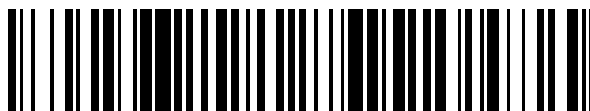


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 732 100**

51 Int. Cl.:

C12N 15/10 (2006.01)

C12N 15/63 (2006.01)

C12N 15/11 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.04.2016** **PCT/EP2016/059346**

87 Fecha y número de publicación internacional: **03.11.2016** **WO16174056**

96 Fecha de presentación y número de la solicitud europea: **27.04.2016** **E 16722555 (6)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019** **EP 3289081**

54 Título: **Composiciones y métodos para el tratamiento de trastornos por expansión de repeticiones de nucleótidos**

30 Prioridad:

27.04.2015 EP 15305646

28.04.2015 US 201562153633 P

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

20.11.2019

73 Titular/es:

GENETHON (33.3%)

1 bis rue de l'Internationale

91000 Evry, FR;

INSERM (INSTITUT NATIONAL DE LA SANTÉ ET

DE LA RECHERCHE MÉDICALE) (33.3%) y

UNIVERSITE D'EVRY VAL D'ESSONNE (33.3%)

72 Inventor/es:

BUJ BELLO, ANA MARIA y

LO SCRUDATO, MIRELLA

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 732 100 T3

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

DESCRIPCIÓN

Composiciones y métodos para el tratamiento de trastornos por expansión de repeticiones de nucleótidos

Campo de la invención

5 La presente invención se refiere a composiciones y métodos para el tratamiento de trastornos por expansión de repeticiones de nucleótidos tales como la distrofia miotónica.

Antecedentes de la invención

10 Las expansiones de repeticiones de nucleótidos, especialmente las expansiones de repeticiones de trinucleótidos, están implicadas en más de dos docenas de trastornos neurológicos y del desarrollo. Un enfoque que se ha propuesto para tratar estas enfermedades es acortar las repeticiones a longitudes no patológicas utilizando nucleasas altamente específicas (véase, para una revisión, Richard G.F., Trends Genet. 31(4):177-186, abril de 2015).

15 Se han utilizado en tales estrategias nucleasas altamente específicas, tales como meganucleasas, y las nucleasas ZFN, TALEN y CRISPR-Cas9. Sin embargo, estas últimas fueron consideradas por el experto en la materia como inapropiadas para la escisión de expansiones de repeticiones de trinucleótidos (véase Richard, citado *supra*). Globalmente, los TALEN se han considerado una herramienta más prometedora para el acortamiento de las repeticiones de nucleótidos.

Contra este fuerte prejuicio los presentes inventores demuestran que el sistema CRISPR-Cas9 puede implementarse para escindir expansiones de repeticiones de nucleótidos del ADN genómico, proporcionando de esta manera una potente e inesperada herramienta para el tratamiento de los trastornos por expansión de repeticiones de nucleótidos.

Compendio de la invención

20 Los inventores han demostrado que, frente al fuerte prejuicio desarrollado anteriormente, el sistema CRISPR-Cas9 puede resultar eficiente para el tratamiento de los trastornos por expansión de repeticiones de nucleótidos.

25 En un aspecto, en la presente memoria se describen moléculas de ARN guía monocatenario (ARNgm) útiles para escindir específicamente una expansión de repeticiones de nucleótidos, especialmente una expansión de repeticiones de trinucleótidos, de una región no codificante de un gen de interés. Las moléculas de ARNgm descritas en la presente memoria son capaces de unir mediante apareamiento de bases un complemento de secuencia a una secuencia (protoespaciador) diana de ADN genómico que se encuentra situada 5' o 3' respecto a la expansión de nucleótidos diana y son capaces de atraer una endonucleasa Cas9 a, o a un sitio próximo a, el sitio de hibridación entre el ARNgm y el ADN genómico. Las moléculas de ARNgm de la invención comprenden la totalidad de los elementos de secuencia apropiados para inducir roturas en la doble cadena mediadas por Cas9 en la vecindad del sitio de complementariedad.

30 En particular, la presente solicitud describe parejas de ARNgm apropiadas para realizar una escisión de la expansión de repeticiones de nucleótidos presente en la región no codificante del gen de interés, en el que la pareja de ARNgm comprende un primer ARNgm que es complementario a una secuencia de ADN genómico situada 5' respecto a la expansión de repeticiones de nucleótidos y un segundo ARNgm que es complementario a una secuencia de ADN genómico situado 3' respecto a la expansión de repeticiones de nucleótidos.

35 En una realización particular, la molécula de ARNgm comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO:1 a NO:4. En otra realización particular, la molécula de ARNgm presenta la secuencia mostrada en cualquiera de las SEQ ID NO:5 a NO:8.

40 Otro aspecto descrito en la presente memoria es la utilización del sistema CRISPR-Cas9 para escindir una expansión de repeticiones de nucleótidos. En particular, la expansión de repeticiones de nucleótidos se encuentra dentro del ADN genómico de una célula diana, más particularmente dentro de una región no codificante de un gen presente en dicho ADN genómico.

45 Según un aspecto adicional, en la presente memoria se describe un método para escindir una expansión de repeticiones de nucleótidos respecto de una región no codificante de un gen en el ADN genómico de una célula, implementando dicho método el sistema CRISPR-Cas9. El método puede comprender la introducción en la célula de una pareja de moléculas de ARNgm y un gen codificante de una endonucleasa Cas9.

En otro aspecto, en la presente memoria se describe un método para el tratamiento de un trastorno por expansión de repeticiones de nucleótidos, en el que dicha repetición se escinde respecto de un gen de interés, que comprende tal expansión de repeticiones de nucleótidos, utilizando al menos una pareja de moléculas de ARNgm tal como se ha descrito anteriormente y una endonucleasa Cas9.

50 Más específicamente, los usos y métodos de la invención pueden comprender la introducción en una célula, tal como una célula de un sujeto que lo necesita:

(i) Una primera molécula de ARNgm,

- (ii) Una segunda molécula de ARNgm, y
- (iii) Una endonucleasa CRISPR/Cas9,

en donde dichos primer y segundo ARNgm son complementarios a una secuencia situada 5' y 3' respecto a dicha expansión de repeticiones de nucleótidos, respectivamente, resultando apropiados de esta manera para escindir dicha expansión de repeticiones de nucleótidos.

En una realización particular, la expansión de repeticiones de nucleótidos se encuentra dentro de un intrón o una región 3' no traducida del gen de interés. En una realización más particular, la expansión de repeticiones de nucleótidos se encuentra dentro de la región 3' no traducida del gen de interés.

En una realización particular adicional, el gen de interés es FMR1, AFF2 o FMR2, AFF3, FXN, ATXN80S/ATXN8, ATXN10, PPP2R2B, BEAN1/TK2, NOP56, C9ORF72, ZN9/CNBP o DMPK. En una realización particular, el gen de interés es DMPK. En una variante de esta realización, la expansión de repeticiones de nucleótidos se encuentra en la región 3' no traducida del gen codificante DMPK.

En algunas otras realizaciones específicas, el trastorno por expansión de repeticiones de nucleótidos es la distrofia miotónica, en particular la distrofia miotónica de tipo 1 (DM1) o de tipo 2 (DM2), más particularmente DM1.

En algunas realizaciones, la molécula de ARNgm comprende una secuencia de ARN guía que presenta 15 a 40 nucleótidos, en particular 20 a 30 nucleótidos, en particular 22 a 26 nucleótidos, tal como 22, 23, 24, 25, 26 nucleótidos y una secuencia de ARN de andamiaje. En una realización particular, la molécula de ARNgm comprende una secuencia de ARN guía que consiste en 24 nucleótidos.

En algunas realizaciones, las moléculas de ARNgm están diseñadas para unir mediante apareamiento de bases el complemento a la secuencia diana de ADN genómico (de otro modo denominada secuencia diana). Esta secuencia diana se denomina protoespaciador y se encuentra situada contiguamente a un motivo de nucleótidos denominado PAM (por sus siglas en inglés, motivo contiguo a protoespaciador) que resulta reconocido por la endonucleasa Cas9 implementada.

En algunas realizaciones, la endonucleasa Cas9 y/o las moléculas de ARNgm se expresan a partir de uno o varios vectores, tales como uno o varios plásmidos o vectores víricos. Por ejemplo, la endonucleasa Cas9 puede expresarse a partir de un primer vector, y la primera y segunda moléculas de ARNgm puede expresarse a partir de un único segundo vector o a de un segundo vector y la otra molécula a partir de un tercer vector. En otra realización, todos los elementos necesarios para la implementación del sistema CRISPR-Cas9 están contenidos en un único vector.

En un aspecto adicional, en la presente memoria se describe un kit que comprende una endonucleasa Cas9 y una primera y segunda moléculas de ARNgm tal como se ha indicado anteriormente. En otro aspecto, en la presente memoria se describe un kit que comprende un vector codificante de una endonucleasa Cas9 y un vector que codifica la primera y/o la segunda moléculas de ARNgm tal como se ha indicado anteriormente, o un único vector que expresa la endonucleasa Cas9 y una o ambas moléculas de ARNgm. Tal como se ha mencionado anteriormente, el vector o vectores en el kit pueden ser un vector plásmido o un vector vírico. Además, el kit según la invención puede incluir cualquier reactivo adicional (tal como uno o más tampones y/o uno o más reactivos de transfección) o dispositivos útiles en la implementación de los métodos y usos descritos en la presente memoria.

Otros aspectos y realizaciones de la invención resultarán evidentes en la siguiente descripción detallada.

Leyendas de las figuras

Figura 1. Casetes de expresión de Cas9 y ARNgm. A y B: Cas9 de *Neisseria meningitidis* (NMCas9) se clonó bajo un promotor sintético específico muscular (C5-12) o un promotor ubicuo (EFS: EF-1 α más corto). Int: intrón quimérico; NLS: señal de localización nuclear; HA: epítipo hemaglutinina de influenza humana; 3XFLAG: 3 epítipos FLAG en tándem; SVpolyA: señal de poliadenilación del virus 40 del simio. C y D: casete de expresión para dos ARNgm, ambos bajo el control del promotor U6 pero que contiene diferentes sitios de clonación para el protoespaciador de ARNgm, *BbsI* y *BspMI*. Una secuencia codificante de GFP nuclear (AcGFP) bajo el promotor desmina (Des) también se encuentra presente en el constructo D.

Figura 2. Detección de NMCas9 en líneas celulares. Análisis de transferencia western de NMCas9 (PM: 130,6 kDa) expresado en células HeLa (carriles 1 a 4) y C2C12 (carriles 5 a 7) transfectadas con plásmido de control (-); carril 1: pC512-Int-smSV40; carriles 2 y 5: pEFS-Int-SVpolyA)) o con un plásmido que aloja NMCas9 (carriles 3 y 6: pC512-Int-NMCas9-smSVpolyA, ver la fig. 1 constructo A; carriles 4 y 7: pEFS-Int-NMCas9-SVpolyA, ver la fig. 1 constructo B). La detección específica de la proteína se llevó a cabo con un anticuerpo dirigido contra los 3 epítipos FLAG en tándem situados en el extremo C-terminal de la proteína. HeLa: línea celular de carcinoma epitelial humano; C2C12: línea celular de mioblastos de ratón; C5-12: promotor específico muscular sintético; EFS: promotor ubicuo más corto EF-1 α .

Figura 3. Región genómica circundante a las repeticiones de CTG del gen DMPK. La secuencia de las dianas de ADN genómico (protoespaciadores) 1, 7, E y N está subrayada; los PAM (motivos contiguos a protoespaciador) respectivos

están circundados por un rectángulo. La repetición de CTG se destaca en negrita. Las posiciones de los cebadores utilizados para amplificar por PCR esta región también están subrayados. Cromosoma: 19; cadena: menos; posición: 46.272.967(inicio) – 46.273.898(final); referencia genómica humana: feb.2009 (GRCh37/hg19).

