

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 997 490**

51 Int. Cl.:

C12N 9/22 (2006.01)

C12N 15/10 (2006.01)

C12N 15/113 (2010.01)

C12N 15/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **11.07.2018 E 21193940 (0)**

97 Fecha y número de publicación de la concesión europea: **18.09.2024 EP 3988656**

54 Título: **Uso de dominios de proteína que interactúan con el nucleosoma para potenciar la modificación dirigida del genoma**

30 Prioridad:

11.07.2017 US 201762531222 P

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

17.02.2025

73 Titular/es:

SIGMA-ALDRICH CO. LLC (100.00%)

3050 Spruce Street

St. Louis, MO 63103, US

72 Inventor/es:

CHEN, FUQIANG;

DING, XIAO;

FENG, YONGMEI y

DAVIS, GREGORY D

74 Agente/Representante:

SÁEZ MAESO, Ana

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 997 490 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

Uso de dominios de proteína que interactúan con el nucleosoma para potenciar la modificación dirigida del genoma

Campo

- 5 La presente divulgación se refiere a composiciones y métodos para aumentar la eficiencia de la modificación genómica dirigida, la regulación transcripcional dirigida o la modificación epigenética dirigida.

Antecedentes

- 10 Las endonucleasas programables se han convertido cada vez más en una herramienta importante para la ingeniería o modificación genómica dirigida en eucariotas. Recientemente, los sistemas de repeticiones palindrómicas cortas agrupadas y regularmente intercaladas (CRISPR) guiados por ARN han surgido como una nueva generación de herramientas de modificación del genoma. Estas nuevas endonucleasas programables proporcionaron una simplicidad y versatilidad sin precedentes en comparación con las generaciones anteriores de nucleasas, tales como las nucleasas de dedo de zinc (ZFN) y las nucleasas efectoras similares a activadores de la transcripción (TALEN). Sin embargo, las barreras de cromatina en las células eucariotas pueden dificultar el acceso a la diana y la escisión por los sistemas CRISPR derivados de procariontes (Hinz et al, Biochemistry, 2015, 54:7063-66; Horlbeck et al., eLife, 2016, 5:e12677).

- 20 De hecho, no se ha observado actividad de edición o se ha observado una actividad de edición baja en determinados sitios genómicos de mamíferos cuando se utilizan *Streptococcus pyogenes* Cas9 (SpCas9), que se considera la nucleasa CRISPR más activa hasta la fecha. Además, muchas de las nucleasas CRISPR que se han caracterizado hasta ahora no muestran actividad en células de mamíferos, aunque sí son activas en bacterias o en sustratos de ADN purificado. Por lo tanto, existe la necesidad de mejorar la capacidad de los sistemas de nucleasas CRISPR y otras proteínas de modificación de ADN programables para superar los obstáculos de la cromatina y aumentar la eficiencia de la modificación genómica o epigenética dirigida en eucariotas.

Sumario

- 30 Entre los diversos aspectos de la presente divulgación se encuentra la provisión de una proteína de fusión que comprende una proteína de repeticiones palindrómicas cortas agrupadas y regularmente intercaladas (CRISPR) unida a al menos un dominio de proteína que interactúa con el nucleosoma, en la que la proteína CRISPR es una nucleasa o nickasa CRISPR/Cas9 de tipo II, o la proteína CRISPR es una nucleasa o nickasa CRISPR/Cpf1 de tipo V, y en la que la proteína de fusión comprende además al menos una señal de localización nuclear, y en la que el al menos un dominio de proteína que interactúa con el nucleosoma es un dominio de unión a ADN de caja de grupo de alta movilidad (HMG) (HMGB), una proteína de unión a nucleosomas HMG (HMGN), un dominio globular central de una variante de histona H1, un dominio de unión a ADN de una proteína del complejo de remodelación de cromatina, o una combinación de los mismos. En determinadas realizaciones, la proteína CRISPR puede ser *Streptococcus pyogenes* Cas9 (SpCas9), *Streptococcus thermophilus* Cas9 (StCas9), *Streptococcus pasteurianus* (SpaCas9), *Campylobacter jejuni* Cas9 (CjCas9), *Staphylococcus aureus* (SaCas9), *Francisella novicida* Cas9 (FnCas9), *Neisseria cinerea* Cas9 (NcCas9), *Neisseria meningitidis* Cas9 (NmCas9), *Francisella novicida* Cpf1 (FnCpf1), *Acidimicrobium* sp. Cpf1 (AsCpf1), o bacteria *Lachnospiraceae* ND2006 Cpf1 (LbCpf1).

- 40 En otras realizaciones, la proteína CRISPR tiene actividad no nucleasa. En tales iteraciones, la proteína CRISPR puede ser una proteína CRISPR/Cas9 de tipo II modificada para carecer de toda actividad de nucleasa y unida a un dominio no nucleasa, o una proteína CRISPR/Cpf1 de tipo V modificada para carecer de toda actividad de nucleasa y unida a un dominio no nucleasa, en la que el dominio no nucleasa puede tener actividad de citosina desaminasa, actividad de histona acetiltransferasa, actividad de activación transcripcional o actividad de represor transcripcional.

- 50 En determinadas realizaciones, el al menos un dominio de proteína que interactúa con el nucleosoma de la proteína de fusión CRISPR puede ser el dominio de caja A de HMGB1, la proteína HMGN1, la proteína HMGN2, la proteína HMGN3a, la proteína HMGN3b, el dominio globular central de la histona H1, el dominio de unión al ADN de la proteína de cambio de imitación (ISWI), el dominio de unión al ADN de la proteína 1 de cromodominio-helicasa-ADN (CHD1) o una combinación de los mismos.

- 55 El al menos un dominio de proteína que interactúa con el nucleosoma se puede unir a la proteína CRISPR directamente mediante un enlace químico, indirectamente mediante un enlazante o una combinación de ambos. El al menos un dominio de proteína que interactúa con el nucleosoma se puede unir al extremo-N, al extremo-C y/o a una ubicación interna de la proteína CRISPR. En algunas realizaciones, la proteína de fusión comprende al menos dos dominios de proteína que interactúan con el nucleosoma vinculados a la proteína CRISPR.

Las proteínas de fusión CRISPR divulgadas en el presente documento pueden comprender además al menos un dominio de penetración celular, al menos un dominio marcador o una combinación de los mismos.

En determinadas realizaciones, la proteína de fusión CRISPR puede tener una secuencia de aminoácidos que tiene al menos aproximadamente un 90 % de identidad de secuencia con SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, o SEQ ID NO:79.

En otras realizaciones, la proteína de fusión CRISPR puede tener una secuencia de aminoácidos como se establece en SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, o SEQ ID NO:79.

Otro aspecto de la presente divulgación abarca complejos proteína-ARN que comprenden al menos una de las proteínas de fusión que contienen CRISPR divulgadas en el presente documento y al menos un ARN guía.

Un aspecto adicional de la presente divulgación proporciona ácidos nucleicos que codifican cualquiera de las proteínas de fusión divulgadas en el presente documento. Los ácidos nucleicos pueden optimizarse mediante codones para su traducción en una célula eucariota. En algunas realizaciones, los ácidos nucleicos pueden ser parte de un vector tal como, por ejemplo, un vector viral, un vector plasmídico o un ARN autorreplicante.

Otro aspecto adicional de la presente divulgación proporciona un método para aumentar la eficiencia de la modificación genómica o epigenética dirigida en una célula eucariota, comprendiendo el método la introducción en la célula eucariota: (a) al menos una proteína de fusión CRISPR como se divulga en el presente documento o un ácido nucleico que codifica dicha(s) proteína(s) de fusión CRISPR, en la que la proteína CRISPR (i) tiene actividad de nucleasa o nickasa o (ii) está modificada para carecer de toda actividad de nucleasa y está vinculada a un dominio que no es de nucleasa; y (b) al menos un ARN guía o un ácido nucleico que codifica al menos un ARN guía; en la que la proteína CRISPR de la al menos una proteína de fusión CRISPR está dirigida a una secuencia cromosómica diana y el al menos un dominio de proteína que interactúa con el nucleosoma de la al menos una proteína de fusión CRISPR altera la estructura nucleosómica o de cromatina de tal manera que la al menos una proteína de fusión CRISPR tiene un mayor acceso a la secuencia cromosómica diana, aumentando así la eficiencia de la modificación genómica o epigenética dirigida, en la que el método no comprende un método para el tratamiento del cuerpo humano o animal mediante cirugía o terapia.

En determinadas realizaciones, los métodos pueden comprender además la introducción en la célula eucariota al menos un polinucleótido donante, comprendiendo el polinucleótido donante al menos una secuencia donante.

Las células eucariotas utilizadas en los métodos divulgados en el presente documento pueden ser células de mamíferos. En algunas realizaciones, las células pueden ser células humanas. Las células pueden ser *in vitro* o *in vivo*.

A continuación se detallan otros aspectos y características de la divulgación.

