

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 780 904**

51 Int. Cl.:

C12N 9/22 (2006.01)

C12N 9/96 (2006.01)

C12N 15/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.08.2015 PCT/US2015/045504**

87 Fecha y número de publicación internacional: **25.02.2016 WO16028682**

96 Fecha de presentación y número de la solicitud europea: **17.08.2015 E 15756300 (8)**

97 Fecha y número de publicación de la concesión europea: **25.12.2019 EP 3180426**

54 Título: **Edición genómica usando nickasas Cas9**

30 Prioridad:

17.08.2014 US 201462038358 P

17.06.2015 US 201562180699 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.08.2020

73 Titular/es:

THE BROAD INSTITUTE, INC. (33.3%)

415 Main Street

Cambridge, MA 02142, US;

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(33.3%) y

PRESIDENT AND FELLOWS OF HARVARD

COLLEGE (33.3%)

72 Inventor/es:

ZHANG, FENG y

RAN, FEI

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 780 904 T3

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

DESCRIPCIÓN

Edición genómica usando nickasas Cas9

5

Solicitudes relacionadas

Esta solicitud reivindica prioridad respecto a la solicitud de patente provisional de Estados Unidos n.º de serie 62/038.358 presentada el 17 de agosto de 2014 y la solicitud de patente provisional de Estados Unidos n.º de serie 62/180.699 presentada el 17 de junio de 2015.

10

Todos los documentos o solicitudes citados en los mismos durante su prosecución ("documentos citados en la solicitud") y todos los documentos citados o remitidos en los documentos citados en la solicitud, y todos los documentos citados o remitidos en este documento ("documentos citados en este documento"), y todos los documentos citados o remitidos en los documentos citados en este documento, junto con cualquier instrucción del fabricante, descripción, especificación de producto y ficha de producto para cualquier producto mencionado en este documento o en cualquier documento citado por referencia en este documento, pueden emplearse en la práctica de la invención. Más específicamente, todos los documentos remitidos se citan en la misma medida que si cada documento individual se indicara específica e individualmente citado.

15

20

Declaración sobre investigación con subvención federal

Esta invención se llevó a cabo con subvención gubernamental según el número de adjudicación MH100706 concedido por el National Institutes of Health. El gobierno tiene determinados derechos en la invención.

25

Campo de la invención

La presente invención se refiere en general al suministro, manipulación y optimización de sistemas, métodos y composiciones usados para el control de la expresión génica que implica dirección a secuencias específicas, tal como perturbación del genoma o edición génica, que están relacionados con repeticiones palindrómicas cortas intercaladas de forma regular agrupadas (CRISPR) y componentes de las mismas.

30

En particular, la presente invención se refiere a la preparación, ensayo y aplicación de enzimas Cas9 mutadas que pueden inducir mellas monocatenarias para una manipulación de precisión del genoma de mamíferos.

35

Antecedentes de la invención

Los recientes avances en las técnicas de secuenciación genómica y los métodos de análisis han acelerado significativamente la capacidad de catalogar y cartografiar los factores genéticos asociados con una gama diversa de funciones biológicas y enfermedades. Se necesitan tecnologías precisas de dirección genómica para posibilitar la manipulación sistémica inversa de las variaciones genéticas causales permitiendo la perturbación selectiva de elementos genéticos individuales, así como avanzar en aplicaciones de biología sintética, biotecnológica y médica. Aunque las técnicas de edición genómica tales como dedos de cinc de diseño, efectores de tipo activador de la transcripción (TALE) o meganucleasas de asentamiento para producir perturbaciones genómicas dirigidas, sigue habiendo una necesidad de nuevas tecnologías de manipulación genómica que se pueden producir, sean fáciles de configurar, se puedan cambiar de escala y sean susceptibles a dirección a múltiples posiciones dentro del genoma eucariótico.

40

45

La edición genómica dirigida, rápida y eficaz usando el sistema de Cas9 guiado por ARN está posibilitando la consulta sistemática de elementos genéticos en una diversidad de células y organismos y tiene enorme potencial como genoterapia de próxima generación (Hsu Lander, y Zhang, 2014). En contraste con otros sistemas dirigidos a ADN basados en proteínas de dedos de cinc (ZFP) (Klug, 2010) y efectores de tipo activador de la transcripción (TALE) (Boch y Bonas, 2010), que dependen de dominios proteínicos para conferir especificidad de unión a ADN, Cas9 forma un complejo con un ARN guía pequeño que dirige la enzima a su diana de ADN mediante emparejamiento de bases de Watson y Crick. Por consiguiente, el sistema es simple y rápido de diseñar y requiere solamente la producción de un oligonucleótido corto para dirigir la unión al ADN a cualquier locus.

50

55

El sistema de CRISPR (repeticiones palindrómicas cortas intercaladas de forma regular agrupadas) microbiano de tipo II (Chylinski, Makarova, Charpentier, y Koonin, 2014), que es el más simple entre los tres tipos de CRISPR conocidos (Barrangou y Marraffini, 2014; Gasiunas, Sinkunas, y Siksnys, 2014; Wiedenheft, Sternberg, y Doudna, 2012), consiste en los genes asociados a CRISPR (Cas) y una serie de elementos repetitivos no codificantes (repeticiones directas) intercalados por secuencias variables cortas (espaciadores). Estos espaciadores cortos de aproximadamente 30 pb a menudo derivan de elementos genéticos exógenos tales como fagos y plásmidos de conjugación, y constituyen la base para una memoria inmunitaria adaptativa de esos elementos invasores (Barrangou *et al.*, 2007). Las secuencias correspondientes en los genomas del fago y los plásmidos se llaman protoespaciadores, y cada protoespaciador está flanqueado por un motivo adyacente de protoespaciador (PAM) corto, que desempeña una función crucial en el

60

65

mecanismo de búsqueda y reconocimiento de la diana de Cas9. La matriz de CRISPR se transcribe y procesa en moléculas de ARN cortas conocidas como ARN de CRISPR (ARNcr) que, junto con un segundo ARN de transactivación (ARNcrtra) corto (Deltcheva *et al.*, 2011), forma complejos con Cas9 para facilitar el reconocimiento y escisión de la diana (Deltcheva *et al.*, 2011; Garneau *et al.*, 2010). Además, el ARNcr y el ARNcrtra pueden fusionarse en un ARN guía simple (ARNgs) para facilitar la dirección de Cas9 (Jinek *et al.*, 2012).

La enzima Cas9 de *Streptococcus pyogenes* (SpCas9), que requiere un PAM 5'-NGG (Mojica, Diez-Villasenor, Garcia-Martinez, y Almendros, 2009), se ha usado ampliamente para aplicaciones de edición genómica (Hsu *et al.*, 2014). Para dirigirse a cualquier locus genómico deseado de interés que cumpla el requisito de PAM, la enzima puede "programarse" simplemente alterando la secuencia guía de 20 pb del ARNgs. Además, la simplicidad de la dirección le confiere una fácil combinación tal como edición simultánea de varios locus incluyendo múltiples ARNgs (Cong *et al.*, 2013; Wang *et al.*, 2013).

Como otras nucleasas de diseño, Cas9 facilita la edición genómica induciendo roturas bicatenarias (DSB) en su sitio diana, que a su vez estimula las rutas de reparación de daños en ADN endógeno que dan lugar a ADN editado: la reparación dirigida por homología (HDR), que requiere un molde homólogo para recombinación, pero que repara DSB con alta fidelidad, y la unión de extremos no homólogos (NHEJ), que funciona sin un molde y produce frecuentemente inserciones o eliminaciones (indel) como consecuencia de la reparación. Pueden diseñarse moldes de HDR exógenos e introducirse junto con Cas9 y ARNgs para promover una alteración de secuencia exacta a un locus diana; sin embargo, este proceso típicamente se produce únicamente en células en división y a baja eficacia.

Determinadas aplicaciones, por ejemplo, edición genómica terapéutica en células madre humanas, demanda edición que no solamente sea eficaz, sino también altamente específica. Las nucleasas con actividad DSB inespecífica podrían inducir mutaciones indeseables con efectos potencialmente perjudiciales, un resultado inaceptable en la mayoría de entornos clínicos. La facilidad notable de dirigir Cas9 ha posibilitado una unión inespecífica extensa y estudios de mutagénesis empleando secuenciación exhaustiva (Fu *et al.*, 2013; Hsu *et al.*, 2013; Pattanayak *et al.*, 2013) e inmunoprecipitación de cromatina (ChIP) en células humanas (Kuscu, Arslan, Singh, Thorpe, y Adli, 2014; Wu *et al.*, 2014). Como resultado, está surgiendo una imagen cada vez más completa de la actividad inespecífica de la enzima. Cas9 tolerará algunos emparejamientos incorrectos entre su guía y un sustrato de ADN, una característica que depende mucho del número, posición (PAM proximal o distal) y la identidad de los emparejamientos incorrectos. La unión inespecífica y la escisión pueden depender además del organismo que se esté editando, el tipo celular y los contextos epigenéticos.

Estos estudios de especificidad, junto con las investigaciones directas del mecanismo catalítico de Cas9, han estimulado la manipulación guiada por homología y estructura para mejorar su especificidad de dirección. La enzima de tipo silvestre hace uso de dos dominios de nucleasa conservados, HNH y RuvC, para escindir ADN por mellado de la hebra complementaria de ARNgs y no complementaria, respectivamente. Una "nickasa" mutante (Cas9n) puede generarse por sustitución con alanina en residuos catalíticos clave dentro de estos dominios - SpCas9 D10A inactiva RuvC (Jinek *et al.*, 2012), mientras que N863A se ha descubierto que inactiva HNH (Nishimasu *et al.*, 2014). Aunque también se ha informado de que una mutación H840A convierte Cas9 en una enzima melladora, este mutante tiene niveles reducidos de actividad en células de mamífero en comparación con N863A (Nishimasu *et al.*, 2014).

Como las mellas monocatenarias en general se reparan mediante la ruta de reparación de escisión de bases no mutagénica (Dianov y Hubscher, 2013), los mutantes Cas9n pueden aprovecharse para mediar la manipulación genómica altamente específica. Una sola mella inducida por Cas9n puede estimular HDR a baja eficacia en algunos tipos celulares, mientras que dos enzimas melladoras, espaciadas apropiadamente y orientadas en el mismo locus, generan de forma eficaz DSB, creando salientes 3' o 5' a lo largo de las dianas, en oposición a una DSB roma como en el caso de tipo silvestre (Mali *et al.*, 2013; Ran *et al.*, 2013). La eficacia de modificación específica de la estrategia de doble mellado es comparable con el tipo silvestre, pero se reducen los indel en sitios inespecíficos predichos por debajo del umbral de detección mediante secuenciación exhaustiva Illumina (Ran *et al.*, 2013).

A pesar del progreso en las tecnologías de manipulación genética dirigida por Cas9, la eficacia de modificaciones génicas satisfactorias, en particular en el contexto de HDR, aún está a niveles bajos, y se necesitan estrategias mejoradas para aumentar la eficacia de HDR para la manipulación genética dirigida por Cas9.

Sumario de la invención y de la presente divulgación

El sistema de CRISPR-Cas no requiere la generación de proteínas personalizadas para abordar secuencias específicas, sino que en su lugar una sola enzima Cas puede programarse mediante una molécula de ARN corta para reconocer una diana de ADN específica. Añadir el sistema de CRISPR-Cas al repertorio de técnicas de secuenciación genómica y métodos de análisis puede simplificar significativamente la metodología y acelerar la capacidad de catalogar y cartografiar factores genéticos asociados con una gama diversa de funciones biológicas y enfermedades. Para utilizar el sistema de CRISPR-Cas de forma eficaz para edición genómica sin efectos perjudiciales, es crucial comprender los aspectos de manipulación y optimización de estas herramientas de manipulación genómica, que son aspectos de la invención reivindicada.

Existe una necesidad apremiante de sistemas alternativos y robustos y técnicas para dirección de secuencias con una amplia serie de aplicaciones. Aspectos de esta invención abordan esta necesidad y proporcionan ventajas relacionadas. Un complejo de CRISPR ejemplar comprende una enzima CRISPR mutada en complejo con una secuencia guía hibridada con una secuencia diana dentro del polinucleótido diana. La secuencia guía se une a una secuencia de acoplamiento de crtra, que a su vez hibrida con una secuencia crtra.

En un aspecto, la divulgación proporciona métodos para usar uno o más elementos de un sistema de CRISPR. El complejo de CRISPR de la invención proporciona un medio eficaz para modificar un polinucleótido diana. El complejo de CRISPR de la invención tiene una amplia diversidad de utilidades, incluyendo la modificación (por ejemplo, eliminación, inserción, translocación, inactivación, activación) de un polinucleótido diana en una multiplicidad de tipo celulares. En un aspecto, la célula es una célula eucariótica. En un aspecto, la célula es una célula procariótica. Por tanto, el complejo de CRISPR de la divulgación tiene un amplio espectro de aplicaciones en, por ejemplo, edición génica o genómica, genoterapia, descubrimiento de fármacos, cribado de fármacos, diagnóstico de enfermedades y pronóstico. Un complejo de CRISPR ejemplar comprende una enzima CRISPR en complejo con una secuencia guía hibridada con una secuencia diana dentro del polinucleótido diana. En un aspecto, la secuencia guía se une a una secuencia de acoplamiento de crtra, que a su vez hibrida con una secuencia crtra. Aspectos de la invención se refieren a enzimas Cas9 que tiene especificidad de diana mejorada en un sistema de CRISPR-Cas9, que tiene ARN guía con actividad óptima, con enzimas Cas9 que son de menor longitud que las enzimas Cas9 de tipo silvestre (y moléculas de ácido nucleico que las codifican), y enzimas Cas9 quiméricas, así como métodos de mejora de la especificidad de diana de una enzima Cas9 o de diseño de un sistema de CRISPR-Cas9 que comprende diseñar o preparar ARN guía que tienen actividad óptima y/o seleccionar o preparar una enzima Cas9 que tenga un menor tamaño o longitud que Cas9 de tipo silvestre, por lo que el empaquetado de un ácido nucleico codificante de dicha construcción en un vector de suministro es ventajoso ya que hay menos codificación para ello en el vector de suministro que para Cas9 de tipo silvestre, y/o generar enzimas Cas9 quiméricas.

También se proporcionan usos de las presentes secuencias, vectores, enzimas o sistemas, en medicina. También se proporcionan usos en los mismos en edición génica o genómica.

La enzima de CRISPR es una nickasa. La nickasa es una Cas9 modificada que comprende una mutación en N863A (de acuerdo con la numeración encontrada en SpCas9 de *S. pyogenes*) o en N580 (de acuerdo con la numeración encontrada en SaCas9 de *S. aureus*) o en un residuo que es equivalente o correspondiente a los residuos en ortólogos de *S. pyogenes* o *S. aureus*. En particular, la mutación del residuo en A (alanina) se prefiere en algunas realizaciones de la divulgación, pero debería ser suficiente cualquier mutación catalíticamente inactiva en estos residuos. Sorprendentemente, los solicitantes encontraron que el uso de esta mutación en un sistema de nickasa doble suprime NHEJ y en su lugar promueve HDR a través de la generación de salientes 3' en el ADN bicatenario mellado.

Como Cas9n (D10A) y Cas9H840A mellan hebras opuestas de ADN como se muestra previamente, la sustitución de Cas9n con Cas9H840A con un par dado de ARNgs debe provocar la inversión del tipo saliente. Por ejemplo, un par de ARNgs que generará un saliente 5' con Cas9n debe generar en principio el saliente 3' correspondiente en su lugar. Por lo tanto, los pares de ARNgs que dan lugar a la generación de un saliente 3' con Cas9n podrían usarse con Cas9H840A para generar un saliente 5'. Inesperadamente, los solicitantes ensayaron Cas9H840A con un conjunto de pares de ARNgs diseñados para generar tanto salientes 5' como 3' (intervalo de compensación de -278 a +58 pb), pero no pudieron observar formación de indel. Sorprendentemente, se observó HDR con la presente nickasa Cas9 mutada (N863A) que crea salientes 3'. Por tanto, de la presente invención, se pueden usar salientes 3', por ejemplo, salientes 3' emparejados para obtener HDR. En SaCas9, los solicitantes han identificado que la mutación correspondiente a N863A en SpCas9 es N580A. Puede tener la ventaja de ser una mutación más predecible para la función de la proteína que el equivalente H840A, que puede cambiar el comportamiento de unión.

En la divulgación, la enzima Cas9 comprende una mutación y puede usarse como proteína de unión a ADN genérico (por ejemplo, la Cas9 mutada puede funcionar o no como nucleasa bicatenaria o como nickasa monocatenaria; puede funcionar simplemente como proteína de unión; pero ventajosamente, la Cas9 es una nickasa); y la Cas9 así mutada puede ser con o sin fusión a un dominio funcional o dominio proteínico. La mutación afecta al dominio catalítico HNH en el residuo N863; la enzima Cas9 es una proteína SpCas9 que comprende la mutación N863A, o cualquier ortólogo mutado que tenga una mutación correspondiente a SpCas9N863A. En un aspecto de la divulgación, la enzima Cas9 mutada puede fusionarse a un dominio proteínico o dominio funcional, por ejemplo, tal como un dominio de activación transcripcional. En un aspecto, el dominio de activación transcripcional puede ser VP64. En otro aspecto, el dominio proteínico o dominio funcional puede ser, por ejemplo, un dominio FokI. Otros aspectos de la divulgación se refieren a la enzima Cas9 mutada que está fusionada a dominios que incluyen, aunque sin limitación, un represor transcripcional, una recombinasa, una transposasa, un remodelador de histonas, una ADN metiltransferasa, un criptocromo, un dominio inducible/controlable por la luz o un dominio inducible/controlable químicamente. En este documento también se describen dominios funcionales adicionales.

En una realización adicional, la divulgación proporciona métodos para generar ARNcrtra mutante y secuencias de repetición directa o secuencias guía quiméricas mutantes que permitan potenciar el rendimiento de estos ARN en células. Aspectos de la divulgación también proporcionan selección de dichas secuencias.

Aspectos de la divulgación también proporcionan métodos de simplificación de la clonación y suministro de componentes del complejo de CRISPR. En la realización preferida de la divulgación, un promotor adecuado, tal como el promotor U6, se amplifica con un oligo de ADN y se coloca contiguo y anterior a una secuencia codificante del ARN guía. El producto de PCR resultante entonces se transfecta en células para dirigir la expresión de ARN guía. Aspectos de la divulgación también se refieren al ARN guía que se transcribe *in vitro* o se solicita desde una empresa de síntesis y se transfiere directamente.

En un aspecto, la divulgación proporciona métodos para mejorar la actividad usando una polimerasa más activa. En un aspecto, puede insertarse un promotor T7 contiguo y anterior a una secuencia codificante de un ARN guía. En una realización preferida, la expresión de ARN guía bajo el control del promotor T7 se dirige mediante la expresión de la polimerasa T7 en la célula. En una realización ventajosa, la célula es una célula eucariótica. En una realización preferida, la célula eucariótica es una célula humana. En una realización más preferida, la célula humana es una célula específica de paciente.

En un aspecto, la divulgación proporciona métodos de reducción de la toxicidad de enzimas Cas. La enzima Cas9 es una nickasa. En una realización, la Cas9 se suministra a la célula en forma de ARNm. Esto permite la expresión transitoria de la enzima reduciendo de ese modo la toxicidad. En otra realización, la Cas9 se suministra a la célula en la construcción nucleotídica que codifica y expresa la enzima Cas9. En otra realización, la divulgación también proporciona métodos de expresión de Cas9 bajo el control de un promotor inducible en las construcciones usadas en los mismos. En otra realización, la Cas9 se suministra a la célula como una proteína. En otra realización y particularmente preferida, la Cas9 se suministra a la célula como una proteína o con una secuencia de nucleótidos que la codifica.

En otro aspecto, la divulgación proporciona métodos de mejora de las aplicaciones *in vivo* del sistema de CRISPR-Cas. Un aspecto ventajoso de la divulgación proporciona la selección de homólogos de Cas9 que se empaquetan fácilmente en vectores víricos para su suministro. Los ortólogos de Cas9 típicamente comparten la organización general de 3-4 dominios RuvC y un dominio HNH. El dominio RuvC más 5' escinde la hebra no complementaria, y el dominio HNH escinde la hebra complementaria. Todas las indicaciones son en referencia a la secuencia guía.

El residuo catalítico en el dominio RuvC 5' se identifica a través de comparación de homología de la Cas9 de interés con otros ortólogos de Cas9 (de locus de CRISPR de tipo II de *S. pyogenes*, locus 1 de CRISPR de *S. thermophilus*, locus 3 de CRISPR de *S. thermophilus* y locus de CRISPR de tipo II de *Franciscilla novicida*), y el residuo de Asp conservado se muta en alanina para convertir Cas9 en una enzima melladora de hebra complementaria. En la presente invención, el residuo de asparagina conservado (por ejemplo, N863 en Cas9 de *S. pyogenes* (Sp)) en el dominio HNH se muta en alanina para convertir Cas9 en una enzima melladora de hebra no complementaria (por ejemplo, SpCas9 N863A).

La presente invención abarca además cualquier ortólogo mutado que corresponda a SpCas9N863A. En algunas realizaciones, el ortólogo es de *Staphylococcus aureus* de modo que la Cas9 es la de o derivada de *Staphylococcus aureus* (denominada SaCas9). En algunas realizaciones, la *Staphylococcus aureus* es *Staphylococcus aureus* subespecie *aureus*. En algunas realizaciones, la mutación correspondiente a N863A en SpCas9 es N580A en *Staphylococcus aureus* o *Staphylococcus aureus* subespecie *aureus*.

La enzima de CRISPR es una enzima de CRISPR de tipo II mutada. Esta enzima de CRISPR de tipo II es una enzima Cas9 mutada. Una enzima Cas puede identificarse Cas9 ya que esto puede referirse a la clase general de enzimas que comparten homología con la nucleasa más voluminosa o grande con múltiples dominios nucleasa del sistema de CRISPR de tipo II. Más preferiblemente, la enzima Cas9 es de, o deriva de, SpCas9 o SaCas9. Por derivada, los solicitantes indican que la enzima derivada se basa en gran medida, en el sentido de tener un alto grado de homología de secuencia con, una enzima de tipo silvestre, pero que se ha mutado (modificado) de alguna manera como se describe en este documento.

Se apreciará que las expresiones enzima Cas y de CRISPR se usan en general indistintamente, salvo que sea evidente lo contrario. Como menciona anteriormente, muchas de las numeraciones de residuos usadas en este documento se refieren a la enzima Cas9 del locus de CRISPR de tipo II en *Streptococcus pyogenes* (anotada como alternativa como SpCas9 o spCas9). Sin embargo, se apreciará que esta invención incluye muchas Cas9 más de otras especies de microbios, tal como SpCas9 derivada de *S. pyogenes*, SaCas9 derivada de *S. aureus*, St1Cas9 derivada de *S. thermophilus* y así sucesivamente. Ejemplos adicionales se proporcionan en este documento. Por tanto, aunque se hacen numerosas referencias en este documento a una enzima Cas o de CRISPR, se apreciará que estas se aplican igualmente a cualquier ortólogo de Cas9 que funcione según se requiere en este documento. En particular, sin embargo, una mención de una enzima Cas o de CRISPR se aplica igualmente a SpCas9 o SaCas9, y viceversa, salvo que sea evidente lo contrario.

Un ejemplo de una secuencia de codones optimizados, en este caso optimizada para seres humanos (es decir, que se optimiza para su expresión en seres humanos) se proporciona en este documento, véase la secuencia de codones optimizados para seres humanos de SaCas9. Aunque se prefiere esto, se apreciará que otros ejemplos son posibles

y la optimización de codones para una especie hospedadora distinta del ser humano, o para la optimización de codones para órganos específicos tal como el cerebro, es conocida.

5 En otras realizaciones, la divulgación proporciona métodos de potenciación de la función de Cas9 generando proteínas Cas9 quiméricas. Estos métodos pueden comprender la fusión de fragmentos N terminales de un homólogo de Cas9 con fragmentos C terminales de otro homólogo de Cas9. Estos métodos también permiten la selección de nuevas propiedades presentadas por las proteínas quiméricas. Se prefieren quimeras de SpCas9 y SaCas9, en algunas realizaciones.

10 Se apreciará que en los presentes métodos, cuando el organismo es un animal o una planta, la modificación puede producirse *ex vivo* o *in vitro*, por ejemplo, en un cultivo celular o una muestra del organismo y, en algunos caso, no *in vivo*. En otras realizaciones, puede producirse *in vivo*.

15 Alguna o todas de la secuencia polinucleotídica que codifica una enzima de CRISPR, secuencia guía, secuencia de acoplamiento de crtra o secuencia crtra puede ser ARN, ADN o una combinación de ARN y ADN. En un aspecto, los polinucleótidos que comprenden la secuencia que codifica una enzima de CRISPR, la secuencia guía, la secuencia de acoplamiento de crtra o la secuencia crtra son ARN. En un aspecto, los polinucleótidos que comprenden la secuencia que codifica una enzima de CRISPR, la secuencia guía, la secuencia de acoplamiento de crtra o la secuencia crtra son ADN. En un aspecto, los polinucleótidos son una mezcla de ADN y ARN, en la que algunos de los polinucleótidos que comprenden la secuencia que codifica una o más de la secuencia de CRISPR, la secuencia guía, la secuencia de acoplamiento de crtra o la secuencia crtra son ADN y algunos de los polinucleótidos son ARN. En un aspecto, el polinucleótido que comprende la secuencia que codifica la enzima de CRISPR es un ADN y la secuencia guía, la secuencia de acoplamiento de crtra o la secuencia crtra son ARN. El uno o más polinucleótidos que comprenden la secuencia que codifica una enzima de CRISPR, la secuencia guía, la secuencia de acoplamiento de crtra o la secuencia crtra pueden suministrarse mediante electroporación, encapsulación en o fijados a partículas, nanopartículas, exosomas o microvesículas; o mediante fijación a, por ejemplo, una partícula de oro y dispararse usando un llamado "cañón de genes".

30 Se apreciará que cuando se hace referencia a un polinucleótido, cuando es polinucleótido es ARN y se dice que "comprende" un elemento tal como una secuencia de acoplamiento de crtra, la secuencia de ARN incluye el elemento. Cuando el polinucleótido es ADN y se dice que comprende un elemento tal como una secuencia de acoplamiento de crtra, la secuencia de ADN es o puede transcribirse en el ARN que comprende el elemento en cuestión. Cuando el elemento es una proteína, tal como la enzima de CRISPR, la secuencia de ADN o ARN mencionada se traduce, o puede traducirse (y en el caso de ADN se transcribe en primer lugar). Además, en casos donde se proporciona un ARN que codifica la enzima de CRISPR a una célula, se entiende que el ARN puede traducirse por la célula en que se suministra.

35 En un aspecto, la divulgación proporciona una composición que no es de origen natural o que está manipulada que comprende:

40 I. dos o más secuencias polinucleotídicas del sistema de CRISPR-Cas que comprenden

(a) una primera secuencia guía que puede hibridar con una primera secuencia diana en un locus polinucleotídico,

45 (b) una segunda secuencia guía que puede hibridar con una segunda secuencia diana en un locus polinucleotídico,

(c) una secuencia de acoplamiento de crtra, y

50 (d) una secuencia de ARNcrtra, y

II. una enzima Cas9 de tipo II o una segunda secuencia polinucleotídica que la codifica,

55 en la que la enzima Cas9 de tipo II es o comprende una enzima SpCas9 que comprende la mutación N863 o N863A, enzima SaCas9 que comprende la mutación N580 o N580A o un ortólogo de la misma, que tiene una mutación correspondiente a SpCas9N863 o N863A,

60 en la que cuando se transcriben, la primera y la segunda secuencia de acoplamiento de crtra hibridan con la primera y segunda secuencia de ARNcrtra respectivamente y la primera y la segunda secuencia guía dirigen la unión específica de secuencia de un primer y un segundo complejo de CRISPR a la primera y segunda secuencia diana respectivamente,

65 en la que el primer complejo de CRISPR comprende la enzima Cas9 en complejo con (1) la primera secuencia guía que puede hibridar con la primera secuencia diana, y (2) la primera secuencia de acoplamiento de crtra que hibrida con la primera secuencia de ARNcrtra,

en la que el segundo complejo de CRISPR comprende la enzima Cas9 en complejo con (1) la segunda secuencia guía que puede hibridar con la segunda secuencia diana, y (2) la segunda secuencia de acoplamiento de crtra que hibrida con la segunda secuencia de ARNcrtra, y

5 en la que la primera secuencia guía dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía dirige la escisión de la otra hebra cerca de la segunda secuencia diana induciendo una rotura bicatenaria, modificando de ese modo el organismo o el organismo no humano o no animal, y en la que la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocan salientes 3'.

10 En una realización preferida, los componentes I y II están unidos de forma funcional a uno o más elementos reguladores. En una realización preferida, el componente (I) comprende una secuencia polinucleotídica del sistema de CRISPR-Cas que comprende la secuencia guía, la secuencia de acoplamiento de crtra y la secuencia de ARNcrtra.
15 En una realización preferida, el componente (I) comprende un primer elemento regulador unido de forma funcional a la secuencia guía y la secuencia de acoplamiento de crtra, y un tercer elemento regulador unido de forma funcional a la secuencia de ARNcrtra.

20 En una realización preferida, la composición comprende un sistema de suministro configurado de forma funcional para suministrar componentes del complejo de CRISPR-Cas o secuencias polinucleotídicas que comprenden o codifican dichos componentes a una célula. En una realización preferida, el sistema de suministro comprende un sistema de vector que comprende uno o más vectores, y en el que los componentes I y II están ubicados en el mismo vector o en diferentes del sistema. En una realización preferida, el uno o más vectores comprenden uno o más vectores víricos. En una realización preferida, el uno o más vectores víricos comprenden uno o más vectores retrovíricos, lentivíricos, adenovíricos, víricos adenoasociados o del virus del herpes simple.

25 En una realización preferida, el sistema de suministro comprende una nanopartícula, liposoma, exosoma, sistema de levadura, microvesícula o cañón de genes.

30 En una realización preferida, la composición comprende uno o más dominios funcionales. En una realización preferida, el uno o más dominios funcionales comprenden un dominio activador de la transcripción. En una realización preferida, el dominio funcional comprende VP64 o KRAB, SID o SID4X, o una recombinasa, una transposasa, un remodelador de histonas, una ADN metiltransferasa, un criptocromo, un dominio inducible/controlable por la luz o un dominio inducible/controlable químicamente.

35 En una realización preferida, la composición de vector comprende un solo vector.

40 En una realización preferida, la célula es una célula eucariótica. En una realización preferida, el uno o más vectores están configurados de forma funcional para dirigir la expresión de transcritos de CRISPR cuando se introducen en una célula eucariótica.

45 En una realización preferida, la secuencia de nucleótidos que codifica la SaCas9 tiene los codones optimizados para su expresión en una célula eucariótica.

50 En una realización preferida, uno o más de los elementos reguladores comprenden un promotor específico de tejido. En una realización preferida, el promotor específico de tejido dirige la expresión de transcritos de CRISPR en músculo, neurona, hueso, piel, sangre, hígado, páncreas o linfocitos.

55 En una realización preferida, la secuencia diana está adyacente a un motivo adyacente de protoespaciador (PAM) reconocido por la enzima Cas9.

60 En una realización preferida, la secuencia diana está flanqueada en su extremo 3' por 5'-NRG (donde N es cualquier nucleótido) para SpCas9 o NNGRR para SaCas9.

65 En una realización preferida, la secuencia guía puede hibridar con una secuencia diana en una célula eucariótica.

En una realización preferida, la secuencia de ARNcrtra es de 30 o más nucleótidos de longitud. En una realización preferida, el ARNcrtra es de 50 o más nucleótidos de longitud.

60 En una realización preferida, la enzima SaCas9 comprende además una o más secuencias de localización nuclear (NLS).

65 En un aspecto, la divulgación proporciona una célula hospedadora o línea celular *in vivo*, *ex vivo* o *in vitro* que comprende o está modificada por la composición o enzima como se describe en este documento, o la descendencia de la misma. En una realización preferida, la célula hospedadora, línea celular o descendencia de la misma es una célula madre o línea de células madre.

En un aspecto, la divulgación proporciona un método *in vivo* o *ex vivo* de modificación de un organismo por manipulación de una o más secuencias dianas en locus genómicos de interés que comprende suministrar al organismo la composición descrita en este documento.

5 En un aspecto, la divulgación proporciona un método *in vivo* o *ex vivo* de modificación de una célula de un organismo por manipulación de una o más secuencias diana en locus genómicos de interés que comprende suministrar a la célula una composición que no es de origen natural o que está manipulada que comprende una composición de vector que codifica de forma funcional una composición como se describe en este documento. En una realización preferida el organismo es una planta o alga.

10 En un aspecto, la invención proporciona una composición o enzima como se describe en este documento para su uso en medicina o para su uso en tratamiento.

15 En un aspecto, la divulgación proporciona el uso de la composición o enzima como se describe en este documento:

en la preparación de un medicamento;

20 en la preparación de un medicamento para edición génica o genómica *ex vivo*; o

en edición génica o genómica *ex vivo*.

25 En un aspecto, la invención proporciona una composición para su uso, método o el uso como se describe en este documento para corregir defectos ocular que surgen de mutaciones genéticas.

30 Por consiguiente, determinadas realizaciones la divulgación proporcionan una composición que no es de origen natural o que está manipulada que comprende: I. dos o más secuencias polinucleotídicas del sistema de CRISPR-Cas que comprenden (a) una primera secuencia guía que puede hibridar con una primera secuencia diana en un locus polinucleotídico, (b) una segunda secuencia guía que puede hibridar con una segunda secuencia diana en un locus polinucleotídico, (c) una secuencia de acoplamiento de *crtra*, y (d) una secuencia de ARN*crtra*, y II. una enzima Cas9 de tipo II o una segunda secuencia polinucleotídica que la codifique, en la que la enzima Cas9 de tipo II es o comprende una enzima SpCas9 que comprende la mutación N863 o N863A, enzima SaCas9 que comprende la mutación N580 o N580An o un ortólogo de la misma, que tiene una mutación correspondiente a SpCas9N863 o N863A, en la que cuando se transcriben, la primera y la segunda secuencia de acoplamiento de *crtra* hibridan con la primera y segunda secuencia de ARN*crtra* respectivamente y la primera y la segunda secuencia guía dirigen la unión específica de secuencia de un primer y un segundo complejo de CRISPR a la primera y segunda secuencia diana respectivamente, en la que el primer complejo de CRISPR comprende la enzima Cas9 en complejo con (1) la primera secuencia guía que puede hibridar con la primera secuencia diana, y (2) la primera secuencia de acoplamiento de *crtra* que hibrida con la primera secuencia de ARN*crtra*, en la que el segundo complejo de CRISPR comprende la enzima Cas9 en complejo con (1) la segunda secuencia guía que puede hibridar con la segunda secuencia diana, y (2) la segunda secuencia de acoplamiento de *crtra* que hibrida con la segunda secuencia de ARN*crtra*, y en la que la primera secuencia guía dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía dirige la escisión de la otra hebra cerca de la segunda secuencia diana induciendo una rotura bicatenaria, modificando de ese modo el organismo o el organismo no humano o no animal, y en la que la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocan salientes 3'. Ventajosamente, los componentes I y II están unidos de forma funcional a uno o más elementos reguladores. Ventajosamente, el componente (I) comprende una secuencia polinucleotídica del sistema de CRISPR-Cas que comprende la secuencia guía, la secuencia de acoplamiento de *crtra* y la secuencia de ARN*crtra*. Ventajosamente, el componente (I) comprende un primer elemento regulador unido de forma funcional a la secuencia guía y la secuencia de acoplamiento de *crtra*, y un tercer elemento regulador unido de forma funcional a la secuencia de ARN*crtra*. Ventajosamente, la composición incluye un sistema de suministro configurado de forma funcional para suministrar componentes del complejo de CRISPR-Cas o secuencias polinucleotídicas que comprenden o codifican dichos componentes De forma ventajosa célula. Ventajosamente, el sistema de suministro comprende un sistema de vector que comprende uno o más vectores, y en el que los componentes I y II están ubicados en el mismo vector o en diferentes del sistema. Ventajosamente, el uno o más vectores comprenden uno o más vectores víricos. Ventajosamente, el uno o más vectores víricos comprenden uno o más vectores retrovíricos, lentivíricos, adenovíricos, víricos adenoasociados o del virus del herpes simple. Ventajosamente, el sistema de suministro comprende una nanopartícula, liposoma, exosoma, sistema de levadura, microvesícula o cañón de genes. Ventajosamente, la composición (por ejemplo, la enzima Cas9 o de CRISPR) incluye uno o más dominios funcionales. Ventajosamente, el uno o más dominios funcionales comprenden un dominio activador de la transcripción. Ventajosamente, el dominio funcional comprende VP64 o KRAB, SID o SID4X, o una recombinasa, una transposasa, un remodelador de histonas, una ADN metiltransferasa, un criptocromo, un dominio inducible/controlable por la luz o un dominio inducible/controlable químicamente. Ventajosamente, en realizaciones que implican uno o más vectores, la composición o sistema de CRISPR-Cas ase suministra mediante un solo vector. Ventajosamente, la célula es una célula eucariótica; o el uno o más vectores están configurados de forma funcional para dirigir la expresión de transcritos

de CRISPR cuando se introducen en una célula eucariótica. Ventajosamente, la secuencia de nucleótidos que codifica la SaCas9 tiene los codones optimizados para su expresión en una célula eucariótica. Ventajosamente, los elementos reguladores comprenden un promotor específico de tejido. Ventajosamente, el promotor específico de tejido dirige la expresión en músculo, neurona, hueso, piel, sangre, hígado, páncreas o linfocitos. Ventajosamente, la secuencia diana está adyacente a un motivo adyacente de protoespaciador (PAM) reconocido por la enzima Cas9. Ventajosamente, la secuencia diana está flanqueada en su extremo 3' por 5'-NRG (donde N es cualquier nucleótido) para SpCas9 o NNGRR para SaCas9. Ventajosamente, la secuencia guía puede hibridar con una secuencia diana en una célula eucariótica. Ventajosamente, la secuencia de ARNcrtra es de 30 o más nucleótidos de longitud. Ventajosamente, el ARNcrtra es de 50 o más nucleótidos de longitud. Ventajosamente, la enzima Cas9, por ejemplo, enzima SaCas9, comprende además una o más secuencias de localización nuclear (NLS). Aspectos ventajosos mencionados en este párrafo pueden aplicarse *mutatis mutandis* a otras realizaciones analizadas en este documento.

La divulgación también comprende una célula hospedadora o línea celular *in vivo*, *ex vivo* o *in vitro* que comprende o está modificada por una composición o enzima o sistema de CRISPR-Cas analizado en este documento, así como la descendencia de la misma, por ejemplo, una célula madre o línea de células madre. La divulgación también comprende un método de modificación de un organismo por manipulación de una o más secuencias diana en locus genómicos de interés, que comprende suministrar al organismo la composición o enzima o sistema de CRISPR-Cas analizado en este documento. La divulgación proporciona además un método *in vivo* o *ex vivo* de modificación de una célula de un organismo por manipulación de una o más secuencias diana en locus genómicos de interés que comprende suministrar a la célula una composición que no es de origen natural o que está manipulada que comprende una composición de vector que codifica de forma funcional una composición o enzima o sistema de CRISPR-Cas analizado en este documento de acuerdo con cualquier realización de este documento. El organismo es una planta o alga. Una composición o enzima o sistema de CRISPR-Cas analizado en este documento de acuerdo con cualquier realización de este documento puede usarse en medicina o para uso en tratamiento, por ejemplo, en la preparación de un medicamento; en la preparación de un medicamento para edición génica o genómica *ex vivo*; o en edición génica o genómica *ex vivo*. La invención también comprende una composición o enzima o sistema de CRISPR-Cas analizado en este documento de acuerdo con cualquier realización de este documento para o en uso, o métodos que implican el uso del mismo o cualquier uso mencionado en este documento para tratar, abordar, minimizar síntomas de, aliviar, o corregir defectos oculares, por ejemplo, que surgen de mutaciones genéticas. La divulgación comprende además productos posibilitados por la presente invención, por ejemplo, células mejoradas o alteradas, productos de expresión tales como productos de expresión mejorados o alterados o plantas o animales no humanos o células que tienen rasgos de la práctica de la invención.

Por consiguiente, en determinadas realizaciones, la divulgación proporciona un método de modificación de un organismo o un organismo no humano por manipulación de una primera y una segunda secuencia diana en hebras opuestas de un ADN bicatenario en un locus genómico de interés en una célula, que comprende

suministrar una composición que no es de origen natural o que está manipulada, que comprende:

I. una primera secuencia polinucleotídica de ARN quimérico (ARNqui) del sistema de CRISPR-Cas, en la que la primera secuencia polinucleotídica comprende:

(a) una primera secuencia guía que puede hibridar con la primera secuencia diana,

(b) una primera secuencia de acoplamiento de crtra, y

(c) una primera secuencia crtra,

II. una segunda secuencia polinucleotídica de ARNqui del sistema de CRISPR-Cas, en la que la segunda secuencia polinucleotídica comprende:

(a) una segunda secuencia guía que puede hibridar con la segunda secuencia diana,

(b) una segunda secuencia de acoplamiento de crtra, y

(c) una segunda secuencia crtra, y

III. una secuencia polinucleotídica que codifica una enzima de CRISPR, en la que la enzima de CRISPR es una proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma (por ejemplo, *S. aureus*) que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* con N580A), y que comprende al menos una o dos o más secuencias de localización nuclear,

en la que (a), (b) y (c) están dispuestas en una orientación 5' a 3',

en la que cuando se transcriben, la primera y la segunda secuencia de acoplamiento de crtra hibridan con la primera y segunda secuencia crtra respectivamente y la primera y la segunda secuencia guía dirigen la unión específica de secuencia de un primer y un segundo complejo de CRISPR a la primera y segunda secuencia diana respectivamente,

5 en la que el primer complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la primera secuencia guía que puede hibridar con la primera secuencia diana, y (2) la primera secuencia de acoplamiento de crtra que hibrida con la primera secuencia crtra,

10 en la que el segundo complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la segunda secuencia guía que puede hibridar con la segunda secuencia diana, y (2) la segunda secuencia de acoplamiento de crtra que hibrida con la segunda secuencia crtra,

en la que la secuencia polinucleotídica que codifica dicha enzima de CRISPR es ADN o ARN, y

15 en la que la primera secuencia guía dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía dirige la escisión de la otra hebra cerca de la segunda secuencia diana induciendo una rotura bicatenaria, modificando de ese modo el organismo o el organismo no humano, y en la que la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocan salientes 3'.

20 La presente invención, por lo tanto, puede considerarse que incluye una estrategia de nickasa doble o doble nickasa. Se apreciará que, aquí y en cualquier otro aspecto o realización de la invención, la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana para provocar salientes 3' pueden permitir lo siguiente: la secuencia de la primera guía y PAM y la segunda guía y PAM se seleccionan conjuntamente y no por aislado de modo que se compensen apropiadamente. En otras palabras, la secuencia de cada una de la primera y la segunda guía y PAM se considera con respecto a las otras para garantizar que provocarán la colocación correcta del primer y segundo complejo de CRISPR en el ADN diana para conseguir (en concierto con la Cas9 mutante) los salientes 3' requeridos. Esto se consigue por comparación de secuencias de los sitios diana en ambas hebras del ADN bicatenario con respecto a la identificación de secuencias guía y PAM adecuadas en ambas hebras y su colocación relativa (es decir, compensación resultante). Habitualmente se da consideración meticulosa a esto en cualquier caso para reducir los efectos inespecíficos, aunque normalmente solo con una única secuencia diana y PAM.

35 Preferiblemente, el vector es un vector vírico, tal como un vector retrovírico, lentivírico o baculovírico o preferiblemente adenovírico/vírico adenoasociado, pero se conocen otros medios de suministro (tales como sistemas de levadura, microvesículas, cañones de genes/medios de fijación de vectores a nanopartículas de oro) y se proporcionan. En algunas realizaciones, uno o más de los vectores víricos o plasmídicos pueden suministrarse mediante nanopartículas, exosomas, microvesículas o un cañón de genes.

40 Por manipulación de una secuencia diana, los solicitantes también indican la manipulación epigenética de una secuencia diana. Esta puede ser del estado de la cromatina de una secuencia diana, tal como por modificación del estado de metilación de la secuencia diana (es decir, adición o eliminación de metilación o patrones de metilación o islas CpG), modificación de histonas, aumento o reducción de la accesibilidad a la secuencia diana, o promoviendo o reduciendo el plegamiento 3D, o a través de activación o represión del gen (su expresión) a través de acción sobre el promotor, potenciador o silenciador.

45 Se apreciará que cuando se hace referencia a un método de modificación de un organismo, incluyendo un organismo procariótico o un organismo eucariótico tal como una planta o un animal, por ejemplo, un mamífero incluyendo ser humano o un mamífero u organismo no humano) por manipulación de una secuencia diana en un locus genómico de interés, esto puede aplicarse al organismo (o mamífero) como un conjunto o solamente una única célula o población de células de ese organismo (si el organismo es multicelular). En el caso de seres humanos, por ejemplo, los solicitantes prevén, entre otras cosas, una sola célula o una población de células y estas pueden modificarse preferiblemente *ex vivo* y después reintroducirse. En este caso, puede ser necesaria una biopsia u otra muestra de tejido o líquido biológico. Por consiguiente, se prevén métodos de tratamiento celular donde, por ejemplo, una sola célula o una población de células se muestrea o cultiva, en los que la célula o células están o se han modificado *ex vivo* como se describe en este documento, y después se reintroducen (células muestreadas) o introducen (células cultivadas) en el organismo. Las células madre, ya sean células madre embrionarias o pluripotentes inducidas o totipotentes, también son particularmente preferidas a este respecto. Pero, por supuesto, también se prevén realizaciones *in vivo*.

60 En determinadas realizaciones, la divulgación proporciona un método de modificación de un organismo o un organismo no humano o no animal por manipulación de una primera y una segunda secuencia diana en hebras opuestas de un ADN bicatenario en un locus genómico de interés en una célula, que comprende

65

suministrar una composición que no es de origen natural o que está manipulada, que comprende un sistema de vector que comprende uno o más vectores que comprenden

I. un primer elemento regulador unido de forma funcional a

(a) una primera secuencia guía que puede hibridar con la primera secuencia diana, y

(b) al menos una o más secuencias de acoplamiento de crtra,

II. un segundo elemento regulador unido de forma funcional a

(a) una segunda secuencia guía que puede hibridar con la segunda secuencia diana, y

(b) al menos una o más secuencias de acoplamiento de crtra,

III. un tercer elemento regulador unido de forma funcional a una secuencia codificante de enzima que codifica una enzima de CRISPR, en la que la enzima de CRISPR es una proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma (por ejemplo, *S. aureus*) que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* con N580A), y

IV. un cuarto elemento regulador unido de forma funcional a una secuencia crtra,

en la que los componentes I, II, III y IV están ubicados en el mismo vector o en diferentes del sistema,

cuando se transcriben, la secuencia de acoplamiento de crtra hibrida con la secuencia crtra y la primera y la segunda secuencia guía dirigen la unión específica de secuencia de un primer y un segundo complejo de CRISPR a la primera y segunda secuencia diana respectivamente,

en la que el primer complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la primera secuencia guía que puede hibridar con la primera secuencia diana, y (2) la secuencia de acoplamiento de crtra que hibrida con la secuencia crtra,

en la que el segundo complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la segunda secuencia guía que puede hibridar con la segunda secuencia diana, y (2) la secuencia de acoplamiento de crtra que hibrida con la secuencia crtra,

en la que la secuencia polinucleotídica que codifica la enzima de CRISPR es ADN o ARN, y

en la que la primera secuencia guía dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía dirige la escisión de la otra hebra cerca de la segunda secuencia diana induciendo una rotura bicatenaria, modificando de ese modo el organismo o el organismo no humano, y en la que la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocan salientes 3'.

Algunos métodos de la divulgación pueden incluir la inducción de la expresión. En algunos métodos de la divulgación, el organismo o sujeto es un eucariota, incluyendo, por ejemplo, una planta o un animal (incluyendo mamífero, incluyendo ser humano) o un eucariota no humano o un animal no humano o un mamífero no humano. En algunos métodos de la divulgación, el organismo o sujeto es una planta. En algunos métodos de la divulgación, el organismo o sujeto es un mamífero o un mamífero no humano. En algunos métodos de la divulgación, el organismo o sujeto es algas. En algunos métodos de la divulgación, el vector vírico es un AAV. En algunos métodos de la divulgación, el vector vírico es un vector derivado de retrovirus o lentivirus. En algunos métodos de la divulgación, el vector vírico es un plásmido Ti o Ri de *Agrobacterium* para su uso en plantas. En los métodos de la divulgación, la enzima de CRISPR es una nickasa Cas9 mutada. En algunos métodos de la divulgación, la expresión de la secuencia guía está bajo el control de un promotor T7 que está dirigido por la expresión de la polimerasa T7. En alguna divulgación de la invención, la expresión de la secuencia guía está bajo el control de un promotor U6.

Por manipulación de una secuencia diana, los solicitantes entienden la alteración de la secuencia diana, que puede incluir la manipulación epigenética de una secuencia diana. Esta manipulación epigenética puede ser del estado de la cromatina de una secuencia diana, tal como por modificación del estado de metilación de la secuencia diana (es decir, adición o eliminación de metilación o patrones de metilación o islas CpG), modificación de histonas, aumento o reducción de la accesibilidad a la secuencia diana, o promoviendo o reduciendo el plegamiento 3D.

Se apreciará que cuando se hace referencia a un método de modificación de un organismo o un organismo no humano por manipulación de una secuencia diana en un locus genómico de interés, esto puede aplicarse al organismo como un conjunto o solamente una única célula o población de células de ese organismo (si el organismo es multicelular).

En el caso de seres humanos, por ejemplo, los solicitantes prevén, entre otras cosas, una sola célula o una población de células y estas pueden modificarse preferiblemente *ex vivo* y después reintroducirse. En este caso, puede ser necesaria una biopsia u otra muestra de tejido o líquido biológico. Las células madre son también particularmente preferidas a este respecto. Pero, por supuesto, también se prevén realizaciones *in vivo*.

Algunos métodos de la divulgación pueden incluir la inducción de la expresión. En algunos métodos de la divulgación, el organismo o sujeto es un eucariota, incluyendo, por ejemplo, una planta o un animal (incluyendo mamífero, incluyendo ser humano) o un eucariota no humano o un animal no humano. En algunos métodos de la divulgación, el organismo o sujeto es una planta. En algunos métodos de la divulgación, el organismo o sujeto es un mamífero o un mamífero no humano. En algunos métodos de la divulgación, el organismo o sujeto es algas. En algunos métodos de la divulgación, el vector vírico es un AAV. En algunos métodos de la divulgación, el vector vírico es un vector retrovírico o lentivírico. En algunos métodos de la divulgación, el vector vírico es un virus del mosaico del tabaco. En la invención la enzima de CRISPR es una nickasa Cas9 mutada. En algunos métodos de la divulgación, la expresión de la secuencia guía está bajo el control del promotor T7 que está dirigido por la expresión de la polimerasa T7. En algunos métodos de la divulgación, la expresión de la secuencia guía está bajo el control de un promotor U6.

La divulgación en algunas realizaciones comprende un método de suministro de una enzima de CRISPR, que comprende suministrar a una célula ARNm que codifica la enzima de CRISPR. La enzima de CRISPR es una Cas9 mutada.

La divulgación en algunas realizaciones comprende un método para preparar el vector AAV de la divulgación, que comprende transfectar uno o más plásmidos que contienen o consisten esencialmente en una o más moléculas de ácido nucleico que codifican el AAV en células que pueden infectar por AAV, y aportar rep y/o cap de AAV obligatorias para la replicación y empaquetado del AAV. En algunas realizaciones, la rep y/o cap de AAV obligatorias para la replicación y empaquetado del AAV se aportan transfectando las células las células con uno o más plásmidos auxiliares o uno o más virus auxiliares. En algunas realizaciones, el virus auxiliar es un poxvirus, adenovirus, herpesvirus o baculovirus. En algunas realizaciones, el poxvirus es un virus de la variolovacuna. En algunas realizaciones, las células son células de mamífero. Y en algunas realizaciones, las células son células de insecto y el virus auxiliar es baculovirus.

En plantas, los patógenos a menudo son específicos de hospedador. Por ejemplo, *Fusarium oxisporum* f. sp. *lycopersici* causa marchitamiento del tomate, pero ataca solamente al tomate, y *F. oxisporum* f. *dianthii Puccinia graminis* f. sp. *tritici* ataca solamente al trigo. Las plantas tienen defensas existentes e inducidas para resistir la mayoría de patógenos. Las mutaciones y eventos de recombinación en todas las generaciones de la planta dan lugar a variabilidad genética que da lugar a susceptibilidad, especialmente ya que los patógenos se reproducen con más frecuencia que las plantas. En plantas, puede no haber resistencia del hospedador, por ejemplo, el hospedador y el patógeno son incompatibles. También puede haber resistencia horizontal, por ejemplo, resistencia parcial contra todas las razas de un patógeno, típicamente controlada por muchos genes y resistencia vertical, por ejemplo, resistencia completa a algunas razas de un patógeno, pero no a otras razas, típicamente controlada por unos pocos genes. A un nivel gen a gen, las plantas y los patógenos evolucionan conjuntamente, y los cambios genéticos en unos equilibra los cambios en los otros. Por consiguiente, usando la variabilidad natural, los obtentores combinan los genes más útiles para producción, calidad, uniformidad, robustez, resistencia. Las fuentes de genes de resistencia incluyen variedades naturales o exógenas, variedades tradicionales, parientes de plantas silvestres y mutaciones inducidas, por ejemplo, tratando material vegetal con agentes mutágenos. Usando la presente divulgación, los obtentores de plantas están provistos de una nueva herramienta para inducir mutaciones. Por consiguiente, un experto en la materia puede analizar el genoma de fuentes de genes de resistencia, y en variedades que tienen características o rasgos deseados emplean la presente invención para inducir la generación de genes de resistencia, con más precisión que agentes mutágenos previos y, por tanto, aceleran y mejoran los programas de fitomejoramiento.

La invención comprende una composición para su uso en medicina, como se define en las reivindicaciones. En la invención, la enzima de CRISPR comprende una mutación en el dominio HNH catalítico (N863A). La enzima de CRISPR es una nickasa Cas9. La invención comprende en algunas realizaciones una composición de la invención o una enzima de CRISPR de la misma (incluyendo o como alternativa ARNm que codifica la enzima de CRISPR), en la que la secuencia diana está flanqueada en su extremo 3' por un motivo 5' denominado motivo adyacente de protoespaciador (PAM), especialmente cuando la Cas9 es (o deriva de) Cas9 de *S. pyogenes* o *S. aureus*. Por ejemplo, un PAM adecuado es 5'-NRG o 5'-NNGRR o 5'-NNGRRT (donde N es cualquier nucleótido) para enzimas SpCas9 o SaCas9 (o enzimas derivadas), respectivamente, como se menciona a continuación.

Se apreciará que SpCas9 o SaCas9 son las de o derivadas de Cas9 de *S. pyogenes* o *S. aureus*.

En algunos métodos de la divulgación, alguna o todas de las secuencias polinucleotídicas que codifican la enzima de CRISPR, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, es o son ARN. En otras realizaciones de la invención, los polinucleótidos que comprenden la secuencia que codifica la enzima de CRISPR, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, son ARN y se suministran mediante nanopartículas, exosomas, microvesículas o un cañón de genes. En determinadas realizaciones de la

invención, la primera y segunda secuencia de acoplamiento de crtra comparten un 100 % de identidad y/o la primera y segunda secuencia crtra comparten un 100 % de identidad. En la invención, la enzima de CRISPR es una nickasa Cas9 mutada, por ejemplo, SpCas9 mutada. En la invención, la enzima de CRISPR comprende una mutación en uno de los dominios catalíticos, en la que la mutación es N863A en el dominio HNH.

5 En las realizaciones de la invención, el saliente 3' es como mucho de 150, 100 o 25 pares de bases o al menos 15, 10 o 1 pares de bases. En realizaciones preferidas, el saliente 3' es de 1-100 pares de bases.

10 La divulgación en algunas realizaciones comprende un método de modificación de un locus genómico de interés introduciendo en una célula que contiene y expresa una molécula de ADN bicatenaria que codifica el producto génico un sistema de CRISPR-Cas manipulado, que no es de origen natural que comprende proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma (por ejemplo, *S. aureus*) que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* con N580A), y dos ARN guía dirigidos a una primera hebra y una segunda hebra de la molécula de ADN respectivamente, por lo que los ARN guía están dirigidos a la molécula de ADN que codifica el producto génico y la proteína Cas mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico; en el que la proteína Cas y los dos ARN guía no se producen de forma natural conjuntamente; y en el que la proteína Cas que mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico provoca salientes 3'.

20 En algunos métodos de la divulgación, alguna o todas de la secuencia polinucleotídica que codifica la enzima de CRISPR, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, es o son ARN. En realizaciones adicionales de la invención, la primera y segunda secuencia de acoplamiento de crtra comparten un 100 % de identidad y/o la primera y segunda secuencia crtra comparten un 100 % de identidad. En la invención, la enzima de CRISPR es una nickasa Cas9 mutada, por ejemplo, SpCas9 mutada (N863A) o SaCas9 (N580A).

30 Aunque se prefiere alanina como residuo de remplazo en el mutante, están disponibles otras alternativas, con la condición de que sean catalíticamente inactivas y también retengan la función nickasa de la Cas9 y provoquen el saliente 3'. A continuación se dan directrices adecuadas, pero las alternativas preferidas a alanina pueden incluir, en algunas realizaciones, otros aminoácidos pequeños y apolares. Estos pueden incluir glicina, valina, leucina o isoleucina, de modo que en SpCas9 esto se traduciría en N863G, N863V, N863L o N863I; y en SaCas9 se traduciría en N580G, N580V, N580L o N580I.

35 En una realización adicional de la invención, uno o más de los vectores víricos se suministran mediante nanopartículas, exosomas, microvesículas o un cañón de genes.

40 La divulgación en algunas realizaciones comprende un método de modificación de un locus genómico de interés minimizando las modificaciones inespecíficas introduciendo en una célula que contiene y que expresa una molécula de ADN bicatenaria que codifica el producto génico un sistema de CRISPR-Cas manipulado, que no es de origen natural que comprende una proteína Cas mutada que tiene una mutación y dos ARN guía dirigidos a una primera hebra y una segunda hebra de la molécula de ADN respectivamente, por lo que los ARN guía están dirigidos a la molécula de ADN que codifica el producto génico y la proteína Cas mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico, por lo que la expresión del producto génico se altera; y, en el que la proteína Cas y los dos ARN guía no se producen de forma natural conjuntamente. El mutante de Cas9 es preferiblemente la N863 SpCas9, N863A SpCas9, N580 SaCas9 o N580A SaCas9, u ortólogos que tienen mutaciones correspondientes.

50 Aspectos de la invención se refieren a la expresión del producto génico que se disminuye o un polinucleótido molde que se introduce adicionalmente en la molécula de ADN que codifica el producto génico o una secuencia intermedia que se escinde permitiendo que los dos salientes 3' se rehibriden y se ligen o la actividad o función del producto génico se altere o la expresión del producto génico se aumente. En una realización de la invención, el producto génico es una proteína. La escisión de la secuencia intermedia puede tener precisión a través del uso de los salientes 3'. Por tanto, la invención prevé una secuencia intermedia que se escinde de forma precisa permitiendo que los dos salientes 3' se rehibriden y se ligen o la actividad o función del producto génico se altere o la expresión del producto génico se aumente.

60 La divulgación también comprende un sistema de CRISPR-Cas manipulado, que no es de origen natural que comprende una proteína Cas que tiene una mutación y dos ARN guía dirigidos a una primera hebra y una segunda hebra respectivamente de una molécula de ADN bicatenaria que codifica un producto génico en una célula, por lo que los ARN guías están dirigidos a la molécula de ADN que codifica el producto génico y la proteína Cas mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico, por lo que la expresión del producto génico se altera; y en el que la proteína Cas mutada y los dos ARN guía no se producen de forma natural conjuntamente. El mutante de Cas9 es preferiblemente la N863 SpCas9, N863A SpCas9, N580 SaCas9 o N580A SaCas9, u ortólogos que tienen mutaciones correspondientes.

65

En aspectos de la invención, los ARN guía pueden comprender una secuencia guía fusionada a una secuencia de acoplamiento de crtra y una secuencia crtra, es decir, una guía quimérica. En otras realizaciones, el ARN guía no es una guía quimérica. Por ejemplo, el ARN guía puede comprender una secuencia guía fusionada (en su extremo 3') a (el extremo 5' de) una secuencia de acoplamiento de crtra, proporcionándose el ARNcrtra por separado.

Se apreciará que las expresiones ARNcrtra y secuencia crtra pueden usarse indistintamente en este documento.

Aspectos de la invención se refieren a la expresión del producto génico que se disminuye o un polinucleótido molde que se introduce adicionalmente en la molécula de ADN que codifica el producto génico o una secuencia intermedia que se escinde permitiendo que los dos salientes 3' se rehibriden y se ligen o la actividad o función del producto génico se altere o la expresión del producto génico se aumente. En una realización de la invención, el producto génico es una proteína. La escisión de la secuencia intermedia puede tener precisión a través del uso de los salientes 3'. Por tanto, la invención prevé una secuencia intermedia que se escinde de forma precisa permitiendo que los dos salientes 3' se rehibriden y se ligen o la actividad o función del producto génico se altere o la expresión del producto génico se aumente.

La invención también comprende un sistema de CRISPR-Cas manipulado, que no es de origen natural que comprende proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A), y dos ARN guía dirigidos a una primera hebra y una segunda hebra respectivamente de una molécula de ADN bicatenaria que codifica un producto génico en una célula, por lo que los ARN guía están dirigidos a la molécula de ADN que codifica el producto génico y la proteína Cas mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico; y en el que la proteína Cas y los dos ARN guía no se producen de forma natural conjuntamente; y en el que la proteína Cas que mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico provoca salientes 3'.

La divulgación también comprende un sistema de vector manipulado, que no es de origen natural que comprende uno o más vectores que comprenden:

a) un primer elemento regulador unido de forma funcional a cada uno de los dos ARN guía del sistema de CRISPR-Cas dirigidos a una primera hebra y una segunda hebra respectivamente de una molécula de ADN bicatenaria que codifica un producto génico,

b) un segundo elemento regulador unido de forma funcional a una secuencia polinucleotídica que codifica proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A),

en el que los componentes (a) y (b) están ubicados en el mismo vector o en diferentes del sistema,

por lo que los ARN guía están dirigidos a la molécula de ADN que codifica el producto génico y la proteína Cas mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico; y en el que la proteína Cas y los dos ARN guía no se producen de forma natural conjuntamente; en el que la proteína Cas que mella cada una de la primera hebra y la segunda hebra de la molécula de ADN que codifica el producto génico provoca salientes 3'.

Aspectos de la divulgación proporcionan métodos de modificación de un organismo, que comprenden una primera y una segunda secuencia diana en hebras opuestas de un ADN bicatenario en un locus genómico de interés en una célula promoviendo la reparación dirigida por homología, que comprende

suministrar una composición que no es de origen natural o que está manipulada, que comprende:

I. una primera secuencia polinucleotídica de ARN quimérico (ARNqui) del sistema de CRISPR-Cas, en la que la primera secuencia polinucleotídica comprende:

(a) una primera secuencia guía que puede hibridar con la primera secuencia diana,

(b) una primera secuencia de acoplamiento de crtra, y

(c) una primera secuencia crtra,

II. una segunda secuencia polinucleotídica de ARNqui del sistema de CRISPR-Cas, en la que la segunda secuencia polinucleotídica comprende:

(a) una segunda secuencia guía que puede hibridar con la segunda secuencia diana,

(b) una segunda secuencia de acoplamiento de crtra, y

(c) una segunda secuencia crtra, y

5 III. una secuencia polinucleotídica que codifica una enzima de CRISPR, en la que la enzima de CRISPR es una proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A), que comprende al menos una o dos o más secuencias de localización nuclear,

10 IV. un molde de reparación que comprende un oligonucleótido monocatenario sintetizado o manipulado, en la que (a), (b) y (c) están dispuestas en una orientación 5' a 3',

15 en la que cuando se transcriben, la primera y la segunda secuencia de acoplamiento de crtra hibridan con la primera y segunda secuencia crtra respectivamente y la primera y la segunda secuencia guía dirigen la unión específica de secuencia de un primer y un segundo complejo de CRISPR a la primera y segunda secuencia diana respectivamente,

20 en la que el primer complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la primera secuencia guía que puede hibridar con la primera secuencia diana, y (2) la primera secuencia de acoplamiento de crtra que hibrida con la primera secuencia crtra,

en la que el segundo complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la segunda secuencia guía que puede hibridar con la segunda secuencia diana, y (2) la segunda secuencia de acoplamiento de crtra que hibrida con la segunda secuencia crtra,

25 en la que la secuencia polinucleotídica que codifica la enzima de CRISPR es ADN o ARN,

30 en la que la primera secuencia guía dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía dirige la escisión de la otra hebra cerca de la segunda secuencia diana induciendo una rotura bicatenaria; en la que la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocan salientes 3' y en la que el molde de reparación se introduce en el ADN bicatenario por recombinación homóloga, por lo que se modifica el organismo.

35 En realizaciones de la invención el molde de reparación puede comprender además un sitio de restricción para endonucleasa de restricción. En otras realizaciones, la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocan un saliente 3'. En realizaciones preferidas de la invención, el saliente 3' es de 1-100 pares de bases. En otros aspectos, alguna o todas de la secuencia polinucleotídica que codifica la enzima de CRISPR mutada, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, es o son ARN. En aspectos adicionales más, los polinucleótidos que comprenden la secuencia que codifica la enzima de CRISPR, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, son ARN y se suministran mediante nanopartículas, exosomas, microvesículas o un cañón de genes. En realizaciones adicionales de la invención, la primera y segunda secuencia de acoplamiento de crtra comparten un 100 % de identidad y/o la primera y segunda secuencia crtra comparten un 100 % de identidad. En las realizaciones de la invención, la enzima de CRISPR mutada es una enzima Cas9 mutada, por ejemplo, SpCas9 mutada (N863A) o *S. aureus* N580A.

50 Aspectos de la divulgación también proporcionan métodos de modificación de un ADN bicatenario a un locus de interés en una célula, comprendiendo el método suministrar a la célula:

I. un primer polinucleótido que comprende:

(a) una primera secuencia guía que puede hibridar con una primera secuencia diana,

55 (b) una primera secuencia de acoplamiento de crtra, y

(c) una primera secuencia crtra;

II. un segundo polinucleótido que comprende:

60 (a) una segunda secuencia guía que puede hibridar con una segunda secuencia diana,

(b) una segunda secuencia de acoplamiento de crtra, y

65 (c) una segunda secuencia crtra;

y

5 III. un tercer polinucleótido que comprende una secuencia que codifica una enzima de CRISPR, en la que la enzima de CRISPR es una proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A), y una o dos o más secuencias de localización nuclear

en el que (a), (b) y (c) en dicho primer y segundo polinucleótido están dispuestas en una orientación 5' a 3';

10 en el que la primera secuencia diana está en una primera hebra del ADN bicatenario y la segunda secuencia diana está en la hebra opuesta del ADN bicatenario, y cuando la primera y segunda secuencia guía hibridan con dichas secuencias diana en la estructura bicatenaria, los extremos 5' del primer polinucleótido y el segundo polinucleótido se compensan entre sí en al menos un par de bases de la estructura bicatenaria;

15 en el que cuando se transcriben, la primera y la segunda secuencia de acoplamiento de crtra hibridan con la primera y segunda secuencia crtra, respectivamente, y la primera y la segunda secuencia guía dirigen la unión específica de secuencia de un primer y un segundo complejo de CRISPR a la primera y segunda secuencia diana respectivamente,

20 en el que el primer complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la primera secuencia guía que puede hibridar con la primera secuencia diana, y (2) la primera secuencia de acoplamiento de crtra que hibrida con la primera secuencia crtra,

25 en el que el segundo complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la segunda secuencia guía que puede hibridar con la segunda secuencia diana, y (2) la segunda secuencia de acoplamiento de crtra que hibrida con la segunda secuencia crtra,

30 y en el que dicha primera hebra del ADN bicatenario se escinde cerca de dicha primera secuencia diana, y dicha hebra opuesta del ADN bicatenario se escinde cerca de dicha segunda secuencia diana, provocando una rotura bicatenaria con salientes 3'.

35 En algunas realizaciones de la invención, el molde de reparación es una estructura bicatenaria de oligonucleótido bicatenario sintetizado o manipulado o, en otras realizaciones, el molde de reparación se genera a partir de un trozo de ADN que se introduce en la célula y se procesa enzimáticamente. Este procesamiento enzimático puede realizarse por enzimas endógenas o por enzimas (por ejemplo, endonucleasas de restricción, nucleasas o un par de nickasas) que se han introducido en la célula de modo que los salientes compatibles se generan en el molde reparación.

40 En un aspecto, la divulgación proporciona un método de modificación de un polinucleótido diana en una célula eucariótica. En algunas realizaciones, el método comprende permitir que un complejo de CRISPR se una al polinucleótido diana para lograr la escisión de dicho polinucleótido diana modificando de ese modo el polinucleótido diana, en el que el complejo de CRISPR comprende una enzima de CRISPR mutada en complejo con una secuencia guía hibridada con una secuencia diana dentro de dicho polinucleótido diana, en el que dicha secuencia guía se une a una secuencia de acoplamiento de crtra que a su vez hibrida con una secuencia crtra. En algunas realizaciones, dicha escisión comprende escindir dos hebras en la ubicación de la secuencia diana por dicha enzima de CRISPR mutada. En algunas realizaciones, dicha escisión provoca transcripción disminuida de un gen diana. En algunas realizaciones, el método comprende además reparar dicho polinucleótido diana escindido por recombinación homóloga con un polinucleótido molde exógeno, en el que dicha reparación provoca una mutación que comprende una inserción, eliminación o sustitución de uno o más nucleótidos de dicho polinucleótido diana. En algunas realizaciones, dicha mutación provoca uno o más cambios aminoacídicos en una proteína expresada a partir de un gen que comprende la secuencia diana. En algunas realizaciones, el método comprende además suministrar uno o más vectores a dicha célula eucariótica, en el que el uno o más vectores dirigen la expresión de una o más de: la enzima de CRISPR mutada, la secuencia guía unida a la secuencia de acoplamiento de crtra y la secuencia crtra. En algunas realizaciones, dichos vectores se suministran a la célula eucariótica en un sujeto. En algunas realizaciones, dicha modificación tiene lugar en dicha célula eucariótica en un cultivo celular. En algunas realizaciones, el método comprende además aislar dicha célula eucariótica de un sujeto antes de dicha modificación. En algunas realizaciones, el método comprende además devolver dicha célula y/o células eucarióticas derivadas del mismo a dicho sujeto.

