

Plataformas de entrega CRISPR objetivo

Campo de la Invención

La presente invención está relacionada con composiciones y métodos para la terapia
5 génica. Varios enfoques descritos en el presente documento utilizan los sistemas *Neisseria meningitidis* Cas9 que proporcionan las plataformas de edición de genes CRISPR hiperprecisas. Además, la invención incorpora mejoras de este sistema Cas9: por ejemplo, el truncamiento de las secuencias de ARN de guía única, y el empaquetamiento de Nme1Cas9 o Nme2Ca9 con su ARN guía en un vector viral adenoasociado que es compatible para la administración *in vivo*.
10 Además, se han identificado ortólogos Cas9 de Tipo II-C que se dirigen a secuencias limitadas de motivo adyacente al protoespaciador de entre uno-cuatro nucleótidos requeridos.

Antecedentes

Las repeticiones palindrómicas cortas agrupadas y regularmente espaciadas (CRISPR)-
15 CRISPR asociada (Cas) es un sistema inmune adaptativo guiado por ARN único que se encuentra en arqueas y bacterias. Estos sistemas proporcionan inmunidad al dirigirse a e inactivar los ácidos nucleicos que se originan a partir de elementos genéticos foráneos. Muchos tipos diferentes de sistemas CRISPR-Cas se han identificado hasta la fecha y se clasifican en dos clases.

20 Dentro de los sistemas CRISPR de clase II, se caracterizan los sistemas CRISPR-Cas de tipo II por una única proteína efectora llamada Cas9, que forma un complejo de ribonucleoproteína (RNP) con ARN CRISPR (crARN) y ARN transactivador (tracrARN) para seleccionar y escindir el ADN. El crARN contiene una secuencia guía programable que puede dirigir Cas9 a casi cualquier secuencia de ADN en organismos vivos.

25 Esta programabilidad de los complejos RNP Cas9 ha sido aprovechada por muchos investigadores para la edición genómica en sistemas eucariontes. Se ha utilizado para editar los genomas de células de mamíferos, embriones humanos, plantas, roedores y otros organismos vivos. Las RNP Cas9 se han utilizado para la edición genómica precisa (con plantilla de donante) e imprecisa, los cuales han encontrado aplicaciones en la terapia génica, agricultura y otros
30 campos. Además, las versiones de nucleasa muerta de los ortólogos Cas9 se están utilizando para la modulación de la transcripción, el etiquetado de ADN específico a sitio y para el perfilado del

proteoma en loci genómicos específicos. Se han utilizado varias Cas9 diferentes para estas aplicaciones. Un aspecto central de la capacidad de programación de Cas9 y, por lo tanto, sus aplicaciones es la capacidad de introducir cualquier secuencia guía en el crARN. El crARN y tracrARN se pueden fusionar para formar un ARN de guía única (sgARN), que es más estable y proporciona una edición genómica mejorada.

Lo que se necesita en la técnica son secuencias mejoradas de Cas9s y sgARN que puedan proporcionar una edición específica y precisa de una gama más amplia de sitios objetivo, especialmente cuando se combinan con plataformas confiables de entrega de ácido nucleico.

10 **Resumen de la Invención**

La presente invención está relacionada con composiciones y métodos para terapia génica. Varios enfoques descritos en el presente documento utilizan los sistemas *Neisseria meningitidis* Cas9 que proporcionan las plataformas de edición de genes CRISPR hiperprecisas. Además, la invención incorpora mejoras de este sistema Cas9: por ejemplo, el truncamiento de las secuencias de ARN de guía única, y el empaquetamiento de Nme1Cas9 o Nme2Cas9 con su ARN guía en un vector viral adenoasociado que es compatible para la administración *in vivo*. Además, se han identificado ortólogos Cas9 de Tipo II-C que se dirigen a secuencias limitadas de motivo adyacente al protoespaciador de entre uno-cuatro nucleótidos requeridos.

En una modalidad, la presente invención contempla una secuencia de ácido ribonucleico de guía única (sgARN) que comprende una repetición truncada: región anti-repetición. En una modalidad, la secuencia de sgARN comprende además una región del vástago 2 truncada. En una modalidad, la secuencia de sgARN comprende además una región espaciadora truncada. En una modalidad, dicha secuencia de sgARN tiene una longitud de 121 nucleótidos. En una modalidad, dicha longitud de secuencia de sgARN se selecciona del grupo que consiste en 111 nucleótidos, 107 nucleótidos, 105 nucleótidos, 103 nucleótidos, 102 nucleótidos, 101 nucleótidos y 99 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud de 100 nucleótidos. En una modalidad, dicha secuencia de sgARN es una secuencia de ácido ribonucleico de guía única Nme1Cas9. En una modalidad, dicha secuencia de sgARN es una secuencia de ácido ribonucleico de guía única Nme2Cas9. En una modalidad, tal secuencia de sgARN es una secuencia de ácido ribonucleico de guía única Nme1Cas9 o una secuencia de ácido ribonucleico de guía única Nme2Cas9.

En una modalidad, la presente invención contempla una secuencia de ácido ribonucleico de guía única (sgARN) que comprende una región del vástago 2 truncada. En una modalidad, la secuencia de sgARN comprende además una repetición truncada: región anti-repetición. En una modalidad, la secuencia de sgARN comprende además una región espaciadora truncada. En una
5 modalidad, dicha secuencia de sgARN tiene una longitud que se selecciona del grupo que consiste en 111 nucleótidos, 107 nucleótidos, 105 nucleótidos, 103 nucleótidos, 102 nucleótidos, 101 nucleótidos y 99 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud de 100 nucleótidos.

En una modalidad, la presente invención contempla un vector viral adenoasociado (AAV)
10 que comprende un vector de ácido nucleico de ácido ribonucleico-*Neisseria meningitidis* Cas9 (sgRNA-Nme1Cas9 o sgRNA-Nme2Cas9) de una sola guía. En una modalidad dicho vector de ácido nucleico de ácido ribonucleico de -*Neisseria meningitidis* Cas9 de una sola guía comprende al menos un promotor. En una modalidad, al menos un promotor se selecciona del grupo que consiste en un promotor U6 y un promotor U1a. En una modalidad, vector de ácido
15 nucleico de ácido ribonucleico de -*Neisseria meningitidis* Cas9 comprende una secuencia de Kozak. En una modalidad, dicho sgARN comprende una secuencia de ácido nucleico que es complementaria a una secuencia del gen de interés. En una modalidad, dicha secuencia del gen de interés se selecciona del grupo que consiste en una secuencia *PCSK9* y una secuencia *ROSA26*. En una modalidad, dicho sgARN comprende una secuencia no truncada que tiene una
20 longitud de 145 nucleótidos. En una modalidad, dicho sgARN comprende una repetición truncada: secuencia anti-repetición. En una modalidad, dicho sgARN comprende además una región del vástago 2 truncada. En una modalidad, dicho sgARN comprende además una región espaciadora truncada. En una modalidad, dicha secuencia de sgARN tiene una longitud de 121 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud seleccionada del
25 grupo que consiste en 111 nucleótidos, 107 nucleótidos, 105 nucleótidos, 103 nucleótidos, 102 nucleótidos, 101 nucleótidos y 99 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud de 100 nucleótidos. En una modalidad, dicho sgARN comprende una región del vástago 2 truncada. En una modalidad, dicho sgARN comprende además una repetición truncada: región anti-repetición. En una modalidad, dicho sgARN comprende además una
30 región espaciadora truncada. En una modalidad, dicha secuencia de sgARN tiene una longitud seleccionada del grupo que consiste en 111 nucleótidos, 107 nucleótidos, 105 nucleótidos, 103

nucleótidos, 101 nucleótidos y 99 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud de 100 nucleótidos. En una modalidad, dicho sgARN comprende una secuencia no truncada que tiene una longitud de 145 nucleótidos.

En una modalidad, la presente invención contempla un método, que comprende: a) proporcionar; i) un paciente que exhibe al menos un síntoma de una afección médica, en donde dicho paciente comprende una pluralidad de genes relacionados con tal afección médica; ii) una plataforma de entrega que comprende un vector de ácido nucleico de ácido ribonucleico-*Neisseria meningitidis* Cas9 (sgARN-Nme1Cas9 o sgARN-Nme2Cas9) de guía única, en donde dicho sgARN comprende una secuencia de ácido nucleico que es complementaria a una porción de al menos una de tal pluralidad de genes; y b) administrar dicho plásmido AAV a dicho paciente en condiciones tales que al menos tal síntoma de dicha afección médica se reduzca. En una modalidad, la plataforma de entrega comprende un vector viral adenoasociado (AAV). En una modalidad, la plataforma de entrega comprende una micropartícula. En una modalidad, dicha afección médica comprende hipercolesterolemia. En una modalidad, dicha afección médica comprende tirosinemia. En una modalidad, al menos una de tal pluralidad de genes es un gen *PCSK9*. En una modalidad, dicho ácido nucleico sgARN es complementario a una porción de dicho gen *PCSK9*. En una modalidad, al menos uno de tal pluralidad de genes es un gen *FAH*. En una modalidad, dicho ácido nucleico sgARN es complementario a una porción de dicho gen *FAH*. En una modalidad, dicho sgARN comprende una repetición truncada-secuencia anti-repetición. En una modalidad, dicho sgARN comprende además una región del vástago 2 truncada. En una modalidad, dicho sgARN comprende además una región espaciadora truncada. En una modalidad, dicha secuencia de sgARN tiene una longitud de 121 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud seleccionada del grupo que consiste en 111 nucleótidos, 107 nucleótidos, 105 nucleótidos, 103 nucleótidos, 101 nucleótidos y 99 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud de 100 nucleótidos. En una modalidad, dicho sgARN comprende una región del vástago 2 truncada. En una modalidad, dicho sgARN comprende además una repetición truncada: región anti-repetición. En una modalidad, dicho sgARN comprende además una región espaciadora truncada. En una modalidad, dicha secuencia de sgARN tiene una longitud seleccionada del grupo que consiste en 111 nucleótidos, 107 nucleótidos, 105 nucleótidos, 103 nucleótidos, 102 nucleótidos, 101 nucleótidos y 99 nucleótidos. En una modalidad, dicha secuencia de sgARN tiene una longitud

de 100 nucleótidos. En una modalidad, dicho sgARN comprende una secuencia no truncada que tiene una longitud de 145 nucleótidos.

En una modalidad, la presente invención contempla un plásmido vírico adenoasociado (AAV) que codifica una proteína nucleasa Cas9 de Tipo II-C en donde dicha proteína comprende un dominio de reconocimiento de motivo adyacente al protoespaciador configurado con un sitio de unión a una secuencia de motivo adyacente al protoespaciador que comprende entre uno a cuatro nucleótidos requeridos. En una modalidad, dicha proteína nucleasa Cas9 de Tipo II-C se selecciona del grupo que consiste en una proteína nucleasa de *Neisseria meningitidis cepa De10444* Nme2Cas9, una proteína nucleasa HpaCas9 *Haemophilus parainfluenzae* y una proteína nucleasa SmuCas9 *Simonsiella muelleri*. En una modalidad, dicha secuencia de motivo adyacente al protoespaciador que comprende de uno a cuatro nucleótidos requeridos se selecciona del grupo que consiste en N₄CN₃, N₄CT, N₄CCN, N₄CCA, y N₄GNT₃. En una modalidad, los uno o cuatro nucleótidos requeridos se seleccionan del grupo que consiste en C, CT, CCN, CCA, CN₃ y GNT₂. En una modalidad, dicha proteína nucleasa Cas9 de Tipo II-C está unida a un sgARN truncado. En una modalidad, el plásmido viral adenoasociado codifica dos secuencias de sgARN. En una modalidad, el plásmido viral adenoasociado codifica una secuencia de poli-adenosina. En una modalidad, el plásmido viral adenoasociado codifica una plantilla de nucleótido donante de reparación dirigida por homología. En una modalidad, el plásmido viral adenoasociado es un plásmido viral adenoasociado todo en uno.

En una modalidad, la presente invención contempla un método que comprende: a) proporcionar; i) un paciente que exhibe al menos un síntoma de una afección médica, en donde dicho paciente comprende una pluralidad de genes relacionados con dicha afección médica, en donde dicha pluralidad de genes comprende un motivo adyacente al protoespaciador que comprende entre uno y cuatro nucleótidos requeridos; ii) una plataforma de entrega que comprende al menos un ácido nucleico que codifica una proteína nucleasa Cas9 de Tipo II-C en donde dicha proteína comprende un dominio de reconocimiento de motivo adyacente al protoespaciador configurado con un sitio de unión a dicha secuencia de motivo adyacente al protoespaciador que comprende entre dos y cuatro nucleótidos requeridos; y b) administrar dicha plataforma de entrega a dicho paciente en condiciones tales que al menos tal síntoma de dicha afección médica se reduzca. En una modalidad, dicha afección médica comprende hipercolesterolemia. En una modalidad, dicha afección médica comprende tirosinemia. En una

modalidad, al menos una de tal pluralidad de genes es un gen *PCSK9*. En una modalidad, dicho ácido nucleico sgARN es complementario a una porción de dicho gen *PCSK9*. En una modalidad, al menos uno de tal pluralidad de genes es un gen *FAH*. En una modalidad, dicho ácido nucleico sgARN es complementario a una porción de dicho gen *FAH*. En una modalidad, dicha plataforma de entrega comprende un plásmido viral adenoasociado. En una modalidad, dicha plataforma de entrega comprende una micropartícula. En una modalidad, dicha proteína nucleasa Cas9 de Tipo II-C se selecciona del grupo que consiste en una proteína nucleasa de *Neisseria meningitidis* cepa *De10444* Nme2Cas9, una proteína nucleasa HpaCas9 *Haemophilus parainfluenzae* y una proteína nucleasa SmuCas9 *Simonsiella muelleri*. En una modalidad, dicha secuencia de motivo adyacente al protoespaciador que comprende de uno a cuatro nucleótidos requeridos se selecciona del grupo que consiste en N₄CN₃, N₄CT, N₄CCN, N₄CCA, y N₄GNT₃. En una modalidad, los uno o cuatro nucleótidos requeridos se seleccionan del grupo que consiste en C, CT, CCN, CCA, CN₃ y GNT₂. En una modalidad, dicha proteína nucleasa Cas9 de Tipo II-C está unida a un sgARN truncado. En una modalidad, el plásmido viral adenoasociado codifica dos secuencias de sgARN. En una modalidad, el plásmido viral adenoasociado codifica una secuencia de poli-adenosina. En una modalidad, el plásmido viral adenoasociado codifica una plantilla de nucleótido donante de reparación dirigida por homología. En una modalidad, el plásmido viral adenoasociado es un plásmido viral adenoasociado todo en uno.

En una modalidad, la presente invención contempla un plásmido viral adenoasociado (AAV) que codifica una proteína nucleasa Cas9 Tipo II-C en donde dicha proteína comprende un dominio de reconocimiento de motivo adyacente al protoespaciador (*ej.*, un Dominio de Interacción-PAM; PID) configurado para unirse con una secuencia de motivo adyacente al protoespaciador (PAM), dicha secuencia PAM comprende un par de dinucleótidos de citosina adyacente. En una modalidad, el par dinucleótido de citosina adyacente está en las posiciones PAM cinco (5) y seis (6). En una modalidad, dicha proteína nucleasa Cas9 de Tipo II-C se deriva de una cepa de *Neisseria meningitidis*. En una modalidad, la cepa de *Neisseria meningitidis* es *De10444*. En una modalidad, la proteína nucleasa Cas9 de Tipo II-C es una proteína nucleasa Nme2Cas9. En una modalidad, la cepa de *Neisseria meningitidis* es 98002. En una modalidad, la proteína nucleasa Cas9 de Tipo II-C es una proteína nucleasa Nme3Cas9. En una modalidad, dicha secuencia PAM se selecciona del grupo que consiste en N₄CC, N₄CCN₃N₄CC, N₄CCN₃, N₄CCA, N₄CC(X), N₄CA₃ y N₁₀. En una modalidad, la secuencia

PAM es N₃CC. En una modalidad, la proteína nucleasa Cas9 de Tipo II-C comprende además una secuencia de sgARN. En una modalidad, la secuencia de sgARN comprende un espaciador que varía en longitud entre aproximadamente diecisiete (17) a veinticuatro (24) nucleótidos.

En una modalidad, la presente invención contempla un método, que comprende: a) proporcionar; i) un paciente que exhibe al menos un síntoma de una afección médica, en donde dicho paciente comprende una pluralidad de genes relacionados con dicha afección médica, en donde dicha pluralidad de genes comprende un motivo adyacente al protoespaciador que comprende un par de dinucleótidos de citosina adyacente; ii) una plataforma de entrega que comprende al menos un ácido nucleico que codifica una proteína nucleasa Cas9 de Tipo II-C en donde dicha proteína comprende un dominio de reconocimiento de motivo adyacente al protoespaciador (*ej.*, un dominio de interacción-PAM; PID) configurado para unirse con dicha secuencia de motivo adyacente al protoespaciador que comprende un par de dinucleótidos de citosina adyacente; y b) administrar dicha plataforma de entrega a dicho paciente en condiciones tales que al menos tal síntoma de dicha afección médica se reduzca. En una modalidad, dicha plataforma de entrega comprende un vector viral adenoasociado. En una modalidad, el vector viral adenoasociado es el vector viral adenoasociado ocho (AAV8). En una modalidad, dicha afección médica comprende hipercolesterolemia. En una modalidad, dicha afección médica comprende tirosinemia. En una modalidad, la afección médica es una enfermedad granulomatosa crónica ligada a x. En una modalidad, la afección médica es aspartilglucosaminuria. En una modalidad, al menos una de tal pluralidad de genes es un gen *PCSK9*. En una modalidad, dicho ácido nucleico sgARN es complementario a una porción de dicho gen *PCSK9*. En una modalidad, al menos uno de tal pluralidad de genes es un gen *FAH*. En una modalidad, dicho ácido nucleico sgARN es complementario a una porción de dicho gen *FAH*. En una modalidad, el plásmido viral adenoasociado codifica al menos una secuencia de sgARN. En una modalidad, el plásmido viral adenoasociado codifica dos secuencias de sgARN. En una modalidad, el plásmido viral adenoasociado codifica una secuencia de poli-adenosina. En una modalidad, el plásmido viral adenoasociado codifica una plantilla de nucleótido donante de reparación dirigida por homología. En una modalidad, el plásmido viral adenoasociado es un plásmido viral adenoasociado todo en uno. En una modalidad, dicha plataforma de entrega comprende una micropartícula. En una modalidad, el par dinucleótido de citosina adyacente está en las posiciones PAM cinco (5) y seis (6). En una modalidad, dicha proteína nucleasa Cas9 de Tipo

II-C se deriva de una cepa de *Neisseria meningitidis*. En una modalidad, la cepa de *Neisseria meningitidis* es *De10444*. En una modalidad, la proteína nucleasa Cas9 de Tipo II-C es una proteína nucleasa Nme2Cas9. En una modalidad, la cepa de *Neisseria meningitidis* es 98002. En una modalidad, la proteína nucleasa Cas9 de Tipo II-C es una proteína nucleasa Nme3Cas9.

5 En una modalidad, dicha secuencia PAM se selecciona del grupo que consiste en N4CC, N4CCN3N4CC, N4CCN3, N4CCA, N4CC(X), N4CA3 y N10. En una modalidad, la secuencia PAM es N3CC. En una modalidad, la proteína nucleasa Cas9 de Tipo II-C comprende además una secuencia de sgARN. En una modalidad, la secuencia de sgARN comprende un espaciador que varía en longitud entre aproximadamente diecisiete (17) a veinticuatro (24) nucleótidos.

10

Definiciones

Para facilitar la comprensión de esta invención, se definen varios términos a continuación. Los términos definidos en este documento tienen significados entendidos comúnmente por una persona de habilidad ordinaria en las áreas relevantes para la presente invención. Términos tales

15 como "un", "una", "el" y "la" no pretenden referirse solo a una entidad singular sino también a entidades plurales y también incluyen la clase general de la cual se puede usar un ejemplo específico para ilustración. La terminología en este documento se usa para describir modalidades específicas de la invención, pero su uso no delimita la invención, excepto como se describe en las reivindicaciones.

20

El término "aproximadamente" o "alrededor" como se usa en el presente documento, en el contexto de cualquiera de las mediciones de ensayo, se refiere a +/- 5% de una medición dada.

Como se usa en el presente documento, el "gen *ROSA26*" o el "gen *Rosa26*" se refiere a un locus humano o de ratón (respectivamente) que se usa ampliamente para lograr la expresión generalizada en el ratón. El direccionamiento al locus *ROSA26* se puede lograr mediante la

25 introducción de un gen deseado en el primer intrón del locus, en un sitio *XbaI* único aproximadamente a 248 pb aguas arriba de la línea original de trampa génica. Se puede construir un constructo mediante el uso de un aceptor de empalme de adenovirus seguido de un gen de interés y un sitio de poliadenilación insertado en el sitio *XbaI* único. También se puede incluir un casete de resistencia a la neomicina en el vector objetivo.

30

Como se usa en el presente documento, el "gen *PCSK9*" o el "gen *Pcsk9*" se refieren a un locus humano o de ratón (respectivamente) que codifica una proteína PCSK9. El gen *PCSK9*

reside en el cromosoma 1 en la banda 1p32.3 e incluye 13 exones. Este gen puede producir al menos dos isoformas a través de empalmes alternativos.

El término "proteína convertasa subutilisina/ kexina tipo 9" y "PCSK9" se refiere a una proteína codificada por un gen que modula los niveles de lipoproteínas de baja densidad. La
5 proteína convertasa subutilisina/kexina tipo 9, también conocida como PCSK9, es una enzima que en los humanos está codificada por el gen *PCSK9*. Seidah, "*The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation*" *Proc. Natl. Acad. Sci. EE.UU.* 100 (3): 928-933 (2003). Se encuentran genes similares (ortólogos) en muchas especies. Muchas enzimas, que incluyen PCSK9, están
10 inactivas cuando se sintetizan por primera vez, porque tienen una sección de cadenas peptídicas que bloquea su actividad; las proteínas convertasas eliminan esa sección para activar la enzima. Se cree que PCSK9 juega un papel regulador en la homeostasis del colesterol. Por ejemplo, PCSK9 puede unirse al dominio de repetición A (EGF-A) de tipo factor de crecimiento epidérmico del receptor de lipoproteínas de baja densidad (LDL-R), lo que da como resultado la
15 internalización y degradación de LDL-R. Claramente, se esperaría que los niveles reducidos de LDL-R den como resultado una disminución del metabolismo de LDL-C, lo que podría conducir a hipercolesterolemia.

El término "hipercolesterolemia", como se usa en el presente documento, se refiere a cualquier afección médica en donde los niveles de colesterol en la sangre están elevados por
20 encima de los niveles clínicamente recomendados. Por ejemplo, si el colesterol se mide mediante el uso de lipoproteínas de baja densidad (LDL), puede existir hipercolesterolemia si los niveles de LDL medidos son superiores, por ejemplo, aproximadamente 70 mg/dl. Alternativamente, si el colesterol se mide al usar colesterol libre en plasma, puede existir hipercolesterolemia si los niveles de colesterol libre medidos son superiores, por ejemplo,
25 aproximadamente 200-220 mg/dl.

Como se usa en el presente documento, el término "CRISPR" o "Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Espaciadas" se refiere a un acrónimo para loci de ADN que contiene repeticiones múltiples, cortas y directas de secuencias de base. Cada repetición contiene una serie de bases seguidas de aproximadamente 30 pares de bases conocidas
30 como secuencia "espaciadora". Los espaciadores son segmentos cortos de ADN de un virus y pueden servir como "memoria" de exposiciones pasadas para facilitar una defensa adaptativa

contra futuras invasiones. Doudna, *Genome editing. The new frontier of genome engineering with CRISPR-Cas9*” *Science* 346(6213):1258096 (2014).

Como se usa en el presente documento, el término "Cas" o “CRISPR asociadas (*cas*)” se refiere a genes a menudo asociados con matrices de repetición-espaciador de CRISPR.

5 Como se usa en el presente documento, el término "Cas9" se refiere a una nucleasa de sistemas CRISPR de tipo II, una enzima especializada para generar cortes de doble cadena en el ADN, con dos sitios de corte activos (los dominios HNH y RuvC), uno para cada hebra de la doble hélice. El tracrARN y el ARN espaciador se pueden combinar en una molécula de "ARN de guía única" (sgARN) que, mezclada con Cas9, podría encontrar y escindir objetivos de ADN a través del apareamiento de Watson-Crick entre la secuencia de guía dentro del sgARN y la secuencia de ADN objetivo, Jinek, “*A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity*” *Science* 337(6096):816-821 (2012).

Como se usa en el presente documento, el término "Cas9 catalíticamente activa" se refiere a una nucleasa Cas9 no modificada que comprende actividad de nucleasa completa.

15 El término "nickasa", como se usa en el presente documento, se refiere a una nucleasa que escinde solo una cadena de ADN, ya sea debido a su función natural o porque ha sido modificada para escindir solo una cadena de ADN. Las variantes de nickasa cas9 que tienen el dominio RuvC o HNH mutado proporcionan control sobre qué cadena de ADN se escinde y cuál permanece intacta. Jinek, “*A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity*” *Science* 337(6096):816-821 (2012) y Cong, “*Multiplex genome engineering using CRISPR/Cas systems*” *Science* 339(6121):819-823 (2013).

20 El término "transactivador crARN", "tracrARN" como se usa en el presente documento, se refiere a un ARN pequeño trans-codificado. Por ejemplo, CRISPR/Cas (repeticiones palindrómicas cortas agrupadas y regularmente espaciadas/ proteínas asociadas a CRISPR) constituye un sistema de defensa mediado por ARN, que protege contra virus y plásmidos. Este camino defensivo tiene tres pasos. Primero, una copia del ácido nucleico invasor se integra en el locus CRISPR. A continuación, los ARN CRISPR (crARN) se transcriben de este locus CRISPR. Los crARN se incorporan a los complejos efectores, donde el crARN guía el complejo hacia el ácido nucleico invasor y las proteínas Cas degradan este ácido nucleico. Hay varias vías de activación CRISPR, una de las cuales requiere un tracrARN, que desempeña un papel en la maduración del crARN. TracrARN es complementario a la secuencia de repetición del pre-

crARN, formando un dúplex de ARN. Esto es escindido por RNase III, una ribonucleasa específica a ARN, para formar un híbrido crARN/tracrARN. Este híbrido actúa como una guía para la endonucleasa Cas9, que escinde el ácido nucleico invasor.

El término "motivo adyacente al protoespaciador" (o PAM) como se usa en este documento, se refiere a una secuencia de ADN que puede ser necesaria para que una Cas9/sgARN forme un bucle R para interrogar una secuencia de ADN específica a través del apareamiento de Watson-Crick de su ARN guía con el genoma. La especificidad de PAM puede ser una función de la especificidad de unión al ADN de la proteína Cas9 (ej., un "dominio de reconocimiento de motivo adyacente al protoespaciador" en el extremo C-terminal de Cas9).

Los términos "dominio de reconocimiento de motivo adyacente al protoespaciador", "Dominio de interacción PAM" o "PID", como se usa en el presente documento, se refiere a una secuencia de aminoácidos Cas9 que comprende un sitio de unión a una secuencia PAM objetivo de ADN.

El término "sitio de unión", como se usa en el presente documento, se refiere a cualquier disposición molecular que tenga una estructura terciaria y/o cuaternaria específica que experimente una unión física o asociación estrecha con un componente de unión. Por ejemplo, la disposición molecular puede comprender una secuencia de aminoácidos. Alternativamente, la disposición molecular puede comprender una secuencia de ácidos nucleicos. Además, la disposición molecular puede comprender una bicapa lipídica u otro material biológico.

Como se usa en el presente documento, el término "sgARN" se refiere al ARN de guía única usado junto con los sistemas CRISPR asociados (*Cas*). Los sgARN son una fusión de crARN y tracrARN y contienen nucleótidos de secuencia complementarios al sitio objetivo deseado. Jinek, "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity" *Science* 337(6096):816-821 (2012). El apareamiento de Watson-Crick del sgARN con el sitio objetivo permite la formación del bucle R, que junto con un PAM funcional permite la escisión del ADN o, en el caso de Cas9 deficiente en nucleasas, se une al ADN en ese locus.

Como se usa en el presente documento, el término "ortogonal" se refiere a objetivos que no se superponen, no están correlacionados o son independientes. Por ejemplo, si se utilizaran dos isoformas ortogonales Cas9, emplearían sgARN ortogonales que solo programan una de las isoformas Cas9 para el reconocimiento y la escisión del ADN. Esvelt, "Orthogonal Cas9 proteins for RNA-guided gene regulation and editing" *Nat Methods* 10(11):1116-1121 (2013).

Por ejemplo, esto permitiría que una isoforma Cas9 (ej., *S. pyogenes* Cas9 o SpyCas9) funcione como una nucleasa programada por un sgARN que puede ser específico para la misma, y otra isoforma Cas9 (ej., *N. meningitidis* Cas9 o NmeCas9) para operar como una nucleasa muerta Cas9 que proporciona ADN dirigido a un sitio de unión a través de su especificidad PAM y sgARN ortogonal. Otras Cas9 incluyen *S. aureus* Cas9 o SauCas9 y *A. naeslundii* Cas9 o AnaCas9.

El término "truncado" como se usa en el presente documento, cuando se usa en referencia a una secuencia de polinucleótidos o una secuencia de aminoácidos significa que al menos una porción de la secuencia de tipo silvestre puede estar ausente. En algunos casos, las secuencias guía truncadas dentro del sgARN o crARN pueden mejorar la precisión de edición de Cas9. Fu, "Improving CRISPR-Cas nuclease specificity using truncated guide RNAs" Nat Biotechnol. 2014 Mar;32(3):279-284 (2014).

El término "pares de bases", como se usa en el presente documento, se refiere a nucleobases específicas (también denominadas bases nitrogenadas), que son los bloques de construcción de secuencias de nucleótidos que forman una estructura primaria de ADN y ARN. El ADN bicatenario puede caracterizarse por patrones específicos de enlace de hidrógeno. Los pares de bases pueden incluir, entre otros, pares de bases guanina-citosina y adenina-timina.

El término "objetivo genómico específico" como se usa en el presente documento, se refiere a cualquier secuencia de nucleótidos predeterminada capaz de unirse a una proteína Cas9 contemplada en el presente documento. El objetivo puede incluir, pero no se limita a, una secuencia de nucleótidos complementaria a un dominio de unión a ADN programable o una proteína Cas9 ortogonal programada con su propio ARN guía, una secuencia de nucleótidos complementaria a un solo ARN guía, una secuencia de reconocimiento de motivo adyacente al protoespaciador, una secuencia de unión con objetivos en lugar y una secuencia de unión con objetivos fuera de lugar.

