

## **CÉLULAS DONANTES UNIVERSALES**

### **REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS**

La presente solicitud reivindica el beneficio de la Solicitud Provisional de EE.UU. N.º 62/896,473, presentada el 5 de septiembre de 2019, y la Solicitud Provisional de EE.UU. N.º 62/979,771, presentada el 21 de febrero de 2020, cuyas divulgaciones se incorporan por la presente como referencia en su totalidad.

### **INCORPORACIÓN POR REFERENCIA DEL LISTADO DE SECUENCIAS**

Esta solicitud contiene un Listado de Secuencias que se ha presentado en formato ASCII a través de EFS-Web y se incorpora por referencia en su totalidad. La copia ASCII, creada el 2 de septiembre de 2020, se llama CT124-PCT-100867-666508-Sequence-Listing\_ST25.txt y tiene un tamaño de aproximadamente 53,000 bytes.

### **CAMPO DE LA INVENCIÓN**

La invención se refiere al campo de la edición génica y, en algunas realizaciones, a modificaciones genéticas con el fin de generar células que sean compatibles con múltiples sujetos, por ejemplo, células donantes universales.

### **ANTECEDENTES**

Se han propuesto diversos enfoques para superar el rechazo alógeno de células trasplantadas o injertadas, incluyendo la histocompatibilidad (*HLA-matching*, siglas del inglés *Human Leukocyte Antigens*, antígenos leucocitarios humanos), las vías de bloqueo que desencadenan la activación de los linfocitos T con anticuerpos, el uso de un cóctel de fármacos inmunosupresores, y la terapia con células autólogas. Otra estrategia para amortiguar el rechazo de los injertos implica la minimización de las diferencias alógenas entre las células trasplantadas o injertadas y el receptor. Los antígenos leucocitarios humanos (HLA) expresados en la superficie celular, moléculas codificadas por genes ubicados en el complejo principal de histocompatibilidad humano en el cromosoma 6, son los principales mediadores del rechazo inmunitario. El emparejamiento erróneo de un solo gen HLA entre el donante y el sujeto puede provocar una fuerte respuesta inmunitaria (Fleischhauer K. *et al.* "Bone marrow-allograft rejection by T lymphocytes recognizing a single amino acid difference in HLA-B44", *N Engl J Med.*, 1990, 323:1818–1822). Los genes HLA se dividen en MHC (complejo principal de histocompatibilidad, por sus siglas en inglés) de clase I (MHC-I) y MHC de clase II (MHC-II). Los genes MHC-I (HLA-A, HLA-B y HLA-C) se expresan en casi todos los tipos de células tisulares, presentando péptidos procesados con antígenos "no propios" a los linfocitos T CD8+, potenciando así su activación en los linfocitos T CD8+ citolíticos. Las células trasplantadas o injertadas que expresan moléculas MHC-I "no propias" provocarán una potente respuesta inmunitaria celular dirigida a estas células y, en última instancia, darán como resultado su desaparición por medio de los linfocitos T CD8+ citolíticos activados. Las proteínas MHC-I están

íntimamente asociadas con la beta-2-microglobulina (B2M) en el retículo endoplásmico, que es esencial para formar moléculas MHC-I funcionales en la superficie celular.

En contraste con la amplia expresión celular de los genes MHC-I, la expresión de los genes MHC-II está restringida a las células presentadoras de antígenos tales como las células dendríticas, los macrófagos y los linfocitos B. Los genes de antígenos HLA son los genes más polimórficos observados en el genoma humano (Rubinstein P., HLA matching for bone marrow transplantation--how much is enough?" *N Engl J Med.*, 2001, 345:1842–1844). La generación de una célula "donante universal" que sea compatible con cualquier genotipo de HLA proporciona una estrategia alternativa que podría resolver el rechazo inmunitario y los costes económicos asociados de las metodologías actuales para la evasión inmunitaria.

Para generar una línea de este tipo de una o más células donantes universales, un enfoque previo ha sido interrumpir funcionalmente la expresión de los genes de clase MHC-I y MHC-II. Esto podría lograrse mediante la interrupción genética, por ejemplo, de ambos alelos genéticos que codifican la cadena ligera de MHC-I, B2M. Cabría esperar que la línea de linfocitos nulos en B2M resultante y sus derivados presenten MHC-I de superficie muy reducida y, por lo tanto, inmunogenicidad reducida hacia los linfocitos T CD8+ alogénicos. El enfoque del direccionamiento de nucleasas efectoras de tipo activador de la transcripción (TALEN, por sus siglas en inglés, *transcription activator-like effector nuclease*) se ha usado para generar líneas de hESC deficientes en B2M mediante la delección de algunos nucleótidos en el exón 2 del gen B2M (Lu, P. *et al.*, "Generating hypoinmunogenic human embryonic stem cells by the disruption of beta 2-microglobulin", *Stem Cell Rev.* 2013, 9:806–813). Aunque las líneas de hESC dirigidas a B2M parecían ser deficientes en HLA-I superficial, se encontró que todavía contenían ARNm específicos de B2M y MHC-I. Los ARNm de B2M y MHC-I se expresaron a niveles equivalentes a los de las hESC no dirigidas (tanto constitutivas como inducidas por IFNg). Por lo tanto, existe la preocupación de que estas líneas de hESC dirigidas a B2M por TALEN puedan expresar MHC-I de superficie celular residual que sería suficiente para provocar rechazo inmunitario, como se ha observado con células de ratón B2M2/2 que también expresan ARNm de B2M (Gross, R. y Rappuoli, R. "Pertussis toxin promoter sequences involved in modulation", *Proc Natl Acad Sci*, 1993, 90:3913–3917). Aunque las líneas de hESC dirigidas a B2M por TALEN no se examinaron para detectar eventos de escisión inespecífica, la aparición de escisiones inespecíficas cuando se usan las TALEN sigue siendo un problema significativo que supondría una preocupación de seguridad importante en su uso clínico (Grau, J. *et al.* "TALENoffer: genome-wide TALEN off-target prediction", *Bioinformatics*, 2013, 29:2931–2932; Guilinger J. P. *et al.* "Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity", *Nat Methods* 2014, 11:429–435). Además, otro informe generó células iPSC que escaparon al reconocimiento alogénico mediante la desactivación de un primer alelo

B2M y la inserción de un gen HLA-E en un segundo alelo B2M, lo que dio como resultado la expresión superficial de dímeros o trímeros de HLA-E en ausencia de expresión superficial de HLA-A, HLA-B o HLA-C (Gornalusse, G.G. *et al.*, "HLA-E-expressing pluripotent stem cells escape allogeneic responses and lysis by NK cells", *Nature Biotechnology*, 2017, 35, 765-773).

Una posible limitación de algunas de las estrategias anteriores es que las células negativas en MHC de clase I son susceptibles a la lisis por los linfocitos citolíticos naturales (NK, *Natural Killer*), ya que las moléculas HLA sirven como inhibidores de ligandos principales para los linfocitos citolíticos naturales (NK). Se ha demostrado que los linfocitos citolíticos naturales (NK) del huésped, eliminan las células donantes B2M-/- injertadas o trasplantadas, y se produce un fenómeno similar *in vitro* con las líneas leucémicas humanas negativas en MHC de clase I (Bix, M. *et al.*, "Rejection of class I MHC-deficient haemopoietic cells by irradiated MHC-matched mice", *Nature*, 1991, 349, 329-331; Zarcone, D. *et al.*, "Human leukemia-derived cell lines and clones as models for mechanistic analysis of natural killer cell-mediated cytotoxicity", *Cancer Res.* 1987, 47, 2674-2682). Por lo tanto, existe la necesidad de mejorar los métodos anteriores para generar células donantes universales que puedan evadir la respuesta inmunitaria, así como la necesidad de generar células que puedan sobrevivir después del injerto. Como se describe en el presente documento, la supervivencia celular después del injerto puede estar mediada por una serie de otras vías independientes de rechazo alógeno, por ejemplo, hipoxia, especies reactivas de oxígeno, privación de nutrientes y estrés oxidativo. También, como se describe en el presente documento, la introducción genética de factores de supervivencia (genes y/o proteínas) puede ayudar a las células a sobrevivir después del injerto. Como se describe en el presente documento, una línea celular donante universal puede combinar propiedades que abordan tanto el rechazo alógeno como la supervivencia después del injerto.

### **SUMARIO**

En algunos aspectos, la presente divulgación abarca un método para generar una célula donante universal. El método comprende suministrar a una célula (a) una nucleasa dirigida al sitio que se dirige a un sitio dentro o cerca de un gen que codifica un factor de supervivencia y (b) un ácido nucleico que comprende una secuencia de nucleótidos que codifica un factor tolerogénico que está flanqueado por (i) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana de (a) y (ii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana de (a), en donde la nucleasa dirigida al sitio escinde el sitio diana de (a) y el ácido nucleico de (b) se inserta en un sitio que se solapa parcialmente con, se solapa completamente con o está contenido dentro del sitio de (a), generando de este modo una célula donante universal, en donde la célula donante universal

tiene una mayor supervivencia celular en comparación con una célula en la que el ácido nucleico de (b) no se ha insertado.

En algunas realizaciones, el factor de supervivencia es TXNIP, ZNF143, FOXO1, JNK o MANF, y el factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47. En realizaciones específicas, el factor de supervivencia es TXNIP y el factor tolerogénico es HLA-E. En realizaciones en las que la nucleasa dirigida al sitio es un sistema CRISPR que comprende una nucleasa CRISPR y un ARN guía (ARNg), la nucleasa CRISPR es una nucleasa Cas9 de tipo II o una nucleasa Cfp1 de tipo V, y la nucleasa CRISPR está unida a al menos una señal de localización nuclear. En algunas realizaciones, el ARNg se dirige a una secuencia polinucleotídica escogida entre las SEQ ID NO: 15-24 o 45-54, y (i) consiste esencialmente en una secuencia de nucleótidos de la SEQ ID NO: 25 y (ii) consiste esencialmente en una secuencia de nucleótidos de la SEQ ID NO: 32.

En algunas realizaciones, el método comprende además suministrar a la célula (c) una nucleasa dirigida al sitio que se dirige a un sitio dentro o cerca de un gen que codifica uno o más de los antígenos leucocitarios humanos MHC-I o MHC-II o un componente o un regulador de la transcripción de un complejo MHC-I o MHC-II, y (d) un ácido nucleico que comprende una secuencia de nucleótidos que codifica un factor tolerogénico que está flanqueado por (iii) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana de (c) y una (iv) secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana de (c), en donde el factor tolerogénico de (d) difiere del factor tolerogénico (b), en donde la nucleasa dirigida al sitio escinde el sitio diana de (c) y el ácido nucleico de (d) se inserta en un sitio que se solapa parcialmente con, se solapa completamente con o está contenido dentro del sitio de (c), en donde la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula en la que el ácido nucleico de (d) no se ha insertado.

En algunas realizaciones, el gen que codifica el uno o más antígenos leucocitarios humanos MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo de MHC-I o MHC-II es un gen MHC-I escogido entre HLA-A, HLA-B o HLA-C, un gen MHC-II escogido entre HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ o HLA-DR, o un gen escogido entre B2M, NLRC5, CIITA, RFX5, RFXAP o RFXANK, y el factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47. En realizaciones específicas, el gen que codifica el uno o más antígenos leucocitarios humanos MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo MHC-I o MHC-II es B2M, y el factor tolerogénico es PD-L1. En realizaciones en las que la nucleasa dirigida al sitio es un sistema CRISPR que comprende una nucleasa CRISPR y un ARNg, la nucleasa CRISPR es una nucleasa Cas9 de tipo II o una nucleasa Cfp1 de tipo V, y la nucleasa CRISPR está unida a al menos una señal de localización

nuclear. En algunas realizaciones, el ARNg se dirige a una secuencia polinucleotídica escogida entre las SEQ ID NO: 1-3 o 35-44, y (iii) consiste esencialmente en una secuencia de nucleótidos de la SEQ ID NO: 7, y (iv) consiste esencialmente en una secuencia de nucleótidos de la SEQ ID NO: 13.

5 En algunas realizaciones, la secuencia de nucleótidos que codifica un factor tolerogénico de (b) y (d) está unida operativamente a un promotor exógeno. El promotor exógeno puede escogerse entre un promotor constitutivo, inducible, temporal, específico del tejido o específico del tipo de célula. En algunas realizaciones, el promotor exógeno es un promotor de CMV, EF1a, PGK, CAG o UBC. En realizaciones específicas, el promotor exógeno es un promotor CAG.

10 La presente divulgación también abarca las células donantes universales generadas mediante los métodos desvelados en el presente documento. En algunas realizaciones, la célula es una célula de mamífero. En algunas realizaciones, la célula es una célula humana. En algunas realizaciones, la célula es una célula madre. En algunas realizaciones, la célula es una célula madre pluripotente (PSC, por sus siglas en inglés), una célula madre embrionaria (ESC, por sus siglas en inglés), una célula madre adulta (ASC, por sus siglas en inglés), una célula madre pluripotente inducida (iPSC, por sus siglas en inglés) o una célula madre o progenitora hematopoyética (HSPC, por sus siglas en inglés) (también denominada una célula madre hematopoyética (HSC, por sus siglas en inglés)). En algunas realizaciones, la célula es una célula diferenciada. En algunas realizaciones, la célula es una célula somática.

20 En general, las células donantes universales desveladas en el presente documento son capaces de diferenciarse en células progenitoras de linaje restringido o células somáticas completamente diferenciadas. En algunas realizaciones, las células progenitoras de linaje restringido son células progenitoras del endodermo pancreático, células progenitoras endocrinas pancreáticas, células progenitoras mesenquimatosas, células progenitoras musculares, blastocitos, células progenitoras hematopoyéticas o células progenitoras neuronales, y las células somáticas completamente diferenciadas son células secretoras endocrinas tales como células beta pancreáticas, células epiteliales, células endodérmicas, macrófagos, hepatocitos, adipocitos, células renales, células sanguíneas o células del sistema inmunitario. En algunas realizaciones, las células somáticas completamente diferenciadas son cardiomiocitos.

30 Un aspecto adicional de la presente divulgación proporciona un método para tratar a un sujeto que lo necesite, en donde el método comprende obtener o haber obtenido las células donantes universales como se desvela en el presente documento después de la diferenciación en células progenitoras de linaje restringido o células somáticas completamente diferenciadas, y administrar al sujeto las células progenitoras de linaje restringido o las células somáticas

35

completamente diferenciadas. También se proporciona un método de obtención de células para su administración a un sujeto que lo necesite, comprendiendo el método obtener o haber obtenido las células donantes universales como se desvela en el presente documento, y mantener las células donantes universales durante un tiempo y en condiciones suficientes para que las células se diferencien en células progenitoras de linaje restringido o células somáticas completamente diferenciadas. En algunas realizaciones, el sujeto es un ser humano que tiene, se sospecha que tiene o está en riesgo de tener una enfermedad. En algunas realizaciones, la enfermedad es una enfermedad hereditaria genéticamente.

Otro aspecto más de la presente divulgación abarca un ARNg dirigido a una secuencia polinucleotídica escogida entre las SEQ ID NO: 15-24 o 45-54.

Si bien la divulgación es susceptible de diversas modificaciones y formas alternativas, las realizaciones específicas de la misma se muestran a modo de ejemplo en los dibujos y se describirán detalladamente en el presente documento. Debe entenderse, sin embargo, que los dibujos y la descripción detallada presentados en el presente documento no pretenden limitar la divulgación a las realizaciones particulares desveladas, sino que, por el contrario, la intención es cubrir todas las modificaciones, equivalentes y alternativas que pertenecen al espíritu y alcance de la presente divulgación tal como se define en las reivindicaciones adjuntas.

Otras características y ventajas de la presente divulgación resultarán evidentes en la siguiente descripción detallada de realizaciones de la presente invención, tomadas con referencia a los dibujos adjuntos.

### **BREVE DESCRIPCIÓN DE LOS DIBUJOS**

La FIG. 1 muestra el análisis por TIDE (*Tracking of Indels by Decomposition*, rastreo de indel mediante descomposición) del corte con ARNg de B2M en células CyT49. Se sometieron a ensayo ARNg de B2M-1, B2M-2 y B2M-3.

Las FIG. 2A-2B muestran la evaluación por citometría de flujo de la expresión de B2M con y sin IFN- $\gamma$  en células CyT49 WT (*wild type*, tipo silvestre) (FIG. 2A) y células CyT49 con KO (*Knock-Out*, inactivación) de B2M (FIG. 2B).

La FIG. 3 muestra el mapa de plásmido del vector donante B2M-CAGGS-PD-L1 para la HDR (*Homology-Directed Repair*, reparación dirigida por homología).

La FIG. 4 muestra el análisis por citometría de flujo para determinar la pluripotencia de células madre CyT49 con KO de B2M/KI (*Knock-In*, inserción) de PD-L1. Los clones derivados fueron > 99 % doble positivos para OCT4 y SOX2, dos factores de transcripción vitales para la pluripotencia. Se utilizó IgG como control negativo.

Las FIG. 5A-5B muestran el análisis por citometría de flujo de CyT49 WT (FIG. 5A) y clones de células madre derivados con KO de B2M/KI de PD-L1 (FIG. 5B). Las células WT regulan positivamente la expresión de B2M en respuesta a IFN $\gamma$ . Los clones con KO de B2M/KI

de PD-L1 expresan completamente PD-L1 y no expresan B2M con o sin tratamiento con IFN $\gamma$ . NT-1 = sin tratamiento. INTG-1 = células tratadas durante 48 horas con IFN $\gamma$  50 ng/ml.

La FIG. **6** muestra la citometría de flujo para FOXA2 y SOX17 en células en fase 1 (endodermo definitivo) diferenciadas a partir de células CyT49 de tipo silvestre, con KI de PD-L1/KO de B2M o CyT49 con KO de B2M.

La FIG. **7** muestra el porcentaje cuantitativo de la expresión de FOXA2 y SOX17 en células en fase 1 (endodermo definitivo) diferenciadas a partir de células de tipo silvestre, con KI de PD-L1/KO de B2M o KO de B2M.

La FIG. **8** muestra el porcentaje cuantitativo de la expresión de CHGA, PDX1 y NKX6.1 en células en fase 4 (PEC) diferenciadas a partir de células de tipo silvestre, con KI de PD-L1/KO de B2M o KO de B2M.

La FIG. **9** muestra poblaciones heterogéneas de células en fase 4 (PEC).

La FIG. **10** muestra la expresión de genes seleccionados durante la evolución temporal de la diferenciación en células diferenciadas a partir de células de tipo silvestre, con KI de PD-L1/KO de B2M o con KO de B2M.

Las FIG. **11A-11F** muestran la expresión de genes seleccionados durante la evolución temporal de la diferenciación en células diferenciadas a partir de células de tipo silvestre, con KI de PD-L1/KO de B2M o con KO de B2M. La FIG. **11A** muestra la expresión de B2M en células de tipo silvestre. La FIG. **11B** muestra la expresión de B2M en células con KO de B2M. La FIG. **11C** muestra la expresión de B2M en células con KI de PD-L1/KO de B2M. La FIG. **11D** muestra la expresión de PD-L1 en células de tipo silvestre. La FIG. **11E** muestra la expresión de PD-L1 en células con KO de B2M. La FIG. **11F** muestra la expresión de PD-L1 en células con KI de PD-L1/KO de B2M.

Las FIG. **12A-12F** muestran la expresión de MHC de clase I y de clase II en la fase de PEC en células diferenciadas a partir de células de tipo silvestre, con KI de PD-L1/KO de B2M o KO de B2M. La FIG. **12A** muestra la expresión de MHC de clase I en células de tipo silvestre. La FIG. **12B** muestra la expresión de MHC de clase I en células con KO de B2M. La FIG. **12C** muestra la expresión de MHC de clase I en células con KI de PD-L1/KO de B2M. La FIG. **12D** muestra la expresión de PD-L1 de MHC de clase II en células de tipo silvestre. La FIG. **12E** muestra la expresión de MHC de clase II en células con KO de B2M. La FIG. **12F** muestra la expresión de MHC de clase II en células con KI de PD-L1/KO de B2M.

La FIG. **13** muestra el análisis por TIDE del corte con ARNg de TXNIP en hiPSC TC1133. La guía T5 parecía ser la mejor para cortar en el exón 1.

La FIG. **14** muestra el mapa de plásmido del vector donante TXNIP-CAGGS-HLA-E para la HDR.

La FIG. **15** muestra el análisis por citometría de flujo para determinar la pluripotencia de células madre CyT49 con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E. Los clones derivados fueron > 99 % doble positivos para OCT4 y SOX2, dos factores de transcripción vitales para la pluripotencia. Los clones tampoco expresan B2M. Los clones no expresan MHC-I.

La FIG. **16** muestra el análisis por citometría de flujo para determinar la pluripotencia de células madre CyT49 con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E. Los clones derivados expresan PD-L1 y HLA-E después de experimentar la diferenciación a la fase 6 (células beta inmaduras). Se utilizó IgG como control negativo.

La FIG. **17** muestra el porcentaje cuantitativo de la expresión de CHGA, PDX1 y NKX6.1 en células en fase 4 (PEC) diferenciadas a partir de células hESC de tipo silvestre, con KO de B2M, con KI de PD-L1/KO de B2M (V1A) o con KO de TXNIP/KI de HLA-E (V1B).

Las FIG. **18A-18B** muestran la expresión de genes seleccionados durante la evolución temporal de la diferenciación en células con KO de TXNIP (FIG. **18A**) o células con KO de TXNIP/KI de HLA-E (V1B) (FIG. **18B**).

Las FIG. **19A-19B** muestran el análisis por citometría de flujo para determinar la activación de linfocitos T usando el ensayo de proliferación de CFSE (*CarboxyFluorescein Succinimidyl Ester*, éster de succinimidil-carboxifluoresceína). Se coincubaron linfocitos T CD3+ primarios humanos con PEC derivadas de clones de CyT49 WT, con KO de B2M, con KO de B2M/KI de PD-L1 o con KO de B2M/KI de PD-L1 + KO de TXNIP/KI de HLA-E (FIG. **19A**). La FIG. **19B** resume la activación de los linfocitos T en las diversas células. ANOVA unidireccional ( $\alpha = 0.05$  con el ensayo de comparaciones múltiples de Dunnett) con "CFSE-T solo" establecido como control. \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ . n.s. = no significativo.

La FIG. **20** muestra la expresión de genes seleccionados a lo largo de un curso temporal de diferenciación de células diferenciadas a partir de células con KO de TXNIP.

La FIG. **21** muestra la evaluación por citometría de flujo de la expresión de PDX1 y NKX6.1 en células PEC diferenciadas a partir de células con KO de TXNIP.

La FIG. **22** muestra la morfología de los diversos clones con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E ("S6-V1B-H9", "S6-V1B-3B11", "S6-V1B-1G7" y "S6-V1B-3C2") en comparación con las células de tipo silvestre ("WT") y las de control con guía no de corte ("NCG#1") después de la diferenciación a la etapa 6.

Las FIG. **23A-23F** muestran la expresión de genes seleccionados de los clones después de la diferenciación a la fase 6. La FIG. **23A** muestra la expresión de genes seleccionados a lo largo de un curso temporal de diferenciación de células diferenciadas a partir de un clon con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E de ejemplo. Las FIG. **23B-23F** muestran la expresión génica de *INS* (FIG. **23B**), *NKX6.1* (FIG. **23C**), *GCG* (FIG. **23D**), *SST* (FIG. **23E**) y

GCK (FIG. 23F) en células de tipo silvestre ("S6-Cyt49 WT"), células de control con guía no de corte ("S6-NCG#1") y diversos clones con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E ("S6-V1B-H9", "S6-V1B-3B11", "S6-V1B-1G7" y "S6-V1B-3C2") que se diferenciaron a la fase 6 con el clon con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E indiferenciados ("ES-V1B-H9") e islotes de tipo silvestre ("Islotes") como controles.

Las FIG. 24A-24B muestran la evaluación por citometría de flujo de la expresión de INS y GCG (FIG. 24A) y de la expresión de INS y NKX6.1 (FIG. 24B) en células en fase 6 diferenciadas a partir de un clon con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E.

Las FIG. 25A-25B muestran el porcentaje de expresión de INS (FIG. 25A) y de expresión de NKX6.1 (FIG. 25B) en las células en Fase 6 diferenciadas a partir de células de tipo silvestre ("S6-WT"), células de control con guía no de corte ("S6-NCG#1") y dos clones con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E ("S6-V1B003" y "V1B-H9").

La FIG. 26A muestra la evaluación por citometría de flujo de la expresión de PDX1 y NKX6.1 en células en fase 4 diferenciadas a partir de células del clon 1 (KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E) con densidades de siembra diferentes.

La FIG. 26B muestra la evaluación por citometría de flujo de la expresión de PD-L1 y HLA-E en células en fase 4 diferenciadas a partir de células del clon 1 (KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E).

Las FIG. 27A-27C muestran el análisis por caracterización de un clon de siembra diferenciado a la fase de PEC. La FIG. 27A muestra la morfología, la FIG. 27B muestra la expresión de genes seleccionados a lo largo de un curso temporal de diferenciación y la FIG. 27C muestra el porcentaje de células que expresan CHGA/NKX6.1<sup>+</sup>/PDX1<sup>+</sup> en la población diferenciada.

La FIG. 28 muestra la expresión de genes seleccionados a lo largo de un curso temporal de diferenciación de células diferenciadas a partir de clones con KO de TXNIP/KI de HLA-E.

## **DESCRIPCIÓN DETALLADA**

### **I. Definiciones**

**Delección:** Como se usa en el presente documento, el término "delección", que puede usarse indistintamente con la expresión "delección genética" o el término "desactivación", se refiere generalmente a una modificación genética en donde se elimina un sitio o una región de ADN genómico mediante cualquier método de biología molecular, por ejemplo, métodos descritos en el presente documento, por ejemplo, mediante la administración a un sitio de ADN genómico de una endonucleasa y al menos un ARNg. Se puede eliminar cualquier número de nucleótidos. En algunas realizaciones, una delección implica la eliminación de al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos diez, al menos quince, al menos veinte o al menos 25 nucleótidos. En algunas realizaciones, una delección implica la

eliminación de 10-50, 25-75, 50-100, 50-200 o más de 100 nucleótidos. En algunas realizaciones, una delección implica la eliminación de un gen diana completo, por ejemplo, un gen B2M. En algunas realizaciones, una delección implica la eliminación de parte de un gen diana, por ejemplo, la totalidad o parte de un promotor y/o una secuencia codificante de un gen B2M. En algunas realizaciones, una delección implica la eliminación de un regulador de la transcripción, por ejemplo, una región promotora, de un gen diana. En algunas realizaciones, una delección implica la eliminación de la totalidad o parte de una región codificante, de manera que el producto normalmente expresado por la región codificante ya no se expresa, se expresa como una forma truncada o se expresa a un nivel reducido. En algunas realizaciones, una delección conduce a una disminución en la expresión de un gen en relación con una célula no modificada.

**Endonucleasa:** Como se usa en el presente documento, el término "endonucleasa" se refiere generalmente a una enzima que escinde enlaces fosfodiéster dentro de un polinucleótido. En algunas realizaciones, una endonucleasa escinde específicamente enlaces fosfodiéster dentro de un polinucleótido de ADN. En algunas realizaciones, una endonucleasa es una nucleasa de dedos de cinc (ZFN, por sus siglas en inglés, *zinc finger nuclease*), una nucleasa efectora de tipo activador de la transcripción (TALEN, por sus siglas en inglés, *transcription activator-like effector nuclease*), una endonucleasa de asentamiento (HE, por sus siglas en inglés, *homologous endonuclease*), una meganucleasa, una MegaTAL o una endonucleasa asociada a CRISPR. En algunas realizaciones, una endonucleasa es una endonucleasa guiada por ARN. En determinados aspectos, la endonucleasa guiada por ARN es una nucleasa CRISPR, por ejemplo, una endonucleasa CRISPR Cas9 de tipo II o una endonucleasa CRISPR Cpf1 de tipo V. En algunas realizaciones, una endonucleasa es una endonucleasa Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocida como Csn1 y Csx12), Cas100, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 o Cpf1, o un homólogo de las mismas, una recombinación de la molécula de origen natural de las mismas, una versión con codones optimizados de las mismas o una versión modificada de las mismas, o combinaciones de las mismas. En algunas realizaciones, una endonucleasa puede introducir una o más roturas monocatenarias (SSB, por sus siglas en inglés) y/o una o más roturas bicatenarias (DSB, por sus siglas en inglés).

**Modificación genética:** Como se usa en el presente documento, la expresión "modificación genética" se refiere generalmente a un sitio de ADN genómico que se ha editado o manipulado genéticamente usando cualquier método de biología molecular, por ejemplo, métodos descritos en el presente documento, por ejemplo, mediante la administración a un sitio

de ADN genómico de una endonucleasa y al menos un ARNg. Las modificaciones genéticas de ejemplo incluyen inserciones, deleciones, duplicaciones, inversiones y translocaciones, y combinaciones de las mismas. En algunas realizaciones, una modificación genética es una deleción. En algunas realizaciones, una modificación genética es una inserción. En otras realizaciones, una modificación genética es una mutación por inserción-deleción (o indel), de modo que el marco de lectura del gen diana se desplaza, conduciendo a un producto génico alterado o a ningún producto génico.

**ARN guía (ARNg):** Como se usa en el presente documento, la expresión "ARN guía" o el término "ARNg" se refiere generalmente a ácido ribonucleico corto que puede interactuar con, por ejemplo, unirse a una endonucleasa y unirse o hibridarse a un sitio genómico o región genómica diana. En algunas realizaciones, un ARNg es un ARN guía de una sola molécula (ARNsg). En algunas realizaciones, un ARNg puede comprender una región de extensión espaciadora. En algunas realizaciones, un ARNg puede comprender una región de extensión de ARNtracr. En algunas realizaciones, un ARNg es monocatenario. En algunas realizaciones, un ARNg comprende nucleótidos de origen natural. En algunas realizaciones, un ARNg es un ARNg modificado químicamente. En algunas realizaciones, un ARNg modificado químicamente es un ARNg que comprende al menos un nucleótido con una modificación química, por ejemplo, una modificación de 2'-O-metil-azúcar. En algunas realizaciones, un ARNg modificado químicamente comprende una cadena principal de ácido nucleico modificado. En algunas realizaciones, un ARNg modificado químicamente comprende un residuo de 2'-O-metil-fosforotioato. En algunas realizaciones, un ARNg puede formar complejo previamente con una ADN endonucleasa.

**Inserción:** Como se usa en el presente documento, el término "inserción" que se puede usar indistintamente con la expresión "inserción genética" o "inserción", se refiere generalmente a una modificación genética en donde un polinucleótido se introduce o se añade a un sitio o una región de ADN genómico mediante cualquier método de biología molecular, por ejemplo, los métodos descritos en el presente documento, por ejemplo, mediante el suministro a un sitio de ADN genómico de una endonucleasa y al menos un ARNg. En algunas realizaciones, una inserción puede producirse dentro o cerca de un sitio de ADN genómico que ha sido el sitio de una modificación genética anterior, por ejemplo, una mutación por deleción o por inserción-deleción. En algunas realizaciones, se produce una inserción en un sitio de ADN genómico que se solapa parcialmente con, se solapa completamente con o está contenido dentro de un sitio de una modificación genética anterior, por ejemplo, una mutación por deleción o por inserción-deleción. En algunas realizaciones, se produce una inserción en un locus seguro. En algunas realizaciones, una inserción implica la introducción de un polinucleótido que codifica una proteína de interés. En algunas realizaciones, una inserción implica la introducción

de un polinucleótido que codifica un factor tolerogénico. En algunas realizaciones, una inserción implica la introducción de un polinucleótido que codifica un factor de supervivencia. En algunas realizaciones, una inserción implica la introducción de un promotor exógeno, por ejemplo, un promotor constitutivo, por ejemplo, un promotor CAG. En algunas realizaciones, una inserción implica la introducción de un polinucleótido que codifica un gen no codificante. En general, un polinucleótido que se va a insertar está flanqueado por secuencias (por ejemplo, brazos de homología) que tienen una homología de secuencia sustancial con el ADN genómico en o cerca del sitio de inserción.

**Complejo principal de histocompatibilidad de clase I (MHC-I):** Como se usan en el presente documento, la expresión "complejo principal de histocompatibilidad de clase I" o el término "MHC-I" se refieren generalmente a una clase de biomoléculas que se encuentran sobre la superficie celular de todas las células nucleadas en vertebrados, incluyendo mamíferos, por ejemplo, seres humanos; y que funcionan para mostrar péptidos de antígenos foráneos o no propios, por ejemplo, proteínas, desde el interior de la célula (es decir, citosólicos) a linfocitos T citotóxicos, por ejemplo, linfocitos T CD8+, con el fin de estimular una respuesta inmunitaria. En algunas realizaciones, una biomolécula MHC-I es un gen MHC-I o una proteína MHC-I. La complejación de las proteínas MHC-I con la proteína beta-2 microglobulina (B2M) es necesaria para la expresión en la superficie celular de todas las proteínas MHC-I. En algunas realizaciones, la disminución de la expresión de un antígeno leucocitario humano (HLA) MHC-I en relación con una célula no modificada implica una disminución (o reducción) en la expresión de un gen MHC-I. En algunas realizaciones, la disminución de la expresión de un antígeno leucocitario humano (HLA) MHC-I en relación con una célula no modificada implica una disminución (o reducción) en la expresión en la superficie celular de una proteína MHC-I. En algunas realizaciones, una biomolécula MHC-I es HLA-A (ID del gen del NCBI N.º: 3105), HLA-B (ID del gen del NCBI N.º: 3106), HLA-C (ID del gen del NCBI N.º: 3107) o B2M (ID del gen del NCBI N.º: 567).

**Complejo principal de histocompatibilidad de clase II (MHC-II):** Como se usa en el presente documento, la expresión "complejo principal de histocompatibilidad de clase II" o el término "MHC-II" se refiere generalmente a una clase de biomoléculas que se encuentran normalmente en la superficie celular de células presentadoras de antígeno en vertebrados, incluyendo mamíferos, por ejemplo, seres humanos; y que funcionan para mostrar péptidos de antígenos no propios o foráneos, por ejemplo, proteínas, desde fuera de la célula (extracelular) a los linfocitos T citotóxicos, por ejemplo, linfocitos T CD8+, con el fin de estimular una respuesta inmunitaria. En algunas realizaciones, una célula presentadora de antígenos es una célula dendrítica, un macrófago o un linfocito B. En algunas realizaciones, una biomolécula MHC-II es un gen MHC-II o una proteína MHC-II. En algunas realizaciones, la disminución de la expresión

de un antígeno leucocitario humano (HLA) MHC-II en relación con una célula no modificada implica una disminución (o reducción) en la expresión de un gen MHC-II. En algunas realizaciones, la disminución de la expresión de un antígeno leucocitario humano (HLA) MHC-II en relación con una célula no modificada implica una disminución (o reducción) en la expresión en la superficie celular de una proteína MHC-II. En algunas realizaciones, una biomolécula MHC-II es HLA-DPA (ID del gen del NCBI N.º: 3113), HLA-DPB (ID del gen del NCBI N.º: 3115), HLA-DMA (ID del gen del NCBI N.º: 3108), HLA-DMB (ID del gen del NCBI N.º: 3109), HLA-DOA (ID del gen del NCBI N.º: 3111), HLA-DOB (ID del gen del NCBI N.º: 3112), HLA-DQA (ID del gen del NCBI N.º: 3117), HLA-DQB (ID del gen del NCBI N.º: 3119), HLA-DRA (ID del gen del NCBI N.º: 3122) o HLA-DRB (ID del gen del NCBI N.º: 3123).

**Polinucleótido:** Como se usa en el presente documento, el término "polinucleótido", que puede usarse indistintamente con la expresión "ácido nucleico", se refiere generalmente a una biomolécula que comprende dos o más nucleótidos. En algunas realizaciones, un polinucleótido comprende al menos dos, al menos cinco, al menos diez, al menos veinte, al menos 30, al menos 40, al menos 50, al menos 100, al menos 200, al menos 250, al menos 500, o cualquier número de nucleótidos. Por ejemplo, los polinucleótidos pueden incluir al menos 500 nucleótidos, al menos aproximadamente 600 nucleótidos, al menos aproximadamente 700 nucleótidos, al menos aproximadamente 800 nucleótidos, al menos aproximadamente 900 nucleótidos, al menos aproximadamente 1000 nucleótidos, al menos aproximadamente 2000 nucleótidos, al menos aproximadamente 3000 nucleótidos, al menos aproximadamente 4000 nucleótidos, al menos aproximadamente 4500 nucleótidos o al menos aproximadamente 5000 nucleótidos. Un polinucleótido puede ser una molécula de ADN o ARN, o una molécula híbrida de ADN/ARN. Un polinucleótido puede ser monocatenario o bicatenario. En algunas realizaciones, un polinucleótido es un sitio o una región de ADN genómico. En algunas realizaciones, un polinucleótido es un gen endógeno que está comprendido dentro del genoma de una célula no modificada o una célula donante universal. En algunas realizaciones, un polinucleótido es un polinucleótido exógeno que no está integrado en el ADN genómico. En algunas realizaciones, un polinucleótido es un polinucleótido exógeno que está integrado en el ADN genómico. En algunas realizaciones, un polinucleótido es un plásmido o un vector vírico adenoasociado. En algunas realizaciones, un polinucleótido es una molécula circular o lineal.

**Locus seguro:** Como se usa en el presente documento, la expresión "locus seguro" se refiere generalmente a cualquier ubicación, sitio o región de ADN genómico que pueda alojar una inserción genética en dicha ubicación, sitio o región sin que se produzcan efectos adversos en una célula. En algunas realizaciones, un locus seguro es una región intragénica o extragénica. En algunas realizaciones, un locus seguro es una región de ADN genómico que normalmente es transcripcionalmente silenciosa. En algunas realizaciones, un locus seguro es

un locus de AAVS1 (PPP1 R12C), ALB, Angptl3, ApoC3, ASGR2, CCR5, FIX (F9), G6PC, Gys2, HGD, Lp(a), Pcsk9, Serpina1, TF o TTR. En algunas realizaciones, un locus seguro se describe en Sadelain, M. *et al.*, "Safe harbours for the integration of new DNA in the human genome" *Nature Reviews Cancer*, 2012, Vol 12, páginas 51-58.

5       **Interrupor de seguridad:** Como se usa en el presente documento, la expresión "interrupor de seguridad" se refiere generalmente a una biomolécula que conduce a una célula a experimentar apoptosis. En algunas realizaciones, un interrupor de seguridad es una proteína o un gen. En algunas realizaciones, un interrupor de seguridad es un gen suicida. En algunas realizaciones, un interrupor de seguridad, por ejemplo, timidina cinasa del virus del herpes  
10 simple (HSV-tk), conduce a una célula a experimentar apoptosis mediante la metabolización de un profármaco, por ejemplo, ganciclovir. En algunas realizaciones, la presencia sobreexpresada de un interrupor de seguridad por sí solo conduce a una célula a experimentar apoptosis. En algunas realizaciones, un interrupor de seguridad es una molécula basada en p53, HSV-tk o caspasa-9 inducible.