50 En un aspecto, la divulgación proporciona un método de modificación de la expresión de un polinucleótido en una célula eucariótica. En algunas realizaciones, el método comprende permitir que un complejo de CRISPR se una al polinucleótido, de modo que dicha unión provoque expresión aumentada o disminuida de dicho polinucleótido; en el que el complejo de CRISPR comprende una enzima de CRISPR mutada en complejo con una secuencia guía hibridada con una secuencia diana dentro de dicho polinucleótido, en el que dicha secuencia guía se une a una secuencia de acoplamiento de crtra que a su vez hibrida con una secuencia crtra. En algunas realizaciones, el método comprende además suministrar uno o más vectores a dichas células eucarióticas, en el que el uno o más vectores dirigen la expresión de una o más de: la enzima de CRISPR mutada, la secuencia guía unida a la secuencia de acoplamiento de crtra y la secuencia crtra.

65

En un aspecto, la divulgación proporciona un método de generación de una célula eucariótica modelo que comprende un gen de enfermedad mutado. En algunas realizaciones, un gen de enfermedad es cualquier gen asociado a un aumento en el riesgo de tener o desarrollar una enfermedad. En algunas realizaciones, el método comprende (a) introducir uno o más vectores en una célula eucariótica, en el que el uno o más vectores dirigen la expresión de una o más de: una enzima de CRISPR, una secuencia guía unida a una secuencia de acoplamiento de crtra y una secuencia crtra; y (b) permitir que un complejo de CRISPR se una a un polinucleótido diana para lograr la escisión del polinucleótido diana dentro de dicho gen de enfermedad, en el que el complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la secuencia guía que puede hibridar con la secuencia diana dentro del polinucleótido diana, y (2) la secuencia de acoplamiento de crtra que hibrida con la secuencia crtra, generando de ese modo una célula eucariótica modelo que comprende un gen de enfermedad mutado. En algunas realizaciones, dicha escisión comprende escindir dos hebras en la ubicación de la secuencia diana por dicha enzima de CRISPR. En algunas realizaciones, dicha escisión provoca transcripción disminuida de un gen diana. En algunas realizaciones, el método comprende además reparar dicho polinucleótido diana escindido por recombinación homóloga con un polinucleótido molde exógeno, en el que dicha reparación provoca una mutación que comprende una inserción, eliminación o sustitución de uno o más nucleótidos de dicho polinucleótido diana. En algunas realizaciones, dicha mutación provoca uno o más cambios aminoacídicos en la expresión de una proteína a partir de un gen que comprende la secuencia diana.

En un aspecto la divulgación proporciona un método de selección de una o más células procarióticas introduciendo una o más mutaciones en un gen en la una o más células procarióticas, comprendiendo el método: introducir uno o más vectores en la una o más células procarióticas, en el que el uno o más vectores dirigen la expresión de una o más de: una enzima de CRISPR, una secuencia guía unida a una secuencia de acoplamiento de crtra, una secuencia crtra y un molde de edición; en el que el molde de edición comprende la una o más mutaciones que anulan la escisión de la enzima de CRISPR; permitir la recombinación homóloga del molde de edición con el polinucleótido diana en la una o más células a seleccionar; permitir que un complejo de CRISPR se una a un polinucleótido diana para lograr la escisión del polinucleótido diana dentro de dicho gen, en el que el complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la secuencia guía que puede hibridar con la secuencia diana dentro del polinucleótido diana, y (2) la secuencia de acoplamiento de crtra que hibrida con la secuencia crtra, en el que la unión del complejo de CRISPR al polinucleótido diana induce muerte celular, permitiendo de ese modo seleccionar una o más células procarióticas en que se ha introducido una o más mutaciones. En una realización preferida, la enzima de CRISPR es Cas9. En otro aspecto de la divulgación, la célula a seleccionar puede ser una célula eucariótica. Aspectos de la invención permiten la selección de células específicas sin que se requiera un marcador de selección o un proceso de dos etapas que pueden incluir un sistema de contraselección.

En un aspecto, la divulgación proporciona métodos de modificación de un polinucleótido diana en una célula eucariótica. En algunas realizaciones, el método comprende permitir que un complejo de CRISPR se una al polinucleótido diana para lograr la escisión de dicho polinucleótido diana modificando de ese modo el polinucleótido diana, en el que el complejo de CRISPR comprende una enzima de CRISPR en complejo con una secuencia guía hibridada con una secuencia diana dentro de dicho polinucleótido diana, en el que dicha secuencia guía se une a una secuencia de acoplamiento de crtra que a su vez hibrida con una secuencia crtra.

En otras realizaciones, esta divulgación proporciona un método de modificación de la expresión de un polinucleótido en una célula eucariótica. El método comprende aumentar o disminuir la expresión de un polinucleótido diana usando un complejo de CRISPR que se une al polinucleótido.

Cuando se desee, para lograr la modificación de la expresión en una célula, se suministra a una célula uno o más vectores que comprenden una secuencia crtra, una secuencia guía unida a la secuencia de acoplamiento de crtra, una secuencia que codifica una enzima de CRISPR. En algunos métodos, el uno o más vectores comprenden un elemento regulador unido de forma funcional a una secuencia codificante de enzima que codifica dicha enzima de CRISPR que comprende una secuencia de localización nuclear; y un elemento regulador unido de forma funcional a una secuencia de acoplamiento de crtra y uno o más sitios de inserción para insertar una secuencia guía anterior a la secuencia de acoplamiento de crtra. Cuando se expresa, la secuencia guía dirige la unión específica de secuencia de un complejo de CRISPR a una secuencia diana en una célula. Típicamente, el complejo de CRISPR comprende una enzima de CRISPR en complejo con (1) la secuencia guía que puede hibridar con la secuencia diana, y (2) la secuencia de acoplamiento de crtra que hibrida con la secuencia crtra.

En algunos métodos, un polinucleótido diana puede inactivarse para lograr la modificación de la expresión en una célula. Por ejemplo, tras la unión de un complejo de CRISPR a una secuencia diana en una célula, el polinucleótido diana se inactiva de modo que la secuencia no se transcriba, la proteína codificada no se produzca o la secuencia no funcione como lo hace la secuencia de tipo silvestre. Por ejemplo, una proteína o secuencia codificante de microARN puede inactivarse de modo que la proteína no se produzca.

En algunas realizaciones, la enzima de CRISPR mutada tiene una mutación en el dominio HNH catalítico (N863A en Sp o N580A en Sa o una mutación correspondiente en un ortólogo), en la que cuando se transcribe, la secuencia de acoplamiento de crtra hibrida con la secuencia crtra y la secuencia guía dirige la unión específica de secuencia de un complejo de CRISPR a la secuencia diana, y en la que la enzima comprende además un dominio funcional. En algunas

realizaciones, la enzima Cas9 mutada puede fusionarse a un dominio proteínico, por ejemplo, tal como un dominio de activación transcripcional. En un aspecto, un dominio de activación transcripcional es VP64. En algunas realizaciones, un dominio de represión de la transcripción es KRAB. En algunas realizaciones, un dominio de represión de la transcripción es SID, o concatámeros de SID (es decir, SID4X). En algunas realizaciones, se proporciona una enzima modificadora epigenética. En algunas realizaciones, se proporciona un dominio de activación, que puede ser el dominio de activación P65. En este documento también se describen dominios funcionales adicionales.

La divulgación también proporciona un método de modificación de un ADN bicatenario en un locus de interés en una célula, comprendiendo el método suministrar a la célula un sistema de vector que comprende uno o más vectores que comprenden:

I. una primera secuencia polinucleotídica que comprende un elemento regulador unido de forma funcional a

(a) una primera secuencia guía que puede hibridar con una primera secuencia diana, y

(b) al menos una o más secuencias de acoplamiento de crtra,

II. una segunda secuencia polinucleotídica que comprende un segundo elemento regulador unido de forma funcional a

(a) una segunda secuencia guía que puede hibridar con una segunda secuencia diana, y

(b) al menos una o más secuencias de acoplamiento de crtra,

III. una tercera secuencia polinucleotídica que comprende un tercer elemento regulador unido de forma funcional a una secuencia que codifica una enzima de CRISPR, en la que la enzima de CRISPR es una proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A), y

IV. una cuarta secuencia polinucleotídica que comprende un cuarto elemento regulador unido de forma funcional a una secuencia crtra,

en el que los componentes I, II, III y IV están ubicados en el mismo vector o en diferentes del sistema

en el que la primera secuencia diana está en una primera hebra del ADN bicatenario y la segunda secuencia diana está en la hebra opuesta del ADN bicatenario, y cuando la primera y segunda secuencia guía hibridan con dichas secuencias diana en la estructura bicatenaria, los extremos 5' del primer polinucleótido y el segundo polinucleótido se compensan entre sí en al menos un par de bases de la estructura bicatenaria;

en el que cuando se transcriben, la primera y la segunda secuencia de acoplamiento de crtra hibridan con una secuencia crtra, y la primera y la segunda secuencia guía dirigen la unión específica de secuencia de un primer y un segundo complejo de CRISPR a la primera y segunda secuencia diana respectivamente,

en el que el primer complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la primera secuencia guía que puede hibridar con la primera secuencia diana, y (2) la primera secuencia de acoplamiento de crtra que hibrida con una secuencia crtra,

en el que el segundo complejo de CRISPR comprende la enzima de CRISPR en complejo con (1) la segunda secuencia guía que puede hibridar con la segunda secuencia diana, y (2) la segunda secuencia de acoplamiento de crtra que hibrida con una secuencia crtra,

y en el que dicha primera hebra del ADN bicatenario se escinde cerca de dicha primera secuencia diana, y dicha hebra opuesta del ADN bicatenario se escinde cerca de dicha segunda secuencia diana, provocando una rotura bicatenaria con salientes 3'.

Ventajosamente en métodos de la divulgación, alguna o todas de la secuencia polinucleotídica que codifica la enzima de CRISPR mutada, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, es o son ARN, y opcionalmente en los que alguna o todas de I, II y III se suministran mediante nanopartículas, exosomas, microvesículas o un cañón de genes. En métodos de la divulgación, ventajosamente la primera y segunda secuencia de acoplamiento de crtra pueden compartir un 100 % de identidad y/o la primera y segunda secuencia crtra comparten un 100 % de identidad. Por ejemplo, cada una de I, II y III puede proporcionarse en un vector, opcionalmente en el que cada una se proporciona en el mismo vector o en uno diferente. El locus de interés en métodos de la invención puede comprender un gen y en el que dicho método provoca un cambio en la expresión de dicho gen, o un cambio en la actividad o función del producto génico. Por ejemplo, el producto génico puede ser una proteína, y/o en el que dicho cambio en la expresión, actividad o función es una reducción en dicha expresión, actividad o función.

Métodos de la divulgación pueden comprender además el suministro de moldes, tales como moldes de reparación, que pueden ser ODNbc o ODNmc, véase a continuación. El suministro de moldes puede ser simultáneo o separado del suministro de alguna o rodas de la enzima de CRISPR, la guía, la secuencia de acoplamiento de crtra o de ARNcrtra y mediante el mismo mecanismo de suministro o uno diferente. En algunas realizaciones, se prefiere que el molde se suministre junto con la guía, la secuencia de acoplamiento de crtra y/o de ARNcrtra y, preferiblemente, también la enzima de CRISPR. Un ejemplo puede ser un vector AAV donde la enzima de CRISPR es SaCas9 (con la mutación N580).

Métodos de la divulgación pueden comprender además: (a) suministrar a la célula un oligodesoxinucleótido bicatenario (ODNbc) que comprende salientes complementarios a los salientes creados por dicha rotura bicatenaria, en los que dicho ODNbc se integra en el locus de interés; o -(b) suministrar a la célula un oligodesoxinucleótido monocatenario (ODNmc), en los que dicho ODNmc actúa como molde para la reparación dirigida por homología de dicha rotura bicatenaria. Los métodos de la invención pueden ser para la prevención o tratamiento de enfermedad en un individuo, opcionalmente en los que dicha enfermedad está causada por un defecto en dichos locus de interés. Los métodos de la invención pueden realizarse *in vivo* en el individuo o *ex vivo* en una célula recogida del individuo, opcionalmente en los que dicha célula se devuelve al individuo.

La divulgación también proporciona un kit o composición que comprende:

I. un primer polinucleótido que comprende:

(a) una primera secuencia guía que puede hibridar con una primera secuencia diana,

(b) una primera secuencia de acoplamiento de crtra, y

(c) una primera secuencia crtra;

II. un segundo polinucleótido que comprende:

(a) una segunda secuencia guía que puede hibridar con una segunda secuencia diana,

(b) una segunda secuencia de acoplamiento de crtra, y

(c) una segunda secuencia crtra;

y

III. un tercer polinucleótido que comprende una secuencia que codifica una enzima de CRISPR, en la que la enzima de CRISPR es una proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A), y una o dos o más secuencias de localización nuclear

en el que (a), (b) y (c) en dicho primer y segundo polinucleótido están dispuestas en una orientación 5' a 3';

en el que la primera secuencia diana está en una primera hebra de un ADN bicatenario y la segunda secuencia diana está en la hebra opuesta del ADN bicatenario, y cuando la primera y segunda secuencia guía hibridan con dichas secuencias diana en la estructura bicatenaria, los extremos 5' del primer polinucleótido y el segundo polinucleótido se compensan entre sí en al menos un par de bases de la estructura bicatenaria,

y opcionalmente en el que cada uno de I, II y III se proporciona en el mismo vector o en uno diferente; y en el que la primera secuencia guía que dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía que dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocan salientes 3'.

La divulgación también proporciona uso de un kit o composición de acuerdo con la invención en un método de la divulgación. La divulgación también proporciona uso de un kit o composición de la divulgación en la fabricación de un medicamento, opcionalmente en el que dicho medicamento es para la prevención o tratamiento de una enfermedad causada por un defecto en dicho locus de interés.

Un ejemplo de una secuencia de codones optimizados, en este caso optimizada para seres humanos (es decir, que se optimiza para su expresión en seres humanos) se proporciona en este documento, véase la secuencia de codones optimizados para seres humanos de SaCas9. Aunque se prefiere esto, se apreciará que otros ejemplos son posibles y la optimización de codones para una especie hospedadora distinta del ser humano, o para la optimización de codones para órganos específicos tal como el cerebro, es conocida.

En un aspecto, el suministro es en forma de un vector. En un aspecto, el vector puede ser un vector vírico, tal como un vector retrovírico, lentivírico o baculovírico o preferiblemente adenovírico/vírico adenoasociado, pero se conocen otros medios de suministro (tales como sistemas de levadura, microvesículas, cañones de genes/medios de fijación de vectores a nanopartículas de oro) y se proporcionan. Un vector puede indicar no solamente un sistema vírico o de levadura (por ejemplo, donde los ácidos nucleicos de interés pueden estar unidos de forma funcional a y bajo el control de (en términos de expresión, tal como proporcionar finalmente un ARN procesado) un promotor), sino también el suministro directo de ácidos nucleicos a una célula hospedadora. Aunque en los métodos de este documento el vector puede ser un vector vírico y este es ventajosamente un AAV, pueden emplearse otros vectores víricos como se analiza en este documento. Por ejemplo, pueden usarse baculovirus para la expresión en células de insecto. Estas células de insecto, a su vez, pueden ser útiles para producir grandes cantidades de vectores adicionales, tales como vectores AAV adaptados para suministro de la presente invención. También se prevé un método de suministro de la presente enzima de CRISPR mutada, que comprende suministrar a una célula ARNm que codifica la enzima de CRISPR mutada. Se apreciará que la enzima de CRISPR está truncada, compuesta de menos de mil aminoácidos o menos de cuatro mil aminoácidos, es una nucleasa o nickasa, tiene los codones optimizados, comprende una o más mutaciones y/o comprende una enzima de CRISPR quimérica, o las otras opciones como se analiza en este documento. Se prefieren vectores víricos AAV, especialmente para el suministro de mutantes de SaCas9.

En determinadas realizaciones, la secuencia diana está flanqueada o seguida, en su extremo 3', de un PAM adecuado para la enzima de CRISPR, típicamente una Cas y en particular una Cas9.

Por ejemplo, un PAM adecuado es 5'-NRG para SpCas9 (o enzimas derivadas), o 5'-NNGRR o 5'-NNGRRT para enzimas SaCas9 (o enzimas derivadas).

Se apreciará que SpCas9 o SaCas9 son las de o derivadas de Cas9 de *S. pyogenes* o *S. aureus*, incluyendo *S. aureus* subespecie *aureus*.

Por consiguiente, en determinadas realizaciones o aspectos de la divulgación, se dice proporcionar un método de modificación de un organismo (o un organismo no humano), por ejemplo, por manipulación de una primera y una segunda secuencia diana en hebras opuestas de un ADN bicatenario en un locus genómico de interés en una célula, que comprende suministrar una composición que no es de origen natural o que está manipulada. En dichas realizaciones o aspectos, se apreciará que el organismo no es un animal. En algunas realizaciones o aspectos, se apreciará que lo proporcionado es una composición para su uso en un método de manipulación genética o genómica o para su uso en un método de modificación de un organismo (o un organismo no humano o no animal), por ejemplo, por manipulación de una primera y una segunda secuencia diana en hebras opuestas de un ADN bicatenario en un locus genómico de interés en una célula. El uso aquí, en algunas realizaciones, puede ser por (o comprender) suministro de la composición.

La expresión "que no es de origen natural o que está manipulada", por ejemplo, con respecto a una composición, es opcional y, cuando está presente, en determinadas realizaciones, puede sustituirse o eliminarse.

La presente invención se basa en varios efectos técnicos, que se definen en general, *inter alia*, por uno o más de: doble mellado optimizado; generación de salientes 3'; inhibición de NHEJ; y eficacia de HDR mejorada.

La presente invención se basa en el efecto técnico de eficacia de HDR mejorada usando el mutante SpCas9N863A o un ortólogo del mismo que tenga una mutación correspondiente a SpCas9N863A, tal como *S. aureus* N580A. Específicamente, la eficacia de HDR mejorada es el resultado de la inhibición de eventos de NHEJ y, por tanto, un sesgo (es decir, aumento) en los eventos de HDR.

La presente invención también se basa en el efecto técnico de doble mellado optimizado debido a la selección de la secuencia diana óptima de modo que las secuencias 5' PAM se enfrenten entre sí. Los PAM enfrentados entre sí se muestran en la figura 1. Cada Cas9 se recluta a un locus genómico por una secuencia guía que se une a una secuencia diana en el genoma. La secuencia diana y el PAM se encuentran cercanos, pero en hebras opuestas del ADN (y, por tanto, los PAM se enfrentan entre sí). Los PAM se asocian con dominios PI (de interacción con PAM) en la Cas9. Las secuencias guía se seleccionan con vistas a la colocación del PAM ya que el PAM también es crucial para un reclutamiento eficaz de Cas9. Por tanto, los PAM (y también las secuencias guía), por lo tanto, se seleccionan de forma óptima en el presente sistema de mellado (es decir, doble mellado) de modo que los PAM estén distales entre sí. Esto es distal en términos de los opuesto a proximal. Se puede considerar que los PAM se enfrentan o están distales entre sí si el número de nucleótidos entre ellos es de al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35 o al menos 40 o más nucleótidos entre ellos. Estos pueden contarse en la dirección 3' a 5' desde un PAM, a lo largo de su hebra de ADN hasta uno opuesto del otro PAM (o hasta que uno llegue a la secuencia complementaria correspondiente al otro PAM (en la otra hebra de ADN)). En la figura 1 puede observarse que los PAM se encuentran en esta disposición distal. Por tanto, la compensación de guía es preferiblemente positiva. Sería negativa si los PAM estuvieran proximales entre sí, por ejemplo, cuando las dos Cas9 de la figura 1 se cambiaran de orden (sometidas a al menos un conjunto de cambios de guía y PAM). En otras palabras, de nuevo con referencia a la figura 1, con cada ARNgs dispuesto 5' a 3' en hebras opuestas, los PAM se enfrentan entre sí ya que de izquierda a derecha una hebra va de 3' a 5' y la otra va de 5' a 3'. Por tanto, los extremos 3' de cada PAM se señalan en dirección

- 5 contraria entre sí, es decir, están distales (lo más lejos) entre sí, mientras que los extremos 5' de cada PAM apuntan unos hacia los otros, es decir, están proximales (lo más cerca entre sí). Si los PAM no se enfrentan entre sí, los extremos 5' de cada PAM apuntarían unos hacia los otros, es decir, estarían distales entre sí, mientras que los extremos 3' de cada PAM apuntarían unos hacia los otros, es decir, estarían proximales. Dicho de otra manera, los PAM se enfrentan si el extremo 3' de un PAM señala en dirección contraria del extremo 3' del otro PAM, mientras que el extremo 5' de un PAM apunta hacia el extremo 5' del otro PAM.
- 10 En este documento se describen longitudes óptimas de salientes 3', pero varían de 1 a 10, 20, 30, 40, 50, 60, 70, 80, 90, o 100, tal como de 1 a 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13,14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100 nucleótidos en cada extremo saliente 3'. La compensación entre el extremo 5' de cada par de guías es, en algunas realizaciones de 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 o 60 nucleótidos.
- 15 También se prefieren intervalos de aproximadamente 15-60, 16-60, 17-60, 18-60, 19-60, 20-60, 21-60, 22-60, 23-60, 24-60, 25-60, 15-55, 16-55, 17-55, 18-55, 19-55, 20-55, 21-55, 22-55, 23-55, 24-, 25-55, 35-60, 15-40, 16-40, 17-40, 18-40, 19-40, 20-40, 21-40, 22-40, 23-40, 24-40, 25-40, 15-45, 16-45, 17-45, 18-45, 19-45, 20-45, 21-45, 22-45, 23-45, 24-45, 25-45, 30-50, 35-55, y especialmente 35-45 en algunas realizaciones.
- 20 En algunas realizaciones, la alteración fenotípica es preferiblemente el resultado de modificación genómica cuando se aborda una enfermedad genética, especialmente en métodos de tratamiento y preferiblemente cuando se proporciona un molde de reparación para corregir o alterar el fenotipo.
- 25 En algunas realizaciones, enfermedades que pueden abordarse incluyen las relacionadas con defectos de corte y empalme que causan enfermedad.
- En algunas realizaciones, las dianas celulares incluyen células madre/progenitoras hematopoyéticas (CD34+); linfocitos T humanos; y células del ojo (células retinianas), por ejemplo, células precursoras fotorreceptoras.
- 30 En algunas realizaciones las dianas génicas incluyen: beta globina humana - HBB (para tratar la anemia falciforme, incluyendo por estimulación de la conversión génica (usando el gen HBD muy relacionado como molde endógeno)); CD3 (linfocitos T); y CEP920 - retina (ojo).
- 35 En algunas realizaciones, las dianas de enfermedad también incluyen: cáncer; anemia falciforme (basada en una mutación puntual); VIH; beta-talasemia; y enfermedad oftálmica, por ejemplo, defecto de corte y empalme que causa amaurosis congénita de Leber (LCA).
- 40 En algunas realizaciones los métodos de suministro incluyen: Suministro "directo" mediado por lípidos catiónicos del complejo de enzima-guía (RiboNucleoProteína) y electroporación de ADN plasmídico.
- 45 Además, la presente invención se basa en el efecto técnico de que la actividad nucleasa del mutante SpCas9N863A o un ortólogo del mismo que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A) siempre provoca escisión de la hebra en la hebra no complementaria. En estas condiciones, se generan salientes 3' si la secuencia diana (como se definen por el ARNGs) en las hebras individuales están dispuestas de modo que las secuencias 5' PAM correspondientes (ubicadas inmediatamente 3' a las secuencias diana) se enfrentan entre sí (están en hebras opuestas, por ejemplo, como se ilustra y analiza en este documento). Dicho ortólogo puede ser una Cas9 de *S. aureus* mutada, es decir, *S. aureus* N580, especialmente N580A.
- 50 La Cas9n N863A (Sp) o N580A (Sa), u ortólogos que tienen mutaciones correspondientes, mellan selectivamente la hebra no complementaria, véanse, por ejemplo, las mellas representadas con triángulos amarillos en la figura 1.
- La compensación de guía, por ejemplo, compensación de ARNGs, se define preferiblemente como la distancia entre el extremo 5' (o PAM-distal) de cada ARNGs.
- 55 Además, la presente invención se basa en un efecto técnico de que los salientes 3' provocan inhibición de NHEJ. El efecto técnico de la generación dirigida de salientes 3' es eficacia de HDR mejorada, ya que dichos salientes provocan inhibición de eventos de NHEJ y, por tanto, un sesgo (es decir, aumento) en los eventos de HDR (es decir, eficacia de HDR mejorada y/o formación de indel reducida).
- 60 Una eficacia de HDR mejorada se considera una mayor frecuencia de eventos de HDR (y/o formación de indel reducida) como resultado de la actividad doble nickasa resultante del uso del mutante SpCas9N863A o un ortólogo que tiene una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A) en comparación con la actividad doble nickasa resultante de una SpCas9 que no comprende la mutación N863A o un ortólogo que no comprende una mutación correspondiente a SpCas9N863A (por ejemplo, *S. aureus* N580A).
- 65

Realizando los métodos de la divulgación de modificación de un organismo o un locus genómico de interés, los expertos en la materia llegan inevitablemente a modificaciones inespecíficas minimizadas. Las composiciones de la invención llegan a modificaciones inespecíficas minimizadas cuando se usan.

5 En algunos aspectos y realizaciones, puede suministrarse una nickasa Cas9 de un solo tipo, por ejemplo, una nickasa SpCas9 o una SaCas9. Esto provoca que el ADN diana se una por las dos (2) SpCas9 o dos (2) SaCas9. Sin embargo, también se prevé que puedan usarse los diferentes ortólogos de Cas9, un ortólogo de Cas9 en la hebra codificante del ADN y otro ortólogo de Cas9 en la hebra de ADN no codificante u opuesta. Por ejemplo, una SpCas9 podría usarse en una hebra y una SaCas9 podría usarse en otra hebra. Como alternativa, una SpCas9 podría usarse en una hebra y un ortólogo de Cas9 podría usarse en otra hebra, o una SaCas9 podría usarse en una hebra y un ortólogo de Cas9 podría usarse en otra hebra. Usar dos Cas9, pero diferentes, requerirá el suministro o expresión generalizada de una Cas9 adicional. Sin embargo, puede ser ventajoso hacerlo ya que dos ortólogos diferentes de Cas9 requieren diferentes PAM y también pueden tener diferentes necesidades de guía, permitiendo, por tanto, un mayor control para el usuario, especialmente si uno o los dos ortólogos son controlables, por ejemplo, inducibles; y más especialmente si cada uno de los dos ortólogos es controlable o inducible por separado, por ejemplo, cada uno se controla o induce mediante un activador diferente (aunque cada uno podría controlarse o inducirse por el mismo activador).

A continuación se proporcionan directrices con respecto a la longitud de la guía (la secuencia espaciadora o guía). En algunas realizaciones, para Sp, la longitud óptima de guía puede variar tan poco como la "guía tru" de 17 nucleótidos de Keith Joung. En algunas realizaciones, para Sa, la longitud óptima de guía puede ser de 20 o 21 o 22 o 23 o 24 nucleótidos de longitud (Ran 2015).

También se proporciona una célula o línea de células hospedadoras. Esta puede ser una célula o línea de células hospedadoras *in vivo*, *ex vivo* o *in vitro*. La célula o línea de células hospedadoras, en algunas realizaciones, puede comprender o haberse modificado por la composición o enzima de acuerdo con la presente invención. También se proporciona la descendencia de dicha célula o línea de células hospedadoras. En algunas realizaciones, las células de la célula hospedadora, línea celular o descendencia son células madre o línea de células madre.

Pueden usarse métodos, productos y usos descritos en este documento con fines no terapéuticos. Además, cualquiera de los métodos descritos en este documento puede aplicarse *in vitro* y *ex vivo*.

En relación con las guías en general, pero específicamente con respecto al ARNGs y el complejo de CRISPR formado con el mismo, es preferible que la guía tenga uno o más de los siguientes elementos. En algunas realizaciones, la secuencia crtra tiene una o más horquillas y es de 30 o más nucleótidos de longitud, más preferiblemente de 40 o más nucleótidos de longitud, o más preferiblemente de 50 o más nucleótidos de longitud. En algunas realizaciones, la secuencia guía es de entre 10 y 30 nucleótidos de longitud. En algunas realizaciones, la enzima de CRISPR/Cas es una enzima Cas9 de tipo II. En algunas realizaciones, la secuencia crtra tiene una o más horquillas y es de 30 o más nucleótidos de longitud, más preferiblemente de 40 o más nucleótidos de longitud, o más preferiblemente de 50 o más nucleótidos de longitud, la secuencia guía es de entre 10 y 30 nucleótidos de longitud y la enzima de CRISPR/Cas es una enzima Cas9 de tipo II.

Breve descripción de los dibujos

Los elementos novedosos de la invención se exponen con particularidad en las reivindicaciones adjuntas. Se obtendrá una mejor comprensión de los elementos y ventajas de la presente invención por referencia a la siguiente descripción detallada que expone realizaciones ilustrativas, en que se utilizan los principios de la invención, y los dibujos adjuntos, de los que:

Figura 1: Diagrama de enzimas Cas9n en una configuración de doble mellado. El mellado de compensación con el mutante D10A, que retiene solamente la actividad catalítica del dominio nucleasa HNH, genera productos con saliente 5' en el genoma diana por mellado de la hebra de ADN complementaria de ARNGs (mellas representadas con triángulos rojos). Como alternativa, Cas9n N863A mella selectivamente la hebra no complementaria (mellas representadas con triángulos amarillos). La compensación de ARNGs se define como la distancia entre el extremo 5' (o PAM-distal) de cada ARNGs. Las secuencias PAM, representadas en verde, están presentes en el genoma diana, pero no en el ARNGs.

Figura 2: El doble mellado reduce la modificación inespecífica. (A) Diagrama de un par de ARNGs de doble mellado con Cas9n D10A para el locus EMX1 humano. Las secuencias guía se muestran en azul, demostrando una compensación de 23 pb. El PAM se muestra en rosa, y los sitios de mellado se representan mediante triángulos rojos. Se enumeran cinco sitios inespecíficos genómicos conocidos (Hsu *et al.*, 2013) para ARNGs1. (B) Resultados de SURVEYOR ejemplares que muestran modificación del locus EMX1 por Cas9 WT y Cas9n junto con ARNGs 1 y/o 2. (C) Cuantificación por secuenciación exhaustiva de modificaciones inespecíficas en cinco sitios inespecíficos conocidos por Cas9 WT y ARNGs 1 o Cas9n con ARNGs 1 y 2.

Figura 3: Diseño general de moldes de HDR ODNmc. El ODNmc consiste en una secuencia de inserción (rojo) flanqueada por brazos de homología en los lados izquierdo y derecho (de al menos 40 pb cada uno). La homología

entre el ODNmc y su región de dirección se indica mediante guiones negros. Los sitios diana de Cas9 de doble mellado se muestran en azul, y sus secuencias PAM correspondientes se muestran en rosa. Los sitios de mellado se representan por triángulos rojos.

5 **Figura 4A-E:** Representación circular del análisis filogenético que revela cinco familias de Cas9, incluyendo tres grupos de Cas9 grandes (~1400 aminoácidos) y dos de Cas9 pequeñas (~1100 aminoácidos); representación circular de familias de CRISPR.

10 **Figura 5A-F:** Representación lineal del análisis filogenético que revela cinco familias de Cas9, incluyendo tres grupos de Cas9 grandes (~1400 aminoácidos) y dos de Cas9 pequeñas (~1100 aminoácidos).

Figura 6: Gráfico que representa la distribución de longitudes de ortólogos de Cas9.

15 **Figura 7A-M:** Secuencia del gen de SpCas9 donde los puntos de mutación están ubicados dentro de la secuencia.

La **figura 8** ilustra ambos salientes 5' y 3' en sistemas de una sola nickasa y de doble nickasa con SpCas9, pero esto se aplica igualmente a SaCas9 y otros ortólogos.

20 La **figura 9A-B** ilustra el resultado de un experimento que muestra la actividad nickasa de los mutantes D10A y N580A de SaCas9. El panel A ilustra la secuencia del locus diana para 5 ARNg indicados en gris y muestra la actividad de las enzimas mutantes con las guías indicadas. El % de NHEJ en el eje de ordenadas representa las tasas de escisión específica medidas por secuenciación TOPO. El panel B muestra Cas9 de *S. aureus* de tipo silvestre con los ARNg indicados dirigidos a cinco locus diferentes. El % de NHEJ en el eje de ordenadas representa las tasas de escisión específica medidas por ensayo de T7E1.

25 **Descripción detallada de la invención y de la presente divulgación**

30 Con respecto a información general sobre sistemas de CRISPR-Cas, componentes de los mismos y suministro de dichos componentes, incluyendo métodos, materiales, vehículos de suministro, partículas, AAV y la preparación y uso de los mismos, incluyendo sobre cantidades y formulaciones, todo ello útil en la práctica de la presente invención, se hace referencia a: patentes de Estados Unidos n.º 8.999.641, 8.993.233, 8.945.839, 8.932.814, 8.906.616, 8.895.308, 8.889.418, 8.889.356, 8.871.445, 8.865.406, 8.795.965, 8.771.945 y 8.697.359; publicaciones de patente de Estados Unidos US 2014-0310830 (sol. de Estados Unidos n.º de serie 14/105.031), US 2014-0287938 A1 (sol. de Estados Unidos n.º de serie 14/213.991), US 2014-0273234 A1 (sol. de Estados Unidos n.º de serie 14/293.674), US 2014-0273232 A1 (sol. de Estados Unidos n.º de serie 14/290.575), US 2014-0273231 (sol. de Estados Unidos n.º de serie 14/259.420), US 2014-0256046 A1 (sol. de Estados Unidos n.º de serie 14/226.274), US 2014-0248702 A1 (sol. de Estados Unidos n.º de serie 14/258.458), US 2014-0242700 A1 (sol. de Estados Unidos n.º de serie 14/222.930), US 2014-0242699 A1 (sol. de Estados Unidos n.º de serie 14/183.512), US 2014-0242664 A1 (sol. de Estados Unidos n.º de serie 14/104.990), US 2014-0234972 A1 (sol. de Estados Unidos n.º de serie 14/183.471), US 2014-0227787 A1 (sol. de Estados Unidos n.º de serie 14/256.912), US 2014-0189896 A1 (sol. de Estados Unidos n.º de serie 14/105.035), US 2014-0186958 (sol. de Estados Unidos n.º de serie 14/105.017), US 2014-0186919 A1 (sol. de Estados Unidos n.º de serie 14/104.977), US 2014-0186843 A1 (sol. de Estados Unidos n.º de serie 14/104.900), US 2014-0179770 A1 (sol. de Estados Unidos n.º de serie 14/104.837) y US 2014-0179006 A1 (sol. de Estados Unidos n.º de serie 14/183.486), US 2014-0170753 (sol. de Estados Unidos n.º de serie 14/183.429); patentes europeas EP 2 784 162 B1 y EP 2 771 468 B1; solicitudes de patente europea EP 2 771 468 (EP1381857.0), EP 2 764 103 (EP13824232.6) y EP 2 784 162 (EP14170383.5); y publicaciones de patente PCT WO 2014/093661 (PCT/US2013/074743), WO 2014/093694 (PCT/US2013/074790), WO 2014/093595 (PCT/US2013/074611), WO 2014/093718 (PCT/US2013/074825), WO 2014/093709 (PCT/US2013/074812), WO 2014/093622 (PCT/US2013/074667), WO 2014/093635 (PCT/US2013/074691), WO 2014/093655 (PCT/US2013/074736), WO 2014/093712 (PCT/US2013/074819), WO2014/093701 (PCT/US2013/074800), WO2014/018423 (PCT/US2013/051418), WO 2014/204723 (PCT/US2014/041790), WO 2014/204724 (PCT/US2014/041800), WO 2014/204725 (PCT/US2014/041803), WO 2014/204726 (PCT/US2014/041804), WO 2014/204727 (PCT/US2014/041806), WO 2014/204728 (PCT/US2014/041808), WO 2014/204729 (PCT/US2014/041809). También se hace referencia a las solicitudes de patente provisional de Estados Unidos 61/758.468; 61/802.174; 61/806.375; 61/814.263; 61/819.803 y 61/828.130, presentadas el 30 de enero de 2013; 15 de marzo de 2013; 28 de marzo de 2013; 20 de abril de 2013; 6 de mayo de 2013 y 28 de mayo de 2013 respectivamente. También se hace referencia a la solicitud de patente provisional de Estados Unidos 61/836.123, presentada el 17 de junio de 2013. Adicionalmente se hace referencia a las solicitudes de patente provisional de Estados Unidos 61/835.931, 61/835.936, 61/836.127, 61/836.101, 61/836.080 y 61/835.973, cada una presentada el 17 de junio de 2013. Se hace referencia adicional a las solicitudes de patente provisional de Estados Unidos 61/862.468 y 61/862.355 presentadas el 5 de agosto de 2013; 61/871.301 presentada el 28 de agosto de 2013; 61/960.777 presentada el 25 de septiembre de 2013 y 61/961.980 presentada el 28 de octubre de 2013. Además se hace referencia también a: solicitudes de patente PCT n.º: PCT/US2014/041803, PCT/US2014/041800, PCT/US2014/041809, PCT/US2014/041804 y PCT/US2014/041806, cada una presentada el 10 de junio de 2014 6/10/14; PCT/US2014/041808 presentada el 11 de junio de 2014; y PCT/US2014/62558 presentada el 28 de octubre de 2014, y solicitudes de patente provisional de Estados Unidos n.º de serie: 61/915.150, 61/915.301, 61/915.267 y 61/915.260, cada una presentada el 12 de diciembre de 2013;

61/757.972 y 61/768.959, presentadas el 29 de enero de 2013 y 25 de febrero de 2013; 61/835.936, 61/836.127, 61/836.101, 61/836.080, 61/835.973, y 61/835.931, presentadas el 17 de junio de 2013; 62/010.888 y 62/010.879, ambas presentadas el 11 de junio de 2014; 62/010.329 y 62/010.441, cada una presentada el 10 de junio de 2014; 61/939.228 y 61/939.242, cada una presentada el 12 de febrero de 2014; 61/980.012, presentada el 15 de abril de 2014; 62/038.358, presentada el 17 de agosto de 2014; 62/054.490, 62/055.484, 62/055.460 y 62/055.487, cada una presentada el 25 de septiembre de 2014; y 62/069.243, presentada el 27 de octubre de 2014. También se hace referencia a solicitudes de patente provisional de Estados Unidos n.º 62/055.484, 62/055.460 y 62/055.487, presentadas el 25 de septiembre de 2014; solicitud de patente provisional de Estados Unidos 61/980.012, presentada el 15 de abril de 2014; y solicitud de patente provisional de Estados Unidos 61/939.242 presentada el 12 de febrero de 2014. Se hace referencia a la solicitud PCT que designa, entre otras, la solicitud de Estados Unidos n.º PCT/US14/41806, presentada el 10 de junio de 2014. Se hace referencia a la solicitud de patente provisional de Estados Unidos 61/930.214 presentada el 22 de enero de 2014. Se hace referencia a solicitudes de patente provisional de Estados Unidos 61/915.251; 61/915.260 y 61/915.267, cada una presentada el 12 de diciembre de 2013. Se hace referencia a la solicitud de patente provisional de Estados Unidos USSN 61/980.012 presentada el 15 de abril de 2014. Se hace referencia a la solicitud PCT que designa, entre otras, la solicitud de Estados Unidos n.º PCT/US14/41806, presentada el 10 de junio de 2014. Se hace referencia a la solicitud de patente provisional de Estados Unidos 61/930.214 presentada el 22 de enero de 2014. Se hace referencia a solicitudes de patente provisional de Estados Unidos 61/915.251; 61/915.260 y 61/915.267, cada una presentada el 12 de diciembre de 2013.

También se hace mención de la solicitud de Estados Unidos 62/091.455, presentada el 12-dic-14, PROTECTED GUIDE RNAS (PGRNAS); solicitud de Estados Unidos 62/096.708, 24-dic-14, PROTECTED GUIDE RNAS (PGRNAS); solicitud de Estados Unidos 62/091.462, 12-dic-14, DEAD GUIDES FOR CRISPR TRANSCRIPTION FACTORS; solicitud de Estados Unidos 62/096.324, 23-dic-14, DEAD GUIDES FOR CRISPR TRANSCRIPTION FACTORS; solicitud de Estados Unidos 62/091.456, 12-dic-14, ESCORTED AND FUNCTIONALIZED GUIDES FOR CRISPR-CAS SYSTEMS; solicitud de Estados Unidos 62/091.461, 12-dic-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR GENOME EDITING AS TO HEMATOPOETIC STEM CELLS (HSCs); solicitud de Estados Unidos 62/094.903, 19-dic-14, UNBIASED IDENTIFICATION OF DOUBLE-STRAND BREAKS AND GENOMIC REARRANGEMENT BY GENOME-WISE INSERT CAPTURE SEQUENCING; solicitud de Estados Unidos 62/096.761, 24-dic-14, ENGINEERING OF SYSTEMS, METHODS AND OPTIMIZED ENZYME AND GUIDE SCAFFOLDS FOR SEQUENCE MANIPULATION; solicitud de Estados Unidos 62/098.059, 30-dic-14, RNA-TARGETING SYSTEM; solicitud de Estados Unidos 62/096.656, 24-dic-14, CRISPR HAVING OR ASSOCIATED WITH DESTABILIZATION DOMAINS; solicitud de Estados Unidos 62/096.697, 24-dic-14, CRISPR HAVING OR ASSOCIATED WITH AAV; solicitud de Estados Unidos 62/098.158, 30-dic-14, ENGINEERED CRISPR COMPLEX INSERTIONAL TARGETING SYSTEMS; solicitud de Estados Unidos 62/151.052, 22-abr-15, CELLULAR TARGETING FOR EXTRACELLULAR EXOSOMAL REPORTING; solicitud de Estados Unidos 62/054.490, 24-sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING PARTICLE DELIVERY COMPONENTS; solicitud de Estados Unidos 62/055.484, 25-sep-14, SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; solicitud de Estados Unidos 62/087.537, 4-dic-14, SYSTEMS, METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; solicitud de Estados Unidos 62/054.651, 24-sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO; solicitud de Estados Unidos 62/067.886, 23-oct-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO; solicitud de Estados Unidos 62/054.675, 24-sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN NEURONAL CELLS/TISSUES; solicitud de Estados Unidos 62/054.528, 24-sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN IMMUNE DISEASES OR DISORDERS; solicitud de Estados Unidos 62/055.454, 25-sep-14, DELIVERY, USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING CELL PENETRATION PEPTIDES (CPP); solicitud de Estados Unidos 62/055.460, 25-sep-14, MULTIFUNCTIONAL-CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES; solicitud de Estados Unidos 62/087.475, 4-dic-14, FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; solicitud de Estados Unidos 62/055.487, 25-sep-14, FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS; solicitud de Estados Unidos 62/087.546, 4-dic-14, MULTIFUNCTIONAL CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES; y solicitud de Estados Unidos 62/098.285, 30-dic-14, CRISPR MEDIATED IN VIVO MODELING AND GENETIC SCREENING OF TUMOR GROWTH AND METASTASIS.

Cada una de estas patentes, publicaciones de patentes y solicitudes, y todos los documentos citados en los mismos o durante su prosecución ("documentos citados en la solicitud") y todos los documentos citados o remitidos en los documentos citados en la solicitud, junto con cualquier instrucción, descripción, especificación de producto y ficha de producto para cualquier producto mencionado en los mismos o en cualquier documento de los mismos y citado en este documento, pueden emplearse en la práctica de la invención. Todos los documentos (por ejemplo, estas patentes, publicaciones de patente y solicitudes y los documentos citados en la solicitud) mencionados en este documento en

cualquier parte o lugar de este documento se citan en este documento de patente por referencia en la misma medida que si cada documento individual se indicara citado específica e individualmente.

También con respecto a información general sobre sistemas de CRISPR-Cas, se hace mención de los siguiente:

- 5
- Multiplex genome engineering using CRISPR/Cas systems. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., y Zhang, F. *Science* Feb 15; 339(6121): 819-23 (2013);
 - RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Jiang W., Bikard D., Cox D., Zhang F, Marraffini LA. *Nat Biotechnol* Mar; 31(3): 233-9 (2013);
 - One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. Wang H., Yang H., Shivaila CS., Dawlaty MM., Cheng AW., Zhang F., Jaenisch R. *Cell* May 9; 153(4): 910-8 (2013);
 - 15 ➤ Optical control of mammalian endogenous transcription and epigenetic states. Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F. *Nature*. 22 de agosto de 2013; 500(7463): 472-6, doi: 10.1038/Nature12466. Epub 23 de agosto de 2013;
 - Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Ran, FA., Hsu, PD., Lin, CY., Gootenberg, JS., Konermann, S., Trevino, AE., Scott, DA., Inoue, A., Matoba, S., Zhang, Y., y Zhang, F. *Cell* Ago 28. pii: S0092-8674(13)01015-5. (2013);
 - 20 ➤ DNA targeting specificity of RNA-guided Cas9 nucleases. Hsu, P., Scott, D., Weinstein, J., Ran, FA., Konermann, S., Agarwala, V., Li, Y., Fine, E., Wu, X., Shalem, O., Cradick, TJ., Marraffini, LA., Bao, G., y Zhang, F. *Nat Biotechnol* doi:10.1038/nbt.2647 (2013);
 - 25 ➤ Genome engineering using the CRISPR-Cas9 system. Ran, FA., Hsu, PD., Wright, J., Agarwala, V., Scott, DA., Zhang, F. *Nature Protocols* Nov; 8(11): 2281-308. (2013);
 - 30 ➤ Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Shalem, O., Sanjana, NE., Hartenian, E., Shi, X., Scott, DA., Mikkelsen, T., Heckl, D., Ebert, BL., Root, DE., Doench, JG., Zhang, F. *Science* Dic 12. (2013). [publicación electrónica antes de impresión];
 - Crystal structure of cas9 in complex with guide RNA and target DNA. Nishimasu, H., Ran, FA., Hsu, PD., Konermann, S., Shehata, SI., Dohmae, N., Ishitani, R., Zhang, F., Nureki, O. *Cell* Feb 27. (2014). 156(5):935-49;
 - 35 ➤ Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Wu X., Scott DA., Kriz AJ., Chiu AC., Hsu PD., Dadon DB., Cheng AW., Trevino AE., Konermann S., Chen S., Jaenisch R., Zhang F., Sharp PA. *Nat Biotechnol*. (2014) Abr 20. doi: 10.1038/nbt.2889,
 - 40 ➤ CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling, Platt *et al.*, *Cell* 159(2): 440-455 (2014) DOI: 10.1016/j.cell.2014.09.014,
 - Development and Applications of CRISPR-Cas9 for Genome Engineering, Hsu *et al.*, *Cell* 157, 1262-1278 (5 de junio de 2014) (Hsu 2014),
 - 45 ➤ Genetic screens in human cells using the CRISPR/Cas9 system, Wang *et al.*, *Science*. 3 de enero de 2014; 343(6166): 80-84. doi:10.1126/science.1246981,
 - 50 ➤ Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation, Doench *et al.*, *Nature Biotechnology* publicado en línea el 3 de septiembre de 2014; doi:10.1038/nbt.3026, y
 - In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9, Swiech *et al.*, *Nature Biotechnology* publicado en línea el 19 de octubre de 2014; doi:10.1038/nbt.3055,
 - 55 ➤ Konermann *et al.*, "Genome-scale transcription activation by an engineered CRISPR-Cas9 complex," doi:10.1038/nature14136,
 - Zetsche *et al.*, "A split-Cas9 architecture for inducible genome editing and transcription modulation," *Nature Biotechnology* 33:139-142, DOI:10.1038/nbt.3149 (publicado en línea el 2 de febrero de 2015),
 - 60

➤ Sidi Chen *et al.*, "Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis," Cell 160, 1246-1260, 12 de marzo de 2015 (cribado combinado en ratón), y

5 ➤ Ran *et al.*, "In vivo genome editing using Staphylococcus aureus Cas9," Nature 520, 186-191 (9 de abril de 2015) doi:10.1038/nature14299 (publicado en línea el 1 de abril de 2015), pueden considerarse en la práctica de la presente invención, y se analizan en resumen a continuación:

10 ➤ Cong *et al.* manipularon sistemas de CRISPR-Cas de tipo II para su uso en células eucarióticas basándose tanto en Cas9 *Streptococcus thermophilus* como también en Cas9 de *Streptococcus pyogenes* y demostraron que las nucleasas Cas9 puede dirigirse por ARN cortos para inducir escisión precisa de ADN en células humanas y de ratón. Su estudio mostró además que Cas9 convertida en una enzima melladora puede usarse para facilitar la reparación dirigida por homología en células eucarióticas con actividad mutágena mínima. Además, su estudio demostró que pueden codificarse múltiples secuencias guía en una sola matriz de CRISPR para posibilitar la edición simultánea de varios sitios de locus genómicos endógenos dentro del genoma de mamífero, lo que demuestra una fácil capacidad de programación y una amplia aplicabilidad de la tecnología de nucleasa guiada por ARN. Esta capacidad de usar ARN para programar escisión de ADN específica de secuencia en células definió una nueva clase de herramientas de manipulación genómica. Estos estudios mostraron además que otros locus de CRISPR probablemente son trasplantables en células de mamífero y también pueden mediar la escisión del genoma de mamífero. De forma importante, puede preverse que varios aspectos del sistema de CRISPR-Cas puedan mejorarse más para aumentar su eficacia y versatilidad.

25 ➤ Jiang *et al.* usaron la endonucleasa Cas9 asociada a repeticiones palindrómicas cotas intercaladas de forma regular agrupadas (CRISPR) en complejo con ARN dobles para introducir mutaciones precisas en los genomas de *Streptococcus pneumoniae* y *Escherichia coli*. La estrategia dependía de escisión dirigida por ARN doble:Cas9 en el sitio genómico diana para destruir células no mutadas y evita la necesidad de marcadores de selección o sistemas de contraselección. El estudio informó de la reprogramación de la especificidad de ARN doble:Cas9 cambiando la secuencia del ARN de CRISPR corto (ARNcr) para hacer cambios de uno solo nucleótido y múltiples nucleótidos portados en moldes de edición. El estudio mostró que el uso simultáneo de dos ARNcr posibilitaba mutagénesis combinada. Además, cuando la estrategia se usó en combinación con recombinación, en *S. pneumoniae*, casi un 100 % de las células que se recuperaban usando la estrategia descrita contenían la mutación deseada y, en *E. coli*, un 65 % de las que se recuperaron contenían la mutación.

35 ➤ Konermann *et al.* abordaron la necesidad en la técnica de tecnologías versátiles y robustas que posibiliten la modulación óptica y química de enzima Cas9 de CRISPR basada en dominios de unión a ADN y también efectores de tipo actividad transcripcional.

40 ➤ La nucleasa Cas9 del sistema microbiano de CRISPR-Cas se dirige a locus genómicos específicos por una secuencia guía, que puede tolerar determinados emparejamientos incorrectos con la diana de ADN y promueve de ese modo mutagénesis inespecífica indeseada. Para abordar esto, Ran *et al.* describieron una estrategia que combinaba una nickasa Cas9 mutante con ARN guía emparejadas para introducir roturas bicatenarias dirigidas. Como las mellas individuales en el genoma se paran con alta fidelidad, el mellado simultáneo mediante compensación apropiada de los ARN guía es necesario para roturas bicatenarias y amplía el número de bases reconocidas específicamente para la escisión de la diana. Los autores demostraron que usar mellado emparejado puede reducir la actividad inespecífica en 50 a 1500 veces en líneas celulares y facilitar la inactivación génica en cigotos de ratón sin sacrificar la eficacia de escisión específica. Esta estrategia versátil posibilita una amplia diversidad de aplicaciones de edición genómica que requieren alta especificidad.

50 ➤ Hsu *et al.* caracterizaron la especificidad de dirección de SpCas9 en células humanas para informar de la selección de sitios diana y evitar los efectos inespecíficos. El estudio evaluó > 700 variantes de ARN guía y los niveles de mutación de indel inducidos por SpCas9 en > 100 locus inespecíficos genómicos predichos en células 293T y 293FT. Los autores mostraron que SpCas9 tolera emparejamientos incorrectos entre el ARN guía y el ADN diana en diferentes posiciones de una manera dependiente de secuencia, sensible al número, posición y distribución de emparejamientos incorrectos. Los autores mostraron además que la escisión mediada por SpCas9 no se ve afectada por la metilación del ADN y que la dosificación de SpCas9 y ARNgs puede valorarse para minimizar la modificación inespecífica. Además, para facilitar las aplicaciones de manipulación genómica en mamíferos, los autores informaron de la provisión de una herramienta de programa informático basado en Internet para guiar la selección y validación de secuencias diana, así como análisis de sitios inespecíficos.

60 ➤ Ran *et al.* describieron un conjunto de herramientas para la edición genómica mediada por Cas9 mediante unión de extremos no homólogos (NHEJ) o reparación dirigida por homología (HDR) en células de mamífero, así como la generación de líneas celulares modificadas para estudios funcionales posteriores. Para minimizar la escisión inespecífica, los autores describieron además una estrategia de doble mellado usando la nickasa Cas9 mutante con ARN guía emparejados. El protocolo proporcionado por los autores obtuvo experimentalmente directrices para la selección de sitios diana, la evaluación de la eficacia de escisión y el análisis de la actividad inespecífica. Los estudios

mostraron que empezando con el diseño de dianas, pueden conseguirse modificaciones génicas en tan poco como en 1-2 semanas, y pueden obtenerse líneas celulares clonales en 2-3 semanas.

5 ➤ Shalem *et al.* describieron una nueva manera para consultar la función génica a una escala genómica. Sus estudios mostraron que el suministro de una colección de inactivación de CRISPR-Cas9 a escala genómica (GeCKO) dirigida a 18 080 genes con 64 751 secuencias guía únicas posibilitaba el cribado de selección negativa y positiva en células humanas. En primer lugar, los autores mostraron el uso de la colección GeCKO para identificar genes esenciales para la viabilidad celular en células madre cancerosas y pluripotentes. Después, en un modelo de melanoma, los autores cribaron genes cuya pérdida está implicada en la resistencia a vemurafenib, un agente terapéutico que inhibe la proteína cinasa BRAF mutante. Sus estudios mostraron que los candidatos de mejor clasificación incluían los genes validados previamente NF1 y MED12, así como los aciertos novedosos NF2, CUL3, TADA2B y TADA1. Los autores observaron un alto nivel de consistencia entre ARN guía independientes dirigidos al mismo gen y una alta tasa de confirmación de aciertos y, por tanto, se mostraron prometedores para cribado a escala genómica con Cas9.

15 ➤ Nishimasu *et al.* informaron de la estructura cristalina de Cas9 de *Streptococcus pyogenes* en complejo con ARNGs y su ADN diana a resolución de 2,5 Å. La estructura reveló una arquitectura bilobulada compuesta de lóbulos de reconocimiento de diana y nucleasa, que acomodan la estructura bicatenaria heterogénea de ARNGs:ADN en un surco cargado positivamente en su superficie de contacto. Aunque el lóbulo de reconocimiento es esencial para la unión de ARNGs y ADN, el lóbulo de nucleasa contiene los dominios de nucleasa HNH y RuvC, que se colocan apropiadamente para la escisión de la hebra complementaria y no complementaria del ADN diana, respectivamente. El lóbulo de nucleasa también contiene un dominio carboxiterminal responsable de la interacción con el motivo adyacente de protoespaciador (PAM). Esta estructura de alta resolución y los análisis funcionales adjuntos han revelado el mecanismo molecular de la dirección a ADN guiada por ARN por parte de Cas9, preparando el camino, por tanto, para el diseño racional de nuevas tecnologías de edición genómica versátiles.

25 ➤ Wu *et al.* cartografiaron sitios de unión en todo el genoma de una Cas9 catalíticamente inactiva (dCas9) de *Streptococcus pyogenes* cargada con ARN guía simples (ARNGs) en células madre embrionarias de ratón (mESC). Los autores mostraron que cada uno de los cuatro ARNGs ensayados dirigen dCas9 a entre decenas y moles de sitios genómicos, frecuentemente caracterizados por una región inicial de 5 nucleótidos en el ARNGs y un motivo adyacente de protoespaciador (PAM) NGG. La inaccesibilidad de la cromatina disminuye la unión de dCas9 a otros sitios con secuencias iniciales coincidentes; por tanto, un 70 % de sitios inespecíficos se asocian con genes. Los autores mostraron que la secuenciación dirigida de 295 sitios de unión de dCas9 en mESC transfectadas con Cas9 catalíticamente activa identificó solamente un sitio mutado por encima de niveles de fondo. Los autores propusieron un modelo de dos estados para la unión de Cas9 y la escisión, en que un acoplamiento inicial activa la unión, pero se requiere emparejamiento generalizado con el ADN diana para la escisión.

40 ➤ Hsu 2014 es un artículo de revisión que analiza en general el historial de CRISPR-Cas9 del yogurt para la edición genómica, incluyendo cribado genético de células. Los contenidos generales de Hsu 2014 no implican los aspectos específicos, por ejemplo, modelos, animales, de la presente invención.

45 ➤ Konermann *et al.*, "Genome-scale transcription activation by an engineered CRISPR-Cas9 complex," doi:10.1038/nature14136 (capacidad de fijar múltiples dominios efectores, por ejemplo, activador transcripcional, reguladores funcionales y epigenómicos en posiciones apropiadas en la guía tales como tallo o tetrabucleo con y sin conectores).

➤ Zetsche *et al.*, "A split-Cas9 architecture for inducible genome editing and transcription modulation," Nature Biotechnology 33:139-142, DOI:10.1038/nbt.3149 (publicado en línea el 2 de febrero de 2015) (capacidad de controlar el ensamblaje de Cas9 para la activación).

50 ➤ Sidi Chen *et al.*, "Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis," Cell 160, 1246-1260, 12 de marzo de 2015 (cribado combinado en ratón).

55 ➤ Tsai *et al.*, "Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing," Nature Biotechnology 32(6): 569-77 (2014) (nucleasas FokI).

60 ➤ Ran *et al.*, "In vivo genome editing usando *Staphylococcus aureus* Cas9," Nature 520, 186-191 (09 de abril de 2015) doi:10.1038/nature14299 (publicado en línea el 1 de abril de 2015) (que se refiere a SaCas9 y que no se puede extrapolar de ensayos bioquímicos). Con respecto a SaCas9, la longitud óptima de la guía para Sa puede ser de 21-24 nucleótidos de longitud; y un PAM puede ser NNGRRT, aunque también puede considerarse NNGRR.

Aspectos de la invención, referidos a sistemas de SaCas9 y otros sistemas de CRISPR-Cas ortólogos, pueden ponerse en práctica usando comparaciones estructurales y funcionales con un sistema de SpCas9 como se usa en los métodos y sistemas descritos adicionalmente en la solicitud de patente internacional PCT/US14/70068 titulada "CRISPR-CAS SYSTEMS AND METHODS FOR ALTERING EXPRESSION OF GENE PRODUCTS, STRUCTURAL INFORMATION

AND INDUCIBLE MODULAR CAS ENZYMES" presentada el 12 de diciembre de 2014, que reivindica prioridad a las solicitudes de patente provisional de Estados Unidos n.º de serie 61/915.267, presentada el 12 de diciembre de 2013 y US 61/939.228 presentada el 12 de febrero de 2014.

5 Homology modeling: Corresponding residues in other Cas9 orthologs can be identified by the methods of Zhang *et al.*, 2012 (Nature; 490(7421): 556-60) y Chen *et al.*, 2015 (PLoS Comput Biol; 11(5): e1004248)-a computational protein-protein interaction (PPI) method to predict interactions mediated by domain-motif interfaces. PrePPI (PPI de predicción), un método de predicción de PPI basado en estructura, combina evidencias estructurales con evidencias no
10 estructurales usando un marco estadístico bayesiano. El método implica tomar un par de proteínas de consulta y usar alineación estructural para identificar representantes estructurales que correspondan a sus estructuras determinadas experimentalmente o modelos de homología. La alineación estructural se usa además para identificar vecinos estructurales tanto cercanos como remotos considerando las relaciones geométricas globales y locales. Siempre que dos vecinos de los representantes estructurales formen un complejo presentado en el Protein Data Bank, esto define un molde para modelar la interacción entre las dos proteínas de consulta. Se crean modelos del complejo
15 superponiendo las estructuras representativas sobre su vecino estructural correspondiente en el molde. Esta estrategia se describe adicionalmente en Dey *et al.*, 2013 (Prot Sci; 22: 359-66).

Cualquiera de las realizaciones de la bibliografía anterior, patentes, publicaciones de patente y solicitudes de patentes que se refieren a CRISPR-Cas puede usarse en la práctica de la presente invención, por ejemplo, cualquier realización
20 de la bibliografía anterior, patentes, publicaciones de patente y solicitudes de patente que se refieren a CRISPR-Cas puede usarse con cualquiera del uno o más salientes de la invención de este documento. La divulgación se refiere a métodos mejorados para el diseño y ensayo de reactivos de nickasa para edición genómica de mamífero de alta precisión usando, en particular, reparación dirigida por homología (HDR), incluyendo selección de dianas, construcción de ARNGs, transfección, detección de mutaciones de indel inducidas por Cas9 usando el ensayo de nucleasa
25 SURVEYOR, y diseño y cuantificación de inserciones dirigidas por homología.

La endonucleasa Cas9 específica de secuencia guiada por ARN se ha adoptado ampliamente como herramienta de manipulación genómica debido a su eficacia y facilidad de uso. Derivada del sistema inmunitario adaptativo CRISPR (repeticiones palindrómicas cortas intercaladas de forma regular agrupadas) microbiano de tipo II, Cas9 ahora se ha
30 manipulado satisfactoriamente para aplicaciones de edición genómica en una diversidad de especies animales y vegetales. Para reducir la mutagénesis inespecífica potencial por Cas9 de tipo silvestre, la mutagénesis guiada por homología y estructura de los dominios catalíticos de Cas9 de *Streptococcus pyogenes* ha producido enzimas "melladoras" (Cas9n) que pueden inducir mellas monocatenarias en lugar de roturas bicatenarias (DSB). Como las mellas en general se reparan con alta fidelidad en células eucarióticas, Cas9n puede aprovecharse para mediar edición genómica altamente específica, mediante unión de extremos no homólogos o reparación dirigida por homología.
35

Específicamente, un efecto técnico de la presente invención puede ser que los productos de saliente 3' por doble mellado mediado por N863A aumentan la eficacia de HDR.

40 Cas9 media la edición genómica a través de introducción dirigida de una rotura bicatenaria (DSB) del ADN. La DSB se reconoce por maquinarias de reparación endógenas y puede dar lugar a edición genómica. Están disponibles dos rutas principales de sistemas de reparación endógenos: NHEJ y HDR. Estas dos rutas están en competición entre sí. Para que proceda HDR, los extremos de ADN rotos tienen que extirparse para generar un saliente 3', que a su vez inhibirá NHEJ.
45

Usando la nickasa D10A Cas9 pueden generarse salientes 5', que pueden procesarse tanto por NHEJ como por HR. Usando la nickasa N863A Cas9 puede generarse saliente 3', que inhibe parcialmente NHEJ y, por lo tanto, desvía el resultado de la edición hacia HR.

50 La HDR en células de mamífero procede mediante la generación de salientes 3' seguida de invasión de hebra de un locus homólogo por el extremo 3'.

Selección de dianas

55 La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder la selección de dianas. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica. Aunque se hace referencia a SpCas9, se apreciará que se aplica igualmente a SaCas9 u otros ortólogos según lo apropiado, apreciando especialmente que los PAM tendrán a variar de un ortólogo a otro.

60 Las dianas de SpCas9 pueden ser cualquier secuencia de ADN de 20 pb seguida en el extremo 3' por 5'-NGG-3'. Está disponible una herramienta en línea que aceptará una región de interés como entrada y producirá una lista de todos los sitios diana potenciales de ARNGs dentro de esa región. Cada sitio diana de ARNGs entonces se asocia con una lista de dianas inespecíficas genómicas predichas (<http://tools.genome-engineering.org>). Esta herramienta en línea puede usarse para ayudar en la selección de dianas, sin embargo, no se considera esencial para la selección de dianas. La selección de dianas también puede basarse en métodos conocidos en la técnica.
65

La herramienta también genera pares de ARNGs de doble mellado automáticamente. La consideración más importante para el diseño de ARNGs de doble mellado es el espaciado entre las dos dianas (Ran *et al.*, 2013). Si la "compensación" entre dos guías se define como la distancia entre los extremos PAM distales (5') de un par de ARNGs, una compensación de -4 a 0 pb es ideal, aunque compensaciones tan grandes como de 100 pb pueden inducir indel mediadas por DSB. Los pares de ARNGs para doble mellado pueden estar dirigidas a hebras de ADN opuestas diana.

Construcción de ARNGs plasmídico

La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder la construcción de ARNGs plasmídico. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica.

Pueden construirse vectores de expresión de ARNGs clonando secuencias diana en una cadena principal plasmídica que codifica, por ejemplo, un casete de expresión de ARNGs dirigido por el promotor U6 humano y, por ejemplo, una Cas9-D10A dirigida por CBh (pSpCas9n(BB), (Addgene n.º 48873) (por ejemplo, secuencias diana de 20 pb). La nickasa N863A puede intercambiarse con, por ejemplo, D10A en todos los casos. Se prefiere preparar este plásmido como una maxipreparación sin endotoxinas. Los oligos generalizados para su uso en la clonación de una nueva diana en, por ejemplo, pSpCas9n(BB) se describen en la tabla 1 y puede producirse fácilmente usando protocolos rutinarios o pueden adquirirse de cualquiera de varios proveedores, por ejemplo, de Integrated DNA Technologies (IDT). Obsérvese que la secuencia de PAM requerida para el reconocimiento de dianas por Cas9 nunca está presente como parte del propio ARNGs.

En general, pueden considerarse los siguientes puntos:

Clonar una secuencia diana en un vector de cadena principal de ARNGs.

Hibridar los oligos.

Diluir los oligos hibridados.

Configurar la digestión/ligamiento con, por ejemplo, pSpCas9n(BB), y los oligos hibridados como inserto de clonación.

Puede realizarse un control negativo usando las mismas condiciones.

Incubar el ligamiento.

Transformar la reacción de ligamiento en una cepa competente.

Selección de colonias positivas a partir de la transformación, inoculado y cultivo.

Aislar el ADN plasmídico y determinar la concentración de ADN por espectrofotometría. Estas construcciones pueden verificarse por secuenciación de Sanger para confirmar la inserción correcta de la secuencia diana. Para condiciones de transfección óptimas posteriores, puede prepararse plásmido sin endotoxinas.

Validación de ARNGs en líneas celulares

La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder la validación de los ARNGs en líneas celulares. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica.

En general, pueden considerarse los siguientes puntos:

Mantener células sanas.

Transfectar, cultivar/mantener y controlar los cultivos.

Recoger las células para la extracción de ADN genómico y/o análisis posterior.

Cuando se trabaja con diferentes tipos celulares, pueden compararse reactivos de transfección alternativos para la eficacia y la toxicidad. También puede ser informativo valorar pSpCas9n(ARNGs) para encontrar la concentración de transfección óptima con la máxima eficacia.

Recolección celular y extracción de ADN

La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder la recolección celular y la extracción de ADN. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica.

En general, pueden considerarse los siguientes puntos:

- 5 Recoger células.
- Dividir en alícuotas, centrifugar y resuspender el sedimento para el lavado.
- Centrifugar y resuspender.
- 10 Extraer ADN genómico usando, por ejemplo, un protocolo de termociclador.
- Centrifugar el producto de reacción para sedimentar los desechos celulares y transferir el sobrenadante aclarado en un tubo nuevo para análisis posterior.
- 15 Determinar la concentración de ADN de la extracción por, por ejemplo, espectrofotometría y normalizar con ddH₂O.

Análisis de indel SURVEYOR

La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder el análisis de indel SURVEYOR. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica.

El ensayo SURVEYOR (Transgenomic 706025) es un método para detectar polimorfismos e indel pequeños. Las muestras de ADN se amplifican por PCR y los productos se calientan para desnaturalizar y se enfrían lentamente para formar estructuras bicatenarias heterogéneas. Las estructuras bicatenarias desapareadas entonces se escinden por la nucleasa SURVEYOR y los productos de escisión se analizan por electroforesis en gel.

En general, pueden considerarse los siguientes puntos:

- 30 Realizar PCR sobre ADN genómico.
- Obsérvese que, como SURVEYOR se diseñó para detectar mutaciones, es crucial usar una polimerasa de alta fidelidad para evitar positivos falsos.
- 35 Procesar el producto de PCR en un gel para asegurar que se ha formado un solo producto de tamaño esperado.
- Purificar el producto de PCR, medir la concentración de ADN y normalizar usando ddH₂O.
- Mezclar el producto de PCR normalizado con tampón de PCR Taq. Fundir y rehibridar los productos gradualmente en un termociclador.
- 40 Mezclar la nucleasa S SURVEYOR y el potenciador S SURVEYOR con todo el producto hibridado de lo anterior. Realizar la digestión. Las muestras que tienen mutaciones dentro de los amplicones de PCR rehibridados se escindirán por SURVEYOR.
- 45 Los productos de digestión pueden mezclarse con un tinte de carga apropiado y visualizarse por electroforesis en un gel de poliacrilamida TBE al 4-20 % (véase el ejemplo, figura 2B).

Pueden estimarse las tasas de modificación genómica calculando en primer lugar las intensidades relativas de los productos de digestión a y b, y la banda c no digerida. La frecuencia de corte f_{cut} entonces se da por $(a+b)/(a+b+c)$. La siguiente fórmula, basada en la distribución de probabilidades binomial de la formación de estructuras bicatenarias estima el porcentaje de indel en la muestra.

$$\%indel = (1 - \sqrt{(1 - f_{cut})}) 100$$

55 Inserción por HDR y sin HDR usando Cas9n

La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder la inserción por HDR y sin HDR usando Cas9n. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica.

- 60 En general, pueden considerarse los siguientes puntos:
- Diseñar brazos de homología ODN_{mc}, que pueden diseñarse para que sean lo más largo posible, con al menos 40 nucleótidos de homología en cada lado de la secuencia a introducir. (Véase el ejemplo de diseño, figura 3).

65

Mezclar plásmidos de Cas9(ARNgs) con ODNmc para nucleofección.