El término "secuencia de unión objetivos en lugar" como se usa en el presente documento, se refiere a una subsecuencia de un objetivo genómico específico que puede ser completamente complementario a un dominio de unión a ADN programable y/o una secuencia de ARN de guía única.

El término "secuencia de unión con objetivos fuera de lugar" como se usa en el presente documento, se refiere a una subsecuencia de un objetivo genómico específico que puede ser

parcialmente complementario a un dominio de unión a ADN programable y/o una secuencia de ARN de guía única.

El término "no se une" como se usa en el presente documento, se refiere a cualquier interacción nucleótido-nucleótido o una interacción nucleótido-aminoácido que exhibe complementariedad parcial, pero tiene complementariedad insuficiente para el reconocimiento para desencadenar la escisión del sitio objetivo por la nucleasa Cas9. Dicha falla de unión puede dar como resultado una unión débil o parcial de dos moléculas, de modo que falla una función biológica esperada (*ej.*, actividad nucleasa).

El término "escisión", como se usa en el presente documento, puede definirse como la generación de un corte en el ADN. Esto podría ser una corte monocatenario o una corte bicatenario dependiendo del tipo de nucleasa que se pueda emplear.

Como se usa en el presente documento, el término "editar" o "editado" se refiere a un método para alterar una secuencia de ácido nucleico de un polinucleótido (*ej.*, una secuencia de ácido nucleico natural o una secuencia de origen natural mutada) mediante la eliminación selectiva de un objetivo genómico específico o la inclusión específica de una nueva secuencia mediante el uso de una plantilla de ADN suministrada de forma exógena. Tal objetivo genómico específico incluye, pero no se limita a, una región cromosómica, ADN mitocondrial, un gen, un promotor, un marco de lectura abierto o cualquier secuencia de ácido nucleico.

Los términos "deleción", "eliminación", "eliminar" o "eliminado" como se usa en el presente documento, pueden definirse como un cambio en la secuencia de nucleótidos o aminoácidos en la que uno o más nucleótidos o residuos de aminoácidos, respectivamente, son, o se vuelven ausentes.

El término "gen de interés", como se usa en el presente documento, se refiere a cualquier gen predeterminado para el que puede desearse la deleción.

El término "alelo", como se usa en el presente documento, se refiere a cualquiera de una serie de formas alternativas del mismo gen o el mismo locus genético.

El término "cantidad efectiva" como se usa en el presente documento, se refiere a una cantidad particular de una composición farmacéutica que comprende un agente terapéutico que logra un resultado clínicamente beneficioso (es decir, por ejemplo, una reducción de los síntomas). La toxicidad y la eficacia terapéutica de tales composiciones pueden determinarse mediante procedimientos farmacéuticos estándar en cultivos celulares o animales

experimentales, *ej.*, para determinar la LD₅₀ (la dosis letal para el 50% de la población) y la ED₅₀ (la dosis terapéuticamente efectiva en el 50% de la población). La relación de dosis entre los efectos tóxicos y terapéuticos es el índice terapéutico, y se puede expresar como la relación, LD₅₀/ED₅₀. Se prefieren los compuestos que exhiben grandes índices terapéuticos. Los datos
5 obtenidos de estos ensayos de cultivo celular y estudios adicionales en animales pueden usarse para formular un rango de dosificación para uso humano. La dosificación de tales compuestos reside preferiblemente dentro de un rango de concentraciones circulantes que incluyen la ED₅₀ con poca o ninguna toxicidad. La dosificación varía dentro de este rango dependiendo de la forma de dosificación empleada, la sensibilidad del paciente y la vía de administración.

10 El término "síntoma", como se usa en el presente documento, se refiere a cualquier evidencia subjetiva u objetiva de enfermedad o alteración física observada por el paciente. Por ejemplo, la evidencia subjetiva generalmente se basa en el autoinforme del paciente y puede incluir, entre otros, dolor, dolor de cabeza, trastornos visuales, náuseas y/o vómito. Alternativamente, la evidencia objetiva generalmente es el resultado de pruebas médicas que
15 incluyen, entre otras, temperatura corporal, recuento sanguíneo completo, paneles de lípidos, paneles de tiroides, presión arterial, frecuencia cardíaca, electrocardiograma, escáner de imágenes corporal y/o de tejidos.

El término "enfermedad" o "afección médica", como se usa en este documento, se refiere a cualquier deterioro del estado normal del cuerpo animal o vegetal vivo o una de sus partes que
20 interrumpe o modifica el desempeño de las funciones vitales. Típicamente se manifiesta al distinguir signos y síntomas, generalmente es una respuesta a: i) factores ambientales (tales como la desnutrición, los riesgos industriales o el clima); ii) agentes infecciosos específicos (tales como gusanos, bacterias o virus); iii) defectos inherentes del organismo (tales como anomalías genéticas); y/o iv) combinaciones de estos factores.

25 Los términos "reducir", "inhibir", "disminuir", "suprimir", "prevenir" y equivalentes gramaticales (incluyendo "inferior", "más pequeño", etc.) cuando se refieren a la expresión de cualquier síntoma en un sujeto no tratado en relación con un sujeto tratado, significa que la cantidad y/o magnitud de los síntomas en el sujeto tratado es menor que en el sujeto no tratado por cualquier cantidad que sea reconocida como clínicamente relevante por cualquier personal
30 médicamente capacitado. En una modalidad, la cantidad y/o magnitud de los síntomas en el sujeto tratado es al menos un 10% más bajo, al menos un 25% más bajo, al menos un 50% más

bajo, al menos un 75% más bajo y/o al menos 90% más bajo que la cantidad y/o magnitud de los síntomas en el sujeto no tratado.

El término "unido", como se usa en el presente documento, se refiere a cualquier interacción entre un medio (o vehículo) y un fármaco. La unión puede ser reversible o

5 irreversible. Tal unión incluye, pero no se limita a, unión covalente, unión iónica, fuerzas o fricción de Van der Waals, y similares. Un medicamento se une a un medio (o vehículo) si está impregnado, incorporado, recubierto, en suspensión con, en solución con, mezclado con, *etc.*

El término "fármaco" o "compuesto", como se usa en el presente documento, se refiere a cualquier sustancia farmacológicamente activa capaz de administrarse que logra un efecto
10 deseado. Los fármacos o compuestos pueden ser sintéticos o de origen natural, no péptidos, proteínas o péptidos, oligonucleótidos o nucleótidos, polisacáridos o azúcares.

El término "administrado" o "administrar", como se usa en el presente documento, se refiere a cualquier método para proporcionar una composición a un paciente de modo que la composición tenga el efecto deseado en el paciente. Un método de ejemplo de administración es
15 mediante un mecanismo directo como la administración local de tejido (es decir, por ejemplo, colocación extravascular), ingestión oral, parche transdérmico, tópica, inhalación, supositorio, *etc.*

El término "paciente" o "sujeto", como se usa en el presente documento, es humano o animal y no necesita ser hospitalizado. Por ejemplo, pacientes ambulatorios, personas en
20 hogares de ancianos son "pacientes". Un paciente puede comprender cualquier edad de un humano o animal no humano y, por lo tanto, incluye tanto adultos como juveniles (es decir, niños). No se pretende que el término "paciente" denote una necesidad de tratamiento médico, por lo tanto, un paciente puede ser voluntaria o involuntariamente parte de la experimentación, ya sea clínica o en apoyo de estudios de ciencias básicas.

25 El término "afinidad", como se usa en el presente documento, se refiere a cualquier fuerza de atracción entre sustancias o partículas que hace que entren y permanezcan en combinación química. Por ejemplo, un compuesto inhibidor que tiene una alta afinidad por un receptor proporcionará una mayor eficacia para evitar que el receptor interactúe con sus ligandos naturales, que un inhibidor con baja afinidad.

30 El término "derivado de" como se usa en el presente documento, se refiere a la fuente de un compuesto o secuencia. En un aspecto, un compuesto o secuencia puede derivarse de un

organismo o especie particular. En otro aspecto, un compuesto o secuencia puede derivarse de un complejo o secuencia más grande.

El término "proteína", como se usa en el presente documento, se refiere a cualquiera de las numerosas sustancias extremadamente complejas de origen natural (como una enzima o anticuerpo) que consisten en residuos de aminoácidos unidos por enlaces peptídicos, que contienen los elementos carbono, hidrógeno, nitrógeno, oxígeno, generalmente azufre. En general, una proteína comprende aminoácidos que tienen un orden de magnitud dentro de los cientos.

El término "péptido", como se usa en el presente documento, se refiere a cualquiera de varias amidas que se derivan de dos o más aminoácidos mediante la combinación del grupo amino de un ácido con el grupo carboxilo de otro y generalmente se obtienen por hidrólisis parcial de proteínas. En general, un péptido comprende aminoácidos que tienen un orden de magnitud con las decenas.

El término "polipéptido" se refiere a cualquiera de varias amidas que se derivan de dos o más aminoácidos mediante la combinación del grupo amino de un ácido con el grupo carboxilo de otro y generalmente se obtienen por hidrólisis parcial de proteínas. En general, un péptido comprende aminoácidos que tienen un orden de magnitud con las decenas o más.

El término "farmacéuticamente" o "farmacológicamente aceptable", como se usa en el presente documento, se refiere a entidades moleculares y composiciones que no producen reacciones adversas, alérgicas u otras reacciones desfavorables cuando se administran a un animal o un humano.

El término "vehículo farmacéuticamente aceptable", como se usa en el presente documento, incluye cualquier y todos los disolventes, o un medio de dispersión que incluye, pero no se limita a, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares), mezclas adecuadas de los mismos y aceites vegetales, recubrimientos, agentes isotónicos y retardantes de la absorción, liposomas, limpiadores disponibles comercialmente y similares. Los ingredientes bioactivos suplementarios también se pueden incorporar en tales vehículos.

"Secuencia de ácido nucleico" y "secuencia de nucleótidos", como se usa en el presente documento, se refieren a un oligonucleótido o polinucleótido, y fragmentos o porciones de los

mismos, y a ADN o ARN de origen genómico o sintético que pueden ser monocatenarios o bicatenarios, y representan la hebra sentido o antisentido.

El término "un ácido nucleico aislado", como se usa en el presente documento, se refiere a cualquier molécula de ácido nucleico que se haya eliminado de su estado natural (*ej.*, eliminado de una célula y, en una modalidad preferida, esté libre de otro ácido nucleico genómico).

Los términos "secuencia de aminoácidos" y "secuencia de polipéptidos", como se usan en el presente documento, son intercambiables y se refieren a una secuencia de aminoácidos.

Como se usa en el presente documento, el término "porción" cuando se refiere a una proteína (como en "una porción de una proteína dada") se refiere a fragmentos de esa proteína. El tamaño de los fragmentos puede variar desde cuatro residuos de aminoácidos hasta la secuencia de aminoácidos completa menos un aminoácido.

El término "porción" cuando se usa en referencia a una secuencia de nucleótidos se refiere a fragmentos de esa secuencia de nucleótidos. El tamaño de los fragmentos puede variar desde 5 residuos de nucleótidos hasta la secuencia completa de nucleótidos menos un residuo de ácido nucleico.

El término "biológicamente activo" se refiere a cualquier molécula que tenga funciones estructurales, reguladoras o bioquímicas. Por ejemplo, la actividad biológica puede determinarse, por ejemplo, mediante la restauración del crecimiento de tipo silvestre en células que carecen de actividad proteica. Las células que carecen de actividad proteica pueden producirse mediante muchos métodos (es decir, por ejemplo, mutación puntual y mutación de cambio de marco). La complementación se logra mediante la transfección de células que carecen de actividad proteica con un vector de expresión que expresa la proteína, un derivado de la misma o una porción de la misma.

Como se usa en el presente documento, los términos "complementario" o "complementariedad" se usan en referencia a "polinucleótidos" y "oligonucleótidos" (son términos intercambiables que se refieren a una secuencia de nucleótidos) relacionados por las reglas de emparejamiento de bases. Por ejemplo, la secuencia "C-A-G-T" es complementaria a la secuencia "G-T-C-A". La complementariedad puede ser "parcial" o "total". La complementariedad "parcial" es cuando una o más bases de ácido nucleico no coinciden de acuerdo con las reglas de emparejamiento de bases. La complementariedad "total" o "completa" entre los ácidos nucleicos es donde cada una de las bases de ácidos nucleicos son apareadas con

otra base bajo las reglas de apareamiento de bases. El grado de complementariedad entre las cadenas de ácido nucleico tiene efectos significativos en la eficiencia y resistencia de la hibridación entre cadenas de ácido nucleico. Esto es de particular importancia en las reacciones de amplificación, así como en los métodos de detección que dependen de la unión entre los ácidos nucleicos.

Como se usa en el presente documento, el término "hibridación" se usa en referencia al emparejamiento de ácidos nucleicos complementarios mediante el uso de cualquier proceso mediante el cual una cadena de ácido nucleico se une con una cadena complementaria a través del emparejamiento de bases para formar un complejo de hibridación. La hibridación y la fuerza de la hibridación (es decir, la fuerza de la asociación entre los ácidos nucleicos) se ve afectada por factores tales como el grado de complementariedad entre los ácidos nucleicos, la rigurosidad de las condiciones involucradas, la T_m del híbrido formado y la relación G:C dentro de los ácidos nucleicos.

Como se usa en el presente documento, el término "complejo de hibridación" se refiere a un complejo formado entre dos secuencias de ácido nucleico en virtud de la formación de enlaces de hidrógeno entre bases complementarias G y C y entre bases complementarias A y T; estos enlaces de hidrógeno pueden estabilizarse aún más mediante interacciones de apilamiento de bases. Las dos secuencias de ácido nucleico complementarias se unen con hidrógeno en una configuración antiparalela. Se puede formar un complejo de hibridación en solución (*ej.*, análisis C_0 t o R_0 t) o entre una secuencia de ácido nucleico presente en la solución y otra secuencia de ácido nucleico inmovilizada en un soporte sólido (*ej.*, una membrana de nylon o un filtro de nitrocelulosa como se emplea en la Southern blotting y Northern blotting, dot blotting o un portaobjetos de vidrio como se emplea en la hibridación *in situ*, incluyendo FISH (hibridación fluorescente *in situ*).

Las señales de control transcripcional en eucariontes comprenden elementos "promotores" y "potenciadores". Los promotores y potenciadores consisten en matrices cortas de secuencias de ADN que interactúan específicamente con proteínas celulares involucradas en la transcripción. Maniatis, T., *Science* 236:1237 (1987). Los elementos promotores y potenciadores se han aislado de una variedad de fuentes eucariotas que incluyen genes en células y virus de plantas, levaduras, insectos y mamíferos (elementos de control análogos, es decir, promotores,

también se encuentran en procariotas). La selección de un promotor y potenciador particular depende del tipo de célula que se utilizará para expresar la proteína de interés.

El término "sitio poli A" o "secuencia poli A" como se usa en el presente documento denota una secuencia de ADN que dirige tanto la terminación como la poliadenilación del transcrito de ARN naciente. La poliadenilación eficiente de la transcripción recombinante es deseable ya que las transcripciones que carecen de una cola de poli A son inestables y se degradan rápidamente. La señal de poli A utilizada en un vector de expresión puede ser "heteróloga" o "endógena". Una señal endógena de poli A es aquella que se encuentra naturalmente en el extremo 3' de la región codificante de un gen dado en el genoma. Una señal de poli A heteróloga es aquella que se aísla de un gen y se coloca en 3' de otro gen. La expresión eficiente de secuencias de ADN recombinante en células eucariotas implica la expresión de señales que dirigen la terminación eficiente y la poliadenilación de la transcripción resultante. Las señales de terminación de la transcripción se encuentran generalmente aguas abajo de la señal de poliadenilación y tienen unos cientos de nucleótidos de longitud.

El término "transfección" o "transfectado" se refiere a la introducción de ADN foráneo en una célula.

Como se usa en el presente documento, los términos "codificación de molécula de ácido nucleico", "codificación de la secuencia de ADN" y "codificación de ADN" se refieren al orden o secuencia de desoxirribonucleótidos a lo largo de una cadena de ácido desoxirribonucleico. El orden de estos desoxirribonucleótidos determina el orden de los aminoácidos a lo largo de la cadena de polipéptidos (proteínas). La secuencia de ADN codifica así la secuencia de aminoácidos.

Como se usa en el presente documento, el término "región codificante" cuando se usa en referencia a un gen estructural se refiere a las secuencias de nucleótidos que codifican los aminoácidos encontrados en el polipéptido naciente como resultado de la traducción de una molécula de ARNm. La región de codificación está limitada, en eucariotas, en el lado 5' por el triplete de nucleótidos "ATG" que codifica la metionina iniciadora y en el lado 3' por uno de los tres tripletes que especifican los codones de parada (es decir, TAA, TAG, TGA)

Como se usa en el presente documento, el término "gen estructural" se refiere a una secuencia de ADN que codifica ARN o una proteína. En contraste, los "genes reguladores" son

genes estructurales que codifican productos que controlan la expresión de otros genes (*ej.*, factores de transcripción).

Como se usa en el presente documento, el término "gen" se refiere a las secuencias de desoxirribonucleótidos que comprenden la región de codificación de un gen estructural e incluyen secuencias ubicadas adyacentes a la región de codificación en ambos extremos 5' y 3' para una distancia de aproximadamente 1 kb en cada extremo que el gen corresponde a la longitud del ARNm de longitud completa. Las secuencias que se encuentran 5' de la región de codificación y que están presentes en el ARNm se denominan secuencias 5' no traducidas. Las secuencias que se encuentran 3' o aguas abajo de la región de codificación y que están presentes en el ARNm se denominan secuencias 3' no traducidas. El término "gen" abarca tanto el ADNc como las formas genómicas de un gen. Una forma genómica o clon de un gen contiene la región codificante interrumpida con secuencias no codificantes denominadas "intrones" o "regiones intervinientes" o "secuencias intervinientes". Los intrones son segmentos de un gen que se transcriben en ARN nuclear heterogéneo (hnARN); los intrones pueden contener elementos reguladores tales como potenciadores. Los intrones se eliminan o "desempalman" de la transcripción nuclear o primaria; por lo tanto, los intrones están ausentes en la transcripción del ARN mensajero (ARNm). El ARNm funciona durante la traducción para especificar la secuencia o el orden de los aminoácidos en un polipéptido naciente.

Además de contener intrones, las formas genómicas de un gen también pueden incluir secuencias ubicadas tanto en el extremo 5' como en el 3' de las secuencias que están presentes en el transcrito de ARN. Estas secuencias se denominan secuencias o regiones "flanqueantes" (estas secuencias flanqueantes están ubicadas 5' o 3' a las secuencias no traducidas presentes en el transcrito de ARNm). La región flanqueante 5' puede contener secuencias reguladoras tales como promotores y potenciadores que controlan o influyen en la transcripción del gen. La región flanqueante 3' puede contener secuencias que dirigen la terminación de la transcripción, la escisión postranscripcional y la poliadenilación.

El término "vector viral" abarca cualquier constructo de ácido nucleico derivado de un genoma de virus capaz de incorporar secuencias de ácido nucleico heterólogo para su expresión en un organismo huésped. Por ejemplo, dichos vectores virales pueden incluir, pero no se limitan a, vectores virales adenoasociados, vectores lentivirales, vectores virales SV40, vectores retrovirales, vectores adenovirales. Aunque los vectores virales se crean ocasionalmente a partir

de virus patógenos, pueden modificarse de tal manera que se minimice su riesgo general para la salud. Esto generalmente implica la eliminación de una parte del genoma viral involucrado con la replicación viral. Tal virus puede infectar eficazmente las células, pero, una vez que se ha producido la infección, el virus puede requerir un virus auxiliar para proporcionar las proteínas faltantes para la producción de nuevos viriones. Preferiblemente, los vectores virales deberían tener un efecto mínimo en la fisiología de la célula que infectan y exhibir propiedades genéticamente estables (*ej.*, no experimentan un reordenamiento genómico espontáneo). La mayoría de los vectores virales están modificados para infectar la mayor variedad posible de tipos de células. Aun así, un receptor viral se puede modificar para dirigir el virus a un tipo específico de célula. Se dice que los virus modificados de esta manera están seudotipados. Los vectores virales a menudo están modificados para incorporar ciertos genes que ayudan a identificar qué células tomaron los genes virales. Estos genes se denominan genes marcadores. Por ejemplo, un gen marcador común confiere resistencia antibiótica a cierto antibiótico.

15 Breve Descripción de las Figuras

El archivo de esta patente contiene al menos un dibujo ejecutado a color. La Oficina de Patentes y Marcas proporcionará copias de esta patente con dibujos a color, previa solicitud y pago de la tarifa necesaria.

La Figura 1 presenta la secuencia representativa de un sgARN Nme1Cas9 y Nme2Cas9 convencional, de 145 nt y de longitud completa.

La figura 2 presenta secuencias de sgARN Nme1Cas9 de ejemplo y una actividad de edición de genes asociada que tiene una repetición truncada: región anti-repetición o una región truncada del vástago 2. Serie de eliminación/truncamiento de sgARNs Nme1Cas9. Parte superior: secuencias alineadas, codificadas por colores como en la Figura 1. Parte inferior: T7E1 ensayos de edición en el sitio 7 objetivo Nme1Cas9 (NTS7), utilizando los sgARN indicados como guías.

La figura 3 presenta secuencias de sgARN Nme1Cas9 de ejemplo y una actividad de edición de genes asociada que tiene una repetición truncada: región anti-repetición o una región truncada del vástago 2. Los sgARN Nme1Cas9 más cortos (# 10 - 101 nt; secuencia guía de 24 nt; y # 11 - 100 nt; secuencia guía de 23 nt) editan eficientemente tres sitios objetivo distintos (NTS7, NTS27 y NTS55) en el genoma humano. Parte superior: secuencias de sgARNs de tipo

silvestre y minimizadas, utilizando el mismo esquema de color que en las figuras anteriores.
Parte inferior: Ensayos T7E1 de eficiencia de edición en los tres sitios objetivo en células HEK293T.

La Figura 4 presenta secuencias de ejemplo (como estructuras secundarias) de sgARN de tipo silvestre Nme1Cas9, y sgARNs truncados 11 y 12 y la edición de genes asociada por la entrega de RNP de Nme1Cas9 y sgARNs. Los tres sitios genómicos (N-TS72, N-TS55 y N-TS40), y un sitio reportero de semáforo fueron dirigidos al genoma humano al usar células HEK293T. Parte superior: secuencias mostradas como estructuras secundarias de sgARN de tipo silvestre y minimizadas. Parte inferior: Las eficiencias de edición medidas mediante el ensayo T7E1 o la citometría de flujo se representan como gráficos de barras.

La Figura 5 presenta la edición de genes en células PLB985 mediante el uso de sgARN 11 minimizado, y el sgARN de tipo silvestre transcrito in vitro. Las células se transfectaron con complejos RNP de Nme1Cas9 y sgARNs y la edición de genes en el sitio genómico N-TS72 medida por TIDE.

La Figura 6 presenta un esquema de una modalidad de un vector AAV que comprende un complejo completo de edición de genes CRISPR/Cas9. Las secuencias representativas de las diversas regiones del vector AAV están codificadas por colores en el Apéndice 1.

La Figura 7 presenta una modalidad de una secuencia codificada por color de ARN de guía única Nme y un promotor como se representa en la Figura 4, en donde la estructura principal se linealiza al usar SapI para insertar un espaciador objetivo 24 nt.

Promotor U6: Turquesa.

ARN de guía única Nme: Morada

Sitios de restricción SapI: Negrita

La Figura 8 presenta una modalidad de una secuencia codificada por colores de un Nme1Cas9 y un promotor como se representa en la Figura 4, en donde los codones de inicio y de parada están subrayados en negrita.

Promotor U1a: Azul

Secuencia de Kozak: Gris

Nme1Cas9 humanizada: Rojo

SV40 NLS: Verde

Nucleoplasmina (NP) NLS: Amarillo

Etiquetas HA (3X): Naranja en negrita

NLS sintética: Turquesa

Señal de poliadenilación de beta-globina: Verde azulado

La Figura 9 presenta datos de ejemplo que muestran la eficiencia de edición de varios
5 sitios objetivo que usan plásmidos AAV con constructos sgARN-Nme1Cas9 guiados al gen
Pcsk9 o al gen *Rosa26* (control).

La Figura 10 presenta una modalidad de secuencias de sitio objetivo codificadas por
color para los constructos sgARN-Nme1Cas9 guiados a un gen *Pcsk9* o un gen *Rosa26* (control).

Espaciador objetivo Nme1Cas9 de 24 nt, color azul en negrita

10 Nme1Cas9 PAM subrayado [NNNNGATT)

Sitios de unión de cebadores T7E1: cursiva verde

Sitios de unión de cebadores TIDE: cursiva morada

La Figura 11 presenta datos de ejemplo que muestran la eficiencia de la edición de genes
después de la inyección hidrodinámica *in vivo* por la vena de la cola de ratón de 30 µg de
15 plásmido sgARN-Nme1Cas9-AAV libre de endotoxina dirigido a *Pcsk9*.

La Figura 12A presenta datos de ejemplo que muestran la eficiencia de la edición de
genes en el hígado en el gen *Pcsk9* y el gen *Rosa26* mediante el vector Nme1-Cas9 empaquetado
en serotipo AAV8 específico a hepatocitos, a una dosis de 4×10^{11} copias genómicas (gc) por
ratón 14 días después de la administración del vector. La Figura 12B presenta datos de ejemplo
20 que muestran la eficiencia de la edición de genes en el hígado en un gen *Pcsk9* y un gen *Rosa26*
por un vector Nme1-Cas9 empaquetado en el serotipo AAV8 específico a hepatocitos, a una
dosis de 4×10^{11} copias genómicas (gc) por ratón 50 días después de la administración del vector.

La Figura 13 presenta datos de ejemplo que muestran la reducción en los niveles de
colesterol de ratón después de la inyección de vectores sgARN-Cas9-AAV dirigidos a un gen
25 *Pcsk9*, un gen *Rosa26* y un grupo de control PBS a los 0, 25 y 50 días.

Las Figuras 14A y 14B presentan datos de ejemplo que muestran una identificación
imparcial de todo el genoma de cortes de doble cadena (DSB) habilitada mediante el ensayo de
secuenciación (*ej.*, GUIDE-Seq) que buscó sitios de edición con objetivos fuera de lugar tanto
para *Pcsk9*-sgARN-Cas9-AAV (A) como para *Rosa26*-sgARN-Cas9-AAV (B).

30 La Figura 15 presenta datos de ejemplo que muestran un análisis TIDE dirigido en
ratones 14 días después de la inyección tanto del *Pcsk9*-sgARN-Cas9-AAV como del *Rosa26*-

sgARN-Cas9-AAV que revelaron una escisión mínima. OnT, sitio con objetivos en lugar; OT1, OT2, etc.: sitios con objetivos fuera de lugar.

La Figura 16 presenta datos de ejemplo que muestran un ensayo de tinción con hematoxilina y eosina en las secciones de hígado de ratones sacrificados el día 14 posterior a la inyección de vectores dirigidos a un gen *Pcsk9* y un gen *Rosa26*. No se observa evidencia de una respuesta inmune del huésped.

La Figura 17 ilustra una modalidad de un flujo de trabajo de identificación de biblioteca PAM *in vitro*. NGS, secuenciación de próxima generación.

La Figura 18 presenta una secuencia putativa de un ensayo de descubrimiento PAM *in vitro* representado en la Figura 17. Cas9 purificada recombinantemente de cada bacteria se incubó con un sgARN y un objetivo con PAM aleatorizado. Nme1Cas9 se utilizó como control.

La Figura 19 presenta datos de ejemplo que muestran el porcentaje de edición del genoma en un único sitio (panel superior) en el genoma humano en células HEK293T. Los porcentajes muestran la formación de indel estimada mediante el uso de un ensayo de endonucleasa T7E1 (Nme2Cas9, HpaCas9) o un ensayo fluorescente (para SmuCas9) basado en el reportero de "semáforo" integrado en el genoma de las células HEK293T.

La Figura 20 presenta datos de ejemplo que muestran la edición del genoma en las células HEK293T de un reportero de semáforo integrado con Nme2Cas9 dirigido a varios protoespaciadores con varios PAM (eje X). Los resultados sugieren un NNNNCC PAM preferido para Nme2Cas9 en células humanas.

La Figura 21 presenta datos de ejemplo que muestran la edición del genoma en células HEK293T en presencia de diversas proteínas anti-CRISPR (Acr). La digestión con T7E1 muestra la edición del genoma después de la transfección del plásmido (para expresar Nme2Cas9 y su sgARN) o la entrega de ARN/proteína (HpaCas9 y su sgARN). Nme2Cas9 es fuertemente inhibido por dos proteínas Acr (AcrIIC3_{Nme} y AcrIIC4_{Hpa}), mientras que HpaCas9 es inhibido por cuatro de los Acrs tipo II-C previamente reportados. Estos resultados muestran que estas dos proteínas Cas9 están sujetas al control de inactivación por anti-CRISPR.

La Figura 22 presenta datos de ejemplo de la edición del gen reportero de semáforo (TLR) al usar el complejo Nme2Cas9-sgARN en los PAM de dinucleótidos "CC". Figura 22A. Las barras azules son el porcentaje de células que exhiben fluorescencia, mientras que las barras

rojas indican el porcentaje de edición con mayor precisión basado en la secuenciación ("análisis TIDE").

La Figura 23 presenta datos de ejemplos de la edición de genes por Nme2Cas9 mediante el uso de ensayos T7E1 en los loci *AAVS1*, Cromosoma 14 NTS4, *VEGF* y *CFTR*.

5 La Figura 24 presenta una modalidad para una secuencia de ADN de marco de lectura abierta bacteriana Nme2Cas9 de tipo silvestre.

La Figura 25 presenta una modalidad de una secuencia de proteína bacteriana Nme2Cas9 de tipo silvestre.

10 La Figura 26 presenta una modalidad de una secuencia de ADN de marco de lectura abierto optimizado con codón humano Nme2Cas9. Amarillo - SV40 NLS; Verde - 3X-HA-Tag; Azul: NLS tipo cMyc.

La Figura 27 presenta una modalidad de una secuencia de proteína humanizada Nme2Cas9. Amarillo - SV40 NLS; Verde - 3X-HA-Tag; Azul: NLS tipo cMyc.

La Figura 28 presenta una modalidad de una secuencia de proteína bacteriana HpaCas9.

