV), virus Sindbis, virus Ross River, virus Chikungunya, S.A. AR86, virus del Bosque de Barmah, virus Middelburg, virus O'nyong-nyong, virus Getah, virus Sagiyama, virus Bebaru, virus Mayaro, virus Una, virus Aura, virus Whataroa, virus Babanki, virus Kyzylagach, virus J de las Highlands, virus Fort Morgan, virus Ndumu y virus Buggy Creek. Los alfavirus particularmente útiles en las construcciones y métodos descritos en la presente memoria son alfavirus del grupo VEE/EEE.

Los términos "replicón de ARN de alfavirus", "ARN de replicón de alfavirus", "replicón de vector de ARN de alfavirus" y "ARN de replicón de vector" se usan de forma intercambiable para referirse a una molécula de ARN que expresa genes de proteína no estructurales, de tal modo que puede dirigir su propia replicación (amplificación) y comprende, como mínimo, secuencias de reconocimiento de replicación de alfavirus 5' y 3', secuencias de codificación para proteínas no estructurales de alfavirus, y un tracto de poliadenilación. Adicionalmente puede contener uno o más elementos (p.ej., secuencias IRES, promotores de núcleo o mini-promotores y similares) para dirigir la expresión, lo que se refiere a la transcripción y la traducción, de una secuencia de ARN heteróloga. El replicón de alfavirus de la descripción puede comprender, en una realización, secuencias de reconocimiento de replicación de alfavirus 5' y 3', secuencias de codificación para proteínas no estructurales de alfavirus, un tracto de poliadenilación y una o más secuencias de codificación seleccionadas del grupo que consiste en SOX-2, *c-Myc*, OCT-3/4, Klf, Glis1 y Nanog.

El término "polinucleótido", "ácido nucleico" o "ácido nucleico recombinante" se refiere a polinucleótidos tales como el ácido desoxirribonucleico (ADN), y, cuando sea apropiado (particularmente en referencia a un replicón), ácido ribonucleico (ARN).

El término "expresión" con respecto a un gen o polinucleótido se refiere a la transcripción del gen o polinucleótido y, según sea apropiado, a la traducción de un transcrito de ARNm a una proteína o polipéptido. Así, como quedará claro a partir del contexto, la expresión de una proteína o polipéptido es el resultado de la transcripción y/o traducción del marco de lectura abierto.

Los especialistas en la técnica reconocerán que, debido a la naturaleza degenerada del código genético, se puede usar una variedad de codones que se diferencian en sus secuencias de nucleótido para codificar un aminoácido dado.

Un polinucleótido o secuencia génica particular que codifica un polipéptido descrito en la presente memoria son referenciados meramente para ilustrar una realización de la descripción, y la descripción incluye polinucleótidos de cualquier secuencia que codifica un polipéptido que comprende la misma secuencia de aminoácidos de los polipéptidos y proteínas de las enzimas utilizadas en los métodos de la descripción. De un modo similar, un polipéptido puede tolerar típicamente una o más sustituciones, eliminaciones e inserciones de aminoácidos en su secuencia de aminoácidos sin perder o sin presentar una pérdida significativa de una actividad deseada. La descripción incluye dichos polipéptidos con secuencias de aminoácidos alternativas, y las secuencias de aminoácidos codificadas por las secuencias de ARN o ADN mostradas en la presente memoria meramente ilustran realizaciones de la descripción.

La descripción proporciona polinucleótidos en la forma de vectores de expresión de ADN recombinante, replicones de ARN o plásmidos, como se describe más detalladamente en otros sitios de la presente memoria, que codifican uno o más polipéptidos.

Un polipéptido de la descripción puede ser amplificado usando ADNc, ARNm o alternativamente, ADN genómico, como plantilla y los cebadores de oligonucleótido apropiados según las técnicas de amplificación PCR estándares y los procedimientos descritos más adelante en la sección de Ejemplos. El ácido nucleico amplificado de este modo puede ser clonado en un vector apropiado y caracterizarse mediante análisis de secuencia. Adicionalmente, se pueden preparar los oligonucleótidos correspondientes a las secuencias de nucleótido mediante técnicas sintéticas estándar, p.ej., empleando un sintetizador de ADN automatizado.

En una realización, un replicón de la descripción comprende una secuencia que es idéntica en un 90%, 95%, 98%, 99% o 100% a la SEQ ID NO:29, 30, 31 o 32, desde aproximadamente la posición 1 hasta aproximadamente la posición 7561 (incluyendo donde "T" de la secuencia puede ser sustituida por "U"), seguida de dos o más RFs seleccionados del grupo que consiste en Oct-3/4, Sox-2, Klf4, c-Myc, Nanog y Glis1. Cuando hay presente más de un RF, las secuencias codificadoras pueden separarse mediante un sitio de entrada de ribosoma (IRES) o un promotor pequeño (p.ej., un núcleo) tal como SP1. El orden de los RFs no es crítico para la descripción; por tanto el orden puede ser Klf4, Oct-3/4, Sox-2, c-Myc o puede ser Sox-2, Klf4, Oct-3/4, c-Myc, u Oct4, Klf4, Sox2, c-Myc o cualquier variación del orden de los RFs. El replicón puede comprender además un marcador seleccionable (p.ej., un marcador de resistencia a antibiótico). Las secuencias codificadoras de los RFs pueden estar separadas por péptidos de auto-ruptura tales como T2A y/o E2A.

El replicón puede comprender desde 5' a 3': (replicasas de ARN de VEE) – (promotor 26S) – (RF₁) – (péptido de auto-ruptura) – (RF₂) – (péptido de auto-ruptura) – (RF₃) – (IRES o promotor central) – (RF₄) – (IRES o promotor opcional) – (marcador seleccionable opcional) – (3'UTR y cola poliA de VEE); donde RF₁₋₄ son factores que inducen la diferenciación de una célula somática a células pluripotentes, donde RF₂₋₃ son opcionales, RF₃₋₄ son opcionales o RF₄ es opcional; donde RF₁₋₄ se seleccionan del grupo que consiste en Oct-4, Klf4, Sox-2, c-Myc, Nanog y Glis1. El replicón de lo anterior es una molécula de ARN. En una realización adicional, el replicón se deriva de VEE e incluye una mutación para reducir la patogenicidad. El VEE puede ser un replicón de ARN basado en la cepa TC-83 (cepa de vacuna) con una mutación puntual (mutación nsP2P₇₇₃ a S), que redujo el efecto citopático del replicón.

En cualquiera de las anteriores realizaciones, los RFs incluyen variantes y secuencias de polinucleótido generadas. Por ejemplo, un RF puede comprender homólogos y variantes de un polipéptido OCT-4, un polipéptido KLF4, un polipéptido SOX-2, un polipéptido c-MYC, un polipéptido NANOG o GLIS1. Por ejemplo, una secuencia codificadora de RF para NANOG útil en cualquiera de las realizaciones de replicón descritas en la presente memoria puede comprender (i) un polinucleótido que codifica un polipéptido de la SEQ ID NO:2; (ii) un polinucleótido que comprende al menos un 95% de identidad con respecto a la SEQ ID NO:1 y que codifica un polipéptido que tiene actividad de NANOG; (iii) un polinucleótido que tiene una secuencia como la establecida en la SEQ ID NO:1 o (iv) un polinucleótido que codifica un polipéptido de la SEQ ID NO:2 que contiene de 1 a 10 sustituciones de aminoácido conservativas y donde el polipéptido presenta actividad Nanog; y donde cualquiera de las anteriores secuencias de ácido nucleico puede tener "T" reemplazado por "U". Por ejemplo, una secuencia codificadora de RF para Oct-4 útil en cualquiera de las realizaciones de replicón descritas en la presente memoria puede comprender (i) un polinucleótido que codifica un polipéptido de la SEQ ID NO:4; (ii) un polinucleótido que comprende al menos un 95% de identidad con respecto a la SEQ ID NO:3 y que codifica un polipéptido que tiene actividad Oct-4; (iii) un polinucleótido que tiene una secuencia como la establecida en la SEQ ID NO:3 o (iv) un polinucleótido que codifica un polipéptido de la SEQ ID NO:4 que contiene de 1 a 10 sustituciones de aminoácido conservativas y donde el polipéptido presenta actividad Oct-4; y donde cualquiera de las anteriores secuencias de ácido nucleico puede tener "T" reemplazado por "U". Por ejemplo, una secuencia codificadora de RF para Sox-2 útil en cualquiera de las realizaciones de replicón descritas en la presente memoria puede comprender (i) un polinucleótido que codifica un polipéptido de la SEQ ID NO:6; (ii) un polinucleótido que comprende al menos un 95% de identidad con respecto a la SEQ ID NO:5 y que codifica un polipéptido que tiene actividad Sox-2; (iii) un polinucleótido que tiene una secuencia como la establecida en la SEQ ID NO:5 o (iv) un polinucleótido que codifica un polipéptido de la SEQ ID NO:6 que contiene de 1 a 10 sustituciones de aminoácido conservativas y donde el polipéptido presenta actividad SOX-2; y donde cualquiera de las anteriores secuencias de ácido nucleico puede tener "T" reemplazado por "U". Por ejemplo, una secuencia codificadora de RF para KLF4 útil en cualquiera de las realizaciones de replicón descritas en la presente memoria puede comprender (i) un polinucleótido que codifica un polipéptido de la SEQ ID NO:8; (ii) un polinucleótido que comprende al menos un 95% de identidad con respecto a la SEQ ID NO:7 y que codifica un polipéptido que tiene actividad KLF4; (iii) un polinucleótido que tiene una secuencia como la establecida en la SEQ ID NO:7 o (iv) un polinucleótido que codifica un polipéptido de la SEQ

ID NO:8 que contiene de 1 a 10 sustituciones de aminoácido conservativas y donde el polipéptido presenta actividad KLF4; y donde cualquiera de las anteriores secuencias de ácido nucleico puede tener "T" reemplazado por "U". Por ejemplo, una secuencia codificadora de RF para c-MYC útil en cualquiera de las realizaciones de replicón descritas en la presente memoria puede comprender (i) un polinucleótido que codifica un polipéptido de la SEQ ID NO:10; (ii) un polinucleótido que comprende al menos un 95% de identidad con respecto a la SEQ ID NO:9 y que codifica un polipéptido que tiene actividad c-MYC; (iii) un polinucleótido que tiene una secuencia como la establecida en la SEQ ID NO:9 o (iv) un polinucleótido que codifica un polipéptido de la SEQ ID NO:10 que contiene de 1 a 10 sustituciones de aminoácido conservativas y donde el polipéptido presenta actividad c-MYC; y donde cualquiera de las anteriores secuencias de ácido nucleico puede tener "T" reemplazado por "U". Por ejemplo, una secuencia codificadora de RF para GLIS1 útil en cualquiera de las realizaciones de replicón descritas en la presente memoria puede comprender (i) un polinucleótido que codifica un polipéptido de la SEQ ID NO:34; (ii) un polinucleótido que comprende al menos un 95% de identidad con respecto a la SEQ ID NO:33 y que codifica un polipéptido que tiene actividad GLIS1; (iii) un polinucleótido que tiene una secuencia como la establecida en la SEQ ID NO:33 o (iv) un polinucleótido que codifica un polipéptido de la SEQ ID NO:34 que contiene de 1 a 10 sustituciones de aminoácido conservativas y donde el polipéptido presenta actividad GLIS1; y donde cualquiera de las anteriores secuencias de ácido nucleico puede tener "T" reemplazado por "U".

Nanog es un gen expresado en células madre embrionarias (ESCs) y desempeña un papel en el mantenimiento de la pluripotencia. Se cree que Nanog actúa con SOX2. En la SEQ ID NO:1 y 2 se presentan un polinucleótido y un polipéptido que codifican Nanog, respectivamente. Adicionalmente, la SEQ ID NO:1 comprende una secuencia de ADN, en la que se reconocerá que "T" puede ser reemplazado por "U". La proteína humana NANOG (véase, p.ej., Número de acceso NP_079141) es una proteína de 305 aminoácidos con una estructura de homeodominio que está localizada en el componente nuclear de las células. De forma similar a NANOG, la región N-terminal de NANOG humana es rica en residuos de Ser, Thr y Pro y el extremo C comprende repeticiones de Trp. El homeodominio en la NANOG humana va desde aproximadamente el residuo 95 hasta aproximadamente el residuo 155. Se conocen homólogos de Nanog humana.

Un "polipéptido Oct" se refiere a cualquiera de los miembros existentes de la familia Octamer de factores de transcripción, o a variantes de los mismos que mantienen la actividad de factor de transcripción, similares (con una actividad de al menos el 50%, 80% o 90%) con respecto al miembro de la familia natural más próximo, o a polipéptidos que comprenden al menos el dominio de unión a ADN del miembro de la familia natural, y además pueden comprender un dominio de activación transcripcional. Los ejemplos de polipéptidos Oct incluyen Oct-1, Oct-2, Oct-3/4, Oct-6, Oct-7, Oct-8, Oct-9 y Oct-11. Por ejemplo, Oct-3/4 (aquí referido como "Oct4") contiene el dominio POU, una secuencia de 150 aminoácidos conservada entre Pit-1, Oct-1, Oct-2 y uric-86. Véase Ryan, A. K. & Rosenfeld, M. G. *Genes Dev.* 11, 1207-1225 (1997). En algunas realizaciones, las variantes presentan al menos una identidad de secuencia de aminoácidos del 85%, 90% o 95% a lo largo de toda su secuencia en comparación con un miembro de la familia de polipéptidos Oct natural, tal como los enumerados anteriormente o tal como los enumerados en el número de acceso del Genbank NP002692.2 (Oct4 humano) o NP038661.1 (Oct4 de ratón). Los polipéptidos Oct (p.ej., Oct3/4) pueden proceder de humano, ratón, rata, bovino, porcino u otros animales. Generalmente, se usará la misma especie de proteína de la especie de células que están siendo manipuladas. Oct-4 (Octamer-4) es un factor de transcripción de homeodominio de la familia POU y regula la expresión de numerosos genes (véase, p.ej., *J. Biol. Chem.*, Vol. 282, Número 29, 21551-21560, 20 de julio, 2007). En la SEQ ID NO:3 y 4 se presentan un polinucleótido y un polipéptido que codifican un Oct4, respectivamente. Adicionalmente, la SEQ ID NO:3 comprende una secuencia de ADN, en la que se reconocerá que "T" puede ser reemplazado por "U". Se conocen homólogos de Oct-4 humano, como se establece en los siguientes números de acceso NP_038661.1 y NM_013633.1 (*Mus musculus*), NP_001009178 y NM_001009178 (*Rattus norvegicus*), y NP_571187 y NM_131112 (*Danio rerio*).

SRY (región Y determinante del sexo)-box 2, también conocido como SOX2, es un factor de transcripción que desempeña una función en la auto-renovación de células madre embrionarias no diferenciadas y en la transactivación de Fgf4, así como en la modulación del doblado de ADN (véase, p.ej., Scaffidi et al., *J. Biol. Chem.*, Vol. 276, Número 50, 47296-47302, 14 de diciembre, 2001). Un "polipéptido Sox" se refiere a cualquiera de los miembros naturales de los factores de transcripción HMG-box relacionados con SRY (Sox), que se caracterizan por la presencia del dominio de grupo de alta movilidad (HMG), o a variantes de los mismos que mantienen una actividad de factor de transcripción similar (con una actividad de al menos el 50%, 80% o 90%) con respecto al miembro de la familia natural más próximo, o a polipéptidos que comprenden al menos el dominio de unión a ADN del miembro de la familia natural, y además pueden comprender un dominio de activación transcripcional. Véase, p.ej., Dang, D. T., et al., *Int. J. Biochem. Cell Biol.* 32: 1103-1121 (2000). Los ejemplos de polipéptidos Sox incluyen, p.ej., Sox1, Sox2, Sox3, Sox4, Sox5, Sox6, Sox7, Sox8, Sox9, Sox10, Sox11, Sox12, Sox13, Sox14, Sox15, Sox17, Sox18, Sox-21 y Sox30. Se ha demostrado que Sox1 produce células iPS con una eficiencia similar a Sox2, y también se ha demostrado que los genes Sox3, Sox15 y Sox18 generan células iPS, aunque con algo menos de eficiencia que el Sox2. Véase Nakagawa et al., *Nature Biotechnology* 26: 101-106 (2007). En algunas realizaciones, las variantes presentan al menos una identidad de secuencia de aminoácidos del 85%, 90% o 95% a lo largo de su secuencia completa en comparación con un miembro de la familia de polipéptidos Sox naturales, tal como los enumerados anteriormente o como el incluido en el número de acceso Genbank CAA83435 (Sox2 humano). Los polipéptidos Sox (p.ej., Sox1, Sox2, Sox3, Sox15 o Sox18) pueden proceder de humano, ratón, rata, bovino, porcino u otros animales. Generalmente, se empleará la misma especie de proteína que la especie de las células que están siendo manipuladas. En la SEQ ID NO:5 y 6 se presentan

un polinucleótido y un polipéptido que codifican un Sox2, respectivamente. Adicionalmente, la SEQ ID NO:5 comprende una secuencia de ADN, en la que se reconocerá que "T" puede ser reemplazado por "U". Se conocen homólogos de Sox2 humano.

El factor 4 de tipo Kruppel, también conocido como KLF4, desempeña una función en el mantenimiento y crecimiento de las células madre. Un "polipéptido Klf" se refiere a cualquiera de los miembros naturales de la familia de factores de tipo Kruppel (Klfs), proteína de dedo de zinc que contienen secuencias de aminoácido similares a las de regulador Kruppel de estructura embrionaria de *Drosophila*, o a variantes de los miembros naturales que mantienen una actividad de factor de transcripción similar (con una actividad de al menos el 50%, 80% o 90%) con respecto al miembro de la familia natural más próximo, o a polipéptidos que comprenden al menos el dominio de unión a ADN del miembro de la familia natural, y además pueden comprender un dominio de activación transcripcional. Véase Dang, D. T., Pevsner, J. & Yang, V. W., *Cell Biol.* 32, 1103-1121 (2000). Los ejemplos de miembros de la familia de Klf incluyen Klf1, Klf2, Klf3, Klf4, Klf5, Klf6, Klf7, Klf8, Klf9, Klf10, Klf11, Klf12, Klf13, Klf14, Klf15, Klf16 y Klf17. Se descubrió que Klf2 y Klf4 son factores capaces de generar células iPS en ratones, así como los genes relacionados Klf1 y Klf5, aunque con una menor eficiencia. Véase Nakagawa, et al., *Nature Biotechnology* 26: 101-106 (2007). En algunas realizaciones, las variantes tienen al menos una identidad de secuencia de aminoácidos de al menos 85%, 90% o 95% a lo largo de toda su secuencia en comparación con un miembro de la familia de polipéptidos de Klf natural tal como los enumerados anteriormente o como los incluidos en el número de acceso de Genbank CAX16088 (Klf4 de ratón) o CAX14962 (Klf4 humano). Los polipéptidos de Klf (p.ej., Klf1, Klf4 y Klf5) pueden proceder de humano, ratón, rata, bovino, porcino u otros animales. Generalmente, se usará la misma especie de proteína que la especie de las células que están siendo manipuladas. Para el alcance en que se describe un polipéptido Klf en la presente memoria, puede ser reemplazado por un polipéptido de receptor beta relacionado con estrógeno (Essrb). De esta manera, se pretende que para cada realización de polipéptido de Klf descrita en la presente memoria, se describa igualmente una correspondiente realización que emplea Essrb en lugar de un polipéptido Klf4. En la SEQ ID NO:7 y 8 se presentan un polinucleótido y un polipéptido que codifican un KLF4, respectivamente. Adicionalmente, la SEQ ID NO:7 comprende una secuencia de ADN, en la que se reconocerá que "T" puede ser reemplazado por "U". Se conocen homólogos de KLF4 humano e incluyen NP_034767, NM_010637 (*Mus musculus*).

La familia MYC de genes celulares está compuesta por c-myc, N-myc y L-myc, tres genes que actúan en la regulación de la proliferación, diferenciación y apoptosis de células (Henriksson y Luscher 1996; Facchini y Penn 1998). Un "polipéptido Myc" se refiere a cualquiera de los miembros naturales de la familia Myc (véase, p.ej., Adhikary, S. & Eilers, M. *Nat. Rev. Mol. Cell Biol.* 6: 635-645 (2005)), o variantes de los mismos que mantienen una actividad similar de factor de transcripción (con una actividad de al menos el 50%, 80% o 90%) con respecto al miembro de la familia natural más próximo, o a polipéptidos que comprenden al menos el dominio de unión a ADN del miembro de la familia natural, y además pueden comprender un dominio de activación transcripcional. Los ejemplos de polipéptidos Myc incluyen, p.ej., c-Myc, N-Myc y L-Myc. En algunas realizaciones, las variantes presentan una identidad de secuencia de aminoácidos de al menos 85%, 90% o 95% a lo largo de su secuencia completa con respecto a un miembro de la familia de polipéptidos Myc natural, tal como los enumerados anteriormente o el incluido en el número de acceso Genbank CAA25015 (Myc humano). Los polipéptidos Myc (p.ej., c-Myc) pueden proceder de humano, ratón, rata, bovino, porcino u otros animales. Generalmente, se usará la misma especie de proteína que la especie de células que están siendo manipuladas. Aunque los genes de la familia myc presentan una estructura y una actividad biológica comunes. El n-Myc es un miembro de la familia MYC que codifica una proteína con un dominio básico hélice-bucle-hélice (bHLH). Las estructuras genómicas de c-myc y N-myc se organizan de forma similar y constan de tres exones. La mayor parte del primer exón y la porción 3' del tercer exón contienen regiones no traducidas que portan secuencias reguladoras transcripcionales o post-transcripcionales. La proteína N-myc se encuentra en el número y dimeriza con otra proteína bHLH para unirse a ADN. En la SEQ ID NO:9 y 10 se presentan un polinucleótido y un polipéptido que codifican un c-Myc, respectivamente. Adicionalmente, la SEQ ID NO:9 comprende una secuencia de ADN, en la que se reconocerá que "T" puede ser reemplazado por "U". En la técnica se conocen homólogos y variantes de la familia Myc de proteínas.

Glis1 (dedo de zinc 1 de la familia Glis) es un gen que codifica una proteína de tipo Krüppel del mismo nombre, cuya localización está en el cromosoma 1p32.3. El gen está enriquecido en óvulos no fertilizados y en embriones en la etapa de una célula y puede usarse para promover la reprogramación directa de células somáticas a células madre pluripotentes inducidas. Glis1 se puede utilizar como uno de los cuatro factores usados en la reprogramación de células somáticas en células madre pluripotentes inducidas. Los otros tres factores de transcripción usados son Oct3/4, Sox2 y Klf4. Un Glis1 humano (NM_147193) se presenta en la SEQ ID NO:33 y 34 (ADNc y polipéptido, respectivamente).

El ADNc que codifica para oct4 (pour5f1), sox2, klf4, c-myc (n-myc o L-myc), Glis1 y nanog humanos, sus variantes y homólogos, puede ser clonado y expresado usando técnicas conocidas en el campo de la técnica. Usando las secuencias establecidas en la presente memoria se pueden clonar polinucleótidos que codifican uno o más factores de des-diferenciación en un vector adecuado para la expresión en un tipo celular de interés.

Una "actividad" RF (p.ej., una actividad de variante de RF) se refiere a la capacidad para des-diferenciar una célula somática cuando se expresa en combinación con otros RFs, tal como es conocido en la técnica. Por ejemplo, una variante de Oct-4 se puede evaluar en términos de actividad de Oct-4 co-expresando la variante de Oct-4 en una célula somática con klf4, Sox-2 y c-myc y determinando si la célula somática se des-diferencia. Si la célula se des-diferencia, entonces se puede decir que la variante de Oct-4 presenta actividad de Oct-4.

En otra realización, el replicón comprende una secuencia como la establecida en la SEQ ID NO:29, 30, 31 o 32. En otra realización adicional, el replicón comprende una secuencia que idéntica en aproximadamente el 90%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% a la SEQ ID NO:29, 30, 31 o 32, y donde cuando el replicón es transfectado en una célula somática, la célula somática es "inducida" para convertirse en una célula madre. Adicionalmente, cualquiera de SEQ ID NO:29, 30, 31 o 32, donde "T" es reemplazado por "U".

En una realización, la SEQ ID NO:29 proporciona un replicón de la descripción. En otra realización, la secuencia de la SEQ ID NO:29 tiene "T" reemplazado por "U". El replicón comprende ARN replicasas de VEE desde el nucleótido 1 hasta aproximadamente el nucleótido 7561, una secuencia Oct-4 humana desde el nucleótido 7592 hasta 8671, una secuencia codificadora para un péptido de auto-ruptura T2A desde el nucleótido 8678-8731, una secuencia Klf4 humana desde 8738-10147, una secuencia codificadora para un péptido de auto-ruptura E2A desde el nucleótido 10154-10213, una secuencia de Sox-2 humana desde 10223-11176, un sitio de entrada de ribosoma interno desde 11195-11805, una secuencia de c-Myc humana desde 11818-13140, un sitio de entrada de ribosoma interno desde 13165-13776, un gen de resistencia de puromicina desde 13777-14376, el 3'UTR y cola poli-A de VEE desde 14383-14510, un gen de resistencia de ampicilina desde 14679-15539 y un promotor SP6 desde 16320-16337.

En una realización, la SEQ ID NO:30 proporciona un replicón de la descripción. En otra realización, la secuencia de la SEQ ID NO:30 tiene "T" reemplazado por "U". El replicón comprende ARN replicasas de VEE desde el nucleótido 1 hasta aproximadamente el nucleótido 7561, una secuencia Oct-4 humana desde el nucleótido 7592 hasta 8671, una secuencia codificadora para un péptido de auto-ruptura T2A desde el nucleótido 8678-8731, una secuencia Klf4 humana desde el nucleótido 8738-10147, una secuencia codificadora para un péptido de auto-ruptura E2A desde el nucleótido 10154-10213, una secuencia de Sox-2 humana desde 10223-11176, un sitio de entrada de ribosoma interno desde 11195-11805, una secuencia de c-Myc humana desde 11818-13140, un sitio de entrada de ribosoma interno desde 13165-13776, un gen de resistencia de puromicina desde 13777-14376, el 3'UTR y cola poli-A de VEE desde 14383-14510, un gen de resistencia de ampicilina desde 14679-15539 y un promotor T7 desde 16319-16336.

En una realización, la SEQ ID NO:31 proporciona un replicón de la descripción. En otra realización, la secuencia de la SEQ ID NO:31 tiene "T" reemplazado por "U". El replicón comprende ARN replicasas de VEE desde el nucleótido 1 hasta aproximadamente el nucleótido 7561, una secuencia Oct-4 humana desde el nucleótido 7592 hasta 8671, una secuencia codificadora para un péptido de auto-ruptura T2A desde el nucleótido 8678-8731, una secuencia Klf4 humana desde el nucleótido 8738-10147, una secuencia codificadora para un péptido de auto-ruptura E2A desde el nucleótido 10154-10213, una secuencia de Sox-2 humana desde 10223-11176, un sitio de entrada de ribosoma interno desde 11195-11805, una secuencia de Glis1 humana desde 11818-13680, un sitio de entrada de ribosoma interno desde 13689-14300, un gen de resistencia de puromicina desde 14301-14900, el 3'UTR y cola poli-A de VEE desde 14907-15034, un gen de resistencia de ampicilina desde 15203-16063 y un promotor SP6 desde 16844-16861.

En una realización, la SEQ ID NO:32 proporciona un replicón de la descripción. En otra realización, la secuencia de la SEQ ID NO:32 tiene "T" reemplazado por "U". El replicón comprende ARN replicasas de VEE desde el nucleótido 1 hasta aproximadamente el nucleótido 7561, una secuencia Oct-4 humana desde el nucleótido 7592 hasta 8671, una secuencia codificadora para un péptido de auto-ruptura T2A desde el nucleótido 8678-8731, una secuencia Klf4 humana desde el nucleótido 8738-10147, una secuencia codificadora para un péptido de auto-ruptura E2A desde el nucleótido 10154-10213, una secuencia de Sox-2 humana desde 10223-11176, un sitio de entrada de ribosoma interno desde 11195-11805, una secuencia de Glis1 humana desde 11818-13680, un sitio de entrada de ribosoma interno desde 13689-14300, un gen de resistencia de puromicina desde 14301-14900, el 3'UTR y cola poli-A de VEE desde 14907-15034, un gen de resistencia de ampicilina desde 15203-16063 y un promotor T7 desde 16843-16860.

En otra característica de la descripción, se puede usar más de un replicón de alfavirus, comprendiendo cada replicón una o más secuencias codificadoras para factores que inducen a una célula somática a convertirse en una célula madre, donde la combinación de los más de uno replicones de alfavirus incluyen toda la secuencia codificadora correspondiente a todos los RFs necesarios para inducir la des-diferenciación en una célula madre.

En realizaciones más específicas, un replicón de alfavirus comprende secuencias codificadoras para la expresión de OCT-3/4, SOX-2, KLF, c-MYC, GLIS1 y/o NANOG. El replicón de alfavirus puede comprender secuencias codificadoras para OCT-4, KLF4, SOX-2, GLIS1 y c-MYC.

El replicón también puede ser diseñado para expresar proteínas estructurales de alfavirus. Las Patentes de EE.UU. nº 7.045.335, 7.078.218, 7.425.337 y 7.442.381 describen numerosas construcciones para dichos replicones de ARN que consisten en las secuencias de reconocimiento de replicación de alfavirus 5' y 3', las secuencias codificadoras para proteínas no estructurales de alfavirus, y un tracto de poliadenilación. Los replicones de ARN de alfavirus pueden contener una o más mutaciones atenuantes, siendo una mutación atenuante una eliminación, adición o sustitución de uno o más nucleótido(s), o una mutación que comprende el reordenamiento o la construcción quimérica que da como resultado una pérdida de la virulencia en un virus vivo que contiene la mutación en comparación con el alfavirus natural apropiado.

Los términos "proteína(s) estructural(es) de alfavirus" se refieren a una, o una combinación de proteínas estructurales codificadas por alfavirus. Éstas son producidas por el virus como una poliproteína y se representan generalmente en la bibliografía como C-E3-E2-6k-E1. E3 y 6k actúan como señales de traslocalización/transporte de membrana para

las dos glicoproteínas, E2 y E1. De esta manera, el uso del término E1 en la presente memoria se puede referir a E1, E3-E1, 6k-E1 o E3-6k-E1, y el uso del término E2 en la presente memoria se puede referir a E2, E3-E2, 6k-E2 o E3-6k-E2. Se pueden introducir mutaciones atenuantes en una cualquiera o más de las proteínas estructurales del alfavirus.

Adicionalmente, y como se ha mencionado anteriormente, los homólogos de enzimas útiles para generar metabolitos son contemplados entre los microorganismos y métodos proporcionados en la presente memoria. El término “homólogos” usado en relación a una enzima o gen original de una primera familia o especie se refiere a distintas enzimas o distintos genes de una segunda familia o especie que son determinados a través de análisis funcional, estructural o genómico para ser una enzima o gen de la segunda familia o especie que corresponde a la enzima o gen originales de la primera familia o especie. Lo más habitual es que los homólogos tengan similitudes funcionales, estructurales o genómicas. Se conocen las técnicas mediante las cuales los homólogos de una enzima o gen pueden ser clonados fácilmente usando sondas genéticas y PCR. La identidad de las secuencias clonadas como homólogas se puede confirmar usando ensayos funcionales y/o mediante mapeado genético de los genes.

Una proteína tiene “homología” o es “homóloga” respecto a una segunda proteína si la secuencia de ácido nucleico que codifica la proteína tiene una secuencia similar a la secuencia de ácido nucleico que codifica la segunda proteína. Alternativamente, una proteína tiene homología respecto a una segunda proteína si las dos proteínas tienen secuencias de aminoácido “similares”. (Por tanto, el término “proteínas homólogas” se define con el significado de que las dos proteínas tienen secuencias de aminoácido similares).

Tal como se usa en la presente memoria, dos proteínas (o una región de las proteínas) son sustancialmente homólogas cuando las secuencias de aminoácido tienen una identidad de al menos 30%, 40%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99%. Para determinar el porcentaje de identidad de dos secuencias de aminoácido, o de dos secuencias de ácido nucleico, las secuencias son alineadas con el objetivo de una comparación óptima (p.ej., se pueden introducir huecos en una o en ambas de una primera y una segunda secuencia de aminoácidos o de ácido nucleico para un alineamiento óptimo y las secuencias no homólogas se pueden descartar para los fines de comparación). La longitud de una secuencia de referencia alineada con fines comparativos puede ser de al menos el 30%, típicamente al menos el 40%, más típicamente al menos el 50%, incluso más típicamente al menos el 60%, e incluso más típicamente al menos el 70%, 80%, 90%, 100% de la longitud de la secuencia de referencia. Entonces se comparan los residuos de aminoácido o los nucleótidos de las correspondientes posiciones de aminoácido y de nucleótido. Cuando una posición de la primera secuencia está ocupada por el mismo residuo de aminoácido o nucleótido que la posición correspondiente de la segunda secuencia, entonces las moléculas son idénticas para dicha posición (tal como se usa en la presente memoria la “identidad” de aminoácido o de ácido nucleico es equivalente a “homología” de aminoácido o ácido nucleico). El porcentaje de identidad entre las dos secuencias es función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos, y la longitud de cada hueco, que es necesario introducir para un alineamiento óptimo de las dos secuencias.

Cuando se usa “homólogo” en referencia a proteínas o péptidos, se reconoce que las posiciones de residuo que no son idénticas a menudo difieren en sustituciones de aminoácidos conservativas. Una “sustitución de aminoácido conservativa” es aquella en la que un residuo de aminoácido es sustituido por otro residuo de aminoácido que tiene una cadena lateral (grupo R) con propiedades químicas similares (p.ej., carga o hidrofobicidad). En general, una sustitución de aminoácido conservativa no cambiará sustancialmente las propiedades funcionales de una proteína. En los casos en que dos o más secuencias de aminoácido difieren entre ellas en sustituciones conservativas, el porcentaje de identidad de secuencia o el grado de homología puede ajustarse al alza para incluir la naturaleza conservativa de la sustitución. Los medios para realizar dicho ajuste son bien conocidos por los especialistas en la técnica (véase, p.ej., Pearson et al., 1994).

Una “sustitución de aminoácido conservativa” es aquella en la que el residuo de aminoácido es reemplazado por un residuo de aminoácido que tiene una cadena lateral similar. Las familias de residuos de aminoácido que tienen cadenas laterales similares han sido definidas en la técnica. Dichas familias incluyen aminoácidos con cadenas laterales básicas (p.ej., lisina, arginina, histidina), cadenas laterales ácidas (p.ej., ácido aspártico, ácido glutámico), cadenas laterales polares no cargadas (p.ej., glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (p.ej., alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales beta-ramificadas (p.ej., treonina, valina, isoleucina) y cadenas laterales aromáticas (p.ej., tirosina, fenilalanina, triptófano, histidina). Los siguientes seis grupos contienen todos aminoácidos que son sustituciones conservativas entre ellos: 1) Serina (S), Treonina (T); 2) Ácido Aspártico (D), Ácido Glutámico (E); 3) Asparagina (N), Glutamina (Q); 4) Arginina (R), Lisina (K); 5) Isoleucina (I), Leucina (L), Metionina (M), Alanina (A), Valina (V); y 6) Fenilalanina (F), Tirosina (Y), Triptófano (W).

La homología de secuencia para polipéptidos, que también se puede referir como porcentaje de identidad de secuencia, se mide típicamente usando un software de análisis de secuencia. Véase, p.ej., el “Sequence Analysis Software Package” del “Genetics Computer Group (GCG)”, Centro de Biotecnología de la Universidad de Wisconsin, 910 University Avenue, Madison, Wis. 53705. El software de análisis de proteínas coteja secuencias similares usando la medida de la homología asignada a las diversas sustituciones, eliminaciones y otras modificaciones, incluyendo las sustituciones de aminoácidos conservativas. Por ejemplo, GCG contiene programas tales como “Gap” y “Bestfit” que

pueden usarse con parámetros por defecto para determinar la homología de secuencia o la identidad de secuencia entre polipéptidos estrechamente relacionados, tal como los polipéptidos homólogos de diferentes especies de organismos, o entre una proteína natural y una muteína de la misma. Véase, p.ej., GCG Versión 6.1.

Un algoritmo típico usado para comparar una secuencia molecular con una base de datos que contiene un gran número de secuencias de diferentes organismos es el programa de ordenador BLAST (Altschul, 1990; Gish, 1993; Madden, 1996; Altschul, 1997; Zhang, 1997), especialmente blastp o tblastn (Altschul, 1997). Los parámetros típicos para BLASTp son: valor de expectación: 10 (por defecto); Filtro: seg (por defecto); Coste para abrir un hueco: 11 (por defecto); Coste para extender un hueco: 1 (por defecto); Alineaciones máximas: 100 (por defecto); Tamaño de la palabra: 11 (por defecto); N° de descripciones: 100 (por defecto); Matriz de penalización: BLOWSUM62.

Cuando se hace una búsqueda en una base de datos que contiene secuencias procedentes de un gran número de organismos diferentes, es típico comparar secuencias de aminoácido. Las búsquedas en bases de datos con secuencias de aminoácido se pueden medir mediante algoritmos diferentes a blastp conocido en la técnica. Por ejemplo, se pueden comparar secuencias de polipéptido con FASTA, un programa en GCG Versión 6.1. FASTA proporciona alineamientos y porcentaje de identidad de secuencia de las regiones del mejor solapamiento entre las secuencias de consulta y de búsqueda (Pearson, 1990). Por ejemplo, el porcentaje de identidad de secuencia entre secuencias de aminoácido se puede determinar usando FASTA con sus parámetros por defecto (con un tamaño de palabra de 2 y con la matriz de puntuación PAM250), según se proporciona en GCG Versión 6.1.

Tal como se describe en la presente memoria, las composiciones y métodos de la descripción proporcionan la capacidad de des-diferenciarse a células somáticas para formar células madre (p.ej., inducir la formación de células madre). Las células madre son células capaces de diferenciación en otros tipos de células, que incluyen aquellas que tienen una función especializada particular (p.ej., células específicas de tejido, células parenquimales y progenitoras de las mismas). Existen diversas clases de células madre, que pueden caracterizarse por su capacidad para diferenciarse en un tipo de célula/tejido deseado. Por ejemplo, las "células progenitoras" pueden ser multipotentes o pluripotentes. Las células progenitoras son células que dan lugar a diferentes tipos de células terminalmente diferenciadas, y células que son capaces de dar lugar a diversas células progenitoras. El término "pluripotente" o "pluripotencia" se refiere a células con la capacidad para dar lugar a células de progenie que pueden someterse a diferenciación, en las condiciones apropiadas, en tipos de células que colectivamente demuestran características asociadas con los linajes celulares de todas las tres capas germinales (endoderma, mesoderma y ectoderma). Las células madre pluripotentes pueden contribuir a todos los tejidos derivados embrionarios de un animal prenatal, postnatal o adulto. Un ensayo estándar aceptado en la técnica, tal como la capacidad para formar un teratoma en ratones SCID de 8-12 semanas de edad, se puede usar para establecer la pluripotencia de una población de células; sin embargo, la identificación de diversas características de las células madre pluripotentes también puede usarse para detectar células pluripotentes. "Características de células madre pluripotentes" se refiere a las características de una célula que distinguen a las células madre pluripotentes de otras células. La capacidad para dar lugar a progenie que puede someterse a diferenciación, en las condiciones apropiadas, en tipos de células que colectivamente demuestran características asociadas con los linajes celulares de las tres capas germinales (endoderma, mesoderma y ectoderma) es una característica de célula madre pluripotente. La expresión o no expresión de determinadas combinaciones de marcadores moleculares también son características de células madre pluripotentes. Por ejemplo, las células madre pluripotentes humanas expresan al menos alguno, y en algunas realizaciones todos, los marcadores de la siguiente lista no limitante: SSEA-3, SSEA-4, TRA-1-60, TRA-1-81, TRA-2-49/6E, ALP, Sox2, E-caderina, UTF-1, Oct4, Rex1 y Nanog. Las morfologías celulares asociadas a las células madre pluripotentes son también características de células madre pluripotentes. En comparación, una célula madre pluripotente es capaz de diferenciarse en un subconjunto de células en comparación con una célula madre pluripotente. Por ejemplo, una célula madre multipotente puede ser capaz de someterse a diferenciación en una o dos de las tres capas germinales. Tal como se usa en la presente memoria, "células no pluripotentes" se refiere a células de mamífero que no son células pluripotentes. Los ejemplos de dichas células incluyen células diferenciadas, así como células multipotentes. Los ejemplos de células diferenciadas incluyen, aunque sin limitación, células de un tejido seleccionado entre médula ósea, piel, músculo esquelético, tejido graso y sangre periférica. Los ejemplos de tipos celulares incluyen, aunque sin limitación, fibroblastos, hepatocitos, mioblastos, neuronas, osteoblastos, osteoclastos y células T.

Otra clase de células incluso más primitiva (es decir, no comprometidas para un destino de diferenciación particular) que las células madre pluripotentes son las denominadas células madre "totipotente" (p.ej., ovocitos fertilizados, células de embriones en los estadios de desarrollo de dos y cuatro células), que tienen la capacidad de diferenciarse en cualquier tipo de células de la especie particular. Por ejemplo, una única célula madre totipotente podría dar lugar a un animal completo, así como a cualquiera de la miríada de tipos celulares que se dan en la especie particular (p.ej., humanos).

Las células madre pluripotentes son un tipo de células que se someten a una auto-renovación a la vez que mantienen la capacidad de dar lugar a todos los tejidos derivados de las tres capas germinales y a todos los linajes de células germinales. Aunque las células madre embrionarias humanas (hES) pluripotentes derivadas de blastocitos humanos son una fuente prometedora para las terapias basadas en células para tratar enfermedades y trastornos tales como la enfermedad de Parkinson, el infarto de miocardio, la lesión de médula espinal, y la diabetes mellitus, su potencial clínico está limitado por su inmunogenicidad y por preocupaciones éticas.

El término “célula precursora”, “célula progenitora” y “célula madre” se usan de forma intercambiable en la técnica y en la presente memoria y se refieren a una célula pluripotente, o sin linaje comprometido, progenitora, que es potencialmente capaz de un número ilimitado de divisiones mitóticas para renovar su línea o para producir células de progenie que se diferenciarán en fibroblastos o en una célula progenitora de linaje comprometido y su progenie, que es capaz de auto-renovación y es capaz de diferenciarse en un tipo celular parenquimal. Al contrario que las células madre pluripotentes, las células progenitoras de linaje comprometido generalmente se consideran incapaces de dar lugar a numerosos tipos celulares que difieren unos de otros fenotípicamente. En su lugar, dan lugar a uno o posiblemente dos tipos celulares de linaje comprometido.