Un oligodesoxinucleótido monocatenario (ODNmc) tiene una alta eficacia como molde para recombinación homóloga, aunque también pueden usarse vectores plasmídicos linealizados. En algunos tipos celulares, una sola nickasa puede estimular un evento de reparación homólogo dirigido en presencia de un molde donador. En otros, tales como células madre embrionarias humanas, puede requerirse una rotura bicatenaria por doble mellado para promover HDR eficaz (Ran *et al.*, 2013). Las consideraciones para elegir pares de ARNgs de doble mellado para HDR son similares a las de atenuación génica por NHEJ, con el requisito adicional de que una de las mellas debe producirse en aproximadamente 20 pb del sitio de inserción de HDR. En células 293FT, la HDR mediada por doble mellado puede ser comparativamente eficaz a HDR mediada por Cas9 de tipo silvestre.

Las enzimas Cas9 melladoras son muy adecuadas para generar modificaciones muy precisas. Como la HDR se produce típicamente a baja eficacia en los mejores casos, también proporcionan plásmidos pSpCas9n que codifican el conector policistrónico 2A seguido por GFP y marcadores de puomicina (Addgene n.º 48140 y 48141) para facilitar el enriquecimiento de las células modificadas.

Los brazos de homología ODNmc pueden diseñarse para que sean lo más largo posible, con al menos 40 nucleótidos de homología en cada lado de la secuencia a introducir. El servicio Ultramer proporcionado por IDT permite la síntesis de oligos de hasta 200 pb de longitud. Los moldes de homología pueden diluirse hasta 10 µM y almacenarse a -20 °C (véase el ejemplo de diseño, figura 3).

El suministro por nucleofección es óptimo para ODNmc. El 4D Nucleofector X Kit S (Lonza V4XC-2032) puede usarse para células HEK293FT sembradas en placas de 6 pocillos tratadas con cultivo hístico. El fabricante proporciona un protocolo óptimo para nucleofección de estos y otros tipos celulares. Se mezclan 500 ng de plásmidos pSpCas9n(ARNgs) totales con 1 µl de ODNmc 10 µM para nucleofección.

El efecto técnico de la presente invención es que los productos de salientes 3' generados por doble mellado mediado por N863A aumenta la eficacia de HDR.

Análisis de HDR y eventos de inserción

La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder el análisis de HDR y eventos de inserción. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica.

En general, pueden considerarse los siguientes puntos:

Preparar medio de FACS.

Preparar placas de pocillos para clasificación de clones.

Disociar las células.

Detener el tratamiento con tripsina, transferir las células y establecer una suspensión monocelular antes de continuar.

Centrifugar las células, aspirar el sobrenadante completamente y resuspender el sedimento minuciosamente.

Filtrar las células para retirar por filtración los agregados celulares.

Clasificar las células individuales usando, por ejemplo, la máquina FACS.

Incubar y expandir las células.

Pase de las poblaciones clonales en placas replicadas, disociación y conservación para extracción de ADN

Puede realizarse genotipado por, por ejemplo, amplificación por PCR del locus de interés, purificación por PCR y secuenciación de Sanger de los productos.

Resolución de problemas

La siguiente descripción es un simple análisis ejemplar de la manera en que puede proceder la resolución de problemas. También pueden aplicarse alternativas equivalentes, así como protocolos conocidos en la técnica.

1. *Colonias de una placa de control negativo mientras se clonan dianas en pSpCas9n.*

a. La presencia de colonias negativas en general indica una digestión de restricción incompleta del plásmido de cadena principal.

5 b. Un simple ejemplo para superar este problema incluye prolongar la reacción Golden Gate durante 20-25 ciclos para aumentar la eficacia de la digestión. La cantidad de enzima de restricción usada también puede aumentarse, aunque el volumen de enzima no puede exceder de un 20 % del volumen de reacción total. Se vuelve a transformar el plásmido de cadena principal de Cas9, se aísla una nueva preparación de ADN plasmídico y se verifica por secuencia el sitio de restricción.

10 2. *La eficacia de transfección de reactivos de Cas9 es baja.*

a. La baja eficacia de transfección puede ser la norma para algunas líneas celulares y especialmente células primarias o líneas de células madre.

15 b. Un simple ejemplo para superar este problema incluye el enriquecimiento de poblaciones celulares para células transfectadas usando plásmidos pSpCas9n(BB)-GFP o pSpCas9n(BB)-Puro para FACS sobre la fluorescencia de GFP o realizar selección con antibiótico.

20 3. *El doble mellado no produce indel.*

a. Los ARNGs de doble mellado individuales pueden ensayarse con el contexto de tipo silvestre para garantizar que cada uno de ellos funciona por separado como una guía de Cas9 válida.

25 b. Se comprueba el espaciado del par de ARNGs. El doble mellado se realiza de forma óptima cuando las guías están espaciadas por 20 pb o menos, y las guías pueden estar orientadas de modo que sus secuencias PAM 5' respectivas estén enfrentadas entre sí.

4. *La eficacia de HDR es baja.*

30 a. Pueden introducirse mutaciones sinónimas dentro del sitio diana en el ODNmc para evitar la escisión del sitio genómico recombinado satisfactoriamente.

35 La invención se refiere a la manipulación y optimización de sistemas, métodos y composiciones usadas para el control de la expresión génica, que implican la dirección de secuencias, tal como la perturbación genómica o la edición génica, que se refieren al sistema de CRISPR-Cas y componentes del mismo. En las realizaciones de la invención, la enzima Cas es una Cas9 mutada (N863A).

40 Una ventaja de los presentes métodos es que el sistema de CRISPR evita la unión inespecífica y sus efectos secundarios resultantes. Esto se consigue usando sistemas dispuestos para que tengan un alto grado de especificidad de secuencia por el ADN diana.

45 La optimización de Cas9 puede usarse para potenciar la función o para desarrollar nuevas funciones, se pueden generar proteínas Cas9 quiméricas. Las proteínas Cas9 quiméricas pueden prepararse combinando fragmentos de diferentes homólogos de Cas9. Por ejemplo, dos proteínas Cas9 quiméricas ejemplares de Cas9 descritas en este documento. Por ejemplo, los solicitantes fusionaron el extremo N de St1Cas9 (el fragmento de esta proteína está en negrita) con el extremo C de SpCas9. El beneficio de preparar Cas9 quiméricas incluye todos y cada uno de:

toxicidad reducida;

50 expresión mejorada en células eucarióticas;

especificidad potenciada;

55 peso molecular reducido de la proteína, hacer la proteína más pequeña combinando los dominios más pequeños de diferentes homólogos de Cas9; y/o

alterar la necesidad de la secuencia PAM.

60 Como se menciona anteriormente, también pueden proporcionarse animales transgénicos, así como plantas transgénicas, especialmente cultivos y algas. El transgénico puede ser útil en aplicaciones distintas de proporcionar un modelo en enfermedad. Estas pueden incluir producción de alimentos o piensos a través de la expresión de, por ejemplo, niveles mayores de proteínas, carbohidratos, nutrientes o vitaminas de lo que se observaría normalmente en el tipo silvestre. A este respecto, se prefieren plantas transgénicas, especialmente legumbres y tubérculos y animales, especialmente mamíferos tales como ganado (vacas, ovejas, cabras y cerdos), pero también aves de corral e insectos comestibles.

65

Las algas transgénicas u otras plantas tales como colza pueden ser particularmente útiles en la producción de aceites vegetales o biocombustibles tales como alcoholes (especialmente metanol y etanol), por ejemplo. Estas pueden manipularse para que expresen o sobreexpresen altos niveles de aceites o alcoholes para su uso en las industrias petroleras o de biocombustibles.

- 5 En términos de suministro *in vivo*, el AAV es ventajoso sobre otros vectores víricos por un par de razones:
- Baja toxicidad (esto puede deberse al método de purificación que no requiere ultracentrifugación de partículas celulares que pueden activar la respuesta inmunitaria)
 - Baja probabilidad de causar mutagénesis por inserción porque no se integra en el genoma hospedador.

15 El AAV tiene un límite de empaquetado de 4,5 o 4,75 Kb. Esto significa que Cas9, así como un promotor y terminador de la transcripción tienen que adaptarse todos en el mismo vector vírico. Construcciones mayores de 4,5 o 4,75 Kb darán lugar a producción de virus significativamente reducida. SpCas9 es bastante grande, el propio gen es de más de 4,1 Kb, lo que hace difícil su empaquetado en AAV. Por lo tanto, realizaciones de la invención incluyen la utilización de homólogos de Cas9 que son más cortos. Por ejemplo:

Especie	Tamaño de Cas9
<i>Corynebacter diphtheriae</i>	3252
<i>Eubacterium ventriosum</i>	3321
<i>Streptococcus pasteurianus</i>	3390
<i>Lactobacillus farciminis</i>	3378
<i>Sphaerochaeta globus</i>	3537
<i>Azospirillum B510</i>	3504
<i>Gluconacetobacter diazotrophicus</i>	3150
<i>Neisseria cinerea</i>	3246
<i>Roseburia intestinalis</i>	3420
<i>Parvibaculum lavamentivorans</i>	3111
<i>Staphylococcus aureus</i>	3159
<i>Nitratifactor salsuginis</i> DSM 16511	3396
<i>Campylobacter lari</i> CF89-12	3009
<i>Streptococcus thermophilus</i> LMD-9	3396

20 Por lo tanto, estas especies son en general especies de Cas9 preferidas, especialmente SaCas9 como se menciona. Los solicitantes han mostrado suministro y datos de expresión de Cas9 en cerebro de ratón *in vivo*.

se prefieren dos maneras de empaquetar moléculas de ácido nucleico codificantes de Cas9, por ejemplo, ADN, en vectores víricos para mediar la modificación genómica *in vivo*:

25 Conseguir inactivación génica mediada por NHEJ:

Vector de un solo virus:

30 Vector que contiene dos o más casetes de expresión:

Promotor-molécula de ácido nucleico que codifica Cas9-terminador

Promotor-ARNg1-terminador

35 Promotor-ARNg2-terminador

Promotor-ARNg(N)-terminador (hasta el límite de tamaño del vector)

40 Vector de doble virus:

Vector 1 que contiene un casete de expresión para dirigir la expresión de Cas9

- Promotor-molécula de ácido nucleico que codifica Cas9-terminador
- 5 Vector 2 que contiene uno o más casetes de expresión para dirigir la expresión de uno o más ARN guía
- Promotor-ARNg1-terminador
- Promotor-ARNg(N)-terminador (hasta el límite de tamaño del vector)
- 10 Para mediar la reparación dirigida por homología.
- Además de las estrategias de vector de un solo virus y doble virus descritas anteriormente, se usa un vector adicional para suministrar un molde de reparación dirigida por homología.
- 15 El promotor usado para dirigir la expresión de la molécula de ácido nucleico codificante de Cas9 puede incluir:
- ITR de AAV puede servir como promotor: esto es ventajoso para eliminar la necesidad de un elemento promotor adicional (que puede ocupar espacio en el vector). El espacio adicional liberado puede usarse para dirigir la expresión de elementos adicionales (ARNg, etc.). También, la actividad de ITR es relativamente más débil, de modo que puede usarse para reducir la toxicidad debida a sobreexpresión de Cas9.
- 20 Para expresión ubicua, pueden usarse promotores: CMV, CAG, CBh, PGK, SV40, cadena pesada o ligera de ferritina, etc.
- 25 Para la expresión en cerebro, pueden usarse promotores: Sinapsina I para todas las neuronas, CaMKIIalfa para neuronas excitadoras, GAD67 o GAD65 o VGAT para neuronas GABAérgicas, etc.
- Para expresión en hígado, puede usarse el promotor de albúmina
- 30 Para la expresión en pulmón, puede usarse SP-B
- Para células endoteliales, puede usarse ICAM
- 35 Para células hematopoyéticas puede usarse IFNbeta o CD45
- Para osteoblastos puede usarse OG-2
- El promotor usado para dirigir el ARN guía puede incluir:
- 40 Promotores de pol III tales como U6 o H1
- Uso del promotor de pol II y casetes intrónicos para expresar ARNg.
- 45 En cuanto a AAV, el AAV puede ser AAV1, AAV2, AAV5 o cualquier combinación de los mismos. Se puede seleccionar el AAV del AAV con respecto a las células a abordar; por ejemplo, se pueden seleccionar los serotipos de AAV 1, 2, 5 o un híbrido o cápside de AAV1, AAV2, AAV5 o cualquier combinación de los mismos para abordar células cerebrales o neuronales; y se puede seleccionar AAV4 para abordar tejido cardíaco. AAV8 es útil para suministro al hígado. Los promotores y vectores anteriores se prefieren individualmente.
- 50 El suministro de ARN también es un método útil para el suministro *in vivo*. Es posible suministrar Cas9 y ARNg (y, por ejemplo, molde de reparación por HR) en células usando liposomas o nanopartículas. Por tanto, el suministro de la enzima de CRISPR, tal como una Cas9 y/o suministro de los ARN de la invención puede ser en forma de ARN y mediante microvesículas, liposomas o nanopartículas. Por ejemplo, el ARNm de Cas9 y el ARNg pueden empaquetarse en partículas liposómicas para suministro *in vivo*. Los reactivos de transfección liposómica tales como Invivofectamine de Life Technologies y otros reactivos en el mercado pueden suministrar de forma eficaz moléculas de ARN al hígado.
- 55 Potenciar la eficacia de NHEJ o de HR también puede ser de ayuda para el suministro. Se prefiere que la eficacia de NHEJ se potencie por coexpresión de enzimas de procesamiento de extremos tales como Trex2 (*Dumitrache et al., Genetics.* agosto de 2011; 188(4): 787-797). Se prefiere que la eficacia de HR se aumente por inhibición transitoria de las maquinarias de NHEJ tales como Ku70 y Ku86. La eficacia de HR también puede aumentarse por coexpresión de enzimas de recombinación homóloga procarióticas o eucarióticas tales como RecBCD, RecA.
- 60 En este documento se describen diversos medios de suministro y se analizan adicionalmente en esta sección.
- 65

Suministro

Suministro de vector, por ejemplo, plásmido, suministro vírico: La enzima de CRISPR, por ejemplo, una Cas9 y/o cualquiera de los presentes ARN, por ejemplo, un ARN guía, pueden suministrarse usando cualquier vector adecuado, por ejemplo, vectores plasmídicos o víricos, tales como tipos de vector de virus adenoasociado (AAV), lentivíricos, adenovíricos u otros virus, o combinaciones de los mismos. Cas9 y uno o más ARN guía pueden empaquetarse en uno o más vectores, por ejemplo, vectores plasmídicos o víricos. En algunas realizaciones, el vector, por ejemplo, vector plasmídico o vírico, se suministra al tejido de interés por, por ejemplo, una inyección intramuscular, mientras que otras veces el suministro es mediante métodos intravenosos, transdérmicos, intranasales, orales, a la mucosa u otros métodos de suministro. Dicho suministro puede ser mediante una sola dosis o múltiples dosis. Un experto en la materia comprende que la dosificación real a suministrar en este documento puede variar enormemente dependo de una diversidad de factores, tales como la elección del vector, la célula diana, el organismo o tejido, el estado general del sujeto a tratar, el grado de transformación/modificación buscado, la vía de administración, el modo de administración, el tipo de transformación/modificación buscada, etc.

Dicha dosificación puede contener además, por ejemplo, un vehículo (agua, solución salina, etanol, glicerol, lactosa, sacarosa, fosfato de calcio, gelatina, dextrano, agar, pectina, aceite de cacahuete, aceite de sésamo, etc.), un diluyente, un vehículo farmacéuticamente aceptable (por ejemplo, solución salina tamponada con fosfato), un excipiente farmacéuticamente aceptable y/u otros compuestos conocidos en la técnica. La dosificación puede contener además una o más sales farmacéuticamente aceptables tales como, por ejemplo, una sal de ácido mineral tal como un clorhidrato, un bromhidrato, un fosfato, un sulfato, etc.; y las sales de ácidos orgánicos tales como acetatos, propionatos, malonatos, benzoatos, etc. Adicionalmente, también pueden estar presentes en este documento sustancias auxiliares, tales como agentes humectantes o emulsionantes, sustancias que tamponan el pH, geles o materiales gelificantes, aromatizantes, colorantes, microesferas, polímeros, agentes de suspensión, etc. Además, también pueden estar presentes uno o más ingredientes farmacéuticos convencionales distintos, tales como conservantes, humectantes, agentes de suspensión, tensioactivos, antioxidantes, agentes antiapelmazantes, rellenos, agentes quelantes, agentes de recubrimiento, estabilizantes químicos, etc., especialmente si la forma galénica es una forma reconstituible. Ingredientes ejemplares adecuados incluyen celulosa microcristalina, carboximetilcelulosa de sodio, polisorbato 80, alcohol feniletílico, clorobutanol, sorbato de potasio, ácido sórbico, dióxido de azufre, galato de propilo, los parabenos, etil vainillina, glicerina, fenol, paraclorofenol, gelatina, albúmina y una combinación de los mismos. Un análisis minucioso de los excipientes farmacéuticamente aceptables está disponible en REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991).

En una realización de este documento, el suministro es mediante un adenovirus, que puede estar en una sola dosis de refuerzo que contiene al menos 1×10^5 partículas (también denominadas unidades de partícula, up) de vector adenovírico. En una realización de este documento, la dosis preferiblemente es de al menos aproximadamente 1×10^6 partículas (por ejemplo, aproximadamente 1×10^6 - 1×10^{12} partículas), más preferiblemente al menos aproximadamente 1×10^7 partículas, más preferiblemente al menos aproximadamente 1×10^8 partículas (por ejemplo, aproximadamente 1×10^8 - 1×10^{11} partículas o aproximadamente 1×10^8 - 1×10^{12} partículas), y mucho más preferiblemente al menos aproximadamente 1×10^9 partículas (por ejemplo, aproximadamente 1×10^9 - 1×10^{10} partículas o aproximadamente 1×10^9 - 1×10^{12} partículas), o incluso al menos aproximadamente 1×10^{10} partículas (por ejemplo, aproximadamente 1×10^{10} - 1×10^{12} partículas) de vector adenovírico. Como alternativa, la dosis comprende no más de aproximadamente 1×10^{14} partículas, preferiblemente no más de aproximadamente 1×10^{13} partículas, incluso más preferiblemente no más de aproximadamente 1×10^{12} partículas, incluso más preferiblemente no más de aproximadamente 1×10^{11} partículas, y mucho más preferiblemente no más de aproximadamente 1×10^{10} partículas (por ejemplo, no más de aproximadamente 1×10^9 partículas). Por tanto, la dosis puede contener una sola dosis de vector adenovírico con, por ejemplo, aproximadamente 1×10^6 unidades de partícula (up), aproximadamente 2×10^6 up, aproximadamente 4×10^6 up, aproximadamente 1×10^7 up, aproximadamente 2×10^7 up, aproximadamente 4×10^7 up, aproximadamente 1×10^8 up, aproximadamente 2×10^8 up, aproximadamente 4×10^8 up, aproximadamente 1×10^9 up, aproximadamente 2×10^9 up, aproximadamente 4×10^9 up, aproximadamente 1×10^{10} up, aproximadamente 2×10^{10} up, aproximadamente 4×10^{10} up, aproximadamente 1×10^{11} up, aproximadamente 2×10^{11} up, aproximadamente 4×10^{11} up, aproximadamente 1×10^{12} up, aproximadamente 2×10^{12} up o aproximadamente 4×10^{12} up de vector adenovírico. Véanse, por ejemplo, los vectores adenovíricos en la patente de Estados Unidos n.º 8.454.972 B2 de Nabel, *et al.*, concedida el 4 de junio de 2013; y las dosificaciones en la columna 29, líneas 36-58 de la misma. En una realización de este documento, el adenovirus se suministra mediante múltiples dosis.

En una realización de este documento, el suministro es mediante un AAV. Una dosificación terapéuticamente eficaz para suministro *in vivo* del AAV a un ser humano se cree que está en el intervalo de aproximadamente 20 a aproximadamente 50 ml de solución salina que contiene de aproximadamente 1×10^{10} a aproximadamente 1×10^{10} de AAV funcional/ml de solución. La dosificación puede ajustarse para equilibrar el beneficio terapéutico frente a cualquier efecto secundario. En una realización de este documento, la dosis de AAV en general está en el intervalo de concentraciones de aproximadamente 1×10^5 a 1×10^{50} genomas de AAV, de aproximadamente 1×10^8 a 1×10^{20} genomas de AAV, de aproximadamente 1×10^{10} a aproximadamente 1×10^{16} genomas o de aproximadamente 1×10^{11} a aproximadamente 1×10^{16} genomas de AAV. Una dosificación humana puede ser de aproximadamente 1×10^{13} genomas de AAV. Dichas concentraciones puede suministrarse en de aproximadamente 0,001 ml a aproximadamente 100 ml, de aproximadamente 0,05 a aproximadamente 50 ml o de aproximadamente 10 a aproximadamente 25 ml de

una solución de vehículo. Otras dosificaciones eficaces pueden establecerse fácilmente por un experto en la materia a través de ensayos rutinarios que establecen las curvas de respuesta a la dosis. Véase, por ejemplo, la patente de Estados Unidos n.º 8.404.658 B2 de Hajjar, *et al.*, concedida el 26 de marzo de 2013, en la columna 27, líneas 45-60,

5 En una realización de este documento, el suministro es mediante un plásmido. En dichas composiciones plasmídicas, la dosificación debe ser una cantidad suficiente de plásmido para provocar una respuesta. Por ejemplo, cantidades adecuadas de ADN plasmídico en composiciones plasmídicas pueden ser de aproximadamente 0,1 a aproximadamente 2 mg, o de aproximadamente 1 µg a aproximadamente 10 µg por individuo de 70 kg. Los plásmidos de la invención en general comprenderán (i) un promotor; (ii) una secuencia que codifica una enzima de CRISPR, unida de forma funcional a dicho promotor; (iii) un marcador de selección; (iv) un origen de replicación; y (v) un terminador de la transcripción posterior y unido funcionalmente a (ii). El plásmido también puede codificar los componentes de ARN de un complejo de CRISPR, pero uno o más de estos puede estar codificado en su lugar en un vector diferente.

10 15 Las dosis en este documento se basan en un individuo promedio de 70 kg. La frecuencia de administración pertenece al ámbito del facultativo médico o veterinario (por ejemplo, médico, veterinario) o científico experto en la materia. También se aprecia que los ratones usados en los experimentos típicamente son de aproximadamente 20 g y de los experimentos en ratones se puede aumentar la escala hasta un individuo de 70 kg.

20 En algunas realizaciones, las moléculas de ARN de la invención se suministran en formulaciones de liposomas o lipofectin y similares y pueden prepararse por métodos bien conocidos por los expertos en la materia. Dichos métodos se describen, por ejemplo, en las patentes de Estados Unidos n.º 5.593.972, 5.589.466 y 5.580.859. Se han desarrollado sistemas de suministro dirigidos específicamente al suministro potenciado y mejorado de ARNip en células de mamífero (véase, por ejemplo, Shen *et al.*, FEBS Let. 2003, 539:111-114; Xia *et al.*, Nat. Biotech. 2002, 20:1006-1010; Reich *et al.*, Mol. Vision. 2003, 9: 210-216; Sorensen *et al.*, J. Mol. Biol. 2003, 327: 761-766; Lewis *et al.*, Nat. Gen. 2002, 32: 107-108 y Simeoni *et al.*, NAR 2003, 31, 11: 2717-2724) y pueden aplicarse a la presente invención. El ARNip se ha usado recientemente de manera satisfactoria para la inhibición de la expresión génica en primates (véase, por ejemplo, Tolentino *et al.*, Retina 24(4):660 que también puede aplicarse a la presente invención.

25 30 De hecho, el suministro de ARN es un método útil de suministro *in vivo*. Es posible suministrar Cas9 y ARNg (y, por tanto, molde de reparación por HR) a células usando liposomas o partículas o nanopartículas. Por tanto, el suministro de la enzima de CRISPR, tal como una Cas9 y/o suministro de los ARN de la invención puede ser en forma de ARN y mediante microvesículas, liposomas o partículas o nanopartículas. Por ejemplo, el ARNm de Cas9 y el ARNg pueden empaquetarse en partículas liposómicas para suministro *in vivo*. Los reactivos de transfección liposómica tales como lipofectamine de Life Technologies y otros reactivos en el mercado pueden suministrar de forma eficaz moléculas de ARN al hígado.

35 40 Un medio de suministro de ARN también preferido incluye suministro de ARN mediante partículas o nanopartículas (Cho, S., Goldberg, M., Son, S., Xu, Q., Yang, F., Mei, Y., Bogatyrev, S. Langer, R. y Anderson, D., Lipid-like nanoparticles for small interfering RNA delivery to endothelial cells, Advanced Functional Materials, 19: 3112-3118, 2010) o exosomas (Schroeder, A., Levins, C., Cortez, C. Langer, R., y Anderson, D., Lipid-based nanotherapeutics for siRNA delivery, Journal of Internal Medicine, 267: 9-21, 2010, PMID: 20059641). De hecho, los exosomas han demostrado ser particularmente útiles en el suministro de ARNip, un sistema con algunos paralelismos al sistema de CRISPR. Por ejemplo, El-Andaloussi S, *et al.*, ("Exosome-mediated delivery of siRNA in vitro and in vivo." Nat Protoc. diciembre de 2012;7(12):2112-26, doi: 10.1038/nprot.2012.131. Epub 15 de noviembre de 2012) describen la manera en que los exosomas son herramientas prometedoras para el suministro de fármacos a través de diferentes barreras biológicas y pueden aprovecharse para el suministro de ARNip *in vitro* e *in vivo*. Su estrategia es generar exosomas dirigidos a través de transfección de un vector de expresión, que comprende una proteína exosómica fusionada con un ligando peptídico. Los exosomas entonces se purifican y caracterizan a partir del sobrenadante celular transfectado, después el ARN se carga en los exosomas. El suministro o administración de acuerdo con la invención puede realizarse con exosomas, en particular, aunque sin limitación, al cerebro. La vitamina E (α-tocoferol) puede conjugarse con Cas de CRISPR y suministrarse al cerebro junto con lipoproteína de alta densidad (HDL), por ejemplo, de una manera similar a lo que hicieron Uno *et al.*, (HUMAN GENE THERAPY 22:711-719 (junio de 2011)) para suministrar ARN interferente pequeño (ARNip) al cerebro. Los ratones se infundieron mediante minibombas osmóticas (modelo 1007D; Alzet, Cupertino, CA) llenadas de solución salina tamponada con fosfato (PBS) o TocsiBACE libre o TocsiBACE/HDL y conectadas con el kit 3 de infusión cerebral (Alzet). Se colocó una cánula de infusión cerebral aproximadamente 0,5 mm posterior a la bregma en la línea media para infusión en el tercer ventrículo dorsal. Uno *et al.* descubrieron que tan poco como 3 nmol de Toc-ARNip con HDL podía inducir una reducción diana en grado comparable por el mismo método de infusión ICV. Una dosificación similar de Cas de CRISPR conjugada con α-tocoferol y coadministrada con HDL dirigida al cerebro puede contemplarse para seres humanos en la presente invención, por ejemplo, de aproximadamente 3 nmol a aproximadamente 3 µmol de Cas de CRISPR Cas dirigida al cerebro puede contemplarse. Zou *et al.* ((HUMAN GENE THERAPY 22:465-475 (abril de 2011)) describen un método de suministro mediado por lentivirus de ARN de horquilla corta dirigidos a PKCy para silenciamiento génico *in vivo* en la médula espinal de ratas. Zou *et al.* administraron aproximadamente 10 µl de un lentivirus recombinante que tenía un valor de 1×10^9 unidades de transducción (UT)/ml mediante un catéter intratecal. Una dosificación similar de Cas de CRISPR expresada en un vector lentivírico dirigido al cerebro puede contemplarse para seres humanos en la

presente invención, por ejemplo, aproximadamente 10-50 ml de Cas de CRISPR dirigida al cerebro en un lentivirus que tiene un valor de 1×10^9 unidades de transducción (UT)/ml puede contemplarse.

5 En términos de suministro local al cerebro, esto puede conseguirse de diversas maneras. Por ejemplo, el material puede suministrarse al interior del estriado, por ejemplo, por inyección. La inyección puede realizarse de manera estereotáctica mediante una craneotomía.

10 Potenciar la eficacia de NHEJ o de HR también puede ser de ayuda para el suministro. Se prefiere que la eficacia de NHEJ se potencie por coexpresión de enzimas de procesamiento de extremos tales como Trex2 (*Dumitrache et al., Genetics.* agosto de 2011; 188(4): 787-797). Se prefiere que la eficacia de HR se aumente por inhibición transitoria de las maquinarias de NHEJ tales como Ku70 y Ku86. La eficacia de HR también puede aumentarse por coexpresión de enzimas de recombinación homóloga procarióticas o eucarióticas tales como RecBCD, RecA.

15 **Empaquetado y promotores**

Las maneras para empaquetar moléculas de ácido nucleico codificantes de Cas9, por ejemplo, ADN, en vectores, por ejemplo, vectores víricos, para mediar la modificación genómica *in vivo* incluyen:

20 Conseguir inactivación génica mediada por NHEJ:

Vector de un solo virus:

Vector que contiene dos o más casetes de expresión:

25 Promotor-molécula de ácido nucleico que codifica Cas9-terminador

Promotor-ARNg1-terminador

30 Promotor-ARNg2-terminador

Promotor-ARNg(N)-terminador (hasta el límite de tamaño del vector)

Vector de doble virus:

35 Vector 1 que contiene un casete de expresión para dirigir la expresión de Cas9

Promotor-molécula de ácido nucleico que codifica Cas9-terminador

40 Vector 2 que contiene uno o más casetes de expresión para dirigir la expresión de uno o más ARN guía

Promotor-ARNg1-terminador

Promotor-ARNg(N)-terminador (hasta el límite de tamaño del vector)

45 Para mediar la reparación dirigida por homología.

Además de las estrategias de vector de un solo virus y de doble virus descritas anteriormente, puede usarse un vector adicional para suministrar un molde de reparación dirigida por homología.

50 El promotor usado para dirigir la expresión de la molécula de ácido nucleico codificante de Cas9 puede incluir:

ITR de AAV puede servir como promotor: esto es ventajoso para eliminar la necesidad de un elemento promotor adicional (que puede ocupar espacio en el vector). El espacio adicional liberado puede usarse para dirigir la expresión de elementos adicionales (ARNg, etc.). Además, la actividad de ITR es relativamente más débil, de modo que puede usarse para reducir la toxicidad potencial debida a sobreexpresión de Cas9,

55 Para la expresión ubicua, puede usarse cualquiera de los siguientes promotores: CMV, CAG, CBh, PGK, SV40, cadena pesada o ligera de ferritina y así sucesivamente.

60 Para la expresión en el cerebro u otras partes del SNC, pueden usarse promotores: Sinapsina I para todas las neuronas, CaMKIIalfa para neuronas excitadoras, GAD67 o GAD65 o VGAT para neuronas GABAérgicas, etc. Para expresión en hígado, se puede usar el promotor de albúmina. Para expresión en pulmón, puede usarse SP-B. Para células endoteliales, puede usarse ICAM. Para células hematopoyéticas puede usarse IFNbeta o CD45. Para osteoblastos puede usarse OG-2.

65 El promotor usado para dirigir el ARN guía puede incluir:

Promotores de pol III tales como U6 o H1

Uso del promotor de pol II y casetes intrónicos para expresar ARNg

5

Virus adenoasociado (AAV)

Cas9 y uno o más ARN guía pueden suministrarse usando tipos de vectores víricos adenoasociados (AAV), lentivíricos, adenovíricos u otros tipos de vectores plasmídicos o víricos, en particular, usando formulaciones y dosis de, por ejemplo, las patentes de Estados Unidos n.º 8.454.972 (formulaciones, dosis para adenovirus), 8.404.658 (formulaciones, dosis para AAV) y 5.846.946 (formulaciones, dosis para plásmidos de ADN) y de ensayos clínicos y publicaciones con respecto a los ensayos clínicos que implican lentivirus, AAV y adenovirus. Por ejemplo, para AAV, la vía de administración, la formulación y la dosis pueden ser como en la patente de Estados Unidos n.º 8.454.972 y como ensayos clínicos que implican AAV. Para adenovirus, la vía de administración, formulación y dosis pueden ser como en la patente de Estados Unidos n.º 8.404.658 y como ensayos clínicos que implican adenovirus. Para el suministro de plásmidos, la vía de administración, la formulación y la dosis pueden ser como en la patente de Estados Unidos n.º 5.846.946 y como estudios clínicos que implican plásmidos. Las dosis pueden basarse en o extrapolarse a un individuo promedio de 70 kg (por ejemplo, un ser humano adulto masculino) y pueden ajustarse para pacientes, sujetos, mamíferos de diferente peso y especie. La frecuencia de administración está dentro del ámbito del facultativo médico o veterinario (por ejemplo, médico, veterinario), dependiendo de factores habituales que incluyen la edad, sexo, salud general, otras afecciones del paciente o sujeto y la afección particular o síntomas que se están abordando. Los vectores víricos pueden inyectarse en el tejido de interés. Para modificación genómica específica de tipo celular, la expresión de Cas9 puede dirigirse por un promotor específico de tipo celular. Por ejemplo, la expresión específica de hígado podría usar el promotor de albúmina y la expresión específica de neuronas (por ejemplo, para abordar trastornos del SNC) podría usar el promotor de sinapsina I.

10

15

20

25

En términos de suministro *in vivo*, AAV es ventajoso sobre otros vectores víricos por un par de razones:

30

Baja toxicidad (esto puede deberse al método de purificación que no requiere ultracentrifugación de partículas celulares que pueden activar la respuesta inmunitaria).

Baja probabilidad de causar mutagénesis por inserción porque no se integra en el genoma hospedador.

35

El AAV tiene un límite de empaquetado de 4,5 o 4,75 Kb. Esto significa que Cas9, así como un promotor y terminador de la transcripción tienen que adaptarse todos en el mismo vector vírico. Construcciones mayores de 4,5 o 4,75 Kb darán lugar a producción de virus significativamente reducida. SpCas9 es bastante grande, el propio gen es de más de 4,1 Kb, lo que hace difícil su empaquetado en AAV. Por lo tanto, realizaciones de la invención incluyen la utilización de homólogos de Cas9 que son más cortos. Por ejemplo:

Especie	Tamaño de Cas9
<i>Corynebacter diphtheriae</i>	3252
<i>Eubacterium ventriosum</i>	3321
<i>Streptococcus pasteurianus</i>	3390
<i>Lactobacillus farciminis</i>	3378
<i>Sphaerochaeta globus</i>	3537
<i>Azospirillum B510</i>	3504
<i>Gluconacetobacter diazotrophicus</i>	3150
<i>Neisseria cinerea</i>	3246
<i>Roseburia intestinalis</i>	3420
<i>Parvibaculum lavamentivorans</i>	3111
<i>Staphylococcus aureus</i>	3159
<i>Nitratifactor salsuginis</i> DSM 16511	3396
<i>Campylobacter lari</i> CF89-12	3009
<i>Streptococcus thermophilus</i> LMD-9	3396

40

Por lo tanto, estas especies son en general especies de Cas9 preferidas.

En cuanto a AAV, el AAV puede ser AAV1, AAV2, AAV5 o cualquier combinación de los mismos. Se puede seleccionar el AAV del AAV con respecto a las células a abordar; por ejemplo, se pueden seleccionar los serotipos de AAV 1, 2, 5 o un híbrido o cápside de AAV1, AAV2, AAV5 o cualquier combinación de los mismos para abordar células cerebrales o neuronales; y se puede seleccionar AAV4 para abordar tejido cardíaco. AAV8 es útil para suministro al hígado. Los promotores y vectores de este documento se prefieren individualmente. Una tabla de determinados serotipos de AAV para estas células (véase Grimm, D. *et al.*, J. Virol. 82: 5887-5911 (2008)) es la siguiente:

Línea celular	AAV-1	AAV-2	AAV-3	AAV-4	AAV-5	AAV-6	AAV-8	AAV-9
Huh-7	13	100	2,5	0,0	0,1	10	0,7	0,0
HEK293	25	100	2,5	0,1	0,1	5	0,7	0,1
HeLa	3	100	2,0	0,1	6,7	1	0,2	0,1
HepG2	3	100	16,7	0,3	1,7	5	0,3	ND
Hep1A	20	100	0,2	1,0	0,1	1	0,2	0,0
911	17	100	11	0,2	0,1	17	0,1	ND
CHO	100	100	14	1,4	333	50	10	1,0
COS	33	100	33	3,3	5,0	14	2,0	0,5
MeWo	10	100	20	0,3	6,7	10	1,0	0,2
NIH3T3	10	100	2,9	2,9	0,3	10	0,3	ND
A549	14	100	20	ND	0,5	10	0,5	0,1
HT1180	20	100	10	0,1	0,3	33	0,5	0,1
Monocitos	1111	100	ND	ND	125	1429	ND	ND
DC inmaduras	2500	100	ND	ND	222	2857	ND	ND
DC maduras	2222	100	ND	ND	333	3333	ND	ND

Lentivirus

Los lentivirus son retrovirus complejos que tienen la capacidad de infectar y expresar sus genes tanto en células mitóticas como en células posmitóticas. El lentivirus más conocido comúnmente es el virus de la inmunodeficiencia humana (VIH), que usa las glucoproteínas de la envuelta de otros virus para abordar una amplia gama de tipos celulares.

Los lentivirus pueden prepararse de la siguiente manera. Después de clonar pCasES10 (que contiene una cadena principal plasmídica de transferencia lentivírica), HEK293FT a bajo pase ($p=5$) se sembraron en un matraz T-75 hasta una confluencia del 50 % el día antes de la transfección en DMEM con suero bovino fetal al 10 % y sin antibióticos. Después de 20 horas, se cambió el medio a medio OptiMEM (sin suero) y se hizo transfección 4 horas después. Las células se transfectaron con 10 μ g de plásmido de transferencia lentivírica (pCasES10) y los siguientes plásmidos de empaquetado: 5 μ g de pMD2.G (seudotipo VSV-g), y 7,5 μ g de psPAX2 (gag/pol/rev/tat). La transfección se hizo en 4 ml de OptiMEM con un agente de suministro lipídico catiónico (50 μ l de Lipofectamine 2000 y 100 μ l de reactivo Plus). Después de 6 horas, se cambió el medio a DMEM sin antibióticos con suero bovino fetal al 10 %. Estos métodos usan suero durante el cultivo celular, pero se prefieren métodos sin suero.

Los lentivirus pueden purificarse de la siguiente manera. Se recogieron sobrenadantes víricos después de 48 horas. Los sobrenadantes se aclararon en primer lugar de los desechos y se filtraron a través de un filtro de baja unión a proteínas de 0,45 μ m (PVDF). Entonces se centrifugaron en una ultracentrífuga durante 2 horas a 24 000 rpm. Los sedimentos víricos se resuspendieron en 50 μ l de DMEM durante una noche a 4 °C. Después se dividieron en alícuotas y se congelaron inmediatamente a -80 °C.

En otra realización, también se contemplan vectores lentivíricos mínimos que no son de primate basados en el virus de la anemia infecciosa equina (EIAV) especialmente para genoterapia ocular (véase, por ejemplo, Balagaan, J Gene Med 2006; 8: 275 - 285). En otra realización, RetinoStat®, un vector de genoterapia lentivírica basado en el virus de la anemia infecciosa equina que expresa las proteínas angiostáticas endostatina y angiostatina que se suministra mediante una inyección subretiniana para el tratamiento de la forma de red de la degeneración macular senil también se contempla (véase, por ejemplo, Binley *et al.*, HUMAN GENE THERAPY 23:980-991 (septiembre de 2012)) y este vector puede modificarse para el sistema de CRISPR-Cas de la presente invención.

En otra realización, vectores lentivíricos de autoinactivación con un ARNip dirigido a un exón común compartido por tat/rev de VIH, señuelo TAR de localización nucleolar y una ribozima en cabeza de martillo específica antiCCR5 (véase, por ejemplo, DiGiusto *et al.*, (2010) *Sci Transl Med* 2:36ra43) pueden usarse y/o adaptarse para el sistema de CRISPR-Cas de la presente invención. Un mínimo de $2,5 \times 10^6$ células CD34+ por kilogramo de peso del paciente pueden recogerse y preestimularse durante 16 a 20 horas en medio X-VIVO 15 (Lonza) que contiene 2 $\mu\text{mol/l}$ glutamina, factor de células madre (100 ng/ml), ligando de Flt-3 (Flt-3L) (100 ng/ml) y trombopoyetina (10 ng/ml) (CellGenix) a una densidad de 2×10^6 células/ml. Las células preestimuladas pueden transducirse con lentivirus a una multiplicidad de infección de 5 durante 16 a 24 horas en matraces de cultivo hístico de 75 cm² recubiertos con fibronectina (25 mg/cm²) (RetroNectin, Takara Bio Inc.).

Los vectores lentivíricos se han divulgado como en el tratamiento para la enfermedad de Parkinson, véanse, por ejemplo, la publicación de patente de Estados Unidos n.º 20120295960 y las patentes de Estados Unidos n.º 7303910 y 7351585. Los vectores lentivíricos también se han divulgado para el tratamiento de enfermedades oculares, véanse, por ejemplo, las publicaciones de patentes de Estados Unidos n.º 20060281180, 20090007284, US20110117189; US20090017543; US20070054961, US20100317109. Los vectores lentivíricos también se han divulgado para el suministro al cerebro, véanse, por ejemplo, las publicaciones de patente de Estados Unidos n.º US20110293571; US20110293571, US20040013648, US20070025970, US20090111106 y la patente de Estados Unidos n.º US7259015.

Suministro de ARN

Suministro de ARN: La enzima de CRISPR, por ejemplo una Cas9, y/o cualquiera de los presentes ARN, por ejemplo, un ARN guía, también pueden suministrarse en forma de ARN. El ARNm de Cas9 puede generarse usando transcripción *in vitro*. Por ejemplo, el ARNm de Cas9 puede sintetizarse usando un casete de PCR que contiene los siguientes elementos: promotor T7-secuencia kozak (GCCACC)-Cas9-3' UTR de beta globina-cola poliA (una cadena de 120 o más adeninas). El casete puede usarse para la transcripción por la polimerasa T7. Los ARN guía también pueden transcribirse usando transcripción *in vitro* a partir de un casete que contiene promotor T7-GG-secuencia de ARN guía.

Para potenciar la expresión y reducir la posible toxicidad, la secuencia codificante de la enzima de CRISPR y/o el ARN guía pueden modificarse para incluir uno o más nucleósidos modificados, por ejemplo, usando pseudo-U o 5-metil-C.

Los métodos de suministro de ARNm son especialmente prometedores para el suministro al hígado actualmente.

Mucho trabajo clínico sobre el suministro de ARN se ha centrado en iARN o antisentido, pero estos sistemas pueden adaptarse para el suministro de ARN para implementar la presente invención. En consecuencia deben leerse las referencias a continuación para iARN, etc.

Nanopartículas

Las nanopartículas son un tipo de partícula.

El ARNm de la enzima de CRISPR y el ARN guía pueden suministrarse simultáneamente usando nanopartículas o envueltas lipídicas.

Por ejemplo, Su X, Fricke J, Kavanagh DG, Irvine DJ ("In vitro and in vivo mRNA delivery using lipid-enveloped pH-responsive polymer nanoparticles" *Mol Pharm.* 6 de junio de 2011;8(3):774-87, doi: 10.1021/mp100390w. Epub 1 de abril de 2011) describe nanopartículas de estructura de núcleo-cubierta biodegradables con un núcleo de poli(β -aminoéster) (PBAE) envuelto por una cubierta de bicapa fosfolipídica. Estas se desarrollaron para suministro de ARNm *in vivo*. El componente PBAE sensible a pH se eligió para promover la alteración del endosoma, mientras que la capa superficial lipídica se seleccionó para minimizar la toxicidad del núcleo policatiónico. Por lo tanto, estas se prefieren para el suministro de ARN de la presente invención.

En una realización, se contemplan nanopartículas basadas en polímeros bioadhesivos de autoensamblaje, que pueden aplicarse al suministro oral de péptidos, suministro intravenoso de péptidos y suministro nasal de péptidos, todos al cerebro. Otras realizaciones, tales como absorción oral y suministro ocular de fármacos hidrófobos también se contemplan. La tecnología de envuelta molecular implica una envuelta de polímero manipulado que se protege y suministra al sitio de la enfermedad (véanse, por ejemplo, Mazza, M. *et al.*, *ACSNano*, 2013. 7(2): 1016-1026; Siew, A., *et al.*, *Mol Pharm*, 2012. 9(1):14-28; Lalatsa, A., *et al.*, *J Contr Rel*, 2012. 161(2):523-36; Lalatsa, A., *et al.*, *Mol Pharm*, 2012. 9(6):1665-80; Lalatsa, A., *et al.*, *Mol Pharm*, 2012. 9(6):1764-74; Garrett, N.L., *et al.*, *J Biophotonics*, 2012. 5(5-6):458-68; Garrett, N.L., *et al.*, *J Raman Spect*, 2012. 43(5):681-688; Ahmad, S., *et al.*, *J Royal Soc Interface* 2010. 7:S423-33; Uchegbu, I.F. *Expert Opin Drug Deliv*, 2006. 3(5):629-40; Qu, X., *et al.*, *Biomacromolecules*, 2006. 7(12):3452-9 y Uchegbu, I.F., *et al.*, *Int J Pharm*, 2001. 224:185-199). Se contemplan dosis de aproximadamente 5 mg/kg, con dosis únicas o múltiples, dependiendo del tejido diana.

En una realización, pueden usarse nanopartículas que pueden suministrar ARN a una célula cancerosa para detener el crecimiento tumoral desarrollado por el laboratorio de Dan Anderson en el MIT y/o adaptarse al sistema de Cas de CRISPR de la presente invención. En particular, el laboratorio de Anderson desarrolló sistemas combinatorios completamente automatizados para la síntesis, purificación, caracterización y formulación de nuevos biomateriales y nanoformulaciones. Véanse, por ejemplo, Alabi *et al.*, Proc Natl Acad Sci U S A. 6 de agosto de 2013;110(32):12881-6; Zhang *et al.*, Adv Mater. 6 de septiembre de 2013;25(33):4641-5; Jiang *et al.*, Nano Lett. 13 de marzo de 2013;13(3):1059-64; Karagiannis *et al.*, ACS Nano. 23 de octubre de 2012;6(10):8484-7; Whitehead *et al.*, ACS Nano. 28 de agosto de 2012;6(8):6922-9 y Lee *et al.*, Nat Nanotechnol. 3 de junio de 2012;7(6):389-93.

La solicitud de patente de Estados Unidos 20110293703 se refiere a compuestos lipídicos que también son particularmente útiles en la administración de polinucleótidos, que pueden aplicarse para suministrar el sistema de Cas de CRISPR de la presente invención. En un aspecto, los compuestos lipídicos aminoalcohólicos se combinan con un agente a suministrar a una célula o a un sujeto para formar micropartículas, nanopartículas, liposomas o micelas. El agente a suministrar por las partículas, liposomas o micelas puede estar en forma de un gas, líquido o sólido, y el agente puede ser un polinucleótido, proteína, péptido o molécula pequeña. Los compuestos lipídicos aminoalcohólicos pueden combinarse con otros compuestos lipídicos aminoalcohólicos, polímeros (sintéticos o naturales), tensioactivos, colesterol, carbohidratos, proteínas, lípidos, etc. para formar las partículas. Estas partículas entonces pueden combinarse opcionalmente con un excipiente farmacéutico para formar una composición farmacéutica.

La publicación de patente de Estados Unidos n.º 20110293703 también proporciona métodos de preparación de los compuestos lipídicos aminoalcohólicos. Se permite que uno o más equivalentes de una amina reaccionen con uno o más equivalentes de un compuesto terminado en epóxido en condiciones adecuadas para formar un compuesto lipídico aminoalcohólico de la presente invención. En determinadas realizaciones, todos los grupos amino de la amina se hacen reaccionar completamente con el compuesto terminado en epóxido para formar aminas terciarias. En otras realizaciones, todos los grupos amino de la amina no se hacen reaccionar completamente con el compuesto terminado en epóxido para formar aminas terciarias, produciendo de ese modo aminas primarias o secundarias en el compuesto lipídico aminoalcohólico. Estas aminas primarias o secundarias se dejan tal cual o pueden hacerse reaccionar con otro electrófilo tal como un compuesto terminado en epóxido diferente. Como apreciará un experto en la materia, la reacción de una amina con menos de un exceso de compuesto terminado en epóxido provocará una pluralidad de diferentes compuestos lipídicos aminoalcohólicos con diversos números de colas. Determinadas aminas pueden funcionalizarse completamente con dos colas de compuesto derivado de epóxido mientras que otras moléculas no se funcionalizarán completamente con colas de compuesto derivado de epóxido. Por ejemplo, una diamina o poliamina puede incluir una, dos, tres o cuatro colas de compuesto derivado de epóxido de los diversos restos amino de la molécula produciendo aminas primarias, secundarias y terciarias. En determinadas realizaciones, todos los grupos amino no se funcionalizan completamente. En determinadas realizaciones, se usan dos de los mismos tipos de compuestos terminados en epóxido. En otras realizaciones, se usan dos o más compuestos terminados en epóxido diferentes. La síntesis de los compuestos lipídicos aminoalcohólicos se realiza con o sin disolvente, y la síntesis puede realizarse a mayores temperaturas que varían de 30-100 °C, preferiblemente a aproximadamente 50-90 °C. Los compuestos lipídicos aminoalcohólicos preparados pueden purificarse opcionalmente. Por ejemplo, la mezcla de compuestos lipídicos aminoalcohólicos puede purificarse para producir un compuesto lipídico aminoalcohólico con un número particular de colas de compuesto derivado de epóxido. O la mezcla puede purificarse para producir un estereoisómero o regioisómero particular. Los compuestos lipídicos aminoalcohólicos pueden alquilarse usando un haluro de alquilo (por ejemplo, yoduro de metilo) u otro agente alquilante y/o pueden acilarse.

La publicación de patente de Estados Unidos n.º 20110293703 también proporciona colecciones de compuestos lipídicos aminoalcohólicos preparados por los métodos de la invención. Estos compuestos lipídicos aminoalcohólicos pueden prepararse y/o cribarse usando técnicas de alto rendimiento que implican manipuladores de líquidos, robots, placas de microvaloración, ordenadores, etc. En determinadas realizaciones, los compuestos lipídicos aminoalcohólicos se criban por su capacidad de transfectar polinucleótidos u otros agentes (por ejemplo, proteínas, péptidos, moléculas pequeñas) en la célula.

La publicación de patente de Estados Unidos n.º 20130302401 se refiere a una clase de poli(beta-aminoalcoholes) (PBAA) que se ha preparado usando polimerización combinatoria. Los PBAA de la invención pueden usarse en biotecnología y aplicaciones biomédicas como recubrimientos (tales como recubrimientos de películas o películas de múltiples capas para dispositivos médicos o implantes), aditivos, materiales, excipientes, agentes que no biocontaminan, agentes de formación de micropatrones y agentes de encapsulación celular. Cuando se usan como recubrimientos superficiales, estos PBAA provocaban diferentes niveles de inflamación, tanto *in vitro* como *in vivo*, dependiendo de sus estructuras químicas. La gran diversidad química de esta clase de materiales nos permitió identificar recubrimientos poliméricos que inhiben la activación de macrófagos *in vitro*. Además, estos recubrimientos reducen el reclutamiento de células inflamatorias y reducen la fibrosis, después del implante subcutáneo de micropartículas de poliestireno carboxiladas. Estos polímeros pueden usarse para formar cápsulas complejas de polielectrolitos para encapsulación celular. La invención también puede tener muchas otras aplicaciones biológicas tales como recubrimientos antimicrobianos, suministro de ADN o ARNip y manipulación de tejidos de células madre. Los contenidos de la publicación de patente de Estados Unidos n.º 20130302401 pueden aplicarse al sistema de Cas de CRISPR Cas de la presente invención.

En otra realización, se contemplan nanopartículas lipídicas (LNP). Se ha encapsulado un ARN interferente pequeño antitransstiretina en nanopartículas lipídicas y se ha suministrado a seres humanos (véase, por ejemplo, Coelho *et al.*, N Engl J Med 2013;369:819-29), y dicho sistema puede adaptarse y aplicarse al sistema de Cas de CRISPR de la presente invención. Se contemplan dosis de aproximadamente 0,01 a aproximadamente 1 mg por kg de peso corporal administradas por vía intravenosa. Se contemplan medicamentos para reducir el riesgo de reacciones relacionadas con infusión, tales como dexametasona, acetaminofeno, difenhidramina o cetirizina y se contemplan de ranitidina. También se contemplan múltiples dosis de aproximadamente 0,3 mg por kilogramo cada 4 semanas durante cinco dosis.

Las LNP han demostrado ser muy eficaces en el suministro de ARNip al hígado (véase, por ejemplo, Taberero *et al.*, Cancer Discovery, abril de 2013, Vol. 3, n.º 4, páginas 363-470) y, por lo tanto, se contemplan para el suministro de ARN que codifica Cas de CRISPR al hígado. Puede contemplarse una dosificación de aproximadamente cuatro dosis de 6 mg/kg de la LNP cada dos semanas. Taberero *et al.* demostraron que se observaba regresión tumoral después de los 2 primeros ciclos de las LNP dosificadas a 0,7 mg/kg, y al final de los 6 ciclos el paciente había conseguido una respuesta parcial con regresión completa de la metástasis del ganglio linfático y reducción sustancial de los tumores hepáticos. Se obtuvo una respuesta completa después de 40 dosis en este paciente, que permaneció en remisión y completó el tratamiento después de recibir dosis durante 26 meses. Dos pacientes con RCC y sitios extrahepáticos de enfermedad incluyendo el riñón, el pulmón y los ganglios linfáticos que estaban progresando después del tratamiento previo con inhibidores de la ruta de VEGF tenían enfermedad estable en todos los sitios durante aproximadamente 8 a 12 meses, y un paciente con PNET y metástasis hepática siguió en el estudio de prolongación durante 18 meses (36 dosis) con enfermedad estable.

Sin embargo, la carga de la LNP debe tenerse en consideración. Según los lípidos catiónicos se combinan con lípidos cargados negativamente inducen estructuras que no son de bicapa que facilitan el suministro intracelular. Como las LNP cargadas se eliminan rápidamente de la circulación después de inyección intravenosa, se desarrollaron lípidos catiónicos ionizables con valores de pKa por debajo de 7 (véase, por ejemplo, Rosin *et al.*, Molecular Therapy, vol. 19, n.º 12, páginas 1286-2200, diciembre de 2011). Polímeros cargados negativamente tales como ARN pueden cargarse en LNP a valores de pH bajos (por ejemplo, pH 4) donde los lípidos ionizables presentan una carga positiva. Sin embargo, a valores de pH fisiológicos, las LNP muestran una baja carga superficial compatible con tiempos en circulación más largos. Nos hemos centrado en cuatro especies de lípidos catiónicos ionizables, concretamente 1,2-dilinoil-3-dimetilamonio-propano (DLinDAP), 1,2-dilinoileiloxi-3-NN-dimetilaminopropano (DLinDMA), 1,2-dilinoileiloxi-ceto-N,N-dimetil-3-aminopropano (DLinKDMA), y 1,2-dilinoileil-4-(2-dimetilaminoetil)-[1,3]-dioxolano (DLinKC2-DMA). Se ha demostrado que los sistemas de ARNip de LNP que contienen estos lípidos muestran propiedades de silenciamiento génico notablemente diferentes en hepatocitos *in vivo*, con potencias que varían de acuerdo con la serie DLinKC2-DMA>DLinKDMA>DLinDMA>>DLinDAP empleando un modelo de silenciamiento génico del factor VII (véase, por ejemplo, Rosin *et al.*, Molecular Therapy, vol. 19, n.º 12, páginas 1286-2200, diciembre de 2011). Una dosificación de 1 µg/ml de LNP o ARN de CRISPR-Cas en o asociado con la LNP puede contemplarse, especialmente para una formulación que contiene DLinKC2-DMA.

La preparación de LNP y la encapsulación de Cas de CRISPR puede usarse y/o adaptarse de Rosin *et al.*, Molecular Therapy, vol. 19, n.º 12, páginas 1286-2200, diciembre de 2011). Los lípidos catiónicos 1,2-dilinoil-3-dimetilamonio-propano (DLinDAP), 1,2-dilinoileiloxi-3-N,N-dimetilaminopropano (DLinDMA), 1,2-dilinoileiloxi-ceto-N,N-dimetil-3-aminopropano (DLinK-DMA), 1,2-dilinoileil-4-(2-dimetilaminoetil)-[1,3]-dioxolano (DLinKC2-DMA), (3-o-[2"-(metoxipoli(etilenglicol)2000) succinil]-1,2-dimiristil-ol-sn-glicol (PEG-S-DMG) y R-3-[(ω-metoxi-poli(etilenglicol)2000) carbamoil]-1,2-dimiristiloxipropil-3-amina (PEG-C-DOMG) pueden proporcionarse por Tekmira Pharmaceuticals (Vancouver, Canadá) o sintetizarse. El colesterol puede adquirirse de Sigma (St Louis, MO). El ARN de Cas de CRISPR específico puede encapsularse en LNP que contienen DLinDAP, DLinDMA, DLinK-DMA y DLinKC2-DMA (lípidos catiónicos:DSPC:CHOL: PEGS-DMG o PEG-C-DOMG a relaciones molares de 40:10:40:10). Cuando se requiere, puede incorporarse SP-DiOC18 al 0,2 % (Invitrogen, Burlington, Canadá) para evaluar la captación celular, el suministro intracelular y la biodistribución. La encapsulación puede realizarse disolviendo mezclas lipídicas compuestas de lípidos catiónicos::DSPC:colesterol:PEG-c-DOMG (relación molar de 40:10:40:10) en etanol hasta una concentración de lípidos final de 10 mmol/l. Esta solución en etanol de lípidos puede añadirse gota a gota a 50 mmol/l de citrato, a pH 4,0 para formar vesículas multilaminares para producir una concentración final de etanol al 30 % vol/vol. Pueden formarse vesículas unilaminares grandes después de extrusión de vesículas multilaminares a través de dos filtros de policarbonato Nuclepore de 80 nm apilados usando la Extruder (Northern Lipids, Vancouver, Canadá). La encapsulación puede conseguirse añadiendo ARN disuelto a 2 mg/ml en 50 mmol/l de citrato, a pH 4,0 que contiene etanol al 30 % vol/vol gota a gota a vesículas unilaminares grandes preformadas extrudidas e incubación a 31 °C durante 30 minutos con mezcla constante hasta una relación ponderal final de ARN/lípido de 0,06/1 p/p. La eliminación de etanol y la neutralización del tampón de formulación se realizaron por diálisis frente a solución salina tamponada con fosfato (PBS), a pH 7,4 durante 16 horas usando membranas de diálisis de celulosa regeneradas Spectra/Por 2. La distribución de tamaños de las nanopartículas puede determinarse por dispersión de luz dinámica usando un medidor de tamaño de partícula NICOMP 370, los modos de vesícula/intensidad y ajuste gaussiano (Nicomp Particle Sizing, Santa Bárbara, CA). El tamaño de partículas para los tres sistemas de LNP puede ser ~70 nm de diámetro. La eficacia de encapsulación de ARN puede determinarse por eliminación del ARN libre usando columnas VivaPureD MiniH (Sartorius Stedim Biotech) de muestras recogidas antes y después de la diálisis. El ARN encapsulado puede

extraerse de las nanopartículas eluidas y cuantificarse a 260 nm. La relación de ARN a lípido se determinó mediante la medición del contenido de colesterol en vesículas usando el ensayo enzimático de colesterol E de Wako Chemicals USA (Richmond, VA). Junto con el análisis de este documento de las LNP y los lípidos PEG, liposomas PEGilados o LNP son asimismo adecuados para el suministro de un sistema de CRISPR-Cas o componentes del mismo.

La preparación de LNP grandes puede usarse y/o adaptarse de Rosin *et al.*, *Molecular Therapy*, vol. 19, n.º 12, páginas 1286-2200, diciembre de 2011. Una solución de premezcla de lípidos (concentración total de lípidos de 20,4 mg/ml) puede prepararse en etanol que contiene DLinKC2-DMA, DSPC y colesterol a relaciones molares de 50:10:38,5. Puede añadirse acetato de sodio a la premezcla de lípidos a una relación molar de 0,75:1 (acetato de sodio:DLinKC2-DMA). Los lípidos posteriormente pueden hidratarse combinando la mezcla con 1,85 volúmenes de tampón citrato (10 mmol/l, pH 3,0) con agitación vigorosa, provocando la formación espontánea de liposomas en tampón acuoso que contiene un 35 % de etanol. La solución de liposomas puede incubarse a 37 °C para permitir un aumento dependiente del tiempo en el tamaño de las partículas. Pueden retirarse alícuotas en diversos momentos durante la incubación para investigar cambios en el tamaño de los liposomas por dispersión de luz dinámica (Zetasizer Nano ZS, Malvern Instruments, Worcestershire, RU). Una vez conseguido el tamaño de partícula deseado, puede añadirse una solución de lípido PEG acuoso (solución madre = 10 mg/ml de PEG-DMG en etanol al 35 % (vol/vol)) a la mezcla de liposomas para producir una concentración molar de PEG final de un 3,5 % de lípidos totales. Tras la adición de PEG-lípidos, los liposomas deben ser de su tamaño, inactivando de forma eficaz el crecimiento adicional. Entonces puede añadirse ARN a los liposomas vacíos a una relación de ARN a lípidos totales de aproximadamente 1:10 (p:p), seguido de incubación durante 30 minutos a 37 °C para formar LNP cargadas. La mezcla posteriormente puede dializarse durante una noche en PBS y filtrarse con un filtro de jeringa de 0,45 µm.

También se contemplan construcciones de ácido nucleico esféricas (SNA™) y otras nanopartículas (particularmente nanopartículas de oro) como un medio para suministrar el sistema de CRISPR-Cas a dianas pretendidas. Los datos significativos muestran que las construcciones de ácido nucleico esféricas (SNA™) de AuraSense Therapeutics, basadas en nanopartículas de oro funcionalizadas con ácido nucleico, son útiles.

La bibliografía que puede emplearse junto con los contenidos de este documento incluyen: Cutler *et al.*, *J. Am. Chem. Soc.* 2011 133:9254-9257, Hao *et al.*, *Small*. 2011 7:3158-3162, Zhang *et al.*, *ACS Nano*. 2011 5:6962-6970, Cutler *et al.*, *J. Am. Chem. Soc.* 2012 134:1376-1391, Young *et al.*, *Nano Lett.* 2012 12:3867-71, Zheng *et al.*, *Proc. Natl. Acad. Sci. USA*. 2012 109:11975-80, Mirkin, *Nanomedicine* 2012 7:635-638 Zhang *et al.*, *J. Am. Chem. Soc.* 2012 134:16488-1691, Weintraub, *Nature* 2013 495:S14-S16, Choi *et al.*, *Proc. Natl. Acad. Sci. USA*. 2013 110(19):7625-7630, Jensen *et al.*, *Sci. Transl. Med.* 5, 209ra152 (2013) y Mirkin, *et al.*, *Small*, 10:186-192.

Pueden construirse nanopartículas de autoensamblaje con ARN con polietilenimina (PEI) que está PEGilada con un ligando peptídico de Arg-Gly-Asp (RGD) fijado al extremo distal del polietilenglicol (PEG). Este sistema se ha usado, por ejemplo, como un medio para abordar la neovasculatura tumoral que expresa integrinas y suministrar ARNip que inhibe la expresión del receptor 2 del factor de crecimiento del endotelio vascular (VEGF R2) y conseguir de ese modo angiogénesis tumoral (véase, por ejemplo, Schiffelers *et al.*, *Nucleic Acids Research*, 2004, Vol. 32, n.º 19). Pueden prepararse nanoplexes mezclando volúmenes iguales de soluciones acuosas de polímero catiónico y ácido nucleico para dar un exceso molar neto de nitrógeno ionizable (polímero) a fosfato (ácido nucleico) sobre el intervalo de 2 a 6. Las interacciones electrostáticas entre polímeros catiónicos y ácido nucleico provocaron la formación de poliplexes con una distribución promedio de los tamaños de partícula de aproximadamente 100 nm, denominados por tanto aquí nanoplexes. Se prevé una dosificación de aproximadamente 100 a 200 mg de Cas de CRISPR para su suministro en las nanopartículas de autoensamblaje de Schiffelers *et al.*

Los nanoplexes de Bartlett *et al.* (*PNAS*, 25 de septiembre de 2007, vol. 104, n.º 39) también pueden aplicarse a la presente invención. Los nanoplexes de Bartlett *et al.* se preparan mezclando volúmenes iguales de soluciones acuosas de polímero catiónico y ácido nucleico para dar un exceso molar neto de nitrógeno ionizable (polímero) a fosfato (ácido nucleico) sobre el intervalo de 2 a 6. Las interacciones electrostáticas entre polímeros catiónicos y ácido nucleico provocaron la formación de poliplexes con una distribución promedio de los tamaños de partícula de aproximadamente 100 nm, denominados por tanto aquí nanoplexes. El DOTA-ARNip de Bartlett *et al.* se sintetizó de la siguiente manera: Se solicitó éster de mono(N-hidroxisuccinimida) del ácido 1,4,7,10-tetraazaciclododecano-1,4,7,10-tetraacético (éster DOTA-NHS) de Macrocyclics (Dallas, TX). La hebra de sentido de ARN modificada con amina con un exceso molar de factor 100 de éster de DOTA-NHS en tampón carbonato (pH 9) se añadió a un tubo de microcentrífuga. Los contenidos se hicieron reaccionar por agitación durante 4 h a temperatura ambiente. El conjugado de DOTA-ARN de sentido se precipitó en etanol, se resuspendió en agua y se hibridó con la hebra de antisentido no modificada para producir DOTA-ARNip. Todos los líquidos se pretrataron con Chelex-100 (Bio-Rad, Hércules, CA) para eliminar los contaminantes de oligoelementos metálicos. Pueden formarse nanopartículas de ARNip dirigidas a Tf y no dirigidas usando policationes que contienen ciclodextrina. Típicamente, se formaron nanopartículas en agua a una relación de carga de 3 (+/-) y una concentración de ARNip de 0,5 g/litro. Un porcentaje de las moléculas de adamantano-PEG sobre la superficie de las nanopartículas dirigidas se modificó con Tf (adamantano-PEG-Tf). Las nanopartículas se suspendieron en un solución de vehículo de glucosa al 5 % (p/vol) para inyección.

Davis *et al.* (*Nature*, Vol. 464, 15 de abril de 2010) realizan un ensayo clínico de ARN que usa un sistema de suministro de nanopartículas dirigido (número de registro de ensayo clínico NCT00689065). A pacientes con cánceres sólidos

resistentes a tratamientos de referencia se les administra dosis de nanopartículas dirigidas en los días 1, 3, 8 y 10 de un ciclo de 21 días por una infusión intravenosa de 30 min. Las nanopartículas consisten en un sistema de suministro sintético que contiene: (1) un polímero basado en ciclodextrina lineal (CDP), (2) un ligando dirigido a la proteína transferrina humana (TF) presentado en el exterior de la nanopartícula para acoplar con los receptores de TF (TFR) en la superficie de las células cancerosas, (3) un polímero hidrófilo (polietilenglicol (PEG) usado para promover la estabilidad de las nanopartículas en líquidos biológicos) y (4) ARNip diseñado para reducir la expresión de RRM2 (la secuencia usada en clínica se indicó previamente como siR2B+5). El TFR se ha sabido desde hace tiempo que está regulado por aumento en células malignas, y RRM2 es un agente antineoplásico establecido. Estas nanopartículas (versión clínica indicada como CALAA-01) han demostrado tolerarse bien en estudios de múltiples dosis en primates no humanos. Aunque a un solo paciente con leucemia mieloide crónica se le ha administrado ARNip por suministro liposómico, el ensayo clínico de Davis *et al.* es el ensayo inicial en seres humanos para suministrar sistémicamente ARNip con un sistema de suministro dirigido y para tratar a pacientes con cáncer sólido. Para averiguar si el sistema de suministro dirigido puede proporcionar suministro eficaz de ARNip funcional a tumores humanos, Davis *et al.* investigaron biopsias de tres pacientes de tres cohortes de dosificación diferentes; los pacientes A, B y C, que tenían todos ellos melanoma metastásico y recibieron dosis CALAA-01 de 18, 24 y 30 mg m⁻² de ARNip, respectivamente. También pueden contemplarse dosis similares para el sistema de Cas de CRISPR de la presente invención. El suministro de la invención puede conseguirse con nanopartículas que contienen un polímero basado en ciclodextrina lineal (CDP), un ligando dirigido a proteína transferrina humana (TF) presentado en el exterior de la nanopartícula para acoplar con receptores de TF (TFR) en la superficie de las células cancerosas y/o un polímero hidrófilo (por ejemplo, polietilenglicol (PEG) usado para promover la estabilidad de las nanopartículas en líquidos biológicos).

En términos de esta invención, se prefiere tener uno o más componentes del complejo de CRISPR, por ejemplo, la enzima de CRISPR o ARNm o ARN guía suministrados usando nanopartículas o envueltas lipídicas. Otros sistemas de suministro o vectores pueden usarse junto con los aspectos de nanopartícula de la invención.

En general, una "nanopartícula" se refiere a cualquier partícula que tenga un diámetro de menos de 1000 nm. En determinadas realizaciones preferidas, las nanopartículas de la invención tienen una dimensión máxima (por ejemplo, diámetro) de 500 nm o menos. En otras realizaciones preferidas, las nanopartículas de la invención tienen una dimensión máxima que varía entre 25 nm y 200 nm. En otras realizaciones preferidas, las nanopartículas de la invención tienen una dimensión máxima de 100 nm o menos. En otras realizaciones preferidas, las nanopartículas de la invención tienen una dimensión máxima que varía entre 35 nm y 60 nm.

Las nanopartículas abarcadas en la presente invención pueden proporcionarse en diferentes formas, por ejemplo, como nanopartículas sólidas (por ejemplo, metal tal como plata, oro, hierro, titanio), no metálicas, sólidos basados en lípidos, polímeros, suspensiones de nanopartículas o combinaciones de los mismos. Pueden prepararse nanopartículas metálicas, dieléctricas y semiconductoras, así como estructuras híbridas (por ejemplo, nanopartículas de núcleo y cubierta). Las nanopartículas hechas de material semiconductor también pueden marcarse con puntos cuánticos si son suficientemente pequeñas (típicamente por debajo de 10 nm) para que se produzca la cuantización de los niveles de energía electrónica. Dichas partículas a nanoescala se usan en aplicaciones biomédicas como vehículos de fármacos o agentes de imágenes y pueden adaptarse para fines similares en la presente invención.

Se han fabricado nanopartículas semisólidas y blandas y están dentro del alcance de la presente invención. Un prototipo de nanopartícula de naturaleza semisólida es el liposoma. Diversos tipos de nanopartículas de liposoma actualmente se usan clínicamente como sistemas de suministro para fármacos antineoplásicos y vacunas. Las nanopartículas con una mitad hidrófila y la otra mitad hidrófoba se denominan partículas Janus y son particularmente eficaces para estabilizar emulsiones. Pueden autoensamblarse en superficies de contacto de agua/aceite y actúan como tensioactivos sólidos.

