

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 741 387**

51 Int. Cl.:

**C12N 5/074** (2006.01)

**C12N 5/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.10.2015 E 15787382 (9)**

97 Fecha y número de publicación de la concesión europea: **26.06.2019 EP 3207124**

54 Título: **Métodos y composiciones para generar o mantener células pluripotentes**

30 Prioridad:

**15.10.2014 US 201462064384 P**

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

**10.02.2020**

73 Titular/es:

**REGENERON PHARMACEUTICALS, INC.  
(100.0%)  
777 Old Saw Mill River Road  
Tarrytown, NY 10591, US**

72 Inventor/es:

**KUNO, JUNKO;  
AUERBACH, WOJTEK y  
VALENZUELA, DAVID M.**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

**ES 2 741 387 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**

Métodos y composiciones para generar o mantener células pluripotentes

5 La copia oficial del listado de secuencias se presenta electrónicamente mediante EFS-Web como un listado de secuencias con formato ASCII con un nombre de archivo 469587SEQLIST.TXT, creado el 14 de octubre, 2015, y que tiene un tamaño de archivo de 757 bites, y se presentó de manera simultánea con las especificaciones. El listado de secuencias que se incluye en este documento con formato ASCII forma parte de las especificaciones.

10 Antecedentes

15 Las células madre pluripotentes inducidas humanas (iPS) pueden mostrar un estado de pluripotencia vírgen o cebado (Nichols y Smith, Cell Stem Cell (2009) Vol. 4(6), pp. 487-492). Las células iPS humanas cebadas expresan características similares a aquellas de células tipo epiblastos posteriores a la implantación, y se destinan a especificación de linaje y diferenciación. En cambio, las células iPS vírgenes humanas expresan características similares a aquellas de células madre embrionarias (ES) de la masa celular interna de un embrión previo a la implantación. En ciertos aspectos, las células iPS vírgenes resultan más pluripotentes que las células cebadas, ya que no se destinan a especificación de linaje. Diversas condiciones de cultivo pueden usarse para mantener iPS humanas en un estado virgen o en un estado cebado. Stem Cells and Development, vol 21, no. 8, páginas 1287-1298, 2012; D’Aiuto et al, Organogenesis, vol 10, no. 4, página 375, 2014 y WO2011044684 divulgan un método de cultivo de células madre pluripotentes inducidas humanas que comprende el uso de medio “si-LIF”.

25 Sumario

La invención es según se define en las reivindicaciones.

30 En un primer aspecto de la invención, se proporciona un método para fabricar una población de células madre pluripotentes inducidas humanas (hiPSCs), comprendiendo el método el cultivo in vitro de una población de células no pluripotentes, que se transforman para expresar un estado pluripotente, en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende:

- (a) un polipéptido de factor de inhibición de leucemia (LIF);
- 35 (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y
- (c) un inhibidor de MEK;

40 en el que el medio de baja osmolaridad tiene una osmolaridad de 200 mOsm/kg a 250 mOsm/kg.

45 En un segundo aspecto de la invención, se proporciona un método para mantener una población de células madre pluripotentes inducidas humanas (hiPSCs) en un cultivo in vitro, comprendiendo el método el cultivo de la población de hiPSCs en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende:

- (a) un polipéptido de factor de inhibición de leucemia (LIF);
- (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y
- 50 (c) un inhibidor de MEK;

en el que el medio de baja osmolaridad tiene una osmolaridad de 200 mOsm/kg a 250 mOsm/kg.

55 En un tercer aspecto de la invención, se proporciona un método para modificar un locus genómico diana en una célula madre pluripotente inducida humana (hiPSC), que comprende:

- (a) introducir en hiPSCs un vector director que comprende un inserto de ácido nucleico de inserción flanqueado por brazos de homología 5’ y 3’ correspondientes a sitios diana 5’ y 3’ en el locus genómico diana, en el que, de manera opcional, el vector director es un vector director grande (LTVEC), en el que la suma total de los brazos de homología 60 5’ y 3’ es de al menos 10 kb; y

- (b) identificar una hiPSC genéticamente modificada que comprende en su genoma el inserto de ácido nucleico que se integra en el locus genómico diana;

65 en el que las hiPSCs se cultivan en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende:

(i) un polipéptido de factor de inhibición de leucemia (LIF);

(ii) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y

(iii) un inhibidor de MEK;

en el que el medio de baja osmolaridad tiene una osmolaridad de 200 mOsm/kg a 250 mOsm/kg.

Se proporcionan métodos para fabricar una población de células madre pluripotentes inducidas humanas (hiPSCs). Tales métodos comprenden el cultivo *in vitro* de una población de células no pluripotentes, que se transforman para expresar un estado pluripotente, en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio tiene una osmolaridad de aproximadamente 175 mOsm/kg a aproximadamente 280 mOsm/kg. Tales medios pueden comprender además el cultivo *in vitro* de una población de células no pluripotentes, que se transforman para expresar un estado pluripotente, en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio base tiene una osmolaridad de aproximadamente 180 mOsm/kg a aproximadamente 250 mOsm/kg.

Se proporcionan además métodos para mantener una población de hiPSCs en un cultivo *in vitro*, comprendiendo el medio el cultivo de la población de hiPSCs en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio tiene una osmolaridad de aproximadamente 175 mOsm/kg a aproximadamente 280 mOsm/kg. Tales métodos pueden comprender además el cultivo de la población de hiPSCs en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio base tiene una osmolaridad de aproximadamente 180 mOsm/kg a aproximadamente 250 mOsm/kg.

En algunos métodos, las hiPSCs comprenden hiPSCs vírgenes o de apariencia virgen. En algunos métodos, las hiPSCs comprenden hiPSCs de tipo vírgenes.

En algunos métodos, el método enriquece una población de hiPSCs vírgenes o de apariencia virgen. En algunos métodos, el método enriquece una población de hiPSCs de tipo virgen.

En algunos métodos, las células transformadas expresan genes de reprogramación que comprenden Oct4, Sox2, Klf4, Myc, o cualquier combinación de estos. En algunos métodos, las células transformadas comprenden hiPSCs cebadas.

En algunos métodos, el medio base tiene una osmolaridad de aproximadamente 200 mOsm/kg. En algunos métodos, el medio base comprende NaCl a aproximadamente 3 mg/ml, bicarbonato de sodio a aproximadamente 2,2 mg/mL, y tiene una osmolaridad de aproximadamente 200 mOsm/kg.

En algunos métodos, el medio base comprende glucosa a aproximadamente 4,5 mg/mL.

En algunos métodos, el medio de baja osmolaridad tiene una osmolaridad de aproximadamente 200 mOsm/kg a aproximadamente 250 mOsm/kg. En algunos métodos, el medio de baja osmolaridad tiene una osmolaridad de aproximadamente 233 mOsm/kg.

En algunos métodos, los complementos comprenden; (a) medio F-12; (b) complemento de N2; (c) medio NEUROBASAL; (d) complemento de B-27; (e) L-glutamina; (f) 2-mercaptoetanol; o (g) cualquier combinación de (a) a (f).

En algunos métodos, el polipéptido de LIF es un polipéptido de LIF humano (hLIF). En algunos métodos el inhibidor de GSK3 comprende CHIR99021. En algunos métodos el inhibidor de MEK comprende PD0325901. En algunos métodos, el medio de baja osmolaridad comprende inhibidores que consisten esencialmente en un inhibidor de GSK3 y un inhibidor de MEK.

En algunos métodos, el medio de baja osmolaridad comprende medio base de aproximadamente 24,75% (v/v), F-12 de aproximadamente 24,75% (v/v), complemento de N2 de aproximadamente 0,5% (v/v), medio NEUROBASAL de aproximadamente 49% (v/v), complemento de B-27 de aproximadamente 1% (v/v), L-glutamina de aproximadamente 2 mM, 2-mercaptoetanol de aproximadamente 0,1 mM, hLIF de aproximadamente 100 unidades/mL, CHIR99021 de aproximadamente 3  $\mu$ M, y PD0325901 de aproximadamente 0,5  $\mu$ M.

- En algunos métodos, el medio de baja osmolaridad no comprende uno o más de los siguientes: complemento de bFGF, complemento de TGF- $\beta$ 1, inhibidor de JNK, inhibidor de p38, inhibidor de ROCK, e inhibidor de PKC. En algunos métodos, el medio de baja osmolaridad no comprende factor de crecimiento de fibroblastos básico (bFGF).
- 5 En algunos métodos, las hiPSCs o las células transformadas se cultivan en MATRIGEL™, células alimentadoras de fibroblastos de prepuccio humano de recién nacidos (NuFF), o GELTREX™.
- 10 En algunos métodos, las hiPSCs expresan uno o más marcadores de pluripotencia. En algunos métodos, los uno o más marcadores de pluripotencia comprenden NANOG, fosfatasa alcalina, o una combinación de estos. En algunos métodos, las hiPSCs tienen un cariotipo normal.
- 15 En algunos métodos, las hiPSCs muestran una morfología que se caracteriza por colonias compactas de forma de cúpula.
- 20 En algunos métodos, las hiPSCs pueden disociarse de manera enzimática en una suspensión de célula única y subcultivarse. En algunos métodos, la disociación enzimática se lleva a cabo usando tripsina. En algunos métodos, la disociación enzimática puede llevarse a cabo en la ausencia de un inhibidor de proteína quinasa asociada con Rho (ROCK). En algunos métodos, las hiPSCs subcultivadas continúan expresando los uno o más marcadores de pluripotencia. En algunos métodos, las hiPSCs subcultivadas mantienen un estado virgen o de apariencia virgen y muestran una morfología que se caracteriza por colonias compactas de forma de cúpula. En algunos métodos, las hiPSCs subcultivadas mantienen un cariotipo normal.
- 25 En algunos métodos, las hiPSCs pueden diferenciarse en células de una cualquiera de las capas germinales de endodermo, ectodermo, o mesodermo.
- 30 En algunos métodos, las hiPSCs tienen un tiempo de duplicación entre aproximadamente 16 horas y aproximadamente 24 horas.
- 35 En algunos métodos, las células transformadas se cultivan primero en un medio de alta osmolaridad antes del cultivo en el medio de baja osmolaridad, en el que el medio de alta osmolaridad comprende bFGF. De manera opcional, el medio de alta osmolaridad tiene una osmolaridad de al menos aproximadamente 290 mOsm/kg.
- 40 En algunos métodos, las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que expresan características de un estado virgen o de apariencia virgen. En algunos métodos, las células transformadas se cultivan primero en el medio de alta osmolaridad durante un período de aproximadamente dos meses. En algunos métodos, las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que muestran una morfología que se caracteriza por grupos de células tridimensionales.
- 45 Se proporcionan además hiPSCs que se fabrican mediante uno cualquiera de los métodos anteriores.
- 50 Se proporcionan además métodos para modificar un locus genómico diana en una hiPSC, que comprende: (a) introducir en la hiPSC un vector director que comprende un inserto de ácido nucleico flanqueado por brazos de homología 5' y 3' correspondientes a sitios diana 5' y 3' en el locus genómico diana; y (b) identificar una hiPSC genéticamente modificada que comprende en su genoma el inserto de ácido nucleico que se integra en el locus genómico diana; en el que la hiPSC se cultiva en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio tiene una osmolaridad de aproximadamente 175 mOsm/kg a aproximadamente 280 mOsm/kg. Tales métodos pueden comprender además: (a) introducir en la hiPSC un vector director que comprende un inserto de ácido nucleico flanqueado por brazos de homología 5' y 3' correspondientes a sitios diana 5' y 3' en el locus genómico diana; y (b) identificar una hiPSC genéticamente modificada que comprende en su genoma el inserto de ácido nucleico que se integra en el locus genómico diana; en el que la hiPSC se cultiva en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio base tiene una osmolaridad de aproximadamente 180 mOsm/kg a aproximadamente 250 mOsm/kg. En algunos métodos, el vector director es un vector de transferencia grande (LTVEC), en el que la suma total de los brazos de homología 5' y 3' es al menos de 10 kb. En algunos métodos, la introducción de la etapa (a) comprende además la introducción de un agente de nucleasa que promueve la recombinación homóloga entre el vector director y el locus genómico diana en la hiPSC. En algunos métodos, la modificación genética diana comprende: (a) eliminación de una secuencia de ácido nucleico humano endógeno; (b) inserción de una secuencia de ácido nucleico exógeno, o (c) reemplazo de la secuencia de ácido nucleico humano endógeno con la secuencia de ácido nucleico exógeno. En algunos métodos, la secuencia de ácido nucleico exógeno comprende uno o más de las siguientes: (a) una secuencia de ácido nucleico que resulta homóloga u ortóloga con respecto a la secuencia de ácido nucleico humano endógeno; (b) una secuencia de ácido nucleico quimérico; (c) un alelo condicional flanqueado por secuencias diana de recombinasa específica de sitio; y (d) un gen reportero que se une de manera manejable a un promotor activo en la hiPSC.
- 65

Tales métodos para modificar un locus genómico diana en una hiPSC pueden comprender además: (a) introducir en la hiPSC uno o más agentes de nucleasa que inducen uno o más mellas o roturas bicatenarias en un sitio de reconocimiento en el locus genómico diana; y (b) identificar al menos una célula que comprende en su genoma una modificación del locus genómico diana; en el que la hiPSC se cultiva en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (i) un polipéptido de factor de inhibición de leucemia (LIF); (ii) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (iii) un inhibidor de MEK; en el que el medio tiene una osmolaridad de aproximadamente 175 mOsm/kg a aproximadamente 280 mOsm/kg. Tales métodos pueden comprender además: (a) introducir en la hiPSC uno o más agentes de nucleasa que inducen uno o más mellas o roturas bicatenarias en un sitio de reconocimiento en el locus genómico diana; y (b) identificar al menos una célula que comprende en su genoma una modificación del locus genómico diana; en el que la hiPSC se cultiva en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (i) un polipéptido de factor de inhibición de leucemia (LIF); (ii) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (iii) un inhibidor de MEK; en el que el medio base tiene una osmolaridad de aproximadamente 180 mOsm/kg a aproximadamente 250 mOsm/kg.

En cualquiera de tales métodos para modificar un locus genómico diana en una hiPSC, las hiPSCs pueden disociarse de manera enzimática en una suspensión de célula única y subcultivarse antes de la etapa (a). De manera opcional, la disociación enzimática se lleva a cabo usando tripsina. De manera opcional, la disociación enzimática se lleva a cabo en la ausencia de un inhibidor de ROCK. En algunos métodos, las hiPSCs subcultivadas continúan expresando uno o más marcadores de pluripotencia. En algunos métodos, las hiPSCs subcultivadas mantienen un estado virgen o de apariencia virgen y muestran una morfología que se caracteriza por colonias compactas de forma de cúpula. En algunos métodos, las hiPSCs subcultivadas mantienen un cariotipo normal.

En algunos métodos, el agente de nucleasa comprende una nucleasa de dedos de zinc (ZFN). En algunos métodos, el agente de nucleasa comprende una nucleasa efectora de tipo activador de transcripción (TALEN). En algunos métodos, el agente de nucleasa comprende una proteína (Cas) asociada a repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR), y un ARN guía (ARNg) que comprende un ARN CRISPR (crARN) que reconoce una secuencia genómica diana y un ARN CRISPR de transactivación (tracrARN). De manera opcional, la proteína Cas es Cas9.

En algunos métodos, la modificación genética dirigida es bialélica.

En algunos métodos las hiPSCs comprenden hiPSCs vírgenes o de apariencia virgen. En algunos métodos, las hiPSCs comprenden hiPSCs de tipo vírgenes. En algunos métodos, las hiPSCs expresan uno o más marcadores de pluripotencia. De manera opcional, los marcadores de pluripotencia comprenden NANOG, fosfatasa alcalina, o una combinación de estos. En algunos métodos, las hiPSCs muestran una morfología que se caracteriza mediante colonias de forma de cúpula. En algunos métodos, las hiPSCs pueden diferenciarse en células de una cualquiera de las capas germinales de endodermo, ectodermo, o mesodermo. En algunos métodos, las hiPSCs tienen un tiempo de duplicación entre de aproximadamente 16 horas y aproximadamente 24 horas. En algunos métodos, las hiPSCs tienen un cariotipo normal.

En algunos métodos, las hiPSCs derivan de células no pluripotentes que se transforman para expresar un estado pluripotente. De manera opcional, las células transformadas expresan genes de reprogramación que comprenden Oct4, Sox2, Klf4, Myc, o cualquier combinación de estos. De manera opcional, las células transformadas comprenden hiPSCs cebadas. En algunos métodos, las células transformadas se cultivan primero en un medio de alta osmolaridad antes del cultivo en el medio de baja osmolaridad, en el que el medio de alta osmolaridad comprende bFGF. De manera opcional, el medio de alta osmolaridad tiene una osmolaridad de al menos 290 mOsm/kg. En algunos métodos, las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que expresan características de un estado virgen o de apariencia virgen. En algunos métodos, las células transformadas se cultivan primero en el medio de alta osmolaridad durante un período de aproximadamente dos meses. En algunos métodos, las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que muestran una morfología que se caracteriza por grupos de células tridimensionales.

En algunos métodos, el medio base tiene una osmolaridad de aproximadamente 200 mOsm/kg. En algunos métodos, el medio base comprende NaCl a aproximadamente 3 mg/ml, bicarbonato de sodio a aproximadamente 2,2 mg/mL, y tiene una osmolaridad de aproximadamente 200 mOsm/kg.

En algunos métodos, el medio base comprende glucosa a aproximadamente 4,5 mg/mL.

En algunos métodos, el medio de baja osmolaridad tiene una osmolaridad de aproximadamente 200 mOsm/kg a aproximadamente 250 mOsm/kg. En algunos métodos, el medio de baja osmolaridad tiene una osmolaridad de aproximadamente 233 mOsm/kg.

En algunos métodos, los complementos comprenden; (i) medio F-12; (ii) complemento de N2; (iii) medio NEUROBASAL; (iv) complemento de B-27; (v) L-glutamina; (vi) 2-mercaptoetanol; o (vii) cualquier combinación de (i)

a (vi). En algunos métodos, el polipéptido de LIF es un polipéptido de LIF humano (hLIF). En algunos métodos, el inhibidor de GSK3 comprende CHIR99021. En algunos métodos, el inhibidor de MEK comprende PD0325901.

5 En algunos métodos, el medio de baja osmolaridad comprende inhibidores que consisten esencialmente en un inhibidor de glucógeno sintasa quinasa 3 (GSK3) y un inhibidor de MEK.

10 En algunos métodos, el medio de baja osmolaridad comprende un medio base de aproximadamente 24,75% (v/v), F-12 de aproximadamente 24,75% (v/v), complemento de N2 de aproximadamente 0,5% (v/v), medio NEUROBASAL de aproximadamente 49% (v/v), complemento de B-27 de aproximadamente 1% (v/v), L-glutamina de aproximadamente 2 mM, 2-mercaptoetanol de aproximadamente 0,1 mM, hLIF de aproximadamente 100 unidades/mL, CHIR99021 de aproximadamente 3  $\mu$ M, y PD0325901 de aproximadamente 0,5  $\mu$ M.

15 En algunos métodos, el medio de baja osmolaridad no comprende uno o más de los siguientes: complemento de bFGF, complemento de TGF- $\beta$ 1, inhibidor de JNK, inhibidor de p38, inhibidor de ROCK; e inhibidor de PKC. En algunos métodos, el medio de baja osmolaridad no comprende complemento de bFGF.

En algunos métodos, las hiPSCs se cultivan en MATRIGEL, células alimentadoras de NuFF, o GELTRES.

20 Se proporcionan además hiPSCs modificadas que se fabrican mediante uno cualquiera de los métodos anteriores.

25 Se proporcionan además cultivos *in vitro* que comprenden: (a) una población de hiPSCs; y (b) un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (i) un polipéptido de factor de inhibición de leucemia (LIF); (ii) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (iii) un inhibidor de MEK; en el que el medio tiene una osmolaridad de aproximadamente 175 mOsm/kg a aproximadamente 280 mOsm/kg. Tales cultivos *in vitro* pueden comprender además (a) una población de hiPSCs; y (b) un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (i) un polipéptido de factor de inhibición de leucemia (LIF); (ii) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (iii) un inhibidor de MEK; en el que el medio base tiene una osmolaridad de aproximadamente 180 mOsm/kg a aproximadamente 250 mOsm/kg.

35 Se proporcionan además poblaciones de hiPSCs que se fabrican o mantienen en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio tiene una osmolaridad de aproximadamente 175 mOsm/kg a aproximadamente 280 mOsm/kg. Tales poblaciones de hiPSCs pueden fabricarse o mantenerse además en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende: (a) un polipéptido de factor de inhibición de leucemia (LIF); (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y (c) un inhibidor de MEK; en el que el medio base tiene una osmolaridad de aproximadamente 180 mOsm/kg a aproximadamente 250 mOsm/kg.

40 En algunas poblaciones o cultivos *in vitro*, las hiPSCs comprenden hiPSCs vírgenes o de apariencia virgen. En algunas poblaciones o cultivos *in vitro*, las hiPSCs comprenden hiPSCs de tipo virgen.

45 En algunas poblaciones o cultivos *in vitro*, las hiPSCs derivan a partir de células no pluripotentes que se transforman para expresar un estado pluripotente. En algunas poblaciones o cultivos *in vitro*, las células transformadas expresan genes de reprogramación que comprenden Oct4, Sox2, Klf4, Myc, o cualquier combinación de estos. En algunas poblaciones o cultivos *in vitro*, las células transformadas comprenden hiPSCs cebadas.

50 En algunas poblaciones o cultivos *in vitro*, el medio base tiene una osmolaridad de aproximadamente 200 mOsm/kg. En algunas poblaciones o cultivos *in vitro*, el medio base comprende NaCl a aproximadamente 3 mg/ml, bicarbonato de sodio a aproximadamente 2,2 mg/mL, y tiene una osmolaridad de aproximadamente 200 mOsm/kg.

55 En algunas poblaciones o cultivos *in vitro*, el medio base comprende glucosa a aproximadamente 4,5 mg/mL.

60 En algunas poblaciones o cultivos *in vitro*, el medio de baja osmolaridad que comprende el medio base y complementos tiene una osmolaridad de aproximadamente 200 mOsm/kg a aproximadamente 250 mOsm/kg. En algunas poblaciones o cultivos *in vitro*, el medio de baja osmolaridad tiene una osmolaridad de aproximadamente 233 mOsm/kg.

65 En algunas poblaciones o cultivos *in vitro*, los complementos comprenden: (a) medio F-12; (b) complemento de N2; (c) medio NEUROBASAL; (d) complemento de B-27; (e) L-glutamina; (f) 2-mercaptoetanol; o (g) cualquier combinación de (a) a (f).

En algunas poblaciones o cultivos *in vitro*, el polipéptido de LIF es un polipéptido de LIF humano (hLIF). En algunas poblaciones o cultivos *in vitro*, el inhibidor de GSK3 comprende CHIR99021. En algunas poblaciones o cultivos *in vitro*,

el inhibidor de MEK comprende PD0325901. En algunas poblaciones o cultivos *in vitro*, el medio de baja osmolaridad comprende inhibidores que consisten esencialmente en un inhibidor de GSK3 y un inhibidor de MEK.

5 En algunas poblaciones o cultivos *in vitro*, el medio de baja osmolaridad comprende medio base de aproximadamente 24,75% (v/v), medio F-12 de aproximadamente 24,75% (v/v), complemento de N2 de aproximadamente 0,5% (v/v), medio NEUROBASAL de aproximadamente 49% (v/v), complemento de B-27 de aproximadamente 1% (v/v), L-glutamina de aproximadamente 2 mM, 2-mercaptoetanol de aproximadamente 0,1 mM, hLIF de aproximadamente 100 unidades/mL, CHIR99021 de aproximadamente 3  $\mu$ M, y PD0325901 de aproximadamente 0,5  $\mu$ M.

10 En algunas poblaciones o cultivos *in vitro*, el medio de baja osmolaridad no comprende uno o más de los siguientes: complemento de bFGF, complemento de TGF- $\beta$ 1, inhibidor de JNK, inhibidor de p38, inhibidor de ROCK, e inhibidor de PKC. En algunas poblaciones o cultivos *in vitro*, el medio de baja osmolaridad no comprende factor de crecimiento de fibroblastos básico (bFGF).

15 En algunas poblaciones o cultivos *in vitro*, las hiPSCs o las células transformadas se cultivan en MATRIGEL™, células alimentadoras de fibroblastos de prepucio humano de recién nacidos (NuFF), o GELTREX™.

20 En algunas poblaciones o cultivos *in vitro*, las hiPSCs expresan uno o más marcadores de pluripotencia. En algunas poblaciones o cultivos *in vitro*, los uno o más marcadores de pluripotencia comprenden NANOG, fosfatasa alcalina, o una combinación de estos. En algunas poblaciones o cultivos *in vitro*, las hiPSCs tienen un cariotipo normal.

25 En algunas poblaciones o cultivos *in vitro*, las hiPSCs muestran una morfología que se caracteriza mediante colonias compacta de forma de cúpula.

30 En algunas poblaciones o cultivos *in vitro*, las hiPSCs pueden disociarse de manera enzimática en una suspensión de célula única y subcultivarse. En algunas poblaciones o cultivos *in vitro*, la disociación enzimática se lleva a cabo usando tripsina. En algunas poblaciones o cultivos *in vitro*, la disociación enzimática puede llevarse a cabo en la ausencia de un inhibidor de proteína quinasa asociada con Rho (ROCK). En algunas poblaciones o cultivos *in vitro*, las hiPSCs continúan expresando los uno o más marcadores de pluripotencia. En algunas poblaciones o cultivos *in vitro*, las hiPSCs subcultivadas mantienen un estado virgen o de apariencia virgen y muestran una morfología que se caracteriza por colonias compactas de forma de cúpula. En algunas poblaciones o cultivos *in vitro*, las hiPSCs subcultivadas mantienen un cariotipo normal.

35 En algunas poblaciones o cultivos *in vitro*, las hiPSCs pueden diferenciarse en células de una cualquiera de las capas germinales de endodermo, ectodermo, o mesodermo.

40 En algunas poblaciones o cultivos *in vitro*, las hiPSCs tienen un tiempo de duplicación entre de aproximadamente 16 horas y aproximadamente 24 horas.

45 En algunas poblaciones o cultivos *in vitro*, las células transformadas se cultivan primero en un medio de alta osmolaridad antes del cultivo en el medio de baja osmolaridad, en el que el medio de alta osmolaridad comprende bFGF. De manera opcional, el medio de alta osmolaridad tiene una osmolaridad de al menos aproximadamente 290 mOsm/kg.

50 En algunas poblaciones o cultivos *in vitro*, las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que expresan características de un estado virgen o de apariencia virgen. En algunas poblaciones o cultivos *in vitro*, las células transformadas se cultivan primero en el medio de alta osmolaridad durante un período de aproximadamente dos meses. En algunas poblaciones o cultivos *in vitro*, las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que muestran una morfología que se caracteriza por grupos de células tridimensionales.

#### Breve descripción de las figuras

55 La Figura 1 representa un esquema para reemplazo de una porción del locus *ADAM6* humano con un ácido nucleico que comprende los loci *Adam6a* de ratón y *Adam6b* de ratón usando un LTVEC y un ARN guía en células iPS humanas. El sitio diana para el ARN guía se indica mediante la flecha.

60 La Figura 2A representa la morfología que se muestra por células iPS humanas que se cultivan durante 8 días en medio 2i.

La Figura 2B representa la morfología que se muestra por células iPS humanas que se cultivan durante 12 días en medio 2i.

65 Las Figuras 3A-3D representan la morfología de células iPS humanas que se cultivan en medio mTeSR™-hLIF o en medio VG2i de baja osmolaridad durante 6 días. Las Figuras 3A y 3B representan la morfología de células iPS humanas que se cultivan en medio mTeSR™-hLIF (Figura 3A) o medio VG2i (Figura 3B) durante 6 días. Las Figuras

3C y 3D representan la morfología de células iPS humanas que se cultivan en células alimentadoras de fibroblastos de prepucio de humano recién nacido (NuFF) en medio mTeSR™-hLIF (Figura 3C) o medio VG2i (Figura 3D) durante 6 días.

5 La Figura 4A representa células iPS humanas reprogramadas que se cultivan en medio VG2i que se ha sometido a coloración para fosfatasa alcalina. Las Figuras 4B y 4C representan células iPS humanas reprogramadas que se cultivan en medio VG2i que se han sometido a inmunocoloración para la expresión de NANOG.

10 Las Figuras 5A-5C representan disociación enzimática y subcultivo de células iPS humanas reprogramadas que se cultivan en medio VG21. La Figura 5A representa células iPS humanas reprogramadas que se cultivan en medio VG2i antes de disociación enzimática con tripsina en la ausencia de un inhibidor de ROCK. La Figura 5B representa células iPS humanas cultivadas en medio VG2i durante 1 día después del subcultivo. La Figura 5C representa células iPS humanas cultivadas en medio VG2i durante 4 días después del subcultivo.

15 Las Figuras 6A y 6B representan los cariotipos de células a partir de dos clones de células iPS humanas diferentes en el pase 10 a continuación de disociación con tripsina para crear una suspensión de célula única.