La descripción demuestra que células humanas terminalmente diferenciadas (p.ej., fibroblastos dérmicos humanos) pueden ser inducidas a des-diferenciarse usando un sistema de expresión de ARNm ectópico (p.ej., un sistema de replicón). La descripción contempla el uso de una variedad de secuencias codificadoras de des-diferenciación (también denominadas Factores de Reprogramación (RFs)) que comprenden, por ejemplo, un polinucleótido que codifica KLF4, OCT4, SOX2, c-MYC o n-MYC (L-MYC), GLIS1, NANOG o cualquier combinación de los mismos (p.ej., KLF4, OCT4, SOX2, c-MYC o n-MYC (L-Myc) y opcionalmente NANOG). La des-diferenciación se puede lograr poniendo en contacto la célula, *in vivo* o *in vitro*, con uno o más vectores de ARN auto-replicantes que permanecen ectópicos para el genoma de la célula hospedante y codifican factores que inducen la des-diferenciación. El vector de ARN auto-replicante ectópico de la descripción puede ser controlado cultivando una célula hospedante transformada con el vector de ARN auto-replicante en presencia de B18R. Los métodos para promover la des-diferenciación proporcionan métodos para promover la regeneración de células y tejidos de mamífero dañados por lesión o enfermedad. La descripción también proporciona métodos para enriquecer células madre inducidas y poblaciones que comprenden dichas células madre enriquecidas.

La generación de células madre pluripotentes específicas de paciente tiene el potencial de acelerar drásticamente la implementación de células madre en usos clínicos para tratar enfermedades degenerativas. La descripción proporciona métodos para emplear fácilmente células estromales donadas, tal como fibroblastos dérmicos, de un paciente y generar células madre pluripotentes inducidas humanas (hiPS o iPS) mediante expresión ectópica de un conjunto de factores de des-diferenciación que comprende ARN que codifica (i) KLF4, OCT4, SOX2, c-MYC o n-MYC (L-Myc), NANOG o cualquier combinación de los mismos; (ii) KLF4, OCT4, SOX2 y GLIS1; y (iii) KLF4, OCT4, SOX2 y NANOG. Las líneas celulares generadas son fisiológica y morfológicamente indistinguibles de las células madre embrionarias humanas (HESC) generadas a partir de la masa celular interior de un embrión humano. Las células hiPS comparten un perfil de expresión génica casi idéntico a dos líneas de HESC establecidas.

El término “des-diferenciación” es familiar para la persona especialista en el campo de la técnica. En general, des-diferenciación significa la regresión de la célula de linaje comprometido al estatus de una célula madre, por ejemplo, “induciendo” un fenotipo des-diferenciado. Por ejemplo, como se describe adicionalmente en la presente memoria, KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, GLIS1 y/o Nanog pueden inducir la des-diferenciación y la inducción de mitosis en células inhibidas mitóticamente de linaje comprometido.

En una realización, la descripción proporciona un cultivo celular que comprende células somáticas humanas que han sido transformadas con un replicón de la descripción. En una realización, las células somáticas son fibroblastos. Las células somáticas también pueden ser queratinocitos. En otra realización, el replicón comprende una secuencia que es idéntica en un 90%, 95%, 98%, 99% o 100% a la SEQ ID NO:29, 30, 31 o 32 desde aproximadamente la posición 1 hasta aproximadamente la posición 7561 (incluyendo donde “T” de la secuencia puede ser sustituido por “U”), seguido de dos o más RFs seleccionados del grupo que consiste en Oct-3/4, Sox-2, Klf4, c-Myc, Nanog y Glis1, seguido de un 3'UTR y cola poliA de VEE. Las más de una secuencias codificadoras de RF pueden estar separadas por un sitio de entrada de ribosoma interno (IRES) o por un promotor pequeño (p.ej., un núcleo) tal como SP1. El orden de los RFs no es crítico para la descripción; por tanto, el orden puede ser Klf4, Oct-3/4, Sox-2, c-Myc, o puede ser Sox-2, Klf4, Oct-3/4, c-Myc, u Oct4, Klf4, Sox2, c-Myc, o cualquier variación del orden de los RFs. En una realización, el replicón comprende una secuencia que es idéntica al menos aproximadamente en un 95%, 98%, 99% o 100% a una secuencia como la establecida en la SEQ ID NO:29, 30, 31 o 32. En otra realización adicional, las células son cultivadas en medio acondicionado que comprende B18R y/o son co-transformadas con un polinucleótido que codifica B18R.

La descripción también proporciona métodos para fabricar una célula madre a partir de una célula somática, que comprenden la transformación de la célula somática con un replicón de ARN como se describe en la descripción, y cultivar la célula somática en las condiciones que promueven la expresión de secuencias codificadoras en el replicón, y cultivar las células durante un periodo de tiempo suficiente para des-diferenciar las células en células madre. En una realización, las células son sometidas a pasaje al menos 5, 10, 15, 20 o más veces. En otra realización, las células son cultivadas durante al menos 10, 20, 30 o más días. En otra realización adicional, las células son cultivadas en medio acondicionado que comprende B18R o son co-transformadas con un polinucleótido que codifica B18R.

La descripción también proporciona cultivos de células madre inducidas obtenidos mediante los métodos descritos en la presente memoria. Las células madre pueden no contener ninguno de los factores RF en el ADN genómico de la célula. Las células madre pueden no contener ningún ADN o ARN retroviral (p.ej., células madre que están libres de ADN o ARN retroviral).

En una realización, la descripción proporciona células madre inducidas aisladas, individualmente o en poblaciones. El término “aisladas” o “purificadas” cuando se refiere a células madre de la descripción significa células que están sustancialmente libres de células que portan marcadores asociados con la dedicación de linaje. Las células madre pluripotentes inducidas humanas pueden estar libres de dichos tipos celulares contaminantes en al menos un 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 99%. Las células madre aisladas también pueden estar sustancialmente libres de moléculas solubles naturales. Tal como se discute más detalladamente a continuación, una célula madre sustancialmente purificada de la descripción puede obtenerse, por ejemplo, mediante extracción (p.ej., vía centrifugación de gradiente de densidad y/o citometría de flujo) de una fuente de cultivo. La pureza se puede medir mediante cualquier método apropiado. Una célula madre de la descripción puede purificarse al 99%-100%, por ejemplo, mediante citometría de flujo (p.ej., análisis FACS), tal como se discute en la presente memoria. Dichas células iPS purificadas carecerán de ADN o ARN retroviral alguno.

La descripción también proporciona una población enriquecida de células madre inducidas. Una “población enriquecida de células madre inducidas” es aquella en la que las células madre inducidas de la descripción han sido separadas parcialmente de otros tipos de células, de tal modo que la población resultante de células tiene una mayor concentración de células madre inducidas que la población original de células. La población enriquecida de células madre inducidas puede presentar una concentración de células madre inducidas que es aproximadamente 10 veces, 100 veces, 500 veces, 1000 veces, 2000 veces, 3000 veces, 4000 veces, 5000 veces, 6000 veces, 7000 veces, 8000 veces, 9000 veces o 10000 veces o más que la que tenía la población original antes de la separación. Las células madre inducidas de la descripción, por ejemplo, pueden suponer al menos el 5%, 10%, 15%, 20%, 35%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% o más de la población enriquecida de células madre. La población enriquecida de células madre inducidas puede obtenerse, por ejemplo, seleccionando contra células que presentan marcadores asociados a células diferenciadas, u otros tipos celulares no deseados, y/o seleccionando para células que presentan marcadores (p.ej., TRA-1-81 y/o TRA-1-60) asociados a células madre pluripotentes inducidas humanas de la descripción, y/o regenerando células madre aisladas en sistemas de cultivo definidos. Alternativamente, o adicionalmente, el enriquecimiento para la expresión de un marcador, la pérdida de expresión de un marcador también pueden usarse para el enriquecimiento. Dichas células iPS enriquecidas carecerán del ARN o ADN retroviral típicamente usado para transformar células con RFs.

La descripción también proporciona líneas celulares de células madre inducidas. Tal como se usa en la presente memoria, una “línea celular” significa un cultivo de células madre de la descripción, o células de progenie de las mismas, que pueden ser reproducidas durante un periodo de tiempo extendido, preferiblemente de forma indefinida, término que incluye, por ejemplo, células que son cultivadas, criopreservadas y re-cultivadas siguiendo la criopreservación. Tal como se usa en la presente memoria, un “cultivo” significa una población de células madre inducidas crecidas en un medio y opcionalmente sometidas al correspondiente pasaje. Un cultivo de células madre puede ser un cultivo primario (p.ej., un cultivo que no ha sido sometido a pasaje) o puede ser un cultivo secundario o posterior (p.ej., una población de células que ha sido subcultivada o sometida a pasaje una o más veces).

La descripción también proporciona células que son des-diferenciadas en células madre (es decir, células madre inducidas) que comprenden características que incluyen la capacidad de auto-renovación y diferenciación en mesoderma, endoderma y epiderma, donde las células des-diferenciadas pueden ser producidas mediante expresión de uno o más RFs ectópicos para el genoma celular del hospedante usando un vector de ARN replicante. El vector de replicón puede ser derivado de un alfavirus (p.ej., virus de encefalitis equina venezolana).

Los usos terapéuticos de las células madre pluripotentes inducidas humanas de la descripción incluyen el trasplante de las células madre pluripotentes inducidas humanas, de poblaciones de células madre, o de su progenie, en individuos para tratar una variedad de estados patológicos que incluyen enfermedades y trastornos que son resultado de cánceres, neoplasmas, lesiones, infecciones víricas, diabetes, etcétera. Las células madre o las poblaciones de células madre (que incluyen células madre alteradas genéticamente) son introducidas en un sujeto que necesite dichas células madre o su progenie, o que necesite un KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, GLIS1 o cualquier combinación de las mismas, proteína o molécula codificada o producida por la célula genéticamente alterada. Por ejemplo, las células madre pluripotentes inducidas humanas se pueden administrar a pacientes de cáncer que han sido sometidos a una quimioterapia que ha matado, reducido o dañado células madre u otras células de un sujeto, donde las células madre inducidas reemplazan a las células dañadas o muertas. En otro ejemplo, las células madre pluripotentes inducidas humanas pueden ser transfectadas o transformadas (además de los factores de des-diferenciación) con al menos un factor terapéutico adicional. Por ejemplo, una vez que las células madre pluripotentes inducidas humanas de la descripción son aisladas u obtenidas mediante los métodos de la descripción, las células madre pueden ser transformadas con un polinucleótido que codifica un polipéptido terapéutico. Dicho método y composiciones pueden proporcionar biorreactores de células madre para la producción de un polipéptido deseado o pueden usarse para la administración génica o la terapia génica. En este ejemplo, las células iPS pueden ser aisladas, transformadas con un polinucleótido que codifica un polipéptido terapéutico y a continuación pueden ser implantadas o administradas a un sujeto, o pueden ser diferenciadas en un tipo celular deseado e implantadas y administradas al sujeto. En estas condiciones, el polinucleótido es expresado en el interior del sujeto para la administración del polipéptido producido.

Si las células humanas son derivadas de una fuente heteróloga (no autóloga/allogénica) con respecto al sujeto receptor, típicamente se administra una terapia de inmunosupresión concomitante, p.ej., la administración del agente

inmunosupresor ciclosporina o FK506. Sin embargo, debido al estado inmaduro de las células madre pluripotentes inducidas humanas de la descripción, dicha terapia inmunosupresora puede no ser requerida. Por consiguiente, las células madre pluripotentes inducidas humanas de la descripción pueden administrarse a un receptor en ausencia de terapia inmunomoduladora (p.ej., inmunosupresora). Alternativamente, las células pueden encapsularse en una membrana, que permite el intercambio de fluidos pero evita el contacto célula/célula. El trasplante de células microencapsuladas es conocido en la técnica, p.ej., Balladur et al., 1995, Surgery 117: 189-94, 1995; y Dixit et al., 1992, Cell Transplantation 1: 275-79.

Las células pueden introducirse directamente en la sangre periférica o pueden depositarse en otras localizaciones del cuerpo, p.ej., en un tejido deseado, o en partículas microportadoras en el peritoneo. Por ejemplo, se pueden trasplantar de 10^2 a 10^9 células en un único procedimiento, y según se requiera se pueden realizar trasplantes adicionales.

La diferenciación de las células madre pluripotentes inducidas humanas o la des-diferenciación de células de linaje comprometido (mitóticamente inhibidas) puede ser inducida *ex vivo*, o alternativamente puede ser inducida por contacto con tejido *in vivo*, (p.ej., por contacto con fibroblastos o componentes de la matriz celular). Opcionalmente, se puede co-administrar o administrar posteriormente al sujeto un agente diferenciador o un agente de des-diferenciación (p.ej., KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, GLIS1, o cualquier combinación de los mismos, o un agonista de los mismos).

Se ha demostrado previamente que el trasplante de células de isleta beta proporciona una terapia para pacientes con diabetes (Shapiro et al., 2000). Las células madre pluripotentes inducidas humanas de la descripción proporcionan una fuente alternativa de células de isleta para prevenir o tratar la diabetes. Por ejemplo, las células madre pluripotentes inducidas de la descripción se pueden generar, aislar y diferenciar en un tipo celular pancreático y administrarse a un sujeto. Alternativamente, las células madre pluripotentes inducidas pueden administrarse al páncreas del sujeto y diferenciarse en células de isleta *in vivo*. Por consiguiente, las células son útiles para el trasplante con el objetivo de prevenir o tratar la aparición de diabetes.

La descripción contempla que se pueden usar los métodos *in vitro* descritos en la presente memoria para el trasplante autólogo de células des-diferenciadas o re-diferenciadas (p.ej., las células son recolectadas y devueltas al mismo individuo). La descripción contempla además que los métodos *in vitro* descritos en la presente memoria puedan usarse para trasplantes no autólogos. En un ejemplo, el trasplante se produce entre un donante y un receptor relacionados genéticamente. En otro ejemplo, el trasplante se produce entre un donante y un receptor no relacionados genéticamente. En cualquiera de los ejemplos anteriores, la descripción contempla que las células des-diferenciadas pueden expandirse en cultivo y almacenarse para una posterior recuperación y uso. De forma similar, la descripción contempla que las células re-diferenciadas pueden expandirse en cultivo y almacenarse para una posterior recuperación y uso.

Las composiciones y métodos de la descripción pueden aplicarse a un procedimiento en el que células diferenciadas (de linaje comprometido) son retiradas de un sujeto, des-diferenciadas en cultivo, y a continuación reintroducidas en dicho individuo o, estando aún en cultivo, ser manipuladas para re-diferenciarse siguiendo rutas de diferenciación específicas (p.ej., células pancreáticas, células neuronales, células hepáticas, células cutáneas, células cardiovasculares, células gastrointestinales, etc.). Dichas células re-diferenciadas pueden introducirse entonces en el individuo. Por ejemplo, se pueden extraer fibroblastos diferenciados, des-diferenciarlos (p.ej., mediante expresión ectópica de un replicón de la descripción que comprende KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, GLIS1, NANOG o cualquier combinación de los mismos) y expandirlos mitóticamente y a continuación re-diferenciarlos (p.ej., con un antagonista de KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, GLIS1, o cualquier combinación de los mismos) o con factores (que incluyen estímulos físicos) que se sepa que producen la diferenciación de hESCs hasta una ruta de linaje comprometido. El método puede comprender la extracción de células diferenciadas de un sujeto lesionado o enfermo. Las células des-diferenciadas procedentes de células tomadas de un sujeto lesionado pueden devolverse posteriormente al sujeto lesionado o enfermo para tratar una lesión o una enfermedad degenerativa. Las células des-diferenciadas pueden reintroducirse en el sitio de la lesión, o las células pueden reintroducirse en un sitio distante de la lesión. De forma similar, las células se pueden recolectar a partir de un sujeto lesionado, des-diferenciarse *in vitro*, re-diferenciarse *in vitro*, y trasplantarse de vuelta al sujeto para tratar una lesión o una enfermedad degenerativa.

Las células madre pluripotentes inducidas humanas de la descripción pueden ser aisladas de una muestra obtenida de un sujeto mamífero. El sujeto puede ser cualquier mamífero (p.ej., bovino, ovino, porcino, canino, felino, equino, primate), que incluye un humano. La muestra de células puede obtenerse de cualquiera de una serie de fuentes diferentes que incluyen, por ejemplo, médula ósea, tejido fetal (p.ej., tejido hepático fetal), sangre periférica, sangre del cordón umbilical, páncreas, etc.

La descripción también proporciona métodos para establecer y/o mantener poblaciones de células madre, o su progenie, así como poblaciones mixtas que comprenden tanto células madre como células de progenie, y las poblaciones de células así producidas. Como con las células madre pluripotentes inducidas humanas de la descripción, una vez que un cultivo de células o un cultivo mixto de células madre es establecido, la población de células es expandida mitóticamente *in vitro* mediante pasaje a medio fresco, puesto que la densidad celular dicta las condiciones que conducen a la proliferación celular, con o sin formación de tejido. Dichos métodos de cultivo pueden

incluir, por ejemplo, el pasaje de células en medio de cultivo que carece de determinados factores de crecimiento que inducen la diferenciación (p.ej., IGF, EGF, FGF, VEGF, y/u otro factor de crecimiento), en presencia de un agente que estimula (p.ej., un agonista) KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, GLIS1 o cualquier combinación de los mismos, en presencia de KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, Glis1 o cualquier combinación de los mismos, o cualquier combinación de los anteriores. Los cultivos que comprenden fibroblastos y células de tipo fibroblasto que comprenden células madre y células de fibroblasto pueden ser transferidos a medio fresco cuando se alcanza una densidad celular suficiente. Algunos tipos de células madre no demuestran el típico contacto de inhibición-apoptosis o se vuelven quiescentes cuando la densidad es máxima. Por consiguiente, se pueden usar técnicas de pasaje apropiadas para reducir la inhibición por contacto y la quiescencia, por ejemplo, transfiriendo una porción de las células a un nuevo recipiente de cultivo con medio fresco. Dicha eliminación o transferencia se puede realizar en cualquier recipiente de cultivo.

Una vez que las células madre pluripotentes inducidas humanas de la descripción han sido establecidas en cultivo, como se ha descrito anteriormente, pueden ser mantenidas o almacenadas en "bancos" celulares que comprenden cultivos *in vitro* continuos de células que requieren una transferencia regular o como células que han sido criopreservadas.

La criopreservación de células madre, o de otras células de la descripción, se puede llevar a cabo según métodos conocidos, tales como los descritos en Doyle et al., (eds.), 1995, *Cell & Tissue Culture: Laboratory Procedures*, John Wiley & Sons, Chichester. Por ejemplo, aunque no a modo de limitación, las células pueden ser suspendidas en un "medio de congelación" tal como, por ejemplo, medio de cultivo que además comprende 15-20% de suero fetal bovino (FBS) y 10% de dimetilsulfóxido (DMSO), con o sin 5-10% de glicerol, a una densidad, por ejemplo, de aproximadamente $4 \cdot 10^6$ células/mL. Las células fueron dispensadas en viales de vidrio o plástico, que a continuación fueron sellados y transferidos a una cámara de congelación de un congelador programable o pasivo. La velocidad óptima de congelación puede determinarse empíricamente. Por ejemplo, se puede usar un programa de congelación que proporciona un cambio en la temperatura de $-1^\circ\text{C}/\text{min}$ a lo largo del calor de fusión. Una vez que los viales que contienen las células han alcanzado los -80°C , son transferidos a un área de almacenamiento en nitrógeno líquido. Las células criopreservadas se pueden almacenar durante un periodo de años, aunque deberían ser comprobadas al menos cada 5 años para mantenimiento de la viabilidad.

Las células criopreservadas de la descripción constituyen un banco de células, de las que se pueden retirar porciones mediante descongelación y a continuación usarse para producir un cultivo de células madre que comprende células madre, según se necesite. La descongelación generalmente debería llevarse a cabo rápidamente, por ejemplo, transfiriendo un vial desde el nitrógeno líquido hasta un baño de agua a 37°C . El contenido descongelado del vial debería ser transferido inmediatamente en condiciones estériles a un recipiente de cultivo que contenga el medio apropiado. Es aconsejable que las células del medio de cultivo sean ajustadas a una densidad inicial de aproximadamente $1 \cdot 10^5$ células/mL. Una vez en el cultivo, las células se pueden examinar diariamente, por ejemplo, con un microscopio invertido, para detectar la proliferación celular, y se pueden subcultivar tan pronto como alcancen una densidad apropiada.

Las células madre pluripotentes inducidas humanas de la descripción se pueden retirar de un banco celular según se necesiten, y usarse para la producción de nuevas células madre, tanto *in vitro*, por ejemplo, como un cultivo de tejido tridimensional, como se describe más adelante, como *in vivo*, por ejemplo, mediante la administración directa de células al sitio en el que se necesitan los fibroblastos o el tejido. Tal como se describe en la presente memoria, las células madre pluripotentes inducidas humanas de la descripción se pueden usar para producir nuevo tejido para uso en un sujeto, donde las células han sido aisladas originalmente de la propia sangre de dicho sujeto, o de otro tejido (es decir, células autólogas). Alternativamente, las células de la descripción se pueden usar como células donantes ubicuas para producir nuevo tejido para uso en cualquier sujeto (es decir, células heterólogas).

Una vez establecido, un cultivo de células madre puede usarse para producir células de progenie y/o fibroblastos capaces de producir nuevo tejido. La diferenciación de células madre en fibroblastos u otros tipos de células, seguida de la producción de tejido a partir de las mismas, se puede activar mediante factores de crecimiento exógenos específicos, o cambiando las condiciones de cultivo (p.ej., la densidad) de un cultivo de células madre. Puesto que las células son pluripotentes, pueden usarse para reconstituir un sujeto irradiado y/o un sujeto tratado con quimioterapia; o como fuente de células para linajes específicos, facilitando su maduración, proliferación y diferenciación en uno o más linajes seleccionados. Los ejemplos de factores que pueden usarse para inducir la diferenciación incluyen eritropoyetina, factores estimulantes de colonia, p.ej., GM-CSF, G-CSF o M-CSF, interleucinas, p.ej., IL-1, -2, -3, -4, -5, -6, -7, -8, etcétera, factor inhibidor de leucemia (LIF), factor Steel (Stl), o similares, cocultivar con células de tejido comprometido, u otros tipos de células de linaje comprometido para inducir a las células madre para que se conviertan en comprometidas para un linaje particular.

En otro ejemplo, las células madre pluripotentes inducidas humanas pueden ser modificadas genéticamente para expresar genes para tipos específicos de factores de crecimiento para una diferenciación exitosa y/o mejorada en fibroblastos, otras células estromales, o células parenquimales y/o el reemplazo tanto pre- como post-implantación.

Las células de la descripción se pueden usar para tratar sujetos que requieren la reparación o el reemplazo de tejido como consecuencia de una enfermedad o un trauma. El tratamiento puede contemplar el uso de las células de la

descripción para producir nuevo tejido, y el uso del tejido así producido, según cualquier método conocido actualmente en la técnica o para ser desarrollado en el futuro. Por ejemplo, las células inducidas (p.ej., células que comprenden un vector de expresión ectópico que expresa KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, Glis1 o cualquier combinación de los mismos) de la descripción se puede implantar, inyectar o administrar de otro modo directamente al sitio del tejido dañado, de tal modo que producirá *in vivo* nuevo tejido. En un ejemplo, la administración incluye la administración de células madre modificadas genéticamente.

En un ejemplo, una formulación que comprende las células de la descripción se prepara para inyección directamente al sitio en el que se desea la producción de nuevo tejido. Por ejemplo, y no a modo de limitación, las células de la descripción se pueden suspender en una disolución de hidrogel para inyección. Alternativamente, la disolución de hidrogel que contiene las células se puede dejar endurecer, por ejemplo, en un molde para formar una matriz que tenga las células dispersas en ella antes de la implantación. Una vez que la matriz se ha endurecido, las formaciones celulares se pueden cultivar de tal modo que las células se expanden mitóticamente antes del implante. Un hidrogel es un polímero orgánico (natural o sintético) que está reticulado vía enlaces covalentes, iónicos o de hidrógeno para crear una estructura abierta tridimensional, que atrapa moléculas de agua para formar un gel. Los ejemplos de materiales que pueden usarse para formar un hidrogel incluyen polisacáridos tales como alginato y sales del mismo, polifosfazinas y poliácridatos, que son reticulados iónicamente, copolímeros de bloque de óxido de polietileno-polipropileno glicol que son reticulados mediante temperatura o pH, respectivamente. Los métodos de síntesis de los materiales de hidrogel, así como los métodos para preparar dichos hidrogeles, son conocidos en la técnica.

Dichas formulaciones celulares pueden comprender además uno o más componentes adicionales, que incluyen componentes de matriz extracelular seleccionados, tales como uno o más tipos de colágeno conocidos en la técnica, y/o factores de crecimiento y fármacos. Los factores de crecimiento que pueden ser incorporados de forma útil en la formulación celular incluyen uno o más factores de crecimiento de tejido conocidos en la técnica tales como, aunque sin limitación, cualquier miembro de la familia TGF- β , IGF-I y -II, hormona de crecimiento, BMPs tales como BMP-13, y similares. Alternativamente, las células de la descripción pueden ser modificadas genéticamente para expresar y producir factores de crecimiento tales como BMP-13 o TGF- β . También se pueden incluir otros componentes en la formulación, que incluyen, por ejemplo, tampones para proporcionar el pH y la isotonicidad apropiados, lubricantes, materiales viscosos para retener las células en el sitio de administración o en las proximidades (p.ej., alginatos, agaros y gomas vegetales) y otros tipos de células que pueden producir un efecto deseado en el sitio de administración (p.ej., potenciamiento o modificación de la formación de tejido o de sus características fisicoquímicas, soporte para la viabilidad de las células, o inhibición de la inflamación o del rechazo). Las células pueden ser cubiertas con un recubrimiento apropiado para heridas para evitar que las células se alejen del sitio. Dichos recubrimientos para heridas son bien conocidos por los especialistas en la técnica.

Alternativamente, las células madre pluripotentes inducidas humanas de la descripción se pueden sembrar en una estructura tridimensional y cultivarse para permitir que las células se diferencien, crezcan y rellenen la matriz o sean implantadas *in vivo* inmediatamente, donde las células sembradas proliferarán sobre la superficie de la estructura y formarán tejido de reemplazo *in vivo* en cooperación con las células del sujeto. Dicha estructura puede ser implantada en combinación con un factor de crecimiento cualquiera, o más de uno, fármacos, tipos de células adicionales, u otros componentes que estimulan la formación, o potencian o mejoran de algún otro modo la práctica de la descripción.

En otro ejemplo adicional, las células madre pluripotentes inducidas humanas de la descripción pueden usarse en conjunción con un sistema de cultivo tridimensional en un "biorreactor" para producir construcciones de tejido que poseen las propiedades bioquímicas, físicas y estructurales críticas del tejido humano nativo cultivando las células y el tejido resultante en las condiciones ambientales que se experimentan habitualmente por parte del tejido nativo. El biorreactor puede incluir una serie de diseños. Típicamente, las condiciones de cultivo incluirán aplicar un estrés fisiológico sobre la construcción que contiene las células que sea similar al que encontrarán *in vivo*.

Las células madre pluripotentes inducidas humanas, su progenie, y el tejido de la descripción se pueden usar en una variedad de aplicaciones. Éstas incluyen, aunque sin limitación, trasplante o implante de las células en una forma diferenciada, una forma no diferenciada, una forma des-diferenciada. Dichas células y tejidos sirven para reparar, reemplazar o aumentar el tejido que ha sido dañado debido a enfermedad o trauma, o que ha fallado en desarrollarse con normalidad.

Las células madre pluripotentes inducidas humanas y el tejido producido según la descripción puede usarse para reparar o reemplazar tejido dañado o destruido o para aumentar el tejido existente.

Adicionalmente, las células o el tejido de la descripción pueden usarse, por ejemplo, para cribar *in vitro* en términos de eficacia y/o citotoxicidad de los compuestos, alérgenos, factores de crecimiento/reguladores, compuestos farmacéuticos, etcétera, en células madre, para elucidar el mecanismo de determinadas enfermedades determinando cambios en la actividad biológica de las células madre (p.ej., cambios en la expresión o la actividad, la capacidad proliferativa, la adhesión de KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, Glis1 o cualquier combinación de los mismos), para estudiar el mecanismo mediante el cual los fármacos y/o factores de crecimiento operan para modular la actividad biológica de las células madre (p.ej., la expresión o actividad de KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, Glis1 o cualquier combinación de los mismos), para diagnosticar y monitorizar el cáncer en

un paciente, para terapia génica, administración de genes o administración de proteínas; y para producir productos biológicamente activos.

Las células madre pluripotentes inducidas humanas también se pueden usar en el aislamiento y evaluación de factores asociados con la diferenciación y maduración de células madre. De esta manera, las células madre pluripotentes inducidas humanas se pueden usar en ensayos para determinar la actividad de los medios, la implicación con dedicación de linajes particulares, o similares. Son aplicables varios sistemas y se pueden diseñar para inducir la diferenciación de las células madre pluripotentes inducidas humanas empleando diversos tipos de estrés fisiológico.

Las células madre humanas pluripotentes inducidas humanas, su progenie, y los tejidos derivados de las mismas de la descripción se pueden usar *in vitro* para cribar una amplia variedad de agentes en términos de su efectividad y citotoxicidad de agentes farmacéuticos, factores de crecimiento/reguladores, agentes anti-inflamatorios, y similares. Para este fin, las células o cultivos de tejido de la descripción se pueden mantener *in vitro* y ser expuestos al agente que va a ser evaluado. La actividad de un agente citotóxico se puede medir por su capacidad de dañar o matar células madre o su progenie en cultivo. Esto se puede determinar fácilmente mediante técnicas de tinción. El efecto de los factores de crecimiento/reguladores se puede determinar analizando el número de células vivas *in vitro*, p.ej., mediante recuentos celulares totales, y mediante recuentos celulares diferenciales. Esto se puede llevar a cabo usando técnicas citológicas y/o histológicas estándar, que incluyen el uso de técnicas inmunocitoquímicas que emplean anticuerpos que definen antígenos celulares específicos de tipo. El efecto de diversos fármacos sobre las células de la descripción se puede determinar en un cultivo en suspensión o en un sistema tridimensional. En un ejemplo se puede analizar el efecto de un agente de ensayo sobre las células madre pluripotentes inducidas humanas de la descripción.

Las células madre que expresan un producto génico de interés, o tejido producido *in vitro* a partir de las mismas, se pueden implantar en un sujeto que sea deficiente en dicho producto génico. Por ejemplo, son de interés particular los genes que expresan productos capaces de prevenir o aliviar síntomas de diversos tipos de enfermedades o trastornos vasculares, o que previenen o promueven trastornos inflamatorios. Por ejemplo, las células de la descripción pueden ser modificadas genéticamente para expresar un producto génico anti-inflamatorio que serviría para reducir el riesgo de fallo de implantación u otro cambio degenerativo en el tejido debido a una reacción de inflamación. Por ejemplo, una célula madre de la descripción puede ser modificada genéticamente para expresar uno o más productos génicos anti-inflamatorios que incluyen, por ejemplo, péptidos o polipéptidos que corresponden al idiotipo de anticuerpos que neutralizan factor estimulante de colonia de granulocitos-macrófagos (GM-CSF), TNF, IL-1, IL-2, u otras citosinas inflamatorias. Se ha demostrado que la IL-1 reduce la síntesis de proteoglicanos y colágenos de tipo II, IX y XI (Tyler et al., 1985, Biochem. J. 227: 69-878; Tyler et al., 1988, Coll. Relat. Res. 82: 393-405; Goldring et al., 1988, J. Clin. Invest. 82: 2026-2037; y Lefebvre et al., 1990, Biophys. Acta. 1052: 366-72). El TNF también inhibe la síntesis de proteoglicanos y de colágeno de tipo II, aunque es mucho menos potente que la IL-1 (Yaron, I., et al., 1989, Arthritis Rheum. 32: 173-80; Ikebe, T., et al., 1988, J. Immunol. 140: 827-31; y Saklatvala, J., 1986, Nature 322: 547-49). Asimismo, por ejemplo, las células de la descripción pueden modificarse para expresar el gen que codifica la proteína reguladora de complemento humana que previene el rechazo de un injerto por parte del hospedante. Véase, por ejemplo, McCurry et al., 1995, Nature Medicine 1: 423-27. En otra realización, las células madre pluripotentes inducidas humanas pueden modificarse para incluir un gen o secuencia de polinucleótidos que expresa o hace que sea expresado un factor angiogénico.

Las células madre inducidas de la descripción expresan uno o más marcadores asociados al fenotipo de célula madre pluripotente humana y/o carecen de uno o más marcadores asociados a una célula diferenciada (p.ej., una célula que tiene una capacidad reducida para la auto-renovación, la regeneración o la diferenciación) y/o una célula de origen neuronal. Una molécula es un "marcador" de un tipo celular deseado si aparece en un porcentaje suficientemente elevado de células del tipo celular deseado, y si aparece en un porcentaje suficientemente bajo de células de un tipo celular no deseado. Se puede obtener un nivel deseado de purificación del tipo celular deseado a partir de una población de células que comprenden tipos celulares tanto deseados como no deseados seleccionando las células de la población de células que tienen el marcador. Un marcador puede presentarse, por ejemplo, en el 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% o más del tipo celular deseado, y puede presentarse en menos del 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, 5%, 1% o menos de un tipo celular no deseado.

Tal como se ha discutido anteriormente, las células madre inducidas de la descripción o las células madre inducidas que han sido diferenciadas se caracterizan por la presencia y/o la ausencia de determinados marcadores que son reconocidos específicamente por una molécula. Por consiguiente, la descripción proporciona métodos para marcar células madre inducidas de la descripción. En un ejemplo, las células madre pluripotentes inducidas humanas pueden ser marcadas con una molécula (p.ej., un anticuerpo) que reconoce específicamente un marcador que está asociado a una célula madre inducida de la descripción. En otro ejemplo, una población de células se puede poner en contacto con una molécula que se une específicamente a un marcador (p.ej., TRA-1-81) en las condiciones que permitan a la molécula unirse al marcador, donde la población de células comprende al menos una célula madre que tiene dicho marcador. En otro ejemplo, una población de células se puede poner en contacto con una molécula que se une específicamente a un marcador en las condiciones que permiten que la molécula se una al marcador, donde la población de células comprende células madre que no tienen el marcador y células no madre que tienen el marcador. La molécula usada puede ser, por ejemplo, un anticuerpo, un derivado de anticuerpo, o un ligando. La molécula opcionalmente puede comprender un resto adicional, por ejemplo, uno que es detectable (p.ej., una marca fluorescente

o colorimétrica) o uno que ayuda al aislamiento de las células marcadas (p.ej., un resto que sea ligado por otra molécula o una partícula magnética).

En una realización, la población de células somáticas transformadas es sometida a una tinción viva para un Antígeno de Rechazo de Tumor 1-61 y 1-81 (TRA-1-60, TRA-1-81). TRA-1-60 y TRA-1-81 se pueden obtener comercialmente, por ejemplo, en Chemicon International, Inc. (Temecula, Calif., EE.UU.). La detección inmunológica de estos antígenos usando anticuerpos monoclonales ha sido usada para caracterizar células madre pluripotentes en combinación con otros marcadores (Shamblott M. J. et al. (1998) PNAS 95: 13726-13731; Schuldiner M. et al. (2000). PNAS 97: 11307-11312; Thomson J. A. et al. (1998). Science 282: 1145-1147; Reubinoff B. E. et al. (2000). Nature Biotechnology 18: 399-404; Henderson J. K. et al. (2002). Stem Cells 20: 329-337; Pera M. et al. (2000). J. Cell Science 113: 5-10). En una realización, una población de células somáticas que han sido transformadas con al menos un vector de ARN ectópico que comprende un KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, y opcional o alternativamente NANOG o Glis1 son enriquecidas en células que comprenden la expresión de TRA-1-81 o TRA-1-60. En una realización adicional, las células también pueden enriquecerse para la pérdida de un marcador detectable asociado a un vector retroviral.

La descripción también proporciona métodos para aislar células madre inducidas de la descripción. Las células madre pluripotentes inducidas humanas de la descripción se pueden aislar, por ejemplo, utilizando moléculas (p.ej., anticuerpos, derivados de anticuerpos, ligandos o moléculas de fusión de péptido de Fc) que se unen a un marcador (p.ej., un TRA-1-81, un TRA-1-60 o una combinación de marcadores) en las células madre pluripotentes inducidas humanas, y seleccionando de este modo positivamente células que se unen a la molécula (es decir, una selección positiva). Otros ejemplos de métodos de selección positiva incluyen métodos para promover preferentemente el crecimiento de un tipo celular deseado en una población mixta de tipos celulares deseados y no deseados. Alternativamente, usando moléculas que se unen a marcadores que no están presentes en el tipo celular deseado, pero que están presentes en un tipo celular no deseado, las células no deseadas que contienen dichos marcadores pueden ser eliminadas de las células deseadas (es decir, una selección negativa). Otros métodos de selección negativa incluyen matar o inhibir el crecimiento de forma preferente de un tipo celular deseado de una población mixta de tipos celulares deseados y no deseados. Por consiguiente, usando una selección negativa, una selección positiva, o una combinación de ambas, se puede preparar una población enriquecida de células madre.

Los procedimientos de separación pueden incluir separación magnética, usando partículas magnéticas recubiertas de anticuerpos, cromatografía de afinidad, agentes citotóxicos unidos a un anticuerpo monoclonal, o agentes como los usados en conjunción con un anticuerpo monoclonal, p.ej., complemento y citotoxinas, y "cribar" con anticuerpo unido a una matriz sólida (p.ej., una placa), u otra técnica conveniente. Las técnicas que proporcionan una separación precisa incluyen los clasificadores celulares activados por fluorescencia, que pueden presentar diversos niveles de sofisticación, p.ej., una pluralidad de canales de color, canales de detección de bajo ángulo o de dispersión de luz obtusa, y canales de impedancia. De forma conveniente, los anticuerpos se pueden conjugar con marcadores, tales como partículas magnéticas, lo que permite una separación directa, biotina, que puede eliminarse con avidina o estreptavidina ligada a un soporte, fluorocromos, que pueden usarse con un clasificador celular activado por fluorescencia, o similares, para permitir una separación fácil del tipo celular particular. Se puede emplear cualquier técnica que no sea indebidamente perjudicial para la viabilidad de las células madre pluripotentes inducidas humanas. En una realización, las células madre son incubadas con un anticuerpo contra un marcador (p.ej., un anticuerpo TRA-1-81) y las células de tinción positiva para el marcador son seleccionadas manualmente y subcultivadas.

Se pueden usar combinaciones de métodos de enriquecimiento para mejorar el tiempo o la eficacia de la purificación o enriquecimiento. Por ejemplo, tras una etapa de enriquecimiento para eliminar células que tienen marcadores que no son indicativos del tipo celular de interés, las células pueden ser separadas o enriquecidas adicionalmente mediante un clasificador celular activado por fluorescencia (FACS) u otra metodología que tenga una elevada especificidad. Con un FACS se puede emplear análisis multi-color. Las células se pueden separar en base al nivel de tinción para un antígeno particular, o la carencia del mismo. Se pueden usar fluorocromos para marcar anticuerpos específicos para un antígeno particular. Dichos fluorocromos incluyen ficobiliproteínas, p.ej., ficoeritrina y aloficocianinas, fluoresceína, rojo Texas, y similares.

Se puede usar cualquier marcador específico de tipo para seleccionar para o contra un tipo celular particular. Los marcadores de células madre inducidas útiles para el enriquecimiento comprenden marcadores expresados tales como TRA-1-81 y pérdida de marcadores (p.ej., GFP) asociada a un vector retroviral u otro vector exógeno.

Una vez aisladas las células madre, opcionalmente se pueden propagar en medio apropiado en presencia o ausencia de una capa cebadora. Adicionalmente, las células madre pluripotentes inducidas humanas de la invención se pueden cultivar en un sistema de biorreactor.

Una vez que las células madre pluripotentes inducidas humanas de la descripción han sido establecidas en cultivo, como se ha descrito anteriormente, pueden ser mantenidas o almacenadas en "bancos" celulares que comprenden cultivos in vitro de células que requieren una transferencia regular o como células que han sido criopreservadas. En algunas realizaciones, las células del banco son usadas para el tratamiento autólogo de un sujeto.

Se puede aislar fácilmente fibroblastos desintegrando un órgano o tejido apropiado, que va a actuar como fuente de fibroblastos. Esto se puede realizar fácilmente usando técnicas conocidas por los especialistas en el arte de la técnica. Por ejemplo, el tejido o el órgano pueden ser desintegrados mecánicamente y/o pueden ser tratados con enzimas digestivas y/o agentes quelantes que debilitan las conexiones entre células colindantes, haciendo posible dispersar el tejido en una suspensión de células individuales sin que se produzca una ruptura celular apreciable. La disociación enzimática se puede conseguir moliendo el tejido y tratando el tejido molido con uno de una cualquiera de una serie de enzimas digestivas, tanto solas como en combinación. Éstas incluyen, aunque sin limitación, tripsina, quimotripsina, colagenasa, elastasa y/o hialuronidasa, ADNasa, pronasa, dispasa, etc. La ruptura mecánica también se puede conseguir mediante una serie de métodos, que incluyen, aunque sin limitación, el uso de molinillos, mezcladores, tamices, homogeneizadores, celdas de presión o insonadores, por nombrar unos pocos. Para una revisión de las técnicas de disgregación de tejido, véase Freshney, *Culture of Animal Cells. A Manual of Basic Techniques*, 2nd Ed., A.R. Liss, Inc., New York, 1987, Ch. 9, pp. 107-126.

Una vez que el tejido ha sido reducido a una suspensión de células individuales, la suspensión se puede fraccionar en subpoblaciones, a partir de la cual se pueden obtener fibroblastos y/u otras células estromales y/o elementos. Esto también se puede llevar a cabo usando técnicas estándar para la separación celular que incluyen, aunque sin limitación, la clonación y selección de tipos celulares específicos, la destrucción selectiva de células no deseadas (selección negativa), la separación basada en la capacidad de aglutinación celular diferencial en la población mixta, procedimientos de congelación-descongelación, propiedades de adherencia diferencial de las células de la población mixta, filtración, centrifugación convencional y zonal, elutriación centrífuga (centrifugación en contracorriente), separación gravitatoria unitaria, distribución contracorriente, electroforesis y clasificación celular activada por fluorescencia. Para una revisión de técnicas de selección clonal y separación celular, véase Freshney, *Culture of Animal Cells. A Manual of Basic Techniques*, 2nd Ed., A.R. Liss, Inc., New York, 1987, Ch. 11 y 12, pp. 137-168.

El aislamiento de fibroblastos puede, por ejemplo, llevarse a cabo se indica a continuación: muestras de tejido fresco son lavadas intensamente y molidas en disolución salina equilibrada de Hanks (HBSS) a fin de eliminar el suero. El tejido molido es incubado durante 1-12 horas en una disolución recién preparada de una enzima disociadora tal como tripsina. Después de esta incubación, las células disociadas son suspendidas, peletizadas por centrifugación y llevadas a discos de cultivo. Todos los fibroblastos se unirán antes que las otras células, por lo tanto, se pueden aislar y crecer las células estromales apropiadas.

Cuando las células des-diferenciadas van a ser usadas para trasplante o implante *in vivo* es útil obtener las células estromales a partir de tejidos del propio paciente.

Se pueden usar sondas y cebadores de oligonucleótido para identificar la expresión de varios factores descritos en la presente memoria, así como en procedimientos de clonación y amplificación. Una sonda de oligonucleótido o un cebador se refieren a una molécula de ácido nucleico de entre 8 y 2000 nucleótidos de longitud. Más particularmente, la longitud de dichos oligonucleótidos puede oscilar entre aproximadamente 8, 10, 15, 20 o de 30 a 100 nucleótidos, pero típicamente será de aproximadamente 10 a 50 (p.ej., de 15 a 30 nucleótidos). La longitud apropiada para los oligonucleótidos en los ensayos de la descripción bajo un conjunto particular de condiciones puede determinarse empíricamente por parte del especialista en la técnica.