La patente de Estados Unidos n.º 8.709.843 proporciona un sistema de suministro de fármacos para suministro dirigido de partículas que contienen agente terapéutico a tejidos, células y compartimentos intracelulares. La invención proporciona partículas dirigidas que comprenden polímero conjugado con un tensioactivo, polímero hidrófilo o lípido.

La patente de Estados Unidos n.º 6.007.845 proporciona partículas que tienen un núcleo de un copolímero multibloque formado por unión covalente de un compuesto multifuncional con uno o más polímeros hidrófobos y uno o más polímeros hidrófilos, y contienen un material biológicamente activo.

La patente de Estados Unidos n.º 5.855.913 proporciona una composición en partículas que tiene partículas aerodinámicamente ligeras que tienen una densidad de menos de 0,4 g/cm³ con un diámetro medio entre 5 µm y 30 µm, que incorporan un tensioactivo en la superficie de las mismas para el suministro de fármacos al sistema pulmonar.

La patente de Estados Unidos n.º 5.985.309 proporciona partículas que incorporan un tensioactivo y/o un complejo hidrófilo o hidrófobo de un agente terapéutico o de diagnóstico cargado positiva o negativamente y una molécula cargada de carga opuesta para su suministro al sistema pulmonar.

La patente de Estados Unidos n.º 5.543.158 proporciona nanopartículas inyectables biodegradables que tienen un núcleo sólido biodegradable que contiene un material biológicamente activo y restos de polialquilenglicol sobre la superficie.

5 El documento WO2012135025 (también publicado como US20120251560), describe polímeros de polietilenimina (PEI) conjugados y azamacrociclos conjugados (denominados colectivamente "lipómero conjugado" o "lipómeros"). En determinadas realizaciones, puede preverse que dichos lipómeros conjugados puedan usarse en el contexto del sistema de CRISPR-Cas para conseguir perturbaciones genómicas *in vitro*, *ex vivo* e *in vivo* para modificar la expresión génica, incluyendo la modulación de la expresión proteínica.

10 En una realización, la nanopartícula puede ser polímero lipídico modificado con epóxido, ventajosamente 7C1 (véase, por ejemplo, James E. Dahlman y Carmen Barnes *et al.*, Nature Nanotechnology (2014) publicado en línea el 11 de mayo de 2014, doi:10.1038/nnano.2014.84). C71 se sintetizó haciendo reaccionar lípidos terminados en epóxido C15 con PEI600 a una relación molar de 14:1, y se formuló con C14PEG2000 para producir nanopartículas (diámetro entre 15 35 y 60 nm) que eran estables en solución de PBS durante al menos 40 días.

Un polímero lipídico modificado con epóxido puede utilizarse para suministrar el sistema de CRISPR-Cas de la presente invención a células pulmonares, cardiovasculares o renales, sin embargo, un experto en la materia puede adaptar el sistema para suministro a otros órganos diana. Se prevén dosificaciones que varían de aproximadamente 20 0,05 a aproximadamente 0,6 mg/kg. También se prevén dosificaciones durante varios días o semanas, con una dosificación total de aproximadamente 2 mg/kg.

Sistemas de suministro de partículas y/o formulaciones:

25 Varios tipos de sistemas de suministro de partículas y/o formulaciones se sabe que son útiles en un diverso espectro de aplicaciones biomédicas. En general, una partícula se define como un objeto pequeño que se comporta como una unidad completa con respecto a su transporte y propiedades. Las partículas se clasifican además de acuerdo con el diámetro. Las partículas gruesas cubren un intervalo entre 2500 y 10 000 nanómetros. Las partículas finas tienen un tamaño entre 100 y 2500 nanómetros. Las partículas ultrafinas, o nanopartículas, en general son entre 1 y 100 30 nanómetros de tamaño. La base del límite de 100 nm es el hecho de que las propiedades novedosas que diferencian las partículas del material en bruto típicamente se desarrollan a una escala de longitud crucial por debajo de 100 nm.

Como se usa en este documento, un sistema de suministro de partículas/formulación se define como cualquier sistema de suministro biológico/formulación que incluya una partícula de acuerdo con la presente invención. Una partícula de 35 acuerdo con la presente invención es cualquier entidad que tenga una dimensión máxima (por ejemplo, diámetro) de menos de 100 micrómetros (µm). En algunas realizaciones, las partículas de la invención tienen una dimensión máxima de menos de 10 µm. En algunas realizaciones, las partículas de la invención tienen una dimensión máxima de menos de 2000 nanómetros (nm). En algunas realizaciones, las partículas de la invención tienen una dimensión máxima de menos de 1000 nanómetros (nm). En algunas realizaciones, las partículas de la invención tienen una 40 dimensión máxima de menos de 900 nm, 800 nm, 700 nm, 600 nm, 500 nm, 400 nm, 300 nm, 200 nm o 100 nm. Típicamente, las partículas de la invención tienen una dimensión máxima (por ejemplo, diámetro) de 500 nm o menos. En algunas realizaciones, las partículas de la invención tienen una dimensión máxima (por ejemplo, diámetro) de 250 nm o menos. En determinadas realizaciones preferidas, las nanopartículas de la invención tienen una dimensión máxima (por ejemplo, diámetro) de 200 nm o menos. En determinadas realizaciones preferidas, las nanopartículas de la invención tienen una dimensión máxima (por ejemplo, diámetro) de 150 nm o menos. En determinadas realizaciones preferidas, las nanopartículas de la invención tienen una dimensión máxima (por ejemplo, diámetro) de 100 nm o 45 menos. Partículas más pequeñas, por ejemplo, que tienen una dimensión máxima de 50 nm o menos se usan en algunas realizaciones de la invención. En algunas realizaciones, las partículas de la invención tienen una dimensión máxima que varía entre 25 nm y 200 nm.

50 La caracterización de partículas (incluyendo, por ejemplo, caracterización de la morfología, dimensión, etc.) se hace usando una diversidad de técnicas diferentes. Las técnicas comunes son microscopía electrónica (TEM, SEM), microscopía de fuerza atómica (AFM), dispersión de luz dinámica (DLS), espectroscopia fotoelectrónica de rayos X (XPS), difracción de polvo de rayos X (XRD), espectroscopia de infrarrojos con transformada de Fourier (FTIR), espectrometría de masas por tiempo de vuelo con desorción/ionización láser asistida por matriz (MALDI-TOF), espectroscopia en el ultravioleta-visible, interferometría de polarización doble y resonancia magnética nuclear (RMN). La caracterización (mediciones de dimensión) puede hacerse en cuanto las partículas naturales (es decir, antes de la carga) o después de cargarse la carga (en este documento, carga se refiere a, por ejemplo, uno o más componentes del sistema de CRISPR-Cas, por ejemplo, la enzima de CRISPR o ARNm o ARN guía, o cualquier combinación de los 60 mismos, y puede incluir vehículos adicionales y/o excipientes) para proporcionar partículas de un tamaño óptimo para su suministro para cualquier aplicación *in vitro*, *ex vivo* y/o *in vivo* de la presente invención. En determinadas realizaciones preferidas, la caracterización de las dimensiones de las partículas (por ejemplo, diámetro) se basa en mediciones usando dispersión de láser dinámica (DLS). Se hace mención a la patente de Estados Unidos n.º 8.709.843; la patente de Estados Unidos n.º 6.007.845; la patente de Estados Unidos n.º 5.855.913; la patente de Estados Unidos n.º 5.985.309; la patente de Estados Unidos n.º 5.543.158; y la publicación de James E. Dahlman y Carmen Barnes 65

et al., Nature Nanotechnology (2014) publicada en línea el 11 de mayo de 2014, doi:10.1038/nnano.2014.84, que se refiere a partículas, métodos de preparación y uso de las mismas y mediciones de las mismas.

Los sistemas de suministro de partículas dentro del alcance de la presente invención pueden proporcionarse en cualquier forma incluyendo, aunque sin limitación, partículas sólidas, semisólidas, en emulsión o coloidales. Por tanto, cualquiera de los sistemas de suministro descritos en este documento incluyendo, aunque sin limitación, por ejemplo, sistemas basados en lípidos, liposomas, micelas, microvesículas, exosomas o cañón de genes pueden proporcionarse como sistemas de suministro de partículas dentro del alcance de la presente invención.

Exosomas

Los exosomas son nanovesículas endógenas que transportan ARN y proteínas, y que pueden suministrar ARN al cerebro y otros órganos diana. Para reducir la inmunogenia, Alvarez-Erviti *et al.*, (2011, Nat Biotechnol 29: 341) usaron células dendríticas autoderivadas para la producción de exosomas. La dirección al cerebro se consiguió manipulando las células dendríticas para que expresaran Lamp2b, una proteína de membrana exosómica, fusionada al péptido RVG específico de neuronas. Los exosomas purificados se cargaron con ARN exógeno por electroporación. Los exosomas dirigidos por RVG inyectados por vía intravenosa suministraron ARNip de GAPDH específicamente a neuronas, microglía, oligodendrocitos en el cerebro, provocando atenuación génica específica. La exposición previa a exosomas de RVG no atenuó la atenuación y no se observó captación no específica en otros tejidos. El potencial terapéutico del suministro de ARNip mediado por exosomas se demostró por la fuerte atenuación de ARNm (60 %) y proteína (62 %) de BACE1, una diana terapéutica en enfermedad de Alzheimer.

Para obtener una combinación de exosomas inmunológicamente inertes, Alvarez-Erviti *et al.* recogieron médula ósea de ratones C57BL/6 endogámicos con un haplotipo homogéneo del complejo principal de histocompatibilidad (MHC). Como las células dendríticas inmaduras producen grandes cantidades de exosomas desprovistos de activadores de linfocitos T tales como MHC-II y CD86, Alvarez-Erviti *et al.* seleccionaron células dendríticas con factor estimulador de colonias de granulocitos/macrófagos (GM-CSF) durante 7 días. Los exosomas se purificaron del sobrenadante de cultivo el siguiente día usando protocolos bien establecidos de ultracentrifugación. Los exosomas producidos eran físicamente homogéneos, con una distribución de tamaños con un máximo a 80 nm de diámetro determinado por análisis de rastreo de nanopartículas (NTA) y microscopía electrónica. Alvarez-Erviti *et al.* obtuvieron 6-12 µg de exosomas (medidos basándose en la concentración de proteína) por 10⁶ células.

A continuación, Alvarez-Erviti *et al.* investigaron la posibilidad de cargar exosomas modificados con cargas exógenas usando protocolos de electroporación adaptados para aplicaciones a nanoescala. Como la electroporación para partículas de membrana a escala nanométrica no está bien caracterizada, se usó ARN marcado con Cy5 no específico para la optimización empírica del protocolo de electroporación. La cantidad de ARN encapsulado se ensayó después de ultracentrifugación y lisis de los exosomas. La electroporación a 400 V y 125 µF produjo la máxima retención de ARN y se usó para todos los experimentos posteriores.

Alvarez-Erviti *et al.* administraron 150 µg de cada ARNip de BACE1 encapsulado en 150 µg de exosomas de RVG a ratones C57BL/6 normales y compararon la eficacia de atenuación con cuatro controles: ratones no tratados, ratones a los que se inyectó exosomas de RVG solamente, ratones a los que se inyectó ARNip de BACE1 en complejo con un reactivo de liposoma catiónico *in vivo* y ratones a los que se inyectó ARNip de BACE1 en complejo con RVG-9R, el péptido RVG conjugado con 9 D-argininas que se une electrostáticamente al ARNip. Se analizaron muestras de tejido cortical 3 d después de la administración y se observó una atenuación significativa de las proteínas (45 %, P < 0,05, frente a 62 %, P < 0,01) tanto en los ratones tratados con ARNip-RVG-9R como en los ratones tratados con exosoma de ARNip RVG, resultante de una disminución significativa en los niveles de ARNm de BACE1 (66 % [+ o -] 15 %, P < 0,001 y 61 % [+ o -] 13 % respectivamente, P < 0,01). Además, los solicitantes demostraron una disminución significativa (55 %, P < 0,05) en los niveles totales de [beta]-amiloides 1-42, un componente principal de las placas amiloides en la patología de Alzheimer, en los animales tratados con exosoma de RVG. La disminución observada fue mayor que la disminución de β-amiloides 1-40 demostrada en ratones normales después de inyección intraventricular de inhibidores de BACE1. Alvarez-Erviti *et al.* realizaron amplificación rápida 5' de los extremos de ADNc (RACE) en el producto de escisión de BACE1, que proporcionó evidencias de atenuación mediada por iARN por el ARNip.

Finalmente, Alvarez-Erviti *et al.* investigaron si los exosomas de ARN-RVG inducían respuestas inmunitarias *in vivo* evaluando las concentraciones en suero de IL-6, IP-10, TNFα e IFN-α. Después del tratamiento con exosomas, se registraron cambios no significativos en todas las citocinas similares al tratamiento con reactivo de transfección de ARNip en contraste con ARNip-RVG-9R, que estimulaba potentemente la secreción de IL-6, confirmando el perfil inmunológicamente inerte del tratamiento con exosomas. Dado que los exosomas encapsulan únicamente un 20 % de ARNip, parece que el suministro con exosoma de RVG es más eficaz que el suministro de RVG-9R ya que se conseguía atenuación de ARNm comparable y mayor atenuación de proteínas con cinco veces menos de ARNip sin el nivel correspondiente de estimulación inmunitaria. Este experimento demostró el potencial terapéutico de la tecnología de exosomas de RVG, que es potencialmente adecuado para silenciamiento a largo plazo de genes relacionados con enfermedades neurodegenerativas. El sistema de suministro de exosomas de Alvarez-Erviti *et al.* puede aplicarse para suministrar el sistema de CRISPR-Cas de la presente invención a dianas terapéuticas, especialmente enfermedades neurodegenerativas. Puede contemplarse una dosificación de aproximadamente 100 a

1000 mg de Cas de CRISPR encapsulado en aproximadamente 100 a 1000 mg de exosomas de RVG para la presente invención.

5 El-Andaloussi *et al.* (Nature Protocols 7, 2112-2126(2012)) divulgan la manera en que exosomas derivados de células cultivadas pueden aprovecharse para el suministro de ARN *in vitro* e *in vivo*. Este protocolo describe por primera vez la generación de exosomas dirigidos a través de transfección de un vector de expresión, que comprende una proteína exosómica fusionada con un ligando peptídico. A continuación, El-Andaloussi *et al.* explican la manera de purificar y caracterizar los exosomas a partir del sobrenadante de células transfectadas. A continuación, El-Andaloussi *et al.* detallan las etapas cruciales para cargar ARN en exosomas. Finalmente, El-Andaloussi *et al.* resumen la manera de usar exosomas para suministrar de forma eficaz ARN *in vitro* e *in vivo* en cerebro de ratón. También se proporcionan ejemplos de resultados previstos en que se evalúa el suministro de ARN mediado por exosomas mediante ensayos funcionales e imágenes. El protocolo completo dura ~3 semanas. El suministro o administración de acuerdo con la invención puede realizarse usando exosomas producidos a partir de células dendríticas autoderivadas. A partir de los contenidos de este documento, esto puede emplearse en la práctica de la invención.

15 En otra realización, se contemplan los exosomas plasmáticos de Wahlgren *et al.*, (Nucleic Acids Research, 2012, Vol. 40, n.º 17 e130). Los exosomas son vesículas de tamaño nanométrico (30-90 nm de tamaño) producidas por muchos tipos celulares, incluyendo células dendríticas (DC), linfocitos B, linfocitos T, mastocitos, células epiteliales y células tumorales. Estas vesículas se forman por gemación hacia el interior de endosomas tardíos y después se liberan al entorno extracelular tras su fusión con la membrana plasmática. Como los exosomas portan de forma natural ARN entre las células, esta propiedad puede ser útil en genoterapia, y a partir de esta divulgación pueden emplearse en la práctica de la presente invención.

20 Pueden prepararse exosomas a partir de plasma por centrifugación de capa leucocitaria a 900 g durante 20 min para aislar el plasma seguido de recogida de los sobrenadantes celulares, centrifugación a 300 g durante 10 min para eliminar las células y a 16 500 g durante 30 min seguido de filtración a través de un filtro de 0,22 mm. Los exosomas se sedimentan por ultracentrifugación a 120 000 g durante 70 min. Se realiza transfección química de ARNip en los exosomas de acuerdo con las instrucciones del fabricante en el RNAi Human/Mouse Starter Kit (Quiagen, Hilden, Alemania). El ARNip se añade a 100 ml de PBS a una concentración final de 2 mmol/ml. Después de añadir reactivo de transfección HiPerFect, la mezcla se incuba durante 10 min. a TA. Para eliminar el exceso de micelas, los exosomas se vuelven a aislar usando microesferas de látex con aldehído/sulfato. La transfección química de Cas de CRISPR en los exosomas puede realizarse de forma similar al ARNip. Los exosomas pueden cocultivarse con monocitos y linfocitos aislados de la sangre periférica de donadores sanos. Por lo tanto, puede contemplarse que los exosomas que contienen Cas de CRISPR pueden introducirse en monocitos y linfocitos de y reintroducirse de forma análoga en un ser humano. Por consiguiente, el suministro o administración de acuerdo con la invención puede realizarse usando exosomas plasmáticos.

Liposomas

40 El suministro o la administración de acuerdo con la invención puede realizarse con liposomas. Los liposomas son estructuras vesiculares esféricas compuestas de una bicapa lipídica uni o multilaminar que rodea compartimentos acuosos internos y una bicapa fosfolipídica lipófila exterior relativamente impermeable. Los liposomas han obtenido una considerable atención como vehículos de suministro de fármacos porque son biocompatibles, atóxicos, pueden suministrar tanto moléculas de fármaco hidrófilas como lipófilas, protegen su carga de la degradación por enzimas plasmáticas y transportan su carga a través de membranas biológicas y la barrera hematoencefálica (BBB) (véase, por ejemplo, Spuch y Navarro, Journal of Drug Delivery, vol. 2011, ID del artículo 469679, 12 páginas, 2011. doi:10.1155/2011/469679 para revisión).

50 Los liposomas pueden prepararse a partir de varios tipos diferentes de lípidos; sin embargo, los fosfolípidos son los más habitualmente usados para generar liposomas como vehículos de fármaco. Aunque la formación de liposomas es espontánea cuando se mezcla una película de lípidos con una solución acuosa, también puede acelerarse aplicando fuerza en forma de agitación usando un homogeneizador, sonicador o un aparato de extrusión (véase, por ejemplo, Spuch y Navarro, Journal of Drug Delivery, vol. 2011, ID del artículo 469679, 12 páginas, 2011. doi:10.1155/2011/469679 para revisión).

55 Pueden añadirse otros varios aditivos a los liposomas para modificar su estructura y propiedades. Por ejemplo, puede añadirse colesterol o esfingomielina a la mezcla liposómica para ayudar a estabilizar la estructura liposómica y para evitar la filtración de la carga interna liposómica. Además, se preparan liposomas a partir de fosfatidilcolina de huevo hidrogenado o fosfatidilcolina de huevo, colesterol y fosfato de dicetilo, y sus tamaños medios de vesícula se ajustaron hasta aproximadamente 50 y 100 nm. (véase, por ejemplo, Spuch y Navarro, Journal of Drug Delivery, vol. 2011, ID del artículo 469679, 12 páginas, 2011. doi:10.1155/2011/469679 para revisión).

60 Una formulación de liposomas puede estar compuesta principalmente de fosfolípidos naturales y lípidos tales como 1,2-diestearoil-sn-glicero-3-fosfatidilcolina (DSPC), esfingomielina, fosfatidilcolinas de huevo y monosialogangliósido. Como esta formulación está compuesta de fosfolípidos únicamente, las formulaciones liposómicas han encontrado muchos problemas, siendo uno de ellos la inestabilidad en plasma. Se han hecho varios intentos por superar estos

problemas, específicamente en la manipulación de la membrana lipídica. Uno de estos intentos se centró en la manipulación del colesterol. La adición de colesterol a formulaciones convencionales reduce la rápida liberación del compuesto bioactivo encapsulado al plasma o la 1,2-dioleoil-sn-glicero-3-fosfoetanolamina (DOPE) aumenta la estabilidad (véase, por ejemplo, Spuch y Navarro, *Journal of Drug Delivery*, vol. 2011, ID del artículo 469679, 12 páginas, 2011. doi:10.1155/2011/469679 para revisión).

En una realización particularmente ventajosa, los liposomas de tipo Caballo de Troya (también denominados Caballos de Troya moleculares) son deseables y pueden encontrarse protocolos en <http://cshprotocols.cshlp.org/content/2010/4/pdb.prot5407.long>. Estas partículas permiten el suministro de un transgén al cerebro completo después de una inyección intravascular. Sin limitarse a teoría alguna, se cree que partículas lipídicas neutras con anticuerpos específicos conjugados en la superficie permiten cruzar la barrera hematoencefálica mediante endocitosis. El solicitante postula la utilización de liposomas de tipo Caballo de Troya para suministrar la familia de nucleasas de CRISPR al cerebro mediante una inyección intravascular, lo que permitiría animales transgénicos en todo el cerebro sin la necesidad de manipulación embrionaria. Pueden contemplarse aproximadamente 1-5 g de ADN o ARN para administración *in vivo* en liposomas.

En otra realización, el sistema de Cas de CRISPR puede administrarse en liposomas, tal como una partícula estable de ácido nucleico-lípido (SNALP) (véase, por ejemplo, Morrissey *et al.*, *Nature Biotechnology*, Vol. 23, n.º 8, agosto de 2005). Se contemplan inyecciones intravenosas diarias de aproximadamente 1, 3 o 5 mg/kg/día de una Cas de CRISPR específica dirigida a una SNALP. El tratamiento diario puede ser durante aproximadamente tres días y después semanalmente durante aproximadamente cinco semanas. En otra realización, también se contempla una Cas de CRISPR específica encapsulada en SNALP administrada por inyección intravenosa a dosis de aproximadamente 1 o 2,5 mg/kg (véase, por ejemplo, Zimmerman *et al.*, *Nature Letters*, Vol. 441, 4 de mayo de 2006). La formulación de SNALP puede contener los lípidos 3-N-[(w-metoxipoli(etilenglicol) 2000) carbamoil]-1,2-dimiristoiloxi-propilamina (PEG-C-DMA), 1,2-dilinoileiloxi-N,N-dimetil-3-aminopropano (DLinDMA), 1,2-diestearoil-sn-glicero-3-fosfocolina (DSPC) y colesterol, en una relación molar porcentual de 2:40:10:48 (véase, por ejemplo, Zimmerman *et al.*, *Nature Letters*, Vol. 441, 4 de mayo de 2006).

En otra realización, se ha demostrado que las partículas estables de ácido nucleico-lípido (SNALP) son moléculas de suministro eficaces a tumores hepáticos derivados de HepG2 muy vascularizados, pero no en tumores hepáticos derivados de HCT-116 poco vascularizados (véase, por ejemplo, Li, *Gene Therapy* (2012) 19, 775-780). Pueden prepararse liposomas de SNALP formulando D-Lin-DMA y PEG-C-DMA con diestearoilfosfatidilcolina (DSPC), colesterol y ARNip usando una relación de lípido/ARNip de 25:1 y una relación molar de 48/40/10/2 de colesterol/D-Lin-DMA/DSPC/PEG-C-DMA. Los liposomas de SNALP resultantes son de aproximadamente 80-100 nm de tamaño.

En otra realización más, un SNALP puede comprender colesterol sintético (Sigma-Aldrich, St Louis, MO, EE.UU.), dipalmitoilfosfatidilcolina (Avanti Polar Lipids, Alabaster, AL, EE.UU.), 3-N-[(w-metoxipoli(etilenglicol)2000)carbamoil]-1,2-dimiristoiloxipropilamina, y 1,2-dilinoileiloxi-3-N,N-dimetilaminopropano catiónico (véase, por ejemplo, Geisbert *et al.*, *Lancet* 2010; 375: 1896-905). Puede contemplarse una dosificación de aproximadamente 2 mg/kg de Cas de CRISPR total por dosis administrada como, por ejemplo, una infusión intravenosa rápida.

En otra realización más, una SNALP puede comprender colesterol sintético (Sigma-Aldrich), 1,2-diestearoil-sn-glicero-3-fosfocolina (DSPC; Avanti Polar Lipids Inc.), PEG-cDMA, y 1,2-dilinoileiloxi-3-(N,N-dimetil)aminopropano (DLinDMA) (véase, por ejemplo, Judge, *J. Clin. Invest.* 119:661-673 (2009)). Las formulaciones usadas para estudios *in vivo* pueden comprender una relación final en masa de lípido/ARN de aproximadamente 9:1,

El perfil de seguridad de nanomedicamentos de iARN se ha revisado por Barros y Gollob de Alnylam Pharmaceuticals (véase, por ejemplo, *Advanced Drug Delivery Reviews* 64 (2012) 1730-1737). La partícula estable de ácido nucleico y lípido (SNALP) está compuesta de cuatro lípidos diferentes, un lípido ionizable (DLinDMA) que es catiónico a pH bajo, un lípido auxiliar neutro, colesterol y un lípido con polietilenglicol (PEG) difundible. La partícula es de aproximadamente 80 nm de diámetro y es de carga neutra a pH fisiológico. Durante la formulación, el lípido ionizable sirve para condensar el lípido con el ARN aniónico durante la formación de las partículas. Cuando se carga positivamente en condiciones endosómicas cada vez más ácidas, el lípido ionizable también media la fusión de SNALP con la membrana endosómica lo que posibilita la liberación del ARN en el citoplasma. El PEG-lípido estabiliza la partícula y reduce la agregación durante la formulación, y posteriormente proporciona un exterior hidrófilo neutro que mejora las propiedades farmacocinéticas.

Hasta la fecha, se han iniciado dos programas clínicos usando formulaciones de SNALP con ARN. Tekmira Pharmaceuticals completó recientemente un estudio de una sola dosis en fase I de SNALP-ApoB en voluntarios adultos con elevado colesterol LDL. ApoB se expresa predominantemente en el hígado y el yeyuno y es esencial para el ensamblaje y la secreción de VLDL y LDL. Diecisiete sujetos recibieron una sola dosis de SNALP-ApoB (aumento de dosis en 7 niveles de dosis). No hubo evidencias de hepatotoxicidad (prevista por la potencial toxicidad limitante de la dosis basada en estudios preclínicos). Un sujeto (de dos) a la dosis máxima experimentó síntomas de tipo gripo coherentes con la estimulación del sistema inmunitario, y se tomó la decisión de concluir el ensayo.

Alnylam Pharmaceuticals ha hecho avanzar asimismo el ALN-TTR01, que emplea la tecnología de SNALP descrita anteriormente y aborda la producción en hepatocitos de TTR tanto mutante como de tipo silvestre para tratar la amiloidosis TTR (ATTR). Se han descrito tres síndromes ATTR: polineuropatía amiloidótica familiar (FAP) y cardiomiopatía amiloidótica familiar (FAC), ambas causadas por mutaciones dominantes autosómicas en TTR; y amiloidosis generalizada senil (SSA) causada por TTR de tipo silvestre. Recientemente se completó un ensayo en fase I de aumento de una sola dosis controlado por placebo de ALN-TTR01 en pacientes con ATTR. ALN-TTR01 se administró como una infusión IV de 15 minutos a 31 pacientes (23 con fármaco del estudio y 8 con placebo) dentro de un intervalo de dosis de 0,01 a 1,0 mg/kg (basada en ARNip). El tratamiento se toleró bien sin aumentos significativos en los ensayos de función hepática. Se anotaron reacciones relacionadas con la infusión en 3 de 23 pacientes a > 0,4 mg/kg; todos respondieron a la ralentización de la tasa de infusión y todos continuaron en el estudio. Se apreciaron elevaciones mínimas y transitorias de las citocinas séricas IL-6, IP-10 y IL-1ra en dos pacientes a la dosis máxima de 1 mg/kg (como se esperaba de los estudios preclínicos y NHP). Se observó reducción de TTR en suero, el efecto farmacodinámico esperado de ALN-TTR01 a 1 mg/kg.

En otra realización más, un SNALP puede prepararse solubilizando un lípido catiónico, DSPC, colesterol y PEG-lípido, por ejemplo, en etanol, por ejemplo, a una relación molar de 40:10:40:10, respectivamente (véase, Semple *et al.*, Nature Nanotechnology, Volumen 28 número 2 febrero de 2010, pág. 172-177). La mezcla de lípidos se añadió a un tampón acuoso (citrato 50 mM, pH 4) con mezcla a una concentración final de etanol y lípidos de un 30 % (vol/vol) y 6,1 mg/ml, respectivamente, y se dejó equilibrar a 22 °C durante 2 min. antes de la extrusión. Los lípidos hidratados se extruyeron a través de dos filtros con tamaño de poro de 80 nm a pilados (Nuclepore) a 22 °C usando una extrusora Lipex (Northern Lipids) hasta que se obtuvo un diámetro de vesícula de 70-90 nm, determinado por análisis de dispersión de luz dinámica. Esto en general requirió 1-3 pases. El ARNip (solubilizado en una solución acuosa de citrato 50 mM, pH 4 que contenía etanol al 30 %) se añadió a las vesículas preequilibradas (35 °C) a una tasa de ~5 ml/min con mezcla. Después de que se alcanzara una relación final deseada de ARNip/lípido de 0,06 (p/p), la mezcla se incubó durante 30 min. más a 35 °C para permitir la reorganización de las vesículas y la encapsulación del ARNip. Entonces se retiró el etanol y el tampón externo se reemplazó con PBS (NaCl 155 mM, Na₂HPO₄ 3 mM, KH₂PO₄ 1 mM, pH 7,5) por diálisis o diafiltración en flujo tangencial. El ARNip se encapsuló en SNALP usando un proceso de dilución por etapas controlado. Los constituyentes lipídicos de KC2-SNALP eran DLin-KC2-DMA (lípido catiónico), dipalmitoilfosfatidilcolina (DPPC; Avanti Polar Lipids), colesterol sintético (Sigma) y PEG-C-DMA usados a una relación molar de 57,1:7,1:34,3:1,4. Tras la formación de las partículas cargadas, las SNALP se dializaron frente a PBS y se esterilizaron en filtro a través de un filtro de 0,2 µm antes de su uso. Los tamaños medios de las partículas fueron 75-85 nm y un 90-95 % del ARNip se encapsuló dentro de las partículas lipídicas. La relación final de ARNip/lípido en las formulaciones usadas para ensayo *in vivo* fue ~0,15 (p/p). Los sistemas de LNP-ARNip que contenían ARNip de factor VII se diluyeron a las concentraciones apropiadas en PBS estéril inmediatamente antes de su uso y las formulaciones se administraron por vía intravenosa a través de la vena lateral de la cola en un volumen total de 10 ml/kg. Este método y estos sistemas de suministro pueden extrapolarse al sistema de Cas de CRISPR de la presente invención.

Otros lípidos

Pueden utilizarse otros lípidos catiónicos, tales como aminolípido 2,2-dilinoil-4-dimetilaminoetil-[1,3]-dioxolano (DLin-KC2-DMA) para encapsular Cas de CRISPR o componentes de la misma o una o más moléculas de ácido nucleico que codifican la misma, por ejemplo, similar a ARNip (véase, por ejemplo, Jayaraman, Angew. Chem. Int. Ed. 2012, 51, 8529 -8533), y, por tanto, pueden emplearse en la práctica de la invención. Puede contemplarse una vesícula preformada con la siguiente composición lipídica: aminolípido, diestearoilfosfatidilcolina (DSPC), colesterol y (R)-2,3-bis(octadeciloxi) propil-1-(metoxi poli(etilenglicol)2000)propilcarbamatato (PEG-lípido) en la relación molar 40/10/40/10 respectivamente, y una relación de ARNip de FVII/lípidos totales de aproximadamente 0,05 (p/p). Para garantizar una estrecha distribución de los tamaños de partícula en el intervalo de 70-90 nm y un bajo índice de polidispersidad de 0,11 + 0,04 (n=56), las partículas pueden extraerse hasta tres veces a través de membranas de 80 nm antes de añadir el ARN de Cas de CRISPR. Pueden usarse partículas que contienen el aminolípido 16 altamente potente, en que la relación molar de los cuatro componentes lipídicos 16, DSPC, colesterol y PEG-lípido (50/10/38,5/1,5) que pueden optimizarse adicionalmente para potenciar la actividad *in vivo*.

Michael S D Kormann *et al.* ("Expression of therapeutic proteins after delivery of chemically modified mRNA in mice: Nature Biotechnology, Volumen:29, páginas: 154-157 (2011)) describen el uso de envueltas lipídicas para suministrar ARN. El uso de envueltas lipídicas también se prefiere en la presente invención.

En otra realización, los lípidos pueden formularse con el sistema de Cas de CRISPR de la presente invención para formar nanopartículas lipídicas (LNP). Los lípidos incluyen, aunque sin limitación, DLin-KC2-DMA4, C12-200 y pueden formularse colípidos diestearoilfosfatidilcolina, colesterol, y PEG-DMG con Cas de CRISPR en lugar de ARNip (véase, por ejemplo, Novobrantseva, Molecular Therapy-Nucleic Acids (2012) 1, e4; doi:10.1038/mtna.2011.3) usando un procedimiento de formación espontánea de vesículas. La relación molar de los componentes puede ser de aproximadamente 50/10/38,5/1,5 (DLin-KC2-DMA o C12-200/diestearoilfosfatidilcolina/colesterol/PEG-DMG). La relación ponderal final de lípido:ARNip puede ser ~12:1 y 9:1 en el caso de nanopartículas lipídicas (LNP) de DLin-KC2-DMA y C12-200, respectivamente. Las formulaciones pueden tener diámetros medios de partícula de ~80 nm con > 90% de eficiencia de atrapamiento. Puede contemplarse una dosis de 3 mg/kg.

Tekmira tiene una cartera de aproximadamente 95 familias de patentes, en Estados Unidos y en el extranjero, que se refieren a diversos aspectos de LNP y formulaciones de LNP (véanse, por ejemplo, las patentes de Estados Unidos n.º 7.982.027; 7.799.565; 8.058.069; 8.283.333; 7.901.708; 7.745.651; 7.803.397; 8.101.741; 8.188.263; 7.915.399; 8.236.943 y 7.838.658 y las patentes europeas n.º 1766035; 1519714; 1781593 y 1664316), que todas ellas pueden usarse y/o adaptarse a la presente invención.

El sistema de Cas de CRISPR o componentes del mismo o una o más moléculas de ácido nucleico que codifican el mismo pueden suministrarse encapsulados en microesferas de PLGA tal como se describe adicionalmente en las solicitudes publicadas de Estados Unidos 20130252281 y 20130245107 y 20130244279 (asignadas a Moderna Therapeutics) que se refieren a aspectos de formulación de composiciones que comprenden moléculas de ácido nucleico modificadas que pueden codificar una proteína, un precursor de proteína o una forma parcial o completamente procesada de la proteína o un precursor de la proteína. La formulación puede tener una relación molar de 50:10:38,5:1,5-3,0 (lípidos catiónicos:lípido fusógeno:colesterol:lípido con PEG). El lípido con PEG puede seleccionarse de, aunque sin limitación, PEG-c-DOMG, PEG-DMG. El lípido fusógeno puede ser DSPC. Véase también, Schrum *et al.*, Delivery y Formulation of Engineered Nucleic Acids, la solicitud publicada de Estados Unidos 20120251618,

La tecnología de Nanomeric aborda los problemas de biodisponibilidad para una amplia gama de agentes terapéuticos, incluyendo fármacos hidrófobos de bajo peso molecular, péptidos y agentes terapéuticos basados en ácido nucleico (plásmidos, ARNip, miARN). Las vías de administración específicas para las que la tecnología ha demostrado ventajas claras incluyen vía oral, transporte a través de la barrera hematoencefálica, suministro a tumores sólidos, así como al ojo. Véanse, por ejemplo, Mazza *et al.*, 2013, ACS Nano. 26 de febrero de 2013;7(2):1016-26; Uchegbu y Siew, 2013, J Pharm Sci. 102(2):305-10 y Lalatsa *et al.*, 2012, J Control Release. 20 de julio de 2012; 161(2):523-36.

La publicación de patente de Estados Unidos n.º 20050019923 describe dendrímeros catiónicos para el suministro de moléculas bioactivas, tales como moléculas polinucleotídicas, péptidos y polipéptidos y/o agentes farmacéuticos a un cuerpo de mamífero. Los dendrímeros son adecuados para dirigir el suministro de las moléculas bioactivas a, por ejemplo, el hígado, bazo, pulmón, riñón o corazón (o incluso al cerebro). Los dendrímeros son macromoléculas tridimensionales sintéticas que se preparan por etapas a partir de unidades monoméricas ramificadas simples, cuya naturaleza y funcionalidad pueden controlarse fácilmente y variarse. Los dendrímeros se sintetizan a partir de la adición repetida de elementos básicos a un núcleo multifuncional (estrategia divergente para la síntesis) o hacia un núcleo multifuncional (estrategia convergente para la síntesis) y cada adición de una cubierta tridimensional de elementos básicos da lugar a la formación de una generación mayor de los dendrímeros. Los dendrímeros de polipropilenimina empiezan a partir de un núcleo de diaminobutano al que se añade dos veces el número de grupos amino mediante una adición doble de Michael de acrilonitrilo a las aminas primarias seguida de la hidrogenación de los nitrilos. Esto provoca una duplicación de los grupos amino. Los dendrímeros de polipropilenimina contienen un 100 % de nitrógenos protonables y hasta 64 grupos amino terminales (generación 5, DAB 64). Los grupos protonables son habitualmente grupos amina que pueden aceptar protones a pH neutro. El uso de dendrímeros como agentes de suministro génico se ha centrado en gran medida en el uso de los compuestos que contienen poliamidoamina y fósforo con una mezcla de amina/amida o N--P(O₂)S como unidades de conjugación respectivamente sin haberse presentado ningún trabajo sobre el uso de los dendrímeros de polipropilenimina de menor generación para suministro génico. Los dendrímeros de polipropilenimina también se han estudiado como sistemas de liberación controlada sensibles al pH para el suministro de fármacos y para la encapsulación de moléculas hospedadas cuando se modifican químicamente por grupos aminoácidos periféricos. La citotoxicidad e interacción de los dendrímeros de polipropilenimina con ADN, así como la eficacia de transfección de DAB 64 también se han estudiado.

La publicación de patente de Estados Unidos n.º 20050019923 se basa en la observación de que, en contra de los informes previos, los dendrímeros catiónicos, tales como dendrímeros de polipropilenimina, presentan propiedades adecuadas, tales como dirección específica y baja toxicidad, para su uso en el suministro dirigido de moléculas bioactivas, tal como material genético. Además, los derivados del dendrímero catiónico también presentan propiedades adecuadas para el suministro dirigido de moléculas bioactivas. Véase también, Bioactive Polymers, solicitud publicada de Estados Unidos 20080267903, que divulga "Diversos polímeros, incluyendo polímeros de poliamina catiónicos y polímeros dendriméricos, se demuestra que poseen actividad antiproliferativa y, por lo tanto, pueden ser útiles para el tratamiento de trastornos caracterizados por proliferación celular indeseable tales como neoplasias y tumores, trastornos inflamatorios (incluyendo trastornos autoinmunitarios), psoriasis y aterosclerosis. Los polímeros pueden usarse en solitario como agentes activos o como vehículos de suministro para otros agentes terapéuticos, tales como moléculas de fármaco o ácidos nucleicos para genoterapia. En dichos casos, la propia actividad antitumoral intrínseca de los polímeros puede complementar la actividad del agente a suministrar". Las divulgaciones de estas publicaciones de patente pueden emplearse junto con los contenidos de este documento para el suministro de uno o más sistemas de Cas de CRISPR o uno o más componentes de los mismos o una o más moléculas de ácido nucleico que codifican los mismos.

Proteínas supercargadas

Las proteínas supercargadas son una clase de proteínas manipuladas o de origen natural con carga teórica neta positiva o negativa infrecuentemente alta y pueden emplearse en el suministro de uno o más sistemas de Cas de

CRISPR o uno o más componentes de los mismos o una o más moléculas de ácido nucleico que codifican los mismos. Tanto las proteínas supercargadas negativamente como supercargadas positivamente muestran una capacidad notable de resistir la agregación inducida térmica o químicamente. Las proteínas supercargadas positivamente también pueden penetrar las células de mamífero. Asociar una carga con estas proteínas, tal como ADN plasmídico, ARN u otras proteínas, puede posibilitar el suministro funcional de estas macromoléculas en células de mamífero tanto *in vitro* como *in vivo*. El laboratorio de David Liu informó de la creación y caracterización de proteínas supercargadas en 2007 (Lawrence *et al.*, 2007, Journal of the American Chemical Society 129, 10110-10112).

El suministro no vírico de ARN y ADN plasmídico en células de mamífero es valioso tanto para aplicaciones de investigación como terapéuticas (Akinc *et al.*, 2010, Nat. Biotech. 26, 561-569). La proteína GFP +36 (u otra proteína supercargada positivamente) se mezcla con ARN en el medio sin suero apropiado y se deja que forme complejos antes de la adición a las células. La inclusión de suero en esta fase inhibe la formación de los complejos de proteína supercargada-ARN y reduce la eficacia del tratamiento. Se ha descubierto que el siguiente protocolo es eficaz para una diversidad de líneas celulares (McNaughton *et al.*, 2009, Proc. Natl. Acad. Sci. USA 106, 6111-6116). Sin embargo, deben realizarse experimentos piloto variando la dosis de proteína y ARN para optimizar el procedimiento para líneas celulares específicas.

(1) Un día antes del tratamiento, se siembran 1×10^5 células por pocillo en una placa de 48 pocillos.

(2) En el día del tratamiento, se diluye la proteína GFP +36 purificada en medio sin suero hasta una concentración final de 200 nM. Se añade ARN a una concentración final de 50 nM. Se agita con vórtice para mezclar y se incuba a temperatura ambiente durante 10 min.

(3) Durante la incubación, se aspira el medio de las células y se lavan una vez con PBS.

(4) Después de la incubación de GFP +36 y ARN, se añaden los complejos de proteína-ARN a las células.

(5) Se incuban las células con los complejos a 37 °C durante 4 h.

(6) Después de la incubación, se aspira el medio y se lava tres veces con 20 U/ml de PBS con heparina. Se incuban las células con medio que contiene suero durante 48 h más o más tiempo dependiendo del ensayo para la actividad.

(7) Se analizan las células por inmunotransferencia, qPCR, ensayo fenotípico u otro método apropiado.

El laboratorio de David Liu ha descubierto además que GFP +36 es un reactivo de suministro plasmídico eficaz en una serie de células. Como el ADN plasmídico es una carga más grande que el ARNip, se requiere proporcionalmente más proteína GFP +36 para formar complejos de forma eficaz con los plásmidos. Para un suministro de plásmidos eficaz los solicitantes han desarrollado una variante de GFP +36 que alberga una marca peptídica HA2 C terminal, un péptido conocido de alteración de endosomas derivado de la proteína hemaglutinina del virus de la gripe. El siguiente protocolo ha sido eficaz en una diversidad de células, pero como anteriormente, se advierte que las dosis de ADN plasmídico y proteína supercargada tienen que optimizarse para líneas celulares específicas y aplicaciones de suministro.

(1) Un día antes del tratamiento, se siembran 1×10^5 por pocillo en una placa de 48 pocillos.

(2) En el día del tratamiento, se diluye proteína GFP +36 purificada en medio sin suero hasta una concentración final de 2 mM. Se añade 1 mg de ADN plasmídico. Se agita con vórtice para mezclar y se incuba a temperatura ambiente durante 10 min.

(3) Durante la incubación, se aspira el medio de las células y se lavan una vez con PBS.

(4) Después de la incubación de GFP +36 y ADN plasmídico, se añaden suavemente los complejos de proteína-ADN a las células.

(5) Se incuban las células con los complejos a 37 °C durante 4 h.

(6) Después de la incubación, se aspira el medio y se lava PBS. Se incuban las células en medio que contiene suero y se incuban durante 24-48 h más.

(7) Se analiza el suministro de plásmido (por ejemplo, por expresión génica dirigida por plásmido) según lo apropiado.

Véanse también, por ejemplo, McNaughton *et al.*, Proc. Natl. Acad. Sci. USA 106, 6111-6116 (2009); Cronican *et al.*, ACS Chemical Biology 5, 747-752 (2010); Cronican *et al.*, Chemistry & Biology 18, 833-838 (2011); Thompson *et al.*, Methods in Enzymology 503, 293-319 (2012); Thompson, D.B., *et al.*, Chemistry & Biology 19 (7), 831-843 (2012). Los métodos de las proteínas supercargadas pueden usarse y/o adaptarse para suministro del sistema de Cas de CRISPR Cas de la presente invención. Estos sistemas del Dr. Lui y los documentos de este documento junto con los contenidos

de este documento pueden emplearse en el suministro de uno o más sistemas de Cas de CRISPR o uno o más componentes de los mismos o una o más moléculas de ácido nucleico que codifican los mismos.

Dispositivos implantables

En otra realización, también se contemplan dispositivos implantables para el suministro del sistema de Cas de CRISPR o uno o más componentes de los mismos o una o más moléculas de ácido nucleico que codifican los mismos. Por ejemplo, la publicación de patente de Estados Unidos 20110195123 divulga un dispositivo médico implantable que eluye un fármaco localmente y se proporciona en un periodo prolongado, incluyendo varios tipos de dicho dispositivo, los modos de implementación del tratamiento y métodos de implante. El dispositivo comprende sustrato polimérico, tal como una matriz, por ejemplo, que se usa como cuerpo del dispositivo, y fármacos, y en algunos casos materiales estructurales adicionales, tales como metales o polímeros adicionales, y materiales para potenciar la visibilidad y la formación de imágenes. Un dispositivo de suministro implantable puede ser ventajoso para proporcionar liberación localmente y durante un periodo prolongado, donde el fármaco se libera directamente a la matriz extracelular (ECM) de la zona enferma tal como tumor, inflamación, degeneración o para objetivos sintomáticos, o a células de músculo liso lesionadas, o para prevención. Un tipo de fármaco es ARN, como se divulga anteriormente, y este sistema puede usarse y/o adaptarse al sistema de Cas de CRISPR de la presente invención. Los modos de implante en algunas realizaciones son procedimientos de implante existentes que se desarrollan y usan actualmente para otros tratamientos, incluyendo braquiterapia y biopsia con aguja. En dichos casos, las dimensiones del nuevo implante descrito en esta invención son similares al implante original. Típicamente, se implantan unos pocos dispositivos durante el mismo procedimiento de tratamiento.

Como se describe en la publicación de patente de Estados Unidos 20110195123, se proporciona un sistema implantable o insertable de suministro de fármacos, incluyendo sistemas aplicables a una cavidad tal como la cavidad abdominal y/o cualquier otro tipo de administración en que el sistema de suministro fármacos no se ancle o fije, que comprende un sustrato polimérico bioestable y/o degradable y/o bioabsorbible, que puede ser, por ejemplo, opcionalmente una matriz. Debe apreciarse que el término "inserción" también incluye implante. El sistema de suministro de fármacos se implementa preferiblemente como un "Loder" como se describe en la publicación de patente de Estados Unidos 20110195123.

El polímero o pluralidad de polímeros son biocompatibles, incorporando un agente y/o pluralidad de agentes, lo que posibilita la liberación de agente a una tasa controlada, en el que el volumen total del sustrato polimérico, tal como una matriz, por ejemplo, en algunas realizaciones es opcional y preferiblemente no mayor de un volumen máximo que permite alcanzar un nivel terapéutico del agente. Como un ejemplo no limitante, dicho volumen está preferiblemente dentro del intervalo de 0,1 m³ a 1000 mm³, según se requiera por el volumen para la carga de agente. El Loder puede ser opcionalmente más grande, por ejemplo, cuando se le incorpora un dispositivo cuyo tamaño se determina por la funcionalidad, por ejemplo, y sin limitación, una articulación de rodilla, un anillo intrauterino o cervicouterino y similares.

El sistema de suministro de fármacos (para suministrar la composición) se diseña en algunas realizaciones para emplear preferiblemente polímeros degradables, en las que el mecanismo de liberación principal es erosión general; o en algunas realizaciones se usan polímeros no degradables, o polímeros que se degradan lentamente, en las que el mecanismo de liberación principal es difusión en lugar de erosión general, de modo que la parte exterior funciona como membrana, y su parte interna funciona como depósito de fármaco, que prácticamente no se ve afectado por los alrededores durante un periodo prolongado (por ejemplo, de aproximadamente una semana a aproximadamente unos pocos meses). También pueden usarse opcionalmente combinaciones de diferentes polímeros con diferentes mecanismos de liberación. El gradiente de concentración en la superficie preferiblemente se mantiene constante de manera eficaz durante un periodo significativo del periodo total de liberación del fármaco y, por lo tanto, la tasa de difusión es constante de forma eficaz (denominada difusión "en modo cero"). Por el término "constante" se entiende una tasa de difusión que se mantiene preferiblemente por encima del umbral inferior de eficacia terapéutica, pero que aún puede caracterizar opcionalmente un estallido inicial y/o puede fluctuar, por ejemplo, aumentando o disminuyendo hasta un determinado grado. La tasa de difusión preferiblemente se mantiene así durante un periodo prolongado, y puede considerarse constante hasta un determinado nivel para optimizar el periodo terapéuticamente eficaz, por ejemplo, el periodo de silenciamiento eficaz.

El sistema de suministro de fármacos opcionalmente y preferiblemente se diseña para proteger el agente terapéutico basado en nucleótido de la degradación, ya sea de naturaleza química o debida al ataque de enzimas y otros factores en el cuerpo del sujeto.

El sistema de suministro de fármacos como se describe en la publicación de patente de Estados Unidos 20110195123 se asocia opcionalmente con aparatos de detección y/o activación que se hacen funcionar en y/o después del implante del dispositivo, mediante métodos no invasivos y/o mínimamente invasivos de activación y/o aceleración/desaceleración, por ejemplo, incluyendo opcionalmente, aunque sin limitación, calentamiento y enfriamiento térmico, rayos láser y ultrasonidos, incluyendo métodos o dispositivos de ultrasonidos enfocados y/o RF (radiofrecuencia).

De acuerdo con algunas realizaciones de la publicación de patente de Estados Unidos 20110195123, el sitio para suministro local puede incluir opcionalmente sitios diana caracterizados por alta proliferación anómala de las células, y apoptosis suprimida, incluyendo tumores, inflamación activa y/o crónica e infección incluyendo enfermedades autoinmunitarias, degeneración de tejidos incluyendo tejido muscular y nervioso, dolor crónico, sitios degenerativos y ubicación de fracturas de hueso y otras ubicaciones de heridas para la potenciación de la regeneración de tejido, y músculo cardíaco, liso o estriado lesionado.

El sitio para el implante de la composición, o sitio diana, preferiblemente caracteriza un radio, área y/o volumen que es suficientemente pequeño para el suministro local dirigido. Por ejemplo, el sitio diana opcionalmente tiene un diámetro en un intervalo de aproximadamente 0,1 mm a aproximadamente 5 cm.

La ubicación del sitio diana se selecciona preferiblemente para una eficacia terapéutica máxima. Por ejemplo, la composición del sistema de suministro de fármacos (opcionalmente con un dispositivo para su implante como se describe anteriormente) se implanta opcionalmente y preferiblemente dentro de o en la proximidad de un entorno tumoral, o el aporte sanguíneo asociado del mismo.

Por ejemplo, la composición (opcionalmente con el dispositivo) se implanta opcionalmente dentro de o en la proximidad del páncreas, próstata, mama, hígado, a través del pezón, dentro del sistema vascular y así sucesivamente.

La ubicación diana se selecciona opcionalmente del grupo que consiste en (como ejemplos no limitantes únicamente, ya que opcionalmente cualquier sitio dentro del cuerpo puede ser adecuado para el implante de un Loder): 1. cerebro en sitios degenerativos como en enfermedad de Parkinson o de Alzheimer en los ganglios basales, materia blanca y gris; 2. médula como en el caso de esclerosis lateral amiotrófica (ALS); 3. cuello uterino para evitar infección por HPV; 4. articulaciones inflamatorias activas y crónicas; 5. dermis como en el caso de psoriasis; 6. sitios de nervios simpáticos y sensitivos para efecto analgésico; 7. implante intraóseo; 8. sitios de infección aguda y crónica; 9. intravaginal; 10. sistema auditivo interno, laberinto del oído interno, sistema vestibular; 11. intratraqueal; 12. intracardiaco; coronario, epicardiaco; 13. vejiga urinaria; 14. sistema biliar; 15. tejido parenquimatoso incluyendo y sin limitación, el riñón, hígado, bazo; 16. ganglios linfáticos; 17. glándulas salivales; 18. encías dentales; 19. intraarticular (en las articulaciones); 20. intraocular; 21. tejido cerebral; 22. ventrículos cerebrales; 23. cavidades, incluyendo cavidad abdominal (por ejemplo, aunque sin limitación, para cáncer de ovario); 24. intraesofágico y 25. intrarrectal.

Opcionalmente, la inserción del sistema (por ejemplo, un dispositivo que contiene la composición) se asocia con inyección de material a la ECM en el sitio diana y las cercanías de ese sitio para influir en el pH local y/o la temperatura y/u otros factores biológicos que influyen en la difusión del fármaco y/o la cinética del fármaco en la ECM, del sitio diana y las cercanías de dicho sitio.

Opcionalmente, de acuerdo con algunas realizaciones, la liberación de dicho agente podría asociarse con aparatos de detección y/o activación que se hacen funcionar antes y/o en y/o después de la inserción, mediante métodos no invasivos y/o mínimamente invasivos y/o también métodos de activación y/o aceleración/desaceleración, incluyendo rayos láser, radiación, calentamiento y enfriamiento térmico, y ultrasonidos, incluyendo métodos o dispositivos de ultrasonidos enfocados y/o RF (radiofrecuencia), y activadores químicos.

De acuerdo con otras realizaciones de la publicación de patente de Estados Unidos 20110195123, el fármaco comprende preferiblemente un ARN, por ejemplo, para casos de cáncer localizado en mama, páncreas, cerebro, riñón, vejiga, pulmón y próstata como se describe a continuación. Aunque se ejemplifica con iARN, son aplicables muchos fármacos para encapsularse en Loder, y pueden usarse en asociación con esta invención, siempre que dichos fármacos puedan encapsularse con el sustrato Loder, tal como una matriz, por ejemplo, y este sistema puede usarse y/o adaptarse al suministro del sistema de Cas de CRISPR de la presente invención.

Como otro ejemplo de una aplicación específica, las enfermedades neurodegenerativas y degenerativas musculares se desarrollan debido a expresión génica anómala. El suministro local de ARN puede tener propiedades terapéuticas para interferir con dicha expresión génica anómala. El suministro local de fármacos antiapoptóticos, antiinflamatorios y antidegenerativos, incluyendo fármacos pequeños y macromoléculas también puede ser opcionalmente terapéutico. En dichos casos, el Loder se aplica para liberación prolongada a tasa constante y/o a través de un dispositivo especializado que se implanta por separado. Todo esto puede usarse y/o adaptarse al sistema de Cas de CRISPR de la presente invención.

Como otro ejemplo más de una aplicación específica, trastornos psiquiátricos y cognitivos se tratan con modificadores génicos. La atenuación génica es una opción de tratamiento. Los Loder que suministran localmente agentes a sitios del sistema nervioso central son opciones terapéuticas para trastornos psiquiátricos y cognitivo, incluyendo, aunque sin limitación, psicosis, enfermedades bipolares, trastornos neuróticos y enfermedades conductuales. Los Loder también podrían suministrar fármacos localmente, incluyendo fármacos pequeños y macromoléculas tras su implante en sitios específicos del cerebro. Todo esto puede usarse y/o adaptarse al sistema de Cas de CRISPR de la presente invención.

Como otro ejemplo de una aplicación específica, el silenciamiento de los mediadores inmunitarios innatos y/o adaptativos en sitios locales posibilita la prevención del rechazo de trasplante de órganos. El suministro local de ARN y reactivos inmunomoduladores con el Loder implantado en el órgano trasplantado y/o el sitio con implante representa inmunosupresión local repeliendo las células inmunitarias tales como CD8 activadas contra el órgano trasplantado. Todo esto puede usarse y/o adaptarse al sistema de Cas de CRISPR de la presente invención.

Como otro ejemplo de una aplicación específica, los factores de crecimiento vasculares, incluyendo VEGF y angiogenina y otros son esenciales para la neovascularización. El suministro local de los factores, péptidos, peptidomiméticos, o la supresión de sus represores es una modalidad terapéutica importante; el silenciamiento de los represores y el suministro local de los factores, péptidos, macromoléculas y fármacos pequeños que estimulan la angiogénesis con el Loder es terapéutico para enfermedad vascular periférica, generalizada y cardíaca.

El método de inserción, tal como implante, opcionalmente puede usarse ya para otros tipos de implante en tejido y/o para inserciones y/o para muestro de tejidos, opcionalmente sin modificaciones, o como alternativa opcionalmente solo con modificaciones no importantes en dichos métodos. Dichos métodos opcionalmente incluyen, aunque sin limitación, métodos de braquiterapia, biopsia, endoscopia con y/o son ultrasonidos, tal como ERCP, métodos estereotácticos en el tejido cerebral, laparoscopia, incluyendo implante con un laparoscopio en articulaciones, órganos abdominales, la pared de la vejiga y cavidades corporales.

Las dosis en este documento se basan en un individuo promedio de 70 kg. La frecuencia de administración pertenece al ámbito del facultativo médico o veterinario (por ejemplo, médico, veterinario) o científico experto en la materia. A este respecto, se puede aumentar la escala de los experimentos que implican ratones sin ninguna experimentación innecesaria, incluyendo tener en consideración que un ratón promedio es de 20 g. Las dosis pueden basarse en o extrapolarse a un individuo promedio de 70 kg, y pueden ajustarse para pacientes, sujetos, mamíferos de diferente peso y especie. La frecuencia de administración está dentro del ámbito del facultativo médico o veterinario (por ejemplo, médico, veterinario), dependiendo de factores habituales que incluyen la edad, sexo, salud general, otras afecciones del paciente o sujeto y la afección particular o síntomas que se están abordando. Vectores víricos pueden inyectarse en el tejido de interés. Para modificación genómica específica de tipo celular, la expresión de Cas9 mutante puede dirigirse por un promotor específico de tipo celular. Por ejemplo, la expresión específica de hígado podría usar el promotor de albúmina y la expresión específica de neuronas podría usar el promotor de sinapsina I. Las concentraciones óptimas de ARNm de la enzima de CRISPR mutada y ARN guía pueden determinarse ensayando diferentes concentraciones en un modelo celular o animal y usando secuenciación exhaustiva para analizar el grado de modificación en posibles locus genómicos inespecíficos. Por ejemplo, para la secuencia guía dirigida a 5'-GAGTCCGAGCAGAAGAAGAA-3' en el gen EMX1 del genoma humano, puede usarse secuenciación exhaustiva para evaluar el nivel de modificación en los dos siguientes locus inespecíficos, 1: 5'-GAGTCCTAGCAGGAGAAGAA-3' y 2: 5'-GAGTCTAAGCAGAAGAAGAA-3'. La concentración que da el nivel más alto de modificación específica mientras se minimiza el nivel de modificación inespecífica debe elegirse para suministro *in vivo*. Para minimizar el nivel de toxicidad y efecto inespecífico, el ARNm de la enzima nickasa de CRISPR mutada (por ejemplo, Cas9 de *S. pyogenes* con una mutación N863A) puede suministrarse con un par de ARN guía dirigidos a un sitio de interés. Los dos ARN guía tienen que estar espaciados de la siguiente manera. Las secuencias guía en rojo (subrayado simple) y azul (subrayado doble) respectivamente (estos ejemplos se basan en el requisito de PAM para Cas9 de *Streptococcus pyogenes*.)

Longitud de saliente (pb)	Diseño de ARN guía (secuencia guía y PAM con código de colores)
	5'-NNNNNNNNNNNNNNNNNNNNCCN <u>NNNNNNNNNNNNNNNNNNNN</u> NGGNNNNNNNNNNNNNNNN-3'
14	3'-NNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNNCCN <u>NNNNNNNNNNNNNNNNNNNN</u> NGGNNNNNNNNNNNNNNNN-3'
13	3'-NNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNNCCN <u>NNNNNNNNNNNNNNNNNNNN</u> NGGNNNNNNNNNNNNNNNN-3'
12	3'-NNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNNCCN <u>NNNNNNNNNNNNNNNNNNNN</u> NGGNNNNNNNNNNNNNNNN-3'
11	3'-NNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNNCCN <u>NNNNNNNNNNNNNNNNNNNN</u> NGGNNNNNNNNNNNNNNNN-3'

ES 2 780 904 T3

Longitud de saliente (pb)	Diseño de ARN guía (secuencia guía y PAM con código de colores)
10	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
9	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
8	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
7	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
6	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
5	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
4	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
3	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
2	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
1	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
romo	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
1	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'
2	3'-NNNNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN-5'
	5'-NNNNNNNNNNNNNNNNNNNN <u>CC</u> NNNNNNNNNNNNNNNNNN <u>GG</u> NNNNNNNNNNNNNNNNNN-3'

con las guías modificadas de Konermann *et al.* (Nature 517, 583-588, 29 de enero de 2015). La referencia a un dominio funcional podría ser un dominio funcional asociado con la enzima de CRISPR o un dominio funcional asociado con la proteína adaptadora. En algunas realizaciones, el uno o más dominios funcionales es una NLS (secuencia de localización nuclear) o una NES (señal de exportación nuclear). En algunas realizaciones, el uno o más dominios funcionales es un dominio de activación transcripcional que comprende VP64, p65, MyoD1, HSF1, RTA, SET7/9 y una histona acetiltransferasa. Otras referencias en este documento a dominio de activación (o activadores) con respecto a los asociados con la enzima de CRISPR incluyen cualquier dominio de activación transcripcional conocido y específicamente VP64, p65, MyoD1, HSF1, RTA, SET7/9 o una histona acetiltransferasa. En algunas realizaciones, el uno o más dominios funcionales es un dominio represor transcripcional. En algunas realizaciones, el dominio represor transcripcional es un dominio KRAB. En algunas realizaciones, el dominio represor transcripcional es un dominio NuE, dominio NcoR, dominio SID o un dominio SID4X. En algunas realizaciones, el uno o más dominios funcionales tienen una o más actividades que comprenden actividad metilasa, actividad desmetilasa, actividad de activación de la transcripción, actividad de represión de la transcripción, actividad de factor de liberación de la transcripción, actividad de modificación de histonas, actividad de escisión de ARN, actividad de escisión de ADN, actividad de integración de ADN o actividad de unión a ácido nucleico. Los dominios de modificación de histonas también son preferidos en algunas realizaciones. A continuación se analizan dominios de modificación de histonas ejemplares. Los dominios transposasa, dominios de maquinaria de HR (recombinación homóloga), dominios recombinasa y/o dominios integrasa también son preferidos como los presentes dominios funcionales. En algunas realizaciones, la actividad de integración de ADN incluye dominios de maquinaria de HR, dominios integrasa, dominios recombinasa y/o dominios transposasa. Se prefieren histona acetiltransferasas en algunas realizaciones. En algunas realizaciones, la actividad de escisión de ADN se debe a una nucleasa. En algunas realizaciones, la nucleasa comprende una nucleasa Fok1. En algunas realizaciones, el uno o más dominios funcionales se fijan a la enzima de CRISPR de modo que, tras la unión al ARNGs y la diana, el dominio funcional esté en una orientación espacial que permita que el dominio funcional funcione en su función atribuida. En algunas realizaciones, el uno o más dominios funcionales se fijan a la proteína adaptadora de modo que, tras la unión de la enzima de CRISPR al ARNGs y la diana, el dominio funcional esté en una orientación espacial que permita que el dominio funcional funcione en su función atribuida. En un aspecto, la divulgación proporciona una composición como se analiza en este documento, en la que el uno o más dominios funcionales se fijan a la enzima de CRISPR o proteína adaptadora mediante un conector, opcionalmente un conector de GlySer, como se analiza en este documento. La represión transcripcional endógena a menudo está mediada por enzimas modificadoras de la cromatina tales como histona metiltransferasas (HMT) y desacetilasas (HDAC). Se conocen dominios efectores de histonas represores y a continuación se proporciona una lista ejemplar. En la tabla ejemplar, se dio preferencia a proteínas y truncamientos funcionales de pequeño tamaño para facilitar el empaquetado vírico eficaz (por ejemplo, mediante AAV). En general, sin embargo, los dominios pueden incluir HDAC, histona metiltransferasas (HMT) e inhibidores de histona acetiltransferasa (HAT), así como proteínas de reclutamiento de HDAC y HMT. El dominio funcional puede ser o incluir, en algunas realizaciones, dominios efectores HDAC, dominios efectores de reclutamiento de HDAC, dominios efectores de histona metiltransferasa (HMT), dominios efectores de reclutamiento de histona metiltransferasa (HMT) o dominios efectores inhibidores de histona acetiltransferasa.

40 Dominios efectores HDAC

Subtipo/complejo	Nombre	Sustrato (si se conoce)	Modificación (si se conoce)	Organismo	Tamaño completo (aa)	Truncamiento seleccionado (aa)	Tamaño final (aa)	Dominio catalítico
HDAC I	HDAC8	-	-	<i>X. laevis</i>	325	1-325	325	1-272: HDAC
HDAC I	RPD3	-	-	<i>S. cerevisiae</i>	433	19-340	322 (Vannier)	19-331: HDAC
HDAC IV	MesoLo4	-	-	<i>M. loti</i>	300	1-300 (Gregoretto)	300	-
HDAC IV	HDAC11	-	-	<i>H. sapiens</i>	347	1-347 (Gao)	347	14-326: HDAC
HD2	HDT1	-	-	<i>A. thaliana</i>	245	1-211 (Wu)	211	-
SIRT I	SIRT3	H3K9Ac H4K16Ac H3K56Ac	-	<i>H. sapiens</i>	399	143-399 (Scher)	257	126-382: SIRT
SIRT I	HST2	-	-	<i>C. albicans</i>	331	1-331 (Hnisz)	331	-
SIRT I	CobB	-	-	<i>E. coli (K12)</i>	242	1-242 (Landry)	242	-
SIRT I	HST2	-	-	<i>S. cerevisiae</i>	357	8-298 (Wilson)	291	-

ES 2 780 904 T3

Subtipo/complejo	Nombre	Sustrato (si se conoce)	Modificación (si se conoce)	Organismo	Tamaño completo (aa)	Truncamiento seleccionado (aa)	Tamaño final (aa)	Dominio catalítico
SIRT III	SIRT5	H4K8Ac H4K16Ac	-	<i>H. sapiens</i>	310	37-310 (Gertz)	274	41-309: SIRT
SIRT III	Sir2A	-	-	<i>P. falciparum</i>	273	1-273 (Zhu)	273	19-273: SIRT
SIRT IV	SIRT6	H3K9Ac H3K56Ac	-	<i>H. sapiens</i>	355	1-289 (Tennen)	289	35-274: SIRT

5 Por consiguiente, los dominios represores de la presente invención pueden seleccionarse de histona metiltransferasas (HMT), histona desacetilasas (HDAC), inhibidores de histona acetiltransferasa (HAT), así como proteínas de reclutamiento de HDAC y HMT.

El dominio HDAC puede ser cualquiera de los de la tabla anterior, concretamente: HDAC8, RPD3, MesoLo4, HDAC11, HDT1, SIRT3, HST2, CobB, HST2, SIRT5, Sir2A o SIRT6.

10 En alguna realización, el dominio funcional puede ser un dominio efector de reclutamiento de HDAC. Ejemplos preferidos incluyen los de la tabla a continuación, concretamente MeCP2, MBD2b, Sin3a, NcoR, SALL1, RCOR1. NcoR se ejemplifica en los presentes ejemplos y, aunque se prefiere, se prevé que también serán útiles otros de la clase.

15 Tabla de dominios efectores de reclutamiento de HDAC

Subtipo/complejo	Nombre	Sustrato (si se conoce)	Modificación (si se conoce)	Organismo	Tamaño completo (aa)	Truncamiento seleccionado (aa)	Tamaño final (aa)	Dominio catalítico
Sin3a	MeCP2	-	-	<i>R. norvegicus</i>	492	207-492 (Nan)	286	-
Sin3a	MBD2b	-	-	<i>H. sapiens</i>	262	45-262 (Boeke)	218	-
Sin3a	Sin3a	-	-	<i>H. sapiens</i>	1273	524-851 (Laherty)	328	627-829: interacción con HDAC1
NcoR	NcoR	-	-	<i>H. sapiens</i>	2440	420-488 (Zhang)	69	-
NuRD	SALL1	-	-	<i>M. musculus</i>	1322	1-93 (Lauberth)	93	-
CoREST	RCOR1	-	-	<i>H. sapiens</i>	482	81-300 (Gu, Ouyang)	220	-

20 En alguna realización, el dominio funcional puede ser un dominio efector metiltransferasa (HMT). Ejemplos preferidos incluyen los de la tabla a continuación, concretamente NUE, vSET, EHMT2/G9A, SUV39H1, dim-5, KYP, SUVR4, SET4, SET1, SETD8 y TgSET8. NUE se ejemplifica en los presentes ejemplos y, aunque se prefiere, se prevé que también serán útiles otros de la clase.

Tabla de dominios efectores de histona metiltransferasa (HMT)

Subtipo/complejo	Nombre	Sustrato (si se conoce)	Modificación (si se conoce)	Organismo	Tamaño completo (aa)	Truncamiento seleccionado (aa)	Tamaño final (aa)	Dominio catalítico
SET	NUE	H2B, H3, H4	-	<i>C. trachomatis</i>	219	1-219 (Pennini)	219	-
SET	vSET	-	H3K27me3	<i>P. bursaria chlorella virus</i>	119	1-119 (Mujtaba)	119	4-112: SET2

ES 2 780 904 T3

Subtipo/complejo	Nombre	Sustrato (si se conoce)	Modificación (si se conoce)	Organismo	Tamaño completo (aa)	Truncamiento seleccionado (aa)	Tamaño final (aa)	Dominio catalítico
familia SUV39	EHMT2/G9A	H1,4K2, H3K9, H3K27	H3K9me1/2, H1K25me1	<i>M. musculus</i>	1263	969-1263 (Tachibana)	295	1025-1233: preSET, SET, postSET
SUV39	SUV39H1	-	H3K9me2/3	<i>H. sapiens</i>	412	79-412 (Snowden)	334	172-412: preSET, SET, postSET
Suvar3-9	dim-5	-	H3K9me3	<i>N. crassa</i>	331	1-331 (Rathert)	331	77-331: preSET, SET, postSET
Suvar3-9 (subfamilia SUVH)	KYP	-	H3K9me1/2	<i>A. thaliana</i>	624	335-601	267 (Jackson)	-
Suvar3-9 (subfamilia SUVR)	SUVR4	H3K9me1	H3K9me2/3	<i>A. thaliana</i>	492	180-492	313 (Thorstensen)	192-462: preSET, SET, postSET
Suvar4-20	SET4	-	H4K20me3	<i>C. elegans</i>	288	1-288 (Vielle)	288	-
SET8	SET1	-	H4K20me1	<i>C. elegans</i>	242	1-242 (Vielle)	242	-
SET8	SETD8	-	H4K20me1	<i>H. sapiens</i>	393	185-393	209 (Couture)	256-382: SET
SET8	TgSET8	-	H4K20me1/2/3	<i>T. gondii</i>	1893	1590-1893 (Sautel)	304	1749-1884: SET

En alguna realización, el dominio funcional puede ser un dominio efector de reclutamiento de histona metiltransferasa (HMT). Ejemplos preferidos incluyen los de la tabla a continuación, concretamente Hp1a, PHF19 y NIPP1.

