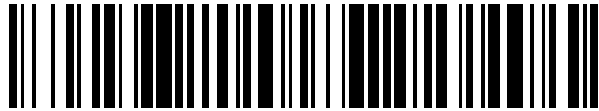


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 587 395**

51 Int. Cl.:

**C12N 5/10** (2006.01)

**C12N 15/869** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **04.06.2009 PCT/US2009/046209**
- 87 Fecha y número de publicación internacional: **10.12.2009 WO09149233**
- 96 Fecha de presentación y número de la solicitud europea: **04.06.2009 E 09759392 (5)**
- 97 Fecha y número de publicación de la concesión europea: **01.06.2016 EP 2297307**

54 Título: **Procedimientos para la producción de células IPS usando un enfoque no vírico**

30 Prioridad:

**04.06.2008 US 58858 P**  
**16.03.2009 US 160584 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**24.10.2016**

73 Titular/es:

**CELLULAR DYNAMICS INTERNATIONAL, INC.**  
**(100.0%)**  
**University Research Park 525 Science Drive,**  
**Suite 200**  
**Madison, WI 53711, US**

72 Inventor/es:

**MACK, AMANDA y**  
**THOMSON, JAMES**

74 Agente/Representante:

**ARIAS SANZ, Juan**

ES 2 587 395 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

**DESCRIPCIÓN**

Procedimientos para la producción de células IPS usando un enfoque no vírico

**ANTECEDENTES**

5 Esta solicitud reivindica la prioridad de la solicitud de EE.UU. n° 61/058.858, presentada el 4 de junio de 2008, y de la solicitud de EE.UU. n° 61/160.584, presentada el 16 de marzo de 2009.

**1. Campo de la invención**

La presente invención se refiere en general al campo de la biología molecular, las células madre y las células diferenciadas. Más particularmente, atañe a la programación o reprogramación de la diferenciación de células somáticas y células indiferenciadas.

**10 2. Descripción de la materia relacionada**

15 En general, las células madre son células indiferenciadas que pueden dar lugar a una sucesión de células funcionales maduras. Por ejemplo, una célula madre hematopoyética puede dar lugar a cualquiera de los diferentes tipos de células sanguíneas diferenciadas terminalmente. Las células madre embrionarias (ES) derivan del embrión y son pluripotentes, poseyendo por tanto la capacidad de desarrollarse hasta cualquier tipo de órgano o tejido o, al menos potencialmente, hasta un embrión completo.

20 Los células madre pluripotentes inducidas, abreviadas comúnmente como iPS o iPSC, son un tipo de célula madre pluripotente derivado artificialmente de una célula no pluripotente, típicamente una célula somática adulta, mediante la inserción de ciertos genes. Las células madre pluripotentes inducidas se cree que son idénticas a las células madre pluripotentes naturales, tales como células madre embrionarias, en muchos aspectos, tales como en términos de la expresión de ciertos genes y proteínas de citoblasto, patrones de metilación de cromatina, tiempo de duplicación, formación de cuerpo embrioide, formación de teratoma, formación de quimera viable y potencia y diferenciabilidad, pero se está valorando todavía la medida total de su relación con las células madre pluripotentes naturales.

25 Las células IPS se produjeron por primera vez en 2006 (Takahashi *et al.*, 2006) a partir de células de ratón y en 2007 a partir de células humanas (Takahashi *et al.*, 2007; Takahashi *et al.*, 2007). Esto se ha citado como un importante avance en la investigación de células madre, ya que puede permitir a los investigadores obtener células madre pluripotentes, que son importantes en la investigación y potencialmente tienen usos terapéuticos, sin el uso controvertido de embriones.

30 Sin embargo, en esta etapa del estudio de estas células madre pluripotentes inducidas (iPS), los investigadores están usando plásmidos víricos integrativos, que insertan los genes en el genoma de células diana, introduciendo potencialmente mutaciones en el sitio de inserción. Por lo tanto, existe la necesidad de desarrollar un procedimiento para inducir células madre pluripotentes esencialmente exentos de componentes víricos exógenos.

35 Debido al potencial médico significativo de la terapia celular y el trasplante de tejidos, existe también la necesidad urgente de producción de cualquier tipo celular deseado mediante la alteración del estado de diferenciación celular de una población celular disponible. Cada tipo celular especializado en un organismo expresa un subconjunto de todos los genes que constituyen el genoma de esa especie. Cada tipo celular se define por su patrón particular de expresión génica regulada. La diferenciación celular es por tanto la transición de una célula de un tipo celular a otro e implica un cambio de un patrón de expresión génica a otro. La diferenciación celular durante el desarrollo puede entenderse como el resultado de una red reguladora génica. Un gen regulador y sus módulos reguladores en cis son los nodos de la red reguladora génica; reciben entradas y crean salidas en otro lugar de la red. Pueden aplicarse también mecanismos similares a la desdiferenciación, por ejemplo induciendo pluripotencia a partir de células somáticas como se mencionan anteriormente, y a la transdiferenciación, que hace referencia específicamente a la transformación de un tipo celular diferenciado en otro. Se han estudiado factores de transcripción que controlan las elecciones de desarrollo para cambiar el estado de diferenciación; sin embargo, se han usado ampliamente también 40 vectores víricos. Por lo tanto, existe la necesidad de procedimientos de programación de la diferenciación exentos de virus mejorados.

**SUMARIO DE LA INVENCION**

La invención se refiere a realizaciones como se definen en las reivindicaciones.

50 La presente invención se refiere a un procedimiento para proporcionar una población celular que tiene un estado de diferenciación alterado respecto a una población celular de partida y que tiene células que están esencialmente exentas de elementos genéticos vectoriales de programación. El procedimiento comprende las etapas de: (a) obtener una población de partida de células que tienen un primer estado de diferenciación, (b) obtener uno o más vectores de programación de la diferenciación, comprendiendo cada vector un origen de replicación y uno más módulos de expresión que codifican uno o más factores de programación de la diferenciación que, en combinación,

5 pueden alterar el estado de diferenciación de la población celular de partida hasta un segundo estado de diferenciación, en el que uno o más de dichos módulos de expresión comprenden una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación para replicar un molde extracromosómico y/o en el que las células de la población de partida expresan dicho factor de acción en trans, (c) introducir el vector o  
 10 vectores de programación de la diferenciación en células de la población de partida, (d) cultivar las células para efectuar la expresión de dicho uno o más factores de programación de la diferenciación de tal modo que surjan los atributos consistentes con el segundo estado de diferenciación en al menos una porción de las células cultivadas y (e) cultivar además células que tienen los atributos durante un número suficiente de generaciones para proporcionar una población celular diana que comprenda células que tienen el segundo estado de diferenciación, pero dichas células están esencialmente exentas de elementos genéticos vectoriales de programación. Además, la invención se refiere a un vector de programación de la diferenciación que comprende un origen de replicación y uno o más módulos de expresión que codifican un factor de acción en trans que se une al origen de replicación para replicar el vector extracromosómicamente, y uno o más factores de programación de la diferenciación.

15 La presente invención supera una importante deficiencia en la técnica al proporcionar células madre pluripotentes inducidas y otros tipos celulares esencialmente exentos de elementos vectoriales exógenos mediante programación de la diferenciación. Por consiguiente, la presente invención se refiere a un procedimiento para proporcionar una población celular que tiene un estado de diferenciación alterado respecto a una población celular de partida y que tiene células que están esencialmente exentas de elementos genéticos vectoriales de programación, comprendiendo el procedimiento las etapas de: (a) obtener una población de partida de células que tienen un primer estado de  
 20 diferenciación; (b) obtener uno o más vectores de programación de la diferenciación, comprendiendo cada vector un origen de replicación y uno o más módulos de expresión que codifican uno o más factores de programación de la diferenciación que, en combinación, pueden alterar el estado de diferenciación de la población celular de partida hasta un segundo estado de diferenciación, en el que uno o más de dichos módulos de expresión comprenden una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación para replicar un molde extracromosómico y/o en el que las células de la población de partida expresan dicho factor de acción en trans; (c) introducir el vector o vectores de programación de la diferenciación en células de la población de partida; (d) cultivar las células para efectuar la expresión de dicho uno o más factores de programación de la diferenciación de tal modo que surjan los atributos consistentes con el segundo estado de diferenciación en al menos una porción de células en las células cultivadas y (e) cultivar además las células que tienen los atributos durante un número  
 25 suficiente de generaciones para proporcionar una población celular diana que comprenda células que tienen el segundo estado de diferenciación, pero estando dichas células esencialmente exentas de elementos genéticos vectoriales de programación. En una primera realización, se proporciona un procedimiento para producir una población celular de células madre pluripotentes inducidas (iPS), comprendiendo el procedimiento las etapas de: a) obtener un vector de reprogramación, elemento vectorial que comprende un origen de replicación y uno o más módulos de expresión que codifican factores de reprogramación de iPS; b) introducir el vector de reprogramación en células de una población de células somáticas; c) cultivar las células para expandir la población; d) seleccionar las células de progenie de dicha población expandida, en el que dicha progenie tiene una o más características de las células madre embrionarias y e) cultivar las células de progenie seleccionadas para proporcionar la población de iPS, en el que uno o más de dichos módulos de expresión comprenden una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación para replicar un molde extracromosómico y/o en el que las células somáticas expresan dicho factor de acción en trans. En un aspecto adicional, la etapa c o la etapa e comprenden además cultivar hasta que las células estén esencialmente exentas de elementos vectoriales o comprendan una etapa de selección adicional como se describe a continuación para facilitar la generación de iPS exentos de vector.

45 En ciertos aspectos, para replicar un molde extracromosómico, uno o más de los módulos de expresión comprenden una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación; como alternativa, las células somáticas comprenden una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación.

50 En realizaciones ejemplares, el origen de replicación puede ser un origen de replicación de un herpesvirus linfotrófico o un gammaherpesvirus, un adenovirus, SV40, un papilomavirus bovino o una levadura, tal como el origen de replicación de un herpesvirus linfotrófico o un gammaherpesvirus correspondiente a oriP de EBV. En un aspecto adicional, el herpesvirus linfotrófico puede ser virus de Epstein Barr (EBV), herpesvirus de sarcoma de Kaposi (KSHV), herpesvirus de Saimiri (HS) o virus de la enfermedad de Marek (MDV). En otro aspecto adicional, el gamaherpesvirus puede ser virus de Epstein Barr (EBV) o herpesvirus de sarcoma de Kaposi (KSHV).

55 En ciertas realizaciones, el factor de acción en trans puede ser un polipéptido correspondiente a, o derivado de, una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, preferiblemente en presencia de un origen de replicación correspondiente a OriP de EBV. El derivado puede tener una capacidad reducida de activar la transcripción a partir de un molde integrado en comparación con EBNA-1 de tipo silvestre y por tanto posibilidades reducidas de activar ectópicamente genes cromosómicos para causar transformación oncogénica. Al mismo tiempo,  
 60 el derivado puede activar la transcripción al menos a un 5 % de la correspondiente proteína de tipo silvestre a partir de un molde extracromosómico después de que el derivado se una al origen de replicación. Dicho derivado puede tener una delección de residuos correspondiente a los residuos aproximadamente 65 a aproximadamente 89 de

- EBNA-1 de tipo silvestre (SEQ ID NO: 1, que hace referencia a la secuencia de proteína EBNA-1 de tipo silvestre que está codificada por la SEQ ID NO: 2), y/o tiene una delección de residuos correspondiente a los residuos aproximadamente 90 a aproximadamente 328 de EBNA-1 (SEQ ID NO: 1), o puede ser un derivado con al menos un 80 % de identidad de secuencia aminoacídica con los residuos 1 a aproximadamente 40 y los residuos aproximadamente 328 a 641 de EBNA-1 (SEQ ID NO: 1). Los aminoácidos 90-328 de EBNA-1 de tipo silvestre comprenden una región rica en repeticiones Gly-Ala que no debería contribuir significativamente a la función de EBNA-1 en la presente invención, y por lo tanto esta región puede ser variable en términos del número de repeticiones presentes (concretamente, la región puede eliminarse toda o en parte). Un derivado ejemplar de EBNA-1 de tipo silvestre puede tener una secuencia de SEQ ID NO: 3, que está codificada por la SEQ ID NO: 4.
- En ciertas realizaciones adicionales, la invención implica una etapa adicional de selección de células de progenie de la población expandida, en la que la progenie está esencialmente exenta de elementos vectoriales. Debido a que los vectores replicados extracromosómicamente, tales como los vectores basados en OriP, se perderán por las células con el tiempo tal como durante las dos semanas después de la transfección, y a que los iPS no necesitan factores de reprogramación exógenos después de entrar en un estado pluripotencial automantenido, esta etapa de selección adicional opcional puede ayudar a acelerar la generación de células madre pluripotentes exentas de vector. Por lo tanto, la etapa adicional puede ser en un momento después de que las células de progenie entren en un estado pluripotente automantenido, tal como al menos de aproximadamente 10 días a al menos 30 días después de introducir los vectores de reprogramación en las células. Para facilitar el proceso de generar iPS exentos de elementos vectoriales, el vector de reprogramación puede comprender además una secuencia nucleotídica que codifica un marcador de selección negativa, y la etapa adicional selecciona células de progenie de la población expandida eliminando las células de progenie que comprenden el marcador de selección con un agente de selección. Por ejemplo, el marcador de selección puede codificar timidina cinasa de herpesvirus simple que permite, por aplicación de un agente de selección tal como ganciclovir, retirar las células con vectores residuales que codifican la cinasa. En ciertos aspectos, la población de iPS generada por los procedimientos está esencialmente exenta del marcador de selección. Es un enfoque alternativo o complementario ensayar la ausencia de elementos genéticos exógenos en las células de progenie usando procedimientos convencionales, tales como PCR-TI, PCR, FISH (hibridación fluorescente *in situ*), matriz génica o hibridación (p.ej. transferencia Southern).
- En algunas realizaciones, la población celular de iPS generada a partir de los procedimientos anteriores puede estar esencialmente exenta de elementos genéticos vectoriales integrados de reprogramación, o esencialmente exenta de elementos genéticos vectoriales.
- En un aspecto adicional, el vector de reprogramación puede introducirse en las células mediante transfección de liposomas, procedimientos de electroporación, bombardeo de partículas, fosfato de calcio, policación o polianión o cualquiera adecuado para introducir elementos genéticos exógenos en las células.
- En otros aspectos adicionales de la invención, las células somáticas pueden ser de mamíferos, o más específicamente de seres humanos. Las células somáticas pueden ser células diferenciadas terminalmente o células madre de tejido incluyendo, pero sin limitación, fibroblastos, células hematopoyéticas o células mesenquimáticas. Por ejemplo, las células somáticas son fibroblastos. Las células somáticas pueden ser de un banco celular de tejido o de un sujeto humano seleccionado, específicamente un ser humano vivo. Los genomas de la progenie de estas células somáticas se considerarán derivados de estas células somáticas de cierta fuente, tal como un individuo humano seleccionado.
- En algunos aspectos adicionales, las células de progenie podrían seleccionarse por una o más características citoblásticas embrionarias, tales como una morfología indiferenciada, un marcador específico de células madre embrionarias o pluripotencia o potencial de diferenciación multilínea o cualquier característica conocida en la materia. Específicamente, las células de progenie pueden seleccionarse por una morfología indiferenciada debido a su conveniencia. El marcador específico de célula madre embrionaria podría ser uno o más marcadores específicos seleccionados del grupo consistente en SSEA-3, SSEA-4, Tra-1-60 o Tra-1-81, Tra-2-49/6E, GDF3, REX1, FGF4, ESG1, DPPA2, DPPA4 y hTERT. Esta etapa de selección puede emplearse en más de un punto temporal para asegurar que las células están en estado pluripotente y no vuelven a un estado diferenciado.
- Además, en ciertos aspectos de la invención, son conocidos en la materia marcadores de selección positiva y pueden usarse en el procedimiento y composiciones de la invención para mejorar la eficacia de transfección o concentración de células transfectadas durante un periodo de tiempo suficiente para el establecimiento de un estado pluripotente automantenido. Por ejemplo, en algunos aspectos, el vector de reprogramación puede comprender además un marcador de selección positiva tal como una secuencia nucleotídica que codifica un factor de resistencia a antibióticos (p.ej., marcador de resistencia a neomicina o higromicina) o una proteína fluorescente o luminiscente (p.ej., GFP, RFP, CFP, etc.). Después de introducir las células somáticas con el vector de reprogramación, el uso del marcador de selección positiva puede ayudar a concentrar las células que tienen vectores de reprogramación. Sin embargo, esta etapa es opcional y depende de la eficacia de transfección y del índice de pérdida vectorial. Si la eficacia de transfección es alta (tal como más de un 90 %) y la pérdida vectorial es suficientemente baja para que las células establezcan un estado pluripotente automantenido, esta selección positiva puede no ser necesaria.
- En otras realizaciones adicionales de la invención, los factores de reprogramación de iPS pueden comprender al

menos un miembro de la familia Sox y al menos un miembro de la familia Oct, específicamente Sox-2 y Oct-4. Se cree que Sox y Oct son básicos en la jerarquía reguladora transcripcional que especifica la identidad celular de ES. Factores adicionales pueden aumentar la eficacia de reprogramación, tales como un conjunto que comprende Sox-2, Oct-4, Nanog y, opcionalmente, Lin-28; o que comprende Sox-2, Oct-4, Klf y opcionalmente c-Myc.

- 5 En algunos aspectos adicionales de los procedimientos anteriores, la etapa d puede oscilar de al menos 8 días a al menos 30 días, o cualquier número de días intermedio de los números anteriores después de la etapa b, como el periodo de tiempo requerido para establecerse un estado pluripotente automantenido.

El especialista en la materia entenderá que los módulos de expresión pueden ligarse operativamente con un elemento regulador transcripcional, tal como un promotor o potenciador.

- 10 En un aspecto adicional, se da a conocer también un vector de reprogramación que comprende un origen de replicación y uno o más módulos de expresión que codifican un factor de acción en trans que se une al origen de replicación para replicar un molde extracromosómico, y factores de reprogramación de iPS. Los factores de reprogramación de iPS pueden comprender Sox y Oct, más específicamente Sox-2 y Oct-4, por ejemplo un conjunto que comprende Sox-2, Oct-4, Nanog y opcionalmente Lin-28; o que comprende Sox-2, Oct-4, Klf y opcionalmente c-Myc,

- 15 En ciertos aspectos del vector de reprogramación, el vector de reprogramación se replica extracromosómicamente y/o carece de la capacidad de integrarse en un genoma de célula hospedadora. En realizaciones ejemplares, el origen de replicación puede ser de un herpesvirus linfotrófico o un gammaherpesvirus, un adenovirus, SV40, un papilomavirus bovino o levadura, tal como un origen de replicación de un herpesvirus linfotrófico o un gammaherpesvirus correspondiente a oriP de EBV. En un aspecto adicional, el herpesvirus linfotrófico puede ser virus de Epstein Barr (EBV), herpesvirus de sarcoma de Kaposi (KSHV), herpesvirus de Saimiri (HS) o virus de la enfermedad de Marek (MDV). El virus de Epstein Barr (EBV) y el herpesvirus de sarcoma de Kaposi (KSHV) son también ejemplos de gammaherpesvirus.

- 20 En ciertas realizaciones del vector de reprogramación, el factor de acción en trans puede ser un polipéptido correspondiente a, o un derivado de, una proteína de tipo silvestre correspondiente a EBNA-1 de EBV. El derivado puede activar la transcripción al menos a un 5 % de la correspondiente proteína de tipo silvestre a partir de un molde extracromosómico después de que el derivado se una al origen de replicación y/o tener una capacidad reducida de activar la transcripción a partir de un molde integrado en comparación con EBNA-1 de tipo silvestre, y por tanto posibilidades reducidas de activar ectópicamente genes cromosómicos para causar una transformación oncogénica.

- 25 Los ejemplos de derivado pueden incluir un derivado que carece de secuencias presentes en la proteína EBNA-1 de tipo silvestre que activan la transcripción a partir de un molde integrado, un derivado que tiene una secuencia de localización nuclear, un derivado que tiene una delección de residuos correspondiente a los residuos de aproximadamente 65 a aproximadamente 89 (SEQ ID NO: 1) y/o que tiene una delección de residuos correspondiente a los residuos de aproximadamente 90 a aproximadamente 328 de EBNA-1 (SEQ ID NO:1), un derivado con al menos un 80 % de identidad de secuencia aminoacídica con los residuos 1 a aproximadamente 40 y los residuos aproximadamente 328 a 641 de EBNA-1 (SEQ ID NO: 1), o un derivado que comprende una primera secuencia nucleotídica que codifica los residuos 1 a aproximadamente 40 de la correspondiente EBNA-1 de tipo silvestre y una segunda secuencia nucleotídica que codifica los residuos de aproximadamente 328 a 641 de la correspondiente EBNA-1 de tipo silvestre.

- 30 En un aspecto adicional, se reivindica también una población celular de iPS producida según el procedimiento anterior. En un otro aspecto adicional, puede darse a conocer también una población celular de iPS que está esencialmente exenta de elementos retrovíricos exógenos o una población celular de iPS esencialmente exenta de elementos víricos exógenos o cualquier elemento de ácido nucleico exógeno tal como elementos genéticos vectoriales; más específicamente, la población celular puede comprender el genoma de un individuo humano seleccionado. En un aspecto adicional, la población celular de iPS puede comprender células cuyo genoma deriva de una célula humana diferenciada terminalmente tal como una célula cutánea primaria (p.ej. un fibroblasto) y esencialmente exenta de elementos retrovíricos exógenos o cualquier ácido nucleico exógeno o elemento genético vectorial. "Esencialmente exento" de elementos de ADN exógenos significa que menos de un 1, 0,5, 0,1, 0,05 % o cualquier porcentaje intermedio de células de la población celular de iPS comprende elementos de ADN exógenos.

- 35 En otro aspecto adicional, puede darse a conocer una célula diferenciada, tejido u órgano que se ha diferenciado de la población celular de iPS como se describe anteriormente. La célula diferenciada puede comprender una célula hematopoyética, un miocito, una neurona, un fibroblasto o una célula epidérmica; el tejido puede comprender tejido de nervio, hueso, intestino, epitelio, músculo, cartílago o cardiaco; el órgano puede comprender cerebro, médula espinal, corazón, hígado, riñón, estómago, intestino o páncreas. En ciertos aspectos, la célula diferenciada, tejido u órgano puede usarse en trasplante de tejido, cribado de fármacos o investigación de desarrollo para reemplazar a células madre embrionarias.

Pueden usarse los procedimientos exentos de virus para inducir cualquier cambio en el estado de diferenciación de una célula. En ciertos aspectos, se proporciona también un procedimiento que proporciona una población celular

que tiene un estado de diferenciación alterado respecto a la población celular de partida y que tiene células que están esencialmente exentas de elementos genéticos vectoriales de programación, comprendiendo el procedimiento las etapas de: a) obtener una población de partida de células que tienen un primer estado de diferenciación, b) obtener uno o más vectores de programación de la diferenciación, comprendiendo cada vector un origen de replicación y uno o más módulos de expresión que codifican uno o más factores de programación de la diferenciación que, en combinación, pueden alterar el estado de diferenciación de la población celular de partida hasta un segundo estado de diferenciación, en el que uno o más de dichos módulos de expresión comprenden una secuencia nucleotídica que codifica un factor de transfección que se une al origen de replicación para replicar un molde extracromosómico y/o en el que las células de la población de partida expresan dicho factor de acción en trans; c) introducir el vector o vectores de programación de la diferenciación en las células de la población celular de partida; d) cultivar las células para efectuar la expresión de uno o más factores de programación de la diferenciación de tal modo que surjan atributos consistentes con el segundo estado de diferenciación en al menos una porción de las células cultivadas y e) cultivar además las células que tienen los atributos durante un número suficiente de generaciones para proporcionar una población celular que comprenda células que tienen el segundo estado de diferenciación, pero estando dichas células esencialmente exentas de elementos genéticos vectoriales de programación.

En ciertos aspectos, para replicar un molde extracromosómico, uno o más de los módulos de expresión en al menos un vector de programación de la diferenciación comprenden una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación; como alternativa, las células de partida pueden comprender una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación.