Breve descripción de los dibujos

La figura 1 presenta la eficiencia de escisión (como porcentaje de indel) de CjCas9 de tipo salvaje (CjCas9), una proteína de fusión que comprende CjCas9 unida a HMGN1 y HMGB1 box A (CjCas9-HN1 HB1), y una proteína de fusión que comprende CjCas9 unida a HMGN1 y al dominio globular central de la histona H1 (CjCas9-HN1H1G) en presencia de un andamio de ARN_g de tipo salvaje o un andamio de ARN_g modificado.

Descripción detallada

La presente divulgación proporciona composiciones y métodos para aumentar la accesibilidad del ADN cromosómico a las proteínas de modificación de ADN programables que son sistemas CRISPR. En particular, la presente divulgación proporciona proteínas de fusión que comprenden al menos un dominio de proteína que interactúa con el nucleosoma unido a una proteína de modificación de ADN programable que es una proteína CRISPR. Los dominios de proteína que interactúan con el nucleosoma alteran o remodelan la estructura nucleosómica y/o de la cromatina de modo que la proteína de modificación de ADN programable tiene mayor acceso a secuencias cromosómicas dirigidas, aumentando así la eficiencia de la modificación genómica dirigida, la regulación transcripcional dirigida o la modificación epigenética dirigida.

(I) Proteínas de fusión

Un aspecto de la presente divulgación proporciona proteínas de fusión, en la que cada proteína de fusión comprende al menos un dominio de proteína que interactúa con el nucleosoma unido a una proteína de modificación de ADN programable. La proteína de modificación de ADN programable puede tener actividad

nucleasa (véase la sección (I)(b)(i), a continuación) o actividad no nucleasa (véase la sección (I)(b)(ii) a continuación). Los dominios de proteínas que interactúan con los nucleosomas se describen a continuación en la sección (I)(a) y los enlaces entre los dominios se describen a continuación en la sección (I)(c).

(a) Dominios de proteína que interactúan con el nucleosoma

- 5 Los dominios de proteína que interactúan con el nucleosoma se refieren a proteínas cromosómicas o fragmentos de las mismas que interactúan con nucleosomas y/o proteínas cromosómicas para facilitar el reordenamiento de nucleosomas y/o la remodelación de la cromatina. En algunas realizaciones, el dominio de proteína que interactúa con el nucleosoma puede derivarse de proteínas de caja de grupo de alta movilidad (HMG) (HMGB). En otras realizaciones, el dominio de proteína que interactúa con el nucleosoma pueden ser
- 10 proteínas de unión a nucleosomas HMG (HMGN) o fragmentos de las mismas. En otras realizaciones, el dominio de proteína que interactúa con el nucleosoma se puede derivar de variantes de la histona H1 enlazante. En otras realizaciones más, el dominio de proteína que interactúa con el nucleosoma puede derivarse de proteínas de complejo de remodelación de la cromatina.

(i) Proteínas HMGB

- 15 En algunas realizaciones, el al menos un dominio de proteína que interactúa con el nucleosoma puede derivarse de una proteína HMGB. Las proteínas HMGB interactúan con los nucleosomas y otras proteínas cromosómicas para regular la estructura y la función de la cromatina. Las proteínas HMGB adecuadas incluyen HMGB1 de mamíferos, HMGB2 de mamíferos y HMGB3 de mamíferos. Por ejemplo, el dominio de proteína que interactúa con el nucleosoma puede derivarse de un HMGB1 humano (RefSeqGene, U51677), HMGB2
- 20 humano (RefSeqGene, M83665) o HMGB3 humano (RefSeqGene, NM_005342). En otras realizaciones, el dominio de proteína que interactúa con el nucleosoma puede derivarse de una proteína HMGB o una proteína similar a HMGB de otros vertebrados, invertebrados (por ejemplo, *Drosophila* DSP1), plantas, levadura u otros eucariotas de células individuales.

- 25 En realizaciones específicas, el al menos un dominio de proteína que interactúa con el nucleosoma puede ser un fragmento de una proteína HMGB. En particular, el fragmento de la proteína HMGB es un dominio de unión al ADN. Las proteínas HMGB por lo general contienen dos dominios de unión al ADN, que se denominan caja A y caja B. En algunas realizaciones, el dominio de interacción del nucleosoma puede ser un dominio de caja A o un dominio de caja B de una proteína HMGB. En realizaciones específicas, el dominio que interactúa con el nucleosoma puede ser un dominio de caja A de HMGB1, un dominio de caja A de HMGB2 o un dominio de
- 30 caja A de HMGB3.

(ii) Proteínas HMGN

- En otras realizaciones, el al menos un dominio de proteína que interactúa con el nucleosoma puede ser una proteína HMGN o un fragmento de la misma. Las proteínas HMGN son proteínas cromosómicas que modulan la estructura y función de la cromatina. Las proteínas HMGN de mamíferos adecuadas incluyen HMGN1,
- 35 HMGN2, HMGN3, HMGN4 y HMGN5. En diversas realizaciones, el dominio de proteína que interactúa con el nucleosoma puede ser HMGN1 humano (RefSeqGene, M21339), HMGN2 humano (RefSeqGene, X13546), HMGN3a humano o HMGN3b humano (RefSeqGene, L40357), HMGN4 humano (RefSeqGene, NM_030763), HMGN5 humano (RefSeqGene, NM_016710), un fragmento del mismo o un derivado del mismo. En otras realizaciones, el dominio de proteína que interactúa con el nucleosoma puede ser una proteína HMGN
- 40 no humana, un fragmento de los mismos o un derivado de los mismos. Las proteínas HMGN son proteínas relativamente pequeñas. De esta forma, toda la proteína HMGN se puede unir a la proteína de modificación de ADN programable. Sin embargo, en algunas realizaciones, un fragmento (por ejemplo, el dominio de unión al nucleosoma ubicado centralmente) de una proteína HMGN se puede unir a la proteína de modificación de ADN programable.

45 (iii) Variantes de la histona H1

- En otras realizaciones más, el al menos un dominio de proteína que interactúa con el nucleosoma se puede derivar de una variante de histona H1 enlazante. Por ejemplo, el dominio de proteína que interactúa con el nucleosoma puede ser un dominio globular central de una variante de la histona H1. Las variantes de la histona H1 se unen al ADN enlazante entre los nucleosomas y el dominio globular central (de aproximadamente 80
- 50 aminoácidos) se une al ADN enlazante en los sitios de entrada y salida de los nucleosomas cerca de la díada de nucleosomas. Las variantes de la histona H1 comprenden una gran familia de proteínas relacionadas con una especificidad distinta para los tejidos, las etapas de desarrollo y los organismos en los que se expresan. Por ejemplo, el ser humano y el ratón contienen 11 variantes de histona H1, el pollo tiene seis variantes (que se llaman histona H5), la rana tiene cinco variantes, el nematodo tiene ocho variantes, las especies de mosca
- 55 de la fruta tienen de una a tres variantes y el tabaco tiene seis variantes. En algunas realizaciones, la variante de histona H1 puede ser una variante humana como se muestra a continuación.

* Talbert et al., *Epigenetics & Chromatin*, 2012, 5:7.

Nombre de la proteína*	Símbolo genético	Acceso en UniProtKB
Histona H1.0	H1F0	P07305
Histona H1.1	HIST1H1A	Q02539
Histona H1.2	HIST1H1C	P16403
Histona H1.3	HIST1H1D	P16402
Histona H1.4	HIST1H1E	P10412
Histona H1.5	HIST1H1B	P16401
Histona H1.6 (específica del testículo)	HIST1H1T	P22492
Histona H1.7 (específica del testículo)	H1FNT	Q75WM6
Histona H1.8 (específica del ovocito)	H1FOO	Q8IZA3
Histona H1.9 (específica del testículo)	HILS1	P60008
Histona H 1.10	H1FX	Q92522

(iv) Proteínas del complejo de remodelación de la cromatina

En realizaciones adicionales, el al menos un dominio de proteína que interactúa con el nucleosoma puede derivarse de una proteína del complejo de remodelación de cromatina. Por ejemplo, el dominio de proteína que interactúa con el nucleosoma puede ser un dominio de unión al ADN de una proteína del complejo de remodelación de cromatina. Los complejos de remodelación de la cromatina son complejos enzimáticos de múltiples subunidades con la capacidad de remodelar la estructura de la cromatina. Estos complejos de remodelación utilizan la energía de la hidrólisis del ATP para mover, desestabilizar, expulsar o reestructurar los nucleosomas.