5

Tabla de dominios efectores de reclutamiento de histona metiltransferasa (HMT)

Subtipo/complejo	Nombre	Sustrato (si se conoce)	Modificación (si se conoce)	Organismo	Tamaño completo (aa)	Truncamiento seleccionado (aa)	Tamaño final (aa)	Dominio catalítico
-	Hp1a	-	H3K9me3	<i>M. musculus</i>	191	73-191	119 (Hathaway)	121-179: chromoshadow
-	PHF19	-	H3K27me3	<i>H. sapiens</i>	580	(1-250) conector GGSG (500-580) +	335 (Ballare) +	163-250: PHD2
-	NIPP1	-	H3K27me3	<i>H. sapiens</i>	351	1-329 (Jin)	329	310-329: EED

En alguna realización, el dominio funcional puede ser dominio efector inhibidor de histona acetiltransferasa. Ejemplos preferidos incluyen SET/TAF-1β enumerado en la tabla a continuación.

10

Tabla de dominios efectores inhibidores de histona acetiltransferasa

Subtipo/complejo	Nombre	Sustrato (si se conoce)	Modificación (si se conoce)	Organismo	Tamaño completo (aa)	Truncamiento seleccionado (aa)	Tamaño final (aa)	Dominio catalítico
-	SET/TAF-1β	-	-	<i>M. musculus</i>	289	1-289 (Cervoni)	289	-

5 También se prefiere abordar elementos de control endógenos (reguladores) (tales como potenciadores y silenciadores) además de un promotor o elementos proximales del promotor. Por tanto, la invención también puede usarse para abordar elementos de control endógenos (incluyendo potenciadores y silenciadores) además de abordar el promotor. Estos elementos de control pueden estar ubicados anteriores o posteriores del sitio de inicio de la transcripción (TSS), partiendo de 200 pb del TSS hasta 100 kb de distancia. Abordar elementos de control conocidos puede usarse para activar o reprimir el gen de interés. En algunos casos, un solo elemento de control puede incluir en la transcripción de múltiples genes diana. Por lo tanto, abordar un solo elemento de control podría usarse para controlar la transcripción de múltiples genes simultáneamente.

15 Abordar supuestos elementos de control, por otro lado (por ejemplo, revistiendo la región del supuesto elemento de control, así como de 200 pb hasta 100 kb alrededor del elemento) puede usarse como medio para verificar dichos elementos (midiendo la transcripción del gen de interés) o para detectar elementos de control novedosos (por ejemplo, revistiendo 100 kb anteriores y posteriores del TSS del gen de interés). Además, abordar supuestos elementos de control puede ser útil en el contexto de la comprensión de las causas genéticas de las enfermedades. Muchas mutaciones y variantes de SNP comunes asociadas con fenotipos de enfermedad están ubicadas fuera de las regiones codificantes. Abordar dichas regiones con los sistemas de activación o represión descritos en este documento puede estar seguido de lectura de la transcripción de a) un conjunto de supuestas dianas (por ejemplo, un conjunto de genes ubicados en máxima proximidad al elemento de control) o b) lectura del todo el transcriptoma por, por ejemplo, secuenciación de ARN o micromatriz. Esto permitiría la identificación de genes candidatos probables implicados en el fenotipo de la enfermedad. Dichos genes candidatos podrían ser útiles como dianas de fármaco novedosas.

25 Se mencionan en este documento inhibidores de histona acetiltransferasa (HAT). Sin embargo, una alternativa en algunas realizaciones es que el uno o más dominios funcionales comprendan una acetiltransferasa, preferiblemente una histona acetiltransferasa. Estos son útiles en el campo de la epigenómica, por ejemplo, en métodos de consulta del epigenoma. Los métodos de consulta del epigenoma pueden incluir, por ejemplo, abordar secuencias epigenómicas. Abordar secuencias epigenómicas puede incluir que la guía se dirija a una secuencia diana epigenómica. La secuencia diana epigenómica puede incluir, en algunas realizaciones, un promotor, silenciador o una secuencia potenciadora.

35 El uso de un dominio funcional unido a una enzima de CRISPR-Cas como se describe en este documento, preferiblemente una nickasa Cas, o una Cas que muestre poca o ninguna o no más de un 5 o 4 o 3 o 2 o 1 % de actividad nucleasa (en comparación con una Cas no mutada), por ejemplo, una dead-Cas, para abordar secuencias epigenómicas diana puede usarse para activar o reprimir promotores, silenciadores o potenciadores. En la presente invención se prefiere que la Cas sea una nickasa, incluyendo que la invención puede implicar nickasas dobles.

40 Se conocen ejemplos de acetiltransferasas, pero pueden incluir, en algunas realizaciones, histona acetiltransferasas. En algunas realizaciones, la histona acetiltransferasa puede comprender el núcleo catalítico de la acetiltransferasa humana p300 (Gerbasch y Reddy, Nature Biotech 6 de abril de 2015).

45 En algunas realizaciones preferidas, el dominio funcional se une a la enzima Cas9 para abordar y activar secuencias genómicas tales como promotores o potenciadores. También puede proporcionarse una o más guías dirigidas a dichos promotores o potenciadores para dirigir la unión de la enzima de CRISPR a dichos promotores o potenciadores.

50 La expresión "asociado con" se usa en esta ocasión en relación con la asociación del dominio funcional a la enzima de CRISPR o la proteína adaptadora. Se usa con respecto a la manera en que una molécula "se asocia" con respecto a otra, por ejemplo, entre una proteína adaptadora y un dominio funcional, o entre la enzima de CRISPR y un dominio funcional. En el caso de dichas interacciones de proteína-proteína, esta asociación puede verse en términos de reconocimiento en la manera en que un anticuerpo reconoce un epítipo. Como alternativa, una proteína puede asociarse con otra proteína mediante una fusión de las dos, por ejemplo, fusionándose una subunidad con otra subunidad. La fusión típicamente se produce mediante la adición de la secuencia de aminoácidos de una a la de la otra, por ejemplo, mediante corte y empalme conjunto de las secuencias de nucleótidos que codifican cada proteína o subunidad. Como alternativa, esto puede verse esencialmente como unión entre dos moléculas o enlace directo, tal como una proteína de fusión. En cualquier caso, la proteína de fusión puede incluir un conector entre las dos subunidades de interés (es decir, entre la enzima y el dominio funcional o entre la proteína adaptadora y el dominio funcional). Por tanto, en algunas realizaciones, la enzima de CRISPR o la proteína adaptadora se asocia con un dominio funcional mediante unión al mismo. En otras realizaciones, la enzima de CRISPR o la proteína adaptadora se

asocia con un dominio funcional porque los dos se fusionan conjuntamente, opcionalmente mediante un conector intermedio.

La fijación puede ser mediante un conector, por ejemplo, un conector flexible de glicina-serina (GlyGlyGlySer) o (GGGS)₃ o un conector rígido alfa-helicoidal tal como (Ala(GluAlaAlaAlaLys)Ala). Conectores tales como (GGGS)₃ se usan preferiblemente en este documento para separar los dominios proteínicos o peptídicos. (GGGS)₃ es preferible porque es un conector relativamente largo (15 aminoácidos). Los residuos de glicina son los más flexibles y los residuos de serina potencian la posibilidad de que el conector esté en el exterior de la proteína. (GGGS)₆ (GGGS)₉ o (GGGS)₁₂ pueden usarse preferiblemente como alternativas. Otras alternativas preferidas son (GGGS)₁, (GGGS)₂, (GGGS)₄, (GGGS)₅, (GGGS)₇, (GGGS)₈, (GGGS)₁₀ o (GGGS)₁₁. Están disponibles conectores alternativos, pero se cree que conectores altamente flexibles funcionan mejor permitiendo una oportunidad máxima de que las 2 partes de la Cas9 queden juntas y, por tanto, reconstituyan la actividad de Cas9. Una alternativa es que pueda usarse la NLS de nucleoplasmina como conector. Por ejemplo, también puede usarse un conector entre la Cas9 y cualquier dominio funcional. De nuevo, puede usarse un conector (GGGS)₃ en esta ocasión (o las versiones de 6, 9 o 12 repeticiones del mismo) o puede usarse la NLS de nucleoplasmina como conector entre Cas9 y el dominio funcional.

Aspectos de la invención se refieren a la expresión del producto génico que se disminuye o un polinucleótido molde que se introduce adicionalmente en la molécula de ADN que codifica el producto génico o una secuencia intermedia que se escinde permitiendo que los dos salientes 3' se rehibriden y se ligen o la actividad o función del producto génico se altere o la expresión del producto génico se aumente. En una realización de la invención, el producto génico es una proteína. La escisión de la secuencia intermedia puede tener precisión a través del uso de los salientes 3'. Por tanto, la invención prevé una secuencia intermedia que se escinde de forma precisa permitiendo que los dos salientes 3' se rehibriden y se ligen o la actividad o función del producto génico se altere o la expresión del producto génico se aumente.

Las opciones de suministro ventajosas para suministro al cerebro incluyen encapsulación de enzima de CRISPR mutada y ARN guía en forma de ADN o ARN en liposomas y conjugación con los Caballos de Troya moleculares para suministro a través de la barrera hematoencefálica (BBB). Se ha mostrado que los Caballos de Troya moleculares son eficaces para suministro de vectores de expresión de B-gal al cerebro de primates no humanos. El mismo procedimiento puede usarse para el suministro de vectores que contienen enzima de CRISPR y ARN guía. Por ejemplo, Xia CF and Boado RJ, Pardridge WM ("Antibody-mediated targeting of siRNA via the human insulin receptor using avidin-biotin technology." *Mol Pharm.* 2009 mayo-junio; 6(3):747-51. doi: 10.1021/mp800194) describe la manera en que el suministro de ARN interferente pequeño (ARNip) a células en cultivo, e *in vivo*, es posible con el uso combinado de un anticuerpo monoclonal específico de receptor (mAb) y tecnología de avidina-biotina. Los autores también informan de que, como el enlace entre el mAb de dirección y el ARNip es estable con la tecnología de avidina-biotina, y se observan efectos de iARN en sitios distantes tales como cerebro *in vivo* después de una administración intravenosa del ARNip dirigido.

Zhang Y, Schlachetzki F, Pardridge WM. ("Global non-viral gene transfer to the primate brain following intravenous administration." *Mol Ther.* 2003 enero; 7(1):11-8) describen la manera en que los plásmidos de expresión que codifican indicadores tales como luciferasa se encapsularon en el interior de un "virus artificial" compuesto de un inmunoliposoma pegilado de 85 nm, que se dirigía al cerebro de macaco *in vivo* con un anticuerpo monoclonal (mAb) contra el receptor de insulina humana (HIR). El HIRmAb posibilita que el liposoma que porta el gen exógeno experimente transcitosis a través de la barrera hematoencefálica y endocitosis a través de la membrana plasmática neuronal después de inyección intravenosa. El nivel de expresión génica de luciferasa en el cerebro fue 50 veces mayor en el macaco en comparación con la rata. Se demostró expresión neuronal generalizada del gen de la beta-galactosidasa en cerebro de primate tanto por histoquímica como por microscopía confocal. Los autores indican que esta estrategia hace factibles transgénicos adultos reversibles en 24 horas. Por consiguiente, se prefiere el uso de inmunoliposoma. Estos pueden usarse junto con anticuerpos para abordar tejidos específicos o proteínas de superficie celular. También se prefieren otros medios de suministro o para suministro de ARN, tales como mediante nanopartículas (Cho, S., Goldberg, M., Son, S., Xu, Q., Yang, F., Mei, Y., Bogatyrev, S., Langer, R. y Anderson, D., Lipid-like nanoparticles for small interfering RNA delivery to endothelial cells, *Advanced Functional Materials*, 19: 3112-3118, 2010) o exosomas (Schroeder, A., Levins, C., Cortez, C. Langer, R., y Anderson, D., Lipid-based nanotherapeutics for siRNA delivery, *Journal of Internal Medicine*, 267: 9-21, 2010, PMID: 20059641). De hecho, los exosomas han demostrado ser particularmente útiles en el suministro de ARNip, un sistema con algunos paralelismos al sistema de CRISPR. Por ejemplo, El-Andaloussi S, *et al.*, ("Exosome-mediated delivery of siRNA in vitro and in vivo." *Nat Protoc.* diciembre de 2012;7(12):2112-26. doi: 10.1038/nprot.2012.131. Epub 15 de noviembre de 2012) describen la manera en que los exosomas son herramientas prometedoras para el suministro de fármacos a través de diferentes barreras biológicas y pueden aprovecharse para el suministro de ARNip *in vitro* e *in vivo*. Su estrategia es generar exosomas dirigidos a través de transfección de un vector de expresión, que comprende una proteína exosómica fusionada con un ligando peptídico. Los exosomas entonces se purifican y caracterizan a partir del sobrenadante celular transfectado, después el ARNip se carga en los exosomas.

Se prefiere eliminación dirigida de genes. Por lo tanto, se prefieren genes implicados en la biosíntesis del colesterol, la biosíntesis de ácidos grasos y otros trastornos metabólicos, genes que codifican proteínas mal plegadas implicadas

en enfermedades por amiloides y otras enfermedades, oncogenes que dan lugar a transformación celular, genes víricos latentes y genes que dan lugar a trastornos dominantes-negativos, entre otros trastornos. Como se ejemplifica en esta ocasión, los solicitantes prefieren suministro génico de un sistema de CRISPR-Cas al hígado, cerebro, tejido ocular, epitelial, hematopoyético u otro de un sujeto o un paciente que lo necesita, que padece trastornos metabólicos, amiloidosis y enfermedades relacionadas con la agregación de proteínas, transformación celular que surge de mutaciones genéticas y translocaciones, efectos negativos dominantes de mutaciones génicas, infecciones víricas latentes y otros síntomas relacionados, usando sistema de suministro vírico o de nanopartículas.

Las aplicaciones terapéuticas del sistema de CRISPR-Cas incluyen glaucoma, amiloidosis y enfermedad de Huntington.

Como ejemplo, la infección crónica por VIH-1 puede tratarse o prevenirse. Para conseguir esto, se pueden generar ARN guía de CRISPR-Cas que estén dirigidos a la inmensa mayoría del genoma de VIH-1 teniendo en cuenta al mismo tiempo variantes de la cepa de VIH-1 para una cobertura y eficacia máximas. Se puede conseguir suministro del sistema de CRISPR-Cas por infección convencional mediada con adenovirus o lentivirus del sistema inmunitario del hospedador. Dependiendo de la estrategia, las células inmunitarias del hospedador podrían a) aislarse, transducirse con CRISPR-Cas, seleccionarse y volver a introducirse en el hospedador o b) transducirse *in vivo* mediante suministro sistémico del sistema de CRISPR-Cas. La primera estrategia permite la generación de una población inmunorresistente mientras que la segunda es más probable que se dirija a depósitos víricos dentro del hospedador.

También se prevé que la presente divulgación genere una colección de células con inactivación génica. Cada célula puede tener un solo gen inactivado.

Se puede preparar una colección de células ES donde cada célula tiene un solo gen inactivado, y la colección completa de células ES tendrá cada gen individual inactivado. Esta colección es útil para cribar la función génica en procesos celulares, así como enfermedades. Para preparar esta colección de células, se puede integrar Cas9 mutada dirigida por un promotor inducible (por ejemplo, promotor inducible por doxiciclina) en la célula ES. Además, se puede integrar un ARN guía simple dirigido a un gen específico en la génica ES. Para preparar la colección de células ES, se pueden mezclar simplemente las células ES con una genoteca que codifique ARN guía dirigidos a cada gen en el genoma humano. Se puede introducir en primer lugar un solo sitio BxB1 attB en el locus AAVS1 de la célula ES humana. Entonces, se puede usar la integrasa BxB1 para facilitar la integración de genes de ARN guía individuales en el sitio BxB1 attB en el locus AAVS1. Para facilitar la integración, cada gen de ARN guía puede estar contenido en un plásmido que porta un solo sitio attP. De esta manera, BxB1 recombinará el sitio attB en el genoma con el sitio attP en el plásmido que contiene ARN guía. Para generar la colección celular, se puede tomar la colección de células que tienen ARN guía simples integrados e inducir la expresión de Cas9 mutada. Después de la inducción, la Cas9 mutada emparejada media el mellado emparejado en sitios especificados por el ARN guía.

La administración crónica de agentes terapéuticos proteínicos puede provocar respuestas inmunitarias inaceptables contra la proteína específica. La inmunogenia de los fármacos proteínicos puede atribuir a unos pocos epítopos de linfocitos T auxiliares (HTL) inmunodominantes. Reducir la afinidad de unión a MHC de estos epítopos HTL contenidos dentro de estas proteínas puede generar fármacos con menor inmunogenia (Tangri S, *et al.* ("Rationally engineered therapeutic proteins with reduced immunogenicity" J Immunol. 15 de marzo de 2005;174(6):3187-96). En la presente divulgación, la inmunogenia de la enzima de CRISPR en particular puede reducirse siguiendo la estrategia expuesta por primera vez en Tangri *et al.* con respecto a la eritropoyetina y desarrollada posteriormente. Por consiguiente, puede usarse evolución dirigida o diseño racional para reducir la inmunogenia de la enzima de CRISPR mutada (por ejemplo, una Cas9 mutada) en la especie hospedadora (ser humano u otra especie).

Los trastornos por repeticiones trinucleotídicas son afecciones preferidas a tratar.

De acuerdo con otro aspecto, se proporciona un método de genoterapia para el tratamiento de un sujeto que tiene una mutación en el gen de CFTR y comprende administrar una cantidad terapéuticamente eficaz de una partícula de genoterapia de CRISPR-Cas, opcionalmente mediante un vehículo farmacéutico biocompatible, a las células de un sujeto. Preferiblemente, el ADN diana comprende la mutación deltaF508. En general, se prefiere que la mutación se repare al tipo silvestre. En este caso, la mutación es una eliminación de los tres nucleótidos que componen el codón de fenilalanina (F) en la posición 508. Por consiguiente, la reparación en este caso requiere la reintroducción del codón ausente en el mutante.

Para implementar esta estrategia de reparación génica, se prefiere que se introduzca un sistema de vector de adenovirus/AAV en la célula, células o paciente hospedador. Preferiblemente, el sistema comprende una Cas9 mutada (o nickasa Cas9 mutada) y el ARN guía junto con un sistema de vector de adenovirus/AAV que comprende el molde de reparación por homología que contiene el residuo F508. Esto puede introducirse en el sujeto mediante uno de los métodos de suministro analizados previamente. El sistema de CRISPR-Cas puede estar guiado por el ARN guía quimérico CFTRdelta 508. Está dirigido a un sitio específico del locus genómico de CFTR a mellar o escindir. Después de la escisión, se inserta el molde de reparación en el sitio de escisión mediante recombinación homóloga que corrige la eliminación que provoca fibrosis quística o causa síntomas relacionados con fibrosis quística. Esta estrategia de

dirigir el suministro y proporcionar introducción sistémica de sistemas de CRISPR con ARN guía apropiados puede emplearse abordar mutaciones genéticas para editar o manipular de otro modo los genes que causan enfermedades y trastornos metabólicos, hepáticos, renales y proteínicos tales como los de la tabla B.

5 Los sujetos tratados en este caso reciben cantidad farmacéuticamente eficaz de sistema de vector AAV en aerosol por pulmón de forma endobronquial suministrado mientras se respira de forma espontánea. Por tanto, el suministro en aerosol se prefiere para el suministro de AAV en general. Puede usarse un adenovirus o una partícula de AAV para al suministro. Pueden clonarse construcciones génicas adecuadas, cada una unida de forma funcional a una o más secuencias reguladoras, en el vector de suministro. En este caso, se proporcionan las siguientes construcciones como
10 ejemplos: promotor Cbh o EFla para Cas9, promotor U6 o H1 para ARN guía quimérico. Una disposición preferida es usar una guía quimérica dirigida a CFTRdelta508, un molde de reparación para la mutación deltaF508 y una nickasa Cas9 mutada de codones optimizados con opcionalmente una o más señales o secuencias de localización nuclear (NLS), por ejemplo, dos (2) o dos (2) o más NLS. También se prevén construcciones sin NLS. Sin embargo, realizaciones ventajosas pueden implicar dos o dos o más NLS.

15 Para identificar el sitio diana de Cas9, los solicitantes analizaron el locus genómico de CFTR e identificaron el sitio diana de Cas9. Preferiblemente, en general y en este caso de CF, el PAM puede contener un motivo NGG o uno NNAGAAW.

20 Alternativas a CF incluyen cualquier trastorno genético y ejemplos de estos son bien conocidos. Otro método preferido o uso de la invención es para corregir defectos en los genes EMP2A y EMP2B que se han identificado como asociados con la enfermedad de Lafora.

25 En algunas realizaciones, una "secuencia guía" puede ser distinta del "ARN guía". Una secuencia guía puede referirse a una secuencia de aprox. 20 pb, dentro del ARN guía, que especifica el sitio diana.

30 Será fácilmente evidente que un hospedador de otras enfermedades puede tratarse de una manera similar. Algunos ejemplos de enfermedades genéticas causadas por mutaciones se proporcionan en este documento, pero se conocen muchas más. La estrategia anterior puede aplicarse a estas enfermedades.

La invención usa ácidos nucleicos para que se unan a secuencias de ADN diana. Esto es ventajoso ya que los ácidos nucleicos son mucho más fáciles y baratos de producir y la especificidad puede variarse de acuerdo con la longitud del tramo donde se busca homología. No se requiere colocación 3-D compleja de múltiples dedos, por ejemplo.

35 Las expresiones "polinucleótido", "nucleótido", "secuencia de nucleótidos", "ácido nucleico" y "oligonucleótido" se usan indistintamente. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sean desoxirribonucleótidos o ribonucleótidos, o análogos de los mismos. Los polinucleótidos pueden tener cualquier estructura tridimensional y pueden realizar cualquier función, conocida o desconocida. Los siguientes son ejemplos no limitantes de polinucleótidos: regiones codificantes o no codificantes de un gen o fragmentos génicos, uno o más locus definidos a partir del análisis de ligamiento, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ARN interferente pequeño (ARNip), ARN de horquilla corta (ARNhc), microARN (miARN), ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. El término también abarca
40 estructuras de tipo ácido nucleico con cadenas principales sintéticas, véase, por ejemplo, Eckstein, 1991; Baserga *et al.*, 1992; Milligan, 1993; documento WO 97/03211; documento WO 96/39154; Mata, 1997; Strauss-Soukup, 1997; y Samstag, 1996. Un polinucleótido puede comprender uno o más nucleótidos modificados, tales como nucleótidos metilados y análogos nucleotídicos. Si están presentes, las modificaciones a la estructura nucleotídica puede conferirse antes o después del ensamblaje del polímero. La secuencia de nucleótidos se puede interrumpir con componentes que no sean nucleótidos. Un polinucleótido puede modificarse además después de la polimerización, tal como por conjugación con un componente de marcaje.

55 Como se usa en este documento, la expresión "tipo silvestre" es una expresión de la técnica comprendida por los expertos en la materia y significa la forma típica de un organismo, cepa, gen o característica como se produce en la naturaleza que se distingue de formas mutantes o variantes.

Como se usa en este documento, debe aceptarse que el término "variante" significa mostrar cualidades que tienen un patrón que se desvía de los que se produce en la naturaleza.

60 Las expresiones "que no es de origen natural" o "que está manipulado" se usan indistintamente e indican la implicación de la mano del hombre. Las expresiones, cuando se refieren a moléculas de ácido nucleico o polipéptidos significan que la molécula de ácido nucleico o el polipéptido está al menos sustancialmente libre de al menos otro componente con el que están asociados de forma natural en la naturaleza y como se encuentran en la naturaleza.

65 "Complementariedad" se refiere a la capacidad de un ácido nucleico de formar enlaces de hidrógeno con otra secuencia de ácido nucleico por emparejamiento de bases tradicional de Watson y Crick u otros tipos no tradicionales. Un porcentaje de complementariedad indica el porcentaje de residuos en una molécula de ácido nucleico que pueden

formar enlaces de hidrógeno (por ejemplo, emparejamiento de bases de Watson y Crick) con una segunda secuencia de ácido nucleico (por ejemplo, 5, 6, 7, 8, 9, 10 de los 10 siendo un 50%, 60%, 70%, 80%, 90%, y 100% de complementariedad). "Perfectamente complementario" significa que todos los residuos contiguos de una secuencia de ácido nucleico formarán enlaces de hidrógeno con el mismo número de residuos contiguos en una segunda secuencia de ácido nucleico. "Sustancialmente complementario", como se usa en este documento, se refiere a un grado de complementariedad que es de al menos un 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, o 100% sobre una región de 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, o más nucleótidos, o se refiere a dos ácidos nucleicos que hibridan en condiciones rigurosas.

Como se usa en este documento, "condiciones rigurosas" para hibridación se refiere a condiciones en que un ácido nucleico que tiene complementariedad con una secuencia diana hibrida predominantemente con la secuencia diana, y sustancialmente no hibrida con secuencias que no son diana. Las condiciones rigurosas en general son dependientes de la secuencia, y varían dependiendo de varios factores. En general, cuanto más larga sea la secuencia, mayor será la temperatura a la que la secuencia hibrida específicamente con su secuencia diana. Ejemplos no limitantes de condiciones rigurosas se describen en detalle en Tijssen (1993) *Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes Parte I, Segundo Capítulo "Overview of principles of hybridization and the strategy of nucleic acid probe assay"*, Elsevier, N.Y. Cuando se hace referencia a una secuencia polinucleotídica, entonces también se prevén secuencias completa o parcialmente complementarias. Estas preferiblemente tienen capacidad de hibridar con la secuencia de referencia en condiciones altamente rigurosas. En general, para maximizar la tasa de hibridación, se seleccionan condiciones de hibridación de rigurosidad relativamente baja: aproximadamente 20 a 25 °C inferiores a punto de fusión térmica (T_m). La T_m es la temperatura a la que un 50 % de la secuencia diana específica hibrida con una sonda perfectamente complementaria en solución a una fuerza iónica y pH definidos. En general, para requerir al menos aproximadamente un 85 % de complementariedad de nucleótidos de las secuencias hibridadas, se seleccionan condiciones de lavado altamente rigurosas para que sean de 5 a 15 °C inferiores a la T_m . Para requerir al menos aproximadamente un 70 % de complementariedad de nucleótidos de las secuencias hibridadas, se seleccionan condiciones de lavado moderadamente rigurosas para que sean de 15 a 30 °C inferiores a la T_m . Las condiciones de lavado altamente permisivas (rigurosidad muy baja) pueden ser tan más como de 50 °C por debajo de la T_m , lo que permite un alto nivel de emparejamiento incorrecto entre secuencias hibridadas. Los expertos en la materia reconocerán que también pueden alterarse otros parámetros físicos y químicos en las fases de hibridación y lavado para influir en el resultado de una señal de hibridación detectable a partir de un nivel específico de homología entre secuencias diana y de sonda. Las condiciones altamente rigurosas preferidas comprenden incubación en formamida al 50 %, SSC 5x y SDS al 1 % a 42 °C, o incubación en SSC 5x y SDS al 1 % a 65 °C, con lavado en SSC 0,2x y SDS al 0,1 % a 65 °C.

"Hibridación" se refiere a una reacción en que uno o más polinucleótidos reaccionan para formar un complejo que se estabiliza mediante formación de enlaces de hidrógeno entre las bases de los residuos nucleotídicos. La formación de enlaces de hidrógeno puede producirse por emparejamiento de bases de Watson y Crick, unión de Hoogsteen, o de cualquier otra manera específica de secuencia. El complejo puede comprender dos hebras que forman una estructura bicatenaria, tres o más hebras que forman un complejo de múltiples hebras, una sola hebra de autohibridación o cualquier combinación de estas. Una reacción de hibridación puede constituir una etapa en un proceso más amplio, tal como el inicio de PCR, o la escisión de un polinucleótido por una enzima. Una secuencia que puede hibridar con una secuencia dada se denomina "complemento" de la secuencia dada.

Como se usa en este documento, la expresión "locus genómico" o "locus" es la ubicación específica de un gen o secuencia de ADN en un cromosoma. Un "gen" se refiere a tramos de ADN o ARN que codifican un polipéptido o una cadena de ARN que tiene que desempeñar una función en un organismo y, por tanto, es la unidad molecular de herencia en organismos vivos. Para el fin de esta invención, puede considerarse que los genes incluyen regiones que regulan la producción del producto génico, estén o no dichas secuencias reguladoras adyacentes a secuencias codificantes y/o transcritas. Por consiguiente, un gen incluye, aunque sin limitarse necesariamente, secuencias promotoras, terminadores, secuencias reguladoras de la traducción tales como sitios de unión a ribosomas y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aislantes, elementos limítrofes, orígenes de replicación, sitios de fijación a la matriz y regiones de control del locus.

Como se usa en este documento, "expresión de un locus genómico" o "expresión génica" es el proceso por el que la información de un gen se usa en la síntesis de un producto génico funcional. Los productos de expresión génica a menudo son proteínas, pero en genes que no codifican proteínas tales como genes de ARNr o genes de ARNt, el producto es ARN funcional. El proceso de expresión génica se usa por todos los seres vivos, eucariotas (incluyendo organismos multicelulares), procariotas (bacterias y arqueas) y virus para generar productos funcionales para sobrevivir. Como se usa en este documento, "expresión" de un gen o ácido nucleico abarca no solamente la expresión génica celular, sino también la transcripción y traducción de uno o más ácidos nucleicos en sistemas de clonación y en cualquier otro contexto. Como se usa en este documento, "expresión" también se refiere al proceso por el que un polinucleótido se transcribe desde un molde de ADN (tal como en un ARNm u otro transcrito de ARN) y/o el proceso por el que un ARNm transcrito se traduce posteriormente en péptidos, polipéptidos o proteínas. Los transcritos y polipéptidos codificados pueden denominarse colectivamente "producto génico". Si el polinucleótido se obtiene a partir de ADN genómico, la expresión puede incluir el corte y empalme del ARNm en una célula eucariótica.

Las expresiones "polipéptido", "péptido" y "proteína" se usan indistintamente en este documento para referirse a polímeros de aminoácidos de cualquier longitud. El polímero puede ser lineal o ramificado, puede comprender aminoácidos modificados, y puede estar interrumpido por elementos que no son aminoácidos. Las expresiones también abarcan un polímero de aminoácidos que se ha modificado; por ejemplo, formación de enlaces disulfuro, glucosilación, lipidación, acetilación, fosforilación o cualquier otra manipulación, tal como conjugación con un componente de marcaje. Como se usa en este documento, el término "aminoácido" incluye aminoácidos naturales y/o no naturales o sintéticos, incluyendo glicina y ambos isómeros ópticos D o L, y análogos de aminoácido y peptidomiméticos.

Como se usa en este documento, el término "domino" o "dominio proteínico" se refiere a una parte de una secuencia de proteína que puede existir y funcionar independientemente del resto de la cadena proteínica.

Como se describe en aspectos de la invención, la identidad de secuencia se refiere a la homología de secuencia. Las comparaciones de homología se pueden llevar a cabo de forma visual o, más habitualmente, con la ayuda de programas de comparación de secuencias que se pueden adquirir fácilmente. Estos programas informáticos disponible en el mercado pueden calcular el porcentaje (%) de homología entre dos o más secuencias y también pueden calcular la identidad de secuencia compartida por dos o más secuencias de aminoácidos o ácido nucleico. En algunas realizaciones preferidas, la región de recubrimiento de los dTALE descritos en este documento tienen secuencias que son al menos un 95 % idénticas o comparten identidad con las secuencias de aminoácidos de la región de recubrimiento proporcionadas en este documento.

Las homologías de secuencia pueden generarse por cualquiera de varios programas informáticos conocidos en la técnica, por ejemplo, BLAST o FASTA, etc. Un programa informático adecuado para realizar dicha alineación es el paquete GCG Wisconsin Bestfit (Universidad de Wisconsin, EE. UU.; Devereux *et al.*, 1984, Nucleic Acids Research 12:387). Ejemplos de otros programas informáticos distintos que pueden realizar comparaciones de secuencia incluyen, aunque sin limitación, el paquete BLAST (véase Ausubel *et al.*, 1999 *ibid* - capítulo 18), FASTA (Atschul *et al.*, 1990, J. Mol. Biol., 403-410) y el paquete GENWORKS de herramientas de comparación. Tanto BLAST como FASTA están disponibles para búsquedas sin conexión y en línea (véase Ausubel *et al.*, 1999 *ibid*, páginas 7-58 a 7-60). Sin embargo, se prefiere usar el programa GCG Bestfit.

Se puede calcular el porcentaje (%) de homología de secuencias a lo largo de secuencias contiguas, es decir, se alinea una secuencia con la otra secuencia y se compara directamente cada aminoácido o nucleótido en una secuencia con el correspondiente aminoácido o nucleótido en la otra secuencia, un residuo de cada vez. Esto se denomina alineación "sin huecos". Normalmente, dichas alineaciones sin huecos se llevan a cabo únicamente a lo largo de un número relativamente corto de residuos.

Aunque este es un método muy simple y lógico, no se puede tener en cuenta que, por ejemplo, en un par de secuencias por lo demás idénticas, una inserción o eliminación puede causar que los siguientes residuos aminoacídicos no estén alineados, lo que potencialmente dará como resultado una gran reducción en el % de homología cuando se lleve a cabo una alineación global. Por consiguiente, la mayoría de los métodos de comparación de secuencias están diseñados para producir alineaciones óptimas que tienen en cuenta posibles inserciones y eliminaciones sin penalizar excesivamente la puntuación global de homología o identidad. Esto se consigue insertando "huecos" en la alineación de secuencias para intentar maximizar la homología o identidad local.

Sin embargo, estos métodos más complejos asignan "penalizaciones por hueco" a cada hueco que se produzca en la alineación de manera que, para el mismo número de aminoácidos idénticos, una alineación de secuencias con el menor número de huecos posible (que refleja una relación mayor entre las dos secuencias comparadas) puede conseguir una puntuación mayor que otra con muchos huecos. Normalmente se usan "costes de hueco de afinidad" que asignan un coste relativamente elevado a la existencia de un hueco y una penalización menor para cada residuo posterior en el hueco. Este es el sistema de puntuación de huecos usado más comúnmente. Por supuesto, altas penalizaciones por hueco pueden producir alineaciones optimizadas con menos huecos. La mayoría de los programas de alineación permiten modificar las penalizaciones por hueco. Sin embargo, se prefiere usar los valores por defecto cuando se emplean dichos programas informáticos para comparaciones de secuencias. Por ejemplo, cuando se usa el paquete GCG Wisconsin Bestfit, la penalización por hueco por defecto para secuencias de aminoácidos es de -12 para un hueco y de -4 para cada prolongación.

El cálculo del % máximo de homología, por lo tanto, requiere en primer lugar la producción de una alineación óptima, teniendo en cuenta las penalizaciones por hueco. Un programa informático adecuado para realizar dicha alineación es el paquete GCG Wisconsin Bestfit (Devereux *et al.*, 1984 Nuc. Acids Research 12 p387). Ejemplos de otros programas informático que pueden realizar comparaciones de secuencia incluyen, aunque sin limitación, el paquete BLAST (véase Ausubel *et al.*, 1999 Short Protocols in Molecular Biology, 4.^a Ed. - capítulo 18), FASTA (Altschul *et al.*, 1990 J. Mol. Biol. 403-410) y el paquete GENWORKS de herramientas de comparación. Tanto BLAST como FASTA están disponibles para búsquedas sin conexión y en línea (véase Ausubel *et al.*, 1999, Short Protocols in Molecular Biology, páginas 7-58 a 7-60). Sin embargo, para algunas aplicaciones, se prefiere usar el programa GCG Bestfit. Una nueva herramienta, llamada BLAST 2 Sequences también está disponible para comparar secuencias de proteínas y

nucleótidos (véase FEMS Microbiol Lett. 1999 174(2): 247-50; FEMS Microbiol Lett. 1999 177(1): 187-8 y el sitio web del National Center for Biotechnology Information en el sitio web del National Institutes for Health).

Aunque el % de homología final se puede medir según la identidad, el propio proceso de alineación típicamente no se basa en una comparación de todo o nada de los pares. En su lugar, por lo general se usa una matriz de puntuación de la similitud graduada que asigna puntuaciones a cada comparación en pares en función de la similitud química o distancia evolutiva. Un ejemplo de una matriz de este tipo usada normalmente es la matriz BLOSUM62 (la matriz por defecto para el conjunto de programas BLAST). Por lo general, los programas GCG Wisconsin usan los valores por defecto públicos o la tabla de comparación de símbolos personalizada si se suministra (véase el manual del usuario para más detalles). Para algunas aplicaciones, se prefiere usar los valores por defecto públicos para el paquete GCG, o en el caso de otro programa informático, la matriz por defecto, tal como BLOSUM62.

Como alternativa, los porcentajes de homología pueden calcularse usando el elemento de múltiples alineaciones en DNASIS™ (Hitachi Software), basado en un algoritmo, análogo a CLUSTAL (Higgins DG y Sharp PM (1988), Gene 73(1), 237-244). Una vez que el programa informático ha producido una alineación óptima, es posible calcular el % de homología, preferiblemente el % de identidad de secuencia. El programa informático realiza esto típicamente como parte de la comparación de secuencias y genera un resultado numérico.

Las secuencias también pueden tener eliminaciones, inserciones o sustituciones de residuos aminoacídicos que producen un cambio sinónimo y dan como resultado una sustancia funcionalmente equivalente. Pueden hacerse sustituciones de aminoácido deliberadas basándose en la similitud en las propiedades de los aminoácidos (tales como polaridad, carga, solubilidad, hidrofobia, hidrofilia y/o la naturaleza anfipática de los residuos) y, por lo tanto, es útil para agrupar aminoácidos juntos en grupos funcionales. Los aminoácidos pueden agruparse juntos basándose en las propiedades de sus cadenas laterales en solitario. Sin embargo, es más útil incluir también datos de mutación. Los conjuntos de aminoácidos así derivados probablemente se tienen que conservar por razones estructurales. Estos conjuntos pueden describirse en forma de un diagrama de Venn (Livingstone C.D. y Barton G.J. (1993) "Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation" Comput. Appl. Biosci. 9: 745-756) (Taylor W.R. (1986) "The classification of amino acid conservation" J. Theor. Biol. 119; 205-218). Pueden hacerse sustituciones conservativas, por ejemplo, de acuerdo con la tabla a continuación que describe una agrupación de aminoácidos de diagrama de Venn generalmente aceptado.

Conjunto		Subconjunto	
Hidrófobo	F W Y H K M I L V A G C	Aromático	F W Y H
		Alifático	I L V
Polar	W Y H K R E D C S T N C	Cargado	H K R E D
		Cargado positivamente	H K R
		Cargado negativamente	E D
Pequeño	V C A G S P T N D	Pequeño	A G S

Realizaciones de la invención incluyen secuencias (polinucleotídicas o polipeptídicas) que pueden comprender sustitución homóloga (sustitución y remplazo se usan ambos en este documento para indicar el intercambio de un residuo aminoacídico o nucleótido existente con un residuo o nucleótido alternativo) que pueda producirse, es decir, una sustitución de similar por similar en el caso de aminoácidos, tal como básico por básico, ácido por ácido, polar por polar, etc. También puede producirse una sustitución no homóloga, es decir, de una clase de residuo a otra o, como alternativa, que implique la inclusión de aminoácidos no naturales, tales como ornitina (en lo sucesivo en este documento denominada Z), ácido diaminobutírico ornitina (en lo sucesivo en este documento denominada B), norleucina ornitina (en lo sucesivo en este documento denominada O), piridilalanina, tienilalanina, naftilalanina y fenilglicina.

Las secuencias de aminoácidos variantes pueden incluir grupos espaciadores adecuados que se pueden insertar entre dos residuos aminoacídicos cualesquiera de la secuencia, incluyendo grupos alquilo tales como grupos metilo, etilo o propilo además de espaciadores aminoacídicos tales como residuos de glicina o β -alanina. Una forma adicional de variación, que implica la presencia de uno o más residuos aminoacídicos en forma peptoide, puede ser muy conocida por los expertos en la materia. Para evitar dudas, se usa "la forma peptoide" para referirse a los residuos aminoacídicos variantes en los que el grupo sustituyente del α -carbono está en el átomo de nitrógeno del residuo en lugar del α -carbono. Los procesos para preparar péptidos en la forma peptoide son conocidos en la técnica, por ejemplo, Simon RJ *et al.*, PNAS (1992) 89(20), 9367-9371 y Horwell DC, Trends Biotechnol. (1995) 13(4), 132-134.

La práctica de la presente invención emplea, salvo que se indique otra cosa, técnicas convencionales de inmunología, bioquímica, química, biología molecular, microbiología, biología celular, genómica y ADN recombinante, que pertenecen a las habilidades de la técnica. Véase Sambrook, Fritsch y Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2.^a edición (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel,

et al. eds., (1987)); la serie METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M.J. MacPherson, B.D. Hames y G.R. Taylor eds. (1995)), Harlow y Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, y ANIMAL CELL CULTURE (R.I. Freshney, ed. (1987)).

5 En un aspecto, la divulgación proporciona vectores que se usan en la manipulación y optimización de sistemas de CRISPR-Cas.

10 Como se usa en este documento, un "vector" es una herramienta que permite o facilita la transferencia de una entidad de un entorno a otro. Es un replicón, tal como un plásmido, fago o cósmido, en que puede insertarse otro segmento de ADN para lograr la replicación del segmento insertado. En general, un vector tiene capacidad de replicación cuando se asocia con los elementos de control apropiados. En general, el término "vector" se refiere a una molécula de ácido nucleico que puede transportar otro ácido nucleico al que se ha unido. Los vectores incluyen, aunque sin limitación, moléculas de ácido nucleico que son monocatenarias, bicatenarias o parcialmente bicatenarias; moléculas de ácido nucleico que comprenden uno o más extremos libres, sin extremos libres (por ejemplo, circulares); moléculas de ácido nucleico que comprenden ADN, ARN o ambos; y otras variedades de polinucleótidos conocidos en la técnica. Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN bicatenario circular en que se pueden insertar segmentos de ADN adicionales, tal como por técnicas de clonación molecular convencionales. Otro tipo de vector es un vector vírico, en el que hay secuencias de ADN o ARN derivadas de virus presentes en el vector para empaquetado en un virus (por ejemplo, retrovirus, retrovirus defectuosos en la replicación, adenovirus, adenovirus defectuosos en la replicación y virus adenoasociados (AAV)). Los vectores víricos también incluyen polinucleótidos portados por un virus para su transfección en una célula hospedadora. Determinados vectores tienen capacidad de replicación autónoma en una célula hospedadora en que se introducen (por ejemplo, vectores bacterianos que tienen un origen de replicación bacteriano y vectores episómicos de mamífero). Otros vectores (por ejemplo, vectores no episómicos de mamífero) se integran en el genoma de la célula hospedadora tras la introducción en la célula hospedadora, y de este modo se replican junto con el genoma del hospedador. Además, determinados vectores pueden dirigir la expresión de genes a los que están unidos de forma funcional. Dichos vectores se denominan en este documento "vectores de expresión". Vectores de expresión comunes de utilidad en técnicas de ADN recombinante a menudo están en forma de plásmidos.

30 Los vectores de expresión recombinantes pueden comprender un ácido nucleico de la invención en una forma adecuada para la expresión del ácido nucleico en una célula hospedadora, lo que significa que los vectores de expresión recombinantes incluyen uno o más elementos reguladores, que pueden seleccionarse basándose en las células hospedadoras a usar para la expresión, que se unen de forma funcional a la secuencia de ácido nucleico a expresar. Dentro de un vector de expresión recombinante, "unido de forma funcional" pretende indicar que la secuencia de nucleótidos de interés se une a uno o más elementos reguladores de una manera que permite la expresión de la secuencia de nucleótidos (por ejemplo, en un sistema de transcripción/traducción *in vitro* o en una célula hospedadora cuando el vector se introduce en la célula hospedadora). Con respecto a los métodos de recombinación y clonación, se hace mención de la solicitud de patente de Estados Unidos 10/815.730, publicada el 2 de septiembre de 2004 como US 2004-0171156 A1.

40 Aspectos de la divulgación se refieren a vectores bicistrónicos para ARN quimérico y Cas9 mutada. Se prefieren vectores de expresión bicistrónicos para ARN quimérico y Cas9 mutada. En general y particularmente en esta realización, la Cas9 mutada está dirigida preferiblemente por el promotor CBh. El ARN quimérico puede estar dirigido preferiblemente por un promotor U6. Idealmente, los dos se combinan. El ARN guía quimérico típicamente consiste en una secuencia guía de 20 pb (Ns) y esta puede unirse a la secuencia crtra (que va desde la primera "U" de la hebra inferior hasta el final del transcrito). La secuencia crtra puede truncarse en diversas posiciones según se indique. Las secuencias guía y crtra están separadas por la secuencia de acoplamiento de crtra, que puede ser GUUUUAGAGCUA. Esto puede estar seguido de la secuencia de bucle GAAA como se muestra. Estos dos son ejemplos preferidos. Los solicitantes han demostrado indel mediados por Cas9 mutada en los locus *EMX1* y *PVALB* humanos por ensayos SURVEYOR. Los ARNqui se indican por su denominación "+n", y el ARNcr se refiere a un ARN híbrido donde las secuencias guía y crtra se expresan como transcritos separados. Durante toda esta solicitud, el ARN quimérico también puede llamarse guía simple o ARN guía sintético (ARNgs). El bucle es preferiblemente GAAA, pero no se limita a esta secuencia o, de hecho, a que sea solamente de 4 pb de longitud. De hecho, las secuencias que forman bucle preferidas para su uso en estructuras de horquilla son de cuatro nucleótidos de longitud, y muchos más preferiblemente tienen la secuencia GAAA. Sin embargo, pueden usarse secuencias de bucle más largas o más cortas, así como secuencias alternativas. Las secuencias preferiblemente incluyen un triplete de nucleótidos (por ejemplo, AAA), y un nucleótido adicional (por ejemplo C o G). Ejemplos de secuencias que forman bucle incluyen CAAA y AAAG.

60 La expresión "elemento regulador" pretende incluir promotores, potenciadores, sitios internos de entrada al ribosoma (IRES) y otros elementos de control de la expresión (por ejemplo, señales de terminación de la transcripción, tales como señales de poliadenilación y secuencias de poli-U). Dichos elementos reguladores se describen, por ejemplo, en Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Los elementos reguladores incluyen aquellos que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de célula hospedadora y aquellos que dirigen la expresión de la secuencia de nucleótidos únicamente en determinadas células hospedadoras (por ejemplo, secuencias reguladoras específicas de tejido). Un promotor específico de tejido puede dirigir la expresión principalmente en un tejido deseado de interés, tal como

músculo, neurona, hueso, piel, sangre, órganos específicos (por ejemplo, hígado, páncreas), o tipos celulares particulares (por ejemplo, linfocitos). Los elementos reguladores también pueden dirigir la expresión de una manera temporal, tal como de una manera dependiente del ciclo celular o dependiente de la fase de desarrollo, que también puede ser o no específica de tejido o tipo celular. En algunas realizaciones, un vector comprende uno o más promotores de pol III (por ejemplo, 1, 2, 3, 4, 5 o más promotores de pol I), uno o más promotores de pol II (por ejemplo, 1, 2, 3, 4, 5 o más promotores de pol II), uno o más promotores de pol I (por ejemplo, 1, 2, 3, 4, 5 o más promotores de pol I), o combinaciones de los mismos. Ejemplos de promotores de pol III incluyen, aunque sin limitación, promotores U6 y H1. Ejemplos de promotores de pol II incluyen, aunque sin limitación, el promotor retrovírico de LTR del virus del sarcoma de Rous (RSV) (opcionalmente con el potenciador de RSV), el promotor de citomegalovirus (CMV) (opcionalmente con el potenciador de CMV) [véase, por ejemplo, Boshart *et al.*, Cell, 41:521-530 (1985)], el promotor de SV40, el promotor de la dihidrofolato reductasa, el promotor de la β -actina, el promotor de la fosfoglicerol cinasa (PGK) y el promotor de EF1 α . También se engloban por la expresión "elemento regulador", elementos potenciadores, tales como WPRE; potenciadores de CMV; el segmento R-U5' en LTR de HTLV-1 (Mol. Cell. Biol., Vol. 8(1), pág. 466-472, 1988); potenciador de SV40; y la secuencia intrónica entre los exones 2 y 3 de la β -globina de conejo (Proc. Natl. Acad. Sci. USA., Vol. 78(3), pág. 1527-31, 1981). Los expertos en la materia apreciarán que el diseño del vector de expresión puede depender de factores tales como la elección de la célula hospedadora a transformar, el nivel de expresión deseada, etc. Un vector puede introducirse en células hospedadoras para producir de ese modo transcritos, proteínas o péptidos, incluyendo proteínas o péptidos de fusión, codificados por ácidos nucleicos como se describe en este documento (por ejemplo, transcritos, proteínas y enzimas de repeticiones palindrómicas cortas intercaladas de forma regular agrupadas (CRISPR), formas mutantes de los mismos, proteínas de fusión de los mismos, etc.). Con respecto a las secuencias reguladoras, se hace mención de la solicitud de patente de Estados Unidos 10/491.026. Con respecto a los promotores, se hace mención de la publicación PCT WO 2011/028929 y la solicitud de Estados Unidos 12/511.940.

Pueden diseñarse vectores para la expresión de transcritos de CRISPR (por ejemplo, transcritos de ácido nucleico, proteínas o enzimas) en células procarióticas o eucarióticas. Por ejemplo, los transcritos de CRISPR pueden expresarse en células bacterianas tales como *Escherichia coli*, células de insecto (usando vectores de expresión de baculovirus), células de levadura o células de mamífero. Se analizan células hospedadoras adecuadas adicionalmente en Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Como alternativa, el vector de expresión recombinante puede transcribirse y traducirse *in vitro*, por ejemplo, usando secuencias reguladores del promotor T7 y la polimerasa T7.

Los vectores pueden introducirse y propagarse en un procariota o célula eucariótica. En algunas realizaciones, un procariota se usa para amplificar las copias de un vector a introducir en una célula eucariótica o como un vector intermedio en la producción de un vector a introducir en una célula eucariótica (por ejemplo, amplificando un plásmido como parte de un sistema de empaquetado de vector vírico). En algunas realizaciones, un procariota se usa para amplificar las copias de un vector y expresar uno o más ácidos nucleicos, tal como para proporcionar una fuente de una o más proteínas para suministro a una célula hospedadora u organismo hospedador. La expresión de proteínas en procariotas muy a menudo se realiza en *Escherichia coli* con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de proteínas de fusión o que no son de fusión. Los vectores de fusión añaden varios aminoácidos a una proteína codificada en los mismos, tal como al extremo amino de la proteína recombinante. Dichos vectores de fusión pueden cumplir uno o más propósitos, tales como: (i) aumentar la expresión de la proteína recombinante; (ii) aumentar la solubilidad de la proteína recombinante; y (iii) ayudar en la purificación de la proteína recombinante actuando como ligando en la purificación por afinidad. A menudo, en vectores de expresión de fusión, se introduce un sitio de escisión proteolítica en la unión del resto de fusión y la proteína recombinante para posibilitar la separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión. Dichas enzimas, y sus secuencias de reconocimiento análogas, incluyen factor Xa, trombina y enterocinasa. Vectores de expresión de fusión ejemplares incluyen pGEX (Pharmacia Biotech Inc; Smith y Johnson, 1988, Gene 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) y pRIT5 (Pharmacia, Piscataway, N.J.) que fusionan glutatión S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente, a la proteína recombinante diana.

Ejemplos de vectores de expresión en *E. coli* sin fusión inducibles adecuados incluyen pTrc (Amrann *et al.*, (1988) Gene 69:301-315) y pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

En algunas realizaciones, un vector es un vector de expresión de levadura. Ejemplos de vectores para su expresión en la levadura *Saccharomyces cerevisiae* incluyen pYepSec1 (Baldari, *et al.*, 1987, EMBO J. 6: 229-234), pMFa (Kuijan y Herskowitz, 1982, Cell 30: 933-943), pJRY88 (Schultz *et al.*, 1987, Gene 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.) y picZ (Invitrogen Corp, San Diego, Calif.).

En algunas realizaciones, un vector dirige la expresión de proteínas en células de insecto usando vectores de expresión de baculovirus. Los vectores de baculovirus disponibles para la expresión de proteínas en células de insecto cultivadas (por ejemplo, células SF9) incluyen la serie pAc (Smith, *et al.*, 1983, Mol. Cell. Biol. 3: 2156-2165) y la serie pVL (Lucklow y Summers, 1989, Virology 170: 31-39).

En algunas realizaciones, un vector puede dirigir la expresión de una o más secuencias en células de mamífero usando un vector de expresión de mamífero. Ejemplos de vectores de expresión de mamífero incluyen pCDM8 (Seed, 1987. Nature 329: 840) y pMT2PC (Kaufman, *et al.*, 1987. EMBO J. 6: 187-195). Cuando se usa en células de mamífero, las funciones de control del vector de expresión se proporcionan típicamente por uno o más elementos reguladores. Por ejemplo, promotores habitualmente usados se obtienen de polioma, adenovirus 2, citomegalovirus, virus del simio 40 y otros divulgados en este documento y conocidos en la técnica. Para otros sistemas de expresión adecuados tanto para células procarióticas como para células eucarióticas véanse, por ejemplo, los capítulos 16 y 17 de Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2.^a ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

En algunas realizaciones, el vector de expresión de mamífero recombinante puede dirigir la expresión del ácido nucleico preferentemente en un tipo particular (por ejemplo, se usan elementos reguladores específicos de tejido para expresar el ácido nucleico). Los elementos reguladores específicos de tejido son conocidos en la técnica. Ejemplos no limitantes de promotores específicos de tejido adecuados incluyen el promotor de albúmina (específico de hígado; Pinkert, *et al.*, 1987. Genes Dev. 1: 268-277), promotores específicos linfoides (Calame y Eaton, 1988. Adv. Immunol. 43: 235-275), en particular promotores de receptores de linfocitos T (Winoto y Baltimore, 1989. EMBO J. 8: 729-733) e inmunoglobulinas (Baneiji, *et al.*, 1983. Cell 33: 729-740; Queen y Baltimore, 1983. Cell 33: 741-748), promotores específicos de neurona (por ejemplo, el promotor de neurofilamentos; Byrne y Ruddle, 1989. Proc. Natl. Acad. Sci. USA 86: 5473-5477), promotores específicos de páncreas (Edlund, *et al.*, 1985. Science 230: 912-916) y promotores específicos de glándula mamaria (por ejemplo, promotor del suero de la leche; patente de Estados Unidos n.º 4.873.316 y publicación de solicitud europea n.º 264.166). También están englobados promotores regulados por el desarrollo, por ejemplo, los promotores de hox murinos (Kessel y Gruss, 1990. Science 249: 374-379) y el promotor de la α -fetoproteína (Campes y Tilghman, 1989. Genes Dev. 3: 537-546). Con respecto a estos vectores procarióticos y eucarióticos, se hace mención de la patente de Estados Unidos 6.750.059. Otras realizaciones de la invención pueden referirse al uso de vectores víricos, con respecto a lo que se hace mención de la solicitud de patente de Estados Unidos 13/092.085. Los elementos reguladores específicos de tejido son conocidos en la técnica y a este respecto, se hace mención de la patente de Estados Unidos 7.776.321.

En algunas realizaciones, un elemento regulador se une de forma funcional a uno o más elementos de un sistema de CRISPR para dirigir la expresión del uno o más elementos del sistema de CRISPR. En general, los CRISPR (repeticiones palindrómicas cortas intercaladas de forma regular agrupadas), también conocidos como SPIDR (repeticiones directas intercaladas SPacer), constituyen una familia de locus de ADN que son habitualmente específicos para una especie bacteriana particular. El locus de CRISPR comprende una clase distinta de repeticiones de secuencia corta intercaladas (SSR) que se reconocen en *E. coli* (Ishino *et al.*, J. Bacteriol., 169:5429-5433 [1987]; y Nakata *et al.*, J. Bacteriol., 171:3553-3556 [1989]), y genes asociados. Se han identificado SSR intercaladas similares en *Haloflex mediterranei*, *Streptococcus pyogenes*, *Anabaena* y *Mycobacterium tuberculosis* (véase Groenen *et al.*, Mol. Microbiol., 10:1057-1065 [1993]; Hoe *et al.*, Emerg. Infect. Dis., 5:254-263 [1999]; Masepohl *et al.*, Biochim. Biophys. Acta 1307:26-30 [1996]; y Mojica *et al.*, Mol. Microbiol., 17:85-93 [1995]). Los locus de CRISPR típicamente difieren de otras SSR en la estructura de las repeticiones, que se han denominado repeticiones espaciadas de forma regular cortas (SRSR) (Janssen *et al.*, OMICS J. Integ. Biol., 6:23-33 [2002]; y Mojica *et al.*, Mol. Microbiol., 36:244-246 [2000]). En general, las repeticiones son elementos cortos que se producen en grupos que están espaciados de forma regular por secuencias intermedias únicas con una longitud sustancialmente constante (Mojica *et al.*, [2000], *supra*). Aunque las secuencias de repetición están muy conservadas entre las cepas, el número de repeticiones intercaladas y las secuencias de las regiones espaciadoras típicamente difieren de una cepa a otra (van Embden *et al.*, J. Bacteriol., 182:2393-2401 [2000]). Se han identificado locus de CRISPR en más de 40 procariotas (véase, por ejemplo, Jansen *et al.*, Mol. Microbiol., 43:1565-1575 [2002]; y Mojica *et al.*, [2005]) incluyendo, aunque sin limitación, *Aeropyrum*, *Pyrobaculum*, *Sulfolobus*, *Archaeoglobus*, *Halocarcula*, *Methanobacterium*, *Methanococcus*, *Methanosarcina*, *Methanopyrus*, *Pyrococcus*, *Picrophilus*, *Thermoplasma*, *Corynebacterium*, *Mycobacterium*, *Streptomyces*, *Aquifex*, *Porphyromonas*, *Chlorobium*, *Thermus*, *Bacillus*, *Listeria*, *Staphylococcus*, *Clostridium*, *Thermoanaerobacter*, *Mycoplasma*, *Fusobacterium*, *Azarcus*, *Chromobacterium*, *Neisseria*, *Nitrosomonas*, *Desulfovibrio*, *Geobacter*, *Myxococcus*, *Campylobacter*, *Wolinella*, *Acinetobacter*, *Erwinia*, *Escherichia*, *Legionella*, *Methylococcus*, *Pasteurella*, *Photobacterium*, *Salmonella*, *Xanthomonas*, *Yersinia*, *Treponema* y *Thermotoga*.

En general, "sistema de CRISPR" se refiere colectivamente a transcritos y otros elementos implicados en la expresión de o en dirigir la actividad de genes asociados con CRISPR ("Cas"), incluyendo secuencias que codifican un gen de Cas, una secuencia crtra (CRISPR transactivadora) (por ejemplo, ARNcrtra o un ARNcrtra parcial activo), una secuencia de acoplamiento de crtra (que abarca una "repetición directa" y una repetición directa parcial procesada por ARNcrtra en el contexto de un sistema de CRISPR endógeno), una secuencia guía (también denominada "espaciador" en el contexto de un sistema de CRISPR endógeno), u otras secuencias y transcritos de un locus de CRISPR. En realizaciones de la invención, las expresiones secuencia guía y ARN guía se usan indistintamente. En algunas realizaciones, uno o más elementos de un sistema de CRISPR se obtienen de un sistema de CRISPR de tipo I, tipo II o tipo III. En algunas realizaciones, uno o más elementos de un sistema de CRISPR se obtienen de un organismo particular que comprende un sistema de CRISPR endógeno, tal como *Streptococcus pyogenes*. En general, un sistema de CRISPR se caracteriza por elementos que promueven la formación de un complejo de CRISPR en el sitio de una secuencia diana (también denominado proteoespaciador en el contexto de un sistema de CRISPR endógeno). En el contexto de la formación de un complejo de CRISPR, "secuencia diana" se refiere a una secuencia

para la que se diseña una secuencia guía que tenga complementariedad, donde la hibridación entre una secuencia diana y una secuencia guía promueve la formación de un complejo de CRISPR. Una secuencia diana puede comprender cualquier polinucleótido, tal como polinucleótidos de ADN o ARN. En algunas realizaciones, una secuencia diana está ubicada en el núcleo o citoplasma de una célula.

En algunas realizaciones, pueden identificarse repeticiones directas *in silico* buscando motivos repetitivos que cumplan alguno o todos los siguientes criterios:

1. se encuentra en una ventana de 2 Kb de secuencia genómica que flanquea el locus de CRISPR de tipo II;
2. abarca de 20 a 50 pb; y
3. está intercalada por 20 a 50 pb.

En algunas realizaciones, pueden usarse 2 de estos criterios, por ejemplo, 1 y 2, 2 y 3, o 1 y 3. En algunas realizaciones, pueden usarse los 3 criterios.

En algunas realizaciones, el ARNcrtra candidato puede predecirse posteriormente por secuencias que cumplen alguno o todos los siguientes criterios:

1. homología de secuencia con repeticiones directas (búsqueda de motivos en Geneious con emparejamientos incorrectos de hasta 18 pb);
2. presencia de un terminador de la transcripción independiente de Rho predicho en la dirección de la transcripción; y
3. estructura secundaria de horquilla estable entre el ARNcrtra y la repetición directa.

En algunas realizaciones, pueden usarse 2 de estos criterios, por ejemplo, 1 y 2, 2 y 3, o 1 y 3. En algunas realizaciones, pueden usarse los 3 criterios.

En algunas realizaciones, los diseños de ARN guía sintéticos quiméricos (ARNgs) pueden incorporar al menos 12 pb de estructura bicatenaria entre la repetición directa y el ARNcrtra.

En la invención, el sistema de CRISPR es un sistema de CRISPR de tipo II y la enzima Cas es Cas9 mutada (N863A), que cataliza la escisión de ADN. La acción enzimática de Cas9 mutada derivada de *Streptococcus pyogenes* o cualquier Cas9 muy relacionada genera roturas bicatenarias en secuencias de sitio diana que hibridan con 20 nucleótidos de la secuencia guía y que tienen una secuencia de motivo adyacente de protoespaciador (PAM) (ejemplos incluyen NGG/NRG o un PAM que puede determinarse como se describe en este documento) después de los 20 nucleótidos de la secuencia diana. La actividad de CRISPR a través de Cas9 mutada para reconocimiento de ADN específico de sitio y escisión se define por la secuencia guía, la secuencia crtra que hibrida en parte con la secuencia guía y la secuencia PAM. Más aspectos del sistema de CRISPR se describen en Karginov y Hannon, *The CRISPR system: small RNA-guided defence in bacteria and archaea*, *Mole Cell* 2010, enero 15; 37(1): 7.

En algunas realizaciones, un vector comprende uno o más sitios de inserción, tal como una secuencia de reconocimiento de endonucleasa de restricción (también denominada "sitios de clonación"). En algunas realizaciones, uno o más sitios de inserción (por ejemplo, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, o más sitios de inserción) están ubicados anteriores y/o posteriores a uno o más elementos de secuencia de uno o más vectores. En algunas realizaciones, un vector comprende un sitio de inserción anterior a la secuencia de acoplamiento de crtra, y opcionalmente posterior a un elemento regulador unidor de forma funcional a la secuencia de acoplamiento de crtra, de modo que después de la inserción de una secuencia guía en el sitio de inserción y tras la expresión, la secuencia guía dirige la unión específica de secuencia de un complejo de CRISPR a una secuencia diana en una célula eucariótica. En algunas realizaciones, un vector comprende dos o más sitios de inserción, estando ubicado cada sitio de inserción entre dos secuencias de acoplamiento de crtra para permitir la inserción de una secuencia guía en cada sitio. En dicha disposición, las dos o más secuencias guía pueden comprender dos o más copias de una secuencia guía, dos o más secuencia guía diferentes, o combinaciones de estas. Cuando se usan múltiples secuencias guía diferentes, puede usarse una sola construcción de expresión para dirigir la actividad de CRISPR a múltiples secuencias diana correspondientes diferentes dentro de una célula. Por ejemplo, un solo vector puede comprender aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, o más secuencias guía. En algunas realizaciones, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, o más de dichos vectores que contienen secuencia guía pueden proporcionarse, y opcionalmente suministrarse a una célula.

En algunas realizaciones, una secuencia que codifica enzima que codifica una enzima de CRISPR tiene los codones optimizados para la expresión en células particulares, tales como células eucarióticas. Las células eucarióticas pueden ser aquellas de o derivadas de un organismo particular, tal como un mamífero incluyendo, aunque sin limitación, ser humano, ratón, rata, conejo, perro o mamífero no humano o primate. En algunas realizaciones, los procesos para modificar la identidad genética de la línea germinar de seres human y/o procesos para modificar la identidad genética

de animales que tienen probabilidad de causarles enfermedad sin ningún beneficio médico sustancial al ser humano o animal, y también animales resultantes de dichos procesos, pueden excluirse.

En general, la optimización de codones se refiere a un proceso de modificación de una secuencia de ácido nucleico para la expresión potenciada en las células hospedadoras de interés reemplazando al menos un codón (por ejemplo, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, o más codones) de la secuencia natural con codones que se usan más frecuentemente o mucho más frecuentemente en los genes de esa célula hospedadora mientras se mantiene la secuencia de aminoácidos natural. Diversas especies muestran un sesgo particular para determinados codones de un aminoácido particular. El sesgo de codones (diferencias en la utilización de codones entre los organismos) a menudo se correlaciona con la eficacia de traducción del ARN mensajero (ARNm), que a su vez se cree que depende de, entre otras cosas, las propiedades de los codones que se están traduciendo y la disponibilidad de moléculas de ARN de transferencia (ARNt) particulares. En general, la predominancia de los ARNt seleccionados en una célula refleja los codones usados con mayor frecuencia en la síntesis de péptidos. En consecuencia, se pueden personalizar los genes para una expresión génica óptima en un organismo dado basándose en la optimización de los codones. Las tablas de utilización de codones están fácilmente disponibles, por ejemplo, en la "Codon Usage Database" disponible en www.kazusa.or.jp/codon/ (visitada el 9 de julio de 2002), y estas tablas pueden adaptarse de varias maneras. Véase, Nakamura, Y., *et al.* "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000). También están disponibles algoritmos informáticos para optimizar los codones de una secuencia particular para su expresión en una célula hospedadora particular, tales como Gene Forge (Aptagen; Jacobus, PA). En algunas realizaciones, uno o más codones (por ejemplo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, o más, o todos los codones) en una secuencia que codifica una enzima de CRISPR corresponden al codón usado más frecuentemente para un aminoácido particular.

En algunas realizaciones, un vector codifica una enzima de CRISPR mutada que comprende una o dos o más secuencias de localización nuclear (NLS), tal como aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más NLS. En algunas realizaciones, la enzima de CRISPR mutada comprende aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, o más NLS en o cerca del extremo amino, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, o más NLS en o cerca del extremo carboxi, o una combinación de estas (por ejemplo, una o más NLS en el extremo amino y una o más NLS en el extremo carboxi). Cuando hay más de una NLS presente, cada una puede seleccionarse independientemente de las otras, de modo que una sola NLS puede estar presente en más de una copia y/o en combinación con uno o más NLS distintos presentes en una o más copias. En una realización preferida de la invención, la enzima de CRISPR mutada comprende como mucho 6 NLS. En algunas realizaciones, una NLS se considera cerca del extremo N o C cuando el aminoácido más cercano de la NLS está en aproximadamente 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, o más aminoácidos a lo largo de la cadena polipeptídica desde el extremo N o C. Ejemplos no limitantes de NLS incluyen una secuencia NLS derivada de: la NLS del antígeno T grande del virus SV40, que tiene la secuencia de aminoácidos PKKKRKV; la NLS de nucleoplasmina (por ejemplo, la NLS bipartita de nucleoplasmina con la secuencia KRPAATKKAGQAKKKK); la NLS de c-myc que tiene la secuencias de aminoácidos PAAKRVKLD o RQRRELKRSP; la NLS de HRNPA1 M9 que tiene la secuencia NQSSNFGPMKGGNFGGRSSGYPYGGGGQYFAKPRNQGGY; la secuencia RMRIZFKNKGKDTAELRRRVEVSVELRKAKKDEQILKRRNV del dominio IBB de importina-alfa; las secuencias VSRKRPRP y PPKKARED de la proteína T de mioma; la secuencia POPKKKPL de p53 humano; la secuencia SALIKKKKMAP de c-abl IV de ratón; las secuencias DRLRR y PKQKKRK de NS1 del virus de la gripe; la secuencia RKLKKIKKL del antígeno delta del virus de la hepatitis; la secuencia REKKKFLKRR de la proteína Mx1 de ratón; la secuencia KRKGDEVDGVEVAKKKSKK de la poli(ADP-ribosa) polimerasa humana; y la secuencia RKCLQAGMNLEARKTKK de los receptores de hormonas esteroideas de glucocorticoesteroides (humanos).

En general, la una o más NLS son de suficiente fuerza para dirigir la acumulación de la enzima de CRISPR mutada en una cantidad detectable en el núcleo de una célula eucariótica. En general, la fuerza de la actividad de localización nuclear puede derivar del número de NLS en la enzima de CRISPR, la o las NLS particulares usadas, o una combinación de estos factores. La detección de acumulación en el núcleo puede realizarse por cualquier técnica adecuada. Por ejemplo, un marcador detectable puede fusionarse a la enzima de CRISPR mutada, de modo que la localización dentro de una célula pueda visualizarse, tal como en combinación con un medio para detectar la localización del núcleo (por ejemplo, un tinte específico para el núcleo tal como DAPI). Los núcleos celulares pueden aislarse de las células, cuyos contenidos después pueden analizarse por cualquier proceso adecuado para detectar proteínas, tal como inmunohistoquímica, transferencia de Western o ensayo de actividad enzimática. La acumulación en el núcleo también puede determinarse indirectamente, tal como mediante un ensayo para el efecto de formación del complejo de CRISPR (por ejemplo, ensayo para la escisión de ADN o mutación en la secuencia diana, o ensayo para actividad de expresión génica alterada influida por la formación del complejo de CRISPR y/o la actividad de la enzima de CRISPR), en comparación con un control no expuesto a la enzima de CRISPR mutada o complejo, o expuesto a una enzima de CRISPR mutada que carece de la una o más NLS.