Breve descripción de las secuencias

20 SEQ ID NO: 1 establece una secuencia de ácido nucleico que comprende ARNg de ADAM6.

SEQ ID NO: 2 establece la secuencia de ácido nucleico de una secuencia diana para un complejo CRISPR/Cas.

25 Definiciones

Los términos “proteína”, “polipéptido”, y “péptido”, que se usan de manera indistinta en la presente incluyen formas poliméricas de aminoácidos de cualquier longitud, incluyendo aminoácidos codificados y no codificados y aminoácidos químicamente o bioquímicamente modificados o derivatizados. Los términos incluyen además polímeros que han sido modificados, tales como polipéptidos que tienen estructura fundamental de péptidos modificada.

30 Los términos “ácido nucleico” y “polinucleótido” que se usan de manera indistinta en la presente incluyen formas poliméricas de nucleótidos de cualquier longitud, incluyendo ribonucleótidos, desoxirribonucleótidos, o análogos o versiones modificadas de estos. Estos incluyen ADN o ARN mono, bicatenarios o de múltiples cadenas, ADN genómico, ADNc, híbridos ADN-ARN, y polímeros que comprenden bases de purina, bases de pirimidina, u otras bases de nucleótidos naturales, químicamente modificadas, bioquímicamente modificadas, no naturales o derivatizadas.

35 “Optimización de codón” incluye, de manera general, un proceso para modificar una secuencia de ácido nucleico para mejorar la expresión en células huésped en particular mediante el reemplazo de al menos un codón de la secuencia nativa con un codón que se usa con mayor frecuencia o con la mayor frecuencia en los genes de la célula huésped mientras se mantiene la secuencia de ácido nucleico nativa. Por ejemplo, un ácido nucleico que codifica una proteína Cas puede modificarse para sustituir codones que tiene una mayor frecuencia de uso en una célula humana. Las tablas de usos de codones se disponen fácilmente, por ejemplo en la "Codon Usage Database." Estas tablas pueden adaptarse en una cantidad de maneras. Véase Nakamura et al. (2000) Nucleic Acids Research 28:292. Algoritmos de ordenador para optimización de codón de una secuencia en particular de una secuencia en particular para expresión en un huésped en particular se encuentran disponibles también (véase, por ejemplo, Gene Forge).

40 “Unión de manera manejable” o “unirse de manera manejable” incluye yuxtaposición de dos o más componentes (por ejemplo, un promotor y otro elemento de secuencia) de manera tal que ambos componentes funcionan de manera normal y permiten la posibilidad de que al menos uno de los componentes pueda mediar una función que se ejerce en al menos uno de los otros componentes. Por ejemplo, un promotor puede unirse de manera operable a una secuencia de codificación si el promotor controla el nivel de transcripción de la secuencia de codificación como respuesta a la presencia o ausencia de uno o más factores de regulación transcripcional.

45 “Complementariedad” de ácidos nucleicos se refiere a que una secuencia de nucleótidos en una cadena de ácido nucleico, debido a la orientación de sus grupos de nucleobases, forma enlaces de hidrógeno con otra secuencia en una cadena de ácido nucleico opuesta. Las bases complementarias en ADN son normalmente A con T y C con G. En ARN, son normalmente C con G y U con A. La complementariedad puede ser perfecta o sustancial/suficiente. La complementariedad perfecta entre dos ácidos nucleicos se refiere a que los dos ácidos nucleicos pueden formar un dúplex en el que cada base en el dúplex se une con una base complementaria mediante apareamiento Watson y Crick. Complementariedad “sustancial” o “suficiente” se refiere a que una secuencia en una cadena no resulta complementaria de manera completa y/o perfecta con respecto a una secuencia en una cadena opuesta, pero existe una unión suficiente entre bases en las dos cadenas para formar un complejo híbrido estable en conjunto con condiciones de hibridación (por ejemplo, concentración de sal y temperatura). Tales condiciones pueden predecirse mediante el uso de las secuencias y cálculos matemáticos estándares para predecir la Tm de cadenas hibridadas, o mediante determinación empírica de Tm usando métodos de rutina. La Tm incluye la temperatura en la que una

población de complejos de hibridación que se forman entre dos ácidos nucleicos son desnaturalizados al 50%. A una temperatura por debajo de la  $T_m$ , se favorece la formación de un complejo de hibridación, mientras que a una temperatura por encima de la  $T_m$ , se favorece la fusión o separación de las cadenas en el complejo de hibridación.

5 La  $T_m$  puede estimarse para un ácido nucleico que tiene un contenido de G+C que se conoce en una solución acuosa de NaCl 1M mediante el uso de, por ejemplo,  $T_m=81,5+0,41(\% \text{ de G+C})$ , a pesar de que otros cálculos de  $T_m$  que se conocen consideran las características estructurales del ácido nucleico.

10 “Condición de hibridación” incluye el medio acumulado en el que una cadena de ácido nucleico se une a una segunda cadena de ácido nucleico mediante interacciones de cadena complementaria y enlace de hidrógeno para producir un complejo de hibridación. Tales condiciones incluyen los componentes químicos y sus concentraciones (por ejemplo, sales, agentes quelantes, formamida) de una solución acuosa u orgánica que contiene los ácidos nucleicos, y la temperatura de la mezcla. Otros factores, tales como la duración del tiempo de incubación o dimensiones de la cámara de reacción pueden contribuir al medio. Véase, por ejemplo, Sambrook et al., *Molecular Cloning, A Laboratory Manual*, 2.sup.nd ed., pp. 1.90-1.91, 9.47-9.51, 1 1.47-11.57 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

20 La hibridación requiere que los dos ácidos nucleicos contengan secuencias complementarias, a pesar de que son posibles las incompatibilidades entre bases. Las condiciones adecuadas para hibridación entre dos ácidos nucleicos dependen de la longitud de los ácidos nucleicos y el grado de complementación, variables que se conocen bien en la técnica. Cuanto mayor resulta el grado de complementación entre dos secuencias de nucleótidos, mayor resulta el valor de la temperatura de fusión ( $T_m$ ) para híbridos de ácidos nucleicos que tienen aquellas secuencias. Para hibridaciones entre ácidos nucleicos con tramos cortos de complementariedad (por ejemplo, complementariedad de hasta 35 o menos, 30 o menos, 25 o menos, 22 o menos, 20 o menos, o 18 o menos nucleótidos) la posición de incompatibilidades se vuelve importante (véase Sambrook *et al.*, supra, 11.7-11.8). Normalmente, la longitud de un ácido nucleico que puede hibridarse resulta de al menos aproximadamente 10 nucleótidos. Las longitudes mínimas ilustrativas para un ácido nucleico que puede hibridarse incluyen al menos aproximadamente 15 nucleótidos, al menos aproximadamente 20 nucleótidos, al menos aproximadamente 22 nucleótidos, al menos aproximadamente 25 nucleótidos, y al menos aproximadamente 30 nucleótidos. Además, la temperatura y concentración de sal de solución de lavado pueden ajustarse tanto como sea necesario de acuerdo con factores tales como longitud de la región de complementación y el grado de complementación.

35 La secuencia de polinucleótido no necesita ser complementaria al 100% con respecto a su ácido nucleico diana para poder hibridarse específicamente. Además, un polinucleótido puede hibridarse en hasta uno o más segmentos de manera tal que segmentos de intervención o adyacentes no se involucran en el evento de hibridación (por ejemplo, una estructura de bucle o estructura de horquilla). Un polinucleótido (por ejemplo, ARNg) puede comprender al menos el 70%, al menos el 80%, al menos el 90%, al menos el 95%, al menos el 99%, o el 100% de complementariedad de secuencia con una región diana en la secuencia de ácido nucleico diana a la que se dirige. Por ejemplo, un ARNg en el que 18 de 20 nucleótidos son complementarios con una región diana, y, por lo tanto, se hibridarían de manera específica, representaría complementariedad del 90%. En este ejemplo, los nucleótidos no complementarios restantes pueden agruparse o entremezclarse con nucleótidos complementarios y no necesitan ser contiguos uno con respecto al otro o con respecto a nucleótidos complementarios.

45 El porcentaje de complementariedad entre segmentos de secuencias de ácidos nucleicos en particular en ácidos nucleicos puede determinarse de manera rutinaria usando programas BLAST (herramientas de búsqueda de alineamientos básicos de tipo local) y programas PowerBLAST que se conocen en la técnica (Altschul et al. (1990) *J. Mol. Biol.* 215:403-410; Zhang y Madden (1997) *Genome Res.* 7:649-656) o mediante el uso de the Gap program (Wisconsin Sequence Analysis Package, versión 8 para Unix, Genetics Computer Group, University Research Park, Madison Wis.), usando configuraciones predeterminadas que usan el algoritmo de Smith y Waterman (*Adv. Appl. Math.*, 1981, 2, 482-489).

55 “Identidad de secuencia” o “identidad” en el contexto de dos polinucleótidos o secuencias de polipéptidos hace referencia a los residuos en las dos secuencias que resultan los mismos cuando se alinean para correspondencia máxima en una ventana de comparación específica. Cuando el porcentaje de identidad de secuencia se usa con referencia a proteínas, se reconoce que las posiciones de residuo que no resultan idénticas difieren frecuentemente por sustituciones de aminoácidos conservadoras, donde los residuos de aminoácidos se sustituyen por otros residuos de aminoácidos con propiedades químicas similares (por ejemplo, carga o hidrofobia) y, por lo tanto, no cambian las propiedades funcionales de la molécula. Cuando las secuencias difieren en sustituciones conservadoras, el porcentaje de identidad de secuencia puede ajustarse hacia arriba para corregir la naturaleza conservadora de la sustitución. Se dice que las secuencias que difieren por tales sustituciones conservadoras tienen “similitud de secuencia” o “similitud”. Los medios para llevar a cabo este ajuste se conocen bien por aquellos expertos en la técnica. Normalmente, esto incluye calificar una sustitución conservadora como una incompatibilidad parcial en lugar de una incompatibilidad completa, aumentando así, el porcentaje de identidad de secuencia. De este modo, por ejemplo, cuando un aminoácido idéntico recibe una calificación de 1 y una sustitución no conservadora reciba una calificación de cero, una sustitución conservadora recibe una calificación entre cero y uno. La calificación de sustituciones conservadoras se calcula, por ejemplo, según se implementa en el programa PC/GENE (Intelligenetics, Mountain View, California).

“Porcentaje de identidad de secuencia” incluye el valor que se determina al comparar dos secuencias que se alinean de manera óptima en una ventana de comparación, en el que la porción de la secuencia de polinucleótidos en la ventana de comparación puede comprender adiciones o eliminaciones (a saber, espacios) en comparación con la secuencia de referencia (que no comprende adiciones o eliminaciones) para alineación óptima de las dos secuencias. El porcentaje se calcula mediante la determinación del número de posiciones en las que la base de ácido nucleico idéntica o residuo de aminoácido tiene lugar en ambas secuencias para dar como resultado el número de posiciones correspondientes, dividir el número de posiciones correspondientes por el número total de posiciones en la ventana de comparación, y multiplicar el resultado por 100 para dar como resultado el porcentaje de identidad de secuencia.

A menos que se establezca lo contrario, los valores de identidad/similitud de secuencia incluyen el valor que se obtiene usando GAP Versión 10, usando los siguientes parámetros: % de identidad y % de similitud para una secuencia de nucleótidos usando GAP de Peso de 50 y Longitud Peso de 3, y la matriz de clasificación nwsgapdna.cmp; % de identidad y % de similitud para una secuencia de aminoácidos usando GAP de Peso de 8 y Longitud Peso de 2, y la matriz de clasificación BLOSUM62; o cualquier programa equivalente de estos. “Programa equivalente” incluye cualquier programa de comparación de secuencias que, para cualquiera de dos secuencias en cuestión, genera una alineación que tiene correspondencias de nucleótidos idénticos o residuos de aminoácidos y una identidad de secuencia de porcentaje idéntico cuando se compara con la alineación correspondiente que se generó por GAP Version 10.

Composiciones o métodos “que comprenden” o “que incluyen” uno o más elementos que se recitan pueden incluir otros elementos que no se recitan de manera específica. Por ejemplo, una composición que “comprende” o “que incluye” una proteína puede contener la proteína sola o en combinación con otros ingredientes.

La designación de un rango de valores incluye todos los enteros en o que definen el rango, y todos los subrangos que se definen por enteros en el rango.

El término “aproximadamente” se refiere a que el valor específico puede variar por algún porcentaje. En algunos ejemplos, el porcentaje puede ser del 1, 2, 3, 4, 8, o 10% del valor específico.

Las formas singulares de los artículos “un”, “una”, y “el” incluyen las referencias en plural a menos que el contexto dicte claramente lo contrario. Por ejemplo, el término “una célula” o “al menos una célula” puede incluir una pluralidad de células, incluyendo mezclas de estas.

#### Descripción detallada

A. Medio de baja osmolaridad para fabricación y mantenimiento de células madre pluripotentes inducidas humanas.

Se proporciona un medio de cultivo celular para uso en los métodos y composiciones de la divulgación. En una realización, el medio resulta adecuado para fabricar una población de células iPS humanas. En otra realización, el medio resulta adecuado para mantener células iPS humanas en cultivo. En algunas realizaciones, las células iPS humanas son vírgenes o de apariencia vírgen.

El medio que se proporciona en la presente comprende al menos un medio base, complementos, un polipéptido de factor de inhibición de leucemia (LIF); un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y un inhibidor de proteína quinasa activado por mitógeno (MEK). Un “medio base” o “medios base” incluye, por ejemplo, un medio base que se conoce en la técnica (por ejemplo, Medio de Eagle modificado por Dulbecco (DMEM)) que resulta adecuado para uso (con complementos agregados) para cultivar o mantener células pluripotentes (por ejemplo, células iPS) en cultivo. El medio base se complementa normalmente con un número de complementos que se conocen en la técnica cuando se usan para mantener células en cultivo.

El presente medio es un medio de baja osmolaridad. En un ejemplo, la osmolaridad se encuentra entre 175-280 mOsm/kg. En ejemplos adicionales, la osmolaridad del medio es de aproximadamente 180-270 mOsm/kg, aproximadamente 200-250 mOsm/kg, aproximadamente 220-240 mOsm/kg, o aproximadamente 225-235 mOsm. En una realización en particular, la osmolaridad del medio es de aproximadamente 233 mOsm/kg.

El medio base que se proporciona para la invención es un medio base de baja osmolaridad al cual se agregan complementos. El presente medio base difiere de los medios base que se usan normalmente para mantener células iPS humanas en cultivo, que incluyen Medio de Eagle modificado por Dulbecco (DMEM), en diversas formas (por ejemplo, Invitrogen DMEM, Cat. No. 1 1971 -025), y un DMEM de bajo contenido de sal que se dispone comercialmente como KO-DMEM™ (Invitrogen Cat. No. 10829-018).

El medio base que se proporciona en la presente es un medio de baja osmolaridad pero exhibe características que no se limitan a baja osmolaridad. Por ejemplo, la formulación de DMEM que se muestra en la Tabla 1 puede constituirse de manera adecuada para los propósitos de la invención mediante alteración de las concentraciones del cloruro de sodio y/o bicarbonato de sodio según se proporcionan en la presente, lo que dará como resultado una osmolaridad

diferente en comparación con el medio base DMEM estándar o medio base DMEM de bajo contenido de sal (KODMEM) que se muestra en la Tabla 1.

5

Tabla 1. Formulación de medio base DMEM.

Componente	Mg/L	mM
Glicina	30	0,4
L-Arginina•HCl	84	0,398
L-Cistina•2HCl	63	0,201
L-Glutamina	584	4
L-Histidina•HCl•H <sub>2</sub> O	42	0,2
L-Isoleucina	105	0,802
L-Leucina	105	0,802
L-Lisina•HCl	146	0,798
L-Metionina	30	0,201
L-Fenilalanina	66	0,4
L-Serina	42	0,4
L-Treonina	95	0,798
L-Triptófano	16	0,0784
Dihidrato de sal disódica de L-tirosina	104	0,398
L-Valina	94	0,803
Cloruro de colina	4	0,0286
Pantotenato de D-calcio	4	8,39 x 10 <sup>-3</sup>
Ácido fólico	4	9,07 x 10 <sup>-3</sup>
Niacinamida	4	0,0328
Piridoxina•HCl	4	0,0196
Riboflavina	0,4	1,06 x 10 <sup>-3</sup>
Tiamina•HCl	4	0,0119
i-Inositol	7,2	0,04
Cloruro de calcio (CaCl <sub>2</sub> ) (anhidro)	200	1,8
Nitrato férrico (Fe (NO <sub>3</sub> ) <sub>3</sub> •9H <sub>2</sub> O)	0,1	2,48 x 10 <sup>-4</sup>
Sulfato de magnesio (MgSO <sub>4</sub> ) (anhid.)	97,67	0,814
Cloruro de potasio (KCl)	400	5,33

D-Glucosa (Dextrosa)	4500	25
Componente	Mg/L	mM
Rojo de fenol	15	0,0399
Contenido de NaCl/NaHCO <sub>3</sub> de DMEM		
Bicarbonato de Sodio (NaHCO <sub>3</sub> )	3700	44,05
Cloruro de Sodio (NaCl)	6400	110,34
Osmolaridad	340 mOsm/kg	
Contenido de NaCl/NaHCO <sub>3</sub> de DMEM de bajo contenido de sal (KO-DMEM)		
Bicarbonato de Sodio (NaHCO <sub>3</sub> )	2200	26
Cloruro de Sodio (NaCl)	5100	87,7
Osmolaridad	275 mOsm/kg	
Contenido de NaCl/NaHCO <sub>3</sub>		
DMEM de Baja Osmolaridad (VG-DMEM)		
Bicarbonato de Sodio (NaHCO <sub>3</sub> )	2200	26
Cloruro de Sodio (NaCl)	3000	50
Osmolaridad	200 mOsm/kg	

El presente medio base puede incluir una sal de un metal alcalino y un haluro, tal como cloruro de sodio (NaCl). Concentraciones de ejemplo de NaCl en el medio base incluyen  $50 \pm 5$  mM o aproximadamente 3 mg/mL. La concentración de una sal de un metal alcalino y un haluro en el medio base o un medio que comprende el medio base y complementos puede ser de, por ejemplo, no más de aproximadamente 100, 90, 80, 70, 60, o 50 mM. Por ejemplo, el medio base o un medio que comprende el medio base y complementos puede comprender una concentración de una sal de un metal alcalino y un haluro de aproximadamente 50-110, 60-105, 70-95, 80-90, 90 mM, o 85 mM. De manera alternativa, la concentración de una sal de un metal alcalino y haluro puede ser de, por ejemplo,  $50 \pm 5$  mM,  $87 \pm 5$  mM,  $110 \pm 5$  mM, aproximadamente 3 mg/mL, aproximadamente 5,1 mg/mL, o aproximadamente 6,4 mg/mL.

En otra realización, el medio base exhibe una concentración de una sal de ácido carbónico. La sal de ácido carbónico puede ser una sal sódica. En un ejemplo como tal, la sal sódica puede ser bicarbonato de sodio. En una realización en particular, el bicarbonato de sodio se presenta en el medio base a una concentración de aproximadamente  $26 \pm 5$  mM o aproximadamente 2,2 mg/mL. La concentración de una sal de ácido carbónico en el medio base o un medio que comprende el medio base y complementos puede ser, por ejemplo, no más de 45, 40, 35, 30, 25, o 20 mM. Por ejemplo, el medio base o un medio que comprende el medio base y complementos puede comprender una concentración de sal de ácido carbónico en el medio base de aproximadamente 10-40, 18-44, 17-30, 18-26, 13-25, 20-30, 25-26, 18, o 26 mM. De manera alternativa, la concentración de sal de ácido carbónico puede ser de, por ejemplo,  $18 \pm 5$  mM,  $26 \pm 5$  mM, de aproximadamente 1,5 mg/mL, o de aproximadamente 2.2 mg/mL.

La suma de la concentración de la sal del metal alcalino y haluro y la sal de ácido carbónico en el medio base o un medio que comprende el medio base y complementos puede ser de, por ejemplo, no más de 140, 130, 120, 110, 100, 90, u 80 mM. Por ejemplo, el medio base o un medio que comprende el medio base y complementos puede comprender una concentración de suma de una sal de un metal alcalino y haluro y una sal de ácido carbónico de aproximadamente 80-140, 85-130, 90-120, 95-120, 100-120, o 115 mM.

La proporción molar de la sal del metal alcalino y haluro y de la sal de ácido carbónico en el medio base o un medio que comprende el medio base y complementos puede ser, por ejemplo, mayor que 2,5. Por ejemplo, el medio base o un medio que comprende el medio base y complementos puede comprender una proporción molar de una sal de un

metal alcalino y haluro y una sal de ácido carbónico de aproximadamente 2,6-4,0, 2,8-3,8, 3,0-3,6, 3,2-3,4, 3,3-3,5, o 3,4.

5 En inclusive otra realización, el medio base es un medio base de baja osmolaridad. La osmolaridad del medio base puede encontrarse en un rango de aproximadamente 175-280 mOsm/kg, aproximadamente 180-250 mOsm/kg, aproximadamente 190-225 mOsm/kg, o aproximadamente 195-205 mOsm/kg. Una osmolaridad de ejemplo del medio base puede ser 200, 214, 216, o 218 mOsm/kg. En un ejemplo en particular, la osmolaridad del medio base es 200 mOsm/kg. La osmolaridad puede determinarse cuando las células se cultivan a diferentes concentraciones de CO<sub>2</sub>.  
 10 En algunos ejemplos, las células se cultivan en CO<sub>2</sub> al 3% o CO<sub>2</sub> al 5%. La osmolaridad del medio base o un medio que comprende el medio base y complementos puede ser, por ejemplo, no más de aproximadamente 330, 320, 310, 300, 290, 280, 275, 270, 260, 250, 240, 230, 220, 210, o 200 mOsm/kg. Por ejemplo, el medio base o el medio que comprende el medio base y complementos puede comprender una osmolaridad de aproximadamente 200-329, 218-322, 240-320, 250-310, 275-295, o 260-300 mOsm/kg. Por ejemplo, el medio base o el medio que comprende el medio base y los complementos puede comprender una osmolaridad de aproximadamente 270 mOsm/kg, aproximadamente 261 mOsm/kg, o aproximadamente 218 mOsm/kg. De manera alternativa, la osmolaridad puede ser de 218 ± 22 mOsm/kg, 261 ± 26 mOsm/kg, 294 ± 29 mOsm/kg, o 322 ± 32 mOsm/kg.

La osmolaridad del medio base puede ser, por ejemplo, de aproximadamente 130-270, 140-260, 150-250, 160-240, 170-230, 180-220, 190-210, 195-205, o 200 mOsm/kg. De manera alternativa, la osmolaridad del medio base puede ser de aproximadamente 200 ± 70, 200 ± 60, 200 ± 50, 200 ± 40, 200 ± 35, 200 ± 30, 200 ± 25, 200 ± 20, 200 ± 15, 200 ± 10, 200 ± 5, o 200 mOsm/kg. De manera alternativa, la osmolaridad del medio base puede ser, por ejemplo de aproximadamente 130-140, aproximadamente 140-150, aproximadamente 150-160, aproximadamente 160-170, aproximadamente 170-180, aproximadamente 180-190, aproximadamente 190-200, aproximadamente 200-210, aproximadamente 210-220, aproximadamente 220-230, aproximadamente 230-240, aproximadamente 240-250, aproximadamente 250-260, aproximadamente 260-270, aproximadamente 270-280, aproximadamente 280-290, aproximadamente 290-300, aproximadamente 300-310, aproximadamente 310-320, o aproximadamente 320-330 mOsm/kg. De manera alternativa, la osmolaridad del medio base puede ser, por ejemplo, de menos de aproximadamente 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, 180, 170, 160, 150, 140, o 130 mOsm/kg.

La osmolaridad del medio que comprende el medio base y complementos puede ser, por ejemplo, de aproximadamente 205-260, 215-250, 225-240, 230-235, o 233 mOsm/kg. De manera alternativa, la osmolaridad del medio que comprende el medio base y complementos puede ser, por ejemplo, de aproximadamente 233 ± 27, 233 ± 25, 233 ± 20, 233 ± 15, 233 ± 10, 233 ± 5, o 233 mOsm/kg. De manera alternativa, la osmolaridad del medio que comprende el medio base y complementos puede ser, por ejemplo, de aproximadamente 200-205, 205-210, 210-215, 215-220, 220-225, 225-230, 230-235, 235-240, 240-245, 245-250, 250-255, o 255-260 mOsm/kg. De manera alternativa, la osmolaridad del medio que comprende el medio base y complementos puede ser, por ejemplo, de menos de 260, 255, 250, 245, 240, 235, 230, 225, 220, 215, 210, 205, o 200 mOsm/kg.

40 En algunos medios de baja osmolaridad, el medio base comprende aproximadamente 87 ± 5 mM de NaCl y aproximadamente 26 ± 5 mM de carbonato. Por ejemplo, el medio base puede comprender aproximadamente 5,1 mg/mL de NaCl, aproximadamente 2,2 mg/mL de carbonato de sodio y una osmolaridad de aproximadamente 275 mOsm/kg.

45 En algunos medios de baja osmolaridad, el medio base comprende aproximadamente 50 ± 5 mM de NaCl y aproximadamente 26 ± 5 mM de carbonato. Por ejemplo, el medio base puede comprender aproximadamente 3,0 mg/mL de NaCl, aproximadamente 2,2 mg/mL de carbonato de sodio y una osmolaridad de aproximadamente 200 mOsm/kg. En una realización preferida, el medio base comprende NaCl a una concentración de 3,0 mg/mL, bicarbonato de sodio a una concentración de aproximadamente 2,2 mg/mL, y tiene una osmolaridad de 200 mOsm/kg.

Otros ejemplos de medios de baja osmolaridad se describen en WO 2011/156723, US 2011/0307968, y US 2015/0067901.

55 Los complementos que se formulan con el medio base de la divulgación resultan adecuados para fabricar, mantener, o enriquecer poblaciones de células iPS humanas que se divulgan en la presente. Tales complementos se indican como "complementos" o "+ complementos" en esta divulgación. El término "complementos" o la frase "+ complementos" incluye uno o más elementos adicionales que se agregan a los componentes del medio base que se describe en la Tabla 1. Por ejemplo, los complementos pueden incluir, sin limitación, medio F-12® (Gibco), complemento N2® (Gibco; solución 100X), medio NEUROBASAL® (Gibco), complemento B-27® (Gibco; solución 50X), L-glutamina, glucosa, 2-mercaptoetanol, un polipéptido de factor de inhibición de leucemia (LIF), un inhibidor de glucógeno sintasa quinasa 3, un inhibidor de MEK, o cualquier combinación de estos. Los complementos pueden incluir además, por ejemplo, suero fetal bovino (FBS), antibiótico(s), penicilina y estreptomina (a saber, penstrep), sales de piruvato (por ejemplo, piruvato de sodio), y aminoácidos no esenciales (por ejemplo, MEM NEAA).

65 En una realización en particular, el polipéptido de LIF es un polipéptido de LIF humano (hLIF). En algunos ejemplos, se usa un polipéptido hLIF a una concentración de aproximadamente 1-1000 unidades/mL, aproximadamente 20-800

unidades/mL, aproximadamente 50-500 unidades/mL, aproximadamente 75-250 unidades/mL, o aproximadamente 100 unidades/mL.

5 Los medios pueden comprender inhibidores, por ejemplo, que consisten esencialmente en un inhibidor de GSK3 y un inhibidor de MEK. Por ejemplo, el medio puede comprender inhibidores que consisten en un inhibidor de GSK3 y un inhibidor de MEK.

10 En otra realización en particular, el inhibidor de GSK3 comprende CHIR99021. En algunos ejemplos, CHIR99021 se usa a una concentración de aproximadamente 0,1 a 10  $\mu\text{M}$ , aproximadamente 1-5  $\mu\text{M}$ , aproximadamente 2-4  $\mu\text{M}$ , o aproximadamente 3  $\mu\text{M}$ .

15 En otra realización en particular, el inhibidor de MEK comprende PD0325901. En algunos ejemplos, PD0325901 se usa a una concentración de aproximadamente 0,1-5  $\mu\text{M}$ , aproximadamente 0,2-1  $\mu\text{M}$ , aproximadamente 0,3-0,7  $\mu\text{M}$ , o aproximadamente 0,5  $\mu\text{M}$ .

20 Un medio de ejemplo comprende un medio base de baja osmolaridad que se describe en la presente de aproximadamente 24,75% (v/v), medio F-12 a aproximadamente 24,75% (v/v), complemento de N2 a aproximadamente 0,5% (v/v), medio NEUROBASAL a aproximadamente 49% (v/v), complemento de B-27 a aproximadamente 1% (v/v), L-glutamina a aproximadamente 2 mM, 2-mercaptoetanol a aproximadamente 0,1 mM, hLIF a aproximadamente 100 unidades/mL, CHIR99021 a aproximadamente 3  $\mu\text{M}$ , y PD0325901 a aproximadamente 0,5  $\mu\text{M}$ .

25 En otra realización en particular, el medio puede comprender o no factor de crecimiento de fibroblastos básico (bFGF, que se conoce también como FGF2 o FGF- $\beta$ ). Preferiblemente, el presente medio no comprende bFGF.