Figura 4. Detección de la región genómica que contiene la delección de CTG. La región 3'UTR de DMPK se amplificó por PCR a partir de ADNg extraído de células HeLa o HEK 293T (no mostradas) con los cebadores F1- y R1-DMPK-3'UTR (mostrados en la fig. 3). Las células han sido transfectadas únicamente con EFS-NMCas9 (Cas9) o cotransfectadas con EFS-NMCas9 y los ARNgm indicados (Cas9 + ARNgm no específicos, β' -> δ' ; Cas9 + ARNgm, β' -> δ') y se recolectaron 48 horas después. Los productos de PCR de longitud completa y los que contiene la delección de CTG se indican con una flecha verde y una flecha roja, respectivamente. Los tamaños esperados de los productos de PCR se muestran en la Tabla 2. Los ARNgm no específicos presentan como diana una secuencia desplazada cuatro nucleótidos en comparación con la que secuencia diana del ARNgm correspondiente (N₂₄ del ARNgm no específico corresponde a N₂₀ del ARNgm).

Figura 5. La delección de la expansión de repeticiones de CTG de células primarias procedentes de pacientes de DM1. Se transfectaron células primarias fetales procedentes de un individuo de control (Ctrl) y de un paciente de DM1 que alojaba 200 repeticiones de CTG (DM200), con un plásmido de control de GFP, carril 1, o se cotransfectaron con los ARNgm y (constructo pAAV-Des-AcGFP-U6sgRNA-NM-7E-DMPK) y EFS-NMCas9 (constructo pEFS-Int-NMCas9-SVpolyA), carriles 2 y 3. Se recolectaron las células dos días después de la transfección y se extrajo el ADNg a partir de las células separadas como positivas para GFP. La región 3'UTR de DMPK se amplificó por PCR tal como en la fig. 4. Los productos de PCR de longitud completa y parcial y los productos de PCR que contienen las delecciones de CTG se indican con una flecha verde, una flecha azul y una flecha roja, respectivamente. Los tamaños esperados de productos de PCR se muestran en la Tabla 2.

Figura 6. Secuenciación de la región genómica que contiene la delección de CTG. El ADNg extraído a partir de las células HEK 293T transfectadas con EFS-NMCas9 (constructo pEFS-Int-NMCas9-SVpolyA) y las parejas de ARNgm indicadas (β' :1 y N; δ' : 7 y N) se amplificaron por PCR y se subclonaron en un plásmido. Los clones individuales se secuenciaron mediante secuenciación estándar (Beckman Coulter Genomics) y se alinearon sus secuencias con la secuencia original de tipo silvestre para mostrar las posiciones de corte. La mayoría de los clones presentaban un corte entre el 3° y 4° nucleótido de las dianas contiguas a la secuencia de PAM (indicadas mediante las flechas negras).

Descripción detallada de la invención

Los inventores en la presente memoria muestran que el sistema CRISPR-Cas9 puede utilizarse eficientemente para escindir repeticiones de nucleótidos del ADN genómico, proporcionando de esta manera una potente herramienta para el tratamiento de los trastornos por expansión de repeticiones de nucleótidos tal como la distrofia miotónica.

Por consiguiente, en un primer aspecto se describe en la presente memoria:

- (i) Una molécula de ARNgm que es capaz de unir mediante apareamiento de bases el complemento de secuencia a una secuencia diana (protoespaciador) en el ADN genómico, que está situado 5' respecto a una repetición de nucleótidos localizada dentro de una región no codificante de un gen de interés,
- (ii) Una molécula de ARNgm que es capaz de unirse mediante apareamiento de bases al complemento de secuencia a una secuencia diana (protoespaciador) en el ADN genómico que está situado 3' respecto a una repetición de nucleótidos localizada dentro de una región no codificante de un gen de interés, y
- (iii) Una pareja de moléculas de ARNgm cada una de las cuales es capaz de unirse mediante apareamiento de bases secuencias que complementan las secuencias diana en el ADN genómico que están situadas 5' y 3', respectivamente, respecto a una repetición de nucleótidos situada dentro de una región no codificante de un gen de interés.

Sistema CRISPR-Cas9

El sistema de repeticiones palindrómicas cortas agrupadas y espaciadas regularmente (CRISPR, por sus siglas en inglés) de tipo II es una tecnología de endonucleasa guiada por ARN que ha emergido recientemente como prometedora herramienta de edición genómica. Existen dos componentes diferenciados en este sistema: (1) un ARN guía y (2) una endonucleasa, en este caso la nucleasa asociada a CRISPR (Cas, por sus siglas en inglés), Cas9. El ARN guía es una combinación de ARN de CRISPR bacteriano (ARNcr) y un ARNcr trans-activador (ARNtracr) construido en un ARN guía monocatenario (ARNgm) quimérico (Jinek et al., Science 337(6096):816-21, 7 de agosto de 2012). El ARNgm combina la especificidad de reconocimiento del ARNcr con las propiedades de andamiaje del ARNtracr en un único transcrito. Al expresarse el ARNgm y Cas9 en la célula, la secuencia diana genómica puede modificarse o interrumpirse permanentemente.

El complejo ARNgm/Cas9 es atraído a la secuencia diana por el apareamiento de bases entre la secuencia guía ARNgm y el complemento de la secuencia diana en el ADN genómico (protoespaciador). Para la unión con éxito de Cas9, la secuencia diana genómica también debe contener la secuencia de motivo contiguo a protoespaciador (PAM) correcta inmediatamente después de la secuencia diana. La unión del complejo ARNgm/Cas9 localiza Cas9 en la secuencia diana genómica de manera que la endonucleasa Cas9 pueda cortar ambas cadenas de ADN, provocando

una rotura de doble cadena (DSB, por sus siglas en inglés). Cas9 cortará 3-4 nucleótidos cadena arriba de la secuencia de PAM. De acuerdo con el sistema implementado en la presente invención, el DSB a continuación puede ser reparada mediante la ruta de reparación de unión de extremos no homólogos (NHEJ, por sus siglas en inglés).

- 5 La presente invención implementa este potente sistema de una manera innovadora, a fin de escindir secuencias de repeticiones que se ha informado que están asociadas a varias enfermedades o trastornos.

Endonucleasa Cas9

Los mecanismos de reconocimiento de ADN del sistema CRISPR-Cas de tipo II implican un ARN guía que dirige la endonucleasa Cas9 para cortar el ADN diana de una manera específica de secuencia, dependiente de la presencia de un motivo contiguo a protoespaciador (PAM) en el ADN diana.

- 10 La secuencia de PAM varía dependiendo de la especie de bacteria a partir de la que se ha obtenido la endonucleasa Cas9. Entre las secuencias de PAM de consenso se incluyen PAM tanto primarias como secundarias, las cuales puede considerarse que identifican secuencias de interés dentro del gen diana de interés. La tabla a continuación proporciona una lista de secuencias de PAM para algunas endonucleasas Cas9 obtenidas de diferentes especies.

Especie	Secuencias de PAM
<i>S. pyogenes</i>	NGG, NAG
<i>S. thermophilus</i>	NNRGAA, NNAGGA, NNANAA, NNGGGA, NGGNG, NNAAAAW
<i>N. meningitidis</i>	NNNNGATT, NNNNGNTT, NNNNGANT, NNNNGTTN, NNNNGANN, NNNNGNNT, NNNNGTNN, NNNNGNTN
<i>S. mutans</i>	NGG
<i>C. jejuni</i>	NNNNACA
<i>P. multocida</i>	GNNNCNNA
<i>F. novicida</i>	NG

- 15 En una realización particular de la invención, la endonucleasa Cas9 utilizada en la invención se deriva de *S. pyogenes*, *S. thermophilus*, *N. meningitidis*, *S. mutans*, *C. jejuni*, *F. novicida*, *S. aureus*, *P. multocida*, *P. bettyae*, *H. parainfluenzae*, *H. pittmaniae* o *L. crispatus*. En una realización específica, en la que la expansión de repeticiones de nucleótidos para escindir se encuentra presente en la región 3' no traducida del gen DMPK, la endonucleasa Cas9 se obtiene de *N. meningitidis*, *S. aureus*, *S. thermophilus* o *C. jejuni* (Zhang et al., Mol. Cell 50(4):488-503, 23 de mayo de 2013; Hou et al., 24 de sept. de 2013, 110(39):15644-9).

ARN guía

En la presente memoria se describen ARN guía monocatenarios (o ARNgm) que están diseñados específicamente para la escisión de una expansión de repeticiones de trinucleótidos.

- 25 Tal como se ha mencionado anteriormente, el ARNgm es la parte del sistema CRISPR-Cas9 que proporciona especificidad de reconocimiento de ADN genómico al sistema. La secuencia de ADN genómico diana comprende 15 a 40 nucleótidos, en particular 20 a 30 nucleótidos, en particular 22 a 26 nucleótidos, tal como 22, 23, 24, 25, 26 nucleótidos, dependiendo de la endonucleasa Cas9 específica utilizada en el sistema, seguido de un motivo contiguo a protoespaciador (PAM) tal como se descrito anteriormente. En una realización particular, la molécula de ARNgm comprende una secuencia de ARN guía que es complementaria a la secuencia del complemento de una secuencia genómica de 15 a 40 nucleótidos, en particular de 20 a 30 nucleótidos, en particular de 22 a 26 nucleótidos, tal como 22, 23, 24, 25, 26 nucleótidos, más específicamente de 24 nucleótidos, que precede a un PAM dentro de la región no codificante diana del gen de interés. En una realización particular, la secuencia de ARN guía es idéntica, o al menos 80% idéntica, preferentemente al menos 85%, 90%, 95%, 96%, 97%, 98% o al menos 99% idéntica a dicha secuencia genómica y es capaz de hibridarse con la secuencia de complemento de dicha secuencia genómica de 15 a 40 nucleótidos, en particular de 20 a 30 nucleótidos, en particular de 22 a 26 nucleótidos, tal como 22, 23, 24, 25, 26 nucleótidos, más específicamente a 24 nucleótidos que preceden a PAM dentro de la región no codificante diana del gen de interés. Tal como es conocido por el experto en la materia, el ARNgm no contiene el motivo PAM y, en consecuencia, no une el complemento de secuencia a PAM. La secuencia diana puede encontrarse en cualquiera de las cadenas del ADN genómico, dentro de una región no codificante de un gen de interés.

En una realización particular, el PAM y al menos 4 nucleótidos cadena arriba de PAM se encuentran en la región no codificante del gen de interés. En una realización particular adicional, la secuencia diana entera y el PAM se encuentran en la región no codificante del gen de interés.

Se encuentran disponibles herramientas bioinformáticas para identificar las secuencias de ADN genómico diana que comprenden el PAM o los PAM apropiados, dependiendo del origen de la endonucleasa Cas9 utilizado en la práctica de la invención, y la secuencia de hibridación, tales como las proporcionadas mediante las herramientas de internet siguientes: CRISPR Design (<http://crispr.mit.edu>), E-CRISP (<http://www.e-crisp.org/E-CRISP/designcrispr.html>), CasFinder (<http://arep.med.harvard.edu/CasFinder/>), y CRISPOR (<http://tefor.net/crispor/crispor.cgi>). El experto en la materia también puede hacer referencia a Doench et al., Nat. Biotechnol. 32(12):1262-7, dic. de 2014, o Prykhodzhiy et al., PLoS One 10(3):e0119372, 5 de marzo de 2015, y pueden encontrar más información y recursos sobre el sistema CRISPR-Cas9 y sobre la identificación de ADN genómico diana que comprende el PAM o los PAM apropiados en el sitio de internet siguiente: <http://www.cnbc.csic.es/~montoliu/CRISPR/>. Las secuencias de PAM pueden identificarse alternativamente mediante la utilización de dicha secuencia como pregunta en herramientas de alineación de secuencias, tales como el algoritmo BLAST o FASTA, dentro de un gen de interés.