Se prefiere, y esto puede aplicarse a cualquiera de los aspectos o realizaciones de la invención, que los métodos, sistemas y composiciones descritos en este documento no incluyan una NLS. En otras palabras, aunque en general es útil, el uso de una NLS puede ser opcional. Por ejemplo, si tiene que abordarse ADN mitocondrial, entonces no se requiere una NLS.

En general, una secuencia guía es cualquier secuencia polinucleotídica que tenga suficiente complementariedad con una secuencia de polinucleótido diana para hibridar con la secuencia diana y dirigir la unión específica de secuencia de un complejo de CRISPR a la secuencia diana. En algunas realizaciones, el grado de complementariedad entre una secuencia guía y su secuencia diana correspondiente, cuando se alinean de forma óptima usando un algoritmo de alineación adecuado, es de aproximadamente o más de aproximadamente un 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97,5%, 99%, o más. La alineación óptima puede determinarse con el uso de cualquier algoritmo adecuado para alinear secuencias, cuyos ejemplos no limitantes incluyen el algoritmo de Smith y Waterman, el algoritmo de Needleman y Wunsch, algoritmos basados en la transformación de Burrows y Wheeler (por ejemplo, el alineador de Burrows y Wheeler), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; disponible en www.novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (disponible en soap.genomics.org.cn) y Maq (disponible en maq.sourceforge.net). En algunas realizaciones, una secuencia guía es de aproximadamente o más de aproximadamente 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, o más nucleótidos de longitud. En algunas realizaciones, una secuencia guía es de menos de aproximadamente 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, o menos nucleótidos de longitud. La capacidad de una secuencia guía de dirigir la unión específica de secuencia de un complejo de CRISPR a una secuencia diana puede evaluarse por cualquier ensayo adecuado. Por ejemplo, los componentes de un sistema de CRISPR suficientes para formar un complejo de CRISPR, incluyendo la secuencia guía a ensayar, pueden proporcionarse a una célula hospedadora que tenga la secuencia diana correspondiente, tal como por transfección con vectores que codifican los componentes de la secuencia de CRISPR, seguida de una evaluación de la escisión preferente dentro de la secuencia diana, tal como por ensayo Surveyor como se describe en este documento. Asimismo, la escisión de una secuencia de polinucleótido diana puede evaluarse en un tubo de ensayo proporcionando la secuencia diana, componentes de un complejo de CRISPR, incluyendo la secuencia guía a ensayar y una secuencia guía de control diferente de la secuencia guía de ensayo, y comparando la unión o tasa de escisión en la secuencia diana entre las reacciones de secuencia guía de ensayo y control. Otros ensayos son posibles, y se les ocurrirá a los expertos en la materia.

Como se menciona anteriormente, la longitud de guía óptima para Sa puede ser, en algunas realizaciones, de 20 o más preferiblemente, 21, 22, 23 o 24 nucleótidos.

Una secuencia guía puede seleccionarse para abordar cualquier secuencia diana. En algunas realizaciones, la secuencia diana es una secuencia dentro de un genoma de una célula. Secuencia diana ejemplares incluyen las que son únicas en el genoma diana. Por ejemplo, para la Cas9 de *S. pyogenes*, una secuencia diana única en un genoma puede incluir un sitio diana de Cas9 de la forma MMMMMMMMNNNNNNNNNNNNXGG donde NNNNNNNNNNNNNXGG (N es A, G, T o C; y X puede ser cualquiera) aparece una sola vez en el genoma. Una secuencia diana única en un genoma puede incluir un sitio diana de Cas9 de *S. pyogenes* de la forma MMMMMMMMNNNNNNNNNNNNXGG donde NNNNNNNNNNNXGG (N es A, G, T o C; y X puede ser cualquiera) aparece una sola vez en el genoma. Para la Cas9 de CRISPR1 de *S. thermophilus*, una secuencia diana única en un genoma puede incluir un sitio diana de Cas9 de la forma MMMMMMMMNNNNNNNNNNNNXXAGAAW donde NNNNNNNNNNNXXAGAAW (N es A, G, T o C; X puede ser cualquiera; y W es A o T) aparece una sola vez en el genoma. Una secuencia diana única en un genoma puede incluir un sitio diana de Cas9 de CRISPR1 de *S. thermophilus* de la forma MMMMMMMMNNNNNNNNNNNNXXAGAAW donde NNNNNNNNNNNXXAGAAW (N es A, G, T o C; X puede ser cualquiera; y W es A o T) aparece una sola vez en el genoma. Para la Cas9 de *S. pyogenes*, una secuencia diana única en un genoma puede incluir un sitio diana de Cas9 de la forma MMMMMMMMNNNNNNNNNNNNXGGXG donde NNNNNNNNNNNXGGXG (N es A, G, T o C; y X puede ser cualquiera) aparece una sola vez en el genoma. Una secuencia diana única en un genoma puede incluir un sitio diana de Cas9 de *S. pyogenes* de la forma MMMMMMMMNNNNNNNNNNNNXGGXG donde NNNNNNNNNNNXGGXG (N es A, G, T o C; y X puede ser cualquiera) aparece una sola vez en el genoma. En cada una de estas secuencias, "M" puede ser A, G, T o C, y no tiene que considerarse en la identificación de una secuencia como única.

En algunas realizaciones, una secuencia guía se selecciona para reducir el grado de estructura secundaria dentro de la secuencia guía. En algunas realizaciones, aproximadamente o menos de aproximadamente un 75%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 1%, o menos de los nucleótidos de la secuencia guía participan en el emparejamiento de bases autocomplementarias cuando se pliega de forma óptima. El plegamiento óptimo puede determinarse mediante cualquier algoritmo de plegamiento polinucleotídico adecuado. Algunos programas se basan en el cálculo de la energía libre de Gibbs mínima. Un ejemplo de uno de dichos algoritmos es mFold, como se describe por Zuker y Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Otro algoritmo de plegamiento ejemplar es el servidor web en línea RNAfold, desarrollado en el Institute for Theoretical Chemistry en la Universidad de Viena, usando el algoritmo de predicción de estructuras centroides (véase, por ejemplo, A.R. Gruber *et al.*, 2008, Cell 106(1): 23-24; y PA Carr y GM Church, 2009, Nature Biotechnology 27(12): 1151-62).

En general, una secuencia de acoplamiento de crtra incluye cualquier secuencia que tenga suficiente complementariedad con una secuencia crtra para promover una o más de: (1) escisión de una secuencia guía flanqueada por secuencias de acoplamiento de crtra en una célula que contiene la secuencia crtra correspondiente; y (2) formación de un complejo de CRISPR en una secuencia diana, en la que el complejo de CRISPR comprende la secuencia de acoplamiento de crtra hibridada con la secuencia crtra. En general, el grado de complementariedad es con referencia a la alineación óptima de la secuencia de acoplamiento de crtra y la secuencia crtra, por toda la longitud de la más corta de las dos secuencias. La alineación óptima puede determinarse mediante cualquier algoritmo de

alineación adecuado, y puede tener en cuenta además las estructuras secundarias, tales como autocomplementariedad dentro de la secuencia crtra o secuencia de acoplamiento de crtra. En algunas realizaciones, el grado de complementariedad entre la secuencia crtra y la secuencia de acoplamiento de crtra por toda la longitud de la más corta de las dos cuando se alinean de forma óptima es de aproximadamente o más de aproximadamente un 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97,5%, 99%, o más. En algunas realizaciones, una secuencia crtra es de aproximadamente o más de aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, o más nucleótidos de longitud. En algunas realizaciones, la secuencia crtra y la secuencia de acoplamiento de crtra están contenidas dentro de un solo transcrito, de modo que la hibridación entre las dos produzca un transcrito que tenga una estructura secundaria, tal como una horquilla. En una realización de la invención, el transcrito o la secuencia polinucleotídica transcrita tiene al menos dos o más horquillas. En realizaciones preferidas, el transcrito tiene dos, tres, cuatro o cinco horquillas. En una realización adicional de la invención, el transcrito tiene como mucho cinco horquillas. En una estructura de horquilla, la parte de la secuencia 5' de la "N" final y anterior al bucle corresponde a la secuencia de acoplamiento de crtra, y la parte de la secuencia 3' del bucle corresponde a la secuencia crtra. Ejemplos no limitantes adicionales de polinucleótidos individuales que comprenden una secuencia guía, una secuencia de acoplamiento de crtra y una secuencia crtra son los siguientes (enumerados de 5' a 3'), donde "N" representa una base de una secuencia guía, el primer bloque de letras en minúscula representa la secuencia de acoplamiento de crtra y el segundo bloque de letras en minúscula representa la secuencia crtra, y la secuencia de poli-T final representa el terminador de transcripción: (1)

NNNNNNNNNNNNNNNNNNNNgtttttgtactctcaagatttGAAAtaaatcttcagagaagctacaaagataa
ggctctatgccgaaatacacaccctgtcattttatggcagggtgttttcgatttttaTTTTTT; (2)

NNNNNNNNNNNNNNNNNNNNgtttttgtactctcaGAAAtgcagaagctacaaagataaggctcatgccg
aaatcaacaccctgtcattttatggcagggtgttttcgatttttaTTTTTT; (3)

NNNNNNNNNNNNNNNNNNNNgtttttgtactctcaGAAAtgcagaagctacaaagataaggctcatgccg
aaatcaacaccctgtcattttatggcagggtgtTTTTTT; (4)

NNNNNNNNNNNNNNNNNNNNgttttagagctGAAAtagcaagttaaataaggctagtcggtatcaact
gaaaaagtgccacccgagtcggtgcTTTTTT; (5)

NNNNNNNNNNNNNNNNNNNNgttttagagctGAAATAGcaagttaaataaggctagtcggtatcaacttgaaaaagtgTTTTTT; y (6)
NNNNNNNNNNNNNNNNNNNNgttttagagctagAAATAGcaagttaaataaggctagtcggtatcaTTTTTT. En algunas realizaciones, las secuencias (1) a (3) se usan en combinación con Cas9 de CRISPR1 de *S. thermophilus*. En algunas realizaciones, las secuencias (4) a (6) se usan en combinación con Cas9 mutada de *S. pyogenes*. En algunas realizaciones, la secuencia crtra es un transcrito separado de un transcrito que comprende la secuencia de acoplamiento de crtra.

En algunas realizaciones, también se proporciona un molde de recombinación. Un molde de recombinación puede ser un componente de otro vector como se describe en este documento, contenido en un vector separado, o proporcionado como un polinucleótido separado. En algunas realizaciones, un molde de recombinación se diseña para que sirva como molde en recombinación homóloga, tal como dentro o cerca de una secuencia diana mellada por una enzima de CRISPR mutada como parte de un complejo de CRISPR. Un polinucleótido molde puede ser de cualquier longitud adecuada, tal como de aproximadamente o más de aproximadamente 10, 15, 20, 25, 50, 75, 100, 150, 200, 500, 1000, o más nucleótidos de longitud. En algunas realizaciones, el polinucleótido molde es complementario a una parte de un polinucleótido que comprende la secuencia diana. Cuando se alinea de forma óptima, un polinucleótido molde podría solapar con uno o más nucleótidos de una secuencia diana (por ejemplo, aproximadamente o más de aproximadamente 1, 5, 10, 15, 20, o más nucleótidos). En algunas realizaciones, cuando una secuencia molde y un polinucleótido que comprende una secuencia diana se alinean de forma óptima, el nucleótido más cercano del polinucleótido molde está dentro de aproximadamente 1, 5, 10, 15, 20, 25, 50, 75, 100, 200, 300, 400, 500, 1000, 5000, 10000, o más nucleótidos de la secuencia diana.

En algunas realizaciones, la enzima de CRISPR mutada es parte de una proteína de fusión que comprende uno o más dominios proteínicos heterólogos (por ejemplo, aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más dominios además de la enzima de CRISPR mutada). Una proteína de fusión de enzima de CRISPR mutada puede comprender cualquier secuencia proteínica adicional, y opcionalmente una secuencia conectora entre dos dominios cualesquiera. Ejemplos de dominios proteínicos que pueden fusionarse con una enzima de CRISPR mutada incluyen, sin limitación, marcas epitópicas, secuencias génicas indicadoras y dominios proteínicos que tienen una o más de las siguientes actividades: actividad metilasa, actividad desmetilasa, actividad de activación de la transcripción, actividad de represión de la transcripción, actividad de factor liberador de la transcripción, actividad de modificación de histonas, actividad de escisión de ARN y actividad de unión a ácido nucleico. Ejemplos no limitantes de marcas epitópicas incluyen marcas de histidina (His), marcas V5, marcas FLAG, marcas de hemaglutinina de la gripe (HA), marcas de Myc, marcas de VSV-G y marcas de tiorredoxina (Trx). Ejemplos de genes indicadores incluyen, aunque sin limitación, glutatión-S-transferasa (GST), peroxidasa de rábano rústico (HRP), cloranfenicol acetiltransferasa (CAT) beta-galactosidasa, beta-glucuronidasa, luciferasa, proteína fluorescente verde (GFP), HcRed, DsRed, proteína fluorescente verde azulada (CFP), proteína fluorescente amarilla (YFP) y proteínas autofluorescentes incluyendo proteína fluorescente azul (BFP). Una enzima de CRISPR mutada puede fusionarse con una secuencia génica que codifica una proteína o un fragmento de una proteína que se une a moléculas de ADN o se une a otras moléculas celulares incluyendo, aunque sin limitación, proteína de unión a maltosa (MBP), marca de S, fusiones de dominio de unión a ADN (DBD) de Lex A, fusiones de dominio de unión a ADN de GAL4 y fusiones de proteína BP16 del virus del herpes simple (HSV). Dominios adicionales que pueden formar parte de una proteína de fusión que comprende una

enzima de CRISPR mutada se describen en el documento US20110059502. En algunas realizaciones, una enzima de CRISPR mutada marcada se usa para identificar la ubicación de una secuencia diana.

En algunas realizaciones, una enzima de CRISPR mutada puede formar un componente de un sistema inducible. La naturaleza inducible del sistema permitiría el control espaciotemporal de la edición génica o expresión génica usando una forma de energía. La forma de energía puede incluir, aunque sin limitación, radiación electromagnética, energía sónica, energía química y energía térmica. Ejemplos de sistemas inducibles incluyen promotores inducibles por tetraciclina (Tet-On o Tet-Off), sistemas de activación de la transcripción de doble híbrido de molécula pequeña (FKBP, ABA, etc.) o sistemas inducibles por la luz (fitocromo, dominios LOV o criptocromo). En una realización, la enzima de CRISPR mutada puede ser parte de un efector transcripcional inducible por la luz (LITE) para dirigir cambios en la actividad transcripcional de una manera específica de secuencia. Los componentes de un de luz pueden incluir una enzima de CRISPR, un heterodímero de citocromo sensible a la luz (por ejemplo, de *Arabidopsis thaliana*) y un dominio de activación/represión transcripcional. Ejemplos adicionales de proteínas de unión a ADN inducibles y métodos para su uso se proporcionan en patentes y solicitudes de patente citadas en este documento, tales como US8889418, US8895308, US20140186919, US20140242700, US20140273234, US20140335620, WO2014093635A1, WO2014093635A9 y US 61/736465 y US 61/721.283.

Como se analiza en este documento, en algunos aspectos, la divulgación proporciona métodos que comprenden suministrar uno o más polinucleótidos, tal como uno o más vectores como se describe en este documento, uno o más transcritos de los mismos, y/o una o más proteínas transcritas a partir de los mismos, a una célula hospedadora. En algunos aspectos, la divulgación proporciona además células producidas por dichos métodos, y animales que comprenden o se producen a partir de dichas células. En algunas realizaciones, una enzima de CRISPR mutada en combinación con (y opcionalmente en complejo con) una secuencia guía se suministra a una célula. Pueden usarse métodos convencionales de transferencia génica víricos y no víricos para introducir ácidos nucleicos en células de mamífero o tejidos diana. Dichos métodos pueden usarse para administrar ácidos nucleicos que codifican componentes de un sistema de CRISPR a células en cultivo, o en un organismo hospedador. Los sistemas de suministro de vector no víricos incluyen plásmidos de ADN, ARN (por ejemplo, un transcrito de un vector descrito en este documento), ácido nucleico desnudo y ácido nucleico en complejo con un vehículo de suministro, tal como un liposoma. Los sistemas de suministro de vector víricos incluyen virus de ADN y ARN, que tienen los genomas episómicos o integrados después del suministro a la célula. Para una revisión de procedimientos de genoterapia, véase Anderson, *Science* 256:808-813 (1992); Nabel y Felgner, *TIBTECH* 11:211-217 (1993); Mitani y Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer y Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada *et al.*, en *Current Topics in Microbiology and Immunology* Doerfler y Böhm (eds) (1995); y Yu *et al.*, *Gene Therapy* 1:13-26 (1994). Como se analiza en este documento, los métodos de suministro no vírico de ácidos nucleicos incluyen electroporación, lipofección, microinyección, biolística, virosomas, liposomas, inmunoliposomas, conjugados de polimerización o lípido:ácido nucleico, ADN desnudo, viriones artificiales y captación potenciada por agente de ADN. La lipofección se describe en, por ejemplo, las patentes de Estados Unidos n.º 5.049.386, 4.946.787; y 4.897.355) y los reactivos de lipofección se venden en el mercado (por ejemplo, Transfectam™ y Lipofectin™). Los lípidos catiónicos y neutros que son adecuados para lipofección eficaz por reconocimiento de receptor de polinucleótidos incluyen los de Felgner, documento WO 91/17424; documento WO 91/16024. El suministro puede ser a células (por ejemplo, administración *in vitro* o *ex vivo*) o tejidos diana (por ejemplo, administración *in vivo*). La preparación de complejos de lípido:ácido nucleico, incluyendo liposomas dirigidos tales como complejos inmunolipídicos, es bien conocida por los expertos en la materia (véase, por ejemplo, Crystal, *Science* 270:404-410 (1995); Blaese *et al.*, *Cancer Gene Ther.* 2:291-297 (1995); Behr *et al.*, *Bioconjugate Chem.* 5:382-389 (1994); Remy *et al.*, *Bioconjugate Chem.* 5:647-654 (1994); Gao *et al.*, *Gene Therapy* 2:710-722 (1995); Ahmad *et al.*, *Cancer Res.* 52:4817-4820 (1992); patentes de Estados Unidos n.º 4.186.183, 4.217.344, 4.235.871, 4.261.975, 4.485.054, 4.501.728, 4.774.085, 4.837.028 y 4.946.787). Como se analiza en este documento, la invención puede implicar sistemas de suministro víricos de ARN o ADN. El uso de sistemas víricos de ARN o ADN para el suministro de ácidos nucleicos aprovechar los procesos altamente evolucionados de dirigir un virus a células específicas en el cuerpo y trasladar la carga vírica al núcleo. Los vectores víricos pueden administrarse directamente a pacientes (*in vivo*) o pueden usarse para tratar células *in vitro*, y las células modificadas opcionalmente pueden administrarse a los pacientes (*ex vivo*). Los sistemas víricos convencionales podrían incluir vectores retrovíricos, lentivíricos, adenovíricos, de virus adenoasociado y del herpes simple para transferencia génica. La integración en el genoma hospedador es posible con los métodos de transferencia génica de retrovirus, lentivirus y virus adenoasociado, produciendo a menudo expresión a largo plazo del transgén insertado. Además, se han observado altas eficacias de transducción en muchos tipos celulares diferentes y tejidos diana. El tropismo de un retrovirus puede alterarse incorporando proteínas de envoltura exógenas, expandiendo la población diana potencial de células diana. Los vectores lentivíricos son vectores retrovíricos que pueden transducir o infectar células que no están en división y típicamente producen altos valores cuantitativos de virus. La selección de un sistema de transferencia génica retrovírico, por lo tanto, dependería del tejido diana. Los vectores retrovíricos están compuestos de repeticiones terminales largas de acción en *cis* con capacidad de empaquetado de hasta 6-10 kb de secuencia exógena. Las LTR de acción en *cis* mínimas son suficientes para la replicación y empaquetado de los vectores, que después se usan para integrar el gen terapéutico en la célula diana para proporcionar expresión transgénica permanente. Los vectores retrovíricos ampliamente usados incluyen los basados en el virus de la leucemia murina (MuLV), el virus de la leucemia del gibón (GaLV), el virus de la inmunodeficiencia del simio (VIS), el virus de la

inmunodeficiencia humana (VIH) y combinaciones de los mismos (véase, por ejemplo, Buchscher *et al.*, J. Virol. 66:2731-2739 (1992); Johann *et al.*, J. Virol. 66:1635-1640 (1992); Sommnerfelt *et al.*, Virol. 176:58-59 (1990); Wilson *et al.*, J. Virol. 63:2374-2378 (1989); Miller *et al.*, J. Virol. 65:2220-2224 (1991); documento PCT/US94/05700). En aplicaciones donde se prefiere expresión transitoria, pueden usarse sistemas adenovíricos. Los vectores adenovíricos tienen capacidad de eficacia de transducción muy alta en muchos tipos celulares y no requieren división celular. Con dichos vectores, se han obtenido altos valores cuantitativos y niveles de expresión. Este vector puede producirse en grandes cantidades en un sistema relativamente simple. Los vectores víricos adenoasociados ("AAV") también pueden usarse para transducir células con ácidos nucleicos diana, por ejemplo, en la producción *in vitro* de ácidos nucleicos y péptidos, y para procedimientos de genoterapia *in vivo* y *ex vivo* (véase, por ejemplo, West *et al.*, Virology 160:38-47 (1987); patente de Estados Unidos n.º 4.797.368; documento WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest. 94:1351 (1994). La construcción de vectores AAV recombinantes se describe en varias publicaciones, incluyendo la patente de Estados Unidos n.º 5.173.414; Tratschin *et al.*, Mol. Cell. Biol. 5:3251-3260 (1985); Tratschin, *et al.*, Mol. Cell. Biol. 4:2072-2081 (1984); Hermonat y Muzyczka, PNAS 81:6466-6470 (1984); y Samulski *et al.*, J. Virol. 63:03822-3828 (1989). Típicamente se usan células de empaquetado para formar partículas víricas que puedan infectar una célula hospedadora. Dichas células incluyen células 293, que empaquetan adenovirus, y células ψ 2 o células PA317, que empaquetan retrovirus. Los vectores víricos usados en genoterapia se generan habitualmente por una línea celular productora que empaqueta un vector de ácido nucleico en una partícula vírica. Los vectores típicamente contienen las secuencias víricas mínimas requeridas para el empaquetado y posterior integración en un hospedador, remplazándose otras secuencias víricas por un casete de expresión para el uno o más polinucleótidos a expresar. Las funciones víricas ausentes típicamente se aportan en *trans* por la línea celular de empaquetado. Por ejemplo, los vectores AAV usados en genoterapia típicamente poseen solamente secuencias ITR del genoma de AAV que se requieren para el empaquetado e integración en el genoma hospedador. El ADN vírico se empaqueta en una línea celular, que contiene un plásmido auxiliar que codifica los otros genes de AAV, concretamente rep y cap, pero que carece de las secuencias ITR. La línea celular también puede infectarse con adenovirus como auxiliar. El virus auxiliar promueve la replicación del vector AAV y la expresión de genes de AAV desde el plásmido auxiliar. El plásmido auxiliar no se empaqueta en cantidades significativas debido a la ausencia de secuencias ITR. La contaminación con adenovirus puede reducirse, por ejemplo, por tratamiento térmico al que el adenovirus es más sensible que el AAV. Por consiguiente, el AAV se considera un candidato ventajoso para su uso como vector de transducción. Dichos vectores de transducción de AAV pueden comprender suficientes funciones de acción en *cis* para replicarse en presencia de funciones auxiliares de adenovirus o herpesvirus o poxvirus (por ejemplo, virus de la variolovacuna) proporcionadas en *trans*. Puede usarse AAV recombinante (rAAV) para portar genes exógenos a las células de una diversidad de linajes. En estos vectores, los genes de cap y/o rep de AAV se eliminan del genoma vírico y se remplazan con un segmento de ADN de elección. Los vectores AAV actuales pueden alojar hasta 4300 bases de ADN insertado. Hay varias maneras de producir rAAV, y la invención proporciona rAAV y métodos para preparar rAAV. Por ejemplo, uno o más plásmidos que contienen o consisten esencialmente en la construcción vírica deseada se transfecan en células infectadas con AAV. Además, se cotransfecta un segundo plásmido auxiliar o adicional en estas células para proporcionar los genes de rep y/o cap de AAV que son obligatorios para la replicación y empaquetado de la construcción vírica recombinante. En estas condiciones, las proteínas rep y/o cap de AAV actúan en *trans* para estimula la replicación y el empaquetado de la construcción de rAAV. De dos a tres días después de la transfección, se recoge rAAV. Tradicionalmente, el rAAV se recoge de las células junto con adenovirus. El adenovirus contaminante entonces se inactiva por tratamiento térmico. En la presente divulgación, el rAAV se recoge ventajosamente no de las propias células, sino del sobrenadante celular. Por consiguiente, en un aspecto inicial, la divulgación proporciona la preparación de rAAV y, además de los anterior, el rAAV puede prepararse por un método que comprende o consiste esencialmente en: infectar células susceptibles con un rAAV que contiene ADN exógeno que incluye ADN para su expresión, y virus auxiliar (por ejemplo, adenovirus, herpesvirus, poxvirus tal como virus de la variolovacuna), en el que el rAAV carece de cap y/o rep funcionales (y el virus auxiliar (por ejemplo, adenovirus, herpesvirus, poxvirus tal como virus de la variolovacuna) proporciona la función cap y/o rev de la que carece rAAV); o infectar células susceptibles con un rAAV que contiene ADN exógeno que incluye ADN para su expresión, en el que el recombinante carece de cap y/o rep funcionales, y transfectar dichas células con un plásmido que aporta la función cap y/o rep de la que carece el rAAV; o infectar células susceptibles con un rAAV que contiene ADN exógeno que incluye ADN para su expresión, en el que el recombinante carece de cap y/o rep funcionales, en el que dichas células aportan la función cap y/o rep de la que carece el recombinante; o transfectar las células susceptibles con un AAV que carece de cap y/o rep funcionales y plásmidos para insertar ADN exógeno en el recombinante de modo que el ADN exógeno se exprese por el recombinante y para aportar las funciones rep y/o cap, por lo que la transfección produce un rAAV que contiene el ADN exógeno que incluye ADN para su expresión que carece de cap y/o rep funcionales. El rAAV puede ser un AAV como se describe en este documento, y ventajosamente puede ser un rAAV1, rAAV2, AAV5 o rAAV que tiene híbrido o cápside que puede comprender AAV1, AAV2, AAV5 o cualquier combinación de los mismos. Se puede seleccionar el AAV del rAAV con respecto a las células a abordar por el rAAV; por ejemplo, se pueden seleccionar los serotipos de AAV 1, 2, 5 o un híbrido o cápside de AAV1, AAV2, AAV5 o cualquier combinación de los mismos para abordar células cerebrales o neuronales; y se puede seleccionar AAV4 para abordar tejido cardíaco. La divulgación proporciona rAAV que contiene o consiste esencialmente en una molécula de ácido nucleico exógena que codifica un sistema de CRISPR (repeticiones palindrómicas cortas intercaladas de forma regular agrupadas), por ejemplo, una pluralidad de casetes que comprenden o consisten en un primer casete que comprende o consisten esencialmente en un promotor, una molécula de ácido nucleico que codifica una proteína asociada a CRISPR (Cas) mutada (proteínas nucleasa o helicasa putativas), por ejemplo, Cas9 mutada (N863A) y un terminador, y uno o dos, o más, ventajosamente hasta el límite de tamaño de empaquetado del vector, por ejemplo, en total (incluyendo el primer

casete) cinco, casetes que comprenden o consisten esencialmente en un promotor, molécula de ácido nucleico que codifica ARN guía (ARNg) y un terminador (por ejemplo, cada casete representado esquemáticamente como promotor-ARNg1-terminador, promotor-ARNg2-terminador ... promotor-ARNg(N)-terminador (donde N es un número que puede insertarse que está en el límite superior del límite de tamaño de empaquetado del vector), o dos o más rAAV individuales, conteniendo, cada uno, uno o más de un casete de un sistema de CRISPR, por ejemplo, un primer rAAV que contiene el primer casete que comprende o consisten esencialmente en un promotor, una molécula de ácido nucleico que codifica Cas, por ejemplo, Cas9 y un terminador, y un segundo rAAV que contiene una pluralidad, cuatro, casetes que comprenden o consisten esencialmente en un promotor, molécula de ácido nucleico que codifica ARN guía (ARNg) y un terminador (por ejemplo, cada casete representado esquemáticamente como promotor-ARNg1-terminador, promotor-ARNg2-terminador ... promotor-ARNg(N)-terminador (donde N es un número que puede insertarse que está en un límite superior del límite de tamaño de empaquetado del vector). Como el rAAV es un virus de ADN, las moléculas de ácido nucleico en el análisis de este documento que se refiere a AAV o rAAV son ventajosamente ADN. El promotor es, en algunas realizaciones, ventajosamente promotor de sinapsina I humana (hSyn). Se conocen métodos adicional para el suministro de ácidos nucleicos a células por los expertos en la materia. Véase, por ejemplo, el documento US20030087817.

En algunas realizaciones, una célula hospedadora se transfecta de forma transitoria o no transitoria con uno o más vectores descritos en este documento. En algunas realizaciones, una célula se transfecta como se produce de forma natural en un sujeto. En algunas realizaciones, una célula que se transfecta se recoge de un sujeto. En algunas realizaciones, la célula se obtiene de células recogidas de un sujeto, tal como una línea celular. Se conoce en la técnica una amplia diversidad de líneas celulares para cultivo hístico. Ejemplos de líneas celulares incluyen, aunque sin limitación, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEK_n, HEKa, MiaPaCell, Panel, PC-3, TF1, CTLL-2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45,01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264,7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 epiteliales de riñón de mono, fibroblastos de embrión de ratón BALB/ 3T3, 3T3 Swiss, 3T3-L1, fibroblastos fetales humanos 132-d5 ; fibroblastos de ratón 10.1, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, células BCP-1, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr ^{-/-}, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10,0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepalcl7, HL-60, HMEC, HT-29, Jurkat, células JY, células K562, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCK II, MDCK II, MOR/0,2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, líneas celulares OPCN / OPCT, Peer, PNT-1A / PNT 2, RenCa, RIN-5F, RMA/RMAS, células Saos-2, Sf-9, SkBr3, T2, T-47D, T84, línea celular THP1, U373, U87, U937, VCaP, células Vero, WM39, WT-49, X63, YAC-1, YAR y variedades transgénicas de las mismas. Hay líneas celulares disponibles de una diversidad de fuentes conocidas por los expertos en la materia (véase, por ejemplo, la American Type Culture Collection (ATCC) (Manassas, Va.)). En algunas realizaciones, una célula transfectada con uno o más vectores descritos en este documento se usa para establecer una nueva línea celular que comprende una o más secuencias derivadas de vector. En algunas realizaciones, una célula transfectada de forma transitoria con los componentes de un sistema de CRISPR como se describe en este documento (tal como por transfección transitoria de uno o más vectores, o transfección con ARN), y modificada a través de la actividad de un complejo de CRISPR, se usa para establecer una nueva línea celular que comprende células que contienen la modificación, pero que carecen de cualquier otra secuencia exógena. En algunas realizaciones, las células transfectadas de forma transitoria o no transitoria con uno o más vectores descritos en este documento, o líneas celulares derivadas de dichas células se usan en la evaluación de uno o más compuestos de ensayo.

En algunas realizaciones, se usa uno o más vectores descritos en este documento para producir un animal no humano transgénico o planta transgénica, por ejemplo, un organismo modelo. En algunas realizaciones, el animal transgénico es un mamífero, tal como ratón, rata o conejo. Se conocen en la técnica métodos para producir plantas y animales transgénicos, y en general emplean con un método de transfección celular, tal como se describe en este documento.

Con respecto al uso del sistema de CRISPR-Cas, en general, se hace mención de los documentos, incluyendo solicitudes de patente, patentes y publicaciones de patente citadas durante toda esta divulgación como realizaciones de la invención que pueden usarse como en esos documentos. El uno o más sistemas de CRISPR-Cas (por ejemplo, individuales o combinados) pueden usarse junto con avances recientes en genómica de cultivos. Dicho uno o más sistemas de CRISPR-Cas pueden usarse para realizar consulta o edición o manipulación génica o genómica eficaz y rentable en plantas, por ejemplo, para una rápida investigación y/o selección y/o consulta y/o comparación y/o manipulación y/o transformación de genes o genomas vegetales; por ejemplo, para crear, identificar, desarrollar, optimizar o conferir uno o más rasgos o características a una o más plantas o para transformar un genoma vegetal. Por consiguiente, puede haber producción mejorada de plantas, nuevas plantas con nuevas combinaciones de rasgos o características o nuevas plantas con rasgos potenciados. Dicho uno o más sistemas de CRISPR-Cas pueden usarse con respecto a plantas en integración dirigida al sitio (SDI) o edición génica (GE) o cualquier técnica de fitomejoramiento inverso cercano (NRB) o fitomejoramiento inverso (RB). Con respecto al uso del sistema de CRISPR-Cas en plantas, se hace mención del sitio web de la Universidad de Arizona "CRISPR-PLANT" (<http://www.genome.arizona.edu/crispr/>) (con asistencia de Penn State y AGI). Realizaciones de la invención pueden

usarse en edición genómica en plantas o donde se han usado previamente técnicas de iARN o de edición genómica similares; véase, por ejemplo, Nekrasov, "Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system," *Plant Methods* 2013, 9:39 (doi:10.1186/1746-4811-9-39); Brooks, "Efficient gene editing in tomato in the first generation using the CRISPR/Cas9 system," *Plant Physiology* septiembre de 2014 pág. 114.247577; Shan, "Targeted genome modification of crop plants using a CRISPR-Cas system," *Nature Biotechnology* 31, 686-688 (2013); Feng, "Efficient genome editing in plants using a CRISPR/Cas system," *Cell Research* (2013) 23:1229-1232, doi:10.1038/cr.2013.114; publicado en línea el 20 de agosto de 2013; Xie, "RNA-guided genome editing in plants using a CRISPR-Cas system," *Mol Plant*. 2013 Nov; 6(6):1975-83, doi: 10.1093/mp/sst119. Epub 17 de agosto de 2013; Xu, "Gene targeting using the *Agrobacterium tumefaciens*-mediated CRISPR-Cas system in rice," *Rice* 2014, 7:5 (2014), Zhou *et al.*, "Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and Redundancy," *New Phytologist* (2015) (Forum) 1-4 (disponible en línea solamente en www.newphytologist.com); Caliando *et al.*, "Targeted DNA degradation using a CRISPR device stably carried in the host genome," *NATURE COMMUNICATIONS* 6:6989, DOI: 10.1038/ncomms7989, www.nature.com/naturecommunications DOI: 10.1038/ncomms7989; patente de Estados Unidos n.º 6.603.061 - *Agrobacterium*-Mediated Plant Transformation Method; patente de Estados Unidos n.º 7.868.149 - Plant Genome Sequences and Uses Thereof y documento US 2009/0100536 - Transgenic Plants with Enhanced Agronomic Traits. En la práctica de la invención, los contenidos y divulgación de Morrell *et al.* "Crop genomics: advances and applications," *Nat Rev Genet*. 29 de diciembre de 2011;13(2):85-96; cada uno de los cuales se cita por referencia en este documento, incluyendo en cuanto a la manera en que pueden usarse las realizaciones de este documento para las plantas. Por consiguiente, una referencia en este documento a células animales también puede aplicarse, *mutatis mutandis*, a células vegetales salvo que sea evidente lo contrario.

En un aspecto, la divulgación proporciona métodos de modificación de un polinucleótido diana en una célula eucariótica, que puede ser *in vivo*, *ex vivo* o *in vitro*. En algunas realizaciones, el método comprende tomar muestras de una célula o población de células de un ser humano o animal no humano o planta (incluyendo microalgas), y modificar la célula o células. El cultivo puede producirse en cualquier fase *ex vivo*. La célula o células incluso pueden reintroducirse en el animal no humano o planta (incluyendo microalgas). Para células reintroducidas, es particularmente preferido que las células sean células madre.

En algunas realizaciones, el método comprende permitir que un complejo de CRISPR se una al polinucleótido diana para lograr la escisión de dicho polinucleótido diana modificando de ese modo el polinucleótido diana, en el que el complejo de CRISPR comprende una enzima de CRISPR en complejo con una secuencia guía hibridada con una secuencia diana dentro de dicho polinucleótido diana, en el que dicha secuencia guía se une a una secuencia de acoplamiento de crtra que a su vez hibrida con una secuencia crtra.

En un aspecto, la divulgación proporciona un método de modificación de la expresión de un polinucleótido en una célula eucariótica. En algunas realizaciones, el método comprende permitir que un complejo de CRISPR se una al polinucleótido, de modo que dicha unión provoque expresión aumentada o disminuida de dicho polinucleótido; en el que el complejo de CRISPR comprende una enzima de CRISPR mutada en complejo con una secuencia guía hibridada con una secuencia diana dentro de dicho polinucleótido, en el que dicha secuencia guía se une a una secuencia de acoplamiento de crtra que a su vez hibrida con una secuencia crtra. Consideraciones y condiciones similares se aplican como anteriormente para métodos de modificación de un polinucleótido diana. De hecho, estas opciones de toma de muestras, cultivo y reintroducción se aplica en todos los aspectos de la presente invención.

De hecho, en cualquier aspecto de la invención, el complejo de CRISPR puede comprender una enzima de CRISPR mutada en complejo con una secuencia guía hibridada con una secuencia diana, en el que dicha secuencia guía puede unirse a una secuencia de acoplamiento de crtra que, a su vez, puede hibridar con una secuencia crtra. Consideraciones y condiciones similares se aplican como anteriormente para métodos de modificación de un polinucleótido diana.

En un aspecto, la divulgación proporciona kits que contienen uno cualquiera o más de los elementos divulgados en los métodos y composiciones anteriores. Los elementos pueden proporcionarse individualmente o en combinaciones, y pueden proporcionarse en cualquier recipiente adecuado, tal como un vial, un frasco o un tubo. En algunas realizaciones, el kit incluye instrucciones en uno o más idiomas, por ejemplo, en más de un idioma.

En algunas realizaciones, un kit comprende uno o más reactivos para su uso en un proceso utilizando uno o más de los elementos descritos en este documento. Los reactivos pueden proporcionarse en cualquier recipiente adecuado. Por ejemplo, un kit puede proporcionar uno o más tampones de reacción o almacenamiento. Los reactivos pueden proporcionarse en una forma que sea útil en un ensayo particular, o en una forma que requiera adición de uno o más componentes distintos antes de su uso (por ejemplo, en forma concentrada o liofilizada). Un tampón puede ser cualquier tampón incluyendo, aunque sin limitación, un tampón de carbonato de sodio, un tampón de bicarbonato de sodio, un tampón de borato, un tampón Tris, un tampón MOPS, un tampón HEPES y combinaciones de los mismos. En algunas realizaciones, el tampón es alcalino. En algunas realizaciones, el tampón tiene un pH de aproximadamente 7 a aproximadamente 10. En algunas realizaciones, el kit comprende uno o más oligonucleótidos correspondientes a una secuencia guía para su inserción en un vector para unir de forma funcional la secuencia guía y un elemento regulador. En algunas realizaciones, el kit comprende un polinucleótido de molde de recombinación homóloga.

En un aspecto, la divulgación proporciona métodos para usar uno o más elementos de un sistema de CRISPR. El complejo de CRISPR de la invención proporciona un medio eficaz para modificar un polinucleótido diana. El complejo de CRISPR de la invención tiene una amplia diversidad de utilidades, incluyendo la modificación (por ejemplo, eliminación, inserción, translocación, inactivación, activación) de un polinucleótido diana en una multiplicidad de tipo celulares. Por tanto, el complejo de CRISPR de la divulgación tiene un amplio espectro de aplicaciones en, por ejemplo, genoterapia, cribado de fármacos, diagnóstico de enfermedades y pronóstico. Un complejo de CRISPR ejemplar comprende una enzima CRISPR mutada en complejo con una secuencia guía hibridada con una secuencia diana dentro del polinucleótido diana. La secuencia guía se une a una secuencia de acoplamiento de crtra, que a su vez hibrida con una secuencia crtra.

La rotura creada por el complejo de CRISPR puede repararse por un proceso de reparación tal como la ruta de unión de extremos no homólogos (NHEJ) propensa a errores o la reparación dirigida por homología (HDR) de alta fidelidad. Durante estos procesos de reparación, un molde polinucleotídico exógeno puede introducirse en la secuencia genómica. En algunos métodos, el proceso de HDR se usa para modificar la secuencia genómica. Por ejemplo, un molde polinucleotídico exógeno que comprende una secuencia a integrar flanqueada por una secuencia anterior y una secuencia posterior se introduce en una célula. Las secuencias anterior y posterior comparte similitud de secuencia con cada lado del sitio de integración en el cromosoma.

Cuando se desea, un polinucleótido donador puede ser ADN, por ejemplo, un plásmido de ADN, un cromosoma artificial bacteriano (BAC), un cromosoma artificial de levadura (YAC), un vector vírico, un trozo lineal de ADN, un fragmento de PCR, un ácido nucleico desnudo o un ácido nucleico en complejo con un vehículo de suministro tal como un liposoma o poloxámero.

El molde polinucleotídico exógeno comprende una secuencia a integrar (por ejemplo, un gen mutado). La secuencia para su integración puede ser una secuencia endógena o exógena a la célula. Ejemplos de una secuencia a integrar incluyen polinucleótidos que codifican una proteína o un ARN no codificante (por ejemplo, un microARN). Por tanto, la secuencia para su integración puede unirse de forma funcional a una secuencia o secuencias de control apropiadas. Como alternativa, la secuencia a integrar puede proporcionar una función reguladora.

Las secuencias anteriores y posteriores en el molde polinucleotídico exógeno se seleccionan para promover la recombinación entre la secuencia cromosómica de interés y el polinucleótido donador. La secuencia anterior es una secuencia de un ácido nucleico que comparte similitud de secuencia con la secuencia genómica anterior al sitio diana para la integración. Asimismo, la secuencia posterior es una secuencia de un ácido nucleico que comparte similitud de secuencia con la secuencia cromosómica posterior al sitio diana de integración. Las secuencias anteriores y posteriores en el molde polinucleotídico exógeno pueden tener un 75%, 80%, 85%, 90%, 95%, o 100% de identidad de secuencia con la secuencia genómica diana. Preferiblemente, las secuencias anteriores y posteriores en el molde polinucleotídico exógeno tienen aproximadamente un 95%, 96%, 97%, 98%, 99%, o 100% de identidad de secuencia con la secuencia genómica diana. En algunos métodos, las secuencias anteriores y posteriores en el molde polinucleotídico exógeno tienen aproximadamente un 99 % o un 100 % de identidad de secuencia con la secuencia genómica diana.

Una secuencia anterior o posterior puede comprender de aproximadamente 20 pb a aproximadamente 2500 pb, por ejemplo, aproximadamente 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, o 2500 pb. En algunos métodos, la secuencia anterior o posterior ejemplar tiene de aproximadamente 200 pb a aproximadamente 2000 pb, de aproximadamente 600 pb a aproximadamente 1000 pb, o más particularmente de aproximadamente 700 pb a aproximadamente 1000 pb.

En algunos métodos, el molde polinucleotídico exógeno puede comprender además un marcador. Dicho marcador puede hacer fácil cribar integraciones dirigidas. Ejemplos de marcadores adecuados incluyen sitios de restricción, proteínas fluorescentes o marcadores de selección. El molde polinucleotídico exógeno de la divulgación puede construirse usando técnicas recombinantes (véase, por ejemplo, Sambrook *et al.*, 2001 y Ausubel *et al.*, 1996).

En otras realizaciones, esta divulgación proporciona un método de modificación de la expresión de un polinucleótido en una célula eucariótica. El método comprende aumentar o disminuir la expresión de un polinucleótido diana usando un complejo de CRISPR que se une al polinucleótido.

En algunos métodos, un polinucleótido diana puede inactivarse para lograr la modificación de la expresión en una célula. Por ejemplo, tras la unión de un complejo de CRISPR a una secuencia diana en una célula, el polinucleótido diana se inactiva de modo que la secuencia no se transcriba, la proteína codificada no se produzca o la secuencia no funcione como lo hace la secuencia de tipo silvestre. Por ejemplo, una proteína o secuencia codificante de microARN puede inactivarse de modo que la proteína no se produzca.

En algunos métodos, una secuencia de control puede inactivarse de modo que ya no funcione como una secuencia de control. Como se usa en este documento, "secuencia de control" se refiere a cualquier secuencia de ácido nucleico

que logre la transcripción, traducción o accesibilidad de una secuencia de ácido nucleico. Ejemplos de una secuencia de control incluyen un promotor, un terminador de la transcripción y un potenciador, que son secuencias de control.

5 La secuencia diana inactivada puede incluir una mutación de eliminación (es decir, eliminación de uno o más nucleótidos), una mutación de inserción (es decir, inserción de uno o más nucleótidos), o una mutación finalizadora (es decir, sustitución de un solo nucleótido por otro nucleótido de modo que se introduce un codón de parada). En algunos métodos, la inactivación de una secuencia diana provoca la "inactivación" de la secuencia diana.

10 Un método de la divulgación puede usarse para crear una planta, un animal o célula que puede usarse como modelo de enfermedad. Como se usa en este documento, "enfermedad" se refiere a una enfermedad, trastorno o indicación en un sujeto. Por ejemplo, un método de la divulgación puede usarse para crear un animal o célula que comprende una modificación en una o más secuencias de ácido nucleico asociada con una enfermedad, o una planta, animal o célula en que la expresión de una o más secuencias de ácido nucleico asociadas con una enfermedad está alterada. Dicha secuencia de ácido nucleico puede codificar una secuencia proteínica asociada a enfermedad o puede ser una secuencia de control asociada a enfermedad. Por consiguiente, se entiende que en realizaciones de la divulgación, una planta, sujeto, paciente, organismo o célula puede ser un sujeto no humano, paciente, organismo o célula. Por tanto, la divulgación proporciona una planta, animal o célula, producido por los presentes métodos, o una descendencia del mismo. La descendencia puede ser un clon de la planta o animal producido, o puede ser el resultado de reproducción sexual por cruce con otros individuos de la misma especie para realizar introgresión de rasgos deseables adicionales en sus descendientes. La célula puede estar *in vivo* o *ex vivo* en los casos de organismos multicelulares, particularmente animales o plantas. En el caso donde la célula está en cultivo, puede establecerse una línea celular si se cumplen condiciones de cultivo apropiadas y preferiblemente si la célula se adapta adecuadamente para este propósito (por ejemplo, una célula madre). También se prevén líneas celulares bacterianas producidas por la divulgación. Por tanto, también se prevén líneas celulares.

25 En algunos métodos, el modelo de enfermedad puede usarse para estudiar los efectos de mutaciones en el animal o célula y el desarrollo y/o progresión de la enfermedad usando medidas habitualmente usadas en el estudio de la enfermedad. Como alternativa, dicho modelo de enfermedad es útil para estudiar el efecto de un compuesto farmacéuticamente activo sobre la enfermedad.

30 En algunos métodos, el modelo de enfermedad puede usarse para evaluar la eficacia de una estrategia de genoterapia potencial. Es decir, puede modificarse un gen o polinucleótido asociado a enfermedad de modo que el desarrollo y/o progresión de la enfermedad se inhiba o reduzca. En particular, el método comprende modificar un gen o polinucleótido asociado a enfermedad de modo que se produzca una proteína alterada y, como resultado, el animal o célula tenga una respuesta alterada. Por consiguiente, en algunos métodos, puede compararse un animal modificado genéticamente con un animal predispuesto al desarrollo de la enfermedad, de modo que pueda evaluarse el efecto del evento de genoterapia.

40 En otra realización, esta divulgación proporciona un método de desarrollo de un agente biológicamente activo que module un evento de señalización celular asociado con un gen de enfermedad. El método comprende poner en contacto un compuesto de ensayo con una célula, que comprende uno o más vectores que dirigen la expresión de una o más de una enzima de CRISPR mutada, una secuencia guía unida a una secuencia de acoplamiento de crtra y una secuencia crtra; y detectar un cambio en una lectura que es indicativo de una reducción o un aumento de un evento de señalización celular asociado con, por ejemplo, una mutación en un gen de enfermedad contenido en la célula.

50 Un modelo celular o modelo animal puede construirse en combinación con el método de la divulgación para cribar un cambio de función celular. Dicho modelo puede usarse para estudiar los efectos de una secuencia genómica modificada por el complejo de CRISPR de la invención sobre una función celular de interés. Por ejemplo, puede usarse un modelo de función celular para estudiar el efecto de una secuencia genómica modificada sobre la señalización intracelular o señalización extracelular. Como alternativa, puede usarse un modelo de función celular para estudiar los efectos de una secuencia genómica modificada sobre la percepción sensitiva. En algunos de dichos modelos, una o más secuencias genómicas asociadas con una ruta bioquímica de señalización en el modelo se modifican.

55 Se han investigado específicamente varios modelos de enfermedad. Estas incluyen los genes de riesgo de autismo *de novo* CHD8, KATNAL2 y SCN2A; y el gen de autismo sindrómico (síndrome de Angelman) UBE3A. Estos genes y modelos de autismo resultantes se prefieren, por supuesto, pero sirven para mostrar la amplia aplicabilidad de la invención entre los genes y los correspondientes modelos.

60 Una expresión alterada de una o más secuencias genómicas asociadas con una ruta bioquímica de señalización puede determinarse evaluando una diferencia en los niveles de ARNm de los correspondientes genes entre la célula de modelo de ensayo y una célula de control, cuando se ponen en contacto con un agente candidato. Como alternativa, la expresión diferencial de las secuencias asociadas con una ruta bioquímica de señalización se determina detectando una diferencia en el nivel del polipéptido codificado o producto génico.

65

Para ensayar una alteración inducida por agente a nivel de transcritos de ARNm o polinucleótidos correspondientes, el ácido nucleico contenido en una muestra en primer lugar se extrae de acuerdo con métodos convencionales en la técnica. Por ejemplo, el ARNm puede aislarse usando diversas enzimas líticas o soluciones químicas de acuerdo con los procedimientos expuestos en Sambrook *et al.* (1989), o extraerse mediante resinas de unión a ácido nucleico siguiendo las instrucciones adjuntas proporcionadas por los fabricantes. El ARNm contenido en la muestra de ácido nucleico extraído se detecta entonces por procedimientos de amplificación o ensayos convencionales de hibridación (por ejemplo, análisis de transferencia de Northern) de acuerdo con métodos ampliamente conocidos o basados en los métodos ejemplificados en este documento.

Para el propósito de esta invención, la amplificación significa cualquier método que emplee un cebador y una polimerasa que puedan replicar una secuencia diana con fidelidad razonable. La amplificación puede realizarse por ADN-polimerasas naturales o recombinantes tales como TaqGold™, ADN-polimerasa T7, fragmento Klenow de la ADN-polimerasa de *E. coli* y retrotranscriptasa. Un método de amplificación preferido es PCR. En particular, el ARN aislado puede someterse a un ensayo de retrotranscripción que se acopla con una reacción en cadena de la polimerasa cuantitativa (RT-PCR) para cuantificar el nivel de expresión de una secuencia asociada con una ruta bioquímica de señalización.

La detección del nivel de expresión génica puede realizarse en tiempo real en un ensayo de amplificación. En un aspecto, los productos amplificados puede visualizarse directamente con agentes fluorescentes de unión a ADN incluyendo, aunque sin limitación, intercalantes en el ADN y agentes de unión al surco de ADN. Como la cantidad de los intercalantes incorporados en las moléculas de ADN bicatenario típicamente es proporcional a la cantidad de los productos de ADN amplificados, se puede determinar convenientemente la cantidad de los productos amplificados cuantificando la fluorescencia del tinte intercalado usando sistemas ópticos convencionales en la técnica. El tinte de unión a ADN adecuado para esta aplicación incluye SYBR green, SYBR blue, DAPI, yodo de propidio, Hoeste, SYBR gold, bromuro de etidio, acridinas, proflavina, naranja de acridina, acriflavina, fluorcumanina, elipticina, daunomicina, cloroquina, distamicina D, cromomicina, homidio, mitramicina, polipiridilos de rutenio, antramycin y similares.

En otro aspecto, pueden emplearse otros marcadores fluorescentes tales como sondas específicas de secuencia en la reacción de amplificación para facilitar la detección y cuantificación de los productos amplificados. La amplificación cuantitativa basada en sonda depende de la detección específica de secuencia de un producto amplificado deseado. Utiliza sondas específicas de diana fluorescentes (por ejemplo, sondas TaqMan®) que provocan una especificidad y sensibilidad aumentadas. Los métodos para realizar amplificación cuantitativa basada en sonda están bien establecidos en la técnica y se muestran en la patente de Estados Unidos n.º 5.210.015.

En otro aspecto más, pueden realizarse ensayos de hibridación convencionales usando sondas de hibridación que comparten homología de secuencia con secuencias asociadas con una ruta bioquímica de señalización. Típicamente, se permite que las sondas formen complejos estables con las secuencias asociadas con una ruta bioquímica de señalización contenida dentro de la muestra biológica derivada del sujeto de ensayo en una reacción de hibridación. Un experto en la materia apreciará que cuando se usa antisentido como ácido nucleico de sonda, los polinucleótidos diana proporcionados en la muestra se eligen para que sean complementarios a secuencias de los ácidos nucleicos de antisentido. A la inversa, cuando la sonda nucleotídica es un ácido nucleico de sentido, el polinucleótido diana se selecciona para que sea complementario a secuencias del ácido nucleico de sentido.

La hibridación puede realizarse en condiciones de diversa rigurosidad. Las condiciones de hibridación adecuadas para la práctica de la presente invención son tales que la interacción de reconocimiento entre la sonda y las secuencias asociadas con una ruta bioquímica de señalización es tanto suficientemente específica como suficientemente estable. Las condiciones que aumentan la rigurosidad de una reacción de hibridación son ampliamente conocidas y están publicadas en la técnica. Véase, por ejemplo, (Sambrook, *et al.*, (1989); Nonradioactive In Situ Hybridization Application Manual, Boehringer Mannheim, segunda edición). El ensayo de hibridación puede formarse usando sondas inmovilizadas en cualquier soporte sólido incluyendo, aunque sin limitación, nitrocelulosa, vidrio, silicio y una diversidad de matrices génicas. Un ensayo de hibridación preferido se realiza en genochips de alta densidad como se describe en la patente de Estados Unidos n.º 5.445.934.

Para una detección conveniente de los complejos de sonda-diana formados durante el ensayo de hibridación, las sondas nucleotídicas se conjugan con un marcador detectable. Los marcadores detectables adecuados para su uso en la presente invención incluyen cualquier composición detectable por medios fotoquímicos, bioquímicos, espectroscópicos, inmunoquímicos, eléctricos, ópticos o químicos. Se conoce en la técnica una amplia diversidad de marcadores detectables apropiados, que incluyen marcadores fluorescentes o quimioluminiscentes, marcadores isotópicos radioactivos, ligandos enzimáticos u otros. En realizaciones preferidas, probablemente se deseará emplear un marcador fluorescentes o una marca enzimática, tal como digoxigenina, β -galactosidasa, ureasa, fosfatasa alcalina o peroxidasa, complejo de avidina/biotina.

Los métodos de detección usados para detectar o cuantificar la intensidad de hibridación típicamente dependerán del marcador seleccionado anteriormente. Por ejemplo, pueden detectarse radiomarcadores usando película fotográfica o un generación de fosfoimágenes. Los marcadores fluorescentes pueden detectarse y cuantificarse usando un fotodetector para detectar la luz emitida. Los marcadores enzimáticos típicamente se detectan proporcionando a la

enzima un sustrato y midiendo el producto de reacción producido por la acción de la enzima sobre el sustrato; y finalmente los marcadores colorimétricos se detectan simplemente visualizando el marcador coloreado.

Un cambio inducido por agente en la expresión de secuencias asociadas con una ruta bioquímica de señalización también puede determinarse examinando los correspondientes productos génicos. Determinar el nivel de proteína típicamente implica a) poner en contacto la proteína contenida en una muestra biológica con un agente que se une específicamente a una proteína asociada con una ruta bioquímica de señalización; y (b) identificar cualquier complejo de agente:proteína así formado. En un aspecto de esta realización, el agente que se une específicamente a una proteína asociada con una ruta bioquímica de señalización es un anticuerpo, preferiblemente un anticuerpo monoclonal.

La reacción se realiza poniendo en contacto el agente con una muestra de las proteínas asociadas con una ruta bioquímica de señalización derivada de las muestras de ensayo en condiciones que permitirán que se forme un complejo entre el agente y las proteínas asociadas con una ruta bioquímica de señalización. La formación del complejo puede detectarse directa o indirectamente de acuerdo con procedimientos convencionales en la técnica. En el método de detección directa, a los agentes se les aporta un marcador detectable y los agentes que no han reaccionado se pueden eliminar del complejo; indicando la cantidad de marcador restante de este modo la cantidad de complejo formado. Para dicho método, es preferible seleccionar marcadores que permanezcan fijados a los agentes incluso durante condiciones de lavado rigurosas. Es preferible que el marcador no impida la reacción de unión. Como, un procedimiento de detección indirecta puede usar un agente que contenga un marcador introducido química o enzimáticamente. Un marcador deseable en general no impide la unión o la estabilidad del complejo de agente:polipéptido resultante. Sin embargo, el marcador típicamente se diseña para que sea accesible para un anticuerpo para una unión eficaz y, por tanto, la generación de una señal detectable.

Se conoce en la técnica una amplia diversidad de marcadores adecuados para detectar niveles de proteína. Ejemplos no limitantes incluyen radioisótopos, enzimas, metales coloidales, compuestos fluorescentes, compuestos bioluminiscentes y compuestos quimioluminiscentes.

La cantidad de complejos de agente:polipéptido formados durante la reacción de unión puede cuantificarse por ensayos cuantitativos convencionales. Como se ilustra anteriormente, la formación de complejo de agente:polipéptido puede medirse directamente por la cantidad de marcador que queda en el sitio de unión. En una alternativa, la proteína asociada con una ruta bioquímica de señalización se ensaya para su capacidad de competir con un análogo marcado para los sitios de unión en el agente específico. En este ensayo competitivo, la cantidad de marcador capturado es inversamente proporcional a la cantidad de secuencias proteínicas asociadas con una ruta bioquímica de señalización presente en una muestra de ensayo.

Varias técnicas para análisis de proteínas basadas en los principios generales resumidos anteriormente están disponibles en la técnica. Estas incluyen, aunque sin limitación, radioinmunoensayos, ELISA (ensayos inmunoradiométricos enzimáticos), inmunoensayos de tipo "sándwich", ensayos inmunoradiométricos, inmunoensayos *in situ* (usando, por ejemplo, oro coloidal, enzima o marcadores radioisotópicos), análisis de transferencia de Western, ensayos de inmunoprecipitación, ensayos inmunofluorescentes y SDS-PAGE.

Los anticuerpos que reconocen específicamente o se unen a proteínas asociadas con una ruta bioquímica de señalización son preferibles para realizar los análisis de proteínas mencionados anteriormente. Cuando se desee, pueden usarse los anticuerpos que reconocen un tipo específico de modificaciones postraduccionales (por ejemplo, modificaciones inducibles por ruta bioquímica de señalización). Las modificaciones postraduccionales incluyen, aunque sin limitación, glucosilación, lipidación, acetilación y fosforilación. Estos anticuerpos pueden adquirirse de proveedores comerciales. Por ejemplo, los anticuerpos antifosfotirosina que reconocen específicamente proteínas fosforiladas en tirosina están disponibles en varios proveedores incluyendo Invitrogen y Perkin Elmer. Los anticuerpos antifosfotirosina son particularmente útiles en detectar proteínas que se fosforilan de forma diferencial en sus residuos de tirosina en respuesta a una sobrecarga de ER. Dichas proteínas incluyen, aunque sin limitación, el factor 2 alfa de inicio de la traducción eucariótico (eIF-2 α). Como alternativa, estos anticuerpos pueden generarse usando tecnologías convencionales de anticuerpos policlonales o monoclonales inmunizando un animal hospedador o una célula productora de anticuerpo con una proteína diana que muestra la modificación postraduccional deseada.

En la práctica del método en cuestión, puede ser deseable discernir el patrón de expresión de una proteína asociada con una ruta bioquímica de señalización en diferente tejido corporal, en diferentes tipos celulares y/o en diferentes estructuras subcelulares. Estos estudios pueden realizarse con el uso de anticuerpos específicos de tejidos, específicos de célula o específicos de estructura subcelular que pueden unirse a marcadores proteínicos que se expresen preferentemente en determinados tejidos, tipos celulares o estructuras subcelulares.

Una expresión alterada de un gen asociado con una ruta bioquímica de señalización también puede determinarse examinando un cambio en la actividad del producto génico con respecto a una célula de control. El ensayo para un cambio inducido por agente en la actividad de una proteína asociada con una ruta bioquímica de señalización dependerá de la actividad biológica y/o la rutas de transducción de señales que esté en investigación. Por ejemplo, cuando la proteína es una cinasa, un cambio en su capacidad de fosforilar el uno o más sustratos posteriores puede

determinarse por una diversidad de ensayos conocidos en la técnica. Los ensayos representativos incluyen, aunque sin limitación, inmunotransferencia e inmunoprecipitación con anticuerpos tales como anticuerpo antifosfotirosina que reconocen proteínas fosforiladas. Además, la actividad cinasa puede detectarse mediante ensayos quimioluminiscentes de alto rendimiento tales como AlphaScreen™ (disponible en Perkin Elmer) y ensayo eTag™ (Chan-Hui, *et al.* (2003) *Clinical Immunology* 111: 162-174).

Cuando la proteína asociada con una ruta bioquímica de señalización es parte de una cascada de señalización que da lugar a fluctuación de la condición de pH intracelular, pueden usarse moléculas sensibles al pH tales como tintes de pH fluorescentes como moléculas indicadoras. En otro ejemplo donde la proteína asociada con una ruta bioquímica de señalización es un canal de iones, pueden controlarse las fluctuaciones en el potencial de membrana y/o la concentración de iones intracelulares. Varios kits comerciales y dispositivos de alto rendimiento son particularmente adecuados para un cribado rápido y robusto de moduladores de canales de iones. Instrumentos representativos incluyen FLIPRTM (Molecular Devices, Inc.) y VIPR (Aurora Biosciences). Estos instrumentos pueden detectar reacciones en más de 1000 pocillos de muestra de una microplaca simultáneamente, y proporcionar medición en tiempo real y datos funcionales en un segundo o incluso un minisegundo.

En la práctica de cualquiera de los métodos divulgados en este documento, un vector adecuado puede introducirse en una célula o un embrión mediante uno o más métodos conocidos en la técnica, incluyendo sin limitación, microinyección, electroporación, sonoporación, biolística, transfección mediada con fosfato de calcio, transfección catiónica, transfección con liposomas, transfección con dendrímeros, transfección por choque térmico, transfección por nucleofección, magnetofección, lipofección, impalefacción, transfección óptica, captación de ácidos nucleicos potenciada por agente patentado y suministro mediante liposomas, inmunoliposomas, virosomas o viriones artificiales. En algunos métodos, el vector se introduce en un embrión por microinyección. El vector o vectores pueden microinyectarse en el núcleo o el citoplasma del embrión. En algunos métodos, el vector o vectores pueden introducirse en una célula mediante nucleofección.

El polinucleótido diana de un complejo de CRISPR puede ser cualquier polinucleótido endógeno o exógeno a la célula eucariótica. Por ejemplo, el polinucleótido diana puede ser un polinucleótido que reside en el núcleo de la célula eucariótica. El polinucleótido diana puede ser una secuencia codificante de un producto génico (por ejemplo, una proteína) o una secuencia no codificante (por ejemplo, un polinucleótido regulador o un ADN de desecho).

Ejemplos de polinucleótidos diana incluyen una secuencia asociada con una ruta bioquímica de señalización, por ejemplo, un gen o polinucleótido asociado a una ruta bioquímica de señalización. Ejemplos de polinucleótidos diana incluyen un gen o polinucleótido asociado a enfermedad. Un gen o polinucleótido "asociado a enfermedad" se refiere a cualquier gen o polinucleótido que esté produciendo productos de transcripción o traducción a un nivel anómalo o en una forma anómala en células derivadas de un tejido afectado por la enfermedad en comparación con tejidos o células de un control sin enfermedad. Puede ser un gen que llegue a expresarse a un nivel anormalmente alto; puede ser un gen que llegue a expresarse a un nivel anormalmente bajo, donde la expresión alterada se correlaciona con la aparición y/o progresión de la enfermedad. Un gen asociado a enfermedad también se refiere a un gen que posee una o más mutaciones o variaciones genéticas que son directamente responsables o está en desequilibrio de ligamiento con uno o más genes que son responsables de la etiología de una enfermedad. Los productos transcritos o traducidos pueden ser conocidos o desconocidos, y pueden estar a un nivel normal o anómalo.

El polinucleótido diana de un complejo de CRISPR puede ser cualquier polinucleótido endógeno o exógeno a la célula eucariótica. Por ejemplo, el polinucleótido diana puede ser un polinucleótido que reside en el núcleo de la célula eucariótica. El polinucleótido diana puede ser una secuencia codificante de un producto génico (por ejemplo, una proteína) o una secuencia no codificante (por ejemplo, un polinucleótido regulador o un ADN de desecho).

El polinucleótido diana de un complejo de CRISPR puede incluir varios genes y polinucleótidos asociados a enfermedad, así como genes y polinucleótidos asociados a ruta bioquímica de señalización enumerados en las patentes de Estados Unidos mencionadas en este documento, incluyendo las patentes de Estados Unidos n.º 8.999.641, 8.993.233, 8.945.839, 8.932.814, 8.906.616, 8.895.308, 8.889.418, 8.889.356, 8.871.445, 8.865.406, 8.795.965, 8.771.945 y 8.697.359 y las solicitudes de patente provisional de Estados Unidos 61/736.527 y 61/748.427, ambas tituladas SYSTEMS METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION presentadas el 12 de diciembre de 2012 y 2 de enero de 2013, respectivamente.

Ejemplos de polinucleótidos diana incluyen una secuencia asociada con una ruta bioquímica de señalización, por ejemplo, un gen o polinucleótido asociado a una ruta bioquímica de señalización. Ejemplos de polinucleótidos diana incluyen un gen o polinucleótido asociado a enfermedad. Un gen o polinucleótido "asociado a enfermedad" se refiere a cualquier gen o polinucleótido que esté produciendo productos de transcripción o traducción a un nivel anómalo o en una forma anómala en células derivadas de un tejido afectado por la enfermedad en comparación con tejidos o células de un control sin enfermedad. Puede ser un gen que llegue a expresarse a un nivel anormalmente alto; puede ser un gen que llegue a expresarse a un nivel anormalmente bajo, donde la expresión alterada se correlaciona con la aparición y/o progresión de la enfermedad. Un gen asociado a enfermedad también se refiere a un gen que posee una o más mutaciones o variaciones genéticas que son directamente responsables o está en desequilibrio de ligamiento

con uno o más genes que son responsables de la etiología de una enfermedad. Los productos transcritos o traducidos pueden ser conocidos o desconocidos, y pueden estar a un nivel normal o anómalo.

5 Ejemplos de genes y polinucleótidos asociados a enfermedad se enumeran en las tablas A y B. Información específica de enfermedades está disponible en McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.) y National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.), disponible en Internet. Ejemplos de genes y polinucleótidos asociados a ruta bioquímica de señalización se enumeran en la tabla C.

10 Mutaciones en estos genes y rutas pueden provocar la producción de proteínas incorrectas o proteínas en cantidades incorrectas que influyen en la función. Ejemplos adicionales de genes, enfermedades y proteínas se citan por la presente por referencia de las patentes de Estados Unidos n.º 8.999.641, 8.993.233, 8.945.839, 8.932.814, 8.906.616, 8.895.308, 8.889.418, 8.889.356, 8.871.445, 8.865.406, 8.795.965, 8.771.945 y 8.697.359 y solicitudes provisionales de Estados Unidos 61/736.527 presentada el 12 de diciembre de 2012 y 61/748.427 presentada el 2 de enero de 2013.