Los cebadores y sondas de oligonucleótido se pueden preparar mediante cualquier método adecuado, que incluye, por ejemplo, la clonación y la restricción de secuencias apropiadas y la síntesis química directa basada en las secuencias conocidas de polinucleótido y polipéptido KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG o cualquier combinación de los mismos. En la técnica se conocen varios ortólogos de otras especies.

Las sondas y cebadores de oligonucleótidos pueden comprender análogos de ácido nucleico tales como, por ejemplo, ácidos nucleicos de péptido, análogos de ácido nucleico bloqueado (LNA), y análogos de morfolino. El extremo 3' de la sonda se puede funcionalizar con una captura o marca detectable para ayudar a la detección de un ácido nucleico KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, Glis1 o cualquier combinación de los mismos.

Cualquiera de los oligonucleótidos o ácidos nucleicos de la descripción se pueden marcar incorporando una etiqueta detectable medible con medios espectroscópicos, fotoquímicos, bioquímicos, inorgánicos o químicos. Por ejemplo, dichas marcas pueden comprender sustancias radiactivas (^{32}P , ^{35}S , ^3H , ^{125}I), colorantes fluorescentes (5-bromodesoxiuridina, fluoresceína, acetilaminofluoreno, digoxigenina), biotina, nanopartículas, y similares. Dichos oligonucleótidos típicamente son marcados en sus extremos 3' y 5'.

Los cebadores y sondas de oligonucleótidos se pueden inmovilizar sobre un soporte sólido. Los especialistas en la técnica conocen los soportes sólidos, que incluyen las paredes de los pocillos de la bandeja de reacción, tubos de ensayo, partículas de poliestireno, partículas magnéticas, tiras de nitrocelulosa, membranas, micropartículas tales como partículas de látex, vidrio y similares. El soporte sólido no es crítico y puede ser seleccionado por el especialista en la técnica. De esta manera, las partículas de látex, las micropartículas, las partículas magnéticas o no magnéticas, las membranas, los tubos de plástico, las paredes de pocillos de microtitulación, el vidrio o las partículas de sílice, y similares, son todos ejemplos adecuados. Los métodos adecuados para inmovilizar oligonucleótidos sobre una fase sólida incluyen interacciones iónicas, hidrofóbicas, covalentes, etcétera. El soporte sólido se puede elegir por su capacidad intrínseca para atraer e inmovilizar el reactivo de captura. Las sondas o cebadores de oligonucleótidos se

pueden unir o inmovilizarse sobre un soporte sólido de forma individual o en grupos de aproximadamente 2-10.000 oligonucleótidos distintos de la descripción en un soporte sólido individual. Se puede usar un sustrato que comprende una pluralidad de cebadores o sondas de oligonucleótido de la descripción para detectar o amplificar KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG, Glis1 o cualquier combinación de los mismos. Por ejemplo, las sondas de oligonucleótidos se pueden usar en una tira de oligonucleótidos tal como las comercializadas por Affymetrix y descritas en la Patente de EE.UU. n° 5.143.854; en las publicaciones PCT WO 90/15070 y 92/10092. Estos sistemas se pueden producir empleando métodos de síntesis mecánica o métodos de síntesis dirigidos por luz, que incorporan una combinación de métodos fotolitográficos y síntesis de oligonucleótidos en fase sólida. La descripción contempla además anticuerpos capaces de unirse específicamente a un polipéptido KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG o Glis1.

Una población de referencia o de control se refiere a un grupo de sujetos o individuos que se predice que son representativos de la población general. Una muestra de ensayo se mide para determinar la cantidad de KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG o Glis1 o de cualquier combinación de los mismos, en la muestra, donde la cantidad se compara con una muestra de control.

En otro ejemplo, la descripción proporciona métodos para diferencia células madre para un linaje comprometido que comprenden inhibir la expresión o la actividad de KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, NANOG o Glis1, o de cualquier combinación de los mismos. Los agentes de diferenciación útiles a este respecto incluyen, por ejemplo, anticuerpos, oligonucleótidos antisentido, construcciones de ARNi, o ribozimas.

Las técnicas de cultivo útiles en los métodos de la descripción se describen en la Publicación de Patente Internacional n° WO 2010/120875.

Los siguientes Ejemplos se proporcionan para ilustrar determinados aspectos de la descripción y para asistir a los especialistas en la técnica a la hora de llevar a la práctica la descripción. Estos Ejemplos no deben considerarse en modo alguno como limitativos del alcance de la descripción de ninguna manera.

Ejemplos

Ejemplo de referencia 1

Células. Se obtuvieron de ATCC fibroblastos BJ de prepucio y la línea celular STO. Los fibroblastos de prepucio humano (HFF) humanos y la línea celular ES humana HUES-9 fueron obtenidos de fuentes existentes. BJ, HFFs y STO fueron cultivadas en DMEM que contenía un 10% de FBS, MEM aminoácidos no esenciales (NEAA), piruvato, penicilina y estreptomycin. Las células HUES-9 e iPS fueron cultivadas con medio de cultivo ES en D-MEM de bloqueo que contenía un 20% de SR de bloqueo, GlutaMAX, NEAA, 2-mercaptoetanol (todos de Invitrogen), penicilina, estreptomycin y bFGF (10 ng/mL). Se prepararon células cebadoras STO mediante tratamiento con mitomicina C (10 µg/mL, Sigma). Para el cultivo libre de cebador de clones de células iPS y HUES-9, las células fueron sometidas a pasaje en pocillos recubiertos de Matrigel™ (BD Bioscience) y se cultivaron en el medio acondicionado preparado a partir de células cebadoras STO con medio de cultivo ES.

Construcción de plásmidos: los ADNcs que codifican para OCT4 (n° de acceso NM_002701), c-MYC (n° de acceso NM_002467) y GLIS1 (n° de acceso BC104911) fueron obtenidos de Open Biosystems. SOX2 (n° de acceso NM_003106), KLF4 (n° de acceso NM_004235), NANOG (n° de acceso BC099704) están disponibles en la ATCC. El B18R (n° de acceso D01019) se obtuvo de Addgene. Los ADNcs se usaron como plantillas para la amplificación PCR para añadir sitios de enzima de restricción y/o secuencia Kozak, y se clonó en vector pBluescript SK+ para comprobar secuencias de ADNc. A continuación los ADNcs fueron clonados en vector pTNF (Promega) para la síntesis de ARNm y pCX4bsr1 para la producción de retrovirus. Para la expresión multicistónica empleando secuencias de péptido 2A víricas, se maduraron oligos F2A, oligos T2A y oligos E2A (Tabla 1) y se clonaron en sitios EcoRI/SpeI, SpeI/XbaI y XbaI/NotI de vector pBluescript SK+, respectivamente. Los ADNcs de factores de reprogramación fueron ligados con secuencias de péptido 2A en marco, y a continuación se clonaron en pVEE-S-IRES-Puro. Los pVEE-S-IRES-Puro fueron contruidos a partir de p5'VEE/S/GFP/Pac3 para clonar factores de reprogramación. Resumidamente, los genes GFP/Pac y 3'UTR parcial en p5'VEE/S/GFP/Pac fueron eliminados con digestión de XbaI/MfeI, y a continuación se introdujeron los múltiples sitios de clonación (MCS; NdeI, AscI, BbvCI, ClaI, MfeI, FseI y NotI) (Tabla 1), IRES y gen de resistencia a Puomicina de pCX4puro. Este vector se renombró como pVEE-IRES-Puro por comodidad. Para generar ARN con ARN polimerasa T7, el promotor SP6 (ATTTAGGTGACACTATAG (véase, p.ej., SEQ ID NO:31 desde 16844-16861)) fue reemplazado por promotor T7 (TAATACGACTGACTATAG (véase, p.ej., SEQ ID NO:32 desde 16843-16860)) mediante PCR (Tabla 1) usando el fragmento SacI/BstZ17I del vector VEE como plantilla (el promotor SP6 está localizado adyacente al sitio SacI).

Tabla 1: Cebadores para clonación PCR

F2A-Directo	5'-AATTCACCGGTGTGAAACAGACTTTGATTTTGACCTTCTCAAGTTGG CGGGAGACGTGGAGTCCAACCCAGGGCCAGATCTA (SEQ ID NO:11)	F2A-oligo
F2A-Inverso	5'-CTAGTAGATCTGGGCCCTGGTTGGACTCCACGTCTCCCGCCAACT TGAGAAGGTCAAAATTCAAAGTCTGTTTACACCCGGTG (SEQ ID NO:12)	F2A-oligo
T2A-F	5'-CTAGTGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTGAGG AGAATCCTGGCCCAACAATTGT (SEQ ID NO:13)	T2A-oligo
T2A-R	5'-CTAGACAATTGTGGGCCAGGATTCTCCTCGACGTCACCGCATGTTA GCAGACTTCCTCTGCCCTCA (SEQ ID NO:14)	T2A-oligo
E2A-F	5'-CTAGACAATGTACTAACTACGCTTTGTTGAAACTCGCTGGCGATGTT GAAAGTAACCCCGGTCTGGCGCGCCCGC (SEQ ID NO:15)	E2A-oligo
E2A-R	5'-GGCCGCGGGCGCGCCAGGACCGGGGTTACTTCAACATCGCCAGC GAGTTTCAACAAGCGTAGTTAGTACATTGt (SEQ ID NO:16)	E2A-oligo
VEE-MCS-F1	5'-CTAGCATATGGGCGCGCCCTCAGCATCGATGGCCGGCCTCTAGAGC GGCCGC (SEQ ID NO:17)	MCS-oligo
VEE-MCS-R1	5'-GGCCGCGGGCGCTCTAGAGGCCGGCCATCGATGCTGAGGGCGCGC CCATATG (SEQ ID NO:18)	MCS-oligo
nsP2a-F1	5'-CAGGACGATCTCATTCTCAC (SEQ ID NO:19)	PCR, nsP2
nsP2a-R1	5'-GCTTGCCACTCCTCTATCGTG (SEQ ID NO:20)	PCR, nsP2
nsP4a-F1	5'-CCACAATACGATCGGCAGTG (SEQ ID NO:21)	PCR, nsP4
nsP4a-R1	5'-ATGTCCTGCAACATATTCAA (SEQ ID NO:22)	PCR, nsP4
hOct4RTa-F1	5'-CGGCGCCAGAAGGGCAAGCG (SEQ ID NO:23)	PCR, OK
hKlf4RTb-R1	5'-CACCTGCTTGACGCAGTGTG (SEQ ID NO:24)	PCR, OK
hKlf4GC2For	5'-GCAGGAGGCGGTCTCTCTGTCACC (SEQ ID NO:35)	PCR, Klf4
hKlf4GC2Rev	5'-CAGGTGTGCCTTGAGATGGGAATC (SEQ ID NO:36)	PCR, Klf4
Bis-Oct-10F	5'-GGAGTAGAAGGATTGTTTTGGTTTA (SEQ ID NO:25)	bisulfito
Bis-Oct-9R	5'-AAACCTTAAAACTTAACCAATCC (SEQ ID NO:26)	bisulfito
Bis-Nanog-4F	5'-AGAGTAGTTGGGATTATAGATTTTA (SEQ ID NO:27)	bisulfito
Bis-Nanog-3R	5'-AACAACAAAACCTAAAACAAACC (SEQ ID NO:28)	bisulfito
EcoR1-Sac1- T7M1-VEE	5'-CGGAATTCGAGCTCTAATACGACTCACTATAGATGGGCGGCGCATGA GAGAAGCCCAG (SEQ ID NO:37)	T7 VEE PCR
Xba1-BstZ17I- VEE	5'-GCTCTAGAGTATACATCCTGGTAAACAGCGACTTGCCC (SEQ ID NO:38)	T7 VEE PCR

ARNm y síntesis de ARN de replicón. Se usó el plásmido pTNT-B18R para la síntesis de ARNm de B18R. El vector pTNT contiene una secuencia líder 5' de β -globina y una cola poli (A) sintética (30 bases) para potenciar la expresión de genes. 30 bases de poli (A) no fueron suficientes para estabilizar el ARNm, por lo que se añadió más cola poli (A) mediante polimerasa de cola poli (A). La síntesis de ARNm de B18R se llevó a cabo con nucleótidos modificados usando el kit de Sistema de Producción de ARN a Gran Escala RiboMAX SP6 (Promega). La modificación se llevó a cabo mediante el reemplazo del 100% de UTP por pseudouridina (Psi) (TriLink Biotechnologies) o del 25% de UTP y CTP con Psi y 5-metil-citidina (5mc) (TriLink Biotechnologies), respectivamente. Tras la reacción de transcripción, la plantilla de ADN se eliminó mediante digestión con ADNasa. El ARNm se purificó mediante extracción con Fenol/Cloroformo/alcohol de isoamilo (PCI) y Cloroformo/alcohol de isoamilo (CI), y a continuación se concentró mediante precipitación en acetato de amonio (2,5 M), que precipita selectivamente ARN, aunque dejando la mayoría de la proteína, ADN y NTPs no incorporados en el sobrenadante, según el protocolo del fabricante (Epicentre). Típicamente, se usaron 10 μ g de plásmido linealizado para una escala de reacción de 100 μ L y se recibieron aproximadamente 400 μ g de ARNm. Para el capping 5' del ARNm, se usó ScriptCap m7G Capping System™ y ScriptCap 2'-O-Metiltransferasa (Epicentre, actualmente disponible en CELLSRIPT) para producir ARN cap 1-capped, que procede a completar cuantitativamente el capping. Tras el Capping 5', el ARNm fue purificado brevemente mediante precipitación en acetato de amonio, y a continuación se añadió cola poli(A) adicional mediante polimerasa de poli(A) (Epicentre, actualmente disponible en CELLSRIPT). El ARNm que porta el Capping 5' y la cola poli(A) fue purificado mediante extracción con PCI y CI, seguido de precipitación en amonio. Para la síntesis de ARN de replicón, se linealizó plásmido plantilla mediante digestión con MluI, y a continuación se usó para la síntesis de ARN del mismo modo que para la síntesis de ARNm. La síntesis de replicón de ARN se llevó a cabo sin modificación de ARN. Después del tratamiento de ADNasa, el ARN sintetizado se purificó mediante precipitación de acetato de amonio sin purificación orgánica debido a que la mayoría del ARN grande fue atrapado en una fase intermedia tras la extracción orgánica. Al ARN de replicón se añadió Capping 5' y cola poli(A) como se ha descrito antes, y a continuación se purificó mediante precipitación en acetato de amonio sin purificación orgánica. Todos los ARNs fueron resuspendidos en la Disolución de Almacenamiento de ARN (Ambion) a una concentración de 1 μ g/ μ L y se almacenó a -80°C hasta su uso.

Preparación de medio acondicionado con B18R (B18R-CM). ARNm de B18R doble modificado al 25% (1 μ g por 1 pocillo de placa de 6 pocillos) fue transfectado en HFFs con Lipofectamine 2000 (Invitrogen). Después de 3 h, las células fueron cultivadas en DMEM avanzado (Invitrogen) que contenía un 15% de FCS (ES cell qualified, Millipore), penicilina y estreptomycin, o medio de cultivo ES. El medio de cultivo fue recolectado al siguiente día, se filtró y se diluyó 5 veces con medio de cultivo celular, y a continuación se usó como B18R-CM (20% de B18R-CM). La actividad de B18R-CM fue medida brevemente a través de la eficacia de una transfección repetida de ARNms.

Generación de iPS mediante transfección de replicón. Se sometió a pasaje células BJ o HFFs en una placa de 6 pocillos en el día 0 y se cultivó hasta una confluencia de ~90-100% (4×10^5 células/pocillo) en el día 1. Se transfectó una mezcla de 1 µg de ARN (ratio 3:1 de replicón de ARN de VEE a ARNm de B18R) con Lipofectamine 2000. Se usó ARNm modificado con B18 doblemente al 25% o ARNm modificado con Psi al 100% para la co-transfección. Después de 3 h, se cambió el medio de transfección al DMEM avanzado (Invitrogen) que contenía un 15% de FCS (ES cell qualified, Millipore), penicilina y estreptomycin. Las células fueron cultivadas en medio que contenía B18R-CM y puromicina (0,8 µg/mL) del día 2. El medio se cambió todos los días y se llevaron a cabo transfecciones cada 3 días (día 1, 4, 7, 10 o 14). Se usó medio ES del día 7. La puromicina se retiró el día 7 o el día 11. Un día después de la transfección final, las células fueron sometidas a pasaje a cebador STO y se cultivaron en medio ES que contenía B18R-CM. El medio ES se cambió todos los días y se cultivó hasta que se generaron las colonias de células iPS. Las colonias fueron seleccionadas mecánicamente para aislamiento de clones o teñidas con el kit de detección de fosfatasa alcalina (Millipore) o con disolución de tinción de AP preparada manualmente que contenía 1 mg/mL de FastRed TR (Sigma) y 0,4 mg/mL de 1-Naftil fosfato (Sigma) en tampón AP (Tris 100 mM, NaCl 100 mM y MgCl₂ 50 mM, pH 9,5).

RT-PCR para la detección de replicón de ARN. Los ARNs totales fueron aislados con el mini kit RNeasy (Qiagen) o con TRIzol (Invitrogen). Los ARNs purificados con TRIzol fueron purificados a continuación mediante precipitación en acetato de amonio. La síntesis de ADNcs se llevó a cabo con el Kit de transcripción QuantiTect Rev. (Qiagen) o con el kit de síntesis de ADNc iScript (Bio-Rad) a partir de 1 µg de ARN total. Se usaron 1-2 µL de 20 µL de reacción a RT para la amplificación PCR. La PCR se llevó a cabo con ADN polimerasa Taq (NEB) suplementada con potenciador PCRx (Invitrogen): 3 minutos a 94°C durante la desnaturalización inicial; 36 ciclos de 94 °C durante 25 s, 56 °C durante 25 s, 68 °C durante 30 s; seguido de 72 °C durante 5 minutos. Las secuencias de cebador usadas en la RT-PCR se han descrito en la Tabla 1.

RT-PCR TaqMan. Los ARNs totales de los cultivos libres de cebador de clones de iPSCs, HUES-9, BJ y HFFs fueron aislados con el mini kit RNeasy. Las reacciones de RT-PCR TaqMan fueron llevadas a cabo usando una reacción en una etapa de ARN-a-Ct (Applied Biosystems) según el protocolo del fabricante. Por reacción se usaron 10 ng de ARN total. Los cebadores y las sondas fueron obtenidos del catálogo de AB TaqMan Gene Expression Assay (GAPDH, Hs99999905_m1; POU5F1 Hs03005111_g1; Sox2 Hs01053049_s1; DNMT3B Hs00171876_m1; TERT Hs00972656_m1; Lin28 Hs00702808_s1; Nanog Hs02387400_g1; TDGF1 Hs02339499_g1). Las reacciones de PCR cuantitativa se llevaron a cabo por triplicado, y las condiciones fueron las siguientes: 20 min a 55 °C, 10 min a 95 °C, 40 ciclos de 95 °C durante 0,15 min, 65 °C durante 1 min. Los datos fueron analizados en el sistema de PCR en tiempo real 7300 (Applied Biosystems) usando el método delta-delta Ct.

Secuenciación genómica de bisulfito. La conversión de citosinas no metiladas en uracilo de ADN genómico se llevó a cabo con el kit EZ DNA Methylation-Gold (Zymo Research) según el protocolo del fabricante. Los ADNcs genómicos convertidos fueron usados a continuación para amplificación mediante PCR de la región de promotor de OCT4 o NANOG con ADN polimerasa ZymoTaq™ (Zymo Research). Los productos de PCR fueron clonados en el vector T de pBluescript SK+, y a continuación fueron secuenciados. Las secuencias de cebador usadas para PCR fueron descritas en la Tabla 1.

Formación de teratoma. Los clones de iPSC fueron cultivados con células cebadoras STO. Las células fueron recolectadas mediante un tratamiento con acutasa, y a continuación fueron inyectadas intramuscular o subcutáneamente en los músculos de la pata trasera o en el flanco dorsal de ratones nude (células cultivadas en disco de aproximadamente 10 cm para 1 inyección). Después de 5 a 8 semanas de inyección, los tumores fueron disecados y fijados con paraformaldehído al 4%. Los tumores fueron embebidos en parafina y se realizó un seccionamiento y después una tinción o inmunotinción con hematoxilina y eosina (H&E) de tres marcadores de capas germinales. Se usó AE1/AE3 (citoqueratina), NF-1 (células neuronales) y GFAP (células neuronales) para marcadores de ectoderma, Desmin (células musculares) para marcador de mesoderma, y AFP (endoderma primitivo y definitivo) para marcador de endoderma.

Tinción de inmunofluorescencia. Las células fueron lavadas dos veces en PBS y fijadas en paraformaldehído al 4% durante 10 minutos. Las células lavadas fueron tratadas con Triton X-100 al 0,1% en PBS durante 10 minutos. Las células fueron bloqueadas con BSA al 2% durante 1 h a temperatura ambiente (RT), y a continuación se incubaron con anticuerpos primarios en PBS a 4 °C durante una noche. Las células fueron lavadas e incubadas con anticuerpos secundarios seguido de la incubación con DAPI o Hoechst 33342, y después se lavaron y clasificaron en PBS. Los anticuerpos primarios, tal como los anticuerpos anti-Oct4 de conejo, anti-Nanog y anti-Sox2 de cabra, anti-SSEA4 de ratón, anti-Tra-1-60 y anti-Tra-1-81, fueron usados en diluciones entre 1:100 y 1:500. Se usaron anticuerpos secundarios Alexa Fluor 488 (BD Biosciences) en diluciones 1:800.

Anticuerpos. Los anticuerpos usados en esta investigación son los indicados a continuación; anti-OCT4 (sc-9081), anti-KLF4 (sc-20691), anti-GLIS1 (sc-67584), anti-c-MYC (sc-42), anti-LIN28 (sc-54030), TRA-1-60 (sc-21705), SSEA1 (sc-21702) y SSEA4 (sc-21704) de Santa Cruz; anti-SOX2 (AF2018) y anti-NANOG (AF1997) de R&D Systems; TRA-1-81 (09-0011) de Stemgent; AE1/AE3 (RB-9010P0), Desmin (MS-376-S0), AFP (RB-365) y GFAP (RB-087) de Labvision; NF-1 (NB-300-155) de Novus Biological.

Secuencia de ARN. Los ARNs totales fueron aislados con el mini kit RNeasy (Qiagen), y la biblioteca de ADNc de cada célula se sintetizó y analizó como se conoce en la técnica.

Para desarrollar una estrategia de generación de iPS basada en ARN, los esfuerzos se focalizaron en una estrategia que: 1) utilizar una única especie de ARN capaz de autorreplicarse por un número limitado de divisiones celulares, reduciendo de este modo el número de transfecciones; 2) fuera capaz de codificar al menos cuatro marcos de lectura abiertos (ORFs) de factor de reprogramación; y 3) expresaran de forma consistente todos los cuatro genes RF a unos niveles umbral elevados a lo largo de múltiples divisiones celulares. Para expresar ectópicamente los cuatro RFs, se usó un replicón de ARN de virus de encefalitis equina venezolana (VEE) modificado, no infeccioso, autorreplicante, que está siendo investigado actualmente como plataforma de expresión para el desarrollo de vacunas. El replicón de VEE es un ARN de especie individual de cadena positiva que imita el ARNm celular con un 5'-Cap y una cola poli(A) que no utiliza un intermedio de ADN, de tal modo que no hay potencial para integración genómica. VEE codifica proteínas complejas de replicación no estructural (nsP) como único ORF en el extremo 5' del ARN que está separado del ORF de la proteína estructural vírica en el extremo 3' (Fig. 1a). Petrakova et al. demostraron la capacidad de expresar proteínas exógenas reemplazando las proteínas estructurales 3' de ORF con GFP. Sin embargo, la exposición de células a ARN de VEE de cadena sencilla induce una fuerte respuesta inmune innata de IFN-alfa/beta que ha limitado de forma grave esta estrategia.

Para evaluar el replicón de ARN de VEE, el ORF 3' fue reemplazado por GFP, seguido de un sitio de entrada ribosomal interno (IRES) y un gen de resistencia a puromicina (Puror) (Fig. 1a). El ARN de VEE-GFP fue producido usando un kit de transcripción *in vitro* de polimerasa SP6 estándar seguido de un capping en 5', y adición de cola poli(A), dando lugar a un transcrito de ARN de 11500 nt de longitud completa con un rendimiento elevado. Para mitigar la respuesta inmune innata al ARN de VEE-GFP, se usó la proteína B18R del virus Vaccinia Occidental, que se une y neutraliza IFNs de tipo I. Se llevó a cabo una comparación de la transfección de fibroblastos de prepucio humano primarios (HFFs) con ARN de VEE-GFP sola, en presencia de la proteína B18R recombinante o con co-transfección de ARNm de B18R. En consonancia con la inducción de una respuesta inmune innata fuerte a las células expuestas a ARN de cadena sencilla, en ausencia de B18R, se observó poca o ninguna expresión de GFP (Fig. 1b). Aunque la adición de proteína B18R recombinante aumentó la expresión de GFP, el nivel de fluorescencia de GFP fue muy bajo. Sin embargo, la co-transfección de replicón de ARN de VEE-GFP con ARNm de B18R dio como resultado niveles elevados de expresión de GFP en HFFs (Fig. 1b-d), demostrando que se requiere B18R para una expresión eficiente de proteínas a partir del replicón de ARN de VEE.

La generación de células iPS requiere una expresión consistente a nivel alto de factores de los factores de reprogramación durante > 7 días; por tanto, se examinó la persistencia del replicón de VEE-GFP en fibroblastos. Se co-transfectaron HFFs con replicón de ARN de VEE-GFP y ARNm de B18R (ratio 3:1) en el día 1, a continuación se cultivaron en presencia o ausencia de medio acondicionado (CM) con B18R más/menos puromicina en el día 2. Aunque las células transfectadas con ARN de VEE-GFP/ARNm de B18R mostraron un nivel elevado de expresión de GFP en el día 1, el nivel de expresión se redujo rápidamente a lo largo de los siguientes días hasta los niveles de línea base en el día 7 (Fig. 1e). Además, en ausencia de exposición continua a B18R-CM, las células transfectadas con ARN de VEE-GFP dejaron de crecer y/o fueron asesinadas por la respuesta inmune innata (Fig. 1d). Por el contrario, las células transfectadas con ARNm de VEE-GFP/ARNm de B18R tratadas con B18R-CM/puro mantuvieron de forma persistente unos niveles elevados de expresión de GFP en >90% de las células con características de crecimiento saludable (Fig. 1d,e). Estos resultados demostraron la capacidad de la exposición a B18R para superar el problema de la respuesta inmune innata inducida por ARN de VEE, y también demostraron la capacidad para retener o degradar de forma selectiva al replicón de ARN de VEE frente a las células por exposición a o retirada de B18R-CM.

El ORF 3' del replicón de ARN de VEE se diseñó para codificar un único ORF combinado de tres factores de reprogramación, OCT4, KLF4, SOX2, separados por péptidos 2A de espaciado ribosomales internos. Los ORFs venían seguidos de un IRES, después de c-MYC (OKS-iM) o GLIS18 (OKS-iG), lo que evita la inestabilidad genómica inducida por c-MYC, seguido de un segundo IRES y el gen de resistencia a puromicina (Puror) (Fig. 1a; Tabla 1). De forma similar al protocolo de ARN de VEE-GFP, los ARNs de VEE-RF fueron producidos mediante transcripción *in vitro* de SP6, capping 5' y adición de cola poli(A), dando lugar a ARN de VEE-OKS-iM de ~14500 nt o ARN de VEE-OKS-iG de ~15000 nt, de longitud completa en un alto rendimiento. La co-transfección de replicones de ARN de VEE-OKS-iM o de ARN de VEE-OKS-iG más ARNm de B18R (ratio 3:1) en BJ o fibroblastos humanos HFF dio como resultados en niveles elevados extendidos de expresión para los cuatro RFs que superaron los niveles de expresión de RF de los retrovirus (Fig. 1f). Estas observaciones demostraron la capacidad para expresar cuatro factores de reprogramación a partir de único replicón de ARN de VEE-RF sintético en células humanas primarias, aunque utilizando B18R para bloquear la respuesta inmune innata.

Para desarrollar un protocolo de generación de células iPS basado en ARN, se evaluaron varios parámetros, que incluyen el número y el calendario de transfecciones de ARN de VEE-RF, la selección para la retención de replicón de ARN de VEE-RF mediante puromicina, y la organización genética del replicón de ARN de VEE-RF (Fig. 1a, 2a). Aunque incluso una transfección sencilla o doble de ARN de RF dio como resultado la generación de células iPS, tres o cuatro transfecciones en presencia de B18R de forma consistente dieron como resultado la mayor generación de colonias positivas para fosfatasa alcalina (AP+) (Fig. 2b-d). Se aislaron mecánicamente >100 colonias de células iPS a partir de los protocolos de ARN de VEE-OKS-iM y VEE-OKS-iG y tuvieron una tasa de éxito >95% en términos de capacidad de los clones de tipo iPS aislados para dividirse continuamente y retener una morfología de célula madre

embrionaria humana (hESC). De los >100 clones de tipo iPS aislados, 30 clones fueron aislados para la expresión de marcadores de células madre mediante inmunofluorescencia. Los 30 clones de iPS de ARN de VEE-RF analizados (6x clones de HFF-OKS-iM, 12x clones de BJ-OKS-iM, 6x clones de HFF-OKS-iG, 6x clones BJ-OKS-iG) mostraron todos una fuerte tinción nuclear de OCT4, SOX2 y NANOG endógenos, y una fuerte tinción de superficie celular de SSEA4, TRA-1-60 y TRA-1-81, con tinción negativa de SSEA1 (Fig. 2e). Para eliminar el replicón de ARN de VEE-RF, todos los protocolos de iPS retiraron B18R-CM y puromicina en el día 7 o 10 durante la reprogramación (Fig. 2a). Para confirmar la pérdida completa de replicones de ARN de VEE-RF, se desarrolló un protocolo altamente sensible y específico de qRT-PCR capaz de detectar <10 femtogramos del replicón de ARN de VEE-RF (Fig. 4). Como era de esperar, el análisis qRT-PCR demostró que todos los clones de células iPS habían perdido el replicón de ARN de VEE-RF (Tabla 2). Además, el análisis de cariotipo de 4 clones de células iPS independientes (BJ-OKS-iM n° 2 y n° 21, BJ-OKS-iG n°5, HFF-OKS-iM n°1) demostró cariotipos diploides normales (Fig. 5).

Tabla 2: Detección de replicón de ARN de RF mediante qRT-PCR

aClones	Pasaje n°								Tfx veces						
	P4	P5	P6	P7	P8	P9	P11								
	bR1R2R3	R1R2R3	R1R2R3	R1R2R3	R1R2R3	R1R2R3	R1R2R3	R1R2R3	cPL	cFD					
BJ-iM-1		+	+	+		-	-	-		-	-	ND	5	2	
BJ-iM-2			-	+	+/-		-	-	-				5	2	
BJ-iM-3			-	-	+/-								5	2	
BJ-iM-14			-	-	-		-	+/-	-		-	-	ND	1	2
BJ-iM-15			-	-	-		-	-	-				1	2	
BJ-iM-16			-	-	-		-	-	-				1	2	
BJ-iM-20			-	-	-					-	-	-	5	0	
BJ-iM-21			-	+	-		-	+/-	+/-	-	-	-	5	0	
BJ-iM-22			-	-	-					-	-	-	5	0	
BJ-iM-23			-	-	-					-	-	-	5	0	
BJ-iM-24			-	+	-		-	-	-	-	-	-	2	0	
BJ-iM-25			-	-	-					-	-	-	2	0	
HFF-iM-1	+	+	ND			-	-	-	-				2	2	
HFF-iM-2	+	+	ND			-	+	+		-	-	-	2	2	
HFF-iM-3	+	+	ND						-	-	-	-	2	2	
HFF-iM-4	-	+	-			-	-	-	-	-	-	-	2	2	
HFF-iM-5	-	-	-			-	-	-	-	-	-	-	2	2	
HFF-iM-6	+	+	+			-	+/-	-		-	-	-	2	2	
HFF-iM-7	+	+	ND						-	-	-		5	2	
HFF-iM-8	+	+	ND	+	+	ND			-	-	-		5	2	
HFF-iM-9	-	+	ND	+/-	+/-	ND			-	-	-		5	2	
HFF-iM-10	+	+	+			-	+/-	+/-	-	-	-		5	2	
HFF-iM-11	-	-	-			-	-	-	-	-	-		5	2	
HFF-iM-12	-	-	-			-	-	-	-	-	-		5	2	
BJ-iG-1	-	-	-			-	-	-	-	-	-		5	0	
BJ-iG-2	-	-	-			-	-	-	-	-	-		5	0	
BJ-iG-3	-	-	-			-	-	-	-	-	-		5	0	
BJ-iG-4	-	-	-			-	-	-	-	-	-		5	0	
BJ-iG-5	-	-	-			-	-	-	-	-	-		5	0	
BJ-iG-6	-	+/-	-			-	-	-	-	-	-		5	0	
HFF-iG-7	-	-	-			-	-	-	-	-	-		4	0	
HFF-iG-8	-	+/-	-			-	-	-	-	-	-		4	0	
HFF-iG-9	-	-	-			-	-	-	-	-	-		4	0	
HFF-iG-10	-	+/-	-			-	+/-	-	-	-	-		4	0	
HFF-iG-11	-	-	+/-			-	+/-	-	-	-	-		4	0	
HFF-iG-12	-	-	-			-	-	-	-	-	-		4	0	

a; indica clones del replicón de ARN OKS-iM, iG indica clones del replicón de ARN OKS-iG. b; regiones correspondientes a RT-PCR, R1; nsP2, R2; nsP4, R3; Oct4-T2A-Klf4 (OK), c; transfección sobre placa (PL) antes de pasaje a células cebadoras, d; transfección después de pasaje a células cebadoras (FD). +; banda positiva detectada, +/-; banda tenue detectada; -; no se detecta banda. ND, no realizado.

Para caracterizar adicionalmente los clones de células iPS establecidos, se analizó la expresión de genes marcadores de ES humanos mediante qRT-PCR. En consonancia con los niveles de expresión en células HUES9 ES humanas, los clones de iPS generados a partir de fibroblastos BJ y HFF originales con el protocolo de ARN de VEE-RF de OKS-iM o de OKS-iG expresaron niveles robustos de OCT4, SOX2, NANOG, LIN28, TDGF1, DNMT3B y TERT endógenos, al contrario que los niveles de expresión bajos o nulos en los fibroblastos de BJ y HFF originales (Fig. 3a). Una característica de la pluripotencia inducida es una metilación de ADN reducida de dinucleótidos CpG en las regiones promotoras de OCT4 y NANOG. El secuenciamiento genómico de bisulfito de ambas regiones promotoras, OCT4 y NANOG, mostró una desmetilación extensiva en los clones de células iPS en comparación con los fibroblastos originales (Fig. 3b). Para investigar perfiles de expresión de ARNm de genoma amplio en los clones de células iPS,

cuyo secuenciamiento de ARN genómico (ARN-seq) se llevó a cabo en los clones de células iPS generadas mediante ARN de VEE-RF de OKS-iM y OKS-iG, las células BJ y HUES-9 ES originales controlan. Los cuatro clones de células iPS analizados mediante ARN-seq mostraron una agrupación jerarquizada no supervisada y marcas de expresión características de las células HUES9 humanas que eran altamente divergentes de los fibroblastos humanos originales (Fig. 3c,d). Finalmente, la pluripotencia *in vivo* de los clones de células iPS humanas fue evaluada para determinar su capacidad para diferenciarse en células de las tres capas germinales mediante formación de teratoma en ratones inmunocomprometidos. Todos los clones de iPS de ARN de VEE-RF analizados formaron teratomas que contenían tipos celulares representativos de las tres capas germinales, se detectaron mediante tinción H&E que se confirmó mediante tinción de inmunohistoquímica (Fig. 3e; Fig. 6). Colectivamente, estas observaciones confirman la capacidad de los replicones de ARN de VEE-RF de OKS-iM y de OKS-iG para generar de forma eficiente células iPS humanas pluripotentes.

La generación de células iPS tiene un gran potencial para el desarrollo de terapias personalizadas de células madre; sin embargo, seguía sin disponerse de un método claro, directo y consistente basado en ARN para generar células iPS. La descripción proporciona una estrategia basada en ARN simple y altamente reproducible para generar células iPS mediante transfección de un único replicón de ARN de VEE-RF sintético que expresa uno, dos, tres, cuatro o más factores de reprogramación independientes. Las células iPS generadas mediante ARN de VEE-RF adquieren una completa pluripotencia a través de criterios biológicos y moleculares *in vivo* que se asemejan a células ES humanas. La generación del transcrito de ARN de VEE-RF utiliza un kit de transcripción *in vitro* SP6 estándar que no requiere condiciones especiales y, por tanto, simplifica adicionalmente la estrategia para un uso amplio. Expresando los cuatro RFs en niveles elevados consistentes a lo largo del tiempo en la misma célula en combinación con la replicación del ARN de VEE-RF durante un número limitado de generaciones múltiples de células, la estrategia de ARN de VEE-RF resuelve los dos problemas de falta de eficiencia principales asociados al intento de generar células iPS mediante transfecciones diarias repetidas diariamente durante >14 días de cuatro ARNms de RF individuales. De forma importante, el ARN de VEE-RF es una estrategia ectópica de golpea-y-corre que no utiliza un intermedio de ADN y, por lo tanto, no existe la posibilidad de que se dé una mutación integrativa como en las estrategias de células iPS basadas en vector de ADN. Además, la pérdida programada de replicón de ARN de VEE-RF por degradación se puede regular retirando B18R del medio. Usando la estrategia de ARN de VEE-RF, se generaron >100 clones de células iPS independientes con de los protocolos de ARN de VEE-RF OCT4/KLF4/SOX2/c-MYC y OCT4/KLF4/SOX2/GLIS1 a partir de poblaciones de fibroblastos humanos originales independientes. Adicionalmente, la estrategia de ARN de VEE-RF se puede diseñar para expresar combinaciones de RF alternativas y/o la inserción de ORFs de RF adicionales en la cadena estructural de ARN de RF para refinar la generación de células iPS a partir de tipos celulares específicos, o para uso en dirigir la trans-diferenciación. En resumen, la estrategia de replicón de ARN de VEE-RF tiene una amplia aplicabilidad para la generación eficiente de células iPS humanas con uso definitivo en terapias de células madre humanas y medicina regenerativa.

Números de acceso. Los datos de ARN-Seq han sido enviados y se puede acceder a los mismos mediante el número de acceso de "Gene Expression Omnibus" (GEO) GSE38265.

Tabla 3. Generación de células iPS con replicón de ARN de VEE-RF.