15 La Figura 29 presenta una modalidad de una secuencia de ADN de marco de lectura abierta bacteriana nativa SmuCas9.

La Figura 30 presenta una modalidad de una secuencia de proteína bacteriana SmuCas9.

20 La Figura 31 presenta una modalidad de una secuencia de ADN de marco de lectura abierto optimizado con codón humano SmuCas9. Amarillo - SV40 NLS; Verde - 3X-HA-Tag; Azul: NLS tipo cMyc.

La Figura 32 presenta una modalidad de una secuencia de proteína humanizada SmuCas9. Amarillo - SV40 NLS; Verde - 3X-HA-Tag; Azul: NLS tipo cMyc.

25 La Figura 33 presenta ejemplos de secuencias de ARN de guía única ortólogas Cas9 de tipo II C compatibles con PAM cortos ricos en C. Amarillo - crARN; Gris - Linker; Púrpura - tracrARN.

La Figura 34 ilustra tres ortólogos *Neisseria meningitidis* Cas9 estrechamente relacionados que tienen PAM distintos.

30 Figura 34A: Esquema que muestra residuos mutados (esferas naranjas) entre Nme2Cas9 (izquierda) y Nme3Cas9 (derecha) mapeados en la estructura predicha de Nme1Cas9, revelando el grupo de mutaciones en el PID (negro).

Figura 34B: Flujo de trabajo experimental del ensayo de descubrimiento de PAM *in vitro* con una secuencia PAM aleatorizada 10 nt aguas abajo de un protoespaciador. Los adaptadores se ligaron al producto escindido y se secuenciaron.

5 Figura 34C: Logos de secuencia del ensayo de descubrimiento de PAM *in vitro* que demuestran un PAM N₄GATT para Nme1Cas9, como se muestra previamente en las células.

10 Figura 34D: Logos de secuencia que muestran Nme1Cas9 con su PID intercambiado con los de Nme2Cas9 (izquierda) y Nme3Cas9 (derecha) reconocen una C en la posición 5. Los nucleótidos restantes se determinaron con menor confianza debido a la eficiencia de escisión moderada de las quimeras de proteínas (Figura 35C).

15 Figura 34E: Logos de secuencia que ilustra que Nme2Cas9 de longitud completa reconoce un PAM N₄CC basado en el ensayo de descubrimiento PAM con una C fija en la posición 5, y PAM nts 1-4 y 6-8 aleatorizados.

La Figura 35 muestra una caracterización de los ortólogos *Neisseria meningitidis* Cas9 con PID de evolución rápida de acuerdo con la Figura 34.

20 Figura 35A: Árbol filogenético no enraizado de ortólogos NmeCas9 que son >80% idénticos a Nme1Cas9. Surgieron tres ramas distintas, con la mayoría de las mutaciones agrupadas en el PID. grupo 1 (azul) con PID >98% de identidad a Nme1Cas9, grupo 2 (naranja) con PID ~ 52% idéntico a Nme1Cas9 y grupo 3 (verde) con PID ~ 86% idéntico a Nme1Cas9. Se marcan tres ortólogos representativos Cas9 de cada grupo (Nme1Cas9, Nme2Cas9 y Nme3Cas9).

25 Figura 35B: Esquema que muestra los loci CRISPR de las cepas que codifican los tres ortólogos Cas9 (Nme1Cas9, Nme2Cas9 y Nme3Cas9) de (A). Se muestran las identidades porcentuales de cada componente CRISPR-Cas con *N. meningitidis* 8013 (que codifica Nme1Cas9).

30 Figura 35C: Número de lecturas de ADN escindido de los ensayos *in vitro* para Nme1Cas9 intacto y para quimeras con PID de Nme1Cas9 intercambiado

con los de Nme2Cas9 y Nme3Cas9. Los recuentos de lectura reducidos indican menores eficiencias de escisión en las quimeras.

Figura 35D: Logos de secuencia del ensayo de descubrimiento de PAM *in vitro* en un PAM aleatorizado NNNNCNNN por Nme1Cas9 con su PID intercambiado por los de Nme2Cas9 (izquierda) o Nme3Cas9 (derecha).

La Figura 36 muestra que el Nme2Cas9 utiliza un espaciador de nucleótidos 22-24 para reconocer y editar sitios adyacentes a un PAM N₄CC. Todos los experimentos se realizaron por triplicado, y las barras de error representan el error estándar de la media (s.e.m.).

Figura 36A: Esquema que muestra el flujo de trabajo de la transfección transitoria en células HEK293T TLR2.0. Los plásmidos Nme2Cas9 y sgARN se transfectaron y las células mCherry+ se detectaron 72 horas después de la transfección.

Figura 36B: Uso de Nme2Cas9 para seleccionar una matriz de PAM en TLR2.0. Todos los sitios con PAM N₄CC se seleccionaron con diversos grados de eficiencia, mientras que no se observó direccionamiento Nme2Cas9 en un PAM N₄GATT o en ausencia de sgARN. SpyCas9 (dirigido a NGG) y Nme1Cas9 (dirigido a N₄GATT) se utilizaron como controles positivos.

Figura 36C: El efecto de la longitud del espaciador en la eficiencia de la edición de Nme2Cas9. Un sgARN dirigido a un sitio TLR2.0 (con un PAM N₄CCA) con distancias espaciadoras que varían de 24 a 20 nts (incluido un G de extremo 5' terminales), que muestra la mayor eficiencia de edición con 22-24 espaciadores de nucleótidos.

Figura 36D: Nickasas Nme2Cas9 (nickasa HNH = Nme2Cas9^{D16A}; RuvC nickasa = Nme2Cas9^{H588A}) se puede usar en tándem para generar indeles en TLR2.0. Los objetivos con sitios de escisión con 32 pares de bases y 64 pares de bases separados se seleccionaron mediante el uso de cualquiera de las nickasas para generar indeles. La nickasa HNH muestra una edición eficiente, particularmente cuando los sitios de escisión se cerraron (32 pb). Nme2Cas9 de tipo silvestre se utilizó como control. El verde es GFP (HDR) y el rojo es mCherry (NHEJ).

La Figura 37 presenta datos de ejemplo con respecto a PAM, espaciador y elementos de semilla para el direccionamiento de Nme2Cas9 en células de mamífero, de acuerdo con la Figura 36. Todos los experimentos se realizaron por triplicado y las barras de error representan el s.e.m.

Figura 37A: Nme2Cas9 dirigido a sitios N₄CD en TLR2.0. Se examinaron cuatro sitios para cada nucleótido no C en la posición probada (N₄CA, N₄CT y N₄CG), y se usó un sitio N₄CC como control positivo.

Figura 37B: Nme2Cas9 dirigido a sitios N₄DC en TLR2.0 [similar a (A)].

Figura 37C: Truncamientos guía en otro sitio TLR2.0, que revela requisitos de longitud similares a los observados en la Figura 36C.

Figura 37D: La eficiencia del direccionamiento de Nme2Cas9 es diferencialmente sensible a los desajustes de un solo nucleótido en la secuencia de semillas. Los datos muestran los efectos de generar desajustes de un solo nucleótido en el sgARN a lo largo del espaciador de 23 nt en un sitio objetivo TLR.

La Figura 38 presenta datos de ejemplo que muestran la eficiencia de la edición del genoma Nme2Cas9 en loci genómicos en células de mamífero a través de múltiples métodos de entrega. Todos los resultados representan 3 réplicas biológicas independientes, y las barras de error representan el s.e.m.

Figura 38A: La edición del genoma Nme2Cas9 que usa las transfecciones transitorias con sgARNs dirigidos a loci a lo largo del genoma humano en células HEK293T. Se seleccionaron 14 sitios con base en el cribado inicial de 38 sitios para demostrar el rango de indeles (según lo detectado por TIDE) en diferentes loci inducidos por Nme2Cas9. Se usó un sitio objetivo Nme1Cas9 (con un PAM N₄GATT) como control negativo.

Figura 38B: Panel izquierdo: Transfección transitoria de un plásmido todo en uno (Nme2Cas9 + sgARN) dirigido al loci *Pcsk9* y *Rosa26* en células de ratón Hepa1-6, según lo detectado por TIDE. Panel derecho: Electroporación de plásmidos de sgARN en células K562 que expresan Nme2Cas9 de forma estable a partir de un lentivector da como resultado una formación indel eficiente en los loci previstos.

Figura 38C: Nme2Cas9 puede electroporarse como un complejo RNP para una edición genómica eficiente. 40 picomoles Cas9 junto con 50 picomoles de sgARNs transcritos *in vitro* dirigidos a tres loci diferentes se electroporaron en células HEK293T. Los indeles se midieron al usar TIDE después de 72 h.

La Figura 39 presenta datos de ejemplo que muestran dependencia de la dosis y eliminaciones de bloque por Nme2Cas9, de acuerdo con la Figura 38.

Figura 39A: El aumento de la dosis de plásmido Nme2Cas9 electroporado (500 ng frente a 200 ng en la Fig. 3A) mejora la eficiencia de edición en dos sitios (TS16 y TS6).

Figura 39B: Nme2Cas9 se puede usar para crear eliminaciones de bloque. Dos objetivos TLR2.0 con sitios de escisión separados de 32 pb se seleccionaron simultáneamente con Nme2Cas9. La mayoría de las lesiones creadas fueron eliminaciones de exactamente 32 pb (verde).

La Figura 40 presenta datos de ejemplo que muestran que las proteínas Anti-CRISPR de Tipo II-C pueden usarse para inhibir la actividad de la edición del gen Nme2Cas9 (*ej.*, como un interruptor de inactivación) *in vitro* e *in vivo*. Todos los experimentos se realizaron por triplicado y las barras de error representan el s.e.m.

Figura 40A: Ensayo de escisión *in vitro* de Nme1Cas9 y Nme2Cas9 en presencia de cinco proteínas anti-CRISPR previamente caracterizadas (relación 10:1 de Acr: Cas9). Parte superior: Nme1Cas9 escinde eficazmente un fragmento que contiene un protoespaciador con un PAM N₄GATT en ausencia de un Acr o en presencia de un Acr de control (AcrE2). Todos los demás Acrs previamente caracterizados inhibieron Nme1Cas9, como se esperaba. Parte inferior: Nme2Cas9 escinde eficientemente un objetivo que contiene un protoespaciador con un PAM N₄CC en presencia de AcrE2 y AcrIIC5_{S_{mu}}, lo que sugiere que AcrIIC5_{S_{mu}} no puede inhibir Nme2Cas9 en una relación molar de 10:1.

Figura 40B: Edición del genoma en presencia de las cinco proteínas anti-CRISPR descritas anteriormente. Los plásmidos que expresan Nme2Cas9, sgARN y cada Acr respectivo (200 ng Cas9, 100 ng sgARN, 200 ng Acr)

se cotransfectaron en células HEK293T, y la edición del genoma se midió al usar TIDE 72 horas después de la transfección. A excepción de AcrE2 y AcrIIC5_{Smu}, todos los demás Acrs inhibieron la edición del genoma, aunque a diferentes eficiencias.

5 Figura 40C: La inhibición de Acr de Nme2Cas9 depende de la dosis con distintas potencias aparentes. AcrIIC1_{Nme} y AcrIIC4_{Hpa} inhiben Nme2Cas9 completamente en relaciones 2:1 y 1:1 de plásmidos cotransfectados, respectivamente.

10 La Figura 41 presenta datos de ejemplo que muestran que un intercambio de PID Nme2Cas9 hace que Nme1Cas9 sea insensible a la inhibición de AcrIIC5_{Smu}, de acuerdo con la Figura 40. La escisión *in vitro* por la quimera Nme1Cas9-Nme2Cas9PID se realizó en presencia de proteínas Acr previamente caracterizadas (Cas9-sgARN 10 uM + Acr 100 uM).

La Figura 42 presenta datos de ejemplo que muestran que Nme2Cas9 no tiene objetivos fuera de lugar detectables en células de mamífero.

15 Figura 42A: Esquema que muestra los sitios duales (DS) a los que pueden dirigir tanto SpyCas9 como Nme2Cas9 en virtud de sus PAM no superpuestos. El PAM Nme2Cas9 (naranja) y el PAM SpyCas9 (azul) están resaltados.

20 Figura 42B: Nme2Cas9 y SpyCas9 inducen indeles en sitios duales. Seis sitios duales en *VEGFA* con secuencias GN₃GN₁₉NGGNCC se seleccionaron para comparaciones directas entre los dos ortólogos. Los plásmidos que expresan cada Cas9 (con el mismo promotor y NLS) se transfectaron junto con la guía relacionada de cada ortólogo en las células HEK293T. Las tasas de indel se determinaron por TIDE 72 horas después de la transfección. La edición de Nme2Cas9 fue detectable en los seis sitios y fue más eficiente que SpyCas9 en dos sitios (DS2 y 6). SpyCas9 editó cuatro de seis sitios (DS1, 2, 4 y 6), con dos sitios que muestran tasas de edición significativamente más altas que Nme2Cas9 (DS1 y 4). DS2, 4 y 6 se seleccionaron para el análisis GUIDE-Seq ya que Nme2Cas9 era
25 igualmente eficiente, menos eficiente y más eficiente que SpyCas9 en
30 estos sitios, respectivamente.

Figura 42C: Nme2Cas9 tiene un perfil con objetivos fuera de lugar limpio en las células humanas. Se muestran los números de sitios con objetivos fuera de lugar detectados por GUIDE-Seq para cada nucleasa en sitios objetivo individuales. Los números con objetivos fuera de lugar de SpyCas9 se muestran en negro. Además de los sitios duales, TS6 (debido a su alta eficiencia y potencial para objetivos con objetivos fuera de lugar) y dos sitios de ratón (para probar la precisión en otro tipo de célula) también mostraron cero o un sitio con objetivos fuera de lugar por guía.

Figura 42D: La secuenciación profunda dirigida confirma la alta precisión de Nme2Cas9 indicada por GUIDE-seq. Los loci con objetivos fuera de lugar superiores detectados por GUIDE-seq se amplificaron y secuenciaron en profundidad. SpyCas9 mostró objetivos fuera de lugar en la mayoría de los loci, mientras que para Nme2Cas9, solo uno (el sitio *Rosa26*) mostró edición en el locus con objetivos fuera de lugar a niveles relativamente bajos (~40% con objetivos en lugar frente a ~ 1% con objetivos fuera de lugar. Tenga en cuenta la escala logarítmica en el eje y.

Figura 42E: Las eficiencias de Nme2Cas9 y SpyCas9 varían según el lugar y el sitio objetivo. Los sitios en todo el genoma (con secuencias GN₃GN₁₉NGGNCC) se seleccionaron para comparaciones directas de edición por los dos ortólogos. Los plásmidos que expresan cada Cas9 (con el mismo promotor, enlazadores, etiquetas y NLS) y su guía relacionada se transfectaron en células HEK293T. La eficacia de indel se determinó por TIDE 72 horas después de la transfección. El diagrama de caja indica las eficiencias de la edición en veintiocho (28) sitios duales por Nme2Cas9 y SpyCas9 (izquierda). Los sitios que no mostraron edición fueron excluidos del análisis. Las eficiencias relativas de Nme2Cas9 y SpyCas9 muestran que Nme2Cas9 es menos eficiente que SpyCas9 (derecha), en promedio. Las eficiencias de edición de ambos ortólogos Cas9 en los veintiocho (28) sitios se incluyeron en el análisis de eficiencias relativas en el panel derecho.

La Figura 42F presenta secuencias de ácidos nucleicos para el sitio con objetivos fuera de lugar validado de la guía *Rosa26*, que muestra la región PAM (subrayada), el dinucleótido CC PAM consenso (en negrita) y tres desajustes en la porción distal PAM del espaciador (rojo).

La Figura 43 presenta datos de ejemplo que muestran la ortogonalidad y la precisión relativa de Nme2Cas9 y SpyCas9 en sitios de doble objetivo, de acuerdo con la Figura 42.

Figura 43A: Las guías Nme2Cas9 y SpyCas9 son ortogonales. Los resultados de TIDE muestran las frecuencias de indeles creadas por ambas nucleasas dirigidas a DS12 con sus sgARNs afines o con los sgARNs del otro ortólogo.

Figura 43B: Nme2Cas9 y SpyCas9 exhiben eficiencias de la edición con objetivos en lugar comparables durante GUIDE-seq. Las barras indican recuentos de lectura con objetivos en lugar de GUIDE-Seq en los tres sitios duales dirigidos por cada ortólogo. Las barras naranjas representan Nme2Cas9 y las barras negras representan SpyCas9.

Figura 43C: Lecturas de SpyCas9 con objetivos en lugar vs. con objetivos fuera de lugar para cada sitio. Las barras naranjas representan las lecturas con objetivos en lugar, mientras que las barras negras representan los objetivos fuera de lugar.

Figura 43D: Lecturas con objetivos en lugar vs. objetivos fuera de lugar de Nme2Cas9 para cada sitio.

Figura 43E: Gráficos de barras que muestran TIDE en los sitios con objetivos fuera de lugar basados en CRISPRseek, sin detectar indeles en los loci con objetivos fuera de lugar.

La Figura 44 presenta datos de ejemplo que muestran la edición del genoma Nme2Cas9 *in vivo* a través de la entrega de AAV todo en uno.

Figura 44A: Flujo de trabajo para la entrega de AAV8.Nme2Cas9 + sgARN para reducir los niveles de colesterol en ratones al dirigirse a *Pcsk9*. Parte superior: esquema del vector AAV todo en uno que expresa Nme2Cas9 y el sgARN. Parte inferior: Cronología para las inyecciones de vena de la

cola de AAV8.Nme2Cas9 +sgARN, seguidas de mediciones de colesterol en el día 14 e análisis de indel, histología y colesterol en el día 28.

Figura 44B: Análisis de secuenciación profunda para medir indeles en el ADN extraído de hígados de ratones inyectados con AAV8.Nme2Cas9 + sgARN dirigido a locis *Pcsk9* y *Rosa26* (control).

Figura 44C: Niveles reducidos de colesterol en suero en ratones inyectados con la guía dirigida a *Pcsk9* en comparación con los controles dirigidos a *Rosa26*. Los valores P se calculan mediante una prueba T no emparejada.

Figura 44D: Tinción H&E de hígados de ratones inyectados con vectores AAV8.Nme2Cas9 + sgRosa26 (izquierda) o AAV8.Nme2Cas9 + sgPcsk9 (derecha). Barra de escala, 25 um.

La figura 45 presenta una modalidad de la estructura principal de AAV minimizada y datos comparativos TLR 2.0 de ejemplo con la estructura principal de AAV de tamaño convencional.

La Figura 46 presenta una comparación de las estructuras Nme2Cas9 de sgARN 11 truncado con sgARN 12 truncado.

La Figura 47 ilustra una modalidad de un AAV todo en uno minimizado con una señal corta de poliA.

La Figura 48 ilustra dos modalidades de una estructura principal de AAV todo en uno minimizada. sgARNs dobles en tándem (parte superior). Plantilla de donante para la reparación directa por homología (parte inferior).

La Figura 49 presenta una validación de un constructo AAV-sgARN-hNme1Cas9 todo en uno.

Figura 49A: Representación esquemática de un único vector rAAV que expresa el codón humano optimizado Nme1Cas9 y su sgARN. La estructura principal está flanqueada por repeticiones terminales invertidas (ITR) de AAV. La señal de poli(a) proviene de la beta-globina de conejo (BGH).

Figura 49B: Diagrama esquemático de los genes de ratón *Pcsk9* (parte superior) y *Rosa26* (parte inferior). Las barras rojas representan exones. Las vistas ampliadas muestran la secuencia protoespaciador (roja) mientras que la secuencia

PAM Nme1Cas9 se resalta en verde. Los sitios de ubicación de corte de doble cadena se indican (con puntas de flecha negras).

Figura 49C: Histograma apilado que muestra una distribución porcentual representativa de inserciones-deleciones (indeles) obtenidas por TIDE después de las transfecciones de plásmidos AAV-sgARN-hNme1Cas9 en células Hepa1–6 dirigidas a los genes *Pcsk9* (sgPcsk9) y *Rosa26* (sgRosa26). Los datos se presentan como valores medios $\pm$ SD de tres réplicas biológicas.

Figura 49D: Histograma apilado que muestra una distribución porcentual representativa de indeles en *Pcsk9* en el hígado de ratones C57Bl/6 obtenidos por TIDE después de la inyección hidrodinámica de plásmido AAV-sgARN-hNme1Cas9.

La Figura 50 presenta datos de ejemplo que muestran que muchos PAM N₄GN₃ están inactivos y no revelaron sitios fuera del objetivo con menos de cuatro desajustes en el genoma del ratón.

La Figura 51 presenta datos de ejemplo que muestran que la inactivación de Hpd mediada por Nme1Cas9 rescata el fenotipo letal en ratones con tirosinemia hereditaria tipo I.

Figura 51A: Diagrama esquemático del gen del ratón Hpd. Las barras rojas representan exones. Las vistas ampliadas muestran las secuencias protoespaciadoras (rojo) que se dirigen al exón 8 (sgHpd1) y al exón 11 (sgHpd2). Las secuencias PAM Nme1Cas9 están en verde y se indican ubicaciones de corte de doble cadena (con puntas de flecha negras).

Figura 51B: Diseño experimental. Se inyectan tres grupos de ratones de tirosinemia hereditaria tipo I Fah^{-/-} con PBS o plásmidos AAV-sgARN-hNme1Cas9 todo en uno sgHpd1 o sgHpd2.

Figura 51C: El peso de los ratones inyectados hidrodinámicamente con PBS (verde), plásmido AAV-sgARN-hNme1Cas9 sgHpd1 dirigido a *Hpd* exón 8 (rojo) o sgHpd2 dirigido a *Hpd* exón 11 (azul) se monitoreó después de la retirada de NTBC. Las barras de error representan tres ratones para los grupos PBS y sgHpd1 y dos ratones para el grupo sgHpd2. Los datos se presentan como media $\pm$ SD.

Figura 51D: Histograma apilado que muestra una distribución porcentual representativa de indeles en *Hpd* en el hígado de ratones *Fah*^{-/-} obtenidos por TIDE después de la inyección hidrodinámica de PBS o plásmidos sgHpd1 y sgHpd2. Los hígados se recolectaron al final del retiro de NTBC (día 43).

5 La Figura 52 presenta datos de ejemplo que muestran las eficiencias indele promedio de las guías presentadas en la Figura 51.

La Figura 53 presenta fotomicrografías histológicas de ejemplo que muestran que el daño hepático es sustancialmente menos grave en los ratones tratados con sgHpd1 y sgHpd2 en comparación con los ratones *Fah*^{mut/mut} inyectados con PBS, como lo indica el menor número de hepatocitos multinucleados en comparación con los ratones inyectados con PBS.

La Figura 54 presenta la entrega de AAV de Nme1Cas9 para la edición del genoma *in vivo*.

Figura 54A: Esquema experimental de las inyecciones de la vena de la cola del vector AAV8-sgARN-hNme1Cas9 para dirigirse a *Pcsk9* (sgPcsk9) y *aRosa26* (sgRosa26) en ratones C57Bl/6. Los ratones se sacrificaron a los 4 (n = 1) o 50 días (n = 5) después de la inyección y se recolectaron los tejidos del hígado. Se recogieron sueros de sangre en los días 0, 25 y 50 después de la inyección para medir el nivel de colesterol.

Figura 54B: Niveles de colesterol en suero, los valores p se calculan mediante la prueba t no emparejada.

Figura 54C: Histograma apilado que muestra una distribución porcentual representativa de indeles en *Pcsk9* o *Rosa26* en hígados de ratones, medido por análisis de secuenciación profunda dirigida. Los datos se presentan como media ± SD de cinco ratones por cohorte.

Figura 54D: Una transferencia de Western anti-PCSK9 representativa que usa la proteína total recogida del homogeneizados de hígado de ratón en el día 50. Se incluyó un total de 2 ng de PCSK9 de ratón recombinante (r-PCSK9) como estándar de movilidad. El asterisco indica una proteína de reacción cruzada que es más grande que la proteína recombinante de control.

La Figura 55 presenta datos de ejemplos que muestran que los ratones inyectados con AAV8-sgARN-hNme1Cas9 generan anticuerpos anti-Nme1Cas9.

La Figura 56 presenta datos de ejemplo que muestran especificidades amplias del genoma GUIDE-seq de Nme1Cas9. Los datos se presentan como media $\pm$ SD.

Figura 56A: Número de lecturas GUIDE-seq para los sitios con objetivos en lugar (OnT) y objetivos fuera de lugar (OT).

5 Figura 56B: Secuenciación profunda dirigida para medir las tasas de lesiones en cada uno de los sitios OT en las células Hepa1–6. Se resalta los desajustes de cada sitio OT con los protoespaciadores de OnT (azul). Los datos se presentan como media $\pm$ SD de tres réplicas biológicas.

10 Figura 56C: Secuenciación profunda dirigida para medir las tasas de lesiones en cada uno de los sitios OT mediante el uso de ADN genómico obtenido de ratones inyectados con AAV8-sgARN-hNme1Cas9 todo en uno sgPcsk9 y sgRosa26, sacrificados en el día 14 (D14) o el día 50 (D50) después de la inyección.

La Figura 57 presenta datos de ejemplo para la edición del gen de la *tirosinasa (Tyr)* *ex vivo* por Nme2Cas9 en cigotos de ratón, en relación con la Figura 58.

15 Figura 57A: Se probaron dos sitios en el gen *Tyr*, cada uno con PAM N₄CC, para su edición en células Hepa1-6. La guía sgTyr2 exhibió una mayor eficiencia de edición y se seleccionó para más pruebas.

20 Figura 57B: Siete ratones sobrevivieron al desarrollo postnatal, y cada uno exhibió fenotipos de color de pelaje, así como la edición con objetivos en lugar, según lo analizado por TIDE.

Figura 57C: Los espectros de indel del ADN de la cola de cada ratón de la Figura 57B, así como un ratón C57BL/6NJ sin editar, como se indica mediante el análisis TIDE. Se indican las eficiencias de inserciones (positivas) y deleciones (negativas) de varios tamaños.

25 La Figura 58 presenta datos de ejemplo de la edición del genoma de Nme2Cas9 *ex vivo* al usar una entrega de AAV todo en uno.

30 Figura 58A: Flujo de trabajo para la edición Nme2Cas9 de AAV único *ex vivo* para generar ratones C57BL/6NJ albinos al dirigirse al gen *Tyr*. Los cigotos se cultivan en KSOM que contiene AAV6.Nme2Cas9: sgTyr durante 5-6 horas, se enjuagan en M2 y se cultivan durante un día antes de ser transferidos al oviducto de receptores pseudo-preñadas.

Figura 58B: Albino (izquierda) y chinchilla o ratones abigarrados (medio) generados por 3 x 10⁹ GC, y chinchilla o ratones abigarrados (derecha) generados por 3 x 10⁸ GC de cigotos con AAV6.Nme2Cas9: sgTyr.

Figura 58C: Resumen de los experimentos de edición de *Tyr ex vivo* de AAV único Nme2Cas9.sgTyr a dos dosis de AAV.

La Figura 59 muestra una alineación de secuencias de nucleótidos Nme1Cas9 y Nme2Cas9. Leyenda: Diferencias no PID aa (sombreado turquesa); diferencias PID aa (sombreado amarillo); residuos del sitio activo (letras rojas).

La Figura 60 muestra una alineación de las secuencias de nucleótidos Nme1Cas9 y Nme3Cas9. Leyenda: Diferencias no PID aa (sombreado turquesa); diferencias PID aa (sombreado amarillo); residuos del sitio activo (letras rojas).

La Figura 61 muestra una modalidad de una secuencia de aminoácidos Nme2Cas9. Leyenda: SV40 NLS (sombreado amarillo); 3X-HA-Tag (sombreado verde); NLS tipo cMyc (sombreado turquesa); Linker (sombreado morado).

La Figura 62 muestra una modalidad de una secuencia de aminoácidos Nme2Cas9. Leyenda: SV40 NLS (sombreado amarillo); 3X-HA-Tag (sombreado verde); NLS de tipo nucleoplasmina (sombreado rojo); NLS c-myc (sombreado turquesa); Linker (sombreado morado).

La Figura 63 muestra una modalidad de una secuencia de aminoácidos recombinante Nme2Cas9 (rNme2Cas9). Leyenda: SV40 NLS (sombreado amarillo); NLS de tipo nucleoplasmina (sombreado rojo); Linker (sombreado morado).

La Figura 64 muestra una modalidad de una secuencia de nucleótidos del plásmido AAV-sgARN-hNmeCas9 todo en uno. Leyenda: andamio sgARN (letras marrones); Secuencia GUIDE (letras negras); *Promotor U6* (letras azules); *Promotor U1a* (letras moradas); NLS NLS (letras verdes); hNmeCas9 (letras rojas); NLS 3X-HA y NLS BGH-pA (letras verdes/negras alternas).

Descripción Detallada de la Invención

La presente invención está relacionada con composiciones y métodos para terapia génica. Varios enfoques descritos en este documento utilizan el sistema *Neisseria meningitidis* Cas9 que proporciona una plataforma de edición de genes CRISPR hiperprecisa. Además, la invención

incorpora mejoras de este sistema Cas9: por ejemplo, el truncamiento de las secuencias de ARN de guía única, y el empaquetamiento de Nme1Cas9 o Nme2Cas9 con su ARN guía en un vector viral adenoasociado que es compatible para la administración *in vivo*. Además, se han identificado ortólogos Cas9 de Tipo II-C que se dirigen a secuencias limitadas de motivo adyacente al protoespaciador de entre uno-cuatro nucleótidos requeridos.

I. Precisión de la edición de genes *Neisseria meningitidis* Cas9 (Nme1Cas9)/CRISPR

Anteriormente, se reportó una versión hiperprecisa de los sistemas CRISPR-Cas9 tipo II-C llamada *Neisseria meningitidis* Cas9 (Nme1Cas9). Además de ser hiperprecisa, Nme1Cas9 también es más pequeña que la Cas9 de *Streptococcus pyogenes* (SpyCas9) ampliamente utilizada, lo que permite que Nme1Cas9 se entregue más fácilmente a través de métodos basados en ARN viral y ARN mensajero (ARNm). La edición genómica con Nme1Cas9 generalmente se ha logrado mediante las transfecciones de plásmidos. Zhang, “*Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis*” *Mol Cell* 50: 488-503 (2013); Hou, “*Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis*” *Proc Natl Acad Sci EE.UU.* 110: 15644-15649 (2013); Esvelt, “*Orthogonal Cas9 proteins for RNA-guided gene regulation and editing*” *Nature Methods* 10:1116-1121 (2013); Zhang, “*DNase H activity of Neisseria meningitidis Cas9*” *Mol Cell* 60:242-255 (2015); Lee, “*The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells*” *Molecular Therapy* 24: 645-654 (2016); Pawluk, “*Naturally occurring off-switches for CRISPR-Cas9*” *Cell* 167: 1829-1838 (2016); y Amrani, “*Nme1Cas9 is an intrinsically high-fidelity genome editing platform*” [biorxiv.org/content/early/2017/08/04/172650](https://www.biorxiv.org/content/early/2017/08/04/172650) (2017).