Los ejemplos de complejos de remodelación de cromatina incluyen SWI/SNF (SWItch/Sacarosa No-Fermentable), ISWI (Imitación SWItch), CHD (Unión cromodominio-helicasa-ADN), Mi-2/NuRD (Remodelación de nucleosomas y desacetilasa), complejos INO80, SWR1, y RSC. En diversas realizaciones, el dominio de proteína que interactúa con el nucleosoma puede derivarse de una ATPasa, una helicasa y/o una proteína de unión al ADN en el complejo de remodelación de cromatina. En algunas realizaciones, el dominio de proteína que interactúa con el nucleosoma puede derivarse de la ATPasa ISWI del complejo ISWI, la proteína de unión al ADN CHD1 del complejo CHD, la helicasa dependiente de ATP SMARCA4 o la ATPasa Snf2 del complejo SWI/SNF, la ATPasa Mi-2 α o la ATPasa Mi2- β del complejo Mi-1/NuRD, la ATPasa AAA 1 similar a RuvB o la ATPasa AAA2 similar a RuvB del complejo INO80, la ATPasa Swr1 del complejo SWR1 o la ATPasa Rsc1 o la ATPasa Rcs2 del complejo RSC. En realizaciones específicas, el dominio de proteína que interactúa con el nucleosoma puede ser un dominio de unión al ADN de la proteína ISWI o un dominio de unión al ADN de la proteína CHD1.

(b) Proteínas programables de modificación del ADN

Una proteína de modificación de ADN programable es una proteína dirigida a unirse a una secuencia específica en el ADN cromosómico, donde modifica el ADN o una proteína asociada con el ADN en la secuencia diana o cerca de ella. De este modo, una proteína de modificación de ADN programable comprende un dominio de unión de ADN programable y un dominio de modificación catalíticamente activo.

El dominio de unión al ADN de las proteínas de modificación de ADN programables es programable, lo que significa que puede diseñarse o fabricarse para reconocer y unirse a diferentes secuencias de ADN. En algunos ejemplos, la unión del ADN está mediada por interacciones entre la proteína de modificación del ADN y el ADN diana. De este modo, el dominio de unión al ADN se puede programar para unirse a una secuencia de ADN de interés mediante ingeniería de proteínas. En otros ejemplos, la unión del ADN está mediada por un ARN guía que interactúa con la proteína de modificación del ADN y el ADN diana. En tales casos, la proteína de unión al

ADN programable se puede dirigir a una secuencia de ADN de interés mediante el diseño del ARN guía apropiado.

Se pueden incluir una variedad de dominios de modificación en la proteína de modificación de ADN programable. En algunas realizaciones, el dominio de modificación tiene actividad nucleasa y puede escindir una o ambas cadenas de una secuencia de ADN de doble cadena. La rotura del ADN puede luego repararse mediante un procedimiento de reparación del ADN celular, tales como la unión de extremos no homólogos (NHEJ) o la reparación dirigida por homología (HDR), de modo que la secuencia de ADN se puede modificar mediante una delección, inserción y/o sustitución de al menos un par de bases. Los ejemplos de proteínas de modificación de ADN programables que tienen actividad de nucleasa incluyen, sin límite, nucleasas CRISPR (o nickasas), nucleasas de dedo de zinc, nucleasas efectoras similares a activadores de la transcripción, meganucleasas y un dominio de unión de ADN programable vinculado a un dominio de nucleasa. Las proteínas de modificación de ADN programables que tienen actividad nucleasa se detallan a continuación en la sección (I)(b)(i).

En otras realizaciones, el dominio de modificación de la proteína de modificación de ADN programable tiene actividad no nucleasa (por ejemplo, actividad de modificación epigenética o actividad de regulación transcripcional) de modo que la proteína de modificación de ADN programable modifica la estructura y/o actividad del ADN y/o proteína(s) asociada(s) con el ADN. De este modo, la proteína de modificación de ADN programable puede comprender un dominio de unión de ADN programable unido a un dominio no nucleasa unido. Tales proteínas se detallan a continuación en la sección (I)(b)(ii).

Las proteínas de modificación de ADN programables pueden comprender dominios de unión y/o modificación de ADN de tipo salvaje o de origen natural, versiones modificadas de dominios de unión y/o modificación de ADN de origen natural, dominios de unión y/o modificación de ADN sintéticos o artificiales y combinaciones de los mismos.

(i) Proteínas programables de modificación del ADN con actividad nucleasa

Los ejemplos de proteínas de modificación de ADN programables que tienen actividad de nucleasa incluyen, sin límite, nucleasas CRISPR, nucleasas de dedo de zinc, nucleasas efectoras de tipo activador de la transcripción, meganucleasas y dominios de nucleasa vinculados a dominios de unión de ADN programables.

Nucleasas *CRISPR*. Las nucleasas CRISPR pueden derivarse de una proteína CRISPR tipo I, tipo II (es decir, Cas9), tipo III, tipo V (es decir, Cpf1) o tipo VI (es decir, Cas13), que están presentes en diversas bacterias y arqueas. En ejemplos adicionales, la nucleasa CRISPR puede derivarse de un sistema CRISPR arqueal, un sistema CRISPR/CasX o un sistema CRISPR/CasY (Burstin et al., *Nature*, 2017, 542(7640):237-241). En diversas realizaciones, la nucleasa CRISPR puede ser de *Streptococcus* sp. (por ejemplo, *S. pyogenes*, *S. thermophilus*, *S. pasteurianus*), *Campylobacter* sp. (por ejemplo, *Campylobacter jejuni*), *Francisella* sp. (por ejemplo, *Francisella novicida*), *Acaryochloris* sp., *Acetohalobium* sp., *Acidaminococcus* sp., *Acidithiobacillus* sp., *Alicyclobacillus* sp., *Allochroomium* sp., *Ammonifex* sp., *Anabaena* sp., *Arthrospira* sp., *Bacillus* sp., *Burkholderiales* sp., *Caldicelulosiruptor* sp., *Candidatus* sp., *Clostridium* sp., *Crocospaera* sp., *Cyanothecae* sp., *Exiguobacterium* sp., *Finnegoldia* sp., *Ktedonobacter* sp., *Lachnospiraceae* sp., *Lactobacillus* sp., *Lyngbya* sp., *Marinobacter* sp., *Methanohalobium* sp., *Microscilla* sp., *Microcoleus* sp., *Microcystis* sp., *Natranaerobius* sp., *Neisseria* sp., *Nitrosococcus* sp., *Nocardiosis* sp., *Nodularia* sp., *Nostoc* sp., *Oscillatoria* sp., *Polaromonas* sp., *Pelotomaculum* sp., *Pseudoalteromonas* sp., *Petrotoga* sp., *Prevotella* sp., *Staphylococcus* sp., *Streptomyces* sp., *Streptosporangium* sp., *Synechococcus* sp., *Thermosiphon* sp., o *Verrucomicrobia* es.

La nucleasa CRISPR puede ser una proteína de tipo salvaje o de origen natural. Como alternativa, la nucleasa CRISPR se puede diseñar para tener una especificidad mejorada, una especificidad PAM alterada, menos efectos fuera de la diana, mayor estabilidad y similares.

En algunas realizaciones, la nucleasa CRISPR puede ser una proteína CRISPR/Cas 9 tipo II. Por ejemplo, la nucleasa CRISPR puede ser *Streptococcus pyogenes* Cas9 (SpCas9), *Streptococcus thermophilus* Cas9 (StCas9), *Streptococcus pasteurianus* (SpaCas9), *Campylobacter jejuni* Cas9 (CjCas9), *Staphylococcus aureus* (SaCas9), *Francisella novicida* Cas9 (FnCas9), *Neisseria cinerea* Cas9 (NcCas9), o *Neisseria meningitidis* Cas9 (NmCas9). En otras realizaciones, la nucleasa CRISPR puede ser una proteína CRISPR/Cpf1 de tipo V, por ejemplo, *Francisella novicida* Cpf1 (FnCpf1), *Acidaminococcus* sp. Cpf1 (AsCpf1), o bacteria *Lachnospiraceae* ND2006 Cpf1 (LbCpf1).

En general, la nucleasa CRISPR comprende al menos un dominio de nucleasa que tiene actividad endonucleasa. Por ejemplo, una nucleasa Cas9 comprende un dominio HNH, que escinde la cadena complementaria del ARN guía, y un dominio RuvC, que escinde la cadena no complementaria, una proteína Cpf1 comprende un dominio RuvC y un dominio NUC, y una nucleasa Cas13a comprende dos dominios HNEPN. En algunas realizaciones, ambos dominios de nucleasa están activos y la nucleasa CRISPR tiene actividad de escisión de doble cadena (es decir, escinde ambas cadenas de una secuencia de ácido nucleico de doble cadena). En otras realizaciones, uno de los dominios de nucleasa se inactiva mediante una o más

mutaciones y/o deleciones, y la variante CRISPR es una nickasa que escinde una cadena de una secuencia de ácido nucleico de doble cadena. Por ejemplo, una o más mutaciones en el dominio RuvC de la proteína Cas9 (por ejemplo, D10A, D8A, E762A y/o D986A) da como resultado una nickasa HNH que corta la cadena complementaria del ARN guía; y una o más mutaciones en el dominio HNH de la proteína Cas9 (por ejemplo, H840A, H559A, N854A, N856A y/o N863A) da como resultado una nickasa RuvC que corta la cadena no complementaria del ARN guía. Mutaciones comparables pueden convertir las nucleasas Cpf1 y Cas13a en nickasas.