15       **Sujeto:** Como se usa en el presente documento, el término "sujeto" se refiere a un mamífero. En algunas realizaciones, un sujeto es un primate no humano o un roedor. En algunas realizaciones, un sujeto es un ser humano. En algunas realizaciones, un sujeto tiene, se sospecha que tiene o está en riesgo de tener una enfermedad o un trastorno. En algunas realizaciones, un sujeto tiene uno o más síntomas de una enfermedad o un trastorno.

20       **Factor de supervivencia:** Como se usa en el presente documento, la expresión "factor de supervivencia" se refiere generalmente a una proteína (por ejemplo, expresada por un polinucleótido como se describe en el presente documento) que, cuando aumenta o disminuye en una célula, permite que la célula, por ejemplo, una célula donante universal, sobreviva después del trasplante o injerto en un sujeto huésped con tasas de supervivencia superiores en  
25 relación con una célula no modificada. En algunas realizaciones, un factor de supervivencia es un factor de supervivencia humano. En algunas realizaciones, un factor de supervivencia es un miembro de una vía crítica implicada en la supervivencia celular. En algunas realizaciones, una vía crítica implicada en la supervivencia celular tiene implicaciones sobre la hipoxia, las especies reactivas de oxígeno, la privación de nutrientes y/o el estrés oxidativo. En algunas realizaciones,  
30 la modificación genética, por ejemplo, delección o inserción, de al menos un factor de supervivencia permite que una célula donante universal sobreviva durante un período de tiempo más largo, por ejemplo, al menos 1.05, al menos 1.1, al menos 1.25, al menos 1.5, al menos 2, al menos 3, al menos 4, al menos 5, al menos 10, al menos 20 o al menos 50 veces más tiempo que una célula no modificada después del injerto. En algunas realizaciones, un factor de  
35 supervivencia es ZNF143 (ID del gen del NCBI N.º: 7702), TXNIP (ID del gen del NCBI N.º: 10628), FOXO1 (ID del gen del NCBI N.º: 2308), JNK (ID del gen del NCBI N.º: 5599) o MANF

(ID del gen del NCBI N.º: 7873). En algunas realizaciones, se inserta un factor de supervivencia en una célula, por ejemplo, una célula donante universal. En algunas realizaciones, se elimina un factor de supervivencia de una célula, por ejemplo, una célula donante universal. En algunas realizaciones, una inserción de un polinucleótido que codifica MANF permite que una célula, por ejemplo, una célula donante universal, sobreviva después del trasplante o injerto en un sujeto huésped con tasas de supervivencia superiores en relación con una célula no modificada. En algunas realizaciones, una mutación por delección o por inserción-delección dentro o cerca de un gen ZNF143, TXNIP, FOXO1 o JNK permite que una célula, por ejemplo, una célula donante universal, sobreviva después del trasplante o injerto en un sujeto huésped con tasas de supervivencia superiores en relación con a una célula no modificada.

**Factor tolerogénico:** Como se usa en el presente documento, la expresión "factor tolerogénico" se refiere generalmente a una proteína (por ejemplo, expresada por un polinucleótido como se describe en el presente documento) que, cuando aumenta o disminuye en una célula, permite que la célula, por ejemplo, una célula donante universal, inhiba o evada el rechazo inmunitario después del trasplante o injerto en un sujeto huésped a tasas superiores en relación con una célula no modificada. En algunas realizaciones, un factor tolerogénico es un factor tolerogénico humano. En algunas realizaciones, la modificación genética de al menos un factor tolerogénico (por ejemplo, la inserción o delección de al menos un factor tolerogénico) permite a una célula, por ejemplo, una célula donante universal, inhibir o evadir el rechazo inmunitario con tasas al menos 1.05, al menos 1.1, al menos 1.25, al menos 1.5, al menos 2, al menos 3, al menos 4, al menos 5, al menos 10, al menos 20 o al menos al menos 50 veces superiores a las de una célula no modificada después del injerto. En algunas realizaciones, un factor tolerogénico es HLA-E (ID del gen del NCBI N.º: 3133), HLA-G (ID del gen del NCBI N.º: 3135), CTLA-4 (ID del gen del NCBI N.º: 1493), CD47 (ID del gen del NCBI N.º: 961) o PD-L1 (ID del gen del NCBI N.º: 29126). En algunas realizaciones, se inserta un factor tolerogénico en una célula, por ejemplo, una célula donante universal. En algunas realizaciones, se elimina un factor tolerogénico de una célula, por ejemplo, una célula donante universal. En algunas realizaciones, una inserción de un polinucleótido que codifica HLA-E, HLA-G, CTLA-4, CD47 y/o PD-L1 permite que una célula, por ejemplo, una célula donante universal, inhiba o evada el rechazo inmunitario después del trasplante o injerto en un sujeto huésped.

**Regulador de la transcripción de MHC-I o MHC-II:** Como se usa en el presente documento, la expresión "regulador de la transcripción de MHC-I o MHC-II" se refiere generalmente a una biomolécula que modula, por ejemplo, aumenta o disminuye, la expresión de un antígeno leucocitario humano MHC-I y/o MHC-II. En algunas realizaciones, una biomolécula es un polinucleótido, por ejemplo, un gen o una proteína. En algunas realizaciones, un regulador de la transcripción de MHC-I o MHC-II aumentará o disminuirá la expresión en la

superficie celular de al menos una proteína MHC-I o MHC-II. En algunas realizaciones, un regulador de la transcripción de MHC-I o MHC-II aumentará o disminuirá la expresión de al menos un gen MHC-I o MHC-II. En algunas realizaciones, el regulador de la transcripción es CIITA (ID del gen del NCBI N.º: 4261) o NLRC5 (ID del gen del NCBI N.º: 84166). En algunas realizaciones, una delección o reducción de la expresión de CIITA o NLRC5 disminuye la expresión de al menos un gen MHC-I o MHC-II.

**Célula donante universal:** Como se usa en el presente documento, la expresión "célula donante universal" se refiere generalmente a una célula modificada genéticamente que es menos susceptible al rechazo alógeno durante un trasplante celular y/o demuestra una mayor supervivencia después del trasplante, en relación con una célula no modificada. En algunas realizaciones, una célula modificada genéticamente como se describe en el presente documento es una célula donante universal. En algunas realizaciones, la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula no modificada. En algunas realizaciones, la célula donante universal tiene una mayor supervivencia celular en comparación con una célula no modificada. En algunas realizaciones, una célula donante universal puede ser una célula madre. En algunas realizaciones, una célula donante universal puede ser una célula madre embrionaria (ESC), una célula madre adulta (ASC), una célula madre pluripotente inducida (iPSC) o una célula madre o progenitora hematopoyética (HSPC) (también denominada una célula madre hematopoyética (HSC)). En algunas realizaciones, una célula donante universal puede ser una célula diferenciada. En algunas realizaciones, una célula donante universal puede ser una célula somática (por ejemplo, células del sistema inmunitario). En algunas realizaciones, se administra a un sujeto una célula donante universal. En algunas realizaciones, se administra una célula donante universal a un sujeto que tiene, se sospecha que tiene o está en riesgo de tener una enfermedad. En algunas realizaciones, la célula donante universal es capaz de diferenciarse en células progenitoras de linaje restringido o células somáticas completamente diferenciadas. En algunas realizaciones, las células progenitoras de linaje restringido son células progenitoras del endodermo pancreático, células progenitoras endocrinas pancreáticas, células progenitoras mesenquimatosas, células progenitoras musculares, blastocitos, células progenitoras hematopoyéticas o células progenitoras neuronales. En algunas realizaciones, las células somáticas completamente diferenciadas son células secretoras endocrinas tales como células beta pancreáticas, células epiteliales, células endodérmicas, macrófagos, hepatocitos, adipocitos, células renales, células sanguíneas o células del sistema inmunitario. En algunas realizaciones, las células somáticas completamente diferenciadas son cardiomiocitos.

**Célula no modificada:** Como se usa en el presente documento, la expresión "célula no modificada" se refiere a una célula que no ha sido sometida a una modificación

genética en la que esté implicado un polinucleótido o un gen que codifica un MHC-I, un MHC-II, un regulador de la transcripción de MHC-I o MHC-II, un factor de supervivencia y/o un factor tolerogénico. En algunas realizaciones, una célula no modificada puede ser una célula madre. En algunas realizaciones, una célula no modificada puede ser una célula madre embrionaria (ESC), una célula madre adulta (ASC), una célula madre pluripotente inducida (iPSC) o una célula madre o progenitora hematopoyética (HSPC) (también denominada una célula madre hematopoyética (HSC)). En algunas realizaciones, una célula no modificada puede ser una célula diferenciada. En algunas realizaciones, una célula no modificada puede seleccionarse de células somáticas (por ejemplo, células del sistema inmunitario, por ejemplo, un linfocito T, por ejemplo, un linfocito T CD8+). Si una célula donante universal se compara "en relación con una célula no modificada", la célula donante universal y la célula no modificada son del mismo tipo de célula o comparten una línea celular parental común, por ejemplo, una iPSC donante universal se compara con una iPSC no modificada.

***Dentro o cerca de un gen:*** Como se usa en el presente documento, la expresión "dentro o cerca de un gen" se refiere a un sitio o una región de ADN genómico que es un componente intrónico o exónico de dicho gen o que está ubicado próximo a dicho gen. En algunas realizaciones, un sitio de ADN genómico está dentro de un gen si comprende al menos una parte de un intrón o exón de dicho gen. En algunas realizaciones, un sitio de ADN genómico ubicado cerca de un gen puede estar en el extremo 5' o 3' de dicho gen (por ejemplo, el extremo 5' o 3' de la región codificante de dicho gen). En algunas realizaciones, un sitio de ADN genómico ubicado cerca de un gen puede ser una región promotora o una región represora que module la expresión de dicho gen. En algunas realizaciones, un sitio de ADN genómico ubicado cerca de un gen puede estar en el mismo cromosoma que dicho gen. En algunas realizaciones, un sitio o región de ADN genómico se encuentra cerca de un gen si se encuentra a menos de 50 Kb, 40 Kb, 30 Kb, 20 Kb, 10 Kb, 5 Kb, 1 Kb, o más cerca del extremo 5' o 3' de dicho gen (por ejemplo, el extremo 5' o 3' de la región codificante de dicho gen).

## **II. Métodos de edición genómica**

La edición genómica se refiere generalmente al proceso de modificar la secuencia de nucleótidos de un genoma, preferentemente de una manera precisa o predeterminada. En algunas realizaciones, los métodos de edición genómica como se describen en el presente documento, por ejemplo, el sistema de endonucleasa CRISPR, pueden usarse para modificar genéticamente una célula como se describe en el presente documento, por ejemplo, para crear una célula donante universal. En algunas realizaciones, los métodos de edición genómica como se describen en el presente documento, por ejemplo, el sistema de endonucleasa CRISPR, pueden usarse para modificar genéticamente una célula como se describe en el presente documento, por ejemplo, para introducir al menos una modificación genética dentro o cerca de

al menos un gen que disminuya la expresión de uno o más antígenos leucocitarios humanos MHC-I y/o MHC-II u otros componentes del complejo de MHC-I o MHC-II en relación con una célula no modificada; para introducir al menos una modificación genética que aumenta la expresión de al menos un polinucleótido que codifica un factor tolerogénico en relación con una célula no modificada; y/o para introducir al menos una modificación genética que aumenta o disminuye la expresión de al menos un gen que codifica un factor de supervivencia en relación con una célula no modificada.

Los ejemplos de métodos de edición genómica descritos en el presente documento incluyen métodos de uso de nucleasas dirigidas al sitio para cortar el ácido desoxirribonucleico (ADN) en ubicaciones diana precisas en el genoma, creando de este modo roturas de ADN monocatenarias o bicatenarias en ubicaciones particulares en el genoma. Dichas roturas pueden repararse y se reparan de manera regular mediante procesos celulares endógenos naturales, tales como la reparación dirigida por homología (HDR) y la unión de extremos no homólogos (NHEJ, por sus siglas en inglés), como se describe en Cox *et al.*, "Therapeutic genome editing: prospects and challenges", *Nature Medicine*, 2015, 21(2), 121-31. Estos dos principales procesos de reparación del ADN consisten en una familia de vías alternativas. La NHEJ (siglas del inglés, *non-homologous DNA end joining*, unión de extremos no homólogos) une directamente los extremos de ADN resultantes de una rotura bicatenaria, en ocasiones con la pérdida o adición de secuencia de nucleótidos, lo que puede interrumpir o potenciar la expresión génica. La HDR utiliza una secuencia homóloga o secuencia donante como molde para insertar una secuencia de ADN definida en el punto de rotura. La secuencia homóloga puede estar en el genoma endógeno, tal como una cromátida hermana. Como alternativa, la secuencia donante puede ser un polinucleótido exógeno, tal como un oligonucleótido monocatenario, un oligonucleótido bicatenario, un oligonucleótido dúplex o un virus, que tiene regiones (por ejemplo, brazo de homología derecho e izquierdo) de alta homología con el locus escindido por nucleasa, pero que también puede contener una secuencia adicional o cambios de secuencia, incluyendo deleciones que pueden incorporarse en el locus diana escindido. Un tercer mecanismo de reparación puede ser la unión de extremos mediada por microhomología (MMEJ, siglas del inglés, *microhomology-mediated end joining*), también denominada "NHEJ alternativa", en la que el resultado genético es similar al de la NHEJ ya que pueden producirse pequeñas deleciones e inserciones en el sitio de escisión. La MMEJ puede hacer uso de secuencias homólogas de unos pocos pares de bases que flanquean al sitio de rotura del ADN para impulsar un resultado de la reparación de unión de extremos de ADN más favorecido y estudios recientes han dilucidado adicionalmente el mecanismo molecular de este proceso; véase, por ejemplo, Cho y Greenberg, *Nature*, 2015, 518, 174-76; Kent *et al.*, *Nature Structural and Molecular Biology*, 2015, 22(3):230-7; Mateos-Gomez *et al.*, *Nature*, 2015, 518, 254-57;

Ceccaldi *et al.*, *Nature*, 2015, 528, 258-62. En algunos casos, puede ser posible predecir los resultados de reparación probables basándose en el análisis de las microhomologías potenciales en el sitio de rotura del ADN.

Cada uno de estos mecanismos de edición genómica se puede usar para crear las modificaciones genéticas deseadas. Una etapa en el proceso de edición genómica puede ser crear una o dos roturas de ADN, siendo las últimas roturas bicatenarias o dos roturas monocatenarias, en el locus diana lo más cerca posible del sitio previsto para la mutación. Esto se puede lograr mediante el uso de endonucleasas, como se describe e ilustra en el presente documento.

#### Sistema de endonucleasa CRISPR

El sistema de endonucleasa CRISPR es un mecanismo de defensa que se produce de manera natural en procariotas que se ha readaptado como una plataforma dirigida al ADN guiado por ARN usada para la edición génica. Los sistemas CRISPR incluyen sistemas de tipos I, II, III, IV, V y VI. En algunos aspectos, el sistema CRISPR es un sistema CRISPR/Cas9 de tipo II. En otros aspectos, el sistema CRISPR es un sistema CRISPR/Cpf1 de tipo V. Los sistemas CRISPR se basan en una endonucleasa de ADN, por ejemplo, Cas9, y dos ARN no codificantes, ARN de CRISPR (ARNcr) y ARN trans-activador (ARNtracr), para dirigirse a la escisión del ADN.

El ARNcr dirige el reconocimiento de la secuencia y la especificidad del complejo de CRISPR-endonucleasa a través de emparejamiento de bases de Watson-Crick normalmente con una secuencia de 20 nucleótidos (nt) en el ADN diana. El cambio de la secuencia de los 20 nt en 5' en el ARNcr permite dirigir el complejo de CRISPR-endonucleasa a locus específicos. El complejo de CRISPR-endonucleasa se une solo a secuencias de ADN que contienen un emparejamiento de secuencia con los primeros 20 nt del ARN simple guía (ARNsg), si la secuencia diana va seguida por un motivo de ADN corto específico (con la secuencia NGG) denominado motivo adyacente de protoespaciador (PAM, por sus siglas en inglés).

El ARNtracr se hibrida con el extremo 3' del ARNcr para formar una estructura de dúplex de ARN a la que se une la endonucleasa para formar el complejo de CRISPR-endonucleasa catalíticamente activo, que entonces puede escindir el ADN diana.

Una vez que el complejo de CRISPR-endonucleasa se une a ADN en un sitio diana, cada uno de dos dominios de nucleasa independientes dentro de la endonucleasa escinden una de las cadenas de ADN a tres bases cadena arriba del sitio de PAM, dejando una rotura bicatenaria (DSB) donde ambas cadenas del ADN terminan en un par de bases (un extremo romo).

En algunas realizaciones, la endonucleasa es una Cas9 (proteína 9 asociada a CRISPR). En algunas realizaciones, la endonucleasa Cas9 es de *Streptococcus pyogenes*,

aunque pueden usarse otros homólogos de Cas9, por ejemplo, Cas9 de *S. aureus*, Cas9 de *N. meningitidis*, CRISPR I Cas9 de *S. thermophilus*, CRISPR 3 Cas9 de *S. thermophilus* o Cas9 de *T. denticola*. En otros casos, la endonucleasa CRISPR es Cpf1, por ejemplo, Cpf1 de *L. bacterium* ND2006 o Cpf1 de *Acidaminococcus* sp. BV3L6. En algunas realizaciones, la endonucleasa es una endonucleasa Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocida como Csn1 y Csx12), Cas100, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 o Cpf1. En algunas realizaciones, se pueden usar variantes de tipo silvestre. En algunas realizaciones, se pueden usar versiones modificadas (por ejemplo, un homólogo de las mismas, una recombinación de la molécula de origen natural de las mismas, una versión con codones optimizados de las mismas o versiones modificadas de las mismas) de las endonucleasas anteriores.

La nucleasa CRISPR se puede unir a al menos una señal de localización nuclear (NLS). La al menos una NLS puede estar ubicada en o a menos de 50 aminoácidos del extremo amino de la nucleasa CRISPR y/o al menos una NLS puede ubicarse en o a menos de 50 aminoácidos del extremo carboxilo de la nucleasa CRISPR.

Los polipéptidos CRISPR/Cas de ejemplo incluyen los polipéptidos Cas9 según lo publicado en Fonfara *et al.*, "Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems, *Nucleic Acids Research*, 2014, 42: 2577-2590. El sistema de nomenclatura del gen CRISPR/Cas ha experimentado una extensa reescritura desde que se descubrieron los genes Cas. Fonfara *et al.* también proporcionan secuencias de PAM para los polipéptidos Cas9 de diversas especies.

#### *Nucleasas de dedos de cinc*

Las nucleasas de dedo de cinc (ZFN, por sus siglas en inglés) son proteínas modulares formadas por un dominio de unión a ADN de dedo de cinc manipulado unido al dominio catalítico de la endonucleasa FokI de tipo II. Debido a que FokI funciona únicamente como un dímero, ha de manipularse un par de ZFN para que se unan a secuencias de "semisitio" diana afines en cadenas opuestas de ADN y con un espaciado preciso entre ellas para permitir la formación del dímero de FokI catalíticamente activo. Tras la dimerización del dominio FokI, que por sí mismo no tiene especificidad de secuencia, se genera una rotura de ADN bicatenaria entre los semisitios de ZFN como etapa de inicio en la edición genómica.

El dominio de unión a ADN de cada ZFN está formado normalmente por 3-6 dedos de cinc con la arquitectura Cys2-His2 abundante, reconociendo cada dedo principalmente un triplete de nucleótidos en una cadena de la secuencia de ADN diana, aunque también puede ser importante la interacción de cadena cruzada con un cuarto nucleótido. La alteración de los

aminoácidos de un dedo en las posiciones que producen contactos clave con el ADN altera la especificidad de secuencia de un dedo específico. Por lo tanto, una proteína de dedos de cinco con cuatro dedos reconocerá de manera selectiva una secuencia diana de 12 pb, en donde la secuencia diana es una composición de las preferencias de triplete aportadas por cada dedo, aunque la preferencia de triplete puede estar influenciada en diversos grados por dedos vecinos. Un aspecto importante de las ZFN es que pueden redirigirse fácilmente a prácticamente cualquier dirección genómica, simplemente modificando dedos individuales. En la mayoría de aplicaciones de las ZFN, se usan proteínas de 4-6 dedos, que reconocen 12-18 pb, respectivamente. Por lo tanto, un par de ZFN reconocerán normalmente una secuencia diana combinada de 24-36 pb, que no incluye el espaciador de 5-7 pb típico entre los semisitios. Los sitios de unión pueden separarse adicionalmente con espaciadores más grandes, incluyendo de 15-17 pb. Es probable que una secuencia diana de esta longitud sea única en el genoma humano, asumiendo que se excluyan secuencias repetitivas u homólogos génicos durante el proceso de diseño. Sin embargo, las interacciones de proteína ZFN-ADN no son absolutas en su especificidad, por lo que no se producen eventos de unión y escisión inespecíficos, ya sea como un heterodímero entre las dos ZFN o como un homodímero de una o la otra de las ZFN. Esta última posibilidad se ha eliminado eficazmente manipulando la interfaz de dimerización del dominio FokI para crear variantes "más" y "menos", también conocidas como variantes de heterodímero obligado, que solo pueden dimerizarse entre sí y no por sí mismas. Forzar al heterodímero obligado impide la formación del homodímero. Esto ha mejorado en gran medida la especificidad de las ZFN, así como cualquier otra nucleasa que adopte estas variantes de FokI.

Se ha descrito en la técnica una variedad de sistemas a base de ZFN, se publican regularmente modificaciones de los mismos y numerosas referencias describen reglas y parámetros que se usan para guiar el diseño de las ZFN; véase, por ejemplo, Segal *et al.*, *Proc Natl Acad Sci*, 1999 96(6):2758-63; Dreier B *et al.*, *J Mol Biol.*, 2000, 303(4):489-502; Liu Q *et al.*, *J Biol Chem.*, 2002, 277(6):3850-6; Dreier *et al.*, *J Biol Chem.*, 2005, 280(42):35588-97; y Dreier *et al.*, *J Biol Chem.* 2001, 276(31):29466-78.

#### *Nucleasas efectoras de tipo activador de la transcripción (TALEN)*

Las TALEN representan otro formato de nucleasas modulares mediante las cuales, como con las ZFN, se enlaza un dominio de unión a ADN manipulado al dominio de nucleasa FokI y un par de TALEN funcionan en tándem para lograr la escisión del ADN diana. La principal diferencia respecto de las ZFN es la naturaleza del dominio de unión de ADN y las propiedades de reconocimiento de la secuencia de ADN diana. El dominio de unión a ADN de TALEN procede de las proteínas TALE, que se describieron originariamente en el patógeno bacteriano de plantas *Xanthomonas sp.* Las proteínas TALE están formadas por series en tándem de 33-

35 repeticiones de aminoácidos, reconociendo cada repetición un solo par de bases en la secuencia de ADN diana que normalmente tiene hasta 20 pb de longitud, proporcionando una longitud de secuencia diana de hasta 40 pb. La especificidad de nucleótidos de cada repetición se determina por el dirresiduo variable de repetición (RVD, por sus siglas en inglés), que incluye solo dos aminoácidos en las posiciones 12 y 13. Las bases guanina, adenina, citosina y timina se reconocen principalmente por los cuatro RVD: Asn-Asn, Asn-Ile, His-Asp y Asn-Gly, respectivamente. Esto constituye un código de reconocimiento mucho más simple que para los dedos de cinc y por lo tanto, representa una ventaja frente a estos últimos para el diseño de nucleasas. Sin embargo, como con las ZFN, las interacciones entre proteína-ADN de las TALEN no son absolutas en su especificidad y las TALEN también se han beneficiado del uso de variantes de heterodímero obligado del dominio FokI para reducir la actividad inespecífica.

Se han creado variantes adicionales del dominio FokI que tienen su función catalítica desactivada. En caso de que una mitad de un par de TALEN o ZFN contenga un dominio de FokI inactivo, solo se producirá una escisión de ADN monocatenario (mellado) en el sitio diana, en lugar de una DSB. El resultado es comparable al uso de las mutantes de "nickasa" de CRISPR/Cas9 o CRISPR/Cpf1 en las que se ha desactivado uno de los dominios de escisión de Cas9. Pueden usarse mellas de ADN para impulsar la edición genómica mediante HDR, pero con una eficacia menor que con una DSB. El principal beneficio es que las mellas inespecíficas se reparan con rapidez y exactitud, a diferencia de la DSB, que es propensa a la reparación errónea mediada por NHEJ.

Se ha descrito en la técnica una variedad de sistemas basados en TALEN y regularmente se publican modificaciones de los mismos; véanse, por ejemplo, Boch, *Science*, 2009 326(5959):1509-12; Mak *et al.*, *Science*, 2012, 335(6069):716-9; y Moscou *et al.*, *Science*, 2009, 326(5959):1501. El uso de las TALEN basadas en la plataforma "Golden Gate", o esquema de clonación, ha sido descrito por múltiples grupos; véanse, por ejemplo, Cermak *et al.*, *Nucleic Acids Res.*, 2011, 39(12):e82; Li *et al.*, *Nucleic Acids Res.*, 2011, 39(14):6315-25; Weber *et al.*, *PLoS One.*, 2011, 6(2):e16765; Wang *et al.*, *J Genet Genomics*, 2014, 41(6):339-47.; y Cermak T *et al.*, *Methods Mol Biol.*, 2015 1239:133-59.

#### *Endonucleasas de asentamiento*

Las endonucleasas de asentamiento (HE) son endonucleasas específicas de secuencia que tienen largas secuencias de reconocimiento (14-44 pares de bases) y escinden el ADN con alta especificidad – normalmente en sitios únicos en el genoma. Hay al menos seis familias conocidas de HE clasificadas por su estructura, que incluyen GIY-YIG, caja His-Cis, HNH, PD-(D/E) xK y de tipo Vsr que derivan de una amplia gama de huéspedes, incluyendo eucariotas, protistas, bacterias, arqueas, cianobacterias y fagos. Al igual que con las ZFN y las TALEN, las HE pueden usarse para crear una DSB en un locus diana como etapa inicial en la edición

genómica. Además, algunas HE naturales y manipuladas cortan únicamente una sola cadena de ADN, actuando así como nickasas específicas del sitio. La secuencia diana grande de las HE y la especificidad que ofrecen han hecho que sean candidatos atractivos para crear DSB específicas del sitio.

5 Se ha descrito en la técnica una variedad de sistemas basados en HE y se comunican regularmente modificaciones de los mismos; véase, por ejemplo, la revisión de Steentoft *et al.*, *Glycobiology*, 2014, 24(8):663-80; Belfort y Bonocora, *Methods Mol Biol.*, 2014, 1123:1-26; y Hafez y Hausner, *Genome*, 2012, 55(8):553-69.

#### *MegaTAL/Tev-mTALEN/MegaTev*

10 Como ejemplos adicionales de nucleasas híbridas, la plataforma MegaTAL y la plataforma Tev-mTALEN usan una fusión de dominios de unión a ADN de TALE y HE catalíticamente activas, aprovechando la unión a ADN ajustable y la especificidad de la TALE, así como la especificidad de la secuencia de escisión de la HE: véase, por ejemplo, Boissel *et al.*, *Nucleic Acids Res.*, 2014, 42: 2591-2601; Kleinstiver *et al.*, *G3*, 2014, 4:1155-65; y Boissel y Scharenberg, *Methods Mol. Biol.*, 2015, 1239: 171-96.

15 En una variación adicional, la arquitectura de MegaTev es la fusión de una meganucleasa (Mega) con el dominio de nucleasa derivado de la endonucleasa de asentamiento GIY-YIG I-TevI (Tev). Los dos sitios activos están ubicados a ~30 pb de distancia en un sustrato de ADN y generan dos DSB con extremos cohesivos no compatibles; véase, por ejemplo, Wolfs *et al.*, *Nucleic Acids Res.*, 2014, 42, 8816-29. Se prevé que evolucionarán otras combinaciones de estrategias basadas en nucleasas existentes y serán útiles para lograr las modificaciones genómicas dirigidas descritas en el presente documento.

#### *dCas9-FokI o dCpf1-FokI y otras nucleasas*

25 La combinación de las propiedades estructurales y funcionales de las plataformas de nucleasa descritas anteriormente ofrece una estrategia adicional para la edición genómica que puede superar potencialmente algunas de las deficiencias inherentes. A modo de ejemplo, el sistema de edición genómica de CRISPR usa normalmente una sola endonucleasa Cas9 para crear una DSB. La especificidad del direccionamiento está impulsada por una secuencia de 20 o 24 nucleótidos en el ARN guía que experimenta emparejamiento de bases de Watson-Crick con el ADN diana (más 2 bases adicionales en la secuencia de PAM de NAG o NGG en el caso de Cas9 de *S. pyogenes*). Dicha secuencia tiene una longitud suficiente para ser única en el genoma humano, aunque la especificidad de la interacción de ARN/ADN no es absoluta, tolerándose en ocasiones una promiscuidad significativa, particularmente en la mitad 5' de la secuencia diana, reduciendo eficazmente el número de bases que impulsan la especificidad.

35 Una solución para esto ha sido desactivar completamente la función catalítica de Cas9 o Cpf1 – conservando únicamente la función de unión a ADN guiada por ARN – y, en su lugar,

fusionando un dominio FokI a la Cas9 desactivada; véanse, por ejemplo, Tsai *et al.*, *Nature Biotech.*, 2014, 32: 569-76; y Guilinger *et al.*, *Nature Biotech.*, 2014, 32: 577-82. Debido a que FokI ha de dimerizarse para volverse catalíticamente activa, se necesitan dos ARN guía para anclar las fusiones de FokI muy próximas para formar el dímero y escindir el ADN. Esto esencialmente duplica el número de bases en los sitios diana combinados, aumentando de este modo la rigurosidad del direccionamiento por los sistemas basados en CRISPR.

Como ejemplo adicional, la fusión del dominio de unión a ADN de TALE a una HE catalíticamente activa, tal como I-TevI, aprovecha tanto la unión a ADN ajustable y la especificidad de la TALE, así como la especificidad de secuencia de escisión de I-TevI, esperándose que pueda reducirse adicionalmente la escisión inespecífica.

#### *Endonucleasas guiadas por ARN*

Los sistemas de endonucleasa guiada por ARN como se usan en el presente documento pueden comprender una secuencia de aminoácidos que tiene al menos un 10 %, al menos un 15 %, al menos un 20 %, al menos un 30 %, al menos un 40 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 %, al menos un 99 % o un 100 % de identidad de secuencia de aminoácidos con respecto a una endonucleasa de tipo silvestre de ejemplo, por ejemplo, Cas9 de *S. pyogenes*, ID de Secuencia N.º 8 del documento US2014/0068797 o Sapranasuskas *et al.*, *Nucleic Acids Res.*, 39(21): 9275-9282 (2011). La endonucleasa puede comprender al menos un 70, 75, 80, 85, 90, 95, 97, 99 o 100 % de identidad con una endonucleasa de tipo silvestre (por ejemplo, Cas9 de *S. pyogenes*, anteriormente citado) sobre 10 aminoácidos contiguos. La endonucleasa puede comprender como máximo: un 70, 75, 80, 85, 90, 95, 97, 99 o 100 % de identidad con una endonucleasa de tipo silvestre (por ejemplo, Cas9 de *S. pyogenes*, anteriormente citado) sobre 10 aminoácidos contiguos. La endonucleasa puede comprender al menos: un 70, 75, 80, 85, 90, 95, 97, 99 o 100 % de identidad con una endonucleasa de tipo silvestre (por ejemplo, Cas9 de *S. pyogenes*, anteriormente citado) sobre 10 aminoácidos contiguos en un dominio nucleasa HNH de la endonucleasa. La endonucleasa puede comprender como máximo: un 70, 75, 80, 85, 90, 95, 97, 99 o 100 % de identidad con una endonucleasa de tipo silvestre (por ejemplo, Cas9 de *S. pyogenes*, anteriormente citado) sobre 10 aminoácidos contiguos en un dominio nucleasa HNH de la endonucleasa. La endonucleasa puede comprender al menos: un 70, 75, 80, 85, 90, 95, 97, 99 o 100 % de identidad con una endonucleasa de tipo silvestre (por ejemplo, Cas9 de *S. pyogenes*, anteriormente citado) sobre 10 aminoácidos contiguos en un dominio nucleasa RuvC de la endonucleasa. La endonucleasa puede comprender como máximo: un 70, 75, 80, 85, 90, 95, 97, 99 o 100 % de identidad con una endonucleasa de tipo silvestre (por ejemplo, Cas9 de *S. pyogenes*, anteriormente citado) sobre 10 aminoácidos contiguos en un dominio nucleasa RuvC de la endonucleasa.

La endonucleasa puede comprender una forma modificada de una endonucleasa de ejemplo de tipo silvestre. La forma modificada de la endonucleasa de ejemplo de tipo silvestre puede comprender una mutación que reduzca la actividad de escisión de ácidos nucleicos de la endonucleasa. La forma modificada de la endonucleasa de ejemplo de tipo silvestre puede tener menos del 90 %, menos del 80 %, menos del 70 %, menos del 60 %, menos del 50 %, menos del 40 %, menos del 30 %, menos del 20 %, menos del 10 %, menos del 5 % o menos del 1 % de la actividad de escisión de ácidos nucleicos de la endonucleasa de ejemplo de tipo silvestre (por ejemplo, Cas9 de *S. pyogenes*, anteriormente citado). La forma modificada de la endonucleasa no puede tener una actividad sustancial de escisión de ácidos nucleicos. Cuando una endonucleasa es una forma modificada que no tiene una actividad de escisión de ácidos nucleicos sustancial, se denomina en el presente documento "enzimáticamente inactiva".

Las mutaciones contempladas pueden incluir sustituciones, adiciones y eliminaciones o cualquier combinación de las mismas. La mutación convierte el aminoácido mutado en alanina. La mutación convierte el aminoácido mutado en otro aminoácido (por ejemplo, glicina, serina, treonina, cisteína, valina, leucina, isoleucina, metionina, prolina, fenilalanina, tirosina, triptófano, ácido aspártico, ácido glutámico, asparaginas, glutamina, histidina, lisina o arginina). La mutación convierte el aminoácido mutado en un aminoácido no natural (por ejemplo, selenometionina). La mutación convierte el aminoácido mutado en miméticos de aminoácidos (por ejemplo, fosfomiméticos). La mutación puede ser una mutación conservadora. Por ejemplo, la mutación convierte el aminoácido mutado en aminoácidos que se asemejan en tamaño, forma, carga, polaridad, conformación y/o rotámeros a los aminoácidos mutados (por ejemplo, mutación de cisteína/serina, mutación de lisina/asparagina, mutación de histidina/fenilalanina). La mutación puede provocar un desplazamiento en el marco de lectura y/o la creación de un codón de terminación prematuro. Las mutaciones pueden provocar cambios en regiones reguladoras de genes o locus que afectan a la expresión de uno o más genes.

#### ARN guía

La presente divulgación proporciona un ARN guía (ARNg) que puede dirigir las actividades de una endonucleasa asociada a un sitio diana específico dentro de un polinucleótido. Un ARN guía puede comprender al menos una secuencia espaciadora que se hibrida con una secuencia de ácido nucleico diana de interés y una secuencia de repetición CRISPR. En los sistemas CRISPR de tipo II, el ARNg también comprende un segundo ARN que se denomina la secuencia de ARNtracr. En el ARN guía de CRISPR de tipo II (ARNg), la secuencia de repetición CRISPR y la secuencia de ARNtracr se hibridan entre sí para formar un dúplex. En los sistemas CRISPR de tipo V, el ARNg comprende un ARNcr que forma un dúplex. En algunas realizaciones, un ARNg puede unirse a una endonucleasa, de modo que el ARNg y la endonucleasa formen un complejo. El ARNg puede proporcionar especificidad diana

al complejo en virtud de su asociación con la endonucleasa. Por lo tanto, el ácido nucleico dirigido al genoma puede dirigir la actividad de la endonucleasa.

Los ARN guía de ejemplo incluyen secuencias espaciadoras que comprenden 15-200 nucleótidos, en donde el ARNg se dirige a una ubicación del genoma basándose en el ensamblaje del genoma humano GRCh38. Como entenderá el experto en la técnica, cada ARNg se puede diseñar para incluir una secuencia espaciadora complementaria a su región o sitio genómico diana. Véanse Jinek *et al.*, *Science*, 2012, 337, 816-821 y Deltcheva *et al.*, *Nature*, 2011, 471, 602-607.

El ARNg puede ser un ARN guía de doble molécula. El ARNg puede ser un ARN guía de una sola molécula.

Un ARN guía de molécula doble puede comprender dos cadenas de ARN. La primera cadena comprende, en la dirección de 5' a 3', una secuencia de extensión de espaciador opcional, una secuencia espaciadora y una secuencia de repetición CRISPR mínima. La segunda cadena puede comprender una secuencia de ARNtracr mínima (complementaria a la secuencia de repetición CRISPR mínima), una secuencia de ARNtracr 3' y una secuencia de extensión de ARNtracr opcional.

Un ARN guía de una sola molécula (ARNsg) puede comprender, en la dirección de 5' a 3', una secuencia de extensión de espaciador opcional, una secuencia espaciadora, una secuencia de repetición de CRISPR mínima, un enlazador de guía de una sola molécula, una secuencia de ARNtracr mínima, una secuencia de ARNtracr 3' y una secuencia de extensión de ARNtracr opcional. La extensión de ARNtracr opcional puede comprender elementos que aportan una funcionalidad adicional (por ejemplo, estabilidad) al ARN guía. El enlazador guía de una sola molécula puede enlazar la repetición CRISPR mínima y la secuencia de ARNtracr mínima para formar una estructura en horquilla. La extensión de ARNtracr opcional puede comprender una o más horquillas.

En algunas realizaciones, un ARNsg comprende un espaciador de 20 nucleótidos en el extremo 5' de la secuencia de ARNsg. En algunas realizaciones, un ARNsg comprende una secuencia espaciadora de menos de 20 nucleótidos en el extremo 5' de la secuencia de ARNsg. En algunas realizaciones, un ARNsg comprende una secuencia espaciadora de más de 20 nucleótidos en el extremo 5' de la secuencia de ARNsg. En algunas realizaciones, un ARNsg comprende una secuencia espaciadora de longitud variable con 17-30 nucleótidos en el extremo 5' de la secuencia de ARNsg. En algunas realizaciones, un ARNsg comprende una secuencia de extensión de espaciador con una longitud de más de 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180 o 200 nucleótidos. En algunas realizaciones, un ARNsg comprende una secuencia de extensión de espaciador con una longitud de menos de 3, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90 o 100 nucleótidos.