Sin embargo, no se ha explorado exhaustivamente la administración de Nme1Cas9 viral, ARN y ribonucleoproteínas (RNP). La entrega basada en ARN y RNP de ortólogos Cas9 para la ingeniería genómica tiene varias ventajas sobre otros métodos de entrega. No solo resulta en una edición más rápida, ya que evita los problemas de expresión relacionados con la entrega basada en ADN de Cas9 y su sgARN, sino que también reduce los efectos con objetivos fuera de lugar asociados con la edición basada en Cas9. La reducción de la actividad con objetivos fuera de lugar resulta de un control más fino de las concentraciones de ARN y RNP de Cas9, y de la degradación relativamente rápida de ARN y RNP de Cas9 en las células. Se ha demostrado que

la presencia prolongada de Cas9 activo dentro de la célula está asociada con mayores efectos con objetivos fuera de lugar. Dado que los ARN y RNP de Cas9 se degradan más rápidamente dentro de las células, Cas9 entregado como ARN o RNP no persiste durante largos períodos de tiempo y, en consecuencia, tiene efectos reducidos con objetivos fuera de lugar.

5 El sgARN Nme1Cas9 de longitud completa de 145 nt usado convencionalmente incluye un crARN de 48 nucleótidos (nt), un enlazador de 4 nt y un tracrARN de 93 nt. La región de crARN del sgARN se compone de una primera secuencia espaciadora de 24 nt, y una segunda secuencia de repetición de 24 nt que se empareja con una región 5' de anti-repetición tracrARN de 24 nt, formando así una repetición: región anti-repetición. La región tracrARN restante de 69
10 nt incluye la región del vástago 1 y la región del vástago 2. Figura 1.

Este sgARN Nme1Cas9 de longitud completa se ha utilizado con éxito para la edición genómica mediante el uso de métodos basados en plásmidos. Además, el sgARN Nme1Cas9 transcrito *in vitro* puede formar un complejo con Nme1Cas9 purificado y usarse para la edición genómica en células humanas. Si bien la edición de genomas de las células humanas ha tenido
15 éxito con sgARN transcritos *in vitro*, la eficiencia de la edición de una RNP Nme1Cas9 se reduce en líneas celulares humanas más difíciles de transfectar, tales como PLB985.

Anteriormente se ha demostrado que la eficiencia de la edición de RNP Cas9 es proporcional a la estabilidad química de sus sgARN. Aunque no es necesario comprender el mecanismo de una invención, se cree que se emplean varios mecanismos celulares para degradar
20 rápidamente los ARN. Por esta razón, los sgARN Cas9 se modifican habitualmente por medios químicos. Algunas de las modificaciones químicas que confieren una mayor estabilidad al sgARN incluyen, pero no se limitan a, enlaces de ribosa 2'-O-metilación y/o fosforotioato. Si bien los ARN modificados químicamente son opciones para mejorar la edición genómica por parte de las RNP Cas9, su efectividad está limitada por el hecho de que la síntesis química de
25 ARN se vuelve cada vez más difícil y costosa a medida que aumenta la longitud del ARN. A 145 nt, la síntesis de sgARN Nme1Cas9 está fuera del alcance de las aplicaciones de la edición genómica de rutina que emplean sgARN sintetizados químicamente.

II. Secuencias de sgARN Nme1Cas9 truncadas

30 Debido a la limitación identificada anteriormente de que un sgARN Nme1Cas9 de 145 nt de longitud completa es demasiado grande para la síntesis química de rutina de sgARNs para la

edición genómica, una modalidad de la presente invención contempla un sgARN Nme1Cas9 truncado. Aunque no es necesario comprender el mecanismo de una invención, se cree que un Nme1Cas-sgARN truncado no compromete la función de una RNP Nme1Cas9. Además, los sgARN para Nme1Cas9 y Nme2Cas9 son idénticos e intercambiables (Figura 35B), por lo que los truncamientos de sgARN son igualmente aplicables a Nme1Cas9 y Nme2Cas9. Las secuencias de ejemplo de sgARN truncados y sitios objetivo asociados se describen a continuación, donde los nucleótidos (nt) de sgARN variables en las regiones guía se dan como residuos "N". En las secuencias objetivo, los 24 nts reconocidos por la región guía de sgARN están subrayados, y la región del motivo adyacente al protoespaciador (PAM) está en negrita.

Tabla 1.

Tabla 1: Secuencias de sgARN truncadas de ejemplo y objetivos genómicos asociados

Descripción	Secuencia
sgARN tipo silvestre (wt)	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCCUUUCUCAUUUCGGAA ACGAAAUGAGAACCGUUGCUACAAUAAGGCCGUCUGAAAAGAUGUGCCGCA ACGCUCUGCCCCUUAAGCUUCUGCUUUAAGGGGCAUCGUUUA
sgARN #1	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAGCUUCUG CUUUAAGGGGCAUCGUUUA
sgARN #2	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAGGGG CAUCGUUUA
sgARN #3	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAGGGG CAU
sgARN #4	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCUUCUGCAUCGUU
sgARN #5	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAAU AAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCUUCUGCAUCGUUUA
sgARN #6	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAAUAA GGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCUUCUGGGGCAUCGUU
sgARN #7	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCUUUCUAGGGGCA UCGUU
sgARN #8	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCUUCUGGGGCAUC GUU
sgARN #9	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCUUCUGGGGCAUCGU U
sgARN #10	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAA UAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCUUCUGGGGCAUCGUU
sgARN #11	NNNNNNNNNNNNNNNNNNNNNNNGUUGUAGCUCUCCGAAACGUUGCUACAAU AAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCUUCUGGGGCAUCGUU
N-TS7 espaciador (24 nt)	GAGGGAGAGAGGUGAGCGGAUGAA
N-TS7 espaciador (23 nt)	GGGGAGAGAGGUGAGCGGAUGAA
N-TS27 espaciador (24 nt)	GUUCUCCAAGCCUCGGACCUCGU
N-TS27 espaciador (23 nt)	GUCUCCAAGCCUCGGACCUCGU
N-TS55 espaciador (24 nt)	GCUGGAUUACUGUGGUAGAGGG
N-TS55 espaciador (23 nt)	GUGGAUUACUGUGGUAGAGGG
N-TS7 sitio genómico objetivo	AGCTTGAGCAAAGGGAGAGAGGTGAGCGGATGAAGGGAGATTGGTGAGTAT C
N-TS27 sitio genómico objetivo	CGCTTCGCGGCTTCTCCAAGCCCTCGGACCTCGTGGGCGTCTTCTCCTGCG T
N-TS55 sitio genómico objetivo	GAATTCAGTAGCTGGATTACTGTGTGGTAGAGGGAGGTGATTAGCACCTGT G

Como se contempla en el presente documento, un sgARN Nme1Cas9 truncado no solo permitiría la síntesis a un costo razonable, sino que también facilita el uso en métodos de entrega basados en virus (*ej.*, plataformas de entrega viral adenoasociadas) donde la longitud permitida de ADN es limitada. En una modalidad, el sgARN truncado reduce el efecto de la edición de Nme1Cas9 con objetivos fuera de lugar. En una modalidad, el sgARN Nme1Cas9 truncado

comprende al menos una modificación química que aumenta la eficacia de la edición de Nme1Cas9.

Como se discutió anteriormente, el sgARN de 145 nt de longitud completa de Nme1Cas9 incluye una región guía, una repetición: región dúplex anti-repetición, una región del vástago 1 y una región del vástago 2. Figura 1. Sin embargo, debido a que la longitud del sgARN es problemática para la edición genómica de rutina, y era altamente deseable desarrollar un sgARN truncado para Nme1Cas9. Actualmente, los métodos de síntesis de ARN disponibles comercialmente requieren que el producto final de ARN no sea superior a ~ 100 nt.

En una modalidad, la presente invención contempla un sgARN Nme1Cas9 que comprende una repetición truncada: dúplex anti-repetición. En una modalidad, la presente invención contempla un sgARN Nme1Cas9 que comprende un del vástago 2 truncado. Figura 2. Además, se ha demostrado previamente que una región de crARN guía variable 5' (ej., región espaciadora) de Nme1Cas9 también puede ser truncada por unos pocos nucleótidos sin pérdida de función. Amrani, "*Nme1Cas9 is an intrinsically high-fidelity genome editing platform*" *biorxiv.org/content/early/2017/08/04/172650* (2017); y Lee, "*The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells*" *Molecular Therapy* 24:645-654 (2016).

En una modalidad, la presente invención contempla un sgARN Nme1Cas9 truncado de 100 nt. Figura 3, Constructo #11. Este constructo de sgARN truncado ng1Cas9 de 100 nt se probó en tres sitios genómicos humanos diferentes mediante transfecciones transitorias en células HEK293T, y en los tres sitios admiten la función Nme1Cas9 en el mismo nivel que, si no mejor, que el sgARN Nme1Cas9 de longitud completa. Figura 3, panel inferior. Además, sgARN 11 y sgARN 13 también se probaron en varios sitios genómicos objetivo al utilizar la entrega de RNP y la eficiencia de edición fue similar o superior a la de sgARN de tipo silvestre. Figura 4. La versión sintética del constructo #11 también se probó en células PLB985, dando como resultado una mayor eficacia de la edición en relación con el sgARN de tipo silvestre transcrito *in vitro*. Figura 5.

III. Plataformas de entrega CRISPR de adenovirus asociado

En comparación con las nucleasas efectoras tipo activador de transcripción (TALEN) y las nucleasas con dedos de zinc (ZFN), las Cas9 se distinguen por su flexibilidad y versatilidad. Komor, "*CRISPR-based technologies for the manipulation of eukaryotic genomes*" *Cell*

2017;168:20–36 . Dichas características los hacen ideales para impulsar el campo de la ingeniería genómica. En los últimos años, CRISPR-Cas9 se ha utilizado para mejorar los productos en la agricultura, la alimentación y la industria, además de las aplicaciones prometedoras en la terapia génica y la medicina personalizada. Barrangou, “*Applications of CRISPR technologies in research and beyond*” *Nat Biotechnol.* 2016;34:933-41. A pesar de la diversidad de sistemas CRISPR de clase 2 que se han descrito, solo unos pocos de ellos se han desarrollado y validado para la edición genómica *in vivo*. Como se muestra en el presente documento, NmeCas9 es una Cas9 compacta y de alta fidelidad que puede considerarse para futuras aplicaciones de la edición genómica *in vivo* al usar rAAV todo en uno. El PAM único de NmeCas9 permite la edición en objetivos adicionales que son inaccesibles para los otros dos ortólogos compactos y validados por rAAV todo en uno (SauCas9 y CjeCas9).

La edición genómica utilizando un sistema CRISPR bacteriano ha abierto una nueva vía para la terapia génica humana. El complejo CRISPR, que recibe el nombre de repeticiones palindrómicas cortas agrupadas y regularmente espaciadas que capturan fragmentos de ácidos nucleicos invasores en bacterias, comprende un ARN guía (*ej.*, sgARN) que dirige una nucleasa Cas9 (proteína asociada a CRISPR 9) para escindir el ADN complementario de doble cadena. La reparación no homóloga de un corte de ADN inducida por Cas9 conduce a pequeñas inserciones o deleciones (indeles) que inactivan los genes objetivo, pero los cortes también pueden repararse mediante plantillas de ADN homólogas que resultan en el reemplazo de genes.

Nelson, “*In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy*” *Science* 351: 403-407 (2016); y Ran, “*In vivo genome editing using Staphylococcus aureus Cas9*” *Nature* 520: 186-191 (2015); e Yin, “*Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype*” *Nature Biotechnology* 32: 551-553 (2014).

La Cas9 de *Streptococcus pyogenes* (Spy) de Tipo II-A actual y ampliamente utilizada como una herramienta flexible para la edición genómica demuestra varias desventajas: i) entrega ineficiente; ii) escisión con objetivos fuera de lugar; y iii) actividad no regulada. Estas desventajas limitan estrictamente la SpyCas9 como una herramienta potencial para la terapia génica. Como se describe en la presente, un complejo Nme1Cas9 o Nme2Cas9 altamente exacto y preciso puede superar estas limitaciones de SpyCas9.

Se ha demostrado que Nme1Cas9 y Nme2Cas9 son una plataforma eficaz para la edición genómica en células de mamíferos y, como una proteína más pequeña que SpyCas9, es más fácil diseñar vectores virales para la entrega *in vivo*. Además, Nme1Cas9 y Nme2Cas9 tienen una edición con objetivos fuera de lugar significativamente menor que SpyCas9 y se han identificado proteínas anti-CRISPR que permiten controlar la actividad de Nme1Cas9 y Nme2Cas9. Esvelt, “*Orthogonal Cas9 proteins for RNA-guided gene regulation and editing*” *Nature Methods* 10: 1116-1121 (2013); Amrani, “*Nme1Cas9 is an intrinsically high-fidelity genome editing platform*” *biorxiv.org/content/early/2017/08/04/172650* (2017); Lee, “*The Neisseria meningitidis CRISPR-Cas9 System Enables Specific Genome Editing in Mammalian Cells*” *Molecular Therapy* 24: 645-654 (2016); Hou, “*Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis*” *Proc Natl Acad Sci EE.UU.* 110: 15644-15649 (2013); y Pawluk, “*Naturally Occurring Off-Switches for CRISPR-Cas9*” *Cell* 167: 1829-38 e9 (2016); y la Figura 21.

El virus adenoasociado (AAV) se ha demostrado como un transporte de entrega con una patogenicidad mínima en entornos preclínicos y clínicos, pero tiene una capacidad de empaquetamiento limitada. Nme1Cas9, codificada por un marco de lectura abierto de ~ 3.3kb, y sus ARN guía están dentro del límite de empaquetamiento de AAV. Nme2Cas9 tiene ventajas similares. A diferencia de SpyCas9, que requiere la entrega por vectores separados para el sgARN y Cas9, Nme1Cas9, Nme2Cas9 y su sgARN son lo suficientemente pequeños como para entregarse con un solo vector AAV.

Otros ortólogos Cas9 han sido entregados *in vivo* con éxito por AAV, como Cas9 de *Campylobacter jejuni* (CjeCas9) y *Staphylococcus aureus* (SauCas9). Kim, “*In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni*” *Nat Commun* 8: 14500 (2017); y Ran, “*In vivo genome editing using Staphylococcus aureus Cas9*” *Nature* 520: 186-191 (2015). Nme1Cas9 generalmente se asocia con un PAM GATT N₄, que es diferente al PAM CjeCas9 (*ej.*, N₄RYAC), o el PAM SauCas9 (*ej.*, NNGRRT) (R = purina (A o G), Y = pirimidina (C o T)).

Nme1Cas9 se ha entregado con éxito como un complejo de ribonucleoproteína (RNP) en células humanas. Figura 2 y Figura 3. Además, los datos presentados en el presente documento muestran que una secuencia de ácido nucleico Nme1Cas9 puede expresarse *in vivo* en ratones

para dirigirse a genes que utilizan un vector sgARN-Nme1Cas9-AAV todo en uno después de una inyección en la vena de la cola.

Los datos presentados en el presente documento demuestran un direccionamiento de un gen de la proproteína convertasa subtilisina/kexina tipo 9 (*Pcsk9*) de ratón. PCSK9 funciona como un antagonista del receptor de lipoproteínas de baja densidad (LDL) y limita la absorción de colesterol LDL. La detección de los niveles reducidos de colesterol en el suero puede proporcionar una lectura funcional directa de la edición eficiente de Nme1Cas9 al usar una plataforma Cas9 dirigida por *PCSK9*.

En una modalidad, la presente invención contempla un vector viral adenoasociado que comprende un complejo Nme1Cas9-sgARN o un complejo Nme2Cas9-sgARN. Aunque no es necesario comprender el mecanismo de una invención, se cree que un complejo AAV/Nme1Cas9-sgARN o un complejo AAV/Nme2Cas9-sgARN son compatibles con una ruta de entrega *in vivo* para proporcionar la edición de genes.

En una modalidad, la presente invención contempla un vector sgARN-Nme1Cas9-AAV que comprende una secuencia de sgARN, una secuencia promotora de ARN polimerasa III U6, una secuencia de Nme1Cas9 optimizada por un codón humano y una secuencia promotora de ARN polimerasa II U1a. Figura 6. U1a es un promotor ubicuo que permite la expresión versátil de Cas9 en varios tejidos de interés. Los genes específicos que se van a editar pueden ser dirigidos insertando una secuencia espaciadora que coincida con un gen objetivo en un casete de sgARN mediante el uso de sitios de restricción convencionales (*ej.*, Sap1). Las secuencias representativas de los diversos elementos del sgARN-Nme1Cas9-AAV se muestran mediante anotaciones coloreadas. Figuras 7 y 8.

Las eficiencias de la edición de varios sitios objetivo al usar un plásmido *Pcsk9*-sgARN-Nme1Cas9-AAV y un plásmido *Rosa26*-sgARN-Nme1Cas9-AAV se estimaron mediante un ensayo T7E1 después de la transfección transitoria en células de hepatoma Hepa1-6 de ratón. Figura 9. Las secuencias representativas del sitio objetivo dentro de un gen *Pcsk9* y un gen *Rosa26* complementario con un plásmido *Pcsk9*-sgARN-Nme1Cas9-AAV y un plásmido *Rosa26*-sgARN-Nme1Cas9-AAV se muestran mediante anotaciones coloreadas. Figura 10.

El diseño del plásmido se validó *in vivo* con ratones mediante la inyección hidrodinámica de 30 µg de plásmido sgARN-Nme1Cas9-AAV libre de endotoxina dirigido a *Pcsk9* a través de la vena de la cola. Se detectó una edición genética significativa en el hígado de ratón 10 días

después de la inyección, medida mediante el seguimiento de indeles mediante descomposición (TIDE), un método basado en secuenciación para evaluar las eficiencias de indel. Figura 11.

Las estructuras principales de plásmido dirigidas a un gen *Pcsk9* y un gen *Rosa26* se empaquetaron en serotipo AAV8 específico a hepatocitos, y se inyectó una dosis de 4×10^{10} de copias genómicas (gc) por ratón a través de la vena de la cola. Los datos preliminares muestran valores de indel de ratones sacrificados a los 14 días después de la inyección a un nivel indel significativo en los genes *Pcsk9* y *Rosa26* del hígado. Figura 12A. Los datos de secuenciación profunda también se han recopilado en el día 50 después de la inyección.

Los tres grupos de ratones se sacrificaron en el día 50 después de la inyección, y se usó ADNg de hígado para medir los valores de indel en *Pcsk9* y *Rosa26* mediante el uso de TIDE. Figura 12B. Los análisis de secuenciación profunda también se han realizado para registrar mediciones precisas de valores de indel.

La proteína PCSK9 "derribada" puede conducir a una disminución significativa de los niveles de colesterol en ratones. El nivel de colesterol en suero se midió mediante el ensayo de punto final colorimétrico Infinity™ (Thermo-Scientific) en 3 grupos de ratones inyectados con vectores dirigidos a un gen *Pcsk9*, un gen *Rosa26* y un grupo de control PBS. Los resultados sugieren que la formación indel inducida por Nme1Cas9 ha llevado a la interrupción del marco de lectura normal del gen *Pcsk9*, como lo demuestran los valores significativamente reducidos de colesterol en suero a los 25 y 50 días después de la inyección. Figura 13. El ensayo de Western blot también se realizó para medir el nivel de proteína PCSK9 en el hígado de ratones en el día 50.

Una identificación imparcial de todo el genoma de cortes de doble cadena (DSB) habilitada por un ensayo de secuenciación (ej., GUIDE-Seq®, Illumina) buscó sitios de edición con objetivos fuera de lugar después de la inyección de vectores dirigidos a un gen *Pcsk9* y un gen *Rosa26*. Los datos revelaron cuatro (4) sitios con objetivos fuera de lugar potenciales fuera del objetivo para *Pcsk9* y seis (6) sitios con objetivos fuera de lugar potenciales para *Rosa26*. Figuras 14A y 14B.

Un análisis TIDE dirigido reveló la edición genómica con objetivos en lugar en las células y en los ratones en el día 14 posterior a la inyección de vectores AAV dirigidos a un gen *Pcsk9* y un gen *Rosa26*. Figura 15. Los análisis de secuenciación profunda para la escisión con

objetivos fuera de lugar en estos sitios también se han realizado a los 50 días después de la inyección.

Un ensayo de tinción con hematoxilina y eosina no mostró signos de infiltración masiva de células inmunes en las secciones del hígado de ratones sacrificados en el día 14 posterior a la inyección de vectores dirigidos a un gen *Pcsk9* y un gen *Rosa26*. Figura 16. Se realizarán ensayos de respuesta inmunitaria específicos a los 50 días después de la inyección.

En una modalidad, la presente invención contempla un método para la edición terapéutica del genoma *in vivo* mediante la entrega de AAV todo en uno de un Nme2Cas9. Aunque no es necesario comprender el mecanismo de una invención, se cree que la compacidad, el pequeño PAM y la alta fidelidad hacen de Nme2Cas9 una herramienta ideal para la edición del genoma *in vivo* al usar AAV. Con este fin, Nme2Cas9 se clonó con su sgARN afín y sus respectivos promotores en un único vector principal AAV. Figura 44A; parte superior este AAV.sgARN.Nme2Cas9 todo en uno fue empaquetado en una cápside AAV8 selectiva de hepatocitos. Se seleccionaron dos genes: i) *Rosa26*, un locus de uso común como control negativo; y ii) la *proproteína convertasa subtilisina/kexina tipo 9* (*Pcsk9*), un importante regulador de la homeostasis del colesterol circulante. Los estudios han demostrado que inactivar la *Pcsk9* mediante el uso de Cas9 resulta en niveles reducidos de colesterol (Ran).

Dos grupos de ratones (n = 5) fueron inyectados con AAV8.sgNA.Nme2Cas9 empaquetado dirigido a *Pcsk9* o *Rosa26*. Se recolectó suero a los 0, 14 y 28 días después de la inyección del vector para medir el nivel de colesterol. Los ratones se sacrificaron a los 28 días después de la inyección y se recolectaron los tejidos del hígado. (Figura 44A, parte inferior. Un análisis de secuenciación profunda mostró un nivel significativamente alto de indeles en *Pcsk9* y *Rosa26*. Figura 44B. Estos valores de indel fueron acompañados por una reducción significativa en el nivel de colesterol en la sangre en ratones inyectados con *sgPcsk9* después de 14 y 28 días; donde los ratones inyectados con *sgRosa26* mantuvieron el nivel normal de colesterol durante todo el estudio. Figura 44C. Un análisis de H&E no mostró signos de toxicidad o daño tisular en ambos grupos después de la expresión de Nme2Cas9. Figura 44D. Estos datos validan que Nme2Cas9 es altamente funcional *in vivo*, y puede ser entregada fácilmente por la plataforma favorable de AAV todo-en-uno.

En una modalidad, la presente invención contempla un constructo AAV.hNmeCas9 minimizado. Véase, la Figura 44A. Como se discutió anteriormente, la presente invención

contempla un constructo AAV.sgARN.hNme1Cas9 todo en uno modificado, que está empaquetado en viriones AAV8 que editó con éxito los genes *Pcsk9* y *Rosa26* en el hígado de ratones.

En una modalidad, la presente invención contempla una estructura principal AAV8 que comprende un casete Nme2Cas9. Similar a Nme1Cas9, Nme2Cas9 también mostró una edición robusta en *Pcsk9* y *Rosa26* en ratones (*infra*). Los datos presentados en este documento muestran que la administración *in vivo* de AAV8-NmeCas9 a ratones se acompaña de una reducción significativa en el nivel de colesterol circulante 28 días después de la inyección del vector.

Con el fin de aumentar la utilidad de esta plataforma de AAV todo en uno, se introdujeron varios truncamientos para minimizar el tamaño de la carga para dejar espacio para funciones adicionales en la cápside de AAV, tales como los sgARN duales o el segmento de ADN donante.

Con el fin de minimizar la carga de la estructura principal de AAV todo en uno, las características adicionales (etiquetas 3x HA y secuencias 2x NLS) se eliminaron sistemáticamente sin comprometer la actividad de la nucleasa del Cas9. Nme1Cas9, que usa el sistema reportero de semáforo (TLR), muestra que este AAV.sgARN.hNme1Cas9 todo en uno minimizado (4.468 kb) es tan potente como la versión anterior más larga con 4 secuencias NLS. Véase, la Figura 45. Los sgARN truncados se construyeron para liberar más espacio mediante el uso de un nuevo sgARN12, que es similar a una versión sgARN11, pero con UA agregado en el extremo 3'. Véase, la Figura 46.

Anteriormente, se ha informado que una secuencia corta de poliA puede ser útil para los constructos Cas9. Platt, (2015). En una modalidad, la presente invención contempla un constructo AAV-Nme2Cas9 que comprende un BGH polyA. Véase, la Figura 47. Aunque no es necesario comprender el mecanismo de una invención, se cree que esta secuencia de poliA reduce aún más el tamaño de la estructura principal del AAV todo en uno.

Además, se cree que esta estructura principal de AAV todo en uno minimizado (4.4 kb) aumenta la utilidad de Nme1Cas9 y Nme2Cas9 al incluir otro sgARN para la inactivación de genes duales o la escisión de fragmentos de ADN. Véase, la Figura 48, parte superior. Esta configuración también proporciona espacio libre en la cápside de AAV para incluir una plantilla de donante (~ 600 pares de bases) para la aplicación de reparación directa por homología. Véase,

la Figura 48, parte inferior. En algunas modalidades, los constructos de sgARN AAV duales se empaquetan dentro de un único vector AAV.

El Nme1Cas9 relativamente compacto está activo en la edición del genoma en una variedad de tipos de células. Para explotar el pequeño tamaño de este ortólogo Cas9, se generó un constructo de AAV todo en uno con Nme1Cas9 optimizado con codón humano bajo la expresión del promotor U1a de ratón y con su sgARN dirigido por el promotor U6. Véase, la Figura 49A. Inicialmente, se seleccionaron dos sitios en el genoma del ratón para probar la actividad de la nucleasa de Nme1Cas9 *in vivo*: el gen *Rosa26* "puerto seguro" (dirigido por sgRosa26); y el gen de proproteína convertasa subtilisina/ kexina tipo 9 (*Pcsk9*) (dirigido por sgPcsk9), un objetivo terapéutico común para reducir el colesterol circulante y reducir el riesgo de enfermedad cardiovascular. Figura 49B. Las predicciones con objetivos fuera de lugar del genoma para estas guías se determinaron computacionalmente mediante el uso del paquete Bioconductor CRISPRseek 1.9.1 con PAM N₄GN₃ y hasta seis desajustes. Zhu, "*CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome editing systems*" *PLoS One* 2014;9:e108424. Muchos PAMS N₄GN₃ están inactivos, por lo tanto, es casi seguro que estos parámetros de búsqueda arrojen una red más amplia que el perfil real fuera del objetivo. A pesar de la naturaleza expansiva de la búsqueda, un análisis no reveló sitios con objetivos fuera de lugar con menos de cuatro desajustes en el genoma del ratón. Véase, la Figura 50. Las eficiencias de edición con objetivos en lugar en estos sitios objetivo se evaluaron en células de hepatoma Hepa1-6 de ratón mediante transfecciones de plásmidos y la cuantificación indel se realizó mediante la descomposición de trazas de secuencia mediante el uso de la herramienta web para el seguimiento de indeles mediante descomposición (TIDE). Brinkman, "*Easy quantitative assessment of genome editing by sequence trace decomposition*" *Nucleic Acids Res.* 2014; 42: e168. Los datos muestran > 25% de valores indel para las guías seleccionadas, la mayoría de las cuales fueron eliminaciones. Véase, la Figura 49C.

Para evaluar la eficacia preliminar del vector construido AAV-sgARN-hNme1Cas9 todo en uno, se administró hidrodinámicamente el plásmido sgPcsk9 libre de endotoxina en los ratones C57Bl/6 mediante la inyección en la vena de la cola. Este método puede entregar ADN plasmídico a ~ 40% de los hepatocitos para la expresión transitoria. Liu, "*Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA*" *Gene Ther.* 1999;6:1258-66. Los análisis de Indel por TIDE mediante el uso de ADN extraído de tejidos hepáticos

revelaron 5-9% de índices 10 días después de la administración del vector, comparable a las eficiencias de la edición obtenidas con pruebas análogas de SpyCas9. Véase, la Figura 49D; y Xue, “*CRISPR-mediated direct mutation of cancer genes in the mouse liver*” *Nature* 2014;514:380–4. Estos resultados sugieren que Nme1Cas9 es capaz de editar células hepáticas *in vivo*.

La tirosinemia hereditaria tipo I (HT-I) es una enfermedad genética fatal causada por mutaciones autosómicas recesivas en el gen *Fah*, que codifica la enzima de hidroxilasa de acetato de fumarato (FAH). Los pacientes con FAH disminuida tienen una vía catabólica de tirosina alterada, lo que conduce a la acumulación de acetoacetato de fumarato y acetoacetato de succinilo tóxico, causando daño hepático y renal. Grompe M., “*The pathophysiology and treatment of hereditary tyrosinemia type I*” *Semin Liver Dis.* 2001;21:563-71. En las últimas dos décadas, la enfermedad ha sido controlada por 2- (2-nitro-4-trifluorometilbenzoil) -1,3-ciclohexanodiona (NTBC), que inhibe la 4-hidroxifenilpiruvato dioxigenasa aguas arriba en la vía de degradación de la tirosina, evitando así la acumulación de los metabolitos tóxicos. Lindstedt, “*Treatment of hereditary tyrosinaemia type I by inhibition of 4-hydroxyphenylpyruvate Dioxygenase*” *Lancet* 1992; 340: 813–7. Sin embargo, este tratamiento requiere un manejo de por vida de la dieta y la medicación y eventualmente puede requerir un trasplante de hígado. Das, AM., “*Clinical utility of nitisinone for the treatment of hereditary tyrosinemia type-1 (HT-I)*” *Appl Clin Genet.* 2017;10:43–8.