(ii) Proteínas programables de modificación del ADN con actividad no nucleasa

En realizaciones alternativas, la proteína de modificación de ADN programable puede ser una proteína quimérica que comprende un dominio de unión de ADN programable unido a un dominio no nucleasa. En tales casos, el dominio de unión al ADN programable es una proteína CRISPR catalíticamente inactiva (muerta). Por ejemplo, la proteína CRISPR catalíticamente inactiva puede ser una Cas9 catalíticamente inactiva (muerta) (dCas9) en la que el dominio RuvC comprende una mutación D10A, D8A, E762A y/o D986A y el dominio HNH comprende una mutación H840A, H559A, N854A, N865A y/o N863A. Alternativamente, la proteína CRISPR catalíticamente inactiva puede ser una proteína Cpf1 catalíticamente inactiva (muerta) que comprende mutaciones comparables en los dominios de nucleasa.

En algunas realizaciones, el dominio no nucleasa de la proteína quimérica puede ser un dominio de modificación epigenética, que altera la estructura del ADN o la cromatina (y puede o no alterar la secuencia de ADN). Los ejemplos no limitantes de dominios de modificación epigenética adecuados incluyen aquellos con actividad de metiltransferasa de ADN (por ejemplo, citosina metiltransferasa), actividad de la ADN desmetilasa, desaminación del ADN (por ejemplo, citosina desaminasa, adenosina desaminasa, guanina desaminasa), aminación del ADN, actividad de oxidación del ADN, actividad de helicasa del ADN, actividad de histona acetiltransferasa (HAT) (por ejemplo, dominio HAT derivado de la proteína de unión a E1A p300), actividad de histona desacetilasa, actividad de histona metiltransferasa, actividad de histona desmetilasa, actividad de histona cinasa, actividad de histona fosfatasa, actividad de histona ubiquitina ligasa, actividad desubiquitinante de histonas, actividad de adenilación de histonas, actividad de deadenilación de histonas, actividad de SUMOilación de histonas, actividad de deSUMOilación de histonas, actividad de ribosilación de histonas, actividad de desribosilación de histonas, actividad de miristoilación de histonas, actividad de demiristoilación de histonas, actividad de citrulinación de histonas, actividad de alquilación de histonas, actividad de desalquilación de histonas o actividad de oxidación de histonas. En realizaciones específicas, el dominio de modificación epigenética puede comprender actividad de citidina desaminasa, actividad de histona acetiltransferasa o actividad de ADN metiltransferasa.

En otras realizaciones, el dominio de modificación no nucleasa de la proteína quimérica puede ser un dominio de activación transcripcional o un dominio represor transcripcional. Los dominios de activación transcripcional adecuados incluyen, sin límite, el dominio VP16 del virus del herpes simple, VP64 (que es un derivado tetramérico de VP16), VP160, dominios de activación p65 de NFκB, dominios de activación p53 1 y 2, dominios de activación de CREB (proteína de unión al elemento de respuesta de AMPc), dominios de activación E2A, dominio de activación del factor de choque térmico humano 1 (HSF1) o dominios de activación NFAT (factor nuclear de células T activadas). Los ejemplos no limitantes de dominios represores transcripcionales adecuados incluyen dominios represores tempranos de AMPc inducibles (ICER), dominios represores de caja asociada a Kruppel (KRAB), dominios represores ricos en glicina YY1, represores similares a Sp1, represores E(spl), represor IκB o represor de proteína 2 de unión a metil-CpG (MeCP2). Los dominios de activación transcripcional o represores transcripcionales pueden fusionarse genéticamente a la proteína de unión al ADN o unirse mediante interacciones no covalentes proteína-proteína, proteína-ARN o proteína-ADN.

En realizaciones particulares, el dominio no nucleasa de la proteína quimérica puede comprender actividad de citidina desaminasa, actividad de histona acetiltransferasa, actividad de activación transcripcional o actividad de represor transcripcional.

En algunas realizaciones, la proteína quimérica que tiene actividad no nucleasa puede comprender además al menos una etiqueta detectable. La etiqueta detectable puede ser un fluoróforo (por ejemplo, FAM, TMR, Cy3, Cy5, Texas Red, Oregon Green, Alexa Fluors, etiquetas Halo o un colorante fluorescente adecuado), una etiqueta de detección (por ejemplo, biotina, digoxigenina y similares), puntos cuánticos o partículas de oro.

(c) Enlaces

Las proteínas de fusión divulgadas en el presente documento comprenden al menos un dominio de proteína que interactúa con el nucleosoma unido a una proteína de modificación de ADN programable. El enlace entre el al menos un dominio de proteína que interactúa con el nucleosoma y la proteína de modificación de ADN programable puede ser directo mediante un enlace químico, o el enlace puede ser indirecto mediante un enlazante.

En algunas realizaciones, el al menos un dominio de proteína que interactúa con el nucleosoma se puede unir directamente a la proteína de modificación de ADN programable mediante un enlace covalente (por ejemplo, enlace peptídico, enlace éster y similares). Alternativamente, el enlace químico puede ser no covalente (por ejemplo, iónico, electrostático, de hidrógeno, hidrofóbico, interacciones de Van der o efectos π).

- 5 En otras realizaciones, el al menos un dominio de proteína que interactúa con el nucleosoma se puede unir a la proteína de modificación de ADN programable mediante un enlazante. Un enlazante es un grupo químico que conecta uno o más grupos químicos mediante al menos un enlace covalente. Los enlazantes adecuados incluyen aminoácidos, péptidos, nucleótidos, ácidos nucleicos, moléculas enlazantes orgánicas (por ejemplo, derivados de maleimida, N-etoxibencilimidazol, ácido bifenil-3,4',5-tricarboxílico, p-aminobenciloxycarbonilo y similares), enlazantes disulfuro y enlazantes poliméricos (por ejemplo, PEG). El enlazante puede incluir uno o
10 más grupos espaciadores que incluyen, pero no se limitan a, alquileo, alquilenilo, alquinileno, alquilo, alquenilo, alquinilo, alcoxi, arilo, heteroarilo, aralquilo, aralquenilo, aralquinilo y similares. El enlazante puede ser neutro o tener carga positiva o negativa. Además, el enlazante puede ser escindible de modo que el enlace covalente del enlazante que conecta el enlazante a otro grupo químico puede romperse o escindirse bajo
15 determinadas condiciones, incluyendo pH, temperatura, concentración de sal, luz, un catalizador o una enzima.

- En otras realizaciones más, el al menos un dominio de proteína que interactúa con el nucleosoma se puede unir a la proteína de modificación de ADN programable mediante enlazantes peptídicos. El enlazante peptídico puede ser un enlazante de aminoácidos flexible (por ejemplo, que comprende aminoácidos pequeños, no polares o polares). Los ejemplos no limitantes de enlazantes flexibles incluyen LEGGGS (SEQ ID NO: 1), TGSG
20 (SEQ ID NO: 2), GSGGGSG (SEQ ID NO: 3), (GGGS)₁₋₄ (SEQ ID NO: 4) y (Gly)₆₋₈ (SEQ ID NO: 5). Alternativamente, el enlazante peptídico puede ser un enlazante de aminoácidos rígido. Tales enlazantes incluyen (EAAAK)₁₋₄ (SEQ ID NO: 6), A(EAAAK)₂₋₅A (SEQ ID NO: 7), PAPAP (SEQ ID NO: 8) y (AP)₆₋₈ (SEQ ID NO: 9). En la técnica se conocen bien ejemplos de enlazantes adecuados y hay programas para diseñar enlazantes fácilmente disponibles (Crasto et al., Protein Eng., 2000, 13(5):309-312).

- 25 El al menos un dominio de proteína que interactúa con el nucleosoma se puede unir al extremo-N, al extremo-C y/o a una ubicación interna de la proteína de modificación de ADN programable. En algunas realizaciones, al menos un dominio de proteína que interactúa con el nucleosoma se puede unir al extremo-N de la proteína de modificación de ADN programable. En otras realizaciones, el al menos un dominio de proteína que interactúa con el nucleosoma se puede unir al extremo-C de la proteína de modificación de ADN programable. En otras
30 realizaciones más, al menos un dominio de proteína que interactúa con el nucleosoma puede estar unido al extremo-N y al menos un dominio de proteína que interactúa con el nucleosoma puede estar unido al extremo-C de la proteína de modificación de ADN programable.

- En algunas realizaciones, la proteína de fusión puede comprender un dominio de proteína que interactúa con el nucleosoma. En otras realizaciones, la proteína de fusión puede comprender dos dominios de proteína que interactúan con el nucleosoma. En otras realizaciones más, la proteína de fusión puede comprender tres dominios de proteína que interactúan con el nucleosoma. En realizaciones adicionales, la proteína de fusión puede comprender cuatro, cinco o más de cinco dominios de proteína que interactúan con el nucleosoma. Los uno o más dominios de proteína que interactúan con el nucleosoma pueden ser iguales o pueden ser diferentes.