En algunas realizaciones, un ARNsg comprende una secuencia de extensión espaciadora que comprende otro resto (por ejemplo, una secuencia de control de la estabilidad, una secuencia de unión a endorribonucleasa, una ribozima). El resto puede disminuir o aumentar la estabilidad de un ácido nucleico dirigido a un ácido nucleico. El resto puede ser un segmento terminador de la transcripción (es decir, una secuencia de terminación de la transcripción). El resto puede actuar en una célula eucariota. El resto puede actuar en una célula procariota. El resto puede actuar tanto en células eucariotas como procariotas. Los ejemplos no limitantes de restos adecuados incluyen: una caperuza 5' (por ejemplo, una caperuza de 7-metilguanilato (m7 G)), una secuencia de ribointerruptor (por ejemplo, para permitir una estabilidad regulada y/o una accesibilidad regulada por proteínas y complejos de proteínas), una secuencia que forma un dúplex de ARNbc (es decir, una horquilla), una secuencia que dirige el ARN a una ubicación subcelular (por ejemplo, núcleo, mitocondria, cloroplastos y similares), una modificación o secuencia que posibilita el seguimiento (por ejemplo, la conjugación directa a una molécula fluorescente, conjugación a un resto que facilita la detección fluorescente, una secuencia que posibilita la detección fluorescente, etc.) y/o una modificación o secuencia que proporciona un sitio de unión para proteínas (por ejemplo, proteínas que actúan sobre el ADN, incluyendo activadores de la transcripción, represores de la transcripción, ADN metiltransferasas, ADN desmetilasas, histona acetiltransferasas, histona desacetilasas y similares).

En algunas realizaciones, un ARNsg comprende una secuencia espaciadora que se hibrida con una secuencia en un polinucleótido diana. El espaciador de un ARNg puede interactuar con un polinucleótido diana de una manera específica de secuencia mediante hibridación (es decir, emparejamiento de bases). La secuencia de nucleótidos del espaciador puede variar dependiendo de la secuencia del ácido nucleico diana de interés.

En un sistema de endonucleasa CRISPR, se puede diseñar una secuencia espaciadora para hibridarse con un polinucleótido diana que se ubica en 5' de un PAM de la endonucleasa utilizada en el sistema. El espaciador puede emparejarse perfectamente con la secuencia diana o puede tener emparejamientos erróneos. Cada endonucleasa, por ejemplo, la nucleasa Cas9, tiene una secuencia PAM particular que reconoce en un ADN diana. Por ejemplo, Cas9 de *S. pyogenes* reconoce un PAM que comprende la secuencia 5'-NRG-3', donde R comprende bien A o G, en donde N es cualquier nucleótido y N se encuentra inmediatamente en 3' de la secuencia de ácido nucleico diana utilizada como diana por la secuencia espaciadora.

Una secuencia polinucleotídica diana puede comprender 20 nucleótidos. El polinucleótido diana puede comprender menos de 20 nucleótidos. El polinucleótido diana puede comprender más de 20 nucleótidos. El polinucleótido diana puede comprender al menos: 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 o más nucleótidos. El polinucleótido diana puede

comprender como máximo: 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 o más nucleótidos. La secuencia polinucleotídica diana puede comprender 20 bases inmediatamente en 5' del primer nucleótido del PAM.

Una secuencia espaciadora que se hibrida con un polinucleótido diana puede tener una longitud de al menos aproximadamente 6 nucleótidos (nt). La secuencia espaciadora puede tener al menos aproximadamente 6 nt, al menos aproximadamente 10 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 aproximadamente 6 nt a aproximadamente 80 nt, de aproximadamente 6 nt a aproximadamente 50 nt, de aproximadamente 6 nt a aproximadamente 45 nt, de aproximadamente 6 nt a aproximadamente 40 nt, de aproximadamente 6 nt a aproximadamente 35 nt, de aproximadamente 6 nt a aproximadamente 30 nt, de aproximadamente 6 nt a aproximadamente 25 nt, de aproximadamente 6 nt a aproximadamente 20 nt, de aproximadamente 6 nt a aproximadamente 19 nt, de aproximadamente 10 nt a aproximadamente 50 nt, de aproximadamente 10 nt a aproximadamente 45 nt, de aproximadamente 10 nt a aproximadamente 40 nt, de aproximadamente 10 nt a aproximadamente 35 nt, de aproximadamente 10 nt a aproximadamente 30 nt, de aproximadamente 10 nt a aproximadamente 25 nt, de aproximadamente 10 nt a aproximadamente 20 nt, de aproximadamente 10 nt a aproximadamente 19 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 ejemplos, la secuencia espaciadora puede comprender 20 nucleótidos. En algunos ejemplos, el espaciador puede comprender 19 nucleótidos. En algunos ejemplos, el espaciador puede comprender 18 nucleótidos. En algunos ejemplos, el espaciador puede comprender 22 nucleótidos.

En algunos ejemplos, el porcentaje de complementariedad entre la secuencia espaciadora y el ácido nucleico diana es de al menos aproximadamente el 30 %, al menos aproximadamente el 40 %, al menos aproximadamente el 50 %, al menos aproximadamente el

60 %, al menos aproximadamente el 65 %, al menos aproximadamente el 70 %, al menos aproximadamente el 75 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 %, al menos aproximadamente el 97 %, al menos aproximadamente el 98 %, al menos aproximadamente el 99 % o del 100 %. En algunos ejemplos, el porcentaje de complementariedad entre la secuencia espaciadora y el ácido nucleico diana es como máximo de aproximadamente el 30 %, como máximo de aproximadamente el 40 %, como máximo de aproximadamente el 50 %, como máximo de aproximadamente el 60 %, como máximo de aproximadamente el 65 %, como máximo de aproximadamente el 70 %, como máximo de aproximadamente el 75 %, como máximo de aproximadamente el 80 %, como máximo de aproximadamente el 85 %, como máximo de aproximadamente el 90 %, como máximo de aproximadamente el 95 %, como máximo de aproximadamente el 97 %, como máximo de aproximadamente el 98 %, como máximo de aproximadamente el 99 % o del 100 %. En algunos ejemplos, el porcentaje de complementariedad entre la secuencia espaciadora y el ácido nucleico diana es del 100 % a lo largo de los seis nucleótidos contiguos más próximos al extremo 5' de la secuencia diana de la cadena complementaria del ácido nucleico diana. El porcentaje de complementariedad entre la secuencia espaciadora y el ácido nucleico diana puede ser de al menos el 60 % a lo largo de aproximadamente 20 nucleótidos contiguos. La longitud de la secuencia espaciadora y el ácido nucleico diana puede diferir en de 1 a 6 nucleótidos, lo que puede considerarse como una protuberancia o protuberancias.

Una secuencia de ARNtracr puede comprender nucleótidos que se hibridan con una secuencia de repetición CRISPR mínima en una célula. Una secuencia de ARNtracr mínima y una secuencia de repetición CRISPR mínima pueden formar un dúplex, es decir, una estructura bicatenaria de bases emparejadas. Juntas, la secuencia de ARNtracr mínima y la repetición CRISPR mínima pueden unirse a una endonucleasa guiada por ARN. Al menos una parte de la secuencia de ARNtracr mínima puede hibridarse con la secuencia de repetición CRISPR mínima. La secuencia de ARNtracr mínima puede ser al menos aproximadamente el 30 %, aproximadamente el 40 %, aproximadamente el 50 %, aproximadamente el 60 %, aproximadamente el 65 %, aproximadamente el 70 %, aproximadamente el 75 %, aproximadamente el 80 %, aproximadamente el 85 %, aproximadamente el 90 %, aproximadamente el 95 % o el 100 % complementaria a la secuencia de repetición CRISPR mínima.

La secuencia de ARNtracr mínima puede tener una longitud de aproximadamente 7 nucleótidos a aproximadamente 100 nucleótidos. Por ejemplo, la secuencia de ARNtracr mínima puede tener de aproximadamente 7 nucleótidos (nt) a aproximadamente 50 nt, de aproximadamente 7 nt a aproximadamente 40 nt, de aproximadamente 7 nt a aproximadamente

30 nt, de aproximadamente 7 nt a aproximadamente 25 nt, de aproximadamente 7 nt a aproximadamente 20 nt, de aproximadamente 7 nt a aproximadamente 15 nt, de aproximadamente 8 nt a aproximadamente 40 nt, de aproximadamente 8 nt a aproximadamente 30 nt, de aproximadamente 8 nt a aproximadamente 25 nt, de aproximadamente 8 nt a aproximadamente 20 nt, de aproximadamente 8 nt a aproximadamente 15 nt, de aproximadamente 15 nt a aproximadamente 100 nt, de aproximadamente 15 nt a aproximadamente 80 nt, de aproximadamente 15 nt a aproximadamente 50 nt, de aproximadamente 15 nt a aproximadamente 40 nt, de aproximadamente 15 nt a aproximadamente 30 nt o de aproximadamente 15 nt a aproximadamente 25 nt de longitud. La secuencia de ARNtracr mínima puede tener aproximadamente 9 nucleótidos de longitud. La secuencia de ARNtracr mínima puede tener aproximadamente 12 nucleótidos. El ARNtracr mínimo puede consistir en los nt 23-48 de ARNtracr descrito en Jinek *et al.*, anteriormente citado.

La secuencia de ARNtracr mínima puede ser al menos aproximadamente un 60 % idéntica a una secuencia de ARNtracr mínima de referencia (por ejemplo, ARNtracr de tipo silvestre de *S. pyogenes*) a lo largo de un tramo de al menos 6, 7 u 8 nucleótidos contiguos. Por ejemplo, la secuencia de ARNtracr mínima puede ser al menos aproximadamente un 65 % idéntica, aproximadamente un 70 % idéntica, aproximadamente un 75 % idéntica, aproximadamente un 80 % idéntica, aproximadamente un 85 % idéntica, aproximadamente un 90 % idéntica, aproximadamente un 95 % idéntica, aproximadamente un 98 % idéntica, aproximadamente un 99 % idéntica o el 100 % idéntica a una secuencia de ARNtracr mínima de referencia a lo largo de un tramo de al menos 6, 7 u 8 nucleótidos contiguos.

El dúplex entre el ARN de CRISPR mínimo y el ARNtracr mínimo puede comprender una doble hélice. El dúplex entre el ARN de CRISPR mínimo y el ARNtracr mínimo puede comprender al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10, o más nucleótidos. El dúplex entre el ARN de CRISPR mínimo y el ARNtracr mínimo puede comprender como máximo aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10, o más nucleótidos.

El dúplex puede comprender un emparejamiento erróneo (es decir, las dos cadenas del dúplex no son 100 % complementarias). El dúplex puede comprender al menos aproximadamente 1, 2, 3, 4 o 5 emparejamientos erróneos. El dúplex puede comprender como máximo aproximadamente 1, 2, 3, 4 o 5 emparejamientos erróneos. El dúplex no puede contener más de 2 emparejamientos erróneos.

En algunas realizaciones, un ARNtracr puede ser un ARNtracr 3'. En algunas realizaciones, una secuencia de ARNtracr 3' puede comprender una secuencia con al menos aproximadamente un 30 %, aproximadamente un 40 %, aproximadamente un 50 %, aproximadamente un 60 %, aproximadamente un 65 %, aproximadamente un 70 %, aproximadamente un 75 %, aproximadamente un 80 %, aproximadamente un 85 %, aproximadamente un 90 %, aproximadamente un 95 %, aproximadamente un 98 %, aproximadamente un 99 % o el 100 % idéntica a una secuencia de ARNtracr 3' de referencia.

aproximadamente un 75 %, aproximadamente un 80 %, aproximadamente un 85 %, aproximadamente un 90 %, aproximadamente un 95 % o un 100 % de identidad de secuencia con una secuencia de ARNtracr de referencia (por ejemplo, un ARNtracr de *S. pyogenes*).

En algunas realizaciones, un ARNg puede comprender una secuencia de extensión de ARNtracr. Una secuencia de extensión de ARNtracr puede tener una longitud de aproximadamente 1 nucleótido a aproximadamente 400 nucleótidos. La secuencia de extensión de ARNtracr puede tener una longitud de más de 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180 o 200 nucleótidos. La secuencia de extensión del ARNtracr puede tener una longitud de aproximadamente 20 a aproximadamente 5000 o más nucleótidos. La secuencia de extensión de ARNtracr puede tener una longitud inferior a 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90 o 100 nucleótidos. La secuencia de extensión del ARNtracr puede comprender menos de 10 nucleótidos de longitud. La secuencia de extensión del ARNtracr puede tener una longitud de 10-30 nucleótidos. La secuencia de extensión del ARNtracr puede tener una longitud de 30-70 nucleótidos.

La secuencia de extensión del ARNtracr puede comprender un resto funcional (por ejemplo, una secuencia de control de la estabilidad, ribozima, secuencia de unión a endorribonucleasa). El resto funcional puede comprender un segmento terminador de la transcripción (es decir, una secuencia de terminación de la transcripción). El resto funcional puede tener una longitud total de aproximadamente 10 nucleótidos (nt) a aproximadamente 100 nucleótidos, de aproximadamente 10 nt a aproximadamente 20 nt, de aproximadamente 20 nt a aproximadamente 30 nt, de aproximadamente 30 nt a aproximadamente 40 nt, de aproximadamente 40 nt a aproximadamente 50 nt, de aproximadamente 50 nt a aproximadamente 60 nt, de aproximadamente 60 nt a aproximadamente 70 nt, de aproximadamente 70 nt a aproximadamente 80 nt, de aproximadamente 80 nt a aproximadamente 90 nt o de aproximadamente 90 nt a aproximadamente 100 nt, de aproximadamente 15 nt a aproximadamente 80 nt, de aproximadamente 15 nt a aproximadamente 50 nt, de aproximadamente 15 nt a aproximadamente 40 nt, de aproximadamente 15 nt a aproximadamente 30 nt o de aproximadamente 15 nt a aproximadamente 25 nt.

En algunas realizaciones, un ARNsg puede comprender una secuencia enlazadora con una longitud de aproximadamente 3 nucleótidos a aproximadamente 100 nucleótidos. En Jinek *et al.*, anteriormente citado, por ejemplo, se usó un "tetrabucle" sencillo de 4 nucleótidos (-GAAA-), (Jinek *et al.*, *Science*, 2012, 337(6096):816-821). Un enlazador ilustrativo tiene una longitud de aproximadamente 3 nucleótidos (nt) a aproximadamente 90 nt, de aproximadamente 3 nt a aproximadamente 80 nt, de aproximadamente 3 nt a aproximadamente 70 nt, de aproximadamente 3 nt a aproximadamente 60 nt, de aproximadamente 3 nt a aproximadamente

50 nt, de aproximadamente 3 nt a aproximadamente 40 nt, de aproximadamente 3 nt a aproximadamente 30 nt, de aproximadamente 3 nt a aproximadamente 20 nt, de aproximadamente 3 nt a aproximadamente 10 nt. Por ejemplo, el enlazador puede tener una longitud de aproximadamente 3 nt a aproximadamente 5 nt, de aproximadamente 5 nt a aproximadamente 10 nt, de aproximadamente 10 nt a aproximadamente 15 nt, de aproximadamente 15 nt a aproximadamente 20 nt, de aproximadamente 20 nt a aproximadamente 25 nt, de aproximadamente 25 nt a aproximadamente 30 nt, de aproximadamente 30 nt a aproximadamente 35 nt, de aproximadamente 35 nt a aproximadamente 40 nt, de aproximadamente 40 nt a aproximadamente 50 nt, de aproximadamente 50 nt a aproximadamente 60 nt, de aproximadamente 60 nt a aproximadamente 70 nt, de aproximadamente 70 nt a aproximadamente 80 nt, de aproximadamente 80 nt a aproximadamente 90 nt o de aproximadamente 90 nt a aproximadamente 100 nt. El enlazador de un ácido nucleico guía de una sola molécula puede tener entre 4 y 40 nucleótidos. El enlazador puede tener al menos aproximadamente 100, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500 o 7000 o más nucleótidos. El enlazador puede tener como máximo 100, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, o 7000 o más nucleótidos.

Los enlazadores pueden comprender cualquiera de una variedad de secuencias, aunque en algunos ejemplos, el enlazador no comprenderá secuencias que tengan regiones de homología extensas con otras partes del ARN guía, que podrían provocar una unión intramolecular que podría interferir con otras regiones funcionales de la guía. En Jinek *et al.*, anteriormente citado, se usó una secuencia sencilla de 4 nucleótidos -GAAA- (Jinek *et al.*, *Science*, 2012, 337(6096):816-821), pero pueden usarse igualmente numerosas secuencias distintas, incluyendo secuencias más largas.

La secuencia enlazadora puede comprender un resto funcional. Por ejemplo, la secuencia enlazadora puede comprender una o más características, incluyendo un aptámero, una ribozima, una horquilla que interactúa con proteínas, un sitio de unión a proteínas, una serie de CRISPR, un intrón o un exón. La secuencia enlazadora puede comprender al menos aproximadamente 1, 2, 3, 4 o 5, o más restos funcionales. En algunos ejemplos, la secuencia enlazadora puede comprender como máximo aproximadamente 1, 2, 3, 4 o 5, o más restos funcionales.

En algunas realizaciones, un ARNsg no comprende un uracilo, por ejemplo, en el extremo 3' de la secuencia de ARNsg. En algunas realizaciones, un ARNsg comprende uno o más uracilos, por ejemplo, en el extremo 3' de la secuencia de ARNsg. En algunas realizaciones, un ARNsg comprende 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 uracilos (U) en el extremo 3' de la secuencia de ARNsg.

Un ARNsg puede modificarse químicamente. En algunas realizaciones, un ARNg modificado químicamente es un ARNg que comprende al menos un nucleótido con una modificación química, por ejemplo, una modificación de 2'-O-metil-azúcar. En algunas realizaciones, un ARNg modificado químicamente comprende una cadena principal de ácido nucleico modificado. En algunas realizaciones, un ARNg modificado químicamente comprende un residuo de 2'-O-metil-fosforotioato. En algunas realizaciones, las modificaciones químicas potencian la estabilidad, reducen la probabilidad o el grado de respuesta inmunitaria innata y/o potencian otros atributos, como se describe en la técnica.

En algunas realizaciones, un ARNg modificado puede comprender una cadena principal modificada, por ejemplo, fosforotioatos, fosfotriésteres, morfolinos, metilfosfonatos, enlaces interazúcares alquilo o cicloalquilo de cadena corta o enlaces interazúcares heteroatómicos o heterocíclicos de cadena corta.

Se describen compuestos a base de morfolino en Braasch y David Corey, *Biochemistry*, 2002, 41(14): 4503-4510; *Genesis*, 2001, Volumen 30, Edición 3; *Heasman, Dev. Biol.*, 2002, 243: 209-214; Nasevicius *et al.*, *Nat. Genet.*, 2000, 26:216-220; Lacerra *et al.*, *Proc. Natl. Acad. Sci.*, 2000, 97: 9591-9596.; y la Pat. de EE.UU. N.º 5,034,506, concedida el 23 de julio de 1991.

Se describen miméticos de oligonucleótidos de ácido nucleico de ciclohexenilo en Wang *et al.*, *J. Am. Chem. Soc.*, 2000, 122: 8595-8602.

En algunas realizaciones, un ARNg modificado puede comprender uno o más restos de azúcar sustituidos, por ejemplo, uno de los siguientes en la posición 2': OH, SH, SCH<sub>3</sub>, F, OCN, OCH<sub>3</sub>, OCH<sub>3</sub> O(CH<sub>2</sub>)<sub>n</sub> CH<sub>3</sub>, O(CH<sub>2</sub>)<sub>n</sub> NH<sub>2</sub> u O(CH<sub>2</sub>)<sub>n</sub> CH<sub>3</sub>, en donde n es de 1 a aproximadamente 10; alquilo inferior C1 a C10, alcoxialcoxi, alquilo inferior sustituido, alcarilo o aralquilo; Cl; Br; CN; CF<sub>3</sub>; OCF<sub>3</sub>; O-, S- o N-alquilo; O-, S- o N-alquenilo; SOCH<sub>3</sub>; SO<sub>2</sub> CH<sub>3</sub>; ONO<sub>2</sub>; NO<sub>2</sub>; N<sub>3</sub>; NH<sub>2</sub>heterocicloalquilo; heterocicloalcarilo; aminoalquilamino; polialquilamino; sililo sustituido; un grupo de escisión de ARN; un grupo indicador; un intercalador; 2'-O-(2-metoxietilo); 2'-metoxi (2'-O-CH<sub>3</sub>); 2'-propoxi (2'-OCH<sub>2</sub> CH<sub>2</sub>CH<sub>3</sub>); y 2'-fluoro (2'-F). También pueden efectuarse modificaciones similares en otras posiciones en el ARNg, en particular la posición 3' del azúcar en el nucleótido del extremo 3' y la posición 5' del nucleótido del extremo 5'. En algunos ejemplos, se pueden reemplazar tanto el azúcar como el enlace internucleosídico, es decir, la cadena principal, de las unidades de nucleótido con nuevos grupos.

Los ARN guía también pueden incluir, además o como alternativa, modificaciones o sustituciones de nucleobases (normalmente citadas en la técnica simplemente como "base"). Como se usan en el presente documento, las nucleobases "no modificadas" o "naturales" incluyen adenina (A), guanina (G), timina (T), citosina (C) y uracilo (U). Las nucleobases modificadas incluyen las nucleobases halladas únicamente de manera infrecuente o transitoria

en ácidos nucleicos naturales, por ejemplo, hipoxantina, 6-metiladenina, 5-Me-pirimidinas, en particular, 5-metilcitosina (también denominada 5-metil-2'-desoxicitosina y normalmente denominada en la técnica 5-Me-C), 5-hidroximetilcitosina (HMC), glucosil-HMC y gentobiosil-HMC, así como nucleobases sintéticas, por ejemplo, 2-aminoadenina, 2-(metilamino)adenina, 2-(imidazoalquil)adenina, 2-(aminoalquilamino)adenina u otras alquiladeninas heterosustituidas, 2-tiouracilo, 2-tiotimina, 5-bromouracilo, 5-hidroximetiluracilo, 8-azaguanina, 7-desazaguanina, N6 (6-aminohehexil)adenina y 2,6-diaminopurina. Kornberg, A., "DNA Replication", W. H. Freeman & Co., San Francisco, pág. 75-77, 1980; Gebeyehu *et al.*, Nucl. Acids Res. 1997, 15:4513. También puede incluirse una base "universal" conocida en la técnica, por ejemplo, inosina. Se ha demostrado que las sustituciones de 5-Me-C aumentan la estabilidad del dúplex de ácido nucleico en 0.6-1.2 °C. (Sanghvi, Y. S., en Crooke, S. T. y Lebleu, B., eds., "Antisense Research and Applications", CRC Press, Boca Ratón, 1993, pág. 276-278) y son aspectos de sustituciones de bases.

Las nucleobases modificadas pueden comprender otras nucleobases sintéticas y naturales, tales como 5-metilcitosina (5-me-C), 5-hidroximetil-citosina, xantina, hipoxantina, 2-aminoadenina, 6-metilo y otros derivados de alquilo de adenina y guanina, 2-propilo y otros derivados de alquilo de adenina y guanina, 2-tiouracilo, 2-tiotimina y 2-tiocitosina, 5-halouracilo y citosina, 5-propinil uracilo y citosina, 6-azo uracilo, citosina y timina, 5-uracilo (pseudouracilo), 4-tiouracilo, 8-halo, 8-amino, 8-tiol, 8-tioalquilo, 8-hidroxilo y otras adeninas y guaninas 8-sustituidas, 5-halo, en particular 5-bromo, 5-trifluorometil y otros uracilos y citosinas 5-sustituidos, 7-metilguanina y 7-metiladenina, 8-azaguanina y 8-azaadenina, 7-desazaguanina y 7-desazaadenina y 3-desazaguanina y 3-desazaadenina.

#### *Complejos de un ácido nucleico dirigido al genoma y una endonucleasa*

Un ARNg interactúa con una endonucleasa (por ejemplo, una nucleasa guiada por ARN tal como Cas9), formando así un complejo. El ARNg guía a la endonucleasa hacia un polinucleótido diana.

La endonucleasa y el ARNg pueden administrarse cada uno por separado a una célula o un sujeto. En algunas realizaciones, la endonucleasa se puede complejar previamente con uno o más ARN guía, o uno o más ARNcr junto con un ARNtracr. El material previamente complejoado puede administrarse posteriormente a una célula o un sujeto. Dicho material previamente complejoado se conoce como una partícula de ribonucleoproteína (RNP). La endonucleasa de la RNP puede ser, por ejemplo, una endonucleasa Cas9 o una endonucleasa Cpf1. La endonucleasa puede estar flanqueada en el extremo N, el extremo C, o tanto en el extremo N como en el extremo C por una o más señales de localización nuclear (NLS). Por ejemplo, una endonucleasa Cas9 puede estar flanqueada por dos NLS, una NLS ubicada en el extremo N y la segunda NLS ubicada en el extremo C. La NLS puede ser cualquier NLS

conocida en la técnica, tal como una NLS de SV40. La relación molar entre el ácido nucleico dirigido al genoma con respecto a la endonucleasa en la RNP puede variar de aproximadamente 1:1 a aproximadamente 10:1. Por ejemplo, la relación molar del ARNsg con respecto a la endonucleasa Cas9 en la RNP puede ser de 3:1.

#### 5 *Ácidos nucleicos que codifican componentes de sistema*

La presente divulgación proporciona un ácido nucleico que comprende una secuencia de nucleótidos que codifica un ácido nucleico dirigido al genoma de la divulgación, una endonucleasa de la divulgación y/o cualquier ácido nucleico o molécula proteica necesario para llevar a cabo los aspectos de los métodos de la divulgación. Los ácidos nucleicos codificantes  
10 pueden ser ARN, ADN o una combinación de los mismos.

El ácido nucleico que codifica un ácido nucleico dirigido al genoma de la divulgación, una endonucleasa de la divulgación y/o cualquier ácido nucleico o molécula proteica necesaria para realizar los aspectos de los métodos de la divulgación puede comprender un vector (por ejemplo, un vector de expresión recombinante).

El término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro  
15 ácido nucleico al que se ha unido. Un tipo de vector es un "plásmido", el cual se refiere a un bucle de ADN bicatenario circular en donde pueden ligarse segmentos de ácido nucleico adicionales. Otro tipo de vector es un vector vírico, en donde pueden ligarse segmentos de ácido nucleico adicionales en el genoma vírico. Determinados vectores son capaces de replicarse de  
20 manera autónoma en una célula huésped en la que se introducen (por ejemplo, vectores bacterianos que tienen un origen de replicación bacteriano y vectores de mamífero episómicos). Otros vectores (por ejemplo, vectores de mamífero no episómicos) se integran en el genoma de una célula huésped tras su introducción en la célula huésped y de este modo, se replican junto con el genoma del huésped.

En algunos ejemplos, los vectores pueden ser capaces de dirigir la expresión de ácidos  
25 nucleicos a los que están unidos operativamente. Dichos vectores se denominan en el presente documento "vectores de expresión recombinante" o de un modo más simple, "vectores de expresión", que cumplen funciones equivalentes.

La expresión "unida operativamente" significa que la secuencia de nucleótido de interés  
30 está unida a una o más secuencias reguladoras de una manera que permite la expresión de la secuencia de nucleótidos. La expresión "secuencia reguladora" pretende incluir, por ejemplo, promotores, potenciadores y otros elementos de control de la expresión (por ejemplo, señales de poliadenilación). Dichas secuencias reguladoras se conocen bien en la técnica y se describen, por ejemplo, en Goeddel; "Gene Expression Technology: Methods in Enzymology",  
35 1990, 185, Academic Press, San Diego, CA. Las secuencias reguladoras incluyen aquellas que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células

huésped y aquellas que dirigen la expresión de la secuencia de nucleótidos únicamente en determinadas células huésped (por ejemplo, secuencias reguladoras específicas de tejido). Los expertos en la materia apreciarán que el diseño del vector de expresión puede depender de factores tales como la elección de la célula diana, el nivel de expresión deseado y similares.

5 Los vectores de expresión contemplados incluyen, pero sin limitación, vectores víricos basados en virus variolovacunal, poliovirus, adenovirus, virus adenoasociado, SV40, virus del herpes simple, virus de la inmunodeficiencia humana, retrovirus (por ejemplo, virus de la leucemia murina, virus de la necrosis del bazo y vectores derivados de retrovirus, tales como el virus del sarcoma de Rous, el virus del sarcoma de Harvey, el virus de la leucosis aviar, un  
10 lentivirus, virus de la inmunodeficiencia humana, virus del sarcoma mieloproliferativo y virus de tumor mamario) y otros vectores recombinantes. Otros vectores contemplados para las células diana eucariotas incluyen, pero sin limitación, los vectores pXT1, pSG5, pSVK3, pBPV, pMSG y pSVLSV40 (Pharmacia). Se pueden usar otros vectores, en tanto que sean compatibles con la célula huésped.

15 En algunos ejemplos, un vector puede comprender uno o más elementos de control de la transcripción y/o traducción. Dependiendo del sistema de huésped/vector utilizado, se puede usar cualquiera de una serie de elementos de control de la transcripción y la traducción adecuados, incluyendo promotores constitutivos e inducibles, elementos potenciadores de la transcripción, terminadores de la transcripción, etc. en el vector de expresión. El vector puede  
20 ser un vector auto-inactivante que inactive las secuencias víricas o los componentes de la maquinaria de CRISPR u otros elementos.

Los ejemplos no limitantes de promotores eucariotas adecuados (es decir, promotores funcionales en una célula eucariota) incluyen los de citomegalovirus (CMV) inmediato temprano, timidina cinasa del virus del herpes simple (VHS), SV40 temprano y tardío, repeticiones  
25 terminales largas (LTR) de retrovirus, promotor del factor de elongación humano 1 $\alpha$  (EF1 $\alpha$ ), promotor de beta-actina de pollo (CBA), promotor de ubiquitina C (UBC), una construcción híbrida que comprende el potenciador de citomegalovirus fusionado con el promotor de beta-actina de pollo (CAG), una construcción híbrida que comprende el potenciador del citomegalovirus fusionado con el promotor, el primer exón y el primer intrón del gen de la beta-actina de pollo (CAG o CAGGS), el promotor del virus de células madre murinas (MSCV), el  
30 promotor del locus fosfoglicerato cinasa 1 (PGK) y el promotor de la metalotioneína de ratón I.

Un promotor puede ser un promotor inducible (por ejemplo, un promotor de choque térmico, promotor regulado por tetraciclina, promotor regulado por esteroides, promotor regulado por metales, promotor regulado por estrógenos, etc.). El promotor puede ser un  
35 promotor constitutivo (por ejemplo, promotor de CMV, promotor de UBC, promotor CAG). En algunos casos, el promotor puede ser un promotor espacialmente restringido y/o temporalmente

restringido (por ejemplo, un promotor específico del tejido, un promotor específico del tipo de célula, etc.).

La introducción de los complejos, polipéptidos y ácidos nucleicos de la divulgación en células puede producirse mediante infección vírica o por bacteriófago, transfección, conjugación, fusión de protoplastos, lipofección, electroporación, nucleofección, precipitación por fosfato de calcio, transfección mediada por polietilenimina (PEI), transfección mediada por DEAE-dextrano, transfección mediada por liposomas, tecnología de cañón de partículas, precipitación por fosfato de calcio, microinyección directa, administración de ácido nucleico mediada por nanopartículas y similares.

### **III. Estrategias para evadir la respuesta inmunitaria y aumentar la supervivencia**

En el presente documento se describen estrategias para permitir que las células modificadas genéticamente, es decir, las células donantes universales, aumenten su supervivencia o viabilidad y/o evadan una respuesta inmunitaria después del injerto en un sujeto. En algunas realizaciones, estas estrategias permiten que las células donantes universales sobrevivan y/o evadan la respuesta inmunitaria con tasas de éxito superiores a las de una célula no modificada. En algunas realizaciones, las células modificadas genéticamente comprenden la introducción de al menos una modificación genética dentro o cerca de al menos un gen que codifica un factor de supervivencia, en donde la modificación genética comprende una inserción de un polinucleótido que codifica un factor tolerogénico. Las células donantes universales pueden comprender además al menos una modificación genética dentro o cerca de un gen que codifica uno o más antígenos leucocitarios humanos MHC-I o MHC-II o un componente o un regulador de la transcripción de un complejo MHC-I o MHC-II, en donde dicha modificación genética comprende una inserción de un polinucleótido que codifica un segundo factor tolerogénico.

En algunas realizaciones, las células modificadas genéticamente comprenden la introducción de al menos una modificación genética dentro o cerca de al menos un gen que disminuye la expresión de uno o más antígenos leucocitarios humanos MHC-I y MHC-II en relación con una célula no modificada; al menos una modificación genética que aumenta la expresión de al menos un polinucleótido que codifica un factor tolerogénico en relación con una célula no modificada; y al menos una modificación genética que altera la expresión de al menos un gen que codifica un factor de supervivencia en relación con una célula no modificada. En otras realizaciones, las células modificadas genéticamente comprenden al menos una mutación por delección o por inserción-delección dentro o cerca de al menos un gen que altera la expresión de uno o más antígenos leucocitarios humanos MHC-I y MHC-II en relación con una célula no modificada; y al menos una inserción de un polinucleótido que codifica al menos un factor tolerogénico en un sitio que se solapa parcialmente con, se solapa completamente con o está

contenido dentro del sitio de una delección de un gen que altera la expresión de uno o más HLA de MHC-I y MHC-II. En otras realizaciones más, las células modificadas genéticamente comprenden al menos una modificación genética que altera la expresión de al menos un gen que codifica un factor de supervivencia en relación con una célula no modificada.

5 Los genes que codifican el complejo principal de histocompatibilidad (MHC) se ubican en el cromosoma humano. 6p21. Las proteínas resultantes codificadas por los genes MHC son una serie de proteínas de superficie que son esenciales para la compatibilidad del donante durante el trasplante celular. Los genes MHC se dividen en MHC de clase I (MHC-I) y MHC de clase II (MHC-II). Los genes MHC-I (HLA-A, HLA-B y HLA-C) se expresan en casi todos los tipos de células tisulares, presentando péptidos procesados con antígenos "no propios" a los linfocitos T CD8+, potenciando así su activación en los linfocitos T CD8+ citolíticos. Las células trasplantadas o injertadas que expresan moléculas MHC-I "no propias" provocarán una potente respuesta inmunitaria celular dirigida a estas células y, en última instancia, darán como resultado su desaparición por medio de los linfocitos T CD8+ citolíticos activados. Las proteínas MHC-I están íntimamente asociadas con la beta-2-microglobulina (B2M) en el retículo endoplásmico, que es esencial para formar moléculas MHC-I funcionales en la superficie celular. Además, hay tres moléculas MHC-Ib no convencionales (HLA-E, HLA-F y HLA-G), que tienen funciones de regulación inmunitaria. Las biomoléculas MHC-II incluyen HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ y HLA-DR. Debido a su función principal en la respuesta inmunitaria, las biomoléculas MHC-I y MHC-II contribuyen al rechazo inmunitario después del injerto celular de células no huésped, por ejemplo, injerto celular con fines de medicina regenerativa.

15 Las moléculas de superficie celular MHC-I están compuestas por cadenas pesadas codificadas por MHC (HLA-A, HLA-B o HLA-C) y la subunidad invariante beta-2-microglobulina (B2M). Por lo tanto, una reducción en la concentración de B2M dentro de una célula permite un método eficaz para reducir la expresión en la superficie celular de las moléculas de la superficie celular MHC-I.

25 En algunas realizaciones, una célula comprende una modificación genómica de uno o más genes MHC-I o MHC-II. En algunas realizaciones, una célula comprende una modificación genómica de una o más secuencias polinucleotídicas que regulan la expresión de MHC-I y/o MHC-II. En algunas realizaciones, una modificación genética de la divulgación se realiza usando cualquier método de edición génica, incluyendo, pero sin limitación, los métodos descritos en el presente documento.

30 En algunas realizaciones, la disminución de la expresión de uno o más antígenos leucocitarios humanos MHC-I y MHC-II en relación con una célula no modificada se logra mediante la manipulación dirigida, por ejemplo, para la eliminación y/o inserción genética de al menos un par de bases, en un gen MHC-I y/o gen MHC-II directamente. En algunas

35

realizaciones, la disminución de la expresión de uno o más antígenos leucocitarios humanos MHC-I y MHC-II en relación con una célula no modificada se logra mediante direccionamiento, por ejemplo, para la delección genética, de un gen CIITA. En algunas realizaciones, la disminución de la expresión de uno o más antígenos leucocitarios humanos MHC-I y MHC-II en relación con una célula no modificada se logra mediante direccionamiento, por ejemplo, para la delección genética, de al menos un regulador de la transcripción de MHC-I o MHC-II. En algunas realizaciones, un regulador de la transcripción de MHC-I o MHC-II es un gen NLRC5 o CIITA. En algunas realizaciones, un regulador de la transcripción de MHC-I o MHC-II es un gen RFX5, RFXAP, RFXANK, NFY-A, NFY-B, NFY-C, IRF-1 y/o TAP1.

En algunas realizaciones, el genoma de una célula se ha modificado para eliminar la totalidad o parte de un gen HLA-A, HLA-B y/o HLA-C. En algunas realizaciones, el genoma de una célula se ha modificado para eliminar la totalidad o parte de una región promotora de un gen HLA-A, HLA-B y/o HLA-C. En algunas realizaciones, el genoma de una célula se ha modificado para eliminar la totalidad o parte de un gen que codifica un regulador de la transcripción de MHC-I o MHC-II. En algunas realizaciones, el genoma de una célula se ha modificado para eliminar la totalidad o parte de una región promotora de un gen que codifica un regulador de la transcripción de MHC-I o MHC-II.

En algunas realizaciones, el genoma de una célula se ha modificado para disminuir la expresión de beta-2-microglobulina (B2M). B2M es un gen no polimórfico que codifica una subunidad de proteína común necesaria para la expresión superficial de todas las cadenas pesadas polimórficas del MHC de clase I. Las proteínas HLA-I están íntimamente asociadas con B2M en el retículo endoplásmico, que es esencial para formar moléculas HLA-I funcionales expresadas en la superficie celular. En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen B2M que comprende una secuencia 5'-GCTACTCTCTCTTCTGGCC-3' (SEQ ID NO: 1). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen B2M que comprende una secuencia 5'-GGCCGAGATGTCTCGCTCCG-3' (SEQ ID NO: 2). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen B2M que comprende una secuencia 5'-CGCGAGCACAGCTAAGGCCA-3' (SEQ ID NO: 3). En realizaciones alternativas, el ARNg se dirige a un sitio dentro del gen B2M que comprende cualquiera de las siguientes secuencias: 5'-TATAAGTGGAGGCGTCGCGC-3' (SEQ ID NO: 35), 5'-GAGTAGCGCGAGCACAGCTA-3' (SEQ ID NO: 36), 5'-ACTGGACGCGTCGCGCTGGC-3' (SEQ ID NO: 37), 5'-AAGTGGAGGCGTCGCGCTGG-3' (SEQ ID NO: 38), 5'-GGCCACGGAGCGAGACATCT-3' (SEQ ID NO: 39), 5'-GCCCCAATGCTGTCAGCTTC-3' (SEQ ID NO: 40), 5'-CTCGCGCTACTCTCTCTTTC-3' (SEQ ID NO: 41), 5'-TCCTGAAGCTGACAGCATTTC-3' (SEQ ID NO: 42), 5'-TTCCTGAAGCTGACAGCATT-3' (SEQ ID NO: 43) o 5'-ACTCTCTCTTCTGGCCTGG-3' (SEQ ID NO: 44). En algunas realizaciones, el ARNg

comprende una secuencia polinucleotídica de una cualquiera de entre la SEQ ID NO: 1, la SEQ ID NO: 2, la SEQ ID NO: 3, la SEQ ID NO: 35, la SEQ ID NO: 36, la SEQ ID NO: 37, la SEQ ID NO: 38, la SEQ ID NO: 39, la SEQ ID NO: 40, la SEQ ID NO: 41, la SEQ ID NO: 42, la SEQ ID NO: 43 o la SEQ ID NO: 44. El complejo de ARNg/nucleasa CRISPR se dirige y escinde un sitio diana en el locus de B2M. La reparación de una rotura bicatenaria mediante NHEJ puede dar como resultado una delección de al menos un nucleótido y/o una inserción de al menos un nucleótido, interrumpiendo o eliminando así la expresión de B2M. Como alternativa, el locus de B2M puede ser dirigido por al menos dos sistemas CRISPR que comprenden cada uno un ARNg diferente, de modo que la escisión en dos sitios en el locus de B2M conduce a una delección de la secuencia entre los dos cortes, eliminando así la expresión de B2M.