Se ha probado varias estrategias de la terapia génica para corregir un gen *Fah* defectuoso mediante el uso de mutagénesis dirigida a sitio o reparación directa por homología mediante CRISPR-Cas9. Paulk, “*Adenoassociated virus gene repair corrects a mouse model of hereditary tyrosinemia in vivo*” *Hepatology* 2010;51:1200–8; Yin, “*Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo*” *Nat Biotechnol.* 2016;34:328–33; y Yin, “*Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype*” *Nat Biotechnol.* 2014;32:551-3. Se ha reportado que la modificación exitosa de solo 1/10,000 de hepatocitos en el hígado es suficiente para rescatar los fenotipos de los ratones *Fah^{mut/mut}*. Recientemente, se ha sugerido un enfoque de reprogramación de la vía metabólica en el que la función de la enzima de la hidroxifenilpiruvato dioxigenasa (HPD) se vio interrumpida

por la eliminación de los exones 3 y 4 del gen *Hpd* en el hígado. Pankowicz, “*Reprogramming metabolic pathways in vivo with CRISPR/Cas9 genome editing to treat hereditary tyrosinaemia*” *Nat Commun.* 2016;7:12642. Esto proporciona un contexto para evaluar la eficacia de la edición de Nme1Cas9, por ejemplo, al dirigirse a *Hpd* y al evaluar el rescate del fenotipo de la enfermedad en ratones mutantes *Fah*. Grompe, “*Loss of fumarylacetoacetate hydrolase is responsible for the neonatal hepatic dysfunction phenotype of lethal albino mice*” *Genes Dev.* 1993;7:2298-307. Para este propósito, dos sitios objetivo (uno en el exón 8 [sgHpd1] y en el exón 11 [sgHpd2]) se seleccionaron e identificaron dentro del marco de lectura abierto de *Hpd*. Véase, la Figura 51A. Estas guías (ej., SgARN) facilitaron las eficiencias promedio inducidas por Nme1Cas9 de 10.8% y 9.1%, respectivamente, mediante transfecciones de plásmidos en células Hepa1-6. Figura 52.

Tres grupos de ratones se trataron mediante inyección hidrodinámica con solución salina tamponada con fosfato (PBS) o con uno de los dos plásmidos AAV-sgARN-hNme1Cas9 todo en uno sgHpd1 y sgHpd2. Un ratón en el grupo sgHpd1 y dos en el grupo sgHpd2 se excluyeron del estudio de seguimiento debido a las inyecciones fallidas de la vena de la cola. Los ratones se retiraron del agua que contenía NTBC siete días después de las inyecciones y se controló su peso durante 43 días después de la inyección. Véase, la Figura 51B. Los ratones inyectados con PBS sufrieron una pérdida de peso severa (un sello distintivo de HT-I) y se sacrificaron después de perder 20% de su peso corporal. En general, todos los ratones sgHpd1 y sgHpd2 mantuvieron con éxito su peso corporal durante 43 días en general y durante al menos 21 días sin NTBC. Véase, la Figura 51C.

El tratamiento con NTBC tuvo que reanudarse durante 2-3 días para dos ratones que recibieron sgHpd1 y uno que recibió sgHpd2 para permitirles recuperar el peso corporal durante la tercera semana después de la inyección de plásmido, tal vez debido a la baja eficiencia de la edición inicial, daño hepático debido a la inyección hidrodinámica, o ambos. Por el contrario, todos los demás ratones tratados con sgHpd1 y sgHpd2 lograron indeles con frecuencias en el rango de 35-60%. Véase, la Figura 51D. Es probable que este nivel de inactivación genética refleje no solo los eventos de edición iniciales, sino también la expansión competitiva de los linajes celulares editados (después del retiro de NTBC) a expensas de sus contrapartes no editadas. La histología hepática reveló que el daño hepático es sustancialmente menos grave en los ratones tratados con sgHpd1 y sgHpd2 en comparación con los ratones *Fah^{mut/mut}* inyectados

con PBS, como lo indica el menor número de hepatocitos multinucleados en comparación con los ratones inyectados con PBS. Véase, la Figura 53.

Los vectores AAV se han utilizado recientemente para la generación de ratones con genoma editado, sin la necesidad de microinyección o electroporación, simplemente empapando los cigotos en medio de cultivo que contiene vector(es) de AAV, seguido por la reimplantación en hembras pseudopreñadas. La edición se obtuvo previamente con un sistema AAV dual en el que SpyCas9 y su sgARN se entregaron en vectores separados. Yoon, “*Streamlined ex vivo and in vivo genome editing in mouse embryos using recombinant adeno-associated viruses*” *Nat. Commun.* 9:412 (2018). Para probar si Nme2Cas9 podría permitir una edición precisa y eficiente en cigotos de ratón con un sistema de entrega de AAV todo en uno, se seleccionó al gen de tirosinasa (*Tyr*), donde una inactivación bi alélica altera la producción de melanina, dando como resultado crías albinas. Yokoyama, “*Conserved cysteine to serine mutation in tyrosinase is responsible for the classical albino mutation in laboratory mice*” *Nucleic Acids Res.* 18:7293-7298 (1990).

Un sgARN *Tyr* eficiente (que escinde el locus *Tyr* a solo 17 pb del sitio de la mutación albina clásica) se validó en células Hepa1-6 mediante transfecciones transitorias. Véase, la Figura 57. A continuación, los cigotos C57BL/6NJ se incubaron durante 5-6 horas en medio de cultivo que contenía 3×10^9 o 3×10^8 GC de un vector AAV6 todo en uno que expresa Nme2Cas9 junto con el sgARN *Tyr*. Después del cultivo nocturno en medios frescos, los cigotos que avanzaron a la etapa de dos células se transfirieron al oviducto de los receptores pseudopreñados y se les permitió desarrollarse a término. Véase, la Figura 58A. El análisis del color del pelaje de las crías reveló ratones albinos, de color gris claro (lo que sugiere un alelo hipomórfico de *Tyr*), o que tenían un color de pelaje abigarrado compuesto de putos albinos y grises claros pero que carecía de pigmentación negra. Véase, las Figuras 58B y 58C. Estos resultados sugieren una alta frecuencia de mutaciones bialélicas ya que la presencia de un solo alelo *Tyr* de tipo silvestre debería producir pigmentación negra. Un total de cinco crías (10%) nacieron del experimento 3×10^9 GC. Todos ellos llevaban indeles; fenotípicamente, dos eran albinos, uno era gris claro y dos tenían pigmentación abigarrada, lo que indica mosaicismo. Del experimento de 3×10^8 GC, se obtuvieron cuatro (4) crías (14%), dos de los cuales murieron al nacer, evitando el color del pelaje o el análisis del genoma. El análisis del color del pelaje de las dos crías restantes reveló una cría gris claro y otro mosaico. Estos resultados indican que la

entrega de un AAV único de Nme2Cas9 y su sgARN se puede usar para generar mutaciones en cigotos de ratón sin microinyección o electroporación.

Para medir la formación de indel con objetivos en lugar en el gen *Tyr*, se aisló el ADN de las colas de cada ratón, se amplificó el locus y se realizó un análisis TIDE. Los datos mostraron que todos los ratones tenían altos niveles de edición con objetivos en lugar mediante Nme2Cas9, entre el rango de 84% al 100%. Véase, las Figuras 57B y 5C. La mayoría de las lesiones en el ratón albino 9-1 fueron una delección de 1 o 4 pb, lo que sugiere mosaicismo o trans-heterocigosidad. Los ratones albinos 9-2 exhibieron una delección 2-bp uniforme. Véase, la Figura 58C. El análisis del ADN de la cola de ratones gris claro reveló la presencia de mutaciones en el marco que son potencialmente una causa del color del pelaje gris claro. La complejidad mutacional limitada sugiere que la edición ocurrió temprano durante el desarrollo embrionario en estos ratones. Una hembra (ratón 9-2) se apareó con un macho albino clásico, y las seis crías resultantes eran albinas, lo que demuestra que las mutaciones generadas por la entrega de AAV zigótico todo en uno de Nme2Cas9+sgARN pueden transmitirse a través de la línea germinal. Estos resultados proporcionan una ruta simplificada hacia la mutagénesis de mamíferos a través de la aplicación de un único vector AAV, en este caso al entregar Nme2Cas9 como su sgARN.

Se considera que los pacientes con mutaciones en el gen *Hpd* tienen tirosinemia de tipo III y exhiben un alto nivel de tirosina en la sangre, pero por lo demás parecen ser en gran medida asintomáticos. Szymanska, “*Tyrosinemia type III in an asymptomatic girl*.” *Mol Genet Metab Rep.* 2015;5:48–50; y Nakamura, “*Animal models of tyrosinemia*” *J Nutr.* 2007;137:1556S–60S. HPD actúa aguas arriba de FAH en la vía del catabolismo de la tirosina y la alteración de *Hpd* mejora los síntomas de HT-I al prevenir la acumulación de metabolitos tóxicos que resulta de la pérdida de FAH. Los análisis estructurales de HPD revelan que el dominio catalítico de la enzima HPD se encuentra en el extremo C-terminal de la enzima y está codificado por los exones 13 y 14. Huang, “*The different catalytic roles of the metal-binding ligands in human 4-hydroxyphenylpyruvate dioxygenase*” *Biochem J.* 2016;473:1179–89. Por lo tanto, los cambios de marco que inducen los indeles aguas arriba del exón 13 deberían hacer que la enzima esté inactiva. Este contexto se utilizó para demostrar que la inactivación de *Hpd* por inyección hidrodinámica del plásmido Nme1Cas9 es un enfoque viable para rescatar ratones HT-I. Nme1Cas9 puede editar sitios que llevan varios PAM diferentes (N₄GATT [consenso],

N₄GCTT, N₄GTTT, N₄GACT, N₄GATA, N₄GTCT, y N₄GACA). Los experimentos de edición de *Hpd* confirmaron uno de los PAM variantes *in vivo* con la guía sgHpd2, que se dirige a un sitio con un PAM N₄GACT.

Aunque las inyecciones hidrodinámicas de plásmidos pueden generar indeles, el desarrollo terapéutico puede requerir estrategias de entrega menos invasivas, tales como el uso de un rAAV. Con este fin, los plásmidos AAV-sgARN-hNme1Cas9 todo en uno se empaquetaron en cápsides de hepatocitos trópicos AAV8 para dirigirse a *Pcsk9* (sgPcsk9) y *Rosa26* (sgRosa26). Véase, la Figura 49B; Gao, “Novel adenoassociated viruses from rhesus monkeys as vectors for human gene therapy” *Proc Natl Acad Sci EE.UU.* 2002; 99: 11854–9; y Nakai, “Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors in mice” *J Virol.* 2005;79:214–24. *Pcsk9* y *Rosa26* se usaron en parte para permitir que la entrega de AAV Nme1Cas9 se comparara con la de otros ortólogos Cas9 entregados de manera similar y dirigidos a los mismos loci. Ran, “In vivo genome editing using *Staphylococcus aureus* Cas9” *Nature* 2015;520:186–91. Los vectores se administraron en ratones C57BL/6 a través de la vena de la cola. Véase, la Figura 54A. Los niveles de colesterol se monitorearon en el suero y se midieron las proteínas PCSK9 y las frecuencias indel en los tejidos del hígado 25 y 50 días después de la inyección.

Mediante el uso de un ensayo de punto final colorimétrico, se determinó que el nivel de colesterol en suero circulante en los ratones administrados con Nme1Cas9/sbPcsk9 disminuyó significativamente ($p < 0.001$) en comparación con los ratones PBS y Nme1Cas9/sbRosa26 a los 25 y 50 días después de la inyección. Véase, la Figura 54B. Los análisis de secuenciación profunda dirigidos en los sitios objetivo de *Pcsk9* y *Rosa26* revelaron índices muy eficientes de 35% y 55%, respectivamente, a los 50 días después de la administración del vector. Figura 54C. Además, un ratón de cada grupo se sacrificó a los 14 días después de la inyección y reveló eficiencias indel con objetivos en lugar del 37% y 46% en *Pcsk9* y *Rosa26*, respectivamente. Como se esperaba, los niveles de proteína PCSK9 en los hígados de los ratones tratados con Nme1Cas9/sbPcsk9 se redujeron sustancialmente en comparación con los ratones inyectados con PBS y Nme1Cas9/sbRosa26. Véase, la Figura 54D. La edición eficiente, la reducción de PCSK9 y la disminución del colesterol en suero indican la entrega exitosa y la actividad de Nme1Cas9 en el locus *Pcsk9*.

SpyCas9 entregado por vectores virales se sabe que provoca respuestas inmunes del huésped. Chew, “*A multifunctional AAV-CRISPR-Cas9 and its host response*” *Nat Methods* 2016;13:868–74; and Wang et al., “*Adenovirus-mediated somatic genome editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses*” *Hum Gene Ther.*

5 2015;26:432-42. Para investigar si los ratones inyectados con AAV8-sgARN-hNme1Cas9 generan anticuerpos anti-Nme1Cas9, se utilizaron sueros de los animales tratados para realizar ELISA IgG1. Estos resultados muestran que Nme1Cas9 provoca una respuesta humoral en estos animales. Véase, la Figura 55. A pesar de la presencia de una respuesta inmune, Nme1Cas9 administrado por rAAV es altamente funcional *in vivo*, sin signos aparentes de anormalidades o
10 daño hepático. Véase, la Figura 16.

Una preocupación importante en la edición terapéutica del genoma CRISPR/Cas9 es la posibilidad de actividad en ediciones con objetivos fuera de lugar. Por ejemplo, se ha descubierto que Nme1Cas9 de tipo silvestre es una plataforma de edición genómica de alta precisión natural en células de mamífero cultivadas. Lee, “*The Neisseria meningitidis CRISPR-*
15 *Cas9 system enables specific genome editing in mammalian cells*” *Mol Ther.* 2016;24:645-54. Para determinar si Nme1Cas9 mantiene su perfil mínimo con objetivos fuera de lugar en células de ratón e *in vivo*, se seleccionaron sitios con objetivos fuera de lugar en el genoma del ratón mediante la identificación imparcial de todo el genoma de DSB habilitados por secuenciación (GUIDE-seq). Tsai, “*Defining and improving the genome-wide specificities of*
20 *CRISPR-Cas9 nucleases*” *Nat Rev Genet.* 2016;17:300-12. Las células Hepa1–6 se transfectaron con plásmidos AAV-sgARN-hNme1Cas9 todo en uno sgPcsk9, sgRosa26, sgHpd1 y sgHpd2 y el ADN genómico resultante se sometió a análisis GUIDE-seq. De acuerdo con las observaciones en células humanas (datos no mostrados), GUIDE-seq reveló muy pocos sitios con objetivos fuera de lugar (OT) en el genoma del ratón. Se identificaron cuatro sitios
25 potenciales de OT para sgPcsk9 y otros seis para sgRosa26. No se detectaron ediciones con objetivos fuera de lugar con sgHpd1 y sgHpd2. Véase, la Figura 56A. Estos datos validan aún más que Nme1Cas9 es intrínsecamente hiperprecisa.

Varios de los supuestos sitios OT para sgPcsk9 y sgRosa26 carecen de las preferencias Nme1Cas9 PAM (es decir, N₄GATT, N₄GCTT, N₄GTTT, N₄GACT, N₄GATA, N₄GTCT y
30 N₄GACA). Véase, la Figura 56B. Para validar estos sitios OT, se realizó una secuenciación profunda dirigida mediante el uso de ADN genómico de células Hepa1-6. Mediante esta lectura

más sensible, los indeles eran indetectables por encima del fondo en todos estos sitios OT, excepto OT1 de *Pcsk9*, que tenía una frecuencia indel < 2%. Véase, la Figura 56B. Para validar la alta fidelidad de Nme1Cas9 *in vivo*, se midió la formación del indel en estos sitios OT en el ADN genómico del hígado de los ratones tratados con AAV8-Nme1Cas9, dirigidos a sgPcsk9 y dirigidos a sgRosa26. Se encontró poca o ninguna edición con objetivos fuera de lugar detectable en el hígado de ratones sacrificados a los 14 días en todos los sitios, excepto sgPcsk9 OT1, que mostró < 2% de la eficiencia de la lesión. Más importante aún, este nivel de edición OT permaneció por debajo de < 2% incluso después de 50 días y también permaneció indetectable o muy bajo para todos los demás sitios OT candidatos. Estos resultados sugirieron que la expresión extendida (50 días) de Nme1Cas9 *in vivo* no compromete su fidelidad de direccionamiento. Véase, la Figura 56C.

Para lograr la entrega dirigida de Nme1Cas9 a varios tejidos *in vivo*, los vectores rAAV son una plataforma de entrega prometedora debido al tamaño compacto del transgen Nme1Cas9, que permite la entrega de Nme1Cas9 y su guía en un formato todo en uno. Los datos presentados en este documento validan este enfoque para el direccionamiento de los genes *Pcsk9* y *Rosa26* en ratones adultos, con una edición eficiente observada incluso a los 14 días después de la inyección. Nme1Cas9 es intrínsecamente precisa, incluso sin la amplia modificación que se requiere para reducir el con objetivos fuera de lugar de SpyCas9. Lee, “*The Neisseria meningitidis CRISPR-Cas9 system enables specific genome editing in mammalian cells*” *Mol Ther.* 2016;24:645–54; Bolukbasi, “*Creating and evaluating accurate CRISPRCas9 scalpels for genomic surgery*” *Nat Methods* 2016;13:41–50; Tsai, “*Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases*” *Nat Rev Genet.* 2016;17:300–12; y Tycko, “*Methods for optimizing CRISPR-Cas9 genome editing specificity*” *Mol Cell.* 2016;63:355-70.

Se realizaron comparaciones lado a lado de la edición OT Nme1Cas9 en células cultivadas e *in vivo* mediante la secuenciación profunda dirigida y se descubrió que el con objetivos fuera de lugar es mínimo en ambos entornos. La edición en el sitio sgPcsk9 OT1 (dentro de un locus no anotado) fue la más alta detectable a ~ 2%.

IV. Ortólogos pequeños Cas9 con PAM ricos en citosina

Como se señaló anteriormente, los sistemas CRISPR se pueden clasificar en al menos seis (6) tipos diferentes. En general, los sistemas de Tipo II se clasifican por la presencia de una proteína nucleasa Cas9. Por ejemplo, se cree que una proteína nucleasa Cas9 es una nucleasa guiada por ARN que puede reutilizarse como plataforma de edición genómica en casi todos los organismos, incluidos los humanos. Los informes han indicado que la edición del genoma Cas9 se ha utilizado en la medicina, agricultura, terapia génica humana y muchas otras aplicaciones.

En general, el direccionamiento de un locus genético específico en el genoma humano se puede lograr mediante una proteína nucleasa Cas9 unida a un ARN de guía única (sgARN) que se dirige al locus mediante una interacción con una secuencia de ácido nucleico específica (*ej.*, un motivo adyacente al protoespaciador; PAM). Los sgARN generalmente comprenden un segmento de 20-24 nucleótidos que es complementario a una secuencia de ácido nucleico objetivo seguido de una región constante que interactúa (*ej.*, se une) con la proteína Cas9. Para que la proteína nucleasa Cas9 realice la edición del genoma, el complejo Cas9:sgARN reconoce primero una secuencia de motivo adyacente al protoespaciador (PAM) que normalmente se encuentra aguas abajo de la secuencia del sitio objetivo. Aunque no es necesario comprender el mecanismo de una invención, se cree que cada proteína nucleasa Cas9 tiene afinidad por un PAM particular (es decir, mediada por un dominio de reconocimiento de motivo adyacente al protoespaciador). En la ausencia del dominio de reconocimiento PAM que se une a una secuencia de ácido nucleico PAM objetivo secuencia aguas abajo de ADN de doble cadena (dsDNA) no puede escindirse por la nucleasa Cas9.

Los reportes sugieren que solo un puñado de ortólogos Cas9 han sido validados para la edición del genoma humano. Tres de los tipos CRISPR-Cas9 reportados incluyen II-A, II-B y II-C. Cas9 de Tipo II-A (*ej.*, *Streptococcus pyogenes* (SpyCas9)), es el Cas9 más utilizado hasta la fecha. Sin embargo, SpyCas9 (y la mayoría de los otros ortólogos de tipo II-A) posee varias características que pueden hacer que no sea adecuado para ciertas aplicaciones. Primero, SpyCas9 es relativamente grande, lo que hace que este Cas9 no sea adecuado para un empaquetamiento eficiente en vectores virales. En segundo lugar, SpyCas9 tiene una alta tasa de actividad con objetivos fuera de lugar (es decir, escinde el ADN en loci no deseados en el genoma humano), aunque se han modificado variantes de mayor especificidad. Finalmente, PAM SpyCas9 (*ej.*, NGG) tiene un uso limitado en algunos sitios del genoma humano, o para aplicaciones en donde se debe reconocer un nucleótido específico durante la edición. Para

superar estas deficiencias, varios grupos han rehabilitado otros ortólogos Cas9 para que funcionen en humanos y otros organismos. Como se discutió anteriormente, los ortólogos Cas9 de tipo II-C (ej., Nme1Cas9) son lo suficientemente pequeños para el empaquetamiento viral todo en uno (ej., vectores de virus adenoasociados (AAV)) que resultan en una mayor actividad de fidelidad en células de mamíferos. Sin embargo, los PAM Cas9 II-C de tipo silvestre suelen tener aproximadamente cuatro (4) nucleótidos de longitud en comparación con una PAM SpyCas9 que suele tener dos (2) nucleótidos de longitud. Esta longitud adicional del PAM puede limitar el número de loci que puede ser dirigido por un PAM Cas9 II-C de tipo silvestre. Esto crea una necesidad en la técnica de identificar más ortólogos Cas9 para la edición genómica.

Si bien hay miles de ortólogos Cas9 en la base de datos de NCBI para elegir, se requiere un proceso empírico para desarrollar pequeños ortólogos Cas9 de tipo II-C con PAM menos restrictivos que proporcionen una funcionalidad mejorada en células de mamíferos. En una modalidad, la presente invención contempla un ortólogo Cas9 de tipo II-C mejorado que permite la edición genómica precisa con una gama más amplia de sitios objetivo. En una modalidad, el ortólogo Cas9 de tipo II-C mejorado tiene un tamaño compacto capaz de una entrega viral eficiente. En una modalidad, el ortólogo Cas9 de tipo II-C mejorado incluye, pero sin limitación, *Haemophilus parainfluenzae* (HpaCas9), *Simonsiella muelleri* (SmuCas9) y *Neisseria meningitidis* de cepa De10444 (Nme2Cas9).

A. PAM cortos asociados con ortólogos Cas9 de tipo II-C

Los datos presentados en este documento muestran la caracterización de objetivos PAM cortos para varios ortólogos Cas9 de tipo II-C. Figura 17. Por ejemplo, los ortólogos Cas9 de tipo II-C pueden interactuar con PAM cortos que comprenden entre uno y cuatro nucleótidos requeridos. Aunque no es necesario comprender el mecanismo de una invención, se cree que estos PAM cortos ricos en C proporcionan una edición genómica mejorada de Cas9 de sitios objetivo que antes no eran accesibles incluso para los ortólogos Cas9 más compactos (ej., Nme1Cas9). En una modalidad, un PAM Nme2Cas9 tiene una secuencia de NNNNCc, en donde "c" es la única preferencia parcial. En una modalidad, un PAM SmuCas9 tiene una secuencia de NNNNCT. Figura 18.

Actualmente se cree que no se han validado los ortólogos Cas9 con PAM cortos ricos en C para la edición del genoma y que Nme2Cas9 es particularmente convincente como un candidato potencial para una actividad de edición de genes altamente eficiente en células

humanas. En una modalidad, la presente invención contempla una nucleasa Nme2Cas9 unida a un sgRNA Nme1Cas9 de tipo silvestre (ej., Cas9 *Neisseria meningitidis* 8013; anteriormente denominado NmeCas9). Nme1Cas9 ha sido descrita previamente. Sontheimer, “*RNA-Directed DNA Cleavage and Gene Editing by Cas9 Enzyme From Neisseria Meningitidis*” Número de publicación de solicitud de patente de Estados Unidos 2014/0349,405 (incorporada aquí como referencia). Aunque Nme1Cas9 puede ser útil para la edición del genoma, su principal limitación es su PAM relativamente largo, que restringe el número de sitios editables en cualquier locus genómico dado.

En algunas modalidades, la presente invención contempla PAM más cortos y menos estrictos para los ortólogos Cas9 de tipo II-C que incluyen, pero no se limitan a, Nme2Cas9. Aunque no es necesario comprender el mecanismo de una invención, se cree que los PAM cortos y menos estrictos alivian parcialmente las limitaciones de restricción del objetivo, al tiempo que dejan muchas, si no la mayoría, de las ventajas de Nme1Cas9 que incluyen, entre otras, tamaño pequeño (ej., compacidad) para una entrega eficiente de AAV todo en uno y una precisión del objetivo mejorada (ej., reducciones en las escisiones con objetivos fuera de lugar). Además, los sgARN minimizados para Nme1Cas9 discutidos anteriormente también son compatibles con los constructos Nme2Cas9. En consecuencia, estos ARN guía truncados probablemente también podrían usarse para la edición del genoma con Nme2Cas9.

En una modalidad, la presente invención contempla un PAM HpaCas9 que tiene una secuencia de NNNNGNTTT. A pesar del hecho de que el PAM largo limita el número de sitios dirigibles en el genoma humano, se cree que el PAM HpaCas9 puede dirigirse a sitios con una precisión muy alta que es similar a la precisión extrema Nme1Cas9 (*supra*).

Los datos presentados en este documento demuestran la capacidad de las nucleasas Cas9 de tipo II-C dirigidas a PAM cortos ricos en C para realizar la edición del genoma en células humanas (HEK293T). Certo, “*Tracking genome engineering outcome at individual DNA breakpoints*” *Nature Methods* 8:671-676 (2011). Por ejemplo, se demostró que HpaCas9 y Nme2Cas9 proporcionan una edición genómica eficiente en loci específicos que demuestran que son activos en células de mamíferos. Figura 19 y Tabla 2.

Tabla 2: Secuencias objetivo de ortólogos Cas9 de tipo II-C representativos en el genoma humano

Cas9	Secuencia espaciadora	PAM	Cromosoma
Nme2	GAATATCAGGAGACTAGGAAGGAG	GAGGCCTA	19
Hpa	GGACAGGAGTCGCCAGAGGCCGGT	GGTGATT	4
Smu	GCACCTGCCTCGTGAATACGGT	AAACCTAC	Reportero de semáforo

Estos datos muestran que tanto Nme2Cas9 como HpaCas9 realizaron la edición del genoma a niveles comparables al Nme1Cas9 previamente validado en el mismo locus genómico. Para SmuCas9, la eficiencia de la edición es relativamente baja, aunque es significativo que la actividad no sea cero, y se esperan mejoras en la eficiencia. Nme2Cas9 se utilizó luego para probar catorce (14) sitios adicionales en el reportero de semáforo (TLR) integrado en el genoma de las células HEK293T. En estos ensayos, cada sitio se ajusta a una plantilla PAM en donde un "C" es el quinto nucleótido de la región PAM (es decir, NNNNCNNN). Sorprendentemente, los catorce sitios fueron editados por Nme2Cas9, lo que indica que esta enzima es constantemente activa con una variedad de guías en células de mamíferos. El ARN guía más exitoso se ajusta al consenso PAM NNNNCCN. Figura 20.

La escisión del ortólogo Cas9 de tipo II-C se probó para determinar la sensibilidad a las proteínas anti-CRISPR. Las proteínas anti-CRISPR son proteínas de origen natural que pueden desactivar Cas9 cuando ya no se desea la actividad de Cas9. Los datos muestran que los tres ortólogos Cas9 de Tipo II-C están inhibidos por ciertos anti-CRISPR. Figura 21. La capacidad de control de estos ortólogos Cas9 por anti-CRISPR podría aumentar su utilidad potencial en la edición del genoma.

B. Edición de genes Nme2Cas9

Los datos presentados en este documento muestran la edición de genes mediante el uso del complejo Nme2Cas9-sgRNA. Los datos emplean el sistema reportero de semáforo (TLR) para demostrar que cualquier dinucleótido CC en una secuencia del gen objetivo puede funcionar como un PAM, dentro del contexto de una secuencia NNNNCC (*supra*). Figura 22. Las barras azules son el porcentaje de células que exhiben fluorescencia, mientras que las barras rojas indican el porcentaje de edición con mayor precisión basado en la secuenciación ("análisis TIDE"). Estos datos confirman que un dinucleótido es suficiente para la unión del PAM Nme2Cas9 en oposición a un requisito para una secuencia de trinucleótidos (*ej.*, la "X" en la secuencia NNNNCCX). Aunque no es necesario comprender el mecanismo de una invención, se cree que esto significa que los sitios objetivo genómicos editables Nme2Cas9 son al menos tan

frecuentes como los sitios editables SpyCas9, y más frecuentes que con SauCas9, Nme1Cas9 o CjeCas9 y otras alternativas actuales.

Además, se emplearon ensayos T7E1 para analizar la edición de sitios genómicos nativos (ej., no un reportero fluorescente artificial integrado). Estos datos sugieren que, en algunas situaciones, la segunda "C" podría no ser necesaria. Véase, la Figura 23. Tenga en cuenta que los sitios objetivo DeTS1 y DeTS4, ambos en el locus AAVS1, permiten la edición en sitios objetivo con PAM candidatos NNNNCA y NNNNCG, respectivamente. Varios de estos sitios objetivo Nme2Cas9 se describen en este documento. Véase, la Tabla 3.