Puede haber tres modos de alterar el estado de diferenciación en la presente invención: desdiferenciación (que puede definirse además como reprogramación), diferenciación o transdiferenciación.

En un cierto aspecto de la invención, es un ejemplo de desdiferenciación la inducción de pluripotencia a partir de células somáticas tales como un fibroblasto, un queratinocito, una célula hematopoyética (p.ej. un linfocito), una célula mesenquimática, una célula hepática, una célula de estómago o una célula  $\beta$ . Los atributos de las células del segundo estado de diferenciación pueden definirse además como una o más características de células madre embrionarias. Para inducir la pluripotencia en los procedimientos de programación de la diferenciación, los factores de programación pueden definirse además como factores de reprogramación que comprenden Sox y Oct, más específicamente Sox-2 y Oct-4, opcionalmente en combinación con uno o más factores adicionales tales como Nanog, Lin-28, Klf, c-Myc o Esrrb. La célula de partida puede ser también una célula menos diferenciada, tal como una célula madre hematopoyética, una célula madre neural o una célula madre mesenquimática, o las correspondientes células progenitoras, que pueden expresar ciertos factores de programación endógenamente y pueden reprogramarse más fácilmente hasta células pluripotentes con necesidad de menos factores. Por ejemplo, las células progenitoras neurales pueden reprogramarse hasta células pluripotentes en ausencia de expresión de Sox-2 exógeno. El procedimiento puede incluir también una etapa adicional de diferenciación de la población celular diana, que se programa hasta un estado más pluripotente basándose en las etapas descritas anteriormente.

En otro aspecto, se incluyen también procedimientos de diferenciación, tales como la inducción de un destino celular más especificado de una célula pluripotente o multipotente; por ejemplo, la diferenciación de una célula madre embrionaria o una célula madre pluripotente inducida hasta una célula más diferenciada, tal como una progenitora hematopoyética, progenitora endodérmica, progenitora pancreática, progenitora endotelial, progenitora de retina o incluso además hasta una célula diferenciada terminalmente, tal como un cardiomiocito, una célula sanguínea, una neurona, un hepatocito, una célula  $\beta$  de islote o una célula de retina; o la diferenciación de una célula multipotente como una célula madre hematopoyética, una célula madre neural o una célula madre mesenquimática, así como una progenitora hematopoyética, progenitora endodérmica, progenitora pancreática o progenitora endotelial. Es un ejemplo específico que una célula pluripotente puede diferenciarse hasta una progenitora endodérmica con los procedimientos que usan SOX, tal como SOX7 o SOX17, como factores de programación. Es otro ejemplo que una progenitora hematopoyética puede diferenciarse hasta un linfocito B con los procedimientos que usan EBF1 como factor de programación de la diferenciación.

En un aspecto adicional, los procedimientos de la presente invención pueden usarse también para la transdiferenciación de un tipo celular diferenciado hasta otro tipo celular diferenciado. La célula de partida y la célula alterada pueden estar ambas terminales o particularmente diferenciadas, por ejemplo, puede programarse un linfocito B hasta un macrófago con factores de programación tales como C/EBP (más específicamente C/EBP $\alpha$  y C/EBP $\beta$ ), o puede programarse una célula exocrina hasta un hepatocito con factores tales como C/EBP $\beta$  o hasta una célula  $\beta$  de islote con factores que comprenden Ngn3 (también conocido como Neurog3), Pdx1 y Mafa.

Además, en ciertos aspectos de la invención, los procedimientos pueden comprender también una etapa adicional de selección de células de las células cultivadas, estando dichas células esencialmente exentas de elementos genéticos vectoriales de programación de la diferenciación, por ejemplo, seleccionando células cultivadas y estando dichas células exentas de un marcador de selección comprendido en el vector de programación de la diferenciación, o ensayando directamente la presencia de elementos genéticos vectoriales mediante procedimientos conocidos en la materia. Por ejemplo, el vector de programación puede comprender un marcador de selección tal como una secuencia nucleotídica que codifica un factor de resistencia a antibióticos (p.ej., marcador de resistencia a neomicina

o higromicina), una proteína fluorescente o luminiscente (p.ej. GFP, RFP, CFP, etc.), o una enzima (p.ej. timidina cinasa). La selección de la pérdida de elementos genéticos vectoriales puede ser en el momento en que se ha establecido el segundo estado de diferenciación o después.

5 En un aspecto adicional, el vector de programación de la diferenciación puede introducirse en las células de partida mediante transfección de liposomas, procedimientos de electroporación, bombardeo de partículas, fosfato de calcio, poliatión o polianión o cualquiera adecuado para introducir elementos genéticos exógenos en las células. Las células de partida pueden ser células de mamífero, más específicamente células humanas. En un aspecto adicional, se proporciona también una célula del segundo estado de diferenciación y esencialmente exenta de elementos genéticos vectoriales producida según los procedimientos anteriores.

10 En otro aspecto adicional, se proporciona también un vector de programación de la diferenciación que comprende un origen de replicación y uno o más módulos de expresión que codifican un factor de acción en trans que se une al origen de replicación para replicar un molde extracromosómico, y uno o más factores de programación de la diferenciación. Los factores de programación de la diferenciación pueden seleccionarse del grupo consistente en Sox (p.ej. Sox-2, Sox-7, Sox-17), Oct (p.ej., Oct-4), Nanog, Lin-28, Klf, c-Myc, Esrrb, EBF1, C/EBP (p.ej. C/EBP $\alpha$ ,  
15 C/EBP $\beta$ ), Ngn3, Pdx y Maf. Pueden ser ejemplos específicos del esqueleto del vector de programación de la diferenciación un vector de expresión episómico, tal como pCEP4, pREP4 o pEBNA DEST de Invitrogen. En un cierto aspecto, el vector de programación de la diferenciación puede definirse además como un vector de reprogramación. El vector de reprogramación puede comprender un miembro de la familia Sox y un miembro de la familia Oct, tales como Sox-2 y Oct-4, y puede comprender además uno o más factores tales como Nanog, Lin-28,  
20 Klf4, c-Myc, o Esrrb.

En ciertos aspectos del vector de programación de la diferenciación, el origen de replicación puede ser un origen de replicación de un herpesvirus linfotrófico o un gammaherpesvirus, un adenovirus, SV40, un papilomavirus bovino o levadura, específicamente un origen de replicación de un herpesvirus linfotrófico o un gammaherpesvirus correspondiente a oriP de EBV. En un aspecto particular, el herpesvirus linfotrófico puede ser virus de Epstein Barr  
25 (EBV), herpesvirus de sarcoma de Kaposi (KSHV), herpesvirus de Saimiri (HS) o virus de la enfermedad de Marek (MDV). El virus de Epstein Barr (EBV) y el herpesvirus de sarcoma de Kaposi (KSHV) son también ejemplos de gammaherpesvirus.

En realizaciones adicionales del vector de programación de la diferenciación, el factor de acción en trans puede ser un polipéptido correspondiente a, o un derivado de, una proteína de tipo silvestre correspondiente a EBNA-1 de EBV. El derivado puede activar la transcripción al menos a un 5 % de la correspondiente proteína de tipo silvestre a partir de un molde extracromosómico después de que el derivado se una al origen de replicación y/o tener una capacidad reducida de activar la transcripción a partir de un molde integrado en comparación con EBNA-1 de tipo silvestre, y por tanto posibilidades reducidas de activar ectópicamente genes cromosómicos para causar la transformación oncogénica. Los ejemplos de derivado pueden incluir un derivado que carece de las secuencias presentes en la proteína EBNA-1 de tipo silvestre que activan la transcripción a partir de un molde integrado, un derivado que tiene una secuencia de localización nuclear, un derivado que tiene una delección de residuos correspondiente a los residuos de aproximadamente 65 a aproximadamente 89 de EBNA-1 y/o que tiene una delección de residuos correspondiente a los residuos de aproximadamente 90 a aproximadamente 328 de EBNA-1,  
30 un derivado con una identidad de secuencia aminoacídica de al menos un 80 % con los residuos 1 a aproximadamente 40 y los residuos aproximadamente 328 a 641 de EBNA-1, o un derivado que comprende una primera secuencia nucleotídica que codifica los residuos 1 a aproximadamente 40 de la correspondiente EBNA-1 de tipo silvestre y una segunda secuencia nucleotídica que codifica los residuos de aproximadamente 328 a 641 de la correspondiente EBNA-1 de tipo silvestre.  
35

Pueden emplearse las realizaciones discutidas en el contexto de los procedimientos y/o composiciones de la invención con respecto a cualquier otro procedimiento o composición descritos en la presente memoria. Por tanto, una realización referente a un procedimiento o composición puede aplicarse a otros procedimientos y composiciones de la invención también.  
40

Como se usa en la presente memoria, los términos “codificar” o “codificando” con referencia a un ácido nucleico se usan para hacer fácilmente comprensible la invención para el especialista en la materia, sin embargo estos términos pueden usarse intercambiamente con “comprender” o “comprendiendo” respectivamente.  
45

Como se usa en la presente memoria, la designación “un” o “una” puede significar uno o más. Como se usan en la presente memoria en las reivindicaciones, cuando se usan junto con la palabra “comprende”, las palabras “un” o “una” pueden significar uno o más de uno.

El uso del término “o” en las reivindicaciones se usa para indicar “y/o” a menos que se indique explícitamente que se hace referencia a las alternativas solo o que las alternativas sean mutuamente excluyentes, aunque la divulgación apoya una definición que hace referencia solo a las alternativas e “y/o”. Como se usa en la presente memoria, “otro” puede significar al menos un segundo o más.  
50

A lo largo de esta solicitud, se usa el término “aproximadamente” para indicar que un valor incluye la variación de

error inherente del dispositivo, del procedimiento que se esté empleando para determinar el valor o de la variación que exista entre los sujetos de estudio.

Resultarán evidentes otros objetos, rasgos y ventajas dados a conocer en la presente memoria a partir de la siguiente descripción detallada.

5 **BREVE DESCRIPCIÓN DE LOS DIBUJOS**

Los siguientes dibujos forman parte de la presente memoria descriptiva y se incluyen para demostrar además ciertos aspectos de la presente invención. La invención puede comprenderse mejor por referencia a uno o más de estos dibujos en combinación con la descripción detallada de realizaciones específicas presentadas en la presente memoria.

10 **FIG. 1A-C:** Genoma de EBV y origen de replicación latente del plásmido (oriP).

**FIG. 2:** Un modelo basado en dominios y representación de la estructura parcial de EBNA-1.

**FIG. 3:** Un ejemplo ilustrativo de un plásmido de esqueleto receptor usado en la presente invención.

**FIG. 4:** Ejemplos de módulos para integrar en un plásmido de esqueleto receptor. Los ejemplos de promotores que podrían usarse para expresión incluyen, pero sin limitación, PGK, CMV, SV40 y EF1a.

15 **FIG. 5:** Un ejemplo ilustrativo de un plásmido de reprogramación que codifica Sox-2, Oct-4, Nanog y Lin28 (opcionalmente).

**DESCRIPCIÓN DE REALIZACIONES ILUSTRATIVAS**

**I. La presente invención**

20 La presente invención supera varios problemas importantes con las tecnologías actuales de reprogramación o programación de la diferenciación de diversas etapas de desarrollo de células, tales como generar células madre pluripotentes inducidas que estén esencialmente exentas de vectores víricos o elementos exógenos. En contraposición con los procedimientos anteriores que usan vectores víricos integrativos, estos procedimientos usan vectores de reprogramación o programación de la diferenciación extracromosómicos, por ejemplo plásmidos basados en elementos de EBV, para transducir los factores de reprogramación o programación de la diferenciación hasta células somáticas o células madre, cultivar estas células y seleccionar en las células de progenie una o más características de células madre embrionarias o en las células atributos consistentes con el estado de diferenciación alterado deseado. Los vectores replicados extracromosómicamente, como vectores basados en elementos de EBV, no se integrarán en el genoma de la célula hospedadora y se perderán con el tiempo después de un periodo suficiente para inducir a las células hasta un estado celular pluripotente o deseado. Un rasgo inherente de estos procedimientos producirá células de progenie esencialmente exentas de elementos genéticos exógenos y una selección negativa puede facilitar el proceso. Estos procedimientos posibilitan el aislamiento de iPS o de cualquier tipo celular deseado esencialmente exento de elementos vectoriales mediante la alteración del estado de diferenciación. Por tanto, las nuevas composiciones y procedimientos posibilitarán la fabricación iPS exentas de vector u otros tipos celulares deseados para terapia sin el riesgo de mutagénesis causado por la inserción aleatoria o expresión persistente de elementos víricos en las células. Se describen a continuación realizaciones y ventajas adicionales de la invención.

**II. Definiciones**

40 “Reprogramación” es un proceso que confiere a una célula una capacidad aumentada mensurablemente de formar progenie de al menos un nuevo tipo celular, tanto en cultivo como *in vivo*, distinto del que tendría en las mismas condiciones sin reprogramación. Más específicamente, la reprogramación es un proceso que confiere a una célula somática el potencial pluripotente. Esto significa, después de una proliferación suficiente, una proporción mensurable de progenie que tiene características fenotípicas del nuevo tipo celular si esencialmente dicha progenie no podría haberse formado antes de reprogramar; de otro modo, la proporción que tiene características del nuevo tipo celular es mensurablemente mayor que antes de reprogramar. En ciertas condiciones, la proporción de progenie con características del nuevo tipo celular puede ser de al menos aproximadamente un 1, 5, 25 % o más en orden de preferencia creciente.

50 “Programación de la diferenciación” es un proceso que cambia una célula formando progenie de al menos un nuevo tipo celular con un nuevo estado de diferenciación, tanto en cultivo como *in vivo*, distinto del que tendría en las mismas condiciones sin reprogramación de la diferenciación. Este proceso incluye diferenciación, desdiferenciación y transdiferenciación. “Diferenciación” es el proceso mediante el cual una célula menos especializada se vuelve un tipo celular más especializado. “Desdiferenciación” es un proceso celular en que una célula parcial o terminalmente diferenciada revierte a una etapa de desarrollo previa, tal como pluripotencia o multipotencia. “Transdiferenciación” es un proceso de transformación de un tipo celular diferenciado en otro tipo celular diferenciado.

Un “origen de replicación” (“ori”) es una secuencia de ADN, p.ej. en un herpesvirus linfotrófico, que cuando está



- presente en un plásmido en una célula es capaz de mantener las secuencias ligadas en el plásmido y/o en un sitio en o cerca de donde se inicia la síntesis de ADN. El ori de EBV incluye secuencias de FR (20 copias imperfectas de una repetición de 30 pb) y preferiblemente secuencias de DS, sin embargo, otros sitios de EBV se unen a EBNA-1, p.ej. las secuencias Rep\* pueden sustituir a DS como origen de replicación (Kirshmaier y Sugden, 1998). Por tanto, un origen de replicación de EBV incluye secuencias FR, DS o Rep\* o cualquier secuencia funcionalmente equivalente mediante modificaciones de ácido nucleico o combinación sintética derivada de las mismas. Por ejemplo, la presente invención puede usar también un origen de replicación genomanipulado de EBV, tal como mediante la inserción o mutación de elementos individuales, como se describe específicamente en Lindner, *et al.*, 2008.
- Un herpesvirus “linfotrófico” es un herpesvirus que se replica en un linfoblasto (p.ej. un linfoblasto B humano) u otro tipo celular y se replica extracromosómicamente durante al menos una parte de su ciclo vital natural. Después de infectar un hospedador, estos virus infectan latentemente el hospedador manteniendo el genoma vírico como un plásmido. El herpesvirus simple (HSV) no es un herpesvirus “linfotrófico”. Los herpesvirus linfotróficos ejemplares incluyen, pero sin limitación, EBV, herpesvirus de sarcoma de Kaposi (KSHV), herpesvirus de Saimiri (HS) y virus de la enfermedad de Marek (MDV).
- Un “vector” o “constructo” (al que a veces se hace referencia como “vehículo” de suministro génico o transferencia génica) hace referencia a una macromolécula o complejo de moléculas que comprende un polinucleótido para suministrar a una célula hospedadora, tanto *in vitro* como *in vivo*.
- Un “plásmido”, un tipo común de vector, es una molécula de ADN extracromosómica separada del ADN cromosómico que es capaz de replicarse independientemente del ADN cromosómico. En ciertos casos, es circular y bicatenario.
- Un “molde”, como se usa en la presente memoria, es una molécula de ADN que se une específicamente a una proteína de tipo silvestre de un herpesvirus linfotrófico, correspondiendo dicha proteína de tipo silvestre a EBNA-1, como resultado de la presencia en ese molde de una secuencia de ADN que se une a la proteína de tipo silvestre con una afinidad que es al menos un 10 % de la de unión de una secuencia de ADN correspondiente al oriP de EBV por la proteína de tipo silvestre, y desde la que se inicia y/o potencia opcionalmente la transcripción del molde después de la unión de la proteína y/o la potenciación del mantenimiento de dicho molde en una célula. Un “molde integrado” es aquel que se mantiene establemente en el genoma de la célula, p.ej. integrado en un cromosoma de esa célula. Un “molde extracromosómico” es aquel que se mantiene establemente en una célula pero que no se integra en el cromosoma.
- Se entiende por “constructo de expresión” o “módulo de expresión” una molécula de ácido nucleico que es capaz de dirigir la transcripción. Un constructo de expresión incluye, al menos, un promotor o una estructura funcionalmente equivalente a un promotor. Pueden incluirse también elementos adicionales tales como un potenciador y/o una señal de terminación de la transcripción.
- El término “exógeno”, cuando se usa con relación a una proteína, gen, ácido nucleico o polinucleótido en una célula u organismo, hace referencia a una proteína, gen, ácido nucleico o polinucleótido que se ha introducido en la célula u organismo por medios artificiales o naturales, o con relación a una célula hace referencia a una célula que se ha aislado e introducido posteriormente en otras células o un organismo por medios artificiales o naturales. Un ácido nucleico exógeno puede ser de un organismo o célula diferente, o puede ser una o más copias adicionales de un ácido nucleico que aparece naturalmente dentro del organismo o célula. Una célula exógena puede ser de un organismo diferente, o puede ser del mismo organismo. A modo de ejemplo no limitante, un ácido nucleico exógeno está en una localización cromosómica diferente de las células naturales, o está flanqueado de otro modo por una secuencia de ácido nucleico diferente a la encontrada en la naturaleza.
- El término “corresponde a” se usa en la presente memoria para indicar que una secuencia polinucleotídica es homóloga (concretamente, es idéntica, y no evolutivamente relacionada estrictamente) de todo o una porción de una secuencia polinucleotídica de referencia, o que una secuencia polipeptídica es idéntica a una secuencia polipeptídica de referencia. En contraposición, el término “complementario de” se usa en la presente memoria para indicar que la secuencia complementaria es homóloga de toda o una porción de una secuencia polinucleotídica de referencia. Para ilustración, la secuencia nucleotídica “TATAC” corresponde a una secuencia de referencia “TATAC” y es complementaria de una secuencia de referencia “GTATA”.
- Un “gen,” “polinucleótido”, “región de codificación”, “secuencia”, “segmento”, “fragmento” o “transgén” que “codifica” una proteína particular es una molécula de ácido nucleico que se transcribe y opcionalmente también se traduce en un producto génico, p.ej. un polipéptido, *in vitro* o *in vivo* cuando se dispone bajo el control de las secuencias reguladoras apropiadas. La región de codificación puede estar presente en forma de ADNc, ADN genómico o ARN. Cuando está presente en forma de ADN, la molécula de ácido nucleico puede ser monocatenaria (concretamente la hebra codificante) o bicatenaria. Los límites de la región de codificación se determinan por un codón de inicio en el extremo 5’ (amino) y un codón de terminación de la traducción en el extremo 3’ (carboxilo). Un gen puede incluir, pero sin limitación, ADNc de ARNm procariótico o eucariótico, secuencias de ADN genómico de ADN procariótico o eucariótico y secuencias de ADN sintético. La secuencia de terminación de la transcripción estará habitualmente

localizada en 3' de la secuencia génica.

5 El término "elementos de control" hace referencia colectivamente a regiones promotoras, señales de poliadenilación, secuencias de terminación de la transcripción, dominios reguladores en dirección 5', orígenes de replicación, sitios de entrada a ribosoma internos ("IRES"), potenciadores, uniones de corte y empalme y similares, que colectivamente proporcionan la replicación, transcripción, procesamiento postranscripcional y traducción de una secuencia de codificación en una célula receptora. No todos estos elementos de control tienen que estar presentes, a condición de que la secuencia de codificación seleccionada sea capaz de replicarse, transcribirse y traducirse en una célula hospedadora adecuada.

10 El término "promotor" se usa en la presente memoria en su sentido ordinario para hacer referencia a una región nucleotídica que comprende una secuencia reguladora de ADN, en el que la secuencia reguladora deriva de un gen que es capaz de unirse a ARN polimerasa e iniciar la transcripción de una secuencia de codificación en dirección 3'.

Se entiende por "potenciador" una secuencia de ácido nucleico que, cuando se dispone próxima a un promotor, confiere una actividad de transcripción aumentada respecto a la actividad de transcripción resultante del promotor en ausencia del dominio potenciador.

15 Se entiende por "ligado operativamente", con referencia a moléculas de ácido nucleico, que dos o más moléculas de ácido nucleico (p.ej. una molécula de ácido nucleico para transcribir, un promotor y un elemento potenciador) se conectan de tal modo que permitan la transcripción de la molécula de ácido nucleico. "Ligado operativamente" con referencia a moléculas de péptido y/o polipéptido significa que dos o más moléculas de péptido y/o polipéptido están conectadas de tal modo que procuran una única cadena polipeptídica, concretamente un polipéptido de fusión, que tiene al menos una propiedad de cada péptido y/o componente polipeptídico de la fusión. El polipéptido de fusión es preferiblemente quimérico, concretamente compuesto por moléculas heterólogas.

20 "Homología" hace referencia al porcentaje de identidad entre dos polinucleótidos o dos polipéptidos. La correspondencia entre una secuencia y la otra puede determinarse mediante técnicas conocidas en la materia. Por ejemplo, la homología puede determinarse mediante una comparación directa de la información de secuencia entre dos moléculas polipeptídicas por alineación de la información de secuencia y uso de programas informáticos fácilmente disponibles. Como alternativa, la homología puede determinarse mediante la hibridación de polinucleótidos en condiciones que formen dúplex estables entre regiones homólogas, seguido de digestión con nucleasa o nucleasas específicas de una cadena y determinación del tamaño de los fragmentos digeridos. Dos secuencias de ADN, o polipeptídicas, son "sustancialmente homólogas" entre sí cuando al menos aproximadamente un 80 %, preferiblemente al menos aproximadamente un 90 % y lo más preferiblemente al menos aproximadamente un 95 % de los nucleótidos o aminoácidos coinciden respectivamente en una longitud definida de las moléculas, como se determina usando los procedimientos anteriores.

35 El término "célula" se usa en la presente memoria en su sentido más amplia en la materia y hace referencia a un cuerpo vivo que es una unidad estructural de tejido de un organismo multicelular, está rodeado por una estructura de membrana que la aísla del exterior, tiene la capacidad de autorreplicación y tiene información genética y un mecanismo para expresarla. Las células usadas en la presente memoria pueden ser células de origen natural o células modificadas artificialmente (p.ej., células de fusión, células modificadas genéticamente, etc.).

40 Como se usa en la presente memoria, el término "célula madre" hace referencia a una célula capaz de autorreplicación y pluripotencia. Típicamente, las células madre pueden regenerar un tejido lesionado. Las células madre de la presente memoria pueden ser, pero sin limitación, células madre embrionarias (ES) o células madre de tejido (también llamados células madre específicas de tejido o células madre somáticas). Cualquier célula producida artificialmente que pueda tener las capacidades anteriormente descritas (p.ej. células de fusión, células reprogramadas o similares usadas en la presente memoria) puede ser un citoblasto.