En algunas realizaciones, el genoma de una célula se ha modificado para disminuir la expresión de la proteína de interacción con tiorredoxina (TXNIP). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GAAGCGTGTCTTCATAGCGC-3' (SEQ ID NO: 15). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TTACTCGTGTCAAAGCCGTT-3' (SEQ ID NO: 16). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TGTCAAAGCCGTTAGGATCC-3' (SEQ ID NO: 17). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GCCGTTAGGATCCTGGCTTG-3' (SEQ ID NO: 18). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GCGGAGTGGCTAAAGTGCTT-3' (SEQ ID NO: 19). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TCCGCAAGCCAGGATCCTAA-3' (SEQ ID NO: 20). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GTTTCGGCTTTGAGCTTCCTC-3' (SEQ ID NO: 21). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GAGATGGTGATCATGAGACC-3' (SEQ ID NO: 22). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TTGTACTCATATTTGTTTCC-3' (SEQ ID NO: 23). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-AACAAATATGAGTACAAGTT-3' (SEQ ID NO: 24). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GAAGCGTGTCTTCATAGCGCAGG-3' (SEQ ID NO: 45). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TTACTCGTGTCAAAGCCGTTAGG-3' (SEQ ID NO: 46). En algunas realizaciones, el ARNg se

dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TGTCAAAGCCGTTAGGATCCTGG-3' (SEQ ID NO: 47). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GCCGTTAGGATCCTGGCTTGCGG-3' (SEQ ID NO: 48). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GCCGAGTGGCTAAAGTGCTTTGG-3' (SEQ ID NO: 49). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TCCGCAAGCCAGGATCCTAACGG-3' (SEQ ID NO: 50). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GTTCGGCTTTGAGCTTCCTCAGG-3' (SEQ ID NO: 51). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-GAGATGGTGATCATGAGACCTGG-3' (SEQ ID NO: 52). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-TTGTACTCATATTTGTTTCCAGG-3' (SEQ ID NO: 53). En algunas realizaciones, el ARNg se dirige a un sitio dentro del gen TXNIP que comprende una secuencia 5'-AACAAATATGAGTACAAGTTCGG-3' (SEQ ID NO: 54). En algunas realizaciones, el ARNg se dirige a un sitio diana dentro del gen TXNIP que comprende una secuencia polinucleotídica de una cualquiera de las SEQ ID NO: 15-24 o 45-54. En algunas realizaciones, el ARNg se dirige a una secuencia polinucleotídica de una cualquiera de entre la SEQ ID NO: 15, la SEQ ID NO: 16, la SEQ ID NO: 17, la SEQ ID NO: 18, la SEQ ID NO: 19, la SEQ ID NO: 20, la SEQ ID NO: 21, la SEQ ID NO: 22, la SEQ ID NO: 23 o la SEQ ID NO: 24. El complejo de ARNg/nucleasa CRISPR se dirige y escinde un sitio diana en el locus del gen TXNIP. La reparación de una rotura bicatenaria mediante NHEJ puede dar como resultado una delección de al menos un nucleótido y/o una inserción de al menos un nucleótido, interrumpiendo o eliminando así la expresión de TXNIP. Como alternativa, la inserción de un polinucleótido que codifica un gen exógeno en el locus del gen TXNIP puede interrumpir o eliminar la expresión de TXNIP.

En algunas realizaciones, el genoma de una célula se ha modificado para disminuir la expresión del transactivador de Clase II (CIITA). CIITA es un miembro de la familia de proteínas de repetición rica en leucina (LRR, *Leucine-Rich Repeat*) LR o de dominio de unión de nucleótidos (NBD, *Nucleotide Binding Domain*) y regula la transcripción de MHC-II al asociarse con el potencialesoma de MHC. La expresión de CIITA se induce en los linfocitos B y las células dendríticas en función de la etapa de desarrollo y es inducible por IFN- $\gamma$  en la mayoría de los tipos de células.

En algunas realizaciones, el genoma de una célula se ha modificado para disminuir la expresión del dominio 5 que contiene CARD de la familia NLR (NLRC5, por sus siglas en inglés). El NLRC5 es un regulador crítico de las respuestas inmunitarias mediadas por MHC-I y, al igual

que CIITA, NLRC5 es altamente inducible por IFN- $\gamma$  y puede translocarse al núcleo. El NLRC5 activa los promotores de genes MHC-I e induce la transcripción de MHC-I, así como genes relacionados implicados en la presentación del antígeno MHC-I.

En algunas realizaciones, se pueden insertar o reinsertar factores tolerogénicos en células modificadas genéticamente para crear células donantes universales con privilegios 5 inmunitarios. En algunas realizaciones, las células donantes universales desveladas en el presente documento se han modificado adicionalmente para expresar uno o más factores tolerogénicos. Los factores tolerogénicos de ejemplo incluyen, sin limitación, uno o más de HLA-C, HLA-E, HLA-F, HLA-G, PD-L1, CTLA-4-Ig, CD47, inhibidor de CI e IL-35. En algunas 10 realizaciones, la modificación genética, por ejemplo, la inserción, de al menos un polinucleótido que codifica al menos un factor tolerogénico permite a una célula donante universal inhibir o evadir el rechazo inmunitario con tasas al menos 1.05, al menos 1.1, al menos 1.25, al menos 1.5, al menos 2, al menos 3, al menos 4, al menos 5, al menos 10, al menos 20 o al menos 50 veces superiores a las de una célula no modificada después del injerto. En algunas 15 realizaciones, una inserción de un polinucleótido que codifica HLA-E, HLA-G, CTLA-4, CD47 y/o PD-L1 permite que una célula donante universal inhiba o evada el rechazo inmunitario después del trasplante o injerto en un sujeto huésped.

El polinucleótido que codifica el factor tolerogénico generalmente comprende brazos de homología izquierdo y derecho que flanquean la secuencia que codifica el factor tolerogénico. 20 Los brazos de homología tienen una homología de secuencia sustancial con el ADN genómico en o cerca del sitio de inserción diana. Por ejemplo, el brazo de homología izquierdo puede ser una secuencia de nucleótidos homóloga con una región ubicada a la izquierda o cadena arriba del sitio diana o del sitio de corte, y el brazo de homología derecho puede ser una secuencia de nucleótidos homóloga con una región ubicada a la derecha o cadena abajo del sitio diana o el 25 sitio de corte. El extremo proximal de cada brazo de homología puede ser homólogo a la secuencia de ADN genómico que linda con el sitio de corte. Como alternativa, el extremo proximal de cada brazo de homología puede ser homólogo a la secuencia de ADN genómico ubicada hasta aproximadamente 10, 20, 30, 40, 50, 60 o 70 nucleobases de distancia del sitio de corte. Como tal, el polinucleótido que codifica el factor tolerogénico se puede insertar en el 30 locus del gen diana a menos de aproximadamente 10, 20, 30, 40, 50, 60 o 70 pares de bases del sitio de corte, y el ADN genómico adicional que bordea el sitio de corte (y que no tiene homología con un brazo homólogo) se puede eliminar. Los brazos de homología pueden variar en longitud de aproximadamente 50 nucleótidos a varios miles de nucleótidos. En algunas realizaciones, los brazos de homología pueden variar en longitud de aproximadamente 35 500 nucleótidos a aproximadamente 1000 nucleótidos. La homología de secuencia sustancial entre los brazos de homología y el ADN genómico puede ser de al menos aproximadamente el

80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 % o al menos aproximadamente el 99 %.

En algunas realizaciones, los brazos de homología se usan con guías de B2M (por ejemplo, ARNg que comprenden una secuencia de nucleótidos de las SEQ ID NO: 1-3, 35-44).

5 En algunas realizaciones, los brazos de homología se diseñan para usarse con cualquier guía de B2M que eliminaría el sitio de inicio del gen B2M. En algunas realizaciones, los brazos de homología de B2M pueden comprender o consistir esencialmente en una secuencia polinucleotídica de la SEQ ID NO: 7 o 13, o una secuencia polinucleotídica que tiene al menos un 85 %, 90 %, 95 % o 99 % de identidad de secuencia con la de la SEQ ID NO: 7 o 13. En  
10 algunas realizaciones, el brazo de homología de B2M izquierdo puede comprender o consistir esencialmente en la SEQ ID NO: 7, o una secuencia polinucleotídica que tiene al menos un 85 %, 90 %, 95 % o 99 % de identidad de secuencia con la de la SEQ ID NO: 7. En algunas realizaciones, el brazo de homología de B2M derecho puede comprender o consistir esencialmente en la SEQ ID NO: 13, o una secuencia polinucleotídica que tiene al menos un  
15 85 %, 90 %, 95 % o 99 % de identidad de secuencia con la de la SEQ ID NO: 13.

En algunas realizaciones, los brazos de homología se usan con guías de TXNIP (por ejemplo, ARNg que comprenden una secuencia de nucleótidos de las SEQ ID NO: 15-24). En algunas realizaciones, los brazos de homología se diseñan para usarse con cualquier guía de TXNIP que se dirija al exón 1 de TXNIP (por ejemplo, ARNg que comprenden una secuencia de  
20 nucleótidos de las SEQ ID NO: 15-20). En algunas realizaciones, los brazos de homología de TXNIP pueden comprender o consistir esencialmente en una secuencia polinucleotídica de la SEQ ID NO: 25 o 32, o una secuencia polinucleotídica que tiene al menos un 85 %, 90 %, 95 % o 99 % de identidad de secuencia con la de la SEQ ID NO: 25 o 32. En algunas realizaciones, el brazo de homología de TXNIP izquierdo puede comprender o consistir esencialmente en la  
25 SEQ ID NO: 25, o una secuencia polinucleotídica que tiene al menos un 85 %, 90 %, 95 % o 99 % de identidad de secuencia con la de la SEQ ID NO: 25. En algunas realizaciones, el brazo de homología de TXNIP derecho puede comprender o consistir esencialmente en la SEQ ID NO: 32, o una secuencia polinucleotídica que tiene al menos un 85 %, 90 %, 95 % o 99 % de identidad de secuencia con la de la SEQ ID NO: 32.

30 El al menos un polinucleótido que codifica al menos un factor tolerogénico puede unirse operativamente a un promotor exógeno. El promotor exógeno puede ser un promotor constitutivo, inducible, temporal, específico del tejido o específico del tipo de célula. En algunas realizaciones, el promotor exógeno es un promotor de CMV, EF1a, PGK, CAG o UBC.

En algunas realizaciones, el al menos un polinucleótido que codifica al menos un factor tolerogénico se inserta en un locus seguro, por ejemplo, el locus de AAVS 1. En algunas  
35 realizaciones, el al menos un polinucleótido que codifica al menos un factor tolerogénico se

inserta en un sitio o una región de ADN genómico que se solapa parcialmente con, se solapa completamente con o está contenido dentro (es decir, está dentro o cerca) de un gen MHC-I, gen MHC-II o un regulador de la transcripción de MHC-I o MHC-II.

5 En algunas realizaciones, un polinucleótido que codifica PD-L1 se inserta en un sitio dentro o cerca de un locus del gen B2M. En algunas realizaciones, un polinucleótido que codifica PD-L1 se inserta en un sitio dentro o cerca de un locus del gen B2M concurrente con o después de una delección de todo o parte de un gen o promotor de B2M. El polinucleótido que codifica PD-L1 está unido operativamente a un promotor exógeno. El promotor exógeno puede ser un promotor de CMV. En algunas realizaciones, el polinucleótido comprende una secuencia de  
10 nucleótidos de la SEQ ID NO: 11.

En algunas realizaciones, un polinucleótido que codifica HLA-E se inserta en un sitio dentro o cerca de un locus del gen B2M. En algunas realizaciones, un polinucleótido que codifica HLA-E se inserta en un sitio dentro o cerca de un locus del gen B2M concurrente con o después de una delección de todo o parte de un gen o promotor de B2M. El polinucleótido que codifica  
15 HLA-E está unido operativamente a un promotor exógeno. El promotor exógeno puede ser un promotor de CMV. En algunas realizaciones, el polinucleótido comprende una secuencia de nucleótidos de la SEQ ID NO: 26, 27, 28, 29, 30 y/o 30. En algunas realizaciones, el polinucleótido comprende una secuencia de nucleótidos de la SEQ ID NO: 55.

En algunas realizaciones, un polinucleótido que codifica HLA-G se inserta en un sitio dentro o cerca de un locus del gen HLA-A, HLA-B o HLA-C. En algunas realizaciones, un polinucleótido que codifica HLA-G se inserta en un sitio dentro o cerca de un locus del gen HLA-A, HLA-B o HLA-C concurrente con o después de una delección de un gen o promotor de HLA-A, HLA-B o HLA-C.  
20

En algunas realizaciones, un polinucleótido que codifica CD47 se inserta en un sitio dentro o cerca de un locus del gen CIITA. En algunas realizaciones, un polinucleótido que codifica CD47 se inserta en un sitio dentro o cerca de un locus del gen CIITA concurrente con o después de una delección de un gen o promotor de CIITA.  
25

En algunas realizaciones, un polinucleótido que codifica HLA-G se inserta en un sitio dentro o cerca de un locus del gen HLA-A, HLA-B o HLA-C concurrente con la inserción de un polinucleótido que codifica CD47 en un sitio dentro o cerca de un locus del gen CIITA.  
30

En algunas realizaciones, el al menos un polinucleótido que codifica al menos un factor tolerogénico puede suministrarse a las células como parte de un vector. Por ejemplo, el vector puede ser un vector plasmídico. En diversas realizaciones, la cantidad de vector plasmídico suministrado a las células puede variar de aproximadamente 0.5 µg a aproximadamente 10 µg (por aproximadamente 10<sup>6</sup> células). En algunas realizaciones, la cantidad de plásmido puede  
35 variar de aproximadamente 1 µg a aproximadamente 8 µg, de aproximadamente 2 µg a

aproximadamente 6 µg o de aproximadamente 3 µg a aproximadamente 5 µg. En realizaciones específicas, la cantidad de plásmido suministrado a las células puede ser de aproximadamente 4 µg.

5 En algunas realizaciones, una célula comprende una expresión aumentada o disminuida de uno o más factores de supervivencia. En algunas realizaciones, una célula comprende una inserción de una o más secuencias polinucleotídicas que codifican un factor de supervivencia. En algunas realizaciones, una célula comprende una delección de uno o más factores de supervivencia. En algunas realizaciones, una modificación genética de la divulgación se realiza usando cualquier método de edición génica, incluyendo, pero sin limitación, los métodos  
10 descritos en el presente documento. En algunas realizaciones, una célula comprende una expresión aumentada o disminuida de al menos un factor de supervivencia en relación con una célula no modificada. En algunas realizaciones, un factor de supervivencia es un miembro o una vía crítica implicada en la supervivencia celular, por ejemplo, la hipoxia, las especies reactivas de oxígeno, la privación de nutrientes y/o el estrés oxidativo. En algunas realizaciones, la modificación genética de al menos un factor de supervivencia permite que una célula donante  
15 universal sobreviva durante un período de tiempo más largo, por ejemplo, al menos 1.05, al menos 1.1, al menos 1.25, al menos 1.5, al menos 2, al menos 3, al menos 4, al menos 5, al menos 10, al menos 20 o al menos 50 veces más tiempo que una célula no modificada después del injerto. En algunas realizaciones, un factor de supervivencia es ZNF143, TXNIP, FOXO1, JNK o MANF.  
20

En algunas realizaciones, una célula que comprende una inserción de un polinucleótido que codifica MANF permite que una célula donante universal sobreviva después del trasplante o injerto en un sujeto huésped con tasas de supervivencia superiores en relación con una célula no modificada. En algunas realizaciones, un polinucleótido que codifica MANF se inserta en un  
25 locus seguro. En algunas realizaciones, un polinucleótido que codifica MANF se inserta en un gen que pertenece a un MHC-I, MHC-II o un regulador de la transcripción de MHC-I o MHC-II.

En algunas realizaciones, el genoma de una célula se ha modificado para eliminar la totalidad o parte de un gen ZNF143, TXNIP, FOXO1 y/o JNK. En algunas realizaciones, el genoma de una célula se ha modificado para eliminar la totalidad o parte de una región promotora de un gen ZNF143, TXNIP, FOXO1 y/o JNK.  
30

En algunas realizaciones, se modifica genéticamente más de un factor de supervivencia dentro de una célula.

En determinadas realizaciones, las células que no tienen expresión de MHC-II y una expresión moderada de MHC-I se modifican genéticamente para que no tengan expresión  
35 superficial de MHC-I o MHC-II. En otra realización, las células sin expresión superficial de MHC-I/II se editan adicionalmente para que tengan expresión de PD-L1, por ejemplo, la inserción de

un polinucleótido que codifica PD-L1. En otra realización más, las células sin expresión superficial de MHC-I/II se editan adicionalmente para que tengan expresión de PD-L1, por ejemplo, la inserción de un polinucleótido que codifica PD-L1, y también se modifican genéticamente para aumentar o disminuir la expresión de al menos un gen que codifica un factor de supervivencia en relación con una célula no modificada.

En algunas realizaciones, las células comprenden además una expresión aumentada o disminuida, por ejemplo, mediante una modificación genética, de uno o más genes adicionales que no participan necesariamente en la inmunoevasión o la supervivencia celular después del injerto. En algunas realizaciones, las células comprenden además un aumento de la expresión de una o más proteínas de interruptor de seguridad en relación con una célula no modificada. En algunas realizaciones, las células comprenden un aumento de la expresión de uno o más genes adicionales que codifican una proteína de interruptor de seguridad. En algunas realizaciones, un interruptor de seguridad también es un gen suicida. En algunas realizaciones, un interruptor de seguridad es la timidina cinasa del virus del herpes simple-1 (HSV-tk) o la caspasa-9 inducible. En algunas realizaciones, un polinucleótido que codifica al menos un interruptor de seguridad se inserta en un genoma, por ejemplo, en un locus seguro. En algunas realizaciones diferentes, el uno o más genes adicionales que están modificados genéticamente codifican una o más de las proteínas de interruptor de seguridad; modalidades de direccionamiento; receptores; moléculas de señalización; factores de transcripción; proteínas o péptidos farmacéuticamente activos; candidatos a fármacos diana; y proteínas que promueven el injerto, el tráfico, el asentamiento, la viabilidad, la autorrenovación, la persistencia y/o la supervivencia de las mismas integradas con la construcción.

Un aspecto de la presente invención proporciona un método de generación de células donantes universales genomanipuladas, en donde una célula donante universal comprende al menos una modificación genómica dirigida en uno o más sitios seleccionados del genoma, comprendiendo el método manipular genéticamente un tipo de célula como se describe en el presente documento mediante la introducción en dichas células de una o más construcciones para permitir la modificación dirigida en el sitio seleccionado; introducir en dichas células de una o más roturas de doble cadena en los sitios seleccionados usando una o más endonucleasas capaces de reconocer el sitio seleccionado; y cultivar las células editadas para permitir la reparación del ADN endógeno para generar inserciones o deleciones dirigidas en los sitios seleccionados; obteniéndose de este modo células donantes universales modificadas genómicamente. Las células donantes universales modificadas genómicamente pueden experimentar rondas sucesivas de modificación genómica de modo que se dirijan y se modifiquen múltiples sitios. Las células con modificación genómica se cultivan, se caracterizan, se seleccionan y se expanden usando técnicas ya conocidas en la técnica. Las células donantes

universales generadas mediante este método comprenderán al menos una modificación genómica dirigida funcional, y en donde las células con modificación genómica, si son células madre, pueden entonces diferenciarse en células progenitoras o células completamente diferenciadas.

5 En algunas otras realizaciones, las células donantes universales genomanipuladas comprenden la expresión introducida o aumentada en al menos uno de HLA-E, HLA-G, CD47 o PD-L1. En algunas realizaciones, las células donantes universales genomanipuladas son deficientes en HLA de clase I y/o clase II. En alguna realización, las células donantes universales genomanipuladas comprenden B2M nulo o bajo. En algunas realizaciones, las células donantes universales genomanipuladas comprenden polinucleótidos exógenos integrados o no integrados que codifican una o más de las proteínas HLA-E, HLA-G y PD-L1. En algunas realizaciones, dicha expresión introducida es una expresión aumentada de genes no expresados o poco expresados comprendidos en dichas células. En algunas realizaciones, los polinucleótidos exógenos no integrados se introducen usando virus Sendai, AAV, episoma o plásmido. En alguna realización, las células donantes universales son nulas para B2M, con expresión introducida de uno o más de HLA-E, HLA-G, PD-L1 y expresión aumentada o disminuida de al menos una proteína de interruptor de seguridad. En otra realización, las células donantes universales son nulas para HLA-A, HLA-B y HLA-C, con expresión introducida de uno o más de HLA-E, HLA-G, PD-L1 y al menos una proteína de interruptor de seguridad. En alguna realización, las células donantes universales son nulas para B2M, con expresión introducida de uno o más de HLA-E, HLA-G PD-L1 y expresión aumentada o disminuida de al menos un factor de supervivencia, por ejemplo, MANF. Se contempla que los métodos para generar cualquiera de las células modificadas genéticamente descritas en el presente documento se realicen usando al menos cualquiera de los métodos de edición génica descritos en el presente documento.

#### **IV. Tipos de células**

Las células como se describen en el presente documento, por ejemplo, las células donantes universales (y las correspondientes células no modificadas) pueden pertenecer a cualquier clase posible de tipo de célula. En algunas realizaciones, una célula, por ejemplo, una célula donante universal (y la correspondiente célula no modificada) puede ser una célula de mamífero. En algunas realizaciones, una célula, por ejemplo, una célula donante universal (y la correspondiente célula no modificada) puede ser una célula humana. En algunas realizaciones, una célula, por ejemplo, una célula donante universal (y la correspondiente célula no modificada) puede ser una célula madre. En algunas realizaciones, una célula, por ejemplo, una célula donante universal (y la correspondiente célula no modificada) puede ser una célula madre pluripotente (PSC). En algunas realizaciones, una célula, por ejemplo, una célula donante

universal (y la correspondiente célula no modificada) puede ser una célula madre embrionaria (ESC), una célula madre adulta (ASC), una célula madre pluripotente inducida (iPSC) o una célula madre o progenitora hematopoyética (HSPC) (también denominada célula madre hematopoyética (HSC)). En algunas realizaciones, una célula, por ejemplo, una célula donante universal (y la correspondiente célula no modificada) puede ser una célula diferenciada. En algunas realizaciones, una célula, por ejemplo, una célula donante universal (y la correspondiente célula no modificada) puede ser una célula somática, por ejemplo, una célula del sistema inmunitario o una célula contráctil, por ejemplo, una célula de músculo esquelético.

Las células, por ejemplo, células madre donantes universales, descritas en el presente documento, pueden diferenciarse en tipos de células pertinentes para evaluar la expresión de HLA, así como la evaluación de la inmunogenicidad de las líneas de células madre universales. En general, la diferenciación comprende mantener las células de interés durante un período de tiempo y en condiciones suficientes para que las células se diferencien en las células diferenciadas de interés. Por ejemplo, las células madre universales desveladas en el presente documento pueden diferenciarse en células progenitoras mesenquimatosas (MPC), cardiomiocitos hipoinmunogénicos, células progenitoras musculares, blastocitos, células endoteliales (EC), macrófagos, hepatocitos, células beta (por ejemplo, células beta pancreáticas), progenitores del endodermo pancreático, progenitores endocrinos pancreáticos, células progenitoras hematopoyéticas o células progenitoras neuronales (NPC). En algunas realizaciones, la célula donante universal puede diferenciarse en células del endodermo definitivo, células del tubo intestinal primitivo, células del intestino posterior-anterior, células del endodermo pancreático (PEC), células endocrinas pancreáticas, células beta inmaduras o células beta en maduración.

Las células madre son capaces tanto de proliferar como de dar lugar a más células progenitoras, que, a su vez, tienen la capacidad de generar un gran número de células madre que, a su vez, pueden dar lugar a células hijas diferenciadas o diferenciables. Las células hijas en sí pueden inducirse para proliferar y producir descendencia que posteriormente se diferencia en uno o más tipos celulares maduros, mientras que también se conservan una o más células con potencial de desarrollo parental. La expresión "célula madre" se refiere, por lo tanto, a una célula con la capacidad o el potencial, en circunstancias particulares, de diferenciarse en un fenotipo más especializado o diferenciado y que conserva la capacidad, en determinadas circunstancias, de proliferar sin diferenciarse sustancialmente. En un aspecto, la expresión célula progenitora o madre se refiere a una célula madre generalizada cuyos descendientes (progenie) se especializan, normalmente en diferentes direcciones, mediante diferenciación, por ejemplo, adquiriendo caracteres completamente individuales, como se produce en la diversificación progresiva de las células y tejidos embrionarios. La diferenciación celular es un

proceso complejo que normalmente se produce mediante diversas divisiones celulares. Una célula diferenciada puede derivar de una célula multipotente que en sí procede de una célula multipotente, y así sucesivamente. Aunque cada una de estas células multipotentes pueden considerarse células madre, la variedad de tipos que pueden generar cada una puede variar considerablemente. Algunas células diferenciadas también tienen la capacidad de dar lugar a células con un mayor potencial de desarrollo. Dicha capacidad puede ser natural o puede inducirse artificialmente tras el tratamiento con diversos factores. En muchos casos biológicos, las células madre también pueden ser "multipotentes" debido a que pueden producir descendencia de más de un tipo de célula distinto, pero esto no es necesario para el carácter de "madre".

Una "célula diferenciada" es una célula que ha progresado más a lo largo de la vía de desarrollo que la célula con la que se está comparando. Por lo tanto, las células madre pueden diferenciarse en células precursoras de linaje restringido (tales como una célula progenitora de miocitos), que a su vez, puede diferenciarse en otros tipos de células precursoras posteriores de la vía (tal como un precursor de miocitos) y posteriormente, en una célula diferenciada terminalmente, tal como un miocito, que desempeña un papel característico en un tipo de tejido concreto y puede conservar o no la capacidad para proliferar adicionalmente. En algunas realizaciones, la célula diferenciada puede ser una célula beta pancreática.

#### *Células madre embrionarias*

Las células descritas en el presente documento pueden ser células madre embrionarias (ESC). Las ESC derivan de blastocitos de embriones de mamíferos y pueden diferenciarse en cualquier tipo de célula y propagarse rápidamente. También se cree que las ESC tienen un cariotipo normal, mantienen una alta actividad de la telomerasa y presentan un notable potencial proliferativo a largo plazo, lo que hace que estas células sean excelentes candidatas para su uso como células donantes universales.

#### *Células madre adultas*

Las células descritas en el presente documento pueden ser células madre adultas (ASC). Las ASC son células indiferenciadas que se pueden encontrar en mamíferos, por ejemplo, seres humanos. Las ASC se definen por su capacidad para autorrenovarse, por ejemplo, pasar a través de diversas rondas de replicación celular mientras mantienen su estado indiferenciado, y su capacidad para diferenciarse en diversos tipos de células distintos, por ejemplo, células gliales. Las células madre adultas son una amplia clase de células madre que pueden abarcar células madre hematopoyéticas, células madre mamarias, células madre intestinales, células madre mesenquimatosas, células madre endoteliales, células madre neuronales, células madre adultas olfativas, células madre de la cresta neuronal y células testiculares.

### *Células madre pluripotentes inducidas*

Las células descritas en el presente documento pueden ser células madre pluripotentes inducidas (iPSC). Una iPSC puede generarse directamente a partir de una célula humana adulta mediante la introducción de genes que codifican factores de transcripción críticos implicados en la pluripotencia, por ejemplo, OCT4, SOX2, cMYC y KLF4. Una iPSC puede derivar del mismo sujeto al que se van a administrar las células progenitoras posteriores. Es decir, se puede obtener una célula somática de un sujeto, reprogramarla en una célula madre pluripotente inducida y a continuación volver a diferenciarla en una célula progenitora para administrarla al sujeto (por ejemplo, células autólogas). Sin embargo, en el caso de las células autólogas, persiste el riesgo de respuesta inmunitaria y de escasa viabilidad tras el injerto.

### *Células progenitoras y madre hematopoyéticas humanas*

Las células descritas en el presente documento pueden ser células progenitoras y madre hematopoyéticas humanas (hHSPC). Este linaje de células madre da lugar a todos los tipos de células sanguíneas, incluyendo eritroides (eritrocitos o glóbulos rojos (GR)), mieloides (monocitos y macrófagos, neutrófilos, basófilos, eosinófilos, megacariocitos/plaquetas y células dendríticas) y linfoides (linfocitos T, linfocitos B, linfocitos NK). Las células sanguíneas se producen por la proliferación y diferenciación de una población muy pequeña de células madre hematopoyéticas (HSC) pluripotentes que también tienen la capacidad de reponerse por sí mismas mediante la autorrenovación. Durante la diferenciación, la progenie de las HSC progresa a través de diversas etapas intermedias de maduración, generando células progenitoras de potencial múltiple y de linaje comprometido antes de alcanzar la madurez. La médula ósea (MO) es el sitio principal de hematopoyesis en los seres humanos y, en condiciones normales, solo se pueden encontrar pequeñas cantidades de células progenitoras y madre hematopoyéticas (HSPC) en la sangre periférica (SP). El tratamiento con citocinas, algunos fármacos mielosupresores usados en el tratamiento del cáncer y compuestos que interrumpen la interacción entre las células hematopoyéticas y del estroma de la MO pueden movilizar rápidamente un gran número de células progenitoras y madre hacia la circulación.

### *Diferenciación de células en otros tipos de células*

Otra etapa de los métodos de la presente divulgación puede comprender diferenciar células en células diferenciadas. La etapa de diferenciación se puede realizar de acuerdo con cualquier método conocido en la técnica. Por ejemplo, las iPSC humanas se diferencian en endodermo definitivo usando diversos tratamientos, incluyendo la activina y el complemento B27 (Life Technologies). El endodermo definitivo se diferencia más en hepatocito, y el tratamiento incluye: FGF4, HGF, BMP2, BMP4, Oncostatina M, Dexametasona, etc. (Duan *et al*, *Stem Cells*, 2010;28:674–686; Ma *et al*, *Stem Cells Translational Medicine*, 2013;2:409–419). En otra realización, la etapa de diferenciación puede realizarse de acuerdo con Sawitza *et al*,

*Sci Rep.* 2015; 5:13320. Una célula diferenciada puede ser cualquier célula somática de un mamífero, por ejemplo, un ser humano. En algunas realizaciones, una célula somática puede ser una célula epitelial secretora exocrina (por ejemplo, célula mucosa de la glándula salival, célula de la glándula prostática), una célula secretora de hormonas (por ejemplo, célula de la hipófisis anterior, célula del tracto intestinal, islotes pancreáticos), una célula epitelial queratinizante (por ejemplo, queratinocito epidérmico), una célula epitelial de barrera estratificada húmeda, una célula transductora sensorial (por ejemplo, un fotorreceptor), una célula neuronal autónoma, una célula de soporte de los órganos sensoriales y de soporte de neuronas periféricas (por ejemplo, célula de Schwann), una neurona del sistema nervioso central, una célula glial (por ejemplo, astrocito, oligodendrocito), una célula del cristalino, un adipocito, una célula de riñón, una célula con función de barrera (por ejemplo, una célula de conducto), una célula de la matriz extracelular, una célula contráctil (por ejemplo, célula del músculo esquelético, célula del músculo cardíaco, célula del músculo liso), una célula sanguínea (por ejemplo, eritrocito), una célula del sistema inmunitario (por ejemplo, megacariocito, célula microglial, neutrófilo, mastocito, un linfocito T, un linfocito B, un linfocito citolítico natural), una célula germinal (por ejemplo, espermátida), una célula nodriza o una célula intersticial.

En general, las poblaciones de las células donantes universales desveladas en el presente documento mantienen la expresión de una o más secuencias de nucleótidos insertadas en el tiempo. Por ejemplo, al menos aproximadamente el 50 %, al menos aproximadamente el 55 %, al menos aproximadamente el 60 %, al menos aproximadamente el 65 %, al menos aproximadamente el 70 %, al menos aproximadamente el 75 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 % o al menos aproximadamente el 99 % de las células donantes universales expresan el uno o más factores tolerogénicos. Además, las poblaciones de células de linaje restringido o completamente diferenciadas derivadas de las células donantes universales descritas en el presente documento mantienen la expresión de una o más secuencias de nucleótidos insertadas en el tiempo. Por ejemplo, al menos aproximadamente el 50 %, al menos aproximadamente el 55 %, al menos aproximadamente el 60 %, al menos aproximadamente el 65 %, al menos aproximadamente el 70 %, al menos aproximadamente el 75 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 % o al menos aproximadamente el 99 % de las células de linaje restringido o completamente diferenciadas expresan el uno o más factores tolerogénicos.

## **V. Formulaciones y administraciones**

### *Formulación y suministro para la edición génica*

Los ARN guía, polinucleótidos, por ejemplo, polinucleótidos que codifican un factor tolerogénico o polinucleótidos que codifican una endonucleasa, y endonucleasas como se describe en el presente documento pueden formularse y suministrarse a las células de cualquier manera conocida en la técnica.

Los ARN guía y/o polinucleótidos pueden formularse con excipientes farmacéuticamente aceptables tales como portadores, disolventes, estabilizadores, adyuvantes, diluyentes, etc., dependiendo del modo particular de administración y de la forma farmacéutica. Las composiciones de ARN guía y/o polinucleótidos se pueden formular para lograr un pH fisiológicamente compatible y varían de un pH de aproximadamente 3 a un pH de aproximadamente 11, de un pH de aproximadamente 3 a un pH de aproximadamente 7, dependiendo de la formulación y de la vía de administración. En algunos casos, el pH se puede ajustar a un intervalo pH de aproximadamente 5.0 a pH de aproximadamente 8. En algunos casos, las composiciones pueden comprender una cantidad terapéuticamente eficaz de al menos un compuesto como se describe en el presente documento, junto con uno o más excipientes farmacéuticamente aceptables. Opcionalmente, las composiciones pueden comprender una combinación de los compuestos descritos en el presente documento o pueden incluir un segundo principio activo en el tratamiento o la prevención del crecimiento bacteriano (por ejemplo y sin limitación, agentes antibacterianos o antimicrobianos) o pueden incluir una combinación de reactivos de la presente divulgación.

Los excipientes adecuados incluyen, por ejemplo, moléculas portadoras que incluyen macromoléculas grandes lentamente metabolizadas, tales como proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos y partículas víricas inactivas. Otros excipientes de ejemplo pueden incluir antioxidantes (por ejemplo y sin limitación, ácido ascórbico), agentes quelantes (por ejemplo y sin limitación, EDTA), carbohidratos (por ejemplo y sin limitación, dextrina, hidroxialquilcelulosa e hidroxialquilmelilcelulosa), ácido esteárico, líquidos (por ejemplo y sin limitación, aceites, agua, suero salino, glicerol y etanol), agentes humectantes o emulsionantes, sustancias tamponadoras del pH y similares.

Los polinucleótidos guía (ARN o ADN) y/o uno o más polinucleótidos de la endonucleasa (ARN o ADN) se pueden suministrar mediante vehículos de suministro víricos o no víricos conocidos en la técnica. Como alternativa, se pueden suministrar uno o más polipéptidos de endonucleasa mediante vehículos de administración víricos o no víricos conocidos en la técnica, tales como electroporación o nanopartículas de lípidos. En aspectos alternativos adicionales, la

endonucleasa de ADN se puede suministrar en forma de uno o más polipéptidos, ya sea sola o previamente complejada con uno o más ARN guía o uno o más ARNcr junto con un ARNtracr.

Los polinucleótidos se pueden suministrar mediante vehículos de administración no víricos que incluyen, pero sin limitación, nanopartículas, liposomas, ribonucleoproteínas, péptidos cargados positivamente, conjugados de ARN de molécula pequeña, quimeras de aptámero-ARN y complejos de proteína de fusión con ARN. Algunos vehículos de suministro no víricos de ejemplo se describen en Peer y Lieberman, *Gene Therapy*, 2011, 18: 1127–1133 (que se centra en los vehículos de suministro no víricos para ARNpi que también son útiles para el suministro de otros polinucleótidos).

Para los polinucleótidos de la divulgación, la formulación puede seleccionarse de cualquiera de los enseñados, por ejemplo, en la solicitud internacional PCT/US2012/069610.

Los polinucleótidos, tales como ARN guía, ARNsg y ARNm que codifican una endonucleasa pueden administrarse a una célula o un sujeto mediante una nanopartícula lipídica (LNP).

Una LNP se refiere a cualquier partícula que tenga un diámetro de menos de 1000 nm, 500 nm, 250 nm, 200 nm, 150 nm, 100 nm, 75 nm, 50 nm o 25 nm. Como alternativa, una nanopartícula puede tener un tamaño en el intervalo de 1-1000 nm, 1-500 nm, 1-250 nm, 25-200 nm, 25-100 nm, 35-75 nm o 25-60 nm.

Las LNP pueden estar formadas por lípidos catiónicos, aniónicos o neutros. Los lípidos neutros, tales como el fosfolípido fusogénico DOPE o el componente de membrana, colesterol, pueden incluirse en las LNP como "lípidos auxiliares" para potenciar la actividad de transfección y la estabilidad de las nanopartículas. Las limitaciones de los lípidos catiónicos incluyen la baja eficacia causada por la mala estabilidad o la rápida eliminación, así como la generación de respuestas inflamatorias o antiinflamatorias.

Las LNP también pueden estar formadas por lípidos hidrófobos, lípidos hidrófilos o lípidos tanto hidrófobos como hidrófilos.

Para producir una LNP se puede usar cualquier lípido o combinación de lípidos que se conozca en la técnica. Los ejemplos de lípidos usados para producir LNP son: DOTMA, DOSPA, DOTAP, DMRIE, DC-colesterol, DOTAP-colesterol, GAP-DMORIE-DPyPE y GL67A-DOPE-DMPE-polietilenglicol (PEG). Los ejemplos de lípidos catiónicos son: 98N12-5, C12-200, DLin-KC2-DMA (KC2), DLin-MC3-DMA (MC3), XTC, MD1 y 7C1. Los ejemplos de lípidos neutros son: DPSC, DPPC, POPC, DOPE y SM. Los ejemplos de lípidos modificados con PEG son: PEG-DMG, PEG-CerC14 y PEG-CerC20.

Los lípidos se pueden combinar en cualquier número de proporciones molares para producir una LNP. Además, el uno o más polinucleótidos pueden combinarse con uno o más lípidos en una gran variedad de relaciones molares para producir una LNP.