10 Tabla 3: Sitios objetivo de PAM representativos para Nme2Cas9

Nombre del sitio objetivo	Locus objetivo	Secuencia objetivo (Espaciador-PAM)
Nme2TS1	AAVS1	ATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACC CCACAGTGGG
Nme2TS4	AAVS1	CAGATAAGGAATCTGCCTAACAGGAGGTGGGGGTAGACG AATATCAGGAGA
Nme2TS5	AAVS1	GGGGTTAGACGAATATCAGGAGACTAGGAAGGAGGAGGC CTAAGGATGGGGG
Nme2TS6	Chr. 14	CCCACCCGGCGGCGCCTCCCTGCAGGGCTGCTCCCGAGCCC AAACCGCCGCG
Nme2TS10	AAVS1	TCCGAGAGCTCAGCTAGTCTTCTTCCTCCAACCCGGGCCCT ATGTCCACTTC
Nme2TS11	AAVS1	TGGGTACTTTTATCTGTCCCCTCCACCCACAGTGGGGCCA CTAGGGACAGG
Nme2TS12	AAVS1	GTAGGGGAGCTGCCAAATGAAAGGAGTGAGAGGTGACC CGAATCCACAGGA
Nme2TS13	AAVS1	TAGCACCTCTCCATCCTCTTGCTTTCTTTGCCTGGACACCCC GTTCTCCTGT
Nme2TS14	AAVS1	GTCTCCCTTGCGTCCCGCCTCCCCTTCTTGTAGGCCTGCATC ATCACCGTTT

Nme2TS15	AAVS1	CCTCACCCAA CCCCATGCCGTGTTCACTCGCTGGGTTCCCT TTTCCTTCTCCT
Nme2TS16	Chr. 14	GCGCAGGACAG GGAGTCGCCAGAGGCCGGTGGTGGATTTC TCCCCGCATCTC
Nme2TS17	Chr. 14	CGCGGGGAC GCCCAGCGGCCGGATATCAGCTGCCACGCCC GCGTGGGCGGA
Nme2TS22	VEGF	GATTCCAATA GATCTGTGTGTCCCTCTCCCCACCCGTCCCT GTCCGGCTCTC
Nme2TS23	VEGF	TGACCCCTGGC TTCTCCCCGCTCCAACGCCCTCAACCCCA CACGCACACAC
Nme2TS24	VEGF	TCCCTCTCCCC ACCGTCCCTGTCCGGCTCTCCGCCTTCCCC TGCCCCCTTC
Nme2TS25	VEGF	ACACGCACACA CTCACTCACCCACACAGACACACACGTCC TCACTCTCGAAG
Nme2TS26	Chr. 7 (CFTR)	TAAGCACAGT GGAAGAATTCATTCTGTTCTCAGTTTTCCT GGATTATGCCT
Nme2TS27	Chr. 7 (CFTR)	TTCATTCTGTT CTCAGTTTTCCTGGATTATGCCTGGCACCAT TAAAGAAAAT

Aunque no es necesario comprender el mecanismo de una invención, se cree que estos datos sugieren que puede haber sitios de edición candidatos en un genoma en cada 4-8 pares de bases, en promedio. Estos datos también sugieren que la mayoría de los sgARN Cas9 tienen alguna

5 funcionalidad, por lo tanto, la necesidad de un cribado de sgARN puede exagerarse en la técnica.

C. Dominios de interacción PAM de rápida evolución

Las aplicaciones *in vivo* de CRISPR-Cas9 tienen el potencial de transformar muchas áreas de biotecnología y terapéutica. Existen miles de ortólogos Cas9 en la naturaleza, de los

10 cuales solo unos pocos han sido validados para la edición genómica *in vivo*. El Cas9 de *Streptococcus pyogenes* (SpyCas9) se ha utilizado ampliamente debido a su motivo adyacente al protoespaciador (PAM) de alta eficiencia y no restrictivo NGG. Sin embargo, el tamaño relativamente grande de SpyCas9 restringe su uso en aplicaciones terapéuticas *in vivo* al usar

transportes de entrega con capacidad de empaquetamiento limitada, como el virus adenoasociado (AAV). Se sabe que varios ortólogos Cas9 más pequeños son activos en células de mamíferos, pero poseen PAM más restrictivos que limitan la densidad del sitio objetivo. La variación natural en los dominios de interacción PAM (PID) de los ortólogos Cas9 estrechamente

relacionados se puede aprovechar para identificar una enzima de edición del genoma que supera estas limitaciones. En algunas modalidades, la presente invención contempla el uso de un complejo Nme2Cas9 que es Cas9 compacto, naturalmente hiperpreciso con un PAM N₄CC. Los datos presentados en este documento muestran que Nme2Cas9 es una plataforma de edición genómica de mamíferos de alta fidelidad que ofrece la misma densidad de sitios objetivo que SpyCas9. La entrega de Nme2Cas9 con su ARN guía a través de un vector AAV todo en uno conduce a una edición del genoma eficiente en ratones adultos, con direccionamiento del gen *Pcsk9* en el hígado induciendo una reducción del colesterol en suero sin objetivos fuera de lugar (*infra*). Nme2Cas9 también proporciona una combinación única de compatibilidad AAV todo en uno, hiperprecisión natural y alta densidad de sitios objetivo para la edición del genoma *in vivo* en mamíferos.

Además de la densidad objetivo, minimizar la actividad con objetivos fuera de lugar (*ej.*, la escisión en loci no deseados) de una Cas9 es altamente deseable para su uso como agente terapéutico seguro. SpyCas9 de tipo silvestre (wt) posee un alto grado de actividad con objetivos fuera de lugar debido a su única cinética de hibridación. (Klein, 2018). En particular, quedan preguntas sobre su eficiencia de edición con objetivos en lugar y estas variantes no superan las limitaciones discutidas anteriormente con respecto al tamaño general. Por el contrario, se ha demostrado en el presente documento que las modalidades de Nme1Cas9 y CjeCas9 comprenden una actividad de edición de genes naturalmente precisa. Aunque no es necesario comprender el mecanismo de una invención, se cree que no se ha reportado previamente ningún ortólogo Cas9 que: i) sea activo en células humanas; ii) exhiba la densidad excepcionalmente alta del sitio objetivo de SpyCas9; iii) sea suficientemente compacto para la capacidad de entrega de AAV todo en uno; y iv) sea naturalmente hiperpreciso. En una modalidad, la presente invención contempla una Nme2Cas9 como una plataforma de edición genómica que comprende todas las características descritas anteriormente. Por ejemplo, Nme2Cas9 comprende un sitio de unión que comprende una alta afinidad por un PAM N₄CC, es hiperpreciso y funciona eficientemente

en células de mamífero. En una modalidad, Nme2Cas9 está empaquetado en una plataforma de entrega de AAV todo en uno para la edición terapéutica del genoma.

1. Ortólogos Nme1Cas9 estrechamente relacionados con PID de evolución rápida

5 Anteriormente se reportó que Nme1Cas9 (de la cepa 8013 de *Neisseria meningitidis*) es una Cas9 pequeña e hiperprecisa para la edición del genoma *in vivo* (Amrani, 2018). Sin embargo, Nme1Cas9 se une a un PAM largo (N₄G₄MTT) que limita su uso en ciertos contextos donde se puede dirigirse a una ventana pequeña. El reconocimiento de PAM por Cas9 ocurre predominantemente a través de la interacción proteína-ADN entre el dominio de interacción de
10 PAM (PID) de Cas9 y los nucleótidos adyacentes al PAM. Los PID están sujetos a una alta presión de selección por fagos y otros elementos genéticos móviles (MGE). Por ejemplo, se ha demostrado que las proteínas anti-CRISPR interactúan con los PID para inhibir Cas9 (*infra*). Esto puede resultar en ortólogos Cas9 estrechamente relacionados que tienen PID que reconocen PAM drásticamente diferentes.

15 Recientemente, este principio se destacó mediante el uso de dos especies de *Geobacillus*. Se determinó que *G. sterothempophilus* comprendía un PID específico para un PAM N₄CRAA, pero cuando se cambió por una cepa LC300 PID su afinidad cambió a un PAM N₄GMAA (Harrington, 2017). Se planteó la hipótesis de que, dado que las cepas de *N. meningitidis* están altamente secuenciadas, se podría encontrar un ortólogo Cas9 estrechamente relacionado con
20 PID de evolución rápida que reconocen diferentes PAM. Se investigaron los ortólogos Cas9 con identidad de secuencia alta (> 80%) para la cepa 8013 NmeCas9 porque este Cas9 se ha caracterizado completamente para la edición del genoma, es pequeño e hiperpreciso. Se identificaron varios ortólogos Cas9 que diferían en sus secuencias de aminoácidos PID en comparación con la cepa 8013. Figura 34A.

25 Se encontraron tres grupos distintos de ortólogos Cas9 con PID drásticamente diferentes. Figura 35A. Se seleccionó una cepa de cada grupo PID, por ejemplo, De11444 del grupo 2 y 98002 del grupo 3. Estos dos loci CRISPR tenían marcos de lectura abiertos Cas9 intactos y matrices CRISPR con varios espaciadores, lo que sugiere que son loci activos. Curiosamente, el crARN y el tracrARN de estos loci CRISPR eran idénticos a los de 8013 y pueden utilizar los
30 mismos sgARN. Figura 35B.

Para probar si estos ortólogos Cas9 tenían PID con afinidad por diferentes PAM, debido a la alta identidad de secuencia en el resto de la proteína de estos ortólogos, el PID 8013 se intercambió con el PID 98002 y el PID De11444. Para identificar los PAM, estas "quimeras" de proteínas se expresaron de forma recombinante, se purificaron y se usaron para la identificación de PAM *in vitro* como se describió anteriormente. Brevemente, un fragmento de ADN que comprende un protoespaciador y una secuencia aleatoria de diez (10) nucleótidos aguas abajo se escindió *in vitro* mediante el uso de Cas9 recombinante y un sgARN dirigido al protoespaciador. Figura 34B. Se usó una longitud espaciadora de nucleótidos G23 para el sgARN, de acuerdo con Nme1Cas9 8013 y otros sistemas de tipo II-C estudiados. El ensayo de identificación de PAM reveló que estas diferentes quimeras Cas9 tenían PID que reconocían los diferentes PAM. Por ejemplo, al reconocer un residuo C en la posición 5 en lugar de un G reconocido por Nme1Cas9 8013 con su PAM N₄GATT. Figura 34C:

Sin embargo, los nucleótidos restantes no pudieron caracterizarse con confianza debido a la baja eficiencia de escisión de las proteínas quiméricas, lo que sugiere que los pocos residuos fuera del PID probablemente estén involucrados en una actividad eficiente. Figura 35C. Para resolver aún más los PAM, se realizó un ensayo *in vitro* en una biblioteca con un PAM aleatorizado de 7 nucleótidos, con una C en la posición 5 (*ej.*, NNNNCNNN). Los resultados sugirieron que NmeCas9-De11444 y NmeCas9-98002 reconocieron PAM NNNNCC(A) y NNNNCAAA, respectivamente. Figura 35D. NmeCas9-De11444 tenía una fuerte preferencia por C en la posición 5, pero menos por los nucleótidos 6 y 7. Como se usa en este documento, el ortólogo Cas9 De11444 se denomina "Nme2Cas9" y el ortólogo Cas9 98002 se denomina "Nme3Cas9".

También se realizó este ensayo mediante el uso de Nme2Cas9 de longitud completa (*ej.*, sin intercambio de PID) y se observó resultados similares. Figura 34E. Estos resultados sugieren que Nme2Cas9 y Nme3Cas9 tienen PID que reconocen PAMs drásticamente diferentes a los de Nme1Cas9.

2. Nme2Cas9 en células humanas

Debido a que el PID Nme2Cas9 se une con una secuencia pequeña de PAM, este ortólogo es útil para la edición del genoma humano, especialmente cuando está involucrada una alta densidad de direccionamiento. Para caracterizar la Nme2Cas9, se clonó una Nme2Cas9 humanizada de longitud completa (sin intercambio PID) en un plásmido dirigido por CMV junto

con NLS para la expresión en mamíferos. Para la caracterización en células humanas, se usó un sistema reportero de semáforo similar al descrito anteriormente (Certo, 2011)

La inducción de indeles de desplazamiento de marco +1 se creó mediante la reparación imperfecta por unión de extremos no homólogos (NHEJ) en el locus TLR 2.0. En ausencia de un ADN donante, resultó una proteína mCherry dentro del marco, que puede cuantificarse mediante la citometría de flujo. Figura 36A. Como prueba inicial, se transfectó un plásmido Nme2Cas9 junto con quince (15) plásmidos de sgARN con espaciadores dirigidos a protoespaciadores con PAM N₄CCX. Como controles, SpyCas9 y Nme1Cas9 se utilizaron junto con sus sgARN afines dirigidos a protoespaciadores NGG y N₄GATT, respectivamente. Las células se recolectaron después de setenta y dos (72) horas y se cuantificó el número de células positivas mCherry para cada sitio objetivo. SpyCas9 y Nme1Cas9 mostraron una edición eficiente en sus respectivos objetivos (~ 28% y 10% mCherry, respectivamente) Figura 36B. Para Nme2Cas9, todos los quince (15) objetivos con PAM N₄CCX fueron funcionales en varios grados (que van del 4% al 20% de mCherry), mientras que los tratamientos con NmeCas9 sin acompañar los controles de sgARN y/o N₄GATT no proporcionaron células mCherry. Figura 36B. Estos datos sugirieron que Nme2Cas9 reconoce un PAM N₄CC en células humanas.

Para resolver aún más los PAM Nme2Cas9, los sitios objetivo también se probaron con N₅CX y N₄CD (D = A, T, G) en células reporteras TLR. No se observó una edición detectable en los sitios objetivo con PAM N₅CX y N₄CD, lo que sugiere que se requieren ambos nucleótidos C en las posiciones 5 y 6 para la actividad de Nme2Cas9 según el reportero TLR 2.0. Figuras 37A y 37B. Estos resultados demuestran que Nme2Cas9 comprende un PID que se une a un PAM N₄CC y es consistentemente funcional en células de mamífero en el locus TLR 2.0.

La longitud de la porción espaciadora del crARN difiere entre los diferentes ortólogos Cas9. La longitud óptima del espaciador de SpyCas9 es de veinte (20) nucleótidos, sin embargo, se toleran truncamientos de hasta diecisiete (17) nucleótidos. Fu, *Nature Biotechnology* 32, 279 (2014). En contraste, Nme1Cas9 comprende sgARN con veinticuatro (24) espaciadores de nucleótidos y tolera truncamientos de hasta dieciocho (18) nucleótidos. (Amrani, 2018). Para probar la longitud del espaciador para Nme2Cas9, se crearon plásmidos de sgARN dirigidos al mismo locus, pero con diferentes longitudes de espaciador. Figura 36C y Figura 37B. Se observaron actividades comparables cuando se usaron espaciadores G23, G22 y G21, con una disminución significativa de la actividad cuando la guía se truncó a G20 y G19. Figura 36C.

Estos resultados sugieren que la longitud óptima del espaciador de Nme2Cas9 es entre 22-24 nucleótidos, similar a la de Nme1Cas9, GeoCas9 y CjeCas9. Por lo tanto, todos los experimentos descritos a continuación se realizaron con 23-24 espaciadores de nucleótidos.

Se cree que los ortólogos Cas9 usan sus dominios HNH y RuvC para inducir un corte bicatenario en las cadenas complementarias y no complementarias del ADN objetivo, respectivamente. Alternativamente, las nickasas Cas9 se han utilizado para mejorar la especificidad de la edición genómica y la reparación directa por homología (HDR) mediante la creación de salientes. (Ran, 2013). Sin embargo, este enfoque solo ha tenido éxito mediante el uso de SpyCas9 debido a su alta densidad objetivo. Para usar Nme2Cas9 como una nickasa, se crearon Nme2Cas9^{D16A} y Nme2Cas9^{H588A} que proporcionan mutaciones en los residuos catalíticos de los dominios RuvC y HNH, respectivamente. Dado que TLR 2.0 también se puede usar para estudiar la eficiencia de HDR, donde un locus reparado expresa GFP cuando se proporciona un donante, se incluyó una secuencia de ADN donante para probar HDR con estas nickasas Nme2Cas9. Los sitios objetivo se seleccionaron dentro del gen TLR 2.0 para probar la funcionalidad de cada nickasa mediante el uso de ARN guías que se dirigen a sitios de escisión espaciados a 32 pb y 64 pb de distancia. Como control, Nme2Cas9 de tipo silvestre dirigida a un solo sitio mostró una edición eficiente, acompañada de la inducción de ambas vías de reparación NHEJ y HDR. Para las nickasas, los sitios de escisión separados por 32 pb y 64 pb de distancia mostraron edición mediante el uso de Nme2Cas9^{D16A} (nickasa HNH), pero ninguno de los objetivos se sustrajo con Nme2Cas9^{H588A}. Figura 36D.

Los ortólogos Cas9 comprenden una secuencia semilla que generalmente se hibrida con una secuencia objetivo entre ocho y doce (8-12) nucleótidos proximales al PAM. Los desajustes (*ej.*, no complementariedad) entre la secuencia semillas y el PAM pueden reducir la actividad de la nucleasa Cas9. Se realizó una serie de transfecciones transitorias que se dirigieron al mismo locus en el gen TLR 2.0 al generar desajustes de nucleótidos individuales a lo largo de un espaciador de veintitrés (23) nucleótidos. Figura 37C. Al igual que otros ortólogos Cas9, los datos sugieren que Nme2Cas9 posee una "secuencia semilla" en los primeros ocho a nueve (8-9) nucleótidos que se hibridan con una secuencia objetivo proximal al PAM, como se deduce de la disminución en el número de células positivas mCherry. Aunque la tolerancia a los desajustes depende en gran medida de la secuencia y el locus objetivo de un sgARN, estos resultados

sugieren que Nme2Cas9 tiene una tolerancia muy baja para los desajustes, particularmente en su secuencia semilla.

3. Eficiencia de la edición genómica de Nme2Cas9

Nme2Cas9 se usó para dirigirse a cuarenta (40) sitios objetivo diferentes en todo el

5 genoma humano en células HEK293T mediante el uso de transfecciones transitorias. Tabla 4.

Tabla 4: Sitios objetivo Nme2Cas9 de células HEK293T representativas

Número	Nombre del sitio	Secuencia espaciadora	PAM	Locus	150ng Cas9	Nombre del cebador TIDE	Cebador TIDE FW	Cebador TIDE RV
1	T51	GGTTCTGGGTACTTTATCTGTCC	CCTCCACC	AAVS1	0.2	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCACTTATTCTG
2	T54	GTCTGCGTAACAGGAGGTGGGGT	TAGACGAA	AAVS1	11	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCACTTATTCTG
3	T55	GAATATCAGGAGACTAGGAAGGAG	GAGGCCCTA	AAVS1	15	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCACTTATTCTG
4	T56	GCCTCCCTGCAGGGCTGCTCCC	CAGCCCAA	LINC01588	20	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
5	T510	GAGTAGTCTTCTTCTCCAAACC	GGGCCCTA	AAVS1	3.5	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCACTTATTCTG
6	T511	GATCTGTCCCTCCACCCACAGT	GGGGCGAC	AAVS1	9	AAVS1_TIDE1	TGGCTTAGCACCTCTCCAT	AGAACTCAGGACCACTTATTCTG
7	T512	GGCCAAATGAAAGAGTGAGAGG	TGACCGGA	AAVS1	10	AAVS1_TIDE2	TCCGCTTCTCTCACTCC	TAGGAAAGGAGGAGGCTTAAAG
8	T513	GCATCCTCTTCTTCTTTCCTG	GACACCCC	AAVS1	2	AAVS1_TIDE2	TCCGCTTCTCTCACTCC	TAGGAAAGGAGGAGGCTTAAAG
9	T516	GGAGTCCACAGAGGCCGCTGTGG	ATTTCCCTC	LINC01588	28	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
10	T517	GCCACGCGCGCGGATATCAGCTGC	CACGCCCG	LINC01588	0.2	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
11	T518	GGAAGGGAACATATTACTATTGC	TTTCCCTC	CYBB	1	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
12	T519	GTGGAGTGGCTGCTATCAGCTAC	GTATCCAA	CYBB	6	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
13	T520	GAGGAAGGGAACATATTACTATTG	CTTTCCCT	CYBB	11.2	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
14	T521	GTGAATTCTCATCAGCTAAATGC	CAAGCCTT	CYBB	1	NT555_TIDE	TAGAGAACTGGGTAGTGTG	CCAATATTGCATGGGATGG
15	T525	GCTCACTCACCCACACAGACACAC	ACGTCCTC	VEGF_A	16.6	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
16	T526	GGAAGAATTTCATTCTGTTCTCAG	TTTTCCTG	CFTR	2	hCFTR_TIDE1	TGGTGATTATGGGAGAACTGGAGC	ACCATTTGAGGAGCTTTGTCTCAC
17	T527	GCTCAGTTTTCTGGAATTATGCT	GGCACCAT	CFTR	4	hCFTR_TIDE1	TGGTGATTATGGGAGAACTGGAGC	ACCATTTGAGGAGCTTTGTCTCAC
18	T531	GCCTTGGAGCGGGGAGAGGCCAG	GGGTCACT	VEGFA	9	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
19	T534	GGGCCGCGGAGATGCTGCAGGGC	GGGGCCCC	LINC01588	0	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
20	T535	GCCACCGCGCGCGGCTCCTCTGC	AGGGCTGC	LINC01588	0	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
21	T536	GCCTGGCAGCTGATATCCGCGCCG	TGGGCTCT	LINC01588	0	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
22	T537	GCCCGCGCGGACGCTGGAGCCAGC	CCCGCAA	LINC01588	0.5	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
23	T538	GCTCTCCCGAGCCCAACCGCGCG	GGCGGAC	LINC01588	2	LINC01588_TIDE	AGAGGAGCCTTCTGACTGCTGCAGA	ATGACAGACACAACCAAGGGCA
24	T541	GTCAGATTGGCTTGTCTGGGAATTG	CCAGCGAA	AGA	3	AGA_TIDE1	GGCATGAAGGAATCGAAGGTC	CATGTCCTCAAGTCAAGAACAAAG
25	T544	GCTGGGTGAATGGAGCGAGCAAGC	TCTTCGAG	VEGFA	3	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
26	T545	GTCTGGAGTGACCCCTGACCTTC	TCCCCTCT	VEGFA	7.4	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
27	T546	GATCTCGAGTGACCCCTGGGCTT	CTCCCGGC	VEGFA	6	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
28	T547	GTGTGTCCCTCTCCCAACCGTCC	CTGTCCGG	VEGFA	23.1	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
29	T548	GTGGAGCGGGGAGAGGCCAGGG	GTCACTCC	VEGFA	2	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
30	T549	GCCTTGGAGCGGGGAGAGGCCAG	GGGTCACT	VEGFA	4	VEGF_TIDE3	GTACATGAAGCAACTCCAGTCCCA	ATCAAAATCCAGCACCGAGGCG
31	T550	GTACCCCTCAATAATTGGCTGGC	AATTCCGA	AGA	6	AGA_TIDE1	GGCATGAAGGAATCGAAGGTC	CATGTCCTCAAGTCAAGAACAAAG
32	T551	GATAATTGGCTGGCAATTCGGAG	CAAGCGAA	AGA	4.5	AGA_TIDE1	GGCATGAAGGAATCGAAGGTC	CATGTCCTCAAGTCAAGAACAAAG
33	T558 (DS11)	GCAGGGGCCAGGTGCTTCTCTG	GGGGCTCT	VEGFA	5	VEGF_TIDE4	ACACGGGCGAGCATGGAAATGTC	GTAGGGGTGTGATGGGAGGCTAAGC
34	T559 (DS12)	GAATGCGAGCGGAGGTTGACTG	GGGGCCAG	VEGFA	11.5	VEGF_TIDE3	ECTGTGTGCTTTGCTTTGTGCTG	GTAGGGGTGTGATGGGAGGCTAAGC
35	T560 (DS13)	GAGTGAGAGAGTGAGAGAGAGACA	CGGGCCAG	VEGFA	3	VEGF_TIDE5	ECTGTGTGCTTTGCTTTGTGCTG	GTAGGGGTGTGATGGGAGGCTAAGC
36	T561 (DS14)	GTAGAGGAGGACCTGTGCCACAT	GGGCCCGC	VEGFA	3.5	VEGF_TIDE5	ECTGTGTGCTTTGCTTTGTGCTG	GTAGGGGTGTGATGGGAGGCTAAGC
37	T562 (DS15)	GGTGGGGGCTCCGTGCCACGCG	GGGTCCAT	VEGFA	3.4	VEGF_TIDE6	GGAGGAGAGTAGTCTGCCGAGG	AGACCGAGTGGCAGTGACAGCAAG
38	T563 (DS16)	GCATGGGCGAGGCTGGGGTGCAE	AAGCCCAAG	VEGFA	10	VEGF_TIDE7	AGGAGAGGGGAAGTGGGGGAGG	GTCTTCTGCTGTGCGGCAAGC
39	T564	GAAATGTGATTTCAGATCCAC	AAGCCCA	FANCI	7	FANCI_TIDE5	GTGGGGGCTCTAAGTTATGTAT	CTTCATCTGTATCTCAGGATCA
40	T565	GAGCAGAAAATTGTGATTTC	AGATCCAC	FANCI	0	FANCI_TIDE5	GTGGGGGCTCTAAGTTATGTAT	CTTCATCTGTATCTCAGGATCA

72 horas después de la transfección, las células se recolectaron seguido de la extracción de ADN y la amplificación selectiva del locus objetivo. Se utilizó un análisis de Seguimiento de Indeles mediante descomposición (TIDE) para medir las tasas de indeles en cada locus. Se observó una edición eficiente por Nme2Cas9, a pesar de que las tasas de indel variaron significativamente dependiendo de la secuencia objetivo y el locus. Figura 38A. Además, la afinidad de Nme2Cas9 por los sitios objetivo cerca a/en loci terapéuticamente relevantes tales como *CYBB* (mutaciones causan la enfermedad granulomatosa crónica ligada al cromosoma X) y

AGA (mutaciones causan aspartilglucosaminuria) sugiere que Nme2Cas9 tiene potencial terapéutico. Además, la eficiencia de la edición podría aumentarse al aumentar la cantidad del plásmido Nme2Cas9. Figura 39A. Tomados en conjunto, estos resultados demuestran que Nme2Cas9 puede construirse para editar selectivamente sitios genómicos objetivo específicos en células HEK293T.

Además de las células HEK293T, la eficiencia de la edición de genes de Nme2Cas9 se determinó en varias otras células de mamíferos, incluidas las células de leucemia humana K562, células de osteosarcoma humano U2OS y células de hepatoma hepático Hepa1-6 de ratón. Se creó un constructo lentiviral que expresa Nme2Cas9 y se transdujeron células K562 para expresar Nme2Cas9 de forma estable bajo el control del promotor SFFV. Esta línea celular estable no mostró diferencias significativas con respecto al crecimiento y la morfología en comparación con las células no tratadas, lo que sugiere que Nme2Cas9 no es tóxico cuando se expresa de forma estable. Estas células se electroporaron de forma transitoria con plásmidos que expresaban sgARN dirigidos a varios sitios objetivo y se analizaron después de setenta y dos (72) horas para determinar las tasas indel por TIDE. Se observó una edición eficiente en los tres sitios probados, lo que demuestra la capacidad de Nme2Cas9 para funcionar en las células K562. Para las células Hepa1-6, los plásmidos que codifican Nme2Cas9 y sgARN se cotransfectaron mediante el uso de técnicas similares a la transducción HEK293T descrita anteriormente. Estos datos también muestran que Nme2Cas9 editó eficientemente los sitios *Pcsk9* y *Rosa26* en esta línea celular de ratón. Figura 38B.

El trabajo previo sugiere que la administración de Cas9s por ribonucleoproteína (RNP), en lugar de la transfección de plásmidos, puede ser una opción alternativa para algunas aplicaciones de la edición genómica. Por ejemplo, los efectos con objetivos fuera de lugar de SpyCas9 pueden reducirse significativamente con las electroporaciones RNP en comparación con la entrega de plásmidos. Kim, *Genome Research* 24: 1012-1019 (2014). Para probar si Nme2Cas9 es funcional mediante la administración de RNP, se clonó un Nme2Cas9 marcado con His junto con tres (3) señales de localización nuclear (NLS) y una proteína recombinante purificada en un constructo de expresión bacteriana. Los sgARN dirigidos a varios sitios objetivo validados se generaron mediante la transcripción *in vitro* de T7. La electroporación de un complejo Nme2Cas9: sgARN indujo una edición exitosa en los sitios objetivo, según lo detectado por TIDE. Figura 38C. Estos resultados sugieren que Nme2Cas9 puede administrarse

como un plásmido o como un complejo RNP. En general, estos resultados demuestran que Nme2Cas9 es funcional en varios tipos de células con diferentes modos de administración.

4. Inhibición de proteínas anti-CRISPR

Se ha reportado que cinco (5) familias de proteínas anti-CRISPR (Acr) contra Nme1Cas9 de diversas especies bacterianas inhiben Nme1Cas9 *in vitro* y en células humanas. (Pawluk, 2016, MBio, en prensa). Teniendo en cuenta la alta identidad de secuencia entre Nme1Cas9 y Nme2Cas9, parecía probable que al menos algunas especies dentro de estas familias de Acr también pudieran inhibir Nme2Cas9. Las cinco familias de Acr se expresaron de forma recombinante, se purificaron y se probó la capacidad de Nme2Cas9 para escindir una secuencia objetivo *in vitro* (relación molar 10:1 Acr: Cas9). Como control negativo, se usó un inhibidor para el sistema CRISPR de tipo I-E en *E. coli* (AcrE2). Como se esperaba, todas las familias de Acr inhibieron Nme1Cas9, mientras que AcrE2 no pudo hacerlo. En particular, Acrs IIC1_{Nme}, -IIC2_{Nme}, -IIC3_{Nme} y -IIC4_{Hpa} inhibió la actividad de la edición del gen Nme2Cas9. Figura 40A, parte superior.

Sorprendentemente, AcrIIC5_{Smu} no inhibió Nme2Cas9 *in vitro* incluso con un exceso de 10 veces, lo que sugiere que probablemente inhibe Nme1Cas9 al interactuar con un PID. Para confirmar adicionalmente esto, se realizó el mismo ensayo de escisión *in vitro* mediante el uso de una versión híbrida de NmeCas9 (ej., Nme1Cas9 con el PID de Nme2Cas9). Debido a la actividad reducida de este híbrido, se usó una concentración más alta (~30X) de Cas9 para lograr un perfil de escisión similar mientras se mantenía la relación molar Cas9: Acr 10:1. De acuerdo con los resultados iniciales, no se observó inhibición por AcrIIC5_{Smu} en esta proteína quimérica. Figura 41: La incapacidad de AcrIIC5_{Smu} para inhibir la proteína híbrida sugiere además que AcrIIC5_{Smu} probablemente interactúa con el PID de Nme1Cas9.