Tal como es bien conocido, un ARNgm es una fusión de un ARNcr y un ARNtracr que proporciona tanto especificidad de reconocimiento de diana (proporcionada por el apareamiento de bases de la secuencia guía con la secuencia del complemento de la secuencia diana de ADN genómico) como andamiaje/capacidad de unión a una nucleasa Cas9. En una realización, la fracción ARNtracr y la endonucleasa Cas9 seleccionada se derivan del mismo subgrupo filogénico. Se describe un árbol filogénico de ortólogos de Cas9 representativos en particular en Fonfara et al., Nucleic Acids Res. 42(4):2577-90, feb. de 2014. A título de realización ilustrativa, pueden citarse componentes derivados de un sistema CRISPR-Cas de tipo II bacteriano II-A, II-B o II-C. Por ejemplo, el ARNtracr puede obtenerse de *N. meningitidis*, *C. jejuni* o *P. multocida* y la endonucleasa Cas9 implementada se deriva de *N. meningitidis*. En una realización específica, la fracción de ARNtracr y la endonucleasa Cas9 seleccionada se corresponden, en el sentido de que ambas se derivan de la misma especie para funcionar juntas. Por ejemplo, el ARNtracr y la endonucleasa Cas9 se derivan ambas de *S. pyogenes*, *S. thermophilus*, *N. meningitidis*, *S. mutans*, *C. jejuni*, *F. novicida*, *S. aureus*, *P. multocida*, *P. bettyae*, *H. parainfluenzae*, *H. pittmaniae* o *L. crispatus* en una realización particular de la invención. Más particularmente, tanto la fracción de ARNtracr como la endonucleasa Cas9 se obtienen de *N. meningitidis*.

Se encuentran disponibles kits y herramientas de biología molecular, tales como los plásmidos apropiados, para producir fácilmente un ARNgm de la especificidad deseada en términos de tanto la secuencia diana de ADN genómico como de la endonucleasa Cas9. Por ejemplo, se encuentran disponibles varios plásmidos y herramientas de Addgene. En una realización particular, el ARNgm o la pareja de ARNgm se expresan a partir de un plásmido bajo el control de un promotor U6. En una realización particular, ambos ARNgm de la pareja de ARNgm de la invención se expresan a partir de un único casete de expresión que contiene los dos andamiajes de ARNgm, cada uno bajo el control de un promotor, en particular el promotor U6, en el mismo vector (por ejemplo, en el mismo plásmido o en el mismo genoma vírico recombinante, tal como en un genoma de VAA). En una realización particular, los dos andamiajes de ARNgm se proporcionan en posición inversa o en tándem, en particular en tándem. En otra realización, el gen codificante de la endonucleasa Cas9 se liga operablemente a un promotor, tal como un promotor inducible o constitutivo, en particular un promotor ubicuo o específico de tejido, en particular un promotor específico muscular. Entre los promotores ubicuos se incluyen, por ejemplo, el promotor EFS, CMV o CAG. Entre los promotores específicos musculares se incluyen, aunque sin limitarse a ellos, el promotor creatín-quinasa muscular (MCK, por sus siglas en inglés), el promotor desmina o el promotor C5.12 sintético, tal como es bien conocido de la técnica. Además, el promotor utilizado para la expresión de la endonucleasa Cas9 puede ser un promotor inducible, tal como un promotor inducible por tetraciclina, por tamoxifeno o por ecdisona.

En una realización particular, cada uno de la primera y segunda moléculas de ARNgm es complementario a una región situada 5' y 3' respecto a la expansión de repeticiones de nucleótidos que debe extraerse, respectivamente. De esta manera, las moléculas de ARNgm se diseñan para unirse específicamente a regiones situadas cadena arriba y cadena debajo de la expansión de repeticiones de nucleótidos, encontrándose el PAM, y al menos 4 nucleótidos situados cadena arriba del PAM, en la región no codificante del gen de interés y preferentemente en donde la secuencia diana entera y el PAM se encuentran dentro de una región no codificante del gen de interés. La distancia de la secuencia diana (región de homología + PAM) de la región escindida puede no ser crítica, aunque a fin de minimizar la desestabilización de la estructura génica, la secuencia diana puede seleccionarse de manera que se encuentre a menos de 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20 o menos de 10 nucleótidos del extremo considerado de la expansión de repeticiones de nucleótidos. Por ejemplo, considerando el ARNgm que se diseña para dirigir la inducción de un DSB situado 5' respecto a la expansión de repeticiones de nucleótidos, el nucleótido más 3' del PAM de la secuencia diana se encuentra a menos de 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20 o a menos de 10 nucleótidos del nucleótido más 5' del primer (en dirección 5' a 3') nucleótido de la expansión de repeticiones de nucleótidos que debe escindirse. Además, considerando el ARNgm que está diseñado para dirigir la inducción de un DSB 3' respecto a la expansión de repeticiones de nucleótidos, el nucleótido más 5' de la región de homología de la secuencia diana se encuentra a menos de 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20 o a menos de 10 nucleótidos del nucleótido más 3' del último (en dirección 5' a 3') nucleótido de la expansión de repeticiones de nucleótidos que debe escindirse. Preferentemente, en la práctica de la presente invención, al diseñar el ARNgm o los ARNgm descritos en la presente memoria, el experto en la materia evitará la escisión de elementos reguladores

conocidos presentes en la secuencia de ADN genómico no codificante en torno a la expansión de repeticiones de nucleótidos que debe escindirse, tal como regiones para el procesamiento/estabilidad del ARN, tal como poliA, regiones de corte y empalme en intrones, etc.

5 En una realización particular, las moléculas de ARNgm están diseñadas para escindir una expansión de repeticiones de trinucleótidos situada dentro de la región 3' no traducida del gen DMPK. En una variante particular de la presente realización, la invención se refiere a una molécula de ARNgm que comprende una secuencia guía seleccionada del grupo que consiste en GCGCUGCCUGAACCCUAGAACUGU (SEQ ID NO:1), ACGGGGCUUGAAGGGUCCUUGUAG (SEQ ID NO:2), UGGGAGCGUCUGGCGCGAUCUCU (SEQ ID NO:3) y GUCGGGGUCUCAGUGCAUCCAAAA (SEQ ID NO:4).

10 La invención se refiere además a un vector tal como se ha definido anteriormente, que comprende una secuencia codificante de una molécula de ARNgm que comprende una secuencia guía seleccionada del grupo que consiste en SEQ ID NO:1 a NO:4.

15 En otra realización particular, la pareja de moléculas de ARNgm es una pareja de un ARNgm que comprende una secuencia guía seleccionada de SEQ ID NO:1 y SEQ ID NO:2 y de un ARNgm que comprende una secuencia guía seleccionada de SEQ ID NO:3 y SEQ ID NO:4.

En una realización particular adicional, el ARNgm utilizado para inducir DSB cadena arriba (o 5') de la región de repeticiones de trinucleótidos se selecciona del grupo que consiste en:

- SEQ ID NO:5 (también denominada ARNgm 1 en la parte experimental, secuencia de unión a ADN, es decir, la secuencia guía, subrayada):

20 GCGCUGCCUGAACCCUAGAACUGUUGUAGCUGCCCUUUCGAAAGAACCGUUG
CUACAAUAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAGCU
UCUGCUUUAAGGGGCAUCGUUUUUUUUUUU

- SEQ ID NO:6 (también denominada ARNgm 7 en la parte experimental, secuencia de unión a ADN subrayada):

GACGGGGCUCGAAGGGUCCUUGUAGGUUGUAGCUGCCCUUUCGAAAGAACCGUU
25 GCUACAAUAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAGC
UUCUGCUUUAAGGGGCAUCGUUUUUUUUUUU

En una realización particular adicional, el ARNgm utilizado para inducir DSB cadena abajo (o 3') de la región de repeticiones de trinucleótidos se selecciona del grupo que consiste en:

- SEQ ID NO:7 (también denominado ARNgm E en la parte experimental, secuencia de unión a ADN subrayada):

30 GUGGGGAGCGUCUGGCGCGAUCUCUGUUGUAGCUGCCCUUUCGAAAGAACCGUU
GCUACAAUAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAGC
UUCUGCUUUAAGGGGCAUCGUUUUUUUUUUU

- SEQ ID NO:8 (también denominado ARNgm N en la parte experimental, secuencia de unión a ADN subrayada):

GUCGGGGUCUCAGUGCAUCCAAAAGUUGUAGCUGCCCUUUCGAAAGAACCGUUG
35 CUACAAUAAGGCCGUCUGAAAAGAUGUGCCGCAACGCUCUGCCCCUUAAGCU
UCUGCUUUAAGGGGCAUCGUUUUUUUUUUU

La invención se refiere además a un vector tal como se ha definido anteriormente, que comprende una secuencia que codifica una molécula de ARNgm que presenta una secuencia seleccionada del grupo que consiste en SEQ ID NO:5 a 8.

40 En otra realización, la pareja de moléculas de ARNgm de la invención es una pareja seleccionada del grupo que consiste en SEQ ID NO:5 y SEQ ID NO:7; SEQ ID NO:5 y SEQ ID NO:8; SEQ ID NO:6 y SEQ ID NO:7, y SEQ ID NO:6 y SEQ ID NO:8. En una realización preferente, la pareja de moléculas de ARNgm se selecciona del grupo que consiste en SEQ ID NO:5 y SEQ ID NO:8; SEQ ID NO:6 y SEQ ID NO:7, y SEQ ID NO:6 y SEQ ID NO:8.

45 Tal como se ha mencionado anteriormente, el ARNgm de la invención puede expresarse a partir de un casete de expresión. La expresión del ARNgm puede controlarse en particular mediante un promotor, tal como un promotor U6. Por consiguiente, la invención incluye además un casete para la expresión de un ARNgm, que comprende una

secuencia codificante de ARN_gm situada bajo el control de un promotor, tal como el promotor U6 mostrado en la SEQ ID NO:9.

En una realización particular, el casete de expresión comprende la siguiente secuencia para la expresión del ARN_gm mostrado en la SEQ ID NO:5 a partir de un promotor U6:

AGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACG
ATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATA
TTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTT
TAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTT
GATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCGCTCCCTGAAC
CCTAGAACTGTGTTGTAGCTCCCTTTCGAAAGAACCGTTGCTACAATAAGGCCGT
CTGAAAAGATGTGCCGCAACGCTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCAT
CGTTTATTTTTTTTAA (SEQ ID NO:10)

5

En una realización particular, el casete de expresión comprende la siguiente secuencia para la expresión del ARN_gm mostrado en la SEQ ID NO:6 a partir de un promotor U6:

AGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACG
ATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATA
TTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTT
TAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTT
GATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGACGGGGCTCGAA
GGGTCCTTGTAGTTGTAGCTCCCTTTCGAAAGAACCGTTGCTACAATAAGGCCG
TCTGAAAAGATGTGCCGCAACGCTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCA
TCGTTTATTTTTTTTAA (SEQ ID NO:11)

10

En una realización particular, el casete de expresión comprende la siguiente secuencia para la expresión del ARN_gm mostrado en la SEQ ID NO:7 a partir de un promotor U6:

AGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACG
ATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATA
TTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTT
TAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTT
GATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGTGGGGAGCGTCT
GGCGCGATCTCTGTTGTAGCTCCCTTTCGAAAGAACCGTTGCTACAATAAGGCCG
TCTGAAAAGATGTGCCGCAACGCTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCA
TCGTTTATTTTTTTTAA (SEQ ID NO :12).

En una realización particular, el casete de expresión comprende la siguiente secuencia para la expresión del ARN_gm mostrado en la SEQ ID NO:8 a partir de un promotor U6:

AGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACG
 ATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATA
 TTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTT
 TAAAATTATGTTTTAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTT
 GATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGTCGGGGTCTCAG
 TGCATCCAAAAGTTGTAGCTCCCTTTGAAAGAACCGTTGCTACAATAAGGCCGT
 CTGAAAAGATGTGCCGCAACGCTCTGCCCTTAAAGCTTCTGCTTTAAGGGGCAT
 CGTTTATTTTTTTTAA (SEQ ID NO :13).

En las SEQ ID NO:11 y 12 tal como se han representado anteriormente, se introdujo la base G subrayada debido a que se requiere una G para iniciar la transcripción a partir del promotor U6. Sin embargo, el experto en la materia entenderá que otros promotores pueden no requerir una G en esta posición inmediatamente antes de la secuencia codificante guía, o pueden requerir una o más otras bases nucleótidas, tal como es bien conocido en la técnica.

Métodos y usos de la invención

La presente invención contempla diversas maneras de alcanzar una secuencia diana de ADN genómico con una endonucleasa Cas9 y moléculas de ARNgm. En algunas realizaciones, la endonucleasa Cas9 se introduce dentro de una célula en una forma polipeptídica. En una variante, la endonucleasa Cas9 se conjuga o se fusiona con un péptido de penetración celular, que es un péptido que facilita la incorporación de una molécula en una célula. Las moléculas de ARNgm pueden administrarse además en la célula como oligonucleótido aislado, directamente o utilizando reactivos de transfección, tal como derivados lipídicos, liposomas, fosfato de calcio, nanopartículas, microinyección o electroporación.