Para el suministro, se puede usar un vector de virus adenoasociado (AAV) recombinante. Son comunes en la técnica las técnicas para producir partículas de rAAV, en donde se proporcionan a una célula el genoma de AAV que se ha de empaquetar que incluye el polinucleótido que se ha de suministrar, los genes rep y cap y funciones de virus auxiliares. La producción de rAAV típicamente requiere que estén presentes los siguientes componentes en una sola célula (indicada en el presente documento como una célula empaquetadora): un genoma de rAAV, genes rep y cap de AAV separados de (es decir, no en) el genoma de rAAV y funciones de virus auxiliares. Los genes rep y cap de AAV pueden ser de cualquier serotipo de AAV para el que pueda obtenerse virus recombinante y puede ser de un serotipo de AAV diferente al de las ITR del genoma de rAAV, incluyendo, pero sin limitación, los serotipos de AAV descritos en el presente documento. La producción de rAAV pseudotipado se desvela en, por ejemplo, la publicación de solicitud de patente internacional número WO 01/83692.

*Formulación y administración de células, por ejemplo, células donantes universales*

Las células modificadas genéticamente, por ejemplo, células donantes universales, como se describe en el presente documento, pueden formularse y administrarse a un sujeto de cualquier manera conocida en la técnica.

Los términos "administrar", "introducir", "implantar", "injertar" y "trasplantar" se usan indistintamente en el contexto de la colocación de células, por ejemplo, células progenitoras, en un sujeto, mediante un método o una vía que da como resultado la ubicación al menos parcial de las células introducidas en un sitio deseado. Las células, por ejemplo, células progenitoras o su descendencia diferenciada pueden administrarse por cualquier vía adecuada que dé como resultado el suministro a una ubicación deseada en el sujeto, donde al menos una parte de las células implantadas o componentes de las células permanecen viables. El período de viabilidad de las células después de la administración a un sujeto puede ser tan corto como unas pocas horas, por ejemplo, veinticuatro horas, unos pocos días, hasta varios años, o incluso la vida del sujeto, es decir, injerto a largo plazo.

Una célula modificada genéticamente, por ejemplo, una célula donante universal, como se describe en el presente documento, puede ser viable después de la administración a un sujeto durante un período que es más largo que el de una célula no modificada.

En algunas realizaciones, una composición que comprende células como se describe en el presente documento puede administrarse mediante una vía adecuada, que puede incluir la administración intravenosa, por ejemplo, en embolada o mediante infusión continua durante un período de tiempo. En algunas realizaciones, la administración intravenosa puede realizarse por vías intramuscular, intraperitoneal, intracerebroespinal, subcutánea, intraarticular, intrasínovial o intratecal. En algunas realizaciones, una composición puede estar en forma sólida, en forma acuosa o en forma líquida. En algunas realizaciones, se puede nebulizar o

liofilizar una forma acuosa o líquida. En algunas realizaciones, una forma nebulizada o liofilizada puede reconstituirse con una solución acuosa o líquida.

También puede emulsionarse o presentarse una composición celular como una composición de liposomas, a condición de que el procedimiento de emulsión no afecte negativamente a la viabilidad celular. Pueden mezclarse las células y cualquier otro principio activo con excipientes que son farmacéuticamente aceptables y compatibles con el principio activo y en cantidades adecuadas para su uso en los métodos terapéuticos descritos en el presente documento.

Los agentes adicionales incluidos en una composición celular pueden incluir sales farmacéuticamente aceptables de los componentes de la misma. Las sales farmacéuticamente aceptables incluyen sales de adición de ácidos (formadas con los grupos amino libres del polipéptido) que se forman con ácidos inorgánicos, tales como, por ejemplo, los ácidos clorhídrico o fosfórico o ácidos orgánicos tales como acético, tartárico, mandélico y similares. Las sales formadas con los grupos carboxilo libres también pueden derivar de bases orgánicas, tales como hidróxidos de sodio, potasio, amonio, calcio o férricos y bases orgánicas, tales como isopropilamina, trimetilamina, 2-etilamino etanol, histidina, procaína y similares.

Se conocen bien en la técnica portadores fisiológicamente tolerables. Los portadores líquidos de ejemplo son soluciones acuosas estériles que no contienen materiales además de los principios activos y agua o que contienen un tampón, tal como fosfato de sodio a un valor de pH fisiológico, solución salina fisiológica o ambos, tal como solución salina tamponada con fosfato. Además, los portadores acuosos pueden contener más de una sal tampón, así como sales tales como cloruros de sodio y potasio, dextrosa, polietilenglicol y otros solutos. Las composiciones líquidas también pueden contener fases líquidas además de o excluyendo el agua. Son ejemplos de dichas fases líquidas adicionales la glicerina, aceites vegetales, tales como aceite de semilla de algodón y emulsiones de aceite-agua. La cantidad de un compuesto activo usada en las composiciones celulares que es eficaz en el tratamiento de un trastorno o afección particular dependerá de la naturaleza del trastorno o de la afección, y puede determinarse mediante técnicas clínicas convencionales.

En algunas realizaciones, se puede administrar una composición que comprenda células a un sujeto, por ejemplo, un sujeto humano, que tenga, se sospeche que tenga o esté en riesgo de padecer una enfermedad. En algunas realizaciones, se puede administrar una composición a un sujeto que no tenga, que no se sospeche que tenga o que no esté en riesgo de padecer una enfermedad. En algunas realizaciones, un sujeto es un ser humano sano. En algunas realizaciones, un sujeto, por ejemplo, un sujeto humano, que tiene, se sospecha que tiene o está en riesgo de tener una enfermedad hereditaria genéticamente. En algunas realizaciones, el sujeto padece o corre el riesgo de desarrollar síntomas indicativos de una

enfermedad. En algunas realizaciones, la enfermedad es diabetes, por ejemplo, diabetes de tipo I o diabetes de tipo II.

#### **VI. Composiciones y métodos específicos de la divulgación**

Por consiguiente, la presente divulgación se refiere en particular a las siguientes composiciones y métodos no limitantes.

En una primera composición, la Composición 1, la presente divulgación proporciona una composición que comprende una célula donante universal que comprende una secuencia de nucleótidos que codifica un primer factor tolerogénico insertado dentro o cerca de un gen que codifica un factor de supervivencia, en donde la célula donante universal expresa el factor tolerogénico y tiene la expresión interrumpida del factor de supervivencia, y la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula de control.

En otra composición, la Composición 2, la presente divulgación proporciona una composición, como se proporciona en la Composición 1, en donde la célula de control es una célula de tipo silvestre o una célula que no comprende la secuencia de nucleótidos insertada.

En otra composición, la Composición 3, la presente divulgación proporciona una composición, como se proporciona en las Composiciones 1 o 2, en donde la expresión interrumpida del factor de supervivencia comprende una expresión reducida o eliminada.

En otra composición, la Composición 4, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 1 a 3, en donde el primer factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47.

En otra composición, la Composición 5, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 1 a 4, en donde el factor de supervivencia es TXNIP, ZNF143, FOXO1, JNK o MANF.

En otra composición, la Composición 6, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones de las reivindicaciones 1 a 5, en donde el primer factor tolerogénico es HLA-E y el factor de supervivencia es TXNIP.

En otra composición, la Composición 7, la presente divulgación proporciona una composición, como se proporciona en las Composiciones 5 o 6, en donde la secuencia de nucleótidos que codifica HLA-E comprende una secuencia que codifica un trímero de HLA-E que comprende un péptido señal de B2M fusionado a un péptido de presentación de HLA-G fusionado a una proteína de membrana B2M fusionada a HLA-E sin su péptido señal.

En otra composición, la Composición 8, la presente divulgación proporciona una composición, como se proporciona en la Composición 7, en donde la secuencia que codifica el trímero de HLA-E consiste esencialmente en la SEQ ID NO: 55.

En otra composición, la Composición 9, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 1 a 8, en donde la secuencia de nucleótidos que codifica el primer factor tolerogénico está unida operativamente a un promotor exógeno,

5 En otra composición, la Composición 10, la presente divulgación proporciona una composición, como se proporciona en la Composición 9, en donde el promotor exógeno es un promotor de CMV, EF1 $\alpha$ , PGK, CAG o UBC.

10 En otra composición, la Composición 11, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones de las reivindicaciones 1 a 10, que comprende además una secuencia de nucleótidos que codifica un segundo factor tolerogénico insertado dentro o cerca de un gen que codifica un antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un complejo MHC-I o MHC-II, en donde la célula donante universal expresa el factor tolerogénico y tiene la expresión interrumpida del antígeno leucocitario humano MHC-I o MHC-II o el  
15 componente o el regulador de la transcripción del complejo MHC-I o MHC-II.

En otra composición, la Composición 12, la presente divulgación proporciona una composición, como se proporciona en la Composición 11, en donde la expresión interrumpida del antígeno leucocitario humano MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo MHC-I o MHC-II comprende una expresión reducida o eliminada.

20 En otra composición, la Composición 13, la presente divulgación proporciona una composición, como se proporciona en las Composiciones 11 o 12, en donde el segundo factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47.

En otra composición, la Composición 14, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 11 a 13, en donde  
25 el antígeno leucocitario humano MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo MHC-I o MHC-II es HLA-A, HLA-B, HLA-C, HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ, HLA-DR, B2M, NLRC5, CIITA, RFX5, RFXAP o RFXANK

En otra composición, la Composición 15, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 11 a 14, en donde  
30 el segundo factor tolerogénico es PD-L1 y el antígeno leucocitario humano MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo MHC-I o MHC-II es B2M.

En otra composición, la composición 16, la presente divulgación proporciona una composición, como se proporciona en la Composición 15, en donde la secuencia de nucleótidos que codifica PD-L1 consiste esencialmente en la SEQ ID NO: 11.

35 En otra composición, la Composición 17, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 11 a 16, en donde

la secuencia de nucleótidos que codifica el segundo factor tolerogénico está unida operativamente a un promotor exógeno,

5 En otra composición, la Composición 18, la presente divulgación proporciona una composición, como se proporciona en la Composición 17, en donde el promotor exógeno es un promotor de CMV, EF1 $\alpha$ , PGK, CAG o UBC.

10 En otra composición, la Composición 19, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 11 a 18, en donde el primer factor tolerogénico es HLA-E, el factor de supervivencia es TXNIP, el segundo factor tolerogénico es PD-L1 y el antígeno leucocitario humano MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo MHC-I o MHC-II es B2M.

En otra composición, la Composición 20, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 1 a 19, en donde la célula es una célula madre.

15 En otra composición, la Composición 21, la presente divulgación proporciona una composición, como se proporciona en la Composición 20, en donde la célula madre es una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

20 En otra composición, la Composición 22, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 1 a 19, en donde la célula es una célula diferenciada o una célula somática.

En otra composición, la Composición 23, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 1 a 19, en donde la célula es capaz de diferenciarse en células progenitoras de linaje restringido o células somáticas completamente diferenciadas.

25 En otra composición, la Composición 24, la presente divulgación proporciona una composición, como se proporciona en la Composición 23, en donde las células progenitoras de linaje restringido son células progenitoras del endodermo pancreático, células progenitoras endocrinas pancreáticas, células progenitoras mesenquimatosas, células progenitoras musculares, blastocitos, células progenitoras hematopoyéticas o células progenitoras neuronales.

30 En otra composición, la Composición 25, la presente divulgación proporciona una composición, como se proporciona en la Composición 23, en donde las células somáticas completamente diferenciadas son células beta pancreáticas, células epiteliales, células endodérmicas, macrófagos, hepatocitos, adipocitos, células renales, células sanguíneas, cardiomiocitos o células del sistema inmunitario.

35

En otra composición, la Composición 26, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 1 a 25, en donde la composición comprende una pluralidad de células donantes universales.

5 En otra composición, la Composición 27, la presente divulgación proporciona una composición, como se proporciona en la Composición 26, en donde la composición comprende una población de células progenitoras de linaje restringido o células somáticas completamente diferenciadas derivadas de la pluralidad de células donantes universales.

10 En otra composición, la Composición 28, la presente divulgación proporciona una composición, como se proporciona en la Composición 27, en donde las células progenitoras de linaje restringido son células progenitoras del endodermo pancreático, células progenitoras endocrinas pancreáticas, células progenitoras mesenquimatosas, células progenitoras musculares, blastocitos, células progenitoras hematopoyéticas o células progenitoras neuronales, y las células somáticas completamente diferenciadas son células beta pancreáticas, células epiteliales, células endodérmicas, macrófagos, hepatocitos, adipocitos, 15 células renales, células sanguíneas, cardiomiocitos o células del sistema inmunitario.

En otra composición, la Composición 29, la presente divulgación proporciona una composición, como se proporciona en las Composiciones 6 o 19, en donde la composición comprende una pluralidad de células donantes universales.

20 En otra composición, la Composición 30, la presente divulgación proporciona una composición, como se proporciona en la Composición 29, en donde la composición comprende una población de células progenitoras de linaje restringido o células somáticas completamente diferenciadas derivadas de la pluralidad de células donantes universales.

25 En otra composición, la Composición 31, la presente divulgación proporciona una composición, como se proporciona en la Composición 30, en donde las células progenitoras de linaje restringido son células del endodermo definitivo, células del tubo intestinal primitivo, células del intestino anterior posterior, progenitores del endodermo pancreático, progenitores endocrinos pancreáticos, células beta inmaduras o células beta en maduración, y las células somáticas completamente diferenciadas son células beta pancreáticas.

30 En otra composición, la Composición 32, la presente divulgación proporciona una composición, como se proporciona en la Composición 26 o 29, en donde al menos aproximadamente el 50 %, al menos aproximadamente el 70 % o al menos aproximadamente el 90 % de las células expresan el primer factor tolerogénico, el segundo factor tolerogénico o el primer y el segundo factores tolerogénicos.

35 En otra composición, la Composición 33, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 27, 28, 30 o 31, en donde al menos aproximadamente el 50 %, al menos aproximadamente el 70 % o al menos

aproximadamente el 90 % de las células expresan el primer factor tolerogénico, el segundo factor tolerogénico o el primer y el segundo factores tolerogénicos.

5 En otra composición, la Composición 34, la presente divulgación proporciona una composición que comprende la pluralidad de células de la Composición 26 o la población de células de las Composiciones 27 o 28.

En otra composición, la Composición 35, la presente divulgación proporciona una composición como se proporciona en la Composición 34 para su uso en el tratamiento de un sujeto que lo necesite.

10 En otra composición, la Composición 36, la presente divulgación proporciona una composición, como se proporciona en la Composición 35, en donde el sujeto tiene, se sospecha que tiene o está en riesgo de tener una enfermedad.

En otra composición, la Composición 37, la presente divulgación proporciona una composición, como se proporciona en la Composición 36, en donde la enfermedad es una enfermedad hereditaria genéticamente.

15 En otra composición, la Composición 38, la presente divulgación proporciona una composición que comprende la pluralidad de células de la Composición 29 o la población de células de las Composiciones 30 o 31.

20 En otra composición, la Composición 39, la presente divulgación proporciona una composición, como se proporciona en la Composición 38 para tratar la diabetes en un sujeto que lo necesite.

En otra composición, la Composición 40, la presente divulgación proporciona una composición, como se proporciona en la Composición 39, en donde el sujeto tiene diabetes de tipo I o diabetes de tipo II.

25 En otra composición, la Composición 41, la presente divulgación proporciona una composición, como se proporciona en una cualquiera de las Composiciones 35 a 40, en donde el sujeto es un ser humano.

30 En un primer método, el Método 1, la presente divulgación proporciona un método de obtención de células para su administración a un sujeto que lo necesite, comprendiendo el método: (a) obtener o haber obtenido la pluralidad de células donantes universales de una cualquiera de las Composiciones 26, 29 o 32, y (b) mantener la pluralidad de células donantes universales durante un tiempo y en condiciones suficientes para que las células se diferencien en células progenitoras de linaje restringido o células somáticas completamente diferenciadas.

35 En otro método, el Método 2, la presente divulgación proporciona un método para tratar a un sujeto que lo necesite, comprendiendo el método: (a) obtener o haber obtenido la pluralidad de células donantes universales de una cualquiera de las Composiciones 26, 29 o 32 después de la diferenciación en células progenitoras de linaje restringido o células somáticas

completamente diferenciadas; y (b) administrar al sujeto las células progenitoras de linaje restringido o las células somáticas completamente diferenciadas.

5 En otro método, el Método 3, la presente divulgación proporciona un método como se proporciona en el Método 2, en donde la administración comprende implantar en el sujeto un dispositivo que comprende las células progenitoras de linaje restringido o las células somáticas completamente diferenciadas.

10 En otro método, el Método 4, la presente divulgación proporciona un método como se proporciona en uno cualquiera de los Métodos 1 a 3, en donde las células progenitoras de linaje restringido son células progenitoras del endodermo pancreático, células progenitoras endocrinas pancreáticas, células progenitoras mesenquimatosas, células progenitoras musculares, blastocitos, células progenitoras hematopoyéticas o células progenitoras neuronales, y las células somáticas completamente diferenciadas son células beta pancreáticas, células epiteliales, células endodérmicas, macrófagos, hepatocitos, adipocitos, células renales, células sanguíneas, cardiomiocitos o células del sistema inmunitario.

15 En otro método, el Método 5, la presente divulgación proporciona un método como se proporciona en uno cualquiera de los Métodos 1 a 4, en donde el sujeto tiene, se sospecha que tiene o está en riesgo de tener un cáncer.

20 En otro método, el Método 6, la presente divulgación proporciona un método como se proporciona en el Método 5, en donde la enfermedad es una enfermedad hereditaria genéticamente.

En otro método, el Método 7, la presente divulgación proporciona un método como se proporciona en uno cualquiera de los Métodos 1 a 6, en donde el sujeto es un ser humano.

25 En otro método, el Método 8, la presente divulgación proporciona un método para tratar la diabetes en un sujeto que lo necesite, comprendiendo el método: (a) obtener o haber obtenido la pluralidad de células donantes universales de la Composición 29 o 32 después de la diferenciación en células del endodermo pancreático, células endocrinas pancreáticas, células beta inmaduras, células beta en maduración o células beta pancreáticas; y (b) administrar al sujeto las células del endodermo pancreático, las células endocrinas pancreáticas, las células beta inmaduras, las células beta en maduración o las células beta pancreáticas.

30 En otro método, el Método 9, la presente divulgación proporciona un método como se proporciona en el Método 8, en donde la administración comprende implantar en el sujeto un dispositivo que comprende las células del endodermo pancreático, las células endocrinas pancreáticas, las células beta inmaduras, las células beta en maduración o las células beta pancreáticas.

35 En otro método, el Método 10, la presente divulgación proporciona un método como se proporciona en el Método 8 o 9, en donde el sujeto tiene diabetes de tipo I o diabetes de tipo II.

En otro método, el Método 11, la presente divulgación proporciona un método como se proporciona en uno cualquiera de los Métodos 8 a 10, en donde el sujeto es un ser humano.

En otra composición, la Composición 41, la presente divulgación proporciona una composición que comprende una célula donante universal que comprende una secuencia de nucleótidos que codifica el antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E) insertada dentro o cerca de un gen que codifica la proteína que interactúa con la tiorredoxina (TXNIP), en donde la célula donante universal expresa HLA-E y tiene la expresión interrumpida de TXNIP, y la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con un control.

En otra composición, la Composición 42, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde la célula de control es una célula de tipo silvestre o una célula que no comprende la secuencia de nucleótidos insertada.

En otra composición, la Composición 43, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde la expresión interrumpida de TXNIP comprende una expresión reducida o eliminada.

En otra composición, la Composición 44, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde la secuencia de nucleótidos que codifica HLA-E comprende una secuencia que codifica un trímero de HLA-E que comprende un péptido señal de B2M fusionado a un péptido de presentación de HLA-G fusionado a una proteína de membrana B2M fusionada a HLA-E sin su péptido señal.

En otra composición, la Composición 45, la presente divulgación proporciona una composición, como se proporciona en la Composición 44, en donde la secuencia que codifica el trímero de HLA-E consiste esencialmente en la SEQ ID NO: 55.

En otra composición, la Composición 46, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde la secuencia de nucleótidos que codifica HLA-E está unida operativamente a un promotor exógeno,

En otra composición, la Composición 47, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde el promotor exógeno es un promotor CAG.

En otra composición, la Composición 48, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde la célula es una célula madre.

En otra composición, la Composición 49, la presente divulgación proporciona una composición, como se proporciona en la Composición 48, en donde la célula madre es una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

En otra composición, la Composición 50, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde la célula es una célula diferenciada o una célula somática.

5 En otra composición, la Composición 51, la presente divulgación proporciona una composición, como se proporciona en la Composición 41, en donde la célula es capaz de diferenciarse en células progenitoras de linaje restringido o células somáticas completamente diferenciadas.

10 En otra composición, la Composición 52, la presente divulgación proporciona una composición, como se proporciona en la Composición 51, en donde las células progenitoras de linaje restringido son células del endodermo definitivo, células del tubo intestinal primitivo, células del intestino anterior posterior, progenitores del endodermo pancreático, progenitores endocrinos pancreáticos, células beta inmaduras o células beta en maduración, y las células somáticas completamente diferenciadas son células beta pancreáticas.

15 En otra composición, la Composición 53, la presente divulgación proporciona una composición que comprende una pluralidad de células donantes universales como se proporciona en la Composición 41.

En otra composición, la Composición 54, la presente divulgación proporciona una composición, como se proporciona en la Composición 53, en donde al menos aproximadamente el 50 % de las células expresan HLA-E.

20 En otra composición, la Composición 55, la presente divulgación proporciona una composición, como se proporciona en la Composición 53, en donde al menos aproximadamente el 70 % de las células expresan HLA-E.

25 En otra composición, la Composición 56, la presente divulgación proporciona una composición, como se proporciona en la Composición 53, en donde al menos aproximadamente el 90 % de las células expresan HLA-E.

En otra composición, la Composición 57, la presente divulgación proporciona una composición que comprende una población de células progenitoras de linaje restringido o células somáticas completamente diferenciadas derivadas de la pluralidad de células donantes universales de la Composición 53.

30 En otra composición, la Composición 58, la presente divulgación proporciona una composición, como se proporciona en la Composición 57, en donde las células progenitoras de linaje restringido son células del endodermo definitivo, células del tubo intestinal primitivo, células del intestino anterior posterior, progenitores del endodermo pancreático, progenitores endocrinos pancreáticos, células beta inmaduras o células beta en maduración, y las células somáticas completamente diferenciadas son células beta pancreáticas.

35

En otra composición, la Composición 59, la presente divulgación proporciona una composición, como se proporciona en la Composición 58, en donde al menos aproximadamente el 50 % de las células expresan HLA-E.

5 En otra composición, la Composición 60, la presente divulgación proporciona una composición, como se proporciona en la Composición 59, en donde al menos aproximadamente el 70 % de las células expresan HLA-E.

En otra composición, la Composición 61, la presente divulgación proporciona una composición, como se proporciona en la Composición 59, en donde al menos aproximadamente el 90 % de las células expresan HLA-E.

10 En otra composición, la Composición 62, la presente divulgación proporciona una composición que comprende una célula modificada genéticamente que tiene la expresión introducida o aumentada del antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E) y tiene la expresión interrumpida de la proteína que interactúa con la tiorredoxina (TXNIP), en donde la célula modificada genéticamente tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula no modificada.

15 En otra composición, la Composición 63, la presente divulgación proporciona una composición, como se proporciona en la Composición 62, que comprende una secuencia de nucleótidos que codifica HLA-E insertada dentro o cerca de un gen que codifica TXNIP, interrumpiendo de este modo el gen TXNIP.

20 En otra composición, la Composición 64, la presente divulgación proporciona una composición, como se proporciona en la Composición 62, en donde la expresión interrumpida de TXNIP comprende una expresión reducida o eliminada

25 En otro método, el Método 12, la presente divulgación proporciona un método para tratar la diabetes en un sujeto que lo necesite, comprendiendo el método: obtener o haber obtenido la pluralidad de células donantes universales de la Composición 53 después de la diferenciación en células del endodermo pancreático, células endocrinas pancreáticas, células beta inmaduras, células beta en maduración o células beta pancreáticas; y (b) administrar al sujeto las células del endodermo pancreático, las células endocrinas pancreáticas, las células beta inmaduras, las células beta en maduración o las células beta pancreáticas.

30 En otro método, el Método 13, la presente divulgación proporciona un método como se proporciona en el Método 12, en donde la administración comprende implantar en el sujeto un dispositivo que comprende las células del endodermo pancreático, las células endocrinas pancreáticas, las células beta inmaduras, las células beta en maduración o las células beta pancreáticas.

35 En otro método, el Método 14, la presente divulgación proporciona un método, como se proporciona en el Método 12, en donde el sujeto tiene diabetes de tipo I o diabetes de tipo II.

En otro método, el Método 15, la presente divulgación proporciona un método, como se proporciona en el Método 12, en donde el sujeto es un ser humano.

En otra composición, la Composición 65, la presente divulgación proporciona una composición que comprende una célula donante universal que comprende (a) una secuencia de nucleótidos que codifica el ligando de muerte programada 1 (PD-L1) insertada dentro o cerca de un gen que codifica la microglobulina beta-2 (B2M) y (b) una secuencia de nucleótidos que codifica el antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E) insertada dentro o cerca de un gen que codifica la proteína que interactúa con la tiorredoxina (TXNIP), en donde la célula donante universal expresa PD-L1 y HLA-E y tiene la expresión interrumpida de B2M y TXNIP, y la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula de control.

En otra composición, la Composición 66, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la célula de control es una célula de tipo silvestre o una célula que no comprende la secuencia de nucleótidos insertada.

En otra composición, la Composición 67, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la expresión interrumpida de B2M comprende la expresión reducida o eliminada de B2M y la expresión interrumpida de TXNIP comprende la expresión reducida o eliminada de TXNIP.

En otra composición, la composición 68, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la secuencia de nucleótidos que codifica PD-L1 consiste esencialmente en la SEQ ID NO: 11.

En otra composición, la Composición 69, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la secuencia de nucleótidos que codifica HLA-E comprende una secuencia que codifica un trímero de HLA-E que comprende un péptido señal de B2M fusionado a un péptido de presentación de HLA-G fusionado a una proteína de membrana B2M fusionada a HLA-E sin su péptido señal.

En otra composición, la Composición 70, la presente divulgación proporciona una composición, como se proporciona en la Composición 69, en donde la secuencia que codifica el trímero de HLA-E consiste esencialmente en la SEQ ID NO: 55.

En otra composición, la Composición 71, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la secuencia de nucleótidos que codifica PD-L1 está unida operativamente a un promotor exógeno y la secuencia de nucleótidos que codifica HLA-E está unida operativamente a un promotor exógeno.

En otra composición, la Composición 72, la presente divulgación proporciona una composición, como se proporciona en la Composición 71, en donde el promotor exógeno es un promotor CAG.

En otra composición, la Composición 73, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la célula es una célula madre.

5 En otra composición, la Composición 74, la presente divulgación proporciona una composición, como se proporciona en la Composición 73, en donde la célula madre es una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

10 En otra composición, la Composición 75, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la célula es una célula diferenciada o una célula somática.

En otra composición, la Composición 76, la presente divulgación proporciona una composición, como se proporciona en la Composición 65, en donde la célula es capaz de diferenciarse en células progenitoras de linaje restringido o células somáticas completamente diferenciadas.

15 En otra composición, la Composición 77, la presente divulgación proporciona una composición, como se proporciona en la Composición 76, en donde las células progenitoras de linaje restringido son células del endodermo definitivo, células del tubo intestinal primitivo, células del intestino anterior posterior, progenitores del endodermo pancreático, progenitores endocrinos pancreáticos, células beta inmaduras o células beta en maduración, y las células somáticas completamente diferenciadas son células beta pancreáticas.

20 En otra composición, la Composición 78, la presente divulgación proporciona una composición que comprende una pluralidad de células donantes universales como se proporciona en la Composición 65.

25 En otra composición, la Composición 79, la presente divulgación proporciona una composición, como se proporciona en la Composición 78, en donde al menos aproximadamente el 50 % de las células expresan PD-L1 y/o al menos aproximadamente el 50 % de las células expresan HLA-E.

30 En otra composición, la Composición 80, la presente divulgación proporciona una composición, como se proporciona en la Composición 78, en donde al menos aproximadamente el 70 % de las células expresan PD-L1 y/o al menos aproximadamente el 70 % de las células expresan HLA-E.

35 En otra composición, la Composición 81, la presente divulgación proporciona una composición, como se proporciona en la Composición 78, en donde al menos aproximadamente el 90 % de las células expresan PD-L1 y/o al menos aproximadamente el 90 % de las células expresan HLA-E.

En otra composición, la Composición 82, la presente divulgación proporciona una composición que comprende una población de células progenitoras de linaje restringido o células somáticas completamente diferenciadas derivadas de la pluralidad de células donantes universales, como se proporciona en la Composición 78.

5 En otra composición, la Composición 83, la presente divulgación proporciona una composición, como se proporciona en la Composición 82, en donde las células progenitoras de linaje restringido son células del endodermo definitivo, células del tubo intestinal primitivo, células del intestino anterior posterior, progenitores del endodermo pancreático, progenitores endocrinos pancreáticos, células beta inmaduras o células beta en maduración, y las células somáticas completamente diferenciadas son células beta pancreáticas.

10 En otra composición, la Composición 84, la presente divulgación proporciona una composición, como se proporciona en la Composición 83, en donde al menos aproximadamente el 50 % de las células expresan PD-L1 y/o al menos aproximadamente el 50 % de las células expresan HLA-E.

15 En otra composición, la Composición 85, la presente divulgación proporciona una composición, como se proporciona en la Composición 83, en donde al menos aproximadamente el 70 % de las células expresan PD-L1 y/o al menos aproximadamente el 70 % de las células expresan HLA-E.

20 En otra composición, la Composición 86, la presente divulgación proporciona una composición, como se proporciona en la Composición 83, en donde al menos aproximadamente el 90 % de las células expresan PD-L1 y/o al menos aproximadamente el 90 % de las células expresan HLA-E.

25 En otra composición, la Composición 87, la presente divulgación proporciona una composición que comprende una célula modificada genéticamente que tiene la expresión introducida o aumentada de PD-L1 y HLA, y la expresión interrumpida de B2M y TXNIP, en donde la célula modificada genéticamente tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula no modificada.

30 En otra composición, la Composición 88, la presente divulgación proporciona una composición, como se proporciona en la Composición 87, que comprende una secuencia de nucleótidos que codifica PD-L1 insertada dentro o cerca de un gen que codifica B2M, interrumpiendo de este modo el gen B2M, y una secuencia de nucleótidos que codifica HLA-E insertada dentro o cerca de un gen que codifica TXNIP, interrumpiendo de este modo el gen TXNIP.

35 En otra composición, la Composición 89, la presente divulgación proporciona una composición, como se proporciona en la Composición 87, en donde la expresión interrumpida de B2M y TXNIP comprende la expresión reducida o eliminada de B2M y TXNIP.

En otro método, el Método 16, la presente divulgación proporciona un método para tratar la diabetes en un sujeto que lo necesite, comprendiendo el método: (a) obtener o haber obtenido la pluralidad de células donantes universales de la Composición 78 después de la diferenciación en células del endodermo pancreático, células endocrinas pancreáticas, células beta inmaduras, células beta en maduración o células beta pancreáticas; y (b) administrar al sujeto las células del endodermo pancreático, las células endocrinas pancreáticas, las células beta inmaduras, las células beta en maduración o las células beta pancreáticas.

En otro método, el Método 17, la presente divulgación proporciona un método como se proporciona en el Método 16, en donde la administración comprende implantar en el sujeto un dispositivo que comprende las células del endodermo pancreático, las células endocrinas pancreáticas, las células beta inmaduras, las células beta en maduración o las células beta pancreáticas.

En otro método, el Método 18, la presente divulgación proporciona un método, como se proporciona en el Método 16, en donde el sujeto tiene diabetes de tipo I o diabetes de tipo II.

En otro método, el Método 19, la presente divulgación proporciona un método, como se proporciona en el Método 16, en donde el sujeto es un ser humano.

En otro método, el Método 20, la presente divulgación proporciona un método para generar una célula donante universal, comprendiendo el método suministrar a una célula: (a) una primera nucleasa dirigida al sitio que se dirige a un sitio dentro o cerca de un gen que codifica un factor de supervivencia; y (b) un primer ácido nucleico que comprende una secuencia de nucleótidos que codifica un primer factor tolerogénico que está flanqueado por (i) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana de (a) y (ii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana de (a), en donde la primera nucleasa dirigida al sitio escinde el sitio diana de (a) y el primer ácido nucleico de (b) se inserta en un sitio que se solapa parcialmente con, se solapa completamente con o está contenido dentro del sitio de (a), generando de este modo una célula donante universal, en donde la célula donante universal tiene una mayor supervivencia celular en comparación con una célula en la que el ácido nucleico de (b) no se ha insertado.

En otro método, el Método 21, la presente divulgación proporciona un método, como se proporciona en el Método 20, en donde el factor de supervivencia es TXNIP, ZNF143, FOXO1, JNK o MANF.

En otro método, el Método 22, la presente divulgación proporciona un método, como se proporciona en los Métodos 20 o 21, en donde el primer factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47.

En otro método, el Método 23, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 22, en donde el factor de supervivencia es TXNIP.

5 En otro método, el Método 24, la presente divulgación proporciona un método, como se proporciona en el Método 23, en donde el primer factor tolerogénico es HLA-E.

En otro método, el Método 25, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 24, en donde la primera nucleasa dirigida al sitio es un sistema CRISPR que comprende una nucleasa CRISPR y un ARN guía (ARNg).

10 En otro método, el Método 26, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 25, en donde la nucleasa CRISPR es una nucleasa Cas9 de Tipo II o una nucleasa Cfp1 de Tipo V, y la nucleasa CRISPR está unida a al menos una señal de localización nuclear.

15 En otro método, el Método 27, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 26, en donde el ARNg comprende una secuencia espaciadora correspondiente a una secuencia diana que consiste en las SEQ ID NO: 15-24.

20 En otro método, el Método 28, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 25 a 27, en donde la secuencia de nucleótidos de (b)(i) consiste esencialmente en la SEQ ID NO: 25, y la secuencia de nucleótidos de (b)(iii) consiste esencialmente en la SEQ ID NO: 32.

25 En otro método, el Método 29, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 28, en donde el método comprende además suministrar a la célula: (c) una segunda nucleasa dirigida al sitio que se dirige a un sitio dentro o cerca de un gen que codifica uno o más de los antígenos leucocitarios humanos MHC-I o MHC-II o un componente o un regulador de la transcripción de un complejo MHC-I o MHC-II; y (d) un segundo ácido nucleico que comprende una secuencia de nucleótidos que codifica un segundo factor tolerogénico que está flanqueado por (iii) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana de (c) y una (iv) secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana de (c), en donde el  
30 segundo factor tolerogénico de (d) difiere del primer factor tolerogénico (b), en donde la segunda nucleasa dirigida al sitio escinde el sitio diana de (c) y el segundo ácido nucleico de (d) se inserta en un sitio que se solapa parcialmente con, se solapa completamente con o está contenido dentro del sitio de (c), en donde la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula en la que el segundo ácido nucleico  
35 de (d) no se ha insertado.

En otro método, el Método 30, la presente divulgación proporciona un método, como se proporciona en el Método 29, en donde el antígeno leucocitario humano MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo MHC-I o MHC-II es HLA-A, HLA-B, HLA-C, HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ, HLA-DR, B2M, NLRC5, CIITA, RFX5, RFXAP o RFXANK.

En otro método, el Método 31, la presente divulgación proporciona un método, como se proporciona en los Métodos 29 o 30, en donde el segundo factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47.

En otro método, el Método 32, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 29 a 31, en donde el antígeno leucocitario humano MHC-I o MHC-II o el componente o el regulador de la transcripción del complejo MHC-I o MHC-II es B2M.

En otro método, el Método 33, la presente divulgación proporciona un método, como se proporciona en el Método 32, en donde el segundo factor tolerogénico es PD-L1.

En otro método, el Método 34, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 29 a 33, en donde la segunda nucleasa dirigida al sitio es un sistema CRISPR que comprende una nucleasa CRISPR y un ARNg.

En otro método, el Método 35, la presente divulgación proporciona un método, como se proporciona en el Método 34, en donde la nucleasa CRISPR es una nucleasa Cas9 de Tipo II o una nucleasa Cfp1 de Tipo V, y la nucleasa CRISPR está unida a al menos una señal de localización nuclear.

En otro método, el Método 36, la presente divulgación proporciona un método, como se proporciona en los Métodos 34 o 35, en donde el ARNg comprende una secuencia espaciadora correspondiente a una secuencia diana que consiste en las SEQ ID NO: 1-3 o 35-44.

En otro método, el Método 37, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 34 a 36, en donde la secuencia de nucleótidos de (d)(iii) consiste esencialmente en la SEQ ID NO: 7, y la secuencia de nucleótidos de (d)(iv) consiste esencialmente en la SEQ ID NO: 13.

En otro método, el Método 38, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 25 a 28 o 34 a 37, en donde la nucleasa CRISPR y el ARNg están presentes en una relación molar de 1:3.

En otro método, el Método 39, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 38, en donde la secuencia de nucleótidos que codifica el primer factor tolerogénico está unida operativamente a un promotor exógeno y la secuencia de nucleótidos que codifica el segundo factor tolerogénico está unida operativamente a un promotor exógeno.

En otro método, el Método 40, la presente divulgación proporciona un método, como se proporciona en el Método 39, en donde el promotor exógeno es un promotor constitutivo, inducible, temporal, específico del tejido o específico del tipo de célula, opcionalmente en donde el promotor exógeno es un promotor de CMV, EF1 $\alpha$ , PGK, CAG o UBC.

5        En otro método, el Método 41, la presente divulgación proporciona un método para generar una célula donante universal, comprendiendo el método suministrar a una célula: (a) una primera nucleasa dirigida al sitio que se dirige a un sitio dentro o cerca de un gen que codifica un factor de supervivencia; (b) un primer ácido nucleico que comprende una secuencia de nucleótidos que codifica un primer factor tolerogénico que está flanqueado por (i) una  
10        secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana de (a) y (ii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana de (a), en donde la primera nucleasa dirigida al sitio escinde el sitio diana de (a) y, a través de un proceso de recombinación homóloga, el primer ácido nucleico de (b) se usa utiliza como molde para insertar la secuencia de nucleótidos que codifica el primer factor tolerogénico en un  
15        sitio que se solapa parcialmente con, se solapa completamente con o está contenido dentro del sitio de (a), interrumpiendo de este modo el gen de (a); (c) una segunda nucleasa dirigida al sitio dentro o cerca de un gen que codifica uno o más de un antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un complejo MHC-I o MHC-II; y (d) un segundo ácido nucleico que comprende una secuencia de nucleótidos que codifica  
20        un segundo factor tolerogénico que está flanqueado por (iii) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana de (c) y una (iv) secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana de (c), en donde el factor tolerogénico de (d) difiere del factor tolerogénico (b), en donde la segunda nucleasa dirigida al sitio escinde el sitio diana de (c) y, a través de un proceso de recombinación  
25        homóloga, el segundo ácido nucleico de (d) se utiliza como molde para insertar la secuencia de nucleótidos que codifica el segundo factor tolerogénico en un sitio que se solapa parcialmente con, se solapa completamente con o está contenido dentro del sitio de (c), interrumpiendo de este modo el gen de (c), generando de este modo una célula donante universal, en donde la célula donante universal tiene una mayor supervivencia celular en comparación con una célula en la que el primer ácido nucleico de (b) y el segundo ácido nucleico de (d) no se han insertado.  
30

En otro método, el Método 42, la presente divulgación proporciona un método, como se proporciona en el Método 41, en donde el factor de supervivencia es TXNIP, el primer factor tolerogénico es HLA-E, el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un complejo MHC-I o MHC-II es B2M, y el segundo factor tolerogénico es PD-L1.  
35

En otro método, el Método 43, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 42, en donde la célula es una célula de mamífero, opcionalmente en donde la célula es una célula humana.