Los datos *in vitro* anteriores sugirieron que Acrs -IIC1_{Nme}, -IIC2_{Nme}, -IIC3_{Nme} y -IIC4_{Hpa} podrían usarse como interruptores de inactivación para la edición del genoma Nme2Cas9. Para probar esto, las transfecciones se realizaron como se describió anteriormente en presencia o ausencia de plásmidos que codifican Acrs conducidos por promotores de mamíferos. Aproximadamente se transfectó 150 ng de cada plásmido (ej., con una relación 1:1:1 de sgARN: Cas9: Acr), ya que se ha informado que la mayoría de los ACR inhiben Nme1Cas9 en esas proporciones. (Pawluk, 2016). Como se esperaba del experimento *in vitro*, AcrIIC1_{Nme}, -IIC2_{Nme}, -IIC3_{Nme} y -IIC4_{Hpa} inhibieron la edición del genoma de Nme2Cas9, mientras que

AcrIIC5_{S_{mu}} no pudo hacerlo. (Figura 40B). Además, se observó que la inhibición completa estaba por debajo de los niveles de detección por Acr3Nme y Acr4Hpa, lo que sugiere su alta potencia en comparación con AcrsIIC1Nme y AcrIIC2Nme. Para comparar aún más la potencia de AcrIIC1Nme y AcrIIC4_{Hpa}, se realizaron experimentos en varias relaciones de Acr a Cas9.

- 5 Figura 40C. En consecuencia, AcrIIC4_{Hpa} es un inhibidor altamente potente contra Nme2Cas9, con concentraciones tan bajas como 25 ng:100 ng Acr: Cas9 que inhibe 4 veces Nme2Cas9. Juntos, estos datos sugieren que las proteínas Acr se pueden usar como interruptores de apagado para las aplicaciones basadas en Nme2Cas9.

10 5. **Hiper-precisión de Nme2Cas9**

Los efectos con objetivos fuera de lugar podrían confundir potencialmente las aplicaciones terapéuticas durante la terapia génica humana *ex vivo* e *in vivo* al crear las mutaciones no deseadas. Dado que SpyCas9 de tipo silvestre tiene un número relativamente alto de sitios con objetivos fuera de lugar en células humanas, se han realizado varios esfuerzos para

15 modificar variantes de SpyCas9 de alta fidelidad con éxito variable. Por el contrario, Nme1Cas9 es naturalmente hiperprecisa, lo que demuestra una notable fidelidad en células y modelos de ratón. El trabajo anterior muestra que la cinética de la hibridación, que no está determinada por el PID, puede determinar la fidelidad de una Cas9, lo que sugiere que Nme2Cas9 también puede ser hiperprecisa.

20 Para evaluar empíricamente los perfiles fuera del objetivo NmeCas9, se utilizaron técnicas genéticas de identificación imparcial de rupturas bicatenarias habilitadas por secuenciación (GUIDE-Seq) para determinar posibles sitios con objetivos fuera de lugar de manera imparcial. GUIDE-Seq se basa en la incorporación de oligodesoxinucleótidos bicatenarios (dsODN) en los sitios de corte bicatenarios de ADN en todo el genoma. Estos sitios

25 de escisión se detectan por amplificación y secuenciación de alto rendimiento.

Como punto de referencia para GUIDE-Seq, se utilizó SpyCas9 de tipo silvestre. En particular, SpyCas9 y Nme2Cas9 pudieron clonarse en estructuras principales idénticas impulsadas por el mismo promotor, y se usaron para dirigirse a los mismos sitios debido a sus PAM no superpuestos. Esta técnica permite la comparación lado a lado de las dos nucleasas.

30 Seis (6) sitios duales (DS) fueron dirigidos en *VEGFA* con una secuencia NGGNCCN. Figura 42A. Setenta y dos (72) horas después de la transfección, el análisis TIDE se realizó en los sitios

objetivo. Nme2Cas9 indujo indeles en los seis (6) sitios, aunque con baja eficiencia en dos de ellos, mientras que SpyCas9 indujo indeles en 4/6 sitios. Figura 42B. En dos de esos 4 sitios (DS1 y DS4), SpyCas9 indujo ~7 veces más indeles que Nme2Cas9, mientras que Nme2Cas9 inducida por un aumento de ~3 veces en indeles en DS6. Para GUIDE-seq, se seleccionaron los objetivos DS2, DS4 y DS6 para determinar la escisión con objetivos fuera de lugar en sitios donde Nme2Cas9 es tan eficiente, menos eficiente o más eficiente que SpyCas9, respectivamente.

Además de los tres sitios de doble objetivo, un sitio objetivo de TS6 con una tasa indel del 30-50% (dependiendo del tipo de célula) junto con los genes *Pcsk9* y *Rosa26* de ratón se sometieron al análisis GUIDE-Seq. Se consideró que los perfiles con objetivos fuera de lugar serían más prominentes porque se sabe que el objetivo TS6 experimenta una edición de genes altamente eficiente. Además, las pruebas de los sitios *Pcsk9* y *Rosa26* de ratón revelarían la fidelidad de Nme2Cas9 en una línea celular diferente y los loci candidatos para la edición del genoma *in vivo*. En consecuencia, se realizaron transfecciones para cada Cas9 junto con sus sgARN afines y se prepararon las bibliotecas dsODNs y GUIDE-Seq. El análisis GUIDE-Seq demostró una edición con objetivos en lugar eficiente con ambos ortólogos Cas9 con patrones similares observados por TIDE. Para la identificación con objetivos fuera de lugar, el análisis reveló que, si bien los tres sitios SpyCas9 tenían el alto número esperado de sitios con objetivos fuera de lugar (*ej.*, que oscilaban entre aproximadamente 10 y 1000). Nme2Cas9 tenía un perfil con objetivos fuera de lugar sorprendentemente limpio. Específicamente, Nme2Cas9 dirigido al mismo sitio dual mostró, como máximo, un sitio con objetivos fuera de lugar. Véase, la Figura 42C.

Para validar los sitios con objetivos fuera de lugar detectados por GUIDE-seq, se realizó una secuenciación profunda dirigida para medir la formación de indeles en los loci con objetivos fuera de lugar superiores seguido de la edición independiente de GUIDE-seq (es decir, sin co-transfección del dsODN). Si bien SpyCas9 mostró una edición considerable en la mayoría de los sitios con objetivos fuera de lugar probados (en algunos casos, más eficientes que eso en el sitio con objetivos en lugar correspondiente), Nme2Cas9 no mostró indeles detectables en los únicos sitios con objetivos fuera de lugar candidatos a DS2 y DS6. Con el sgARN *Rosa26*, Nme2Cas9 indujo ~ 1% de la edición en el sitio *Rosa26*-OT1 en las células Hepa1-6, en comparación con ~30% de la edición con objetivos en lugar. Figura 42D.

A continuación, para permitir el uso de SpyCas9 como punto de referencia para GUIDE-seq, debido a que SpyCas9 y Nme2Cas9 tienen PAM no superpuestos, pueden por lo tanto editar potencialmente cualquier sitio dual (DS) flanqueado por una secuencia 5'-NGGNCC-3', que cumple simultáneamente los requisitos de PAM de ambas propiedades de unión de Cas9. Esto permite comparaciones lado a lado de los con objetivos fuera de lugar con sgARNs que se unen exactamente al mismo sitio con objetivos en lugar. Mediante el uso de plásmidos emparejados que expresan cada Cas9 y sus respectivos sgARNs, veintiocho (28) DS se dirigieron a múltiples loci en todo el genoma humano. Setenta y dos (72) horas después de la administración del plásmido, se realizó un análisis TIDE en los sitios dirigidos por cada nucleasa. Nme2Cas9 indujo indeles en diecinueve (19) sitios objetivo, aunque a bajas eficiencias (<5%) en cuatro de ellos, mientras que SpyCas9 indujo indeles en veintitrés (23) de los sitios objetivo, en un caso con <5% de eficiencia. Tres sitios de doble objetivo fueron recalcitrantes a la edición por ambas nucleasas. Si bien SpyCas9 es claramente más eficiente en general, ambas enzimas tienen eficiencias similares en muchos de los sitios, y en dos de los diecisiete sitios que se editaron por ambas nucleasas, Nme2Cas9 fue más eficiente en estas condiciones. Véase, la Figura 42E.

Es de destacar que este sitio con objetivos fuera de lugar tiene un consenso Nme2Cas9 PAM (ACTCCCT) con solo 3 desajustes en el extremo distal PAM de la región complementaria de la guía (es decir, fuera de la semilla). Véase, la Figura 42F. Estos datos apoyan y refuerzan nuestros resultados GUIDE-seq que indican un alto grado de precisión para la edición del genoma Nme2Cas9 en células de mamíferos.

Se compararon los objetivos en lugar vs, objetivos fuera de lugar en estos lados mediante la amplificación dirigida de cada locus seguido del análisis TIDE. Figura 43A. Curiosamente, TIDE no pudo detectar indeles en esos sitios con objetivos fuera de lugar para ninguno de los sgARN, mientras que se observó una edición eficiente con objetivos en lugar. Además, los recuentos de lectura para estos con objetivos fuera de lugar fueron insignificantes en comparación con los observados en el caso de SpyCas9, lo que sugiere que Nme2Cas9 es altamente específico. (Figura 43C, izquierda versus derecha, respectivamente). Para corroborar aún más estos resultados de GUIDE-Seq, CRISPRseek se usó para predecir computacionalmente sitios potenciales con objetivos fuera de lugar para dos de los sgARN más activos con sitios muy similares en el genoma. (Zhu, 2014). Estos se realizaron con PAM N₄CX

y 2-5 desajustes, principalmente en la región distal de PAM. Figura 43D. Tomados en conjunto, estos datos sugieren que Nme2Cas9 es una nucleasa de alta fidelidad en células de mamíferos.

6. Aplicaciones clínicas

En una modalidad, la presente invención contempla un complejo Nme2Cas9 como el primer Cas9 compacto e hiperpreciso con un pequeño PAM no restrictivo para la edición terapéutica del genoma mediante la entrega de AAV. Aunque pequeños, los ortólogos Cas9 hiperprecisos previamente reportados tienen PAM más largos que los descritos en este documento, lo que restringe su uso terapéutico debido a los sitios objetivo limitados en un gen dado (y el perfil con objetivos fuera de lugar en el caso de SauCas9). Esta desventaja se exacerba en los loci donde solo se puede dirigirse a una ventana específica, o se requiere una eliminación precisa del bloque.

La plataforma de entrega de AAV todo en uno establecida en este documento puede usarse para dirigirse a cualquier gen en cualquier tejido. Además, la hiperprecisión de Nme2Cas9 permite la edición precisa de los genes objetivo, por lo tanto, mejora las preocupaciones por la seguridad planteadas debido a actividades con objetivos fuera de lugar previamente observadas. Con este fin, Nme2Cas9 tiene el potencial no solo de complementar las herramientas existentes, sino de convertirse en una opción preferida para la edición terapéutica del genoma mediante la administración viral.

Además, la inhibición de Nme2Cas9 por varios Acrs sugiere una posible presión evolutiva impuesta sobre Cas9 para evolucionar rápidamente un dominio particular. Específicamente, la falta de inhibición de Nme2Cas9 por AcrIIC5_{smu} aumenta la posibilidad de que su mecanismo de inhibición sea a través de un PID. Tras tener en cuenta que AcrIIC5_{smu} es el inhibidor más potente de Nme1Cas9 hasta la fecha, se contempla aquí donde AcrIIC5_{smu} se puede usar para apagar Nme1Cas9 de manera robusta pero no Nme2Cas9. Esto es de particular interés en contextos celulares donde la multiplexación se vería mejorada por la capacidad de controlar un ortólogo específico.

Finalmente, si bien hay miles de ortólogos Cas9 en la base de datos pública, solo unos pocos se han caracterizado. Algunas modalidades contempladas en el presente documento aprovechan la variación natural en los ortólogos Cas9 estrechamente relacionados para crear dos nuevas nucleasas Cas9, a saber, Nme2Cas9 y Nme3Cas9, con PAM N4CC y N4CAA, respectivamente. Los datos presentados en este documento demuestran que incluso los ortólogos

estrechamente relacionados pueden tener propiedades muy diferentes. Por ejemplo, estos ortólogos usan exactamente el mismo sgARN que Nme1Cas9, lo que evita las dificultades en la predicción de tracrARN y determina la longitud del espaciador correcto para cada ortólogo.

Además, es probable que los sgARN más cortos y más estables (tales como las modificaciones químicas) se puedan modificar para expandirse a las tres nucleasas. Estas características pueden facilitar los esfuerzos de edición del genoma y reducir los costos asociados con la ingeniería de proteínas y ARN.

Debería ser evidente para un experto en la técnica que las modalidades descritas en el presente documento no están restringidas a Cas9s y pueden aplicarse a otras proteínas Cas tales como Cas12 y Cas13. También se debe tener en cuenta que la hiper variabilidad de Cas9 no se limita a los PID. Se considera aquí que existen cepas que comparten un alto grado de homología con una Cas9 dada pero que difieren en otros dominios debido a otros tipos de presión selectiva. En conjunto, Nme2Cas9 es una nucleasa novedosa que mejora las plataformas CRISPR actuales para la edición terapéutica del genoma.

V. Plataformas de entrega de nucleótidos

Además de los sistemas de entrega de nucleótidos de AAV descritos anteriormente, la presente invención contempla varios sistemas de entrega compatibles con ácidos nucleicos que proporcionan una distribución más o menos uniforme y tienen tasas de liberación controlables.

Algunas modalidades de la presente invención contemplan los sistemas de entrega de ácido nucleico que codifican los complejos Cas9-sgRNA de Tipo II-C como se describe en el presente documento.

A continuación, se describe una variedad de diferentes medios que son útiles para crear sistemas de entrega de ácido nucleico. No se pretende que ningún medio o vehículo sea limitante para la presente invención. Tenga en cuenta que cualquier medio o vehículo puede combinarse con otro medio o vehículo; por ejemplo, en una modalidad, un vehículo de micropartículas de polímero unido a un compuesto puede combinarse con un medio de gel.

Los vehículos o medios contemplados por esta invención comprenden un material seleccionado del grupo que comprende gelatina, colágeno, ésteres de celulosa, sulfato de dextrano, polisulfato de pentosano, quitina, sacáridos, albúmina, selladores de fibrina, polivinilpirrolidona sintética, óxido de polietileno, óxido de polipropileno, polímeros de bloque

de óxido de polietileno y óxido de polipropileno, polietilenglicol, acrilatos, acrilamidas, metacrilatos que incluyen, pero no se limitan a, metacrilato de 2-hidroxietilo, poli(orto ésteres), cianoacrilatos, bioadhesivos de tipo gelatina-resorcina-aldehído, ácido poliacrílico y copolímeros y bloques de copolímeros de los mismos.

5

Micropartículas

Una modalidad de la presente invención contempla un sistema de entrega de ácido nucleico que comprende una micropartícula. Preferiblemente, las micropartículas comprenden liposomas, nanopartículas, microesferas, nanoesferas, microcápsulas y nanocápsulas.

10 Preferiblemente, algunas micropartículas contempladas por la presente invención comprenden poli(lactida-co-glicólido), poliésteres alifáticos que incluyen, pero no se limitan a, ácido poliglicólico y ácido poliláctico, ácido hialurónico, polisacáridos modificados, quitosano, celulosa, dextrano, poliuretanos, ácidos poliacrílicos, psuedo-poli (aminoácidos), copolímeros relacionados con polihidroxibutrato, polianhídridos, polimetilmetacrilato, poli (óxido de etileno),
15 lecitina y fosfolípidos.

Liposomas

Una modalidad de la presente invención contempla liposomas capaces de unir y liberar ácidos nucleicos como se describe en la presente. Los liposomas son bicapas lipídicas esféricas microscópicas que rodean un núcleo acuoso que están hechas de moléculas anfifílicas tales como
20 los fosfolípidos. Por ejemplo, un liposoma puede atrapar un ácido nucleico entre las colas hidrófobas de la micela de fosfolípidos. Los agentes solubles en agua pueden quedar atrapados en el núcleo y los agentes solubles en lípidos pueden disolverse en la bicapa en forma de concha. Los liposomas tienen una característica especial porque permiten que los productos químicos solubles e insolubles en agua se usen juntos en un medio sin el uso de tensioactivos u otros
25 emulsionantes. Los liposomas pueden formarse espontáneamente al mezclar con fuerza fosfolípidos en medios acuosos. Los compuestos solubles en agua se disuelven en una solución acuosa capaz de hidratar los fosfolípidos. Tras la formación de los liposomas, por lo tanto, estos compuestos quedan atrapados dentro del centro liposómico acuoso. La pared del liposoma, al ser una membrana de fosfolípidos, contiene materiales solubles en grasa como los aceites. Los
30 liposomas proporcionan una liberación controlada de compuestos incorporados. Además, los liposomas pueden recubrirse con polímeros solubles en agua, tales como el polietilenglicol para

aumentar la vida media farmacocinética. Una modalidad de la presente invención contempla una tecnología ultra de alta velocidad para refinar la producción de liposomas, dando como resultado liposomas estables, unilamelares (de una sola capa) que tienen características estructurales específicamente diseñadas. Estas propiedades únicas de los liposomas permiten el

5 almacenamiento simultáneo de compuestos normalmente inmiscibles y la capacidad de su liberación controlada.

En algunas modalidades, la presente invención contempla liposomas catiónicos y aniónicos, así como liposomas que tienen lípidos neutros. Preferiblemente, los liposomas catiónicos comprenden materiales cargados negativamente al mezclar los materiales y los
10 componentes liposomales de ácidos grasos y al permitir que se asocien con la carga. Claramente, la elección de un liposoma catiónico o aniónico depende del pH deseado de la mezcla final de liposomas. Los ejemplos de liposomas catiónicos incluyen lipofectina, lipofectamina y lipofectace.

Una modalidad de la presente invención contempla un sistema de entrega de ácido
15 nucleico que comprende liposomas que proporciona la liberación controlada de al menos un ácido nucleico. Preferiblemente, los liposomas que son capaces de una liberación controlada: i) son biodegradables y no tóxicos; ii) transportan tanto compuestos solubles en agua como en aceite; iii) solubilizan compuestos recalcitrantes; iv) prevén la oxidación del compuesto; v) promueven la estabilización de proteínas; vi) controlan la hidratación; vii) controlan la liberación
20 del compuesto por variaciones en la composición de la bicapa, tales como, pero sin limitarse a, longitud de la cadena de ácidos grasos, composición de lípidos de ácidos grasos, cantidades relativas de ácidos grasos saturados e insaturados y configuración física; viii) tienen dependencia disolvente; iv) tienen dependencia del pH y v) tiene dependencia de la temperatura.

Las composiciones de liposomas se clasifican ampliamente en dos clasificaciones. Los
25 liposomas convencionales son generalmente mezclas de lecitina natural estabilizada (PC) que pueden comprender fosfolípidos sintéticos de cadena idéntica que pueden contener o no glicolípidos.

Los liposomas especiales pueden comprender: i) ácidos grasos bipolares; ii) la capacidad de unir anticuerpos para las terapias dirigidas a tejidos; iii) estar recubiertos con materiales tales como,
30 pero sin limitación, lipoproteínas y carbohidratos; iv) encapsulación múltiple y v) compatibilidad de emulsión.

Los liposomas pueden fabricarse fácilmente en el laboratorio mediante métodos tales como, entre otros, sonicación y vibración. Alternativamente, los liposomas de entrega de compuestos están disponibles comercialmente. Por ejemplo, se sabe que *Collaborative Laboratories, Inc.* fabrica liposomas diseñados a medida para requisitos de entrega específicos.

5 Microesferas, Micropartículas y Microcápsulas

Las microesferas y microcápsulas son útiles debido a su capacidad para mantener una distribución generalmente uniforme, proporcionar una liberación de compuesto controlada estable y son económicas de producir y dispensar. Preferiblemente, un gel de entrega asociado o el gel impregnado de compuesto es transparente o, alternativamente, dicho gel está coloreado para una fácil visualización por parte del personal médico.

Las microesferas se pueden obtener comercialmente (Prolease®, Alkerme's: Cambridge, Mass.). Por ejemplo, un medio liofilizado que comprende al menos un agente terapéutico se homogeneiza en un disolvente adecuado y se pulveriza para fabricar microesferas en el intervalo de 20 a 90 μm . Luego, siguen técnicas que mantienen la integridad de la liberación sostenida durante las fases de purificación, encapsulación y almacenamiento. Scott, *Improving Protein Therapeutics With Sustained Release Formulations, Nature Biotechnology, Volumen 16: 153-157 (1998)*.

La modificación de la composición de microesferas mediante el uso de polímeros biodegradables puede proporcionar la capacidad de controlar la tasa de liberación de ácido nucleico. Miller, *Degradation Rates of Oral Resorbable Implants {Polylactates and Polyglycolates: Rate Modification and Changes in PLA/PGA Copolymer Ratios, J. Biomed. Mater. Res., Vol. II:711-719 (1977)}*.

Alternativamente, se prepara una preparación de microesferas de liberación sostenida o controlada mediante el uso de un método de secado en agua, donde primero se prepara una solución de disolvente orgánico de una sal metálica de polímero biodegradable. Posteriormente, se agrega un medio disuelto o disperso de un ácido nucleico a la solución de sal metálica de polímero biodegradable. La relación en peso de un ácido nucleico a la sal metálica de polímero biodegradable puede ser, por ejemplo, de aproximadamente 1:100000 a aproximadamente 1:1, preferiblemente de aproximadamente 1:20000 a aproximadamente 1:500 y más preferiblemente de aproximadamente 1:10000 a aproximadamente 1:500. A continuación, la solución de disolvente orgánico que contiene la sal metálica de polímero biodegradable y el ácido nucleico se

vierte en una fase acuosa para preparar una emulsión de aceite/agua. El disolvente en la fase oleosa se evapora para proporcionar las microesferas. Finalmente, estas microesferas se recuperan, se lavan y se liofilizan. Posteriormente, las microesferas pueden calentarse a presión reducida para eliminar el agua residual y el disolvente orgánico.

5 Otros métodos útiles en la producción de microesferas que son compatibles con una sal de polímero de metal biodegradable y una mezcla de ácido nucleico son: i) separación de fases durante una adición gradual de un agente coacervante; ii) un método de secado en agua o un método de separación de fases, donde se agrega un antifloculante para evitar la aglomeración de partículas y iii) mediante un método de secado por pulverización.

10 En una modalidad, la presente invención contempla un medio que comprende una microesfera o microcápsula capaz de administrar una liberación controlada de un ácido nucleico para una duración de aproximadamente entre 1 día y 6 meses. En una modalidad, la microesfera o micropartícula puede colorearse para permitir al médico la capacidad de ver el medio claramente a medida que se dispensa. En otra modalidad, la microesfera o microcápsula puede
15 ser transparente. En otra modalidad, la microesfera o micropartícula está impregnada con un tinte fluoroscópico radiopaco.

Las microcápsulas de liberación controlada pueden producirse mediante el uso de técnicas de encapsulación conocidas, tales como extrusión centrífuga, recubrimiento en bombo y suspensión de aire. Tales microesferas y/o microcápsulas pueden modificarse para lograr las
20 tasas de liberación deseadas. Por ejemplo, Oliosphere® (Macromed) es un sistema de microesferas de liberación controlada. Estas microesferas particulares están disponibles en tamaños uniformes que oscilan entre 5 y 500 μm y están compuestas de polímeros biocompatibles y biodegradables. Las composiciones poliméricas específicas de una microesfera pueden controlar la tasa de liberación de ácido nucleico de tal manera que son posibles las
25 microesferas diseñadas a medida, incluido el manejo efectivo del efecto de explosión. ProMaxx® (Epic Therapeutics, Inc.) es un sistema de administración de matriz de proteínas. El sistema es de naturaleza acuosa y se puede adaptar a los modelos de administración farmacéutica estándar. En particular, ProMaxx® son microesferas proteicas bioerosionables que administran tanto fármacos pequeños como macromoleculares, y pueden personalizarse con respecto al
30 tamaño de la microesfera y las características de liberación deseadas.

En una modalidad, una microesfera o micropartícula comprende un material de encapsulación sensible al pH que es estable a un pH menor que el pH del mesenterio interno. El rango típico en el mesenterio interno es de pH 7.6 a pH 7.2. En consecuencia, las microcápsulas deben mantenerse a un pH inferior a 7. Sin embargo, si se espera una variabilidad del pH, el

5 material sensible al pH se puede seleccionar en función de los diferentes criterios de pH necesarios para la disolución de las microcápsulas. El ácido nucleico encapsulado, por lo tanto, se seleccionará para el entorno de pH en el que se desea la disolución y se almacenará en un pH preseleccionado para mantener la estabilidad. Ejemplos de material sensible al pH útil como encapsulantes son Eudragit® L-100 o S-100 (Rohm GMBH), ftalato de

10 hidroxipropilmetilcelulosa, succinato de acetato de hidroxipropilmetilcelulosa, ftalato de acetato de polivinilo, ftalato de acetato de celulosa y trimelitato de acetato de celulosa. En una modalidad, los lípidos comprenden el recubrimiento interno de las microcápsulas. En estas composiciones, estos lípidos pueden ser, pero no se limitan a, ésteres parciales de ácidos grasos y anhídridos de hexitol, y grasas comestibles tales como triglicéridos. Lew C. W., *Controlled-*

15 *Release pH Sensitive Capsule And Adhesive System And Method*. Patente de los Estados Unidos N° 5.364.634 (incorporada aquí como referencia).

En una modalidad, la presente invención contempla una micropartícula que comprende una gelatina u otro catión polimérico que tiene una densidad de carga similar a la gelatina (es decir, poli-L-lisina) y se usa como un complejo para formar una micropartícula primaria. Se

20 produce una micropartícula primaria como una mezcla de la siguiente composición: i) Gelatina (60 floraciones, tipo A de piel porcina), ii) condroitina 4-sulfato (0.005% - 0.1%), iii) glutaraldehído (25%, grado 1), y iv) clorhidrato de 1-etil-3- (3-dimetilaminopropil)-carbodiimida (clorhidrato de EDC), y sacarosa ultra pura (Sigma Chemical Co., St. Louis, MO). No se cree que la fuente de gelatina sea crítica; puede ser de origen bovino, porcino, humano u otro animal.

25 Típicamente, el catión polimérico está entre 19,000-30,000 daltons. Luego se agrega sulfato de condroitina al complejo con sulfato de sodio o etanol como agente de coacervación.

Después de la formación de una micropartícula, un ácido nucleico se une directamente a la superficie de la micropartícula o se une indirectamente mediante el uso de un "puente" o "espaciador". Los grupos amino de los grupos de gelatina lisina se derivatizan fácilmente para

30 proporcionar sitios para el acoplamiento directo de un compuesto. Alternativamente, los espaciadores (es decir, moléculas de unión y fracciones de derivación en ligandos con objetivos

en lugar) tales como avidina-biotina también son útiles para acoplar indirectamente ligandos de direccionamiento a las micropartículas. La estabilidad de la micropartícula está controlada por la cantidad de reticulación del separador de glutaraldehído inducida por el clorhidrato de EDC. Un medio de liberación controlada también está determinado empíricamente por la densidad final de los enlaces cruzados de espaciador de glutaraldehído.

En una modalidad, la presente invención contempla micropartículas formadas mediante secado por pulverización de una composición que comprende fibrinógeno o trombina con un ácido nucleico. Preferiblemente, estas micropartículas son solubles y la proteína seleccionada (es decir, fibrinógeno o trombina) crea las paredes de las micropartículas. En consecuencia, los ácidos nucleicos se incorporan dentro y entre las paredes proteicas de la micropartícula. Heath, *Microparticles And Their Use In Wound Therapy*. Patente de los Estados Unidos N° 6.113.948 (incorporada aquí como referencia). Después de la aplicación de las micropartículas a los tejidos vivos, la reacción posterior entre el fibrinógeno y la trombina crea un sellador de tejidos que libera el compuesto incorporado en el área circundante inmediata.

Un experto en la técnica comprenderá que la forma de las microesferas no necesita ser exactamente esférica; solo como partículas muy pequeñas capaces de pulverizarse o extenderse dentro o sobre un sitio quirúrgico (es decir, abierto o cerrado). En una modalidad, las micropartículas están compuestas de un material biocompatible y/o biodegradable seleccionado del grupo que consiste en polilactida, poliglicólido y copolímeros de lactida/glicólido (PLGA), ácido hialurónico, polisacáridos modificados y cualquier otro material bien conocido.

Sección Experimental

Ejemplo I

Construcción del plásmido de vector sgARN-Nme1Cas9-AAV todo en uno

El gen bacteriano Nme1Cas9 se ha optimizado con codones para la expresión en humanos, y se ha clonado en un plásmido AAV2 bajo el promotor ubicuo U1a. El ARN guía está bajo el promotor U6. El gen *cas9* contiene cuatro señales de localización nuclear y tres secuencias de etiqueta HA en tándem. Se insertaron secuencias espaciadoras en el casete de crARN al digerir el plásmido con la enzima de restricción SapI mediante el uso de

oligonucleótidos sintéticos recocidos para generar un dúplex con salientes compatibles con los generados por la estructura principal digerida con SapI.

El gen Nme1Cas9 optimizado con codón humano bajo el control del promotor U1a y un casete de sgARN dirigido por el promotor U6 se clonaron en una estructura principal de plásmido AAV2. El ORF NmeCas9 estaba flanqueado por cuatro señales de localización nuclear, dos en cada terminal, además de una etiqueta de epítipo de triple HA. Este plásmido está disponible a través de Addgene (plásmido ID 112139). Véase, la Figura 64. Los oligonucleótidos con secuencias espaciadoras dirigidas a *Hpd*, *Pcsk9* y *Rosa26* se insertaron en el casete de sgARN mediante ligadura en un sitio de clonación SapI.

La producción de vectores AAV se realizó en el Centro de Terapia Genética Horae de la Facultad de Medicina de la Universidad de Massachusetts. Brevemente, los plásmidos se empaquetaron en cápsides AAV8 por transfección de plásmidos triples en células HEK293 y se purificaron por sedimentación como se describió previamente. Gao, “*Introducing genes into mammalian cells: viral vectors*” en: Green MR, Sambrook J, editores. *Molecular cloning: a laboratory manual*. Volumen 2. 4^{ta} ed. New York: Cold Spring Harbor Laboratory Press; 2012. p. 1209–13. Los perfiles con objetivos fuera de lugar de estos espaciadores se predijeron computacionalmente mediante el uso del paquete CRISPRseek de Bioconductor. Los parámetros de búsqueda se adaptaron a la configuración de Nme1Cas9: gARN.tamaño = 24, PAM= “NNNNGATT,” PAM.- tamaño = 8, RNA.PAM.patrn = “NNNNGNNN\$,” pesos = c(0, 0, 0, 0, 0, 0, 0.014, 0, 0, 0.395, 0.317, 0, 0.389, 0.079, 0.445, 0.508, 0.613, 0.851, 0.732, 0.828, 0.615, 0.804, 0.685, 0.583), max.desajuste = 6, desajuste.permitido.PAM= 7, topN = 10,000, puntaje mínimo = 0.