45 Los "células madre embrionarias (ES)" son células madre pluripotentes derivadas de embriones tempranos. Se estableció por primera vez un ES en 1981, que se ha aplicado también a la producción de ratones con desactivación génica desde 1989. En 1998, se estableció un ES humano, que se está poniendo actualmente a disposición para medicina regenerativa.

50 Al contrario que los ES, las células madre de tejido tienen un potencial de diferenciación limitado. Las células madre de tejido están presentes en localizaciones particulares en tejidos y tienen una estructura intracelular indiferenciada. Por lo tanto, la pluripotencia de las células madre de tejido es típicamente baja. Las células madre de tejido tienen una alta relación de núcleo/citoplasma y tienen pocos orgánulos intracelulares. La mayoría de células madre de tejido tienen una baja pluripotencia, un ciclo celular largo y capacidad proliferativa más allá de la vida del individuo. Las células madre de tejido se separan en categorías basándose en los sitios de los que derivan las células, tales como el sistema dérmico, el sistema digestivo, el sistema medular óseo, el sistema nervioso y similares. Las células madre de tejido en el sistema dérmico incluyen células madre epidérmicas, células madre de folículo piloso y similares. Las células madre de tejido del sistema digestivo incluyen células madre pancreáticas (comunes), células madre hepáticas y similares. Los células madre de tejido del sistema medular óseo incluyen células madre hematopoyéticas, células madre mesenquimáticas y similares. Los células madre de tejido del sistema nervioso

incluyen células madre neurales, células madre retinianas y similares.

“Células madre pluripotentes inducidas”, abreviado comúnmente como iPS o iPSC, hace referencia a un tipo de célula madre pluripotente preparada artificialmente a partir de una célula no pluripotente, típicamente una célula somática adulta, o célula terminalmente diferenciada tal como un fibroblasto, una célula hematopoyética, un miocito, una neurona, una célula epidérmica o similares, mediante la inserción de ciertos genes a los que se hace referencia como factores de reprogramación.

“Pluripotencia” hace referencia a una célula madre que tiene el potencial de diferenciarse en todas las células constituyentes de uno o más tejidos u órganos, o preferiblemente cualquiera de las tres capas germinales: endodermo (revestimiento interior del estómago, tracto gastrointestinal, pulmones), mesodermo (músculo, hueso, sangre, sistema urogenital) y ectodermo (tejidos epidérmicos y sistema nervioso). Las “células madre pluripotentes” usados en la presente memoria hacen referencia a células que pueden diferenciarse en células derivadas de cualquiera de las tres capas germinales, por ejemplo descendientes directas de células totipotentes o células pluripotentes inducidas.

Como se usa en la presente memoria, “células madre totipotentes” hace referencia a células que tienen la capacidad de diferenciarse en todas las células constituyentes de un organismo, tales como las células que se producen por la fusión de un óvulo y un espermatozoide. Las células producidas por las primeras pocas divisiones del óvulo fertilizado son también totipotentes. Estas células pueden diferenciarse en tipos celulares embrionarios y extraembrionarios. Las células madre pluripotentes pueden dar lugar a cualquier tipo celular fetal o adulto. Sin embargo, solos no pueden desarrollarse hasta un animal fetal o adulto porque carecen del potencial de contribuir al tejido extraembrionario, tal como la placenta.

En contraposición, muchas células progenitoras son multipotentes, concretamente son capaces de diferenciación en un número limitado de destinos celulares. Las células progenitoras multipotentes pueden dar lugar a varios otros tipos celulares, pero esos tipos están limitados en número. Es un ejemplo de una célula madre multipotente una célula hematopoyética, una célula madre sanguínea que puede desarrollarse hasta varios tipos de células sanguíneas, pero no puede desarrollarse hasta células cerebrales u otros tipos de células. Al final de la larga serie de divisiones celulares que forman el embrión, están las células que están terminalmente diferenciadas, o que se considera que están permanentemente dedicadas a una función específica.

“Autorrenovación” hace referencia a la capacidad de pasar por numerosos ciclos de división celular manteniendo el estado indiferenciado.

Como se usa en la presente memoria, el término “célula somática” hace referencia a cualquier célula distinta de las células germinales, tales como óvulo, esperma o similares, que no transfiere directamente su ADN a la siguiente generación. Típicamente, las células somáticas tienen una pluripotencia limitada o nula. Las células somáticas usadas en la presente memoria pueden ser de origen natural o modificadas genéticamente.

Las células están “sustancialmente exentas” de elementos genéticos exógenos, como se usa en la presente memoria, cuando tienen menos de un 10 % del elemento o elementos y están “esencialmente exentas” de elementos genéticos exógenos cuando tienen menos de un 1 % del elemento o elementos. Sin embargo, son aún más deseables las poblaciones celulares en las que menos de un 0,5 % o menos de un 0,1 % de la población celular total comprende elementos genéticos exógenos. Por tanto, poblaciones de iPS en las que menos de un 0,1 a un 10 % (incluyendo todos los porcentajes intermedios) de las células de la población comprenden elementos genéticos exógenos.

### III. Antecedentes generales para células madre pluripotentes inducidas

En ciertas realizaciones de la invención, se dan a conocer procedimientos de reprogramación de células somáticas mediante la introducción de factores de reprogramación en las células somáticas con un sistema basado en vector extracromosómico. La progenie de estas células podría ser idéntica a las células madre embrionarios en diversos aspectos como se describen a continuación, pero esencialmente exentas de elementos genéticos exógenos. La comprensión de las características de las células madre embrionarios podría ayudar a seleccionar las células madre pluripotentes. Podrían usarse para estos procedimientos novedosos factores de reprogramación conocidos de los estudios de reprogramación de células madre. Se contempla además que estas células madre pluripotentes inducidas podrían usarse potencialmente para reemplazar a células madre embrionarias para aplicaciones terapéuticas y de investigación debido a la traba ética de usar estos últimos,

#### A. Células madre

Los células madre son células encontradas en la mayoría, si no todos, los organismos multicelulares. Se caracterizan por la capacidad de autorrenovarse mediante división celular mitótica y diferenciación en un diverso intervalo de tipos celulares especializados. Los dos tipos genéricos de células madre de mamífero son: células madre embrionarias que se encuentran en los blastocitos y células madre adultas que se encuentran en tejidos adultos. En un embrión en desarrollo, las células madre pueden diferenciarse en todos los tejidos embrionarios especializados. En organismos adultos, las células madre y células progenitoras actúan como un sistema de

reparación para el cuerpo, reponiendo las células especializadas pero manteniendo también el recambio normal de órganos regenerativos, tales como sangre, piel o tejidos intestinales.

5 Como las células madre pueden crecer y transformarse en células especializadas con características consistentes con las células de diversos tejidos tales como músculos o nervios mediante cultivo celular, se ha propuesto su uso en terapias médicas. En particular, las líneas celulares embrionarias, células madre embrionarias autólogas generados mediante clonación terapéutica y células madre adultas altamente plásticas de la sangre del cordón umbilical o de la médula ósea se publicitan como candidatos prometedores. Más recientemente, la reprogramación de células adultas hasta células madre pluripotentes inducidas tiene un enorme potencial para reemplazar a los células madre embrionarias.

10 **B. Células madre embrionarias**

Las líneas celulares de células madre embrionarias (líneas celulares de ES) son cultivos de células derivadas del tejido epiblastico de la masa celular interna (ICM) de embriones de etapa de blastocito o mórula temprana. Un blastocito es un embrión de etapa temprana de aproximadamente 4 a 5 días en seres humanos y que consiste en 50-150 células. Los ES son pluripotentes y dan lugar durante el desarrollo a todos los derivados de las tres capas germinales primarias: ectodermo, endodermo y mesodermo. En otras palabras pueden desarrollarse hasta cualquiera de los más de 200 tipos celulares del cuerpo adulto cuando se les da la estimulación suficiente y necesaria para un tipo celular específico. No contribuyen a las membranas extraembrionarias o a la placenta.

20 Casi toda la investigación hasta la fecha ha tenido lugar usando células madre embrionarias de ratón (mES) o células madre embrionarias humanas (hES). Ambos tienen las características de células madre esenciales, aunque requieren entornos muy diferentes para mantener un estado indiferenciado. Los ES de ratón pueden crecer en una capa de gelatina y requieren la presencia del factor inhibidor de leucemia (LIF). Los ES humanos podrían crecer en una capa alimentadora de fibroblastos embrionarios de ratón (MEF) y requieren a menudo la presencia de factor de crecimiento fibroblástico básico (bFGF o FGF-2). Sin condiciones de cultivo óptimas o manipulación genética (Chambers *et al.*, 2003), las células madre embrionarias se diferenciarán rápidamente.

25 Una célula madre embrionaria humana puede definirse también por la presencia de varios factores de transcripción y proteínas de superficie celular. Los factores de transcripción Oct-4, Nanog y Sox-2 forman la red reguladora principal que asegura la supresión de los genes que conducen a la diferenciación y el mantenimiento de la pluripotencia (Boyer *et al.*, 2005). Los antígenos de superficie celular más comúnmente usados para identificar hES incluyen los glicolípidos SSEA3 y SSEA4 y los antígenos de sulfato de queratano Tra-1-60 y Tra-1-81.

30 Después de 20 años de investigación, no hay tratamientos aprobados ni ensayos humanos que usen células madre embrionarias. Los ES, al ser células pluripotentes, requieren señales específicas para una correcta diferenciación, si se inyectan directamente al cuerpo, los ES se diferenciarán en muchos tipos diferentes de células, causando un teratoma. Diferenciar los ES en células utilizables evitando el rechazo de trasplante es solo una de las trabas a las que los investigadores de células madre embrionarias se enfrentan. Muchos países tienen actualmente moratorias sobre la investigación de ES o la producción de nuevas líneas de ES. Debido a sus capacidades combinadas de expansión ilimitada y pluripotencia, las células madre embrionarias siguen siendo una fuente teóricamente potencial de medicina regenerativa y reemplazo de tejido después de lesión o enfermedad. Sin embargo, un modo de evitar estos problemas es inducir el estado pluripotente en células somáticas mediante reprogramación directa.

**C. Factores de reprogramación**

40 La generación de iPS es crucial en los genes usados para inducción. Los siguientes factores o combinaciones de los mismos podrían usarse en el sistema vectorial dado a conocer en la presente invención. En ciertos aspectos, los ácidos nucleicos que codifican Sox y Oct (preferiblemente Oct3/4) se incluirán en el vector de reprogramación. Por ejemplo, un vector de reprogramación puede comprender módulos de expresión que codifican Sox-2, Oct-4, Nanog y opcionalmente Lin-28, o módulos de expresión que codifican Sox-2, Oct-4, Klf4 y opcionalmente c-myc, o módulos de expresión que codifican Sox-2, Oct-4 y opcionalmente Esrrb. Los ácidos nucleicos que codifican estos factores de reprogramación pueden estar comprendidos en el mismo módulo de expresión, diferentes módulos de expresión, el mismo vector de reprogramación o diferentes vectores de reprogramación.

50 Se ha identificado a Oct-3/4 y ciertos miembros de la familia del gen Sox (Sox-1, Sox-2, Sox-3 y Sox-15) como reguladores transcripcionales cruciales implicados en el proceso de inducción cuya ausencia hace imposible la inducción. Sin embargo, se han identificado genes adicionales, que incluyen ciertos miembros de la familia de Klf (Klf-2, Klf-2, Klf-4 y Klf-5), la familia de Myc (C-myc, L-myc y N-myc), Nanog y LIN28, que aumentan la eficacia de inducción.

55 Oct-3/4 (Pou5f1) es uno de la familia de factores de transcripción octaméricos ("Oct") y desempeña un papel crucial en el mantenimiento de la pluripotencia. La ausencia de Oct-3/4 en las células Oct-3/4+, tales como blastómeros y células madre embrionarias, conduce a una diferenciación trofoblástica espontánea, y la presencia de Oct-3/4 da lugar por tanto a la pluripotencia y el potencial de diferenciación de las células madre embrionarias. Diversos otros genes de la familia de "Oct", incluyendo parientes cercanos a Oct-3/4, Oct1 y Oct6, no consiguen desencadenar la inducción, demostrando por tanto la exclusividad de Oct-3/4 para el proceso de inducción.

La familia de genes Sox está asociada al mantenimiento de la pluripotencia de forma similar a Oct-3/4, aunque está asociada a células madre multipotentes y unipotentes en contraposición con Oct-3/4, que se expresa exclusivamente en células madre pluripotentes. Aunque Sox-2 era el gen inicial usado para inducción por Yamanaka *et al.*, Jaenisch *et al.* y Thomson *et al.*, se ha encontrado que otros genes de la familia de Sox funcionan también en el proceso de inducción. Sox1 procura iPS con una eficacia similar a Sox-2, y los genes Sox3, Sox15 y Sox18 generan también iPS, aunque con eficacia disminuida.

En células madre embrionarias es necesario Nanog, junto con Oct-3/4 y Sox-2, para promover la pluripotencia. Por lo tanto, fue sorprendente cuando Yamanaka *et al.* reseñó que Nanog no era necesario para inducción, aunque Thomson *et al.* ha reseñado que es posible generar iPS con Nanog como uno de los factores

LIN28 es una proteína de unión a ARNm expresada en células madre embrionarias y células de carcinoma embrionario asociadas a la diferenciación y proliferación. Thomson *et al.* demostró que es un factor en la generación de iPS, aunque no es necesario.

Se identificó inicialmente Klf4 de la familia de genes Klf por Yamanaka *et al.* y se confirmó por Jaenisch *et al.* como un factor para la generación de iPS de ratón, y se demostró por Yamanaka *et al.* como un factor para la generación de iPS humanos. Sin embargo, Thomson *et al.* reseñó que Klf4 era innecesario para la generación de iPS humanos y de hecho no consiguió generar iPS humanos. Se ha encontrado que Klf2 y Klf4 eran factores capaces de generar iPS, y los genes relacionados Klf1 y Klf5 lo hacían también, aunque con eficacia reducida.

La familia de genes Myc son protooncogenes implicados en el cáncer. Yamanaka *et al.* y Jaenisch *et al.* demostraron que c-myc es un factor implicado en la generación de iPS de ratón y Yamanaka *et al.* demostró que era un factor implicado en la generación de iPS humanos. Sin embargo, Thomson *et al.* y Yamanaka *et al.* reseñaron que c-myc era innecesario para la generación de iPS humanos. El uso de la familia de genes "myc" en la inducción de iPS humanos es problemático por la eventualidad de iPS en terapias clínicas, ya que un 25 % de los ratones trasplantados con iPS inducidos por c-myc desarrollaban teratomas mortales. Se ha identificado que N-myc y L-myc inducen en lugar de c-myc con eficacia similar.

**D. Inducción de células madre pluripotentes usando vectores integrativos**

Los IPS derivan típicamente por transfección de ciertos genes asociados a células madre en células no pluripotentes, tales como fibroblastos adultos. La transfección se consigue típicamente mediante integración de vectores víricos en la práctica actual, tales como retrovirus. Los genes transfectados incluyen los reguladores transcripcionales principales Oct-3/4 (Pou5f1) y Sox-2, aunque se sugiere que otros genes potencian la eficacia de inducción. Después de un periodo crítico, pequeños números de células transfectadas empiezan a volverse morfológica y bioquímicamente similares a las células madre pluripotentes, y se aíslan típicamente mediante selección morfológica, tiempo de duplicación o mediante un gen informador e infección antibiótica.

En noviembre de 2007, se consiguió un hito creando iPS a partir de células humanas adultas en dos estudios de equipos de investigación independientes (Yu *et al.*, 2007; Yamanaka *et al.*, 2007). Con el mismo principio usado anteriormente en modelos de ratón, Yamanaka había transformado exitosamente fibroblastos humanos en células madre pluripotentes usando los mismos cuatro genes fundamentales: Oct3/4, Sox-2, Klf4 y c-Myc con un sistema retrovírico, pero c-Myc es oncogénico. Thomson y colaboradores usaron Oct-4, Sox-2, NANOG y un gen diferente LIN28 usando un sistema lentivírico que evita el uso de c-Myc.

Sin embargo, los sistemas de transfección vírica usados insertan los genes en localizaciones aleatorias en el genoma del hospedador; esto es un problema para las aplicaciones terapéuticas potenciales de estos iPSC, porque las células creadas podrían ser susceptibles de cáncer. Miembros de ambos equipos consideran por lo tanto necesario desarrollar nuevos procedimientos de suministro.

Por otro lado, la expresión persistente forzada de factores de reprogramación ectópicos puede estar ligada con una frecuencia elevada de formación tumoral, y la solución final a este problema será la generación de iPS exentos de transgenes. Puede ser necesario un juego de genes introducidos víricamente para iniciar el proceso de reprogramación, pero gradualmente los genes de pluripotencia endógenos de la propia célula se vuelven activos y los genes víricos se silenciarán con potencial de reactivación estocástica. Investigadores recientes han demostrado que pueden requerirse factores exógenos durante un mínimo de aproximadamente 10-16 días para que las células entren en un estado pluripotente automantenido (Brambrink *et al.*, 2008; Stadtfeld *et al.*, 2008). La determinación de la longitud mínima de la expresión transgénica permite el desarrollo de procedimientos de suministro no retrovíricos para derivar iPS, una ventaja conseguida por los procedimientos e iPS presentes dados a conocer como se describe a continuación.

**IV. Vectores extracromosómicos para generar células madre pluripotentes inducidas exentas de vector y otros tipos celulares**

Como se describe anteriormente, se ha conseguido la inducción de células madre pluripotentes a partir de células somáticas humanas usando retrovirus o vectores lentivíricos para expresión ectópica de genes de reprogramación. Los retrovirus recombinantes tales como el virus de la leucemia de murino Moloney tienen la capacidad de integrarse

en el genoma hospedador de modo estable. Contienen una transcriptasa inversa que permite la integración en el genoma del hospedador. Los lentivirus son una subclase de retrovirus. Están ampliamente adaptados como vectores gracias a su capacidad de integrarse en el genoma de células no en división así como en división. Estos vectores víricos se han usado ampliamente también en un contexto más amplio: la programación de la diferenciación de células, incluyendo desdiferenciación, diferenciación y transdiferenciación. El genoma vírico en forma de ARN se transcribe de forma inversa cuando el virus entra en la célula produciendo ADN, que se inserta entonces en el genoma en una posición aleatoria por la enzima integrasa vírica. Por lo tanto, la tecnología actual de reprogramación exitosa depende de enfoques víricos basados en la integración.

Sin embargo, con la presente tecnología, la integración orientada no es aún rutinaria (Bode *et al.*, 2000b) y la alternativa convencional, integración aleatoria, puede conducir a mutagénesis de inserción con consecuencias impredecibles sobre las células madre pluripotentes inducidas. Por las mismas razones, no puede controlarse la expresión del transgén, puesto que depende del contexto de cromatina del sitio de integración (Baer *et al.*, 2000). La expresión de alto nivel puede conseguirse solo en loci genómicos favorables, pero existe el peligro de que la integración en sitios altamente expresados interfiera con las funciones celulares vitales de las células madre pluripotentes inducidas.

Además, existen evidencias crecientes de la existencia de mecanismos de defensa celular contra el ADN extraño que funcionan regulando negativamente transgenes en un proceso que está acompañado por la metilación de ADN (Bingham, 1997, Garrick *et al.*, 1998). Además, los componentes víricos pueden actuar junto con otros factores para transformar células. Acompañada de una expresión continua de una serie de genes víricos, la persistencia de al menos parte del genoma vírico en la célula puede causar la transformación celular. Estos genes pueden interferir con una ruta de señalización celular, causando los cambios fenotípicos observados de la célula que conducen a una célula transformada que muestra una división celular aumentada, que es favorable para el virus.

Por lo tanto, en ciertas realizaciones, la presente invención desarrolla procedimientos para generar células madre pluripotentes inducidas y otros tipos celulares deseados esencialmente exentos de elementos genéticos exógenos, tales como el vector retrovírico o lentivírico usado en los procedimientos anteriores. Estos procedimientos hacen uso de vectores de replicación extracromosómica, o vectores capaces de replicarse episómicamente. Una serie de plásmidos que contienen virus de ADN, tales como adenovirus, virus vacuolante de simio 40 (SV40) o papilomavirus bovino (BPV) o ARS de levadura incipiente (secuencias de replicación autónoma), se replican extracromosómicamente en células de mamífero. Estos plásmidos episómicos están intrínsecamente exentos de todas aquellas desventajas (Bode *et al.*, 2001) asociadas a vectores integrativos, pero nunca se han dado a conocer públicamente para generar células madre pluripotentes inducidas. Uno basado en herpesvirus linfotrófico, incluyendo virus de Epstein Barr (EBV) como se define anteriormente, puede replicarse también extracromosómicamente y ayudar a suministrar genes de reprogramación a células somáticas. Aunque los orígenes de replicación de estos virus o elementos de ARS están bien caracterizados, nunca se han conocido públicamente para reprogramar células diferenciadas hasta esta divulgación.

Por ejemplo, el enfoque basado en plásmido usado en la invención extrae elementos robustos necesarios para la replicación exitosa y mantenimiento de un sistema basado en elementos de EBV sin comprometer la manejabilidad del sistema en un entorno clínico como se describe con detalle a continuación. Los elementos de EBV esenciales son oriP y EBNA-1 o sus variantes o equivalentes funcionales. Es una ventaja adición de este sistema que estos elementos exógenos se perderán con el tiempo después de introducirse en células, conduciendo a iPS automantenidos esencialmente exentos de estos elementos.

#### A. Virus de Epstein-Barr

El virus de Epstein-Barr (EBV), también llamado herpesvirus humano 4 (HHV-4), es un virus de la familia del herpes (que incluye herpesvirus simple y citomegalovirus) y es uno de los virus más comunes en seres humanos. El EBV mantiene su genoma extracromosómicamente y funciona en colaboración con la maquinaria de la célula hospedadora para una replicación y mantenimiento eficaces (Lindner y Sugden, 2007), basándose solamente en dos rasgos esenciales para su replicación y retención en las células durante la división celular (Yates *et al.* 1985; Yates *et al.* 1984). Existe un elemento, al que se hace referencia comúnmente como oriP, en cis y sirve como origen de replicación. El otro factor, EBNA-1, funciona en trans mediante la unión a secuencias en oriP, promoviendo la replicación y mantenimiento de ADN de plásmido. Como ejemplo no limitante, los inventores extraen estos dos rasgos y los usan en el contexto de un plásmido para lanzar los genes necesarios para reprogramar células somáticas, para facilitar la replicación y expresión mantenida de estos genes frente a plásmidos convencionales.

#### B. OriP

OriP es el sitio en que o cerca del que se inicia la replicación de ADN y está compuesto por dos secuencias de acción en cis separadas por aproximadamente 1 kilopar de bases conocidas como la familia de repeticiones (FR) y la simetría diádica (DS).

FR está compuesta por 21 copias imperfectas de una repetición de 30 pb y contiene 20 sitios de unión de alta afinidad a EBNA-1 (FIG. 1). Cuando FR se une a EBNA-1, sirve tanto como potenciador transcripcional de

promotores en cis de hasta 10 kb de distancia (Reisman y Sugden, 1986; Yates, 1988; Sugden y Warren, 1989; Wysokenski y Yates, 1989; Gahn y Sugden, 1995; Kennedy y Sugden, 2003; Altmann *et al.*, 2006) como para contribuir a la retención nuclear y el mantenimiento fiel de plásmidos que contienen FR (Langle-Rouault *et al.*; 1998; Kirchmaier y Sugden, 1995; Wang *et al.*, 2006; Nanbo y Sugden, 2007). El reparto eficaz de los plásmidos de oriP es también atribuible probablemente a FR. Aunque el virus ha evolucionado para mantener 20 sitios de unión a EBNA-1 en FR, el mantenimiento de plásmido eficaz requiere solo 7 de estos sitios, y puede reconstituirse por un polímero de tres copias de DS, que tiene un total de 12 sitios de unión a EBNA-1 (Wysokenski y Yates, 1989).