5 En otro método, el Método 44, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 43, en donde la célula es una célula madre.

En otro método, el Método 45, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 43, en donde la célula es una célula madre pluripotente, una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

10 En otro método, el Método 46, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 43, en donde la célula es una célula diferenciada o una célula somática.

En otro método, el Método 47, la presente divulgación proporciona un método, como se proporciona en uno cualquiera de los Métodos 20 a 43, en donde la célula donante universal es capaz de diferenciarse en células progenitoras de linaje restringido o células somáticas completamente diferenciadas.

15 En otro método, el Método 48, la presente divulgación proporciona un método, como se proporciona en el Método 47, en donde las células progenitoras de linaje restringido son células progenitoras del endodermo pancreático, células progenitoras endocrinas pancreáticas, células progenitoras mesenquimatosas, células progenitoras musculares, blastocitos, células progenitoras hematopoyéticas o células progenitoras neuronales.

20 En otro método, el Método 49, la presente divulgación proporciona un método como se proporciona en el Método 47, en donde las células somáticas completamente diferenciadas son células secretoras endocrinas tales como células beta pancreáticas, células epiteliales, células endodérmicas, macrófagos, hepatocitos, adipocitos, células renales, células sanguíneas o células del sistema inmunitario.

En otro método, el Método 49A, la presente divulgación proporciona un método, como se proporciona en el Método 47, en donde las células somáticas completamente diferenciadas son cardiomiocitos o células del sistema inmunitario.

30 En otra composición, la Composición 90, la presente divulgación proporciona una composición que comprende una pluralidad de células donantes universales generadas mediante uno cualquiera de los Métodos 20 a 49.

En otra composición, la Composición 91, la presente divulgación proporciona una composición, como se proporciona por la Composición 90, mantenida durante un tiempo y en condiciones suficientes para que las células experimenten diferenciación.

35

En otra composición, la Composición 92, la presente divulgación proporciona una composición, como se proporciona por la Composición 90 o 91, para su uso en el tratamiento de un sujeto que lo necesite.

5 En otra composición, la Composición 93, la presente divulgación proporciona una composición, como se proporciona por la Composición 92, en donde el sujeto es un ser humano que tiene, se sospecha que tiene o está en riesgo de tener una enfermedad.

En otro método, el Método 50, la presente divulgación proporciona un método que comprende administrar a un sujeto la pluralidad de células donantes universales de la Composición 90 o 91.

10 En otro método, el Método 51, la presente divulgación proporciona un método para tratar a un sujeto que lo necesite, comprendiendo el método: (a) obtener o haber obtenido la pluralidad de células donantes universales de la Composición 90 después de la diferenciación en células progenitoras de linaje restringido o células somáticas completamente diferenciadas; y (b) administrar al sujeto las células progenitoras de linaje restringido o las células somáticas completamente diferenciadas.

15 En otro método, el Método 52, la presente divulgación proporciona un método de obtención de células para su administración a un sujeto que lo necesite, comprendiendo el método: (a) obtener o haber obtenido las células donantes universales de la reivindicación 31; y (b) mantener las células donantes universales durante un tiempo y en condiciones suficientes para que las células se diferencien en células progenitoras de linaje restringido o células somáticas completamente diferenciadas.

20 En otro método, el Método 53, la presente divulgación proporciona un método, como se proporciona por los Métodos 51 o 52, en donde las células progenitoras de linaje restringido son células progenitoras del endodermo pancreático, células progenitoras endocrinas pancreáticas, células progenitoras mesenquimatosas, células progenitoras musculares, blastocitos, células progenitoras hematopoyéticas o células progenitoras neuronales.

25 En otro método, el Método 54, la presente divulgación proporciona un método como se proporciona por los Métodos 51 o 52, en donde las células somáticas completamente diferenciadas son células secretoras endocrinas tales como células beta pancreáticas, células epiteliales, células endodérmicas, macrófagos, hepatocitos, adipocitos, células renales, células sanguíneas o células del sistema inmunitario.

30 En otro método, el Método 54A, la presente divulgación proporciona un método, como se proporciona en el Método 51 o 52, en donde las células somáticas completamente diferenciadas son cardiomiocitos.

En otro método, el Método 55, la presente divulgación proporciona un método, como se proporciona en el Método 50 a 54, en donde el sujeto es un ser humano que tiene, se sospecha que tiene o está en riesgo de tener una enfermedad.

5 En otro método, el Método 56, la presente divulgación proporciona un método, como se proporciona por el Método 55, en donde la enfermedad es una enfermedad hereditaria genéticamente.

En otra composición, la Composición 93, la presente divulgación proporciona un ARN guía que comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 15-24.

10 En otro método, el Método 57, la presente divulgación proporciona un método *in vitro* para generar una célula donante universal, comprendiendo el método suministrar a una célula madre: (a) una nucleasa guiada por ARN; (b) un ARN guía (ARNg) que se dirige a un sitio diana en un locus del gen de la proteína que interactúa con la tiorredoxina (TXNIP); y (c) un vector que comprende un ácido nucleico, comprendiendo el ácido nucleico: (i) una secuencia de  
15 nucleótidos que codifica un factor tolerogénico; (ii) una secuencia de nucleótidos que consiste esencialmente en la SEQ ID NO: 25 y que tiene homología de secuencia con una región genómica ubicada a la izquierda y a menos de 50 nucleobases del sitio diana; y (iii) una secuencia de nucleótidos que consiste esencialmente en la SEQ ID NO: 32 y que tiene homología de secuencia con una región genómica ubicada a la derecha y a menos de 50  
20 nucleobases del sitio diana, en donde (i) está flanqueada por (ii) y (iii); en donde el locus del gen TXNIP se escinde en el sitio diana y el ácido nucleico se inserta en el locus del gen TXNIP, interrumpiendo de este modo el gen TXNIP y generando una célula donante universal, en donde la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula de control.

25 En otro método, el Método 57A, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde el ácido nucleico se inserta en el locus del gen TXNIP a menos de 50 pares de bases del sitio diana.

30 En otro método, el Método 58, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde la célula de control es una célula de tipo silvestre o una célula que no comprende el ácido nucleico insertado.

En otro método, el Método 59, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde el gen TXNIP interrumpido tiene la expresión reducida o eliminada de TXNIP.

35 En otro método, el Método 60, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde el ARNg comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 15-24.

En otro método, el Método 61, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde el ARNg comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 20.

5 En otro método, el Método 62, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde el vector es un vector plasmídico.

En otro método, el Método 63, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde el factor tolerogénico es el antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E).

10 En otro método, el Método 64, la presente divulgación proporciona un método, como se proporciona por el Método 63, en donde la secuencia de nucleótidos que codifica HLA-E comprende una secuencia que codifica un trímero de HLA-E que comprende un péptido señal de B2M fusionado a un péptido de presentación de HLA-G fusionado a una proteína de membrana B2M fusionada a HLA-E sin su péptido señal.

15 En otro método, el Método 65, la presente divulgación proporciona un método, como se proporciona por el Método 63, en donde la secuencia que codifica el trímero de HLA-E consiste esencialmente en la SEQ ID NO: 55.

En otro método, el Método 66, la presente divulgación proporciona un método, como se proporciona por el Método 65, en donde la secuencia que codifica el trímero de HLA-E está unida operativamente a un promotor exógeno.

20 En otro método, el método 67, la presente divulgación proporciona un método, como se proporciona por el Método 66, en donde el promotor exógeno es un promotor de CMV, EF1 $\alpha$ , PGK, CAG o UBC.

En otro método, el Método 68, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde la nucleasa guiada por ARN es una nucleasa Cas9.

25 En otro método, el método 69, la presente divulgación proporciona un método, como se proporciona por el Método 68, en donde la nucleasa Cas9 está unida a al menos una señal de localización nuclear.

30 En otro método, el Método 70, la presente divulgación proporciona un método, como se proporciona por el Método 69, en donde la nucleasa Cas9 y el ARNg están presentes en una relación molar de 1:3.

En otro método, el Método 71, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde la célula madre es una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

35 En otro método, el Método 72, la presente divulgación proporciona un método, como se proporciona por el Método 57, en donde la célula madre es una célula madre humana.

En otro método, el Método 73, la presente divulgación proporciona un método *in vitro* para generar una célula donante universal, comprendiendo el método suministrar a una célula madre: (a) una nucleasa guiada por ARN; (b) un ARN guía (ARNg) que se dirige a un sitio diana en un locus del gen de la proteína que interactúa con la tiorredoxina (TXNIP) y (c) un vector que comprende un ácido nucleico, comprendiendo el ácido nucleico: (i) una secuencia de nucleótidos que codifica un factor tolerogénico; (ii) una secuencia de nucleótidos que tiene homología de secuencia con una región genómica ubicada a la izquierda y a menos de 50 nucleobases del sitio diana; y (iii) una secuencia de nucleótidos que tiene homología de secuencia con una región genómica ubicada a la derecha y a menos de 50 nucleobases del sitio diana, en donde (i) está flanqueada por (ii) y (iii), y el vector comprende una secuencia de nucleótidos que consiste en la SEQ ID NO: 34 o 56; en donde el locus del gen TXNIP se escinde en el sitio diana y el ácido nucleico se inserta en el locus del gen TXNIP, interrumpiendo de este modo el gen TXNIP y generando una célula donante universal, en donde la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula de control.

En otro método, el Método 73A, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde el ácido nucleico se inserta en el locus del gen TXNIP a menos de 50 pares de bases del sitio diana.

En otro método, el Método 74, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde la célula de control es una célula de tipo silvestre o una célula que no comprende el ácido nucleico insertado.

En otro método, el Método 75, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde el gen TXNIP interrumpido tiene la expresión reducida o eliminada de TXNIP.

En otro método, el Método 76, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde el ARNg comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 15-24.

En otro método, el Método 77, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde el ARNg comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 20.

En otro método, el Método 78, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde el vector es un vector plasmídico.

En otro método, el Método 79, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde el factor tolerogénico es el antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E).

En otro método, el Método 80, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde la nucleasa guiada por ARN es una nucleasa Cas9.

En otro método, el método 81, la presente divulgación proporciona un método, como se proporciona por el Método 80, en donde la nucleasa Cas9 está unida a al menos una señal de localización nuclear.

En otro método, el Método 82, la presente divulgación proporciona un método, como se proporciona por el Método 80, en donde la nucleasa Cas9 y el ARNg están presentes en una relación molar de 1:3.

En otro método, el Método 83, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde la célula madre es una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

En otro método, el Método 84, la presente divulgación proporciona un método, como se proporciona por el Método 73, en donde la célula madre es una célula madre humana.

En otro método, el Método 85, la presente divulgación proporciona un método *in vitro* para generar una célula donante universal, comprendiendo el método suministrar a una célula madre: (a) un primer complejo de ribonucleoproteína (RNP) que comprende una nucleasa guiada por ARN y un ARN guía (ARNg) que se dirige a un sitio diana en un locus del gen de la beta-2 microglobulina (B2M); (b) un primer vector que comprende un ácido nucleico, comprendiendo el ácido nucleico: (i) una secuencia de nucleótidos que codifica un primer factor tolerogénico; (ii) una secuencia de nucleótidos que consiste esencialmente en la SEQ ID NO: 7 y que tiene homología de secuencia con una región genómica ubicada a la izquierda y a menos de 50 nucleobases del sitio diana en el locus del gen B2M; y (iii) una secuencia de nucleótidos que consiste esencialmente en la SEQ ID NO: 13 y que tiene homología de secuencia con una región genómica ubicada a la derecha y a menos de 50 nucleobases del sitio diana en el locus del gen B2M, en donde (i) está flanqueada por (ii) y (iii); en donde el locus del gen B2M se escinde en el sitio diana y el ácido nucleico que comprende la secuencia de nucleótidos que codifica el primer factor tolerogénico se inserta en el locus del gen B2M, interrumpiendo de este modo el gen B2M; (c) un segundo complejo de RNP que comprende una nucleasa guiada por ARN y un ARNg dirigido a un sitio diana en un locus del gen de la proteína que interactúa con la tiorredoxina (TXNIP); y (d) un segundo vector que comprende un ácido nucleico, comprendiendo el ácido nucleico: (i) una secuencia de nucleótidos que codifica un segundo factor tolerogénico; (ii) una secuencia de nucleótidos que consiste esencialmente en la SEQ ID NO: 25 y que tiene homología de secuencia con una región genómica ubicada a la izquierda y a menos de 50 nucleobases del sitio diana en el locus del gen TXNIP; y (iii) una secuencia de nucleótidos que consiste esencialmente en la SEQ ID NO: 32 y que tiene homología de secuencia con una región genómica ubicada a la derecha y a menos de 50 nucleobases del

sitio diana en el locus del gen TXNIP, en donde (i) está flanqueada por (ii) y (iii); en donde el locus del gen TXNIP se escinde en el sitio diana y el ácido nucleico que comprende la secuencia de nucleótidos que codifica el segundo factor tolerogénico se inserta en el locus del gen TXNIP, interrumpiendo de este modo el gen TXNIP y generando una célula donante universal, en donde la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula de control.

En otro método, el Método 85A, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde el ácido nucleico en (b) se inserta en el locus del gen B2M a menos de 50 pares de bases del sitio diana y/o en donde el ácido nucleico en (d) se inserta en el locus del gen TXNIP a menos de 50 pares de bases del sitio diana.

En otro método, el Método 86, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde la célula de control es una célula de tipo silvestre o una célula que no comprende el ácido nucleico insertado.

En otro método, el Método 87, la presente divulgación proporciona un método, como se proporciona por el Método 85, el gen B2M interrumpido tiene la expresión reducida o eliminada de B2M, y el gen TXNIP interrumpido tiene la expresión reducida o eliminada de TXNIP.

En otro método, el Método 88, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde el ARNg del primer complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 1-3 o 35-44, y el ARNg del segundo complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 15-24.

En otro método, el Método 89, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde el ARNg del primer complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en la SEQ ID NO: 2 y el ARNg del segundo complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en la SEQ ID NO: 20.

En otro método, el Método 90, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde el primer vector es un vector plasmídico y el segundo vector es un vector plasmídico.

En otro método, el Método 91, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde el primer factor tolerogénico es el ligando de muerte programada 1 (PD-L1) y el segundo factor tolerogénico es el antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E).

En otro método, el Método 92, la presente divulgación proporciona un método, como se proporciona por el Método 91, en donde la secuencia de nucleótidos que codifica PD-L1 consiste esencialmente en la SEQ ID NO: 11.

En otro método, el Método 93, la presente divulgación proporciona un método, como se proporciona por el Método 91, en donde la secuencia de nucleótidos que codifica HLA-E comprende una secuencia que codifica un trímero de HLA-E que comprende un péptido señal de B2M fusionado a un péptido de presentación de HLA-G fusionado a una proteína de membrana B2M fusionada a HLA-E sin su péptido señal y la secuencia que codifica el trímero de HLA-E consiste esencialmente en la SEQ ID NO: 55.

En otro método, el Método 94, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde la secuencia de nucleótidos que codifica el primer factor tolerogénico está unida operativamente a un promotor exógeno y la secuencia de nucleótidos que codifica el segundo factor tolerogénico está unida operativamente a un promotor exógeno.

En otro método, el método 95, la presente divulgación proporciona un método, como se proporciona por el Método 94, en donde el promotor exógeno es un promotor de CMV, EF1 $\alpha$ , PGK, CAG o UBC.

En otro método, el Método 96, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde cada uno de entre el primer complejo de RNP y el segundo complejo de RNP comprende una relación molar de nucleasa guiada por ARNg de 1:3.

En otro método, el Método 97, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde la nucleasa guiada por ARN de cada uno de entre el primer complejo de RNP y el segundo complejo de RNP es una nucleasa Cas9.

En otro método, el método 98, la presente divulgación proporciona un método, como se proporciona por el Método 97, en donde la nucleasa Cas9 está unida a al menos una señal de localización nuclear.

En otro método, el Método 99, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde la célula madre es una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

En otro método, el Método 100, la presente divulgación proporciona un método, como se proporciona por el Método 85, en donde la célula madre es una célula madre humana.

En otro método, el Método 101, la presente divulgación proporciona un método *in vitro* para generar una célula donante universal, comprendiendo el método suministrar a una célula madre: (a) un primer complejo de ribonucleoproteína (RNP) que comprende una nucleasa guiada por ARN y un ARN guía (ARNg) que se dirige a un sitio diana en un locus del gen de la beta-2 microglobulina (B2M); (b) un primer vector que comprende un ácido nucleico, comprendiendo el ácido nucleico: (i) una secuencia de nucleótidos que codifica un primer factor tolerogénico; (ii) una secuencia de nucleótidos que tiene homología de secuencia con una región genómica ubicada a la izquierda y a menos de 50 nucleobases del sitio diana en el locus del gen B2M; y (iii) una secuencia de nucleótidos que tiene homología de secuencia con una región

genómica ubicada a la derecha y a menos de 50 nucleobases del sitio diana en el locus del gen B2M, en donde (i) está flanqueada por (ii) y (iii) y el primer vector comprende una secuencia de nucleótidos que consiste en la SEQ ID NO: 33; en donde el locus del gen B2M se escinde en el sitio diana y el ácido nucleico que comprende la secuencia de nucleótidos que codifica el primer factor tolerogénico se inserta en el locus del gen B2M, interrumpiendo de este modo el gen B2M; (c) un segundo complejo de RNP que comprende una nucleasa guiada por ARN y un ARNg dirigido a un sitio diana en un locus del gen de la proteína que interactúa con la tiorredoxina (TXNIP); y (d) un segundo vector que comprende un ácido nucleico, comprendiendo el ácido nucleico: (i) una secuencia de nucleótidos que codifica un segundo factor tolerogénico; (ii) una secuencia de nucleótidos que tiene homología de secuencia con una región genómica ubicada a la izquierda y a menos de 50 nucleobases del sitio diana en el locus del gen TXNIP; y (iii) una secuencia de nucleótidos que tiene homología de secuencia con una región genómica ubicada a la derecha y a menos de 50 nucleobases del sitio diana en el locus del gen TXNIP, en donde (i) está flanqueada por (ii) y (iii) y el segundo vector que comprende una secuencia de nucleótidos que consiste en las SEQ ID NO: 34 o 56 en donde el locus del gen TXNIP se escinde en el sitio diana y el ácido nucleico que comprende la secuencia de nucleótidos que codifica el segundo factor tolerogénico se inserta en el locus del gen TXNIP, interrumpiendo de este modo el gen TXNIP y generando una célula donante universal, en donde la célula donante universal tiene una mayor evasión inmunitaria y/o supervivencia celular en comparación con una célula de control.

En otro método, el Método 101A, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde el ácido nucleico en (b) se inserta en el locus del gen B2M a menos de 50 pares de bases del sitio diana y/o en donde el ácido nucleico en (d) se inserta en el locus del gen TXNIP a menos de 50 pares de bases del sitio diana.

En otro método, el Método 102, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde la célula de control es una célula de tipo silvestre o una célula que no comprende el ácido nucleico insertado.

En otro método, el Método 103, la presente divulgación proporciona un método, como se proporciona por el Método 10, en donde el gen B2M interrumpido tiene la expresión reducida o eliminada de B2M, y el gen TXNIP interrumpido tiene la expresión reducida o eliminada de TXNIP.

En otro método, el Método 104, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde el ARNg del primer complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 1-3 o 35-44, y el ARNg del segundo complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en las SEQ ID NO: 15-24.

En otro método, el Método 105, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde el ARNg del primer complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en la SEQ ID NO: 2 y el ARNg del segundo complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia que consiste en la SEQ ID NO: 20.

En otro método, el Método 106, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde el primer vector es un vector plasmídico y el segundo vector es un vector plasmídico.

En otro método, el Método 107, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde el primer factor tolerogénico es el ligando de muerte programada 1 (PD-L1) y el segundo factor tolerogénico es el antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E).

En otro método, el Método 108, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde cada uno de entre el primer complejo de RNP y el segundo complejo de RNP comprende una relación molar de nucleasa guiada por ARNg de 1:3.

En otro método, el Método 109, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde la nucleasa guiada por ARN de cada uno de entre el primer complejo de RNP y el segundo complejo de RNP es una nucleasa Cas9.

En otro método, el método 110, la presente divulgación proporciona un método, como se proporciona por el Método 109, en donde la nucleasa Cas9 está unida a al menos una señal de localización nuclear.

En otro método, el Método 111, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde la célula madre es una célula madre embrionaria, una célula madre adulta, una célula madre pluripotente inducida o una célula madre hematopoyética.

En otro método, el Método 112, la presente divulgación proporciona un método, como se proporciona por el Método 101, en donde la célula madre es una célula madre humana.

En un primer proceso, el Proceso 1, la presente divulgación proporciona un proceso para generar células donantes universales, comprendiendo el proceso: (a) modificar las células madre insertando una secuencia de nucleótidos que codifica el ligando de muerte programada 1 (PD-L1) dentro o cerca de un gen que codifica la microglobulina beta-2 (B2M), generando de este modo células positivas para PD-L1; (b) enriquecer las células positivas para PD-L1; (c) modificar las células positivas para PD-L1 de (b) insertando una secuencia de nucleótidos que codifica el antígeno de histocompatibilidad HLA de clase I, la cadena alfa E (HLA-E) dentro o cerca de un gen que codifica la proteína que interactúa con la tiorredoxina (TXNIP), generando de este modo células doblemente positivas para PD-L1 y HLA-E; (d) enriquecer las células

doblemente positivas para PD-L1 y HLA-E; (e) clasificar células individuales para seleccionar células doblemente positivas para PD-L1 y HLA-E; (f) caracterizar las células de (e) como células donantes universales; y (g) congelar las células donantes universales para su almacenamiento a largo plazo.

5           En otro proceso, el Proceso 2, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, en donde la modificación en (a) comprende suministrar a las células madre (1) un primer complejo de ribonucleoproteína (RNP) que comprende una nucleasa guiada por ARN y un ARN guía (ARNg) que se dirige a un sitio diana en el locus del gen B2M y (2) un primer vector que comprende un ácido nucleico, comprendiendo el ácido  
10   nucleico (i) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana en el locus del gen B2M, (ii) la secuencia de nucleótidos que codifica PD-L1 y (iii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana en el locus del gen B2M, en donde el locus del gen B2M se escinde en el sitio diana y el ácido nucleico que comprende la secuencia de nucleótidos que codifica PD-L1 se inserta en el locus del gen  
15   B2M, interrumpiendo de este modo el gen B2M.

          En otro proceso, el Proceso 2A, la presente divulgación proporciona un método, como se proporciona por el Proceso 2, en donde el ácido nucleico se inserta en el locus del gen B2M a menos de 50 pares de bases del sitio diana.

          En otro proceso, el Proceso 3, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 2, en donde la nucleasa guiada por ARN del primer complejo de RNP es una nucleasa Cas9 y el ARNg del primer complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia diana que consiste en la SEQ ID NO: 2.  
20

          En otro proceso, el Proceso 4, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 3, en donde la nucleasa Cas9 se une a al menos una señal de localización nuclear.  
25

          En otro proceso, el Proceso 5, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 2, en donde la primera RNP comprende una relación molar de ARNg:nucleasa guiada por ARN de 3:1.

          En otro proceso, el Proceso 6, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 2, en donde la secuencia de nucleótidos de (a)(2)(i) consiste esencialmente en la SEQ ID NO: 7, y la secuencia de nucleótidos de (a)(2)(iii) consiste esencialmente en la SEQ ID NO: 13.  
30

          En otro proceso, el Proceso 7, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 2, en donde la secuencia de nucleótidos que codifica PD-L1 consiste esencialmente en la SEQ ID NO: 11.  
35

En otro proceso, el Proceso 8, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 2, en donde la secuencia de nucleótidos que codifica PD-L1 está unida operativamente a un promotor CAG.

5 En otro proceso, el Proceso 9, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 2, en donde el primer vector es un vector plasmídico y comprende una secuencia de nucleótidos que consiste en la SEQ ID NO: 33.

En otro proceso, el Proceso 10, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 2, en donde el suministro de (a)(1) y (a)(2) comprende la electroporación.

10 En otro proceso, el Proceso 11, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, en donde el enriquecimiento en células positivas para PD-L1 en (b) comprende la clasificación celular asistida magnéticamente (MACS, por sus siglas en inglés), la clonación de células individuales, la expansión de dichas células positivas para PD-L1 o una combinación de las mismas.

15 En otro proceso, el Proceso 12, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, en donde la modificación en (c) comprende suministrar a las células positivas para PD-L1 (1) un segundo complejo de RNP que comprende una nucleasa guiada por ARN y un ARNg dirigido a un sitio diana en el locus del gen TXNIP y (2) un segundo vector que comprende un ácido nucleico, comprendiendo el ácido nucleico (i) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana en el locus del gen TXNIP, (ii) la secuencia de nucleótidos que codifica HLA-E y (iii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana en el locus del gen TXNIP, en donde el locus del gen TXNIP se escinde en el sitio diana y el ácido nucleico que comprende la secuencia de nucleótidos que codifica HLA-E se inserta en el locus del gen TXNIP,  
20 interrumpiendo de este modo el gen TXNIP.

25 En otro proceso, el Proceso 12A, la presente divulgación proporciona un método, como se proporciona por el Proceso 12, en donde el ácido nucleico se inserta en el locus del gen TXNIP a menos de 50 pares de bases del sitio diana.

30 En otro proceso, el Proceso 13, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 12, en donde la nucleasa guiada por ARN del segundo complejo de RNP es una nucleasa Cas9 y el ARNg del segundo complejo de RNP comprende una secuencia espaciadora correspondiente a una secuencia diana que consiste en la SEQ ID NO: 20.

35 En otro proceso, el Proceso 14, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 13, en donde la nucleasa Cas9 se une a al menos una señal de localización nuclear.

En otro proceso, el Proceso 15, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 12, en donde la segunda RNP comprende una relación molar de ARNg:nucleasa guiada por ARN de 3:1.

5 En otro proceso, el Proceso 16, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 12, en donde la secuencia de nucleótidos de (c)(2)(i) consiste esencialmente en la SEQ ID NO: 25, y la secuencia de nucleótidos de (c)(2)(iii) consiste esencialmente en la SEQ ID NO: 32.

10 En otro proceso, el Proceso 17, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 12, en donde la secuencia de nucleótidos que codifica HLA-E comprende una secuencia que codifica un trímero de HLA-E que comprende un péptido señal de B2M fusionado a un péptido de presentación de HLA-G fusionado a una proteína de membrana B2M fusionada a HLA-E sin su péptido señal.

15 En otro proceso, el Proceso 18, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 17, en donde la secuencia que codifica el trímero de HLA-E consiste esencialmente en la SEQ ID NO: 55.

En otro proceso, el Proceso 19, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 12, en donde la secuencia de nucleótidos que codifica HLA-E está unida operativamente a un promotor CAG.

20 En otro proceso, el Proceso 20, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 12, en donde el segundo vector es un vector plasmídico y comprende una secuencia de nucleótidos que consiste en las SEQ ID NO: 34 o 56.

En otro proceso, el Proceso 21, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 12, en donde el suministro de (c)(1) y (c)(2) comprende la electroporación.

25 En otro proceso, el Proceso 22, la presente divulgación proporciona un proceso, como se proporciona en el Proceso de la reivindicación 1, en donde el enriquecimiento en células doblemente positivas para PD-L1 y HLA-E en (d) comprende la clasificación celular asistida magnéticamente, la clonación de células individuales, la expansión de dichas células doblemente positivas para PD-L1 y HLA-E, o una combinación de las mismas.

30 En otro proceso, el Proceso 23, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, en donde la clasificación de células individuales en (e) comprende la clasificación de células activadas por fluorescencia (FACS, por sus siglas en inglés), la clonación de células individuales, la expansión de dichas células clasificadas en células individuales o una combinación de las mismas.

En otro proceso, el Proceso 24, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, en donde la caracterización en (f) comprende el análisis de ADN para determinar el perfil de cigosidad y/o indel.

5 En otro proceso, el Proceso 25, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, en donde la caracterización en (f) comprende análisis celulares para determinar la morfología, viabilidad, cariotipo, niveles de endotoxina, niveles de micoplasma, análisis de dianas activadas/inactivadas, inserción de vectores aleatoriamente, Cas9 residual, vector residual, estado de pluripotencia, capacidad de diferenciación o una combinación de los mismos.

10 En otro proceso, el Proceso 26, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, en donde el proceso comprende además la congelación antes de la caracterización en (f).

15 En otro proceso, el Proceso 27, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 1, que comprende además en (a) expandir las células positivas para PD-L1 generadas, en (c) expandir las células doblemente positivas para PD-L1 y HLA-E generadas, en (e) expandir las células doblemente positivas para PD-L1 y HLA-E seleccionadas o una combinación de los mismos.

20 En otro proceso, el Proceso 28, la presente divulgación proporciona un proceso para generar células donantes universales, comprendiendo el proceso: (a) modificar las células madre insertando una secuencia de nucleótidos que codifica un primer factor tolerogénico dentro o cerca de un gen que codifique un antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un complejo MHC-I o MHC-II, generando de este modo células positivas para el primer factor tolerogénico; (b) enriquecer las células positivas para el primer factor tolerogénico; (c) modificar las células positivas para el primer factor tolerogénico de (b) insertando una secuencia de nucleótidos que codifica un segundo factor tolerogénico dentro o cerca de un gen que codifica un factor de supervivencia, generando de este modo células positivas para el primer factor tolerogénico/positivas para el segundo factor tolerogénico (d) enriquecer las células positivas para el primer factor tolerogénico/positivas para el segundo factor tolerogénico; (e) clasificar las células individualmente para seleccionar las células positivas para el primer factor tolerogénico/positivas para el segundo factor tolerogénico; (f) caracterizar las células de (e) como células donantes universales; y (g) congelar las células donantes universales para su almacenamiento a largo plazo.

35 En otro proceso, el Proceso 29, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 28, en donde el enriquecimiento en células positivas para el primer factor tolerogénico en (b) comprende la clasificación celular asistida magnéticamente (MACS),

la clonación de células individuales, la expansión de dichas células positivas para el primer factor tolerogénico o una combinación de las mismas.

En otro proceso, el Proceso 30, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 28 o 29, en donde el enriquecimiento en células positivas para el primer factor tolerogénico/positivas para el segundo factor tolerogénico en (b) comprende la  
5 clasificación celular asistida magnéticamente, la clonación de células individuales, la expansión de dichas células positivas para el primer factor tolerogénico/positivas para el segundo factor tolerogénico o una combinación de las mismas.

En otro proceso, el Proceso 31, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 30, que comprende además en (a)  
10 expandir las células positivas para el primer factor tolerogénico generadas, en (c) expandir las células positivas para el primer factor tolerogénico/positivas para el segundo factor tolerogénico generadas, en (e) expandir las células positivas para el primer factor tolerogénico/positivas para el segundo factor tolerogénico seleccionadas o una combinación de los mismos.

En otro proceso, el Proceso 32, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 31, en donde la modificación en (a) comprende suministrar a las células madre (1) una primera nucleasa guiada por ARN y un primer ARN guía (ARNg) que se dirige a un sitio diana en un antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II y (2) un primer vector que comprende un primer ácido nucleico, el  
20 primer ácido nucleico que comprende (i) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana en los antígenos leucocitarios humanos MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II, (ii) la secuencia de nucleótidos que codifica el primer factor tolerogénico y (iii) una  
25 secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana en los antígenos leucocitarios humanos MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II, en donde el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II se escinde en el sitio diana y el primer ácido nucleico que comprende la secuencia de nucleótidos que codifica el primer factor tolerogénico se inserta  
30 en el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II, interrumpiendo de este modo el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un gen del complejo MHC-I o MHC-II.

En otro proceso, el Proceso 32A, la presente divulgación proporciona un método, como se proporciona por el Proceso 32, en donde el ácido nucleico se inserta en el antígeno  
35

leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II a menos de 50 pares de bases del sitio diana.

En otro proceso, el Proceso 33, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 32, en donde la primera nucleasa guiada por ARN y el primer ARNg forman un primer complejo de ribonucleoproteína (RNP).

En otro proceso, el Proceso 34, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 33, en donde la modificación en (a) comprende suministrar a las células madre (1) un primer complejo de ribonucleoproteína (RNP) que comprende una primera nucleasa guiada por ARN y un primer ARN guía (ARNg) que se dirige a un sitio diana en un antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II y (2) un primer vector que comprende un primer ácido nucleico, el primer ácido nucleico que comprende (i) una secuencia de nucleótidos homóloga con una región ubicada a la izquierda del sitio diana en el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II, (ii) la secuencia de nucleótidos que codifica el primer factor tolerogénico y (iii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana en el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II, en donde el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II se escinde en el sitio diana y el primer ácido nucleico que comprende la secuencia de nucleótidos que codifica el primer factor tolerogénico se inserta en el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II, interrumpiendo de este modo el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un gen del complejo MHC-I o MHC-II.

En otro proceso, el Proceso 34A, la presente divulgación proporciona un método, como se proporciona por el Proceso 34, en donde el ácido nucleico se inserta en el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un locus del gen del complejo MHC-I o MHC-II a menos de 50 pares de bases del sitio diana.

En otro proceso, el Proceso 35, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 34, en donde el antígeno leucocitario humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un gen del complejo MHC-I o MHC-II es HLA-A, HLA-B, HLA-C, HLA-DP, HLA-DM, HLA-DOA, HLA-DOB, HLA-DQ, HLA-DR, B2M, NLRC5, CIITA, RFX5, RFXAP o RFXANK.

En otro proceso, el Proceso 36, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 35, en donde el antígeno leucocitario

humano MHC-I o MHC-II o un componente o un regulador de la transcripción de un gen del complejo MHC-I o MHC-II es B2M.

5 En otro proceso, el Proceso 37, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 36, en donde la secuencia de nucleótidos de (a)(2)(i) consiste esencialmente en la SEQ ID NO: 7, y la secuencia de nucleótidos de (a)(2)(iii) consiste esencialmente en la SEQ ID NO: 13.

En otro proceso, el Proceso 38, la presente divulgación proporciona un proceso, como se proporciona en los Procesos 36 o 37, en donde el primer ARNg comprende una secuencia espaciadora correspondiente a una secuencia diana que consiste en la SEQ ID NO: 2.

10 En otro proceso, el Proceso 39, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 32 a 38, en donde la primera nucleasa guiada por ARN es una nucleasa Cas9.

En otro proceso, el Proceso 40 la presente divulgación proporciona un proceso, como se proporciona en el Proceso 39, en donde la nucleasa Cas9 se une a al menos una señal de localización nuclear.

15 En otro proceso, el Proceso 41, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 32 a 40, en donde la primera RNP comprende una relación molar de primer ARNg:primera nucleasa guiada por ARN de 3:1.

En otro proceso, el Proceso 42, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 41, en donde el primer factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47.

En otro proceso, el Proceso 43, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 42, en donde la secuencia de nucleótidos que codifica el primer factor tolerogénico está unida operativamente a un promotor exógeno.

25 En otro proceso, el Proceso 44, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 43, en donde el promotor exógeno es un promotor de CMV, EF1 $\alpha$ , PGK, CAG o UBC.

En otro proceso, el Proceso 45, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 44, en donde el primer factor tolerogénico es PD-L1.

30 En otro proceso, el Proceso 46, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 45, en donde la secuencia de nucleótidos que codifica PD-L1 consiste esencialmente en la SEQ ID NO: 11.

En otro proceso, el Proceso 47, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 46, en donde la secuencia de nucleótidos que codifica PD-L1 está unida operativamente a un promotor CAG.

35

En otro proceso, el Proceso 48, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 45 a 47, en donde el primer vector comprende una secuencia de nucleótidos que consiste en la SEQ ID NO: 33.

5 En otro proceso, el Proceso 49, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 48, en donde la modificación en (c) comprende suministrar a las células madre (1) una segunda nucleasa guiada por ARN y un segundo ARN guía (ARNg) que se dirige a un sitio diana en un locus del gen del factor de supervivencia y (2) un segundo vector que comprende un segundo ácido nucleico, comprendiendo el segundo ácido nucleico (i) una secuencia de nucleótidos homóloga con una  
10 región ubicada a la izquierda del sitio diana en el locus del gen del factor de supervivencia, (ii) la secuencia de nucleótidos que codifica el segundo factor tolerogénico y (iii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana en el locus del gen del factor de supervivencia, en donde el locus del gen del factor de supervivencia se escinde en el sitio diana y el segundo ácido nucleico que comprende la secuencia de nucleótidos que  
15 codifica el segundo factor tolerogénico se inserta en el locus del gen del factor de supervivencia, interrumpiendo de este modo el gen del factor de supervivencia.

En otro proceso, el Proceso 49A, la presente divulgación proporciona un método, como se proporciona por el Proceso 49, en donde el ácido nucleico se inserta en el locus del gen del factor de supervivencia a menos de 50 pares de bases del sitio diana.

20 En otro proceso, el Proceso 50, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 49, en donde la segunda nucleasa guiada por ARN y el segundo ARNg forman un segundo complejo de ribonucleoproteína (RNP).

En otro proceso, el Proceso 51, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 48, en donde la modificación en (c)  
25 comprende suministrar a las células positivas para el primer factor tolerogénico (1) un segundo complejo de ribonucleoproteína (RNP) que comprende una segunda nucleasa guiada por ARN y un segundo ARN guía (ARNg) que se dirige a un sitio diana en un locus del gen del factor de supervivencia y (2) un segundo vector que comprende un segundo ácido nucleico, comprendiendo el segundo ácido nucleico (i) una secuencia de nucleótidos homóloga con una  
30 región ubicada a la izquierda del sitio diana en el locus del gen del factor de supervivencia, (ii) la secuencia de nucleótidos que codifica el segundo factor tolerogénico y (iii) una secuencia de nucleótidos homóloga con una región ubicada a la derecha del sitio diana en el locus del gen del segundo factor de supervivencia, en donde el locus del gen del factor de supervivencia se escinde en el sitio diana y el segundo ácido nucleico que comprende la secuencia de nucleótidos  
35 que codifica el segundo factor tolerogénico se inserta en el locus del gen del factor de supervivencia, interrumpiendo de este modo el gen del factor de supervivencia.

En otro proceso, el Proceso 51A, la presente divulgación proporciona un método, como se proporciona por el Proceso 51, en donde el ácido nucleico se inserta en el locus del gen del factor de supervivencia a menos de 50 pares de bases del sitio diana.