Ejemplo II

Cultivo celular y transfección

Se cultivaron células de hepatoma de ratón Hepa1-6 en DMEM con FBS al 10% y penicilina/estreptomicina (Gibco) al 1% en una incubadora a 37° C con CO₂ al 5%. Se cultivaron células HEK293T humanas y células PLB985 en medios DMEM y RPMI respectivamente. Ambos se complementaron con 10% de FBS y 1% de penicilina/estreptomicina (Gibco). Las transfecciones transitorias de las células Hepa 1-6 se realizaron al usar Lipofectamine LTX mientras que el reactivo de transfección Polyfect (Qiagen) se usó para las células HEK293T.

Para la transfección transitoria, se cultivaron aproximadamente 1×10^5 células por pozo en una placa de 24 pocillos 24 horas antes de la transfección. Cada pocillo se transfectó con plásmidos sgARN-Nme1Cas9-AAV todo en uno de 500 ng, al usar Lipofectamine LTX con reactivo Plus (ThermoFisher) de acuerdo con el protocolo del fabricante. Las células HEK293T se

5 transfectaron con 400 ng de plásmido todo en uno que expresa Nme1Cas9 y sgARN en una placa de 24 pozos de acuerdo con las directrices del fabricante (ej., *Psck9* y *Rosa26*).

Todas las líneas celulares se mantuvieron en una incubadora a 37 °C con CO₂ al 5%. Se cultivaron hepatoma de ratón Hepa1-6 y células HEK293T en DMEM con FBS al 10% y penicilina/estreptomicina al 1% (Gibco). Las células K562 se cultivaron en las mismas
10 condiciones, pero usando IMDM. Las células IMR-90 se cultivaron en EMEM y FBS al 10%. Finalmente, las células HDFa se cultivaron en DMEM y FBS al 20%.

Ejemplo III

Expresión y purificación de Nme1Cas9

15 Nme1Cas9 se clonó en un vector pMCSG7 que contenía un promotor T7 seguido de etiqueta His 6X y luego un sitio de escisión de la proteasa del virus del grabado del tabaco (TEV). Este constructo se transformó en la cepa Rosetta2 DE3 de E. coli y se expresó Nme1Cas9. Brevemente, se cultivó las bacterias a 37 °C hasta alcanzar OD₆₀₀ de 0,6. En este punto, la temperatura se redujo a 18 °C seguido de la adición de 1 mM de β-D-1-
20 tiogalactopiranosido de isopropilo (IPTG). Las células se cultivaron durante la noche y luego se recolectaron para la purificación.

La purificación de Nme1Cas9 se realizó en tres pasos: Cromatografía de afinidad de níquel, cromatografía de intercambio catiónico y luego cromatografía de exclusión por tamaño. Los protocolos detallados para estos se pueden encontrar en publicaciones anteriores (Jinek,
25 Science 337, 816-821, 2012).

Ejemplo IV

Administración de ribonucleoproteína (RNP) de Nme1Cas9

La entrega de RNP de Nme1Cas9 se realizó al utilizar el sistema de transfección de neón
30 (ThermoFisher). Aproximadamente 20 picomoles de Nme1Cas9 y 25 picomoles de sgARN se mezclaron en buffer R y se incubaron a temperatura ambiente durante 20-30 minutos. Este

complejo premontado se mezcló luego con 50,000-100,000 células y se sometió a electroporación usando puntas de neón de 10 µL. Después de la electroporación, las células se sembraron en placas de 24 pozos que contenían los medios de cultivo apropiados sin antibióticos.

5

Ejemplo V

Aislamiento del ADN de células y tejidos

El ADN genómico se aisló 72 horas después de la transfección de las células a través del kit DNeasy® de Sangre y Tejido (Qiagen) de acuerdo con el protocolo del fabricante. Se sacrificaron los ratones y se recolectó el tejido hepático 10 días después de la inyección hidrodinámica o 50 días después de la administración del vector por la vena de la cola, y el ADN genómico se aisló con un kit DNeasy® de Sangre y Tejido (Qiagen) de acuerdo con el protocolo del fabricante.

15

Ejemplo VI

Análisis Indel

Se usaron 50 ng de ADN genómico para la amplificación por PCR con cebadores genómicos específicos a sitio y High Fidelity® 2X PCR Master Mix (New England Biolabs). Para el análisis TIDE, se purificaron 30 µl de producto de PCR mediante el uso del kit de purificación de PCR QIAquick® (Qiagen) y se sometieron a secuenciación Sanger. Los valores indel se obtuvieron al usar la herramienta web TIDE (tide-calculator.nki.nl/) como se describió anteriormente. Brinkman, *Nucl. Acids Res.* (2014).

20

Para el ensayo T7 de Endonucleasa I (T7EI), se hibridaron 10 µl del producto de PCR y se trataron con 0,5 µl de T7 Endonucleasa I (New England Biolabs) en 1X NEB Buffer 2 durante 1 hora. Las muestras se procesaron en un gel de agarosa al 2,5% y se cuantificaron con el programa ImageMaster-TotalLab®. Los porcentajes Indel se calcularon como se describió anteriormente. Guschin, *Engineered Zinc Finger Proteins: Methods and Protocols* (2010).

25

Ejemplo VII

GUIDE-Seq para el análisis con objetivos fuera de lugar

30

El análisis GUIDE-seq se realizó como se describió anteriormente. Tsai, *Nature Biotechnology* (2014), Bolukbasi, *Nature Methods* (2015a); Amrani, [biorxiv.org/content/early/2017/08/04/172650](https://doi.org/10.1101/172650) (2017).

5 Brevemente, las células Hepa1-6 se transfectaron con 500 ng de plásmidos sgARN-Nme1Cas9-AAV todo en uno y 7,5 pmol de oligonucleótido GUIDE-seq recocido mediante el uso de Lipofectamina LTX® con reactivo Plus® (ThermoFisher), para los dos espaciadores dirigidos a genes *Pcsk9* y *Rosa26*. El ADN genómico se extrajo con un kit DNeasy® de Sangre y Tejido (Qiagen) 72 horas después de la transfección siguiendo el protocolo del fabricante. Los preparativos de la biblioteca, la secuenciación profunda y el análisis de lecturas se realizaron
10 como se describió anteriormente. Tsai, *Nature Biotechnology* (2014), Bolukbasi, *Nature Methods* (2015a); Amrani, [biorxiv.org/content/early/2017/08/04/172650](https://doi.org/10.1101/172650) (2017).

Ejemplo IX

Producción de vectores AAV

15 Los plásmidos se empaquetaron en AAV8 por transfección de plásmidos triples en células HEK 293 y se purificaron por sedimentación como se describió previamente en el Centro de Terapia Genética Horae en la Facultad de Medicina de la Universidad de Massachusetts. Gao GP, Sena-Esteves M. *Introducing Genes into Mammalian Cells: Viral Vectors*. In: Green MR, Sambrook J, eds. *Molecular Cloning, Volumen 2: A Laboratory Manual*. New York: Cold Spring
20 Harbor Laboratory Press; 2012:1209–1313.

Ejemplo X

Animales, inyecciones de vectores AAV y procesamiento de tejido hepático

25 Todos los experimentos con animales fueron aprobados bajo las pautas del Comité Institucional de Cuidado y Uso de Animales de la Facultad de Medicina de la Universidad de Massachusetts. Para las inyecciones hidrodinámicas, se inyectaron 2,5 ml de 30 µg de plásmido sgARN-Nme1Cas9-AAV libre de endotoxina dirigido a *Pcsk9*, o PBS como control, a través de la vena de la cola en ratones hembra C57BL/6 de 9-18 semanas de edad. Para las inyecciones del vector AAV8, se inyectaron ratones hembra C57BL/6 de 9-18 semanas de edad con 4×10^{11}
30 copias del genoma por ratón a través de la vena de la cola. Se usaron ratones hembra C57BL/6NJ de 8 semanas de edad para experimentos de edición del genoma *in vivo*. Para los

experimentos *ex vivo*, los embriones que avanzaron a la etapa de dos células se transfirieron al oviducto de ratones hembra pseudopreñadas E0.5.

Los ratones se sacrificaron por CO₂ y se recolectó el hígado. Los tejidos se fijaron en paraformaldehído al 4% durante la noche, y se embebieron en parafina, se seccionaron y se

5 tiñeron con hematoxilina y eosina (H&E).

Ejemplo XI

Análisis de suero

Se extrajo sangre (~ 200 µl) de la vena facial a los 0, 25 y 50 días después de la

10 administración del vector. El suero se aisló mediante el uso de un separador de suero (BD, Cat. 365967) y se almacenó a - 80 °C hasta el ensayo. Los niveles de colesterol en suero se midieron mediante el ensayo de punto final colorimétrico Infinity™ (Thermo-Scientific) siguiendo el protocolo del fabricante. Brevemente, se prepararon diluciones en serie del Calibrador de

15 Química Data-Cal™ en PBS. En una placa de 96 pozos, se mezclaron 2 µl de suero de ratón o dilución de calibrador con 200 µl de reactivo líquido de colesterol Infinity™, luego se incubaron a 37 °C durante 5 min. La absorbancia se midió a 500 nm mediante el uso de un lector de microplacas BioTek Synergy® HT.

Ejemplo XII

Descubrimiento de ortólogos Cas9 con PID hiper-evolucionados

20

La secuencia Nme1Cas9 se destruyó para encontrar todos los ortólogos Cas9 en especies de *Neisseria*. Se seleccionaron ortólogos con > 80% de identidad para Nme1Cas9 para el resto de este análisis. Luego se alinearon los PID de cada uno al usar ClustalW2 con el de Nme1Cas9 (desde el aminoácido 820 hasta el 1082) y se seleccionaron aquellos con grupos de mutaciones

25 en el PID.

La secuencia del péptido Nme1Cas9 se usó como una consulta en búsquedas BLAST para encontrar todos los ortólogos Cas9 en cepas de *Neisseria meningitidis*. Los ortólogos con > 80% de identidad para Nme1Cas9 se seleccionaron para el estudio. Los PID se alinearon luego con los de Nme1Cas9 (residuos 820-1082) al usar ClustalW2® y aquellos con grupos de

30 mutaciones en el PID se seleccionaron para un análisis posterior. Se construyó un árbol

filogenético no enraizado de ortólogos NmeCas9 al usar FigTree (tree.bio.ed.ac.uk/software/figtree).

Ejemplo XIII

5 Clonación y purificación de los ortólogos de Acr y Cas9, Nme2 y Nme3

Los PID de Nme2Cas9 y Nme3Cas9 se ordenaron como gBlocks (IDT) para reemplazar el PID de Nme1Cas9 mediante el uso de Gibson Assembly (NEB) en un plásmido de expresión bacteriano pMSCG7 con etiqueta de His 6X. El constructo se transformó en *E. coli*, se expresó y se purificó como se describió previamente.

10 Brevemente, las células de Rosetta (DE3) que contienen los plásmidos Cas9 respectivos se cultivaron a 37 °C hasta una densidad óptica de 0,6 y la expresión de proteínas se indujo mediante IPTG 1 mM durante 16 horas a 18 °C. Las células se recolectaron y se lisaron mediante sonicación en tampón de lisis (Tris 50 mM pH 7,5, NaCl 500 mM, imidazol 5 mM, DTT 1 mM) suplementado con cóctel inhibidor de la lisozima y la proteasa (Sigma).

15 El lisado se pasó luego a través de una columna de agarosa Ni-NTA (Qiagen), la proteína unida se eluyó con imidazol 300 mM y se dializó en tampón de almacenamiento (HEPES 20 mM pH 7,5, NaCl 250 mM, DTT 1 mM). Para las proteínas Acr, se expresaron proteínas marcadas con His 6x en la cepa BL21 de *E. coli* Rosetta (DE3). Las células se cultivaron a 37 °C hasta una densidad óptica (OD_{600 nm}) de 0,6 en una incubadora con agitación. Los cultivos bacterianos se
20 enfriaron a 18 °C y se indujo la expresión de proteínas mediante la adición de IPTG 1 mM para la expresión durante la noche. Al día siguiente, las células se recolectaron y se volvieron a suspender en tampón de lisis (Tris 50 mM pH 7,5, NaCl 500 mM, imidazol 5 mM, DTT 1 mM) suplementado con 1 mg/ml de lisozima y cóctel de inhibidor de proteasa (Sigma) y la proteína se purificó mediante el uso del mismo protocolo que para Cas9. La etiqueta His 6X se eliminó por
25 incubación con la proteasa del virus del grabado del tabaco (TEV) durante la noche a 4 °C para aislar Acrs sin etiquetar y escindido con éxito.

Ejemplo IVX

Ensayo de descubrimiento de PAM *in vitro*

Se generó una biblioteca de protoespaciadores con secuencias de PAM aleatorizadas mediante el uso de PCR superpuestas, con el cebador directo que contiene el PAM aleatorizado de 10 nucleótidos.

La biblioteca se purificó en gel y se sometió a reacción de escisión *in vitro* mediante Cas9
5 purificado junto con sgARN transcritos *in vitro*. Se utilizó el complejo Cas9: sgARN 300 nM para escindir 300 nM del fragmento objetivo en 1X NE Buffer 3.1 (NEB) a 37 °C durante 1 hora. La reacción se trató luego con proteinasa K a 50 °C durante 10 minutos y se ejecutó en un gel de agarosa al 4% con 1X TAE. El producto de escisión se purificó y se sometió a preparación de biblioteca. La biblioteca se secuenció al usar la plataforma de secuenciación Illumina
10 NextSeq500® y se analizó. Los logs de secuencia se generaron al usar R.

Ejemplo XV

Transfecciones y edición del genoma de mamíferos

El Nme2Cas9 humanizado se clonó en el plásmido pCDest2 usado previamente para la
15 expresión de Nme1Cas9 y SpyCas9 mediante el uso de Gibson Assembly. La transfección de las células HEK293T y HEK293T-TLR se realizó como se describió previamente (Amrani. 2018). Para las transfecciones de Hepa1-6, se usó Lipofectamine LTX para transfectar 500 ng de plásmido AAV.sgRNA.Nme2Cas9 todo en uno en aproximadamente 1×10^5 células por pozo que se habían cultivado en placas de 24 pozos 24 horas antes de la transfección. Para las células
20 K562 que expresan Nme2Cas9 de forma estable, se electroporaron 50,000 - 150,000 células con plásmido sgARN 500 mediante el uso de puntas de neón de 10 μ L.

Para medir los indeles en todas las células, 72 horas después de las transfecciones, se recolectaron células y se extrajo el ADN genómico mediante el uso del kit DNaesy® de Sangre y Tejido (Qiagen). El locus dirigido se amplificó por PCR, secuenciado por Sanger (Genewiz®) y
25 analizado por TIDE (Brinkman. 2014).

Ejemplo XVI

Transducción lentiviral de células K562 para expresar Nme2Cas9 de forma estable

Las células K562 que expresan Nme2Cas9 de forma estable se generaron como se
30 describió previamente. Para la producción del lentivirus, el vector lentiviral se cotransfectó en células HEK293T junto con los plásmidos de empaquetamiento (Addgene 12260 y 12259) en

placas de 6 pozos mediante el uso del reactivo de transfección TransIT-LT1 (Mirus Bio) según lo recomendado por el fabricante. Después de 24 horas, el medio se aspiró de las células transfectadas y se reemplazó con 1 ml de medio fresco de DMEM.

Al día siguiente, se recolectó el sobrenadante que contenía el virus de las células transfectadas y se filtró a través de un filtro de 0,45 μm . Se usaron 10 μL del sobrenadante sin diluir junto con 2,5 μg de Polibreno para transducir ~ 1 millón de células K562 en placas de 6 pozos. Las células transducidas se seleccionaron al usar 2,5 $\mu\text{g/mL}$ de medio que contiene puromicina.

Ejemplo XVII

Entrega de RNP para la edición del genoma de mamíferos

Para los experimentos de RNP, se usó un sistema de electroporación de neón. Se ensamblaron 40 picomoles de 3X NLS Nme2Cas9 junto con 50 picomoles de sgARN transcrito *in vitro* en buffer R, y se electroporaron al usar puntas de neón de 10 μL . Después de la electroporación, las células se colocaron en placas precalentadas de 24 pozos que contenían los medios de cultivo apropiados sin antibióticos. Los parámetros de electroporación (voltaje, ancho, número de pulsos) fueron 1150 v, 20 ms, 2 pulsos para células HEK293T; 1000 v, 50 ms, 1 pulso para células K562.

Ejemplo XVIII

GUIDE-Seq

Los experimentos GUIDE-Seq se realizaron como se describió anteriormente con modificaciones menores (Amrani, 2018).

Brevemente, las células HEK293T se transfectaron con 200 ng de Cas9, 200 ng de sgARN y 7,5 pmol de oligonucleótido de GUIDE-seq recocido mediante el uso de Polyfect (Qiagen) para guías dirigidas a sitios duales con SpyCas9 o Nme2Cas9. Las células Hepa1-6 se transfectaron como se describió anteriormente.

El ADN genómico se extrajo con un kit DNeasy® de Sangre y Tejido (Qiagen) 72 horas después de la transfección según el protocolo del fabricante. La preparación de la biblioteca y la secuenciación se realizaron exactamente como se describió anteriormente.

Para el análisis, los sitios que se emparejaron con una secuencia con diez desajustes con el sitio objetivo se consideraron sitios con objetivos fuera de lugar potenciales. Los datos se analizaron mediante el uso del paquete Bioconductor GUIDEseq versión 1.1.17 (Zhu, 2017).

5

Ejemplo XIX

Secuenciación profunda dirigida y análisis

La secuenciación profunda dirigida se usó para confirmar los resultados de GUIDE-Seq y medir más cuantitativamente las tasas de indel. Se usó una amplificación por PCR de dos pasos para producir fragmentos de ADN para cada sitio con objetivos en lugar y objetivos fuera de lugar. Para SpyCas9, se seleccionaron las principales ubicaciones con objetivos fuera de lugar.

10

En el primer paso, los cebadores específicos de locus con salientes universales con extremos complementarios a los adaptadores se mezclaron con la mezcla maestra de PCR 2X Phusion® (NEB) para generar fragmentos que llevan los salientes. En el segundo paso, los productos de PCR purificados se amplificaron con un cebador directo universal y cebadores inversos indexados.

15

Los productos de tamaño completo (~ 250 pb de longitud) se extrajeron en gel y se secuenciaron al usar una ejecución MiSeq de extremo emparejado. El análisis de datos MiSeq se realizó exactamente como se describió anteriormente (Amrani 2018).

20

Ejemplo XX

Análisis con objetivos fuera de lugar mediante el uso de CRISPRseek

Los análisis globales con objetivos fuera de lugar para TS25 y TS47 se realizaron al usar el paquete CRISPRseek de Bioconductor.

Se hicieron cambios menores para acomodar las características de Nme2Cas9 no compartidas con SpyCas9. Específicamente, se utilizaron los siguientes cambios: gARN.tamaño = 24, PAM = "NNNNCC", PAM.tamaño = 6, ARN.PAM.patcón= "NNNNCN", se recolectaron sitios con objetivos fuera de lugar con menos de 6 desajustes. Se seleccionaron los principales sitios potenciales con objetivos fuera de lugar en función del número y la posición de los desajustes. Se usó ADNg de células dirigidas por cada sgARN respectivo para amplificar cada locus con objetivos fuera de lugar y se analizó por TIDE.

30

Ejemplo XXI

Entrega de AAV8.Nme2Cas9 in vivo y procesamiento de tejido hepático

Todos los procedimientos con animales fueron revisados y aprobados por el Comité Institucional de Cuidado y Uso de Animales (IACUC) en la Facultad de Medicina de la
5 Universidad de Massachusetts.

Para las inyecciones del vector AAV8, se inyectaron ratones hembra C57BL/6 de 8 semanas de edad con 4×10^{11} copias del genoma por ratón a través de la vena de la cola dirigidas a *Pcsk9* o *Rosa26*. Se sacrificaron los ratones 28 días después de la administración del vector y se recolectaron los tejidos del hígado para su análisis. Los tejidos hepáticos se fijaron en
10 formalina al 4% durante la noche, y se embebieron en parafina, se seccionaron y se tiñeron con hematoxilina y eosina (H&E). Se extrajo sangre de la vena facial a los 0, 14 y 28 días después de la inyección, y se aisló el suero mediante el uso de un separador de suero (BD, Cat. No. 365967) y se almacenó a -80°C hasta el ensayo. El nivel de colesterol en suero se midió al usar el ensayo de punto final colorimétrico Infinity™ (Thermo-Scientific) siguiendo el protocolo del
15 fabricante y como se describió anteriormente (Ibraheim, 2018).

Ejemplo XXII

Procesamiento de animales y tejido hepático

Para las inyecciones hidrodinámicas, se inyectaron 2,5 ml de 30 μg de plásmido AAV-sgRNA-hNme1Cas9 libre de endotoxina dirigido a *Pcsk9* o 2,5 ml de PBS en una vena de la cola
20 en ratones hembra C57BL/6 de 9 a 18 semanas de edad. Los ratones se sacrificaron 10 días después y se recolectó el tejido hepático. Para las inyecciones del vector AAV8, se inyectaron ratones hembra C57BL/6 de 12 a 16 semanas de edad con 4×10^{11} copias del genoma por ratón a través de la vena de la cola, mediante el uso de vectores dirigidos a *Pcsk9* o *Rosa26*. Se
25 sacrificaron los ratones 14 y 50 días después de la administración del vector y se recolectaron los tejidos del hígado para su análisis.

Para el direccionamiento de Hpd, se administraron 2 ml de PBS o 2 ml de 30 μg de plásmido AAV-sgARN-hNme1Cas9 sin endotoxina en ratones knockout *Fah* (Fahneo) de tirosinemia de tipo 1 de 15 a 21 semanas de edad. Los sgARN codificados se dirigieron a sitios
30 en el exón 8 (sgHpd1) o el exón 11 (sgHpd2). Los ratones HT1 se alimentaron con 10 mg/L de NTBC (2- (2-nitro-4-trifluorometilbenzoil) -1,3-ciclohexanodiona) (Sigma-Aldrich, Cat. No.

PHR1731-1G) en agua potable cuando se indique. Ambos sexos fueron utilizados en estos experimentos. Los ratones se mantuvieron en agua NTBC durante siete días después de la inyección y luego se cambiaron a agua normal. El peso corporal se controló cada 1-3 días. Los ratones de control inyectados con PBS se sacrificaron cuando quedaron moribundos después de perder el 20% de su peso corporal después de la eliminación del tratamiento con NTBC.

Los ratones se sacrificaron de acuerdo con nuestro protocolo y se cortó el tejido hepático y se almacenaron fragmentos a -80 °C. Algunos tejidos hepáticos se fijaron en formalina al 4% durante la noche, se embebieron en parafina, se seccionaron y se tiñeron con hematoxilina y eosina (H&E).

XXIII

Transferencia Western Blot

Las fracciones de tejido hepático se molieron y se resuspendieron en 150 µl de buffer de lisis RIPA. El contenido total de las proteínas se estimó mediante el kit de ensayo de proteínas BCA Pierce TM (Thermo- Scientific) siguiendo el protocolo del fabricante. Se cargaron un total de 20 µg de proteína de tejido o 2 ng de proproteína recombinante proteasa convertasa 9 / PCSK9 de ratón (R&D Systems, 9258-SE-020) en un gel prefabricado 4-20% Mini--Rotean® TGX TM (Bio -Rad). Las bandas separadas se transfirieron a una membrana de PVDF y se bloquearon con una solución de Blocking-Grade Blocker al 5% (Bio-Rad) durante 2 horas a temperatura ambiente. Las membranas se incubaron con anticuerpos anti-GAPDH de conejo (Abcam ab9485, 1:2000) o anticuerpos anti-PCSK9 de cabra (R&D Systems AF3985, 1:400) durante la noche a 4 °C. Las membranas se lavaron cinco veces en TBST y se incubaron con anti-conejo de cabra conjugados con peroxidasa de rábano picante (HRP) (Bio-Rad 1,706,515, 1:4000) y anticuerpos secundarios anti-cabra de burro (R&D Systems HAF109, 1: 2000) durante 2 horas a temperatura ambiente. Las membranas se lavaron cinco veces en TBST y se visualizaron con sustrato Clarity TM western ECL (Bio-Rad) mediante el uso de un procesador M35A X-OMAT (Kodak).

Ejemplo XXIV

Respuesta inmune humoral

La respuesta inmune de IgG1 humoral a Nme1Cas9 se midió mediante ELISA (Bethyl; Mouse IgG1 ELISA Kit, E99-105) siguiendo el protocolo del fabricante con algunas modificaciones. Brevemente, se realizó la expresión y purificación en tres pasos de Nme1Cas9 y SpyCas9. Se utilizaron un total de 0,5 µg de proteínas recombinantes Nme1Cas9 o SpyCas9 suspendidas en 1X de tampón de recubrimiento (Bethyl) para recubrir placas de 96 pozos (Corning) y se incubaron durante 12 horas a 4 °C con agitación. Los pozos se lavaron tres veces mientras se agitaban durante 5 minutos al usar 1X Wash Buffer. Las placas se bloquearon con solución de bloqueo 1X BSA (Bethyl) durante 2 horas a temperatura ambiente, luego se lavaron tres veces. Las muestras de suero se diluyeron 1:40 mediante el uso de PBS y se añadieron a cada pozo por duplicado. Después de incubar las muestras a 4 °C durante 5 h, las placas se lavaron 3 veces durante 5 min y se agregó 100 µL de anticuerpo IgG1 anti-ratón biotinilado (Bethyl; 1: 100,000 en solución de bloqueo 1xBSA) a cada pozo. Después de incubar durante 1 hora a temperatura ambiente, las placas se lavaron cuatro veces y se añadieron 100 µl de sustrato TMB a cada pozo. Las placas se dejaron desarrollarse en la oscuridad durante 20 minutos a temperatura ambiente y luego se agregó 100 µL de solución de parada ELISA por pozo. Después del desarrollo de la solución amarilla, se registró la absorbancia a 450 nm mediante el uso de un lector de microplacas BioTek Synergy® HT.

Ejemplo XXV

Incubación y transfección del cigoto

Cepas de ratones y recolección de embriones

Todos los experimentos con animales se realizaron bajo la guía del Comité Institucional de Cuidado y Uso de Animales (IACUC) de la Facultad de Medicina de la Universidad de Massachusetts. Se obtuvieron ratones C57BL/6NJ (código 005304) del Laboratorio Jackson. Todos los animales se mantuvieron en un ciclo de luz de 12 h. La mitad del ciclo de luz del día en que se observó un tapón de acoplamiento se consideró el día embrionario 0.5 (E0.5) de gestación. Los cigotos se recolectaron en E0.5 rasgando la ampolla con pinzas e incubación en medio M2 que contiene hialuronidasa para eliminar las células del cúmulo.

Entrega de AAV8.Nme2Cas9+ sgARN *in vivo* y procesamiento de tejido hepático

Para las inyecciones del vector AAV8, se inyectó ratones hembra C57BL/6NJ de 8 semanas de edad con 4×10^{11} copias del genoma por ratón a través de la vena de la cola, con el sgARN dirigido a un sitio validado en *Pcsk9* o *Rosa26*. Se sacrificaron los ratones 28 días después de la administración del vector y se recolectaron los tejidos del hígado para su análisis. Los tejidos hepáticos se fijaron en formalina al 4% durante la noche, se embebieron en parafina, se seccionaron y se tiñeron con hematoxilina y eosina (H&E). Se extrajo sangre de la vena facial a los 0, 14 y 28 días después de la inyección, y se aisló el suero mediante el uso de un separador de suero (BD, Cat. 365967) y se almacenó a -80°C hasta el ensayo. El nivel de colesterol en suero se midió al usar el ensayo de punto final colorimétrico Infinity™ (Thermo-Scientific) siguiendo el protocolo del fabricante y como se describió anteriormente. Ibraheim, “*All-in-One Adeno-associated Virus Delivery and Genome Editing by Neisseria meningitidis Cas9 in vivo*” *Genome Biology* 19:137 (2018).

Para una transferencia Western blot anti-PCSK9, se cargaron 40 µg de proteína de tejido o 2 ng de proteína PCSK9 de ratón recombinante (R&D Systems, 9258-SE-020) en un gel prefabricado MiniProtean® TGX™ (Bio-Rad). Las bandas separadas se transfirieron a una membrana de PVDF y se bloquearon con una solución de Blocking-Grade Blocker® al 5% (Bio-Rad) durante 2 horas a temperatura ambiente. A continuación, la membrana se incubó con anticuerpos anti-GAPDH de conejo (Abcam ab9485, 1:2,000) o anti-PCSK9 de cabra (R&D Systems AF3985, 1: 400) durante la noche. Las membranas se lavaron en TBST y se incubaron con anti-conejo de cabra conjugados con peroxidasa de rábano picante (HRP) (Bio-Rad 1706515, 1: 4,000) y anticuerpos secundarios anti-cabra de burro (R&D Systems HAF109, 1: 2,000) durante 2 horas a temperatura ambiente. Las membranas se lavaron nuevamente en TBST y se visualizaron mediante el uso del sustrato Clarity™ western ECL (Bio-Rad) al usar un procesador M35A XOMAT (Kodak).

Entrega de AAV6.Nme2Cas9 *ex vivo* en cigotos de ratón

Los cigotos se incubaron en gotas de 15 µl de KSOM (Medio Optimizado Simplex Complementado con Potasio, Millipore, Cat. No. MR-106-D) que contiene 3×10^9 o 3×10^8 GCs del vector AAV6.Nme2Cas9.sgTyr durante 5-6 h (4 cigotos en cada gota). Después de la incubación, los cigotos se enjuagaron en M2 y se transfirieron a KSOM fresco para cultivo

durante la noche. Al día siguiente, los embriones que avanzaron a la etapa de 2 células se transfirieron al oviducto de receptores hembra pseudopreñadas y se les permitió desarrollarse a término.

5

Ejemplo XXVI

Cuantificación y análisis estadísticos

Se realizó un análisis de los datos de descubrimiento PAM *in vitro* mediante el uso de R. GraphPad Prism 6® para todos los análisis estadísticos. Para los experimentos con células de mamíferos que usan Nme2Cas9, se realizaron 3 réplicas independientes y se calcularon los porcentajes indel al usar el software TIDE, con barras de error que representan la s.e.m. Los parámetros TIDE se establecieron para cuantificar indeles < 20 nucleótidos para todas las figuras. Para las comparaciones lado a lado de Nme2Cas9 y SpyCas9, se calcularon porcentajes indel promedio mediante el uso de Microsoft Excel. Para experimentos *in vivo* en ratones, n = 5 para sujetos de control y prueba. Los valores de P se calcularon mediante la prueba t de dos colas no apareados.

10

15