En otra realización, la presente invención contempla introducir la endonucleasa Cas9 y/o moléculas de ARNgm en la célula diana en forma de un vector que expresa dicha endonucleasa y/o moléculas de ARNgm. Los métodos de introducción y expresión de genes en una célula son conocidos en la técnica. El vector de expresión puede transferirse dentro de una célula anfitriona por medios físicos, químicos o biológicos. El vector de expresión puede introducirse en la célula utilizando métodos físicos conocidos, tales como la precipitación con fosfato de calcio, la lipofección, el bombardeo con partículas, la microinyección y la electroporación. Los medios químicos para introducir un polinucleótido en una célula anfitriona incluyen sistemas de dispersión coloidal, tales como complejos de macromoléculas, nanocápsulas, microesferas, perlas y derivados lipídicos y liposomas. En otras realizaciones, la endonucleasa Cas9 y/o las moléculas de ARNgm se introducen por medios biológicos, en particular mediante un vector vírico. Entre los vectores víricos representativos útiles en la práctica de la invención se incluyen, aunque sin limitarse a ellos, un vector derivado de adenovirus, retrovirus, en particular lentivirus, poxvirus, virus del herpes simplex tipo 1 y virus adenoasociados (VAA). La selección del vector vírico apropiado evidentemente dependerá de la célula diana y del tropismo vírico.

En una realización, la endonucleasa Cas9 y las moléculas de ARNgm se proporcionan dentro de diferentes vectores (tales como dos vectores: uno que contiene un gen codificante de la endonucleasa Cas9, y un segundo codificante de ambas moléculas de ARNgm, o tres vectores, uno codificante de la endonucleasa Cas9 y un vector para cada molécula de ARNgm). En otra realización, la totalidad de los vectores del sistema CRISPR-Cas9, incluyendo la endonucleasa Cas9 y ambas moléculas de ARNgm requeridas para la escisión de la expansión de repeticiones de trinucleótidos, se expresan a partir de un único vector de expresión.

El sistema de la presente invención se utiliza para escindir una expansión de repeticiones de nucleótidos dentro de una región no codificante de un gen de interés. En una realización, el motivo de nucleótidos repetidos es un binucleótido, trinucleótido, tetranucleótido, pentanucleótido o hexanucleótido repetido, tal como una repetición CAG, CTG, CGG, GAA, AGG, CCG, CCTG, ATTCT, TGAA, GGCCTG o GGGGCC. En una realización particular, la repetición se encuentra dentro de un gen de interés seleccionado de FMR1, AFF2 o FMR2, AFF3, FXN, ATXN80S/ATXN8, ATXN10, PPP2R2B, BEAN1/TK2, NOP56, C9ORF72, ZN9/CNBP o DMPK. En una realización particular, la expansión de repeticiones de nucleótidos es una repetición de trinucleótido, tal como una repetición CAG, CTG, CGG, GAA, AGG o CCG. En una realización particular adicional, la repetición de nucleótidos, en particular la repetición de trinucleótidos, se encuentra presente dentro de la región 3' no traducida del gen DMPK. En una realización particular, la expansión de repeticiones de nucleótidos (p.ej., una expansión de repeticiones de trinucleótidos) comprende entre 20 y 10000 repeticiones del motivo nucleótido, más particularmente de 50 a 5000 repeticiones. Por ejemplo, la expansión de repeticiones de nucleótidos que debe escindirse (p.ej., una expansión de repeticiones de trinucleótidos) puede comprender cualquier número de repeticiones, tal como al menos 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o al menos más de 2000 repeticiones del motivo de nucleótidos. Más específicamente, el número de repeticiones es un número patológico de repeticiones, que significa que dicha repetición de nucleótidos (p.ej., una repetición de trinucleótidos) está asociada, o puede estar asociada, a

un estado de enfermedad. En una realización particular, la repetición es una repetición de CTG dentro de la región 3' no traducida del gen DMPK y es patológica a partir de 50 o más repeticiones.

Tal como se utiliza en la presente memoria, los términos “tratar” y “tratamiento” se refieren a administrar en un sujeto una cantidad eficaz de una composición de manera que el sujeto presenta una reducción de por lo menos un síntoma de la enfermedad o una mejora de la enfermedad, por ejemplo, resultados clínicos beneficiosos o deseados. Para los fines de la presente invención, entre los resultados clínicos beneficiosos o deseados se incluyen, aunque sin limitarse a ellos, el alivio de uno o más síntomas, la reducción de la extensión de la enfermedad, la estabilización (es decir, el no agravamiento) del estado de la enfermedad, el retardo o enlentecimiento de la progresión de la enfermedad, la mejora o paliación del estado de la enfermedad, y la remisión (parcial o total), detectable o no detectable. El tratamiento puede referirse a la prolongación de la supervivencia en comparación con la supervivencia esperada si no recibe tratamiento. Alternativamente, el tratamiento resulta “eficaz” en el caso de que la progresión de la enfermedad se reduzca o se detenga. “Tratamiento” también puede significar la prolongación de la supervivencia en comparación con la supervivencia esperada en caso de no recibir tratamiento. Aquellos que necesitan tratamiento incluyen los pacientes en los que ya se ha diagnosticado un trastorno asociado a la expresión de una secuencia polinucleótida, así como aquellos que es probable que desarrollen tal trastorno debido a susceptibilidad genética u otros factores. Tal como se utiliza en la presente memoria, los términos “tratar” y “tratamiento” también se refieren a la prevención de una enfermedad o trastorno, que significa retardar o evitar la aparición de tal enfermedad o trastorno.

Tal como se utiliza en la presente memoria, un “trastorno por expansión de repeticiones de nucleótidos” es una enfermedad o trastorno que está causado o ligado a una expansión de repeticiones de nucleótidos donde las repeticiones de nucleótidos en determinados genes exceden el umbral estable normal, que difiere según gen. Entre los trastornos por expansión de repeticiones de nucleótidos se incluyen el síndrome del X frágil (FXS, por sus siglas en inglés) y el síndrome de temblor y ataxia asociado al X frágil (FXTAS, por sus siglas en inglés), ataxia espinocerebelar tipos 8, 10, 12, 31 y 36, distrofia miotónica de tipo 1 y de tipo 2, ataxia de Friedreich, enfermedad similar a Huntington de tipo 2 (HDL-2) y esclerosis lateral amiotrófica (C9-ELA).

En una realización particular, los métodos de la presente invención se refieren al tratamiento de un trastorno por expansión de repeticiones de nucleótidos en el que la expansión de repeticiones de nucleótidos asociada al trastorno se localiza dentro de una región no codificante de un gen. Entre ellos se incluyen, por ejemplo, expansiones de repeticiones de nucleótidos localizadas dentro de un intrón, una región 5' no traducida o una región 3' no traducida de un gen codificante de proteína.

En una realización particular, el trastorno de expansión de repeticiones de nucleótidos es la distrofia miotónica, asociada a una expansión de repetición trinucleotido (tal como un CTG) dentro de la región 3' no traducida del gen DMPK.

En una realización particular, la expansión de repetición de nucleótidos dentro del gen DMPK se escinde utilizando una pareja apropiada de moléculas de ARNgm y una endonucleasa Cas9 apropiada, correspondiente a los ARNgm seleccionados.

La molécula de ARNgm, la pareja de moléculas de ARNgm, el vector (codificante de una o más moléculas de ARNgm y/o una endonucleasa Cas9) y la célula según la invención pueden formularse y administrarse para tratar una diversidad de estados de enfermedad por expansión de repeticiones de nucleótidos por cualesquiera medios que produzcan el contacto de la molécula de ARNgm, la pareja de moléculas de ARNgm, el vector y la célula con su sitio de acción en el sujeto que lo necesita.

La presente invención proporciona además composiciones farmacéuticas que comprenden un ARNgm o una pareja de ARNgm de la invención, o el vector de la invención (codificante de un ARNgm de la invención, o una pareja de ARNgm solos o junto con una secuencia codificante de una endonucleasa Cas9), o la célula de la invención. Tales composiciones comprenden una cantidad terapéuticamente eficaz del terapéutico (el o los ARNgm, el vector o la célula de la invención) y un portador farmacéuticamente aceptable. En una realización específica, la expresión “farmacéuticamente aceptable” se refiere a aprobado por una agencia reguladora del gobierno federal o estatal o listada en la Farmacopea estadounidense o europea u otra farmacopea generalmente reconocida para el uso en animales y seres humanos. El término “portador” se refiere a un diluyente, adyuvante, excipiente o vehículo con el que se administra el terapéutico. Tales portadores farmacéuticos pueden ser líquidos estériles, tales como agua y aceites, incluyendo los originados en petróleo o de origen animal, vegetal o sintético, tales como aceite de cacahuete, aceite de soja, aceite mineral, aceite de sésamo y similar. El agua es un portador preferente en el caso de que la composición farmacéutica se administre por vía intravenosa. También pueden utilizarse soluciones salinas y soluciones acuosas de dextrosa y glicerol como portadores líquidos, particularmente para soluciones inyectables. Entre los excipientes farmacéuticos adecuados se incluyen almidón, glucosa, lactosa, sucrosa, estearato sódico, monoestearato de glicerol, talco, cloruro sódico, leche desnatada seca, glicerol, propilenglicol, agua, etanol y similares.

La composición, si se desea, también puede contener cantidades menores de agentes humectantes o emulsionantes, o agentes tamponadores del pH. Estas composiciones pueden adoptar la forma de soluciones, suspensiones, emulsiones, tabletas, píldoras, cápsulas, polvos, formulaciones de liberación sostenida y similares. La formulación oral puede incluir portadores estándares, tales como grados farmacéuticos de manitol, lactosa, almidón, estearato de

magnesio, sacarina sódica, celulosa, carbonato de magnesio, etc. Se describen ejemplos de portadores farmacéuticos adecuados en "Remington's Pharmaceutical Sciences", por E.W. Martin. Tales composiciones contendrán una cantidad terapéuticamente eficaz del terapéutico, preferentemente en forma purificada, junto con una cantidad adecuada de portador, de manera que proporcione la forma para la administración apropiada en el sujeto.

- 5 En una realización preferida, la composición se formula de acuerdo con procedimientos rutinarios como composición farmacéutica adaptada para la administración intravenosa en seres humanos. Típicamente, las composiciones para la administración intravenosa son soluciones en tampón acuoso isotónico estéril. En caso necesario, la composición puede incluir además un agente solubilizador y un anestésico local, tal como lignocaína para aliviar el dolor en el sitio de la inyección.
- 10 La cantidad del terapéutico de la invención que resultará eficaz en el tratamiento de una expansión de repeticiones de nucleótidos puede determinarse mediante técnicas clínicas estándares. Además, pueden utilizarse opcionalmente ensayos *in vivo* y/o *in vitro* para ayudar a predecir los intervalos de dosis óptimos. La dosis precisa para la utilización en la formulación también dependerá de la vía de administración y de la gravedad de la enfermedad, y debería decidirse de acuerdo con el criterio del profesional y las circunstancias de cada paciente. La dosis de uno o más
- 15 ARNgm, vector o célula administrada en el sujeto que lo necesita variará según varios factores, entre ellos, aunque sin limitación, la vía de administración, la enfermedad específica tratada, la edad del sujeto o el nivel de expresión necesario para obtener el efecto terapéutico requerido. El experto en la materia puede determinar fácilmente, basándose en sus conocimientos en el campo, el intervalo de dosis requerido basándose en estos factores y otros factores.