5 En otro proceso, el Proceso 52, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 51, en donde el gen de supervivencia es TXNIP, ZNF143, FOXO1, JNK o MANF.

En otro proceso, el Proceso 53, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 52, en donde el gen de supervivencia es TXNIP.

10 En otro proceso, el Proceso 54, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 53, en donde el segundo ARNg comprende una secuencia espaciadora correspondiente a una secuencia diana que consiste en la SEQ ID NO: 20.

En otro proceso, el Proceso 55, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 52 o 53, en donde la secuencia de nucleótidos de (c)(2)(i) consiste esencialmente en la SEQ ID NO: 25, y la secuencia de nucleótidos de (c)(2)(iii) consiste esencialmente en la SEQ ID NO: 32.

15 En otro proceso, el Proceso 56, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 49 a 55, en donde la segunda nucleasa guiada por ARN es una nucleasa Cas9.

20 En otro proceso, el Proceso 57, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 56, en donde la nucleasa Cas9 se une a al menos una señal de localización nuclear.

En otro proceso, el Proceso 58, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 49 a 57, en donde la segunda RNP comprende una relación molar de segundo ARNg:segunda nucleasa guiada por ARN de 3:1.

25 En otro proceso, el Proceso 59, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 49 a 58, en donde el segundo factor tolerogénico es PD-L1, HLA-E, HLA-G, CTLA-4 o CD47.

30 En otro proceso, el Proceso 60, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 59, en donde la secuencia de nucleótidos que codifica el segundo factor tolerogénico está unida operativamente a un promotor exógeno.

En otro proceso, el Proceso 61, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 60, en donde el promotor exógeno es un promotor de CMV, EF1 $\alpha$ , PGK, CAG o UBC.

35 En otro proceso, el Proceso 62, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 61, en donde el segundo factor tolerogénico es HLA-E.

En otro proceso, el Proceso 63, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 62, en donde la secuencia de nucleótidos que codifica HLA-E comprende una secuencia que codifica un trímero de HLA-E que comprende un péptido señal de B2M fusionado a un péptido de presentación de HLA-G fusionado a una proteína de membrana B2M fusionada a HLA-E sin su péptido señal.

En otro proceso, el Proceso 64, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 63, en donde la secuencia que codifica el trímero de HLA-E consiste esencialmente en la SEQ ID NO: 55.

En otro proceso, el Proceso 65, la presente divulgación proporciona un proceso, como se proporciona en el Proceso 63 o 64, en donde la secuencia de nucleótidos que codifica HLA-E está unida operativamente a un promotor CAG.

En otro proceso, el Proceso 66, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 62 a 65, en donde el segundo vector comprende una secuencia de nucleótidos que consiste en las SEQ ID NO: 34 o 56.

En otro proceso, el Proceso 67, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 66, en donde la clasificación de células individuales en (e) comprende la clasificación de células activadas por fluorescencia (FACS), la clonación de células individuales, la expansión de dichas células clasificadas en células individuales o una combinación de las mismas.

En otro proceso, el Proceso 68, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 67, en donde la caracterización en (f) comprende el análisis de ADN para el perfil de cigosidad y/o indel.

En otro proceso, el Proceso 69, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 68, en donde la caracterización en (f) comprende análisis celulares para determinar la morfología, viabilidad, cariotipo, niveles de endotoxina, niveles de micoplasma, análisis de dianas activadas/inactivadas, inserción de vectores aleatoriamente, Cas9 residual, vector residual, estado de pluripotencia, capacidad de diferenciación o una combinación de los mismos.

En otro proceso, el Proceso 70, la presente divulgación proporciona un proceso, como se proporciona en uno cualquiera de los Procesos 28 a 69, en donde el proceso comprende además la congelación antes de la caracterización en (f).

## **VII. Ejemplos**

Los siguientes ejemplos describen la generación y caracterización de células donantes universales específicas de acuerdo con la presente divulgación.

### ***Ejemplo 1: Mantenimiento y expansión celular***

**Mantenimiento de hESC/hiPSC.** Se mantuvieron, se cultivaron, se hicieron pases de, se hicieron proliferar y se sembraron células de la línea de células madre embrionarias humanas CyT49 (línea celular hES patentada, ViaCyte, Inc., San Diego, CA) como se desvela en Schulz *et al.* (2012) PLoS ONE 7(5): e37004. Las células CyT49 se disociaron usando ACCUTASE® (Stemcell Technologies 07920 o equivalente).

Se mantuvieron células madre pluripotentes inducidas humanas (hiPSC), tales como la línea celular TC1133 (Lonza), en StemFlex Complete (Life Technologies, A3349401) sobre placas de cultivo tisular recubiertas con BIOLAMININA 521 CTG (BioLamina n.º de cat. CT521). Las placas se recubrieron previamente con una dilución 1:10 o 1:20 de BIOLAMININA en DPBS, calcio, magnesio (Life Technologies, 14040133) durante 2 horas a 37 °C. Las células se alimentaron a diario con medio StemFlex. Para los pases de las células, se usaron las mismas densidades de células que para CyT49. Para la siembra en placa de las células como células individuales, las células se sembraron en placa con Suplemento RevitaCell™ al 1 % (100X) (Thermofisher n.º de cat. A2644501) en StemFlex sobre placas recubiertas con BIOLAMININA.

**Clonación de células hPSC individuales.** Para la clonación de células individuales, se alimentaron hPSC (hESC o hiPSC) con StemFlex Complete con Revitacell (para una concentración final de 1X Revitacell) 3-4 horas antes de la disociación con ACCUTASE®. Después de la disociación, las células se clasificaron como una célula individual por pocillo de una placa de cultivo tisular de 96 pocillos recubierta con BIOLAMININA. Se usó el clasificador WOLF FACS (Nanocollect) para clasificar las células individuales en los pocillos. Las placas se llenaron previamente con 100-200 µl de StemFlex Complete con Revitacell. Tres días después de la siembra celular, las células se alimentaron con StemFlex nuevo y se siguieron alimentando cada dos días con 100-200 µl de medio. Después de 10 días de crecimiento, las células se alimentaron a diario con StemFlex hasta el día 12-14. En este momento, las placas se disociaron con ACCUTASE® y las suspensiones de células recogidas se dividieron 1:2 con una mitad en una nueva placa de 96 pocillos para mantenimiento y la otra mitad en una solución de extracción de ADN QuickExtract™ (Lucigen). Después de la extracción del ADN, se realizó una PCR para evaluar la presencia o ausencia de ediciones génicas deseadas en los locus de ADN diana. Se usó la secuenciación de Sanger para verificar las ediciones deseadas.

**Expansión de clones de hPSC derivados de células individuales.** Para CyT49 (ViaCyte), los clones dirigidos con éxito se pasaron a placas de 24 pocillos con medio XF KSR A10H10 al 10 % puro, pero sobre placas recubiertas con BIOLAMININA. Después de la fase de 24 pocillos, los clones de CyT49 se pasaron como se describe en Schulz *et al.* (2012) PLoS ONE 7(5): e37004.

Para las hiPSC (TC1133), las células se mantuvieron en StemFlex Complete durante los procesos de clonación y mantenimiento regular sobre placas recubiertas con BIOLAMININA con Revitacell en las fases de pases.

**Ejemplo 2: Generación de células madre pluripotentes humanas (hPSC) con desactivación de B2M**

**Selección de ARN guía (ARNg) para B2M en hPSC.** Se diseñaron tres ARNg dirigidos a B2M para dirigirse al exón 1 de la secuencia codificante de B2M. Estos ARNg tenían bajas puntuaciones inespecíficas predichas basadas en la predicción de la homología de secuencia usando el software de diseño de ARNg. Las secuencias diana de los ARNg se presentan en la Tabla 1. Un ARNg comprende una secuencia de ARN correspondiente a la secuencia de ADN diana.

**Tabla 1. Secuencias diana del ARNg de B2M**

| Nombre                     | Secuencia diana (5'-3') | SEQ ID NO: | PAM |
|----------------------------|-------------------------|------------|-----|
| ARNg de B2M-1 (Exón 1_T12) | GCTACTCTCTCTTTCTGGCC    | 1          | TGG |
| ARNg de B2M-2 (Exón 1_T2)  | GGCCGAGATGTCTCGCTCCG    | 2          | TGG |
| ARNg de B2M-3 (Exón 1_T8)  | CGCGAGCACAGCTAAGGCCA    | 3          | CGG |
| Exón 1_T1                  | TATAAGTGGAGGCGTCGCGC    | 35         | TGG |
| Exón 1_T3                  | GAGTAGCGCGAGCACAGCTA    | 36         | AGG |
| Exón 1_T4                  | ACTGGACGCGTCGCGCTGGC    | 37         | GGG |
| Exón 1_T5                  | AAGTGGAGGCGTCGCGCTGG    | 38         | CGG |
| Exón 1_T6                  | GGCCACGGAGCGAGACATCT    | 39         | CGG |
| Exón 1_T7                  | GCCCGAATGCTGTCAGCTTC    | 40         | AGG |
| Exón 1_T9                  | CTCGCGCTACTCTCTTTTC     | 41         | TGG |
| Exón 1_T10                 | TCCTGAAGCTGACAGCATTC    | 42         | GGG |
| Exón 1_T11                 | TTCCTGAAGCTGACAGCATT    | 43         | CGG |
| Exón 1_T13                 | ACTCTCTCTTTCTGGCCTGG    | 44         | AGG |

Para evaluar su eficacia de corte en hPSCs, se electroporaron células CyT49 (línea celular hES patentada de ViaCyte) usando el electroporador Neon Electroporator (sistema de transfección Neon Transfection System de ThermoFisher n.º de cat. MPK5000) con una mezcla de ribonucleoproteína (RNP) de proteína Cas9 (Biomay) y ARN guía (Synthego) (Véase la Tabla

3 para consultar las secuencias de ARNg) en una relación molar de 3:1 (ARNg:Cas9) con valores absolutos de 125 pmol de Cas9 y 375 pmol de ARNg. Para formar el complejo de RNP, se combinaron ARNg y Cas9 en un recipiente con tampón R (kit de sistema de transfección Neon de 100 µl ThermoFisher n.º de cat. MPK10096) hasta un volumen total de 25 µl y se incubó durante 15 min a TA. Las células se disociaron usando ACCUTASE®, a continuación se volvieron a suspender en medio DMEM/F12 (Gibco, n.º de cat. 11320033), se contaron usando un NC-200 (Chemometec) y se centrifugaron. Se volvió a suspender un total de  $1 \times 10^6$  células con el complejo de RNP y se añadió tampón R hasta un volumen total de 125 µl. A continuación, esta mezcla electroporó con 2 pulsos durante 30 ms a 1100 V. Después de la electroporación, las células se pipetearon en un tubo Eppendorf lleno de medio StemFlex media con RevitaCell. A continuación, esta suspensión celular se sembró en placas de cultivo tisular recubiertas previamente con BIOLAMININA 521 CTG a una dilución 1:20. Las células se cultivaron en una incubadora normóxica (37 °C, CO<sub>2</sub> al 8 %) durante 48 horas. Después de 48 horas, el ADN genómico se recogió de las células usando QuickExtract (Lucigen, Middleton, WI; n.º de cat. QE09050).

Se realizó una PCR para determinar la secuencia de B2M diana y se evaluó la eficacia de corte del ADN amplificado resultante mediante análisis por TIDE. La PCR para las regiones pertinentes se realizó usando Platinum Taq Supermix (Invitrogen, n.º de cat. 125320176 y n.º de cat. 11495017). Las secuencias de los cebadores de PCR se presentan en la Tabla 2; y las condiciones de ciclado se proporcionan en la Tabla 3.

**Tabla 2. Cebadores de TIDE de B2M**

| Nombre | Tipo    | Secuencia (5'-3')     | SEQ ID NO: |
|--------|---------|-----------------------|------------|
| B2MF2  | directo | CAGACAGCAAACCTCACCCAG | 4          |
| B2MR2  | inverso | AAACTTTGTCCCGACCCTCC  | 5          |

**Tabla 3. Parámetros de ciclado de PCR de B2M**

| Etapas            | Temperatura | Tiempo   | Ciclos |
|-------------------|-------------|----------|--------|
| Desnaturalización | 94 °C       | 2 min    | 1      |
| Desnaturalización | 94 °C       | 15 s     | 38     |
| Hibridación       | 55 °C       | 30 s     |        |
| Extensión         | 68 °C       | 45 s     |        |
| Elongación        | 68 °C       | 5 min    | 1      |
| Mantener          | 4           | mantener |        |

Los amplicones resultantes se enviaron para limpieza por PCR y secuenciación de Sanger. Los resultados de la secuenciación de Sanger se introdujeron en el software Tsunami junto con la secuencia guía. Los porcentajes y las identidades de indel se calcularon con el software. A continuación, se seleccionaron ARNg particulares basándose en su frecuencia de

indel en las hPSC. La FIG. 1 muestra la eficiencia de corte de ARNg de B2M-1, B2M-2 y B2M-3.

Se evaluaron los efectos inespecíficos de los ARNg seleccionados en el ADN derivado de células madre usando un análisis de captura de híbridos de los sitios predichos de similitud de secuencia. Las guías de B2M-2 y B2M-3 no mostraron efectos inespecíficos detectables. Se escogió ARNg de B2M-2 para la generación de clones adicional debido a su actividad específica elevada y a su actividad inespecífica indetectable.

**Generación y caracterización de clones de hPSC con KO de B2M.** Usando ARNg de B2M-2, se electroporaron hESC CyT49 y se clasificaron en células individuales 3 días después de la electroporación usando el clasificador WOLF FACS (Nanocollect) en placas de 96 pocillos recubiertas con BIOLAMININA 521 CTG con StemFlex y Revitacell. Se cultivaron células individuales en placa en una incubadora normóxica (37 °C, CO<sub>2</sub> al 8 %) con cambios de medio cada dos días hasta que las colonias fueron lo suficientemente grandes para volver a sembrarse como células individuales. Cuando alcanzaron la confluencia, las muestras se dividieron para mantenimiento y extracción del ADN genómico.

El estado KO de B2M de los clones se confirmó mediante PCR y secuenciación de Sanger. Las secuencias de ADN resultantes de la región B2M diana se alinearon en el software Snapgene para determinar la identidad de los indel y la cigosidad. Los clones con las ediciones deseadas se expandieron y se verificaron adicionalmente a través de evaluación por citometría de flujo para determinar la expresión de B2M (véase la Tabla 4 para consultar la lista de anticuerpos utilizados). Los clones se evaluaron con o sin tratamiento con interferón-gamma (25 ng/ml, R & D Systems, 285-IF). La FIG. 2A muestra la expresión de B2M en células de tipo silvestre y la FIG. 2B presenta la expresión de B2M en células con KO de B2M. El estado cariotípico de los clones se evaluó a través del servicio Cell Line Genetics (Madison, WI) y se publicó el cariotipo normal.

**Tabla 4. Anticuerpos para la citometría de flujo de pluripotencia**

| Antígeno    | Clon    | Fluoróforo | Fabricante    | N.º de catálogo |
|-------------|---------|------------|---------------|-----------------|
| Oct3/4      | 40/3    | Alexa 647  | BD Bioscience | 560329          |
| SOX2        | 030-678 | PE         | BD Bioscience | 562195          |
| B2M         | 2M2     | PE         | Biolegend     | 316305          |
| HLA-ABC     | W6/32   | Alexa 488  | Biolegend     | 311415          |
| mIgG1 kappa | N/A     | PE         | BD Bioscience | 555749          |
| PD-L1       | B7-H1   | Alexa-488  | ThermoFisher  | 53-5983-42      |
| HLA-E       | 3D12    | PE         | ThermoFisher  | 12-9953-42      |

Se confirmó que los clones conservaron la pluripotencia mediante citometría de flujo intracelular para determinar los marcadores de pluripotencia OCT4 y SOX2. Los clones pluripotentes confirmados se diferenciaron en progenitores endocrinos pancreáticos usando métodos previamente establecidos (Schulz *et al.* (2012) PLoS ONE 7(5): e37004).

### **Ejemplo 3: Generación de células madre pluripotentes humanas (hPSC) con KO de B2M/inserción (KI) de PD-L1**

**Diseño de la estrategia de KO de B2M/KI de PD-L1.** El diseño del plásmido para insertar PD-L1 (CD274) en el locus de B2M se realizó de modo que el codón iniciador de B2M se eliminó después de experimentar reparación dirigida por homología (HDR) para insertar PD-L1, anulando cualquier posibilidad de expresión parcial de B2M. La FIG. 3 presenta un esquema del plásmido y la Tabla 5 identifica los elementos y las ubicaciones en el mismo. El plásmido donante contenía un ADNc impulsado por promotor CAGGS de PD-L1 flanqueado por brazos de homología de 800 pares de bases con una secuencia idéntica al locus de B2M alrededor del exón 1. La secuencia completa del plásmido comprende la SEQ ID NO: 33.

**Tabla 5. Elementos del plásmido donante B2M-CAGGS-PD-L1**

| Elemento                             | Ubicación (tamaño en pb) | SEQ ID NO: |
|--------------------------------------|--------------------------|------------|
| ITR izquierda                        | 1-130 (130)              | 6          |
| LHA-B2M                              | 145-944 (800)            | 7          |
| potenciador de CMV                   | 973-1352 (380)           | 8          |
| promotor de $\beta$ -actina de pollo | 1355-1630 (276)          | 9          |
| intrón quimérico                     | 1631-2639 (1009)         | 10         |
| PD-L1                                | 2684-3556 (873)          | 11         |
| señal de poli(A) de bGH              | 3574-3798 (225)          | 12         |
| RHA-B2M                              | 3805-4604 (800)          | 13         |
| ITR derecha                          | 4646-4786 (141)          | 14         |
| Plásmido completo                    | 7133 pb                  | 33         |

El ARNg de B2M-2 se usó para facilitar la inserción del transgén PD-L1 en el locus de B2M diana. El plásmido donante de PD-L1 se introdujo junto con el complejo de RNP formado por el ARNg dirigido a B2M y la proteína Cas9. Por 1 millón de células CyT49 (ViaCyte), se suministraron 4  $\mu$ g de ADN plasmídico junto con la RNP. La electroporación se realizó como se describe en el Ejemplo 2. Siete días después de la electroporación, las células se clasificaron por la expresión superficial de PD-L1 usando el clasificador WOLF FACS (Nanocollect) en placas de 96 pocillos recubiertas con BIOLAMININA 521 CTG con StemFlex con Revitacell. Para la clasificación por FACS, las células sin editar sirvieron como control negativo. Se seleccionaron células positivas en PD-L1 para la clasificación y clonación de células individuales.

Para detectar la expresión superficial de PD-L1, se usaron anticuerpos fluorescentes anti PD-L1 (véase la Tabla 4). Se cultivaron células individuales en placa en una incubadora normóxica (37 °C, CO<sub>2</sub> al 8 %) con cambios de medio cada dos días hasta que las colonias fueron lo suficientemente grandes para volver a sembrarse como células individuales. Cuando alcanzaron la confluencia, las muestras se dividieron para mantenimiento y extracción del ADN genómico.

Los clones dirigidos correctamente se identificaron mediante PCR para la inserción de PD-L1 (KI) usando cebadores que amplifican una región desde fuera de los brazos de homología del plásmido a la inserción de ADNc de PD-L1, lo que permite la amplificación del DNA integrado con KI únicamente. Se sometió a ensayo la inserción específica para determinar la cigosidad mediante PCR para evaluar si la KI se produjo de manera heterocigota u homocigota. Si se identificaba un clon heterocigoto, el alelo negativo para KI se enviaba a

secuenciación de Sanger para verificar que contenía un indel interruptor de B2M en el alelo no KI. Los clones con KI correctos con alteración completa de B2M (ya sea mediante inserción de KI o formación de indel) se expandieron en formatos de cultivo tisular crecientes hasta alcanzar un tamaño de población de 30 millones de células. Se expandieron aproximadamente 10 clones de esta manera, y se confirmó que eran pluripotentes mediante ensayos para determinar OCT4 y SOX2 mediante citometría de flujo intracelular (FIG. 4). Los clones que pasaron los ensayos anteriores se sometieron a continuación a ensayos adicionales para el análisis cariotípico (Cell Line Genetics), como se describe a continuación. Además, los clones se sometieron a ensayo a continuación para determinar su competencia para diferenciarse en precursores endodérmicos pancreáticos (PEC) mediante el protocolo establecido (Schulz *et al.* (2012) PLoS ONE 7(5): e37004), como se describe a continuación. La pérdida de B2M se confirmó además por la falta de expresión de B2M con o sin tratamiento con interferón-gamma (25 ng/ml, R & D Systems, 285-IF) a través de citometría de flujo. Las FIG. 5A y 5B muestran la expresión de PD-L1 en células de tipo silvestre y con KO de B2M/KI de PD-L1, respectivamente.

#### **Ejemplo 4: Análisis del cariotipo de los clones editados**

**Análisis por cariotipado de la banda G de células madre embrionarias (ES) editadas.** Se hicieron pases de 1 millón de células ES editadas a un matraz de cultivo T-25 con medio de cultivo (DMEM/F12+ KSR libre de Xeno al 10 % con 10 ng/ml de activina y 10 ng/ml de heregulina). Después de cultivar durante una noche, se enviaron tres matraces de cultivo T25 al Laboratorio de Citogenética (Cell Line Genetics, Inc.) para el análisis de cariotipado; análisis FISH para los cromosomas 1, 12, 17, 20; y análisis de hibridación genómica comparativa de matrices (aCGH, por sus siglas en inglés) con la matriz patrón de 8x60K. Los resultados de la banda G de las células seleccionadas electroporadas con guías no de corte ("NCG"), clones con KO de B2M y clones con KO de B2M/KI de PD-L1 ("V1-A") se muestran en la Tabla 6.

**Tabla 6. Resultados del cariotipado de la banda G**

| Línea celular | Tipo             | Pase | Análisis de cariotipado | Análisis FISH | Análisis de matrices de aCGH |
|---------------|------------------|------|-------------------------|---------------|------------------------------|
| NCG#1         | guía no de corte | P36  | Normal                  | Normal        | APTA                         |
| NCG#2         | guía no de corte | P36  | Normal                  | Normal        | APTA                         |
| KO de B2M#1   | KO de B2M        | P38  | Normal                  | Normal        | APTA                         |
| KO de B2M#2   | KO de B2M        | P36  | Normal                  | Normal        | APTA                         |

|    |             |                       |     |        |        |      |
|----|-------------|-----------------------|-----|--------|--------|------|
|    | KO de B2M#3 | KO de B2M             | P36 | Normal | Normal | APTA |
|    | V1-A003     | KO de B2M/KI de PD-L1 | P37 | Normal | Normal | APTA |
| 5  | V1-A004     | KO de B2M/KI de PD-L1 | P38 | Normal | Normal | APTA |
|    | V1-A007     | KO de B2M/KI de PD-L1 | P37 | Normal | Normal | APTA |
| 10 | V1-A008     | KO de B2M/KI de PD-L1 | P38 | Normal | Normal | APTA |

**Ejemplo 5: Diferenciación de células madre embrionarias humanas editadas a células del endodermo pancreático (PEC)**

**Mantenimiento de células madre embrionarias humanas (ES) editadas.** Las células madre embrionarias humanas editadas en diversos pases (P38-42) se sembraron en el orden de 33,000 células/cm<sup>2</sup> para un pase de 4 días o 50,000 células/cm<sup>2</sup> para un pase de 3 días con medio hESM (DMEM/F12+KSR al 10 % + 10 ng/ml de activina A y 10 ng/ml de heregulina) y suero AB humano al 10 % final.

**Agregación de células madre embrionarias humanas editadas para la diferenciación de PEC.** Las ES editadas se disociaron en células individuales con ACCUTASE®, a continuación se centrifugaron y se volvieron a suspender en StemPro al 2 % (n.º de cat. A1000701, Invitrogen, CA) en medio DMEM/F12 en el orden de 1 millón de células por ml, y se sembraron un total de 350-400 millones de células en un frasco rotatorio de 850 cm<sup>2</sup> (n.º de cat. 431198, Corning, NY) con una velocidad de rotación de 8 rpm  $\pm$  0.5 rpm durante 18-20 horas antes de la diferenciación. Los agregados de células madre embrionarias humanas editadas se diferenciaron en linajes pancreáticos usando frascos rotatorios como se describe en Schulz *et al.* (2012) PLoS ONE 7(5): e37004.

**Ejemplo 6: Caracterización de células del endodermo pancreático diferenciadas (PEC)**

**Citometría de flujo para FOXA2 y SOX17 en la fase 1 (DE) y CHGA, PDX1 y NKX6.1 en la fase de PEC.** Se lavaron los agregados de la fase 1 derivados de hESC, o los agregados pancreáticos derivados de hESC, con PBS y a continuación se disociaron enzimáticamente en una suspensión de células individuales a 37 °C utilizando ACCUMAX™ (nº de catálogo A7089, Sigma, MO). Se añadió tampón de separación MACS (n.º de cat. 130-091-221, Miltenyi Biotec, Renania del Norte-Westfalia, Alemania) y la suspensión se pasó a través de un filtro de 40  $\mu$ m y se sedimentó. Para la tinción de marcadores intracelulares, se fijaron las células durante 30 min en paraformaldehído al 4 % (p/v), se lavaron en tampón FACS (PBS, BSA al 0.1 % (p/v), NaN<sub>3</sub> al 0.1 % (p/v)) y a continuación las células se permeabilizaron con tampón Perm (PBS,

Triton X-100 al 0.2 % (v/v) (n.º de cat. A16046, Alfa Aesar, MA), suero de burro normal al 5 % (v/v), NaN<sub>3</sub> al 0.1 % (p/v)) durante 30 min en hielo y a continuación se lavaron con tampón de lavado (PBS, BSA al 1 % (p/v), NaN<sub>3</sub> al 0.1 % (p/v)). Las células se incubaron con anticuerpos primarios (Tabla 7) diluidos con tampón de bloqueo (PBS, Triton X-100 al 0.1 % (v/v), suero de burro normal al 5 % (v/v), NaN<sub>3</sub> al 0.1 % (p/v)) durante una noche a 4 °C. Las células se lavaron en tampón IC y a continuación se incubaron con anticuerpos secundarios apropiados durante 60 minutos a 4 °C. Las células se lavaron en tampón IC y a continuación en tampón FACS. Los datos de citometría de flujo se adquirieron con el citómetro de flujo NovoCyte (ACEA Biosciences, Bruselas). Los datos se analizaron usando el software FlowJo (Tree Star, Inc.). Las células intactas se identificaron basándose en la dispersión de luz hacia adelante (ángulo bajo) y lateral (ortogonal, 90°). El fondo se estimó usando controles de anticuerpos y células indiferenciadas. En las figuras, se muestra un gráfico de citometría de flujo representativo para una de las subpoblaciones. Los números presentados en las figuras representan el porcentaje de células totales de la ventana de flujo de las células intactas.

**Tabla 7. Anticuerpos para la citometría de flujo para la caracterización de PEC diferenciadas**

| Antígeno | Fluoróforo | Fuente                                           | Dilución |
|----------|------------|--------------------------------------------------|----------|
| SOX17    | AF647      | BD Bioscience (n.º de catálogo 562594)           | 1:50     |
| FOXA2    | PE         | Miltenyi Biotechnology (n.º de cat. 130-107-773) | 1:10     |
| PDX1     | PE         | BD Bioscience (n.º de cat. 562161)               | 1:2.5    |
| NKX6.1   | AF647      | BD Bioscience (n.º de cat. 563338)               | 1:2.5    |
| CHGA     | AF405      | Novus (n.º de cat. NBP2-33198AF405)              | 1:1000   |

En la fase de DE, la población de células doble positivas para FOXA2 y SOX17 fue más del 90 % del total de las células diferenciadas a partir de los tipos silvestres de CyT49. Las células con KI de PD-L1/KO de B2M y KO de B2M mostraron un porcentaje comparable de DE en comparación con las células de tipo silvestre (FIG. 6 y FIG. 7).

En la fase de PEC, se realizó una citometría de flujo para la cromogranina (CHGA), PDX1 y NKX6.1. La población heterogénea en la fase de PEC incluye progenitores pancreáticos y células endocrinas tempranas (FIG. 8). A la luz del gráfico circular de la población heterogénea (FIG. 9), la distribución de las poblaciones celulares de las células editadas diferenciadas (KI de PD-L1/KO de B2M o KO de B2M) fue muy similar a las células de tipo silvestre.

**Secuenciación de ARN dirigida.** La secuenciación de ARN dirigida para el análisis de la expresión génica se realizó usando Illumina TruSeq y un panel personalizado de oligos dirigidos a 111 genes. El panel contenía principalmente genes que son marcadores de las fases de desarrollo durante la diferenciación pancreática. Al final de cada fase de diferenciación, se

recogieron y extrajeron 10 µl de APV (volumen de sedimentos agregados) usando el protocolo de columna de centrifugación Qiagen RNeasy o RNeasy 96, incluyendo el tratamiento con ADNasa sobre la columna. La cuantificación y el control de calidad se realizaron usando TapeStation combinado con Qubit o usando Qiagen QIAxcel. Se procesaron 50-200 ng de ARN de acuerdo con el protocolo de preparación de genotecas Illumina TruSeq, que consistía en la síntesis de ADNc, la hibridación del grupo de oligonucleótidos personalizado, el lavado, la extensión, la ligadura de los oligonucleótidos unidos, la amplificación por PCR de las genotecas y la limpieza de las genotecas, antes de la cuantificación y el control de calidad de las genotecas de ADNbc resultantes usando TapeStation combinado con Qubit o usando Qiagen QIAxcel. Las genotecas se diluyeron posteriormente a una concentración de 4 nM y se combinaron, seguido de la desnaturalización, la adición de control PhiX y una dilución adicional a 10-12 pM antes de cargarlas en el secuenciador Illumina MiSeq. Después de la ejecución de la secuenciación, se realizó automáticamente el análisis de datos inicial a través de BaseSpace, generando recuentos de lectura sin procesar para cada una de las sondas personalizadas. Para cada gen, se sumaron estos recuentos de lectura para todas las sondas correspondientes a ese gen, con la adición de 1 recuento de lectura (para evitar divisiones entre 0 cadena abajo). Se realizó la normalización para el gen SF3B2, y las lecturas se visualizaron normalmente como factor de cambio frente a la fase 0. Cuando los datos se procesaron para el análisis de componentes principales, la normalización se realizó usando el método DEseq.

La expresión de genes seleccionados se mostró en la FIG. 10. El patrón de expresión cinética de *FOXA2*, *CHGA*, *PDX1* y *NKX6.1* a partir de células con KI de PD-L1/KO de B2M o KO de B2M fue similar al de las células de tipo silvestre.

**Confirmación de la expresión de B2M y PD-L1 en la fase de PEC.** En la fase de PEC, se estimularon agregados diferenciados con o sin interferón-gamma (50 ng/ml) durante 48 horas. Los agregados se lavaron con PBS y a continuación se disociaron enzimáticamente en una suspensión de células individuales a 37 °C usando ACCUMAX™ (n.º de catálogo A7089, Sigma, MO). Se añadió tampón de separación MACS (n.º de cat. 130-091-221, Miltenyi Biotec, Renania del Norte-Westfalia, Alemania) y la suspensión se pasó a través de un filtro de 40 µm y se sedimentó. Para la tinción del marcador de superficie, se incubaron las células disociadas con anticuerpo conjugado con fluorescencia diluido en tampón de separación MACS durante 20 minutos y a continuación se lavaron en tampón de separación MACS. Las células se volvieron a suspender en tampón FACS para la adquisición del flujo. Los datos de citometría de flujo se adquirieron con el citómetro de flujo NovoCyte. Como se muestra en las FIG. 11A-11F, la expresión de B2M estaba por debajo del límite de detección en las PEC diferenciadas a partir de KO de B2M (FIG. 11B) o KI de PD-L1/KO de B2M (FIG. 11C) y se expresó PD-L1 en las PEC diferenciadas a partir de KI de PD-L1/KO de B2M (FIG. 11F). En general, más de

aproximadamente el 90 % de las PEC expresaron PD-L1, lo que indica una población homogénea de células. Con frecuencia, se produce una pérdida de expresión del transgén con el tiempo después de la diferenciación de las células madre editadas génicamente (Hong *et al.*, Mol. Ther., 2017, 25(1):44-53).

**Fenotipo inmunitario de las células PEC.** En la fase de PEC, se estimularon agregados diferenciados con o sin interferón-gamma (50 ng/ml) durante 48 horas. Se recogieron los agregados para tinción con MHC de clase I y II. No hay expresión de MHC de clase II en la fase de PEC a partir de células de tipo silvestre o editadas (KI de PD-L1/KO de B2M y KO de B2M) (FIG. 12D-12F). La expresión de HLA-ABC (MHC de clase I) fue baja (1.3 % a partir de células de tipo silvestre) y fue regulada a un nivel alto mediante la estimulación con IFN- $\gamma$ . Sin embargo, HLA-ABC no se expresó ni siquiera con estimulación con IFN- $\gamma$  en las células editadas (KI de PD-L1/KO de B2M y KO de B2M) (FIG. 12A-12C).

**Ejemplo 7: Generación de células madre pluripotentes humanas (hPSC) con KO de TXNIP**

**Selección de ARN guía (ARNg) para TXNIP.** Se diseñaron diez ARNg dirigidos a TXNIP para dirigirse al exón 1 y al exón 2 de la secuencia codificante de TXNIP (Tabla 8). Las secuencias PAM se presentan en negrita en las secuencias diana presentadas en la Tabla 8, y las secuencias de ADN correspondientes a las secuencias guía se presentan en la Tabla 8. Estos ARNg tenían bajas puntuaciones inespecíficas predichas basadas en la predicción de la homología de secuencia usando el software de diseño de ARNg.

**Tabla 8. Secuencias diana de TXNIP y secuencias de ARNg seleccionadas**

| Nombre              | Secuencia diana (5'-3')<br>(secuencia PAM en<br>negrita) | SE<br>Q<br>ID<br>NO<br>: | Versión de ADN de la secuencia<br>guía (5'-3') | SE<br>Q<br>ID<br>NO<br>: |
|---------------------|----------------------------------------------------------|--------------------------|------------------------------------------------|--------------------------|
| TXNIP_Exón<br>1_T1  | GAAGCGTGTCTTCATAG<br>CGC <b>AGG</b>                      | 45                       | GAAGCGTGTCTTCATAGCGC                           | 15                       |
| TXNIP_Exón<br>1_T21 | TTACTCGTGTCAAAGCC<br>GTT <b>AGG</b>                      | 46                       | TTACTCGTGTCAAAGCCGTT                           | 16                       |
| TXNIP_Exón<br>1_T22 | TGTCAAAGCCGTTAGGA<br>TCCT <b>TGG</b>                     | 47                       | TGTCAAAGCCGTTAGGATCC                           | 17                       |
| TXNIP_Exón<br>1_T23 | GCCGTTAGGATCCTGGC<br>TTG <b>CGG</b>                      | 48                       | GCCGTTAGGATCCTGGCTTG                           | 18                       |
| TXNIP_Exón<br>1_T25 | GCGGAGTGGCTAAAGT<br>GCTTT <b>TGG</b>                     | 49                       | GCGGAGTGGCTAAAGTGCTT                           | 19                       |

|   |                    |                                      |    |                       |    |
|---|--------------------|--------------------------------------|----|-----------------------|----|
| 5 | TXNIP_Exón<br>1_T5 | TCCGCAAGCCAGGATCC<br>TAAC <b>GG</b>  | 50 | TCCGCAAGCCAGGATCCTAA  | 20 |
|   | TXNIP_Exón<br>2_T4 | GTTCCGGCTTTGAGCTTC<br>CTC <b>AGG</b> | 51 | GTTCCGGCTTTGAGCTTCCTC | 21 |
|   | TXNIP_Exón<br>2_T2 | GAGATGGTGATCATGAG<br>ACCT <b>TGG</b> | 52 | GAGATGGTGATCATGAGACC  | 22 |
|   | TXNIP_Exón<br>2_T1 | TTGTACTCATATTTGTTT<br>CC <b>AGG</b>  | 53 | TTGTACTCATATTTGTTTCC  | 23 |
|   | TXNIP_Exón<br>2_T3 | AACAAATATGAGTACAA<br>GTT <b>CGG</b>  | 54 | AACAAATATGAGTACAAGTT  | 24 |

**Generación y caracterización de clones de hiPSC KO de TXNIP.** Para evaluar la eficiencia de corte de estos ARNg en las hiPSC, se electroporaron células TC1133 hiPSC usando el electroporador Neon Electroporator (sistema de transfección Neon Transfection System de ThermoFisher n.º de cat. MPK5000) con una mezcla de RNP de proteína Cas9 (Biomay) y ARN guía (Synthego) en una relación molar de 3:1 (ARNg:Cas9) con valores absolutos de 125 pmol de Cas9 y 375 pmol de ARNg. Para formar el complejo de RNP, se combinaron ARNg y Cas9 en un recipiente con tampón R hasta un volumen total de 25 µl y se incubaron durante 15 min a TA. Las células se disociaron usando ACCUTASE®, a continuación se volvieron a suspender en medio DMEM/F12 (Gibco, n.º de cat. 11320033), se contaron usando un NC-200 (Chemometec) y se centrifugaron. Se volvió a suspender un total de  $1 \times 10^6$  células con el complejo de RNP y se añadió tampón R hasta un volumen total de 125 µl. A continuación, esta mezcla se sometió a electroporación usando los parámetros: 2 pulsos, 30 ms, 1100 V. Después de la electroporación, las células se pipetearon en un tubo Eppendorf lleno de medio StemFlex media con RevitaCell. A continuación, esta suspensión celular se sembró en placas de cultivo tisular recubiertas previamente con BIOLAMININA 521 CTG. Las células se cultivaron en una incubadora normóxica (37 °C, CO<sub>2</sub> al 8 %) durante 48 horas. Después de 48 horas, el ADN genómico se recogió de las células usando QuickExtract.

Se realizó una PCR para la secuencia TXNIP diana y el ADN amplificado resultante se secuenció por Sanger. Se usó el análisis por TIDE para analizar los datos de secuenciación de salida para los porcentajes de indel usando el software Tsunami. La FIG. 13 muestra la eficiencia de corte para los ARNg de TXNIP. A continuación, se seleccionaron ARNg particulares basándose en su frecuencia de indel en las hPSC.

Se evaluaron los efectos inespecíficos de los ARNg con mayor eficiencia de corte en el ADN derivado de células madre usando un análisis de captura de híbridos de los sitios predichos de similitud de secuencia. Se realizaron experimentos adicionales con ARNg T5 de TXNIP ya que no mostró efectos inespecíficos detectables y demostró una actividad específica elevada.

**Generación y caracterización de clones de hPSC con KO de TXNIP.** Usando ARNg T5 de TXNIP, se electroporaron hESC CyT49 (Viacyte) y se clasificaron en células individuales 3 días después de la electroporación usando el clasificador WOLF FACS (Nanocollect) en placas de 96 pocillos BIOLAMININA 521 CTG con StemFlex y Revitacell. Se cultivaron células individuales en placa en una incubadora normóxica (37 °C, CO<sub>2</sub> al 8 %) con cambios de medio cada dos días hasta que las colonias fueron lo suficientemente grandes para volver a sembrarse como células individuales. Cuando alcanzaron la confluencia, las muestras se dividieron para mantenimiento y extracción del ADN genómico.