El elemento de simetría diádica (DS) es suficiente para la iniciación de la síntesis de ADN en presencia de EBNA-1 (Aiyar *et al.* 1998; Yates *et al.*, 2000), y aparece la iniciación en o cerca de DS (Gahn y Schildkraut, 1989; Niller *et al.*, 1995). La terminación de la síntesis de ADN vírico se cree que aparece en FR, porque cuando FR se une a EBNA-1, funciona como barrera de la horquilla de replicación como se observa por electroforesis en gel 2D (Gahn y Schildkraut, 1989; Ermakova *et al.*, 1996; Wang *et al.*, 2006). La iniciación de la síntesis de ADN a partir de DS se permite en un ciclo de una vez por célula (Adams, 1987; Yates y Guan, 1991) y se regula por los componentes del sistema de replicación celular (Chaudhuri *et al.*, 2001; Ritzi *et al.*, 2003; Dhar *et al.*, 2001; Schepers *et al.*, 2001; Zhou *et al.*, 2005; Julien *et al.*, 2004). DS contiene cuatro sitios de unión a EBNA-1, aunque con menor afinidad que aquellos encontrados en FR (Reisman *et al.*, 1985). La topología de DS es tal que los cuatro sitios de unión están dispuestos como dos pares de sitios, con un espaciador de centro a centro de 21 pb entre cada par y un espaciador de centro a centro de 33 pb entre los dos sitios de unión interna no apareados (FIG. 1c) (Baer *et al.*, 1984; Rawlins *et al.*, 1985).

Los papeles funcionales de los elementos en DS se han confirmado por estudios de otra región del genoma de EBV, denominada Rep\*, que se ha identificado como un elemento que puede sustituir a DS ineficazmente (Kirchmaier y Sugden, 1998). Polimerizar Rep\* 8 veces procuraba un elemento tan eficaz como DS en su apoyo de la replicación (Wang, *et al.*, 2006). La disección bioquímica de Rep\* identificaba un par de sitios de unión a EBNA-1 con un espaciador de centro a centro de 21 pb crítico para su función replicativa (*ibid.*). Se encontró que el replicador mínimo de Rep\* era el par de sitios de unión a EBNA-1, ya que la función replicativa se retenía incluso después de reemplazar todas las secuencias flanqueantes del polímero por secuencias derivadas del fago lambda. Las comparaciones de DS y Rep\* han revelado un mecanismo común: estos replicadores apoyan la iniciación de la síntesis de ADN agrupando la maquinaria replicativa celular mediante un par de sitios apropiadamente espaciados, plegados y unidos a EBNA-1.

Hay otros plásmidos extracromosómicos permitidos que se replican en células de mamífero que no están relacionados con EBV, y en algunos modos parecen similares a la zona de iniciación en la cepa Raji de EBV. Hans Lipps y colaboradores han desarrollado y estudiado plásmidos que contienen "regiones de enlazamiento con el soporte/matriz nuclear" (S/MAR) y una unidad transcripcional robusta (Piechaczek *et al.*, 1999; Jenke *et al.*, 2004). Su S/MAR deriva del gen de interferón beta humano, es rica en A/T y se define operativamente por su asociación con la matriz nuclear y su relajación preferencial a baja fuerza iónica o cuando se embebe en ADN superenrollado (Bode *et al.*, 1992). Estos plásmidos se replican semiconservativamente, se unen a proteínas ORC y apoyan la iniciación de la síntesis de ADN eficazmente de forma aleatoria a lo largo de su ADN (Schaarschmidt *et al.*, 2004). Se mantienen eficazmente en células de hámster y humanas en proliferación sin selección de fármaco y, cuando se introducen en embriones porcinos, pueden apoyar la expresión de GFP en la mayoría de tejidos de animales fetales (Manzini *et al.*, 2006).

### C. EBNA-1

El antígeno nuclear 1 de Epstein Barr (EBNA-1) es una proteína de unión a ADN que se une a FR y DS de oriP o Rep\*, facilitando la replicación y reparto fiel del plásmido de EBV a las células hija independientemente de, pero concertadamente con, los cromosomas celulares durante cada división celular.

Los 641 aminoácidos (AA) de EBNA-1 se han clasificado en dominios asociados a sus funciones variadas mediante análisis mutacional y delecional (FIG. 2). Dos regiones entre AA40-89 y AA329-378 son capaces de ligar dos elementos de ADN en cis o en trans cuando se unen a EBNA-1, y se han denominado por tanto región de ligamiento 1 y 2 (LR1, LR2) (Middleton y Sugden, 1992; Frappier y O'Donnell, 1991; Su *et al.*, 1991; Mackey *et al.*, 1995). Fusionar estos dominios de EBNA-1 con GFP dirige a GFP a los cromosomas mitóticos (Marechal *et al.*, 1999; Kanda *et al.*, 2001). LR1 y LR2 son funcionalmente redundantes para replicación; una delección de cualquiera procura un derivado de EBNA-1 capaz de apoyar la replicación de ADN (Mackey y Sugden, 1999; Sears *et al.*, 2004). LR1 y LR2 son ricas en residuos de arginina y glicina, y se parecen al motivo de gancho AT que se une a ADN rico en A/T (Aravind y Landsman, 1998), (Sears *et al.*; 2004). Un análisis *in vitro* de LR1 y LR2 de EBNA-1 ha demostrado su capacidad de unirse a ADN rico en A/T (Sears *et al.*, 2004). Cuando LR1, que contiene uno de dichos ganchos AT, se fusionó con el dominio de unión a ADN y dimerización de EBNA-1, se encontró que era suficiente para la replicación de ADN de plásmidos de oriP, aunque menos eficazmente que la EBNA-1 de tipo silvestre (*ibid.*).

LR1 y LR2 difieren, sin embargo. La mitad C-terminal de LR1 está compuesta por aminoácidos distintos que los Arg-Gly repetidos de la mitad N-terminal, y se denomina región única 1 (UR1). La UR1 es necesaria para que EBNA-1 active la transcripción eficazmente a partir de ADN informadores transfectados e integrados que contienen FR (Wu *et al.*, 2002; Kennedy y Sugden, 2003; Altmann *et al.*, 2006). La UR1 es también esencial para la transformación

eficaz de células B infectadas por EBV. Cuando un derivado de EBNA-1 que carece de este dominio reemplaza a la proteína de tipo silvestre en el contexto del virus completo, estos virus derivados tienen un 0,1 % de la capacidad de transformación del virus de tipo silvestre (Altmann *et al.*, 2006).

5 La LR2 no se requiere para el apoyo por EBNA-1 de la replicación de *oriP* (Shire *et al.*, 1999; Mackey y Sugden, 1999; Sears *et al.*, 2004). Adicionalmente, la mitad N-terminal de EBNA-1 puede reemplazarse por proteínas celulares que contienen motivos de gancho AT, tales como HMGA1a, y seguir reteniendo función replicativa (Hung *et al.*, 2001; Sears *et al.*, 2003; Altmann *et al.*, 2006). Estos hallazgos indican que es probable que se requieran las actividades del gancho de AT de LR1 y LR2 para el mantenimiento de *oriP* en células humanas.

10 Un tercio de los residuos de EBNA-1 (AA91-328) consiste en repeticiones de glicina-glicina-alanina (GGA), implicadas en la capacidad de EBNA-1 de escapar del sistema inmunitario del hospedador al inhibir la degradación y presentación proteosómica (Levitskaya *et al.*, 1995; Levitskaya *et al.*, 1997). Se ha encontrado también que estas repeticiones inhiben la traducción de EBNA-1 *in vitro* e *in vivo* (Yin *et al.*, 2003). Sin embargo, la delección de gran parte de este dominio no tiene efecto aparente sobre las funciones de EBNA-1 en cultivo celular, haciendo el papel que desempeña este dominio difícil de elucidar.

15 Se codifica una señal de localización nuclear (NLS) por los AA379-386, que se asocia también con la maquinaria de importación nuclear celular (Kim, *et al.*, 1997; Fischer *et al.*, 1997). Las secuencias con las regiones ricas en Arg-Gly de LR1 y LR2 pueden funcionar también como NLS debido a su alto contenido básico.

20 Por último, el extremo C (AA458-607) codifica los dominios de unión a ADN y dimerización superpuestos de EBNA-1. La estructura de estos dominios unidos a ADN se ha resuelto por cristalografía de rayos X, y se ha encontrado que es similar al dominio de unión a ADN de la proteína E2 de papilomavirus (Hegde *et al.*, 1992; Kim *et al.*, 2000; Bochkarev *et al.*, 1996).

25 En realizaciones específicas de la invención, un vector de reprogramación contendrá tanto *oriP* como una secuencia abreviada que codifica una versión de EBNA-1 competente para apoyar la replicación de plásmido y su mantenimiento apropiado durante la división celular. La secuencia altamente repetitiva en el tercio aminoterminal de EBNA-1 de tipo silvestre y la retirada de una región de 25 aminoácidos que ha demostrado toxicidad en diversas células son prescindibles para la función de acción en trans de EBNA-1 asociada a *oriP* (Yates *et al.* 1985; Kennedy *et al.* 2003). Por lo tanto, podría usarse un derivado ejemplar, la forma abreviada de EBNA-1 conocida como deltaUR1 (el derivado con una secuencia proteica de SEQ ID NO:3, que está codificada por la SEQ ID NO: 4), junto con *oriP* en este sistema basado en plásmido. Más ejemplos de derivados de EBNA-1 pueden activar la transcripción a partir de un molde extracromosómico (véanse, por ejemplo, Kirchmaier y Sugden, 1997 y Kennedy y Sugden, 2003).

30 Un derivado de EBNA-1 usado en la invención es un polipéptido que, respecto al correspondiente polipéptido de tipo silvestre, tiene una secuencia aminoacídica modificada. Las modificaciones incluyen la delección, inserción o sustitución de al menos un residuo aminoacídico en una región correspondiente a la región única (residuos de aproximadamente 65 a aproximadamente 89) de LR1 (residuos de aproximadamente 40 a aproximadamente 89) de EBNA-1, y pueden incluir una delección, inserción o sustitución de uno o más residuos aminoacídicos en las regiones correspondientes a otros residuos de EBNA-1, p.ej. de aproximadamente el residuo 1 a aproximadamente el residuo 40, de aproximadamente el residuo 90 a aproximadamente 328 (región repetida "Gly-Gly-Ala"), de aproximadamente el residuo 329 a aproximadamente 377 (LR2), de aproximadamente el residuo 379 a aproximadamente 386 (NLS), de aproximadamente el residuo 451 a aproximadamente 608 (unión a ADN y dimerización) o de aproximadamente el residuo 609 a aproximadamente 641, a condición de que el derivado resultante tenga las propiedades deseadas, p.ej. dimerice y se una a ADN que contiene un ori correspondiente a *oriP*, se localice en el núcleo, no sea citotóxico y active la transcripción a partir de un molde extracromosómico pero no active sustancialmente la transcripción a partir de uno integrado. Las sustituciones incluyen sustituciones que utilizan la forma D en lugar de L, así como otros análogos aminoacídicos conocidos, p.ej. aminoácidos no naturales tales como aminoácidos  $\alpha$ -disustituidos, N-alquilaminoácidos, ácido láctico y similares. Estos análogos incluyen fosfoserina, fosfotreonina, fosfotirosina, hidroxiprolina, gamma-carboxiglutamato, ácido hipúrico, ácido octahidroindol-2-carboxílico, estatina, ácido 1,2,3,4-tetrahydroisoquinolin-3-carboxílico, penicilamina, ornitina, citrulina,  $\alpha$ -metilalanina, para-benzoilfenilalanina, fenilglicina, propargilglicina, sarcosina,  $\epsilon$ -N,N,N-trimetil-lisina,  $\epsilon$ -N-acetil-lisina, N-acetilserina, N-formilmetionina, 3-metilhistidina, 5-hidroxilisina, omega-N-metilarginina y otros aminoácidos e iminoácidos similares y *tert*-butilglicina.

35 Las sustituciones aminoacídicas conservativas son preferidas, es decir, por ejemplo, aspártico-glutámico como aminoácidos ácidos polares; lisina/arginina/histidina como aminoácidos básicos polares; leucina/isoleucina/metionina/valina/alanina/glicina/prolina como aminoácidos no polares o hidrófobos; serina/treonina como aminoácidos hidrófilos polares o no cargados. La sustitución aminoacídica conservativa incluye también agrupamientos basados en las cadenas laterales. Por ejemplo, es un grupo de aminoácidos que tienen cadenas laterales alifáticas glicina, alanina, valina, leucina e isoleucina; es un grupo de aminoácidos que tienen cadenas laterales hidroxilo alifático serina y treonina; es un grupo de aminoácidos que tienen cadenas laterales que contienen amida asparagina y glutamina; es un grupo de aminoácidos que tienen cadenas laterales aromáticas fenilalanina, tirosina y triptófano; es un grupo de aminoácidos que tienen cadenas laterales básicas lisina, arginina e histidina y es un grupo de aminoácidos que tienen cadenas laterales que contienen azufre cisteína y metionina. Por ejemplo, es



razonable esperar que el reemplazo de una leucina por isoleucina o valina, de un aspartato por un glutamato, de una treonina por una serina o un reemplazo similar de un aminoácido por un aminoácido estructuralmente relacionado, no tendrá un gran efecto sobre las propiedades del polipéptido resultante. Puede determinarse fácilmente si el cambio aminoacídico da como resultado un polipéptido funcional ensayando la actividad específica del polipéptido.

5 Las sustituciones aminoacídicas que entran dentro del alcance de la divulgación se logran, en general, seleccionando sustituciones que no difieran significativamente en su efecto sobre el mantenimiento de (a) la estructura del esqueleto peptídico en la zona de sustitución, (b) la carga o hidrofobicidad de la molécula en el sitio diana o (c) el volumen de la cadena lateral. Los residuos de origen natural se dividen en grupos basados en las propiedades de cadena lateral comunes:

- 10 1) hidrófobos: norleucina, met, ala, val, leu, ile;
- 2) hidrófilos neutros: cys, ser, thr;
- 3) ácidos: asp, glu;
- 4) básicos: asn, gln, his, lys, arg;
- 5) residuos que influyen en la orientación de la cadena: gly, pro y
- 15 6) aromáticos; trp, tyr, phe.

La invención concibe también polipéptidos con sustituciones no conservativas. Las sustituciones no conservativas conllevan el intercambio de un miembro de una de las clases descritas anteriormente por otra.

20 Pueden prepararse sales de adición de ácido del polipéptido o residuos aminoacídicos del polipéptido poniendo en contacto el polipéptido o amina con uno o más equivalentes del ácido inorgánico u orgánico deseado, tal como por ejemplo ácido clorhídrico. Pueden prepararse también ésteres de grupos carboxilo de los polipéptidos mediante cualquiera de los procedimientos habituales conocidos en la materia.

25 Los análogos incluyen estructuras que tienen uno o más ligamientos peptídicos opcionalmente reemplazados por un ligamiento seleccionado del grupo consistente en:  $-\text{CH}_2\text{NH}-$ ,  $-\text{CH}_2\text{S}-$ ,  $-\text{CH}_2\text{-CH}_2-$ ,  $-\text{CH}=\text{CH}-$  (cis y trans),  $-\text{CH}=\text{CF}-$  (trans),  $-\text{COCH}_2-$ ,  $-\text{CH}(\text{OH})\text{CH}_2-$  y  $-\text{CH}_2\text{SO}-$  mediante procedimientos conocidos en la materia y descritos además en las siguientes referencias: Spatola, 1983; Spatola, 1983; Morley, 1980; Hudson *et al.*, 1979 ( $-\text{CH}_2\text{NH}-$ ,  $\text{CH}_2\text{CH}_2-$ ); Spatola *et al.*, 1986 ( $-\text{CH}_2\text{S}-$ ); Hann, 1982 ( $-\text{CH}-\text{CH}-$ , *cis* y *trans*); Almquist *et al.*, 1980 ( $-\text{COCH}_2-$ ); Jennings-White *et al.*, 1982 ( $-\text{COCH}_2-$ ); Szelke *et al.*, solicitud europea EP 45665 ( $-\text{CH}(\text{OH})\text{CH}_2-$ ); Holladay *et al.*, 1983 ( $-\text{C}(\text{OH})\text{CH}_2-$ ) y Hruby; 1982 ( $-\text{CH}_2\text{S}-$ ). Es un ligamiento no peptídico particularmente preferido  $-\text{CH}_2\text{NH}-$ . Dichos análogos pueden tener mayor estabilidad química, propiedades farmacológicas potenciadas (semivida, absorción, potencia, eficacia, etc.), especificidad alterada (p.ej. un amplio espectro de actividades biológicas), antigenicidad reducida y prepararse económicamente.

#### D. Rasgo de exención de residuos

35 De forma importante, la replicación y mantenimiento de plásmidos basados en *oriP* son imperfectos y se pierden rápidamente (25 % por división celular) de las células al cabo de las primeras dos semanas de su introducción en las células; sin embargo, aquellas células que retienen el plásmido los pierden menos frecuentemente (3 % por división celular) (Leight y Sugden, 2001; Nanbo y Sugden, 2007). Una vez se retira la selección de células que albergan el plásmido, los plásmidos se perderán durante cada división celular hasta que se hayan eliminado todos ellos con el tiempo, sin dejar huella de su existencia anterior dentro de las células hija resultantes. Es este rasgo de ausencia de huellas lo que subraya el atractivo del sistema basado en *oriP* como alternativa al enfoque asociado a virus actual para suministrar genes para generar iPS y otras células deseadas con programación de la diferenciación. Se perderán también otros vectores extracromosómicos durante la replicación y propagación de células hospedadoras, y podrían emplearse también en la presente invención.

#### V. Construcción y suministro de vector

45 En ciertas realizaciones, podrían construirse vectores de reprogramación o de programación de la diferenciación para comprender elementos adicionales además de las secuencias de ácido nucleico que codifican factores de reprogramación o factores de programación de la diferenciación como se describen anteriormente para expresar estos factores de reprogramación en células. Los rasgos novedosos de estos procedimientos son el uso de vectores de replicación extracromosómica, que no se integrarán en el genoma de la célula hospedadora y pueden perderse durante las generaciones de replicación. Se dan a conocer a continuación los detalles de los componentes de estos vectores y procedimientos de suministro.

#### A. Vector

El uso de vectores extracromosómicos basados en plásmido o liposoma, p.ej. vectores basados en *oriP* y/o vectores que codifican un derivado de EBNA-1, permite introducir grandes fragmentos de ADN en una célula y mantenerlos

extracromosómicamente, replicarlos una vez por ciclo celular, repartirlos a las células hija eficazmente y no desencadenar sustancialmente respuesta inmunitaria. En particular, EBNA-1, la única proteína vírica requerida para la replicación del vector de expresión basado en oriP, no desencadena una respuesta inmunitaria celular debido a que ha desarrollado un mecanismo eficaz para evitar el procesamiento requerido para la presentación de sus antígenos en moléculas de MHC de clase I (Levitskaya *et al.*, 1997). Además, EBNA-1 puede actuar en *trans* para potenciar la expresión del gen clonado, induciendo la expresión de un gen clonado hasta 100 veces en algunas líneas celulares (Langle-Rouault *et al.*, 1998; Evans *et al.*; 1997). Finalmente, la fabricación de dichos vectores de expresión basados en oriP es barata.

Otros vectores extracromosómicos incluyen otros vectores basados en herpesvirus linfotróficos. El herpesvirus linfotrófico es un herpesvirus que se replica en un linfoblasto (p.ej. un linfoblasto B humano) y se vuelve un plásmido durante parte de su ciclo vital natural. El herpesvirus simple (HSV) no es un herpesvirus "linfotrófico". Los herpesvirus linfotróficos ejemplares incluyen, pero sin limitación, EBV, herpesvirus de sarcoma de Kaposi (KSHV), herpesvirus de Saimiri (HS) y virus de la enfermedad de Marek (MDV). Están contempladas todas las demás fuentes de vectores basados en episoma, tales como ARS de levadura, adenovirus, SV40 o BPV.

Un especialista en la materia estaría bien capacitado para construir un vector mediante técnicas recombinantes estándares (véanse, por ejemplo, Maniatis *et al.*, 1988 y Ausubel *et al.*; 1994).

Los vectores pueden comprender también otros componentes o funcionalidades que modulen además el suministro génico y/o la expresión génica, o que proporcionen de otro modo propiedades beneficiosas a las células diana. Dichos otros componentes incluyen, por ejemplo, componentes que influyen en la unión u orientación a células (incluyendo componentes que median la unión específica de tipo celular o tejido), componentes que influyen en la captación del ácido nucleico vectorial por la célula, componentes que influyen en la localización del polinucleótido en la célula después de la captación (tales como agentes que median la localización nuclear) y componentes que influyen en la expresión del polinucleótido.

Dichos componentes podrían incluir también marcadores, tales como marcadores detectables y/o de selección, que pueden usarse para detectar o seleccionar células que han captado y están expresando el ácido nucleico suministrado por el vector. Dichos componentes pueden proporcionarse como un rasgo natural del vector (tal como el uso de ciertos vectores víricos que tienen componentes o funcionalidades que median la unión y captación), o los vectores pueden modificarse para proporcionar dichas funcionalidades. Son conocidos en la materia una gran variedad de dichos vectores y están generalmente disponibles. Cuando un vector se mantiene en una célula hospedadora, el vector puede replicarse establemente por las células durante la mitosis como una estructura autónoma incorporada al genoma de la célula hospedadora o mantenerse en el núcleo o citoplasma de la célula hospedadora.

**B. Elementos reguladores:**

Los módulos de expresión eucariótica incluidos en los vectores contienen preferiblemente (en dirección 5' a 3') un promotor transcripcional eucariótico ligado operativamente a una secuencia de codificación de proteína, señales de corte y empalme que incluyen secuencias intermedias y una secuencia de terminación transcripcional/poliadenilación.

**i. Promotor/potenciadores**

Un "promotor" es una secuencia de control que es una región de una secuencia de ácido nucleico en que se controlan la iniciación e índice de transcripción. Puede contener elementos genéticos a los que pueden unirse proteínas y moléculas reguladoras, tales como ARN polimerasa y otros factores de transcripción, para iniciar la transcripción específica de una secuencia de ácido nucleico. Las frases "operativamente situado", "operativamente ligado", "bajo el control" y "bajo el control transcripcional" significan que un promotor está en una localización funcional y/u orientación correcta con relación a una secuencia de ácido nucleico para controlar la iniciación transcripcional y/o la expresión de esa secuencia.

Los promotores adecuados para uso en el vector que codifica EBNA-1 de la invención son aquellos que dirigen la expresión de los módulos de expresión que codifican la proteína EBNA-1, dando como resultado niveles en estado estacionario suficientes de proteína EBNA-1 para mantener establemente vectores que contienen oriP de EBV. Los promotores se usan también para una expresión eficaz de módulos de expresión que codifican factores de reprogramación.

Un promotor comprende generalmente una secuencia que funciona situando el sitio de inicio de la síntesis de ARN. El mejor ejemplo conocido de esto es la secuencia TATA, pero en algunos promotores que carecen de secuencia TATA, tales como por ejemplo el promotor del gen desoxinucleotidil transferasa terminal de mamífero y el promotor de los genes tardíos de SV40, un elemento discreto que se superpone con el sitio de inicio mismo ayuda a fijar el lugar de iniciación. Elementos promotores adicionales regulan la frecuencia de la iniciación transcripcional. Típicamente, estos se localizan en la región a 30-110 pb en dirección 5' del sitio de inicio, aunque se ha mostrado que una serie de promotores contienen elementos funcionales en dirección 3' del sitio de inicio también. Para llevar una secuencia de codificación "bajo el control de" un promotor, se sitúa el extremo 5' del sitio de iniciación de la

transcripción del marco de lectura transcripcional en dirección 3' del promotor elegido. El promotor "en dirección 5'" estimula la transcripción del ADN y promueve la expresión del ARN codificado.

El espaciador entre los elementos promotores frecuentemente es flexible, de modo que la función promotora se conserva cuando se invierten o mueven elementos entre sí. En el promotor tk, el espaciador entre los elementos promotores puede aumentarse a 50 pb antes de que la actividad empiece a reducirse. Dependiendo del promotor, parece que los elementos individuales pueden funcionar cooperativa o independientemente para activar la transcripción. Un promotor puede usarse o no junto con un "potenciador", que hace referencia a una secuencia reguladora de acción en cis implicada en la activación transcripcional de una secuencia de ácido nucleico.