Replicón de ARN	Célul.	Tfx Días	Selección de puromicina	Colonias AP+ por pocillo inicial
OKS-iM	BJ	d1,	d2-d7	6
OKS-iM	BJ	d1, 2	d2-d7	32
OKS-iM	BJ	d1, 2, 3	d2-d7	221
OKS-iM	BJ	d1, 4, 7, 10	d2-d7	140
OKS-iM	BJ	d1	ninguna	6
OKS-iM	BJ	d1, 2	ninguna	12
OKS-iM	BJ	d1, 2, 3	ninguna	8
OKS-iM	HFF	d1, 5, 9	d2-d10	179
OKS-iM	HFF	d1, 4, 7, 10	d2-d4	189
OKS-iM	HFF	d1, 4, 7, 10	d2-d7	308
OKS-iM	HFF	d1, 4, 7, 10	d2-d10	338
OKS-iG	BJ	d1, 4, 7, 10	d2-d7	282
OKS-iG	BJ	d1, 4, 7, 10	d2-d10	122
OKS-iG	HFF	d1, 4, 7, 10	d2-d7	267
OKS-iG	HFF	d1, 4, 7, 10	d2-d10	248

LISTA DE SECUENCIAS

	<110> The Regents of the University of California Dowdy, Steven F. Yoshioka, Naohisa	
5	<120> Generación de células iPS humanas mediante un ARN autoreplicante sintético	
	<130> 00015-219WO1	
10	<140> Sin asignar aún <141> 21-05-2013	
	<150> US 61/649,876 <151> 21-05-2012	
15	<150> US 61/798,229 <151> 15-03-2013	
	<160> 40	
20	<170> PatentIn versión 3.5	
	<210> 1 <211> 918 <212> ADN <213> Homo sapiens	
25	<220> <221> CDS <222> (1)..(918)	
30	<400> 1 atg agt gtg gat cca gct tgt ccc caa agc ttg cct tgc ttt gaa gca Met Ser Val Asp Pro Ala Cys Pro Gln Ser Leu Pro Cys Phe Glu Ala 1 5 10 15 tcc gac tgt aaa gaa tct tca cct atg cct gtg att tgt ggg cct gaa Ser Asp Cys Lys Glu Ser Ser Pro Met Pro Val Ile Cys Gly Pro Glu 20 25 30 gaa aac tat cca tcc ttg caa atg tct tct gct gag atg cct cac acg Glu Asn Tyr Pro Ser Leu Gln Met Ser Ser Ala Glu Met Pro His Thr 35 40 45 gag act gtc tct cct ctt cct tcc tcc atg gat ctg ctt att cag gac Glu Thr Val Ser Pro Leu Pro Ser Ser Met Asp Leu Leu Ile Gln Asp 50 55 60 agc cct gat tct tcc acc agt ccc aaa ggc aaa caa ccc act tct gca Ser Pro Asp Ser Ser Thr Ser Pro Lys Gly Lys Gln Pro Thr Ser Ala 65 70 75 80 gag aag agt gtc gca aaa aag gaa gac aag gtc ccg gtc aag aaa cag Glu Lys Ser Val Ala Lys Lys Glu Asp Lys Val Pro Val Lys Lys Gln 85 90 95 aag acc aga act gtg ttc tct tcc acc cag ctg tgt gta ctc aat gat Lys Thr Arg Thr Val Phe Ser Ser Thr Gln Leu Cys Val Leu Asn Asp 100 105 110	48 96 144 192 240 288 336

ES 2 698 527 T3

```

aga ttt cag aga cag aaa tac ctc agc ctc cag cag atg caa gaa ctc      384
Arg Phe Gln Arg Gln Lys Tyr Leu Ser Leu Gln Gln Met Gln Glu Leu
      115                      120                      125

tcc aac atc ctg aac ctc agc tac aaa cag gtg aag acc tgg ttc cag      432
Ser Asn Ile Leu Asn Leu Ser Tyr Lys Gln Val Lys Thr Trp Phe Gln
      130                      135                      140

aac cag aga atg aaa tct aag agg tgg cag aaa aac aac tgg ccg aag      480
Asn Gln Arg Met Lys Ser Lys Arg Trp Gln Lys Asn Asn Trp Pro Lys
      145                      150                      155                      160

aat agc aat ggt gtg acg cag aag gcc tca gca cct acc tac ccc agc      528
Asn Ser Asn Gly Val Thr Gln Lys Ala Ser Ala Pro Thr Tyr Pro Ser
      165                      170                      175

ctt tac tct tcc tac cac cag gga tgc ctg gtg aac ccg act ggg aac      576
Leu Tyr Ser Ser Tyr His Gln Gly Cys Leu Val Asn Pro Thr Gly Asn
      180                      185                      190

ctt cca atg tgg agc aac cag acc tgg aac aat tca acc tgg agc aac      624
Leu Pro Met Trp Ser Asn Gln Thr Trp Asn Asn Ser Thr Trp Ser Asn
      195                      200                      205

cag acc cag aac atc cag tcc tgg agc aac cac tcc tgg aac act cag      672
Gln Thr Gln Asn Ile Gln Ser Trp Ser Asn His Ser Trp Asn Thr Gln
      210                      215                      220

acc tgg tgc acc caa tcc tgg aac aat cag gcc tgg aac agt ccc ttc      720
Thr Trp Cys Thr Gln Ser Trp Asn Asn Gln Ala Trp Asn Ser Pro Phe
      225                      230                      235                      240

tat aac tgt gga gag gaa tct ctg cag tcc tgc atg cag ttc cag cca      768
Tyr Asn Cys Gly Glu Glu Ser Leu Gln Ser Cys Met Gln Phe Gln Pro
      245                      250                      255

aat tct cct gcc agt gac ttg gag gct gcc ttg gaa gct gct ggg gaa      816
Asn Ser Pro Ala Ser Asp Leu Glu Ala Ala Leu Glu Ala Gly Glu
      260                      265                      270

ggc ctt aat gta ata cag cag acc act agg tat ttt agt act cca caa      864
Gly Leu Asn Val Ile Gln Gln Thr Thr Arg Tyr Phe Ser Thr Pro Gln
      275                      280                      285

acc atg gat tta ttc cta aac tac tcc atg aac atg caa cct gaa gac      912
Thr Met Asp Leu Phe Leu Asn Tyr Ser Met Asn Met Gln Pro Glu Asp
      290                      295                      300

gtg tga
Val
305

```

<210> 2
 <211> 305
 <212> PRT
 <213> Homo sapiens

<400> 2
 Met Ser Val Asp Pro Ala Cys Pro Gln Ser Leu Pro Cys Phe Glu Ala
 1 5 10 15

ES 2 698 527 T3

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

<210> 3
 <211> 1083
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1083)

<400> 3
 atg gcg gga cac ctg gct tcg gat ttc gcc ttc tcg ccc cct cca ggt 48
 Met Ala Gly His Leu Ala Ser Asp Phe Ala Phe Ser Pro Pro Pro Gly
 1 5 10 15
 ggt gga ggt gat ggg cca ggg ggg ccg gag ccg ggc tgg gtt gat cct 96
 Gly Gly Gly Asp Gly Pro Gly Gly Pro Glu Pro Gly Trp Val Asp Pro
 20 25 30
 cgg acc tgg cta agc ttc caa ggc cct cct gga ggg cca gga atc ggg 144
 Arg Thr Trp Leu Ser Phe Gln Gly Pro Pro Gly Gly Pro Gly Ile Gly
 35 40 45
 ccg ggg gtt ggg cca ggc tct gag gtg tgg ggg att ccc cca tgc ccc 192

ES 2 698 527 T3

Pro	Gly	Val	Gly	Pro	Gly	Ser	Glu	Val	Trp	Gly	Ile	Pro	Pro	Cys	Pro	
50						55					60					
ccg	ccg	tat	gag	ttc	tgt	ggg	ggg	atg	gcg	tac	tgt	ggg	ccc	cag	gtt	240
Pro	Pro	Tyr	Glu	Phe	Cys	Gly	Gly	Met	Ala	Tyr	Cys	Gly	Pro	Gln	Val	
65					70					75					80	
gga	gtg	ggg	cta	gtg	ccc	caa	ggc	ggc	ttg	gag	acc	tct	cag	cct	gag	288
Gly	Val	Gly	Leu	Val	Pro	Gln	Gly	Gly	Leu	Glu	Thr	Ser	Gln	Pro	Glu	
				85					90					95		
ggc	gaa	gca	gga	gtc	ggg	gtg	gag	agc	aac	tcc	gat	ggg	gcc	tcc	ccg	336
Gly	Glu	Ala	Gly	Val	Gly	Val	Glu	Ser	Asn	Ser	Asp	Gly	Ala	Ser	Pro	
			100					105					110			
gag	ccc	tgc	acc	gtc	acc	cct	ggt	gcc	gtg	aag	ctg	gag	aag	gag	aag	384
Glu	Pro		Thr	Val	Thr	Pro	Gly	Ala	Val	Lys	Leu	Lys	Lys	Glu	Lys	
		115					120					125				
ctg	gag	caa	aac	ccg	gag	gag	tcc	cag	gac	atc	aaa	gct	ctg	cag	aaa	432
Leu	Glu	Gln	Asn	Pro	Glu	Glu	Ser	Gln	Asp	Ile	Lys	Ala	Leu	Gln	Lys	
		130				135					140					
gaa	ctc	gag	caa	ttt	gcc	aag	ctc	ctg	aag	cag	aag	agg	atc	acc	ctg	480
Glu	Leu	Glu	Gln	Phe	Ala	Lys	Leu	Leu	Lys	Gln	Lys	Arg	Ile	Thr	Leu	
145					150					155					160	
gga	tat	aca	cag	gcc	gat	gtg	ggg	ctc	acc	ctg	ggg	gtt	cta	ttt	ggg	528
Gly	Tyr	Thr	Gln	Ala	Asp	Val	Gly	Leu	Thr	Leu	Gly	Val	Leu	Phe	Gly	
				165				170						175		
aag	gta	ttc	agc	caa	acg	acc	atc	tgc	cgc	ttt	gag	gct	ctg	cag	ctt	576
Lys	Val	Phe	Ser	Gln	Thr	Thr	Ile	Cys	Arg	Phe	Glu	Ala	Leu	Gln	Leu	
			180					185					190			
agc	ttc	aag	aac	atg	tgt	aag	ctg	cgg	ccc	ttg	ctg	cag	aag	tggt	gtg	624
Ser	Phe	Lys	Asn	Met	Cys	Lys	Leu	Arg	Pro	Leu	Leu	Gln	Lys	Trp	Val	
		195				200						205				
gag	gaa	gct	gac	aac	aat	gaa	aat	ctt	cag	gag	ata	tgc	aaa	gca	gaa	672
Glu	Glu	Ala	Asp	Asn	Asn	Glu	Asn	Leu	Gln	Glu	Ile	Cys	Lys	Ala	Glu	
		210				215					220					
acc	ctc	gtg	cag	gcc	cga	aag	aga	aag	cga	acc	agt	atc	gag	aac	cga	720
Thr	Leu	Val	Gln	Ala	Arg	Lys	Arg	Lys	Arg	Thr	Ser	Ile	Glu	Asn	Arg	
225					230					235					240	
gtg	aga	ggc	aac	ctg	gag	aat	ttg	ttc	ctg	cag	tgc	ccg	aaa	ccc	aca	768
Val	Arg	Gly	Asn	Leu	Glu	Asn	Leu	Phe	Leu	Gln	Cys	Pro	Lys	Pro	Thr	
				245					250					255		
ctg	cag	cag	atc	agc	cac	atc	gcc	cag	cag	ctt	ggg	ctc	gag	aag	gat	816
Leu	Gln	Gln	Ile	Ser	His	Ile	Ala	Gln	Gln	Leu	Gly	Leu	Glu	Lys	Asp	
			260					265					270			
gtg	gtc	cga	gtg	tggt	ttc	tgt	aac	cgg	cgc	cag	aag	ggc	aag	cga	tca	864
Val	Val	Arg	Val	Trp	Phe	Cys	Asn	Arg	Arg	Gln	Lys	Gly	Lys	Arg	Ser	
		275					280					285				
agc	agc	gac	tat	gca	caa	cga	gag	gat	ttt	gag	gct	gct	ggg	tct	cct	912
Ser	Ser	Asp	Tyr	Ala	Gln	Arg	Glu	Asp	Phe	Glu	Ala	Ala	Gly	Ser	Pro	

ES 2 698 527 T3

290	295	300	
ttc tca ggg gga cca gtg tcc ttt cct ctg gcc cca ggg ccc cat ttt			960
Phe Ser Gly Gly Pro Val Ser Phe Pro Leu Ala Pro Gly Pro His Phe			
305	310	315	320
ggg acc cca ggc tat ggg agc cct cac ttc act gca ctg tac tcc tcg			1008
Gly Thr Pro Gly Tyr Gly Ser Pro His Phe Thr Ala Leu Tyr Ser Ser			
	325	330	335
gtc cct ttc cct gag ggg gaa gcc ttt ccc cct gtc tcc gtc acc act			1056
Val Pro Phe Pro Glu Gly Glu Ala Phe Pro Pro Val Ser Val Thr Thr			
	340	345	350
ctg ggc tct ccc atg cat tca aac tga			1083
Leu Gly Ser Pro Met His Ser Asn			
	355	360	

<210> 4
 <211> 360
 <212> PRT
 <213> Homo sapiens

<400> 4

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

Ser Ser Asp Tyr Ala Gln Arg Glu Asp Phe Glu Ala Ala Gly Ser Pro	
290	295 300
Phe Ser Gly Gly Pro Val Ser Phe Pro Leu Ala Pro Gly Pro His Phe	
305	310 315 320
Gly Thr Pro Gly Tyr Gly Ser Pro His Phe Thr Ala Leu Tyr Ser Ser	
	325 330 335
Val Pro Phe Pro Glu Gly Glu Ala Phe Pro Pro Val Ser Val Thr Thr	
	340 345 350
Leu Gly Ser Pro Met His Ser Asn	
	355 360

<210> 5

ES 2 698 527 T3

<211> 954
<212> ADN
<213> Homo sapiens

5

<220>
<221> CDS
<222> (1)..(954)

<400> 5
atg tac aac atg atg gag acg gag ctg aag ccg ccg ggc ccg cag caa 48
Met Tyr Asn Met Met Glu Thr Glu Leu Lys Pro Pro Gly Pro Gln Gln
1 5 10 15
act tcg ggg ggc ggc ggc ggc aac tcc acc gcg gcg gcg gcc ggc ggc 96
Thr Ser Gly Gly Gly Gly Gly Asn Ser Thr Ala Ala Ala Ala Gly Gly
20 25 30
aac cag aaa aac agc ccg gac cgc gtc aag ccg ccc atg aat gcc ttc 144
Asn Gln Lys Asn Ser Pro Asp Arg Val Lys Arg Pro Met Asn Ala Phe
35 40 45
atg gtg tgg tcc cgc ggg cag cgg cgc aag atg gcc cag gag aac ccc 192
Met Val Trp Ser Arg Gly Gln Arg Arg Lys Met Ala Gln Glu Asn Pro
50 55 60
aag atg cac aac tcg gag atc agc aag cgc ctg ggc gcc gag tgg aaa 240
Lys Met His Asn Ser Glu Ile Ser Lys Arg Leu Gly Ala Glu Trp Lys
65 70 75 80
ctt ttg tcg gag acg gag aag cgg ccg ttc atc gac gag gct aag cgg 288
Leu Leu Ser Glu Thr Glu Lys Arg Pro Phe Ile Asp Glu Ala Lys Arg
85 90 95
ctg cga gcg ctg cac atg aag gag cac ccg gat tat aaa tac cgg ccc 336
Leu Arg Ala Leu His Met Lys Glu His Pro Asp Tyr Lys Tyr Arg Pro
100 105 110
cgg cgg aaa acc aag acg ctc atg aag aag gat aag tac acg ctg ccc 384
Arg Arg Lys Thr Lys Thr Leu Met Lys Lys Asp Lys Tyr Thr Leu Pro
115 120 125
ggc ggg ctg ctg gcc ccc ggc ggc aat agc atg gcg agc ggg gtc ggg 432
Gly Gly Leu Leu Ala Pro Gly Gly Asn Ser Met Ala Ser Gly Val Gly
130 135 140
gtg ggc gcc ggc ctg ggc gcg ggc gtg aac cag cgc atg gac agt tac 480
Val Gly Ala Gly Leu Gly Ala Gly Val Asn Gln Arg Met Asp Ser Tyr
145 150 155 160
10 gcg cac atg aac ggc tgg agc aac ggc agc tac agc atg atg cag gac 528

ES 2 698 527 T3

Ala	His	Met	Asn	Gly	Trp	Ser	Asn	Gly	Ser	Tyr	Ser	Met	Met	Gln	Asp		
				165					170					175			
cag	ctg	ggc	tac	ccg	cag	cac	ccg	ggc	ctc	aat	gcg	cac	ggc	gca	gcg	576	
Gln	Leu	Gly	Tyr	Pro	Gln	His	Pro	Gly	Leu	Asn	Ala	His	Gly	Ala	Ala		
			180					185					190				
cag	atg	cag	ccc	atg	cac	cgc	tac	gac	gtg	agc	gcc	ctg	cag	tac	aac	624	
Gln	Met	Gln	Pro	Met	His	Arg	Tyr	Asp	Val	Ser	Ser	Ala	Leu	Gln	Tyr	Asn	
			195				200						205				
tcc	atg	acc	agc	tcg	cag	acc	tac	atg	aac	ggc	tcg	ccc	acc	tac	agc	672	
Ser	Met	Thr	Ser	Ser	Gln	Thr	Tyr	Met	Asn	Gly	Ser	Pro	Thr	Tyr	Ser		
	210						215					220					
atg	tcc	tac	tcg	cag	cag	ggc	acc	cct	ggc	atg	gct	ctt	ggc	tcc	atg	720	
Gln	Ser	Tyr	Ser	Gln	Gly	Thr	Pro	Gly	Met	Ala	Leu	Gly	Ser	Met			
	225				230				235					240			
ggc	tcg	gtg	gtc	aag	tcc	gag	gcc	agc	tcc	agc	ccc	cct	gtg	gtt	acc	768	
Gly	Ser	Val	Val	Lys	Ser	Glu	Ala	Ser	Ser	Ser	Pro	Pro	Val	Val	Thr		
				245					250					255			
tct	tcc	tcc	cac	tcc	agg	gcg	ccc	tgc	cag	gcc	ggg	gac	ctc	cgg	gac	816	
Ser	Ser	Ser	His	Ser	Arg	Ala	Pro	Cys	Gln	Ala	Gly	Asp	Leu	Arg	Asp		
			260					265					270				
atg	atc	agc	atg	tat	ctc	ccc	ggc	gcc	gag	gtg	ccg	gaa	ccc	gcc	gcc	864	
Met	Ile	Ser	Met	Tyr	Leu	Pro	Gly	Ala	Glu	Val	Pro	Glu	Pro	Ala	Ala		
	275						280					285					
ccc	agc	aga	ctt	cac	atg	tcc	cag	cac	tac	cag	agc	ggc	ccg	gtg	ccc	912	
Pro	Ser	Arg	Leu	His	Met	Ser	Gln	His	Tyr	Gln	Ser	Gly	Pro	Val	Pro		
	290					295					300						
ggc	acg	gcc	att	aac	ggc	aca	ctg	ccc	ctc	tca	cac	atg	tga			954	
Gly	Thr	Ala	Ile	Asn	Gly	Thr	Leu	Pro	Leu	Ser	His	Met					
	305				310				315								

<210> 6

<211> 317

<212> PRT

<213> Homo sapiens

<400> 6

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

ES 2 698 527 T3

```

      115      120      125
Gly Gly Leu Leu Ala Pro Gly Gly Asn Ser Met Ala Ser Gly Val Gly
  130      135      140
Val Gly Ala Gly Leu Gly Ala Gly Val Asn Gln Arg Met Asp Ser Tyr
  145      150      155      160
Ala His Met Asn Gly Trp Ser Asn Gly Ser Tyr Ser Met Met Gln Asp
      165      170      175
Gln Leu Gly Tyr Pro Gln His Pro Gly Leu Asn Ala His Gly Ala Ala
      180      185      190
Gln Met Gln Pro Met His Arg Tyr Asp Val Ser Ala Leu Gln Tyr Asn
      195      200      205
Ser Met Thr Ser Ser Gln Thr Tyr Met Asn Gly Ser Pro Thr Tyr Ser
      210      215      220
Met Ser Tyr Ser Gln Gln Gly Thr Pro Gly Met Ala Leu Gly Ser Met
      225      230      235      240
Gly Ser Val Val Lys Ser Glu Ala Ser Ser Ser Pro Pro Val Val Thr
      245      250      255
Ser Ser Ser His Ser Arg Ala Pro Cys Gln Ala Gly Asp Leu Arg Asp
      260      265      270
Met Ile Ser Met Tyr Leu Pro Gly Ala Glu Val Pro Glu Pro Ala Ala
      275      280      285
Pro Ser Arg Leu His Met Ser Gln His Tyr Gln Ser Gly Pro Val Pro
      290      295      300
Gly Thr Ala Ile Asn Gly Thr Leu Pro Leu Ser His Met
      305      310      315

```

<210> 7
 <211> 1440
 <212> ADN
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1440)

```

<400> 7
atg agg cag cca cct ggc gag tct gac atg gct gtc agc gac gcg ctg      48
Met Arg Gln Pro Pro Gly Glu Ser Asp Met Ala Val Ser Asp Ala Leu
1      5      10      15

ctc cca tct ttc tcc acg ttc gcg tct ggc ccg gcg gga agg gag aag      96
Leu Pro Ser Phe Ser Thr Phe Ala Ser Gly Pro Ala Gly Arg Glu Lys
20      25      30

aca ctg cgt caa gca ggt gcc ccg aat aac cgc tgg cgg gag gag ctc      144
Thr Leu Arg Gln Ala Gly Ala Pro Asn Asn Arg Trp Arg Glu Glu Leu
35      40      45

tcc cac atg aag cga ctt ccc cca gtg ctt ccc ggc cgc ccc tat gac      192
Ser His Met Lys Arg Leu Pro Pro Val Leu Pro Gly Arg Pro Tyr Asp
50      55      60

ctg gcg gcg gcg acc gtg gcc aca gac ctg gag agc ggc gga gcc ggt      240
Leu Ala Ala Ala Thr Val Ala Thr Asp Leu Glu Ser Gly Gly Ala Gly
65      70      75      80

gcg gct tgc ggc ggt agc aac ctg gcg ccc cta cct cgg aga gag acc      288
Ala Ala Cys Gly Gly Ser Asn Leu Ala Pro Leu Pro Arg Arg Glu Thr
85      90      95

gag gag ttc aac gat ctc ctg gac ctg gac ttt att ctc tcc aat tcg      336
Glu Glu Phe Asn Asp Leu Leu Asp Leu Asp Phe Ile Leu Ser Asn Ser

```

ES 2 698 527 T3

																100								105								110													
ctg	acc	cat	cct	ccg	gag	tca	gtg	gcc	gcc	acc	gtg	tcc	tcg	tca	gcg																	384													
Leu	Thr	His	Pro	Pro	Glu	Ser	Val	Ala	Ala	Thr	Val	Ser	Ser	Ser	Ala																														
																115								120								125													
tca	gcc	tcc	tct	tcg	tcg	tcg	ccg	tcg	agc	agc	ggc	cct	gcc	agc	gcg																	432													
Ser	Ala	Ser	Ser	Ser	Ser	Ser	Pro	Ser	Ser	Ser	Gly	Pro	Ala	Ser	Ala																														
																130								135								140													
ccc	tcc	acc	tgc	agc	ttc	acc	tat	ccg	atc	cgg	gcc	ggg	aac	gac	ccg																	480													
Pro	Ser	Thr	Cys	Ser	Phe	Thr	Tyr	Pro	Ile	Arg	Ala	Gly	Asn	Asp	Pro																														
																145								150								155													
ggc	gtg	gcg	ccg	ggc	ggc	acg	ggc	gga	ggc	ctc	ctc	tat	ggc	agg	gag																	528													
Gly	Val	Ala	Pro	Gly	Gly	Thr	Gly	Gly	Gly	Leu	Leu	Tyr	Gly	Arg	Glu																														
																165								170								175													
tcc	gct	ccc	cct	ccg	acg	gct	ccc	ttc	aac	ctg	gcg	gac	atc	aac	gac																	576													
Ser	Ala	Pro	Pro	Pro	Thr	Ala	Pro	Phe	Asn	Leu	Ala	Asp	Ile	Asn	Asp																														
																180								185								190													
gtg	agc	ccc	tcg	ggc	ggc	ttc	gtg	gcc	gag	ctc	ctg	cgg	cca	gaa	ttg																	624													
Val	Ser	Pro	Ser	Gly	Gly	Phe	Val	Ala	Glu	Leu	Leu	Arg	Pro	Glu	Leu																														
																195								200								205													
gac	ccg	gtg	tac	att	ccg	ccg	cag	cag	ccg	cag	ccg	cca	ggt	ggc	ggg																	672													
Asp	Pro	Val	Tyr	Ile	Pro	Pro	Gln	Gln	Pro	Gln	Pro	Pro	Gly	Gly	Gly																														
																210								215								220													
ctg	atg	ggc	aag	ttc	gtg	ctg	aag	gcg	tcg	ctg	agc	gcc	cct	ggc	agc																	720													
Leu	Met	Gly	Lys	Phe	Val	Leu	Lys	Ala	Ser	Leu	Ser	Ala	Pro	Gly	Ser																														
																225								230								235								240					
gag	tac	ggc	agc	ccg	tcg	gtc	atc	agc	gtc	agc	aaa	ggc	agc	cct	gac																	768													
Glu	Tyr	Gly	Ser	Pro	Ser	Val	Ile	Ser	Val	Ser	Lys	Gly	Ser	Pro	Asp																														
																245								250								255													
ggc	agc	cac	ccg	gtg	gtg	gtg	gcg	ccc	tac	aac	ggc	ggg	ccg	ccg	cgc																	816													
Gly	Ser	His	Pro	Val	Val	Val	Ala	Pro	Tyr	Asn	Gly	Gly	Pro	Pro	Arg																														
																260								265								270													
acg	tgc	ccc	aag	atc	aag	cag	gag	gcg	gtc	tct	tcg	tgc	acc	cac	ttg																	864													
Thr	Cys	Pro	Lys	Ile	Lys	Gln	Glu	Ala	Val	Ser	Ser	Cys	Thr	His	Leu																														
																275								280								285													
ggc	gct	gga	ccc	cct	ctc	agc	aat	ggc	cac	cgg	ccg	gct	gca	cac	gac																	912													
Gly	Ala	Gly	Pro	Pro	Leu	Ser	Asn	Gly	His	Arg	Pro	Ala	Ala	His	Asp																														
																290								295								300													
ttc	ccc	ctg	ggg	cgg	cag	ctc	ccc	agc	agg	act	acc	ccg	acc	ctg	ggt																	960													
Phe	Pro	Leu	Gly	Arg	Gln	Leu	Pro	Ser	Arg	Thr	Thr	Pro	Thr	Leu	Gly																														
																305								310								315								320					
ctt	gag	gaa	gtg	ctg	agc	agc	agg	gac	tgt	cac	cct	gcc	ctg	ccg	ctt																	1008													
Leu	Glu	Glu	Val	Leu	Ser	Ser	Arg	Asp	Cys	His	Pro	Ala	Leu	Pro	Leu																														
																325								330								335													
cct	ccc	ggc	ttc	cat	ccc	cac	ccg	ggg	ccc	aat	tac	cca	tcc	ttc	ctg																	1056													
Pro	Pro	Gly	Phe	His	Pro	His	Pro	Gly	Pro	Asn	Tyr	Pro	Ser	Phe	Leu																														
																340								345								350													

ES 2 698 527 T3

```

ccc gat cag atg cag ccg caa gtc ccg ccg ctc cat tac caa gag ctc      1104
Pro Asp Gln Met Gln Pro Gln Val Pro Pro Leu His Tyr Gln Glu Leu
      355                      360                      365

atg cca ccc ggt tcc tgc atg cca gag gag ccc aag cca aag agg gga      1152
Met Pro Pro Gly Ser Cys Met Pro Glu Glu Pro Lys Pro Lys Arg Gly
      370                      375                      380

aga cga tcg tgg ccc ccg aaa agg acc gcc acc cac act tgt gat tac      1200
Arg Arg Ser Trp Pro Arg Lys Arg Thr Ala Thr His Thr Cys Asp Tyr
      385                      390                      395                      400

gcg ggc tgc ggc aaa acc tac aca aag agt tcc cat ctc aag gca cac      1248
Ala Gly Cys Gly Lys Thr Tyr Thr Lys Ser Ser His Leu Lys Ala His
      405                      410                      415

ctg cga acc cac aca ggt gag aaa cct tac cac tgt gac tgg gac ggc      1296
Leu Arg Thr His Thr Gly Glu Lys Pro Tyr His Cys Asp Trp Asp Gly
      420                      425                      430

tgt gga tgg aaa ttc gcc cgc tca gat gaa ctg acc agg cac tac cgt      1344
Cys Gly Trp Lys Phe Ala Arg Ser Asp Glu Leu Thr Arg His Tyr Arg
      435                      440                      445

aaa cac acg ggg cac cgc ccg ttc cag tgc caa aaa tgc gac cga gca      1392
Lys His Thr Gly His Arg Pro Phe Gln Cys Gln Lys Cys Asp Arg Ala
      450                      455                      460

ttt tcc agg tcg gac cac ctc gcc tta cac atg aag agg cat ttt taa      1440
Phe Ser Arg Ser Asp His Leu Ala Leu His Met Lys Arg His Phe
      465                      470                      475

```

<210> 8

<211> 479

<212> PRT

<213> Homo sapiens

<400> 8

```

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

```

ES 2 698 527 T3

```

165      170      175
Ser Ala Pro Pro Pro Thr Ala Pro Phe Asn Leu Ala Asp Ile Asn Asp
180      185      190
Val Ser Pro Ser Gly Gly Phe Val Ala Glu Leu Leu Arg Pro Glu Leu
195      200      205
Asp Pro Val Tyr Ile Pro Pro Gln Gln Pro Gln Pro Pro Gly Gly Gly
210      215      220
Leu Met Gly Lys Phe Val Leu Lys Ala Ser Leu Ser Ala Pro Gly Ser
225      230      235      240
Glu Tyr Gly Ser Pro Ser Val Ile Ser Val Ser Lys Gly Ser Pro Asp
245      250      255
Gly Ser His Pro Val Val Val Ala Pro Tyr Asn Gly Gly Pro Pro Arg
260      265      270
Thr Cys Pro Lys Ile Lys Gln Glu Ala Val Ser Ser Cys Thr His Leu
275      280      285
Gly Ala Gly Pro Pro Leu Ser Asn Gly His Arg Pro Ala Ala His Asp
290      295      300
Phe Pro Leu Gly Arg Gln Leu Pro Ser Arg Thr Thr Pro Thr Leu Gly
305      310      315      320
Leu Glu Glu Val Leu Ser Ser Arg Asp Cys His Pro Ala Leu Pro Leu
325      330      335
Pro Pro Gly Phe His Pro His Pro Gly Pro Asn Tyr Pro Ser Phe Leu
340      345      350
Pro Asp Gln Met Gln Pro Gln Val Pro Pro Leu His Tyr Gln Glu Leu
355      360      365
Met Pro Pro Gly Ser Cys Met Pro Glu Glu Pro Lys Pro Lys Arg Gly
370      375      380
Arg Arg Ser Trp Pro Arg Lys Arg Thr Ala Thr His Thr Cys Asp Tyr
385      390      395      400
Ala Gly Cys Gly Lys Thr Tyr Thr Lys Ser Ser His Leu Lys Ala His
405      410      415
Leu Arg Thr His Thr Gly Glu Lys Pro Tyr His Cys Asp Trp Asp Gly
420      425      430
Cys Gly Trp Lys Phe Ala Arg Ser Asp Glu Leu Thr Arg His Tyr Arg
435      440      445
Lys His Thr Gly His Arg Pro Phe Gln Cys Gln Lys Cys Asp Arg Ala
450      455      460
Phe Ser Arg Ser Asp His Leu Ala Leu His Met Lys Arg His Phe
465      470      475

```

<210> 9

<211> 1323

<212> ADN

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1323)

<400> 9

```

atg ggc ccc ctc aac gtt agc ttc acc aac agg aac tat gac ctc gac      48
Met Gly Pro Leu Asn Val Ser Phe Thr Asn Arg Asn Tyr Asp Leu Asp
1      5      10      15

tac gac tcg gtg cag ccg tat ttc tac tgc gac gag gag gag aac ttc      96
Tyr Asp Ser Val Gln Pro Tyr Phe Tyr Cys Asp Glu Glu Glu Asn Phe
20      25      30

tac cag cag cag cag cag agc gag ctg cag ccc ccg gcg ccc agc gag      144
Tyr Gln Gln Gln Gln Ser Glu Leu Gln Pro Pro Ala Pro Ser Glu
35      40      45

```

ES 2 698 527 T3

gat atc tgg aag aaa ttc gag ctg ctg ccc acc ccg ccc ctg tcc cct Asp Ile Trp Lys Lys Phe Glu Leu Leu Pro Thr Pro Pro Leu Ser Pro 50 55 60	192
agc cgc cgc tcc ggg ctc tgc tgc ccc tcc tac gtt gcg gtc aca ccc Ser Arg Arg Ser Gly Leu Cys Ser Pro Ser Tyr Val Ala Val Thr Pro 65 70 75 80	240
ttc tcc ctt cgg gga gac aac gac ggc ggt ggc ggg agc ttc tcc acg Phe Ser Leu Arg Gly Asp Asn Asp Gly Gly Gly Ser Phe Ser Thr 85 90 95	288
gcc gac cag ctg gag atg gtg acc gag ctg ctg gga gga gac atg gtg Ala Asp Gln Leu Glu Met Val Thr Glu Leu Leu Gly Gly Asp Met Val 100 105 110	336
aac cag agt ttc atc tgc gac ccg gac gac gag acc ttc atc aaa aac Asn Gln Ser Phe Ile Cys Asp Pro Asp Asp Glu Thr Phe Ile Lys Asn 115 120 125	384
atc atc atc cag gac tgt atg tgg agc ggc ttc tgc gcc gcc gcc aag Ile Ile Ile Gln Asp Cys Met Trp Ser Gly Phe Ser Ala Ala Ala Lys 130 135 140	432
ctc gtc tca gag aag ctg gcc tcc tac cag gct gcg cgc aaa gac agc Leu Val Ser Glu Lys Leu Ala Ser Tyr Gln Ala Ala Arg Lys Asp Ser 145 150 155 160	480
ggc agc ccg aac ccc gcc cgc ggc cac agc gtc tgc tcc acc tcc agc Gly Ser Pro Asn Pro Ala Arg Gly His Ser Val Cys Ser Thr Ser Ser 165 170 175	528
ttg tac ctg cag gat ctg agc gcc gcc gcc tca gag tgc atc gac ccc Leu Tyr Leu Leu Asp Leu Ser Ala Ala Ser Glu Cys Ile Asp Pro 180 185 190	576
tgc gtg gtc ttc ccc tac cct ctc aac gac agc agc tgc ccc aag tcc Ser Val Val Phe Pro Tyr Pro Leu Asn Asp Ser Ser Ser Pro Lys Ser 195 200 205	624
tgc gcc tgc caa gac tcc agc gcc ttc tct ccg tcc tgc gat tct ctg Cys Ala Ser Gln Asp Ser Ser Ala Phe Ser Pro Ser Ser Asp Ser Leu 210 215 220	672
ctc tcc tgc acg gag tcc tcc ccg cag ggc agc ccc gag ccc ctg gtg Leu Ser Ser Thr Glu Ser Ser Pro Gln Gly Ser Pro Glu Pro Leu Val 225 230 235 240	720
ctc cat gag gag aca ccg ccc acc acc agc agc gac tct gag gag gaa Leu His Glu Glu Thr Pro Pro Thr Thr Ser Ser Asp Ser Glu Glu Glu 245 250 255	768
caa gaa gat gag gaa gaa atc gat gtt gtt tct gtg gaa aag agg cag Gln Glu Asp Glu Glu Glu Ile Asp Val Val Ser Val Glu Lys Arg Gln 260 265 270	816
gct cct ggc aaa agg tca gag tct gga tca cct tct gct gga ggc cac Ala Pro Gly Lys Arg Ser Glu Ser Gly Ser Pro Ser Ala Gly Gly His 275 280 285	864
agc aaa cct cct cac agc cca ctg gtc ctc aag agg tgc cac gtc tcc	912

ES 2 698 527 T3

Ser	Lys	Pro	Pro	His	Ser	Pro	Leu	Val	Leu	Lys	Arg	Cys	His	Val	Ser		
290						295					300						
aca	cat	cag	cac	aac	tac	gca	gcg	cct	ccc	tcc	act	cgg	aag	gac	tat		960
Thr	His	Gln	His	Asn	Tyr	Ala	Ala	Pro	Pro	Ser	Thr	Arg	Lys	Asp	Tyr		
305					310					315					320		
cct	gct	gcc	aag	agg	gtc	aag	ttg	gac	agt	gtc	aga	gtc	ctg	aga	cag		1008
Pro	Ala	Ala	Lys	Arg	Val	Lys	Leu	Asp	Ser	Val	Arg	Val	Leu	Arg	Gln		
				325					330					335			
atc	agc	aac	aac	cga	aaa	tgc	acc	agc	ccc	agg	tcc	tcg	gac	acc	gag		1056
Ile	Ser	Asn	Val	Arg	Lys	Cys	Thr	Ser	Pro	Arg	Ser	Ser	Asp	Thr	Glu		
				340					345					350			
gag	aat	gtc	aag	agg	cga	aca	cac	aac	gtc	ttg	gag	cgc	cag	agg	agg		1104
Glu	Asn	Val	Lys	Arg	Arg	Thr	His	Asn	Val	Leu	Glu	Arg	Gln	Arg	Arg		
			355					360					365				
aac	gag	cta	aaa	cgg	agc	ttt	ttt	gcc	ctg	cgt	gac	cag	atc	ccg	gag		1152
Asn	Glu	Leu	Lys	Arg	Ser	Phe	Phe	Ala	Leu	Arg	Asp	Gln	Ile	Pro	Glu		
	370					375					380						
ttg	gaa	aac	aat	gaa	aag	gcc	ccc	aag	gta	gtt	atc	ctt	aaa	aaa	gcc		1200
Leu	Glu	Asn	Asn	Glu	Lys	Ala	Pro	Lys	Val	Val	Ile	Leu	Lys	Lys	Ala		
	385				390						395				400		
aca	gca	tac	atc	ctg	tcc	gtc	caa	gca	gag	gag	caa	aag	ctc	att	tct		1248
Thr	Ala	Tyr	Ile	Leu	Ser	Val	Gln	Ala	Glu	Glu	Gln	Lys	Leu	Ile	Ser		
				405					410					415			
gaa	gag	gac	ttg	ttg	cgg	aaa	cga	cga	gaa	cag	ttg	aaa	cac	aaa	ctt		1296
Glu	Glu	Asp	Leu	Leu	Arg	Lys	Arg	Arg	Glu	Gln	Leu	Lys	His	Lys	Leu		
			420					425					430				
gaa	cag	cta	cgg	aac	tct	tgt	gcg	taa									1323
Glu	Gln	Leu	Arg	Asn	Ser	Cys	Ala										
			435				440										

<210> 10
 <211> 440
 <212> PRT
 <213> Homo sapiens

<400> 10
 Met Gly Pro Leu Asn Val Ser Phe Thr Asn Arg Asn Tyr Asp Leu Asp
 1 5 10 15
 Tyr Asp Ser Val Gln Pro Tyr Phe Tyr Cys Asp Glu Glu Glu Asn Phe
 20 25 30
 Tyr Gln Gln Gln Gln Ser Glu Leu Gln Pro Pro Ala Pro Ser Glu
 35 40 45
 Asp Ile Trp Lys Lys Phe Glu Leu Leu Pro Thr Pro Pro Leu Ser Pro
 50 55 60
 Ser Arg Arg Ser Gly Leu Cys Ser Pro Ser Tyr Val Ala Val Thr Pro
 65 70 75 80
 Phe Ser Leu Arg Gly Asp Asn Asp Gly Gly Gly Gly Ser Phe Ser Thr
 85 90 95
 Ala Asp Gln Leu Glu Met Val Thr Glu Leu Leu Gly Gly Asp Met Val
 100 105 110
 Asn Gln Ser Phe Ile Cys Asp Pro Asp Asp Glu Thr Phe Ile Lys Asn

ES 2 698 527 T3

```

115      120      125
Ile Ile Ile Gln Asp Cys Met Trp Ser Gly Phe Ser Ala Ala Ala Lys
130      135      140
Leu Val Ser Glu Lys Leu Ala Ser Tyr Gln Ala Ala Arg Lys Asp Ser
145      150      155      160
Gly Ser Pro Asn Pro Ala Arg Gly His Ser Val Cys Ser Thr Ser Ser
165      170      175
Leu Tyr Leu Gln Asp Leu Ser Ala Ala Ala Ser Glu Cys Ile Asp Pro
180      185      190
Ser Val Val Phe Pro Tyr Pro Leu Asn Asp Ser Ser Ser Pro Lys Ser
195      200      205
Cys Ala Ser Gln Asp Ser Ser Ala Phe Ser Pro Ser Ser Asp Ser Leu
210      215      220
Leu Ser Ser Thr Glu Ser Ser Pro Gln Gly Ser Pro Glu Pro Leu Val
225      230      235      240
Leu His Glu Glu Thr Pro Pro Thr Thr Ser Ser Asp Ser Glu Glu Glu
245      250      255
Gln Glu Asp Glu Glu Glu Ile Asp Val Val Ser Val Glu Lys Arg Gln
260      265      270
Ala Pro Gly Lys Arg Ser Glu Ser Gly Ser Pro Ser Ala Gly Gly His
275      280      285
Ser Lys Pro Pro His Ser Pro Leu Val Leu Lys Arg Cys His Val Ser
290      295      300
Thr His Gln His Asn Tyr Ala Ala Pro Pro Ser Thr Arg Lys Asp Tyr
305      310      315      320
Pro Ala Ala Lys Arg Val Lys Leu Asp Ser Val Arg Val Leu Arg Gln
325      330      335
Ile Ser Asn Asn Arg Lys Cys Thr Ser Pro Arg Ser Ser Asp Thr Glu
340      345      350
Glu Asn Val Lys Arg Arg Thr His Asn Val Leu Glu Arg Gln Arg Arg
355      360      365
Asn Glu Leu Lys Arg Ser Phe Phe Ala Leu Arg Asp Gln Ile Pro Glu
370      375      380
Leu Glu Asn Asn Glu Lys Ala Pro Lys Val Val Ile Leu Lys Lys Ala
385      390      395      400
Thr Ala Tyr Ile Leu Ser Val Gln Ala Glu Glu Gln Lys Leu Ile Ser
405      410      415
Glu Glu Asp Leu Leu Arg Lys Arg Arg Glu Gln Leu Lys His Lys Leu
420      425      430
Glu Gln Leu Arg Asn Ser Cys Ala
435      440

```

<210> 11
 <211> 84
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador oligonucleótido - F2A directo

<400> 11
 aattcaccgg tgtgaaacag actttgaatt ttgaccttct caagttggcg ggagacgtgg 60
 agtccaaccc agggcccaga tcta 84

<210> 12
 <211> 84
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador oligonucleótido - F2A inverso

<400> 12
 ctagtagatc tgggccctgg gttggactcc acgtctcccg ccaacttgag aaggtcaaaa 60
 ttcaaagtct gtttcacacc ggtg 84

<210> 13
 <211> 66
 <212> ADN

	<213> Secuencia Artificial	
	<220>	
5	<223> Cebador oligonucleótido - T2A directo	
	<400> 13	
	ctagtgaggg cagaggaagt ctgctaacaat gcggtgacgt cgaggagaat cctggcccac	60
	aattgt	66
10	<210> 14	
	<211> 66	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
15	<223> Cebador oligonucleótido - T2A inverso	
	<400> 14	
	ctagacaatt gtgggccagg attctcctcg acgtcaccgc atgttagcag acttcctctg	60
	cctca	66
20	<210> 15	
	<211> 76	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
25	<223> Cebador oligonucleótido - E2A directo	
	<400> 15	
	ctagacaatg tactaactac gctttgttga aactcgttg cgatgttgaa agtaaccccg	60
	gtcctggcgc gccgc	76
30	<210> 16	
	<211> 76	
	<212> ADN	
	<213> Secuencia Artificial	
35	<220>	
	<223> Cebador oligonucleótido - E2A inverso	
	<400> 16	
	ggccgcgggc gcgccaggac cggggttact ttcaacatcg ccagcgagtt tcaacaaagc	60
	gtagttagta cattgt	76
40	<210> 17	
	<211> 52	
	<212> ADN	
	<213> Secuencia Artificial	
45	<220>	
	<223> Cebador oligonucleótido - VEE-MCS directo	
	<400> 17	
50	ctagcatatg ggccgcacct cagcatcgat ggccggcctc tagacggcc gc	52
	<210> 18	
	<211> 52	
	<212> ADN	
55	<213> Secuencia Artificial	
	<220>	
	<223> Cebador oligonucleótido - VEE-MCS inverso	

<400> 18
 ggccgcggcc gctctagagg ccggccatcg atgctgaggg cgcgccata tg 52

5 <210> 19
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> Cebador oligonucleótido - nsP2a directo

<400> 19
 caggacgatc tcattctcac 20

15 <210> 20
 <211> 21
 <212> ADN
 <213> Secuencia Artificial

20 <220>
 <223> Cebador oligonucleótido - nsP2a inverso

<400> 20
 gcttgccact cctctatcgt g 21

25 <210> 21
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

30 <220>
 <223> Cebador oligonucleótido - nsP4a directo

<400> 21
 ccacaatacg atcggcagtg 20

35 <210> 22
 <211> 21
 <212> ADN
 <213> Secuencia Artificial

40 <220>
 <223> Cebador oligonucleótido - nsP4a inverso

45 <400> 22
 atgtcctgca acatattcaa a 21

50 <210> 23
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

55 <220>
 <223> Cebador oligonucleótido - Oct4RTa directo

<400> 23
 cggcgccaga agggcaagcg 20

60 <210> 24
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

65 <220>
 <223> Cebador oligonucleótido - Klf4RTb

<400> 24
 cacctgcttg acgcagtgtc 20
 5 <210> 25
 <211> 25
 <212> ADN
 <213> Secuencia Artificial
 10 <220>
 <223> Cebador oligonucleótido - Bis-Oct-10F
 <400> 25
 ggagtagaag gattgttttg gttta 25
 15 <210> 26
 <211> 25
 <212> ADN
 <213> Secuencia Artificial
 20 <220>
 <223> Cebador oligonucleótido - Bis-Oct-9R
 <400> 26
 25 aaacctaaa aacttaacca aatcc 25
 <210> 27
 <211> 26
 <212> ADN
 30 <213> Secuencia Artificial
 <220>
 <223> Cebador oligonucleótido - Bis-Nanog-4F
 35 <400> 27
 agagtagttg ggattataga tattta 26
 <210> 28
 <211> 24
 40 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Cebador oligonucleótido - Bis-Nanog-3R
 45 <400> 28
 aacaacaaaa ctaaaaaca aacc 24
 <210> 29
 50 <211> 16337
 <212> ADN
 <213> Secuencia Artificial
 <220>
 55 <223> VEE-Oct-Klf-Sox-cMyc
 <400> 29