- 40 En realizaciones en las que la proteína de fusión comprende dos o más dominios de proteína que interactúan con el nucleosoma, los dos o más dominios que interactúan con el nucleosoma pueden estar unidos a cualquiera de los extremos, a ambos extremos y/o a una ubicación interna de la proteína de modificación de ADN programable. Los dos o más dominios de proteína que interactúan con el nucleosoma pueden ser iguales o pueden ser diferentes. Por ejemplo, el complejo puede comprender al menos dos dominios de unión al ADN de HMG, al menos dos proteínas HMGN, al menos un dominio de unión al ADN de HMG y al menos una
45 proteína HMGN, al menos un dominio de unión al ADN de HMG o proteína HMGN y al menos un dominio central de una variante de histona H1, al menos un dominio de unión al ADN de HMG o proteína HMGN y al menos un dominio de una proteína del complejo de remodelación de cromatina, al menos un dominio de unión al ADN de HMG o proteína HMGN, al menos un dominio central de una variante de histona H1 y al menos un dominio de una proteína del complejo de remodelación de cromatina, y similares.

- 50 (d) Señal de localización nuclear, dominio de penetración celular y/o dominio marcador

Las proteínas de fusión divulgadas en el presente documento comprenden al menos una señal de localización nuclear y pueden comprender además al menos un dominio de penetración celular y/o un dominio marcador.

- Los ejemplos no limitantes de señales de localización nuclear incluyen PKKKRKV (SEQ ID NO:10), PKKKRRV (SEQ ID NO:11), KRPAATKKAGQAKKKK (SEQ ID NO:12), YGRKKRRQRRR (SEQ ID NO:13), RKKRRQRRR (SEQ ID NO:14), PAAKRVKLD (SEQ ID NO:15), RQRRNELKRSP (SEQ ID NO:16), VSRKRPRP (SEQ ID NO:17), PPKKARED (SEQ ID NO:18), PQPKKKPL (SEQ ID NO:19), SALIKKKKKMAP (SEQ ID NO:20), PKQKKRK (SEQ ID NO:21), RKLKKIKKL (SEQ ID NO:22), REKKKFLKRR (SEQ ID NO:23), KRGDEVGDGVDEVAKKSKK (SEQ ID NO:24), RKCLQAGMN-LEARKTKK (SEQ ID NO:25),
55

NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO:26), y
RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO:27).

- Ejemplos de dominios de penetración celular adecuados incluyen, sin limitación, GRKKRRQRRRPPQPKKKRKV (SEQ ID NO:28), PLSSIFSRIGDPPKKRKV (SEQ ID NO:29), GALFLGWLGAAGSTMGAPKKRKV (SEQ ID NO:30), GALFLGLGAAGSTMGAWSQPKKRKV (SEQ ID NO:31), KETWWETWWTEWSQPKKRKV (SEQ ID NO:32), YARAAARQARA (SEQ ID NO:33), THRLPRRRRRR (SEQ ID NO:34), GGRRARRRRRR (SEQ ID NO:35), RRQRRTSKLMKR (SEQ ID NO:36), GWTLSAGYLLGKINLKALAALAKKIL (SEQ ID NO:37), KALAWKALAKALAKALAKHLAKALAKA-LKCEA (SEQ ID NO:38), y RQIKIWFQNRRMKWKK (SEQ ID NO:39).
- Los dominios marcadores incluyen proteínas fluorescentes y etiquetas de purificación o epítopos. Las proteínas fluorescentes adecuadas incluyen, sin limitación, proteínas fluorescentes verdes (por ejemplo, GFP, eGFP, GFP-2, tagGFP, turboGFP, Emerald, Azami Green, Azami Green monomérico, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarillas (por ejemplo, YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), proteínas fluorescentes azules (por ejemplo, BFP, EBFP, EBFP2, Azurita, mKalama1, GFPuv, Zafiro, Zafiro T), proteínas fluorescentes cian (por ejemplo, ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes rojas (por ejemplo, mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred) y proteínas fluorescentes anaranjadas (por ejemplo, mOrange, mKO, Kusabira-Orange, Kusabira-Orange monomérico, mTangerine, tdTomato). Los ejemplos no limitantes de etiquetas de purificación o epítopos adecuados incluyen 6xHis, FLAG®, HA, GST, Myc y similares.

La al menos una señal de localización nuclear, dominio de penetración celular y/o dominio marcador se puede ubicar en el extremo-N, el extremo-C y/o en una ubicación interna de la proteína de fusión.

(e) Proteínas de fusión específicas

- En general, el al menos un dominio de proteína que interactúa con el nucleosoma de la proteína de fusión se elige entre el dominio de caja A de HMGB1, la proteína HMGN1, la proteína HMGN2, la proteína HMGN3a, la proteína HMGN3b, el dominio globular central de la histona H1, el dominio de unión al ADN de la proteína de cambio de imitación (ISWI), el dominio de unión al ADN de la proteína 1 de cromodominio-helicasa-ADN (CHD1) o combinaciones de los mismos.
- La proteína de modificación de ADN programable de la proteína de fusión es una proteína CRISPR. Por ejemplo, la proteína CRISPR puede ser *Streptococcus pyogenes* Cas9 (SpCas9), *Streptococcus thermophilus* Cas9 (StCas9), *Streptococcus pasteurianus* (SpaCas9), *Campylobacter jejuni* Cas9 (CjCas9), *Staphylococcus aureus* (SaCas9), *Francisella novicida* Cas9 (FnCas9), *Neisseria cinerea* Cas9 (NcCas9), *Neisseria meningitis* Cas9 (NmCas9), *Francisella novicida* Cpf1 (FnCpf1), *Acidminococcus* sp. Cpf1 (AsCpf1), o bacteria *Lachnospiraceae* ND2006 Cpf1 (LbCpf1).
- En algunas realizaciones, la proteína de fusión tiene una secuencia de aminoácidos que tiene al menos aproximadamente un 80 % de identidad de secuencia con cualquiera de las SEQ ID NOS: 61-79. En general, cualquier sustitución de aminoácidos es conservadora, es decir, limitado a los intercambios entre miembros del grupo 1: glicina, alanina, valina, leucina e isoleucina; grupo 2: serina, cisteína, treonina y metionina; grupo 3: prolina; grupo 4: fenilalanina, tirosina y triptófano; y grupo 5: aspartato, glutamato, asparagina y glutamina. En diversas realizaciones, la secuencia de aminoácidos de la proteína de fusión tiene al menos aproximadamente 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 o 99 % de identidad de secuencia con cualquiera de las SEQ ID NOS:61-79. En algunas realizaciones, la proteína de fusión tiene una secuencia de aminoácidos como se establece en SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, o SEQ ID NO:79.

(II) Complejos

- Otro aspecto de la presente divulgación abarca complejos que comprenden al menos un sistema CRISPR (es decir, proteína CRISPR y ARN guía) y al menos un dominio de proteína que interactúa con el nucleosoma. El al menos un dominio de proteína que interactúa con el nucleosoma está unido a la proteína CRISPR del sistema CRISPR (es decir, el complejo comprende una proteína de fusión CRISPR como se describe en la sección (I) anterior). En otros ejemplos, el al menos un dominio de proteína que interactúa con el nucleosoma se puede unir al ARN guía del sistema CRISPR. El enlace puede ser directo o indirecto, esencialmente como se describe anteriormente en la sección (I)(c). Por ejemplo, un dominio de proteína que interactúa con el nucleosoma se puede unir a una proteína de unión de aptámeros de ARN, y el ARN guía puede comprender secuencias de aptámeros, de modo que la unión de la proteína de unión de aptámeros de ARN a la secuencia de aptámeros de ARN vincula el dominio de proteína que interactúa con el nucleosoma al ARN guía.

Los dominios de proteínas que interactúan con los nucleosomas se describen anteriormente en la sección (I)(a), y las proteínas CRISPR se detallan anteriormente en la sección (I)(b). La proteína CRISPR puede tener actividad nucleasa o nickasa (por ejemplo, puede ser un CRISPR/Cas9 tipo II o un CRISPR/Cpf1 tipo V). Por ejemplo, un complejo puede comprender una nucleasa CRISPR, o un complejo puede comprender dos nickasas CRISPR. Alternativamente, la proteína CRISPR puede modificarse para que carezca de toda actividad de nucleasa y unirse a dominios no nucleasa (por ejemplo, dominios que tienen actividad de citosina desaminasa, actividad de acetiltransferasa de histona, actividad de activación transcripcional o actividad de represor transcripcional). En algunas realizaciones, el dominio no nucleasa también puede estar unido a una proteína de unión a aptámero de ARN.