Un promotor puede ser aquel asociado naturalmente a una secuencia de ácido nucleico, como puede obtenerse aislando las secuencias no codificantes 5' localizadas en dirección 5' del segmento de codificación y/o exón. Puede hacerse referencia a dicho promotor como "endógeno". De forma similar, un potenciador puede ser aquel asociado naturalmente a una secuencia de ácido nucleico, localizado en dirección 3' o 5' de esa secuencia. Como alternativa, se obtendrán ciertas ventajas al situar el segmento de ácido nucleico codificante bajo el control de un promotor recombinante o heterólogo, que hace referencia a un promotor que no está normalmente asociado a una secuencia de ácido nucleico en su entorno natural. Un potenciador recombinante o heterólogo hace referencia también a un potenciador no asociado normalmente a una secuencia de ácido nucleico en su entorno natural. Dichos promotores o potenciadores pueden incluir promotores o potenciadores de otros genes y promotores o potenciadores aislados de cualquier otro virus, o célula procarionótica o eucariótica, y promotores o potenciadores "de origen no natural", concretamente que contienen diferentes elementos de diferentes regiones reguladoras transcripcionales y/o mutaciones que alteran la expresión. Por ejemplo, los promotores que se usan más comúnmente en la construcción de ADN recombinante incluyen los sistemas promotores de  $\beta$ -lactamasa (penicilinas), lactosa y triptófano (trp). Además de producir secuencias de ácido nucleico de promotores y potenciadores sintéticamente, las secuencias pueden producirse usando clonación recombinante y/o tecnología de amplificación de ácido nucleico, incluyendo PCR™, en conexión con las composiciones dadas a conocer en la presente memoria (véanse las patentes de EE.UU. nº 4.683.202 y 5.928.906). Además, se contempla que pueden emplearse también las secuencias de control que dirigen la transcripción directa y/o expresión de secuencias en orgánulos no nucleares tales como mitocondrias, cloroplastos y similares.

Naturalmente, será importante emplear un promotor y/o potenciador que dirija eficazmente la expresión de los segmentos de ADN en el orgánulo, tipo celular, tejido, órgano u organismo elegido para expresión. Los especialistas en la materia de la biología molecular conocen generalmente el uso de promotores, potenciadores y combinaciones de tipo celular para la expresión de proteínas (véase por ejemplo Sambrook *et al.* 1989). Los promotores empleados pueden ser constitutivos, específicos de tejido, inducibles y/o útiles en las condiciones apropiadas para dirigir un alto nivel de expresión del segmento de ADN introducido, tal como es ventajoso en la producción a gran escala de proteínas y/o péptidos recombinantes. El promotor puede ser heterólogo o endógeno.

Adicionalmente, podría usarse también cualquier combinación de promotor/potenciador (como por ejemplo de la base de datos de promotores eucarióticos EPDB, mediante [www.epd.isb-sib.ch/](http://www.epd.isb-sib.ch/)) para impulsar la expresión. El uso de un sistema de expresión citoplasmático T3, T7 o SP6 es otra posible realización. Las células eucarióticas pueden apoyar la transcripción citoplasmática a partir de ciertos promotores bacterianos si se proporciona la polimerasa bacteriana apropiada, como parte del complejo de suministro o como constructo de expresión genética adicional.

Los ejemplos no limitantes de promotores incluyen promotores víricos tempranos o tardíos tales como promotores tempranos o tardíos de SV40, promotores tempranos inmediatos de citomegalovirus (CMV), promotores tempranos de virus de sarcoma de Rous (RSV), promotores celulares eucarióticos tales como, p.ej., promotor de beta-actina (Ng, S.Y., *Nuc. Acid. Res.* 17: 601-615, 1989, Quitsche *et al.*, *J. Biol. Chem.* 264: 9539-9545, 1989), promotor de GADPH (Alexander *et al.*, *Proc. Nat. Acad. Sci. USA* 85: 5092-5096, 1988, Ercolani *et al.*, *J. Biol. Chem.* 263: 15335-15341, 1988), promotor de metalotioneína (Karin *et al.*, *Cell* 36: 371-379, 1989; Richards *et al.*, *Cell* 37: 263-272, 1984) y promotores de elementos de respuesta concatenados tales como promotores de elemento de respuesta de AMP cíclico (cre), promotor de elemento de respuesta de suero (ser), promotor de éster de forbol (TPA) y promotores de elementos de respuesta cerca de una secuencia TATA mínima (tre). También es posible usar secuencias promotoras de hormona de crecimiento humana (p.ej. el promotor mínimo de hormona de crecimiento humana descrito en Genbank, nº de acceso X05244, nucleótidos 283-341) o un promotor de tumor mamario de ratón (disponible en la ATCC, nº de cat. ATCC 45007). Un ejemplo específico podría ser un promotor de fosfoglicerato cinasa (PGK).

## ii. Señales de iniciación y sitios de unión a ribosoma internos

Puede requerirse también una señal de iniciación específica para una traducción eficaz de secuencias de codificación. Estas señales incluyen el codón de iniciación ATG o secuencias adyacentes. Pueden tener que proporcionarse señales de control traduccional exógenas, incluyendo el codón de iniciación ATG. Un especialista en la materia sería capaz de determinar fácilmente esto y proporcionar las señales necesarias. Es bien conocido que el codón de iniciación debe estar "en fase" con el marco de lectura de la secuencia de codificación deseada para asegurar la traducción de todo el inserto. Las señales de control traduccional exógenas y codones de iniciación pueden ser naturales o sintéticos. La eficacia de la expresión puede potenciarse por la inclusión de elementos

potenciadores de la transcripción apropiados.

En ciertas realizaciones de la invención, se usan elementos de sitio de entrada a ribosoma interno (IRES) para crear mensajes multigénicos o policistrónicos. Los elementos de IRES son capaces de evitar el modelo de barrido ribosómico de traducción dependiente de Cap metilado en 5' y de empezar la traducción en sitios internos (Pelletier y Sonenberg, 1988). Se han descrito los elementos de IRES de dos miembros de la familia de picornavirus (polio y encefalomiocarditis) (Pelletier y Sonenberg, 1988) así como un IRES de un mensaje de mamífero (Macejak y Sarnow, 1991). Los elementos de IRES pueden ligarse con marcos de lectura abiertos heterólogos. Pueden transcribirse conjuntamente múltiples marcos de lectura abiertos, separado cada uno por un IRES, creando mensajes policistrónicos. En virtud del elemento de IRES, cada marco de lectura abierto es accesible a ribosomas para una traducción eficaz. Pueden expresarse eficazmente múltiples genes usando un solo promotor/potenciador para transcribir un solo mensaje (véanse las patentes de EE.UU. n° 5.925.565 y 5.935.819).

### iii. Sitios de clonación múltiple

Los vectores pueden incluir un sitio de clonación múltiple (MCS), que es una región de ácido nucleico que contiene múltiples sitios de enzima de restricción, cualquiera de los cuales puede usarse junto con tecnología recombinante estándar para digerir el vector (véanse, por ejemplo, Carbonelli *et al.*, 1999, Levenson *et al.*, 1998 y Cocea, 1997). "Digestión con enzima de restricción" hace referencia a la escisión catalítica de una molécula de ácido nucleico con una enzima que funciona solo en localizaciones específicas de una molécula de ácido nucleico. Muchas de estas enzimas de restricción están comercialmente disponibles. El uso de dichas enzimas es ampliamente comprendido por los especialistas en la materia. Frecuentemente, se linealiza o fragmenta un vector usando una enzima de restricción que corta en el MCS, posibilitando que se ligan con el vector secuencias exógenas. "Ligamiento" hace referencia al proceso de formación de enlaces fosfodiéster entre dos fragmentos de ácido nucleico, que pueden ser o no contiguos entre sí. Las técnicas que implican enzimas de restricción y reacciones de ligamiento son bien conocidas por los especialistas en la materia de la tecnología recombinante.

### iv. Sitios de corte y empalme

La mayoría de moléculas de ARN eucarióticas experimentarán corte y empalme de ARN para retirar los intrones de los transcritos primarios. Los vectores que contienen secuencias eucarióticas genómicas pueden requerir sitios de corte y empalme donantes y/o aceptores para asegurar un procesamiento apropiado del transcrito para expresión de proteína (véase, por ejemplo, Chandler *et al.*, 1997).

### v. Señales de terminación

Los vectores o constructos de la presente invención pueden comprender al menos una señal de terminación. Una "señal de terminación" o "terminador" comprende secuencias de ADN implicadas en la terminación específica de un transcrito de ARN por una ARN polimerasa. Por tanto, en ciertas realizaciones se contempla una señal de terminación que finaliza la producción de un transcrito de ARN. Puede ser necesario un terminador *in vivo* para conseguir niveles de mensaje deseables.

En sistemas eucarióticos, la región terminadora puede comprender también secuencias de ADN específicas que permitan la escisión específica de sitio del nuevo transcrito para exponer un sitio de poliadenilación. Esto señala a una polimerasa endógena especializada para añadir un tramo de aproximadamente 200 residuos de A (poliA) al extremo 3' del transcrito. Las moléculas de ARN modificadas con esta cola de poliA parecen más estables y se traducen más eficazmente. Por tanto, en otras realizaciones que implican eucariotas, se prefiere que el terminador comprenda una señal para la escisión del ARN, y es más preferido que la señal terminadora promueva la poliadenilación del mensaje. Los elementos de sitio terminador y/o de poliadenilación pueden servir para potenciar los niveles de mensaje y para minimizar la lectura del módulo en otras secuencias.

Los terminadores contemplados para uso en la invención incluyen cualquier terminador de la transcripción conocido descrito en la presente memoria o conocido por un especialista en la materia incluyendo, pero sin limitación, por ejemplo las secuencias de terminación de genes, tales como por ejemplo las secuencias terminadoras de hormona de crecimiento bovino, o de terminación vírica, tales como por ejemplo el terminador de SV40. En ciertas realizaciones, la señal de terminación puede ser la falta de secuencia transcribible o traducible, tal como debido a truncamiento de la secuencia.

### vi. Señales de poliadenilación

En la expresión, particularmente la expresión eucariótica, se incluirá típicamente una señal de poliadenilación para efectuar una poliadenilación apropiada del transcrito. La naturaleza de la señal de poliadenilación no se cree que sea crucial para la práctica exitosa de la invención, y puede emplearse cualquiera de dichas secuencias. Las realizaciones preferidas incluyen la señal de poliadenilación de SV40 o la señal de poliadenilación de hormona de crecimiento bovino, conveniente y conocida por funcionar bien en diversas células diana. La poliadenilación puede aumentar la estabilidad del transcrito o puede facilitar el transporte citoplasmático.

### vii. Orígenes de replicación

Para propagar un vector en una célula hospedadora, puede contener uno o más sitios de origen de replicación (a menudo denominados "ori"), por ejemplo una secuencia de ácido nucleico correspondiente a oriP de EBV como se describe anteriormente o un oriP genomanipulado con una función similar o elevada de programación de la diferenciación, que es una secuencia de ácido nucleico específica en que se inicia la replicación. Como alternativa, puede emplearse un origen de replicación de otro virus de replicación extracromosómica como se describe anteriormente o una secuencia de replicación autónoma (ARS).

#### viii. Marcadores de selección y rastreables

En ciertas realizaciones de la invención, las células que contienen un constructo de ácido nucleico de la presente invención pueden identificarse *in vitro* o *in vivo* mediante la inclusión de un marcador en el vector de expresión. Dichos marcadores conferirían un cambio identificable a la célula, permitiendo la fácil identificación de las células que contienen el vector de expresión. Generalmente, un marcador de selección es aquel que confiere una propiedad que permite la selección. Un marcador de selección positiva es aquel en que la presencia del marcador permite su selección, mientras que un marcador de selección negativa es aquel en que su presencia evita su selección. Es un ejemplo de un marcador de selección positiva un marcador de resistencia a fármaco.

Habitualmente, la inclusión de un marcador de selección de fármaco ayuda a la clonación e identificación de transformantes, por ejemplo, los genes que confieren resistencia a neomicina, puromicina, higromicina, DHFR, GPT, zeocina e histidinol son marcadores de selección útiles. Además de marcadores que confieren un fenotipo que permite la discriminación de transformantes basándose en la implantación de condiciones, se contemplan también otros tipos de marcadores que incluyen marcadores rastreables tales como GFP, cuya base es el análisis colorimétrico. Como alternativa, pueden utilizarse enzimas rastreables como marcadores de selección negativa tales como timidina cinasa (*tk*) o cloranfenicol acetiltransferasa (CAT) de herpesvirus simple. Un especialista en la materia sabría también cómo emplear marcadores inmunológicos, posiblemente junto con análisis de FACS. No se cree que sea importante el marcador usado, a condición de que pueda expresarse simultáneamente con el ácido nucleico que codifica un producto génico. Son bien conocidos ejemplos adicionales de marcadores de selección y rastreables para un especialista en la materia. Un rasgo de la presente invención incluye usar marcadores de selección y rastreables para seleccionar células exentas de vector después de que los factores de programación de la diferenciación hayan producido estado de diferenciación alterado deseado en esas células.

#### C. Suministro de vector

La introducción de un vector de reprogramación o programación de la diferenciación en células somáticas con la actual invención puede usar cualquier procedimiento adecuado para el suministro de ácido nucleico para transformación de una célula, como se describe en la presente memoria o como sería conocido por un especialista en la materia. Dichos procedimientos incluyen, pero sin limitación, el suministro directo de ADN tal como mediante transfección *ex vivo* (Wilson *et al.*, 1989, Nabel *et al.*, 1989); mediante inyección (patentes de EE.UU. n° 5.994.624, 5.981.274, 5.945.100, 5.780.448, 5.736.524, 5.702.932, 5.656.610, 5.589.466 y 5.580.859), incluyendo microinyección (Harland y Weintraub, 1985; patente de EE.UU. n° 5.789.215); mediante electroporación (patente de EE.UU. n° 5.384.253; Tur-Kaspa *et al.*, 1986; Potter *et al.*, 1984); mediante precipitación con fosfato de calcio (Graham y Van Der Eb, 1973; Chen y Okayama, 1987; Rippe *et al.*, 1990); mediante el uso de DEAE-dextrano seguido de polietilenglicol (Gopal, 1985); mediante carga sónica directa (Fechheimer *et al.*, 1987); mediante transfección mediada por liposoma (Nicolau y Sene, 1982; Fraley *et al.*, 1979; Nicolau *et al.*, 1987; Wong *et al.*, 1980; Kaneda *et al.*, 1989; Kato *et al.*, 1991) y transfección mediada por receptor (Wu y Wu, 1987; Wu y Wu, 1988); mediante bombardeo de microproyectiles (solicitudes PCT n° WO 94/09699 y 95/06128; patentes de EE.UU. n° 5.610.042, 5.322.783, 5.563.055, 5.550.318, 5.538.877 y 5.538.880); mediante agitación con fibras de carburo de silicio (Kaepler *et al.*, 1990; patentes de EE.UU. n° 5.302.523 y 5.464.765); mediante transformación mediada por *Agrobacterium* (patentes de EE.UU. n° 5.591.616 y 5.563.055); mediante transformación mediada por PEG de protoplastos (Omirulleh *et al.*, 1993; patentes de EE.UU. n° 4.684.611 y 4.952.500); mediante captación de ADN mediada por desecación/inhibición (Potrykus *et al.*, 1985) y cualquier combinación de dichos procedimientos. Mediante la aplicación de técnicas tales como estas, pueden transformarse estable o transitoriamente un orgánulo u orgánulos, una célula o células, un tejido o tejidos, un organismo u organismos.

##### i. Transfección mediada por liposoma

En una cierta realización de la invención, puede atraparse un ácido nucleico en un complejo lipídico tal como, por ejemplo, un liposoma. Los liposomas son estructuras vesiculares caracterizadas por una membrana de bicapa fosfolipídica y un medio acuoso interno. Los liposomas multilamelares tienen múltiples capas lipídicas separadas por medio acuoso. Se forman espontáneamente cuando se suspenden fosfolípidos en un exceso de solución acuosa. Los componentes lipídicos experimentan una autorredistribución antes de la formación de estructuras cerradas y atrapan el agua y los solutos disueltos entre las bicapas lipídicas (Ghosh y Bachhawat, 1991). Se contempla también un ácido nucleico complejado con Lipofectamine (Gibco BRL) o Superfect (Qiagen). La cantidad de liposomas usada puede variar con la naturaleza del liposoma así como de la célula usada, por ejemplo, pueden contemplarse de aproximadamente 5 a aproximadamente 20 µg de ADN de vector por 1 a 10 millones de células.

El suministro de ácido nucleico mediado por liposoma y la expresión de ADN extraño *in vitro* han sido muy exitosos

(Nicolau y Sene, 1982; Fraley *et al.*, 1979; Nicolau *et al.*, 1987). Se ha demostrado también la viabilidad del suministro mediado por liposoma y expresión de ADN extraño en embrión de pollo cultivado, células HeLa y de hepatoma (Wong *et al.*, 1980).

5 En ciertas realizaciones de la invención, un liposoma puede complejarse con un virus hemaglutinante (HVJ). Se ha mostrado que esto facilita la fusión con la membrana celular y promueve la entrada de ADN encapsulado en liposoma (Kaneda *et al.*, 1989). En otras realizaciones, un liposoma puede complejarse o emplearse junto con proteínas cromosómicas nucleares no histónicas (HMG-1) (Kato *et al.*, 1991). En aún otras realizaciones, un liposoma puede complejarse o emplearse junto con tanto HVJ como HMG-1. En otras realizaciones, un vehículo de suministro puede comprender un ligando y un liposoma.

10 **ii. Electroporación**

En ciertas realizaciones de la presente invención, se introduce un ácido nucleico en un orgánulo, célula, tejido u organismo mediante electroporación. La electroporación implica la exposición de una suspensión de células y ADN a una descarga eléctrica de alto voltaje. Las células receptoras pueden hacerse más susceptibles de transformación mediante lesión mecánica. También la cantidad de vectores usada puede variar con la naturaleza de las células usadas, por ejemplo, puede contemplarse de aproximadamente 5 a aproximadamente 20 µg de ADN de vector por 1 a 10 millones de células.

La transfección de células eucarióticas usando electroporación ha sido bastante exitosa. Se han transfectado prelinfocitos B de ratón con genes de kappa-inmunoglobulina humanos (Potter *et al.*, 1984) y se han transfectado hepatocitos de rata con el gen de cloranfenicol acetiltransferasa (Tur-Kaspa *et al.*, 1986) de esta manera.

20 **iii. Fosfato de calcio**

En otras realizaciones de la presente invención, se introduce un ácido nucleico en las células usando precipitación con fosfato de calcio. Se han transfectado células KB humanas con ADN de adenovirus 5 (Graham y Van Der Eb, 1973) usando esta técnica. También de esta manera, se transfectaron células L(A9) de ratón, C127 de ratón, CHO, CV-1, BHK, NIH3T3 y HeLa con un gen marcador de neomicina (Chen y Okayama, 1987) y se transfectaron hepatocitos de rata con una variedad de genes marcadores (Rippe *et al.*, 1990).

**iv. DEAE-dextrano**

En otra realización, se suministra un ácido nucleico a una célula usando DEAE-dextrano seguido de polietilenglicol. De esta manera, se introdujeron plásmidos informadores en células de mieloma y eritroleucemia de ratón (Gopal, 1985).

30 **v. Carga por sonicación**

Realizaciones adicionales de la presente invención incluyen la introducción de un ácido nucleico mediante carga sónica directa. Los fibroblastos LTK se han transfectado con el gen de timidina cinasa mediante carga por sonicación (Fechheimer *et al.*, 1987).

**vi. Transfección mediada por receptor**

35 Aún más, un ácido nucleico puede suministrarse a una célula diana mediante vehículos de suministro mediado por receptor. Estos aprovechan la captación selectiva de macromoléculas por endocitosis mediada por receptor que aparecerá en una célula diana. En vista de la distribución específica de tipo celular de diversos receptores, este procedimiento de suministro suma otro grado de especificidad a la presente invención.

40 Ciertos vehículos de orientación génica mediados por receptor comprenden un ligando específico de receptor celular y un agente de unión a ácido nucleico. Otros comprenden un ligando específico de receptor celular al que se ha enlazado operativamente el ácido nucleico para suministrar. Se han usado varios ligandos para la transferencia génica mediada por receptor (Wu y Wu, 1987; Wagner *et al.*, 1990; Perales *et al.*, 1994; Myers, EPO 0273085), que establecen la operatividad de la técnica. Se ha descrito el suministro específico en el contexto de otro tipo celular de mamífero (Wu y Wu, 1993). En ciertos aspectos de la presente invención, se elegirá un ligando para corresponder con un receptor expresado específicamente en la población de células diana.

45 En otras realizaciones, un componente de vehículo de suministro de ácido nucleico de un vehículo orientador de ácido nucleico específico de célula puede comprender un ligando de unión específica en combinación con un liposoma. El ácido o ácidos nucleicos para suministrar se albergan en el liposoma y el ligando de unión específica se incorpora funcionalmente a la membrana del liposoma. El liposoma por tanto se unirá específicamente al receptor o receptores de una célula diana y suministrará los contenidos a una célula. Se ha mostrado que dichos sistemas son funcionales usando sistemas en que, por ejemplo, se usa factor de crecimiento epidérmico (EGF) en el suministro mediado por receptor de un ácido nucleico a células que exhiben regulación positiva del receptor de EGF.

En aún otras realizaciones, el componente de vehículo de suministro de ácido nucleico de un vehículo de suministro orientado puede ser el liposoma mismo, que comprenderá preferiblemente uno o más lípidos o glicoproteínas que

dirigen la unión específica de célula. Por ejemplo, se ha incorporado lactosilceramida, un asialgangliósido galactosa-terminal, a liposomas y se ha observado un aumento de la captación del gen de insulina por hepatocitos (Nicolau *et al.*, 1987). Se contempla que los constructos transformantes específicos de tejido dados a conocer en la presente memoria puedan suministrarse específicamente a una célula diana de manera similar.

## 5 **vii Bombardeo de microprojectiles**

Las técnicas de bombardeo de microprojectiles pueden usarse para introducir un ácido nucleico en al menos un orgánulo, célula, tejido u organismo (patente de EE.UU. n° 5.550.318, patente de EE.UU. n° 5.538.880, patente de EE.UU. n° 5.610.042 y solicitud PCT WO 94/09699). Este procedimiento depende de la capacidad de acelerar microprojectiles recubiertos con ADN a una alta velocidad, permitiéndoles atravesar membranas celulares y entrar en células sin destruirlas (Klein *et al.*, 1987). Hay una amplia variedad de técnicas de bombardeo de microprojectiles conocidas en la materia, muchas de las cuales son aplicables a la invención.

En este bombardeo de microprojectiles, una o más partículas pueden recubrirse con al menos un ácido nucleico y suministrarse a células mediante una fuerza propulsora. Se han desarrollado varios dispositivos para acelerar partículas pequeñas. Uno de dichos dispositivos se basa en una descarga de alto voltaje para generar una corriente eléctrica que, a su vez, proporciona la fuerza motora (Yang *et al.*, 1990). Los microprojectiles usados han consistido en sustancias biológicamente inertes tales como partículas o perlas de wolframio u oro. Las partículas ejemplares incluyen aquellas que comprenden wolframio, platino y preferiblemente oro. Se contempla que, en algunos casos, la precipitación de ADN sobre partículas metálicas no sería necesaria para el suministro de ADN a una célula receptora usando bombardeo de microprojectiles. Sin embargo, se contempla que las partículas pueden contener ADN en lugar de recubrirse con ADN. Las partículas recubiertas con ADN pueden aumentar el nivel de suministro de ADN a través del bombardeo de partículas pero no son necesarias en sí mismas.