ES 2 698 527 T3

atgggcgggcg catgagagaa gccagacca attacctacc caaaatggag aaagttcacg	60
ttgacatcga ggaagacagc ccattcctca gagctttgca gcgagcttc ccgcagtttg	120
aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcac	180
tggcttcaaa actgatcgaa acggaggtg acccatccga cacgatcctt gacattggaa	240
gtgcgcccgc ccgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat	300
gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg	360
aaataactga taaggaattg gacaagaaaa tgaaggagct ggccgcccgc atgagcgacc	420
ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc	480
aagtcgtgtg ttaccaggat gtatacgcgg ttgacggacc gacaagtctc tatcaccaag	540
ccaataaggg agttagagtc gcctactgga taggctttga caccaccctt tttatgttta	600
agaacttggc tggagcatat ccatacact ctaccaactg ggccgacgaa accgtgttaa	660
cggctcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggtcga	780
ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttact	840
tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgtgcta	960
cgatgcaccg cgagggattc ttgtgctgca aagtacaga cacattgaac ggggagaggg	1020
tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac	1080
tggcaacaga tgtcagtgcg gacgacgcg aaaaactgct ggttgggctc aaccagcgta	1140
tagtcgtcaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccg	1200
tagtggccca ggcatttgct aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320

ES 2 698 527 T3

acaagataac atctatztat aagcgcccg ataccctaac catcatcaaa gtgaacagcg 1380
 atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa 1440
 caagaatcag gaaaatgtta gaggagcaca aggagccgctc acctctcatt accgccgag 1500
 acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgctgtaa gccgaggagt 1560
 tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg 1620
 tcgacttgat gttacaagag gctggggccg gctcagtga gacacctcgt ggcttgataa 1680
 aggttaccag ctacgatggc gaggacaaga tcggctctta cgctgtgctt tctccgag 1740
 ctgtactcaa gaggtaaaaa ttatcttgca tccacctct cgctgaacaa gtcatagtga 1800
 taacacactc tggccgaaaa gggcggttatg ccgtggaacc ataccatggt aaagtagtgg 1860
 tgccagaggg acatgaata cccgtccag actttcaagc tctgagtga agtgccacca 1920
 ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag 1980
 gagcgtgaa cactgatgaa gaattattaca aaactgtcaa gccagcgag cagcagggcg 2040
 aatacctgta cgacatcgac aggaacagc gcgtcaagaa agaactagtc actgggctag 2100
 ggctcacagg cgagctggtg gatcctcct tccatgaatt cgcctacgag agtctgagaa 2160
 cagcaccagc cgctccttac caagtacaa ccataggggt gtatggcgtg ccaggatcag 2220
 gcaagtctgg catcattaaa agcgcagtca ccaaaaaaga tctagtgggtg agcgccaaga 2280
 aagaaaactg tgcagaaatt ataaggagc tcaagaaaat gaaagggtg gacgtcaatg 2340
 ccagaactgt ggactcagt ctcttgatg gatgcaacaa ccccgtagag accctgtata 2400
 ttgacgaagc ttttgcctgt catgcaggta ctctcagagc gctcatagcc attataagac 2460
 ctaaaaaggc agtgctctgc ggggatccca aacagtgcg tttttttaac atgatgtgcc 2520
 tgaaagtga ttttaaccac gagatttgca cacaagtctt ccacaaaagc atctctcgcc 2580
 gttgcactaa atctgtgact tcggtcgtct caacctgtt ttacgacaaa aaaatgagaa 2640
 cgacgaatcc gaaagagact aagattgtga ttgacactac cggcagtacc aaacctaagc 2700
 aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca 2760
 aaggcaacga aataatgacg gcagctgcct ctcaagggt gaccgtaaa ggtgtgtatg 2820
 ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg 2880
 tcctactgac ccgcacggag gaccgcatc gtgggaaaac actagccggc gaccatgga 2940
 taaaaaact gactgccaag taccctggga atttactgc cagatagag gagggtgcaag 3000
 cagagcatga tgccatcatg aggcacatct tggagagacc ggacctacc gacgtcttcc 3060
 agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca 3120
 tagacatgac cactgaacaa tggaactcgt tggattatct tgaaacggac aaagctcact 3180

ES 2 698 527 T3

cagcagagat agtattgaac caactatgcg tgaggttctt tggactcgat ctggactccg	3240
gtctattttc tgcaccact gttccgttat ccattaggaa taatcactgg gataactccc	3300
cgtcgcctaa catgtacggg ctgaataaag aagtgggccg tcagctctct cgcaggtacc	3360
cacaactgcc tcgggcagtt gccactggaa gagtctatga catgaacact ggtacactgc	3420
gcaattatga tccgcgcata aacctagtag ctgtaaacag aagactgcct catgctttag	3480
tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcagc aaattgaagg	3540
gcagaactgt cctgggtggc ggggaaaagt tgcccgcccc aggcaaatg gttgactggg	3600
tgtcagaccg gcctgaggct accttcagag ctccgctgga tttaggcatc ccaggtgatg	3660
tgcccaaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcac	3780
tgaatcccg cggaacctgt gtcagcatag gttatgggta cgctgacagg gccagcgaaa	3840
gcatcattgg tgctatagcg cggcagttca agttttcccg ggtatgcaaa ccgaaatcct	3900
cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
acaattctta caagctttca tcaaccttga ccaacattta tacaggttcc agactccacg	4020
aagccggatg tgcacctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
gagtgtattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgccgagcgc	4140
tgtataagaa attcccgga agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
tggtcaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc aacaaagttt	4260
cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgct aagattgtca	4320
acgataacaa ttacaagtca gtacgattc cactgttgtc caccggcatc tttccggga	4380
acaaagatcg actaacccaa tcattgaacc atttgctgac agcttttagac accactgatg	4440
cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
atgcagagct ggtgaggggt catccgaaga gttctttggc tggaaggaag ggctacagca	4620
caagcgtatg caaaactttc tcatatattg aagggaacca gtttcaccag gcggccaagg	4680
atatagcaga aattaatgcc atgtggcccc ttgcaacgga ggccaatgag caggtatgca	4740
tgtatatcct cggagaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
aagcctccac accacctagc acgctgcctt gcttgtgcat ccatgccatg actccagaaa	4860
gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980

ES 2 698 527 T3

caccgaaagt gctcgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag 5040
 acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac 5100
 cacttataac cgaggatgag accaggacta gaacgcctga gccgatcatc atcgaagagg 5160
 aagaagagga tagcataagt ttgtgtgcag atggcccgcac ccaccagggtg ctgcaagtcg 5220
 aggagacat tcacggggccg ccctctgtat ctagctcatc ctggtccatt cctcatgcat 5280
 ccgactttga tgtggacagt ttatccatac ttgacaccct ggagggagct agcgtgacca 5340
 gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc 5400
 gaccgggtgcc tgcgcctcga acagtattca ggaacctcc acatcccgtc ccgcgcacaa 5460
 gaacaccgtc acttgcaccc agcagggcct gctcgagaac cagcctagtt tccaccccgc 5520
 caggcgtgaa taggggtgatc actagagagg agctcgaggc gcttaccctg tcacgcactc 5580
 ctagcaggtc ggtctcgaga accagcctgg tctccaacc gccaggcgta aataggggtga 5640
 ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg 5700
 catacatctt ttctccgac accgggtcaag ggcatttaca acaaaaatca gtaaggcaaa 5760
 cgggtgctatc cgaagtgggtg ttggagagga ccgaattgga gatttcgtat gccccgcgcc 5820
 tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaatccc acacctgcta 5880
 acagaagcag ataccagtcc aggaaggtgg agaacatgaa agccataaca gctagacgta 5940
 ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacct 6000
 tgcacacctg tctttgtat tcacttagtg tgaaccgtgc cttttcaagc cccaaggctg 6060
 cagtggaaag ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta 6120
 ttattccaga gtacgatgcc tatttgaca tggttgacgg agcttcatgc tgcttagaca 6180
 ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaac 6240
 ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtcttggcag 6300
 ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgtg ttggattcgg 6360
 cggcctttaa tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaacgt 6420
 ttaaagaaaa ccccatcagg cttactgaag aaaacgtggg aaattacatt accaaattaa 6480
 aaggacaaaa agctgctgct ctttttgca agacacataa tttgaatatg ttgcaggaca 6540
 taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa 6600
 aacatactga agaacggccc aaggtacagg tgatccaggc tgccgatccg ctagcaacag 6660
 cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga 6720
 acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact 6780
 tccagcctgg ggattgtgtt ctggaaactg acatcgctc gtttgataaa agtgaggacg 6840

ES 2 698 527 T3

acgccatggc tctgaccgcg ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
tgacgctgat tgaggcggct ttcggcgaaa tttcatcaat acatttgccc actaaaacta	6960
aatttaaatt cggagccatg atgaaatctg gaatgttcct cacactgttt gtgaacacag	7020
tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcaccatgtg	7080
cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
acaggtgctg cacctgggtg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
aagcgcccta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
gtgtggcaga cccctaaaa aggtgtttaa agcttggcaa acctctggca gcagacgatg	7320
aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaaccgta ggaacttcca	7440
tcatagttaa ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
gggcccctat aactctctac ggctaacctg aatggactac gacatagtct agtccgcaa	7560
gtctagcata tgggcgcgtg aattcgccac catggcggga cacctggctt cggatttcgc	7620
cttctgccc cctccaggtg gtggaggta tgggccaggg gggccggagc cgggctgggt	7680
tgatcctcgg acctggctaa gcttccaagg ccctcctgga gggccaggaa tcgggccggg	7740
ggttgggcca ggctctgagg tgtgggggat tccccatgc ccccgccgt atgagttctg	7800
tggggggatg gcgtactgtg ggccccaggt tggagtgggg ctagtcccc aaggcggctt	7860
ggagacctct cagcctgagg gcgaagcagg agtcgggggt gagagcaact ccgatggggc	7920
ctccccggag ccctgcaccg tcaccctcgg tgccgtgaag ctggagaagg agaagctgga	7980
gcaaaacccg gaggagtccc aggcacatcaa agctctgcag aaagaactcg agcaatttgc	8040
caagctcctg aagcagaaga ggatcacctt gggatataca caggccgatg tggggctcac	8100
cctgggggtt ctatttggga aggtattcag ccaaacgacc atctgccgct ttgaggctct	8160
gcagcttagc ttcaagaaca tgtgtaagct gcggcccttg ctgcagaagt ggggtggagga	8220
agctgacaac aatgaaaatc ttcaggagat atgcaaagca gaaaccctcg tgcaggcccc	8280
aaagagaaag cgaaccagta tcgagaaccg agtgagaggc aacctggaga atttgttcct	8340
gcagtgcccc aaaccacac tgcagcagat cagccacatc gccagcagc ttgggctcga	8400
gaaggatgtg gtccgagtgt ggttctgtaa ccgcgccag aagggaagc gatcaagcag	8460
cgactatgca caacgagagg attttgaggc tgctgggtct ctttctcag ggggaccagt	8520
gtcctttcct ctggccccag ggccccattt tggtaccca ggctatggga gccctcactt	8580
cactgcactg tactcctcgg tccctttccc tgagggggaa gcctttcccc ctgtctccgt	8640

ES 2 698 527 T3

caccactctg ggctctccca tgcattcaaa ctctagttag ggcagaggaa gtctgctaac	8700
atgcggtgac gtcgaggaga atcctggccc acaattgatg gctgtcagcg acgcgctgct	8760
cccatctttc tccacgttcg cgtctggccc ggcggaagg gagaagacac tgcgtcaagc	8820
aggtgccccg aataaccgct ggcgggagga gctctccac atgaagcgac ttccccagc	8880
gcttcccggc cgccctatg acctggcggc ggcgaccgtg gccacagacc tggagagcgg	8940
cggagccggt gcggttgcg gcggtagcaa cctggcgccc ctacctcgga gagagaccga	9000
ggagttcaac gatctcctgg acctggactt tattctctcc aattcgctga cccatctcc	9060
ggagtcagt gcccaccacg tgtctcgtc agcgtcagcc tcctcttcgt cgtcgccgtc	9120
gagcagcggc cctgccagcg cgccctccac ctgcagcttc acctatccga tccgggccc	9180
gaacgaccg ggcggtggcg cgggcgccac gggcgaggc ctctctatg gcagggagtc	9240
cgctccccct ccgacggctc cttcaacct ggcgacatc aacgacgtga gcccctcggg	9300
cggcttcgtg gccgagctcc tgcggccaga attggaccg gtgtacattc cgccgcagca	9360
gccgcagccg ccaggtggcg ggctgatggg caagttcgtg ctgaaggcgt cgctgagcgc	9420
ccctggcagc gagtacggca gcccgtcggc catcagcgtc agcaaaggca gccctgacgg	9480
cagccaccg gtggtggtgg cgccctacaa cggcggggcg ccgcgcacgt gcccgaagat	9540
caagcaggag gcggtctctt cgtgcaccca cttggcgct ggacccctc tcagcaatgg	9600
ccaccggccg gctgcacacg acttccccct gggcgggcag ctcccagca ggactacccc	9660
gacctgggt cttgaggaag tgcgtgagcag cagggaactgt caccctgcc tgcgcttcc	9720
tcceggcttc cateccccacc cggggccc aa ttaacctcc ttctgccc atcagatgca	9780
gccgcaagtc ccgcgctcc attaccaaga gctcatgcca cccggttct gcatgccaga	9840
ggagcccaag ccaaagagg gaagacgac gtggccccg aaaaggaccg ccaccacac	9900
ttgtgattac gcgggctgcg gcaaaccta cacaagagt tcccatctca aggcacacct	9960
gcgaaccac acaggtgaga aaccttacc ctgtgactgg gacggctgtg gatggaaatt	10020
cgcccgctca gatgaactga ccaggcacta ccgtaaacac acggggcacc gcccgttcca	10080
gtgccaaaaa tgcgaccgag cattttcag gtcggaccac ctgccttac acatgaagag	10140
gcatttttct agacaatgta ctaactacgc tttgtgaaa ctgctggcg atgtgaaag	10200
taaccccggt cctggcgcg ccatgtacaa catgatggag acggagctga agccgcggg	10260
cccgcagcaa acttcggggg gcggcgcg caactccacc gcggcgcg ccggcgcaa	10320
ccagaaaaac agcccgacc gcgtcaagcg gcccatgaat gccttcatgg tgtggtccg	10380
cgggcagcgg cgcaagatgg ccagagaga cccaagatg cacaactcgg agatcagcaa	10440
gcgcctgggc gccgagtgga aacttttgc ggagacggag aagcgccgt tcacgcagca	10500

ES 2 698 527 T3

ggctaagcgg ctgcgagcgc tgcacatgaa ggagcaccgc gattataaat accggccccg	10560
gcggaacc aagacgctca tgaagaagga taagtacacg ctgcccggcg ggctgctggc	10620
ccccggcggc aatagcatgg cgagcgggt cggggtgggc gccggcctgg gcgcgggcgt	10680
gaaccagcgc atggacagtt acgcgcacat gaacggctgg agcaacggca gctacagcat	10740
gatgcaggac cagctgggct acccgagca cccgggcctc aatgcgcacg gcgcagcgca	10800
gatgcagccc atgcaccgct acgacgtgag cgccctgcag tacaactcca tgaccagctc	10860
gcagacctac atgaacggct cgccaccta cagcatgtcc tactcgagc agggcacc	10920
tggcatggct cttggctcca tgggttcggt ggtcaagtcc gaggccagct ccagcccc	10980
tgtggttacc tcttctccc actccaggc gccctgccag gccggggacc tccgggacat	11040
gatcagcatg tatctccccg gcgcgcaggt gccgaacc gccgccccca gcagacttca	11100
catgtcccag cactaccaga gcggcccggt gcccggcacg gccattaacg gcacactgcc	11160
cctctcacac atgtgagcgg ccatcgatgt cgacaactaa cttaagctag caacggtttc	11220
cctctagcgg gatcaattcc gcccccccc cctaacgtta ctggccgaag ccgcttgaa	11280
taaggccggt gtgcgtttgt ctatatgtta tttccacca tattgccgc ttttggaat	11340
gtgagggccc ggaaacctgg ccctgtcttc ttgacgagca ttctagggg tctttccct	11400
ctgcgcaag gaatgcaagg tctgttgaat gtctgaagg aagcagttcc tctggaagct	11460
tcttgaagac aaacaacgctc tgtagcgacc ctttgcaggc agcggaacc cccacctggc	11520
gacaggtgcc tctgcggcca aaagccacgt gtataagata cacctgcaa ggcggcacaa	11580
ccccagtgcc acgttgtgag ttggatagtt gtgaaagag tcaaatggct ctctcaagc	11640
gtattcaaca aggggtgaa ggatgccag aaggtaccc attgtatgg atctgatctg	11700
gggcctcggg gcacatgctt tacatgtgtt tagtcgaggt taaaaaacg tctaggccc	11760
ccgaaccacg gggacgtggg tttctttga aaaacacgat aataccaatt cgccaccatg	11820
ggccccctca acgttagctt caccaacag aactatgacc tcgactacga ctcggtgcag	11880
ccgtatttct actgcgacga ggaggagaac ttctaccagc agcagcagca gagcgagctg	11940
cagcccccg cgccagcga ggatatctgg aagaaattcg agctgctgcc caccocgcc	12000
ctgtccccta gccgcgcgc cggtctctgc tcgccctcct acgttgcggt cacacccttc	12060
tcccttcggg gagacaacga cggcggtggc gggagcttct ccacggccga ccagctggag	12120
atggtgaccg agctgctggg aggagacatg gtgaaccaga gttcatctg cgaccggac	12180
gacgagacct tcataaaaa catcatcatc caggactgta tgtggagcgg cttctcgcc	12240
gccgccaagc tcgtctcaga gaagctggcc tcctaccag ctgcgcgcaa agacagcggc	12300

ES 2 698 527 T3

agcccggaacc ccgcccgcgg ccacagcgtc tgctccacct ccagcttgta cctgcaggat	12360
ctgagcgccg ccgcctcaga gtgcatcgac ccctcggtgg tcttccccta ccctctcaac	12420
gacagcagct cgcccaagtc ctgcgcctcg caagactcca ggcctttctc tccgtcctcg	12480
gattctctgc tctcctcgac ggagtcctcc ccgcagggca gcccgcagcc cctggtgctc	12540
catgaggaga caccgcccac caccagcagc gactctgagg aggaacaaga agatgaggaa	12600
gaaatcgatg ttgtttctgt ggaaaagagg caggctcctg gcaaaaggtc agagtctgga	12660
tcaccttctg ctggaggcca cagcaaacct cctcacagcc cactggtcct caagagggtc	12720
cacgtctcca cacatcagca caactacgca ggcctccct ccactcgga ggactatcct	12780
gctgccaaga gggtaagtt ggacagtgtc agagtctga gacagatcag caacaaccga	12840
aatgcacca gcccaggtc ctgcgacacc gaggagaatg tcaagaggcg aacacacaac	12900
gtcttgagc gccagaggag gaacgagcta aaacggagct ttttgccct gcgtgaccag	12960
atcccgagtg tggaaaacaa tgaaaaggcc cccaaggtag ttatccttaa aaaagccaca	13020
gcatacatcc tgtccgtcca agcagaggag caaaagctca tttctgaaga ggacttggtg	13080
cggaaacgac gagaacagtt gaaacacaaa ctgaacagc tacggaactc ttgtgcgtaa	13140
tctagagtcg acccgggcgg ccgcaactaa cttaagctag caacggtttc cctctagcgg	13200
gatcaattcc gcccccccc cctaacgtta ctggccgaag ccgcttgga taaggccggt	13260
gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggaat gtgagggccc	13320
ggaaaacctg ccctgtcttc ttgacgagca ttctagggg tcttcccct ctcgccaaag	13380
gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct tcttgaagac	13440
aaacaacgct tgtagcgacc ctttgacggc agcggaaccc cccacctggc gacagggtgc	13500
tctgcgcca aaagccacgt gtataagata cacctgcaaa ggcggcacaa cccagtgcc	13560
acgttgtag ttggatagtt gtggaaagag tcaaatggct ctctcaagc gtattcaaca	13620
aggggctgaa ggatgccag aaggtagccc attgtatggg atctgatctg gggcctcgg	13680
gcacatgctt tacatgtgtt tagtcgaggt taaaaaacg tctaggcccc ccgaaccacg	13740
gggacgtggt tttcctttga aaaacacgat aataccatga ccgagtacaa gccacgggtg	13800
cgctcgcca ccgcgcgacg cgtccccagg gccgtacgca ccctcgccg cggttcgcc	13860
gactaccccg ccacgcgcca caccgtcgat ccggaccgcc acatcgagcg ggtcaccgag	13920
ctgcaagaac tttcctcac gcgcgtcggg ctgcacatcg gcaagggtgtg ggtcgggac	13980
gacggcgccg cggtagcggt ctggaccacg ccggagagcg tcgaagcggg ggcggtgttc	14040
gccgagatcg gcccgcgcat ggccgagttg agcggttccc ggctggccgc gcagcaacag	14100
atggaaggcc tcttgcgcc gcaccggccc aaggagcccg cgtggttctt ggccaccgtc	14160

ES 2 698 527 T3

ggcgctctcgc ccgaccacca gggcaaggggt ctgggcagcgc ccgtcgtgct ccccgagtg 14220
 gagggcgccg agcgcgcccg ggtgcccgcc ttctggaga cctccgcgcc ccgcaacctc 14280
 ccttctacg agcggtcgcg cttcacgcgc accgcccagc tcgaggtgcc cgaaggaccg 14340
 cgcacctggt gcatgaccgc caagcccgtt gcctgagaat tggcaagctg cttacataga 14400
 actcgcggcg attggcatgc cgccttaaaa tttttatttt atttttcttt tcttttccga 14460
 atcggtatttt gtttttaata tttcaaaaaa aaaaaaaaaa aaaaaaaaaa cgcgtcgagg 14520
 ggaattaatt cttgaagacg aaagggccag gtggcacttt tcggggaaat gtgcgcggaa 14580
 cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac 14640
 cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg 14700
 tcgcccttat tccctttttt gcgcattttt gccttcctgt ttttgctcac ccagaaacgc 14760
 tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg 14820
 atctcaacag cggtaagatc cttgagagtt ttcccccga agaacgtttt ccaatgatga 14880
 gcacttttaa agttctgcta tgtggcgcgg tattatcccg tgttgacgcc gggcaagagc 14940
 aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag 15000
 aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga 15060
 gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg 15120
 cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga 15180
 atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt 15240
 tgcgcaaact attaaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 15300
 ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt 15360
 ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg 15420
 ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 15480
 tggatgaacg aaatagacag atcgtcgaga taggtgcctc actgattaag cattggtaac 15540
 tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta 15600
 aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt 15660
 tttcgttcca ctgagcgtca gaccocgtag aaaagatcaa aggatcttct tgagatcctt 15720
 tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt 15780
 gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc 15840
 agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 15900
 tagcaccgcc tacatacctc gctctgctaa tctgttacc agtggctgct gccagtggcg 15960
 ataagtcgtg tcttaccggg ttggaactca gacgatagtt accggataag gcgcagcggg 16020
 cgggctgaac ggggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac 16080
 tgagatacct acagcgtgag cattgagaaa gcgccacgct tcccgaaggg agaaaggcgg 16140
 acaggatatcc ggtaagcggc agggtcggaa caggagagcg cagagggag cttccagggg 16200
 gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat 16260
 ttttgtgatg ctgctcaggg gggcggagcc tatggaaaaa cgccagcaac gcgagctcga 16320
 tttaggtgac actatag 16337

<210> 30
 <211> 16336
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> promotor VEE-Oct-Klf-Sox-cMyc-T7

<400> 30
 atgggcggcg catgagagaa gcccagacca attacctacc caaaatggag aaagttcacg 60
 ttgacatcga ggaagacagc ccattcctca gagctttgca gcgagcttc cgcagtttg 120
 aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcac 180
 tggcttcaaa actgatcgaa acggagggtg acccatccga cagatcctt gacattggaa 240
 gtgcgccgc cgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat 300
 gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg 360
 aaataactga taaggaattg gacaagaaaa tgaaggagct ggccgccgc atgagcgacc 420
 ctgacctgga aactgagact atgtgectcc acgacgacga gtcgtgtcgc tacgaagggc 480
 aagtcgtgt ttaccaggat gtatacgcgg ttgacggacc gacaagtctc tatcaccaag 540
 ccaataaggg agttagagtc gcctactgga taggctttga caccaccct tttatgttta 600
 agaacttggc tggagcatat ccatcatact ctaccaactg ggccgacgaa accgtgttaa 660
 cggctcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt 720
 ccattcttag aaagaagtat ttgaaacct ccaacaatgt tctattctct gttggctcga 780
 ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttact 840
 tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg 900
 tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta 960
 cgatgcaccg cgagggattc ttgtgctgca aagtgcaga cacattgaac ggggagaggg 1020
 tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac 1080
 tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta 1140

10

ES 2 698 527 T3

tagtcgtcaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccg	1200
tagtgggcca ggcatttgct aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320
acaagataac atctatztat aagcgcccgg atacccaaac catcatcaaa gtgaacagcg	1380
atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa	1440
caagaatcag gaaaatgtta gaggagcaca aggagccgtc acctctcatt accgccgag	1500
acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
tgcgcgcagc tctaccacct ttggcagctg atgttgagga gcccactctg gaagccgatg	1620
tcgacttgat gttacaagag gctggggccg gctcagtgga gacacctcgt ggcttgataa	1680
aggttaccag ctacgatggc gaggacaaga tcggtcttta cgctgtgctt tctccgcag	1740
ctgtactcaa gagtgaaaaa ttatcttgca tccacctctc cgctgaacaa gtcatagtga	1800
taacacactc tggccgaaaa gggcgttatg ccgtggaacc ataccatggt aaagtagtgg	1860
tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtga agtgccacca	1920
ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
gagcgtgaa cactgatgaa gaattattaca aaactgtcaa gccagcgag cacgacggcg	2040
aatacctgta cgacatcgac aggaacacgt gcgtcaagaa agaactagtc actgggctag	2100
ggctcacagg cgagctggtg gatcctccct tccatgaatt cgctacgag agtctgagaa	2160
cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
gcaagtctgg catcattaaa agcgcagtc ccaaaaaaga tctagtgggt agcgccaaga	2280
aagaaaactg tgcagaaatt ataagggacg tcaagaaaat gaaagggctg gacgtcaatg	2340
ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
ttgacgaagc ttttgcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
ctaaaaggc agtgctctgc ggggatccca aacagtgcgg tttttttaac atgatgtgcc	2520
tgaaagtgca ttttaaccac gagatttgca cacaagtctt ccacaaaagc atctctcgcc	2580
gttgactaa atctgtgact tcggtcgtct caaccttggt ttacgacaaa aaaatgagaa	2640
cgacgaatcc gaaagagact aagattgtga ttgacactac cggcagtacc aaacctaagc	2700
aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
aaggcaacga aataatgacg gcagctgcct ctcaagggct gacctgtaa ggtgtgtatg	2820
ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
tcctactgac ccgcacggag gaccgcacg tgtggaaaac actagccggc gacccatgga	2940
taaaaaact gactgccaaag taccctggga atttactgc cacgatagag gagtggcaag	3000

ES 2 698 527 T3

cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gacgtcttcc	3060
agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
tagacatgac cactgaacaa tggaaactg tggattatct tgaacaggac aaagctcact	3180
cagcagagat agtattgaac caactatgcy tgaggttctt tggactcgat ctggactccg	3240
gtctattttc tgcaccact gttccgttat ccattaggaa taatcactgg gataactccc	3300
cgtcgcctaa catgtacggg ctgaataaag aagtgggtccg tcagctctct cgcagggtacc	3360
cacaactgcc tcgggcagtt gccactggaa gagtctatga catgaacact ggtacactgc	3420
gcaattatga tccgcgcata aacctagtac ctgtaaacag aagactgcct catgctttag	3480
tcctccacca taatgaacac ccacagagt acttttcttc attcgtcagc aaattgaagg	3540
gcagaactgt cctggtgggc ggggaaaagt tgtccgtccc aggcaaatg gttgactggt	3600
tgtcagaccg gcctgaggct acctcagag ctcggtgga tttaggcatc ccagggtgatg	3660
tgcccaaata tgacataata ttgttaatg tgaggacccc atataaatac catcactatc	3720
agcagtgtga agacatgcc attaagctta gcatgttgac caagaaagct tgtctgcac	3780
tgaatcccg cggaacctgt gtcagcatag gttatgggta cgctgacagg gccagcgaaa	3840
gcatcattgg tgctatagcg cggcagttca agttttcccg ggtatgcaaa ccgaaatcct	3900
cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
acaattctta caagctttca tcaaccttga ccaacattta tacagggttc agactccacg	4020
aagccgatg tgcacctca tatcatgtgg tgcaggggga tattgccacg gccaccgaag	4080
gagtgattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
tgtataagaa attcccgaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
tggtcaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc aacaaagttt	4260
cggaggttga aggtgacaaa cagttggcag aggttatga gtccatcgct aagattgtca	4320
acgataacaa ttacaagtca gtacgattc cactgttgtc caccggcatc tttccggga	4380
acaaagatcg actaaccctaa tcattgaacc atttctgac agcttttagac accactgatg	4440
cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagt acagaacctg	4560
atgcagagct ggtgaggggtg catccgaaga gttctttggc tggaaaggaag ggctacagca	4620
caagcgatgg caaaactttc tcataatttg aagggaacaa gtttcaccag gcggccaagg	4680
atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
tgtatatacct cgagaaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800

ES 2 698 527 T3

aagcctccac accacctagc acgtgcctt gcttgtgcat ccatgccatg actccagaaa	4860
gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
caccgaaagt gcctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag	5040
acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
cacttataac cgaggatgag accaggacta gaacgcctga gccgatcatc atcgaagagg	5160
aagaagagga tagcataagt ttgtgtcag atggcccgac ccaccagggtg ctgcaagtcg	5220
aggcagacat tcacgggccc cctctgtat ctagctcatc ctggtocatt cctcatgcat	5280
ccgactttga tgtggacagt ttatccatac ttgacaccct ggagggagct agcgtgacca	5340
gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc	5400
gaccggtgcc tgcgcctcga acagtattca ggaacctcc acatcccgt cgcgcacaa	5460
gaacaccgtc acttgcaccc agcagggcct gtcgagaac cagcctagtt tccacccgc	5520
caggcgtgaa tagggtgatc actagagagg agctcgaggc gcttaccctg tcacgcactc	5580
ctagcaggtc ggtctcgaga accagcctgg tctccaaccc gccaggcgta aatagggtga	5640
ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg	5700
catacatctt ttctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760
cgggtgctatc cgaagtgggtg ttggagagga ccgaattgga gatttcgtat gccccgcgc	5820
tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaattccc acacctgcta	5880
acagaagcag ataccagtcc aggaagggtg agaacatgaa agccataaca gctagacgta	5940
ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacct	6000
tgcctcctgt tcctttgtat tcacttagtg tgaaccgtgc cttttcaagc cccaaggctg	6060
cagtggaaagc ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaac	6240
ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtcctggcag	6300
ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgccgta ttggattcgg	6360
cggcctttaa tgtggaatgc ttcaagaaat atgctgtaaa taatgaatat tgggaaacgt	6420
ttaaagaaaa ccccatcagg cttactgaag aaaacgtggt aaattacatt accaaattaa	6480
aaggacaaaa agctgctgct ctttttgca agacacataa tttgaatatg ttgcaggaca	6540
taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
aacatactga agaacggccc aaggtacagg tgatccaggc tgccgatccg ctagcaacag	6660

ES 2 698 527 T3

cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
tccagcctgg ggatttgttt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
acgccatggc tctgaccgcg ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
tgacgctgat tgaggcggct ttcggcgaaa ttcatcaat acatttgccc actaaaacta	6960
aatttaaatt cggagccatg atgaaatctg gaatgttcct cacactgttt gtgaacacag	7020
tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcacatgtg	7080
cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
acaggtgcgc cacctgggtg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
gtgtggcaga cccctaataa aggtgtttaa agcttggaac acctctggca gcagacgatg	7320
aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaaccgta ggaacttcca	7440
tcatagttat ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
gggcccctat aactctctac ggctaacctg aatggactac gacatagtct agtccgcaa	7560
gtctagcata tgggcgcgtg aattcgccac catggcggga cacctggctt cggatttcgc	7620
cttctcgccc cctccagggtg gtggagggtga tgggccaggg gggccggagc cgggctgggt	7680
tgatcctcgg acctggctaa gcttccaagg cctcctgga gggccaggaa tcgggccggg	7740
ggttgggcca ggctctgagg tgtgggggat tccccatgc ccccgccgt atgagttctg	7800
tggggggatg gcgtactgtg ggccccaggt tggagtgggg ctagtcccc aaggcggctt	7860
ggagacctct cagcctgagg gcgaagcagg agtcgggggt gagagcaact ccgatggggc	7920
ctccccggag cctgcaccg tcacccctgg tgccgtgaag ctggagaagg agaagctgga	7980
gcaaaacccg gaggagtccc aggacatcaa agctctgcag aaagaactcg agcaatttgc	8040
caagctcctg aagcagaaga ggatcacctt gggatataca caggccgatg tggggctcac	8100
cctgggggtt ctatttgga aggtattcag ccaaacgacc atctgccgct ttgaggctct	8160
gcagcttagc ttcaagaaca tgtgtaagct gcggcccttg ctgcagaagt gggaggagga	8220
agctgacaac aatgaaaatc ttcaggagat atgcaaagca gaaaccctcg tgcaggcccg	8280
aaagagaaag cgaaccagta tcgagaaccg agtgagaggc aacctggaga attgttcct	8340
gcagtgcccg aaaccacac tgcagcagat cagccacatc gccagcagc ttgggctcga	8400
gaaggatgtg gtccgagtgt ggttctgtaa ccggcgccag aagggaagc gatcaagcag	8460

ES 2 698 527 T3

cgactatgca caacgagagg attttgaggc tgctgggtct cctttctcag ggggaccagt	8520
gtccttttct ctggccccag ggccccattt tggtagccca ggctatggga gccctcactt	8580
cactgcactg tactcctcgg tccctttccc tgagggggaa gcctttcccc ctgtctccgt	8640
caccactctg ggctctccca tgcattcaaa ctctagttag ggagaggaa gtctgctaac	8700
atgcggtgac gtcgaggaga atcctggccc acaattgatg gctgtcagcg acgcgtgct	8760
cccattcttc tccacgttcg cgtctggccc ggcggaagg gagaagacac tgcgtcaagc	8820
aggtgccccg aataaccgct ggcgggagga gctctccac atgaagcgac ttccccagt	8880
gcttcccgcg cgccttatg acctggcgcg ggcgaccgtg gccacagacc tggagagcgg	8940
cggagccggt gcggttgcg gcggtagcaa cctggcgccc ctacctcgga gagagaccga	9000
ggagtccaac gatctcctgg acctggactt tattctctcc aattcgctga cccatctcc	9060
ggagtcaagt gccgccaccg tgcctcgtc agcgtcagcc tcctcttcgt cgtcgccgtc	9120
gagcagcgcg cctgccagcg cgcctccac ctgcagcttc acctatccga tccggcgccg	9180
gaacgacccg ggctggcgcg cggcgggcac gggcgaggcg ctctctatg gcaggagtc	9240
cgtccccct cgcagcgctc cctcaacct ggcgacatc aacgagctga gccctcggg	9300
cggcttcgtg gccgagctcc tgcggccaga attggaccgg gtgtacattc cgcgcagca	9360
gccgcagccg ccagggtggcg ggctgatggg caagtctgtg ctgaaggcgt cgtgagcgc	9420
ccctggcagc gactacggca gcccgctcgt catcagcgtc agcaaaggca gccctgacgg	9480
cagccacccg gtgtggtggt cgcctacaa cggcgggccg ccgcgcacgt gcccgaagat	9540
caagcaggag gcggtctctt cgtgcaccca cttggcgct ggacccctc tcagcaatgg	9600
ccaccggccg gctgcacacg acttccccct gggcgggcag ctccccagca ggactacccc	9660
gacctgggt cttgaggaag tgctgagcag cagggactgt caccctgcc tgcgcttcc	9720
tcccggttc catccccacc cggggcccaa ttaccatcc ttctgccc atcagatgca	9780
gccgcaagtc ccgcgctcc attaccaaga gctcatgcca ccgggttct gcatgccaga	9840
ggagcccaag ccaaagagg gaagacgac gtggccccg aaaaggaccg ccaccacac	9900
ttgtgattac gcgggtcgc gcaaaccta cacaagagt tccatctca aggcacacct	9960
gcgaaccac acaggtgaga aacctacca ctgtgactgg gacggctgtg gatggaaatt	10020
cgcgcgtca gatgaactga ccaggcacta ccgtaaacac acggggcacc gccggtcca	10080
gtgccaaaaa tgcgaccgag cattttcag gtcggaccac ctgccttac acatgaagag	10140
gcatttttct agacaatgta ctaactacgc tttgtgaaa ctgcgtggcg atgttgaaag	10200
taaccgccgt cctggcgcg ccatgtacaa catgatggag acggagctga agccgcggg	10260
ccgcagcaa acttcggggg gcggcgcgcg caactccacc gcggcgcgcg ccggcggcaa	10320

ES 2 698 527 T3

ccagaaaaac agcccgacc ggcgcaagcg gcccatgaat gccttcatgg tgtggtcccg	10380
cgggcagcgg cgcaagatgg ccagagagaa cccaagatg cacaactcgg agatcagcaa	10440
gcgcttgggc gccgagtgga aacttttgtc ggagacggag aagcggccgt tcacgcacga	10500
ggctaagcgg ctgcgagcgc tgcacatgaa ggagcaccgg gattataaat accggccccc	10560
gcggaaaacc aagacgctca tgaagaagga taagtacacg ctgcccgccg ggctgctggc	10620
ccccggcggc aatagcatgg cgagcggggc cggggtgggc gccggcctgg gcgcggggct	10680
gaaccagcgc atggacagtt acgcgcacat gaacggctgg agcaacggca gctacagcat	10740
gatgcaggac cagctgggct acccgagca cccgggcctc aatgcgcacg gcgcagcgca	10800
gatgcagccc atgcaccgct acgacgtgag cgccctgcag tacaactcca tgaccagctc	10860
gcagacctac atgaacggct cgccaccta cagcatgtcc tactcgcagc agggcacccc	10920
tggcatggct cttggctcca tgggttcggc ggtcaagtcc gaggccagct ccagccccc	10980
tgtggttacc tcttctccc actccagggc gccctgccag gccggggacc tccgggacat	11040
gatcagcatg tatctccccg gcgccgaggt gccggaacc gccgccccca gcagacttca	11100
catgtcccag cactaccaga gcggcccggt gcccggcacg gccattaacg gcacactgcc	11160
cctctcacac atgtgagcgg ccatcgatgt cgacaactaa cttaagctag caacggtttc	11220
cctctagcgg gatcaattcc gcccccccc cctaacttta ctggccgaag ccgcttgga	11280
taaggccggt gtgcgtttgt ctatatgtta tttccacca tattgccgtc ttttggaat	11340
gtgagggccc ggaacactgg cctgtcttc ttgacgagca ttctagggg tcttccct	11400
ctcgccaaag gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct	11460
tcttgaagac aaacaacgtc tgtagcgacc ctttgaggc agcggaacc cccacctggc	11520
gacaggtgcc tctgcggcca aaagccacgt gtataagata cacctgcaa ggcggcacaa	11580
ccccagtgcc acgttgtgag ttggatagtt gtggaagag tcaaatggct ctctcaagc	11640
gtattcaaca aggggctgaa ggatgccag aaggtacccc attgtatggg atctgatctg	11700
gggcctcggc gcacatgctt tacatgtgtt tagtcgaggt taaaaaacg tctaggcccc	11760
ccgaaccacg gggacgtggc tttctttga aaaacacgat aataccaatt cgccaccatg	11820
ggccccctca acgttagctt caccaacagg aactatgacc tcgactacga ctcggtgcag	11880
ccgtatttct actgcgacga ggagagaaac ttctaccagc agcagcagca gagcgagctg	11940
cagcccccg cgccagcga ggatatctgg aagaaattcg agctgctgcc cccccgcc	12000
ctgtccccta gccgcgctc cgggctctgc tcgccctcct acgttgcggt cacacccttc	12060
tcccttcggg gagacaacga cggcggtggc gggagcttct ccacggccga ccagctggag	12120

ES 2 698 527 T3

atggtgaccg agctgctggg aggagacatg gtgaaccaga gtttcatctg cgacccggac	12180
gacgagacct tcatcaaaaa catcatcatc caggactgta tgtggagcgg cttctcggcc	12240
gccgccaagc tcgtctcaga gaagctggcc tcctaccagg ctgcgcgcaa agacagcggc	12300
agcccgaaacc ccgcccgcgg ccacagcgtc tgctccacct ccagcttgta cctgcaggat	12360
ctgagcgccg ccgcctcaga gtgcatcgac ccctcgggtg tcttccccta ccctctcaac	12420
gacagcagct cgcccaagtc ctgcgcctcg caagactcca gcgccttctc tccgtectcg	12480
gattctctgc tctcctcgac ggagtcctcc ccgcagggca gccccgagcc cctggtgctc	12540
catgaggaga caccgcccac caccagcagc gactctgagg aggaacaaga agatgaggaa	12600
gaaatcgatg ttgtttctgt ggaaaagagg caggctcctg gcaaaaggtc agagtctgga	12660
tcaccttctg ctggaggcca cagcaaacct ctcacagcc cactggtcct caagaggtgc	12720
cacgtctcca cacatcagca caactacgca gcgcctccct ccactcggaa ggactatcct	12780
gctgccaaga gggtaagtt ggacagtgtc agagtctga gacagatcag caacaaccga	12840
aatgcacca gcccaggtc ctcgacacc gaggagaatg tcaagaggcg aacacacaac	12900
gtcttgagc gccagaggag gaacgagcta aaacggagct tttttgccct gcgtgaccag	12960
atcccgaggt tggaaaacaa tgaaaaggcc cccaaggtag ttatccttaa aaaagccaca	13020
gcatacatcc tgtccgtcca agcagaggag caaaagctca tttctgaaga ggacttggtg	13080
cggaaacgac gagaacagtt gaaacacaaa cttgaacagc tacggaactc ttgtgcgtaa	13140
tctagagtcg acccgggcgg ccgcaactaa cttagctag caacggtttc cctctagcgg	13200
gateaattcc gcccccccc cctaacgtta ctggccgaag ccgcttgaa taaggccggt	13260
gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggaat gtgagggcc	13320
ggaaaacctg ccctgtcttc ttgacgagca ttctagggg tctttcccct ctgcctaaa	13380
gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct tcttgaagac	13440
aaacaacgtc tgtagcgacc ctttgaggc agcggaacc cccacctggc gacaggtgcc	13500
tctgcggcca aaagccacgt gtataagata cacctgcaaa ggcggcacaa cccagtgcc	13560
acgttgtag ttgatagtt gtggaagag tcaaatggct ctctcaagc gtattcaaca	13620
aggggctgaa ggatgccag aaggtacccc attgtatggg atctgatctg gggcctcgg	13680
gcacatgctt tacatgtgtt tagtcgaggt taaaaaacg tctaggcccc ccgaaccag	13740
gggacgtggt tttcctttga aaaacacgat aataccatga ccgagtacaa gccacgggtg	13800
cgctcgcca ccgcgcgacga cgtccccagg gccgtacgca ccctcgccg cggttcgcc	13860
gactaccccg ccacgcgcca caccgtcgat ccggaaccgc acatcgagcg ggtcaccgag	13920
ctgcaagaac tcttctcac gcgcgtcggg ctgcacatcg gcaaggtgtg ggtcggggac	13980