Un ARN guía comprende (i) un ARN CRISPR (crRNA) que contiene una secuencia guía en el extremo 5' que se hibrida con una secuencia diana y (ii) una secuencia crRNA transaccional (tracrRNA) que interactúa con la proteína CRISPR. La secuencia guía de crRNA de cada ARN guía es diferente (es decir, es específica de la secuencia). La secuencia de tracrRNA es generalmente la misma en los ARN guía diseñados para formar complejos con una proteína CRISPR de una especie bacteriana particular.

La secuencia guía de crRNA está diseñada para hibridar con una secuencia diana (es decir, protoespaciador) que está bordeado por un motivo adyacente al protoespaciador (PAM) en una secuencia de doble cadena. Las secuencias PAM para las proteínas Cas9 incluyen 5'-NGG, 5'-NGGNG, 5'-NNAGAAW y 5'-ACAY, y las secuencias PAM para Cpf1 incluyen 5'-TTN (en la que N se define como cualquier nucleótido, W se define ya sea como A o T, e Y se define como C o T). En general, la complementariedad entre la secuencia guía de crRNA y la secuencia diana es al menos del 80 %, al menos del 85 %, al menos del 90 %, al menos del 95 % o al menos del 99 %. En realizaciones específicas, la complementariedad es completa (es decir, 100 %). En diversas realizaciones, la longitud de la secuencia guía de crRNA puede variar desde aproximadamente 15 nucleótidos hasta aproximadamente 25 nucleótidos. Por ejemplo, la secuencia guía de crRNA puede tener aproximadamente 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 o 25 nucleótidos de longitud. En realizaciones específicas, el crRNA puede tener una longitud de aproximadamente 19, 20, 21 o 22 nucleótidos.

El crRNA y el tracrRNA comprenden secuencias repetidas que forman una o más estructuras de un bucle de tallo, que pueden interactuar con la proteína CRISPR. La longitud de cada bucle y tallo puede variar. Por ejemplo, el uno o más bucles pueden variar en longitud desde aproximadamente 3 a aproximadamente 10 nucleótidos, y el uno o más tallos pueden variar en longitud desde aproximadamente 6 a aproximadamente 20 pares de bases. El uno o más tallos pueden comprender una o más protuberancias de 1 a aproximadamente 10 nucleótidos.

El crRNA puede variar en longitud desde aproximadamente 25 nucleótidos hasta aproximadamente 100 nucleótidos. En diversas realizaciones, el crRNA puede variar en longitud que varía desde aproximadamente 25 a aproximadamente 50 nucleótidos, desde aproximadamente 590 a aproximadamente 75 nucleótidos, o desde aproximadamente 75 a aproximadamente 100 nucleótidos. El tracrRNA puede variar en longitud desde aproximadamente 50 nucleótidos hasta aproximadamente 300 nucleótidos. En diversas realizaciones, el tracrRNA puede variar en longitud desde aproximadamente 50 a aproximadamente 90 nucleótidos, desde aproximadamente 90 a aproximadamente 110 nucleótidos, desde aproximadamente 110 a aproximadamente 130 nucleótidos, desde aproximadamente 130 a aproximadamente 150 nucleótidos, desde aproximadamente 150 a aproximadamente 170 nucleótidos, desde aproximadamente 170 a aproximadamente 200 nucleótidos, desde aproximadamente 200 a aproximadamente 250 nucleótidos, o desde aproximadamente 250 a aproximadamente 300 nucleótidos.

La secuencia de tracrRNA en el ARN guía generalmente se basa en la secuencia codificante del tracrRNA de tipo salvaje en la especie bacteriana de interés. En algunas realizaciones, la secuencia de tracrRNA de tipo salvaje (o la región de repetición constante de crRNA y la región 5' correspondiente del tracrRNA que forma una estructura dúplex con la región de repetición constante de crRNA) se puede modificar para facilitar la formación de la estructura secundaria, aumentar la estabilidad de la estructura secundaria, facilitar la expresión en células eucariotas, aumentar la eficiencia de edición, etcétera. Por ejemplo, se pueden introducir uno o más cambios de nucleótidos en la secuencia de ARN guía constante (véase el ejemplo 8, a continuación).

El ARN guía puede ser una sola molécula (es decir, un único ARN guía o ARN_{gu}), en el que la secuencia crRNA está unida a la secuencia tracrRNA. Alternativamente, el ARN guía puede ser dos moléculas separadas. Una primera molécula que comprende la secuencia guía de crRNA en el extremo 5' y una secuencia adicional en el extremo 3' que es capaz de aparear las bases con el extremo 5' de una segunda molécula, en la que la segunda molécula comprende una secuencia 5' que es capaz de aparear las bases con el extremo 3' de la primera molécula, así como una secuencia de tracrRNA adicional. En algunas realizaciones, el ARN guía de los sistemas CRISPR/Cpf1 tipo V puede comprender solo crRNA.

En algunas realizaciones, la una o más regiones de tallo-bucle del ARN guía se pueden modificar para comprender una o más secuencias de aptámeros (Konermann et al., *Nature*, 2015, 517(7536):583-588; Zalatan et al., *Cell*, 2015, 160(1-2):339-50). Los ejemplos de dominios de proteína de aptámero de ARN adecuados incluyen la proteína de cubierta MS2 (MCP), la proteína de cubierta del bacteriófago PP7 (PCP), la proteína

Com del bacteriófago Mu, la proteína N22 del bacteriófago lambda, la proteína de unión al tallo-bucle (SLBP), la proteína 1 relacionada con el síndrome de retraso mental del cromosoma X frágil (FXR1), proteínas derivadas del bacteriófago tales como AP205, BZ13, f1, f2, fd, fr, ID2, JP34/GA, JP501, JP34, JP500, KU1, M11, M12, MX1, NL95, PP7, ϕ Cb5, ϕ Cb8r, ϕ Cb12r, ϕ Cb23r, Q β , R17, SP- β , TW18, TW19 y VK, fragmentos de los mismos o derivados de los mismos. La longitud de la secuencia de aptámero adicional puede variar desde aproximadamente 20 nucleótidos hasta aproximadamente 200 nucleótidos.

El ARN guía puede comprender ribonucleótidos estándar, ribonucleótidos modificados (por ejemplo, pseudouridina), isómeros de ribonucleótidos y/o análogos de ribonucleótidos. En algunas realizaciones, el ARN guía puede comprender además al menos una etiqueta detectable. La etiqueta detectable puede ser un fluoróforo (por ejemplo, FAM, TMR, Cy3, Cy5, Texas Red, Oregon Green, Alexa Fluors, etiquetas Halo o un colorante fluorescente adecuado), una etiqueta de detección (por ejemplo, biotina, digoxigenina y similares), puntos cuánticos o partículas de oro. Los expertos en la técnica están familiarizados con el diseño y la construcción de gRNA, por ejemplo, las herramientas de diseño de gRNA están disponibles en internet o en fuentes comerciales.

El ARN guía se puede sintetizar químicamente, sintetizar enzimáticamente o mediante una combinación de los mismos. Por ejemplo, el ARN guía se puede sintetizar utilizando métodos de síntesis en fase sólida basados en fosforamidita estándar. Alternativamente, el ARN guía se puede sintetizar *in vitro* mediante la unión operativa del ADN que codifica el ARN guía a una secuencia de control del promotor que es reconocida por una ARN polimerasa del fago. Los ejemplos de secuencias promotoras de fagos adecuadas incluyen secuencias promotoras T7, T3, SP6 o variaciones de las mismas. En realizaciones en las que el ARN guía comprende dos moléculas separadas (es decir., crRNA y tracrRNA), el crRNA se puede sintetizar químicamente y el tracrRNA se puede sintetizar enzimáticamente.

En algunas realizaciones, el complejo puede comprender al menos un dominio de proteína que interactúa con el nucleosoma como se describió anteriormente en la sección (I)(a) unido a cualquiera de las proteínas de modificación de ADN programables como se describió anteriormente en la sección (I)(b).

(III) Ácidos nucleicos

Un aspecto adicional de la presente divulgación proporciona ácidos nucleicos que codifican las proteínas de fusión descritas anteriormente en la sección (I) y los complejos CRISPR descritos en la sección (II). Los complejos CRISPR pueden estar codificados por ácidos nucleicos individuales o múltiples. Los ácidos nucleicos pueden ser ADN o ARN, lineales o circulares, de cadena única o de doble cadena. El ARN o el ADN se pueden optimizar mediante codones para una traducción eficiente en proteína en la célula eucariota de interés. Los programas de optimización de codones están disponibles como software gratuito o de fuentes comerciales.