Para el bombardeo, se concentran células en suspensión sobre filtros o medio de cultivo sólido. Como alternativa, pueden disponerse embriones inmaduros u otras células diana sobre medio de cultivo sólido. Se colocan las células para bombardear a una distancia apropiada por debajo de la placa de detención de macroprojectiles

## 25 **VI. Selección de iPS**

En ciertos aspectos de la invención, después de introducir un vector de reprogramación en células somáticas, las células se cultivarán para expansión (opcionalmente seleccionadas por la presencia de elementos vectoriales como un marcador de selección positiva o rastreable para concentrar células transfectadas) y los vectores de reprogramación expresarán factores de reprogramación en estas células y se replicarán y repartirán junto con la división celular. Estos factores de reprogramación expresados reprogramarán el genoma de células somáticas para establecer un estado pluripotente automantenido, y mientras tanto o después de la retirada de la selección positiva de la presencia de vectores, se perderán gradualmente los elementos genéticos exógenos. Estas células madre pluripotentes inducidas podrían seleccionarse de la progenie derivada de estas células somáticas basándose en las características de las células madre embrionarias, porque se espera que sean sustancialmente idénticos a los células madre embrionarias pluripotentes. Podría emplearse también una etapa de selección negativa adicional para acelerar o ayudar a la selección de iPS esencialmente exentos de elementos genéticos exógenos ensayando la ausencia de ADN de vector de reprogramación o usando marcadores de selección.

### **A. Selección de las características de células madre embrionarias**

Los iPSC generados exitosamente de estudios anteriores eran notablemente similares a las células madre pluripotentes aislados naturales (tales como células madre embrionarias de ratón y humanos, mESC y hESC, respectivamente) en los siguientes aspectos, confirmando por tanto la identidad, autenticidad y pluripotencia de los iPSC con células madre pluripotentes aisladas naturales. Por tanto, las células madre pluripotentes inducidas generadas con los procedimientos dados a conocer en esta invención podrían seleccionarse basándose en una o más de las siguientes características de células madre embrionarias

#### 45 **i. Propiedades biológicas celulares**

**Morfología:** Los iPSC son morfológicamente similares a los ESC. Cada célula puede tener forma redonda, un nucleolo grande y citoplasma escaso. Las colonias de iPSC podrían ser similares también a las de ESC. Los iPSC humanos forman colonias de bordes agudos, planas y estrechamente empaquetadas similares a los hESC y los iPSC de ratón forman colonias similares a mESC, colonias menos planas y más agregadas que las de hESC.

50 **Propiedades de crecimiento:** El tiempo de duplicación y la actividad mitótica son pilares de los ESC, ya que las células madre deben autorrenovarse como parte de su definición. Los iPSC podrían ser mitóticamente activos, autorrenovarse activamente, proliferar y dividirse a un índice igual a los ESC.

**Marcadores de células madre:** Los iPSC pueden expresar marcadores antigénicos de superficie celular expresados en ESC. Los iPSC humanos expresaban los marcadores específicos de hESC incluyendo, pero sin limitación, SSEA-3, SSEA-4, TRA-1-60, TRA-1-81, TRA-2-49/6E y Nanog. Los iPSC de ratón expresaban SSEA-1, pero no SSEA-3 ni SSEA-4, de forma similar a mESC.

**Genes de células madre:** Los iPSC pueden expresar genes expresados en ESC indiferenciados, incluyendo Oct-3/4, Sox2, Nanog, GDF3, REX1, FGF4, ESG1, DPPA2, DPPA4 y hTERT.

5 **Actividad de telomerasa:** Las telomerasas son necesarias para mantener la división celular sin restricciones por el límite de Hayflick de ~50 divisiones celulares. Los hESC expresan una alta actividad de telomerasa para mantener la autorrenovación y proliferación, y los iPSC demuestran también una alta actividad de telomerasa y expresan hTERT (transcriptasa inversa de telomerasa humana), un componente necesario del complejo de telomerasa-proteína.

**Pluripotencia:** Los iPSC serán capaces de diferenciación de forma similar a los ESC hasta tejidos diferenciados totalmente.

10 **Diferenciación neural:** Los iPSC podrían diferenciarse hasta neuronas, expresando  $\beta$ III-tubulina, tirosina hidroxilasa, AADC, DAT, ChAT, LMX1B y MAP2. La presencia de enzimas asociadas a catecolamina puede indicar que los iPSC, como hESC, pueden ser diferenciables hasta neuronas dopaminérgicas. Los genes asociados a células madre se regularán negativamente después de la diferenciación.

15 **Diferenciación cardíaca:** Los iPSC podrían diferenciarse en cardiomiocitos que empezaban espontáneamente a latir. Los cardiomiocitos expresaban TnTc, MEF2C, MYL2A, MYHC $\beta$  y NKX2.5. Los genes asociados a células madre se regularán negativamente después de la diferenciación.

20 **Formación de teratoma:** Los iPSC inyectados en ratones inmunodeficientes pueden formar espontáneamente teratomas después de cierto tiempo, tal como 9 semanas. Los teratomas son tumores de linajes múltiples que contienen tejido derivado de las tres capas germinales endodermo, mesodermo y ectodermo; esto es a diferencia de otros tumores, que son típicamente de un solo tipo celular. La formación de teratoma es emblemática de pluripotencia.

**Cuerpo embriode:** Los hESC en cultivo forman espontáneamente estructuras de tipo embrión de tipo bola denominadas “cuerpos embrioides”, que consisten en un núcleo de hESC mitóticamente activos y diferenciadores y una periferia de células totalmente diferenciadas de las tres capas germinales. Los iPSC pueden formar también cuerpos embrioides y tienen células diferenciadas periféricas.

25 **Inyección de blastocitos:** Los hESC residen naturalmente en la masa celular interna (embrioblasto) de los blastocitos y, en el embrioblasto, se diferencian en el embrión, mientras que la cubierta del blastocito (troboblasto) se diferencia en tejidos extraembrionarios. El troboplasto hueco es incapaz de formar un embrión vivo, y por tanto es necesario que las células madre embrionarias en el embrioblasto se diferencien y formen el embrión. Los iPSC inyectados por micropipeta en un troboplasto, para generar un blastocito transferido a hembras receptoras, pueden dar como resultado crías de ratón vivos quiméricos: ratones con derivados de iPSC incorporados por todos sus cuerpos con un 10-90 % de quimerismo

30

## ii. Reprogramación epigenética

35 **Desmetilación de promotor:** La metilación es la transferencia de un grupo metilo a una base de ADN, típicamente la transferencia de un grupo metilo a una molécula de citosina en un sitio CpG (secuencia adyacente de citosina/guanina). La metilación extendida de un gen interfiere con la expresión al evitar la actividad de las proteínas de expresión o agrupar enzimas que interfieren con la expresión. Por tanto, la metilación de un gen lo silencia eficazmente al evitar la transcripción. Los promotores de genes asociados a la pluripotencia, incluyendo Oct-3/4, Rex1 y Nanog, pueden desmetilarse en iPSC, mostrando su actividad promotora y la promoción y expresión activa de genes asociados a la pluripotencia en iPSC.

40 **Desmetilación de histona:** Las histonas son proteínas compactadoras que se localizan estructuralmente en secuencias de ADN que pueden producir su acción a través de diversas modificaciones relacionadas con la cromatina. Las histonas H3 asociadas a Oct-3/4, Sox2 y Nanog pueden desmetilarse para activar la expresión de Oct-3/4, Sox2 y Nanog.

## B. Selección del rasgo de exento de residuos

45 Un vector de reprogramación tal como un plásmido basado en oriP de esta invención se replicará extracromosómicamente y perderá su presencia en células hospedadoras después de generaciones. Sin embargo, una etapa de selección adicional de células progenie esencialmente exentas de elementos vectoriales exógenos puede facilitar este proceso. Por ejemplo, puede extraerse una muestra de célula de progenie para ensayar la presencia o pérdida de elementos vectoriales exógenos como es conocido en la materia (Leight y Sugden, “Molecular and Cellular Biology”, 2001).

50

Un vector de reprogramación puede comprender además un marcador de selección, más específicamente un marcador de selección negativa tal como un gen que codifica una timidina cinasa para seleccionar las células de progenie esencialmente exentas de dicho marcador de selección. El gen de timidina cinasa de tipo 1 de herpesvirus simple (HSVtk) actúa como un marcador mortal condicional en células de mamífero. La enzima codificada por HSVtk es capaz de fosforilar ciertos análogos nucleosídicos (p.ej. ganciclovir, un fármaco antiherpético), convirtiéndolos así

55



en inhibidores tóxicos de la replicación de ADN. Es un enfoque alternativo o complementario ensayar la ausencia de elementos genéticos exógenos en células de progenie usando procedimientos convencionales, tales como PCR-TI, PCR, FISH (hibridación fluorescente *in situ*), matriz génica o hibridación (p.ej., transferencia Southern).

## VII. Cultivo de iPS

5 Después de introducir en células somáticas un vector de reprogramación usando los procedimientos dados a conocer, estas células pueden cultivarse en un medio suficiente para mantener la pluripotencia. El cultivo de células madre pluripotentes inducidas (iPS) generados en esta invención puede usar diversos medios y técnicas desarrollados para cultivar células madre pluripotentes de primates, más especialmente células madre embrionarias, como se describe en la solicitud de patente de EE.UU. n° 20070238170 y la solicitud de patente de EE.UU. 20030211603.

10 Por ejemplo, como las células madre embrionarias humanas (hES), los iPS pueden mantenerse en 80 % de DMEM (Gibco n° 10829-018 o n° 11965-092); 20 % de suero fetal bovino definido (FBS) no termoinactivado, 1 % de aminoácidos no esenciales, L-glutamina 1 mM y beta-mercaptoetanol 0,1 mM. Como alternativa, los ES pueden mantenerse en medio exento de suero, compuesto por 80 % de Knock-Out DMEM (Gibco n° 10829-018), 20 % de reemplazo de suero (Gibco n° 10828-028), 1 % de aminoácidos no esenciales, L-glutamina 1 mM y beta-mercaptoetanol 0,1 mM. Justo antes del uso, se añade bFGF humano a una concentración final de aproximadamente 4 ng/ml (documento WO 99/20741).

15 Los IPS, como los ES, tienen antígenos característicos que pueden identificarse por inmunohistoquímica o citometría de flujo, usando los anticuerpos de SSEA-1, SSEA-3 y SSEA-4 (Developmental Studies Hybridoma Bank, National Institute of Child Health and Human Development, Bethesda, Md.) y TRA-1-60 y TRA-1-81 (Andrews *et al.* en Robertson E. ed: "Teratocarcinomas and Embryonic Stem Cells", IRL Press, 207-246, 1987). La pluripotencia de las células madre embrionarias puede confirmarse inyectando aproximadamente  $0,5-10 \times 10^6$  células en los músculos de las patas traseras de ratones SCID macho de 8-12 semanas. Se desarrollan teratomas que demuestran al menos un tipo celular de cada una de las tres capas germinales.

## 25 VIII. Ejemplos

Los siguientes ejemplos se incluyen para demostrar realizaciones preferidas de la invención. Debería apreciarse por los especialistas en la materia que las técnicas dadas a conocer en los ejemplos siguientes representan técnicas descubiertas por el inventor que funcionan bien en la práctica de la invención, y por tanto pueden constituir modos preferidos para su práctica.

### 30 Ejemplo 1. Construcción de un plásmido de reprogramación exento de residuos

Los inventores construyen un plásmido de esqueleto receptor que contiene la secuencia oriP que incluye DS y FR separados por aproximadamente 1000 pares de bases derivados de EBV (véase Lindner y Sugden, 2007) y la forma abreviada de EBNA-1 de tipo silvestre conocida como deltaUR1 (a la que se hace referencia como DomNeg2 en Kennedy *et al.* 2003) (FIG. 3). Se construye actualmente el plásmido para expresar deltaUR1 impulsado por el promotor factor de alargamiento 1a (EF1a), que contiene también una secuencia intrónica para maximizar la expresión de deltaUR1. El plásmido de esqueleto se ha configurado actualmente para incluir un marcador de selección para células de mamífero que codifica resistencia a higromicina; sin embargo, la elección del marcador de resistencia sigue siendo flexible según la sensibilidad de la línea celular en que se introduzca el plásmido. De forma similar, el plásmido codifica resistencia a fármacos para selección procariótica y, en este caso, el plásmido codifica resistencia a ampicilina.

40 Los inventores han integrado una serie de módulos en el plásmido receptor descrito anteriormente que codifican los genes requeridos para y que contribuyen a que las células de reprogramación se vuelvan pluripotentes (concretamente, iPS). Un módulo codificaba dos genes esenciales para el proceso de reprogramación, Sox-2 y Oct-4 (FIG. 4). Los inventores podrían usar el promotor de fosfoglicerato cinasa (PGK), promotor del gen inmediato temprano de citomegalovirus (CMV) o promotor de SV40 para impulsar la expresión de Sox-2 y Oct-4, pero esta elección está sometida a cambios dependiendo de la eficacia de esa expresión. Opcionalmente, han incluido también el segundo intrón del gen de beta-globina humana para maximizar también la expresión del transcrito. Ambos genes, por lo tanto, se codificaban por el mismo transcrito, mientras que la traducción podría iniciarse a partir de un ATG canónico para Sox-2 y de un sitio de entrada a ribosoma interno (IRES) derivado del virus de la encefalomiocarditis para Oct-4. De forma similar, otros módulos codificaban un transcrito bicistrónico que codifica Nanog y Lin28 o codificaban un transcrito bicistrónico que codifica Klf4 y c-Myc (que se impulsaban por el promotor PGK o cualquier otro promotor adecuado) y separados por un IRES también (FIG. 4). Podrían usarse variaciones de un transcrito multicistrónico que comprende dos cualesquiera o más de Sox-2, Oct-4, Nanog, Lin28, Klf4, c-Myc y EBNA-1.

55 Podría haber ciertas variaciones del sistema para optimizar su eficacia. La bibliografía actual indica que Lin28 podría ser descartable para el proceso de reprogramación y por lo tanto es probable que los inventores pudieran ajustar este sistema de plásmido para incluir solo Sox-2, Oct-4 y Nanog. Además, el tipo de IRES elegido se ha probado funcional, aunque en el contexto de otros conjuntos génicos. Sin embargo, es posible que el IRES pueda probarse

inadecuado para promover los niveles de expresión requeridos para una apropiada reprogramación y pueda dar como resultado la rotura de los módulos de tal modo que cada gen de reprogramación sea impulsado por su propio promotor humano.

- 5 En resumen, el plásmido lanzadera maestro o plásmido de reprogramación podría codificar Sox-2, Oct-4, Nanog y posiblemente Lin28 (FIG. 5), mientras que su replicación y mantenimiento podrían promoverse por la presencia de oriP y deltaUR1. Este plásmido podría estar preparado también para futuras modificaciones que incluyen un marcador de selección negativa tal como timidina cinasa y un marcador de selección positiva adicional tal como secuencias que codifican la proteína fluorescente verde o roja.

**Ejemplo 2. Uso de un plásmido de reprogramación exento de residuos**

- 10 La reprogramación eficaz dependerá de la introducción eficaz de este gran plásmido (15-20 kb) en células de mamífero. Los inventores están empleando actualmente un enfoque basado en la lipofilia para introducir el ADN en fibroblastos humanos; sin embargo, este enfoque es probable que sea modificado según el tipo celular que se esté transfectando. Por ejemplo, probablemente elegirían electroporación para la introducción de plásmidos de ADN en células hematopoyéticas. Una vez las células se transfectan apropiadamente, los inventores dispondrán estas  
15 células en un lecho de fibroblastos embrionarios de ratón (MEF) irradiados o Matrigel sobre discos de cultivo de 10 cm usando medios adecuados para las células transfectadas. Aproximadamente 6 días después de la transfección, se cambiarán los medios por medios especializados en reprogramación de células y se reemplazarán diariamente cada dos días (Yu *et al.*, 2007).

- 20 Basándose en el procedimiento actual de generación de iPS, los inventores seleccionarán probablemente colonias que se parezcan a células madre aproximadamente 20 días después de la transfección y las transferirán a MEF o Matrigel en placas de cultivo de 6 pocillos alimentando diariamente o cada dos días con medios especializados. Una vez ha tenido lugar una expansión suficiente, se cariotiparán los clones y se probarán los marcadores específicos apropiados para células madre.

- 25 Todos los procedimientos dados a conocer y reivindicados en la presente memoria pueden elaborarse y ejecutarse sin experimentación indebida a la vista de la presente divulgación.

**REFERENCIAS**

- Patente de EE.UU. 4.683.202  
 Patente de EE.UU. 4.684.611  
 Patente de EE.UU. 4.952.500  
 30 Patente de EE.UU. 5.302.523  
 Patente de EE.UU. 5.384.253  
 Patente de EE.UU. 5.384.253  
 Patente de EE.UU. 5.464.765  
 Patente de EE.UU. 5.538.877  
 35 Patente de EE.UU. 5.538.880  
 Patente de EE.UU. 5.550.318  
 Patente de EE.UU. 5.580.859  
 Patente de EE.UU. 5.589.466  
 Patente de EE.UU. 5.591.616  
 40 Patente de EE.UU. 5.610.042  
 Patente de EE.UU. 5.656.610  
 Patente de EE.UU. 5.702.932  
 Patente de EE.UU. 5.736.524  
 Patente de EE.UU. 5.780.448  
 45 Patente de EE.UU. 5.789.215

- Patente de EE.UU. 5.925.565
- Patente de EE.UU. 5.928.906
- Patente de EE.UU. 5.935.819
- Patente de EE.UU. 5.945.100
- 5 Patente de EE.UU. 5.981.274
- Patente de EE.UU. 5.994.624
- Solicitud de patente de EE.UU. 20030211603
- Solicitud de patente de EE.UU. 20070238170
- Adams, J. Virol., 61(5): 1743-1746,1987.
- 10 Aiyar *et al.*, EMBO J., 17(21): 6394-6403, 1998.
- Alexander *et al.*, Proc. Nat. Acad. Sci. USA, 85: 5092-5096, 1988.
- Almquist *et al.*, Med. Chem., 23(12): 1392-1398, 1980.
- Altmann *et al.*, Proc. Natl. Acad. Sci. USA, 103(38): 14188-14193, 2006.
- Andrews *et al.*, en: "Teratocarcinomas and Embryonic Stem Cells", Robertson (Ed.), IRL Press, 207-246, 1987.
- 15 Aravind y Landsman, Nucleic Acids Res., 26(19): 4413-4421, 1998.
- Ausubel *et al.*, "Current Protocols in Molecular Biology", Greene Publ. Assoc. Inc: & John Wiley & Sons, Inc, MA; 1994.
- Baer *et al.*, Biochemistry, 39: 7041-7049, 2000.
- Baer *et al.*, Nature, 310(5974): 207-211, 1984.
- 20 Bingham, Cell, 90(3): 385-387, 1997.
- Bochkarev *et al.*, Cell, 84(5): 791-800, 1996.
- Bode *et al.*, Biol. Chem., 381: 801-813, 2000b.
- Bode *et al.*, Gene Ther. Mol. Biol., 6: 33-46, 2001.
- Bode *et al.*, Science, 255(5041): 195-197,1992.
- 25 Boyer *et al.*, Cell, 122(6): 947-56, 2005.
- Brambrink *et al.*, Cell Stem Cell, 7(2): 151-159, 2008.
- Carbonelli *et al.*, FEMS Microbiol. Lett., 177(1): 75-82, 1999.
- Chambers *et al.*, Cell, 113(5): 643-55, 2003.
- Chandler *et al.*, Proc. Natl. Acad. Sci. USA, 94(8): 3596-601, 1997.
- 30 Chaudhuri *et al.*, Proc. Natl. Acad. Sci. USA, 98(18): 10085-10089, 2001.
- Chen y Okayama, Mol. Cell Biol., 7(8): 2745-2752, 1987.
- Cocea, Biotechniques, 23(5): 814-816, 1997.
- Dhar *et al.*, Cell, 106(3): 287-296, 2001.
- EPO 0273085
- 35 EPO 45665
- Ercolani *et al.*, J. Biol. Chem., 263: 15335-15341, 1988.
- Ermakova *et al.*, J. Biol. Chem., 271(51): 33009-33017, 1996.

- Evans, *et al.*, en: "Cancer Principles and Practice of Oncology", Devita *et al.* (Eds.), Lippincot-Raven, NY, 1054-1087, 1997.
- Fechheimer *et al.*, Proc Natl. Acad. Sci. USA, 84: 8463-8467, 1987.
- Fischer *et al.*, J. Virol., 71: 5148-5146, 1997.
- 5 Fraley *et al.*, Proc. Natl. Acad. Sci. USA, 76: 3348-3352, 1979.
- Frappier y O'Donnell, Proc. Natl. Acad. Sci. USA, 88(23): 10875-10879, 1991.
- Gahn y Schildkraut, Cell, 58(3): 527-535, 1989.
- Gahn y Sugden, J. Virol., 69(4): 2633-2636, 1995.
- Garricket *et al.*, Nat. Genet., 18: 56-59, 1998.
- 10 Ghosh y Bachhawat Gopal, 1985
- Gopal, Mol. Cell Biol.; 5: 1188-1190, 1985.
- Graham y Van Der Eb, Virology, 52: 456-467, 1973.
- Hann, J. Chem. Soc. Perkin Trans., I 307-314, 1982.
- Harland y Weintraub, J. Cell. Biol., 101(3): 1094-1099, 1985.
- 15 Hegde *et al.*, Nature, 359(6395): 505-512, 1992.
- Holladay *et al.*, Tetrahedron. Lett. 24: 4401-4404, 1983.
- Hruby, Life Sci., 31: 189-199, 1982.
- Hudson *et al.*, Int. J. Pept. Prot. Res., 14: 177-185, 1979.
- Hung *et al.*, Proc. Natl. Acad. Sci. USA, 98(4): 1865-1870, 2001.
- 20 Jenke *et al.*, Proc. Natl. Acad. Sci. USA, 101 (31), 11322-11327, 2004.
- Jennings-White *et al.*, Tetrahedron Lett., 23: 2533, 1982.
- Julien *et al.*, Virology, 326(2): 317-328, 2004.
- Kaepler *et al.*, Plant Cell Reports, 9: 415-418, 1990.
- Kanda *et al.*, Mol. Cell. Biol., 21(10): 3576-3588, 2001.
- 25 Kaneda *et al.*, Science, 243: 375-378, 1989.
- Karin *et al.*, Cell, 36: 371-379, 1989.
- Kato *et al.*, J. Biol. Chem., 266: 3361-3364, 1991.
- Kennedy y Sugden, Mol. Cell. Biol., 23(19): 6901-6908, 2003.
- Kennedy *et al.*, Proc. Natl. Acad. Sci. USA, 100: 14269-14274, 2003.
- 30 Kim *et al.*, J. Biol. Chem., 275(40): 31245-31254, 2000.
- Kim *et al.*, Virology, 239(2): 340-351, 1997.
- Kirchmaier y Sugden, J. Virol. 69(2): 1280-1283, 1995.
- Kirchmaier y Sugden, J. Virol., 71(3): 1766-1775, 1997.
- Kirchmaier y Sugden, J. Virol. 72(6): 4657-4666, 1998.
- 35 Klein *et al.* Nature; 327: 70-73, 1987.
- Langle-Rouault *et al.*, J. Virol., 72(7): 6181-6185, 1998.
- Leight y Sugden, Mol. Cell Biol. 21: 4149-61, 2001.