30 El medio puede comprender o no uno o más del complemento de factor de crecimiento transformante beta-1 (TGF- $\beta$ -1), complemento de bFGF, inhibidor de quinasa c-Jun N-terminal (JNK) (por ejemplo, SP600125), inhibidor de proteína quinasa p-38 activada por mitógeno (p-38) (por ejemplo, SB203580), inhibidor de proteína quinasa asociada con Rho (ROCK) (por ejemplo, Y-27632), e inhibidor de proteína quinasa C (PKC) (por ejemplo, Go6983). El medio puede comprender o no forskolina. Por ejemplo, algunos medios no comprenden uno o más de complementos de TGF- $\beta$ 1, complemento de bFGF, inhibidor de JNK (por ejemplo, SP600125), inhibidor de p38 (por ejemplo, SB203580), inhibidor de ROCK (por ejemplo, Y-27632), e inhibidor de PKC (por ejemplo, Go6983). Algunos medios no comprenden uno o más de inhibidor de p38 e inhibidor de JNK. Algunos medios no comprenden complemento de bFGF o complemento de TGF- $\beta$ 1. Algunos medios no comprenden complemento de TGF- $\beta$ 1. Algunos medios no comprenden uno cualquiera de complemento de TGF- $\beta$ 1, complemento de bFGF, inhibidor de JNK (por ejemplo, SP600125), inhibidor de p38 (por ejemplo, SB203580), inhibidor de ROCK (por ejemplo, Y-27632), e inhibidor de PKC (por ejemplo, Go6983). Algunos medios no comprenden forskolina.

#### 40 B. Células madre pluripotentes inducidas humanas

Se proporcionan métodos y composiciones en la presente para fabricar una población de células iPS humanas. Se proporcionan además métodos y composiciones para mantener células iPS humanas en cultivo. Se proporcionan, además, células iPS humanas que se producen o se mantienen en cultivo.

45 El término "células pluripotente" o "célula madre pluripotente" incluye una célula no diferenciada que posee la capacidad de desarrollarse en más de un tipo de célula diferenciada. Tales células pluripotentes pueden ser, por ejemplo, una célula madre embrionaria de mamífero (célula ES) o una célula madre pluripotente inducida de mamífero (célula iPS). Ejemplos de células pluripotentes incluyen células iPS humanas.

50 El término "célula madre embrionaria" o "célula ES" se refiere a una célula madre totipotente o pluripotente que deriva de embrión, deriva a partir de masa celular interna de un blastocito, que puede mantenerse en un cultivo *in vitro* en condiciones adecuadas. Las células ES son capaces de diferenciarse en células de cualquiera de las tres capas germinales de vertebrados, por ejemplo, el endodermo, el ectodermo, o el mesodermo. Las células ES se caracterizan además por su capacidad de propagarse de manera indefinida en condiciones de cultivo *in vitro* adecuadas. Véase, por ejemplo, Thomson et al. (Science (1998) Vol. 282(5391), pp. 1145-1147).

55 El término "célula madre pluripotente inducida" o "célula iPS" incluye una célula madre pluripotente que puede derivar directamente a partir de una célula adulta diferenciada. Las células iPS humanas pueden generarse mediante la introducción de conjuntos específicos de factores de reprogramación en una célula no pluripotente que pueden incluir, por ejemplo, Oct3/4, factores de transcripción de familia Sox (por ejemplo, Sox1, Sox2, Sox3, Sox15), factores de transcripción de familia Myc (por ejemplo, c-Myc, 1-Myc, n-Myc), factores de transcripción de la familia tipo Krüppel (KLF) (por ejemplo, KLF1, KLF2, KLF4, KLF5), y/o factores de transcripción relacionados, tales como NANOG, LIN28, y/o Glis1. Las células iPS humanas pueden generarse además, por ejemplo, mediante el uso de miARNs, pequeñas moléculas que imitan las acciones de factores de transcripción, o especificadores de linaje. Las células iPS humanas se caracterizan por su capacidad de diferenciarse en cualquier célula de las tres capas germinales de vertebrados, por ejemplo, el endodermo, el ectodermo, o el mesodermo. Las células iPS humanas se caracterizan además por su

capacidad de propagarse de manera indefinida en condiciones de cultivo *in vitro*. Véase, por ejemplo, Takahashi and Yamanaka (Cell (2006) Vol. 126(4), pp. 663-676).

5 Los términos “virgen” y “cebadas” identifican diferentes estados de pluripotencia de las células iPS humanas. El término “de apariencia virgen” identifica una célula que expresa un estado pluripotente que exhibe una o más características de una célula pluripotente virgen. Las células iPS humanas de apariencia virgen pueden denominarse también como células iPS humanas “de tipo virgen”. Los términos “de apariencia virgen” y “de tipo virgen” están previstos para ser equivalentes. En algunas realizaciones, las células iPS humanas de apariencia virgen exhiben una o más características morfológicas de células iPS humanas vírgenes, tales como una morfología que se caracteriza por colonias compactas de forma de cúpula. En algunas realizaciones, las células iPS humanas de apariencia virgen expresan uno o más de los marcadores de pluripotencia que se describen en la presente. En algunas realizaciones, las células iPS humanas vírgenes o de apariencia virgen son células iPS humanas vírgenes. En otras realizaciones, las células iPS humanas vírgenes o de apariencia virgen son células iPS de apariencia virgen.

15 Las características de células iPS vírgenes y cebadas se describen en la técnica. Véase, por ejemplo, Nichols y Smith (Cell Stem Cell (2009) Vol. 4(6), pp. 487-492). Las células iPS humanas vírgenes exhiben un estado de pluripotencia similar a aquel de células ES de la masa celular interna de un embrión de previo a la implantación. Tales células vírgenes no se ceban para especificación de linaje y destino. Las células iPS vírgenes femeninas se caracterizan por dos cromosomas X activos. En cultivo, la autorenovación de células iPS humanas vírgenes depende del factor de inhibición de leucemia (LIF) y otros inhibidores. Las células iPS humanas vírgenes cultivadas muestran una morfología clonal que se caracteriza por colonias redondeadas de forma de cúpula y una falta de polaridad apicobasal. Las células vírgenes cultivadas pueden mostrar además uno o más marcadores de pluripotencia según se describe en otra parte en la presente. En condiciones adecuadas, el tiempo de duplicación de células iPS humanas vírgenes en cultivo puede ser de entre 16 y 24 horas.

20 Las células iPS humanas cebadas expresan un estado de pluripotencia similar a aquel de células tipo epiblastos posteriores a la implantación. Tales células se ceban para especificación de linaje y destino. Las células iPS cebadas femeninas se caracterizan por un cromosoma X activo y un cromosoma X inactivo. En cultivo, la autorenovación de células iPS humanas cebadas depende del factor de crecimiento de fibroblastos (FGF) y activina. Las células iPS humanas cebadas que se cultivan muestran una morfología clonal que se caracteriza por una monocapa epitelial y muestran polaridad apicobasal. En condiciones apropiadas, el tiempo de duplicación de células iPS humanas en cultivo puede ser de 24 horas o más.

30 En una realización, las células iPS humanas pueden derivar a partir de células no pluripotentes que se transforman para expresar un estado pluripotente. Tales células transformadas incluyen, por ejemplo, células que se han transformado para expresar genes de reprogramación que inducen pluripotencia. Un estado pluripotente puede incluir, por ejemplo, la expresión de uno o más de los marcadores de pluripotencia que se describen en la presente. Tales células (tales como fibroblastos de prepucio humano) pueden transformarse para expresar genes de reprogramación, o cualquiera de los genes adicionales de interés, mediante cualquier medio que se conoce en la técnica. Véase, por ejemplo, Takahashi and Yamanaka (Cell (2006) Vol. 126(4), pp. 663-676). Por ejemplo, pueden introducirse en las células usando uno o más plásmidos, vectores lentivirales, o vectores retrovirales. En algunos casos, los vectores se integran en el genoma y pueden retirarse después de que la reprogramación se ha completado. En realizaciones en particular, las células no pluripotentes se transforman con genes de reprogramación que comprenden Oct4, Sox2, Klf4, Myc, o cualquier combinación de estos. En algunos ejemplos, las células transformadas comprenden células iPS humanas cebadas.

35 En algunas realizaciones, las células iPS humanas que se cultivan en el medio de baja osmolaridad que se describe en la presente expresan uno o más fenotipos, perfiles de expresión génica, o marcadores característicos de un estado virgen. En un ejemplo, las células iPS humanas expresan uno o más marcadores de pluripotencia cuya expresión indica un estado virgen. Tales marcadores de pluripotencia pueden incluir fosfatasa, NANOG, 5T4, ABCG2, Activina RIB/ALK-4, Activina RIIB, Cadherina-E, Cbx2, CD9, CD30/TNFRSF8, CD117/c-kit, CDX2, CHD1, Cripto, DNMT3B, DPPA2, DPPA4, DPPA5/ESG1, EpCAM/TROP1, ERR beta/NR3B2, ESGP, proteína F-box 15/FBXO15, FGF-4, FGF-5, FoxD3, GBX2, GDNF/NR6A1, GDF-3, Gi24/VISTA/ B7-H5, integrina alfa 6/CD49f, integrina alfa 6 beta 1, integrina alfa 6 beta 4, integrina beta 1/CD29, KLF4, KLF5, L1TD1, Lefty, Lefty-1, Lefty-A, LIN-28A, LIN-28B, LIN-41, cMaf, cMyc, Oct-3/4, Oct-4A, Podocalixina, Rex-1/ZFP42, Smad2, Smad2/3, SOX2, SSEA-1, SSEA-3, SSEA-4, STAT3, Stella/Dppa3, SUZ12, TBX2, TBX3, TBX5, TERT, TEX19, TEX19.1, THAP11, TRA-1-60(R), TROP-2, UTF1, y/o ZIC3. En un ejemplo específico, el marcador de pluripotencia expresado es fosfatasa alcalina, NANOG, o ambos.

40 En otra realización, células iPS humanas que se cultivan en el medio de baja osmolaridad que se describe en la presente muestran características morfológicas que indican un estado virgen. Una morfología de ejemplo se caracteriza por células que tienen colonias compactas de forma de cúpula en cultivo.

45 Las células iPS humanas que se cultivan en el medio de baja osmolaridad que se describe en la presente pueden tener un cariotipo normal. Un cariotipo normal incluye un cariotipo en el que todos los cromosomas normalmente característicos de las especies se presentan y no se han alterado de manera notable o un estado de células que carecen de cualquier anomalía cromosómica numérica o estructural visible que puede detectarse con análisis de

bandeo de cromosomas. Las células iPS humanas que se cultivan en el medio de baja osmolaridad que se describe en la presente puede tener un cariotipo normal, por ejemplo, después de aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, o 100 pases en el medio de baja osmolaridad que se describe en la presente.

En otra realización, las células iPS humanas que se cultivan en el medio de baja osmolaridad que se describe en la presente pueden disociarse de manera mecánica o enzimática en una suspensión de célula única, pasarse, y/o subcultivarse. Tales células iPS humanas que se cultivan en el medio de baja osmolaridad que se describe en la presente pueden tener un cariotipo normal y pueden mantener el cariotipo normal después de disociarse de manera mecánica o enzimática en una suspensión de célula única, pasarse y/o subcultivarse. Por ejemplo, tales células iPS humanas que se cultivan en el medio de baja osmolaridad que se describe en la presente pueden tener un cariotipo normal y pueden mantener el cariotipo normal después de disociarse de manera mecánica o enzimática en una suspensión de célula única, modificarse en un locus genómico diana usando los métodos que se describen en otra parte en la presente, y subcultivarse. En un ejemplo, la disociación enzimática puede llevarse a cabo usando tripsina.

Cuando se cultivan en el presente medio de baja osmolaridad, las células iPS humanas pueden proporcionar mayor eficiencia de transformación debido a la disociación mejorada en una suspensión de célula única. Con otros tipos de medios (por ejemplo, medio mTeSR™ o medio 2i) que se usan normalmente para mantener células iPS humanas en cultivo, la disociación de células iPS humanas debe realizarse de manera mecánica o con enzimas tales como colagenasa que son menos severas que la tripsina. Se recomienda de manera general no pasar células iPS humanas como células únicas, ya que esta práctica ha demostrado colocar presiones selectivas no deseadas en poblaciones celulares que puede conducir a, por ejemplo, aberraciones genéticas en cultivo. Las células iPS humanas resultan vulnerables a apoptosis al momento de separación y disociación celular, y sufren normalmente muerte celular masiva después de disociación completa. Véase Watanabe et al. (2007) Nature 25(6):681-686. De este modo, la disociación de células iPS humanas se realiza normalmente con reactivos o métodos que minimizan la rotura de colonias cuando sucede el pase y no crean suspensiones de célula única. Por consiguiente, las células no se disocian de manera tan efectiva o tan completa. Sin embargo, la disociación completa puede resultar importante para procedimientos tales como aislamiento clonal a continuación de la transferencia génica o generación de una modificación genética dirigida, particularmente cuando se trata de aislar clones relativamente raros tales como aquellos que sufren recombinación homóloga para producir una modificación dirigida conveniente. En cambio, con el presente medio de baja osmolaridad, puede usarse tripsina para disociar las células, y la disociación mejorada da como resultado eficiencia de transformación aumentada. Por ejemplo, tal disociación puede crear suspensiones de célula única que dan como resultado mayores eficiencias dirigidas cuando se dirigen, por ejemplo, mediante electroporación o usando los métodos para fabricar modificaciones genéticas dirigidas que se describen en otra parte en la presente. Además, a diferencia de otros tipos de medio que se usan normalmente para mantener células iPS humanas en cultivo (por ejemplo, medio mTeSR™ o medio 2i), la disociación enzimática de células iPS humanas que se cultivan con el presente medio de baja osmolaridad (preferiblemente un medio de baja osmolaridad que no comprende bFGF) puede realizarse en la ausencia de uno o más inhibidores que resultan generalmente necesarios para el pase de tales células. Un inhibidor de ejemplo que puede omitirse es un inhibidor de proteína quinasa asociada con Rho (ROCK). Un inhibidor de ROCK resulta generalmente necesario cuando se produce el pase de células iPS humanas para inhibir la activación de vías proapoptóticas. En particular, la adición de un inhibidor de ROCK se recomienda de manera general cuando las suspensiones de célula única de células iPS humanas se colocan en placas, ya que se ha informado que esto aumenta la supervivencia de células. Véase Watanabe et al. (2007) Nature 25(6):681-686. Cuando se usa el medio de baja osmolaridad que se divulga en la presente, sin embargo, tales inhibidores de ROCK no resultan necesarios, incluso durante el pase de suspensiones de célula única. Tales suspensiones de célula única pueden mantener la pluripotencia y un cariotipo normal a continuación del tratamiento con tripsina y recolocación en placa cuando se usa el medio de baja osmolaridad que se divulga en la presente.

En una realización adicional, las células iPS humanas subcultivadas en el medio de baja osmolaridad que se describe en la presente pueden mantener un estado virgen o de apariencia virgen a continuación de disociación enzimática y subcultivo. Las células iPS humanas subcultivadas que se cultivan en el medio de baja osmolaridad que se describe en la presente pueden mantener un estado virgen o de apariencia virgen a continuación de disociación enzimática y subcultivo incluso cuando se pasan como suspensiones de célula única y/o cuando sucede modificación en un locus genómico diana usando los métodos que se describen en otra parte en la presente. En algunos ejemplos, las células iPS humanas pueden continuar mostrando una morfología que se caracteriza por colonias compactas de forma de cúpula. Las células iPS humanas subcultivadas pueden continuar expresando además uno o más marcadores de pluripotencia según se describe en la presente.

#### C. Métodos de fabricación y mantenimiento de una población de células madre pluripotentes inducidas humanas

Se proporcionan métodos y composiciones para fabricar células iPS humanas en un cultivo *in vitro*. Se proporcionan además métodos y composiciones para mantener células iPS humanas en un cultivo *in vitro*.

El término "fabricar" incluye cultivar células no pluripotentes que se transforman para expresar uno o más factores de reprogramación según se describe en la presente, en condiciones adecuadas para inducir un cambio de fenotipo celular, expresión génica, o ambos, de manera tal que las células muestran un estado virgen o de apariencia virgen,

5 a saber, expresan una o más características de células iPS humanas vírgenes. Un estado virgen o de apariencia virgen puede expresarse como respuesta a condiciones de cultivo en particular, por ejemplo, cultivo en un medio de baja osmolaridad según se describe en la presente. En algunos ejemplos, la proporción de células que expresan un estado virgen o de apariencia virgen es de al menos aproximadamente el 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, y hasta el 100% de las células en cultivo.

10 En una realización, el método enriquece un cultivo *in vitro* para una población de células iPS humanas vírgenes o de apariencia virgen. En una realización como tal, las células iPS humanas vírgenes o de apariencia virgen pueden propagarse en cultivo preferiblemente en células que no expresan un estado virgen o de apariencia virgen. En otra realización, las células iPS humanas vírgenes o de apariencia virgen pueden seleccionarse a partir de un cultivo, disociarse de manera enzimática, y subcultivarse para producir una población mejorada de células iPS humanas vírgenes o de apariencia virgen.

15 En una realización, las células no pluripotentes que se transforman para expresar un estado pluripotente, se cultivan *in vitro* en un medio que se proporciona en la presente, que resulta adecuado para inducir expresión de un estado virgen o de apariencia virgen durante un período de al menos 1, 2, 5, 7, 10, 14, 21, o 28 días, o cualquier período de tiempo suficiente para inducir expresión de un estado virgen o de apariencia virgen en cultivo. Las células transformadas pueden cultivarse en el presente medio durante al menos 1, 2, 3, o 4 semanas. Algunas veces las células transformadas se cultivan durante 1-4 semanas. La expresión de un estado virgen o de apariencia virgen puede determinarse mediante observación de características morfológicas o la expresión de marcadores pluripotentes, característicos de un estado virgen o de apariencia virgen, que se describen en otra parte en la presente.

25 En una realización, las células no pluripotentes que se transforman para expresar un estado pluripotente, se cultivan en el presente medio de baja osmolaridad hasta que expresan características de un estado virgen o de apariencia virgen. Las células pueden cultivarse luego en el presente medio para mantener un estado virgen o de apariencia virgen. En otra realización, las células no pluripotentes que se transforman para expresar un estado pluripotente, se cultivan primero en un medio de alta osmolaridad antes del cultivo en el presente medio de baja osmolaridad. Tal medio de alta osmolaridad exhibe una osmolaridad mayor que el presente medio de baja osmolaridad y puede comprender bFGF. La osmolaridad del medio de alta osmolaridad puede ser, por ejemplo, de aproximadamente 300-380, 310-370, 320-360, 330-350, o 340 mOsm/kg. De manera alternativa, la osmolaridad del medio de alta osmolaridad puede ser, por ejemplo,  $340 \pm 70$ ,  $340 \pm 60$ ,  $340 \pm 50$ ,  $340 \pm 40$ ,  $340 \pm 30$ ,  $340 \pm 20$ , o  $340 \pm 10$  mOsm/kg. Por ejemplo, la osmolaridad del medio de alta osmolaridad puede ser de aproximadamente 270-280, 280-290, 290-300, 300-310, 310-320, 320-330, 330-340, 340-350, 350-360, 360-370, 370-380, 380-390, 390-400, o 400-410 mOsm/kg. De manera alternativa, la osmolaridad del medio de alta osmolaridad puede ser al menos de aproximadamente 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, o 410 mOsm/kg. Algunos medios de alta osmolaridad comprenden uno o más de albúmina de suero bovino, bFGF, factor de crecimiento transformante  $\beta$  (TGF $\beta$ ), cloruro de litio, ácido pipecólico, y ácido gamma aminobutírico (GABA). Ejemplos de un medio de alta osmolaridad incluyen medio mTeSR™ (Stemcell Technologies).

40 En algunas realizaciones, las células no pluripotentes que se transforman para expresar un estado pluripotente, pueden cultivarse primero en medio de alta osmolaridad que comprende bFGF hasta que comienzan a expresar características de un estado virgen o de apariencia virgen, en ese momento las células se cultivan en el presente medio de baja osmolaridad. En un ejemplo, las células pueden cultivarse en medio de alta osmolaridad que comprende bFGF durante un período de al menos 1, 2, 5, 10, 30, 60, o 90 días, un período de 1, 2, 4, 8, o 12 semanas, o un período entre 1 día a 3 meses. Un período de tiempo de ejemplo para cultivo en un medio de alta molaridad que comprende bFGF es 2 meses.

50 En otra realización, las células no pluripotentes que se transforman para expresar un estado pluripotente, pueden cultivarse primero en un medio de alta osmolaridad que comprende bFGF hasta que comienzan a mostrar una morfología que se caracteriza por grupos de células tridimensionales, en ese momento las células se cultivan en el presente medio de baja osmolaridad. En tales realizaciones, las células que muestran grupos tridimensionales pueden seleccionarse, disociarse (por ejemplo, mediante tripsina), y transferirse a un nuevo cultivo en el medio de baja osmolaridad que se describe en la presente.

55 Los términos “mantiene”, “que mantiene” incluyen la conservación de al menos una o más de las características o fenotipos de las células iPS humanas que se describen en la presente. Tales características pueden incluir mantener la pluripotencia, morfología celular, perfiles de expresión génica, y/u otras características funcionales de células vírgenes. Los términos “mantiene”, “que mantiene” y “mantenimiento” pueden incluir además la propagación de células y/o un aumento del número de células vírgenes que se cultivan. Los términos incluyen condiciones de cultivo que impiden que las células se conviertan a un estado cebado o no pluripotente. Los términos incluyen además condiciones de cultivo que permiten que las células permanezcan pluripotentes y/o vírgenes, mientras que las células pueden continuar o no con división y aumento de número.

65 En una realización, las células iPS humanas se cultivan *in vitro* en un medio que se proporciona en la presente que resulta adecuado para mantener tales células en un estado virgen o de apariencia virgen. En un ejemplo en particular, las células iPS humanas pueden cultivarse en un medio adecuado durante un período de 1, 2, 5, 7, 10, 14, 21, o 28

días, o durante un período de aproximadamente 2 semanas, aproximadamente 3 semanas, aproximadamente 4 semanas, o más, siempre que las células cultivadas se mantengan en un estado virgen o de apariencia virgen. Las células pueden cultivarse durante al menos 1, 2, 3 o 4 semanas. Algunas veces las células se cultivan durante 1-4 semanas. Las células iPS humanas pueden mantenerse, por ejemplo, durante cualquier período de tiempo suficiente para propagación de las células en cultivo, modificación genética de las células, y/o subcultivo de las células.

En otra realización, las células iPS humanas o células no pluripotentes que se transforman para expresar un estado pluripotente, pueden cultivarse en un sustrato o capa de células alimentadoras adecuada para cultivo *in vitro*. En un ejemplo en particular, las células se cultivan en MATRIGEL™ (BD Biosciences). En otro ejemplo, las células se cultivan en células alimentadoras de fibroblastos de prepucio de humanos recién nacidos (NuFF). En otro ejemplo, las células se cultivan en GELTRES™ (Life Technologies). En otro ejemplo, las células se cultivan en vitronectina (por ejemplo, VITRONECTIN XF™ (STEMCELL Technologies).

En una realización adicional, el tiempo de duplicación de células iPS humanas en el presente medio de baja osmolaridad se reduce en comparación con las células iPS humanas cebadas o células no pluripotentes que se transforman para expresar un estado pluripotente. En un ejemplo en particular, el tiempo de duplicación de las presentes células iPS humanas se encuentra entre aproximadamente 16-24 horas.

#### D. Modificaciones genéticas y métodos para fabricar modificaciones genéticas dirigidas

En algunas realizaciones, los métodos para fabricar y mantener células iPS humanas comprenden introducir una modificación genética en las células iPS humanas. Asimismo, la divulgación proporciona células iPS humanas que comprenden una modificación genética.

En realizaciones en particular, la modificación genética comprende una modificación de uno o más ácidos nucleicos endógenos, una sustitución de uno o más ácidos nucleicos endógenos, un reemplazo de un ácido nucleico endógeno con un ácido nucleico heterólogo, un bloqueo o un knock-in. En ejemplos específicos, la modificación genética se introduce mediante la introducción de un vector director grande (LTVEC) en las células iPS humanas o las células no pluripotentes que se transforman para expresar un estado pluripotente. En un ejemplo como tal, el LTVEC puede comprender ADN para insertarse en el genoma de las células.

Pueden usarse diversos métodos para fabricar modificaciones genéticas dirigidas en células iPS humanas. Por ejemplo, pueden usarse diversos métodos para fabricar modificaciones genéticas dirigidas que modifican el nivel y/o la actividad de proteínas en células iPS humanas. Por ejemplo, en un caso, la modificación genética dirigida emplea un sistema que generará una modificación genética dirigida mediante un evento de recombinación homóloga (HR). La reparación dirigida por homología (HDR) o HR incluye una forma de reparación de ácido nucleico que puede requerir homología de secuencia de nucleótidos, usa una molécula "donante" para reparación modelo de una molécula "diana" (a saber, la que experimentó la rotura bicatenaria), y conduce a la transferencia de información genética a partir del donante a diana. Sin ánimos de apegarse a ninguna teoría en particular, tal transferencia puede incluir corrección de incompatibilidades de ADN heterodúplex que se forma entre la diana rota y el donante, y/o apareamiento de cadena dependiente de síntesis, en el que el donante se usa para resintetizar información genética que se volverá parte de la diana, y/o procesos relacionados. En algunos casos, el polinucleótido donante, una porción del polinucleótido donante, una copia del polinucleótido donante, o una porción de una copia del polinucleótido donante se integra en el ADN diana. En otros casos, una célula puede modificarse usando agentes de nucleasa que generan una rotura monocatenaria o bicatenaria en una locación genómica dirigida. La rotura monocatenaria o bicatenaria se repara luego mediante la ruta de unión de extremos no homólogos (NHEJ). NHEJ incluye la reparación de roturas bicatenarias en un ácido nucleico mediante ligamientos directos de los extremos de rotura uno con respecto al otro sin la necesidad de un modelo homólogo. El ligamiento de secuencias no contiguas mediante NHEJ puede dar como resultado frecuentemente eliminaciones, inserciones, o translocaciones próximas al sitio de rotura bicatenaria. Tales sistemas utilizan, por ejemplo, generación de pérdida dirigida de modificaciones genéticas funcionales. Métodos no limitantes para generar tal modificación genética dirigida se evalúan en detalle en otra parte en la presente, incluyendo, por ejemplo, el uso de plásmidos directores o vectores directores grandes. Véase, además, Wang et al. (2013) Cell 153:910-918, Mandalos et al. (2012) PLOS ONE 7:e45768:1-9, y Wang et al. (2013) Nat Biotechnol. 31:530-532.

Se reconoce que en realizaciones específicas, la modificación genética dirigida de cualquier polinucleótido de interés puede ocurrir mientras que la célula pluripotente (a saber, célula iPS humana) se mantiene en el medio de cultivo que se describe en la presente. De manera alternativa, la modificación genética dirigida de cualquier polinucleótido de interés puede ocurrir mientras que la célula pluripotente (a saber, la célula iPS humana) se mantiene en medio de cultivo diferente, y se transfiere posteriormente a un medio de cultivo de baja osmolaridad que se divulga en la presente.

De manera general, el nivel y/o actividad de una proteína se modifica si el nivel de proteína y/o el nivel de actividad de la proteína resulta estadísticamente mayor o menor que el nivel de proteína en una célula de control adecuada que no se ha modificado genéticamente o mutagenizado para alterar la expresión y/o actividad de la proteína. En realizaciones específicas, la concentración y/o actividad se altera en al menos el 1%, 5%, 10%, 20%, 30%, 40%, 50%,

60%, 70%, 80%, o 90% con respecto a una célula de control que no se ha modificado para tener un nivel y/o actividad modificados de la proteína.

5 Una "célula objeto" es una en la que se ha efectuado una alteración genética, tal como una modificación genética que se divulga en la presente, o es una célula que desciende a partir de una célula alterada de esta manera y que comprende la alteración. Un "control" o "célula de control" proporciona un punto de referencia para medir cambios de fenotipo de la célula objeto. En una realización, una célula de control se corresponde tan estrechamente como resulta posible con la célula con actividad de proteína reducida a excepción de que carece de modificación genética o  
10 mutación dando como resultado la actividad modificada (por ejemplo, las células respectivas pueden originarse a partir de la misma línea celular). En otros casos, la célula de control puede comprender, por ejemplo: (a) una célula de tipo salvaje, a saber, del mismo genotipo que el material de partida para la alteración genética que dio como resultado la célula objeto; (b) una célula del mismo genotipo que el material de partida que se ha modificado genéticamente con una construcción nula (a saber, con una construcción que no tiene efecto conocido en el rasgo de interés, tal como  
15 una construcción que comprende un gen marcador); (c) una célula que es una progenie de una célula objeto que no se ha modificado genéticamente (a saber, la célula de control y la célula objeto se originan a partir de la misma línea celular); (d) una célula genéticamente idéntica a la célula objeto pero que no se expone a condiciones o estímulos que inducirían expresión del gen de interés; o (e) la propia célula objeto, en condiciones en que la modificación genética no da como resultado una alteración de la expresión del polinucleótido de interés.

20 El nivel de expresión del polipéptido puede medirse directamente, por ejemplo, mediante el análisis del nivel del polipéptido en la célula u organismo, o de manera indirecta, por ejemplo, mediante la medición de la actividad del polipéptido.