15 Dichos genes, proteínas y rutas pueden ser el polinucleótido diana de un complejo de CRISPR.

Tabla A

ENFERMEDAD/TRASTORNOS	GEN(ES)
Neoplasia	PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4;
	Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF;
	HIF1a; HIF3a; Met; HRG; Bcl2; PPAR alfa; PPAR
	gamma; WT1 (tumor de Wilms); familia de receptores de FGF
	miembros (5 miembros: 1, 2, 3, 4, 5); CDKN2a; APC; RB
	(retinoblastoma); MEN1; VHL; BRCA1; BRCA2; AR
	(receptor de andrógeno); TSG101; IGF; receptor de IGF; Igf1 (4
	variantes); Igf2 (3 variantes); receptor de Igf 1; receptor de Igf 2;
	Bax; Bcl2; familia de caspasas (9 miembros:
	1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Apc
Degeneración macular senil	Abcr; Ccl2; Cc2; cp (ceruloplasmina); Timp3; catepsinaD;
	Vldlr; Ccr2
Esquizofrenia	Neuregulina1 (Nrg1); Erb4 (receptor de neuregulina);
	Complexina1 (Cplx1); Tph1 triptófano hidroxilasa; Tph2
	triptófano hidroxilasa 2; Neurexina 1; GSK3; GSK3a;
	GSK3b
Trastornos	5-HTT (Slc6a4); COMT; DRD (Drd1a); SLC6A3; DAOA;
	DTNBP1; Dao (Dao1)
Trastornos por repeticiones trinucleotídicas	HTT (enfermedad de Huntington); SBMA/SMAX1/AR (enfermedad de Kennedy); FXN/X25 (ataxia de Friedrich); ATX3 (enfermedad de Machado-Joseph); ATXN1 y ATXN2 (ataxias medulocerebelosas); DMPK (distrofia miotónica); Atrofin-1 y Atn1 (enfermedad de DRPLA); CBP (Creb-BP - inestabilidad global); VLDLR (enfermedad de Alzheimer); Atxn7; Atxn10
Síndrome del X frágil	FMR2; FXR1; FXR2; mGLUR5

ES 2 780 904 T3

ENFERMEDAD/TRASTORNOS	GEN(ES)
Trastornos relacionados con secretasa	APH-1 (alfa y beta); presenilina (Psen1); nicastrina (Ncstn); PEN-2
Otros	Nos1; Parp1; Nat1; Nat2
Trastornos relacionados con priones	Prp
ALS	SOD1; ALS2; STEX; FUS; TARDBP; VEGF (VEGF-a; VEGF-b; VEGF-c)
Drogodependencia	Prkce (alcohol); Drd2; Drd4; ABAT (alcohol); GRIA2; Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Gria1 (alcohol)
Autismo	Mecp2; BZRAP1; MDGA2; Sema5A; Neurexina 1; X frágil (FMR2 (AFF2); FXR1; FXR2; Mglur5)
Enfermedad de Alzheimer	E1; CHIP; UCH; UBB; Tau; LRP; PICALM; Clusterina; PS1; SORL1; CR1; Vldlr; Uba1; Uba3; CHIP28 (Aqp1, Acuaporina 1); Uchl1; Uchl3; APP
Inflamación	IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL-17b; IL-17c; IL-17d; IL-17f); IL-23; Cx3cr1; ptpn22; TNFa; NOD2/CARD15 for IBD; IL-6; IL-12 (IL-12a; IL-12b); CTLA4; Cx3cl1
Enfermedad de Parkinson	x-Sinucleína; DJ-1; LRRK2; Parkina; PINK1

Tabla B:

Enfermedades y trastornos de coagulación de la sangre	Anemia (CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, UMPH1, PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB, ABCB7, ABC7, ASAT); síndrome de linfocitos desnudos (TAPBP, TPSN, TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP, RFX5), trastornos hemorrágicos (TBXA2R, P2RX1, P2X1); factor H y factor H de tipo 1 (HF1, CFH, HUS); factor V y factor VIII (MCFD2); deficiencia de factor VII (F7); deficiencia de factor X (F10); deficiencia de factor XI (F11); deficiencia de factor XII (F12, HAF); deficiencia de factor XIIIa (F13A1, F13A); deficiencia de factor XIIIb (F13B); anemia de Fanconi (FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064, FANCB, FANCC, FACC, BRCA2, FANCD1, FANCD2, FANCD, FACD, FAD, FANCE, FACE, FANCF, XRCC9, FANCG, BRIP1, BACH1, FANCJ, PHF9, FANCL, FANCM, KIAA1596); trastornos de linfocitosis hemofagocítica (PRF1, HPLH2, UNC13D, MUNC13-4, HPLH3, HLH3, FHL3); hemofilia A (F8, F8C, HEMA); hemofilia B (F9, HEMB), trastornos hemorrágicos (PI, ATT, F5); deficiencias y trastornos de leucocitos (ITGB2, CD18, LCAMB, LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LVWM, CACH, CLE, EIF2B4); anemia falciforme (HBB); talasemia (HBA2, HBB, HBD, LCRB, HBA1).
Desregulación celular y enfermedades y trastornos oncológicos	Linfoma no hodgkiniano de linfoma de linfocitos B (BCL7A, BCL7); leucemia (TAL1, TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1, HOXD4, HOX4B, BCR, CML, PHL, ALL, ARNT, KRAS2, RASK2, GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH, CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214, D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1 I1, NSD3, FLT3, AF1Q, NPM1, NUMA1, ZNF145, PLZF, PML, MYL, STAT5B, AF10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML, PHL, ALL, GRAF,

ES 2 780 904 T3

	NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2, NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1, NFE1, ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN).
Inflamación y enfermedades y trastornos inmunitarios	SIDA (KIR3DL1, NKAT3, NKB1, AMB11, KIR3DS1, IFNG, CXCL12, SDF1); síndrome linfoproliferativo autoinmunitario (TNFRSF6, APT1, FAS, CD95, ALPS1A); inmunodeficiencia combinada (IL2RG, SCIDX1, SCIDX, IMD4); VIH-1 (CCL5, SCYA5, D17S136E, TCP228), susceptibilidad a infección por VIH (IL10, CSIF, CMKBR2, CCR2, CMKBR5, CCCR5 (CCR5)); inmunodeficiencias (CD3E, CD3G, AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4, TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX, TNFRSF14B, TACI); inflamación (IL-10, IL-1 (IL-1a, IL-1b), IL-13, IL-17 (IL-17a (CTLA8), IL-17b, IL-17c, IL-17d, IL-17f), IL-23, Cx3cr1, ptpn22, TNFa, NOD2/CARD15 para IBD, IL-6, IL-12 (IL-12a, IL-12b), CTLA4, Cx3c11); inmunodeficiencias combinadas graves (SCID) (JAK3, JAKL, DCLRE1C, ARTEMIS, SCIDA, RAG1, RAG2, ADA, PTPRC, CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCIDX1, SCIDX, IMD4).
Enfermedades y trastornos metabólicos, hepáticos, renales y proteínicos	Neuropatía amiloide (TTR, PALB); amiloidosis (APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); cirrosis (KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988); fibrosis quística (CFTR, ABCC7, CF, MRP7); enfermedades de almacenamiento de glucógeno (SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM); adenoma hepático, 142330 (TCF1, HNF1A, MODY3), insuficiencia hepática, de aparición temprana y trastorno neurológico (SCOD1, SCO1), deficiencia de lipasa hepática (LIPC), hepatoblastoma, cáncer y carcinomas (CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5; enfermedad renal quística medular (UMOD, HNFJ, FJHN, MCKD2, ADMCKD2); fenilcetonuria (PAH, PKU1, QDPR, DHPR, PTS); enfermedad renal y hepática poliquística (FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63).
Enfermedades y trastornos musculares/esqueléticos	Distrofia muscular de Becker (DMD, BMD, MYF6), distrofia muscular de Duchenne (DMD, BMD); distrofia muscular de Emery-Dreifuss (LMNA, LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1, EMD2, FPLD, CMD1A); distrofia muscular facioescapulohumeral (FSHMD1A, FSHD1A); distrofia muscular (FKRP, MDC1C, LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD, TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C, DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB, LGMD2E, SGCD, SGD, LGMD2F, CMD1 I, TCAP, LGMD2G, CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD2I, TTN, CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPNI, SELN, RSMD1, PLEC1, PLTN, EBS1); osteopetrosis (LRP5, BMND1, LRP7, LR3, OPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1, TIRC7, OC116, OPTB1); atrofia muscular (VAPB, VAPC, ALS8, SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1, CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1).
Enfermedades y trastornos neurológicos y neuronales	ALS (SOD1, ALS2, STEX, FUS, TARDBP, VEGF (VEGF-a, VEGF-b, VEGF-c); enfermedad de Alzheimer (APP, AAA, CVAP, AD1, APOE, AD2, PSEN2, AD4, STM2, APBB2, FE65L1, NOS3,

	<p>PLAU, URK, ACE, DCP1, ACE1, MPO, PACIP1, PAXIP1 I, PTIP, A2M, BLMH, BMH, PSEN1, AD3); autismo (Mecp2, BZRAP1, MDGA2, Sema5A, Neurexina 1, GLO1, MECP2, RTT, PPMX, MRX16, MRX79, NLGN3, NLGN4, KIAA1260, AUTSX2); síndrome del X frágil (FMR2, FXR1, FXR2, mGLUR5); enfermedad de Huntington y trastornos parecidos a la enfermedad (HD, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17); enfermedad de Parkinson (NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, SNCA, NACP, PARK1, PARK4, DJ1, PARK7, LRRK2, PARK8, PINK1, PARK6, UCHL1, PARK5, SNCA, NACP, PARK1, PARK4, PRKN, PARK2, PDJ, DBH, NDUFV2); síndrome de Rett (MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, x-Sinucleína, DJ-1); esquizofrenia (Neuregulina1 (Nrg1), Erb4 (receptor de neuregulina), Complejina1 (Cplx1), Tph1 triptófano hidroxilasa, Tph2, triptófano hidroxilasa 2, Neurexina 1, GSK3, GSK3a, GSK3b, 5-HTT (Slc6a4), COMT, DRD (Drd1a), SLC6A3, DAOA, DTNBP1, Dao (Dao1)); trastornos relacionados con secretasa (APH-1 (alfa y beta), Presenilina (Psen1), nicastrina (Ncstn), PEN-2, Nos1, Parp1, Nat1, Nat2); trastornos por repetición trinucleotídica (HTT (enfermedad de Huntington), SBMA/SMAX1/AR (enfermedad de Kennedy), FXN/X25 (ataxia de Friedrich), ATX3 (enfermedad de Machado-Joseph), ATXN1 y ATXN2 (ataxias medulocerebelosas), DMPK (distrofia miotónica), Atrofin-1 y Atn1 (enfermedad de DRPLA), CBP (Creb-BP - inestabilidad global), VLDLR (enfermedad de Alzheimer), Atxn7, Atxn10).</p>
Enfermedades y trastornos oculares	<p>Degeneración macular senil (Abcr, Ccl2, Cc2, cp (ceruloplasmina), Timp3, catepsinaD, Vldlr, Ccr2); cataratas (CRYAA, CRYA1, CRYBB2, CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6, AN2, MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19, CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM, MIP, AQP0, CRYAB, CRYA2, CTPP2, CRYBB1, CRYGD, CRYG4, CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1, GJA8, CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1); opacidad y distrofia corneal (APOA1, TGFBI, CSD2, CDGG1, CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD, PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD); córnea plana congénita (KERA, CNA2); glaucoma (MYOC, TIGR, GLC1A, JOAG, GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1, NTG, NPG, CYP1B1, GLC3A); amaurosis congénita de Leber (CRB1, RP12, CRX, CORD2, CRD, RPRIP1, LCA6, CORD9, RPE65, RP20, AIPL1, LCA4, GUCY2D, GUC2D, LCA1, CORD6, RDH12, LCA3); distrofia macular (ELOVL4, ADMD, STGD2, STGD3, RDS, RP7, PRPH2, PRPH, AVMD, AOFMD, VMD2).</p>

Tabla C:

FUNCIÓN CELULAR	GENES
Señalización de PI3K/AKT	PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2;
	PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1;
	AKT2; IKBKB; PIK3CA; CDK8; CDKN1B; NFKB2; BCL2;
	PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2;
	ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3;
	PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7;
	YWHAZ; ILK; TP53; RAF1; IKBKG; RELB; DYRK1A;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
	CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1;
	CHUK; PDPK1; PPP2R5C; CTNNA1; MAP2K1; NFKB1;
	PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2;
	TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK;
	HSP90AA1; RPS6KB1
Señalización de ERK/MAPK	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2;
	EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6;
	MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1;
	PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A;
	PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN;
	EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC;
	CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ;
	PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1;
	PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1;
	CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK
Receptor de glucocorticoesteroides	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1;
Señalización	MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I;
	PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2;
	MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1;
	MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13;
	RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1;
	PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3;
	MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP;
	CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2;
	PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1;
	ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1;
	STAT1; IL6; HSP90AA1
Señalización de guía axonal	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12;
	IGF1; RAC1; RAP1A; EIF4E; PRKCZ; NRP1; NTRK2;
	ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
	PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2;
	CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11;
	PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1;
	FYN; HGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1;
	GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3;
	CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B;
	AKT3; PRKCA
Señalización del receptor de efrina	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1;
	PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2;
	MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2;
	DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14;
	CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1;
	KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2;
	PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4; AKT1; JAK2; STAT3; ADAM10;
	MAP2K1; PAK3; ITGB3; CDC42; VEGFA; HGA2;
	EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4;
	AKT3; SGK
Actina del citoesqueleto	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1;
Señalización	PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6;
	ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8;
	PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8;
	F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD;
	PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7;
	PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3;
	ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL;
	BRAF; VAV3; SGK
Enfermedad de Huntington	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2;
Señalización	MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2;
	PIK3CA; HDAC5; CREB1; PRKCI; HSPA5; REST;
	GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
	GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2;
	HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A;
	HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1;
	PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX;
	ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3
Señalización de apoptosis	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1;
	BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB;
	CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8;
	BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF;
	RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2;
	CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2;
	BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK;
	CASP3; BIRC3; PARP1
Señalización del receptor de linfocitos B	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11;
	AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1;
	MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9;
	EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB;
	MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1;
	NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN;
	GSK3B; ATF4; AKT3; VAV3; RPS6KB1
Extravasación de leucocitos	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA;
Señalización	RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11;
	MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12;
	PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB;
	MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK;
	MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2;
	CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK;
	CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
Señalización de integrina	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A;
	TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2;
	CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8;
	CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP;
	RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1;
	TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2;
	CRKL; BRAF; GSK3B; AKT3
Respuesta de fase aguda	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11;
Señalización	AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14;
	PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS;
	MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1;
	TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1;
	IKBKG; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1;
	CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN;
	AKT3; IL1R1; IL6
Señalización de PTEN	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11;
	MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA;
	CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1;
	MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR;
	RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2;
	AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1;
	NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2;
	GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
Señalización de p53	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A;
	BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2;
	PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1;
	PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9;
	CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A;
	HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1;
	SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN;
	SNAI2; GSK3B; BAX; AKT3
Receptor de hidrocarburos arilo	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1;
Señalización	NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
	SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1;
	MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1;
	SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF;
	CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1;
	CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1;
	HSP90AA1
Metabolismo xenobiótico	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1;
Señalización	NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A;
	PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1;
	ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD;
	GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL;
	NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1;
	CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1;
	NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1;
	HSP90AA1
Señalización de SAPK/JNK	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1;
	GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA;
	FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1;
	GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A;
	TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2;
	PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1;
	CRKL; BRAF; SGK
Señalización de PPAR/RXR	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN;
	RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2;
	ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8;
	IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A;
	NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7;
	CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1;
	TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1;
	ADIPOQ
Señalización de NF-KB	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ; TRAF6;
	TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
	MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2;
	KRAS; RELA; PIK3C2A; TRAF2; TLR4; PDGFRB; TNF;
	INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1;
	PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10;
	GSK3B; AKT3; TNFAIP3; IL1R1
Señalización de neuregulina	ERBB4; PRKCE; ITGAM; ITGA5; PTEN; PRKCZ; ELK1;
	MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI;
	CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS;
	PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2;
	ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3;
	EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL;
	AKT3; PRKCA; HSP90AA1; RPS6KB1
Wnt y beta catenina	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO;
Señalización	AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A;
	WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2; ILK;
	LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1;
	PPP2R5C; WNT5A; LRP5; CTNNB1; TGFBR1; CCND1;
	GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B;
	AKT3; SOX2
Señalización del receptor de insulina	PTEN; INS; EIF4E; PTPN11; PRKCZ; MAPK1; TSC1;
	PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3;
	MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1;
	SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN;
	MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1;
	GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK;
	RPS6KB1
Señalización de IL-6	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK3;
	MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1;
	MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG;
	RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
	MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6
Colestasis hepática	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA;
	RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8;
	PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1;
	TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8;
	CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4;
	JUN; IL1R1; PRKCA; IL6
Señalización de IGF-1	IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2;
	PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8;
	IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A;
	YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1;
	PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3;
	FOXO1; SRF; CTGF; RPS6KB1
Respuesta de agresión oxidativa mediada por NRF2	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1;
	NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL;
	NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP;
	MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1;
	GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
Fibrosis hepática/activación de células hepáticas estrelladas	EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF;
	SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9;
	IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8;
	PDGFRA; NFKB1; TGFBR1; SMAD4; VEGFA; BAX;
	IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
Señalización de PPAR	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB;
	NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3;
	NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2;
	PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG;
	RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA;
	MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
Señalización de Fc épsilon RI	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11;
	AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD;
	MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN;
	MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3;
	VAV3; PRKCA
Señalización del receptor acoplado a proteína G	PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB;
	PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB;
	PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1;
	IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK;
	PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3;
	PRKCA
Metabolismo de inositol fosfato	PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6;
	MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2;
	PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A; PIK3R1;
	MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK
Señalización de PDGF	EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB;
	PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC;
	PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2;
	PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC;
	JUN; CRKL; PRKCA; SRF; STAT1; SPHK2
Señalización de VEGF	ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF;
	AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3;
	BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN;
	RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN;
	VEGFA; AKT3; FOXO1; PRKCA
Señalización de linfocitos citolíticos naturales	PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11;
	KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB;
	PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6;
	PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1;
	PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
Ciclo celular: G1/S	HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC;
Regulación de puntos de control	ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11;
	HDAC9; CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1;
	E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1;
	GSK3B; RBL1; HDAC6
Señalización del receptor de linfocitos T	RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS;
	NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	RELA; PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB; FYN;
	MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10;
	JUN; VAV3
Señalización del receptor de muerte	CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD;
	FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8;
	DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB;
	CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3;
	BIRC3
Señalización de FGF	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11;
	AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8;
	MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1;
	AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4;
	AKT3; PRKCA; HGF
Señalización de GM-CSF	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A;
	STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3;
	ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3;
	STAT1
Señalización de esclerosis lateral amiotrófica	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2;
	PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1;
	PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1;
	APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
Señalización de JAK/Stat	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B;
	PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A;
	PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3;
	STAT1
Nicotinato y nicotinamida	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1;
Metabolismo	PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1;
	PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2;
	MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK
Señalización de quimiocinas	CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ;
	CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13;
	RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1;
	MAP2K2; MAP2K1; JUN; CCL2; PRKCA
Señalización de IL-2	ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS;
	STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2;
	JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3
Depresión sináptica	PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS;
a largo plazo	PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3;
	KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA;
	YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA
Receptor de estrógenos	TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2;
Señalización	SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1;
	HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP;
	MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2
Ruta de ubiquitinación	TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4;
de proteínas	CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7;
	USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8;
	USP1; VHL; HSP90AA1; BIRC3
Señalización de IL-10	TRAF6; CCR1; ELK1; IKKB; SP1; FOS; NFKB2;
	MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF;
	IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1;
	JUN; IL1R1; IL6
Activación de VDR/RXR	PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1;
	NCOR2; SP1; PRKCI; CDKN1B; PRKD1; PRKCD;
	RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1;
	LRP5; CEBPB; FOXO1; PRKCA

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
Señalización de TGF-beta	EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1;
	FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2;
	SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2;
	MAP2K1; TGFBR1; SMAD4; JUN; SMAD5
Señalización del receptor de tipo Toll	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13;
	RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK;
	NFKB1; TLR2; JUN
Señalización de MAPK p38	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS;
	CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2;
	MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1;
	SRF; STAT 1
Señalización de neurotrofina/TRK	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS;
	PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A;
	RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1;
	CDC42; JUN; ATF4
Activación de FXR/RXR	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8;
	APOB; MAPK10; PPARG; MTPP; MAPK9; PPARGC1A;
	TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
Potenciación sináptica	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1;
a largo plazo	PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS;
	PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1;
	ATF4; PRKCA
Señalización de calcio	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1;
	CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11;
	HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4;
	HDAC6
Señalización de EGF	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1;
	STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
Señalización de hipoxia en el sistema cardiovascular	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT;
	HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM;
	VEGFA; JUN; ATF4; VHL; HSP90AA1
Inhibición mediada por LPS/IL-1	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1;

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
de la función de RXR	MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2;
	TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1
Activación de LXR/RXR	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA;
	NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1;
	SREBF1; IL1R1; CCL2; IL6; MMP9
Procesamiento de amiloide	PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2;
	CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1;
	PSEN1; CSNK1A1; GSK3B; AKT3; APP
Señalización de IL-4	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1;
	PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1;
	FRAP1; AKT3; RPS6KB1
Ciclo celular: G2/M	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC;
Punto de control de daño en el ADN	CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A;
Regulación	PRKDC; ATM; SFN; CDKN2A
Señalización de óxido nítrico en el sistema cardiovascular	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3;
	CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1;
	VEGFA; AKT3; HSP90AA1
Metabolismo de purina	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4;
	PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C;
	NT5E; POLD1; NME1
Señalización mediada por AMPc	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3;
	SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
Disfunción mitocondrial	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9;
	PARK7; PSEN1; PARK2; APP; CASP3
Señalización de Notch	HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2;
	PSEN1; NOTCH3; NOTCH1; DLL4
Retículo endoplásmico	HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4;
Ruta de sobrecarga	EIF2AK3; CASP3
Metabolismo de pirimidina	NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B;
	NT5E; POLD1; NME1
Señalización de enfermedad de Parkinson	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7;
	PARK2; CASP3
Señalización cardíaca y beta adrenérgica	GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC;
	PPP2R5C
Glucólisis/gluconeogénesis	HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1
Señalización de interferón	IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
Señalización de Sonic Hedgehog	ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRK1B
Glicerofosfolípidos	PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2
Metabolismo	
Degradación de fosfolípidos	PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2
Metabolismo de triptófano	SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1
Degradación de lisina	SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C
Reparación de escisión nucleotídica	ERCC5; ERCC4; XPA; XPC; ERCC1
Ruta	
Almidón y sacarosa	UCHL1; HK2; GCK; GPI; HK1
Metabolismo	
Metabolismo de aminoglúcidos	NQO1; HK2; GCK; HK1
Ácido araquidónico	PRDX6; GRN; YWHAZ; CYP1B1
Metabolismo	
Señalización del ritmo circadiano	CSNK1E; CREB1; ATF4; NR1D1
Sistema de coagulación	BDKRB1; F2R; SERPINE1; F3
Receptor de dopamina	PPP2R1A; PPP2CA; PPP1CC; PPP2R5C
Señalización	
Metabolismo de glutatión	IDH2; GSTP1; ANPEP; IDH1
Metabolismo de glicerolípidos	ALDH1A1; GPAM; SPHK1; SPHK2
Metabolismo de ácido linoleico	PRDX6; GRN; YWHAZ; CYP1B1
Metabolismo de metionina	DNMT1; DNMT3B; AHCY; DNMT3A
Metabolismo de piruvato	GLO1; ALDH1A1; PKM2; LDHA
Arginina y prolina	ALDH1A1; NOS3; NOS2A
Metabolismo	
Señalización de eicosanoides	PRDX6; GRN; YWHAZ
Fructosa y manosa	HK2; GCK; HK1
Metabolismo	
Metabolismo de galactosa	HK2; GCK; HK1
Biosíntesis de estilbeno, cumarina y lignina	PRDX6; PRDX1; TYR
Presentación de antígenos	CALR; B2M
Ruta	
Biosíntesis de esteroides	NQO1; DHCR7
Metabolismo de butanoato	ALDH1A1; NLGN1
Ciclo del citrato	IDH2; IDH1
Metabolismo de ácidos grasos	ALDH1A1; CYP1B1
Glicerofosfolípidos	PRDX6; CHKA
Metabolismo	
Metabolismo de histidina	PRMT5; ALDH1A1
Metabolismo de inositol	ERO1L; APEX1

ES 2 780 904 T3

FUNCIÓN CELULAR	GENES
Metabolismo de xenobióticos por citocromo p450	GSTP1; CYP1B1
Metabolismo de metano	PRDX6; PRDX1
Metabolismo de fenilalanina	PRDX6; PRDX1
Metabolismo de propanoato	ALDH1A1; LDHA
Selenoaminoácido	PRMT5; AHCY
Metabolismo	
Metabolismo de esfingolípidos	SPHK1; SPHK2
Aminofosfonato	PRMT5
Metabolismo	
Andrógenos y estrógenos	PRMT5
Metabolismo	
Ascorbato y aldarato	ALDH1A1
Metabolismo	
Biosíntesis de ácidos biliares	ALDH1A1
Metabolismo de cisteína	LDHA
Biosíntesis de ácidos grasos	FASN
Receptor de glutamato	GNB2L1
Señalización	
Respuesta de agresión oxidativa mediada por NRF2	PRDX1
Pentosa fosfato	GPI
Ruta	
Pentosa y glucuronato	UCHL1
Interconversiones	
Metabolismo de retinol	ALDH1A1
Metabolismo de riboflavina	TYR
Metabolismo de tirosina	PRMT5, TYR
Biosíntesis de ubiquinona	PRMT5
Degradación de valina, leucina e isoleucina	ALDH1A1
Metabolismo de glicina, serina y treonina	CHKA
Degradación de lisina	ALDH1A1
Dolor/gusto	TRPM5; TRPA1
Dolor	TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2;
	Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5;
	Prkaca;
	Prkacb; Prkar1a; Prkar2a
Función mitocondrial	AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2
Neurología del desarrollo	BMP-4; Cordinina (Chrd); Noggin (Nog); WNT (Wnt2);

FUNCIÓN CELULAR	GENES
	Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b;
	Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); beta-catenina;
	Dkk-1; proteínas relacionadas con Frizzled; Otx-2; Gbx2; FGF-8;
	Reelina; Dab1; unc-86 (Pou4f1 o Brn3a); Numb; Reln

Realizaciones de la divulgación también se refieren a métodos y composiciones (invención) relacionadas con la inactivación de genes, amplificación de genes y reparación de mutaciones particulares asociadas con inestabilidad de repeticiones de ADN y trastornos neurológicos (Robert D. Wells, Tetsuo Ashizawa, Genetic Instabilities and Neurological Diseases, segunda edición, Academic Press, 13 de octubre de 2011 - Medical). Aspectos específicos de secuencias de repetición en tándem se ha descubierto que son responsables de más de veinte enfermedades humanas (New insights in repeat instability: role of RNA-DNA hybrids. Mclvor EI, Polak U, Napierala M. RNA Biol. 2010 sep-oct; 7(5):551-8). El sistema de CRISPR-Cas puede aprovecharse para corregir estos defectos de inestabilidad genómica.

Un aspecto adicional de la invención se refiere a utilizar el sistema de CRISPR-Cas para corregir defectos en los genes EMP2A y EMP2B que se han identificado como asociados con enfermedad de Lafora. La enfermedad de Lafora es una afección recesiva autosómica que se caracteriza por epilepsia mioclónica progresiva que puede empezar como ataques epilépticos en la adolescencia. Unos pocos casos de la enfermedad pueden estar causados por mutaciones en genes aún por identificar. La enfermedad causa ataques, espasmos musculares, dificultad para caminar, demencia y finalmente muerte. No hay tratamiento actualmente que se haya mostrado eficaz contra la progresión de la enfermedad. Otras anomalías genéricas asociadas con la epilepsia también puede abordarse mediante el sistema de CRISPR-Cas y la genética subyacente se describe adicionalmente en Genetics of Epilepsy and Genetic Epilepsies, editado por Giuliano Avanzini, Jeffrey L. Noebels, Mariani Foundation Paediatric Neurology;20; 2009).

En otro aspecto más de la invención, el sistema de CRISPR-Cas puede usarse para corregir defectos oculares que surgen de varias mutaciones genéticas descritas adicionalmente en Genetic Diseases of the Eye, segunda edición, editado por Elias I. Traboulsi, Oxford University Press, 2012.

Varios aspectos adicionales de la invención se refieren a corregir defectos asociados con una amplia gama de enfermedades genéticas que se describen adicionalmente en el sitio web del National Institutes of Health bajo la subsección del tema Genetic Disorders (sitio web en health.nih.gov/topic/GeneticDisorders). Las enfermedades cerebrales genéticas pueden incluir, aunque sin limitación, adrenoleucodistrofia, agenesia del cuerpo caloso, síndrome de Aicardi, enfermedad de Alpers, enfermedad de Alzheimer, síndrome de Barth, enfermedad de Batten, CADASIL, degeneración cerebelosa, enfermedad de Fabry, enfermedad de Gerstmann-Straussler-Scheinker, enfermedad de Huntington y otros trastornos por repetición de tripletes, enfermedad de Leigh, síndrome de Lesch-Nyhan, enfermedad de Menkes, miopatías mitocondriales y colpocefalia por NINDS. Estas enfermedades se describen adicionalmente en el sitio web del National Institutes of Health bajo la subsección Genetic Brain Disorders.

En algunas realizaciones, la afección es neoplasia. En algunas realizaciones, cuando la afección es neoplasia, los genes a abordar son cualquiera de los enumerados en la tabla A (en este caso PTEN y así sucesivamente). En algunas realizaciones, la afección puede ser degeneración macular senil. En algunas realizaciones, la afección puede ser un trastorno esquizofrénico. En algunas realizaciones, la afección puede ser un trastorno por repetición trinucleotídica. En algunas realizaciones, la afección puede ser síndrome del X frágil. En algunas realizaciones, la afección puede ser un trastorno relacionado con secretasa. En algunas realizaciones, la afección puede ser un trastorno relacionado con priones. En algunas realizaciones, la afección puede ser ALS. En algunas realizaciones, la afección puede ser drogodependencia. En algunas realizaciones, la afección puede ser autismo. En algunas realizaciones, la afección puede ser enfermedad de Alzheimer. En algunas realizaciones, la afección puede ser inflamación. En algunas realizaciones, la afección puede ser enfermedad de Parkinson.

Ejemplos de proteínas asociadas con enfermedad de Parkinson incluyen, aunque sin limitación, α -sinucleína, DJ-1, LRRK2, PINK1, Parkina, UCHL1, Sinfilina-1 y NURR1.

Ejemplos de proteínas relacionadas con adicción pueden incluir ABAT, por ejemplo.

Ejemplos de proteínas relacionadas con inflamación pueden incluir la proteína-1 quimioatrayente de monocitos (MCP1) codificada por el gen Ccr2, el receptor de tipo 5 de quimiocinas C-C (CCR5) codificado por el gen Ccr5, el receptor IIB de IgG (FCGR2b, también denominado CD32) codificado por el gen Fcgr2b o la proteína Fc épsilon R1g (FCER1g) codificada por el gen Fcer1g, por ejemplo.

Ejemplos de proteínas asociadas a enfermedades cardiovasculares pueden incluir IL1B (interleucina 1, beta), XDH (xantina deshidrogenasa), TP53 (proteína tumoral p53), PTGIS (prostaglandina I2 (prostaciclina) sintasa), MB (mioglobina), IL4 (interleucina 4), ANGPT1 (angiopoyetina 1), ABCG8 (dominio de unión a ATP, subfamilia G (WHITE), miembro 8) o CTSK (catepsina K), por ejemplo.

5 Ejemplos de proteínas asociadas a enfermedad de Alzheimer pueden incluir la proteína receptora de lipoproteína de muy baja densidad (VLDLR) codificada por el gen VLDLR, la enzima 1 activadora de modificador de tipo ubiquitina (UBA1) codificada por el gen UBA1 o la proteína de la subunidad catalítica de la enzima E1 activadora de NEDD8 (UBE1C) codificada por el gen UBA3, por ejemplo.

10 Ejemplos de proteínas asociadas con trastornos del espectro autista pueden incluir la proteína 1 asociada al receptor de benzodiazepina (periférica) (BZRAP1) codificada por el gen BZRAP1, la proteína del miembro 2 de la familia AF4/FMR2 (AFF2) codificada por el gen AFF2 (también denominado MFR2), la proteína del homólogo 1 autosómico de retraso mental por X frágil (FXR1) codificada por el gen FXR1 o la proteína del homólogo 2 autosómico de retraso mental por X frágil (FXR2) codificada por el gen FXR2, por ejemplo.

15 Ejemplos de proteínas asociadas a degeneración macular pueden incluir el dominio de unión a ATP, la proteína del miembro 4 de la subfamilia A (ABC1) (ABCA4) codificada por el gen ABCR, la proteína de apolipoproteína E (APOE) codificada por el gen APOE o la proteína de ligando 2 de quimiocina (motivo C-C) (CCL2) codificada por el gen CCL2, por ejemplo.

20 Ejemplos de proteínas asociadas a esquizofrenia pueden incluir NRG1, ErbB4, CPLX1, TPH1, TPH2, NRXN1, GSK3A, BDNF, DISC1, GSK3B y combinaciones de las mismas.

25 Ejemplos de proteínas implicadas en supresión tumoral pueden incluir ATM (ataxia telangiectasia mutada), ATR (ataxia telangiectasia y Rad3 relacionada), EGFR (receptor del factor de crecimiento epidérmico), ERBB2 (homólogo 2 del oncogén vírico de leucemia linfoblástica v-erb-b2), ERBB3 (homólogo 3 del oncogén vírico de leucemia linfoblástica v-erb-b2), ERBB4 (homólogo 4 del oncogén vírico de leucemia linfoblástica v-erb-b2), Notch 1, Notch2, Notch 3 o Notch 4, por ejemplo.

30 Ejemplos de proteínas asociadas con un trastorno por secretasa pueden incluir PSENEN (homólogo de potenciador 2 de presenilina (*C. elegans*)), CTSB (catepsina B), PSEN1 (presenilina 1), APP (proteína precursora de amiloide beta (A4)), APH1B (homólogo B deficiente en faringe anterior 1 (*C. elegans*)), PSEN2 (presenilina 2 (enfermedad de Alzheimer 4)), o BACE1 (enzima 1 de escisión de APP en sitio beta), por ejemplo.

35 Ejemplos de proteínas asociadas con esclerosis lateral amiotrófica pueden incluir SOD1 (superóxido dismutasa 1), ALS2 (esclerosis lateral amiotrófica 2), FUS (fusionado en sarcoma), TARDBP (proteína de unión a ADN de TAR), VAGFA (factor A de crecimiento del endotelio vascular), VAGFB (factor B de crecimiento del endotelio vascular) y VAGFC (factor C de crecimiento del endotelio vascular), y cualquier combinación de las mismas.

40 Ejemplos de proteínas asociadas con enfermedades por priones pueden incluir SOD1 (superóxido dismutasa 1), ALS2 (esclerosis lateral amiotrófica 2), FUS (fusionado en sarcoma), TARDBP (proteína de unión a ADN de TAR), VAGFA (factor A de crecimiento del endotelio vascular), VAGFB (factor B de crecimiento del endotelio vascular) y VAGFC (factor C de crecimiento del endotelio vascular), y cualquier combinación de las mismas.

45 Ejemplos de proteínas relacionadas con afecciones neurodegenerativas en trastornos por priones pueden incluir A2M (alfa-2-macroglobulina), AATF (factor de transcripción antagonizante de la apoptosis), ACPP (fosfatasa ácida de próstata), ACTA2 (actina alfa 2 de músculo liso de aorta), ADAM22 (dominio de metalopeptidasa ADAM), ADORA3 (receptor de adenosina A3) o ADRA1D (receptor alfa-1D adrenérgico para alfa-1D adrenorreceptor), por ejemplo.

50 Ejemplos de proteínas asociadas con inmunodeficiencia pueden incluir A2M [alfa-2-macroglobulina]; AANAT [arilalquilamina N-acetiltransferasa]; ABCA1 [dominio de unión a ATP, subfamilia A (ABC1), miembro 1]; ABCA2 [dominio de unión a ATP, subfamilia A (ABC1), miembro 2]; o ABCA3 [dominio de unión a ATP, subfamilia A (ABC1), miembro 3]; por ejemplo.

55 Ejemplos de proteínas asociadas con trastornos por repeticiones trinucleotídicas incluyen AR (receptor de andrógenos), FMR1 (1 de retraso mental por X frágil), HTT (huntingtina), o DMPK (proteína cinasa de distrofia miotónica), FXN (frataxina), ATXN2 (ataxina 2), por ejemplo.

60 Ejemplos de proteínas asociadas con trastornos de neurotransmisión incluyen SST (somatostatina), NOS1 (óxido nítrico sintasa 1 (neuronal)), ADRA2A (receptor adrenérgico, alfa-2A), ADRA2C (receptor adrenérgico, alfa-2C), TACR1 (receptor 1 de taquicinina) o HTR2c (receptor 2C de 5-hidroxitriptamina (serotonina)), por ejemplo.

65 Ejemplos de secuencias asociadas al neurodesarrollo incluyen A2BP1 [proteína 1 de unión a ataxina 2], AADAT [aminoadipato aminotransferasa], AANAT [arilalquilamina N-acetiltransferasa], ABAT [4-aminobutirato

aminotransferas], ABCA1 [dominio de unión a ATP, subfamilia A (ABC1), miembro 1] o ABCA13 [dominio de unión a ATP, subfamilia A (ABC1), miembro 13], por ejemplo.

5 Ejemplos adicionales de afecciones preferidas tratables con el presente sistema incluyen las que pueden seleccionarse de: síndrome de Aicardi-Goutières; enfermedad de Alexander; síndrome de Allan-Herndon-Dudley; trastornos relacionados con POLG; alfa-manosidosis (tipo II y III); síndrome de Alström; síndrome de Angelman; ataxia-telangiectasia; lipofuscinosis cerioidea neuronal; beta-talasemia; atrofia óptica bilateral y atrofia óptica de tipo I (infantil); retinoblastoma (bilateral); enfermedad de Canavan; síndrome cerebrooculofacioesquelético 1 [COFS1]; xantomatosis cerebrotendinosa; síndrome de Cornelia de Lange; trastornos relacionados con MAPT; enfermedades genéticas por priones; síndrome de Dravet; enfermedad de Alzheimer familia de aparición temprana; ataxia de Friedreich [FRDA];
 10 síndrome de Fryns; fucosidosis; distrofia muscular congénita de Fukuyama; galactosialidosis; enfermedad de Gaucher; acidemias orgánicas; linfocitosis hemofagocítica; síndrome de progeria de Hutchinson-Gilford; mucopolidosis II; enfermedad por almacenamiento de ácido silícico libre infantil; neurodegeneración asociada a PLA2G6; síndrome de Jervell y Lange-Nielsen; epidermólisis ampollosa de uniones; enfermedad de Huntington; enfermedad de Krabbe (infantil); síndrome de Leigh asociado a ADN mitocondrial y NARP; síndrome de Lesch-Nyhan; lisencefalia asociada a LIS1; síndrome de Lowe; enfermedad urinaria de jarabe de arce; síndrome de duplicación de MECP2; trastorno del transporte de cobre relacionados con ATP7A; distrofia muscular relacionada con LAMA2; deficiencia de arilsulfatasa A; mucopolisacaridosis de tipo I, II o III; trastornos de la biogénesis de peroxisomas, espectro del síndrome de Zellweger; neurodegeneración con trastornos por acumulación de hierro cerebral; deficiencia de esfingomielinasa
 20 ácida; enfermedad de Niemann-Pick de tipo C; encefalopatía por glicina; trastornos relacionados con ARX; trastornos del ciclo de urea; osteogénesis imperfecta relacionada con COL1A1/2; síndromes de eliminación de ADN mitocondrial; trastornos relacionados con PLP1; síndrome de Perry; síndrome de Phelan-McDermid; enfermedad de almacenamiento de glucógeno de tipo II (enfermedad de Pompe) (infantil); trastornos relacionados con MAPT; trastornos relacionados con MECP2; condrodistrofia punteada rizomélica de tipo 1; síndrome de Roberts; enfermedad de Sandhoff; enfermedad de Schindler - tipo 1; deficiencia de adenosina desaminasa; síndrome de Smith-Lemli-Opitz; atrofia muscular de la columna; ataxia espinoocerebelosa de aparición infantil; deficiencia de hexosaminidasa A; displasia tanatóforica de tipo 1; trastornos relacionados con colágeno de tipo VI; síndrome de Usher de tipo I; distrofia muscular congénita; síndrome de Wolf-Hirschhorn; deficiencia de lipasa ácida lisosómica; y xerodermia pigmentosa.

30 Como será evidente, se prevé que el presente sistema pueda usarse para abordar cualquier secuencia polinucleotídica de interés. Algunos ejemplos de afecciones o enfermedades que podrían tratarse de forma útil usando el presente sistema se incluyen en las tablas anteriores y ejemplos de genes actualmente asociados con estas afecciones también se proporcionan ahí. Sin embargo, los genes ejemplificados no son exhaustivos.

35 Por ejemplo, "StCas9 de tipo silvestre" se refiere a Cas9 de tipo silvestre de *S. thermophilus*, cuya secuencia proteínica se da en la base de datos SwissProt con el número de acceso G3ECR1. Asimismo, la Cas9 de *S. pyogenes* se incluye en SwissProt con el número de acceso Q99ZW2. SaCas9 también está bien caracterizada.

40 La capacidad de usar sistemas de CRISPR-Cas para realizar edición y manipulación génica eficaz y rentable permitirá una rápida selección y comparación de manipulaciones genéticas individuales y combinadas para transformar dichos genomas para producción mejorada y rasgos potenciados. A este respecto, se hace referencia a las patentes y publicaciones de Estados Unidos: patente de Estados Unidos n.º 6.603.061 - Agrobacterium-Mediated Plant Transformation Method; patente de Estados Unidos n.º 7.868.149 - Plant Genome Sequences and Uses Thereof and documento US 2009/0100536 - Transgenic Plants with Enhanced Agronomic Traits. En la práctica de la invención, los contenidos y divulgación de Morrell *et al.* "Crop genomics: advances and applications," Nat Rev Genet. 29 de diciembre de 2011;13(2):85-96.

50 La presente invención se basa en varios efectos técnicos, que se definen en general, *inter alia*, por doble mellado optimizado, generación de salientes 3', inhibición de NHEJ y eficacia de HDR mejorada.

55 La presente invención se basa en el efecto técnico de eficacia de HDR mejorada usando el mutante SpCas9N863A o un ortólogo del mismo que tenga una mutación correspondiente a SpCas9N863A. Específicamente, la eficacia de HDR mejorada es el resultado de la inhibición de eventos de NHEJ y, por tanto, un sesgo (es decir, aumento) en los eventos de HDR. Dicho ortólogo puede ser una Cas9 de *S. aureus* mutada.

La presente invención también se basa en el efecto técnico de doble mellado optimizado debido a la selección de la secuencia diana óptima de modo que las secuencias 5' PAM se enfrenten entre sí.

60 Además, la presente invención se basa en el efecto técnico de que la actividad nucleasa del mutante SpCas9N863A o un ortólogo del mismo que tiene una mutación correspondiente a SpCas9N863A siempre provoca escisión de la hebra en la hebra no complementaria. En estas condiciones, se generan salientes 3' si las secuencia diana (como se definen por el ARNs) en las hebras individuales están dispuestas de modo que las secuencias 5' PAM correspondientes (ubicadas inmediatamente 3' a las secuencias diana) se enfrentan entre sí. Dicho ortólogo puede ser una Cas9 de *S. aureus* mutada.

65

Además, la presente invención se basa en un efecto técnico de que los salientes 3' provocan inhibición de NHEJ. El efecto técnico de la generación dirigida de salientes 3' es eficacia de HDR mejorada, ya que dichos salientes provocan inhibición de eventos de NHEJ y, por tanto, un sesgo (es decir, aumento) en los eventos de HDR (es decir, eficacia de HDR mejorada y/o formación de indel reducida).

Una eficacia de HDR mejorada se considera una mayor frecuencia de eventos de HDR (y/o formación de indel reducida) como resultado de la actividad doble nickasa resultante del uso del mutante SpCas9N863A o un ortólogo que tiene una mutación correspondiente a SpCas9N863A en comparación con la actividad doble nickasa resultante de una SpCas9 que no comprende la mutación N863A o un ortólogo que no comprende una mutación correspondiente a SpCas9N863A. Dicho ortólogo puede ser una Cas9 de *S. aureus* mutada.

Realizando los métodos de la invención de modificación de un organismo o un locus genómico de interés, los expertos en la materia llegan inevitablemente a modificaciones inespecíficas minimizadas.

EJEMPLOS

Los siguientes ejemplos se dan con el fin de ilustrar diversas realizaciones de la invención y no pretenden limitar la presente invención de ninguna manera. Los presentes ejemplos, junto con los métodos descritos en este documento son actualmente representativos de las realizaciones preferidas, son ejemplares y no se pretenden como limitaciones del alcance de la invención. Cambios en los mismos y otros usos que están englobados dentro del alcance de las reivindicaciones se les ocurrirán a los expertos en la materia.

Ejemplo 1: Selección de dianas

Las dianas de SpCas9 son cualquier secuencia de ADN de 20 pb seguida en el extremo 3' por 5'-NGG-3'. El sitio web CRISPR DESIGN del Lab de Zhang, MIT, proporciona una herramienta en línea (<http://tools.genome-engineering.org>) que acepta una región de interés como entrada y proporciona como salida una lista de todos los sitios diana de ARNGs potenciales dentro de esa región. Cada sitio diana de ARNGs entonces se asocia con una lista de dianas inespecíficas genómicas predichas.

La herramienta también genera pares de ARNGs de doble mellado automáticamente. La consideración más importante para el diseño de ARNGs de doble mellado es el espaciado entre las dos dianas (Ran *et al.*, 2013). Si la "compensación" entre dos guías se define como la distancia entre los extremos PAM distales (5') de un par de ARNGs, una compensación de -4 a 0 pb es ideal, aunque compensaciones tan grandes como de 100 pb pueden inducir indel mediadas por DSB. Los pares de ARNGs para doble mellado están dirigidas a hebras de ADN opuestas diana.

Ejemplo 2: Construcción de ARNGs plasmídico

Se construyen vectores de expresión de ARNGs clonando secuencias diana de 20 pb en una cadena principal plasmídica que codifica un casete de expresión de ARNGs dirigido por el promotor U6 humano y una Cas9-D10A dirigida por CBh (pSpCas9n(BB), Addgene n.º 48873). La nickasa N863A se intercambia con D10A en todos los casos. Se recomienda preparar este plásmido como una maxipreparación sin endotoxinas. Los oligos generalizados usados para clonar una nueva diana en pSpCas9n(BB) se describen en la tabla 1 y se adquieren en Integrated DNA Technologies (IDT). Obsérvese que la secuencia de PAM requerida para el reconocimiento de dianas por Cas9 no está presente como parte del propio ARNGs.

Tabla 1:

Cebador	Secuencia (5' a 3')	Descripción
ARNGs-dir	CACCGNNNNNNNNNNNNNNNNNNNN	Saliente cohesivo más diana genómica de 20 pb específica a clonar en cadenas principales de ARNGs
ARNGs-inv	AAACNNNNNNNNNNNNNNNNNNNNC	Oligo de hibridación complementario para clonación de nueva diana en cadenas principales de ARNGs.

1. Una secuencia diana se clona en un vector de cadena principal de ARNGs. Los oligos ARNGs-dir y ARNGs-inv se resuspenden hasta 100 uM. Obsérvese que estos oligos incluyen una guanina adjunta (minúscula) no presente en el sitio diana para aumentar la transcripción a partir del promotor U6.

2. 1 µl de cada oligo se combina con 1 µl de tampón de ligamiento T4, 10x (New England Biolabs (NEB) B0202S), 0,5 µl de PNK T4 (NEB M0201S) y 6,5 µl de ddH₂O para una reacción total de 10 µl. La solución se trata con polinucleótido cinasa para añadir fosfato 5' y los oligos se hibridan en un termociclador con el siguiente protocolo: 37 °C durante 30 min, 95 °C durante 5 min, disminución en rampa hasta 25 °C a 5 °C/min.

3. Los oligos hibridados se diluyen (10 µl de reacción) añadiendo 90 µl de ddH₂O.

4. Se realiza digestión/ligamiento Golden Gate con pSpCas9n(BB) y los oligos hibridados como inserto de clonación. El plásmido resultante contiene sitios de restricción BsmBI gemelos en lugar de la secuencia diana de ARNGs de modo que la digestión deja salientes complementarios a los salientes de los oligos hibridados. En una reacción de 25 µl, se combinan 25 ng de pSpCas9n(BB), 1 µl de oligos hibridados diluidos de la etapa (3), 12,5 µl de tampón de ligamiento rápido, 2x (Enzymatics L6020L), 1 µl de BsmBI FastDigest (ThermoScientific FD1014), 2,5 µl de BSA 10x (NEB B9001), 0,125 µl de ligasa T7 (Enzymatics L6020L) y 7 µl de ddH₂O.

5. Se realiza un control negativo usando las mismas condiciones que anteriormente y sustituyendo los oligos de inserto con ddH₂O.

6. El ligamiento se incuba en un termociclador durante 6 ciclos de 37 °C durante 5 min, 20 °C durante 5 min. El ligamiento es estable para almacenamiento a -20 °C.

7. Se transforman 2 µl de la reacción de ligamiento en una cepa de *E. coli* competente usando el protocolo apropiado, se prefiere la cepa Stbl3, se siembra en placas de selección por ampicilina (100 µg/ml de ampicilina) y se incuba durante una noche a 37 °C. Típicamente la transformación se produce a alta eficacia; no se forman colonias en la placa de control negativo, y se forman cientos cuando los oligos de ARNGs se han clonado satisfactoriamente en la cadena principal.

8. Se cogen dos o más colonias 14 horas después, de la transformación con una punta de pipeta estéril y las bacterias se usan para inocular 3 ml de caldo LB o TB con 100 µg/ml de ampicilina. El cultivo se agita a 37 °C durante 14 horas.

9. El ADN plasmídico se aísla de los cultivos usando el kit Qiagen Spin Miniprep (27104) y la concentración de ADN se determina por espectrofotometría. Estas construcciones se verifican por secuenciación de Sanger a través del armazón de ARNGs para confirmar la correcta inserción de la secuencia diana. Para condiciones de transfección óptimas posteriores, se prepara plásmido sin endotoxinas.

Ejemplo 3: Validación de ARNGs en líneas celulares

Este ejemplo describe la validación funcional de ARNGs en células HEK293FT; las condiciones de cultivo y transfección pueden variar para otros tipos de células.

1. Células HEK293FT (Life Technologies R700-07) se mantienen en medio D10 estéril (DMEM, alto en glucosa (Life Technologies 10313-039) complementado con suero bovino fetal al 10 % vol/vol (Seradigm 1500-500) y HEPES 10 mM (Life Technologies 15630-080)). Para una salud óptima, las células se pasa cada día a una relación de 1:2-2,5 y se mantienen en un 80 % de confluencia.

2. Las células se siembran en placa para transfección. Se siembran 120 000 células por pocillo de una placa tratada con cultivo hístico de 24 pocillos en un volumen total de 500 µl. Los cultivos y las transfecciones se aumentan o disminuyen de escala proporcionalmente para diferentes formatos basándose en el área superficial de crecimiento. Para muchos tipos celulares adherentes, plástico recubierto con poli-D-lisina puede mejorar la adherencia y viabilidad.

3. Después de 18 horas, las placas se comprueban para determinar la confluencia de las células, en general un 90 % es ideal. Se usa reactivo Lipofectamine 2000 (Life Technologies 11668109) para transfectar ADN de acuerdo con el protocolo del fabricante. Para una placa de 24 pocillos, la transfección es de no más de 500 ng de ADN total.

4. Para el suministro de plásmido de pSpCas9n(ARNGs) de un mellado, se transfectan 500 ng; para construcciones de múltiple mellado, por ejemplo, se suministran dos ARNGs para doble mellado, se mezclan diferentes construcciones hasta 500 ng a relaciones equimolares antes de la transfección.

5. Se añaden controles de transfección (pocillos no transfectados y plásmido de GFP), así como controles experimentales (Cas9n sin guías o guías en solitario), en estos experimentos. La transfección por triplicado técnico facilita el análisis.

6. En 6 horas desde la transfección, el medio se cambia a 2 ml de medio D10 reciente precalentado por pocillo. A las 24 horas, se estima la eficacia de transfección examinando los pocillos transfectados con GFP. > 80 % de las células son GFP-positivas.

7. Las células se recogen para la extracción de ADN genómico y/o análisis posterior a las 48-72 horas. Para la recogida en el punto temporal de 72 horas, el medio se cambia de nuevo a las 48 horas para mantener la salud celular óptima.

Cuando se trabaja con diferentes tipos celulares, deben compararse reactivos de transfección alternativos para la eficacia y la toxicidad. También puede ser informativo valorar pSpCas9n(ARNGs) para encontrar la concentración de transfección óptima con la máxima eficacia.

Ejemplo 4: Recolección celular y extracción de ADN

1. Las células se recogen en formato de placa de 24 pocillos aspirando el medio completamente y añadiendo 100 µl de reactivo TrypLE Express (Life Technologies 12604013) para facilitar la disociación.
2. La suspensión celular se recoge en un tubo Eppendorf de 1,5 ml y se centrifuga durante 5 min a 1500 g, el sobrenadante se aspira completamente, y el sedimento celular se resuspende en 200 µl de DPBS (Life Technologies 14190-250) para el lavado.
3. La suspensión celular se centrifuga de nuevo durante 5 min a 1500 g y se resuspende en 50 µl de QuickExtract (Epicentre QE09050).
4. La suspensión de QuickExtract se transfiere a un tubo de PCR de 0,2 ml y se extrae el ADN genómico de acuerdo con el siguiente protocolo de termociclador adaptado de las instrucciones del fabricante: 65 °C durante 15 min, 98 °C durante 10 min.
5. El producto de reacción se centrifuga para sedimentar los desechos celulares y el sobrenadante aclarado se transfiere a un tubo nuevo para análisis adicional.
6. La concentración de ADN de la extracción se determina por espectrofotometría y se normaliza a 100-200 ng/µl con ddH₂O.

Ejemplo 5: Análisis de indel SURVEYOR

- El ensayo SURVEYOR (Transgenomic 706025) es un método para detectar polimorfismos e indel pequeños. Las muestras de ADN se amplifican por PCR y los productos se calientan para desnaturalizar y se enfrían lentamente para formar estructuras bicatenarias heterogéneas. Las estructuras bicatenarias desapareadas entonces se escinden por la nucleasa SURVEYOR y los productos de escisión se analizan por electroforesis en gel.
1. Se realiza PCR sobre el ADN genómico. Los cebadores para PCR SURVEYOR producen un amplicón limpio de aproximadamente 500 pb en muestras de células no transfectadas. Los cebadores de PCR genómicos se diseñan usando un programa informático tal como Primer3. Se configura una reacción de 50 µl que contiene 1 µl de cada cebador SURVEYOR 10 µM, 10 µl de tampón de reacción Herculase II 5x (Agilent 600675), 0,5 µl de dNTP 100 mM, 0,5 µl de polimerasa de fusión Herculase II, 2 µl de MgCl₂ 25 mM y 36 µl de ddH₂O. La solución se desnaturaliza durante 20 s a 95 °C, se hibrida durante 20 s a 60 °C, y se prolonga durante 20 s a 72 °C.
2. Se usa polimerasa de alta fidelidad para evitar positivos falsos.
3. Se procesan 2 µl del producto de PCR en un gel de agarosa al 1 % para asegurar que se ha formado un solo producto de tamaño esperado.
4. El producto de PCR se purifica usando el kit de purificación de PCR QIAquick (28104) de acuerdo con las instrucciones proporcionadas. La concentración de ADN del eluido se mide por espectrofotometría y se normaliza a 20 ng/µl usando ddH₂O.
5. Se mezclan 18 µl de producto de PCR normalizado con 2 µl de tampón de PCR Taq, 10x, para una reacción total de 20 µl. Los productos se funden y rehibridan gradualmente en un termociclador: se funden a 95 °C durante 10 min, la temperatura se disminuye en rampa hasta 85 °C a una tasa de -0,3 °C/s. Se mantienen a 85 °C 1 min, después se llevan en rampa hasta 75 °C a 0,3 °C/s. Se mantienen a 75 °C 1 min, después se llevan en rampa hasta 65 °C... y así sucesivamente, hasta que la temperatura alcanza 25 °C. Desde 25 °C, se disminuye en rampa hasta 4 °C a 0,3 °C/s y se mantiene.
6. Se mezclan 2,5 µl de MgCl₂ 0,15 M, 0,5 µl de ddH₂O, 1 µl de nucleasa S SURVEYOR y 1 µl de potenciador S SURVEYOR con todo el producto hibridado de la etapa (5) para un volumen de reacción total de 25 µl. La digestión se realiza incubando la reacción a 42 °C durante 30 minutos. Las muestras que tienen mutaciones dentro de los amplicones de PCR rehibridados se escinden por SURVEYOR.
7. Los productos de digestión se mezclan con un tinte de carga apropiado y se visualizan por electroforesis en un gel de poliacrilamida TBE al 4-20 % (véase, por ejemplo, **figura 2B**).
8. Se estiman las tasas de modificación genómica calculando en primer lugar las intensidades relativas de los productos de digestión a y b, y la banda c no digerida. La frecuencia de corte f_{cut} entonces se da por $(a + b)/(a + b + c)$. La siguiente fórmula, basada en la distribución de probabilidades binomial de la formación de estructuras bicatenarias estima el porcentaje de indel en la muestra.

$$\%indel = (1 - \sqrt{(1 - f_{cut})})100$$

Ejemplo 6: Inserción por HDR y sin HDR usando Cas9n

- 5 1. Se diseñan brazos de homología ODNmc para que sean lo más largos posible, con al menos 40 nucleótidos de homología en cada lado de la secuencia a introducir. El servicio Ultramer proporcionado por IDT permite la síntesis de oligos de hasta 200 pb de longitud. Los moldes de homología se diluyen hasta 10 μM y se almacenan a -20 °C (véase el ejemplo de diseño, **figura 3**).
- 10 2. El suministro por nucleofección es óptimo para ODNmc. El 4D Nucleofector X Kit S (Lonza V4XC-2032) se usa para células HEK293FT sembradas en placas de 6 pocillos tratadas con cultivo hístico. El fabricante proporciona un protocolo óptimo para nucleofección de estos y otros tipos celulares.
- 15 3. Se mezclan 500 ng de plásmidos pSpCas9n(ARNgs) totales con 1 μl de ODNmc 10 μM para nucleofección.
- La HDR en células de mamífero procede mediante la generación de salientes 3' seguida de invasión de hebra de un locus homólogo por el extremo 3'. El efecto técnico es que la generación de productos de saliente 3' por doble mellado mediado por N863A aumenta la eficacia de HDR.

Ejemplo 7: Análisis de HDR y eventos de inserción

- Los resultados de HDR puede evaluarse y utilizarse de una diversidad de maneras. Aquí, se describe el aislamiento por FACS de células 293FT clonales para pSpCas9n(ARNgs)-GFP. Es importante observa que los procedimientos de FACS puede variar entre tipos celulares.
- 25 1. Se prepara medio de FACS (D10 sin rojo fenol para facilitar la clasificación de fluorescencia): DMEM, alto en glucosa, sin rojo fenol (Life Technologies 31053-028) complementado con suero bovino fetal al 10 % vol/vol y HEPES 10 mM complementado con penicilina al 1 %-estreptomicina (Life Technologies 15140122).
- 30 2. Se preparan placas de 96 pocillos para la clasificación de los clones añadiendo 100 μl de medio D10 convencional a cada pocillo.
3. Unas 24 horas después de la transfección en la sección (7), el medio se aspira completamente y las células se disocian usando suficiente TrypLE Express para cubrir la superficie de crecimiento mínimamente.
- 35 4. Se detiene el tratamiento con tripsina añadiendo medio D10, las células se transfieren a un tubo nuevo de 15 ml y se continúa la trituración suavemente 20 veces. Es crucial que las células están en suspensión de una sola célula antes de proceder.
- 40 5. Las células se centrifugan durante 5 min a 200 g, el sobrenadante se aspira completamente, y el sedimento se resuspende minuciosa y cuidadosamente en 200 μl de medio de FACS.
- 45 6. Las células se filtran a través de un tamiz celular (BD Falcon 352235) para separar por filtración los agregados de células y las células se colocan en hielo.
7. Se clasifican células individuales en las placas preparadas en (2). La máquina FACS puede seleccionar para células GFP+ para enriquecer las células transfectadas. Los pocillos pueden inspeccionarse visualmente para comprobar la presencia de una célula.
- 50 8. Las células se incuban y se expanden durante 2-3 semanas, se cambian de medio a D10 reciente según lo necesario.
9. Cuando las células exceden un 60 % confluencia, las poblaciones clonales se pasan a placas replicadas que contienen medio D10 reciente. Las células se disocian, un 20 % de las células se pasan a placas replicadas y un 80 % se conservan para la extracción de ADN como se describe en la sección (5).
- 55 10. Se realiza genotipado por amplificación por PCR del locus de interés, purificación por PCR y secuenciación de Sanger de los productos.

Ejemplo 7: Nickasas SaCas9 D10A y N580A

60 Editas Medicine presentó datos que demuestran la actividad nickasa de los mutantes D10A y N580A de SaCas9. (Figura 9).

Referencias

- 5 Ding, Q. et al. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* 12, 238-251 (2013).
- Soldner, F. et al. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 146, 318-331 (2011).
- 10 Carlson, D.F. et al. Efficient TALEN-mediated gene knockout in livestock. *Proc Natl Acad Sci U S A* 109, 17382-17387 (2012).
- Geurts, A.M. et al. Knockout Rats via Embryo Microinjection of Zinc-Finger Nucleases. *Science* 325, 433-433 (2009).
- 15 Takasu, Y. et al. Targeted mutagenesis in the silkworm *Bombyx mori* using zinc finger nuclease mRNA injection. *Insect Biochem Molec* 40, 759-765 (2010).
- Watanabe, T. et al. Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. *Nat Commun* 3 (2012).
- 20 Porteus, M.H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* 300, 763 (2003).
- Miller, J.C. et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 25, 778-785 (2007).
- 25 Sander, J.D. et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods* 8, 67-69 (2011).
- Wood, A.J. et al. Targeted genome editing across species using ZFNs and TALENs. *Science* 333, 307 (2011).
- 30 Christian, M. et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757-761 (2010).
- Zhang, F. et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* 29, 149-153 (2011).
- 35 Miller, J.C. et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 29, 143-148 (2011).
- Reyon, D. et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* 30, 460-465 (2012).
- 40 Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509-1512 (2009).
- Moscou, M.J. & Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501 (2009).
- Sanjana, N.E. et al. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc* 7, 171-192 (2012).
- 45 Deveau, H., Garneau, J.E. & Moineau, S. CRISPR-Cas system and its role in phage-bacteria interactions. *Annu Rev Microbiol* 64, 475-493 (2010).
- Horvath, P. & Barrangou, R. CRISPR-Cas, the immune system of bacteria and archaea. *Science* 327, 167-170 (2010).
- 50 Makarova, K.S. et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* 9, 467-477 (2011).
- Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet* 45, 273-297 (2011).
- 55 Cong, L. et al. Multiplex genome engineering using CRISPR-Cas systems. *Science* 339, 819-823 (2013).
- Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* 339, 823-826 (2013b).
- 60 Jinek, M. et al. RNA-programmed genome editing in human cells. *eLife* 2, e00471 (2013).
- Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* 31, 230-232 (2013).

- Garneau, J.E. et al. The CRISPR-Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67-71 (2010).
- 5 Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012).
- Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* 109, E2579-2586 (2012).
- 10 Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* 11, 636-646 (2010).
- Hsu, P.D. & Zhang, F. Dissecting neural function using targeted genome engineering technologies. *ACS Chem Neurosci* 3, 603-610 (2012).
- 15 Perez, E.E. et al. Establishment of HIV-1 resistance in CD4(+) T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* 26, 808-816 (2008).
- Cong, L. et al. Multiplex Genome Engineering Using CRISPR-Cas Systems. *Science* 339, 819-823 (2013).
- 20 Chen, F.Q. et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat Methods* 8, 753-U796 (2011).
- Bedell, V.M. et al. In vivo genome editing using a high-efficiency TALEN system. *Nature* 491, 114-U133 (2012).
- 25 Saleh-Gohari, N. & Helleday, T. Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res* 32, 3683-3688 (2004).
- Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607 (2011).
- 30 Sapranaukas, R. et al. The *Streptococcus thermophilus* CRISPR-Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39, 9275-9282 (2011).
- Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31, 227-229 (2013).
- 35 Wang, H. et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR-Cas-Mediated Genome Engineering. *Cell* 153, 910-918 (2013).
- 40 Shen, B. et al. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res* 23, 720-723 (2013).
- Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173-1183 (2013).
- 45 Tuschl, T. Expanding small RNA interference. *Nat Biotechnol* 20, 446-448 (2002).
- Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. & Kucherlapati, R.S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* 317,230-234 (1985).
- 50 Thomas, K.R., Folger, K.R. & Capecchi, M.R. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* 44, 419-428 (1986).
- Hasty, P., Rivera-Perez, J. & Bradley, A. The length of homology required for gene targeting in embryonic stem cells. *Mol Cell Biol* 11, 5586-5591 (1991).
- 55 Wu, S., Ying, G.X., Wu, Q. & Capecchi, M.R. A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond. *Nat Protoc* 3, 1056-1076 (2008).
- 60 Guschin, D.Y. et al. A rapid and general assay for monitoring endogenous gene modification. *Methods Mol Biol* 649, 247-256 (2010).
- Oliveira, T.Y. et al. Translocation capture sequencing: a method for high throughput mapping of chromosomal rearrangements. *J Immunol Methods* 375, 176-181 (2012).
- 65

- Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607 (2011).
- 5 Bogenhagen, D.F. & Brown, D.D. Nucleotide sequences in *Xenopus* 5S DNA required for transcription termination. *Cell* 24, 261-270 (1981).
- Bultmann, S. et al. Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res* 40, 5368-5377 (2012).
- 10 Valton, J. et al. Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. *J Biol Chem* 287, 38427-38432 (2012).
- Christian, M. et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757-761 (2010).
- 15 Mussolino, C. et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic acids research* 39, 9283-9293 (2011).
- Bobis-Wozowicz, S., Osiak, A., Rahman, S.H. & Cathomen, T. Targeted genome editing in pluripotent stem cells using zinc-finger nucleases. *Methods* 53, 339-346 (2011).
- 20 Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31, 233-239 (2013).
- Michaelis, L.M., Maud "Die kinetik der invertinwirkung.". *Biochem. z* (1913).
- 25 Mahfouz, M.M. et al. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc Natl Acad Sci U S A* 108, 2623-2628 (2011).
- Wilson, E.B. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 22, 209-212 (1927).
- 30 Tangri S, et al., Rationally engineered therapeutic proteins with reduced immunogenicity, *J Immunol.* 2005 Mar 15;174(6):3187-96.
- REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991)
- 35 Lopes, V.S., etc al., Retinal gene therapy with a large MYO7A cDNA using adeno-associated virus. *Gene Ther*, 2013 Jan 24. doi: 10.1038/gt.2013.3.[Epub ahead of print]
- Kaplitt, M.G., et al., Safety and tolerability of gene therapy with an adeno-associated virus (AAV) borne GAD gene for Parkinson's disease: an open label, phase I trial. *Lancet.* 2007 Jun 23;369(9579):2097-105.
- 40 Nathwani, A.C., et al., Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med.* 2011 Dec 22;365(25):2357-65. doi: 10.1056/NEJMoa1108046. Epub 2011 Dec 10.
- 45 Gray SJ, Foti SB, Schwartz JW, Bachaboina L, Taylor-Blake B, Coleman J, Ehlers MD, Zylka MJ, McCown TJ, Samulski RJ. Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors. *Hum Gene Ther.* 2011 Sep;22(9):1143-53. doi: 10.1089/hum.2010.245.
- 50 Liu D, Fischer I. Two alternative promoters direct neuron-specific expression of the rat microtubule-associated protein 1B gene. *J Neurosci.* 1996 Aug 15;16(16):5026-36.
- Levitt N. Briggs D. Gil A. Proudfoot N.J. Definition of an efficient synthetic poly(A) site. *Genes Dev.* 1989;3:1019-1025.
- 55 McClure C, Cole KL, Wulff P, Klugmann M, Murray AJ. Production and titrating of recombinant adeno-associated viral vectors. *J Vis Exp.* 2011 Nov 27;(57):e3348. doi: 10.3791/3348.
- Banker G, Goslin K. Developments in neuronal cell culture. *Nature.* 1988 Nov 10;336(6195):185-6.
- 60 Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell research* 23, 465-472.
- Dianov, G.L., and Hubscher, U. (2013). Mammalian base excision repair: the forgotten archangel. *Nucleic acids research* 41, 3483-3490.
- 65

- Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., and Calarco, J.A. (2013). Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nature methods* 10, 741-743.
- 5 Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology*.
- Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome Engineering of *Drosophila* with the CRISPR RNA-Guided Cas9 Nuclease. *Genetics* 194, 1029-1035.
- 10 Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*.
- 15 Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature biotechnology*.
- Maresca, M., Lin, V.G., Guo, N., and Yang, Y. (2013). Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome research* 23, 539-546.
- 20 Pattanayak, V., Lin, S., Guillinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature biotechnology*.
- 25 Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., ... Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315(5819), 1709-1712. doi: 10.1126/science.1138140.
- Barrangou, R., & Marraffini, L. A. (2014). CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell*, 54(2), 234-244. doi: 10.1016/j.molcel.2014.03.011.
- 30 Boch, J., & Bonas, U. (2010). *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol*, 48, 419-436. doi: 10.1146/annurev-phyto-080508-081936.
- 35 Chylinski, K., Makarova, K. S., Charpentier, E., & Koonin, E. V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res*, 42(10), 6091-6105. doi: 10.1093/nar/gku241.
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., ... Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121), 819-823. doi: 10.1126/science.1231143.
- 40 Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., ... Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471(7340), 602-607. doi: 10.1038/nature09886.
- 45 Dianov, G. L., & Hubscher, U. (2013). Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res*, 41(6), 3483-3490. doi: 10.1093/nar/gkt076.
- 50 Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*, 31(9), 822-826. doi: 10.1038/nbt.2623.
- Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., ... Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 468(7320), 67-71. doi: 10.1038/nature09523.
- 55 Gasiunas, G., Sinkunas, T., & Siksnys, V. (2014). Molecular mechanisms of CRISPR-mediated microbial immunity. *Cell Mol Life Sci*, 71(3), 449-465. doi: 10.1007/s00018-013-1438-6.
- Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157(6), 1262-1278. doi: 10.1016/j.cell.2014.05.010.
- 60 Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., ... Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*, 31(9), 827-832. doi: 10.1038/nbt.2647.
- 65 Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096), 816-821. doi: 10.1126/science.1225829.