En algunas realizaciones, el ácido nucleico que codifica la proteína de fusión o los componentes proteicos del complejo CRISPR pueden ser ARN. El ARN puede sintetizarse enzimáticamente *in vitro*. Para ello, el ADN que codifica la proteína de interés se puede unir operativamente a una secuencia promotora que es reconocida por una ARN polimerasa del fago para la síntesis de ARN *in vitro*. Por ejemplo, la secuencia promotora puede ser una secuencia promotora T7, T3 o SP6 o una variación de una secuencia promotora T7, T3 o SP6. El ADN que codifica la proteína puede ser parte de un vector, como se detalla a continuación. En tales realizaciones, el ARN transcrito *in vitro* se puede purificar, adicionar la caperuza y/o poliadenilar. En otras realizaciones, el ARN que codifica la proteína de fusión o el componente proteico del complejo puede ser parte de un ARN autorreplicante (Yoshioka et al., Cell Stem Cell, 2013, 13:246-254). El ARN autorreplicante puede derivarse de un replicón de ARN del virus de la encefalitis equina venezolana (EEV) autorreplicante y no infeccioso, que es un ARN monocatenario de sentido positivo que es capaz de autorreplicarse durante un número limitado de divisiones celulares y que puede modificarse para codificar proteínas de interés (Yoshioka et al., Cell Stem Cell, 2013, 13:246-254).

En otras realizaciones, el ácido nucleico que codifica la proteína de fusión o la proteína CRISPR y el ARN guía del complejo puede ser ADN. La secuencia codificante de ADN se puede unir operativamente a al menos una secuencia de control del promotor para la expresión en la célula de interés. En determinadas realizaciones, la secuencia codificante de ADN se puede unir operativamente a una secuencia promotora para la expresión de la proteína o ARN en células de bacterias (por ejemplo, *E. coli*) o eucariotas (por ejemplo, células de levadura, de insecto o de mamífero). Los promotores bacterianos adecuados incluyen, sin límite, promotores T7, promotores *lac* del operón, promotores *trp*, promotores *tac* (que son híbridos de promotores *trp* y *lac*), variaciones de cualquiera de los anteriores y combinaciones de cualquiera de los anteriores. Los ejemplos no limitantes de promotores Pol II eucariotas adecuados incluyen promotores constitutivos, regulados o específicos de células o tejidos. Las secuencias de control del promotor constitutivo eucariota adecuadas incluyen, pero no se limitan a, el promotor temprano inmediato del citomegalovirus (CMV), el promotor del virus de los simios (SV40), el promotor tardío principal del adenovirus, el promotor del virus del sarcoma de Rous (RSV), el promotor del virus del tumor mamario del ratón (MMTV), el promotor de la fosfoglicerato cinasa (PGK), el promotor del factor de elongación (ED1)-alfa, promotores de ubiquitina, promotores de actina, promotores de tubulina, promotores de inmunoglobulina, fragmentos de los mismos o combinaciones de cualquiera de los

anteriores. Los ejemplos de secuencias de control de promotor reguladas eucariotas adecuadas incluyen, sin limitación, aquellos regulados por choque térmico, metales, esteroides, antibióticos o alcohol. Los ejemplos no limitantes de promotores específicos de tejido incluyen el promotor B29, el promotor CD14, el promotor CD43, el promotor CD45, el promotor CD68, el promotor de desmina, el promotor de elastasa-1, el promotor de endoglina, el promotor de fibronectina, el promotor Flt-1, el promotor GFAP, el promotor GPIIb, el promotor ICAM-2, el promotor INF- β , el promotor Mb, el promotor Nphsl, el promotor OG-2, el promotor SP-B, el promotor SYN1 y el promotor WASP. La secuencia promotora puede ser de tipo salvaje o puede modificarse para una expresión más eficiente o eficaz. En algunas realizaciones, la secuencia codificante de ADN también puede estar unida a una señal de poliadenilación (por ejemplo, señal poliA de SV40, señal poliA de la hormona de crecimiento bovino (BGH), *etc.*) y/o al menos una secuencia de terminación transcripcional. La secuencia que codifica el ARN guía está unida operativamente a una secuencia de control del promotor Pol III para su expresión en células eucariotas. Los ejemplos de promotores Pol III adecuados incluyen, pero no se limitan a, los promotores de ARN de mamíferos U6, U3, H1 y 7SL. En algunas situaciones, la proteína de fusión o los componentes del complejo pueden purificarse a partir de células bacterianas o eucariotas.

En diversas realizaciones, el ácido nucleico que codifica la proteína de fusión o la proteína CRISPR y el ARN guía del complejo pueden estar presentes en un vector. Los vectores adecuados incluyen vectores plasmídicos, vectores virales y ARN autorreplicante (Yoshioka et al., Cell Stem Cell, 2013, 13:246-254). En algunas realizaciones, el ácido nucleico que codifica la proteína de fusión o los componentes del complejo pueden estar presentes en un vector plasmídico. Los ejemplos no limitantes de vectores plasmídicos adecuados incluyen pUC, pBR322, pET, pBluescript y variantes de los mismos. En otras realizaciones, el ácido nucleico que codifica la proteína de fusión o los componentes del complejo o puede ser parte de un vector viral (por ejemplo, vectores lentivirales, vectores virales adenoasociados, vectores adenovirales, *etc.*). El plásmido o vector viral puede comprender secuencias de control de expresión adicionales (por ejemplo, secuencias potenciadoras, secuencias Kozak, secuencias de poliadenilación, secuencias de terminación transcripcional, *etc.*), secuencias marcadoras seleccionables (por ejemplo, genes de resistencia a antibióticos), orígenes de replicación y similares. Información adicional sobre los vectores y el uso de los mismos se puede encontrar en "Current Protocols in Molecular Biology" Ausubel et al., John Wiley & Sons, New York, 2003 o "Molecular Cloning: A Laboratory Manual" Sambrook & Russell, Cold Spring Harbor Press, Cold Spring Harbor, NY, 3rd edition, 2001.

(IV) Kits

Un aspecto adicional de la presente divulgación proporciona kits que comprenden la al menos una de las proteínas de fusión detalladas anteriormente en la sección (I), al menos uno de los complejos CRISPR descritos anteriormente en la sección (II), y/o al menos uno de los ácidos nucleicos descritos anteriormente en la sección (III). Los kits pueden comprender además reactivos de transfección, medios de crecimiento celular, medios de selección, reactivos de transcripción *in vitro*, reactivos de purificación de ácidos nucleicos, reactivos de purificación de proteínas, soluciones reguladoras y similares. Los kits proporcionados en el presente documento generalmente incluyen instrucciones para llevar a cabo los métodos que se detallan a continuación. Las instrucciones incluidas en los kits pueden estar adheridas al material de embalaje o pueden incluirse como un prospecto. Si bien las instrucciones por lo general son materiales escritos o impresos, no se limitan a estos. Esta divulgación contempla cualquier medio capaz de almacenar tales instrucciones y comunicarlás a un usuario final. Tales medios incluyen, pero no se limitan a, medios de almacenamiento electrónicos (por ejemplo, discos magnéticos, cintas, cartuchos, chips), medios ópticos (por ejemplo, CD ROM) y similares. Como se utiliza en el presente documento, el término "instrucciones" puede incluir la dirección de un sitio de Internet que proporciona las instrucciones.

(V) Células

La presente divulgación también proporciona células que comprenden la al menos una de las proteínas de fusión detalladas anteriormente en la sección (I), al menos uno de los complejos CRISPR descritos anteriormente en la sección (II), y/o al menos uno de los ácidos nucleicos descritos anteriormente en la sección (III). En general, la célula es una célula eucariota. Por ejemplo, la célula puede ser una célula humana, una célula de un mamífero no humano, una célula de un vertebrado no mamífero, una célula de un invertebrado, una célula de un insecto, una célula vegetal, una célula de levadura o un organismo eucariota unicelular.

(VI) Métodos para aumentar la eficiencia de la modificación genómica, transcripcional o epigenética dirigida

Otro aspecto de la presente divulgación abarca métodos para aumentar la eficiencia de la modificación genómica dirigida, la modificación transcripcional dirigida o la modificación epigenética dirigida en células eucariotas aumentando la accesibilidad de una proteína de modificación de ADN programable a su secuencia diana en el ADN cromosómico. En algunas realizaciones, el método comprende introducir en la célula eucariota de interés al menos una de las proteínas de fusión descritas anteriormente en la sección (I), al menos uno de los complejos CRISPR descritos anteriormente en la sección (II), o un ácido nucleico que codifica la al menos una proteína de fusión o complejo CRISPR como se describió anteriormente en la sección (III), y opcionalmente, un polinucleótido donante.

La proteína de modificación de ADN programable de la proteína de fusión está diseñada para reconocer y unirse a una secuencia diana en el ADN cromosómico, y el uno o más dominios de proteína que interactúan con el nucleosoma de la proteína de fusión interactúan con nucleosomas en o cerca de la secuencia diana para alterar o remodelar la estructura nucleosómica y/o de la cromatina. Como consecuencia, la proteína de modificación del ADN tiene mayor acceso a la secuencia cromosómica diana, de manera que aumenta la eficiencia de la modificación por parte de la proteína de modificación del ADN. En realizaciones específicas, la proteína de fusión comprende al menos un dominio de proteína que interactúa con el nucleosoma unido a una nucleasa CRISPR, de manera que las interacciones entre el o los dominios de proteína que interactúan con el nucleosoma y los nucleosomas/cromatina en o cerca de la secuencia diana aumentan la eficiencia de las modificaciones genómicas dirigidas (véanse los ejemplos 1-8).