25 En otros casos, las células que tienen la modificación genética dirigida se seleccionan usando métodos que incluyen, pero sin limitación, análisis de Southern blot, secuenciación de ADN, análisis de PCR, o análisis fenotípico. Tales células se emplean luego en los diversos métodos y composiciones que se describen en la presente.

30 Una modificación genética dirigida puede comprender una alteración dirigida a un polinucleótido de interés. Tales modificaciones dirigidas incluyen, pero sin limitación, adiciones de uno o más nucleótidos, eliminaciones de uno o más nucleótidos, sustituciones de uno o más nucleótidos, un bloqueo del polinucleótido de interés o una porción de este, un knock-in del polinucleótido de interés o una porción de este, un reemplazo de una secuencia de ácido nucleico endógeno con una secuencia de ácido nucleico heterólogo, o una combinación de estos. En realizaciones específicas, al menos 1, 2, 3, 4, 5, 7, 8, 9, 10, 100, 500 o más nucleótidos o al menos 10 kb a 500 kb o más se cambian para formar  
35 la modificación genómica dirigida.

40 En otras realizaciones, la actividad y/o nivel de un polipéptido se modifica mediante la introducción en la célula de un polinucleótido que altera el nivel o actividad del polipéptido. El polinucleótido puede modificar la expresión del polipéptido directamente, mediante alteración de la traducción del ARN mensajero, o de manera indirecta, mediante la codificación de un polipéptido que altera la transcripción o traducción del gen que codifica una proteína. En otras realizaciones, la actividad de un polipéptido se modifica mediante la introducción en la célula de una secuencia que codifica un polipéptido que altera la actividad del polipéptido diana.

45 En una realización, las células iPS humanas pueden comprender un alelo condicional que modifica la actividad y/o el nivel de una proteína. Un "alelo condicional" incluye un gen modificado que se designa para tener el nivel modificado y/o actividad de la proteína en un período de desarrollo conveniente y/o en un tejido de interés conveniente. El nivel y/o actividad modificados puede compararse con una célula de control que carece de la modificación dando lugar al alelo condicional, o en el caso de actividad modificada en un período de desarrollo conveniente con períodos anteriores y/o siguientes, o en el caso de un tejido conveniente, con una actividad media de todos los tejidos. En una realización,  
50 el alelo condicional comprende un alelo condicional nulo del gen que puede desactivarse o activarse en un período de desarrollo conveniente y/o en tejidos específicos.

55 En una realización no limitante, el alelo condicional es un alelo multifuncional, según se describe en US 2011/0104799. En realizaciones específicas, el alelo condicional comprende: (a) una secuencia de actuación en orientación sentido con respecto de la transcripción de un gen diana, y un cassette de selección de medicamento (DSC) en orientación sentido o antisentido; (b) en orientación antisentido, una secuencia de nucleótidos de interés (NSI) y un condicional por módulo de inversión (COIN, que utiliza un intrón que separa exones y un módulo de tipo trampa génica invertible; véase, por ejemplo, US 2011/0104799); y (c) unidades recombinantes que se recombinan al exponerse a una primera recombinasa para formar un alelo condicional que (i) carece de la secuencia de actuación y el DSC, y (ii) contiene el  
60 NSI en orientación sentido y el COIN en orientación antisentido.

65 La presente divulgación permite modificar un locus genómico diana en un cromosoma en una célula. En ejemplos en particular, los métodos que se proporcionan en la presente para el reconocimiento de un locus genómico en un cromosoma mediante el empleo de un vector director en la ausencia de, o en combinación con, un agente de nucleasa.

Los métodos para llevar a cabo modificaciones genéticas dirigidas pueden comprender, por ejemplo, el uso de un vector director (por ejemplo, un LTVEC), ya sea solo o en combinación con una o más nucleasas, según se describe

en otra parte en la presente. Véase, por ejemplo, US 2015/0159175, US 2015/0159174, US 2014/0310828, US 2014/0309487, y US 2013/0309670. Asimismo, métodos para llevar a cabo modificaciones genéticas dirigidas pueden comprender el uso de una o más nucleasas ya sea por sí solas o en combinación con un vector director.

5 Por ejemplo, se proporcionan métodos para modificar un locus genómico diana en una célula iPS humana, que comprende: (a) introducir en la célula uno o más agentes de nucleasa que inducen uno o más mellas o roturas bicatenarias en un sitio de reconocimiento en o próximo al locus genómico diana; y (b) identificar al menos una célula que comprende en su genoma una modificación en el locus genómico diana. Tales métodos pueden dar como resultado una interrupción del locus genómico diana. La interrupción de la secuencia de ácido nucleico endógeno puede surgir, por ejemplo, cuando una rotura bicatenaria que se crea por una nucleasa se repara mediante reparación de ADN mediada por unión de extremos no homólogos (NHEJ), que genera un alelo mutante que comprende un inserto o una eliminación de una secuencia de ácido nucleico y origina así interrupción de ese locus genómico. Ejemplos de interrupción incluyen alteración de un elemento regulador (por ejemplo, promotor o potenciador), una mutación de sentido erróneo, una mutación sin sentido, una mutación con desplazamiento del marco de lectura, una mutación por truncación, una mutación nula, o un inserto o eliminación de un número pequeño de nucleótidos (por ejemplo, causando una mutación con desplazamiento del marco de lectura). La interrupción puede dar como resultado la inactivación (a saber, pérdida de función) o pérdida del alelo.

20 Otros métodos para modificar un locus genómico diana en una célula iPS humana comprenden: (a) introducir en la célula un vector director que comprende un ácido nucleico de inserción flanqueado por brazos de homología 5' y 3' correspondientes a sitios diana 5' y 3'; y (b) identificar al menos una célula que comprende en su genoma el inserto de ácido nucleico que se integra al locus genómico diana.

25 Otros métodos para modificar un locus genómico diana en una célula iPS humana comprenden: (a) introducir en la célula: (i) un agente de nucleasa, en los que el agente de nucleasa induce una mella o rotura bicatenaria en un sitio de reconocimiento en el locus genómico diana; y (ii) un vector director que comprende un inserto de ácido nucleico flanqueado por brazos de homología 5' y 3' correspondientes a sitios diana 5' y 3' que se ubican lo suficientemente próximos con respecto al sitio de reconocimiento; y (c) identificar al menos una célula que comprende una modificación (por ejemplo, integración del inserto de ácido nucleico) en el locus genómico diana. Tales métodos pueden dar como resultado diversos tipos de modificaciones genéticas dirigidas. Tales modificaciones genéticas dirigidas pueden incluir, por ejemplo, adiciones de uno o más nucleótidos, eliminaciones de uno o más nucleótidos, sustituciones de uno o más nucleótidos, una mutación puntual, un bloqueo de un polinucleótido de interés o una porción de este, un knock-in de un polinucleótido de interés o una parte de este, un reemplazo de una secuencia de ácido nucleico endógeno con una secuencia de ácido nucleico heterólogo, exógeno, u ortólogo, un intercambio de dominio, un intercambio de exón, un intercambio de intrón, un intercambio de secuencia reguladora, un intercambio de genes, o una combinación de estos. Por ejemplo, al menos, 1, 2, 3, 4, 5, 7, 8, 9, 10 o más nucleótidos pueden cambiarse para formar la modificación genómica dirigida. Las eliminaciones, inserciones o reemplazos pueden ser de cualquier tamaño, según se divulga en otra parte en la presente.

40 a. Agentes de nucleasa y sitios de reconocimiento para agentes de nucleasa

45 El término "sitio de reconocimiento para un agente de nucleasa" incluye una secuencia de ADN en la que una mella o rotura bicatenaria se induce mediante un agente de nucleasa. El sitio de reconocimiento para un agente de nucleasa puede ser endógeno (o nativo) con respecto a la célula o el sitio de reconocimiento puede ser exógeno con respecto a la célula. En realizaciones específicas, el sitio de reconocimiento es exógeno con respecto a la célula y no ocurre así de manera natural en el genoma de la célula. En otras realizaciones adicionales, el sitio de reconocimiento es exógeno con respecto a la célula y con respecto a los polinucleótidos de interés que se consideran convenientes para posicionarse en el locus diana. En realizaciones adicionales, el sitio de reconocimiento exógeno o endógeno se presenta solo una vez en el genoma de la célula huésped. En realizaciones específicas, se identifica un sitio endógeno o nativo que ocurre solo una vez en el genoma. Un sitio como tal puede usarse luego para designar agentes de nucleasa que producirán una mella o rotura bicatenaria en el sitio de reconocimiento endógeno.

55 La longitud del sitio de reconocimiento puede variar e incluye, por ejemplo, sitios de reconocimiento que son de aproximadamente 30-36 pb para par de nucleasas con dedos de zinc (ZFN) (a saber, aproximadamente 15-18 pb para cada ZFN), aproximadamente 36 pb para una nucleasa efectora de tipo activador de transcripción (TALEN), o aproximadamente 20 pb para un ARN guía CRISPR/Cas9.

60 En una realización, cada monómero del agente de nucleasa reconoce un sitio de reconocimiento de al menos 9 nucleótidos. En otras realizaciones, el sitio de reconocimiento es de aproximadamente 9 a aproximadamente 12 nucleótidos de largo, de aproximadamente 12 a aproximadamente 15 nucleótidos de largo, de aproximadamente 15 a aproximadamente 18 nucleótidos de largo, o de aproximadamente 18 a aproximadamente 21 nucleótidos de largo, y cualquier combinación de tales subrangos (por ejemplo, 9-18 nucleótidos). Se reconoce que un agente de nucleasa dado puede unirse al sitio de reconocimiento y escindir el sitio de unión o, de manera alternativa, el agente de nucleasa puede unirse a una secuencia que resulta diferente a partir del sitio de reconocimiento. Además, el término sitio de reconocimiento comprende tanto el sitio de unión del agente de nucleasa como el sitio de mella/escisión independientemente de si el sitio de mella/escisión se encuentra en o fuera del sitio de unión de agente de nucleasa.

En otra variación, la escisión por el agente de nucleasa puede ocurrir en posiciones de nucleótidos inmediatamente opuestas una con respecto a la otra para producir un corte de extremo romo o, en otros casos, las incisiones pueden escalonarse para producir salientes monocatenarias, que se denominan también “extremos pegajosos”, que pueden ser ya sea salientes 5' o salientes 3'.

Cualquier agente de nucleasa que induce una mella o rotura bicatenaria en un sitio de reconocimiento conveniente puede usarse en los métodos y composiciones que se divulgan en la presente. Un agente de nucleasa nativo o que ocurre de manera natural puede emplearse siempre que el agente de nucleasa induzca una mella o rotura bicatenaria en un sitio de reconocimiento conveniente. De manera alternativa, puede emplearse un agente de nucleasa modificado o diseñado. Un “agente de nucleasa diseñado” incluye una nucleasa que se diseña (modifica o deriva) a partir de su forma nativa para reconocer de manera específica e inducir una mella o rotura bicatenaria en el sitio de reconocimiento conveniente. De este modo, un agente de nucleasa diseñado puede derivar a partir de un agente de nucleasa nativo que ocurre de manera natural o puede crearse o sintetizarse de manera artificial. La modificación del agente de nucleasa puede ser tan pequeña como un aminoácido en un agente de escisión de proteína o un nucleótido en un agente de escisión de ácido nucleico. En algunas realizaciones, la nucleasa diseñada induce una mella o rotura bicatenaria en un sitio de reconocimiento, en el que el sitio de reconocimiento no fue una secuencia que no se habría reconocido por un agente de nucleasa nativo (no diseñado o no modificado). La producción de una mella o rotura bicatenaria en un sitio de reconocimiento u otro ADN puede denominarse en la presente como “corte” o escisión” del sitio de reconocimiento u otro ADN.

Se proporcionan también variantes activas y fragmentos de los sitios de reconocimiento de ejemplo. Tales variantes activas puede comprender identidad de secuencia de al menos 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más con respecto al sitio de reconocimiento, en el que las variantes activas retienen actividad biológica y, por lo tanto, son capaces de ser reconocidas y escindidas por un agente de nucleasa de una manera de secuencia específica. Análisis para medir la rotura bicatenaria de un sitio de reconocimiento por un agente de nucleasa se conocen en la técnica (por ejemplo, TaqMan® qPCR assay, Frendewey D. et al., *Methods in Enzymology*, 2010, 476:295-307).

El sitio de reconocimiento del agente de nucleasa puede posicionarse en cualquier lugar en o próximo al locus diana. El sitio de reconocimiento puede ubicarse en una región de codificación de un gen, o en regiones reguladoras que influyen la expresión del gen. Un sitio de reconocimiento del agente de nucleasa puede ubicarse en un intrón, un exón, un promotor, un potenciador, una región reguladora, o cualquier región codificadora de no proteína. En realizaciones específicas, el sitio de reconocimiento se posiciona en el polinucleótido que codifica el marcador de selección. Una posición como tal puede ubicarse en la región de codificación del marcador de selección o en las regiones reguladoras, que influyen en la expresión del marcador de selección. De este modo, un sitio de reconocimiento del agente de nucleasa puede ubicarse en un intrón del marcador de selección, un promotor, un potenciador, una región reguladora, o cualquier región codificadora de no proteína del polinucleótido que codifica el marcador de selección. En realizaciones específicas, una mella o rotura bicatenaria en el sitio de reconocimiento interrumpe la actividad del marcador de selección. Los métodos para evaluar la presencia o ausencia de un marcador de selección funcional se conocen.

En una realización, el agente de nucleasa es una nucleasa efectora de tipo activador de transcripción (TALEN). Nucleasas efectoras TAL son una clase de nucleasa de secuencia específica que pueden usarse para crear roturas bicatenarias en secuencias diana específicas en el genoma de un organismo procariontario o eucariótico. Las nucleasas efectoras TAL se crean mediante fusión de un efector (TAL) de tipo activador de transcripción nativo o diseñado, o parte funcional de este, con un dominio catalítico de una endonucleasa, tal como, por ejemplo, FokI. El dominio de unión a ADN del efector TAL modular único permite el diseño de proteínas con especificidad de reconocimiento de potencialmente cualquier ADN dado. De este modo, los dominios de unión a ADN de las nucleasa efectoras TAL pueden diseñarse para reconocer sitios diana de ADN específicos y, de este modo, usarse para fabricar roturas bicatenarias en secuencias diana convenientes. Veanse, WO 2010/079430; Morbitzer et al. (2010) PNAS 10.1073/pnas.1013133107; Scholze & Boch (2010) *Virulence* 1:428-432; Christian et al. *Genetics* (2010) 186:757-761; Li et al. (2010) *Nuc. Acids Res.* (2010) doi:10.1093/nar/gkq704; y Miller et al. (2011) *Nature Biotechnology* 29:143-148.

Ejemplos de nucleasas TAL adecuadas, y métodos para preparar nucleasas TAL adecuadas se divulgan en, por ejemplo, la solicitud de patente US nro. 2011/0239315 A1, 2011/0269234 A1, 2011/0145940 A1, 2003/0232410 A1, 2005/0208489 A1, 2005/0026157 A1, 2005/0064474 A1, 2006/0188987 A1, y 2006/0063231 A1. En diversas realizaciones, las nucleasas efectoras TAL se diseñan para cortar en o próximas a una secuencia de ácido nucleico en, por ejemplo, un locus de interés o un locus genómico de interés, en el que la secuencia de ácido nucleico diana se encuentra en o próxima a una secuencia para modificarse mediante un vector director. Las nucleasas TAL adecuadas para uso con los diversos métodos y composiciones que se proporcionan en la presente incluyen aquellas que se diseñan específicamente para unirse en o próximas a secuencias de ácido nucleico diana para modificarse mediante vectores directores según se describe en la presente.

En una realización, cada monómero de la TALEN comprende 12-25 repeticiones TAL, en la cada repetición TAL se une a un subsitio de 1 pb. En algunas TALENs, cada monómero de la TALEN comprende 33-35 repeticiones TAL que reconocen un par de bases único mediante dos residuos hipervariables. En una realización, el agente de nucleasa es

una proteína quimérica que comprende un dominio de unión a ADN basado en repetición TAL que se une de manera manejable con una nucleasa independiente. En una realización, la nucleasa independiente es una endonucleasa FokI. En una realización, el agente de nucleasa comprende un primer dominio de unión a ADN basado en repetición de TAL y un segundo dominio de unión a ADN basado en repetición de TAL, en el que cada uno de los dominios de ADN basados en repetición de TAL primero y segundo se operan unidos con una nucleasa FokI, en el que los dominios de unión a ADN basado en repetición de TAL primero y segundo reconocen dos secuencias de ADN diana contiguas en cada cadena de la secuencia de ADN diana que se separa por un sitio de escisión de aproximadamente 6 pb a aproximadamente 40 pb, y en el que las nucleasas FokI dimerizan y fabrican una rotura bicatenaria en una secuencia diana. Por ejemplo, el agente de nucleasa puede comprender un primer dominio de unión a ADN basado en repetición de TAL y un segundo dominio de ADN basado en repetición de TAL, en el que cada uno de los dominios de ADN basados en repetición de TAL primero y segundo se unen de manera manejable con una nucleasa FokI, en la que el dominio de unión a ADN basado en repetición TAL primero y segundo reconoce dos secuencias de ADN diana contiguas en cada cadena de la secuencia de ADN diana que se separan por una secuencia separadora de longitud variable (12-20 pb), y en la que las subunidades de nucleasa FokI dimerizan para crear una nucleasa activa que crea una rotura bicatenaria en una secuencia diana.

El agente de nucleasa que se emplea en los diversos métodos y composiciones que se divulgan en la presente puede comprender además una nucleasa con dedos de zinc (ZFN). En una realización, cada monómero de la ZFN comprende 3 o más dominios de unión de ADN basados en dedos de zinc, en los que cada dominio de unión de ADN basado en dedo de zinc se une a subunidades de 3 pb. En otras realizaciones, la ZFN es una proteína quimérica que comprende un dominio de unión a ADN basado en dedos de zinc que se une de manera operable con una nucleasa independiente. En una realización, la endonucleasa independiente es una endonucleasa FokI. En una realización, el agente de nucleasa comprende una primera ZFN y una segunda ZFN, en la que cada una de la ZFN primera y ZFN segunda se unen de manera manejable con una nucleasa FokI, en la que la primera y la segunda ZFN reconocen dos secuencias de ADN diana contiguas en cada cadena de la secuencia de ADN diana que se separa por sitio de escisión de aproximadamente 6 pb a aproximadamente 40 pb, y en la que las nucleasas FokI dimerizan y crean una rotura bicatenaria. Por ejemplo, el agente de nucleasa puede comprender una primera ZFN y una segunda ZFN, en la que cada una de la primera ZFN y de la segunda ZFN se unen de manera manejable con una subunidad de nucleasa FokI, en la que la ZFN primera y segunda reconocen dos secuencias de ADN diana contiguas en cada cadena de la secuencia de ADN diana que se separan por separador de aproximadamente 5-7 pb, y en la que las subunidades de nucleasa FokI dimerizan para crear una nucleasa activa que crea una rotura bicatenaria. Véanse, por ejemplo, US20060246567; US20080182332; US20020081614; US20030021776; WO/2002/057308A2; US20130123484; US20100291048; y, WO/2011/017293A2.

En otra realización más, el agente de nucleasa es una meganucleasa. Las meganucleasas se han clasificado en cuatro familias que se basan en motivos de secuencias conservadas, las familias son las familias LAGLIDADG, GIY-YIG, H-N-H, y His-Cys box. Estos motivos participan en la coordinación de iones metálicos e hidrólisis de enlaces de fosfodiéster. Las meganucleasas se reconocen por sus largos sitios de reconocimiento, y por tolerar algunos polimorfismos de secuencia en sus sustratos de ADN. Los dominios de meganucleasas, estructura, y funciones se conocen, véase por ejemplo, Guhan y Muniyappa (2003) *Crit Rev Biochem Mol Biol* 38:199-248; Lucas et al., (2001) *Nucleic Acids Res* 29:960-9; Jurica y Stoddard, (1999) *Cell Mol Life Sci* 55:1304-26; Stoddard, (2006) *Q Rev Biophys* 38:49-95; y Moure et al., (2002) *Nat Struct Biol* 9:764. En algunos ejemplos, se usan una variante que ocurre naturalmente y/o meganucleasa derivada diseñada. Los métodos para modificar la cinética, interacciones de cofactor, expresión, condiciones óptimas y/o especificidad de sitio de reconocimiento, y análisis de actividad se conocen, véanse, por ejemplo, Epinat et al., (2003) *Nucleic Acids Res* 31:2952-62; Chevalier et al., (2002) *Mol Cell* 10:895-905; Gimble et al., (2003) *Mol Biol* 334:993-1008; Seligman et al., (2002) *Nucleic Acids Res* 30:3870-9; Sussman et al., (2004) *J Mol Biol* 342:31-41; Rosen et al., (2006) *Nucleic Acids Res* 34:4791-800; Chames et al., (2005) *Nucleic Acids Res* 33:e178; Smith et al., (2006) *Nucleic Acids Res* 34:e149; Gruen et al., (2002) *Nucleic Acids Res* 30:e29; Chen y Zhao, (2005) *Nucleic Acids Res* 33:e154; WO2005105989; WO2003078619; WO2006097854; WO2006097853; WO2006097784; y WO2004031346.

Cualquier meganucleasa puede usarse en la presente, pero sin limitación, I-Scel, I-Scell, I-Scelll, I-ScelV, I-SceV, I-SceVI, I-SceVII, I-CeuI, I-CeuAIIP, I-Crel, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-TliI, I-Ppol, PI-Pspl, FScel, F-Scell, F-SuvI, F-TevI, F-TevII, I-Amal, I-Anil, I-Chul, I-Cmoel, I-Cpal, I-Cpall, I-CsmI, I-Cvul, I-CvuAIP, I-Ddil, IDdil, I-Dirl, I-Dmol, I-Hmul, I-Hmull, I-HsNIP, I-Llal, I-Msol, I-Naal, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-Njal, I-Nsp236IP, IPakI, I-PbolIP, I-PculP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PpbIP, I-SpBetaIP, I-Scal, I-SexIP, I-SnelP, I-SpomI, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiSTe3pb, I-TdelP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPAIP, I-UarHGPA13P, I-ViniP, I-ZbilP, PI-Mtul, PI-MtuHIP, PI-MtuHIIP, PI-Pful, PI-Pfull, PI-Pkol, PI-Pkoll, PI-Rma43812IP, PI-SpBetaIP, PI-Scel, PI-Tful, PI-Tfull, PI-Thyl, PI-TliI, PI-TliII, o cualquiera de las variantes activas o fragmentos de estos.

En una realización, la meganucleasa reconoce secuencias de ADN bicatenarias de 12 a 40 pares de bases. En una realización, la meganucleasa reconoce una secuencia diana que corresponde perfectamente en el genoma. En una realización, la meganucleasa es una nucleasa buscadora de objetivos. En una realización, la nucleasa buscadora de objetivos es una familia LAGLIDADG de nucleasa buscadora de objetivos. En una realización, la familia LAGLIDADG de nucleasa buscadora de objetivos se selecciona de I-Scel, I-Crel, e I-Dmol.

Los agentes nucleasa pueden comprender además endonucleasas de restricción que incluyen endonucleasas Tipo I, Tipo II, Tipo III y Tipo IV. Las endonucleasas de restricción Tipo I y Tipo III reconocen sitios de reconocimiento específicos, pero escinden normalmente en una posición variable a partir del sitio de unión de nucleasa, que puede encontrarse a cientos de pares de bases alejado con respecto al sitio de escisión (sitio de reconocimiento). En los sistemas Tipo II, la actividad de restricción es independiente de cualquier actividad de metilasa, y la escisión ocurre normalmente en sitios específicos en o próximos con respecto al sitio de unión. La mayoría de enzimas Tipo II cortan secuencias palindrómicas, sin embargo, las enzimas Tipo IIa reconocen sitios de reconocimientos no palindrómicos y escinden por fuera del sitio de reconocimiento. Las enzimas Tipo IIb cortan secuencias dos veces con ambos sitios por fuera del sitio de reconocimiento, y las enzimas Tipo IIc reconocen un sitio de reconocimiento asimétrico y escinden en un lado y a una distancia definida de aproximadamente 1-20 nucleótidos a partir del sitio de reconocimiento. Las enzimas de restricción Tipo IV se dirigen a ADN metilado. Las enzimas de restricción se describen y se clasifican de manera adicional, por ejemplo, en la base de datos REBASE (webpage at rebase.neb.com; Roberts et al., (2003) *Nucleic Acids Res* 31:418-20), Roberts et al., (2003) *Nucleic Acids Res* 31:1805-12, y Belfort et al., (2002) in *Mobile DNA II*, pp. 761-783, Eds. Craigie et al., (ASM Press, Washington, DC).

El agente de nucleasa que se emplea en los diversos métodos y composiciones puede comprender además un sistema de repeticiones palindrómicas cortas agrupadas y regularmente interesparadas (CRISPR)/asociado a CRISPR (Cas) o componentes de un sistema como tal. Los sistemas CRISPR/Cas incluyen transcripciones y otros elementos que se involucran en la expresión de, o dirigen la actividad de, genes Cas. Un sistema CRISPR/Cas puede ser un sistema tipo I, uno tipo II, o uno tipo III. Los métodos y composiciones que se divulgan en la presente emplean sistemas CRISPR/Cas mediante la utilización de complejos CRISPR (que comprenden un ARN guía (ARNg) que forma un complejo con una proteína Cas) para escisión de sitio dirigido de ácidos nucleicos.

Algunos sistemas CRISPR/Cas que se usan en los métodos que se divulgan en la presente no ocurren de manera natural. Un sistema "que no ocurre de manera natural" incluye cualquier cosa que indica la implicación de la mano del hombre, tal como uno o más componentes del sistema que se alteran o se mutan a partir de su estado de ocurrencia natural, siendo al menos sustancialmente libres de al menos uno de otro componente con el que se asocian de manera natural en la naturaleza, o asociándose con al menos uno de otro componente con el que no se asocian de manera natural. Por ejemplo, los sistemas CRISPR/Cas emplean complejos CRISPR/Cas que no ocurren de manera natural que comprenden un ARNg y una proteína Cas que no ocurren naturalmente en conjunto.

Las proteínas Cas comprenden de manera general al menos un sitio de reconocimiento de ARN o dominio de unión. Tales dominios pueden interactuar con ARNs guías (ARNgs, que se describen con mayor detalle a continuación). Las proteínas Cas pueden comprender además dominios de nucleasa (por ejemplo, dominios de DNasa y RNasa), dominios de unión a ADN, dominios de helicasa, dominios de interacción proteína-proteína, dominios de dimerización, y otros dominios. Un dominio de nucleasa posee actividad catalítica para escisión de ácido nucleico. La escisión incluye la rotura de los enlaces covalentes de una molécula de ácido nucleico. La escisión puede producir extremos romos o extremos escalonados, y puede ser monocatenaria o bicatenaria.

Ejemplos de proteínas Cas incluyen Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e (CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9 (Csn1 o Csx12), Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (CasA), Cse2 (CasB), Cse3 (CasE), Cse4 (CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, y Cu1966, y homólogos o versiones modificadas de estas.

En algunas instancias, una proteína Cas existe a partir de un sistema CRISPR/Cas tipo II. Por ejemplo, la proteína Cas puede ser una proteína Cas9, o puede derivar a partir de una proteína Cas9. Las proteínas Cas9 comparten normalmente cuatro motivos clave con una arquitectura conservada. Los motivos 1, 2 y 4 son motivos de tipo RuvC, y los motivos 3 son un motivo HNH. La proteína Cas9 puede existir a partir de, por ejemplo, *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus sp.*, *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycoides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas sp.*, *Crocospaera watsonii*, *Cyanotheca sp.*, *Microcystis aeruginosa*, *Synechococcus sp.*, *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccsii*, *Candidatus Desulfurudis*, *Clostridium botulinum*, *Clostridium difficile*, *Finexgoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium vinosum*, *Marinobacter sp.*, *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc sp.*, *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira sp.*, *Lyngbya sp.*, *Microcoleus chthonoplastes*, *Oscillatoria sp.*, *Petrogona mobilis*, *Thermosiphon africanus*, o *Acaryochloris marina*. La proteína Cas9 puede existir a partir de *Staphylococcus aureus*. Ejemplos adicionales de los miembros de la familia de Cas9 incluyen aquellos que se describen en WO 2014/131833. En un ejemplo específico, la proteína Cas9 es una proteína Cas9 a partir de *S. pyogenes* o deriva a partir de esta. La secuencia de aminoácidos de una proteína Cas9 a partir de *S. pyogenes* puede encontrarse, por ejemplo, en la base de datos SwissProt con el número de acceso Q99ZW2.

Las proteínas Cas pueden ser proteínas de tipo salvaje (a saber, aquellas que ocurren en la naturaleza), proteínas Cas modificadas (a saber, variantes de proteína Cas), o fragmentos de proteínas Cas de tipo salvaje o modificadas.