El estado KO de TXNIP de los clones se confirmó mediante PCR y secuenciación de Sanger. Las secuencias de ADN resultantes de la región TXNIP diana se alinearon en el software Snapgene para determinar la identidad de los indel y la cigosidad. Los clones con las ediciones deseadas se expandieron y se verificaron adicionalmente a través de evaluación por citometría de flujo para determinar la expresión de TXNIP. El estado cariotípico de los clones se evaluó a través del servicio Cell Line Genetics y se publicó el cariotipo normal (Tabla 9).

**Tabla 9. Análisis del cariotipo**

| Línea celular  | Pase | Análisis de cariotipado | Análisis FISH | Análisis de matrices de aCGH |
|----------------|------|-------------------------|---------------|------------------------------|
| KO de TXNIP#2  | P31  | Normal                  | Normal        | APTA                         |
| KO de TXNIP#13 | P31  | Normal                  | Normal        | APTA                         |

Se confirmó que los clones conservaron la pluripotencia mediante citometría de flujo intracelular para determinar los marcadores de pluripotencia OCT4 y SOX2. Los clones pluripotentes confirmados se diferenciaron en progenitores endocrinos pancreáticos usando métodos previamente establecidos (Schulz *et al.* (2012) PLoS ONE 7(5): e37004).

La secuenciación de ARN dirigida para el análisis de la expresión génica se realizó usando Illumina TruSeq y un panel personalizado de oligos, como se ha descrito anteriormente. La expresión de genes seleccionados se mostró en la FIG. 20. El patrón de expresión cinética de *FOXA2*, *CHGA*, *PDX1* y *NKX6.1* a partir de células con KO de TXNIP fue similar al de las células de tipo silvestre. En la fase de PEC, también se realizó una citometría de flujo para la cromogranina (CHGA), PDX1 y NKX6.1. La población heterogénea en la fase de PEC incluía un 30.6 % de células progenitoras pancreáticas (es decir, CHGA<sup>+</sup>/NKX6.1<sup>+</sup>/PDX1<sup>+</sup>) (FIG. 21).

**Ejemplo 8: Generación de Células Madre Pluripotentes Humanas (hPSC) con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E**

**Diseño de la estrategia de KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E.** Se generaron células en las que se insertó la secuencia codificante de PD-L1 en el locus de B2M (desactivando de este modo el gen B2M) y la secuencia codificante de HLA-E en el locus de TXNIP (desactivando de este modo el gen TXNIP).

El diseño del plásmido para insertar PD-L1 (CD274) en el locus de B2M se representó en el Ejemplo 3. El plásmido donante contiene un ADNc impulsado por promotor CAGGS de PD-L1 flanqueado por brazos de homología de 800 pares de bases con una secuencia idéntica al locus de B2M alrededor del exón 1. El ARNg de B2M-2 se usó para facilitar la inserción del transgén PD-L1 en el locus de B2M diana. El plásmido donante de PD-L1 se introdujo junto con el complejo de RNP formado por el ARNg dirigido a B2M y la proteína Cas9. Por 1 millón de células CyT49 (ViaCyte), se suministraron 4 µg de ADN plasmídico junto con la RNP. La electroporación se realizó como se describe en el Ejemplo 2. Siete días después de la electroporación, las células se enriquecieron en células positivas para PD-L1 mediante clasificación celular asistida magnéticamente (MACS) usando reactivos Miltenyi (Microperlas de anti IgG de ratón n.º de cat. 130-048-401, Columnas LS n.º de cat. 130-042-401, y separador MidiMACS n.º de cat. 130-042-302) o reactivos Thermofisher (Imán DynaMag™-15 n.º de cat. 12301D, kit de IgG de ratón CELLection™ Pan n.º de cat. 11531D, IgG de ratón Dynabeads™ Pan n.º de cat. 11042).

Después de expandir la población positiva para PD-L1 enriquecida, se insertó un transgén de ADNc de trímero de HLA-E en el locus genómico de TXNIP mediante HDR inducida por CRISPR usando un plásmido donante que comprendía la secuencia de HLA-E. El ADNc del trímero de HLA-E estaba compuesto por un péptido de señal B2M fusionado a un péptido de presentación de HLA-G fusionado a la proteína de membrana B2M fusionada a la proteína HLA-E sin su péptido de señal. Este diseño de trímero se ha publicado anteriormente (Gornalusse *et al.* (2017) *Nat. Biotechnol.* 35(8): 765-772). La secuencia codificante del trímero de HLA-E (incluyendo los enlazadores) es la SEQ ID NO: 55 (es decir, las SEQ ID NO: 26-31). El plásmido donante para el suministro de HLA-E contiene un promotor CAGGS que impulsa la expresión del trímero de HLA-E flanqueado por brazos de homología de 800 pares de bases con una secuencia idéntica al locus de TXNIP alrededor del exón 1 (FIG. 14, Tabla 10 y Tabla 11). En algunas realizaciones, el plásmido donante comprende las SEQ ID NO: 34 o 56.

**Tabla 10. Elementos del plásmido 1 donante TXNIP-CAGGS-HLA-E**

| Elemento      | Ubicación (tamaño en pb) | SEQ ID NO: |
|---------------|--------------------------|------------|
| ITR izquierda | 1-130 (130)              | 6          |
| LHA-TXNIP     | 145-944 (800)            | 25         |

|    |                                      |                  |    |
|----|--------------------------------------|------------------|----|
| 5  | potenciador de CMV                   | 973-1352 (380)   | 8  |
|    | promotor de $\beta$ -actina de pollo | 1355-1630 (276)  | 9  |
|    | intrón quimérico                     | 1631-2639 (1009) | 10 |
|    | secuencia señal de B2M               | 2684-2743 (60)   | 26 |
|    | péptido HLA-G                        | 2744-2770 (27)   | 27 |
| 10 | Enlazador GS                         | 2771-2815 (45)   | 28 |
|    | proteína de membrana B2M             | 2816-3112 (297)  | 29 |
|    | Enlazador GS                         | 3113-3172 (60)   | 30 |
|    | HLA-E                                | 3173-4183 (1011) | 31 |
|    | señal de poli(A) de bGH              | 4204-4428 (225)  | 12 |
| 15 | RHA-TXNIP                            | 4435-5234 (800)  | 32 |
|    | ITR derecha                          | 5276-5416 (141)  | 14 |
|    | Plásmido completo                    | 7763 pb          | 34 |

**Tabla 11. Elementos del plásmido 2 donante TXNIP-CAGGS-HLA-E**

|    |                                      |                                 |                   |
|----|--------------------------------------|---------------------------------|-------------------|
| 15 | <b>Elemento</b>                      | <b>Ubicación (tamaño en pb)</b> | <b>SEQ ID NO:</b> |
| 20 | ITR izquierda                        | 1-130 (130)                     | 6                 |
|    | LHA-TXNIP                            | 145-944 (800)                   | 25                |
|    | potenciador de CMV                   | 973-1352 (380)                  | 8                 |
|    | promotor de $\beta$ -actina de pollo | 1355-1630 (276)                 | 9                 |
|    | intrón quimérico (truncado)          | 1631-2336 (706)                 | 57                |
| 25 | secuencia señal de B2M               | 2381-2440 (60)                  | 26                |
|    | péptido HLA-G                        | 2441-2467 (27)                  | 27                |
|    | Enlazador GS                         | 2468-2512 (45)                  | 28                |
|    | proteína de membrana B2M             | 2513-2809 (297)                 | 29                |
|    | Enlazador GS                         | 2810-2869 (60)                  | 30                |
| 30 | HLA-E                                | 2870-3880 (1011)                | 31                |
|    | señal de poli(A) de bGH              | 3901-4125 (225)                 | 12                |
|    | RHA-TXNIP                            | 4132-4931 (800)                 | 32                |
|    | ITR derecha                          | 4973-5113 (141)                 | 14                |
|    | Plásmido completo                    | 7460 pb                         | 56                |

El ARNg de TXNIP-T5 se usó para facilitar la inserción del transgén HLA-E en el locus de TXNIP diana. El plásmido donante de HLA-E se introdujo junto con el complejo de RNP formado por el ARNg de TXNIP-T5 y la proteína Cas9. Por cada 1 millón de células PD-L1+, se suministraron 4  $\mu$ g de ADN del plásmido donante de HLA-E (SEQ ID NO: 56) junto con la RNP.

Como alternativa, puede usarse ADN de plásmido donante de HLA-E (SEQ ID NO: 34). La electroporación se realizó como se describe en el Ejemplo 2. Siete días después de la electroporación, las células se enriquecieron en células positivas para HLA-E mediante MACS usando reactivos Miltenyi o reactivos Thermofisher. Después de el enriquecimiento en HLA-E, las células se clasificaron individualmente usando el clasificador FACS WOLF (Nanocollect) en placas de 96 pocillos recubiertas con BIOLAMININA 521 CTG con StemFlex y Revitacell. Se cultivaron células individuales en placa en una incubadora normóxica (37 °C, CO<sub>2</sub> al 8 %) con cambios de medio cada dos días hasta que las colonias fueron lo suficientemente grandes para volver a sembrarse como células individuales. Cuando alcanzaron la confluencia, las muestras se dividieron para mantenimiento y extracción del ADN genómico. Los anticuerpos anti-PD-L1 y anti-HLA-E (Tabla 4) se usaron para el enriquecimiento por MACS y la clasificación por FACS en placas de 96 pocillos con la selección ajustada para células doblemente positivas para HLA-E y PD-L1. Para la clasificación por FACS, las células sin editar sirvieron como control negativo.

Los clones dirigidos correctamente se identificaron mediante PCR para la inserción de KI de PD-L1 y la inserción de KI de HLA-E usando cebadores que amplifican una región desde fuera de los brazos de homología del plásmido hasta la inserción de ADNc de PD-L1 o la inserción de ADNc de HLA-E, respectivamente, permitiendo la amplificación del ADN integrado en KI únicamente. Se sometió a ensayo la inserción específica para determinar la cigosidad mediante PCR para evaluar si la KI se produjo de manera heterocigota u homocigota. Si se identificaba un clon heterocigoto, el alelo negativo para KI se enviaba a secuenciación de Sanger para verificar que contenía un indel interruptor de B2M o un indel interruptor de TXNIP, respectivamente. Los clones con KI correctos con interrupciones completas de B2M y TXNIP (ya sea mediante inserción de KI o formación de indel) se expandieron en formatos de cultivo tisular crecientes hasta alcanzar un tamaño de población de 30 millones de células. Se expandieron aproximadamente 10 clones de esta manera, y se confirmó que eran pluripotentes mediante ensayos para determinar OCT4 y SOX2 mediante citometría de flujo intracelular (FIG. 15).

Los clones que pasaron los ensayos anteriores se sometieron a continuación a ensayos adicionales para el análisis cariotípico (Cell Line Genetics), como se ha descrito anteriormente. En la Tabla 12 se muestran los resultados de la banda G de los clones con KO de B2M/KI de PD-L1 + KO de TXNIP/KI de HLA-E seleccionados ("V1-B"). Además, los clones V1-B se sometieron a ensayo a continuación para determinar su competencia para diferenciarse en precursores del endodermo pancreático (PEC).

Tabla 12. Resultados de la banda G

| Línea celular | Pase | Análisis de cariotipado | Análisis FISH | Análisis de matrices de aCGH |
|---------------|------|-------------------------|---------------|------------------------------|
| V1-B003       | P37  | Normal                  | Normal        | APTA                         |
| V1-B007       | P37  | Normal                  | Normal        | APTA                         |
| V1-B008       | P36  | Normal                  | Normal        | APTA                         |

PD-L1 y HLA-E continuaron expresándose después de la diferenciación en células en Fase 6 según el protocolo endocrino pancreático publicado anteriormente (Rezania *et al.* (2014) Nat. Biotechnol. 32(11): 1121-1133) (FIG. 16). La población de células diferenciadas es homogénea en términos de la expresión del transgén, por ejemplo, el 94.4 % de las células expresan PD-L1 y el 97.0 % de las células expresan HLA-E. La FIG. 22A muestra una morfología similar de las diversas células clonadas ("S6-V1B-H9", "S6-V1B-3B11", "S6-V1B-1G7" y "S6-V1B-3C2") diferenciadas a la Fase 6 en comparación con las células de tipo silvestre y las células de control con guía no de corte. La expresión de genes seleccionados de clones con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E se muestra en las FIG. 23A-FIG. 23F. El patrón de expresión cinética de *INS*, *NKX6.1*, *GCK*, *GCG* y *SST* a partir de células clonadas con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E fue similar al de las células de tipo silvestre (FIG. 23A). Los niveles de expresión de los marcadores de Fase 6 *INS* (FIG. 23B), *NKX6.1* (FIG. 23C), *GCG* (FIG. 23D), *SST* (FIG. 23E) y *GCK* (FIG. 23F) a partir de diversos clones con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E diferenciados ("S6-V1B-H9", "S6-V1B-3B11", "S6-V1B-1G7" y "S6-V1B-3C2") fueron similares a los niveles en células de tipo silvestre en Fase 6 e islotes de tipo silvestre. Como control negativo se usó un clon con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E indiferenciado ("ES-V1B-H9").

Las FIG. 24A-24B muestran la evaluación por citometría de flujo de la expresión de *INS* y *GCG* (FIG. 24A) y de la expresión de *INS* y *NKX6.1* (FIG. 24B) en células en fase 6 diferenciadas a partir de un clon con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E. Las FIG. 25A-25B muestran el porcentaje de expresión de *INS* (FIG. 25A) y de expresión de *NKX6.1* (FIG. 25B) en las células en Fase 6 diferenciadas a partir de dos clones con KO de B2M/KI de PD-L1 y KO de TXNIP/KI de HLA-E ("S6-V1B003" y "V1B-H9"). La expresión en ambas fue similar a la de las células de control de tipo silvestre y de guía no de corte.

En la fase de PEC, se realizó una citometría de flujo para la cromogranina (CHGA), PDX1 y *NKX6.1*. La población heterogénea en la fase de PEC incluye progenitores pancreáticos, endocrinos tempranos (FIG. 17). La secuenciación de ARN dirigida para el análisis de la expresión génica se realizó como se ha descrito anteriormente. La expresión de genes seleccionados para el clon con KO de TXNIP se muestra en la FIG. 18A y la expresión de genes

seleccionados para el clon V1-B se muestra en la FIG. **18B**. El patrón de expresión cinética de *FOXA2*, *CHGA*, *PDX1* y *NKX6.1* a partir de células clonadas V1-B o con KO de TXNIP fue similar al de las células de tipo silvestre.

Se generaron células en las que se insertó la secuencia codificante de HLA-E en el locus de TXNIP (desactivando de este modo el gen TXNIP) usando el vector donante de HLA-E que comprende la secuencia de nucleótidos de la SEQ ID NO: 56. La secuenciación de ARN dirigida para el análisis de la expresión génica se realizó como se ha descrito anteriormente. La expresión de genes seleccionados para el clon con KO de TXNIP/KI de HLA-E se muestra en la FIG. **28**. El patrón de expresión cinética de *FOXA2*, *CHGA*, *PDX1* y *NKX6.1* a partir de células con KO de TXNIP/KI de HLA-E fue similar al de las células de tipo silvestre.

Como alternativa, se generaron células en las que se insertó la secuencia codificante de HLA-E en el locus de TXNIP usando el vector donante de HLA-E que comprende la secuencia de nucleótidos de la SEQ ID NO: 34. Se diferenciaron células editadas en masa a la fase de PEC y expresaron HLA-E en al menos el 75 % de la población de células (datos no mostrados). La evaluación por citometría de flujo de la expresión de *PDX1* y *NKX6.1* en células PEC diferenciadas a partir de células con KO de TXNIP fue similar a la de células PEC diferenciadas a partir de células de tipo silvestre (datos no mostrados).

#### **Ejemplo 9: Ensayo de activación/proliferación de linfocitos T**

Se sometieron a ensayo células diferenciadas en PEC para determinar su capacidad para desencadenar una respuesta inmunitaria mediante ensayos de activación/proliferación de linfocitos T humanos *in vitro*. Se adquirieron PBMC donantes nuevas en Hemacare y se purificaron los linfocitos T CD3+ usando el kit de aislamiento de linfocitos T Pan, humano (Miltenyi n.º de cat. 130-096-535). Los linfocitos T aislados se marcaron con el protocolo del kit de proliferación celular CellTrace™ CFSE (ThermoFisher n.º de cat. C34554) según las instrucciones del fabricante, y se incubaron conjuntamente con PEC diferenciadas durante 5 días. Se usaron perlas Dynabeads™ Human T-Activator CD3/CD28 para la expansión y activación de los linfocitos T (ThermoFisher n.º de cat. 11161D) como control positivo para activar los linfocitos T. Los linfocitos T solos se marcaron con CFSE y se usaron como control negativo. Se midió el porcentaje de células CD3+ CFSE+ para evaluar el porcentaje de proliferación de linfocitos T (FIG. **19A-19B**). Las PEC WT generaron la proliferación de linfocitos T por encima del control de linfocitos T solo. Las PEC derivadas de CyT49 con KO de B2M, KO de B2M/KI de PD-L1 y KO de B2M/KI de PD-L1 + KO de TXNIP/KI de HLA-E no desencadenaron la proliferación de linfocitos T por encima del control de linfocitos T únicamente, lo que demuestra la naturaleza hipoinmunogénica de las células editadas.

#### **Ejemplo 10: Estudio de la eficacia in vivo de líneas clonales dirigidas a genes**

Se produjeron células del endodermo pancreático a partir de líneas celulares clonales hES derivadas de CyT49 con las siguientes modificaciones genéticas: 1) la supresión dirigida de la expresión de B2M y la expresión forzada de PD-L1, 2) la supresión dirigida de la expresión de B2M y la expresión forzada de HLA-E o 3) la supresión dirigida de TXNIP. Además, se obtuvo una línea celular no modificada clonal a partir de la transfección con un ARN-de guía no de corte (NCG).

Siguiendo los procedimientos convencionales, los agregados de endodermo pancreático derivados de las líneas clonales indicadas se cargaron en dispositivos perforados (DP) para producir artículos de ensayo o de control. Los DP permiten la vascularización directa tras el trasplante subcutáneo, y las células progenitoras pancreáticas encapsuladas maduran *in vivo* en células endocrinas pancreáticas funcionales, incluidas las células productoras de insulina sensibles a la glucosa.

Como se resume en la Tabla 13, a cinco grupos de ratas desnudas atímicas se les implantaron por vía subcutánea con dos artículos, cada uno de los cuales contenía aproximadamente  $7 \times 10^6$  células de endodermo pancreático obtenidas a partir de diferenciaciones de las cuatro líneas clonales descritas anteriormente, o células hES CyT49 de tipo silvestre (ViaCyte).

**Tabla 13. Diseño del estudio**

| Número de grupo | ID del grupo | Origen de hESC   | Modificación genética      |                         | Número de Animales | Punto final |
|-----------------|--------------|------------------|----------------------------|-------------------------|--------------------|-------------|
|                 |              |                  | Desactivación (Pérdida de) | Inserción (Ganancia de) |                    |             |
| 1               | Control      | CyT49            | Ninguno                    | Ninguno                 | 6 por grupo        | 20 Semanas  |
| 2               | NCG          | Subclón de CyT49 | Ninguno                    | Ninguno                 |                    |             |
| 3               | KO de TXNIP  | Subclón de CyT49 | TXNIP                      | Ninguno                 |                    |             |
| 4               | KO de B2M/   | Subclón de CyT49 | B2M                        | PD-L1                   |                    |             |
| 5               | KO de B2M/   | Subclón de CyT49 | B2M                        | HLA-E                   |                    |             |

A partir de las 12 semanas, todos los animales supervivientes se sometieron a evaluación de la eficacia mediante pruebas de secreción de insulina estimulada por glucosa (GSIS, por sus siglas en inglés). Se obtuvieron muestras de sangre de animales que no habían ayunado antes y después de la administración intraperitoneal de glucosa 3 g/kg. Las concentraciones séricas de péptido C humano se determinaron mediante ensayos inmunoabsorbentes ligados a enzimas convencionales.

El ensayo de GSIS se realizó a las 12, 16 y 20 semanas. Los resultados indicaron que no hubo diferencias sustanciales entre los grupos experimentales, especialmente más allá del punto temporal de las 12 semanas. En comparación con los niveles de péptido C detectados en el grupo de control (Grupo 1, <40 pM a 2.0 nM, media 1.1 nM) los niveles de péptido C fueron elevados en 2 de los 6 animales del Grupo 3 (KO de TXNIP, media 1.5 nM). Los otros grupos, el Grupo 2 (NCG, media 0.5 nM), el Grupo 4 (KO de B2M/KI de PD-L1, media 0.5 nM) y el Grupo 5 (KO de B2M/KI de HLA-E, media 0.4 nM), presentaron un intervalo similar de niveles de péptido C en comparación con el grupo de control, pero con más animales cerca del extremo inferior del intervalo. Sin embargo, estas diferencias no fueron estadísticamente significativas. Estos resultados indicaron que ni las modificaciones genéticas que se introdujeron ni las manipulaciones necesarias para generar líneas clonales afectaron a la capacidad de las líneas celulares en cuestión para diferenciarse en células del endodermo pancreático *in vitro* y posteriormente generar células beta funcionales *in vivo*.

A las 20 semanas, después del ensayo de GSIS, los animales se sacrificaron y los artículos de ensayo explantados se fijaron en formol tamponado neutro, se procesaron en portaobjetos y se tiñeron con HyE y mediante inmunohistoquímica para determinar la insulina y el glucagón.

Las evaluaciones de eficacia *in vivo* a través de ensayos de GSIS no mostraron diferencias sustanciales entre los artículos de control no editados y los artículos de ensayo editados formulados con células del endodermo pancreático derivadas de líneas celulares clonales, cada una de ellas portadora de un subconjunto de modificaciones genéticas. Los resultados sugieren que las modificaciones genéticas individuales y el proceso por el que se introducen pueden ser tolerados *in vivo*.

**Ejemplo 11. Estudio de eficacia *in vivo* de líneas celulares con KO de B2M/KI de PD-L1, KO de TXNIP/KI de HLA-E**

Se generaron cuatro líneas clonales esencialmente como se ha descrito anteriormente en el Ejemplo 8 y se cargaron en dispositivos perforados para formar artículos de ensayo. Los artículos de control contenían células CyT49 no modificadas (ViaCyte). Se implantaron por vía subcutánea artículos que comprendían aproximadamente  $7 \times 10^6$  células del endodermo pancreático en ratas desnudas atímicas (2 artículos/rata, 8 ratas/grupo).

A las 12, 16, 20 y 24 semanas, todos los animales supervivientes se sometieron a ensayos de secreción de insulina estimulada por glucosa (GSIS). Se obtuvieron muestras de sangre de animales en ayunas antes y después de la administración intraperitoneal de glucosa 3 g/kg. Las concentraciones séricas de péptido C humano se determinaron mediante ensayos inmunoabsorbentes ligados a enzimas convencionales. Se detectó péptido C en suero en la mayoría de los animales 12 semanas después del implante. Los niveles de péptido C en suero

16, 20 y 24 semanas después del implante se presentan en la Tabla 14. No se observaron diferencias estadísticamente significativas entre los grupos de animales implantados con células editadas génicamente frente a los de control.

**Tabla 14. Niveles de péptido C en suero *in vivo*.**

| 5  | Grupos                                    | Péptido C en suero (pmol) |               |                 |
|----|-------------------------------------------|---------------------------|---------------|-----------------|
|    |                                           | Media                     | Inferior 95 % | Superior (95 %) |
| 10 | Células editadas génicamente - 16 semanas | 729                       | 40            | 1418            |
|    | Células editadas génicamente - 20 semanas | 1080                      | 391           | 1769            |
|    | Células editadas génicamente - 24 semanas | 1676                      | 987           | 2365            |
| 15 | Células de control - 16 semanas           | 1075                      | 386           | 1764            |
|    | Células de control - 20 semanas           | 1883                      | 1193          | 2572            |
|    | Células de control - 24 semanas           | 2466                      | 1777          | 3155            |

A las 25 semanas, los animales supervivientes serán sometidos a una provocación de insulina (ensayo de tolerancia a la insulina, ITT) para evaluar los cambios del péptido C humano en suero en respuesta a la disminución de la glucosa en sangre en ausencia de acceso a alimentos. Se obtendrán muestras de sangre de animales en ayunas antes y en múltiples puntos temporales (15, 30, 60 minutos) después de la administración intraperitoneal de 1 unidad de insulina por kg de peso corporal. Las concentraciones séricas de péptido C humano se determinarán mediante ensayos inmunoabsorbentes ligados a enzimas convencionales.

A las 26 semanas, los animales supervivientes se sacrificarán y los artículos de ensayo explantados se procesarán en portaobjetos y se teñirán con HyE y mediante inmunohistoquímica (IHC) para determinar la insulina y el glucagón e identificar las células endocrinas pancreáticas humanas. Se realizará una IHC adicional para el marcador nuclear NuMA1 específico de seres humanos para identificar la posible ubicación de células derivadas del injerto fuera del lumen del explante del artículo de ensayo.

#### 30 **Ejemplo 12. Estudio de eficacia *in vivo* de línea celular con KO de B2M/KI de PD-L1, KO de TXNIP/KI de HLA-E**

Se formularán agregados de células del endodermo pancreático KO de B2M/KI de PD-L1, KO de TXNIP/KI de HLA-E (que comprenden aproximadamente  $7 \times 10^6$  células) en artículos de ensayo. Se implantarán por vía subcutánea cuarenta y seis ratas desnudas atímicas con dos artículos de ensayo cada una. Los animales en estudio se evaluarán para GSIS, ITT y glucosa en sangre sin ayuno (NFBG, por sus siglas en inglés). Se sacrificarán diez animales por grupo

en los puntos temporales de terminación programados de 13, 17, 26 y 39 semanas, mientras que 6 animales adicionales estarán en estudio para tener en cuenta posibles terminaciones precoces no programadas. De cada animal, los dos artículos de ensayo explantados se asignarán aleatoriamente a la evaluación histológica o a la evaluación del contenido total de péptido C. La Tabla 15 presenta el diseño del estudio.

**Tabla 15. Diseño del estudio**

|    | Número de grupo | Número de Artículos de ensayo | Número de animales | Puntos temporales de GSIS (Semanas) | Puntos temporales de ITT (Semanas) | Punto final (Semanas) | Análisis de Explantes  |
|----|-----------------|-------------------------------|--------------------|-------------------------------------|------------------------------------|-----------------------|------------------------|
| 10 | 1               | 20                            | 10 machos          | 12                                  | NA                                 | 13                    | Para cada Grupo:       |
|    | 2               | 20                            | 10 machos          | 12, 16                              | NA                                 | 17                    | Histología             |
| 15 | 3               | 20                            | 10 machos          | 12, 16, 20, 24                      | 25                                 | 26                    | 5 animales             |
|    | 4               | Hasta 32                      | Hasta 16 machos    | 12, 16, 20, 24, 30, 36              | 25, 33                             | 39                    | Contenido de péptido C |
| 20 | Total           | 92                            | 46 machos          |                                     |                                    |                       | 5 animales             |

A las 12, 16, 20, 24, 30 y 36 semanas, todos los animales supervivientes se someterán a evaluaciones de eficacia a través de ensayos de secreción de insulina estimulada por glucosa (GSIS). Se obtendrán muestras de sangre de animales en ayunas antes y después de la administración intraperitoneal de glucosa 3 g/kg. Las concentraciones séricas de péptido C humano se determinarán mediante ensayos inmunoabsorbentes ligados a enzimas convencionales.

A las 25 y 33 semanas, los animales supervivientes se someterán a una provocación de insulina (ensayo de tolerancia a la insulina, ITT) para evaluar los cambios del péptido C humano en suero en respuesta a la disminución de la glucosa en sangre en ausencia de acceso a alimentos. Se obtendrán muestras de sangre de animales en ayunas antes y en múltiples puntos temporales (15, 30, 60 minutos) después de la administración intraperitoneal de 1 unidad de insulina por kg de peso corporal. Las concentraciones séricas de péptido C humano se determinarán mediante ensayos inmunoabsorbentes ligados a enzimas convencionales.

La glucosa en sangre sin ayuno (NFBG, *non-fasting blood glucose*) se medirá antes del inicio del ayuno para los ensayos de GSIS e ITT, aproximadamente a las 12, 16, 20, 24, 25, 30, 33 y 36 semanas.

En los puntos finales programados identificados en la Tabla 13, los animales se sacrificarán. La eutanasia se realizará por inhalación de CO<sub>2</sub> seguida de una toracotomía bilateral. Se realizará una necropsia general en todas las terminaciones programadas y no programadas y se registrarán las anomalías macroscópicas.

5 Los explantes designados se congelarán seguido de homogeneización del contenido del lumen. El contenido total de péptido C del homogeneizado se determinará mediante ensayos inmunoenzimáticos convencionales. El contenido total de péptido C del explante se usará para proyectar la dosis clínica.

10 Los artículos de ensayo explantados designados se fijarán en formol tamponado neutro, se procesarán en portaobjetos y se teñirán con HyE y mediante inmunohistoquímica (IHC) para determinar la insulina y el glucagón e identificar las células endocrinas pancreáticas humanas. Se realizará una IHC adicional para el marcador nuclear NuMA1 específico de seres humanos para identificar la posible ubicación de células derivadas del injerto fuera del lumen del explante del artículo de ensayo.

15 ***Ejemplo 13: Generación de iPSC humanas con KO de B2M/KI de PD-L1, KO de TXNIP/KI de HLA-E***

Se generaron iPSC humanas (iPSC 0025) en las que se insertó la secuencia codificante de PD-L1 en el locus de B2M. Se formó un complejo de RNP combinando ARNg de B2M-2 (SEQ ID NO: 2) y proteína Cas9 en una relación molar de 3:1 (ARNg:Cas9). Para formar el  
20 complejo de RNP, se combinaron ARNg y Cas9 en un recipiente con tampón R hasta un volumen total de 25 µl y se incubaron durante 15 min a TA. Las células se disociaron usando ACCUTASE®, a continuación se volvieron a suspender en medio DMEM/F12 (Gibco, n.º de cat. 11320033), se contaron usando un NC-200 (Chemometec) y se centrifugaron. Se resuspendieron un total de 1x10<sup>6</sup> células con el complejo de RNP. Se añadieron cuatro µg del  
25 plásmido donante B2M-CAGGS-PD-L1 (SEQ ID NO: 33) y tampón R para un volumen total de 125 µl. A continuación, esta mezcla se sometió a electroporación usando los parámetros: 2 pulsos, 30 ms, 1100 V. Siete días después de la electroporación, las células se enriquecieron en células positivas para PD-L1 mediante MACS usando reactivos Miltenyi o reactivos Thermofisher esencialmente como se ha descrito anteriormente en el Ejemplo 8.

30 Después de que se expandiera la población positiva en PD-L1 enriquecida, las células se electroporaron con un complejo de RNP que comprendía ARNg de TXNIP-T5 (SEQ ID NO: 20) y la proteína Cas9 en una relación molar de 3:1 (ARNg:Cas9) y 4 µg del plásmido 2 donante TXNIP-CAGGS-HLA-E (SEQ ID NO: 56) esencialmente como se ha descrito anteriormente. Siete días después de la electroporación, las células se enriquecieron en células positivas para  
35 HLA-E mediante MACS usando reactivos Miltenyi o reactivos Thermofisher. Después de el enriquecimiento en HLA-E, las células se clasificaron individualmente usando el clasificador

FACS WOLF (Nanocollect) en placas de 96 pocillos recubiertas con BIOLAMININA 521 CTG con StemFlex y Revitacell. Se cultivaron células individuales en placa en una incubadora normóxica (37 °C, CO<sub>2</sub> al 8 %) con cambios de medio cada dos días hasta que las colonias fueron lo suficientemente grandes para volver a sembrarse como células individuales. Cuando alcanzaron la confluencia, las muestras se dividieron para mantenimiento y extracción del ADN genómico. Los anticuerpos anti-PD-L1 y anti-HLA-E (Tabla 4) se usaron para el enriquecimiento por MACS y la clasificación por FACS en placas de 96 pocillos con la selección ajustada para células doblemente positivas para HLA-E y PD-L1. Para la clasificación por FACS, las células sin editar sirvieron como control negativo.

Los clones dirigidos correctamente se identificaron mediante PCR para la inserción de KI de PD-L1 y la inserción de KI de HLA-E usando cebadores que amplifican una región desde fuera de los brazos de homología del plásmido hasta la inserción de ADNc de PD-L1 o la inserción de ADNc de HLA-E, respectivamente, permitiendo la amplificación del ADN integrado en KI únicamente. Se sometió a ensayo la inserción específica para determinar la cigosidad mediante PCR para evaluar si la KI se produjo de manera heterocigota u homocigota. Si se identificaba un clon heterocigoto, el alelo negativo para KI se enviaba a secuenciación de Sanger para verificar que contenía un indel interruptor de B2M o un indel interruptor de TXNIP, respectivamente. Los clones con KI correctos con interrupciones completas de B2M y TXNIP (ya sea mediante inserción de KI o formación de indel) se expandieron en formatos de cultivo tisular crecientes hasta alcanzar un tamaño de población de 30 millones de células. Los clones seleccionados se expandieron de esta manera y se confirmó que eran pluripotentes mediante ensayos para determinar OCT4 y SOX2 mediante citometría de flujo intracelular.

Se diferenciaron cuatro clones de hiPSC editadas (VI-B) usando el protocolo endocrino pancreático de Rezania *et al.* (Nat Biotechnol. Noviembre de 2014; 32(11):1121-33). En la Etapa 4, se realizó una citometría de flujo para la cromogranina (CHGA), PDX1 y NKX6.1. Los resultados para PDX1 y NKX6.1 de un clon (clon 1) sembrado a diferentes densidades representativas se muestran en la FIG. 26A. La CHGA fue negativa para los cuatro clones. También se realizó una citometría de flujo para PD-L1 y HLA-E. Los resultados para PD-L1 y HLA-E de un clon (clon 1) se muestran en la FIG. 26B.

**Ejemplo 14: Proceso para fabricar Bancos de Células Criogénicas de Células Madre Pluripotentes Humanas (hPSCs) de KO de B2M/KI de PD-L1, KO de TXNIP/KI de HLA-E**

Se electroporaron hESC CyT49 (ViaCyte) con un complejo de RNP que comprendía ARNg de B2M-2 (SEQ ID NO: 2) y proteína Cas9 en una relación molar de 3:1 (ARNg:Cas9) y 4 µg del plásmido donante B2M-CAGGS-PD-L1 (SEQ ID NO: 33) durante 2 pulsos de 30 ms a 1100 V. Después de la electroporación, las células se pipetearon en un tubo Eppendorf lleno de medio StemFlex con RevitaCell. A continuación, esta suspensión celular se sembró en placas

de cultivo tisular recubiertas previamente con BIOLAMININA 521 CTG a una dilución 1:20. Las células se cultivaron en una incubadora normóxica (37 °C, 8 % de CO<sub>2</sub>).

Siete días después de la electroporación, las células se enriquecieron en células positivas para PD-L1 mediante MACS usando anticuerpos anti-PD-L1 marcados con Alexa-488 y perlas magnéticas (DYNABEADS® Pan Mouse IgG; Thermo Fisher). Las células positivas para PD-L1 se expandieron cultivándolas en medio de expansión XF-KSR (Gibco) durante 7 días.

A continuación, las células positivas para PD-L1 se electroporaron con un complejo de RNP que comprendía ARNg de TXNIP-T5 (SEQ ID NO: 20) y la proteína Cas9 en una relación molar de 3:1 (ARNg:Cas9) y 4 µg del plásmido 2 donante TXNIP-CAGGS-HLA-E (SEQ ID NO: 56) durante 2 pulsos de 30 ms a 1100 V. Después de la electroporación, las células se pipetearon en un tubo Eppendorf lleno de medio StemFlex con RevitaCell. A continuación, esta suspensión celular se sembró en placas de cultivo tisular recubiertas previamente con BIOLAMININA 521 CTG a una dilución 1:20. Las células se cultivaron en una incubadora normóxica (37 °C, 8 % de CO<sub>2</sub>).

Siete días después de la electroporación, las células se enriquecieron en células positivas para HLA-E mediante MACS usando anticuerpos anti-HLA-E marcados con PE y perlas magnéticas (DYNABEADS® Pan Mouse IgG; Thermo Fisher). Las células doblemente positivas para PD-L1 y HLA-E se expandieron cultivándolas en medio de expansión XF-KSR (Gibco) durante aproximadamente 5 días.

Las células doblemente positivas para PD-L1 y HLA-E se clasificaron individualmente. Para ello, las células se alimentaron con StemFlex Complete con Revitacell (para una concentración final de 1X Revitacell) 3-4 horas antes de la disociación con ACCUTASE®. Después de la disociación, las células individuales se clasificaron en pocillos individuales de una placa de cultivo tisular de 96 pocillos recubierta con BIOLAMININA. Se usó el clasificador FACS WOLF (Nanocollect) para clasificar las células individuales en los pocillos usando los anticuerpos anti-PD-L1 y anti-HLA-E descritos anteriormente. Las placas se llenaron previamente con 100-200 µl de StemFlex Complete con Revitacell. Tres días después de la siembra celular, las células se alimentaron con StemFlex nuevo y se siguieron alimentando cada dos días con 100-200 µl de medio. Después de 10 días de crecimiento, las células se alimentaron a diario con StemFlex hasta el día 12-14. En este momento, las placas se disociaron con ACCUTASE® y las suspensiones celulares recogidas se dividieron 1:2 en dos placas de 96 pocillos, que se cultivaron durante aproximadamente 4 días.

Se cosechó una porción de las células para el análisis visual (morfología) y el análisis de ADN (PCR y secuenciación de ADN para el análisis de la cigosidad y el perfil de indel), y el resto de las células se cultivaron y se expandieron para su cultivo en matraces T175. Después

de aproximadamente dos semanas de cultivo, se seleccionaron los clones para su congelación. Las células se caracterizaron antes y después de la congelación para determinar la morfología, viabilidad, endotoxinas, micoplasma, cariotipo, pluripotencia, capacidad de diferenciación, análisis de dianas activadas/inactivadas, integración de plásmidos aleatoriamente, Cas9/plásmido residual usando procedimientos convencionales. Las células se congelaron en medios criogénicos y se almacenaron en viales criogénicos a -80 °C o nitrógeno líquido.

Mediante dicho proceso se fabricó y aisló un clon con KO de B2M/KI de PD-L1 + KO de TXNIP/KI de HLA-E particular ("clon de siembra"). El clon de siembra se diferenció a la fase de PEC y se caracterizó. La FIG. **27A** muestra que la morfología del clon de siembras en la fase PEC era similar a la de las células de tipo silvestre. La FIG. **27B** muestra que el patrón de expresión cinética de *FOXA2*, *CHGA*, *PDX1* y *NKX6.1* a lo largo de un curso temporal de diferenciación en las células diferenciadas a partir del clon de siembra fue similar al de las células de tipo silvestre. La FIG. **27C** muestra el porcentaje de células CHGA<sup>+</sup>/NKX6.1<sup>+</sup>/PDX1<sup>+</sup> en la población diferenciada.