- Klug, A. (2010). The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem*, 79,213-231. doi: 10.1146/annurev-biochem-010909-095056.
- 5 Kuscu, C., Arslan, S., Singh, R., Thorpe, J., & Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol*. doi: 10.1038/nbt.2916.
- Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., ... Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*, 31(9), 833-838. doi: 10.1038/nbt.2675.
- 10 Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 155(Pt 3), 733-740. doi: 10.1099/mic.0.023960-0.
- 15 Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., ... Nureki, O. (2014). Crystal structure of cas9 in complex with guide RNA and target DNA. *Cell*, 156(5), 935-949. doi: 10.1016/j.cell.2014.02.001.
- Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E., Doudna, J. A., & Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*, 31(9), 839-843. doi: 10.1038/nbt.2673.
- 20 Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., ... Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, 154(6), 1380-1389. doi: 10.1016/j.cell.2013.08.021.
- 25 Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F., & Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, 153(4), 910-918. doi: 10.1016/j.cell.2013.04.025.
- 30 Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, 482(7385), 331-338. doi: 10.1038/nature10886.
- Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., ... Sharp, P. A. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol*. doi: 10.1038/nbt.2889.
- 35 Ran et al., In vivo genome editing using *Staphylococcus aureus* Cas9, Ran et al, *Nature*, 1st April 2015 doi:10.1038/nature14299 (Ran 2015),

REIVINDICACIONES

1. Una composición que no es de origen natural o que está manipulada para su uso en tratamiento, por manipulación de una primera y segunda secuencia diana en hebras opuestas de un ADN bicatenario en un locus genómico de interés en una célula promoviendo la reparación dirigida por homología, en la que la composición comprende:
- 5 I. una primera secuencia polinucleotídica de ARN quimérico (ARNqui) del sistema de CRISPR-Cas, en la que la primera secuencia polinucleotídica comprende:
- (a) una primera secuencia guía que puede hibridar con la primera secuencia diana,
- (b) una primera secuencia de acoplamiento de crtra, y
- 10 (c) una primera secuencia crtra,
- II. una segunda secuencia polinucleotídica de ARNqui del sistema de CRISPR-Cas, en la que la segunda secuencia polinucleotídica comprende:
- (a) una segunda secuencia guía que puede hibridar con la segunda secuencia diana,
- (b) una segunda secuencia de acoplamiento de crtra, y
- 15 (c) una segunda secuencia crtra,
- III. una enzima Cas9 de tipo II o una secuencia polinucleotídica que la codifica, en la que la enzima Cas9 es una proteína SpCas9 que comprende la mutación N863A, o un ortólogo de la misma que tiene una mutación correspondiente a SpCas9N863A, que comprende cero, o al menos una o dos o más secuencias de localización nuclear, y
- 20 IV. un molde de reparación que comprende un oligonucleótido monocatenario sintetizado o manipulado, en la que (a), (b) y (c) están dispuestas en una orientación 5' a 3',
- en la que la secuencia polinucleotídica que codifica la enzima Cas9 es ADN o ARN, en la que la primera secuencia guía dirige la escisión de una hebra del ADN bicatenario cerca de la primera secuencia diana y la segunda secuencia guía dirige la escisión de la otra hebra cerca de la segunda secuencia diana provocando salientes 3', y
- 25 en la que el molde de reparación se introduce en el ADN bicatenario.
2. La composición para su uso de acuerdo con la reivindicación 1, en la que el molde de reparación comprende además un sitio de restricción para endonucleasa de restricción.
3. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 o 2, en la que alguna o todas de la secuencia polinucleotídica que codifica la enzima Cas9, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, es o son ARN; o
- 30 en la que los polinucleótidos que comprenden la secuencia que codifica la enzima Cas9, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra o la primera y la segunda secuencia crtra, son ARN y se suministran mediante nanopartículas, exosomas, microvesículas o un cañón de genes.
4. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 3,
- 35 en la que alguno o todos de la secuencia polinucleotídica que codifica la enzima Cas9, la primera y la segunda secuencia guía, la primera y la segunda secuencia de acoplamiento de crtra, la primera y la segunda secuencia crtra y el molde de reparación, se suministran mediante un sistema de vector,
- en la que el sistema de vector comprende uno o más vectores,
- preferiblemente en la que el uno o más vectores comprenden uno o más vectores víricos.
- 40 5. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en la que la primera y segunda secuencia de acoplamiento de crtra comparten un 100 % de identidad y/o en la que la primera y segunda secuencia crtra comparten un 100 % de identidad.
6. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en la que el saliente 3' es de 1-100 pares de bases.
- 45 7. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en la que la composición es para su uso *ex vivo*.
8. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en la que la composición es para su uso *in vivo*.
9. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 8, en la que la Cas9 es una Cas9 de *S. aureus* mutada que comprende la mutación N580A.
- 50

10. La composición para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 9, en la que el tratamiento es un tratamiento de una enfermedad genética y en la que el molde de reparación se introduce en el ADN bicatenario, por lo que se obtiene alteración fenotípica,

opcionalmente la enfermedad genética es un trastorno ocular.

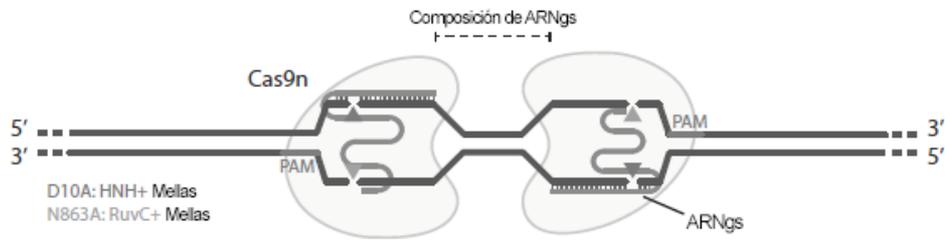
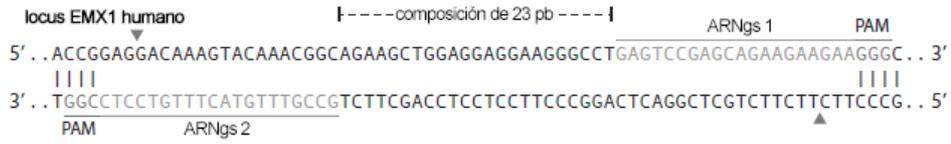


Figura 1

A



sitios inespecíficos conocidos para ARNgs 1

inespecífico (OT)-1 GAGGCCGAGCAGAAGAAGACGG

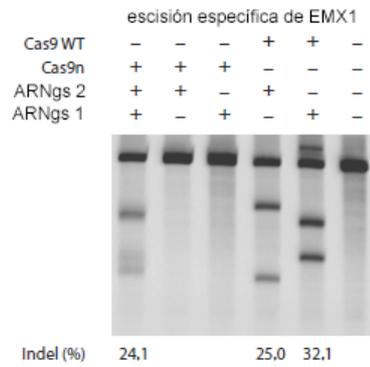
inespecífico (OT)-2 AAGTCTGAGCACAGAAGAATGG

inespecífico (OT)-3 GAGTCCTAGCAGGAGAAGAAGAG

inespecífico (OT)-4 GAGTCTAAGCAGAAGAAGAAGAG

inespecífico (OT)-5 GAGTTAGAGCAGAAGAAGAAGG

B



C

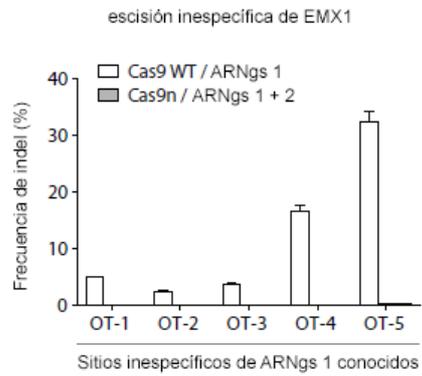


Figura 2

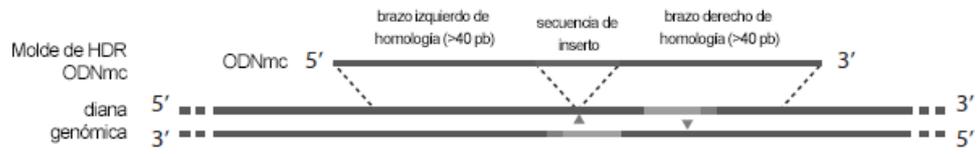
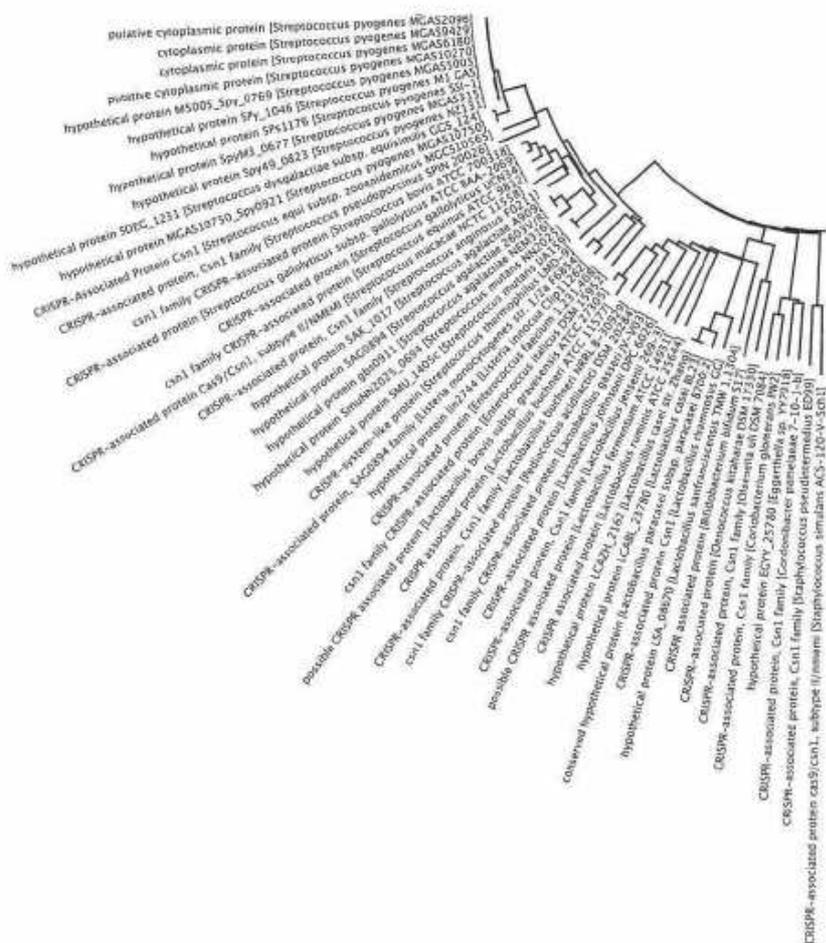


Figura 3



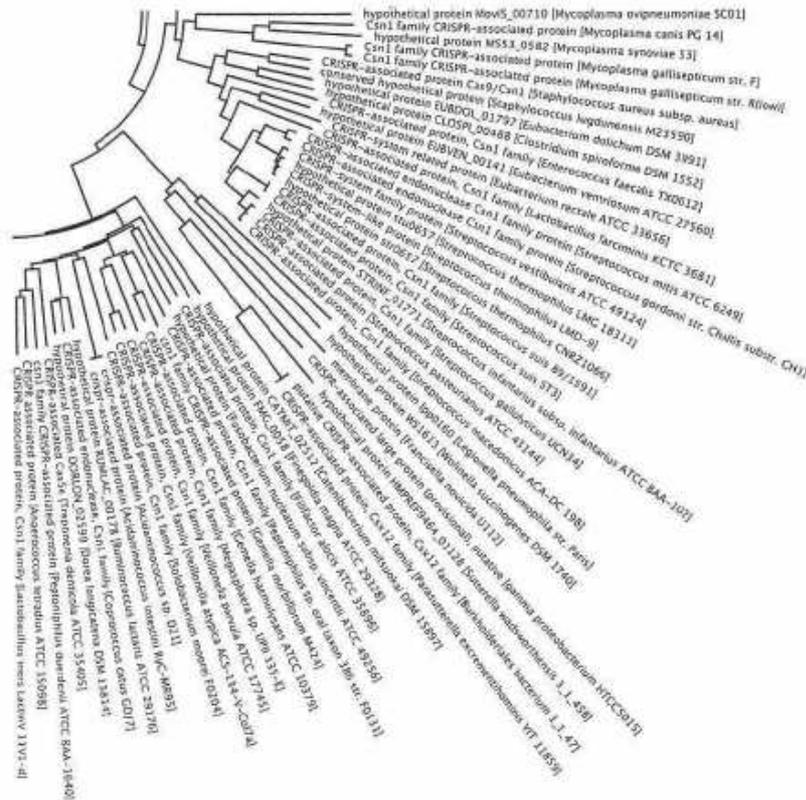
hypothetical protein = proteína hipotética; CRISPR-associated protein, Csn1 family = familia Csn1 de proteína asociada a CRISPR;
 CRISPR-associated endonuclease Csn1 family protein, = proteína de familia Csn1 de endonucleasa asociada a CRISPR;
 CRISPR-associated protein Cas9/Csn1, subtype II= proteína Cas9/Csn1 asociada a CRISPR, subtipo II; CRISPR-associated CasSe
 family protein =proteína de la familia CasSe asociada a CRISPR;

Figura 4B



Putative cytoplasmic protein = proteína citoplasmática putativa; cytoplasmic protein = proteína citoplasmática.
hypothetical protein = proteína hipotética; CRISPR-associated Protein Csn1 = Csn1 Proteína asociada a CRISPR.
CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated Protein
Cas9/Csn1, subtype II = Proteína Cas9/Csn1 asociada a CRISPR, subtipo II; possible CRISPR-associated protein =
Proteína asociada a CRISPR posible; conserved hypothetical protein = proteína hipotética conservada

Figura 4C



Putative CRISPR-associated Protein Csx12 family = familia Csx12 de proteína asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated Protein Cas9/Csn1, subtype II = Proteína Cas9/Csn1 asociada a CRISPR, subtipo II; conserved hypothetical protein = proteína hipotética conservada; CRISPR-system related protein = proteína relacionada con el sistema CRISPR; CRISPR-system like protein = proteína de tipo sistema CRISPR; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; membrane protein = proteína de membrana; CRISPR-associated large protein (provisional), putative = proteína grande asociada a CRISPR (provisional) putativa; CRISPR-associated CasSe = CasSe asociada a CRISPR.

Figura 4D

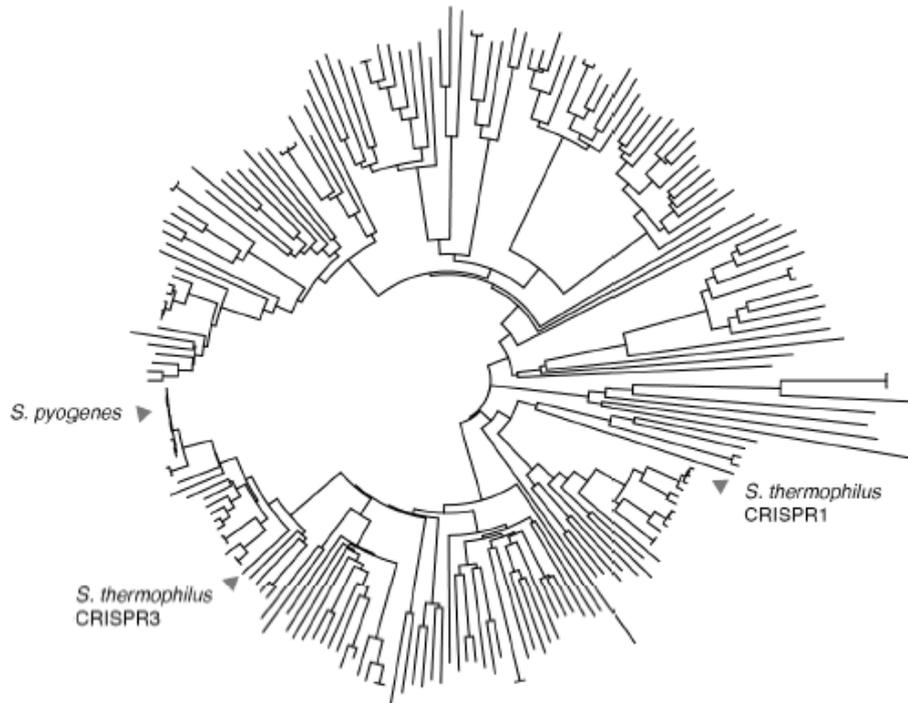
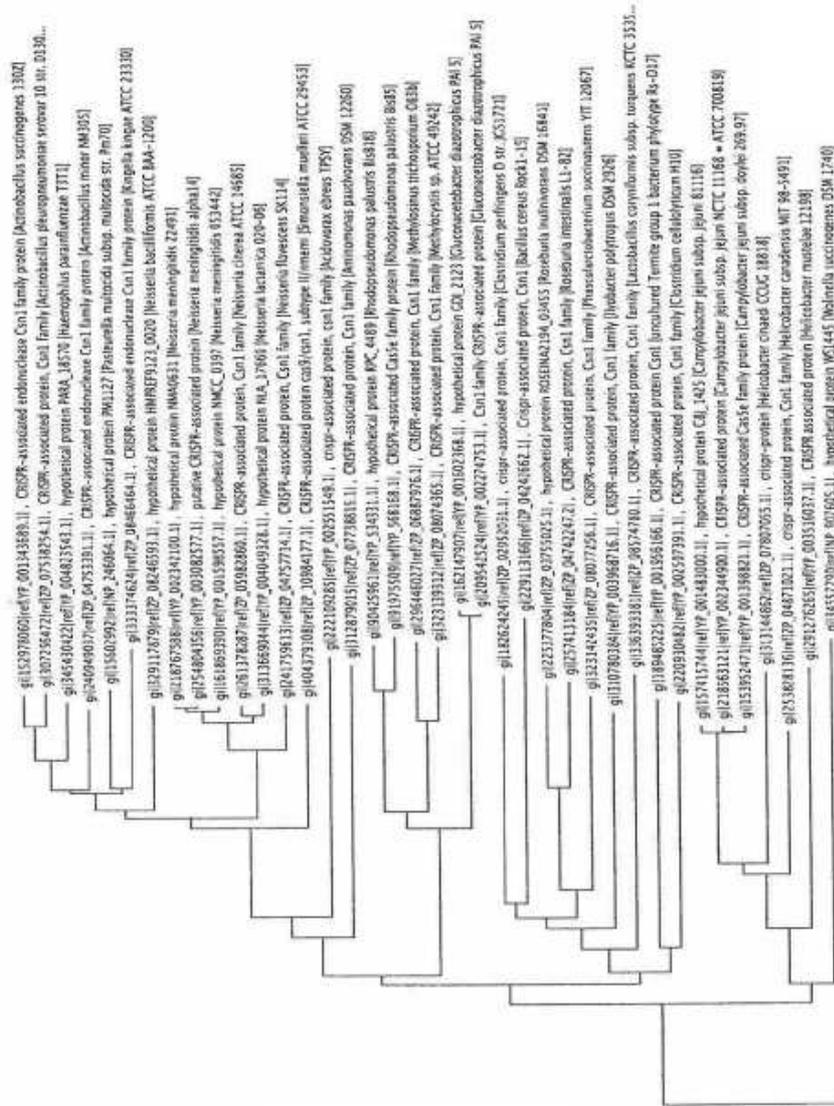
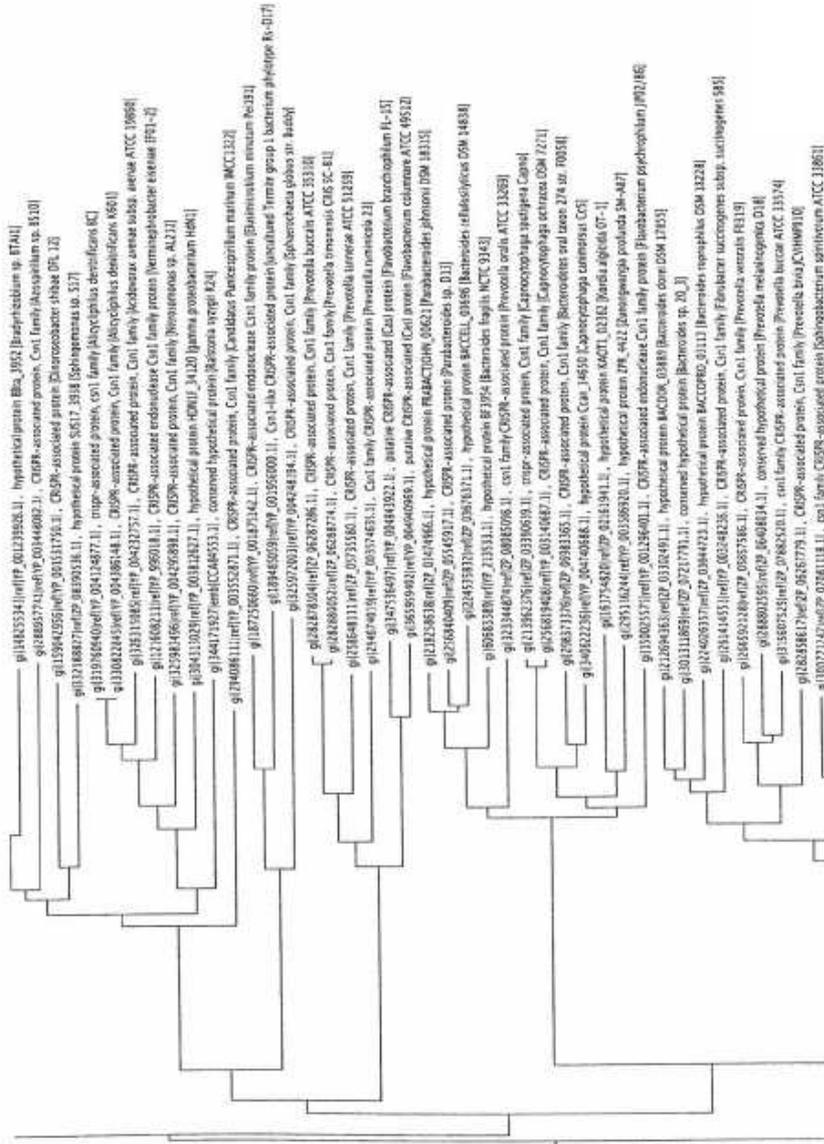


Figura 4E



Putative CRISPR-associated Protein = proteína asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated Protein Cas9/Csn1, subtype II = Proteína Cas9/Csn1 asociada a CRISPR, subtipo II; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; CRISPR-associated CasSe family protein = proteína de la familia CasSe asociada a CRISPR

Figura 5A



Putative CRISPR-associated (Cas) Protein = proteína (Cas) asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada; Csn1-like CRISPR-associated protein = proteína asociada a CRISPR de tipo Csn1

Figura 5B

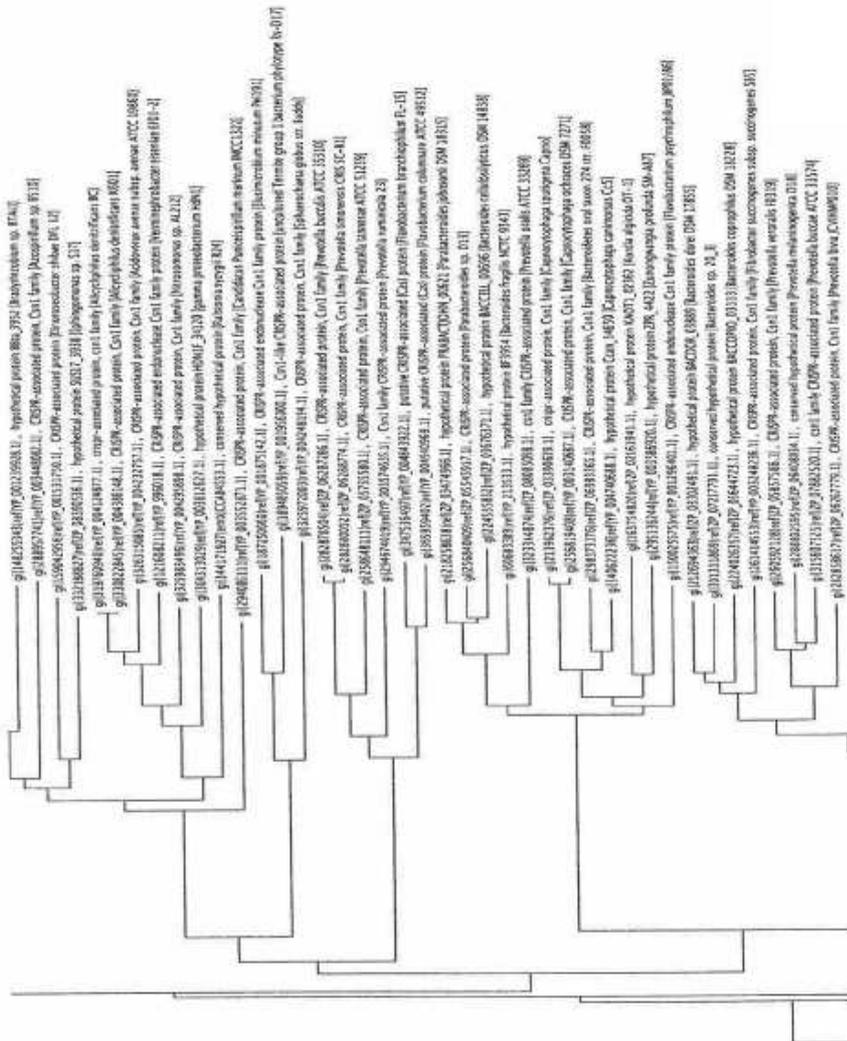
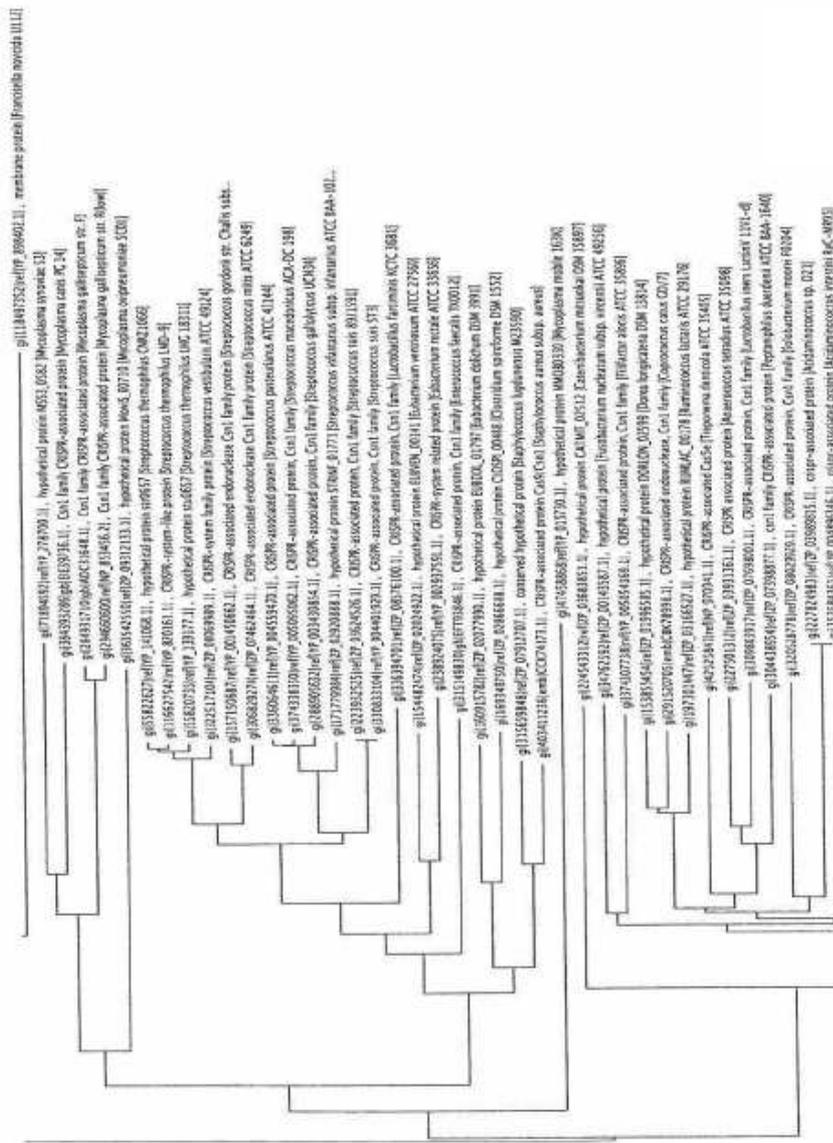


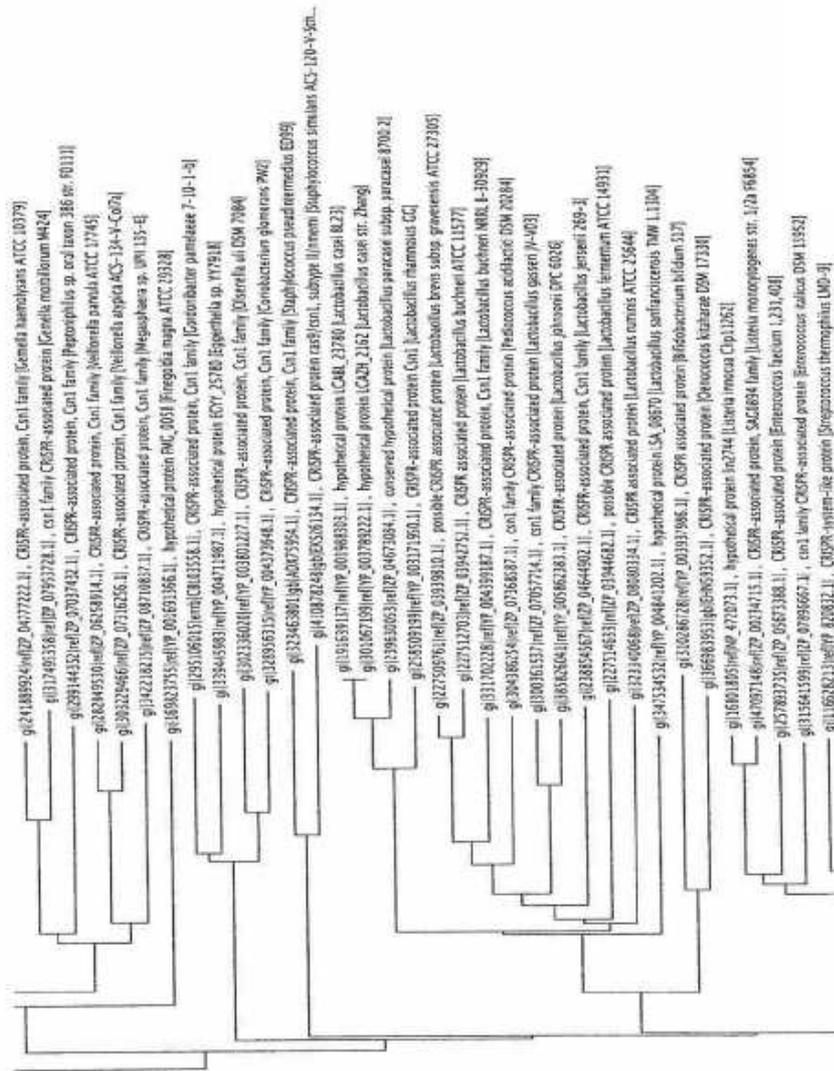
Figura 5C

Putative CRISPR-associated (Cas) Protein = proteína (Cas) asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada; Csn1-like CRISPR-associated protein = proteína asociada a CRISPR de tipo Csn1



Hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada; Csn1-like CRISPR-associated protein = proteína asociada a CRISPR de tipo Csn1; CRISPR-system-like protein = Proteína de tipo sistema CRISPR; CRISPR-associated protein Cas9/Csn1 = proteína Cas9/Csn1 asociada a CRISPR

Figura 5D



Hypothetical protein = proteína hipotética; CRISPR-associated Protein Csn1 family = familia Csn1 Proteína asociada a CRISPR; possible CRISPR-associated protein = Proteína asociada a CRISPR posible; conserved hypothetical protein = proteína hipotética conservada

Figura 5E

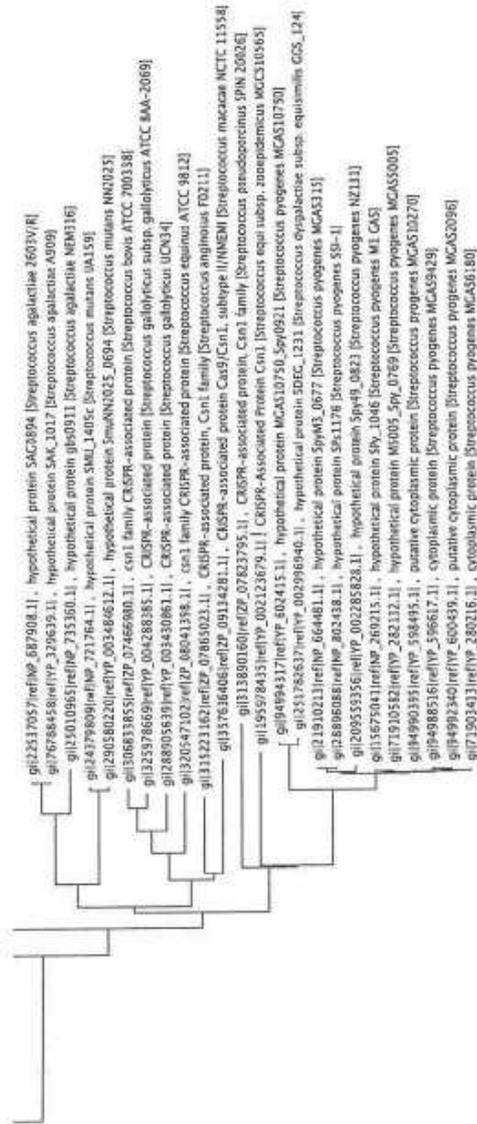


Figura 5F

Hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated protein Cas9/Csn1, subtype II = proteína Cas9/Csn1 asociada a CRISPR, subtipo II; Putative cytoplasmic protein = proteína citoplasmática putativa; cytoplasmic protein = proteína citoplasmática

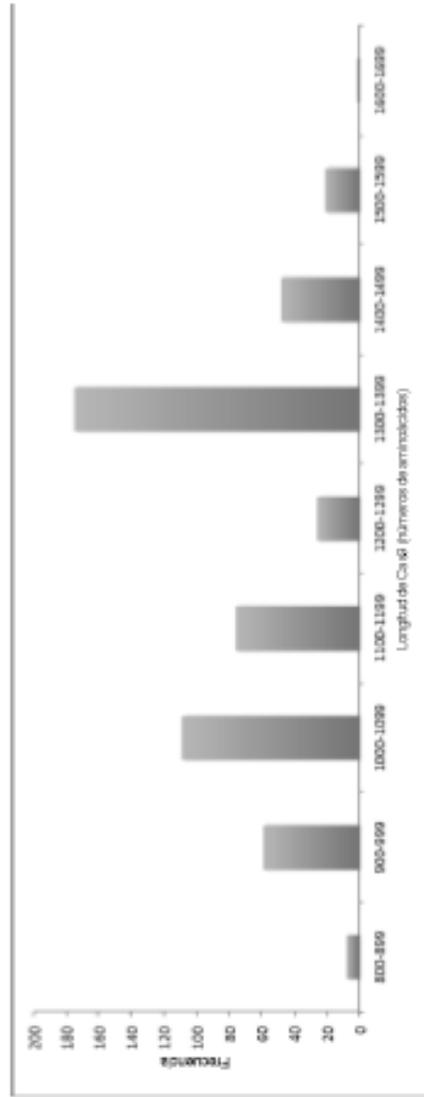


Figura 6

Posiciones de mutación de SpCas9

hSpCas9

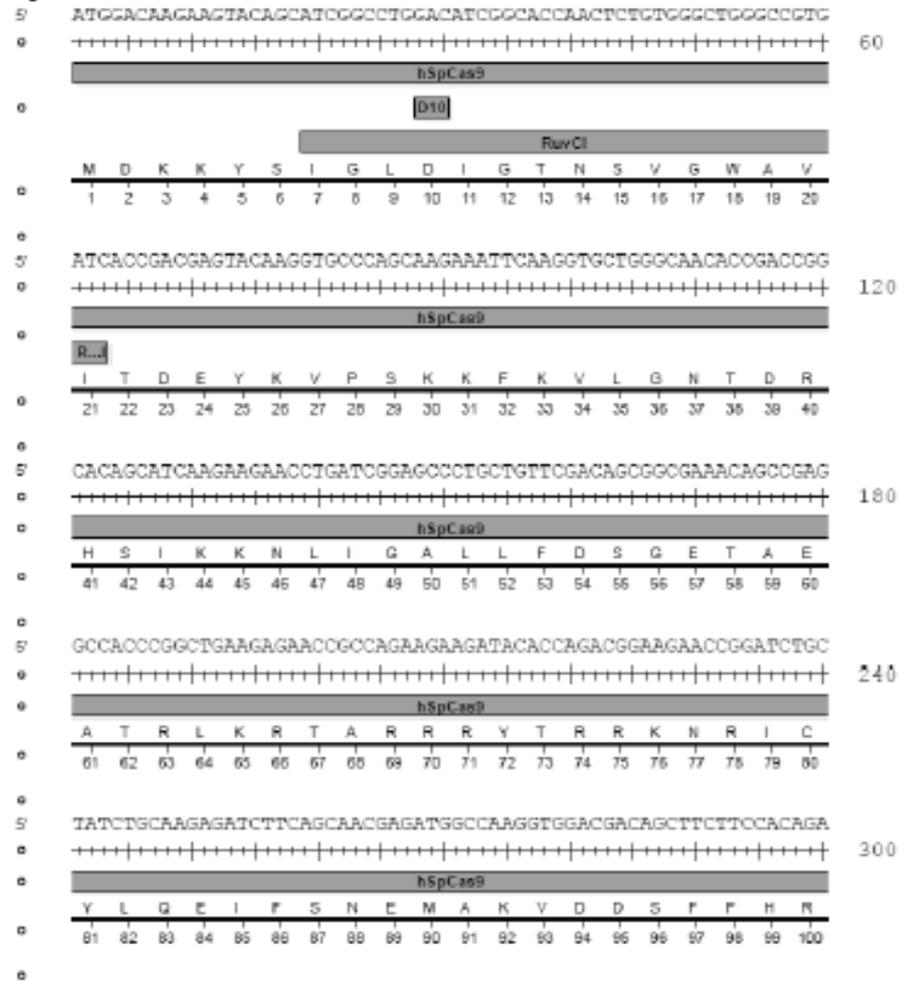


Figura 7A

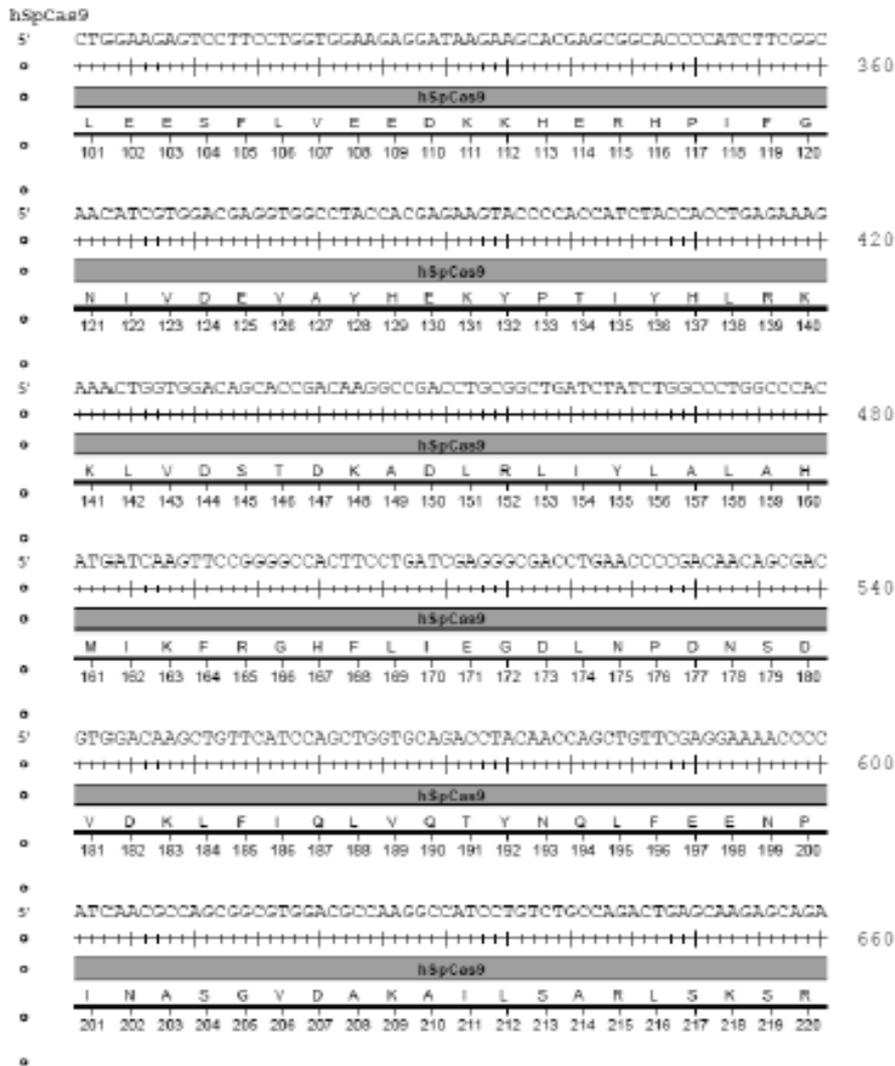


Figura 7B

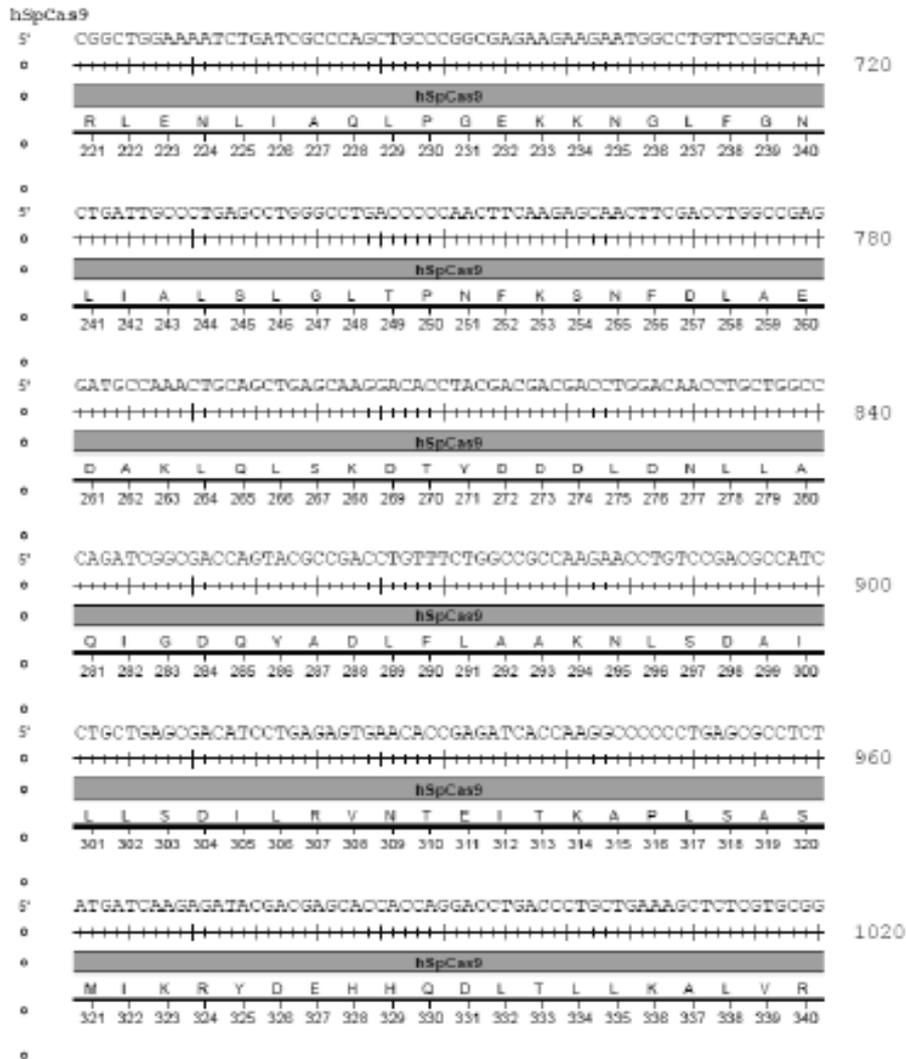


Figura 7C

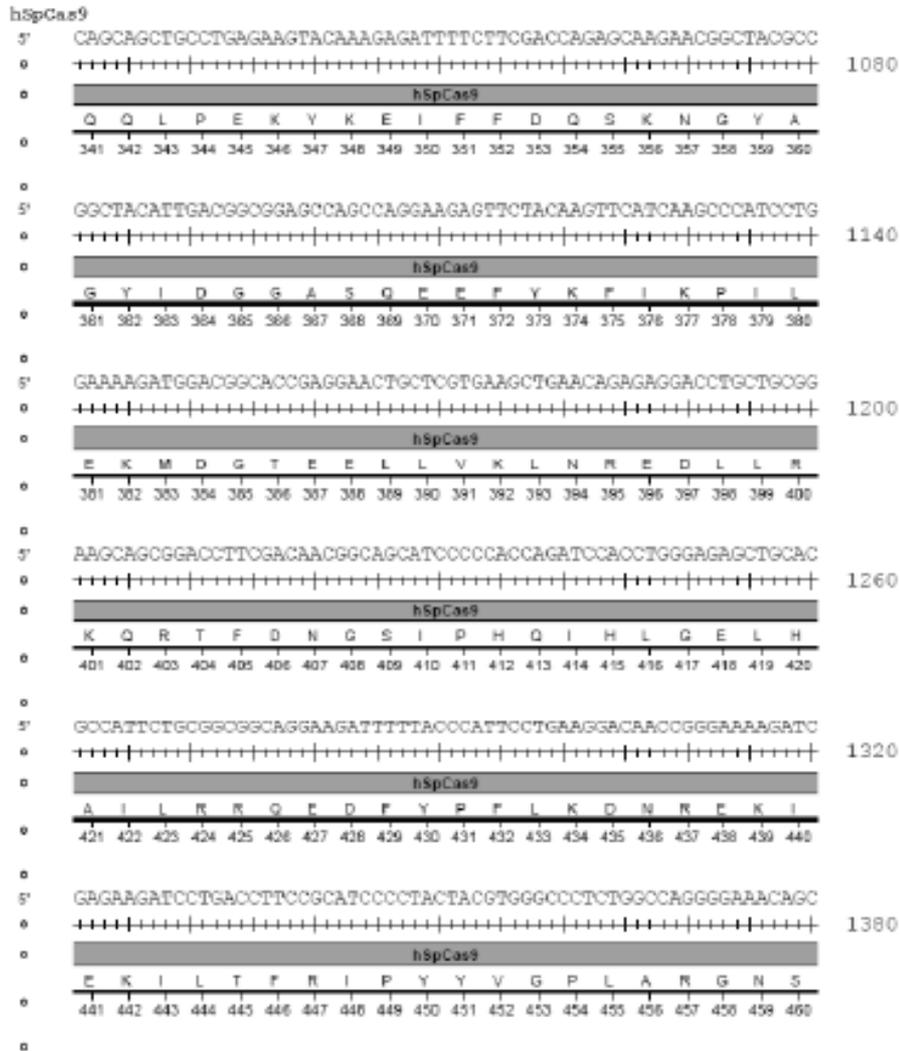


Figura 7D

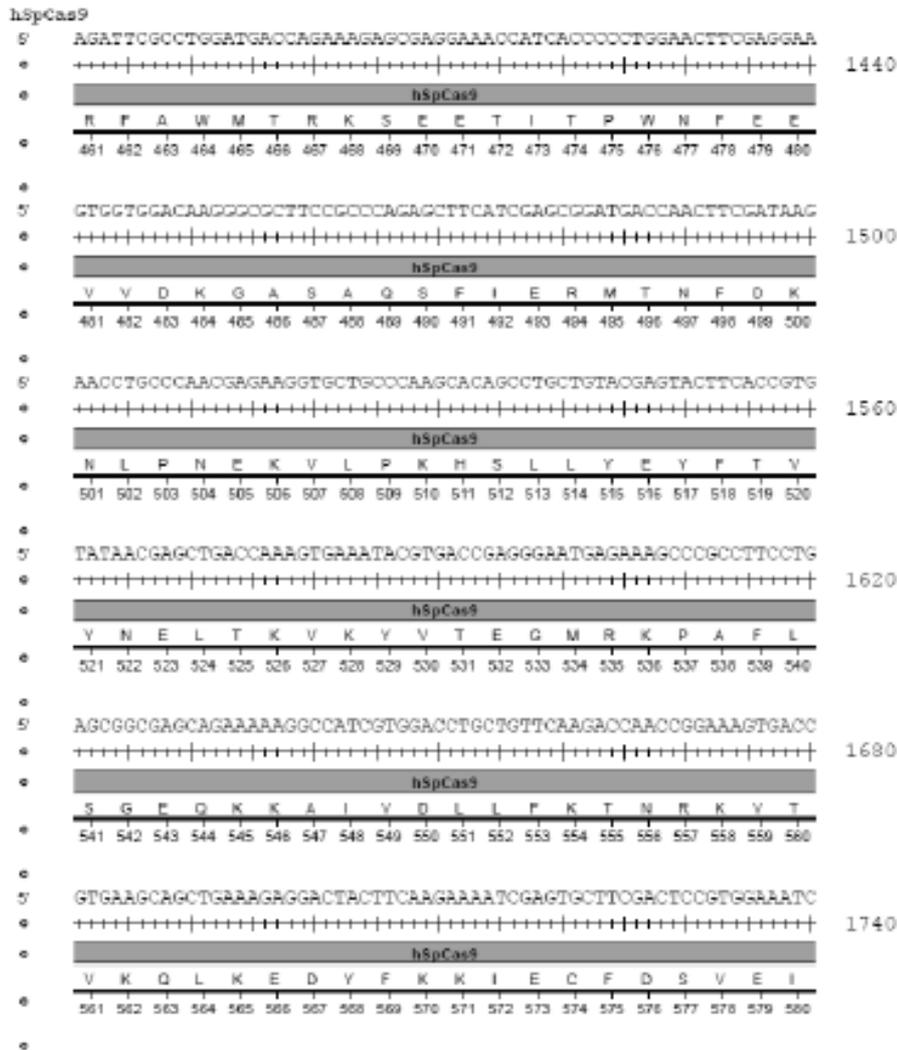


Figura 7E

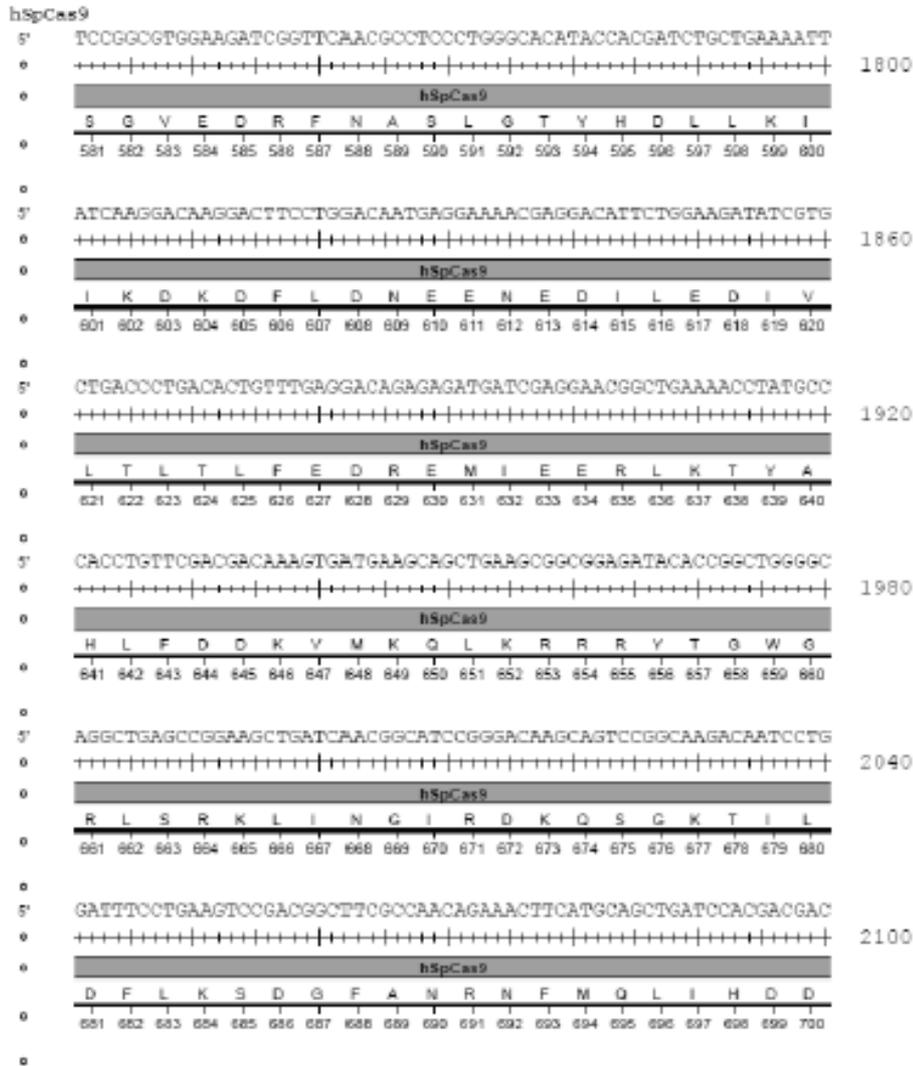


Figura 7F

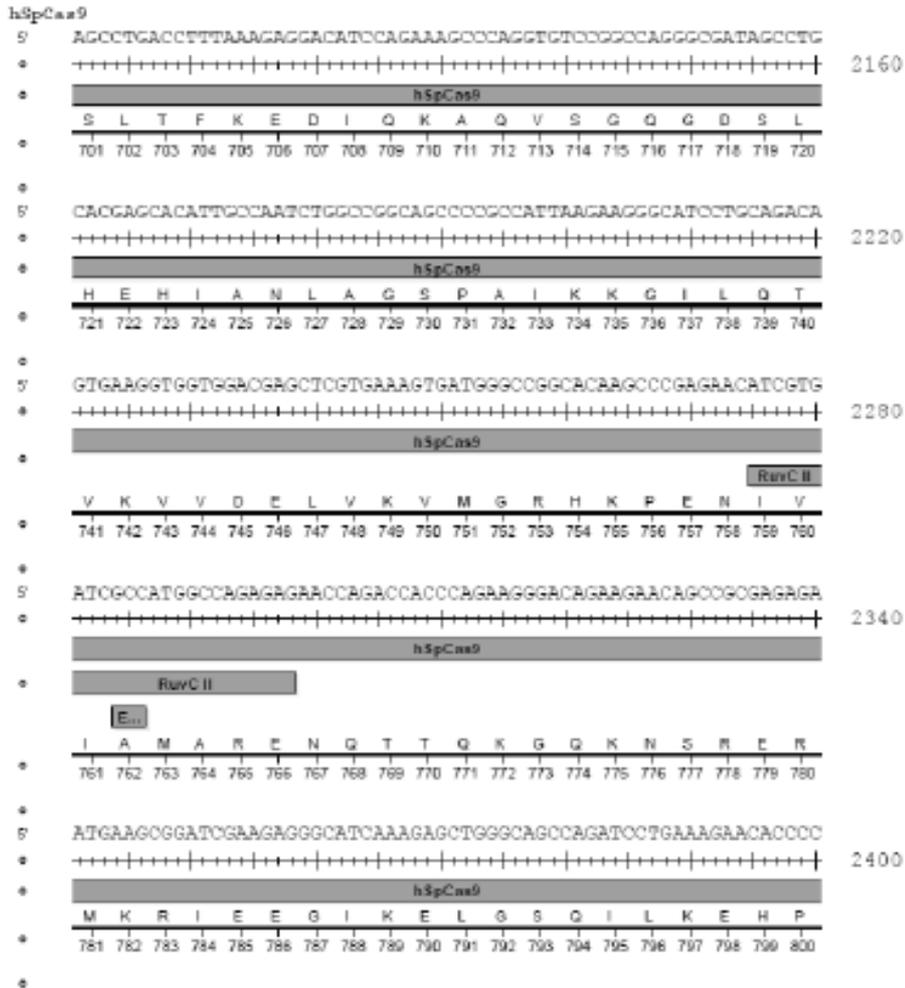


Figura 7G

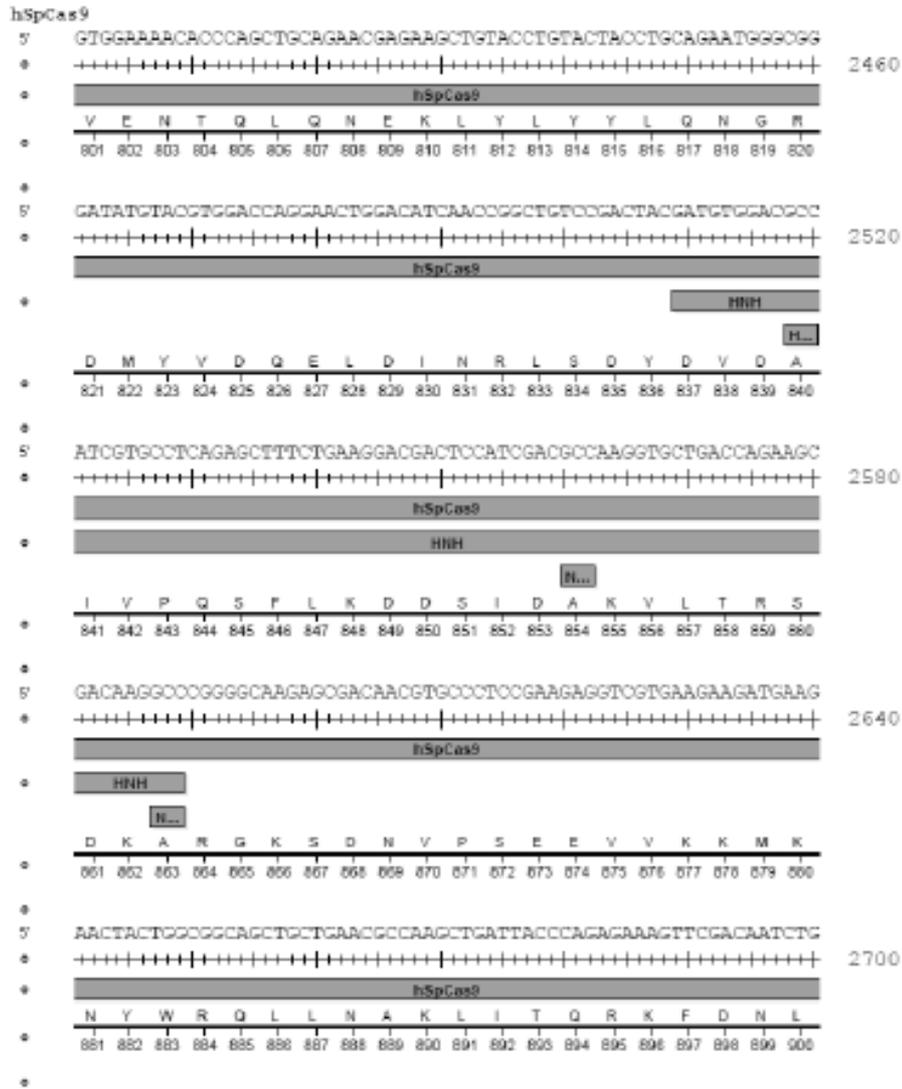


Figura 7H

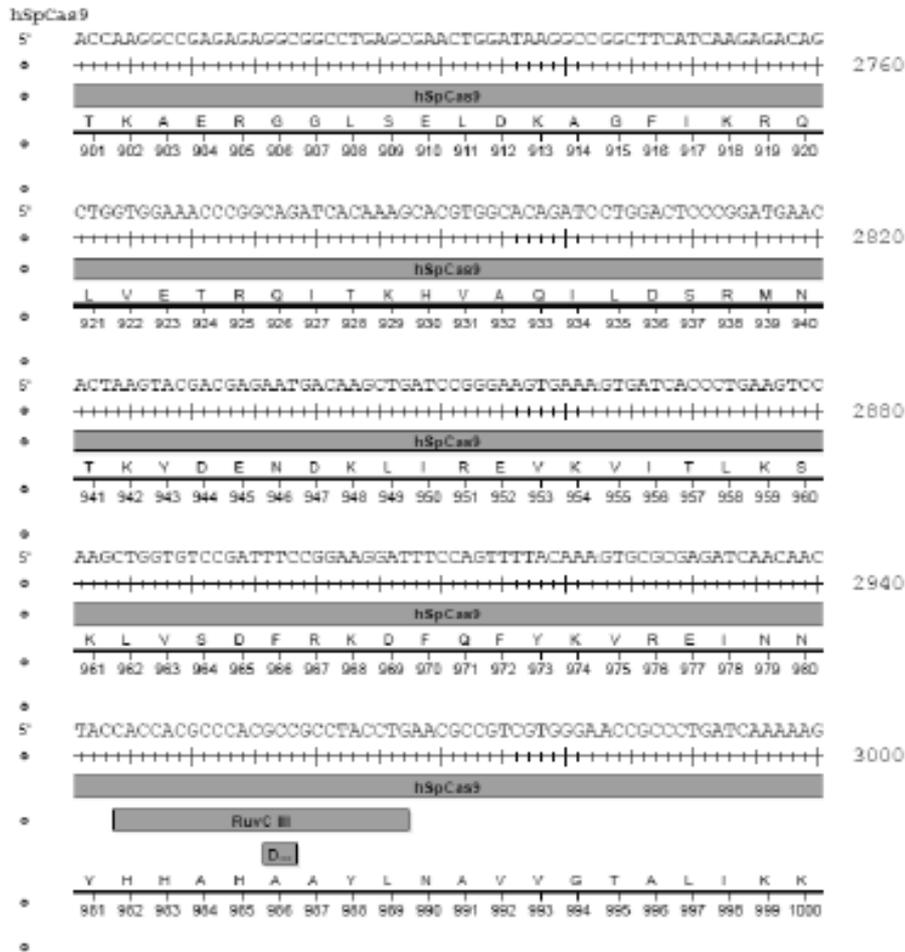


Figura 71

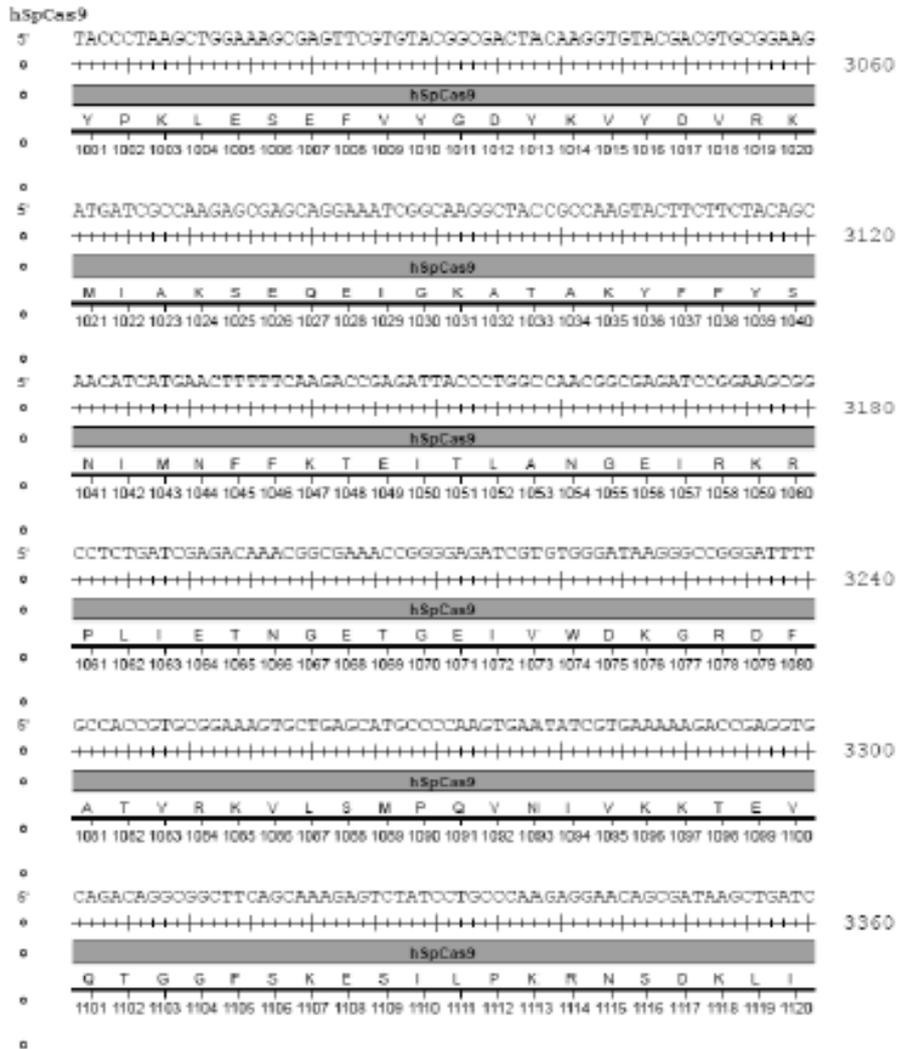


Figura 7J

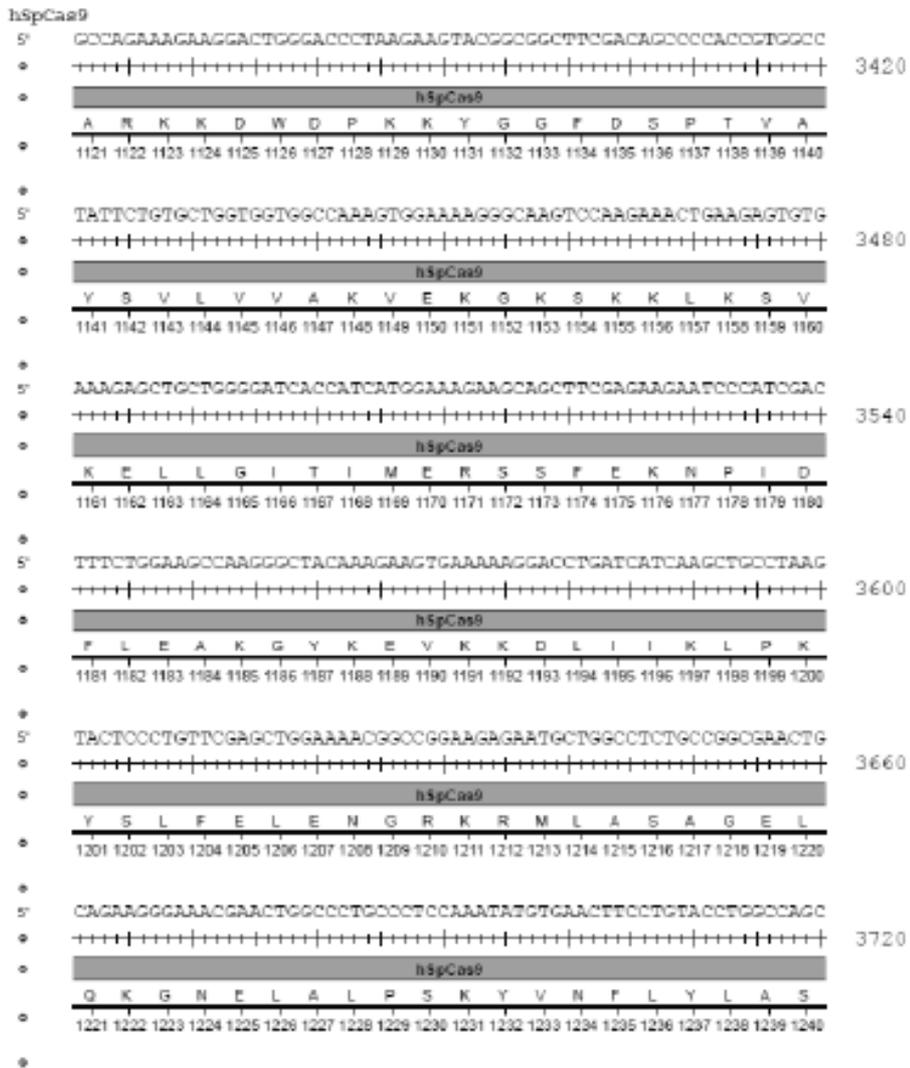


Figura 7K

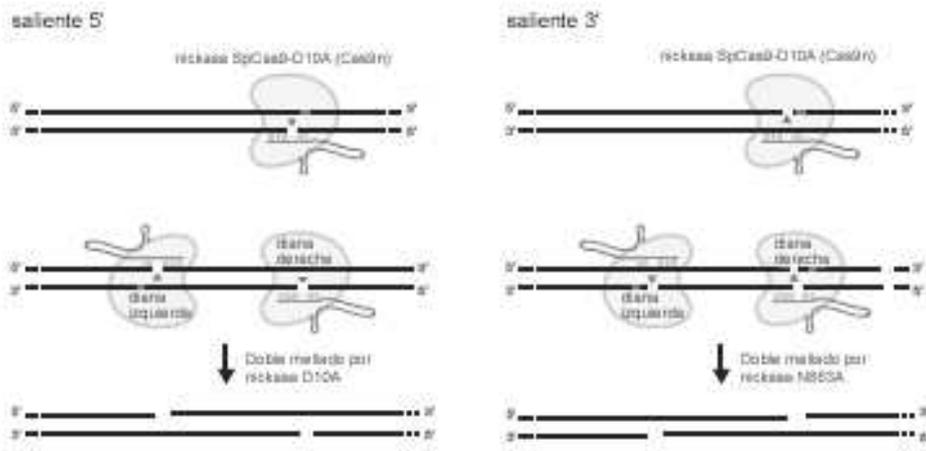


Figura 8

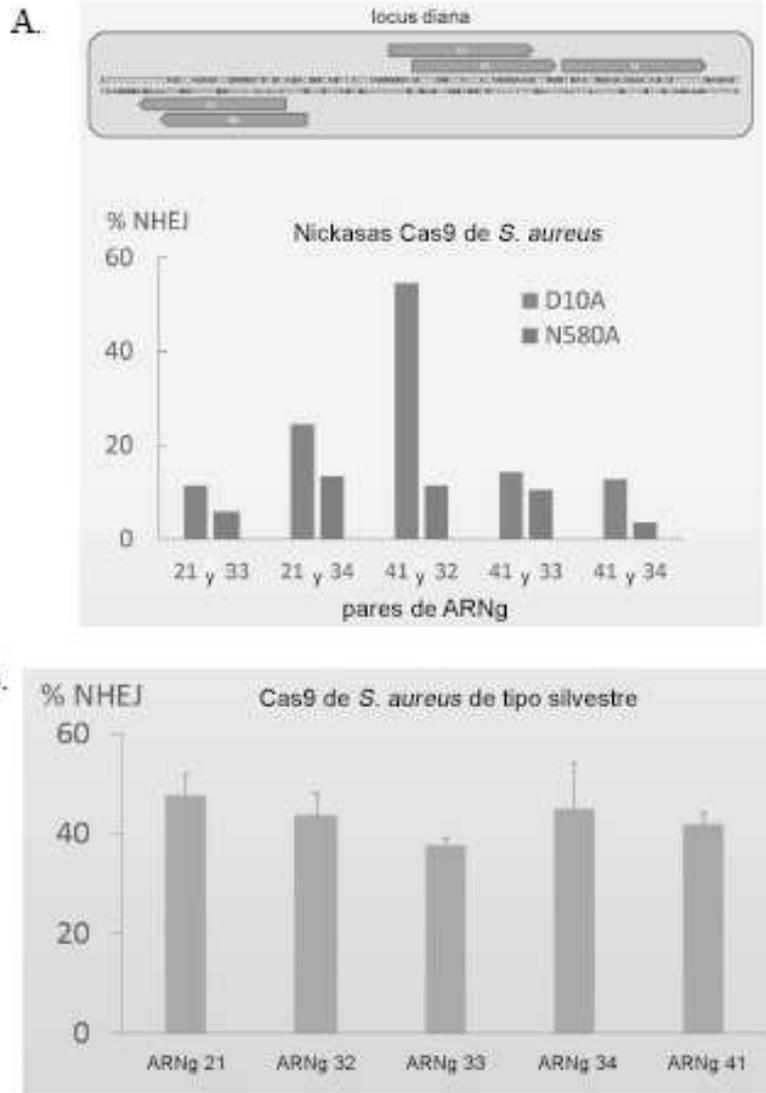


Figura 9A-B