5 Las proteínas Cas pueden ser además variantes activas o fragmentos de proteínas Cas de tipo salvaje o modificadas. Las variantes activas o fragmentos pueden comprender al menos el 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad de secuencia con respecto a la proteína Cas de tipo salvaje o modificada o una porción de esta, en las que las variantes activas retienen la capacidad de cortar en un sitio de escisión conveniente y, por lo tanto, retienen actividad que induce mellas o que induce rotura bicatenaria. Los ensayos para actividad que induce mellas o rotura bicatenaria se conocen y miden, de manera general, la actividad general y especificidad de la proteína Cas en sustratos de ADN que contienen el sitio de escisión.

Las proteínas Cas pueden modificarse para aumentar o reducir la afinidad de unión al ácido nucleico, especificidad de unión al ácido nucleico y/o actividad enzimática. Las proteínas Cas pueden modificarse además para cambiar cualquier otra actividad o propiedad de la proteína, tal como estabilidad. Por ejemplo, uno o más dominios de nucleasa de la proteína Cas pueden modificarse, eliminarse, o inactivarse, o una proteína Cas puede truncarse para retirar dominios que no resultan esenciales para la función de la proteína o para optimizar (por ejemplo, mejorar o reducir) la actividad de la proteína Cas.

20 Algunas proteínas Cas comprenden al menos dos dominios de nucleasa, tales como dominios DNasa. Por ejemplo, una proteína Cas9 puede comprender un dominio de nucleasa de tipo RuvC y un dominio de nucleasa de tipo HNH. Los dominios RuvC y HNH pueden cada uno de ellos cortar una cadena diferente de ADN bicatenario para crear una rotura bicatenaria en el ADN. Véase, por ejemplo, Jinek et al. (2012) *Science* 337:816-821.

25 Uno o ambos de los dominios de nucleasa pueden eliminarse o mutarse de manera tal que dejan de ser funcionales o tienen actividad de nucleasa reducida. Si uno de los dominios de nucleasa se elimina o muta, la proteína Cas resultante (por ejemplo, Cas9) puede denominarse como una nickasa y puede generar una rotura monocatenaria en una secuencia diana en un ADN bicatenario pero no una rotura bicatenaria (a saber, puede escindir la cadena complementaria o la cadena no complementaria, pero no ambas). Si ambos dominios de nucleasa se eliminan o se mutan, la proteína Cas resultante (por ejemplo, Cas9) tendrá capacidad reducida para escindir ambas cadenas de un ADN bicatenario. Un ejemplo de una mutación que convierte Cas9 en una nickasa es una mutación D10A (aspartato a alanina en posición 10 de Cas9) en el dominio RuvC de Cas9 a partir de *S. pyogenes*. Asimismo, H939A (histidina a alanina en posición de aminoácido 839) o H840A (histidina a alanina en posición de aminoácido 840) en el dominio HNH de Cas9 a partir de *S. pyogenes* puede convertir la Cas9 en una nickasa. Otros ejemplos de mutaciones que convierten Cas9 en una nickasa incluyen las mutaciones correspondientes a Cas9 a partir de *S. thermophilus*. Véase, por ejemplo, Sapranaukas et al. (2011) *Nucleic Acids Research* 39:9275-9282 y WO 2013/141680. Tales mutaciones pueden generarse usando métodos que se conocen bien tales como mutagénesis de sitio dirigido, mutagénesis mediada por PCR, o síntesis génica total. Ejemplos de otras mutaciones que crean nickasas pueden encontrarse en, por ejemplo, WO/2013/176772A1 y WO/2013/142578A1.

40 Las proteínas Cas pueden ser además proteínas de fusión. Por ejemplo, una proteína Cas puede fusionarse con un sitio de escisión, un dominio de modificación epigenética, un dominio de activación transcripcional, o un dominio de supresión transcripcional. Véase WO 2014/089290. Las proteínas Cas pueden fusionarse además con un polipéptido heterólogo proporcionando estabilidad aumentada o reducida. El dominio fusionado o polipéptido heterólogo puede ubicarse en el N-terminal, el C-terminal, o internamente en la proteína Cas.

Un ejemplo de una proteína Cas de fusión es una proteína Cas que se fusiona con un polipéptido heterólogo que proporciona una localización subcelular. Tales secuencias pueden incluir, por ejemplo, una señal de localización nuclear (NLS) tal como la NLS SV40 para dirigirse al núcleo, una señal de localización mitocondrial para dirigirse a la mitocondria, una señal de retención ER, y similares. Véase, por ejemplo, Lange et al. (2007) *J. Biol. Chem.* 282:5101-5105. Tales señales de localización subcelular pueden ubicarse en el N-terminal, el C-terminal o en cualquier parte en la proteína Cas. Una NLS puede comprender un tramo de aminoácidos básicos, y puede ser una secuencia monopartita o una secuencia bipartita.

55 Las proteínas Cas pueden comprender además un dominio de penetración de célula. Por ejemplo, el dominio de penetración de célula puede derivar a partir de la proteína HIV-1 TAT, el motivo TLM de penetración de célula a partir del virus humano de hepatitis B, MPG, Pep-1, VP22, un péptido de penetración de célula a partir del virus del herpes simple, o una secuencia de péptidos de poliarginina. Véase, por ejemplo, WO 2014/089290. El dominio de penetración de célula puede ubicarse en el N-terminal, el C-terminal, o en cualquier lugar en la proteína Cas.

60 Las proteínas Cas pueden comprender además un polipéptido heterólogo para facilitar el rastreo o purificación, tal como una proteína fluorescente, una etiqueta de purificación, o una etiqueta de epítipo. Ejemplos de proteínas fluorescentes incluyen proteínas fluorescentes verdes (por ejemplo, GFP, GFP-2, tagGFP, turboGFP, eGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarillas (por ejemplo, YFP, eYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), proteínas fluorescentes azules (por ejemplo, eBFP, eBFP2, Azurite, mKalamal, GFPuv, Sapphire, T-sapphire), proteínas fluorescentes cian (por ejemplo, eCFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes rojas (por ejemplo, mKate, mKate2, mPlum, DsRed monomer,

5 mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), proteínas fluorescentes naranjas (por ejemplo, mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato), y cualquier otra proteína fluorescente adecuada. Ejemplos de etiquetas incluyen glutatión-S-transferasa (GST), proteína de unión a quitina (CPB), proteína de unión a maltosa, tiorredoxina (TRX), poli(NANP), etiqueta de purificación por afinidad en tándem (TAP), myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, hemaglutinina (HA), nus, Softag 1, Softag 3, Strep, SPB, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, histidina (His), proteína transportadora de biotina carboxilasa (BCCP), y calmodulina.

10 Las proteínas Cas pueden proporcionarse en cualquier forma. Por ejemplo, una proteína Cas puede proporcionarse en la forma de una proteína, tales como una proteína Cas que forma un complejo con un ARNg. De manera alternativa, una proteína Cas puede proporcionarse en la forma de un ácido nucleico que codifica la proteína Cas, tal como un ARN (por ejemplo, ARN mensajero (mARN)) o ADN. De manera opcional, el ácido nucleico que codifica la proteína Cas puede ser un codón que se optimiza para traducción eficiente en proteína en una célula u organismo en particular (a saber, una célula humana). Cuando un ácido nucleico que codifica la proteína Cas se introduce en la célula, la proteína Cas puede expresarse de manera transitoria, condicional, o constitutiva en la célula.

20 Los ácidos nucleicos que codifican proteínas Cas pueden integrarse de manera estable en el genoma de la célula y unirse de manera manejable con un promotor activo en la célula. De manera alternativa, los ácidos nucleicos que codifican proteínas Cas pueden unirse de manera manejable con un promotor en una construcción de expresión. Las construcciones de expresión incluyen construcciones de ácido nucleico capaces de dirigir expresión de un gen u otra secuencia de ácido nucleico de interés (por ejemplo, un gen Cas) y que pueden transferir una secuencia de ácido nucleico de interés como tal a una célula diana. Por ejemplo, el ácido nucleico que codifica la proteína Cas puede encontrarse en un vector director que comprende un inserto de ácido nucleico y/o un vector que comprende un ADN que codifica el ARNg. De manera alternativa, puede encontrarse en un vector o un plásmido que se separa con respecto al vector director que comprende el inserto de ácido nucleico y/o por separado con respecto al vector que comprende el ADN que codifica el ARNg. Pueden usarse promotores en una construcción de expresión que incluye, por ejemplo, promotores activos en una célula iPS humana o una célula no pluripotente que se transforma para expresar un estado virgen. Tales promotores pueden ser, por ejemplo, promotores condicionales, promotores inducibles, promotores constitutivos, o promotores de tejido específico.

35 Un "ARN guía" o ARNg" incluye una molécula de ARN que se une con una proteína Cas y dirige la proteína Cas a una ubicación específica en un ADN diana. Los ARNs guías pueden comprender dos segmentos: un "segmento dirigido a ADN" y un "segmento de unión a proteína". "Segmento" incluye un segmento, sección, o región de una molécula, tal como un tramo contiguo de nucleótidos en un ARN. Algunos ARNs comprenden dos moléculas de ARN por separado: un "ARN-activador" y un "ARN-director". Otros ARNs constituyen una molécula de ARN única (polinucleótido de ARN único), que puede denominarse también un "ARNg de molécula única", un "ARN guía único", o un "ARNgs". Véase, por ejemplo, WO/2013/176772A1, WO/2014/065596A1, WO/2014/089290A1, WO/2014/093622A2, WO/2014/099750A2, WO/2013142578A1, y WO 2014/131833A1. Los términos "ARN guía" y "ARNg" son inclusivos, incluyendo tanto ARNs de doble molécula como ARNs de molécula única.

45 Un ARNg de dos moléculas de ejemplo comprende una molécula de tipo crARN ("ARN CRISPR" o "ARN-director" o "crARN" o "repetición de crARN") y una molécula correspondiente de tipo tracrARN ("ARN CRISPR de transactivación" o "ARN-activador" o "tracrARN" o "matriz"). Un crARN comprende tanto el segmento dirigido a ADN (monocatenario) del ARNg como un tramo de nucleótidos que forma una mitad del dúplex de dsARN del segmento de unión a proteína del ARNg.

50 Un tracrARN correspondiente (ARN-activador) comprende un tramo de nucleótidos que forman la otra mitad del dúplex de dsARN del segmento de unión a proteína del ARNg. Un tramo de nucleótidos de un crARN resulta complementario con y se hibrida con un tramo de nucleótidos de un tracrARN para formar el dúplex de dsARN del dominio de unión a proteína del ARNg. Como tal, puede decirse que cada crARN tiene su tracrARN correspondiente.

55 El crARN y el tracrARN correspondiente hibridan para formar un ARNg. El crARN proporciona de manera adicional un segmento dirigido a ADN monocatenario que hibrida con una secuencia diana. Si se usa para modificación en una célula, la secuencia exacta de una molécula de crARN o tracrARN puede designarse para ser específica con respecto a las especies en las que las moléculas de ARN se usarán. Véase, por ejemplo, Mali et al. (2013) Science 339:823-826; Jinek et al. (2012) Science 337:816-821; Hwang et al. (2013) Nat. Biotechnol. 31:227-229; Jiang et al. (2013) Nat. Biotechnol. 31:233-239; y Cong et al. (2013) Science 339:819-823.

60 El segmento dirigido a ADN (crARN) de un ARNg dado comprende una secuencia de nucleótidos que resulta complementaria con respecto a una secuencia en un ADN diana. El segmento dirigido a ADN de un ARNg interactúa con un ADN diana de una manera de secuencia específica mediante hibridación (a saber, apareamiento de bases). Como tal, la secuencia de nucleótidos del segmento dirigido a ADN puede variar y determina la ubicación en el ADN diana con el cual el ARNg y el ADN diana interactuarán. El segmento dirigido a ADN de un ARNg objeto puede modificarse para hibridar con respecto a cualquier secuencia conveniente en un ADN diana. Los crARN que ocurren de manera natural difieren dependiendo del sistema de Cas9 y organismo pero contienen frecuentemente un segmento director de entre 21 a 72 nucleótidos de largo, flanqueado por dos repeticiones directas (DR) de una longitud

de entre 21 a 46 nucleótidos (véase, por ejemplo, WO2014/131833). En el caso de *S. pyogenes*, las DRs son de 36 nucleótidos de largo y el segmento director es de 30 nucleótidos de largo. La DR que se ubica 3' resulta complementaria e híbrida con el tracrARN correspondiente, que a su vez se une con la proteína Cas9.

5 El segmento dirigido a ADN pueden tener una longitud de aproximadamente 12 nucleótidos a aproximadamente 100 nucleótidos, el segmento dirigido a ADN puede tener una longitud de aproximadamente 12 nucleótidos (nt) a aproximadamente 80 nt, de aproximadamente 12 nt a aproximadamente 50 nt, de aproximadamente 12 nt a aproximadamente 40 nt, de aproximadamente 12 nt a aproximadamente 30 nt, de aproximadamente 12 nt a aproximadamente 25 nt, de aproximadamente 12 nt a aproximadamente 20 nt, o de aproximadamente 12 nt a aproximadamente 19 nt, De manera alternativa, el segmento dirigido a ADN puede tener una longitud de aproximadamente 19 nt a aproximadamente 20 nt, de aproximadamente 19 nt a aproximadamente 25 nt, de aproximadamente 19 nt a aproximadamente 30 nt, de aproximadamente 19 nt a aproximadamente 35 nt, de aproximadamente 19 nt a aproximadamente 40 nt, de aproximadamente 19 nt a aproximadamente 45 nt, de aproximadamente 19 nt a aproximadamente 50 nt, de aproximadamente 19 nt a aproximadamente 60 nt, de aproximadamente 19 nt a aproximadamente 70 nt, de aproximadamente 19 nt a aproximadamente 80 nt, de aproximadamente 19 nt a aproximadamente 90 nt, de aproximadamente 19 nt a aproximadamente 100 nt, de aproximadamente 20 nt a aproximadamente 25 nt, de aproximadamente 20 nt a aproximadamente 30 nt, de aproximadamente 20 nt a aproximadamente 35 nt, de aproximadamente 20 nt a aproximadamente 40 nt, de aproximadamente 20 nt a aproximadamente 45 nt, de aproximadamente 20 nt a aproximadamente 50 nt, de aproximadamente 20 nt a aproximadamente 60 nt, de aproximadamente 20 nt a aproximadamente 70 nt, de aproximadamente 20 nt a aproximadamente 80 nt, de aproximadamente 20 nt a aproximadamente 90 nt, o de aproximadamente 20 nt a aproximadamente 100 nt.

25 La secuencia de nucleótidos del segmento dirigido a ADN que resulta complementaria a una secuencia de nucleótidos (secuencia diana) del ADN diana puede tener una longitud de al menos 12 nt. Por ejemplo, la secuencia dirigida a ADN (a saber, la secuencia en el segmento dirigido a ADN que resulta complementario con respecto a una secuencia diana en el ADN diana) puede tener una longitud de al menos aproximadamente 12 nt, al menos aproximadamente 15 nt, al menos aproximadamente 18 nt, al menos aproximadamente 19 nt, al menos aproximadamente 20 nt, al menos aproximadamente 25 nt, al menos aproximadamente 30 nt, al menos aproximadamente 35 nt, o al menos aproximadamente 40 nt. De manera alternativa, la secuencia dirigida a ADN puede tener una longitud de aproximadamente 12 nucleótidos (nt) a aproximadamente 80 nt, de aproximadamente 12 nt a aproximadamente 50 nt, de aproximadamente 12 nt a aproximadamente 45 nt, de aproximadamente 12 nt a aproximadamente 40 nt, de aproximadamente 12 nt a aproximadamente 35 nt, de aproximadamente 12 nt a aproximadamente 30 nt, de aproximadamente 12 nt a aproximadamente 25 nt, de aproximadamente 12 nt a aproximadamente 20 nt, de aproximadamente 12 nt a aproximadamente 19 nt, de aproximadamente 19 nt a aproximadamente 20 nt, de aproximadamente 19 nt a aproximadamente 25 nt, de aproximadamente 19 nt a aproximadamente 30 nt, de aproximadamente 19 nt a aproximadamente 35 nt, de aproximadamente 19 nt a aproximadamente 40 nt, de aproximadamente 19 nt a aproximadamente 45 nt, de aproximadamente 19 nt a aproximadamente 50 nt, de aproximadamente 19 nt a aproximadamente 60 nt, de aproximadamente 20 nt a aproximadamente 25 nt, de aproximadamente 20 nt a aproximadamente 30 nt, de aproximadamente 20 nt a aproximadamente 35 nt, de aproximadamente 20 nt a aproximadamente 40 nt, de aproximadamente 20 nt a aproximadamente 45 nt, de aproximadamente 20 nt a aproximadamente 50 nt, o de aproximadamente 20 nt a aproximadamente 60 nt. En algunos casos, la secuencia dirigida a ADN puede tener una longitud de hasta alrededor 20 nt.

45 Los tracrARNs pueden ser de cualquier forma (por ejemplo, tracrARNs de longitud completa o tracrARNs parcialmente activos) y de diversas longitudes. Estos pueden incluir transcripciones primarias o formas procesadas. Por ejemplo, tracrARNs (como parte de un ARN monocatenario o como una molécula por separado como parte de un ARNg de dos moléculas) puede comprender o consistir en la totalidad o una porción de una secuencia de tracrARN de tipo salvaje (por ejemplo, de aproximadamente o más de aproximadamente 20, 26, 32, 45, 48, 54, 63, 67, 85, o más nucleótidos de una secuencia de tracrARN de tipo salvaje). Ejemplos de secuencias de tracrARN de tipo salvaje de *S. pyogenes* incluyen versiones de 171 nucleótidos, 89 nucleótidos, 75 nucleótidos, y 65 nucleótidos. Véasen, por ejemplo, Deltcheva et al. (2011) Nature 471:602-607; WO 2014/093661. Ejemplos de tracrARNs en ARNs guías únicos (ARNgss) incluyen los segmentos de tracrARN que se encuentran en +48, +54, +67, y +85 versiones de ARNgss, donde "+n" indica que hasta el nucleótido +n de tracrARN de tipo salvaje se incluye en el ARNgss. Véase US 8,697,359.

50 El porcentaje de complementariedad entre la secuencia dirigida a ADN y secuencia diana en el ADN diana puede ser de al menos el 60% (por ejemplo, al menos el 65%, al menos el 70%, al menos el 75%, al menos el 80%, al menos el 90%, al menos el 95%, al menos el 97%, al menos el 98%, al menos el 99%, o el 100%). En algunos casos, el porcentaje de complementariedad entre la secuencia dirigida a ADN y la secuencia diana en el ADN diana es de al menos el 60% en aproximadamente 20 nucleótidos contiguos. En un ejemplo, el porcentaje de complementariedad entre la secuencia dirigida a ADN y la secuencia diana en el ADN diana es del 100% en 14 nucleótidos contiguos en el extremo 5' de la secuencia diana en la cadena complementaria del ADN diana y tan bajo como el 0% en el resto. En un caso como tal, la secuencia dirigida a ADN puede considerarse que es de 14 nucleótidos de largo. En otro ejemplo, el porcentaje de complementariedad entre la secuencia dirigida a ADN y la secuencia diana en el ADN diana es del 100% en los siete nucleótidos contiguos en el extremo 5' de la secuencia diana en la cadena complementaria

del ADN diana y tan bajo como el 0% en el resto. En un caso como tal, la secuencia dirigida a ADN puede considerarse que es de 7 nucleótidos de largo.

5 El segmento de unión a proteína de un ARNg puede comprender dos tramos de nucleótidos que resultan complementarios uno con respecto al otro. Los nucleótidos complementarios del segmento de unión a proteína hibridan para formar un dúplex de ARN bicatenario (dsARN). El segmento de unión a proteína de un ARNg objeto interactúa con una proteína Cas, y el ARNg dirige la proteína Cas de unión a una secuencia de nucleótidos específica en ADN diana mediante el segmento dirigido a ADN.

10 Los ARNs guías pueden incluir modificaciones o secuencias que proporcionan características adicionales convenientes (por ejemplo, estabilidad modificada o regulada; dirección subcelular; rastreo con una etiqueta fluorescente; un sitio de unión para una proteína o complejo de proteínas; y similares). Ejemplos de tales modificaciones incluyen, por ejemplo, una caperuza 5' (por ejemplo, una caperuza 7-metilguanilato (m7G)); una cola 3' poliadenilada (a saber, una cola de poli(A) en 3'); una secuencia de riboswitch (por ejemplo, para permitir la estabilidad regulada y/o accesibilidad regulada mediante proteínas y/o complejos de proteínas); una secuencia de control de estabilidad; una secuencia que forma un dúplex de dsARN (a saber, una horquilla); una modificación o secuencia que dirige el ARN a una ubicación subcelular (por ejemplo, núcleo, mitocondria, cloroplastos, y similares); una modificación o secuencia que se proporciona para rastreo (por ejemplo, conjugación directa con una molécula fluorescente, conjugación con una fracción que facilita la detección fluorescente, una secuencia que permite la detección fluorescente, etc.); una modificación o secuencia que proporciona un sitio de unión para proteínas (por ejemplo, proteínas que actúan en ADN, incluyendo activadores transcripcionales, supresores transcripcionales, metiltransferasas de ADN, demetilasas de ADN, histonas acetiltransferasas, histonas deacetilasas, y similares); y combinaciones de estos.

25 Los ARNs guías pueden proporcionarse en cualquier forma. Por ejemplo, el ARNg puede proporcionarse en la forma de ARN, ya sea como dos moléculas (crARN y tracrARN por separado) o como una molécula (ARNgs), y, de manera opcional, en la forma de un complejo con una proteína Cas. El ARNg puede proporcionarse además en la forma de ADN que codifica el ARN. El ADN que codifica el ARNg puede codificar una molécula de ARN única (sgARN) o moléculas de ARN por separado (por ejemplo, crARN y tracrARN por separado). En el último caso, el ADN que codifica el ARNg puede proporcionarse como moléculas de ADN por separado que codifican el crARN y tracrARN, respectivamente.

35 Cuando un ADN que codifica un ARNg se introduce en la célula, el ARNg puede expresarse de manera transitoria, condicional, o constitutiva en la célula. Los ADNs que codifican ARNs pueden integrarse de manera estable en el genoma de la célula y unirse de manera manejable con un promotor activo en la célula. De manera alternativa, los ADNs que codifican ARNs pueden unirse de manera manejable con un promotor en una construcción de expresión. Por ejemplo, el ADN que codifica el ARNg puede encontrarse en un vector director que comprende un inserto de ácido nucleico y/o un vector que comprende un ácido nucleico que codifica una proteína Cas. De manera alternativa, puede ser un vector o un plásmido que se separa a partir del vector director que comprende el inserto de ácido nucleico y/o se separa a partir del vector que comprende el ácido nucleico que codifica la proteína Cas. Tales promotores pueden encontrarse activos, por ejemplo, en una célula iPS humana o una célula no pluripotente que se transforma para expresar un estado pluripotente. Tales promotores puede ser, por ejemplo, promotores condicionales, promotores inducibles, promotores constitutivos, o promotores de tejido específico. En algunos casos, el promotor es un promotor de ARN polimerasa III, tal como un promotor U6 humano.

50 De manera alternativa, pueden prepararse ARNs mediante diversos métodos diferentes. Por ejemplo, pueden prepararse ARNs mediante transcripción *in vitro* usando, por ejemplo, ARN polimerasa T7 (véanse, por ejemplo, WO 2014/089290 y WO 2014/065596). Los ARNs guías pueden ser además una molécula que se produce sintéticamente que se prepara mediante síntesis química.

Una secuencia diana para un sistema CRISPR/Cas incluye secuencias de ácido nucleico que se presentan en un ADN diana al cual un segmento dirigido a ADN de un ARNg se unirá, siempre que existan condiciones suficientes para unión. Por ejemplo, secuencias diana incluyen secuencias a las que un ARN guía se designa para tener complementariedad, donde la hibridación entre una secuencia diana y una secuencia dirigida a ADN promueve la formación de un complejo CRISPR. No se requiere necesariamente de complementariedad completa, siempre que exista complementariedad suficiente para originar hibridación y promover la formación de un complejo CRISPR. Las secuencias diana incluyen además sitios de escisión para proteínas Cas, que se describen con mayor detalle a continuación. Una secuencia diana puede comprender cualquier polinucleótido, que puede ubicarse, por ejemplo, en el núcleo o citoplasma de una célula o en un orgánulo de una célula, tal como una mitocondria o cloroplasto.

65 La secuencia diana en un ADN diana puede dirigirse por (a saber, unirse con, o hibridar con, o ser complementaria con) una proteína Cas o un ARNg. Condiciones de unión a ADN/ARN adecuadas incluyen condiciones fisiológicas que se presentan normalmente en una célula. Otras condiciones de unión a ADN/ARN adecuadas (por ejemplo, condiciones en un sistema libre de células) se conocen en la técnica (véase, por ejemplo, Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook et al., Harbor Laboratory Press 2001)). La cadena del ADN diana que resulta complementaria con e hibrida con la proteína Cas o ARNg puede denominarse la "cadena complementaria", y la

cadena del ADN diana que resulta complementaria con la "cadena complementaria" (y no resulta, por lo tanto, complementaria con la proteína Cas o ARNg) puede denominarse "cadena no complementaria" o "cadena modelo".

5 La proteína Cas puede escindir el ácido nucleico en un sitio o por fuera de una secuencia de ácido nucleico que se presenta en un ADN diana a la que el segmento dirigido a ADN o un ARNg se unirá. El "sitio de escisión" incluye la posición de un ácido nucleico en la que una proteína Cas produce una rotura monocatenaria o una rotura bicatenaria. Por ejemplo, la formación de un complejo CRISPR (que comprende un ARNg hibridado con una secuencia diana y que forma un complejo con una proteína Cas) puede dar como resultado la escisión de una o ambas cadenas en o  
10 próxima a (por ejemplo, en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50, o más pares de bases de) la secuencia de ácido nucleico presente en un ADN diana a la que un segmento dirigido a ADN de un ARNg se unirá. Si el sitio de escisión se encuentra por fuera de la secuencia de ácido nucleico presente en un ADN diana a la que un segmento dirigido a ADN o un ARNg se unirá, el sitio de escisión se considera todavía que se encuentra en la "secuencia diana". El sitio de escisión puede encontrarse en una cadena sola o en ambas cadenas de un ácido nucleico. Los sitios de escisión  
15 pueden encontrarse en la misma posición en ambas cadenas del ácido nucleico (produciendo extremos romos) o pueden encontrarse en sitios diferentes en cada cadena (produciendo extremos escalonados). Los extremos escalonados pueden producirse, por ejemplo, mediante el uso de dos proteínas Cas que producen una rotura monocatenaria en diferentes sitios de escisión en cada cadena. Por ejemplo, una primera nickasa puede crear una rotura monocatenaria en la primera cadena de ADN bicatenario (dsDNA), mientras que una segunda nickasa puede crear una rotura monocatenaria en la segunda cadena de dsDNA de manera tal que se crean secuencias sobresalientes. En algunos casos, la secuencia diana de la nickasa en la primera cadena se separa de la secuencia diana de la nickasa en la segunda cadena por al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 75, 100, 250, 500, o 1.000 bases de pares.

25 La escisión de sitio específico de ADN diana mediante Cas9 puede ocurrir en ubicaciones que se determinan por ambos (i) complementariedad de pares de bases entre el ARNg y el ADN diana y (ii) un motivo corto, que se denomina el motivo adyacente de protoespaciador (PAM), en el ADN diana. El PAM puede flanquear la secuencia diana. De manera opcional, la secuencia diana puede flanquearse en el extremo 3' por el PAM. Por ejemplo, el sitio de escisión de Cas9 puede ser de aproximadamente 1 a aproximadamente 10 o de aproximadamente 2 a aproximadamente 5  
30 pares de bases (por ejemplo, 3 pares de bases) hacia arriba o hacia abajo con respecto a la secuencia de PAM. En algunos casos, (por ejemplo, cuando se usan una Cas9 de *S. pyogenes* o una Cas9 estrechamente relacionada), la secuencia de PAM de la cadena no complementaria puede ser 5'-XGG-3', donde X es cualquier nucleótido de ADN y se encuentra inmediatamente 3' de la secuencia diana de la cadena no complementaria del ADN diana. Como tal, la secuencia PAM de la cadena complementaria sería 5'-CCY-3', donde Y es cualquier nucleótido de ADN y se encuentra inmediatamente 5' de la secuencia diana de la cadena complementaria del ADN diana. En algunos casos como tales, X e Y puede ser complementarios y el par de bases X-Y puede ser cualquier par de bases (por ejemplo, X=C y Y=G; X=G y Y=C; X=A y Y=T, X=T, y Y=A).

40 Ejemplos de secuencias diana incluyen una secuencia de ADN complementaria con el segmento dirigido a ADN de un ARNg, o una secuencia de ADN como tal que se agrega a una secuencia de PAM. Un ejemplo de una secuencia diana comprende la secuencia de nucleótidos GNNNNNNNNNNNNNNNNNNNGG (GN<sub>1-20</sub>GG; SEQ ID NO: 2). Otras secuencias diana puede tener entre 4 y 22 nucleótidos de largo de SEQ ID NO: 2, incluyendo el 5' G y el 3' GG. Inclusive otras secuencias diana pueden tener entre 14 y 20 nucleótidos de largo de SEQ ID NO: 2.

45 La secuencia diana puede ser cualquier secuencia de ácido nucleico endógeno o exógeno con respecto a una célula. La secuencia diana puede ser una secuencia que codifica un producto génico (por ejemplo, una proteína) o una secuencia no codificadora (por ejemplo, una secuencia reguladora o ADN basura) o puede incluir ambos.