- Levenson *et al.*, Hum. Gene Ther.; 9(8): 1233-1236, 1998.
- Levitskaya *et al.*, Nature, 375(6533): 685-688, 1995.
- Levitskaya *et al.*, Proc. Natl. Acad. Sci. USA, 94(23): 12616-12621, 1997.
- Lindner y Sugden, Plasmid, 58:1-12, 2007.
- 5 Lindner *et al.*, J. Virol., 82(12): 5693-702, 2008.
- Macejak y Sarnow, Nature, 353: 90-94, 1991.
- Mackey y Sugden, Mol. Cell. Biol., 19(5): 3349-3359, 1999.
- Mackey *et al.*, J. Virol., 69(10): 6199-6208, 1995.
- Maniatis, *et al.*, "Molecular Cloning, A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 1958.
- 10 Manzini *et al.*, Proc. Natl. Acad. Sci. USA, 103(47): 17672-17677, 2006.
- Marechal *et al.*, J. Virol., 73(5): 4385-4392, 1999.
- Middleton y Sugden, J. Virol., 66(1): 489-495, 1992.
- Morley, Trends Pharm. Sci., 463-468, 1980.
- Nabel *et al.*, Science, 244(4910): 1342-1344, 1989.
- 15 Nanbo y Sugden, EMBO J., 26: 4252-62, 2007.
- Ng, Nuc. Acid Res., 17: 601-615, 1989.
- Nicolau y Sene, Biochim. Biophys. Acta, 721: 185-190, 1982.
- Nicolau *et al.*, Methods Enzymol., 149: 157-176, 1987.
- Niller *et al.*, J. Biol. Chem., 270(21): 12864-12868, 1995.
- 20 Omirulleh *et al.*, Plant Mol. Biol., 21(3): 415-428, 1993.
- Pelletier y Sonenberg, Nature, 334(6180): 320-325, 1988.
- Perales *et al.*, Proc. Natl. Acad. Sci. USA, 91: 4086-4090, 1994.
- Piechaczek *et al.*, Nucleic Acids Res., 27(2): 426-428, 1999.
- Potrykus *et al.*, Mol. Gen. Genet., 199(2): 169-177, 1985.
- 25 Potter *et al.*, Proc. Natl. Acad. Sci. USA, 81: 7161-7165, 1984.
- Quitsche *et al.*, J. Biol. Chem., 264: 9539-9545, 1989.
- Rawlins *et al.*, Cell, 42(3): 859-868, 1985.
- Reisman y Sugden, Mol. Cell. Biol., 6(11): 3838-3846, 1986.
- Reisman *et al.*, Mol. Cell. Biol., 5(8): 1822-1832, 1985.
- 30 Richards *et al.*, Cell, 37: 263-272, 1984.
- Rippe, *et al.*, Mol. Cell. Biol. 10: 689-695, 1990.
- Ritzi *et al.*, J. Cell Sci., 116(pt 19): 3971-3984, 2003.
- Sambrook *et al.*, en: "Molecular cloning: a laboratory manual", 2<sup>a</sup> Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor; NY, 1989:
- 35 Schaarschmidt *et al.*, EMBO J., 23(1):191-201, 2004.
- Schepers *et al.*; EMBO J., 20(16): 4588-4602, 2001.
- Sears *et al.*, J. Virol., 77(21): 11767-11780, 2003.

- Sears *et al.*, J. Virol., 78(21): 11487-11505, 2004.
- Shire *et al.*, J. Virol., 73(4): 2587-2595, 1999.
- Spatola *et al.*, Life Sci., 38: 1243-1249, 1986.
- Spatola, en: "Peptide Backbone Modifications", 1:3, 1983.
- 5 Spatola, en: "Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins", Weinstein (Ed.), Marcel Dekker, NY, 267, 1983.
- Stadtfeld *et al.*, Cell Stem Cell, 2: 230-240; 2008.
- Su *et al.*; Proc. Natl. Acad. Sci. USA; 88(23): 10870-19874, 1991.
- Sugden y Warren, J. Virol., 63(6): 2644-2649, 1989.
- 10 Takahashi *et al.*, Cell 126(4): 663-676, 2006.
- Takahashi *et al.*, Cell, 126(4): 663-76, 2007.
- Torchia *et al.*; Curr. Opin. Cell Biol., 10: 373-383, 1998.
- Tur-Kaspa *et al.*, Mol. Cell Biol., 6: 716-718, 1986.
- Wagner *et al.*, Proc. Natl. Acad. Sci. USA 87(9): 3410-3414, 1990.
- 15 Wang *et al.*, Mol. Cell. Biol., 26(3): 1124-1134, 2006.
- Wilson *et al.*, Science, 244: 1344-1346, 1989.
- WO 94/09699
- WO 95/06128
- WO 99/20741
- 20 Wongetal., Gene, 10: 87-94, 1980.
- Wu y Wu, Adv. Drug Delivery Rev., 12: 159-167, 1993.
- Wu y Wu, Biochemistry, 27: 887-892, 1988.
- Wu y Wu, J. Biol. Chem., 262: 4429-4432, 1987.
- Wu *et al.*, J. Virol.; 76(5): 2480-2490, 2002.
- 25 Wysokenski y Yates, J. Virol., 63(6): 2657-2666, 1989.
- Yamanaka *et al.*, Cell, 131(5): 861-72, 2007.
- Yang y Russell, Proc. Natl. Acad. Sci. USA, 87: 4144-4148, 1990.
- Yates y Guan, J. Virol., 65(1): 483-488, 1991.
- Yates *et al.*, J. Virol., 74(10): 4512-4522, 2000.
- 30 Yates *et al.*, Nature, 313: 812-815, 1985.
- Yates *et al.*, Proc. Natl. Acad. Sci. USA, 81: 3806-3810, 1984.
- Yates, Cancer Cells, (6)197-205, 1988.
- Yin *et al.*, Science, 301(5638): 1371-1374, 2003.
- Yu *et al.*, Science, 318: 1917-1920, 2007.
- 35 Zhou *et al.*, EMBO J., 24(7): 1406-1417, 2005.

**LISTADO DE SECUENCIAS**

<110> MACK, AMANDA

THOMSON, JAMES

- <120> PROCEDIMIENTOS PARA LA PRODUCCIÓN DE CÉLULAS IPS
- <130> CDIN:005WO
- <140> DESCONOCIDO
- 5 <141> 03-06-2009
- <150> 61/058.858
- <151> 04-06-2008
- <150> 61/160.584
- <151> 16-03-2009
- 10 <160> 4
- <170> PatentIn versión 3.5
- <210> 1
- <211> 641
- <212> PRT
- 15 <213> Herpesvirus humano 4
- <400> 1

```

Met Ser Asp Glu Gly Pro Gly Thr Gly Pro Gly Asn Gly Leu Gly Glu
1          5          10          15

Lys Gly Asp Thr Ser Gly Pro Glu Gly Ser Gly Gly Ser Gly Pro Gln
          20          25          30

Arg Arg Gly Gly Asp Asn His Gly Arg Gly Arg Gly Arg Gly Arg Gly
          35          40          45

Arg Gly Gly Gly Arg Pro Gly Ala Pro Gly Gly Ser Gly Ser Gly Pro
          50          55          60

Arg His Arg Asp Gly Val Arg Arg Pro Gln Lys Arg Pro Ser Cys Ile
65          70          75          80

Gly Cys Lys Gly Thr His Gly Gly Thr Gly Ala Gly Ala Gly Ala Gly
          85          90          95

Gly Ala Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Ala Gly
          100          105          110

Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly
          115          120          125
    
```

Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala  
 130 135 140

Gly Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly Ala Gly  
 145 150 155 160

Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly  
 165 170 175

Ala Gly Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly  
 180 185 190

Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Ala Gly Gly Ala Gly  
 195 200 205

Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala  
 210 215 220

Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala  
 225 230 235 240

Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly  
 245 250 255

Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly  
 260 265 270

Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly  
 275 280 285

Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Gly Ala Gly  
 290 295 300

Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Ala Gly  
 305 310 315 320

Gly Ala Gly Ala Gly Gly Gly Gly Arg Gly Arg Gly Gly Ser Gly Gly  
 325 330 335

Arg Gly Arg Gly Gly Ser Gly Gly Arg Gly Arg Gly Gly Ser Gly Gly  
 340 345 350

Arg Arg Gly Arg Gly Arg Glu Arg Ala Arg Gly Gly Ser Arg Glu Arg  
 355 360 365

Ala Arg Gly Arg Gly Arg Gly Arg Gly Glu Lys Arg Pro Arg Ser Pro



370						375										380
Ser 385	Ser	Gln	Ser	Ser	Ser 390	Ser	Gly	Ser	Pro	Pro 395	Arg	Arg	Pro	Pro	Pro	400
Gly	Arg	Arg	Pro	Phe 405	Phe	His	Pro	Val	Gly 410	Glu	Ala	Asp	Tyr	Phe 415	Glu	
Tyr	His	Gln	Glu 420	Gly	Gly	Pro	Asp	Gly 425	Glu	Pro	Asp	Val	Pro 430	Pro	Gly	
Ala	Ile	Glu 435	Gln	Gly	Pro	Ala	Asp 440	Asp	Pro	Gly	Glu	Gly 445	Pro	Ser	Thr	
Gly	Pro 450	Arg	Gly	Gln	Gly	Asp 455	Gly	Gly	Arg	Arg	Lys 460	Lys	Gly	Gly	Trp	
Phe 465	Gly	Lys	His	Arg	Gly 470	Gln	Gly	Gly	Ser	Asn 475	Pro	Lys	Phe	Glu	Asn 480	
Ile	Ala	Glu	Gly	Leu 485	Arg	Ala	Leu	Leu	Ala 490	Arg	Ser	His	Val	Glu 495	Arg	
Thr	Thr	Asp	Glu 500	Gly	Thr	Trp	Val	Ala 505	Gly	Val	Phe	Val	Tyr 510	Gly	Gly	
Ser	Lys	Thr 515	Ser	Leu	Tyr	Asn 520	Leu	Arg	Arg	Gly	Thr	Ala 525	Leu	Ala	Ile	
Pro	Gln	Cys	Arg	Leu	Thr	Pro 535	Leu	Ser	Arg	Leu	Pro 540	Phe	Gly	Met	Ala	
Pro 545	Gly	Pro	Gly	Pro	Gln 550	Pro	Gly	Pro	Leu	Arg 555	Glu	Ser	Ile	Val	Cys 560	
Tyr	Phe	Met	Val	Phe 565	Leu	Gln	Thr	His	Ile 570	Phe	Ala	Glu	Val	Leu 575	Lys	
Asp	Ala	Ile	Lys 580	Asp	Leu	Val	Met	Thr 585	Lys	Pro	Ala	Pro	Thr 590	Cys	Asn	
Ile	Arg	Val 595	Thr	Val	Cys	Ser	Phe 600	Asp	Asp	Gly	Val	Asp 605	Leu	Pro	Pro	
Trp 610	Phe	Pro	Pro	Met	Val	Glu 615	Gly	Ala	Ala	Ala	Glu 620	Gly	Asp	Asp	Gly	

Asp Asp Gly Asp Glu Gly Gly Asp Gly Asp Glu Gly Glu Glu Gly Gln  
 625 630 635 640

Glu

<210> 2

<211> 1926

<212> ADN

5 <213> Herpesvirus humano 4

<400> 2

atgtctgacg aggggccagg tacaggacct ggaatggcc taggagagaa gggagacaca 60  
 tctggaccag aaggctccgg cggcagtgga cctcaaagaa gagggggtga taaccatgga 120  
 cgaggacggg gaagaggacg aggacgagga ggcggaagac caggagcccc gggcggctca 180  
 ggatcagggc caagacatag agatggtgtc cggagacccc aaaaacgtcc aagttgcatt 240  
 ggctgcaaag ggaccacagg tggaacagga gcaggagcag gagcgggagg ggcaggagca 300  
 ggaggggagc gagcaggagg aggggcagga gcaggaggag gggcaggagg ggcaggaggg 360  
 gcaggagggg caggagcagg aggaggggca ggagcaggag gaggggcagg aggggcagga 420  
 ggggcaggag caggaggagg ggcaggagca ggaggagggg caggaggggc aggagcagga 480  
 ggaggggagc gaggggcagg aggggcagga gcaggaggag gggcaggagc aggaggaggg 540  
 gcaggagggg caggagcagg aggaggggca ggaggggagc gaggggcagg agcaggagga 600  
 ggggcaggag caggaggggc aggaggggca ggaggggagc gagcaggagg ggcaggagca 660  
 ggaggagggg caggaggggc aggaggggca ggagcaggag gggcaggagc aggaggggca 720  
 ggagcaggag gggcaggagc aggaggggca ggaggggagc gagcaggagg ggcaggaggg 780  
 gcaggagcag gaggggcagg aggggcagga gcaggaggag gggcaggagg ggcaggagca 840  
 ggaggagggg caggaggggc aggagcagga ggggcaggag gggcaggagc aggaggggca 900  
 ggaggggagc gagcaggagg ggcaggaggg gcaggagcag gaggaggggc aggagcagga 960  
 ggggcaggag caggaggtgg aggccggggt cgaggaggca gtggaggccg gggtcgagga 1020  
 ggtagtggag gccggggtcg aggaggtagt ggaggccgcc ggggtagagg acgtgaaaga 1080  
 gccagggggg gaagtcgtga aagagccagg gggagaggtc gtggacgtgg agaaaagagg 1140  
 cccaggagtc ccagtagtca gtcacatca tccgggtctc caccgcgag gccccctcca 1200  
 ggtagaaggc catttttcca cctgtaggg gaagccgatt attttgaata ccaccaagaa 1260  
 ggtggcccag atggtgagcc tgacgtgcc cgggagcga tagagcaggg ccccgagat 1320  
 gaccaggag aaggccaag cactggacce cggggtcagg gtgatggagg caggcgcaaa 1380

ES 2 587 395 T3

aaaggagggt ggtttgaaa gcatcgtggt caaggagggt ccaacccgaa atttgagaac 1440  
 attgcagaag gtttaagagc tctcctggct aggagtcaag tagaaaggac taccgacgaa 1500  
 ggaacttggg tcgccggtgt gttcgtatat ggaggtagta agacctcct ttacaaccta 1560  
 aggcgaggaa ctgcccttgc tattccacaa tgtcgtctta caccattgag tcgtctcccc 1620  
 tttggaatgg cccctggacc cggcccacaa cctggcccgc taaggagtc cattgtctgt 1680  
 tatttcatgg tctttttaca aactcatata tttgctgagg ttttgaagga tgcgattaag 1740  
 gaccttgta tgacaaagcc cgctcctacc tgcaatatca gggtgactgt gtgcagcttt 1800  
 gacgatggag tagatttggc tccctggttt ccacctatgg tggaaggggc tgccgaggag 1860  
 ggtgatgacg gagatgacgg agatgaagga ggtgatggag atgaggggtga ggaagggcag 1920  
 gagtga 1926

<210> 3

<211> 392

<212> PRT

5 <213> Herpesvirus humano 4

<400> 3

Met Ser Asp Glu Gly Pro Gly Thr Gly Pro Gly Asn Gly Leu Gly Glu  
 1 5 10 15

Lys Gly Asp Thr Ser Gly Pro Glu Gly Ser Gly Gly Ser Gly Pro Gln  
 20 25 30

Arg Arg Gly Gly Asp Asn His Gly Arg Gly Arg Gly Arg Gly Arg Gly  
 35 40 45

Arg Gly Gly Gly Arg Pro Gly Ala Pro Gly Gly Ser Gly Ser Gly Ala  
 50 55 60

Gly Ala Gly Ala Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Gly Arg  
 65 70 75 80

Gly Arg Gly Gly Ser Gly Gly Arg Gly Arg Gly Gly Ser Gly Gly Arg  
 85 90 95

Gly Arg Gly Gly Ser Gly Gly Arg Arg Gly Arg Gly Arg Glu Arg Ala  
 100 105 110

Arg Gly Gly Ser Arg Glu Arg Ala Arg Gly Arg Gly Arg Gly Arg Gly  
 115 120 125

Glu Lys Arg Pro Arg Ser Pro Ser Ser Gln Ser Ser Ser Ser Gly Ser  
 130 135 140

Pro Pro Arg Arg Pro Pro Pro Gly Arg Arg Pro Phe Phe His Pro Val  
 145 150 155 160

Gly Glu Ala Asp Tyr Phe Glu Tyr His Gln Glu Gly Gly Pro Asp Gly  
 165 170 175

Glu Pro Asp Val Pro Pro Gly Ala Ile Glu Gln Gly Pro Ala Asp Asp  
 180 185 190

Pro Gly Glu Gly Pro Ser Thr Gly Pro Arg Gly Gln Gly Asp Gly Gly  
 195 200 205

Arg Arg Lys Lys Gly Gly Trp Phe Gly Lys His Arg Gly Gln Gly Gly  
 210 215 220

Ser Asn Pro Lys Phe Glu Asn Ile Ala Glu Gly Leu Arg Ala Leu Leu  
 225 230 235 240

Ala Arg Ser His Val Glu Arg Thr Thr Asp Glu Gly Thr Trp Val Ala  
 245 250 255

Gly Val Phe Val Tyr Gly Gly Ser Lys Thr Ser Leu Tyr Asn Leu Arg  
 260 265 270

Arg Gly Thr Ala Leu Ala Ile Pro Gln Cys Arg Leu Thr Pro Leu Ser  
 275 280 285

Arg Leu Pro Phe Gly Met Ala Pro Gly Pro Gly Pro Gln Pro Gly Pro  
 290 295 300

Leu Arg Glu Ser Ile Val Cys Tyr Phe Met Val Phe Leu Gln Thr His  
 305 310 315 320

Ile Phe Ala Glu Val Leu Lys Asp Ala Ile Lys Asp Leu Val Met Thr  
 325 330 335

Lys Pro Ala Pro Thr Cys Asn Ile Arg Val Thr Val Cys Ser Phe Asp  
 340 345 350

Asp Gly Val Asp Leu Pro Pro Trp Phe Pro Pro Met Val Glu Gly Ala  
 355 360 365

Ala Ala Glu Gly Asp Asp Gly Asp Asp Gly Asp Glu Gly Gly Asp Gly

ES 2 587 395 T3

370

375

380

Asp Glu Gly Glu Glu Gly Gln Glu  
385 390

<210> 4

<211> 1179

<212> ADN

5 <213> Herpesvirus humano 4

<400> 4

```

atgtctgacg aggggccagg tacaggacct ggaaatggcc taggagagaa gggagacaca      60
tctggaccag aaggctccgg cggcagtgga cctcaaagaa gagggggtga taaccatgga      120
cgaggacggg gaagaggacg aggacgagga ggcggaagac caggagcccc gggcggctca      180
ggatcagggg ccggcgcagg agcaggagcg ggaggggcag gagcaggagg tggaggccgg      240
ggtcgaggag gcagtggagg ccggggtcga ggaggtagtg gaggccgggg tcgaggaggt      300
agtggaggcc gccggggtag aggacgtgaa agagccaggg ggggaagtcg tgaaagagcc      360
agggggagag gtcgtggacg tggagaaaag aggcccagga gtcccagtag tcagtcatca      420
tcatccgggt ctccaccgcg caggccccct ccaggtagaa ggccattttt ccaccctgta      480
ggggaagccg attattttga ataccaccaa gaaggtggcc cagatggtga gcctgacgtg      540
ccccgggag cgatagagca gggccccgca gatgaccag gagaaggccc aagcactgga      600
ccccggggtc agggatgatg aggcaggcgc aaaaaaggag ggtggttttg aaagcatcgt      660
ggtcaaggag gttccaacct gaaatttgag aacattgcag aaggtttaag agctctcctg      720
gctaggagtc acgtagaaag gactaccgac gaaggaactt gggtcgccgg tgtgttcgta      780
tatggaggta gtaagacctc cttttacaac ctaaggcgag gaactgcctt tgctattcca      840
caatgtcgtc ttacaccatt gagtcgtctc ccttttgaa tggcccctgg acccggccca      900
caacctggcc cgctaaggga gtccattgtc tgttatttca tggctttttt acaaactcat      960
atatttgctg aggttttgaa ggatgcgatt aaggaccttg ttatgacaaa gcccgctcct     1020
acctgcaata tcagggtgac tgtgtgcagc tttgacgatg gagtagattt gcctccctgg     1080
ttccaccta tgggtgaagg ggctgccgcg gagggatgat acggagatga cggagatgaa     1140
ggaggtgatg gagatgaggg tgaggaaggg caggagtga                               1179

```

**REIVINDICACIONES**

1. Un procedimiento para proporcionar una población celular que tiene un estado de diferenciación alterado respecto a una población celular de partida y que tiene células que están esencialmente exentas de elementos genéticos vectoriales de programación, comprendiendo el procedimiento las etapas de:
  - 5 a) obtener una población de partida de células que tienen un primer estado de diferenciación,
  - b) obtener uno o más vectores de programación de la diferenciación, comprendiendo cada vector un origen de replicación y uno más módulos de expresión que codifican uno o más factores de programación de la diferenciación que, en combinación, pueden alterar el estado de diferenciación de la población celular de partida hasta un segundo estado de diferenciación, en el que uno o más de dichos módulos de expresión comprenden una secuencia nucleotídica que codifica un factor de acción en trans que se une al origen de replicación para replicar un molde extracromosómico y/o en el que las células de la población de partida expresan dicho factor de acción en trans,
  - 10 c) introducir el vector o vectores de programación de la diferenciación en células de la población de partida,
  - d) cultivar las células para efectuar la expresión de dicho uno o más factores de programación de la diferenciación de tal modo que surjan los atributos consistentes con el segundo estado de diferenciación en al menos una porción de las células en las células cultivadas y
  - 15 e) cultivar además células que tienen los atributos durante un número suficiente de generaciones para proporcionar una población celular diana que comprenda células que tienen el segundo estado de diferenciación, pero estando dichas células esencialmente exentas de elementos genéticos vectoriales de programación.
  - 20
2. El procedimiento de la reivindicación 1, que comprende además una etapa adicional seleccionada del grupo consistente en:
  - (a) seleccionar células de dichas células cultivadas en la etapa d) o e), estando dichas células esencialmente exentas de elementos genéticos vectoriales de programación de la diferenciación y
  - 25 (b) seleccionar células de dichas células cultivadas en la etapa d) o e), estando dichas células esencialmente exentas de un marcador de selección comprendido en el vector de programación de la diferenciación.
3. El procedimiento de la reivindicación 2(b), en el que el marcador de selección es timidina cinasa de herpesvirus simple, un factor de resistencia a antibiótico o una proteína fluorescente.
4. El procedimiento de una cualquiera de las reivindicaciones 1 a 3, en el que dicha alteración del estado de diferenciación se selecciona del grupo consistente en reprogramación, diferenciación y transdiferenciación.
- 30 5. El procedimiento de una cualquiera de las reivindicaciones 1 a 4, en el que dicha alteración del estado de diferenciación es reprogramación y en la que dicha célula de partida se selecciona del grupo consistente en
  - (a) una célula somática y dichos atributos se definen como una o más características de células madre embrionarias, y
  - 35 (b) un fibroblasto, un queratinocito, una célula hematopoyética, una célula mesenquimática, una célula hepática, una célula de estómago o una célula  $\beta$ .
6. El procedimiento de la reivindicación 5, que comprende además una etapa de diferenciación de la población de células diana.
7. El procedimiento de una cualquiera de las reivindicaciones 1 a 5, en el que dicha alteración del estado de diferenciación es reprogramación y en el que los factores de programación de la diferenciación se definen además como factores de reprogramación que comprenden Sox-2 y Oct-4.
- 40 8. El procedimiento de la reivindicación 7, en el que los factores de reprogramación comprenden además Nanog, Lin28, Klf4 o c-Myc.
9. El procedimiento de la reivindicación 4, en el que dicha alteración del estado de diferenciación es diferenciación y en el que la célula de partida es una célula madre embrionaria, una célula madre pluripotente inducida, una célula madre hematopoyética, una célula madre neural, una célula madre mesenquimática, una progenitora hematopoyética, una progenitora endodérmica, una progenitora pancreática o una progenitora endotelial.
- 45 10. El procedimiento de la reivindicación 4, en el que dicha alteración del estado de diferenciación es transdiferenciación y en el que el primer y segundo estados de diferenciación están diferenciados terminalmente.
- 50