20 Objetos no limitativos de la invención

La presente solicitud proporciona los siguientes objetos no limitativos:

1. Una pareja de moléculas de ARNgm, en donde dicha pareja comprende una primera y una segunda moléculas de ARNgm que es capaz de unir mediante apareamiento de bases un complemento de secuencia a una secuencia diana de ADN genómico localizadas en los lados 5' y 3' de una expansión de repeticiones de nucleótidos localizada dentro de una región no codificante de un gen de interés, respectivamente, resultando apropiada de esta manera para escindir dicha expansión de nucleótidos con un sistema CRISPR/Cas9.
- 25 2. Un ARNgm que comprende una secuencia que es capaz de unir mediante apareamiento de bases el complemento de secuencia a una secuencia diana de ADN genómico localizada en el lado 5' o 3' de una expansión de repeticiones de nucleótidos dentro de una región no codificante de un gen de interés.
- 30 3. La pareja de moléculas de ARNgm según el objeto 1, o el ARNgm según el objeto 2, en donde dicha repetición de nucleótidos está localizada dentro de un intrón o dentro de la región 5' o 3' no traducida (5'UTR o 3'UTR, por sus siglas en inglés) de dicho gen de interés.
4. La pareja de moléculas de ARNgm o el ARNgm según el objeto 3, en donde la repetición de nucleótidos está localizada dentro de la 3'-UTR de dicho gen de interés.
- 35 5. Un vector codificante del ARNgm o una pareja de moléculas de ARNgm según cualquiera de los objetos 1 a 4, siendo el vector preferentemente un plásmido o un vector vírico, tal como un vector VAAr.
6. Una célula diana, que se transfecta o transduce con el vector según el objeto 5.
7. Un método para la producción de un ARNgm o pareja de ARNgm, que comprende cultivar la célula diana según el objeto 6 bajo condiciones que permiten la producción de dicho ARNgm o pareja de ARNgm, y la recuperación de dicho
- 40 ARNgm o dicha pareja de moléculas de ARNgm a partir de dicha etapa de cultivo.
8. Un método *in vitro* para escindir una repetición de nucleótidos localizada dentro de una región no codificante de un gen en una célula, que comprende introducir en dicha célula una pareja de moléculas de ARNgm según cualquiera de los objetos 1 y 3 a 4, o un vector según el objeto 5, y una endonucleasa CRISPR/Cas9.
9. El método según el objeto 8, en donde la endonucleasa Cas9 se deriva de *S. pyogenes*, *S. thermophilus*, *S. aureus*, *N. meningitidis*, *S. mutans*, *C. jejuni*, *F. novicida*, *P. multocida*, *P. bettyae*, *H. parainfluenzae*, *H. pittmaniae* o *L. crispatus*, preferiblemente de *N. meningitidis*.
- 45 10. Una pareja de ARNgm según cualquiera de los objetos 1 y 3 a 4, o un vector según el objeto 5, para el uso en combinación con una endonucleasa Cas9 en un método para el tratamiento de un trastorno por expansión de repeticiones de nucleótidos, en donde las moléculas de ARNgm de la pareja de ARNgm se diseñan para escindir una expansión de repeticiones de nucleótidos asociada a dicho trastorno respecto de una región no codificante de un gen de interés.
- 50

11. La pareja de ARNgm para el uso según el objeto 10, en donde la expansión de repeticiones de nucleótidos es una expansión de repeticiones de binucleótidos, trinucleótidos, tetranucleótidos, pentanucleótidos o hexanucleótidos, en particular una expansión de repeticiones de trinucleótidos.

5 12. La pareja de ARNgm para el uso según el objeto 10 o 11, en donde la expansión de repeticiones de nucleótidos se localiza dentro de la 3'-UTR de dicho gen de interés.

13. La pareja de ARNgm para el uso según cualquiera de los objetos 10 a 12, en donde dicho trastorno es síndrome de X frágil y síndrome de temblor y ataxia asociado al X frágil, ataxia espinocerebelar de tipo 8, 10, 12, 31 y 36, distrofia miotónica de tipo 1 y 2, ataxia de Friedreich, enfermedad similar a Huntington tipo 2 y esclerosis lateral amiotrófica tipo C9-ELA.

10 14. El ARNgm según los objetos 2 a 4, o la pareja de ARNgm para el uso según cualquiera de los objetos 10 a 13, en donde el gen de interés es FMR1, AFF2 o FMR2, AFF3, FXN, ATXN80S/ATXN8, ATXN10, PPP2R2B, BEAN1/TK2, NOP56, C9ORF72, ZN9/CNBP o DMPK.

15. La pareja de ARNgm para el uso según cualquiera de los objetos 10 a 14, en donde el trastorno es distrofia miotónica tipo 1 y el gen de interés es el gen DMPK.

15 Ejemplos

Materiales y métodos

Construcción de plásmidos

La lista de plásmidos utilizados y contruidos en el presente estudio se proporciona en la Tabla 3. Se han llevado a cabo experimentos de clonación en *Escherichia coli* DH10B químicamente competentes (Invitrogen). Los casetes de expresión para la proteína Cas9 de *Neisseria meningitidis* (NM) se listan como pC512-Int-NMCas9-smSVpolyA y pEFS-Int-NMCas9-SVpolyA. Ambos contienen la secuencia optimizada humana de NMCas9 bajo el control del promotor específico muscular sintético C5-12 o el promotor EFS ubicuo (versión más corta del promotor EF-1α). La secuencia de NMCas9 contiene además dos señales de localización nuclear (SLN), uno por extremo, el epítipo hemaglutinina (HA) de influenza humana y tres epítopos FLAG en tándem (3XFLAG) en el extremo C-terminal. La secuencia de NM Cas9 se amplificó por PCR a partir de plásmido JDS246NMCas9 (proporcionado por Jean Paul Concordet) y se clonó en los plásmidos de VAA pC512-Int-smSVpolyA y pEFS-Int-SVpolyA, cada uno de los cuales contiene el promotor indicado, un intrón quimérico y una señal de poliadenilación de SV40 (sm: versión más corta).

Casete de expresión para el ARN guía monocatenario (ARNgm) NM, pBlue_Double_U6sgRNAs_NM, se diseñó con el fin de contener dos veces en tándem la secuencia optimizada humana para el andamiaje (parte constante) del ARNgm NM bajo el control del promotor U6. El constructo contiene dos sitios de clonación por restricción diferentes donde se introduce la cadena del ARNgm específico para la diana genómica (protoespaciador), *BbsI* y *BspMI*. La secuencia del andamiaje de ARNgm fue proporcionada por Jean Paul Concordet y modificada en los sitios de clonación del protoespaciador. El constructo Double_U6sgRNAs_NM se sintetizó sintéticamente y fue clonada en el plásmido pBluescript SK(+) por GeneCust.

Los protoespaciadores ARNgm 1, 7, E y N se diseñaron tal como se describe en la sección "Diseño de los ARNgm" (ver posteriormente). En primer lugar, se clonaron los ARNgm 1 y 7 en el sitio *BbsI* del plásmido pBlue_Double_U6sgRNAs_NM para originar los plásmidos pBlue-U6gRNA-NM-1-DMPK y pBlue-U6gRNA-NM-7-DMPK. Se clonaron los ARNgm E y N en el sitio *BspMI* de este último, originando las siguientes combinaciones: pBlue-U6gRNA-NM-1N-DMPK, pBlue-U6gRNA-NM-7E-DMPK y pBlue-U6gRNA-NM-7N-DMPK. Con el fin de seleccionar mediante separación las células transfectadas con los ARNgm, los constructos indicados anteriormente se digirieron enzimáticamente en *SnaBI* y se clonaron en un plásmido que contenía un GFP nuclear, pAAV-Des-AcGFP.

Diseño de los ARNgm

La región genómica de la 3'UTR de *DMPK* circundante a las repeticiones de CTG se cribó manualmente para la presencia de motivos contiguos a protoespaciador (PAM) específicos para *Neisseria meningitidis* (NM). Aparte de la secuencia de consenso del NM PAM (NNNNGATT) [Zhang Y. et al., Mol. Cell., 2013], también se consideraron las siguientes variantes, sometidas a ensayo por Esvelt y colaboradores en *Escherichia coli* [Esvelt K.M. et al., Nature Methods, 2013]: NNNNGNTT, NNNNGANT, NNNNGTTN. Considerando ambos, las posiciones de las dianas y los PAM dentro de la región 3' UTR y también el número de potenciales fuera-de-diana (calculado por CasOFFinder fijando el número de corte de no correspondencias en ≤ 6), se seleccionaron cuatro ARNgm, dos con diana en la región cadena arriba de la repetición de CTG (ARNgm 1 y 7) y los otros con diana en la región cadena debajo de la repetición de CTG (ARNgm E y N). Sólo uno de ellos es respecto a un PAM de consenso (ARNgm N); en efecto los demás son respecto, dos a la variante NNNNGANT (ARNgm 1 y 7) y uno respecto a la variante NNNNGNTT (ARNgm E) (fig. 3). Se diseñaron ARNgm de manera que su diana fuese una secuencia genómica de 24 nt y con el fin de clonarla en el sitio *BbsI* (ARNgm 1 y 7) o en el sitio *BspMI* (ARNgm E y N) del plásmido pBlue_Double_U6sgRNA_NM. Además, se añadió el nucleótido G al 5' de ARNgm 7 y E (ya presente en el 5' del otro ARNgm) para optimizar la transcripción controlada por U6. Todos los ARNgm se sintetizaron sintéticamente como cebadores directos e inversos simples (ver

la lista de cebadores) y después se hibridaron *in vitro* antes de la ligación con el plásmido pBlue_Double_U6sgRNA_NM (o derivados) digeridos con los enzimas de restricción apropiados.

Cultivo celular y ensayo de transfección para someter a ensayo la actividad de nucleasa

5 Se sembraron células 1 día antes de la transfección, se recolectaron 2 días después de la transfección y se mantuvieron a -80°C hasta la extracción del ADN genómico, a menos que se indique lo contrario. Se utilizó una temperatura estándar de 37°C y 5% de CO₂ para cultivar y mantener las células en cultivo. Se proporcionan los datos para los ensayos de transección posteriormente.

10 Se cultivaron células HeLa y células HEK 293T en medio de Eagle modificado por Dulbecco (DMEM) rico en glucosa y con GlutaMAX (Invitrogen), complementado con suero de feto bovino al 10% (FBS, Invitrogen). Las células se sembraron en placa de 12 pocillos, 0,5-1,0x10⁵/pocillo para las células HeLa, 0,25x10⁵ células/pocillo para las HEK 293T, en un volumen final=1 ml (~80% de confluencia el día de la transfección). Las reacciones de transfección se prepararon de la manera siguiente: 3 µl de reactivo de transfección FuGENE HD y 1 µg de ADN total (proporción FuGENE-ADN de 3:1) en 50 µl de medio sin suero. La proporción de NMCas 9 y ARNgm era de 3:1.

15 Se cultivaron mioblastos fetales humanos primarios en medio de crecimiento de células musculares esqueléticas (PromoCell) complementadas con FBS al 15%. Con el fin de alcanzar ~50% de confluencia, el día de la transfección se sembraron ~0,3-0,5x10⁵ células/placa Petri de 10 cm (volumen final=8 ml). Las células se transfectaron con reactivo de transfección JetPEI (transfección Polyplus) en una proporción 2:1 con el ADN en un volumen final de 500 µl de NaCl 150 mM (12 µl de reactivo de transfección para 6 µg de ADN total; proporción NMCas9-ARNgm de 1:1). Las reacciones de transfección se incubaron durante 30 min a TA y se cambió el medio de cultivo antes de su adición.

20 Ensayo de PCR para someter a ensayo la delección genómica

Se extrajo el ADN genómico de células HeLa y de células HEK 293T utilizando el kit de purificación de ADN genómico GeneJET (Thermo Scientific) y se eluyeron en un volumen final de 150 µl.

Se extrajo el ADN genómico de mioblastos fetales humanos primarios separados y no separados, utilizando el kit QIAamp DNA Micro and Mini (QIAGEN) y se eluyeron en un volumen final de 30 y 200 µl, respectivamente.

25 Se llevó a cabo la PCR utilizando la ADN polimerasa Taq de alta fidelidad Platinum (Invitrogen), DMSO al 10%, 100-150 ng de ADN total y los cebadores F1- y R1-DMPK-3'UTR. Los productos de PCR se separaron mediante electroforesis: 15-20 µl de cada reacción de PCR se cargaron en un gel de agarosa al 1,5% que contenía la tinción de ADN GelRed (la tanda de electroforesis se realizó a 90-100 voltios durante 1 h 15-30 min). Tras adquirir las imágenes de los geles de electroforesis tras la exposición a UV, se ajustó el brillo y el contraste.