50 Variantes activas y fragmentos de agentes de nucleasa (a saber, un agente de nucleasa diseñado) se proporcionan también. Tales variantes activas pueden comprender al menos el 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad de secuencia con respecto al agente de nucleasa nativo, en las que las variantes activas retienen la capacidad de cortar en un sitio de reconocimiento conveniente y retener, por lo tanto, actividad de inducir mella o rotura bicatenaria. Por ejemplo, cualquiera de los agentes de nucleasa que se describen en la presente pueden modificarse a partir de una secuencia de endonucleasa nativa y diseñarse para  
55 reconocer e inducir una mella o rotura bicatenaria en un sitio de reconocimiento que no se reconoció por el agente de nucleasa nativo. De este modo, en algunas realizaciones, la nucleasa diseñada tiene una especificidad para inducir una mella o rotura bicatenaria en un sitio de reconocimiento que resulta diferente con respecto al sitio de reconocimiento del agente de nucleasa nativo. Los ensayos para actividad que induce mella o rotura bicatenaria se conocen y miden, de manera general, la actividad y especificidad generales de la endonucleasa en sustratos de ADN que contienen el sitio de reconocimiento.

60 El agente de nucleasa puede introducirse en la célula pluripotente mediante cualquier método que se conoce en la técnica. El polipéptido que codifica el agente de nucleasa puede introducirse directamente en la célula. De manera alternativa, un polinucleótido que codifica el agente de nucleasa puede introducirse en la célula. Cuando un polinucleótido que codifica el agente de nucleasa se introduce en la célula, el agente de nucleasa puede expresarse de manera transitoria, condicional o constitutiva. De este modo, el polinucleótido que codifica el agente de nucleasa puede contenerse en un cassette de expresión y unirse de manera manejable con un promotor condicional, un

promotor inducible, un promotor constitutivo, o un promotor de tejido específico. De manera alternativa, el agente de nucleasa se introduce en la célula como un mRNA que codifica un agente de nucleasa.

- 5 En realizaciones específicas, el polinucleótido que codifica el agente de nucleasa se integra de manera estable en el genoma de la célula pluripotente y se une de manera manejable con un promotor activo en la célula. En otras realizaciones, el polinucleótido que codifica el agente de nucleasa se encuentra en el mismo vector director que comprende el inserto de ácido nucleico, mientras que en otros casos, el polinucleótido que codifica el agente de nucleasa se encuentra en un vector o un plásmido que se encuentra por separado con respecto al vector director que comprende el inserto de ácido nucleico.

10 Cuando el agente de nucleasa se proporciona a la célula pluripotente a través de la introducción de un polinucleótido que codifica el agente de nucleasa, un polinucleótido como tal que codifica un agente de nucleasa puede modificarse para sustituir codones que tienen una mayor frecuencia de uso en la célula de interés, en comparación con la secuencia de polinucleótidos que ocurre de manera natural que codifica el agente de nucleasa. Por ejemplo, el polinucleótido que codifica el agente de nucleasa puede modificarse para sustituir codones que tienen una frecuencia mayor de uso en una célula humana, en comparación con la secuencia de polinucleótidos que ocurre de manera natural.

#### 20 b. Marcadores de selección

Diversos marcadores de selección pueden usarse en los métodos y composiciones que se divulgan en la presente que se proporcionan para modificar un locus genómico diana en un cromosoma. Tales marcadores se divulgan en otra parte en la presente e incluyen, pero sin limitación, marcadores de selección que imparten resistencia a un antibiótico tal como G418, higromicina, blasticidina, neomicina, o puromicina. El polinucleótido que codifica los marcadores de selección se une de manera manejable con un promotor activo en una célula iPS humana o una célula no pluripotente que se transforma para expresar un estado virgen.

#### 30 c. Locus genómico diana

Se proporcionan diversos métodos y composiciones que permiten la integración de al menos un inserto de ácido nucleico en un locus genómico diana en un cromosoma. Un "locus genómico diana en un cromosoma" comprende cualquier segmento o región de ADN en un cromosoma a la que se desea integrar un inserto de ácido nucleico. El locus genómico en un cromosoma siendo dirigido puede ser nativo con respecto a células iPS humanas o una célula no pluripotente que se transforma para expresar un estado pluripotente, o, de manera alternativa, puede comprender un segmento de ADN heterólogo o exógeno que se integró en un cromosoma de la célula. Tales segmentos de ADN heterólogos o exógenos pueden incluir transgenes, cassettes de expresión, polinucleótidos que codifican marcadores de selección, o regiones de ADN genómico heterólogo o exógeno. El locus genómico diana en el cromosoma puede comprender cualquiera del sistema de integración genómica dirigido incluyendo, por ejemplo, el sitio de reconocimiento, el marcador de selección, insertos de ácido nucleico previamente integradas, agentes de nucleasa que codifican polinucleótidos, promotores, etc. De manera alternativa, el locus genómico diana en el cromosoma puede localizarse en un cromosoma artificial de levadura (YAC), cromosoma artificial de bacteria (BAC), un cromosoma artificial de humano, o cualquier otra región genómica diseñada contenida en una célula huésped adecuada. De este modo, en realizaciones específicas, el locus genómico dirigido en el cromosoma puede comprender secuencia genómica nativa de una célula humana o secuencia de ácido nucleico heterólogo o exógeno de un mamífero no humano, una célula no humana, un roedor, un humano, una rata, un ratón, un hámster, un conejo, un cerdo, un bovino, un ciervo, una oveja, una cabra, un pollo, un gato, un perro, un hurón, un primate (por ejemplo, mono tífi, Rhesus), mamífero domesticado o un mamífero agrícola, o cualquier otro organismo de interés o una combinación de estos.

#### 50 d. Vectores directores e insertos de ácidos nucleicos

Según se indicó anteriormente, los métodos y composiciones que se proporcionan en la presente emplean vectores directores solos o en combinación con un agente de nucleasa. "Recombinación homóloga" se usa de manera convencional para hacer referencia al intercambio de fragmentos de ADN entre dos moléculas de ADN en sitios de entrecruzamiento en las regiones de homología.

#### i. Inserto de ácido nucleico

60 Una o más de insertos de ácido nucleico por separado pueden emplearse en los métodos que se divulgan en la presente, y pueden introducirse en una célula mediante vectores directores por separado o en el mismo vector director. Los insertos de ácido nucleico incluyen segmentos de ADN para integrarse en loci genómicos diana. La integración de un inserto de ácido nucleico en un locus diana puede dar como resultado una adición de una secuencia de ácido nucleico de interés con respecto al locus diana, eliminación de una secuencia de ácido nucleico de interés en el locus diana, y/o reemplazo de una secuencia de ácido nucleico de interés en el locus diana.

El inserto de ácido nucleico o el ácido nucleico correspondiente en el locus diana que se reemplaza puede ser una región de codificación, un intrón, un exón, una región no traducida, una región reguladora, un promotor, un potenciador, o cualquier combinación de estos. Además, el inserto de ácido nucleico o el ácido nucleico correspondiente en el locus diana que se reemplaza puede ser de cualquier longitud conveniente, incluyendo, por ejemplo, entre 10-100 nucleótidos de largo, 100-500 nucleótidos de largo, 500 nucleótidos-1 kb de largo, 1 kb a 1,5 kb de nucleótidos de largo, 1,5 kb a 2 kb de nucleótidos de largo, 2 kb a 2,5 kb de nucleótidos de largo, 2,5 kb a 3 kb de nucleótidos de largo, 3 kb a 5 kb de nucleótidos de largo, 5 kb a 8 kb de nucleótidos de largo, 8 kb a 10 kb de nucleótidos de largo o más. En otros casos, la longitud puede ser de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, de aproximadamente 350 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 800 kb, de aproximadamente 800 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, de aproximadamente 2,5 Mb a aproximadamente 2,8 Mb, de aproximadamente 2,8 Mb a aproximadamente 3 Mb. En inclusive otros casos, la longitud puede ser de al menos 100, 200, 300, 400, 500, 600, 700, 800, o 900 nucleótidos o de al menos 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 15 kb, 16 kb o mayor.

En algunos vectores directores, el inserto de ácido nucleico puede ser de aproximadamente 5 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, de aproximadamente 190 kb a aproximadamente 200 kb. De manera alternativa, el inserto de ácido nucleico puede ser de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.

En algunos casos, el reemplazo del ácido nucleico en el locus diana da como resultado la eliminación de una secuencia diana que varía de aproximadamente 1 kb a aproximadamente 200 kb, de aproximadamente 2 kb a aproximadamente 20 kb, o de aproximadamente 0,5 kb a aproximadamente 3 Mb. En algunos, casos, la magnitud de la eliminación resulta mayor que una longitud total del brazo de homología 5' y el brazo de homología 3'.

En algunos casos, la magnitud de la eliminación de la secuencia diana varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 70 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, de aproximadamente 190 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, de aproximadamente 350 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 800 kb, de aproximadamente 800 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, de aproximadamente 2,5 Mb a aproximadamente 2,8 Mb, de aproximadamente 2,8 Mb a aproximadamente 3 Mb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb.

En otros casos, el inserto de ácido nucleico o el ácido nucleico correspondiente en el locus diana que se reemplaza puede ser de al menos 10 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, al menos 300 kb, al menos 350 kb, al menos 400 kb, al menos 450 kb, o al menos 500 kb o mayor.

El inserto de ácido nucleico puede comprender ADN genómico o cualquier otro tipo de ADN. Por ejemplo, el inserto de ácido nucleico puede ser una procariota, una eucariota, una levadura, un ave (por ejemplo, pollo), un mamífero no humano, un roedor, un humano, una rata, un ratón, un hámster, un conejo, un cerdo, un bovino, un ciervo, una oveja, una cabra, un gato, un perro, un hurón, un primate (por ejemplo, mono tití, Rhesus), un mamífero domesticado, un mamífero agrícola, o cualquier otro organismo de interés. En un ejemplo, el polinucleótido de inserción puede comprender cualquier locus humano o no humano.

El inserto de ácido nucleico y/o ácido nucleico en el locus diana puede comprender una secuencia de codificación o una secuencia de no codificación, tal como un elemento regulador (por ejemplo, un promotor, un potenciador, o un elemento de unión a supresor transcripcional). Por ejemplo, el inserto de ácido nucleico puede comprender un alelo con reemplazo de al menos un exón de un gen endógeno, o un alelo con reemplazo del gen endógeno completo (a saber, "reemplazo para intercambio de genes").

El inserto de ácido nucleico puede comprender además un alelo condicional. El alelo condicional puede ser un alelo multifuncional, según se describe en US 2011/0104799. Por ejemplo, el alelo condicional puede comprender: (a) una secuencia de actuación en orientación sentido con respecto a la transcripción de un gen diana, y (b) un cassette de selección de medicamento (DSC) en orientación sentido o antisentido; (c), una secuencia de nucleótidos de interés (NSI) en orientación antisentido; y (d) un condicional por módulo de inversión (COIN, que utiliza un intrón que separa exones y un módulo de tipo trampa génica invertible) en orientación inversa. Véase, por ejemplo, US 2011/0104799. El alelo condicional puede comprender además unidades que pueden recombinarse que se recombinan al exponerse a una primera recombinasa para formar un alelo condicional que (i) carece de la secuencia de actuación y el DSC; y (ii) contiene el NSI en orientación sentido y el COIN en orientación antisentido. Véase US 2011/0104799.

Algunas inserciones de ácido nucleico comprenden un polinucleótido que codifica un marcador de selección. El marcador de selección puede contenerse en un cassette de selección. Tales marcadores de selección incluyen, pero sin limitación, una neomicina fosfotransferasa (neo<sup>r</sup>), higromicina B fosfotransferasa (hyg<sup>r</sup>), puromicina-N-acetiltransferasa (puro<sup>r</sup>), blastidina deaminasa S (bsr<sup>r</sup>), xantina/guanina fosforribosil transferasa (gpt), o quinasa timidina de virus de herpes simple (HSV-k), o una combinación de estos. El polinucleótido que codifica el marcador de selección puede unirse de manera manejable con un promotor activo en una célula siendo dirigida.

En algunos vectores directores, el inserto de ácido nucleico comprende un gen reportero. Ejemplos de genes reporteros son genes que codifican luciferasa, β-galactosidasa, proteína fluorescente verde (GFP), proteína fluorescente verde mejorada (eGFP), proteína fluorescente cian (CFP), proteína fluorescente amarilla (YFP), proteína fluorescente amarilla mejorada (eYFP), proteína fluorescente azul (BFP), proteína fluorescente azul (eBFP), DsRed, ZsGreen, MmGFP, mPlum, mCherry, tdTomato, mStrawberry, J-Red, mOrange, mKO, mCitrine, Venus, YPet, Emerald, CyPet, Cerulean, T-Sapphire, fosfatasa alcalina, y una combinación de estos. Tales genes reporteros pueden unirse de manera manejable con un promotor activo en una célula siendo dirigida.

En algunos vectores directores, el inserto de ácido nucleico comprende uno o más cassettes de expresión o cassettes de eliminación. Un cassette dado puede comprender una secuencia de nucleótidos de interés, un ácido nucleico que codifica un marcador de selección y/o un gen reportero, junto con diversos componentes reguladores que influyen en la expresión. Ejemplos de marcadores que pueden seleccionarse y genes reporteros que pueden incluirse se analizan con mayor detalle en otra parte en la presente.

En algunos vectores directores, el inserto de ácido nucleico comprende un ácido nucleico flanqueado con secuencias diana de recombinación de sitio específico. A pesar de que el inserto de ácido nucleico entero puede flanquearse por tales secuencias diana de recombinación de sitio específico, cualquier región o polinucleótido individual de interés en el inserto de ácido nucleico puede flanquearse además por tales sitios. Las secuencias diana de recombinación de sitio específico, que pueden flanquear el inserto de ácido nucleico o cualquier polinucleótido de interés en el inserto de ácido nucleico pueden incluir, por ejemplo, loxP, lox511, lox2272, lox66, lox71, loxM2, lox5171, FRT, FRT11, FRT71, attP, att, FRT, rox, o una combinación de estos. En un ejemplo, los sitios de recombinación de sitio específico flanquean un polinucleótido que codifica un marcador de selección y/o un gen reportero contenido en el inserto de ácido nucleico. A continuación de la integración del inserto de ácido nucleico en un locus dirigido, las secuencias entre los sitios de recombinación de sitio específico pueden retirarse.

## ii. Vectores directores

Los vectores directores pueden emplearse para introducir el inserto de ácido nucleico en un locus genómico diana y comprenden el inserto de ácido nucleico y brazos de homología que flanquean el inserto de ácido nucleico. Los vectores directores pueden ser de forma lineal o de forma circular, y pueden ser monocatenarios o bicatenarios. Para

facilitar la referencia, los brazos de homología se denominan en la presente como brazos de homología 5' y 3' (a saber, hacia arriba y hacia abajo). Esta terminología se refiere a la posición relativa de los brazos de homología con respecto al inserto de ácido nucleico en el vector director. Los brazos de homología 5' y 3' corresponden a regiones en el locus dirigido, que se denominan en la presente como "secuencia diana 5'" y "secuencia diana 3'", respectivamente.

Un brazo homólogo y una secuencia diana "corresponden" o "se corresponden" uno con respecto al otro cuando las dos regiones comparten un nivel suficiente de identidad de secuencia una con respecto a la otra para actuar como sustratos para una reacción de recombinación homóloga. El término "homología" incluye secuencias de ADN que son ya sea idénticas o comparten identidad de secuencia con una secuencia correspondiente. La identidad de secuencia entre una secuencia diana dada y el brazo de homología correspondiente que se encuentra en el vector director puede ser de cualquier grado de identidad de secuencia que permite que la recombinación homóloga ocurra. Por ejemplo, la cantidad de identidad de secuencia compartida por el brazo de homología del vector director (o un fragmento de este) y la secuencia diana (o un fragmento de esta) puede ser de identidad de secuencia de al menos el 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de identidad de secuencia, de manera tal que las secuencias se someten a recombinación homóloga. Además, una región correspondiente de homología entre el brazo de homología y la secuencia diana correspondiente puede ser de cualquier longitud que resulta suficiente para promover recombinación homóloga en el sitio de reconocimiento escindido. Por ejemplo, un brazo de homología dado y/o secuencia diana correspondiente pueden comprender regiones correspondientes de homología que son al menos de aproximadamente 5-10 kb, 5-15 kb, 5-20 kb, 5-25 kb, 5-30 kb, 5-35 kb, 5-40 kb, 5-45 kb, 5-50 kb, 5-55 kb, 5-60 kb, 5-65 kb, 5-70 kb, 5-75 kb, 5-80 kb, 5-85 kb, 5-90 kb, 5-95 kb, 5-100 kb, 100-200 kb, o 200-300 kb de largo o más (tal como se describe en los vectores LTVEC que se describen en otra parte en la presente) de manera tal que el brazo de homología tiene suficiente homología para someterse a recombinación homóloga con las secuencias diana correspondientes en el genoma de la célula.

Los brazos de homología pueden corresponder a un locus que resulta nativo con respecto a una célula (por ejemplo, el locus dirigido), o de manera alternativa, pueden corresponder a una región de un segmento de ADN heterólogo o exógeno que se integra en el genoma de la célula, incluyendo, por ejemplo, transgenes, cassettes de expresión, o regiones de ADN heterólogo o exógeno. De manera alternativa, los brazos de homología del vector director pueden corresponder a una región de un cromosoma artificial de levadura (YAC), un cromosoma artificial de bacteria (BAC), un cromosoma artificial de humano, o cualquier otra región diseñada contenida en una célula huésped adecuada. Aún más, los brazos de homología del vector director pueden corresponder o derivar a partir de una región de una librería BAC, una librería cósmida, o una librería de fago P1. En ciertos casos, los brazos de homología del vector director corresponden a un locus que resulta nativo, heterólogo, o exógeno con respecto a una célula iPS humana o una célula no pluripotente que se transforma para expresar un estado virgen. En algunos casos, los brazos de homología corresponden a un locus de la célula que no se puede dirigir usando un método convencional o que puede dirigirse solo de manera incorrecta o solo con eficiencia significativamente baja en la ausencia de una mella o rotura bicatenaria inducida por un agente de nucleasa (por ejemplo, una proteína Cas). En algunos casos, los brazos de homología derivan a partir de ADN sintético.

En algunos vectores directores, los brazos de homología 5' y 3' corresponden a un locus seguro. Interacciones entre ADN exógeno integrado y un genoma huésped pueden limitar la confiabilidad y seguridad de integración y pueden conducir a poner en evidencia efectos fenotípicos que no se deben a la modificación genética dirigida pero que en cambio se deben a efectos no previstos de la integración en genes endógenos circundantes. Por ejemplo, transgenes insertados al azar pueden someterse a efectos de posición y silenciamiento, lo que vuelve su expresión no confiable y no previsible. Asimismo, la integración de ADN exógeno en un locus cromosómico puede afectar genes endógenos circundantes y cromatina, alterando así el comportamiento celular y fenotipos. Los loci seguros incluyen loci cromosómicos donde transgenes u otros insertos de ácidos nucleicos exógenos pueden expresarse de manera estable y confiable en todos los tejidos de interés sin alterar abiertamente el comportamiento celular o fenotipo. Vease, por ejemplo, Sadelain et al. (2012) Nat. Rev. Cancer 12:51-58. Por ejemplo, los loci seguros pueden incluir loci cromosómicos donde ADN exógeno puede integrarse y funcionar de una manera predecible sin afectar de manera adversa la estructura o expresión de gen endógeno. Los loci seguros pueden incluir regiones extragénicas o regiones intragénicas tales como, por ejemplo, loci sin genes que no son esenciales, prescindibles, o capaces de alterarse sin poner en evidencia consecuencias fenotípicas.

Por ejemplo, el locus *Rosa26* y su equivalente en humanos ofrece una configuración de cromatina abierta en todos los tejidos y se expresa de manera ubicua durante desarrollo embrionario y en adultos. Véase Zambrowicz et al. (1997) Proc. Natl. Acad. Sci. USA 94:3789-3794. De manera adicional, el locus *Rosa26* puede dirigirse con alta eficiencia, y la interrupción del gen *Rosa26* no produce fenotipo evidente. Otro ejemplo, de un locus adecuado es el locus *Ch25h*.

Un brazo de homología de un vector director puede ser de cualquier longitud que resulta suficiente para promover un evento de recombinación homóloga con una secuencia diana correspondiente, incluyendo, por ejemplo, al menos, 5-10 kb, 5-15 kb, 5-20 kb, 5-25 kb, 5-30 kb, 5-35 kb, 5-40 kb, 5-45 kb, 5-50 kb, 5-55 kb, 5-60 kb, 5-65 kb, 5-70 kb, 5-75 kb, 5-80 kb, 5-85 kb, 5-90 kb, 5-95 kb, 5-100 kb, 100-200 kb, o 200-300 kb de largo o mayor. Según se describe con mayor detalle a continuación, los vectores directores grandes pueden emplear brazos directores de mayor longitud.

Los agentes de nucleasas (por ejemplo, sistemas CRISPR/Cas) pueden emplearse en combinación con vectores directores para asistir en la modificación de un locus diana. Tales agentes de nucleasas pueden promover recombinación homóloga entre el vector director y el locus diana. Tales agentes de nucleasas, se emplean en combinación con un vector director, el vector director puede comprender brazos de homología 5' y 3' correspondientes a secuencias diana 5' y 3' que se ubican lo suficientemente próximas con respecto a un sitio de escisión de nucleasa de manera tal que se promueve la ocurrencia de un evento de recombinación homóloga entre las secuencias diana y los brazos de homología al momento de una mella o rotura bicatenaria en el sitio de escisión de nucleasa. El término "sitio de escisión nucleasa" incluye una secuencia ADN en la que se crea una mella o rotura bicatenaria mediante un agente de nucleasa (por ejemplo, un sitio de escisión de Ca9). Las secuencias diana en el locus dirigido que corresponden a brazos de homología 5' y 3' del vector director se "ubican lo suficientemente próximas" con respecto a un sitio de escisión de nucleasa si la distancia resulta tal como para promover la ocurrencia de un evento de recombinación homóloga entre las secuencias diana 5' y 3' y los brazos de homología al momento de una mella o rotura bicatenaria en el sitio de reconocimiento. De este modo, en casos específicos, las secuencias diana que corresponden a los brazos de homología 5' y/o 3' del vector director se encuentran en al menos 1 nucleótido de un sitio de reconocimiento dado o se encuentran en al menos 10 nucleótidos a aproximadamente 14 kb de un sitio de reconocimiento dado. En algunos casos, el sitio de escisión de nucleasa, se encuentra inmediatamente adyacente con respecto a al menos, una o ambas de las secuencias diana.

La relación espacial de las secuencias diana que corresponden a los brazos de homología del vector director y el sitio de escisión de nucleasa puede variar. Por ejemplo, las secuencias diana pueden localizarse 5' con respecto al sitio de escisión de nucleasa, las secuencias diana pueden localizarse 3' con respecto al sitio de reconocimiento, o las secuencias diana pueden flanquear el sitio de escisión de nucleasa.

El uso combinado del vector director (incluyendo, por ejemplo, un vector director grande) con un agente de nucleasa puede dar como resultado una eficacia dirigida aumentada en comparación con el uso del vector director por sí solo. Por ejemplo, cuando un vector director se usa en conjunto con un agente de nucleasa, la eficiencia dirigida del vector director puede aumentarse al menos dos veces, al menos, tres veces, al menos 4 veces, o al menos 10 veces cuando se compara con el uso del vector director por sí solo.

### iii. Vectores directores grandes

Algunos vectores directores son "vectores directores grandes" o "LTVECs", que incluyen vectores directores que comprenden brazos de homología que corresponden y derivan a partir de secuencias de ácidos nucleicos más grandes con respecto a aquellas que se usan normalmente por otros enfoques previstos para realizar recombinación homóloga en células. Ejemplos de generación de modificaciones genéticas dirigidas usando LTVECs se divulgan en, por ejemplo, WO 2015/088643, US 2015/0159175, US 2015/0159174, US 2014/0310828, US 2014/0309487, y US 2013-0309670. Los LTVECs incluyen además vectores directores que comprenden insertos de ácidos nucleicos que tienen secuencias de ácidos nucleicos más grandes con respecto a aquellas que se usan normalmente por otros enfoques previstos para realizar recombinación homóloga en células. Por ejemplo, los LTVECs posibilitan la modificación de grandes loci que no pueden acomodarse mediante vectores directores basados en plásmidos tradicionales debido a sus limitaciones de tamaño. Por ejemplo, el locus dirigido puede ser (a saber, los brazos de homología 5' y 3' pueden corresponderse con) un locus de la célula que no se puede dirigir usando un método convencional o que puede dirigirse solo de manera incorrecta o solo con eficiencia significativamente baja en la ausencia de una mella o rotura bicatenaria inducida por un agente de nucleasa (por ejemplo, una proteína Cas).

Ejemplos de LTVECs incluyen vectores que derivan a partir de un cromosoma artificial de bacteria (BAC), un cromosoma artificial de humano, o un cromosoma artificial de levadura (YAC). Ejemplos no limitantes de LTVECs y métodos para realizarlos se describen en, por ejemplo, US Pat. Nro. 6,586,251; US Pat. Nro. 6,596,541; US Pat. Nro. 7,105,348; y WO 2002/036789 (PCT/US01/45375). Los LTVECs pueden ser de forma lineal o de forma circular.

Los LTVECs pueden ser de cualquier longitud, incluyendo, por ejemplo, de aproximadamente 50 kb a aproximadamente 300 kb, de aproximadamente 50 kb a aproximadamente 75 kb, de aproximadamente 75 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 125 kb, de aproximadamente 125 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 175 kb, de aproximadamente 175 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 225 kb, de aproximadamente 225 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 275 kb o de aproximadamente 275 kb a aproximadamente 300 kb. De manera alternativa, un LTVEC puede ser de al menos 10 kb, al menos 15 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, al menos 300 kb, al menos 350 kb, al menos 400 kb, al menos 450 kb, o al menos 500 kb o mayor. El tamaño de un LTVEC puede ser demasiado grande para permitir análisis de eventos dirigidos mediante ensayos convencionales, por ejemplo, southern blotting y PCR de largo alcance (por ejemplo, 1 kb a 5 kb).

En algunos casos, un LTVEC comprende un inserto de ácido nucleico que varía de aproximadamente 5 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a

5 aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, o de aproximadamente 190 kb a aproximadamente 200 kb. En otros casos, el inserto de ácido nucleico puede variar de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb.

20 En algunos LTVECs, la suma total del brazo de homología hacia arriba y el brazo de homología hacia abajo es de al menos 10 kb. En otros LTVECs, el brazo de homología hacia arriba varía de aproximadamente 5 kb a aproximadamente 100 kb y/o el brazo de homología hacia abajo varía de aproximadamente 5 kb a aproximadamente 100 kb. La suma total de los brazos de homología hacia arriba y hacia abajo puede ser, por ejemplo, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 70 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, o de aproximadamente 190 kb a aproximadamente 200 kb.

35 En algunos casos, el LTVEC e inserto de ácido nucleico se designan para permitir una eliminación en el locus diana de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, o de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 400 kb, de aproximadamente 400 kb a aproximadamente 500 kb, de aproximadamente 500 kb a aproximadamente 1 Mb, de aproximadamente 1 Mb a aproximadamente 1,5 Mb, de aproximadamente 1,5 Mb a aproximadamente 2 Mb, de aproximadamente 2 Mb a aproximadamente 2,5 Mb, o de aproximadamente 2,5 Mb a aproximadamente 3 Mb. De manera alternativa, la eliminación puede ser de al menos 10 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, al menos 300 kb, al menos 350 kb, al menos 400 kb, al menos 450 kb, o al menos 500 kb o mayor.

50 En otros casos, el LTVEC e inserto de ácido nucleico se diseñan para permitir un inserto en el locus diana de una secuencia de ácido nucleico exógeno que varía de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 60 kb, de aproximadamente 60 kb a aproximadamente 80 kb, de aproximadamente 80 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 200 kb, de aproximadamente 200 kb a aproximadamente 250 kb, de aproximadamente 250 kb a aproximadamente 300 kb, de aproximadamente 300 kb a aproximadamente 350 kb, o de aproximadamente 350 kb a aproximadamente 400 kb. De manera alternativa, el inserto puede ser de al menos 10 kb, al menos 20 kb, al menos 30 kb, al menos 40 kb, al menos 50 kb, al menos 60 kb, al menos 70 kb, al menos 80 kb, al menos 90 kb, al menos 100 kb, al menos 150 kb, al menos 200 kb, al menos 250 kb, al menos 300 kb, al menos 350 kb, al menos 400 kb, al menos 450 kb, o al menos 500 kb o mayor.

60 En inclusive otros casos, el inserto de ácido nucleico y/o la región del locus endógeno que se elimina es de al menos 100, 200, 300, 400, 500, 600, 700, 800, o 900 nucleótidos o al menos 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 11 kb, 12 kb, 13 kb, 14 kb, 15 kb, 16 kb o mayor.