ES 2 698 527 T3

gacggcgccg cggtagcggt ctggaccacg ccggagagcg tcgaagcggg ggcggtgttc	14040
gccgagatcg gcccgcgcat ggccgagttg agcggttccc ggctggccgc gcagcaacag	14100
atggaaggcc tcctggcgcc gcaccggccc aaggagcccg cgtggttcct ggccaccgtc	14160
ggcgtctcgc ccgaccacca gggcaagggt ctgggcagcg ccgtcgtgct ccccgagtg	14220
gaggcgggccg agcgcgcccg ggtgcccgcc ttcttgaga cctccgcgc ccgcaacctc	14280
cccttctacg agcggctcgg cttcaccgtc accgcccagc tcgaggtgcc cgaaggaccg	14340
cgcacctggt gcatgacctg caagcccggt gcctgagaat tggcaagctg cttacataga	14400
actcgcgcg atttggcatgc cgccttaaaa tttttatttt atttttcttt tcttttccga	14460
atcggatttt gtttttaata tttcaaaaa aaaaaaaaa aaaaaaaaa cgcgtcgagg	14520
ggaattaatt cttgaagacg aaagggccag gtggcacttt tcggggaaat gtgcgcggaa	14580
cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac	14640
cctgataaat gcttcaataa tattgaaaa ggaagagtat gagtattcaa catttcctg	14700
tcgcccttat tccctttttt gcggcatttt gccttctgt ttttgetcac ccagaaacgc	14760
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg	14820
atctcaacag cggtaagatc cttgagagtt ttcccccga agaacgtttt ccaatgatga	14880
gcacttttaa agttctgcta tgtggcgcg tattatcccg tgttgacgcc gggcaagagc	14940
aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag	15000
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtgtcc ataaccatga	15060
gtgataacac tgccggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg	15120
cttttttgca caacatgggg gatcatgtaa ctccgcttga tcgttgggaa ccggagctga	15180
atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt	15240
tgcgcaaact attaaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact	15300
ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt	15360
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg	15420
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta	15480
tggtgaacg aaatagacag atcgtgaga taggtgcctc actgattaag cattggtaac	15540
tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta	15600
aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt	15660
tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt	15720
ttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt	15780

ES 2 698 527 T3

gtttgcccga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc 15840
agataccaaa tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 15900
tagcaccgcc tacatacttc gctctgctaa tcctgttacc agtggctgct gccagtggcg 15960
ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg 16020
cgggctgaac ggggggttcg tgcacacagc ccagcttga gcgaacgacc tacaccgaac 16080
tgagatacct acagcgtgag cattgagaaa gcgccacgct tcccgaaggg agaaaggcgg 16140
acaggtatcc ggtaagcggc agggctcgaa caggagagcg cacgaggag cttccagggg 16200
gaaacgcctg gtatctttat agtctgtcg ggtttcgcca cctctgactt gagcgtcgat 16260
ttttgtgatg ctgcagggg gggcggagcc tatggaaaaa cgccagcaac gcgagctcta 16320
atacgactca ctatag 16336

<210> 31
<211> 16861
<212> ADN
<213> Secuencia Artificial

<220>
<223> VEE-Oct-Klf-Sox-Glis-SP6

<400> 31
atgggcggcg catgagagaa gcccagacca attacctacc caaaatggag aaagttcacg 60
ttgacatcga ggaagacagc ccattcctca gagctttgca gcggagcttc ccgcagtttg 120
aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcac 180
tggtctcaaa actgatcgaa acggagggtg acccatccga cagatcctt gacattggaa 240
gtgcgcccgc ccgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat 300
gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg 360
aaataactga taaggaattg gacaagaaaa tgaaggagct ggccgccgtc atgagcgacc 420
ctgacctgga aactgagact atgtgectcc acgacgacga gtcgtgtcgc tacgaagggc 480
aagtcgctgt ttaccaggat gtatacgcgg ttgacggacc gacaagtctc tatcaccaag 540
ccaataaggg agttagagtc gcctactgga taggttttga caccaccct tttatgttta 600
agaacttggc tggagcatat ccatacact ctaccaactg ggccgacgaa accgtgttaa 660
cggctcgtaa catagggcta tgcagctctg acgttatgga gcggtcacgt agagggatgt 720
ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga 780
ccatctacca cgagaagag gacttactga ggagctggca cctgccgtct gtatttcaact 840
tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg 900
tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta 960

ES 2 698 527 T3

cgatgcaccg cgagggattc ttgtgctgca aagtgcacaga cacattgaac ggggagaggg	1020
tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaatg actggcatac	1080
tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta	1140
tagtctgtaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgccc	1200
tagtggccca ggcatttgct aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320
acaagataac atctatztat aagcgcccgg atacccaaac catcatcaaa gtgaacagcg	1380
atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa	1440
caagaatcag gaaaatgtta gaggagcaca aggagccgtc acctctcatt accgccgagg	1500
acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
tcgacttgat gttacaagag gctggggccg gctcagtgga gacacctcgt ggcttgataa	1680
aggttaccag ctacgatggc gaggacaaga tcggctctta cgctgtgctt tctccgcagg	1740
ctgtactcaa gagtgaaaaa ttatcttgca tccaccctct cgctgaacaa gtcatagtga	1800
taacacactc tggccgaaaa gggcgttatg ccgtggaacc ataccatggt aaagtagtgg	1860
tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtga agtgccacca	1920
ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
gagcgctgaa cactgatgaa gaatattaca aaactgtcaa gcccagcgag cagcagggcg	2040
aatacctgta cgacatcgac aggaacacgt gcgtcaagaa agaactagtc actgggctag	2100
ggctcacagg cgagctggtg gatcctccct tccatgaatt cgctacgag agtctgagaa	2160
cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
gcaagtcttg catcattaaa agcgcagtc ccaaaaaaga tctagtgggt agcgccaaga	2280
aagaaaactg tgcagaaatt ataagggacg tcaagaaaaat gaaagggctg gacgtcaatg	2340
ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
ttgacgaagc ttttgcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
ctaaaaaggc agtgctctgc ggggatccca aacagtgcgg tttttttaac atgatgtgcc	2520
tgaaagtgca ttttaaccac gagatttgca cacaagtctt ccacaaaagc atctctcgcc	2580
gttgactact atctgtgact tcggtcgtct caaccttggt ttacgacaaa aaaatgagaa	2640
cgacgaatcc gaaagagact aagattgtga ttgacactac cggcagttacc aaacctaagc	2700
aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
aaggcaacga aataatgacg gcagctgcct ctcaagggct gacccgtaaa ggtgtgtatg	2820

ES 2 698 527 T3

ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
tcctactgac ccgcacggag gaccgcatcg tgtggaaaac actagccggc gacccatgga	2940
taaaaaact gactgccaag taccttgga atttactgc cacgatagag gagtggcaag	3000
cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gacgtcttcc	3060
agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
tagacatgac cactgaacaa tggaaactg tggattatct tgaacggac aaagctcact	3180
cagcagagat agtattgaac caactatgag tgaggttctt tggactcgat ctggactccg	3240
gtctatcttc tgcaccact gttccgttat ccattaggaa taatcactgg gataactccc	3300
cgctgcctaa catgtacggg ctgaataaag aagtggccg tcagctctct cgcaggctacc	3360
cacaactgcc tcgggcagtt gccactggaa gactctatga catgaacact ggtacactgc	3420
gcaattatga tccgcgcata aacctagtag ctgtaaacag aagactgcct catgctttag	3480
tcctccacca taatgaacac ccacagagt acttttcttc attcgtcagc aaattgaagg	3540
gcagaactgt cctggtggc ggggaaaagt tgtccgtccc aggcataatg gttgactggt	3600
tgtcagaccg gcctgaggct acctcagag ctccgctgga tttaggcatc ccagggtgatg	3660
tgcccaaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcac	3780
tgaatcccg cggaacctgt gtcagcatag gttatgggta cgctgacagg gccagcgaaa	3840
gcatcatttg tgctatagcg cggcagttca agttttcccg ggtatgcaaa ccgaaatcct	3900
cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
acaattctta caagctttca tcaaccttga ccaacattta tacagggttc agactccacg	4020
aagccgatg tgcacctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
gagtgattat aaatgctgct aacagcaaag gacaacctg cgagggggtg tgcggagcgc	4140
tgtataagaa attcccgaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
tggtaaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc acaaaagttt	4260
cggaggttga aggtgacaaa cagttggcag aggttatga gtccatcgct aagattgtca	4320
acgataaaa ttacaagtca gtagcgattc cactgttgct caccggcatc ttttcggga	4380
acaaagatcg actaaccctaa tcattgaacc atttgctgac agcttttagac accactgatg	4440
cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
ctaggagaga agcagtggag gagatagca tatccgacga ctcttcagt acagaacctg	4560
atgcagagct ggtgaggggt catccgaaga gttctttggc tggaaggaag ggctacagca	4620

ES 2 698 527 T3

caagcgatgg caaaactttc tcatatatttg aagggaacaa gtttcaccag gcggccaagg	4680
atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggatatgca	4740
tgtatatcct cggagaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
aagcctccac accacctagc acgctgcctt gcttgtgcat ccatgccatg actccagaaa	4860
gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctccagcct atattgttct	4980
caccgaaagt gcctgcgtat attcatcaa ggaagtatct cgtggaaaca ccaccggtag	5040
acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
cacttataac cgaggatgag accaggacta gaacgcctga gccgatcatc atcgaagagg	5160
aagaagagga tagcataagt ttgtgtcag atggcccgac ccaccagggtg ctgcaagtgc	5220
aggcagacat tcacgggccc cctctgtat ctagctcatc ctggtccatt cctcatgcat	5280
ccgactttga tgtggacagt ttatccatac ttgacaccct ggaggggagct agcgtgacca	5340
gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc	5400
gaccggtgcc tgcgcctcga acagtattca ggaaccctcc acatcccgtc ccgcgcacaa	5460
gaacaccgtc acttgcaccc agcagggcct gctcgagaac cagcctagtt tccacccgc	5520
caggcgtgaa tagggtgac actagagagg agctcgaggc gcttaccctg tcacgcactc	5580
ctagcaggtc ggtctcgaga accagcctgg tctccaaccc gccaggcgta aatagggtga	5640
ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg	5700
catacatctt ttctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760
cgggtgctatc cgaagtgggtg ttggagagga ccgaattgga gatttcgtat gccccgcgcc	5820
tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaattccc acacctgcta	5880
acagaagcag ataccagtcc aggaagggtg agaacatgaa agccataaca gctagacgta	5940
ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacct	6000
tgcacctgt tcctttgtat tcacttagtg tgaaccgtgc cttttcaagc cccaaggctg	6060
cagtggaaag ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
ttattccaga gtacgatgcc tatttggaca tgggtgacgg agcttcatgc tgcttagaca	6180
ctgccagttt ttgccctgca aagctgcga gctttccaaa gaaacactcc tatttggaac	6240
ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtccctggcag	6300
ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgtg ttggattcgg	6360
cggcctttta tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaacgt	6420
ttaaagaaaa ccccatcagg cttactgaag aaaacgtggt aaattacatt accaaattaa	6480

ES 2 698 527 T3

aaggacaaaa agctgctgct ctttttgcca agacacataa tttgaatatg ttgcaggaca	6540
taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
aacatactga agaacggccc aaggtacagg tgatccaggc tgccgatccg ctagcaacag	6660
cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
tccagcctgg ggatttgtgt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
acgccatggc tctgaccgcg ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
tgacgctgat tgaggcggct ttcggcgaaa tttcatcaat acatttgccc actaaaacta	6960
aatttaaatt cggagccatg atgaaatctg gaatgttcct cacactgttt gtgaacacag	7020
tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcaccatgtg	7080
cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
acaggtgcgc cacctgggtg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
gtgtggcaga cccctaaaaa aggctgttta agcttgcaa acctctggca gcagacgatg	7320
aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaaccgta ggaacttcca	7440
tcatagttat ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
ggggccctat aactctctac ggctaacctg aatggactac gacatagtct agtccgccaa	7560
gtctagcata tgggcgcgtg aattcgccac catggcggga cacctggctt cggatttcgc	7620
cttctcgccc cctccaggtg gtggaggatg tgggccaggg gggccggagc cgggctgggt	7680
tgatcctcgg acctggctaa gcttccaagg cctcctgga gggccaggaa tcgggccggg	7740
ggttgggcca ggctctgagg tgtgggggat tccccatgc ccccgcgct atgagttctg	7800
tggggggatg gcgtactgtg ggccccaggt tggagtgggg ctagtcccc aaggcggctt	7860
ggagacctct cagcctgagg gcgaagcagg agtcgggggt gagagcaact ccgatggggc	7920
ctccccggag cctgcaccg tcaccctgg tgccgtgaag ctggagaagg agaagctgga	7980
gcaaaacccg gaggagtccc aggacatcaa agctctgcag aaagaactcg agcaatttgc	8040
caagctcctg aagcagaaga ggatcacctt gggatataca caggccgatg tggggctcac	8100
cctgggggtt ctatttggga aggtattcag ccaaacgacc atctgccgct ttgaggctct	8160
gcagcttagc ttcaagaaca tgttaagct gcggcccttg ctgcagaagt ggggtggagga	8220
agctgacaac aatgaaaatc ttcaggagat atgcaaagca gaaaccctcg tgcaggcccg	8280

ES 2 698 527 T3

aaagagaaag cgaaccagta tgcagaaccg agtgagaggc aacctggaga atttgttcct	8340
gcagtgcccc aaaccacac tgcagcagat cagccacatc gcccagcagc ttgggctcga	8400
gaaggatgtg gtccgagtgt ggttctgtaa ccgcgccag aagggaagc gatcaagcag	8460
cgactatgca caacgagag attttgaggc tgctgggtct cctttctcag ggggaccagt	8520
gtcctttcct ctggccccag ggccccattt tggtagccca ggctatggga gccctcactt	8580
cactgcactg tactcctcgg tccctttccc tgaggggaa gcctttcccc ctgtctccgt	8640
caccactctg ggctctccca tgcattcaaa ctctagttag ggagaggaa gtctgctaac	8700
atgcggtgac gtcgaggaga atcctggccc acaattgatg gctgtcagcg acgcgtgct	8760
cccatctttc tccacgttcg cgtctggccc ggcgggaag gagaaacac tgcgtcaagc	8820
aggtgccccg aataaccgct ggcgggagga gctctccac atgaagcagc tccccagc	8880
gcttccggc cgccctatg acctggcgcc ggcgaccgtg gccacagacc tggagagcgg	8940
cggagccggt gcggttcg gcggtagcaa cctggcgccc ctacctcgga gagagaccga	9000
ggagttcaac gatctcctgg acctggactt tattctctcc aattcgtga cccatctcc	9060
ggagtcaagt gccgccaccg tgtcctcgtc agcgtcagcc tctctctcgt cgtcgcgtc	9120
gagcagcgcc cctgccagcg cgccctccac ctgcagcttc acctatccga tccgggcgg	9180
gaacgaccgg ggcgtggcgc cggcgccac ggcgggagc ctctctatg gcaggagtc	9240
cgtccccct ccgacggctc cttcaacct ggcggacatc aacgacgtga gccctcggg	9300
cggcttcgtg gccgagctcc tgcggccaga attggaccgg gtgtacattc cgcgcagca	9360
gccgcagccg ccagggtggc ggctgatggg caagtctgtg ctgaaggcgt cgtgagcgc	9420
ccctggcagc gactacggca gcccgctcgt catcagcgtc agcaaaggca gccctgacgg	9480
cagccacccg gtggtggtg cgccctacaa cggcgggccg ccgcgcacgt gcccgaagat	9540
caagcaggag gcggtctctt cgtgcaccca cttggcgct ggacccctc tcagcaatgg	9600
ccaccggccg gctgcacacg acttccccct gggcgggcag ctccccagca ggactaccc	9660
gacctgggt cttgaggaag tgctgagcag cagggactgt caccctgcc tgccgcttc	9720
tcccggttc catccccacc cggggccaa ttacctcc ttctgccc atcagatgca	9780
gccgaagtc ccgccgctc attaccaaga gtcatgcca ccggttcct gcatgccaga	9840
ggagcccaag ccaaagagg gaagacgac gtggccccg aaaaggaccg ccaccacac	9900
ttgtgattac gcgggtgct gcaaaccta cacaagagt tccatctca aggcacacct	9960
gcgaaccac acagggtgaga aacctacca ctgtgactgg gacggtgtg gatggaaatt	10020
cgcgcgtca gatgaactga ccaggcacta ccgtaaacac acggggcacc gccggtcca	10080
gtgcaaaaa tgcgaccgag cattttcag gtcggaccac ctgccttac acatgaagag	10140

ES 2 698 527 T3

gcatttttct agacaatgta ctaactacgc tttgttgaaa ctgctggcg atgttgaaag	10200
taaccccggt cctggcgcg ccatgtacaa catgatggag acggagctga agccgccggg	10260
cccgcagcaa acttcggggg ggcgcggcg caactccacc ggcgcggcg cgcgcggcaa	10320
ccagaaaaac agcccgacc ggcgcaagcg gcccatgaat gccttcatgg tgtggtccc	10380
cgggcagcgg cgcaagatgg ccagggagaa cccaagatg cacaactcgg agatcagcaa	10440
gcgctgggc gccgagtga aacttttgtc ggagacggag aagcgccgt tcacgacga	10500
ggctaagcgg ctgcgagcg tgcacatgaa ggagcaccgg gattataaat accggcccc	10560
gcggaacc aagacgctca tgaagaagga taagtacacg ctgccggcg ggctgctggc	10620
ccccggcggc aatagcatgg cgagcgggt cgggtgggc gccggcctgg gcgcggcgt	10680
gaaccagcgc atggacagtt acgcgcacat gaacggctgg agcaacggca gctacagcat	10740
gatgcaggac cagctgggct accgcagca cccggcctc aatgcgcacg gcgcagcga	10800
gatgcagccc atgcaccgt acgacgtgag cgcctgcag tacaactcca tgaccagtc	10860
gcagacctac atgaacggct cgcacaccta cagcatgtcc tactcgcagc agggcacccc	10920
tggcatggct cttggtcca tgggttcggg ggtcaagtcc gaggccagct ccagcccc	10980
tgtggttacc tcttctccc actccagggc gccctgccag gccggggacc tccgggacat	11040
gatcagcatg tatctcccc ggcgcgaggt gccggaaccc gccgccccca gcagaattca	11100
catgtcccag cactaccaga ggcgccgggt gcccggcacg gccattaacg gcacactgcc	11160
cctctcacac atgtgagcgg ccatcgatgt cgacaactaa cttaagctag caacggtttc	11220
cctctagcgg gatcaattcc gcccccccc cctaacgtta ctggccgaag ccgcttgga	11280
taaggccggt gtgcgtttgt ctatatgtta tttccacca tattgccgtc ttttggaat	11340
gtgagggccc ggaaacctgg ccctgtcttc ttgacgagca ttctagggg tcttcccct	11400
ctcgccaaag gaatgcaag tctgttgaat gtcgtgaag aagcagttcc tctggaagct	11460
tcttgaagac aaacaacgtc tgtagcgacc ctttgcaggc agcggaaccc cccacctggc	11520
gacaggtgcc tctgcggcca aaagccacgt gtataagata cacctgcaa ggcggcacaa	11580
ccccagtgcc acgttgtgag ttggatagtt gtggaaagag tcaaatggct ctctcaagc	11640
gtattcaaca aggggctgaa ggatgccag aaggtacccc attgtatggg atctgatctg	11700
gggcctcggg gcacatgctt tacatgtgtt tagtcgaggt taacaaaacg tctaggcccc	11760
ccgaaccacg gggacgtgg tttcctttga aaacacgat aataccaatt cgcaccatg	11820
gcagagggcc gcacatccct gtctgccac tgcggggc cgctggccac tggcctgcac	11880
ccagacctgg acctccggg ccgaagcctc gccaccctg cgccttctg ctacctctg	11940

ES 2 698 527 T3

ggcagcgaac ccagctcttg cctgggcctc cagcccgaga cccacctccc cgagggcagc	12000
ctgaagcggg gctgcgtctt gggcctaccc cccacctccc cagcctcctc ctcacctgt	12060
gcctcctccg acgtcacctc catcatccgc tctcccaga cgtctctggt cacctgtgta	12120
aatggactcc ggagcccccc tctgacggga gatctggggg gcccttccaa gcgggcccgg	12180
cctggccctg catcgacgga cagccatgag ggcagcttgc aacttgaagc ctgccggaag	12240
gcgagcttcc tgaagcagga acccgcgat gagttttcag agctctttg gcctcaccag	12300
cagggcctgc ccccccccta tccccgtct cagttgccgc ctggcccaag ccttgaggc	12360
ctggggctg gcttgccagg cagggtggtg gccggggcgc aggcgtgccg ctgggtggac	12420
tgctgtgcag cctatgagca gcaggaggag ctggtgcggc acatcgagaa gagccacatc	12480
gaccagcgca agggcgagga cttcacctgc ttctgggctg gctgcgtgcg ccgtacaag	12540
cccttcaacg ccgctacaa gctgctcatc cacatgcgag tgcactcggg cgagaagccc	12600
aacaagtgca tgtttgaagg ctgcagcaag gccttctcac ggctggagaa cctcaagatc	12660
cacctgagga gccacacggg cgagaagccg tacctgtgcc agcaccggg ttgccagaag	12720
gccttcagca actccagcga ccgcgccaa caccagcgca cccacctaga cacgaagccg	12780
tacgcctgtc agatccctgg ctgctccaag cgctacacag accccagctc cctccgcaag	12840
cacgtcaagg ccatttcagc caaagagcag caggtgcgta agaagctgca tgcgggacct	12900
gacaccgagg ccgacgtcct gaccgagtggt ctggtcctgc agcagctcca cacgtccaca	12960
cagctggctg ccagcgacgg caagggtggc tgtggcctgg gccaggagct gctcccaggt	13020
gtgtatcctg gctecatcac cccccataac ggacttgcct cgggcctcct gccccagcg	13080
cacgacgtac cttccaggca ccaccgctg gatgccacca ccagttccca ccaccatctg	13140
tccccctctg ccatggctga gagcaaccgg gatgggttgg ggcccgccct cctctcacca	13200
atagtcagcc ccctgaaggg gctggggcca ccgcgctgc ccccatcctc tcagagccat	13260
tctccggggg gccagccctt cccacactc cccagcaagc cgtcctaccc acccttccag	13320
agccctccac cccgcctct gccagccca caaggttacc agggcagttt ccactccatc	13380
cagagttgct tcccctatgg cgactgctac cggatggctg aaccagcagc cgggtgggac	13440
ggactggtcg gggagacca cggtttcaac cccctgcggc ccaatggcta ccacagcctc	13500
agcacgcctc tgctgccac aggtatgag gccctggctg aggcctcatg cccacagcg	13560
ctgccacagc agccatctga agatgtggtg tccagcggcc ccgaggactg tggtttcttc	13620
cccaatggag cttttgacca ctgcctgggc cacatcccct ccatctacac agacacctga	13680
gcggccgcaa ctaacttaag ctagcaacgg tttccctcta gcgggatcaa ttccgcccc	13740
ccccctaac gttactggcc gaagcgctt ggaataaggc cgggtgtgct ttgtctatat	13800

ES 2 698 527 T3

gttattttcc accatattgc cgtcttttgg caatgtgagg gcccggaac ctggccctgt	13860
cttcttgacg agcattccta ggggtctttc ccctctcgcc aaaggaatgc aaggtctgtt	13920
gaatgtcgtg aaggaagcag ttcctctgga agcttcttga agacaaacaa cgtctgtagc	13980
gaccttttgc aggcagcgga acccccacc tggcgacagg tgctctcgcg gccaaaagcc	14040
acgtgtataa gatacacctg caaaggcggc acaaccccag tgccacgttg tgagttggat	14100
agttgtggaa agagtcaa at ggctctcctc aagcgtattc aacaaggggc tgaaggatgc	14160
ccagaaggta cccatttga tgggatctga tctggggcct cgggtgcacat gctttacatg	14220
tgtttagtcg aggttaaaaa aacgtctagg cccccgaac cacggggacg tggttttcct	14280
ttgaaaaaca cgataatacc atgaccgagt acaagcccac ggtgcgcctc gccaccgcg	14340
acgacgtccc cagggccgta cgcaccctcg ccgcgcgtt ccgcgactac cccgccacgc	14400
gccacaccgt cgatccggac cgccacatcg agcgggtcac cgagctgcaa gaactcttcc	14460
tcacgcgcgt cgggctcgac atcggcaagg tgtgggtcgc ggacgacggc gccgcggtgg	14520
cggctcggac cagccggag agcgtcgaag cgggggcggg gttcgccgag atcggcccgc	14580
gcattggcca gttgagcggg tcccggctgg ccgcgcagca acagatggaa ggcctcctgg	14640
cgccgcaccg gcccaaggag cccgcgtggt tcttgccac cgtcggcgtc tcgcccacc	14700
accagggcaa gggctctggc agcgcgcgtc tgctccccgg agtgaggcgc gccgagcgcg	14760
ccgggggtgcc cgcttctctg gagacctcgc cgcccgcga cctccccctc tacgagcggc	14820
tcggcttcac cgtcacccgc gacgtcgagg tgcccgaagg accgcgcacc tgggtcatga	14880
cccgaagcc cgggtgcctga gaattggcaa gctgcttaca tagaactcgc ggcgattggc	14940
atgcgcctt aaaaattttt ttttattttt cttttctttt ccgaatcgga ttttgttttt	15000
aatatttcaa aaaaaaaaaa aaaaaaaaaa aaaacgcgtc gaggggaatt aattcttgaa	15060
gacgaaaggg ccagggtggca cttttcgggg aaatgtgcgc ggaacccta tttgtttatt	15120
tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca	15180
ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt	15240
ttttgcggca ttttgccctc ctgtttttgc tcacccagaa acgctggtga aagtaaaaga	15300
tgctgaagat cagttgggtg cacgagtggt ttacatcgaa ctggatctca acagcggtaa	15360
gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagttct	15420
gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg gtcgccgat	15480
acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga	15540
tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata aactgcggc	15600

ES 2 698 527 T3

caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat 15660
 gggggatcat gtaactcgcc ttgatcggtt ggaaccggag ctgaatgaag ccataccaaa 15720
 cgacgagcgt gacaccacga tgctgtagc aatggcaaca acgttgcgca aactattaac 15780
 tggcgaacta cttactctag cttcccggca acaattaata gactggatgg aggcggataa 15840
 agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc 15900
 tggagccggt gagcgtgggt ctgcgggtat cattgcagca ctggggccag atggttaagcc 15960
 ctcccgatc gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag 16020
 acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag accaagttta 16080
 ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa 16140
 gatccttttt gataatctca tgacaaaaat cccttaacgt gagttttcgt tccactgagc 16200
 gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat cttttttttc tgcgcgtaat 16260
 ctgctgcttg caaacaaaaa aaccaccgct accagcgggtg gtttgtttgc cggatcaaga 16320
 gctaccaact ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt 16380
 ccttctagt tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata 16440
 cctcgtcttg ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac 16500
 cgggttgga tcaagacgat agttaccgga taaggcgcag cggtcgggct gaacgggggg 16560
 ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg 16620
 tgagcattga gaaagcgcca cgcttccga agggagaaaag gcggacaggt atccggttaag 16680
 cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct 16740
 ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc 16800
 agggggggcg agcctatgga aaaacgccag caacgcgagc tcgatttagg tgacactata 16860
 g 16861

<210> 32
 <211> 16860
 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> VEE-Oct-Klf-Sox-Glis-T7

<400> 32
 atgggcggcg catgagagaa gccagacca attacctacc caaaatggag aaagttcacg 60
 ttgacatcga ggaagacagc ccattcctca gagctttgca gcggagcttc ccgcagtttg 120
 aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatc 180
 tggcttcaaa actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa 240

ES 2 698 527 T3

gtgcgcccgc cgcgagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat	300
gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg	360
aaataactga taaggaattg gacaagaaaa tgaaggagct ggccgccgtc atgagcgacc	420
ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc	480
aagtcgctgt ttaccaggat gtatacgcg gtgacggacc gacaagtctc tatcaccaag	540
ccaataaggg agttagagtc gcctactgga taggctttga caccaccctc tttatgttta	600
agaacttggc tggagcatat ccatcatact ctaccaactg ggccgacgaa accgtgttaa	660
cggctcgtaa catagggcta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttact	840
tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
tcgttaaaag aatagetatc agtccaggcc tgtatgggaa gccttcaggc tatgtctcta	960
cgatgcaccg cgagggattc ttgtgtcga aagtgcaga cacattgaac ggggagaggg	1020
tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgacaaatg actggcatac	1080
tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta	1140
tagtgcgcaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccg	1200
tagtggccca ggcatttgct aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320
acaagataac atctatttat aagcgcccg atacccaaac catcatcaaa gtgaacagcg	1380
atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa	1440
caagaatcag gaaaatgtta gaggagcaca aggagccgtc acctctcatt accgccgagg	1500
acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
tgcgcgcagc tctaccacct ttggcagctg atgttgagga gcccactctg gaagccgatg	1620
tcgacttgat gttacaagag gctggggccg gctcagtgga gacacctcgt ggcttgataa	1680
aggttaccag ctacgatggc gaggacaaga tcggtcttta cgctgtgctt tctccgagg	1740
ctgtactcaa gagtgaaaaa ttatcttgca tccacctctc cgctgaacaa gtcatagtga	1800
taacacactc tggccgaaaa gggcggttatg ccgtggaacc ataccatggt aaagtagtgg	1860
tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtga agtgccacca	1920
ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
gagcgtgaa cactgatgaa gaatattaca aaactgtcaa gcccgagcag cgcgacggcg	2040
aatacctgta cgacatcgac aggaacaggt gcgtcaagaa agaactagtc actgggctag	2100

ES 2 698 527 T3

ggctcacagg cgagctggtg gatcctccct tccatgaatt cgctacgag agtctgagaa	2160
cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
gcaagtcttg catcattaaa agcgagtc ccaaaaaaga tctagtgggt agcgccaaga	2280
aagaaaactg tgcagaaatt ataagggagc tcaagaaaat gaaagggctg gacgtcaatg	2340
ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
ttgacgaagc ttttgcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
ctaaaaaggc agtgctctgc ggggatccca aacagtgcgg tttttttaac atgatgtgcc	2520
tgaagtgc ttttaaccac gagatttga cacaagtctt ccacaaaagc atctctcgcc	2580
gttgactaa atctgtgact tcggtcgtct caacctgtt ttacgacaaa aaaatgagaa	2640
cgacgaatcc gaaagagact aagattgtga ttgacactac cggcagtagc aaacctaagc	2700
aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
aaggcaacga aataatgacg gcagctgcct ctcaagggct gacccgtaaa ggtgtgtatg	2820
ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
tcctactgac ccgcacggag gaccgcatcg tgtggaaaac actagccggc gacccatgga	2940
taaaaaact gactgccaag taccctggga atttactgc cacgatagag gagtggcaag	3000
cagagcatga tgccatcatg aggacatct tggagagacc ggaccctacc gacgtcttcc	3060
agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
tagacatgac cactgaacaa tggaacactg tggattatct tgaacaggac aaagctcact	3180
cagcagagat agtattgaac caactatgcg tgaggttctt tggactcgat ctggactccg	3240
gtctatcttc tgcaccact gttccgttat ccattaggaa taatcactgg gataactccc	3300
cgtcgcctaa catgtacggg ctgaataaag aagtgggtccg tcagctctct cgcaggtagc	3360
cacaactgcc tcgggcagtt gccactggaa gagtctatga catgaacact ggtactactg	3420
gcaattatga tccgcgcata aacctagtac ctgtaaacag aagactgcct catgctttag	3480
tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcagc aaattgaagg	3540
gcagaactgt cctggtggtc ggggaaaagt tgcgcgtccc aggcaaatg gttgactggt	3600
tgtcagaccg gcctgaggct acctcagag ctcggtgga tttaggcatc ccagggtgatg	3660
tgcccaaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcatc	3780
tgaatcccg cggaacctgt gtcagcatag gttatgggta cgctgacagc gccagcgaaa	3840
gcatcatttg tgctatagcg cggcagttca agttttcccg ggtatgcaaa ccgaaatcct	3900

ES 2 698 527 T3

cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagcg	3960
acaattctta caagctttca tcaaccttga ccaacattta tacaggttcc agactccacg	4020
aagccggtatg tgcacctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
gagtgattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
tgtataagaa attcccgaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
tggtaaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc aaaaagttt	4260
cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgct aagattgtca	4320
acgataacaa ttacaagtca gtacgcatc cactgttgtc caccggcatc ttttcoggga	4380
aaaaagatcg actaaccctaa tcattgaacc atttgtgac agcttttagac accactgatg	4440
cagatgtacg catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
atgcagagct ggtgaggggtg catccgaaga gttctttggc tggaagggaag ggctacagca	4620
caagcgatgg caaaactttc tcataatttg aagggaacca gtttcaccag gcggccaagg	4680
atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
tgtatatcct cggagaaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
aagcctccac accacctagc acgtgcctt gcttgtgcat ccatgccatg actccagaaa	4860
gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgtcga tcctttccat	4920
tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctccagcct atattgttct	4980
caccgaaagt gcctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag	5040
acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
cacttataac cgaggatgag accaggacta gaacgcctga gccgatcatc atcgaagagg	5160
aagaagagga tagcataagt ttgtgtcag atggcccgac ccaccagggtg ctgcaagtcg	5220
aggcagacat tcacgggccc cctctgtat ctagctcatc ctggtccatt cctcatgcat	5280
ccgactttga tgtggacagt ttatccatac ttgacaccct ggagggagct agcgtgacca	5340
gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc	5400
gaccgggtgc tgcgcctcga acagtattca ggaaccctcc acatcccgtc ccgcgcacaa	5460
gaacaccgtc acttgacccc agcagggcct gtcgagaac cagcctagtt tccacccgc	5520
caggcgtgaa tagggtgatc actagagagg agctcgaggc gcttaccctg tcacgcactc	5580
ctagcaggtc ggtctcgaga accagcctgg tctccaaccc gccaggcgta aataggggtga	5640
ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg	5700
catacatctt ttctccgac accggtcaag ggcatttaca aaaaaatca gtaaggcaaa	5760

ES 2 698 527 T3

cggtgctatc cgaagtgggtg ttggagagga ccgaattgga gatttcgtat gccccgcgcc	5820
tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaattccc acacctgcta	5880
acagaagcag ataccagtcc aggaaggtgg agaacatgaa agccataaca gctagacgta	5940
ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacct	6000
tgcatacctgt tcctttgtat tcacttagtg tgaaccgtgc cttttcaagc cccaaggctc	6060
cagtgggaagc ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaac	6240
ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtcctggcag	6300
ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgta ttggattcgg	6360
cggcctttaa tgtggaatgc ttcaagaaat atgctgttaa taatgaatat tgggaaacgt	6420
ttaaagaaaa ccccatcagg cttactgaag aaaacgtggg aaattacatt accaaattaa	6480
aaggacaaaa agctgctgct ctttttgca agacacataa tttgaatatg ttgcaggaca	6540
taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
aacatactga agaacggccc aaggtacagg tgatccaggc tgccgatccg ctagcaacag	6660
cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
tccagcctgg ggattgtgtt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
acgccatggc tctgaccgcg ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
tgacgctgat tgaggcggct ttcggcgaaa tttcatcaat acatttgccc actaaaacta	6960
aatttaaatt cggagccatg atgaaatctg gaatgttcct cacactgttt gtgaacacag	7020
tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcacatgtg	7080
cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
acaggtgcgc cacctggttg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
gtgtggcaga cccctaaaa aggctgttta agcttgcaa acctctggca gcagacgatg	7320
aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaaccgta ggaacttcca	7440
tcatagttat ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
gggcccctat aactctctac ggctaacctg aatggactac gacatagtct agtccgcaa	7560

ES 2 698 527 T3

gtctagcata tgggcgcgtg aattcgccac catggcggga cacctggctt cggatttcgc	7620
cttctcgccc cctccagggtg gtggagggtga tgggccaggg gggccggagc cgggctgggt	7680
tgatcctcgg acctggctaa gcttccaagg cctcctgga gggccaggaa tcgggccggg	7740
ggttgggcca ggctctgagg tgtgggggat tccccatgc ccccgcctt atgagttctg	7800
tggggggatg gcgtactgtg ggccccagggt tggagtgggg ctagtgcctc aaggcggctt	7860
ggagacctct cagcctgagg gcgaagcagg agtcgggggtg gagagcaact ccgatggggc	7920
ctccccggag cctgcaccg tcacccttg tgccgtgaag ctggagaagg agaagctgga	7980
gcaaaacccg gaggagtccc aggacatcaa agctctgcag aaagaactcg agcaatttgc	8040
caagctcctg aagcagaaga ggatcacctt gggatataca caggccgatg tggggctcac	8100
cctgggggtt ctatttggga aggtattcag ccaaacgacc atctgcctt ttgaggctct	8160
gcagcttagc ttcaagaaca tgtgtaagct gcggcccttg ctgcagaagt gggaggagga	8220
agctgacaac aatgaaaatc ttcaggagat atgcaaagca gaaacctcg tgcaggcccg	8280
aaagagaaag cgaaccagta tcgagaaccg agtgagaggc aacctggaga atttgctcct	8340
gcagtgcctg aaaccacac tgcagcagat cagccacatc gccagcagc ttgggctcga	8400
gaaggatgtg gtccgagtgt ggttctgtaa ccggcgccag aagggaagc gatcaagcag	8460
cgactatgca caacgagagg attttgaggc tgctgggtct cctttctcag ggggaccagt	8520
gtcctttcct ctggccccag ggccccattt tggtagccca ggctatggga gccctcactt	8580
cactgcactg tactcctcgg tccctttccc tgagggggaa gcctttcccc ctgtctcct	8640
caccactctg ggctctccca tgcattcaaa ctctagttag ggcagaggaa gtctgctaac	8700
atgcgggtgac gtgcaggaga atcctggccc acaattgatg gctgtcagcg acgcgtgct	8760
cccatctttc tccacgttcg cgtctggccc ggcggaagg gagaagacac tgcgtcaagc	8820
aggtgccccg aataaccgct ggcgggagga gctctccac atgaagcagc ttccccagt	8880
gcttcccggc cgccttatg acctggcggc ggcgaccgtg gccacagacc tggagagcgg	8940
cggagccggt gcggttgcg gcggtagcaa cctggcgccc ctacctcgga gagagaccga	9000
ggagttcaac gatctcctgg acctggactt tattctctcc aattcgctga cccatcctcc	9060
ggagtcagtg gccgccaccg tgtcctcgtc agcgtcagcc tcctcttctg cgtcgcctc	9120
gagcagcggc cctgccagcg cgcctccac ctgcagcttc acctatccga tccgggccc	9180
gaacgacccg ggcgtggcgc cggcgggcac gggcgaggc ctctctatg gcagggagtc	9240
cgtccccct cgcagcgtc ccttcaacct ggcgagacatc aacgacgtga gccctcggg	9300
cggcttcgtg gccgagctcc tgcggccaga attggacccg gtgtacattc cgcgcagca	9360
gccgcagccg ccagggtggc ggctgatggg caagttcgtg ctgaaggcgt cgtgagcgc	9420

ES 2 698 527 T3

ccctggcagc gagtacggca gcccgtcggt catcagcgtc agcaaaggca gccctgacgg	9480
cagccacccg gtggtggtgg cgccctacaa cggcgggccg ccgcgcacgt gcccgaagat	9540
caagcaggag gcggtctctt cgtgcaccca cttgggcgct ggaccccctc tcagcaatgg	9600
ccaccggccg gctgcacacg acttccccct gggcgggcag ctccccagca ggactacccc	9660
gaccttgggt cttgaggaag tgctgagcag cagggaactgt caccctgccc tgccgcttcc	9720
tcccggttcc catccccacc cggggcccaa ttacccatcc ttcttgccc atcagatgca	9780
gccgcaagtc ccgcgcctcc attaccaaga gctcatgcca cccggttctc gcatgccaga	9840
ggagcccaag ccaagagggt gaagacgacg gtggccccg aaaaggaccg ccaccacac	9900
ttgtgattac gcgggtgctg gcaaaccta cacaagagt tcccatctca aggcacacct	9960
gcgaaccac acaggtgaga aaccttacca ctgtgactgg gacggctgtg gatggaaatt	10020
cggcgctca gatgaactga ccaggcacta ccgtaaacac acggggcacc gcccgttcca	10080
gtgcaaaaaa tgcgaccgag cattttccag gtcgaccac ctgccttac acatgaagag	10140
gcatttttct agacaatgta ctaactacgc tttgtgaaa ctgctggcg atgttgaaag	10200
taaccccggt cctggcgctc ccatgtacaa catgatggag acggagctga agccgcggg	10260
ccgcagcaa acttcggggg gcggcgggcg caactccacc gcggcgggcg ccggcgcaa	10320
ccagaaaaac agcccgacc gcgtcaagcg gcccatgaat gccttcatgg tgttggtccc	10380
cgggcagcgg cgcaagatgg ccagggagaa cccaagatg cacaactcgg agatcagcaa	10440
gcgcctgggc gccgagtga aacttttgc ggagacggag aagcgccgt tcacgcagca	10500
ggctaagcgg ctgcgagcgc tgcacatgaa ggagcaccg gattataaat accggccccg	10560
gcggaaaacc aagacgtca tgaagaagga taagtacacg ctgcccggcg ggctgctggc	10620
ccccggcggc aatagcatgg cgagcggggt cggggtgggc gccggcctgg gcgcgggct	10680
gaaccagcgc atggacagtt acgcgcacat gaacggctgg agcaacggca gctacagcat	10740
gatgcaggac cagctgggct acccgagca cccgggcctc aatgcgcacg gcgcagcgca	10800
gatgcagccc atgcaccgct acgacgtgag cgccctgcag tacaactcca tgaccagctc	10860
gcagacctac atgaacggct cgccaccta cagcatgtcc tactcgacg agggcacccc	10920
tgcatggct cttggctcca tgggttcggt ggtcaagtcc gaggccagct ccagccccc	10980
tgtggttacc tcttctccc actccaggc gccctgccag gccggggacc tccgggacat	11040
gatcagcatg tatctccccg gcgccgaggt gccggaacc gccgccccca gcagacttca	11100
catgtcccag cactaccaga gcggcccggt gcccggcacg gccattaacg gcacactgcc	11160
cctctcacac atgtgagcgg ccatcgatgt cgacaactaa cttaagctag caacggtttc	11220