11. El procedimiento de una cualquiera de las reivindicaciones 1 a 10, en el que el origen de replicación es un origen de replicación de un herpesvirus linfotrófico, un adenovirus, SV40, un papilomavirus bovino o una levadura.
12. El procedimiento de la reivindicación 11, en el que el origen de replicación es un origen de replicación de herpesvirus linfotrófico y corresponde a oriP de EBV.
- 5 13. El procedimiento de la reivindicación 11, en el que el herpesvirus linfotrófico es virus de Epstein Barr (EBV), herpesvirus de sarcoma de Kaposi (KSHV), herpesvirus de Saimiri (HS) o virus de la enfermedad de Marek.
14. El procedimiento de una cualquiera de las reivindicaciones 1 a 13, en el que el factor de acción en trans
- (a) corresponde a EBNA-1 de EBV,
- (b) es un derivado de una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, teniendo dicho derivado una capacidad reducida de activar la transcripción a partir de un molde integrado en comparación con EBNA-1 de tipo silvestre,
- 10 (c) es un derivado de una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, activando dicho derivado la transcripción al menos a un 5 % de los niveles de la correspondiente proteína de tipo silvestre a partir de un molde extracromosómico después de que se une el derivado al origen de replicación, o
- 15 (d) es un derivado de una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, teniendo dicho derivado una delección de residuos correspondiente a los residuos 65 a 89 de EBNA-1, y/o una delección de residuos correspondiente a los residuos 90 a 328 de EBNA-1.
15. El procedimiento de la reivindicación 14(b) o 14(c), en el que el derivado carece de las secuencias presentes en la proteína EBNA-1 de tipo silvestre que activan la transcripción a partir de un molde integrado.
- 20 16. El procedimiento de la reivindicación 14(b) o 14(c), en el que el derivado tiene una delección de residuos correspondiente a los residuos 65 a 89 de EBNA-1 y/o una delección de residuos correspondiente a los residuos 90 a 328 de EBNA-1.
17. El procedimiento de la reivindicación 14(b) o 14(c), en el que el derivado codifica una proteína con al menos un 80 % de identidad de secuencia aminoacídica con los residuos 1 a 40 y los residuos 328 a 641 de EBNA-1.
- 25 18. El procedimiento de una cualquiera de las reivindicaciones 1 a 17, en el que los módulos de expresión se ligan operativamente con un elemento regulador transcripcional.
19. El procedimiento de una cualquiera de las reivindicaciones 1 a 18, en el que se expresan eficazmente múltiples genes usando un solo promotor/potenciador para transcribir un solo mensaje.
- 30 20. El procedimiento de la reivindicación 19, en el que se transcriben conjuntamente múltiples marcos abiertos de lectura, separado cada uno por un IRES, creando mensajes policistrónicos.
21. Un vector de programación de la diferenciación que comprende un origen de replicación y uno o más módulos de expresión que codifican un factor de acción en trans que se une al origen de replicación para replicar el vector extracromosómicamente, y uno o más factores de programación de la diferenciación.
- 35 22. El vector de programación de la diferenciación de la reivindicación 21, en el que el factor de acción en trans es un derivado de una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, activando dicho derivado la transcripción al menos a un 5 % de la correspondiente proteína de tipo silvestre a partir de un molde extracromosómico después de unirse al origen de replicación, y tiene una capacidad reducida de activar la transcripción a partir de un molde integrado en comparación con EBNA-1 de tipo silvestre.
- 40 23. El vector de programación de la diferenciación de la reivindicación 22, en el que el derivado comprende una primera secuencia nucleotídica que codifica los residuos 1 a 40 de la correspondiente EBNA-1 de tipo silvestre y una segunda secuencia nucleotídica que codifica los residuos 328 a 641 de la correspondiente EBNA-1 de tipo silvestre.
24. El vector de programación de la diferenciación de una cualquiera de las reivindicaciones 21 a 23, en el que los factores de programación de la diferenciación se seleccionan del grupo consistente en Sox-2, Sox-7, Sox-17, Oct-4, Nanog, Lin-28, c-Myc, Klf4, Esrrb, EBF1, C/EBP $\alpha$ , C/EBP $\beta$ , Ngn3, Pdx y Mafa.
- 45 25. El vector de programación de la diferenciación de una cualquiera de las reivindicaciones 21 a 23, definido además como un vector de reprogramación que comprende un miembro de la familia Sox y un miembro de la familia Oct.
26. El vector de programación de la diferenciación de la reivindicación 25, en el que los factores de programación de la diferenciación comprenden además uno o más seleccionados del grupo consistente en Nanog, Lin-28, Klf4 y c-Myc.
- 50

27. El vector de programación de la diferenciación de la reivindicación 21, en el que el vector de programación de la diferenciación carece de la capacidad de integrarse en un genoma de célula hospedadora.
28. El vector de programación de la diferenciación de una cualquiera de las reivindicaciones 21 a 27, en el que el origen de replicación es un origen de replicación de un herpesvirus linfotrófico, un adenovirus, SV40, un papilomavirus bovino o una levadura.
29. El vector de programación de la diferenciación de la reivindicación 28, en el que el origen de replicación es un origen de replicación de un herpesvirus linfotrófico y corresponde a oriP de EBV.
30. El vector de programación de la diferenciación de la reivindicación 28, en el que el herpesvirus linfotrófico es virus de Epstein Barr (EBV), herpesvirus de sarcoma de Kaposi (KSHV), herpesvirus de Saimiri (HS) o virus de la enfermedad de Marek (MDV).
31. El vector de programación de la diferenciación de una cualquiera de las reivindicaciones 21 a 30, en el que el factor de acción en trans
- corresponde a EBNA-1 de EBV,
  - es un derivado de una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, teniendo dicho derivado una capacidad reducida de activar la transcripción a partir de un molde integrado en comparación con EBNA-1 de tipo silvestre,
  - es un derivado de una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, activando dicho derivado la transcripción al menos a un 5 % de los niveles de la correspondiente proteína de tipo silvestre a partir de un molde extracromosómico después de que el derivado se una al origen de replicación, o
  - es un derivado de una proteína de tipo silvestre correspondiente a EBNA-1 de EBV, teniendo dicho derivado una delección de residuos correspondiente a los residuos 65 a 89 de EBNA-1, y/o una delección de residuos correspondiente a los residuos 90 a 328 de EBNA-1.
32. El vector de programación de la diferenciación de la reivindicación 31(b), 31(c) o 22, en el que el derivado carece de las secuencias presentes en la proteína EBNA-1 de tipo silvestre que activan la transcripción a partir de un molde integrado.
33. El vector de programación de la diferenciación de la reivindicación 31(b), 31(c) o 22, en el que el derivado tiene una delección de residuos correspondiente a los residuos 65 a 89 de EBNA-1 y/o una delección de residuos correspondiente a los residuos 90 a 328 de EBNA-1.
34. El vector de programación de la diferenciación de la reivindicación 31(b), 31(c) o 22, en el que el derivado codifica una proteína con al menos un 80 % de identidad de secuencia aminoacídica con los residuos 1 a 40 y los residuos 328 a 641 de EBNA-1.
35. El vector de programación de la diferenciación de una cualquiera de las reivindicaciones 21 a 34, en el que los módulos de expresión se ligan operativamente con un elemento regulador transcripcional.
36. El vector de programación de la diferenciación de una cualquiera de las reivindicaciones 21 a 35, en el que se expresan eficazmente múltiples genes usando un solo promotor/potenciador para transcribir un solo mensaje.
37. El vector de programación de la diferenciación de la reivindicación 36, en el que se transcriben conjuntamente múltiples marcos abiertos de lectura, separado cada uno por un IRES, creando mensajes policistrónicos.



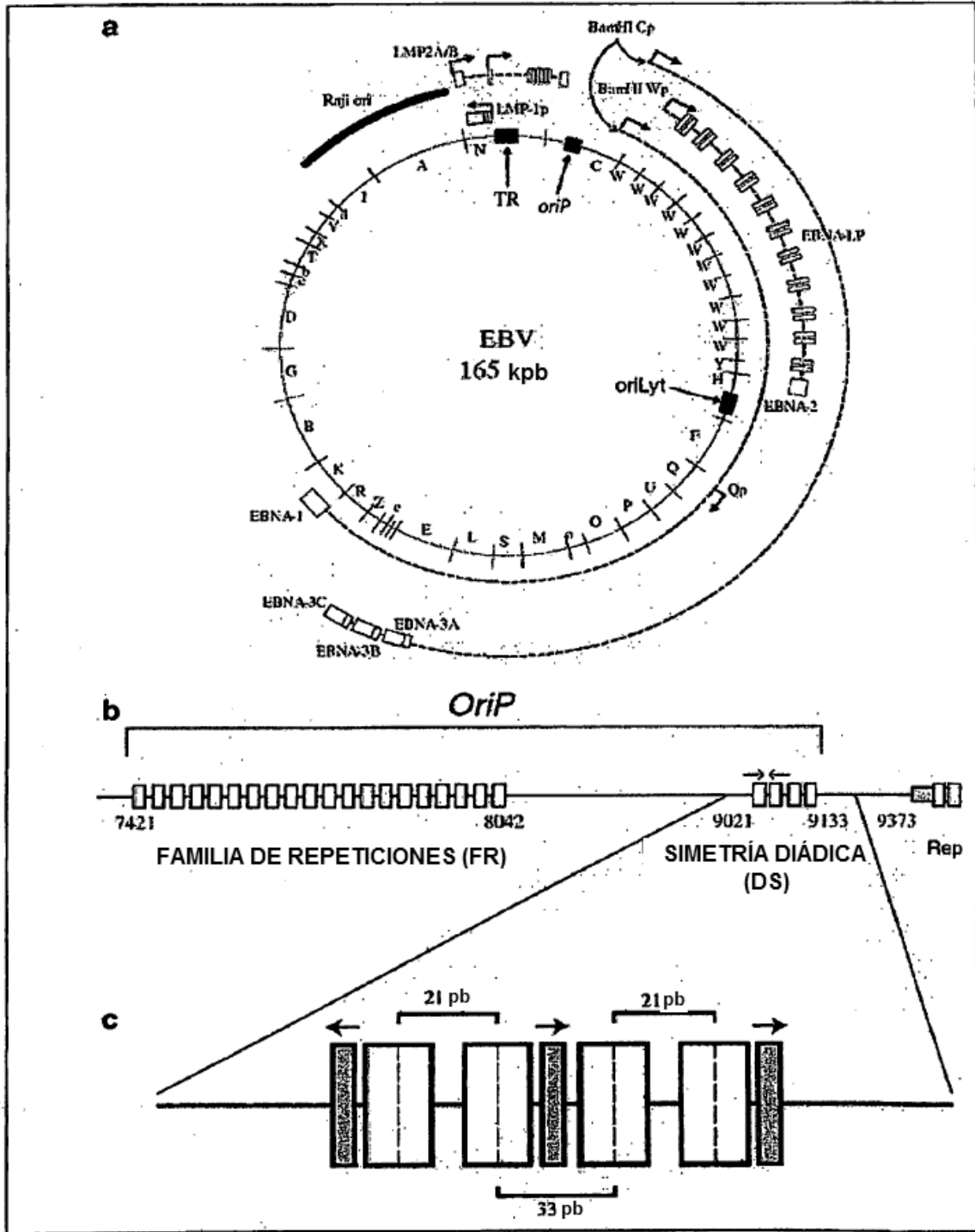


FIG. 1

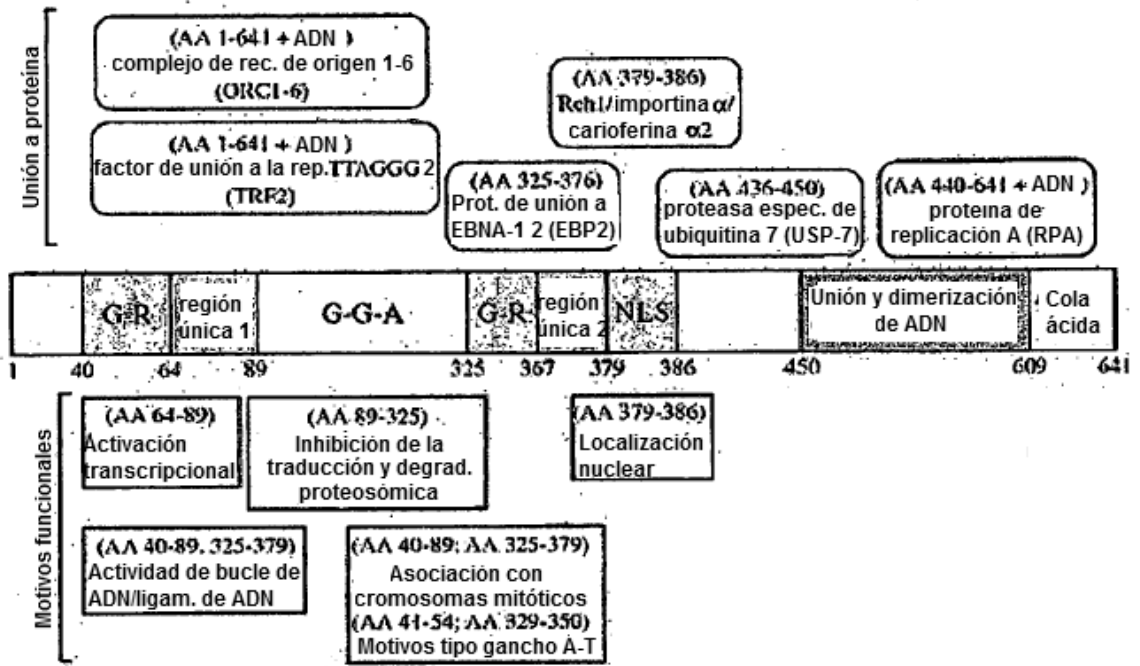


FIG. 2

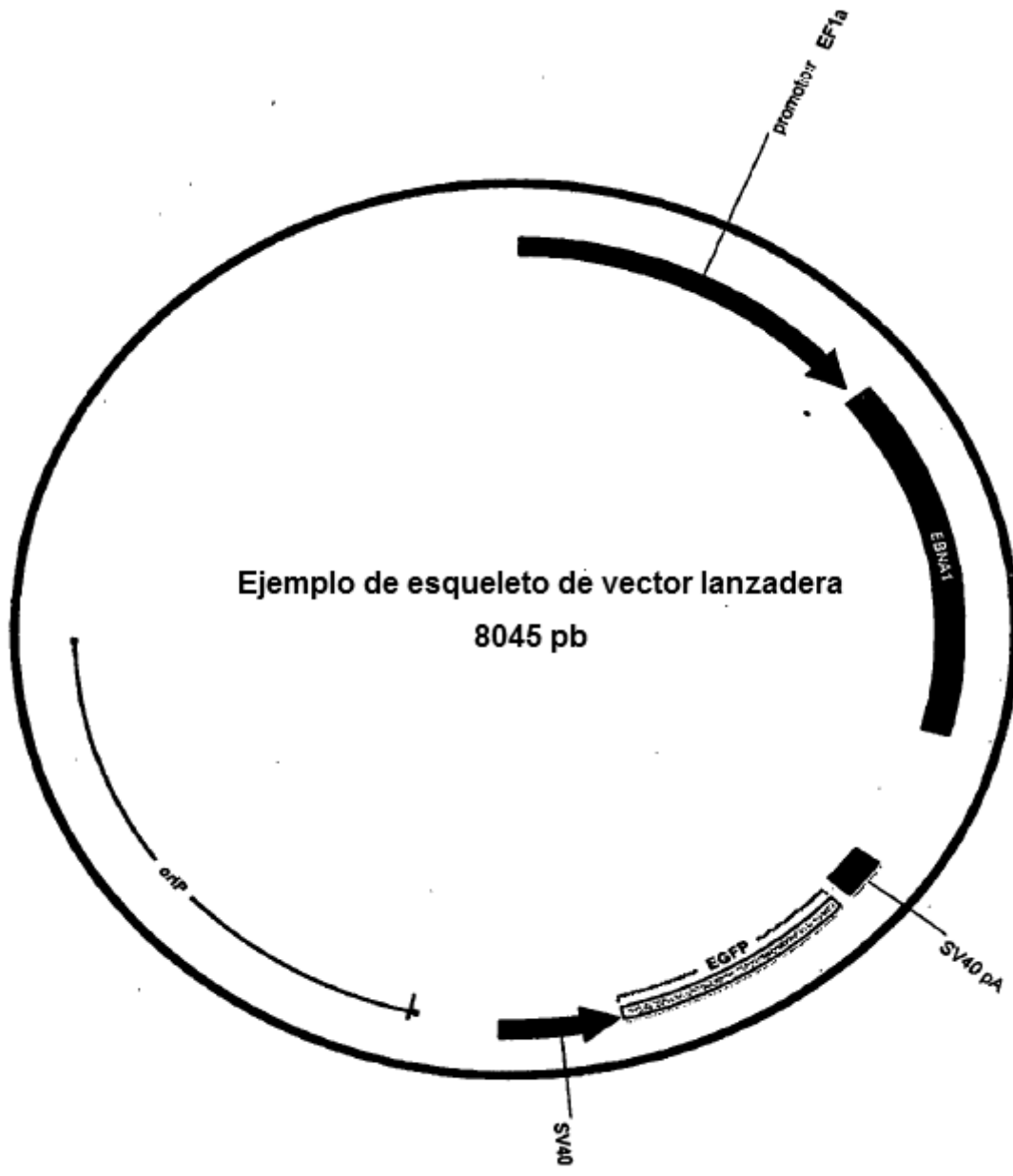


FIG. 3

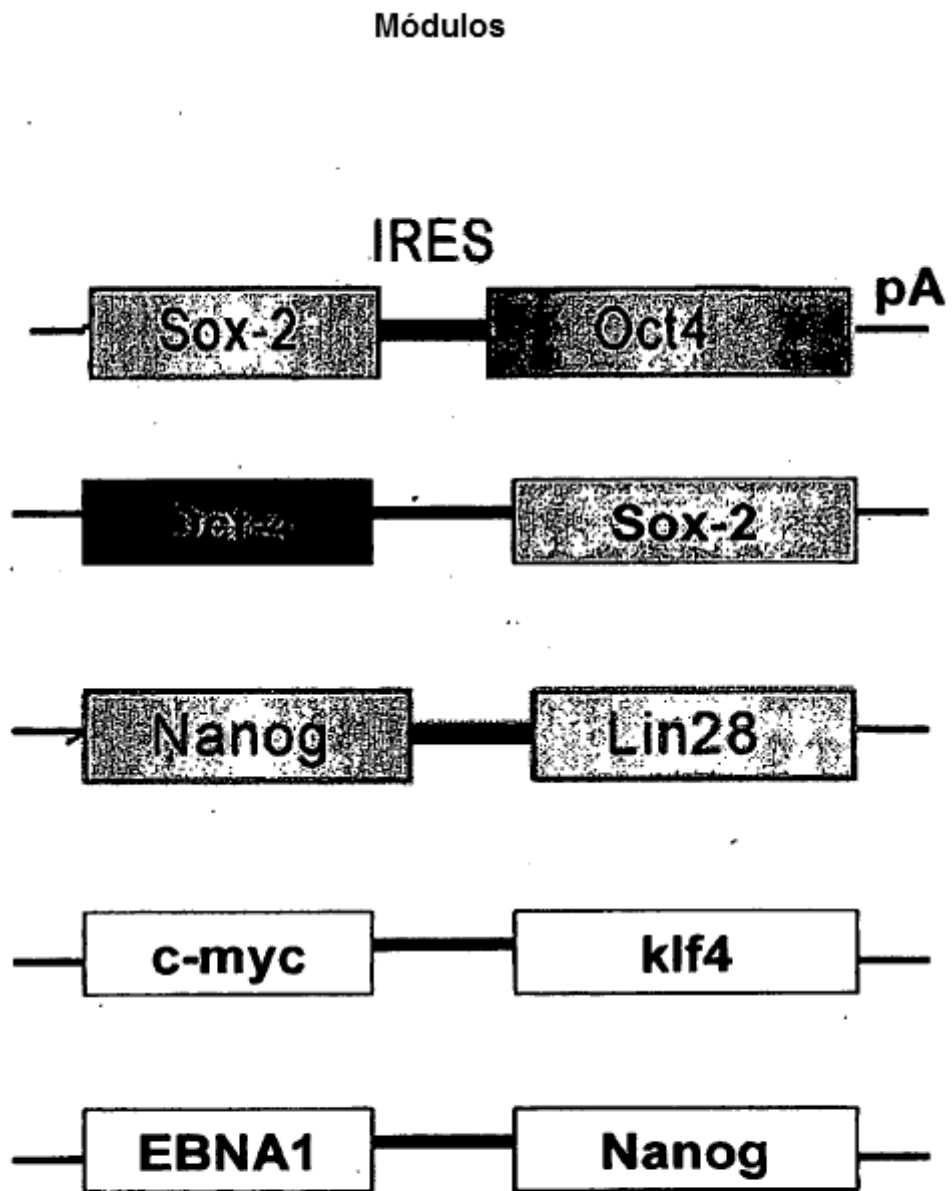


FIG. 4

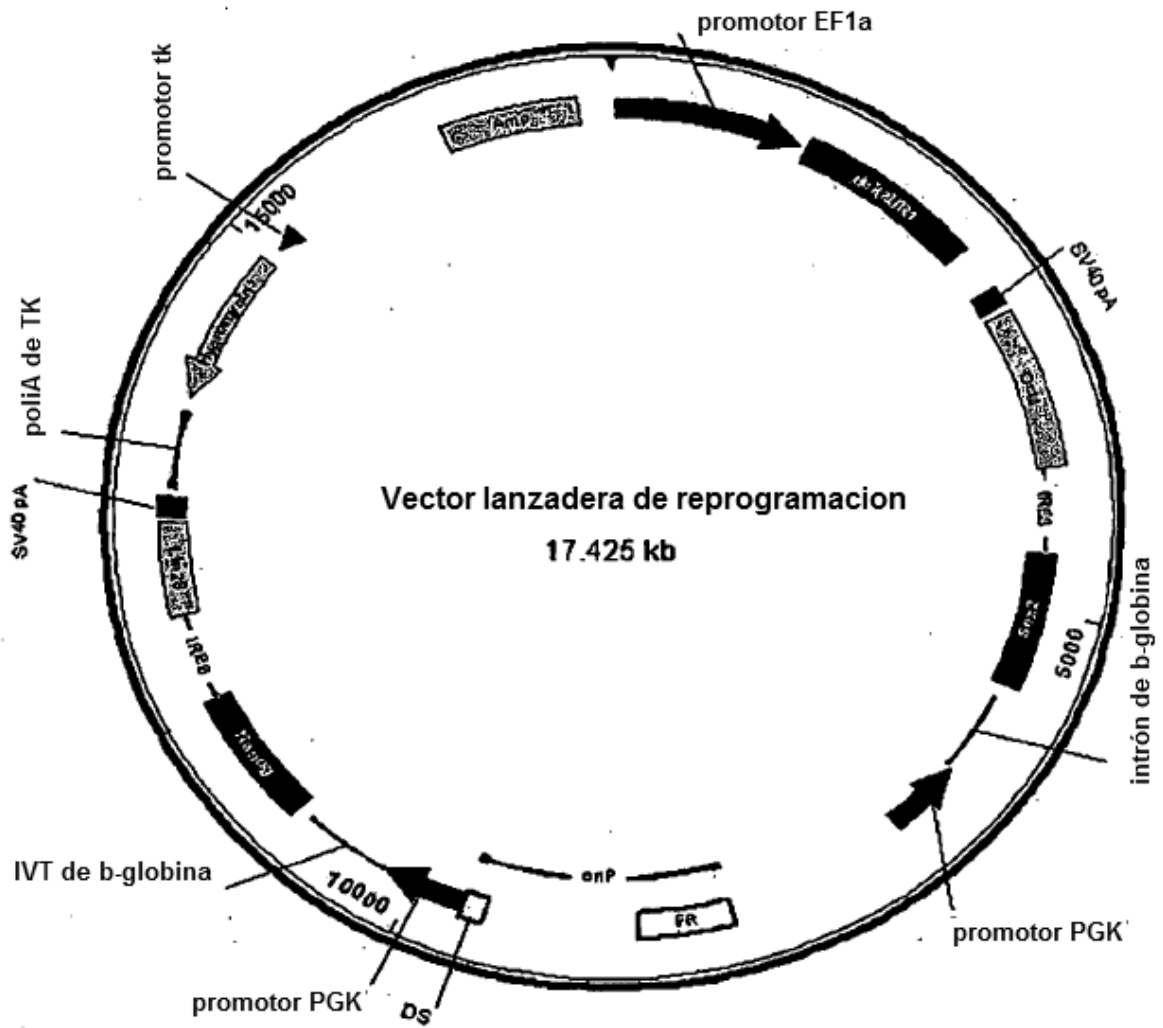


FIG. 5