65 iv. Métodos de integración de un inserto de ácido nucleico cerca del sitio de reconocimiento en un cromosoma mediante recombinación homóloga

En algunos ejemplos, los métodos para modificar un locus genómico diana en un cromosoma en una célula pluripotente pueden comprender: (a) proporcionar una célula que comprende un locus genómico diana en un

5 cromosoma, (b) introducir en la célula un primer vector director que comprende un primer inserto de ácido nucleico  
 10 flanqueado por brazos de homología 5' y 3' correspondientes a secuencias diana 5' y 3'; y (c) identificar una menos  
 una célula que comprende en su genoma el primer inserto de ácido nucleico que se integra en el locus genómico  
 diana en el cromosoma. Según se analiza en detalle en otra parte en la presente, en ejemplos específicos, la suma  
 total del primer brazo de homología y el segundo brazo de homología del vector director es de aproximadamente 4  
 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, de aproximadamente 4 kb a aproximadamente 5 kb, de aproximadamente 5 kb a  
 aproximadamente 6 kb, de aproximadamente 6 kb a aproximadamente 7 kb, de aproximadamente 7 kb a  
 aproximadamente 8 kb, de aproximadamente 8 kb a aproximadamente 9 kb, o es al menos de 10 kb o al menos de 10 kb y menos de 150 kb. En realizaciones específicas,  
 se emplea un LTVEC. En una realización no limitante, tales métodos se llevan a cabo empleando el medio de cultivo  
 que se proporciona en la presente.

15 En otros ejemplos, los métodos para modificar un locus genómico diana en un cromosoma en una célula pluripotente  
 pueden comprender: (a) proporcionar una célula que comprende un locus genómico diana en un cromosoma que  
 comprende un sitio de reconocimiento para un agente de nucleasa, (b) introducir en la célula (i) el agente de nucleasa,  
 en los que el agente de nucleasa induce una mella o rotura bicatenaria en el primer sitio de reconocimiento; y (ii) un  
 primer vector director que comprende un primer inserto de ácido nucleico flanqueado por brazos de homología 5' y 3'  
 20 correspondientes a secuencias diana 5' y 3' que se ubican lo suficientemente próximas con respecto al primer sitio de  
 reconocimiento; y (c) identificar al menos una célula que comprende en su genoma el primer inserto de ácido nucleico  
 que se integra en el locus genómico diana en el cromosoma. Según se analiza en detalle en otra parte en la presente,  
 en ejemplos específicos, la suma total del primer brazo de homología y el segundo brazo de homología del vector  
 director es de aproximadamente 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, de aproximadamente 4 kb a aproximadamente 5 kb,  
 de aproximadamente 5 kb a aproximadamente 6 kb, de aproximadamente 6 kb a aproximadamente 7 kb, de  
 aproximadamente 7 kb a aproximadamente 8 kb, de aproximadamente 8 kb a aproximadamente 9 kb, o es al menos de 10 kb o al menos de 10 kb y menos de 150 kb. En  
 25 realizaciones específicas, se emplea un LTVEC. En una realización no limitante, tales métodos se llevan a cabo  
 empleando el medio de cultivo que se proporciona en la presente.

30 Diversos métodos pueden emplearse para identificar células pluripotentes que tienen el inserto de ácido nucleico que  
 se integra en el locus genómico diana. La inserción del inserto de ácido nucleico en el locus genómico diana da como  
 resultado una "modificación de alelo". El término "modificación de alelo" o "MOA" incluye la modificación de la  
 secuencia de ADN exacta de un alelo de un gen(es) o locus cromosómico (loci) en un genoma. Ejemplos de  
 "modificación de alelo (MOA)" incluyen, pero sin limitación, eliminaciones, sustituciones, o inserciones de tan poco  
 como un solo nucleótido o eliminaciones de muchas kilobases que abarcan un gen(s) o locus cromosómico (loci) de  
 35 interés, así como también todo tipo de modificaciones posibles entre estos dos extremos.

40 En diversos ejemplos, para facilitar la identificación de la modificación dirigida, se emplea un ensayo cuantitativo de  
 alto rendimiento, a saber, ensayo de modificación de alelo (MOA). El ensayo de MOA que se describe en la presente  
 permite un análisis a gran escala de un alelo(s) modificado en un cromosoma parental a continuación de una  
 modificación genética. El ensayo de MOA puede llevarse a cabo mediante diversas técnicas analíticas, incluyendo,  
 pero sin limitación, una PCR cuantitativa, por ejemplo, una PCR en tiempo real (qPCR). Por ejemplo, la PCR en tiempo  
 real comprende un primer conjunto de cebador-sonda que reconoce el locus diana y un segundo conjunto de cebador-  
 sonda que reconoce un locus de referencia no dirigido. De manera adicional, el primer conjunto de cebador-sonda  
 comprende una sonda fluorescente que reconoce la secuencia amplificada. En ensayo cuantitativo puede llevarse a  
 45 cabo además mediante una variedad de técnicas analíticas, incluyendo, pero sin limitación, hibridación in situ mediada  
 por fluorescencia (FISH), hibridación genómica comparativa, amplificación de ADN isotérmica, hibridación cuantitativa  
 en una sonda(s) inmovilizada, Invader Probes®, MMP assays®, TaqMan® Molecular Beacon, y tecnología de sonda  
 Eclipse™. Véase, por ejemplo, US2005/0144655.

50 En diversas realizaciones, en la presencia de la mella o rotura bicatenaria, la eficiencia directora de un vector director  
 (tal como un LTVEC) en el locus genómico diana es de al menos aproximadamente 2 veces mayor, al menos 3 veces  
 mayor, al menos 4 veces mayor que en la ausencia de la mella o rotura bicatenaria (usando, por ejemplo, el mismo  
 vector director y los mismos brazos de homología y sitios diana correspondientes en el locus genómico de interés pero  
 en la ausencia de un agente de nucleasa agregado que crea la mella o rotura bicatenaria).

55 Los diversos métodos que se establecen anteriormente pueden repetirse de manera secuencial para permitir la  
 integración dirigida de cualquier número de insertos de ácido nucleico en un locus genómico dirigido dado en un  
 cromosoma. De este modo, se proporcionan los diversos métodos para la inserción de al menos 1, 2, 3, 4, 5, 6, 7, 8,  
 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más insertos de ácido nucleico en el locus genómico diana en un  
 cromosoma. En realizaciones en particular, tales métodos de teselación secuencial permiten la reconstrucción de  
 60 regiones cromosómicas grandes a partir de una célula animal o a partir de una célula de mamífero (a saber, un  
 humano, un no humano, un roedor, un ratón, un mono, una rata, un hámster, un mamífero domesticado, o un animal  
 agrícola) en un locus genómico dirigido en un cromosoma. En tales casos, la transferencia y reconstrucción de  
 regiones genómicas que incluyen ambas regiones codificantes y no codificantes permite que la complejidad de una  
 región dada se conserve mediante retención de, al menos en parte, las regiones de codificación, las regiones de no  
 65 codificación y las variaciones del número de copias en la región genómica nativa. De este modo, los diversos métodos  
 proporcionan, por ejemplo, métodos para generar regiones genómicas "heterólogas" o exógenas" en una célula iPS  
 humana o una célula no pluripotente que se transforma para expresar un estado pluripotente.

v. Polinucleótidos de interés

5 Cualquier polinucleótido de interés puede contenerse en los diversos insertos de ácido nucleico y se integran así en el locus genómico diana en un cromosoma. Los métodos que se divulgan en la presente proporcionan al menos 1, 2, 3, 4, 5, 6, o más polinucleótidos de interés para integrarse en el locus genómico diana.

10 El polinucleótido de interés en el inserto de ácido nucleico cuando se integra en el locus genómico diana en un cromosoma puede introducir una o más modificaciones genéticas en la célula pluripotente. La modificación genética puede comprender una eliminación de una secuencia de ácido nucleico endógeno y/o la adición de un polinucleótido exógeno o heterólogo u ortólogo en el locus genómico diana. En una realización, la modificación genética comprende un reemplazo de una secuencia de ácido nucleico endógeno con un polinucleótido exógeno de interés en el locus genómico diana. De este modo, los métodos que se proporcionan en la presente permiten la generación de una  
15 modificación genética que comprende un bloqueo, una eliminación, una inserción, un reemplazo (knock-in), una mutación puntual, un intercambio de dominios, un intercambio de exones, un intercambio de intrones, un intercambio de secuencia reguladora, un intercambio de genes, o una combinación de estos en un locus genómico diana en un cromosoma. Tales modificaciones pueden ocurrir al momento de integrar el primero, segundo, tercero, cuarto, quinto, sexto, séptimo, o cualquier inserto de ácido nucleico posterior en el locus genómico diana.

20 El polinucleótido de interés en el inserto de ácido nucleico y/o que se integra en el locus genómico diana puede comprender una secuencia que resulta nativa u homóloga con respecto a la célula pluripotente en la que se introduce; el polinucleótido de interés puede ser heterólogo con respecto a la célula en la que se introduce; el polinucleótido de interés puede ser exógeno con respecto a la célula en la que se introduce; el polinucleótido de interés puede ser ortólogo con respecto a la célula en la que se introduce; o el polinucleótido de interés puede ser de una especie diferente con respecto a la célula en la que se introduce. "Homólogo" con referencia a una secuencia incluye una secuencia que resulta nativa con respecto a la célula. "Heteróloga" con referencia a una secuencia incluye una secuencia que se origina a partir de una especie foránea, o, si es a partir de la misma especie, se modifica sustancialmente a partir de su forma nativa en composición y/o locus genómico mediante intervención humana  
25 deliberada. "Exógena" con referencia a una secuencia incluye una secuencia que se origina a partir de una especie foránea. "Ortólogo" incluye un polinucleótido a partir de una especie que resulta funcionalmente equivalente con respecto a una secuencia de referencia que se conoce en otra especie (a saber, una variante de especie). El polinucleótido de interés puede ser a partir de cualquier organismo de interés, incluyendo, pero sin limitación, no humano, un roedor, un hámster, un ratón, una rata, un humano, un mono, un ave, un mamífero agrícola, o un mamífero no agrícola. El polinucleótido de interés puede comprender además una región codificante, una región no codificante, una región reguladora, o un ADN genómico. De este modo, el 1<sup>era</sup>, 2<sup>da</sup>, 3<sup>ra</sup>, 4<sup>ta</sup>, 5<sup>ta</sup>, 6<sup>ta</sup>, 7<sup>ma</sup>, y/o cualquiera de los insertos de ácido nucleico posteriores pueden comprender tales secuencias.

30 En una realización, el polinucleótido de interés en el inserto de ácido nucleico y/o que se integra en el locus genómico diana en un cromosoma resulta homólogo con respecto a un ácido nucleico humano. En otra realización adicional, el polinucleótido de interés que se integra en el locus diana es un fragmento de un ácido nucleico genómico. En una realización, el ácido nucleico genómico es un ácido nucleico genómico de ratón, un ácido nucleico genómico de humano, un ácido nucleico no humano, un ácido nucleico de roedor, un ácido nucleico de rata, un ácido nucleico de hámster, un ácido nucleico de mono, un ácido nucleico de mono, un ácido nucleico de mamífero agrícola, o un ácido nucleico de mamífero no agrícola o una combinación de este.  
35

40 En una realización, el polinucleótido de interés puede variar de aproximadamente 500 nucleótidos a aproximadamente 200 kb según se describe anteriormente. El polinucleótido de interés puede ser de aproximadamente 500 nucleótidos a aproximadamente 5 kb, de aproximadamente 5 kb a aproximadamente 200 kb, de aproximadamente 5 kb a aproximadamente 10 kb, de aproximadamente 10 kb a aproximadamente 20 kb, de aproximadamente 20 kb a aproximadamente 30 kb, de aproximadamente 30 kb a aproximadamente 40 kb, de aproximadamente 40 kb a aproximadamente 50 kb, de aproximadamente 50 kb a aproximadamente 60 kb a aproximadamente 70 kb, de aproximadamente 70 kb a aproximadamente 80 kb a aproximadamente 90 kb, de aproximadamente 90 kb a aproximadamente 100 kb, de aproximadamente 100 kb a aproximadamente 110 kb, de aproximadamente 110 kb a aproximadamente 120 kb, de aproximadamente 120 kb a aproximadamente 130 kb, de aproximadamente 130 kb a aproximadamente 140 kb, de aproximadamente 140 kb a aproximadamente 150 kb, de aproximadamente 150 kb a aproximadamente 160 kb, de aproximadamente 160 kb a aproximadamente 170 kb, de aproximadamente 170 kb a aproximadamente 180 kb, de aproximadamente 180 kb a aproximadamente 190 kb, o de aproximadamente 190 kb a aproximadamente 200 kb.  
45

50 El polinucleótido de interés en el inserto de ácido nucleico y/o que se inserta en el locus genómico diana en un cromosoma pueden codificar un polipéptido, puede codificar un miARN, puede codificar un ARN no codificante largo, o puede comprender cualquiera de las regiones reguladoras o regiones no codificantes de interés incluyendo, una secuencia reguladora, una secuencia promotora, una secuencia potenciadora, una secuencia de unión a represor transcripcional, o una eliminación de una secuencia de no codificación de proteína, pero no comprende una eliminación de una secuencia de codificación de proteína. De manera adicional, el polinucleótido de interés en el inserto de ácido nucleico y/o que se inserta en el locus genómico diana en un cromosoma puede codificar una proteína que se expresa en el sistema nervioso, el sistema esquelético, el sistema digestivo, el sistema circulatorio, el sistema muscular, el  
55 60 65

sistema respiratorio, el sistema cardiovascular, el sistema linfático, el sistema endocrino, el sistema urinario, el sistema reproductivo, o una combinación de estos.

5 El polinucleótido de interés en un inserto de ácido nucleico y/o que se integra en el locus genómico diana en un cromosoma puede comprender una modificación genética de una secuencia de codificación. Tales modificaciones genéticas incluyen, pero sin limitación, una mutación de eliminación de una secuencia de codificación o la fusión de dos secuencias de codificación.

10 El polinucleótido de interés en el inserto de ácido nucleico y/o que se integra en el locus genómico diana en un cromosoma puede comprender un polinucleótido que codifica una proteína mutante. En una realización, la proteína mutante se caracteriza mediante una característica de unión alterada, localización alterada, expresión alterada, y/o patrón de expresión alterada. En una realización, el polinucleótido de interés en el inserto de ácido nucleico y/o que se integra en el locus genómico diana en un cromosoma comprende al menos un alelo de enfermedad. En tales casos, 15 el alelo de enfermedad puede ser un alelo dominante o el alelo de enfermedad es un alelo recesivo. Además, el alelo de enfermedad puede comprender un alelo de polimorfismo de un solo nucleótido (SNP). El polinucleótido de interés que codifica la proteína mutante, puede ser de cualquier organismo, incluyendo, pero sin limitación, un polinucleótido de mamífero, un mamífero no humano, roedor, ratón, rata, un humano, un mono, un mamífero agrícola, o un mamífero doméstico que codifica una proteína mutante.

20 El polinucleótido de interés en el inserto de ácido nucleico y/o que se integra en el locus genómico diana en un cromosoma puede comprender, además, una secuencia reguladora, incluyendo, por ejemplo, una secuencia promotora, una secuencia potenciadora, una secuencia de unión de represor transcripcional, o una secuencia 25 terminadora transcripcional. En realizaciones específicas, el polinucleótido de interés en el inserto de ácido nucleico y/o que se integra en el locus genómico diana en un cromosoma comprende un polinucleótido que tiene una eliminación de una secuencia no codificadora de proteína, pero que no comprende una eliminación de una secuencia codificadora de proteína. En una realización, la eliminación de la secuencia no codificadora de proteína comprende una eliminación de una secuencia reguladora. En otra realización, la eliminación del elemento regulador comprende una eliminación de una secuencia promotora. En una realización, la eliminación del elemento regulador comprende una 30 eliminación de una secuencia potenciadora. Un polinucleótido de interés como tal puede ser a partir de cualquier organismo, incluyendo, pero sin limitación, un polinucleótido de mamífero, un mamífero no humano, roedor, ratón, rata, un humano, un mono, un mamífero agrícola, o un mamífero doméstico que codifica una proteína mutante.

35 Si diferentes versiones de una secuencia se asocian con un número de acceso en tiempos diferentes, se hace referencia a la versión asociada con el número de acceso en la fecha de presentación efectiva de esta aplicación. La fecha de presentación efectiva se refiere a la primera fecha que tuvo lugar entre la fecha de presentación actual o fecha de presentación de una solicitud prioritaria que hace referencia al número de acceso si corresponde. Asimismo, si diferentes versiones de una publicación, sitio web o similares se publican en momentos diferentes, se hace referencia a la versión más recientemente publicada con respecto a la fecha de presentación efectiva de la solicitud a 40 menos que se indique lo contrario. Cualquier característica, etapa, elemento, realización, o aspecto de la invención pueden usarse en combinación con cualquier otro a menos que se indique específicamente lo contrario. Muchas modificaciones y otras realizaciones de los métodos y composiciones que se establecen en la presente vendrán a la mente de una persona experta en la técnica a la cual estos métodos y composiciones corresponden teniendo el beneficio de las enseñanzas que se presentan en las descripciones anteriores y los dibujos asociados. Por lo tanto, 45 debe comprenderse que los métodos y composiciones no se dirigen a limitarse a las realizaciones específicas que se divulgan y que se incluyen modificaciones y otras realizaciones en el alcance de las reivindicaciones adjuntas. A pesar de que se emplean términos específicos en la presente, estos se usan en un sentido genérico y descriptivo y no con fines limitativos.

50 Los siguientes ejemplos se ofrecen a modo de ilustración y no a modo de limitación.

#### Ejemplos

##### Ejemplo 1. Generación de Células iPS humanas

55 Este ejemplo describe la generación de células iPS humanas a partir de células humanas no pluripotentes. Se muestra un protocolo de muestra en la Tabla 2. Vectores PiggyBac (System Biosciences) (PB-600A\_CAGGGS Bst XI (0,64 µg/µL) y PB-200 (0,99 µg/µL) que comprenden los genes que codifican cuatro factores de reprogramación (hOct4, hSox2, hKLF-4, hMYC) que se unen de manera manejable a un promotor CM7 se introdujeron en fibroblastos de prepucio humano de recién nacidos usando reactivos de transfección RED and BLUE GeneIn™ (GlobalStem). Las células transfectadas se incubaron en células alimentadoras de NuFF1 en medio E7 (Life Technologies) para permitir 60 la incorporación de los vectores y expresión de los factores de reprogramación. El medio E7 comprendió DMEM/F-12, NaHCO<sub>3</sub>, ácido L-ascórbico, insulina, transferrina, selenio, y FGF-2.

65 La selección de puomicina comenzó 10 días después de la transfección usando 2 µg/mL de puomicina en medio E7. En el día 21, se seleccionaron y cultivaron las colonias en medio mTeSR™ (medio mTeSR™ 1 a partir de STEMCELL Technologies), que comprende DMEM/F-12, NaHCO<sub>3</sub>, ácido L-ascórbico, insulina, transferrina, selenio, FGF-2, TGF-

5  $\beta$ -1, glutatión, L-glutamina, lípidos definidos, tiamina, elementos traza B y C,  $\beta$ -mercaptoetanol, albúmina de suero bovino, ácido pipercolico, cloruro de litio, y GABA. Una comparación de componentes del medio mTeSR™ y el medio E7 se muestran en la Tabla 3. En los días 29 a 57, las células se propagaron y se pasaron en medio mTeSR™ hasta que se alcanzó ~50% de confluencia en 6 placas de pozos. En los días 65 a 73, la propagación y pase continuaron usando medio mTeSR™ y Reactivo de Disociación Celular Leve (Stem Cell Technologies). En el día 76, el medio se cambió a medio VG2i de baja osmolaridad para propagación adicional, pase, y mantenimiento de las células que comprenden hiPSCs vírgenes o de apariencia virgen. El tiempo de pase en el método anterior se determinó mediante morfología celular.

10

Tabla 2. Protocolo de muestra para generación de células iPS humanas.

Día	Acción	Medio
0	Colocar en placas células NuFF1 en placas de 6 pozos (1 x 10 <sup>5</sup> por pozo)	Medio de fibroblasto
1	Transfección piggyBac	Medio E7
10	Selección de puromicina de partida ( 2ug/mL)	
21	Recolectar colonias mediante corte	Medio mTeSR
29-57	Pase de células hasta que alcanzan ~50% de pozo (pase mediante Reactivo de Disociación Celular Leve)	
65-73	Continuar pase células y congelar 2/3 cuando se pasan (pase mediante Reactivo de Disociación Celular Leve)	
76	Cambiar medio a VG2i	VG2i

Tabla 3. Comparación de componentes de medio mTeSR y E7.

Componente	mTeSR	E7
DMEM / F-12	+	+
Ácido L-Ascórbico	+	+
Insulina	+	+
Transferrina	+	+
Selenio	+	+
FGF2	+	+
TGF $\beta$ 1	+	
Glutatión	+	
L-Glutamina	+	
Lípidos definidos	+	
Tiamina	+	
Elementos traza B	+	
Elementos traza C	+	
B-Mercaptoetanol	+	

Albúmina (BSA)	+	
Ácido pipercolico	+	
LiCl	+	
GABA	+	

Ejemplo 2. LTVEC director en células iPS humanas

5 Este ejemplo describe el uso de LTVEC director en células iPS humanas. Según se muestra en la Figura 1, introdujimos mediante electroporación en células iPS humanas que se propagaron en medio VG2i las siguientes moléculas de ácidos nucleicos: (1) un LTVEC (0,67 µg); (2) un plásmido que codifica una endonucleasa Cas9 (5 µg); y (3) un plásmido que codifica un solo ARN guía CRISPR (ARNg) (10 µg). En un conjunto de muestras, se excluyeron la Cas9 y ARNg. De manera específica, 3 x 10<sup>6</sup> células se sometieron a electroporación a un voltaje de 700V, una capacitancia de 25 µF, y una resistencia de 400 ohms. El LTVEC comprende un ácido nucleico de 16,7 kb que comprende genes *Adam6a* y *Adam6b* de ratón flanqueados por brazos de homología que contienen 34 kb y 105 kb de ADN genómico que deriva de regiones genómicas que flanquean la secuencia de 4,1 kb del locus *ADAM6* humano previsto para eliminación. El LTVEC transportó además un cassette de selección de medicamentos que dirige la expresión de una enzima que imparten resistencia con respecto a un medicamento antibiótico (higromicina). El ARNg *ADAM6* humano que se usó tuvo la siguiente secuencia: GTATAGCCCTGTTACACATT (SEQ ID NO: 1).

Células que tomaron el LTVEC y lo incorporaron en sus genomas fueron capaces de crecer y formar colonias en un plato de cultivo de tejido revestido de GELTRES™ en un medio de desarrollo que contiene el medicamento antibiótico. Debido a que introdujimos de 500 a 1.000 veces más moléculas de ácido nucleico que codifica CRISPR/Cas9 en comparación con moléculas de LTVEC, la mayoría de las colonias resistentes al medicamento que contienen LTVEC contuvieron también, al menos de manera transitoria, los componentes CRISPR/Cas9. Recolectamos las colonias resistentes al medicamento y las evaluamos mediante el método de pérdida de alelo (Valenzuela et al. (2003) Nat. Biotech. 21:652-660; Frenthewey et al. (2010) Methods Enzymol. 476:295-307) para identificar clones que tuvieron el alelo dirigido de manera correctamente.

25 Los resultados del LTVEC director asistido por CRISPR/Cas9 del locus *ADAM6* se muestran en la Tabla 4.

Tabla 4. Dirección de LTVEC asistida por CRISPR/Cas9.

Condición directora	Eficiencia directora
Solo LTVEC	3,1%
LTVEC + CRISPR	7,3%

30 Cuando el LTVEC solo se introdujo en células iPS humanas, se observó una eficiencia directora del 3,1%. En cambio, la combinación del LTVEC con Cas9 guiado por el ARNg de *ADAM6* dio como resultado una eficiencia directora del 7,3%.

Ejemplo 3. Efecto del medio de baja osmolaridad en morfología de célula iPS humana.

35 Este ejemplo describe el efecto de concentración de sal, fuerza iónica, y/o osmolaridad en el estado de pluripotencia de células iPS humanas en cultivo. Células iPS humanas se cultivaron en un sustrato MATRIGEL™ o GELTRES™ en un medio que se describe en la Tabla 5 (concentración hLIF final de 100 U/mL; concentración de CHIR99021 final de 3 µM, y concentración de PD0325901 final de 0,5 µM) o en medio mTeSR™-hLIF.

Tabla 5. Medio para cultivo de célula iPS.

Componente	Cantidad (v/v)
Medio base	24,75
Medio de F-12	24,75
Complemento N2®	0,5

Medio Neurobasal®	49
Complemento B-27®	1
Penicilina/ Estreptomina	1
L-Glutamina (200 mM)	1
2-Mercaptoetanol (55 mM)	0,1836
hLIF (1 x 10 <sup>4</sup> unidades/mL)	0,001
CHIR99021 (10 mM)	0,03
PD0325901 (10 mM)	0,005

Tabla 6. Osmolaridad de medio y componentes de medio.

Medio o componente de medio	Osmolaridad (mOsm/kg)
Medio 2i	261
Medio VG2i	233
Neurobasal	216
DMEM/F-12	305
DMEM	340
F-12	290
VG-DMEM	200

5 Cuando se usó el medio base DMEM, este medio se denominó medio 2i. Cuando se usó el medio base VG-DMEM, este medio de baja osmolaridad se denominó medio VG2i. La osmolaridad del medio VG2i (233 mOsm/kg) es más baja que la osmolaridad del medio 2i tradicional (261 mOsm/kg). La Tabla 6 muestra las osmolaridades para estos medios así como también la osmolaridad de diferentes medios base y componentes de medios,

10 Según se muestra en la Figura 2, las células iPS humanas que se cultivan en MATRIGEL™ en medio 2i durante un período de 8 días (Figura 2A) o 12 días (Figura 2B) mostraron una morfología característica de células iPS en un estado cebado, en particular crecimiento en una monocapa epitelial y la aparición de polaridad apicobasal.

15 El medio mTeSR™-hLIF y medio VG2i se evaluaron además en cuanto a sus efectos en la morfología y estado de pluripotencia de células iPS humanas. En este estudio, las células iPS humanas fueron cultivadas en MATRIGEL™ o células alimentadoras de NuFF en medio mTeSR™-hLIF (Figuras 3A y 3C) o en medio VG2i (Figuras 3B y 3D) durante un período de 6 días. Cuando se cultivaron en medio mTeSR™-hLIF en MATRIGEL™ o células alimentadoras de NuFF, las células iPS humanas mostraron una morfología característica de un estado de pluripotencia cebado, en particular crecimiento en una monocapa epitelial y la aparición de polaridad apicobasal. Algunas células cultivadas en

20 medio mTeSR™-hLIF comenzaron a mostrar una morfología caracterizada por agrupamientos tridimensionales. En cambio, cuando se cultivaron en medio VG2i en MATRIGEL™ o células alimentadoras NuFF, las células iPS humanas mostraron una morfología característica de un estado de pluripotencia virgen, en particular crecimiento en colonias redondas, de forma de cúpula y una falta de polaridad apicobasal.

25 Ejemplo 4. Efecto de medio de baja osmolaridad en la expresión de marcadores de pluripotencia en células iPS humanas.

Este ejemplo describe el efecto de concentración de sal, fuerza iónica, y/o osmolaridad en la expresión de marcadores de pluripotencia en células iPS humanas que han sido reprogramadas a partir de un estado cebado a un estado virgen.

30 A continuación de 24 días de cultivo en medio VG2i en un sustrato MATRIGEL™, células iPS humanas vírgenes reprogramadas se tiñeron en cuanto a la expresión de fosfatasa alcalina o NANOG. Se observó que las células

reprogramadas expresaron fuertemente tanto fosfatasa alcalina (Figura 4A) y NANONG (Figuras 4B y 4C), lo que resulta indicativo de un estado de pluripotencia virgen.

5 Ejemplo 5. Efecto de medio de baja osmolaridad en disociación enzimática y subcultivo de células iPS humanas.

En este ejemplo, las células iPS humanas que se reprogramaron a un estado virgen usando medio VG2i de baja osmolaridad se disociaron de manera enzimática usando tripsina para crear una suspensión de célula única (Figura 5A). La suspensión celular se pasó en nuevas placas recubiertas por GELTRES™ para subcultivo en medio VG2i. Se observó que después de 1 día (Figura 5B) y 4 días (Figura 5C) las células subcultivadas continuaron mostrando una morfología característica de células en un estado de pluripotencia virgen. En particular, las células crecieron como colonias redondeadas de forma de cúpula y no exhibieron una polaridad apicobasal. Cabe destacar que la disociación enzimática podría llevarse a cabo en la ausencia de un inhibidor de ROCK, que resulta normalmente necesario para impedir activación de rutas proapoptóticas. Esto sugiere que las rutas proapoptóticas no se activan fuertemente durante disociación enzimática y subcultivo en células iPS humanas vírgenes que se cultivan en condiciones que se identifican en la presente. De manera adicional, las células iPS humanas que se cultivan en VG2i y se pasan como células únicas a continuación de disociación enzimática con tripsina mantienen un cariotipo normal. Dos pases de 10 células iPS humanas que derivan a partir de clones diferentes y que se producen a continuación de disociación con tripsina para crear una suspensión de célula única se cariotiparon, y ambas tuvieron un cariotipo 46 XY normal (Figuras 6A y 6B).