ES 2 698 527 T3

cctctagcgg gatcaattcc gccccccccc cctaacgtta ctggccgaag ccgcttgga	11280
taaggccggt gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggcaat	11340
gtgagggccc ggaaacctgg cctgtcttc ttgacgagca ttctagggg tctttcccct	11400
ctcgccaaag gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct	11460
tcttgaagac aaacaacgtc tgtagcgacc ctttgaggc agcggaaccc cccacctggc	11520
gacaggtgcc tctgcggcca aaagccacgt gtataagata cacctgcaa ggcggcacaa	11580
ccccagtgcc acgttgtgag ttggatagtt gtggaagag tcaaatggct ctctcaagc	11640
gtattcaaca aggggctgaa ggatgccag aaggtagccc attgtatggg atctgatctg	11700
gggcctcggg gcacatgctt tacatgtgtt tagtcgaggt taaaaaacg tctaggcccc	11760
ccgaaccacg gggagctggg tttccttga aaaacacgat aataccaatt cgccaccatg	11820
gcagaggccc gcacatccct gtctgccac tgtcggggc cgctggccac tggcctgcac	11880
ccagacctgg acctcccggg ccgaagcctc gccaccctg cgcttctctg ctaccttctg	11940
ggcagcgaac ccagctctgg cctgggcctc cagcccgaga cccacctccc cgagggcagc	12000
ctgaagcggg gctgcgtctt gggcctaccc cccacctccc cagcctctc ctcacctgt	12060
gcctcctccg acgtcacctc catcatccgc tctcccaga cgtctctggg cacctgtgta	12120
aatggactcc ggagccccc tctgacggga gatctggggg gcccttcaa gggggccgg	12180
cctggccctg catcgacgga cagccatgag ggcagcttgc aactgaagc ctgccggaag	12240
gcgagcttcc tgaagcagga acccgggat gagtttccag agctctttgg gcctaccag	12300
cagggcctgc cgcccccta tcccctgtct cagttgccgc ctggcccaag ccttgagggc	12360
ctggggctgg gcctggcagg cagggtggtg gccggggcgc aggcgtgccg ctgggtggac	12420
tgctgtgcag cctatgagca gcaggaggag ctggtgcggc acatcgagaa gagccacatc	12480
gaccagcgca agggcgagga cttcacctgc ttctgggctg gctgcgtgcg ccgctacaag	12540
cccttcaacg cccgctacaa gctgctcatc cacatgcgag tgactcggg cgagaagccc	12600
aacaagtgca tgtttgaagg ctgcagcaag gccttctcac ggctggagaa cctcaagatc	12660
cacctgagga gccacacggg cgagaagccg tacctgtgcc agcaccggg ttgccagaag	12720
gccttcagca actccagcga ccgcgccaa caccagcgca cccacctaga cacgaagccg	12780
tacgcctgtc agatccctgg ctgctccaag cgctacacag acccagctc cctccgcaag	12840
cacgtcaagg ccatttcagc caaagagcag caggtgcgta agaagctgca tgcgggccct	12900
gacaccgagg ccgacgtcct gaccgagtgt ctggtcctgc agcagctcca cacgtccaca	12960
cagctggctg ccagcgacgg caagggtggc tgtggcctgg gccaggagct gctcccaggt	13020
gtgtatcctg gctccatcac cccccataac ggacttgcac cgggcctcct gccccagcg	13080

ES 2 698 527 T3

cacgacgtac cttccaggca ccaccgctg gatgccacca ccagttccca ccaccatctg	13140
tcccctctgc ccatggctga gagcaccgg gatgggttg ggcccgccct cctctacca	13200
atagtcagcc cctgaaggg gctggggcca ccgccgctgc ccccatcctc tcagagccat	13260
tctccggggg gccagccctt ccccacactc cccagcaagc cgtcctaccc acccttccag	13320
agccctccac ccccgctct gcccagccca caaggttacc agggcagttt ccactccatc	13380
cagagttgct tcccctatgg cgactgctac cggatggctg aaccagcagc cgggtggggac	13440
ggactggctg gggagaccca cggtttcaac ccctgcggc ccaatggcta ccacagcctc	13500
agcacgcctt tgctgccac aggtatgag gccctggctg aggcctcatg cccacagcg	13560
ctgccacagc agccatctga agatgtggtg tccagcggcc ccgaggactg tggcttcttc	13620
cccaatggag ctttgacca ctgctgggc cacatcccct ccatctacac agacacctga	13680
gcggccgcaa ctaacttaag ctagcaacgg tttccctcta gcgggatcaa ttccgcccc	13740
ccccctaac gttactggc gaagcgcctt ggaataaggc cgggtgctg ttgtctatat	13800
gttattttcc accatattgc cgtcttttg caatgtgagg gcccgaaac ctggccctgt	13860
cttcttgacg agcattccta ggggtctttc ccctctgcc aaaggaatgc aaggtctgtt	13920
gaatgtcgtg aaggaagcag ttcctctgga agcttcttga agacaaaca cgtctgtagc	13980
gaccctttgc aggcagcggg accccccacc tggcgacagg tgcctctgcg gccaaaagcc	14040
acgtgtataa gatacacctg caaaggcggc acaaccacag tgccacgttg tgagttggat	14100
agttgtggaa agagtcaaat ggctctcctc aagcgtattc aacaaggggc tgaaggatgc	14160
ccagaaggta cccattgta tgggatctga tctggggcct cggtgacat gctttacatg	14220
tgtttagtcg aggttaaaaa aacgtctag ccccccgaac cacggggacg tggttttcct	14280
ttgaaaaaca cgataatacc atgaccgagt acaagccac ggtgcgcctc gccaccgcg	14340
acgacgtccc cagggcgta cgcaccctcg ccgccgctt cgcgcactac cccgccacgc	14400
gccacaccgt cgatccggac cgccacatcg agcgggtcac cgagctgcaa gaactcttcc	14460
tcacgcgcgt cgggctcgac atcggaagg tgtgggtcgc ggacgacgac gccgcggtg	14520
cggctcggac cagcgggag agcgtcgaag cggggcggt gtccgcccag atcgccccg	14580
gcatggccga gttgagcgt tcccggctgg ccgcgcagca acagatggaa ggcctctgg	14640
cgccgcaccg gcccaaggag ccgcgctggt tccctggcac cgtcggcgtc tcgcccgacc	14700
accagggcaa ggtctgggc agcgcgctg tgctccccg agtgaggcg gccgagcgcg	14760
ccgggggtgc cgccttctg gagacctcg cgccccgaa cctccccctc tacgagcggc	14820
tcggcttcac cgtcaccgcc gacgtcagag tgcccgaag accgcgcacc tgggtcatga	14880

ES 2 698 527 T3

ccgcgaagcc cggtgcctga gaattggcaa gctgcttaca tagaactcgc ggcgattggc 14940
 atgccgcctt aaaattttta ttttattttt cttttctttt ccgaatcgga ttttgttttt 15000
 aatattttcaa aaaaaaaaaa aaaaaaaaaa aaaacgcgtc gaggggaatt aattcttgaa 15060
 gacgaaaggg ccaggtggca cttttcgggg aaatgtgcgc ggaaccccta tttgtttatt 15120
 tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca 15180
 ataatatga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt 15240
 ttttgcgga ttttgcttc ctgtttttgc tcaccagaa acgctgggta aagtaaaaga 15300
 tgctgaagat cagttgggtg cacgagtggt ttacatcgaa ctggatctca acagcggtaa 15360
 gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagttct 15420
 gctatgtggc gcgtattat cccgtgttga cgccgggcaa gagcaactcg gtcgccgat 15480
 aactattct cagaatgact tggttgagta ctaccagtc acagaaaagc atcttacgga 15540
 tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtata aactgcggc 15600
 caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat 15660
 gggggatcat gtaactcgcc ttgatcggtt ggaaccggag ctgaatgaag ccataccaaa 15720
 cgacgagcgt gacaccacga tgctgtagc aatggcaaca acgttgcgca aactattaac 15780
 tggcgaacta cttactctag cttcccgga acaattaata gactggatgg aggcggataa 15840
 agttgcagga ccacttctgc gctcgccct tccggctggc tggtttattg ctgataaatc 15900
 tggagccggt gagcgtgggt ctgcggtat cattgcagca ctggggccag atggttaagc 15960
 ctcccgatc gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag 16020
 acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag accaagttta 16080
 ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa 16140
 gatccttttt gataatctca tgacaaaaat ccctaacgt gagttttcgt tccactgagc 16200
 gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat cttttttttc tgcgcgtaat 16260
 ctgctgcttg caaacaaaaa aaccaccgct accagcgggt gtttgtttgc cggtacaaga 16320
 gctaccaact ctttttccga aggttaactg cttcagcaga gcgcagatac caaatactgt 16380
 cttctagtg tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata 16440
 cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac 16500
 cgggttgac tcaagacgat agttaccgga taaggcgag cggtcgggct gaacgggggg 16560
 ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg 16620
 tgagcattga gaaagcgcca cgcttccga agggagaaaag gcggacaggt atccggttaag 16680
 cggcagggtc ggaacaggag agcgacagag ggagcttcca gggggaaacg cctggtatct 16740
 ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc 16800
 aggggggcgg agcctatgga aaaacgccag caacgcgagc tctaatacga ctactatag 16860

5 <210> 33
 <211> 1863
 <212> ADN
 <213> Homo sapiens

10 <220>
 <221> CDS

ES 2 698 527 T3

<222> (1)..(1863)

<400> 33

atg gca gag gcc cgc aca tcc ctg tct gcc cac tgt cgg ggc ccg ctg Met Ala Glu Ala Arg Thr Ser Leu Ser Ala His Cys Arg Gly Pro Leu 1 5 10 15	48
gcc act ggc ctg cac cca gac ctg gac ctc ccg ggc cga agc ctc gcc Ala Thr Gly Leu His Pro Asp Leu Asp Leu Pro Gly Arg Ser Leu Ala 20 25 30	96
acc cct gcg cct tcc tgc tac ctt ctg ggc agc gaa ccc agc tct ggc Thr Pro Ala Pro Ser Cys Tyr Leu Leu Gly Ser Glu Pro Ser Ser Gly 35 40 45	144
ctg ggc ctc cag ccc gag acc cac ctc ccc gag ggc agc ctg aag cgg Leu Gly Leu Gln Pro Glu Thr His Leu Pro Glu Gly Ser Leu Lys Arg 50 55 60	192
tgc tgc gtc ttg ggc cta ccc ccc acc tcc cca gcc tcc tcc tca ccc Cys Cys Val Leu Gly Leu Pro Pro Thr Ser Pro Ala Ser Ser Ser Pro 65 70 75 80	240
tgt gcc tcc tcc gac gtc acc tcc atc atc cgc tcc tcc cag acg tct Cys Ala Ser Ser Asp Val Thr Ser Ile Ile Arg Ser Ser Gln Thr Ser 85 90 95	288
ctg gtc acc tgt gta aat gga ctc cgg agc ccc cct ctg acg gga gat Leu Val Thr Cys Val Asn Gly Leu Arg Ser Pro Pro Leu Thr Gly Asp 100 105 110	336
ctg ggg ggc cct tcc aag cgg gcc cgg cct ggc cct gca tcg acg gac Leu Gly Gly Pro Ser Lys Arg Ala Arg Pro Gly Pro Ala Ser Thr Asp 115 120 125	384
agc cat gag ggc agc ttg caa ctt gaa gcc tgc cgg aag gcg agc ttc Ser His Glu Gly Ser Leu Gln Leu Glu Ala Cys Arg Lys Ala Ser Phe 130 135 140	432
ctg aag cag gaa ccc gcg gat gag ttt tca gag ctc ttt ggg cct cac Leu Lys Gln Glu Pro Ala Asp Glu Phe Ser Glu Leu Phe Gly Pro His 145 150 155 160	480
cag cag ggc ctg ccg ccc ccc tat ccc ctg tct cag ttg ccg cct ggc Gln Gln Gly Leu Pro Pro Pro Tyr Pro Leu Ser Gln Leu Pro Pro Gly 165 170 175	528
cca agc ctt gga ggc ctg ggg ctg ggc ctg gca ggc agg gtg gtg gcc Pro Ser Leu Gly Gly Leu Gly Leu Gly Leu Ala Gly Arg Val Val Ala 180 185 190	576

ES 2 698 527 T3

ggg cgg cag gcg tgc cgc tgg gtg gac tgc tgt gca gcc tat gag cag Gly Arg Gln Ala Cys Arg Trp Val Asp Cys Cys Ala Ala Tyr Glu Gln 195 200 205	624
cag gag gag ctg gtg cgg cac atc gag aag agc cac atc gac cag cgc Gln Glu Glu Leu Val Arg His Ile Glu Lys Ser His Ile Asp Gln Arg 210 215 220	672
aag ggc gag gac ttc acc tgc ttc tgg gct ggc tgc gtg cgc cgc tac Lys Gly Glu Asp Phe Thr Cys Phe Trp Ala Gly Cys Val Arg Arg Tyr 225 230 235 240	720
aag ccc ttc aac gcc cgc tac aag ctg ctc atc cac atg cga gtg cac Lys Pro Phe Asn Ala Arg Tyr Lys Leu Leu Ile His Met Arg Val His 245 250 255	768
tcg ggc gag aag ccc aac aag tgc atg ttt gaa ggc tgc agc aag gcc Ser Gly Glu Lys Pro Asn Lys Cys Met Phe Glu Gly Cys Ser Lys Ala 260 265 270	816
ttc tca cgg ctg gag aac ctc aag atc cac ctg agg agc cac acg ggc Phe Ser Arg Leu Glu Asn Leu Lys Ile His Leu Arg Ser His Thr Gly 275 280 285	864
gag aag ccg tac ctg tgc cag cac ccg ggt tgc cag aag gcc ttc agc Glu Lys Pro Tyr Leu Cys Gln His Pro Gly Cys Gln Lys Ala Phe Ser 290 295 300	912
aac tcc agc gac cgc gcc aag cac cag cgc acc cac cta gac acg aag Asn Ser Ser Asp Arg Ala Lys His Gln Arg Thr His Leu Asp Thr Lys 305 310 315 320	960
ccg tac gcc tgt cag atc cct ggc tgc tcc aag cgc tac aca gac ccc Pro Tyr Ala Cys Gln Ile Pro Gly Cys Ser Lys Arg Tyr Thr Asp Pro 325 330 335	1008
agc tcc ctc cgc aag cac gtc aag gcc cat tca gcc aaa gag cag cag Ser Ser Leu Arg Lys His Val Lys Ala His Ser Ala Lys Glu Gln Gln 340 345 350	1056
gtg cgt aag aag ctg cat gcg ggc cct gac acc gag gcc gac gtc ctg Val Arg Lys Lys Leu His Ala Gly Pro Asp Thr Glu Ala Asp Val Leu 355 360 365	1104
acc gag tgt ctg gtc ctg cag cag ctc cac acg tcc aca cag ctg gct Thr Glu Cys Leu Val Leu Gln Gln Leu His Thr Ser Thr Gln Leu Ala 370 375 380	1152
gcc agc gac ggc aag ggt ggc tgt ggc ctg ggc cag gag ctg ctc cca Ala Ser Asp Gly Lys Gly Gly Cys Gly Leu Gly Gln Glu Leu Leu Pro 385 390 395 400	1200
ggt gtg tat cct ggc tcc atc acc ccc cat aac gga ctt gca tcg ggc Gly Val Tyr Pro Gly Ser Ile Thr Pro His Asn Gly Leu Ala Ser Gly 405 410 415	1248
ctc ctg ccc cca gcg cac gac gta cct tcc agg cac cac ccg ctg gat Leu Leu Pro Pro Ala His Asp Val Pro Ser Arg His His Pro Leu Asp 420 425 430	1296

ES 2 698 527 T3

```

gcc acc acc agt tcc cac cac cat ctg tcc cct ctg ccc atg gct gag      1344
Ala Thr Thr Ser Ser His His His Leu Ser Pro Leu Pro Met Ala Glu
      435                        440                        445

agc acc cgg gat ggg ttg ggg ccc ggc ctc ctc tca cca ata gtc agc      1392
Ser Thr Arg Asp Gly Leu Gly Pro Gly Leu Leu Ser Pro Ile Val Ser
      450                        455                        460

ccc ctg aag ggg ctg ggg cca ccg ccg ctg ccc cca tcc tct cag agc      1440
Pro Leu Lys Gly Leu Gly Pro Pro Pro Leu Pro Pro Ser Ser Gln Ser
      465                        470                        475                        480

cat tct ccg ggg ggc cag ccc ttc ccc aca ctc ccc agc aag ccg tcc      1488
His Ser Pro Gly Gly Gln Pro Phe Pro Thr Leu Pro Ser Lys Pro Ser
      485                        490                        495

tac cca ccc ttc cag agc cct cca ccc ccg cct ctg ccc agc cca caa      1536
Tyr Pro Pro Phe Gln Ser Pro Pro Pro Pro Pro Leu Pro Ser Pro Gln
      500                        505                        510

ggg tac cag ggc agt ttc cac tcc atc cag agt tgc ttc ccc tat ggc      1584
Gly Tyr Gln Gly Ser Phe His Ser Ile Gln Ser Cys Phe Pro Tyr Gly
      515                        520                        525

gac tgc tac ccg atg gct gaa cca gca gcc ggt ggg gac gga ctg gtc      1632
Asp Cys Tyr Arg Met Ala Glu cca Ala Ala Gly Gly Asp Gly Leu Val
      530                        535                        540

ggg gag acc cac ggt ttc aac ccc ctg ccg ccc aat ggc tac cac agc      1680
Gly Glu Thr His Gly Phe Asn Pro Leu Arg Pro Asn Gly Tyr His Ser
      545                        550                        555                        560

ctc agc acg ccc ttg cct gcc aca ggc tat gag gcc ctg gct gag gcc      1728
Leu Ser Thr Pro Leu Pro Ala Thr Gly Tyr Glu Ala Leu Ala Glu Ala
      565                        570                        575

tca tgc ccc aca gcg ctg cca cag cag cca tct gaa gat gtg gtg tcc      1776
Ser Cys Pro Thr Ala Leu Pro Gln Gln Pro Ser Glu Asp Val Val Ser
      580                        585                        590

agc ggc ccc gag gac tgt ggc ttc ttc ccc aat gga gcc ttt gac cac      1824
Ser Gly Pro Glu Asp Cys Gly Phe Phe Pro Asn Gly Ala Phe Asp His
      595                        600                        605

tgc ctg ggc cac atc ccc tcc atc tac aca gac acc tga      1863
Cys Leu Gly His Ile Pro Ser Ile Tyr Thr Asp Thr
      610                        615                        620

```

<210> 34
 <211> 620
 <212> PRT
 <213> Homo sapiens

<400> 34
 Met Ala Glu Ala Arg Thr Ser Leu Ser Ala His Cys Arg Gly Pro Leu
 1 5 10 15
 Ala Thr Gly Leu His Pro Asp Leu Asp Leu Pro Gly Arg Ser Leu Ala
 20 25 30
 Thr Pro Ala Pro Ser Cys Tyr Leu Leu Gly Ser Glu Pro Ser Ser Gly
 35 40 45

ES 2 698 527 T3

Leu Gly Leu Gln Pro Glu Thr His Leu Pro Glu Gly Ser Leu Lys Arg
 50 55 60
 Cys Cys Val Leu Gly Leu Pro Pro Thr Ser Pro Ala Ser Ser Ser Pro
 65 70 75 80
 Cys Ala Ser Ser Asp Val Thr Ser Ile Ile Arg Ser Ser Gln Thr Ser
 85 90 95
 Leu Val Thr Cys Val Asn Gly Leu Arg Ser Pro Pro Leu Thr Gly Asp
 100 105 110
 Leu Gly Gly Pro Ser Lys Arg Ala Arg Pro Gly Pro Ala Ser Thr Asp
 115 120 125
 Ser His Glu Gly Ser Leu Gln Leu Glu Ala Cys Arg Lys Ala Ser Phe
 130 135 140
 Leu Lys Gln Glu Pro Ala Asp Glu Phe Ser Glu Leu Phe Gly Pro His
 145 150 155 160
 Gln Gln Gly Leu Pro Pro Pro Tyr Pro Leu Ser Gln Leu Pro Pro Gly
 165 170 175
 Pro Ser Leu Gly Gly Leu Gly Leu Gly Leu Ala Gly Arg Val Val Ala
 180 185 190
 Gly Arg Gln Ala Cys Arg Trp Val Asp Cys Cys Ala Ala Tyr Glu Gln
 195 200 205
 Gln Glu Glu Leu Val Arg His Ile Glu Lys Ser His Ile Asp Gln Arg
 210 215 220
 Lys Gly Glu Asp Phe Thr Cys Phe Trp Ala Gly Cys Val Arg Arg Tyr
 225 230 235 240
 Lys Pro Phe Asn Ala Arg Tyr Lys Leu Leu Ile His Met Arg Val His
 245 250 255
 Ser Gly Glu Lys Pro Asn Lys Cys Met Phe Glu Gly Cys Ser Lys Ala
 260 265 270
 Phe Ser Arg Leu Glu Asn Leu Lys Ile His Leu Arg Ser His Thr Gly
 275 280 285
 Glu Lys Pro Tyr Leu Cys Gln His Pro Gly Cys Gln Lys Ala Phe Ser
 290 295 300
 Asn Ser Ser Asp Arg Ala Lys His Gln Arg Thr His Leu Asp Thr Lys
 305 310 315 320
 Pro Tyr Ala Cys Gln Ile Pro Gly Cys Ser Lys Arg Tyr Thr Asp Pro
 325 330 335
 Ser Ser Leu Arg Lys His Val Lys Ala His Ser Ala Lys Glu Gln Gln
 340 345 350
 Val Arg Lys Lys Leu His Ala Gly Pro Asp Thr Glu Ala Asp Val Leu
 355 360 365
 Thr Glu Cys Leu Val Leu Gln Gln Leu His Thr Ser Thr Gln Leu Ala
 370 375 380
 Ala Ser Asp Gly Lys Gly Gly Cys Gly Leu Gly Gln Glu Leu Leu Pro
 385 390 395 400
 Gly Val Tyr Pro Gly Ser Ile Thr Pro His Asn Gly Leu Ala Ser Gly
 405 410 415
 Leu Leu Pro Pro Ala His Asp Val Pro Ser Arg His His Pro Leu Asp
 420 425 430
 Ala Thr Thr Ser Ser His His His Leu Ser Pro Leu Pro Met Ala Glu
 435 440 445
 Ser Thr Arg Asp Gly Leu Gly Pro Gly Leu Leu Ser Pro Ile Val Ser
 450 455 460
 Pro Leu Lys Gly Leu Gly Pro Pro Pro Leu Pro Pro Ser Ser Gln Ser
 465 470 475 480
 His Ser Pro Gly Gly Gln Pro Phe Pro Thr Leu Pro Ser Lys Pro Ser
 485 490 495
 Tyr Pro Pro Phe Gln Ser Pro Pro Pro Pro Pro Leu Pro Ser Pro Gln
 500 505 510
 Gly Tyr Gln Gly Ser Phe His Ser Ile Gln Ser Cys Phe Pro Tyr Gly
 515 520 525
 Asp Cys Tyr Arg Met Ala Glu Pro Ala Ala Gly Gly Asp Gly Leu Val
 530 535 540
 Gly Glu Thr His Gly Phe Asn Pro Leu Arg Pro Asn Gly Tyr His Ser
 545 550 555 560
 Leu Ser Thr Pro Leu Pro Ala Thr Gly Tyr Glu Ala Leu Ala Glu Ala
 565 570 575
 Ser Cys Pro Thr Ala Leu Pro Gln Gln Pro Ser Glu Asp Val Val Ser
 580 585 590
 Ser Gly Pro Glu Asp Cys Gly Phe Phe Pro Asn Gly Ala Phe Asp His
 595 600 605
 Cys Leu Gly His Ile Pro Ser Ile Tyr Thr Asp Thr
 610 615 620

5 <210> 35
 <211> 25
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Cebador oligonucleótido - hKlfGC2 directo
 10 <400> 35
 gcaggaggcg gtctcttcgt gcacc 25
 <210> 36
 <211> 25
 <212> ADN
 15 <213> Secuencia Artificial
 <220>
 <223> Cebador oligonucleótido - hKlf4GC2 inverso
 20 <400> 36
 caggtgtgcc ttgagatggg aactc 25
 <210> 37
 <211> 58
 25 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Cebador oligonucleótido - EcoR1-Sac1-T7M1-VEE
 30 <400> 37
 cggaattcga gctctaatac gactcactat agatgggacg cgcattgagag aagccacg 58
 <210> 38
 35 <211> 38
 <212> ADN
 <213> Secuencia Artificial
 <220>
 40 <223> Cebador oligonucleótido - Xba1-BstZ17I-VEE
 <400> 38
 gctctagagt atacatcctg gtaaacagcg acttgccc 38
 45 <210> 39
 <211> 1056
 <212> ADN
 <213> Vaccinia virus
 50 <220>
 <221> CDS
 <222> (1)..(1056)
 <400> 39

ES 2 698 527 T3

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

ES 2 698 527 T3

Tyr	Asp	Asp	Val	Arg	Ile	Lys	Asn	Asp	Ile	Val	Val	Ser	Arg	Cys	Lys		
225					230					235					240		
ata	ctt	acg	gtt	ata	ccg	tca	caa	gac	cac	agg	ttt	aaa	cta	ata	cta		768
Ile	Leu	Thr	Val	Ile	Pro	Ser	Gln	Asp	His	Arg	Phe	Lys	Leu	Ile	Leu		
			245					250					255				
gat	cca	aaa	atc	aac	gta	acg	ata	gga	gaa	cct	gcc	aat	ata	aca	tgc		816
Asp	Pro	Lys	Ile	Asn	Val	Thr	Ile	Gly	Glu	Pro	Ala	Asn	Ile	Thr	Cys		
		260						265					270				
act	gct	gtg	tca	acg	tca	tta	ttg	att	gac	gat	gta	ctg	att	gaa	tgg		864
Thr	Ala	Val	Ser	Thr	Ser	Leu	Leu	Ile	Asp	Asp	Val	Leu	Ile	Glu	Trp		
		275					280					285					
gaa	aat	cca	tcc	gga	tgg	ctt	ata	gga	ttc	gat	ttt	gat	gta	tac	tct		912
Glu	Asn	Pro	Ser	Gly	Trp		Ile	Gly	Phe	Asp	Phe	Asp	Val	Tyr	Ser		
290						295				300							
ggt	tta	act	agt	aga	ggc	ggt	att	acc	gag	gcg	acc	ttg	tac	ttt	gaa		960
Val	Leu	Thr	Ser	Arg	Gly	Gly	Ile	Thr	Glu	Ala	Thr	Leu	Tyr	Phe	Glu		
305					310				315					320			
aat	ggt	act	gaa	gaa	tat	ata	ggt	aat	aca	tat	aaa	tgt	cgt	gga	cac		1008
Asn	Val	Thr	Glu	Glu	Tyr	Ile	Gly	Asn	Thr	Tyr	Lys	Cys	Arg	Gly	His		
			325					330						335			
aac	tat	tat	ttt	gaa	aaa	acc	ctt	aca	act	aca	gta	gta	ttg	gag	taa		1056
Asn	Tyr	Tyr	Phe	Glu	Lys	Thr	Leu	Thr	Thr	Thr	Val	Val	Leu	Glu			
		340						345					350				

<210> 40
 <211> 351
 <212> PRT
 <213> Vaccinia virus

<400> 40

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

ES 2 698 527 T3

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

REIVINDICACIONES

1. Un ARN replicón de alfavirus que comprende:

al menos un dominio de replicasa no estructural procedente de un alfavirus y al menos una secuencia heteróloga no de alfavirus que codifica al menos dos factores de reprogramación (RFs) para inducir la generación de células madre pluripotentes cuando se expresa en una célula somática, en donde los al menos dos RFs son seleccionados de la lista que consiste en KLF4, OCT-4, SOX-2, c-MYC o n-MYC o l-MYC, GLIS1 y NANOG.

2. El ARN replicón de alfavirus de la reivindicación 1, en donde el replicón comprende secuencias obtenidas de un alfavirus seleccionado del grupo que consiste en virus de la encefalitis equina oriental (EEE), virus de la encefalitis equina venezolana (VEE), virus Everglades, virus Mucambo, virus Pixuna y virus de la encefalitis equina occidental (WEE).

3. El ARN replicón de alfavirus de la reivindicación 1, en donde el replicón comprende secuencias de un alfavirus seleccionado del grupo que consiste en virus de Sindbis, virus de Bosque de Semliki, virus Middelburg, virus Chikungunya, virus O'nyong-nyong, virus Ross River, virus del Bosque de Barmah, virus Getah, virus Sagiyama, virus Bebaru, virus Mayaro, virus Una, virus Aura, virus Whataroa, virus Babanki, virus Kyzylagach, virus J de las Highlands, virus Fort Morgan, virus Ndumu y virus Buggy Creek.

4. El ARN replicón de alfavirus de cualquier reivindicación precedente, en donde el replicón comprende secuencias obtenidas del virus de la encefalitis equina venezolana.

5. El ARN replicón de alfavirus de la reivindicación 1, en donde la al menos una secuencia heteróloga no de alfavirus comprende al menos 2, 3, 4 o 5 secuencias heterólogas no de alfavirus.

6. El ARN replicón de alfavirus de una cualquiera de las reivindicaciones 1-5,

(i) en donde el polinucleótido que codifica el polipéptido KLF4 codifica un polipéptido KLF4 que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:8; o

(ii) en donde el polinucleótido que codifica el polipéptido KLF4 codifica un polipéptido KLF4 que tiene una secuencia de la SEQ ID NO:8; o

(iii) en donde el polinucleótido que codifica el polipéptido KLF4 comprende una secuencia como la establecida en la SEQ ID NO:7; en donde "T" es "U"; o

(iv) en donde el polinucleótido que codifica el polipéptido SOX-2 codifica un polipéptido SOX-2 que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:6; o

(v) en donde el polinucleótido que codifica el polipéptido SOX-2 codifica un polipéptido SOX-2 que tiene una secuencia de la SEQ ID NO:6; o

(vi) en donde el polinucleótido que codifica el polipéptido Sox-2 comprende una secuencia como la establecida en la SEQ ID NO:5; en donde "T" es "U"; o

(vii) en donde el polinucleótido que codifica el polipéptido OCT-4 codifica un polipéptido OCT-4 que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:4; o

(viii) en donde el polinucleótido que codifica el polipéptido OCT-4 codifica un polipéptido OCT-4 que tiene una secuencia de la SEQ ID NO:4; o

(ix) en donde el polinucleótido que codifica el polipéptido OCT-4 comprende una secuencia como la establecida en la SEQ ID NO:3; en donde "T" es "U"; o

(x) en donde el polinucleótido que codifica el polipéptido c-MYC codifica un polipéptido c-MYC que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:10; o

(xi) en donde el polinucleótido que codifica el polipéptido c-MYC codifica un polipéptido c-MYC que tiene una secuencia de la SEQ ID NO:10; o

(xii) en donde el polinucleótido que codifica el polipéptido c-MYC comprende una secuencia como la establecida en la SEQ ID NO:9; en donde "T" es "U"; o

(xiii) en donde el polinucleótido que codifica el polipéptido GLIS1 codifica un polipéptido GLIS1 que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:34; o

(xiv) en donde el polinucleótido que codifica el polipéptido GLIS1 codifica un polipéptido GLIS1 que tiene una secuencia de la SEQ ID NO:34; o

(xv) en donde el polinucleótido que codifica el polipéptido GLIS1 comprende una secuencia como la establecida en la SEQ ID NO:33; en donde "T" es "U"; o

(xvi) en donde el polinucleótido que codifica el polipéptido NANOG codifica un polipéptido NANOG que tiene al menos un 95% de identidad con respecto a una secuencia de la SEQ ID NO:2; o

5 (xvii) en donde el polinucleótido que codifica el polipéptido NANOG codifica un polipéptido NANOG que tiene una secuencia de la SEQ ID NO:2; o

(xviii) en donde el polinucleótido que codifica el polipéptido NANOG comprende una secuencia como la establecida en la SEQ ID NO:1; en donde "T" es "U".

10 **7.** El ARN replicón de alfavirus de la reivindicación 1, en donde el replicón comprende de 5' a 3': (replicasa de ARN de VEE) – (promotor) – (RF₁) – (péptido de auto-ruptura) – (RF₂) – (péptido de auto-ruptura) – (RF₃) – (IRES o promotor central) – (RF₄) – (IRES o promotor opcional) – (marcador seleccionable opcional) – (3'UTR y cola poliA de VEE) – (marcador seleccionable opcional) – promotor; donde RF₁₋₄ son factores que inducen la des-diferenciación de una célula somática en células pluripotentes, donde RF₂₋₃ son opcionales, RF₃₋₄ son opcionales o RF₄ es opcional; en donde RF₁₋₄ se seleccionan del grupo que consiste en Oct-4, Klf4, Sox-2, c-Myc, Nanog y Glis1.

15 **8.** El ARN replicón de alfavirus de la reivindicación 1 o 7, en donde el replicón comprende una secuencia que es idéntica en un 90%, 95%, 98%, 99% o 100% con respecto a la SEQ ID NO:29, 30, 31 o 32 desde aproximadamente la posición 1 hasta aproximadamente la posición 7561, en donde "T" de la secuencia es sustituido por "U", seguido de dos o más RFs, seguido de un 3'UTR y cola poliA, donde los dos o más RFs se seleccionan del grupo que consiste en Oct-4, Sox-2, Klf4, c-Myc, Nanog y Glis1; en donde las secuencias codificadoras de los dos o más RFs pueden estar separadas por un sitio de entrada ribosomal interno (IRES) o un promotor pequeño, preferiblemente en donde el replicón comprende una secuencia que es idéntica en al menos un 95%, 98%, 99% o 100% a una secuencia seleccionada del grupo que consiste en la SEQ ID NO:29, 30, 31 o 32, en donde "T" es "U".

20 **9.** Una composición que comprende células humanas que comprenden un replicón de una cualquiera de las reivindicaciones 1-8, que preferiblemente comprende además medio acondicionado con B18R.

25 **10.** La composición de la reivindicación 9 en medio acondicionado con B18R, en donde las células humanas son células somáticas, o en donde las células humanas son fibroblastos.

11. Un método *in vitro* para preparar células madre que comprende cultivar la composición de la reivindicación 9, durante al menos 30 días en las condiciones que expresen las secuencias codificadoras del replicón y aislar las células madre.

30 **12.** Un método *in vitro* para preparar células madre que comprende transformar células somáticas con un replicón de cualquiera de las reivindicaciones 1-8 *in vitro*, cultivar las células somáticas en medio acondicionado con B18R en las condiciones que promueven la expresión del replicón y aislar las células madre.

13. El método de la reivindicación 11 o 12, en donde el medio acondicionado con B18R es producido mediante transfección de ARNm de B18R en fibroblastos humanos primarios.

35 **14.** Un método *in vitro* para preparar células madre que comprende:

poner en contacto una célula somática humana con un replicón de ARN de alfavirus autorreplicante ectópico que comprende un polinucleótido que codifica al menos de dos a cuatro factores de des-diferenciación seleccionados del grupo que consiste en un (i) KLF4, (ii) OCT4, (iii) SOX2, (iv) c-MYC o n-MYC o L-MYC, (v) GLIS1 y (vi) NANOG, *in vitro*,

40 cultivar la célula somática en medio acondicionado con B18R para expresar el factor de des-diferenciación;

seleccionar las células que presentan una morfología de célula madre y/o marcadores de célula madre; y

subcultivar las células para obtener una población de células madre inducidas, preferiblemente en donde las células son seleccionadas detectando la expresión de un antígeno de rechazo tumoral 1-60 y/o 1-81.

15. Un sistema de vector para producir células madre humanas, que comprende:

45 al menos un replicón de ARN autorreplicante derivado de un alfavirus, preferiblemente en donde el alfavirus es VEE, en donde el replicón comprende un polinucleótido que codifica al menos dos factores de des-diferenciación seleccionados del grupo que consiste en KLF4, OCT4, SOX2, c-MYC o n-MYC o L-MYC, GLIS1, NANOG o cualquier combinación de los mismos, preferiblemente en donde el replicón comprende:

- (a) Oct4, Sox2, Klf4 y c-Myc, o
- (b) Oct4, Sox2, Klf4 y Glis1.

50

- 5 **16.** Una célula o población de células somáticas humanas aisladas que comprenden un replicón de ARN de alfavirus ectópico que comprende dos o más secuencias de polinucleótido de des-diferenciación que codifican factores de des-diferenciación seleccionados del grupo que consiste en KLF4, OCT-4, SOX-2, c-MYC o n-MYC o L-MYC, GLIS1 y NANOG, preferiblemente en donde, en las condiciones de cultivo que expresan los polinucleótidos de des-diferenciación en el replicón de ARN de alfavirus ectópico, la célula somática se des-diferencia.
- 17.** Un medio que comprende una población de células de la reivindicación 16, en donde el medio inhibe la degradación del replicón de ARN de alfavirus ectópico.

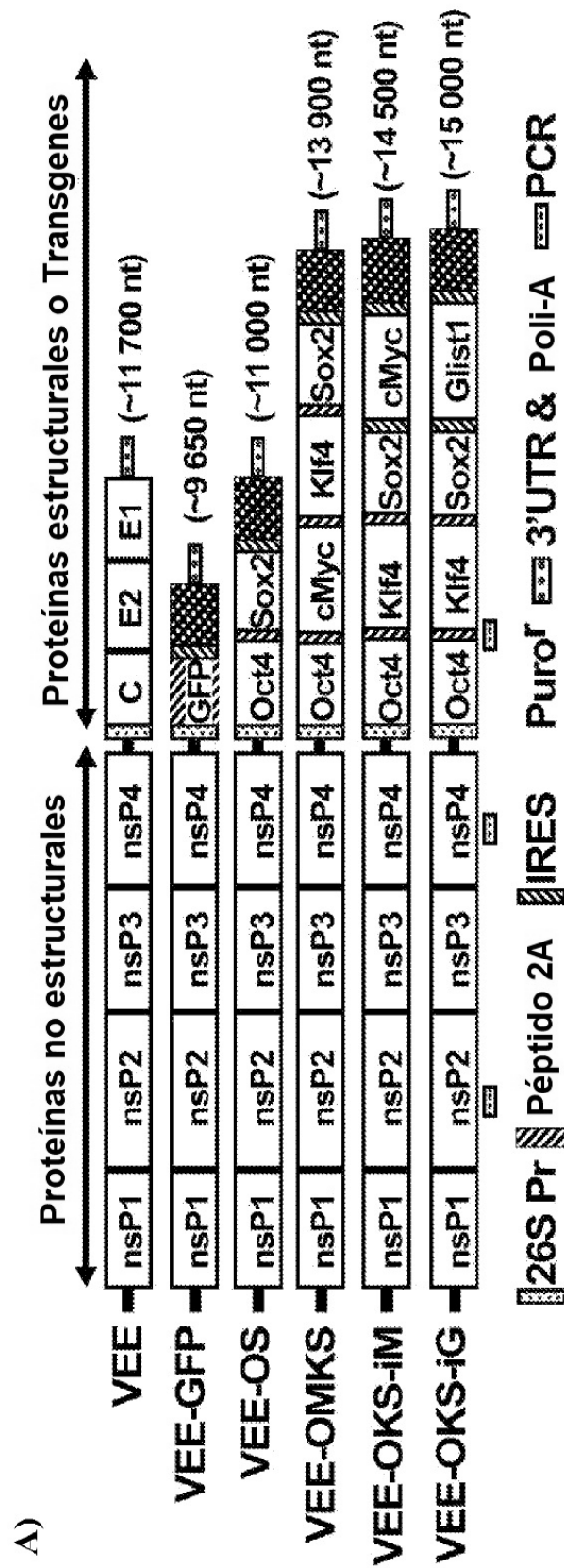


FIGURA 1A

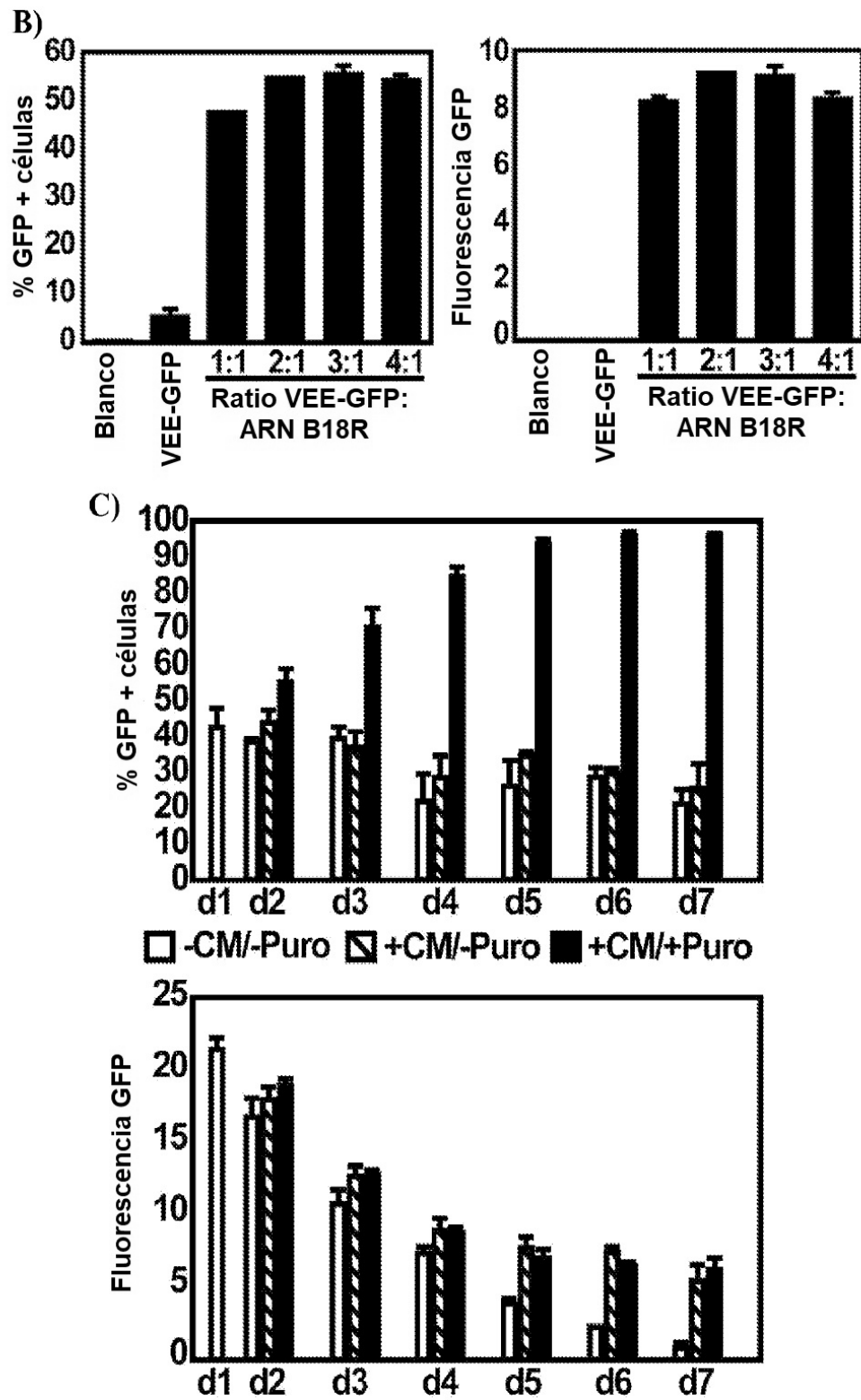


FIGURA 1B-C

D)