30 Se llevó a cabo la secuenciación de los productos de PCR que contenían la delección de la repetición de CTG, subclonando en orden para secuenciar clones individuales. Se llevó a cabo la PCR tal como se ha indicado anteriormente con los cebadores F1-DMPK-3UTR-KpnI y R1-DMPK-3UTR-XbaI. Los productos de PCR se separaron mediante electroforesis en un gel de agarosa a baja temperatura de fusión al 2% y se extrajeron en gel bandas de ADN respecto a los productos de PCR que contenían la delección de las repeticiones de CTG, se digirieron con *KpnI* y *XbaI* y se clonaron en el plásmido pBluescript II SK(+) digerido con los mismos enzimas de restricción. Los plásmidos se extrajeron de los clones positivos y se enviaron para secuenciación con el cebador pBlue_MCS_before2.

Transferencia western

40 Las células se lisaron sobre hielo en 100 µl de tampón de lisis que contenía NaCl 150 mM, Tris-HCl 10 mM (pH 7,4), EGTA 1 mM, EDTA 1 mM, fluoruro sódico 100 mM, pirofosfato sódico 4 mM, ortovanatado sódico 2 mM, Triton X-100 al 1% e IGEPAL al 0,5% complementado con un cóctel inhibidor de proteasas completo (Roche). Las concentraciones del extracto de proteínas totales se determinaron mediante el kit de ensayo de proteínas de Bio-Rad (Bio-Rad). Las proteínas se separaron mediante electroforesis en SDS-PAGE en gel premoldeado al 10% (Bio-Rad) y después se transfirieron a membrana de nitrocelulosa. Con el fin de evitar la unión no específica, las membranas se sumergieron en una solución de bloqueo (solución Odyssey al 50%, solución TBST al 50% [Tween al 0,1% en Tris-HCl 1 M, pH 7,6) durante 2 horas antes de la incubación durante la noche con anticuerpos dirigidos contra epítipo Flag (anticuerpo monoclonal M2 anti-Flag, 1:3.000, Sigma). El día después, se lavó el exceso de anticuerpos en TBST y la membrana se incubó con un anticuerpo secundario conjugado con Alexa (anticuerpo de cabra antiratón-Alexa Fluor 680, 1:10.000, Invitrogen) durante 1 hora. Las bandas de proteínas se visualizaron mediante fluorescencia de infrarrojos utilizando el sistema de adquisición de imágenes Odyssey (LICOR Biotechnology, Inc.).

50 Resultados

Constructos de expresión de Cas9 y ARNgm

Entre todos los sistemas de CRISPR/Cas9 disponibles, los presentes inventores seleccionaron uno originado en la bacteria *Neisseria meningitidis* (NM) debido a que NMCas9 es de pequeño tamaño y puede caber dentro de un vector virus adenoasociado (VAA). Además, los presentes inventores encontraron mediante cribado manual la presencia de

5 sitios diana potenciales para NMCas9 dentro de la región 3' no traducida (3'-UTR) del gen DMPK humano. Los presentes inventores generaron dos constructos de VAA que contenían la secuencia codificante de NMCas 9 optimizada humana (3,25 Kb) bajo el promotor ubicuo EFS (versión más corta del factor 1 alfa de elongación) o el promotor específico muscular C5-12. Los presentes inventores incluyeron además otras secuencias reguladoras en los constructos, un intrón quimérico situado cadena debajo de cada promotor para estabilizar el ARNm de Cas9, y una señal de poliadenilación del virus 40 del simio (poliA de SV40). Se fusionó Cas9 con dos señales de localización nuclear (NLS), uno en cada extremo, con el fin de transportar la proteína hasta dentro del núcleo y HA (derivado de hemaglutinina de influenza humano) y epitopos Flag, los cuales resultan útiles para la detección de proteínas. El esquema de los constructos se representa en la fig. 1, A y B.

10 Los presentes inventores sometieron a ensayo la funcionalidad de los casetes de expresión de Cas9 *in vitro* en las líneas celulares (fig. 2) e *in vivo* en ratones C57B1/6 de tipo salvaje (datos no mostrados). Brevemente, las células se transfectaron con constructos de Cas9 y, dos días después de la transfección, se lisaron para la extracción de las proteínas totales o se fijaron para experimentos de marcaje. Se detectó NMCas9 en células HeLa y C2C12 en el tamaño molecular esperado mediante análisis de transferencia western (fig. 2) y en el núcleo de las células mediante inmunofluorescencia (datos no mostrados) con anticuerpos dirigidos contra el epítipo Flag.

15 Con el fin de deletar la repetición de CTG de la región 3'UTR del gen DMPK humano, los presentes inventores diseñaron ARNm situados cadena arriba y cadena debajo de esta región de repetición de CTG, que podría controlar el corte doble de ADN mediado por NMCas9 después de la secuencia del codón de parada de DMPK y antes de la señal de poliA. De esta manera, los presentes inventores diseñaron un casete de expresión que contenía dos andamiajes de ARNm, cada uno bajo el control del promotor U6 y localizado en tándem en el mismo plásmido (fig. 1, constructo C). Los presentes inventores seleccionaron cuatro protoespaciadores ARNm que minimizaban los efectos fuera-de-diana (ver las dianas genómicas respectivas en la fig. 3 y en la Tabla 1) y se clonaron en los sitios de clonación *Bsl* y *Bspml* que precedían a los andamiajes de ARNm.

Tabla 1: secuencias de ADN de protoespaciador y de motivo contiguo a protoespaciador (PAM)

Nombre de diana	Secuencia del ADN diana (protoespaciador) + secuencia de ADN de PAM (subrayado)	Posición# (inicio-final)	Cadena
1	GCGCTCCCTGAACCCTAGAACTGTCTTCGACT	46273725- 46273756	-
7	ACGGGGCTCGAAGGGTCCTTGAGCCGGGAAT	46273524- 46273555	-
E	TGGGGAGCGTCTGGCGCGATCTCTGCCTGCTT	46273258- 46273289	-
N	GTCGGGGTCTCAGTGCATCCAAAACGTGGATT	46273057- 46273088	+

Referencia del genoma humano: feb. de 2009 (GRCh37/hg19)

Los presentes inventores generaron tres parejas de ARNm diferentes que potencialmente podrían controlar la delección de la expansión de repeticiones de CTG (β , γ y δ , Tabla 2).

Tabla 2: parejas de ARNm y tamaño de delección esperado

	Ningún ARNm	Cas9+ARNm		
Pareja de ARNm		B (1+N)	γ (7+E)	δ (7+N)
Número total de potenciales fuera-de-diana		9+8=17	6+6=12	6+8=14
Tamaño del fragmento de ADN escindido		658#	266#	457#

Tamaño esperado de producto de PCR (pb) ⁺⁺⁺	891*	233#	625#	434#
--	------	------	------	------

⁺ Considerando un número de no correspondencias ≤ 6 , comprobado mediante Cas-OFFinder <http://www.rgenome.net/cas-offinder/>

⁺⁺⁺ Con cebadores F1-DMPK-3UTR y R1-DMPK-3UTR

* Tamaño calculado para un número de repeticiones de CTG (n CTG)=20

5 (861* para células HeLa [n CTG=10]; 846* para células 293T [n CTG=5])

#tamaño calculado considerando un corte preciso entre el 3º y 4º nucleótido cadena arriba de la secuencia de PAM

Estos constructos también se subclonaron en un plásmido de VAA que alojaba la secuencia para una proteína fluorescente verde (fig. 1, constructo D y materiales y métodos).

Deleción mediada por Cas9 de la expansión de repeticiones de CTG en el gen *DMPK* humano

10 Con el fin de someter a ensayo si los ARNgm diseñados eran capaces de dirigir NMCas9 a las regiones diana de ADN genómico correspondientes, los presentes inventores cotransfectaron células HeLa (o HEK 293T) con el plásmido que contenía el casete de expresión EFS-NMCas9 y cada uno de los tres plásmidos que alojaban una pareja de ARNgm y controles apropiados (Tabla 2 y figura 4). A continuación, se extrajo el ADN de las células y se utilizó como molde en una reacción de PCR para amplificar una región de 891 pb de longitud que incluía todas las dianas de ADN genómico (cebadores F1- y R1-DMPK-3'UTR, fig. 3). El gel de agarosa de los productos de PCR se muestra en la fig. 15 4. Las parejas de ARNgm (β , γ y δ) resultó en la amplificación por PCR del fragmento de ADN de 891 pb y un fragmento más pequeño correspondiente a la región de ADN esperada después de la escisión mediada por CRISPR/Cas9 (fig. 4, flecha roja). Estos productos de PCR de tamaño más pequeño se subclonaron y se secuenciaron para verificar su identidad. Estos productos de PCR de tamaño más pequeño se subclonaron y se secuenciaron para verificar su identidad. Los resultados de la secuenciación de los productos de PCR obtenidos con las parejas de ARNgm β (1 más N) y δ (7 más N) se representan en la fig. 6 y demuestran que, en la mayoría de los clones, los sitios de corte de NMCas9 se localizan en el nucleótido N3 o N4 de la secuencia diana contiguos al PAM (#n). Por tanto, estos resultados demuestran la eficacia de los ARNgm diseñados en el control de la deleción mediada por NMCas9 del ADN genómico flanqueante de la repetición de CTG de *DMPK*.

25 Los presentes inventores también sometieron a ensayo la eficacia de este sistema CRISPR/Cas9 en el locus DMPK humano en presencia de una expansión de CTG patológica. Con este fin, células primarias derivadas de un paciente de DM1 portador de una repetición de 200 CTG (DM200) se cotransfectaron con los plásmidos EFS-NMCas9 y ARNgm y (ARNgm 7 más E, constructo D) y se utilizaron células primarias procedentes de un individuo normal (Ctrl) como control (fig. 1). Se extrajo ADN genómico de las células Ctrl y DM200 separadas si GFP-positivas con el fin de incrementar la proporción de ADN originado de células que expresaban los ARNgm. Se llevó a cabo una PCR tal como se ha indicado anteriormente y resultó en la amplificación de pequeños fragmentos de ADN en células Ctrl y DM200 que correspondían en tamaño a los fragmentos originados de la región de ADN genómico escindida esperada (fig. 5, flecha roja).

35 Los presentes inventores observaron, además, productos de PCR de longitud completa y de longitud parcial correspondientes a ADN genómico no cortado (fig. 5, flechas verde y azul, respectivamente). Los últimos es probable que sean una amplificación parcial de la región del ADN genómico completo del paciente de DM1, ya que las repeticiones de 200 CTG no pueden amplificarse fácilmente mediante PCR.

Globalmente estos resultados demuestran que el sistema Cas9-ARNgm descrito resulta adecuado para escindir las repeticiones de CTG de la región 3'UTR del gen *DMPK* humano.