Listado de secuencias

<110> Kuno, Junko Auerbach, Wojtek Valenzuela, David M.  
 25 <120> MÉTODOS Y COMPOSICIONES PARA GENERAR O MANTENER CÉLULAS PLURIPOTENTES  
 <130> 57766/469587  
 30 <150> US 62/064,384  
 <151> 2014-10-15  
 <160> 2  
 35 <170> FastSEQ for Windows Version 4.0  
 <210> 1  
 40 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 45 <220>  
 <223> Sintetizada  
 50 <400> 1  
 gtatagccct gttacacatt 20  
 <210> 2  
 55 <211> 23  
 <212> ADN  
 <213> Secuencia artificial  
 60 <220>  
 <223> Sintetizada  
 65 <220>  
 <221> misc\_característica

<222> (2) ... (21)

5 <223> n = A,T,C o G

<400> 2

gooooooooo nnnnnnnnnn ngg 23

**REIVINDICACIONES**

- 5 1. Un método para fabricar una población de células madre pluripotentes inducidas humanas (hiPSCs), comprendiendo el método el cultivo *in vitro* de una población de células no pluripotentes, que se transforman para expresar un estado pluripotente, en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende:
- 10 (a) un polipéptido de factor de inhibición de leucemia (LIF);
- (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y
- (c) un inhibidor de MEK;
- 15 en el que el medio de baja osmolaridad tiene una osmolaridad de 200 mOsm/kg a 250 mOsm/kg.
2. El método de la reivindicación 1, en el que las células transformadas se cultivan primero en una medio de alta osmolaridad antes del cultivo en el medio de baja osmolaridad, en el que el medio de alta osmolaridad comprende bFGF, de manera opcional, el medio de alta osmolaridad tiene una osmolaridad de al menos aproximadamente 290 mOsm/kg, y de manera opcional, en el que:
- 20 (a) las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que expresan características de un estado virgen;
- 25 (b) las células transformadas se cultivan primero en el medio de alta osmolaridad durante un período de aproximadamente dos meses;
- (c) las células transformadas se cultivan primero en el medio de alta osmolaridad hasta que muestran una morfología que se caracteriza por grupos de células tridimensionales; o
- 30 (d) una combinación de estos.
3. Un método para mantener una población de células madre pluripotentes inducidas humanas (hiPSCs) en un cultivo *in vitro*, comprendiendo el método el cultivo de la población de hiPSCs en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende:
- 35 (a) un polipéptido de factor de inhibición de leucemia (LIF);
- (b) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y
- 40 (c) un inhibidor de MEK;
- en el que el medio de baja osmolaridad tiene una osmolaridad de 200 mOsm/kg a 250 mOsm/kg.
- 45 4. El método de una cualquiera de las reivindicaciones 1-3, que comprende además la disociación de manera enzimática de las hiPSCs en una suspensión de célula única y subcultivo de las hiPSCs disociadas de manera enzimática, en el que, de manera opcional, la disociación enzimática:
- 50 (a) se lleva a cabo usando tripsina;
- (b) se lleva a cabo en la ausencia de un inhibidor de ROCK; o
- (c) una combinación de estos; y
- 55 en el que las hiPSCs subcultivadas:
- (a) continúan expresando uno o más marcadores de pluripotencia;
- 60 (b) mantienen un estado virgen y muestran una morfología que se caracteriza por colonias compactas de forma de cúpula; o
- (c) una combinación de estos; y
- en el que las hiPSCs subcultivadas mantienen un cariotipo normal.
- 65 5. Un método para modificar un locus genómico diana en una célula madre pluripotente inducida humana (hiPSC), que comprende:

5 (a) introducir en hiPSCs un vector director que comprende un inserto de ácido nucleico flanqueado por brazos de homología 5' y 3' correspondientes a sitios diana 5' y 3' en el locus genómico diana, en el que, de manera opcional, el vector director es un vector director grande (LTVEC), en el que la suma total de los brazos de homología 5' y 3' es al menos de 10 kb; y

10 (b) identificar una hiPSC genéticamente modificada que comprende en su genoma el inserto de ácido nucleico que se integra en el locus genómico diana; en el que las hiPSCs se cultivan en un medio de baja osmolaridad que comprende un medio base y complementos, en el que el medio de baja osmolaridad comprende:

15 (i) un polipéptido de factor de inhibición de leucemia (LIF);

(ii) un inhibidor de glucógeno sintasa quinasa 3 (GSK3); y

(iii) un inhibidor de MEK;

20 en el que el medio de baja osmolaridad tiene una osmolaridad de 200 mOsm/kg a 250 mOsm/kg.

6. El método de la reivindicación 5, en el que la modificación genética dirigida comprende:

25 (a) eliminación de una secuencia de ácido nucleico endógeno humano;

(b) inserción de una secuencia de ácido nucleico exógeno, en el que, de manera opcional, la secuencia de ácido nucleico exógeno comprende uno o más de los siguientes:

30 (i) una secuencia de ácido nucleico que resulta homólogo u ortólogo con respecto a la secuencia de ácido nucleico humano;

(ii) una secuencia de ácido nucleico quimérico;

35 (iii) un alelo condicional flanqueado por secuencias diana de recombinasa específica de sitio; y

(iv) un gen reportero que se une de manera manejable a un promotor activo en la hiPSC; o

(c) reemplazo de la secuencia de ácido nucleico endógeno humano con la secuencia de ácido nucleico exógeno,

40 en el que, de manera opcional, la modificación genética dirigida es bialélica.

7. El método de las reivindicaciones 5 o 6, en el que la introducción de la etapa (a) comprende además la introducción de un agente de nucleasa que promueve la recombinación homóloga entre el vector director y el locus genómico diana en la hiPSC.

45 8. El método de la reivindicación 7, en el que el agente de nucleasa comprende

(a) una nucleasa con dedos de zinc (ZFN);

50 (b) una nucleasa efectora de tipo activador de transcripción (TALEN);

(c) una meganucleasa; o

55 (d) una proteína (Cas) asociada a repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR), y un ARN guía (ARNg) que comprende un ARN CRISPR (crARN) que reconocen una secuencia genómica diana y un ARN CRISPR de transactivación (tracrARN),

60 de manera opcional, en el que el agente de nucleasa comprende la proteína Cas y el ARN guía, en el que la proteína Cas es Cas9.

9. El método de acuerdo con una cualquiera de las reivindicaciones 5-8, en el que con anterioridad a la etapa (a), las hiPSCs se disocian de manera enzimática en una suspensión de célula única y subcultivo, en el que, de manera opcional, la disociación enzimática:

65 (a) se lleva a cabo usando tripsina;

(b) se lleva a cabo en la ausencia de un inhibidor de ROCK; o

- (c) una combinación de estos; y
- 5 de manera opcional, en el que las hiPSCs subcultivadas:
- (a) continúan expresando uno o más marcadores de pluripotencia;
- 10 (b) mantienen un estado virgen y muestran una morfología que se caracteriza por colonias compactas de forma de cúpula; o
- (c) una combinación de estos; y
- 15 de manera opcional, en el que las hiPSCs subcultivadas mantienen un cariotipo normal.
10. El método de una cualquiera de las reivindicaciones 1-9, en el que las hiPSCs tienen un cariotipo normal, y en el que las hiPSCs:
- 20 (a) comprenden hiPSCs vírgenes;
- (b) expresan uno o más marcadores de pluripotencia, en el que, de manera opcional, los marcadores de pluripotencia comprenden NANOG, fosfatasa alcalina, o una combinación de estos;
- 25 (c) muestran una morfología que se caracteriza por colonias compactas de forma de cúpula;
- (d) pueden diferenciarse en células de una cualquiera de las capas germinales de endodermo, ectodermo, o mesodermo;
- 30 (e) tienen un tiempo de duplicación de entre aproximadamente 16 horas y aproximadamente 24 horas; o
- (f) cualquier combinación de (a) a (e).
11. El método de una cualquiera de las reivindicaciones 1-10, en el que las hiPSCs derivan a partir de células no pluripotentes que se transforman para expresar un estado pluripotente, en el que las células transformadas expresan genes de reprogramación que comprenden Oct4, Sox2, Klf4, Myc, o cualquier combinación de estos y, de manera opcional, las células transformadas comprenden hiPSCs cebadas.
- 35 12. El método de una cualquiera de las reivindicaciones 1-11, en el que el medio base tiene una osmolaridad de aproximadamente 180 mOsm/kg a aproximadamente 250 mOsm/kg, en el que, de manera opcional, el medio base comprende cloruro de sodio a aproximadamente 3 mg/mL, bicarbonato de sodio a aproximadamente 2,2 mg/mL, y tiene una osmolaridad de aproximadamente 200 mOsm/kg, y en el que, de manera opcional, el medio comprende glucosa a aproximadamente 4,5 mg/mL.
- 40 13. El método de una cualquiera de las reivindicaciones 1-12, en el que el medio de baja osmolaridad tiene una osmolaridad de alrededor 220-240 mOsm/kg, en el que, de manera opcional, el medio de baja osmolaridad tiene una osmolaridad de 225-235 mOsm/kg, y en el que, de manera opcional, el medio de baja osmolaridad tiene una osmolaridad de 233 mOsm/kg.
- 45 14. El método de una cualquiera de las reivindicaciones 1-13, en el que:
- 50 (a) los complementos comprenden:
- (i) medio F-12;
- 55 (ii) complemento de N2;
- (iii) medio NEUROBASAL;
- (iv) complemento de B-27;
- 60 (v) L-glutamina;
- (vi) 2- mercaptoetanol; o
- 65 (vii) cualquier combinación de (i) a (vi).
- (b) el polipéptido de LIF es un polipéptido de LIF humano (hLIF);

- (c) el inhibidor de GSK3 comprende CHIR99021;
- 5 (d) el inhibidor de MEK comprende PD0325901; o
- (e) cualquier combinación de (a) a (d), y
- 10 de manera opcional, en el que las hiPSCs se cultivan en MATRIGEL™, células alimentadoras de fibroblastos de prepucio humano de recién nacidos (NuFF), o GELTREX.
- 15 15. El método de una cualquiera de las reivindicaciones 1-14, en el que el medio de baja osmolaridad comprende inhibidores que consisten esencialmente en un inhibidor de glucógeno sintasa quinasa 3 (GSK3) y un inhibidor de MEK.
- 20 16. El método de una cualquiera de las reivindicaciones 1-15, en el que el medio de baja osmolaridad comprende medio base de aproximadamentel 24,75% (v/v), medio F-12 de aproximadamentel 24,75% (v/v), complemento de N2 de aproximadamentel 0,5% (v/v), medio NEUROBASAL de aproximadamentel 49% (v/v), complemento de B-27 de aproximadamentel 1% (v/v), L-glutamina de aproximadamente 2 mM, 2-mercaptoetanol de aproximadamente 0,1 mM, hLIF de aproximadamente 100 unidades/mL, CHIR99021 de aproximadamente 3  $\mu$ M, y PD0325901 de aproximadamente 0,5  $\mu$ M, en el que, de manera opcional, las hiPSCs se cultivan en MATRIGEL, células alimentadoras de fibroblastos de prepucio humano de recién nacidos (NuFF), o GELTREX.
- 25 17. El método de una cualquiera de las reivindicaciones 1-16, en el que el medio de baja osmolaridad no comprende uno o más de los siguientes: complemento de bFGF; complemento de TGF- $\beta$ 1; inhibidor de JNK; inhibidor de p38; inhibidor de ROCK; e inhibidor de PKC; en el que, de manera opcional, el medio de baja osmolaridad no comprende complemento de bFGF.

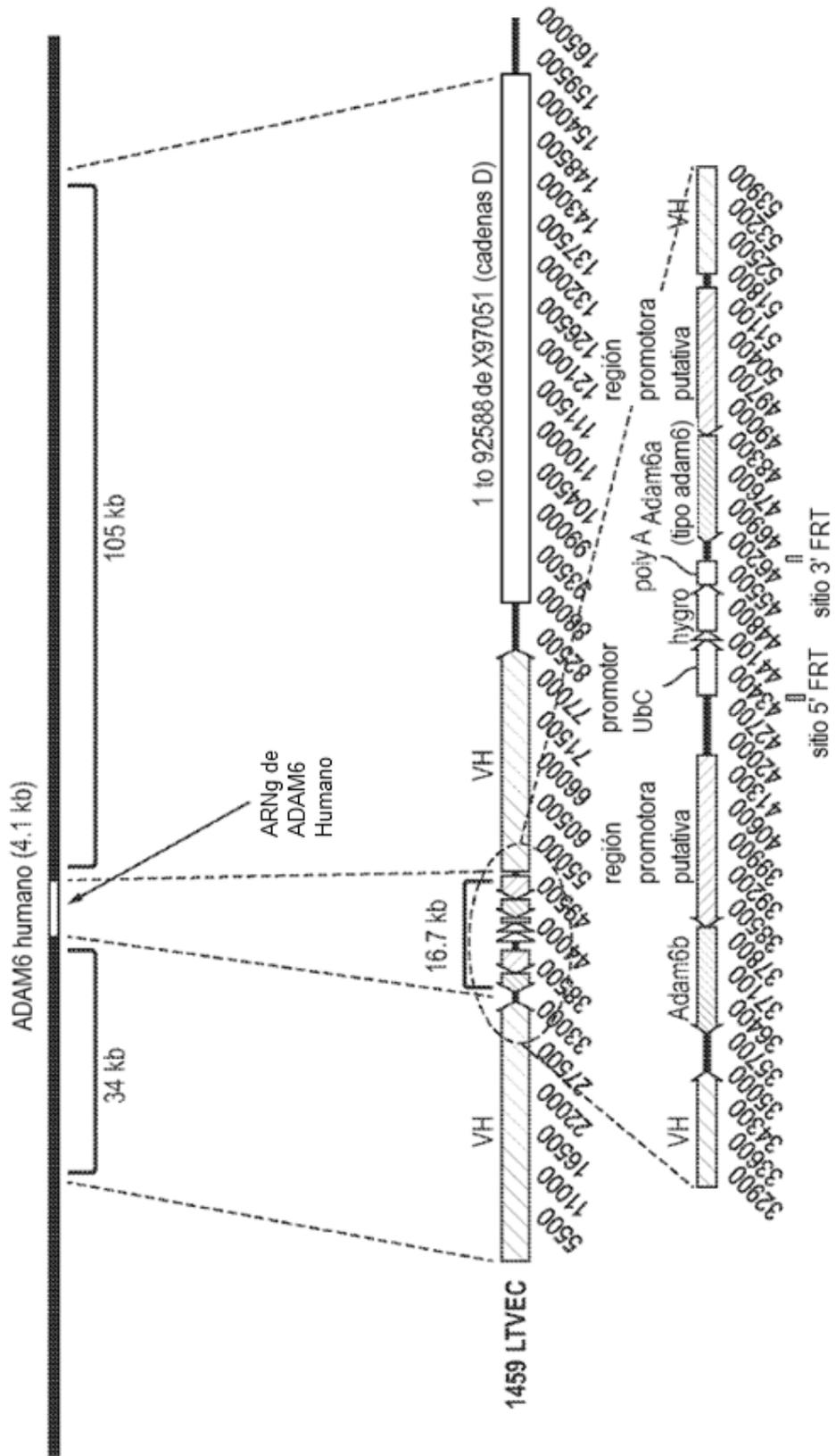
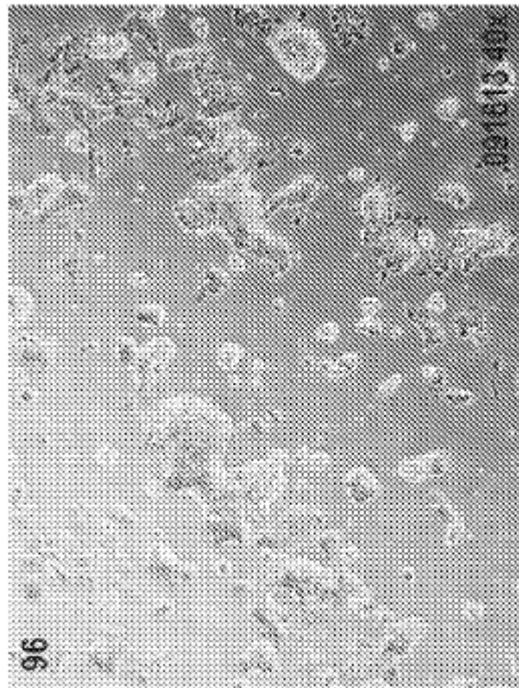
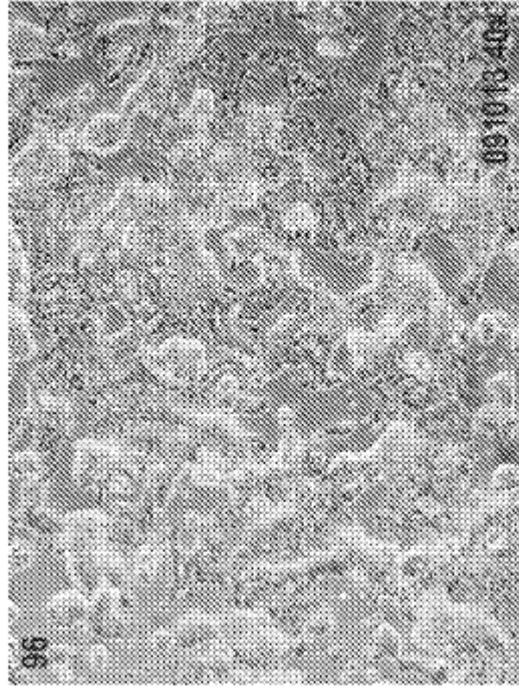


FIG. 1



8 dias en 2i

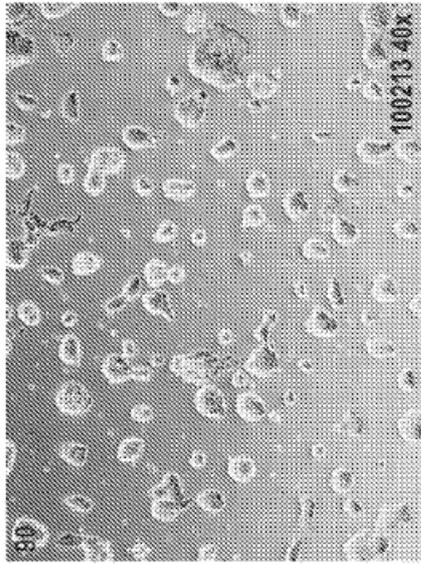
**FIG. 2A**



12 dias en 2i

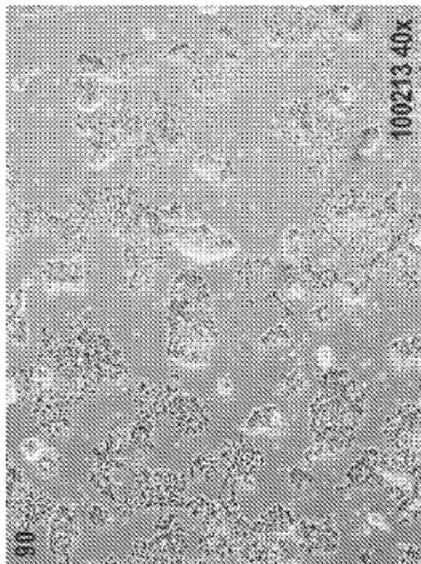
**FIG. 2B**





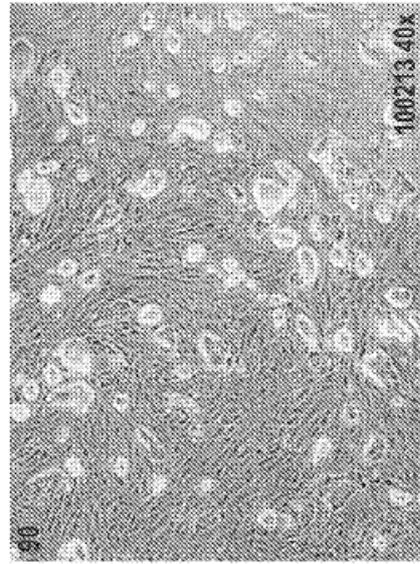
hVG2i Matrigel

**FIG. 3B**



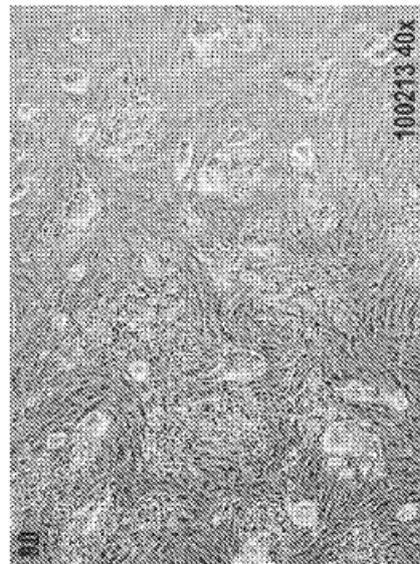
mTeSR-hLIF Matrigel

**FIG. 3A**



hVG2i alimentadoras de NuFF

**FIG. 3D**



mTeSR-hLIF alimentadoras de NuFF

**FIG. 3C**

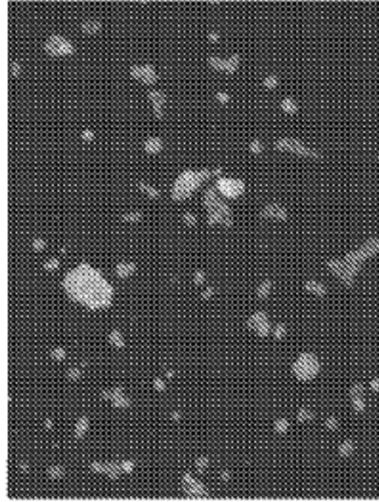


FIG. 4C

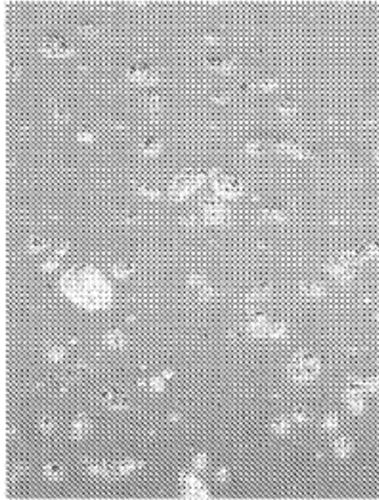


FIG. 4B

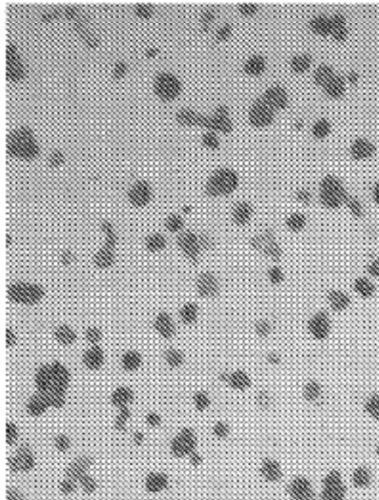


FIG. 4A

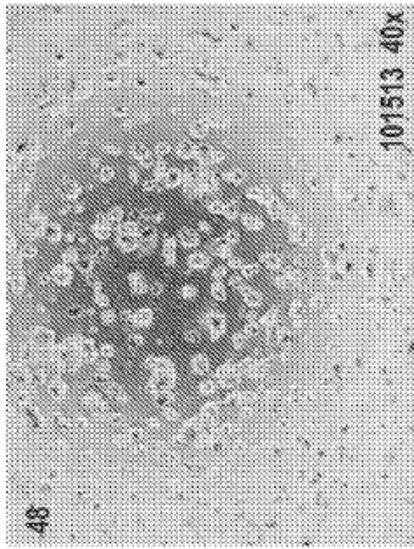
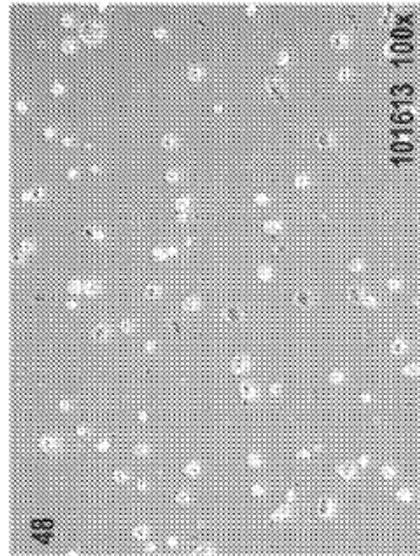


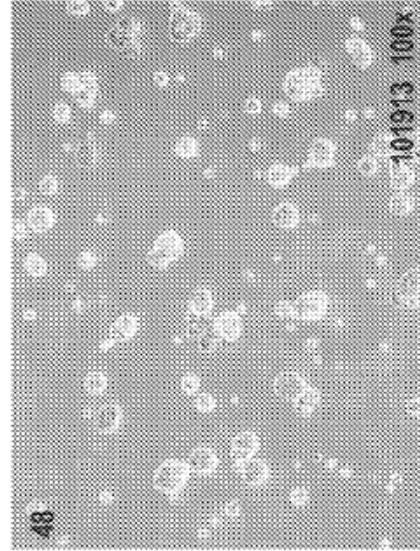
FIG. 5A

Pase de célula  
única mediante  
Trypsina sin  
inhibidor de Rock



1 día después de pase

FIG. 5B



4 días después de pase

FIG. 5C



FIG. 6A

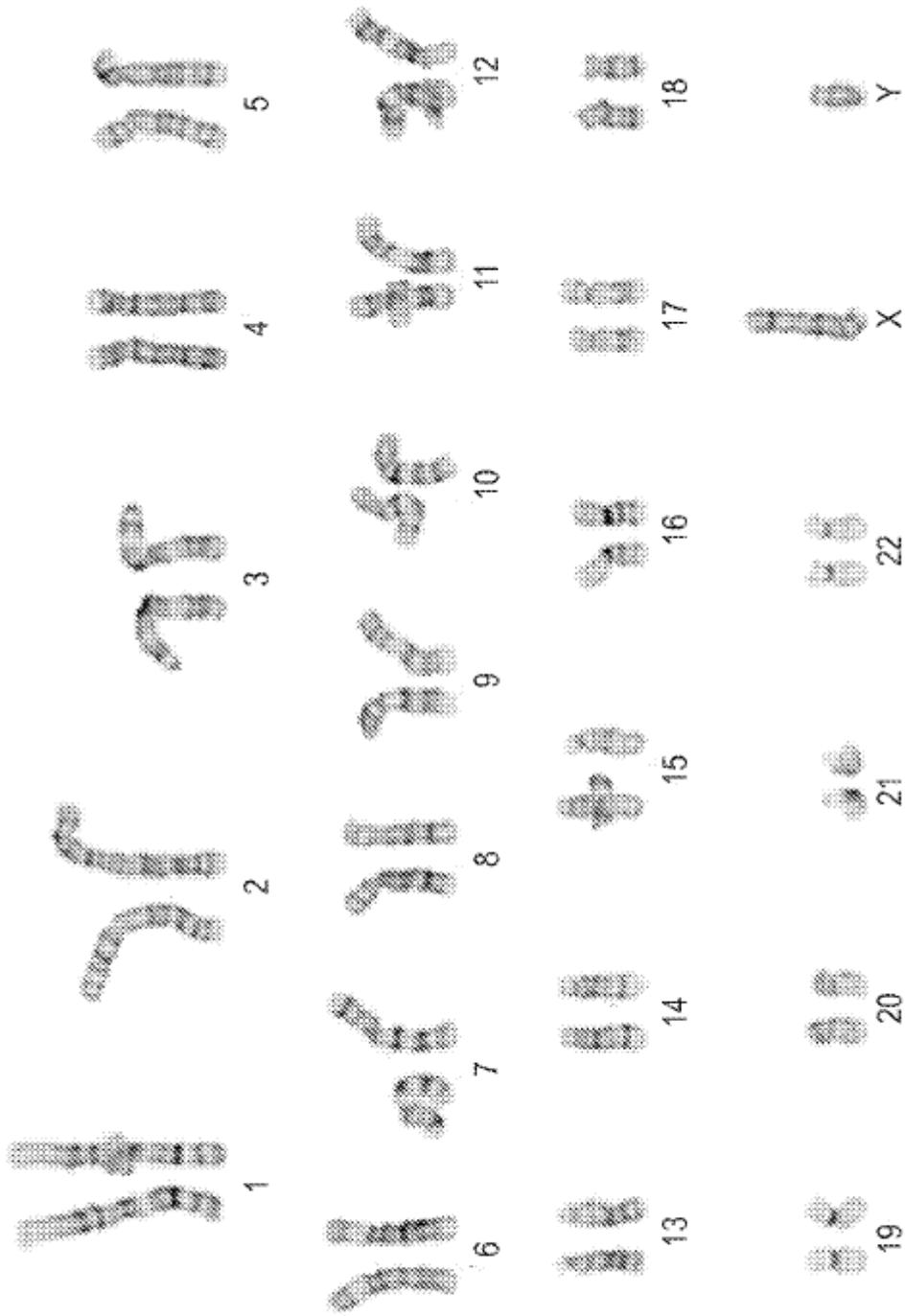


FIG. 6B