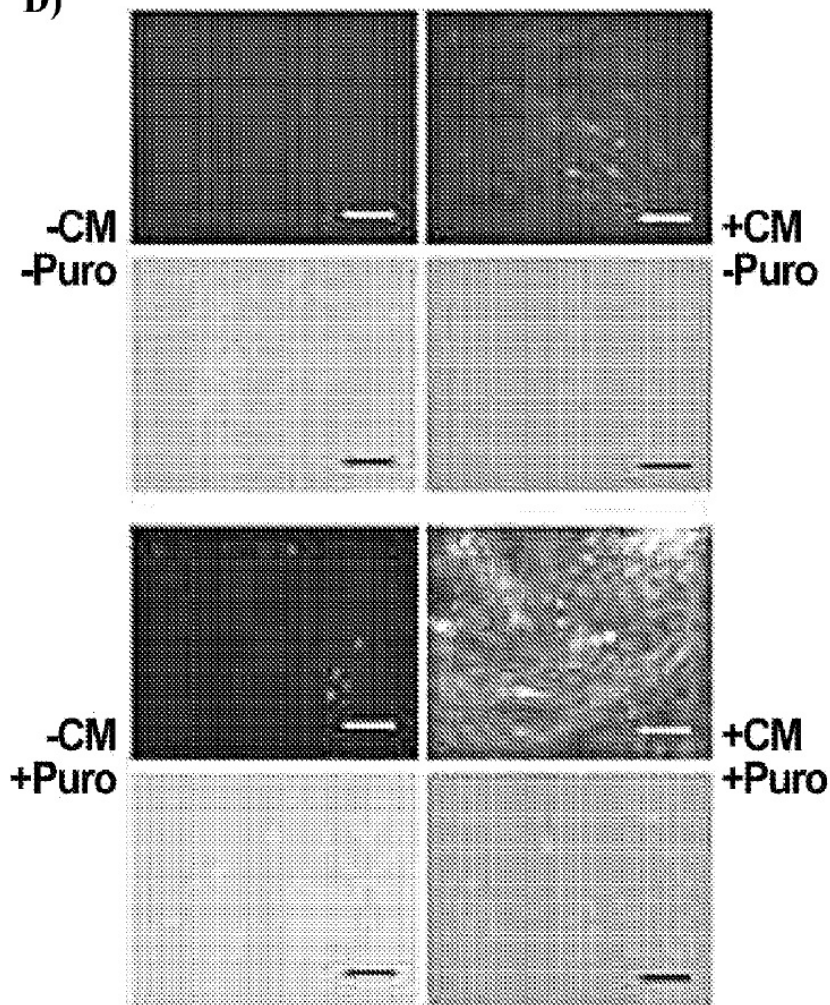


FIGURA 1D

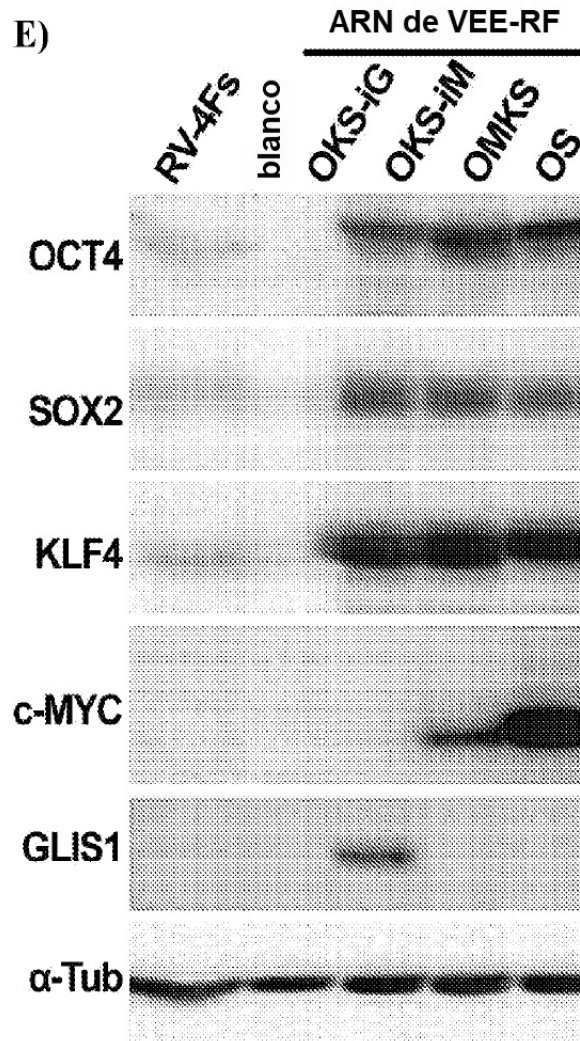


FIGURA 1E

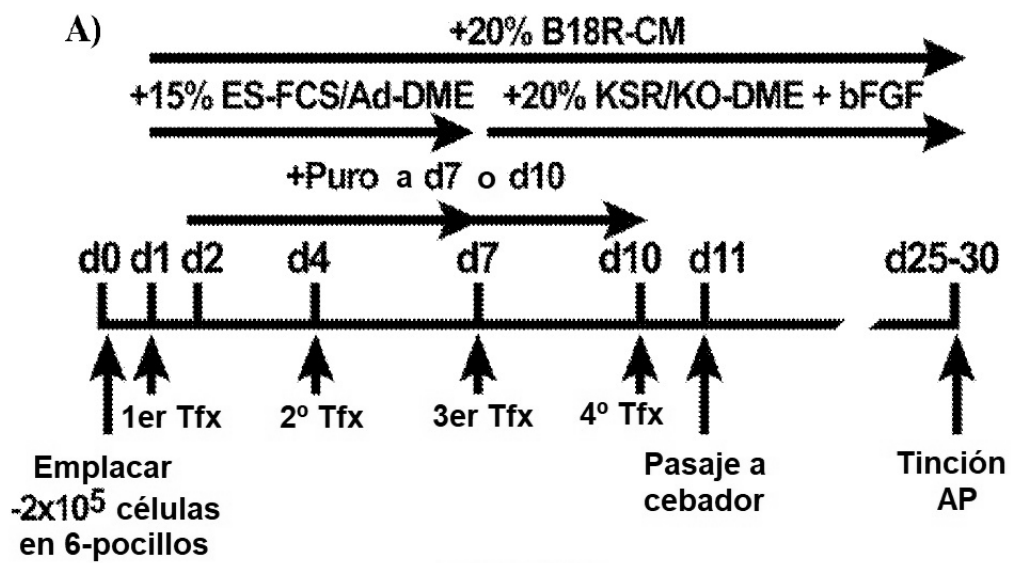


FIGURA 2A

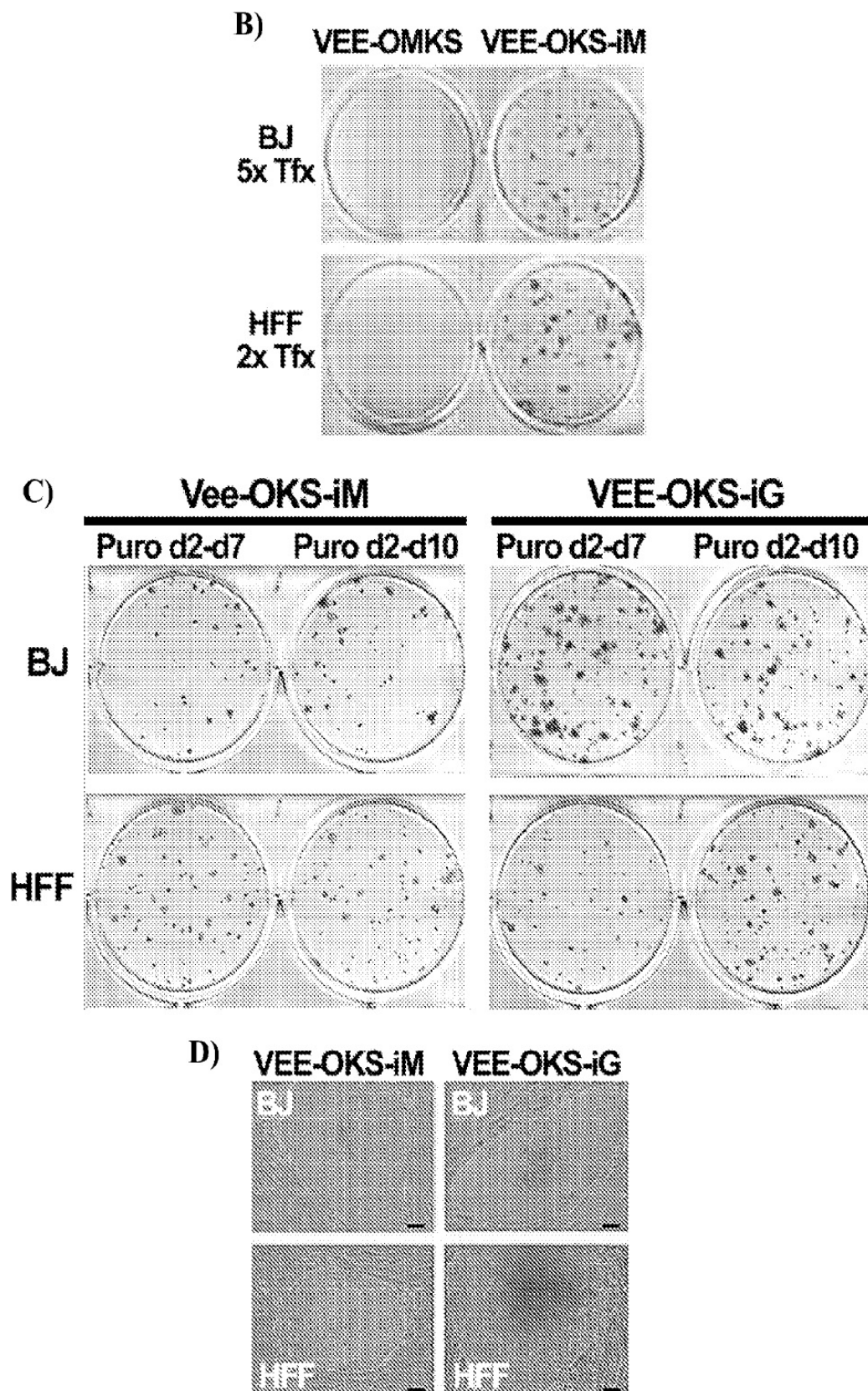


FIGURA 2B-D

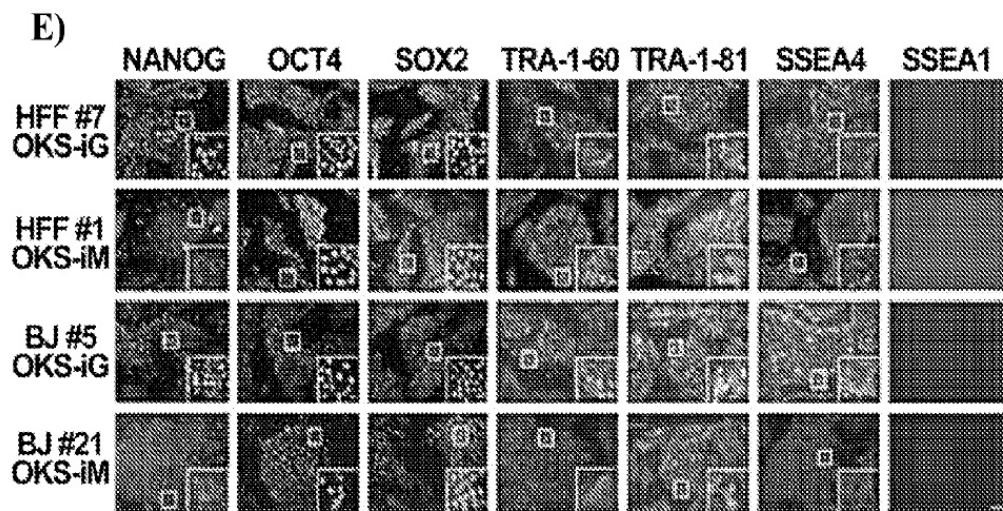


FIGURA 2E

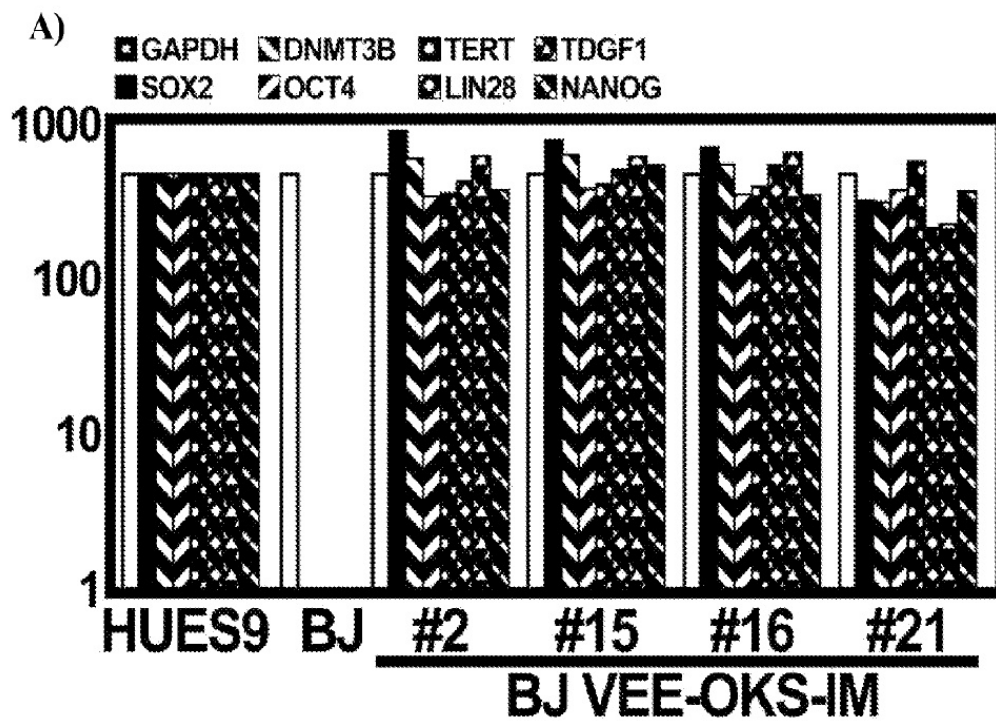


FIGURA 3A

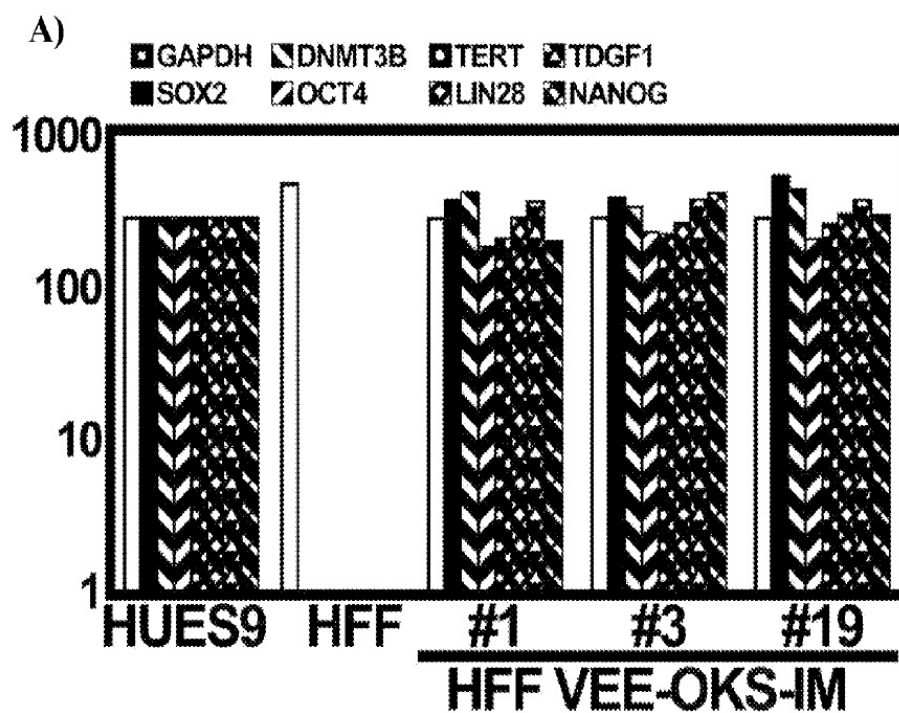


FIGURA 3A (continuación)

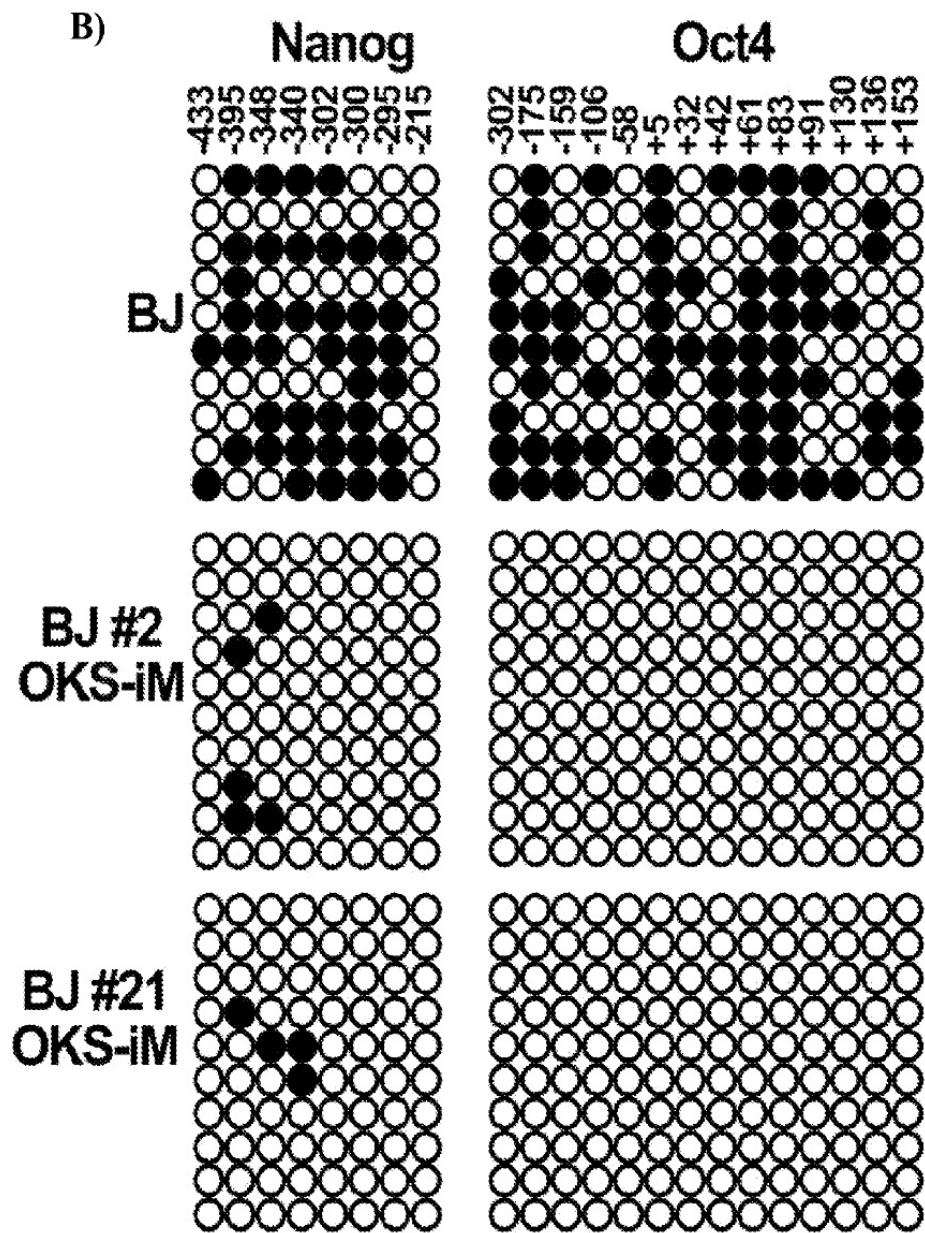


FIGURA 3B

C)

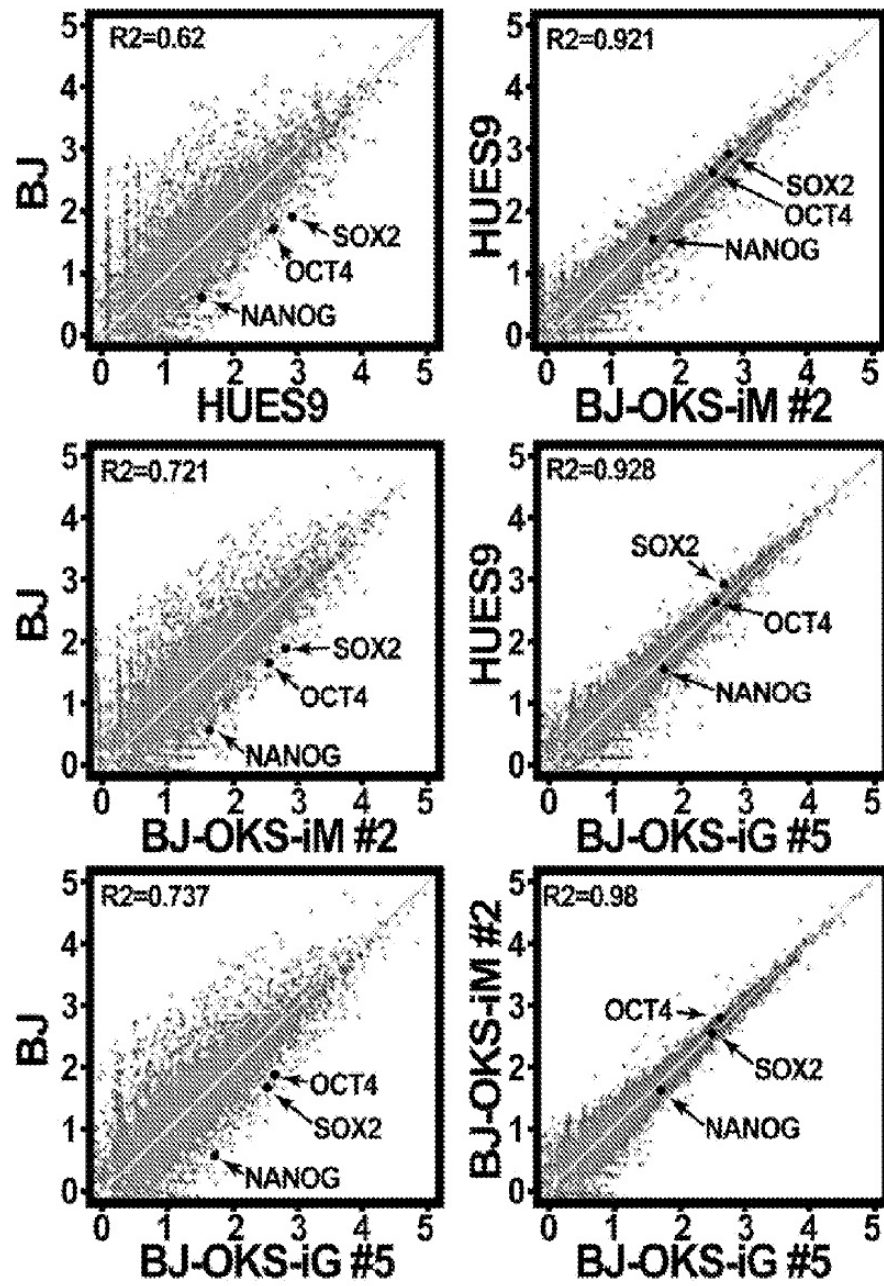
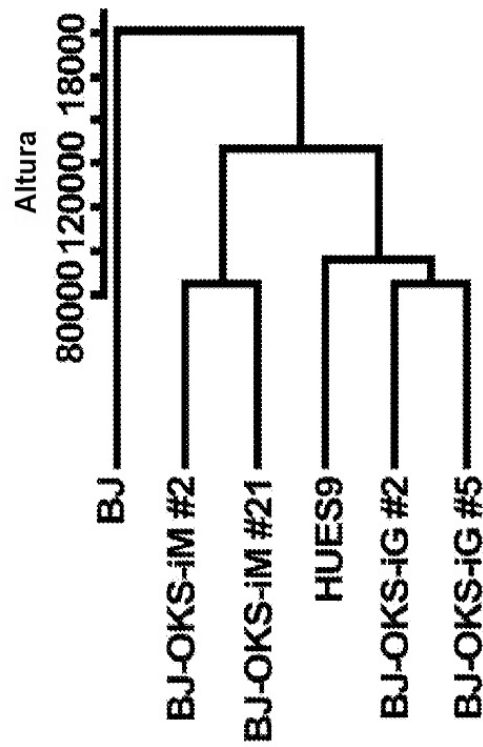


FIGURA 3C

D)



E)

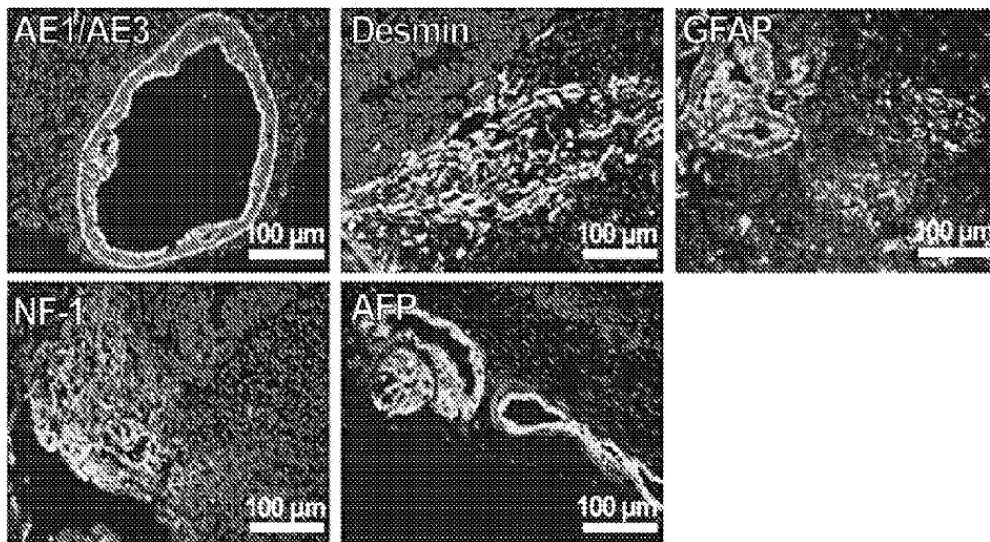


FIGURA 3D-E

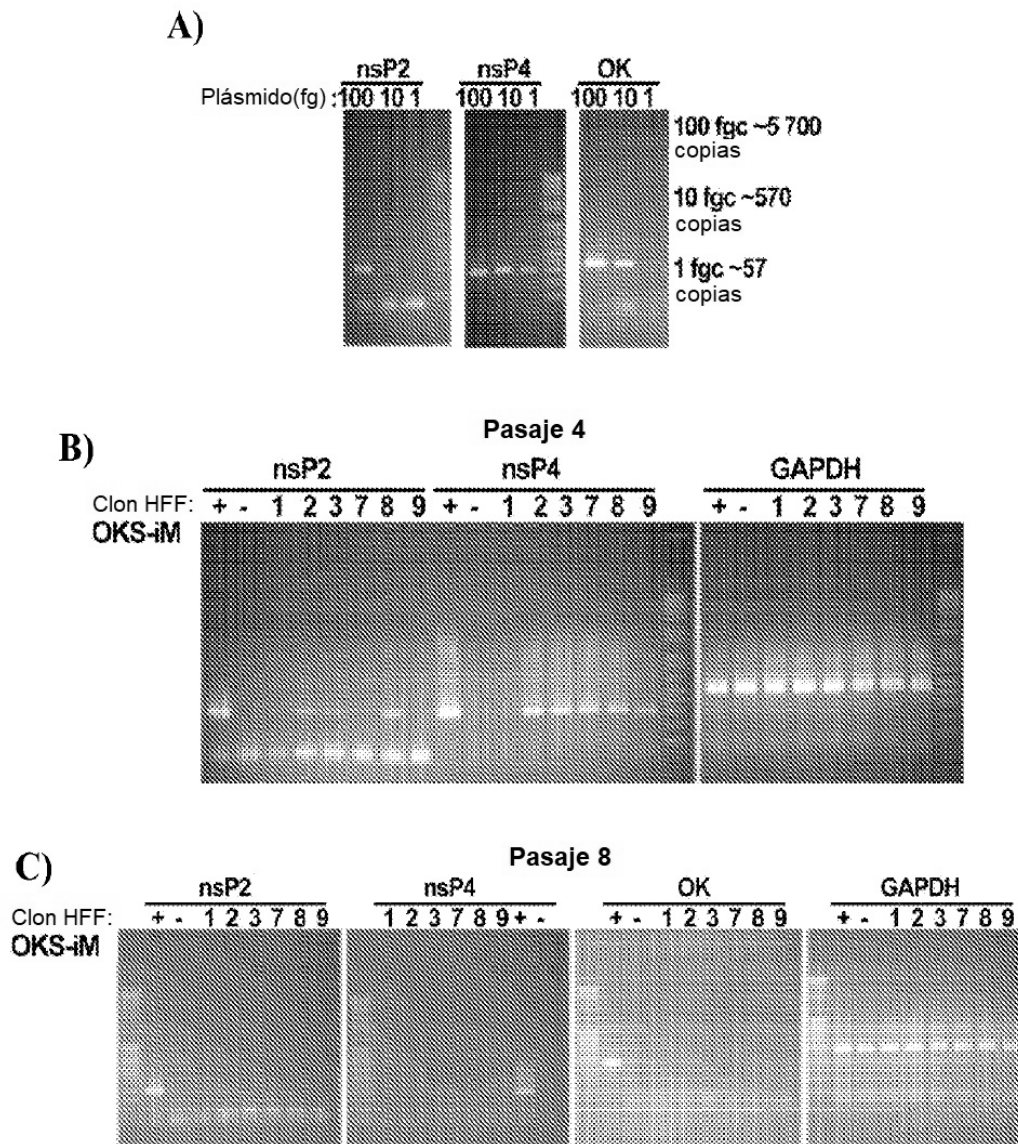


FIGURA 4A-C

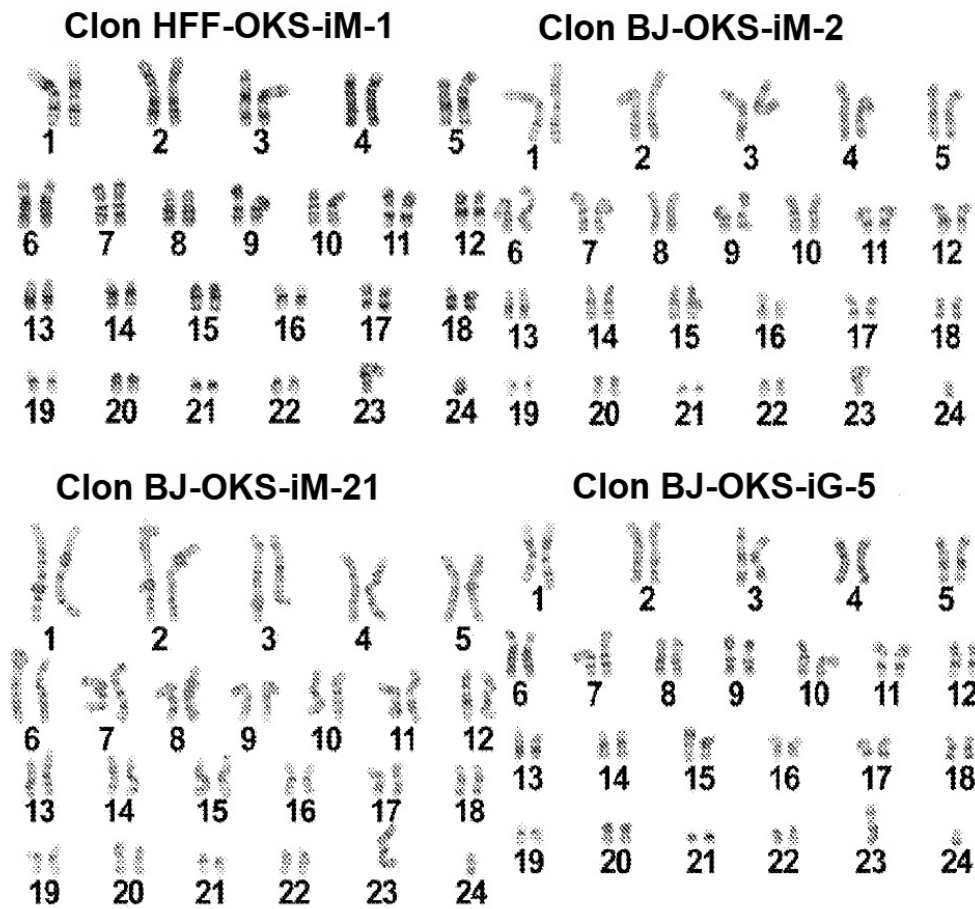


FIGURA 5

A)

Clon HFF-OKS-iM n°1

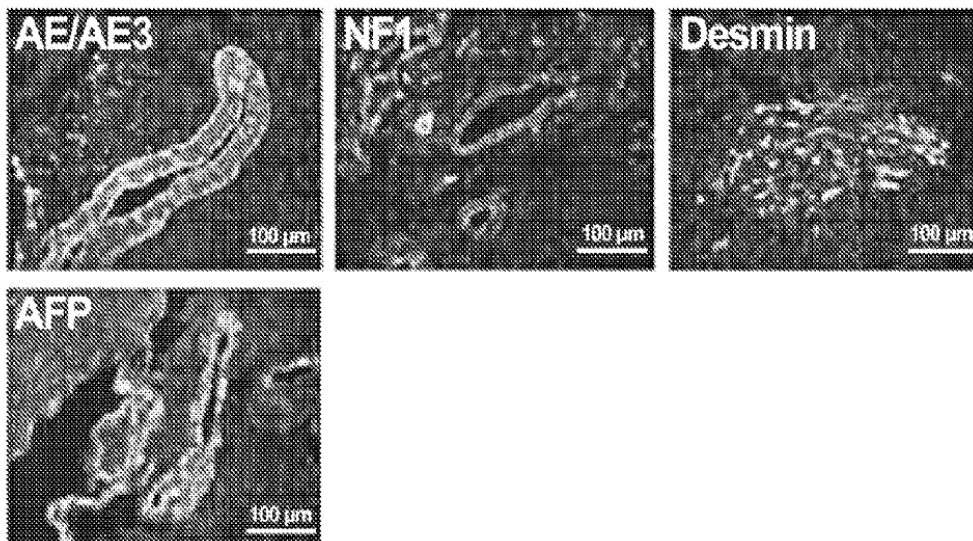


FIGURA 6A

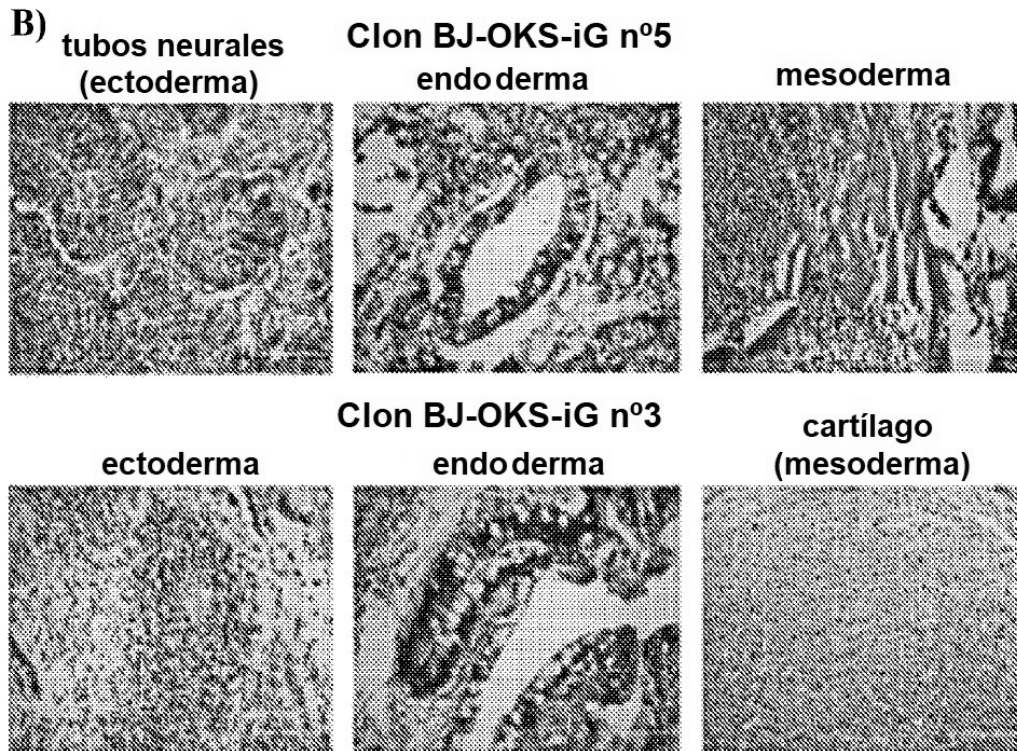


FIGURA 6B

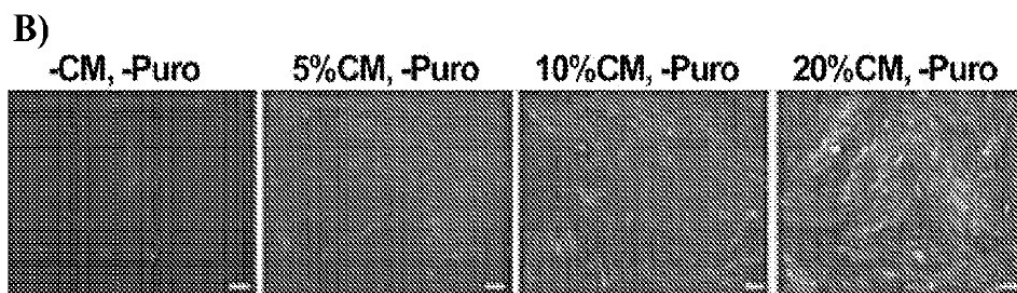
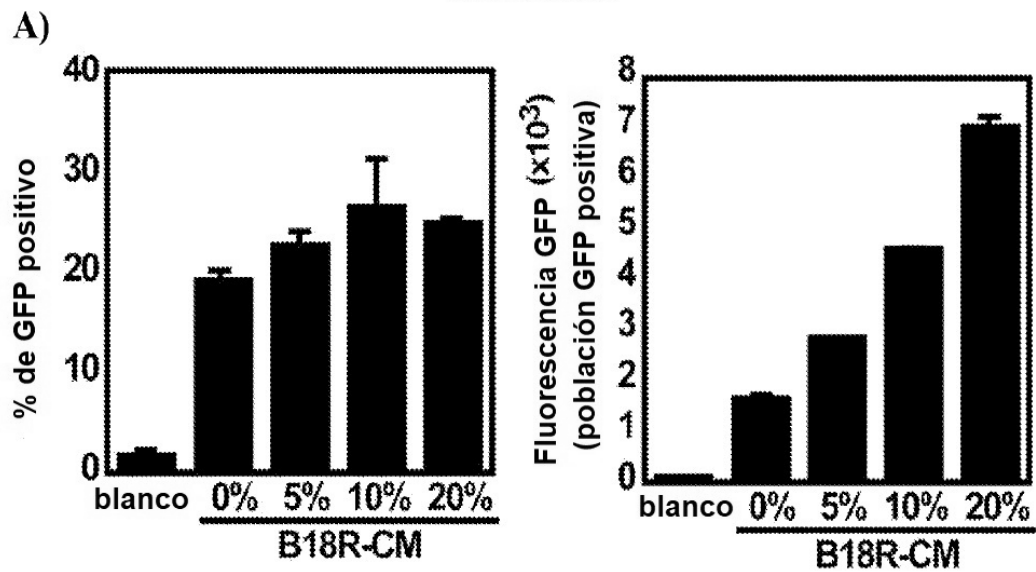
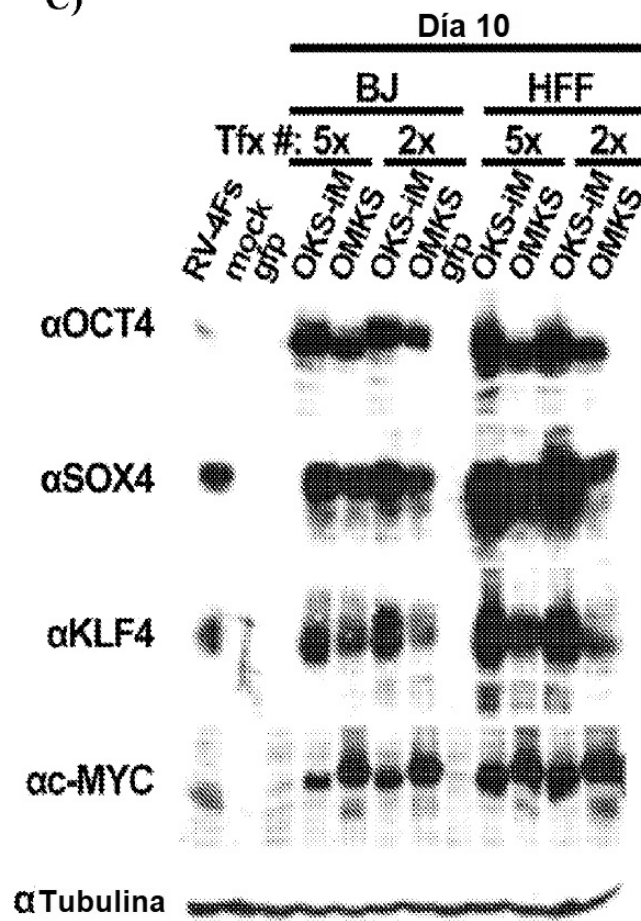


FIGURA 7A-B

C)



D)

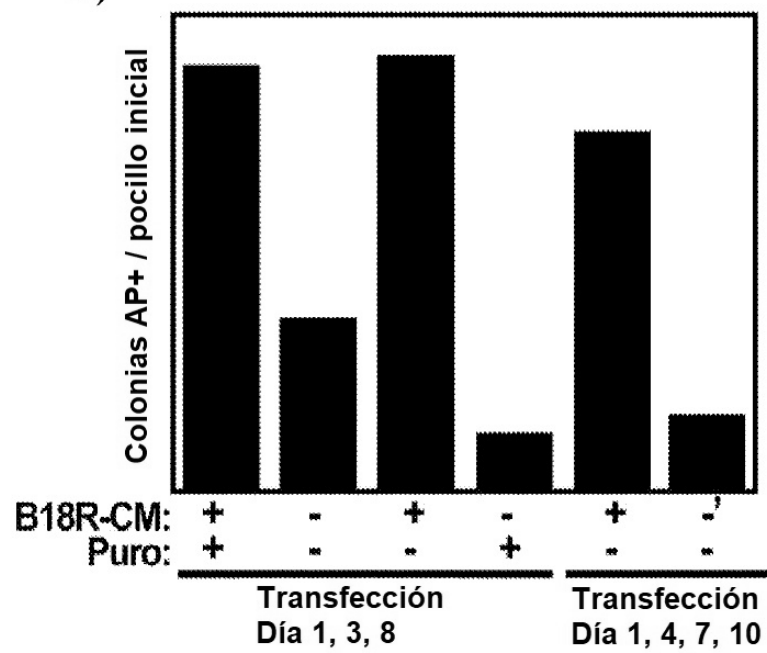


FIGURA 7C-D