40

LISTA DE SECUENCIAS

	<110> GENETHON et al.	
5	<120> COMPOSICIONES Y MÉTODOS PARA EL TRATAMIENTO DE TRASTORNOS POR EXPANSIÓN DE REPETICIONES DE NUCLEÓTIDOS	
	<130> B2019PC00	
	<160> 14	
10	<170> PatentIn versión 3.5	
	<210> 1	
	<211> 24	
15	<212> ARN	
	<213> artificial	
	<220>	
20	<223> secuencia de ARNgm 1 guía	
	<400> 1	
	gcgcucccug aacccuagaa cugu	24
	<210> 2	
25	<211> 24	
	<212> ARN	
	<213> artificial	
	<220>	
30	<223> secuencia de ARNgm 7 guía	
	<400> 2	
	acggggcucg aaggguccuu guag	24
35	<210> 3	
	<211> 24	
	<212> ARN	
	<213> artificial	
40	<220>	
	<223> secuencia de ARNgm E guía	
	<400> 3	
45	uggggagcgu cuggcgcgau cucu	24
	<210> 4	
	<211> 24	
	<212> ARN	
	<213> artificial	
50	<220>	
	<223> secuencia de ARNgm N guía	
	<400> 4	
55	gucggggucu cagugcaucc aaaa	24
	<210> 5	
	<211> 137	
	<212> ARN	
60	<213> artificial	
	<220>	
	<223> ARNgm 1	

ES 2 732 100 T3

	<400> 5		
	gcgcucccug aaccuagaa cuguguugua gcucccuuc gaaagaaccg uugcuacaa	60	
	aaggccgucu gaaaagaugu gccgcaacgc ucugcccuu aaagcuucug cuuaagggg	120	
	caucguuuau uuuuuuu	137	
5	<210> 6 <211> 138 <212> ARN <213> artificial		
10	<220> <223> ARNgm 7		
	<400> 6		
	gacggggcuc gaaggguccu uguagguugu agcuccuuu cgaaagaacc guugcuacaa	60	
	uaaggccguc ugaaaagaug ugccgcaacg cucugcccu uaaagcuucu gcuuaaggg	120	
	gcaucguua uuuuuuu	138	
15	<210> 7 <211> 138 <212> ARN <213> artificial		
20	<220> <223> ARNgm E		
	<400> 7		
	guggggagcg ucuggcgcga ucucuguugu agcuccuuu cgaaagaacc guugcuacaa	60	
	uaaggccguc ugaaaagaug ugccgcaacg cucugcccu uaaagcuucu gcuuaaggg	120	
25	gcaucguua uuuuuuu	138	
	<210> 8 <211> 137 <212> ARN <213> artificial		
30	<220> <223> ARNgm N		
35	<400> 8		
	gucggggucu cagugcaucc aaaaguugua gcucccuuc gaaagaaccg uugcuacaa	60	
	aaggccgucu gaaaagaugu gccgcaacgc ucugcccuu aaagcuucug cuuaagggg	120	
	caucguuuau uuuuuuu	137	
	<210> 9 <211> 263 <212> ADN <213> artificial		
40	<220> <223> promotor U6		
45	<400> 9		

ES 2 732 100 T3

	aggctcgggca ggaagagggc ctatttccca tgattccttc atatttgcat atacgataca	60
	aggctgttag agagataatt agaattaatt tgactgtaaa cacaaagata ttagtacaaa	120
	atacgtgacg tagaaagtaa taatttcttg ggtagtttgc agttttaaaa ttatgtttta	180
	aatggacta tcatatgctt accgtaactt gaaagtattt cgatttcttg gctttatata	240
	tcttgtggaa aggacgaaac acc	263
5	<210> 10 <211> 402 <212> ADN <213> artificial	
10	<220> <223> casete de ARNgm 1	
	<400> 10	
	aggctcgggca ggaagagggc ctatttccca tgattccttc atatttgcat atacgataca	60
	aggctgttag agagataatt agaattaatt tgactgtaaa cacaaagata ttagtacaaa	120
	atacgtgacg tagaaagtaa taatttcttg ggtagtttgc agttttaaaa ttatgtttta	180
	aatggacta tcatatgctt accgtaactt gaaagtattt cgatttcttg gctttatata	240
	tcttgtggaa aggacgaaac accgcgctcc ctgaacccta gaactgtgtt gtagctccct	300
	ttcgaaagaa ccgttgctac aataaggccg tctgaaaaga tgtgccgcaa cgctctgccc	360
	cttaaagctt ctgctttaag gggcatcgtt tatttttttt aa	402
15	<210> 11 <211> 403 <212> ADN <213> artificial	
20	<220> <223> casete de ARNgm 7	
	<400> 11	
	aggctcgggca ggaagagggc ctatttccca tgattccttc atatttgcat atacgataca	60
	aggctgttag agagataatt agaattaatt tgactgtaaa cacaaagata ttagtacaaa	120
	atacgtgacg tagaaagtaa taatttcttg ggtagtttgc agttttaaaa ttatgtttta	180
	aatggacta tcatatgctt accgtaactt gaaagtattt cgatttcttg gctttatata	240
	tcttgtggaa aggacgaaac accgacgggg ctggaagggt ccttgtaggt ttagctccc	300
	tttcgaaaga accgttgcta caataaggcc gtctgaaaag atgtgccgca acgctctgcc	360
25	ccttaaagct tctgctttaa gggcatcgtt ttattttttt taa	403
30	<210> 12 <211> 403 <212> ADN <213> artificial	
	<220> <223> casete de ARNgm E	
35	<400> 12	

ES 2 732 100 T3

	aggtcgggca ggaagagggc ctatttccca tgattccttc atatttgcat atacgataca	60
	aggctgttag agagataatt agaattaatt tgactgtaaa cacaaagata ttagtacaaa	120
	atacgtgacg tagaaagtaa taatttcttg ggtagtttgc agttttaaaa ttatgtttta	180
	aatggacta tcatatgctt accgtaactt gaaagtattt cgatttcttg gctttatata	240
	tcttgtggaa aggacgaaac accgtgggga gcgtctggcg cgatctctgt ttagctccc	300
	tttcgaaaga accgttgcta caataaggcc gtctgaaaag atgtgccgca acgctctgcc	360
	ccttaaagct tctgctttaa ggggcatcgt ttattttttt taa	403
5	<210> 13 <211> 402 <212> ADN <213> artificial	
10	<220> <223> casete de ARNgm N	
	<400> 13	
	aggtcgggca ggaagagggc ctatttccca tgattccttc atatttgcat atacgataca	60
	aggctgttag agagataatt agaattaatt tgactgtaaa cacaaagata ttagtacaaa	120
	atacgtgacg tagaaagtaa taatttcttg ggtagtttgc agttttaaaa ttatgtttta	180
	aatggacta tcatatgctt accgtaactt gaaagtattt cgatttcttg gctttatata	240
	tcttgtggaa aggacgaaac accgtcgggg tctcagtgca tccaaaagtt gtagctccct	300
	ttcgaaagaa ccgttgctac aataaggccg tctgaaaaga tgtgccgcaa cgctctgccc	360
	cttaaagctt ctgctttaag gggcatcgtt tatttttttt aa	402
15	<210> 14 <211> 932 <212> ADN <213> homo sapiens	
	<400> 14	
	gtccctaggc ctggcctatc ggaggcgctt tccctgctcc tgttcgccgt tgttctgtct	60
	cgtgccgccg ccctgggctg cattgggttg gtggcccacg ccggccaact caccgcagtc	120
20	tggcgccgcc caggagccgc ccgcgctccc tgaaccctag aactgtcttc gactccgggg	180

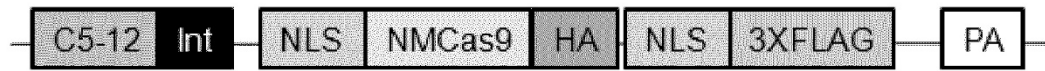
ES 2 732 100 T3

ccccgttgga agactgagtg cccgggggcac ggcacagaag ccgcgcccac cgcctgccag	240
ttcacaaccg ctccgagcgt gggctctccgc ccagctccag tcctgtgatc cgggcccgcc	300
ccctagcggc cggggaggga ggggccgggt ccgcggccgg cgaacggggc tcgaagggtc	360
cttgtagccg ggaatgctgc tgctgctgct gctgctgctg ctgctgctgc tgctgctgct	420
gctgctgctg ctgctggggg gatcacagac catttctttc tttcggccag gctgaggccc	480
tgacgtggat gggcaaactg caggcctggg aaggcagcaa gccgggcccgt ccgtgttcca	540
tcctccacgc acccccacct atcgttggtt cgcaaagtgc aaagctttct tgtgcatgac	600
gccctgctct ggggagcgtc tggcgcgatc tctgcctgct tactcgggaa atttgctttt	660
gccaaaccgc ctttttcggg gatcccgcg cccctcctc acttgcgctg ctctcggagc	720
cccagccggc tccgcccgct tcggcggttt ggatatttat tgacctcgtc ctccgactcg	780
ctgacaggct acaggacccc caacaacccc aatccacgtt ttggatgcac tgagaccccg	840
acattcctcg gtatttattg tctgtcccca cctaggaccc ccacccccga ccctcgcgaa	900
taaaaggccc tccatctgcc caaagctctg ga	932

REIVINDICACIONES

1. Una pareja de moléculas de ARNgm, en donde dicha pareja comprende una primera y una segunda moléculas de ARNgm que son capaces de unir mediante apareamiento de bases un complemento de secuencia a una secuencia diana de ADN genómico que está localizada en el lado 5' y 3' de una expansión de repeticiones de nucleótidos localizada dentro de una región no codificante de un gen de interés, respectivamente, resultando apropiada de esta manera para escindir dicha expansión de nucleótidos con un sistema CRISPR/Cas9,
5
en donde el gen de interés es el gen *DMPK*,
en donde la pareja de moléculas de ARNgm es una pareja de:
10
 - un ARNgm que comprende una secuencia guía seleccionada de SEQ ID NO:1 y SEQ ID NO:2 y
 - un ARNgm que comprende una secuencia guía seleccionada de SEQ ID NO:3 y SEQ ID NO:4.
2. La pareja de ARNgm según la reivindicación 1, en donde la pareja de moléculas de ARNgm es una pareja de:
15
 - un ARNgm seleccionado del grupo que consiste en SEQ ID NO:5 o SEQ ID NO:6, y
 - un ARNgm seleccionado del grupo que consiste en SEQ ID NO:7 o SEQ ID NO:8.
3. Un ARNgm que comprende una secuencia que es capaz de unir mediante apareamiento de bases el complemento de secuencia a una secuencia diana de ADN genómico que está localizada en el lado 5' o 3' de una expansión de repeticiones de nucleótidos dentro de una región no codificante de un gen de interés,
20en donde el gen de interés es el gen *DMPK*,
en donde dicho ARNgm comprende una secuencia guía seleccionada del grupo que consiste en SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 y SEQ ID NO:4.
4. El ARNgm según la reivindicación 3, en donde dicho ARNgm se selecciona del grupo que consiste en SEQ ID NO:5 a SEQ ID NO:8.
5. Un vector codificante del ARNgm o una pareja de moléculas de ARNgm según cualquiera de las reivindicaciones 1 a 4, siendo el vector preferiblemente un plásmido o un vector vírico, tal como un vector VAAr.
25
6. Una célula diana, que está transfectada o transducida con el vector según la reivindicación 5.
7. Un método para la producción de un ARNgm o una pareja de ARNgm, que comprende cultivar la célula diana según la reivindicación 6 bajo condiciones que permiten la producción de dicho ARNgm o pareja de ARNgm, y recuperar dicho ARNgm o dicha pareja de moléculas de ARNgm a partir de dicha etapa de cultivo.
30
8. Un método *in vitro* para escindir una repetición de nucleótidos localizada dentro de una región no codificante de un gen *DMPK* en una célula, que comprende introducir en dicha célula una pareja de moléculas de ARNgm según la reivindicación 1 o 2, o un vector según la reivindicación 5, y una endonucleasa CRISPR/Cas9 derivada de *N. meningitidis*.
- 35 9. Una pareja de ARNgm según la reivindicación 1 o 2, o un vector según la reivindicación 5, para el uso en combinación con una endonucleasa Cas9 derivada de *N. meningitidis* en un método para tratar la distrofia miotónica de tipo 1.
10. La pareja de ARNgm para el uso según la reivindicación 9, en donde la expansión de repeticiones de nucleótidos es una expansión de repeticiones de binucleótido, trinucleótido, tetranucleótido, pentanucleótido o hexanucleótido localizada dentro de la 3'-UTR del gen *DMPK*, en particular una expansión de repeticiones de trinucleótido.
40

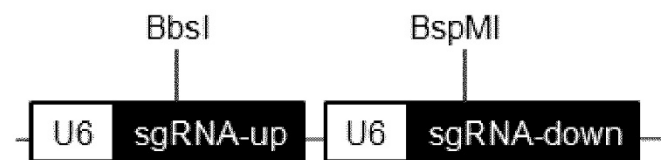
A. C512-NMCas9



B. EFS-NMCas9



C. Doble-ARNgmU6-NM



D. DesGFP-Doble-ARNgmU6

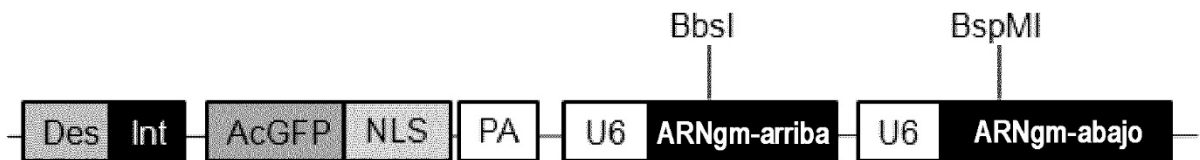


Figura 1

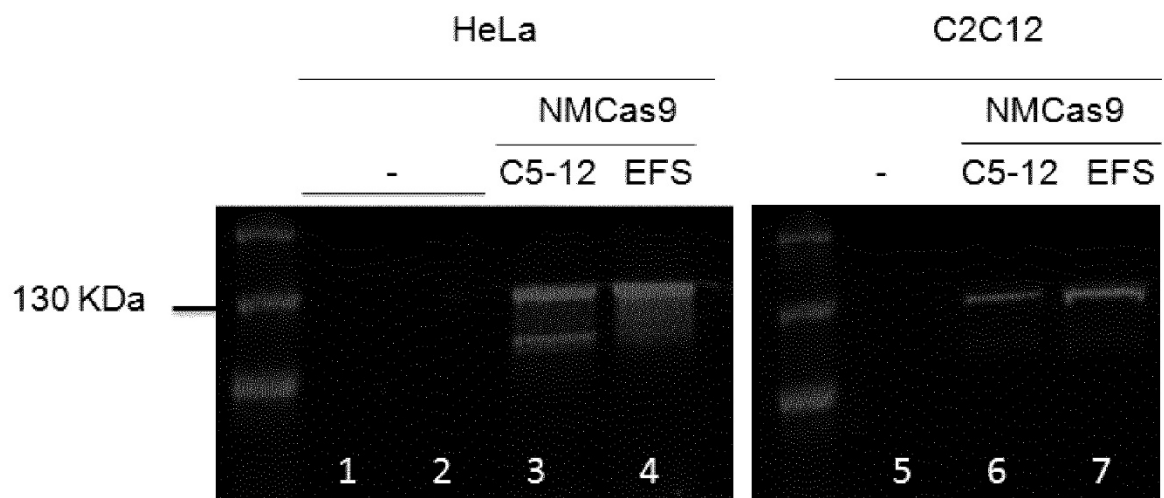


Figura 2

Región genómica* circundante a la repetición de CTG de la 3'UTR de *DMPK* y localización de los ARNm y motivos contiguos a protoespaciador (PAM)

***Referencia del genoma humano:
feb.2009 (GRCh37/hg19)
cr.19, cadena negativa
posición: 46.272.967(inicio) - 46.273.898(final)**

LEYENDA	
NNNNGATT (AATC/NNNN en la cadena negativa)	PAM de consenso para <i>Neisseria meningitidis</i> (NM) sistema CRISPR/Cas9 ¹
NNNNGANT y NNNNGNTT	PAM alternativos ²
TGA	codón de parada del gen <i>DMPK</i>
AATAAAA	poliA gen <i>DMPK</i>

¹Zhang Y et al, Mol Cell 2013

²Esvelt KM et al, Nat Methods 2013

F1-DMPK-3' UTR

[illegible]

Figura 3

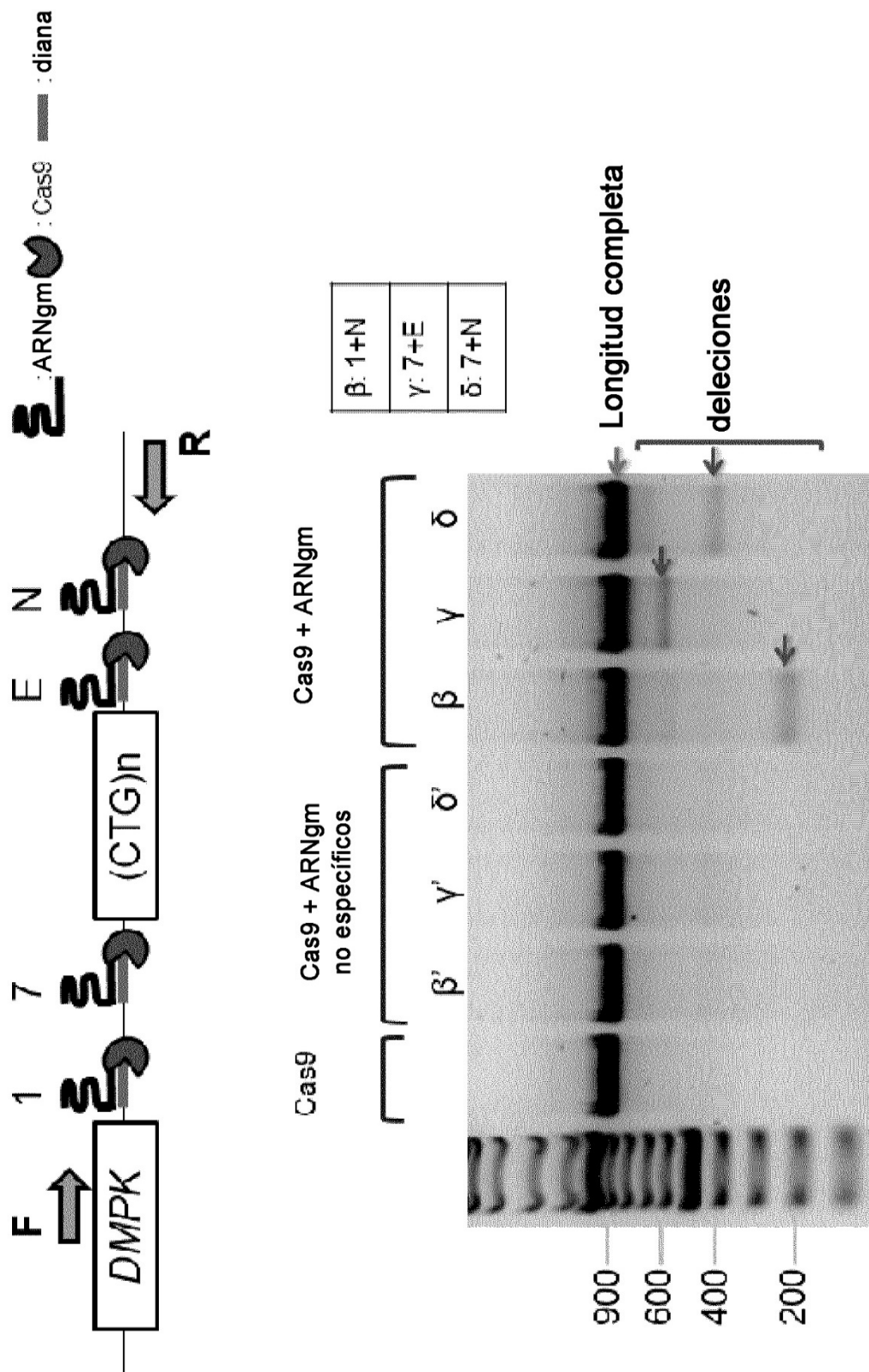


Figura 4

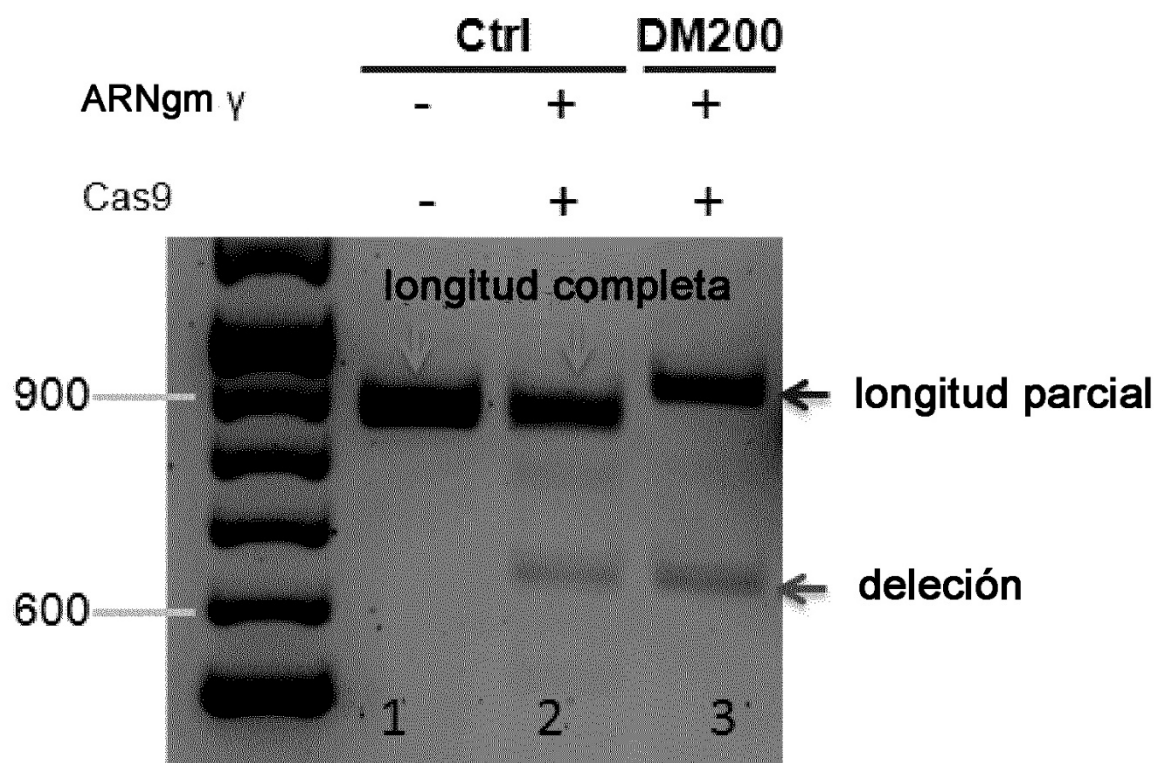


Figura 5

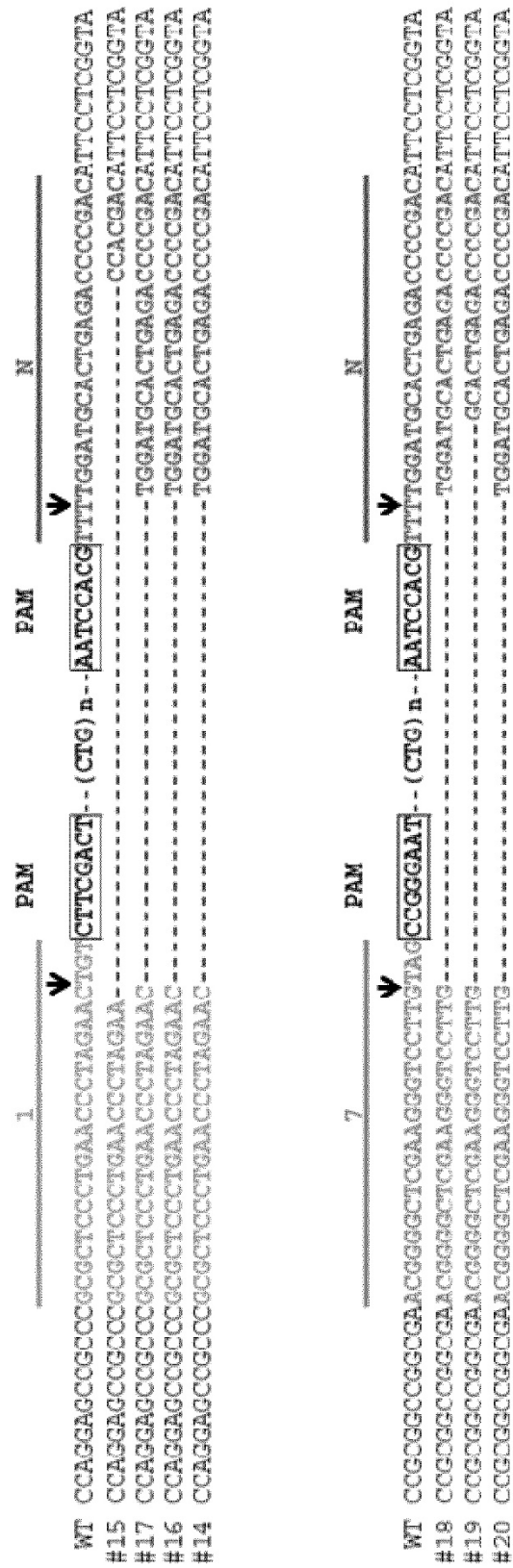


Figura 6