De este modo, los métodos divulgados en el presente documento pueden aumentar la eficiencia de la edición genómica dirigida (por ejemplo, correcciones genéticas, inactivaciones genéticas, inserciones genéticas y similares), modificaciones epigenéticas dirigidas y regulación transcripcional dirigida.

(a) Introducción en la célula

Como se mencionó anteriormente, el método comprende introducir en la célula al menos una proteína de fusión, al menos un complejo CRISPR o ácido(s) nucleico(s) que codifican dicha proteína de fusión o complejo CRISPR (y, opcionalmente, un polinucleótido donante). La al menos una proteína de fusión, complejo CRISPR o ácido(s) nucleico(s) se pueden introducir en la célula de interés mediante una variedad de medios.

En algunas realizaciones, la célula puede transfectarse con las moléculas apropiadas (es decir, proteína, ADN y/o ARN). Los métodos de transfección adecuados incluyen la nucleofección (o electroporación), la transfección mediada por fosfato de calcio, la transfección con polímeros catiónicos (por ejemplo, DEAE-dextrano o polietilenimina), transducción viral, transfección de virosomas, transfección de viriones, transfección de liposomas, transfección de liposomas catiónicos, transfección de inmunoliposomas, transfección de lípidos no liposomales, transfección de dendrímeros, transfección de choque térmico, magnetofección, lipofección, administración de pistola genética, impalefacción, sonoporación, transfección óptica y captación de ácidos nucleicos potenciada por agentes patentados. Los métodos de transfección son bien conocidos en la técnica (véase, por ejemplo, "Current Protocols in Molecular Biology" Ausubel et al., John Wiley & Sons, New York, 2003 o "Molecular Cloning: A Laboratory Manual" Sambrook & Russell, Cold Spring Harbor Press, Cold Spring Harbor, NY, 3rd edition, 2001). En otras realizaciones, las moléculas pueden introducirse en la célula mediante microinyección. Por ejemplo, las moléculas pueden inyectarse en el citoplasma o en los núcleos de las células de interés. La cantidad de cada molécula introducida en la célula puede variar, pero los expertos en la técnica están familiarizados con los medios para determinar la cantidad apropiada.

Las diversas moléculas pueden introducirse en la célula de forma simultánea o secuencial. Por ejemplo, la proteína de fusión o el complejo CRISPR (o ácidos nucleicos codificantes) y el polinucleótido donante pueden introducirse al mismo tiempo. Alternativamente, uno puede introducirse primero y luego el otro puede introducirse más tarde en la célula.

En general, la célula se mantiene en condiciones apropiadas para el crecimiento y/o mantenimiento celular. Las condiciones adecuadas de cultivo celular son bien conocidas en la técnica y se describen, por ejemplo, en Santiago et al., Proc. Natl. Acad. Sci. USA, 2008, 105:5809-5814; Moehle et al. Proc. Natl. Acad. Sci. USA, 2007, 104:3055-3060; Urnov et al., Nature, 2005, 435:646-651; y Lombardo et al., Nat. Biotechnol., 2007, 25:1298-1306. Los expertos en la técnica apreciarán que los métodos para cultivar células son conocidos en la técnica y pueden variar y variarán dependiendo del tipo de célula. La optimización de rutina se puede utilizar, en todos los casos, para determinar las mejores técnicas para un tipo de célula particular.

(b) Modificación genómica o epigenética dirigida

El uno o más dominios de proteína que interactúan con el nucleosoma de la proteína de fusión o el complejo CRISPR interactúan con nucleosomas y/o ADN cromosómico en o cerca de la secuencia cromosómica diana de tal manera que la estructura nucleosómica y/o de la cromatina se altera/remodela, aumentando así la accesibilidad de la proteína de modificación de ADN programable de la proteína de fusión o la proteína CRISPR del complejo CRISPR a la secuencia cromosómica diana. Un mayor acceso a la secuencia cromosómica diana da como resultado una mayor frecuencia/eficiencia de la modificación genómica, transcripcional o epigenética dirigida.

En realizaciones en las que la proteína de fusión comprende una proteína de modificación de ADN programable que tiene actividad de nucleasa, la proteína de fusión puede escindir una o ambas cadenas de la secuencia cromosómica diana. Las roturas de doble cadena se pueden reparar mediante un procedimiento de reparación de unión de extremos no homólogos (NHEJ). Debido a que NHEJ es propenso a errores, las indeles (es decir, deleciones o inserciones) de al menos un par de bases, sustituciones de al menos un par de bases o combinaciones de las mismas pueden ocurrir durante la reparación de la rotura. De acuerdo con lo anterior, la secuencia cromosómica diana puede modificarse, mutarse o inactivarse. Por ejemplo, una delección, inserción

o sustitución en el marco de lectura de una secuencia codificante puede dar lugar a un producto proteico alterado o a ningún producto proteico (que se denomina "inactivación"). En algunas iteraciones, el método puede comprender además la introducción en la célula de un polinucleótido donante (véase a continuación) que comprende una secuencia donante que está flanqueada por una secuencia que tiene una identidad de secuencia sustancial con las secuencias ubicadas a cada lado de la secuencia cromosómica diana, de modo que durante la reparación de la rotura de doble cadena mediante un procedimiento de reparación dirigida por homología (HDR), la secuencia donante en el polinucleótido donante puede intercambiarse con o integrarse en la secuencia cromosómica en la secuencia cromosómica diana. La integración de una secuencia exógena se denomina "inserción".

En diversas iteraciones, por lo tanto, la eficiencia de la modificación genómica dirigida se puede aumentar al menos aproximadamente 0.1 vez, al menos aproximadamente 0.5 veces, al menos aproximadamente 1 vez, al menos aproximadamente 2 veces, al menos aproximadamente 5 veces, al menos aproximadamente 10 veces, o al menos aproximadamente 20 veces, al menos aproximadamente 50 veces, al menos aproximadamente 100 veces, o más de aproximadamente 100 veces en relación con la proteína de modificación de ADN programable parental que no está vinculada a al menos un dominio de proteína que interactúa con el nucleosoma.

En realizaciones en las que la proteína de fusión comprende una proteína de modificación de ADN programable que tiene actividad no nucleasa, la proteína de fusión puede modificar el ADN o las proteínas asociadas en la secuencia cromosómica diana o modificar la expresión de la secuencia cromosómica diana. Por ejemplo, cuando la proteína de modificación de ADN programable comprende actividad de modificación epigenética, el estado de acetilación, metilación, fosforilación y adenilación de las histonas, etc. se puede modificar o se puede modificar el estado de metilación, aminación del ADN, etc. A modo de ejemplo, en realizaciones en las que la proteína de modificación de ADN programable comprende actividad de citidina desaminasa, uno o más residuos de citidina en la secuencia cromosómica diana pueden convertirse en residuos de uridina. Alternativamente, cuando la proteína de modificación de ADN programable comprende activación transcripcional o actividad represora, la transcripción en la secuencia cromosómica diana puede aumentarse o disminuirse.

La modificación epigenética o regulación transcripcional resultante se puede incrementar al menos aproximadamente 0.1 vez, al menos aproximadamente 0.5 veces, al menos aproximadamente 1 vez, al menos aproximadamente 2 veces, al menos aproximadamente 5 veces, al menos aproximadamente 10 veces, o al menos aproximadamente 20 veces, al menos aproximadamente 50 veces, al menos aproximadamente 100 veces, o más de aproximadamente 100 veces en relación con la proteína de modificación de ADN programable parental que no está vinculada a al menos un dominio de proteína que interactúa con el nucleosoma.

Las modificaciones genómicas, transcripcionales y epigenéticas dirigidas detalladas anteriormente se pueden realizar de forma individual o multiplexada (es decir, se pueden seleccionar dos o más secuencias cromosómicas simultáneamente).

(c) Polinucleótido donante opcional

En realizaciones en las que la proteína de fusión comprende una proteína de modificación de ADN programable que tiene actividad de nucleasa, el método puede comprender además la introducción de al menos un polinucleótido donante en la célula. El polinucleótido donante puede ser monocatenario o de doble cadena, lineal o circular, y/o ARN o ADN. En algunas realizaciones, el polinucleótido donante puede ser un vector, por ejemplo, un vector plasmídico.

El polinucleótido donante comprende al menos una secuencia donante. En algunos aspectos, la secuencia donante del polinucleótido donante puede ser una versión modificada de una secuencia cromosómica endógena o nativa. Por ejemplo, la secuencia donante puede ser esencialmente idéntica a una porción de la secuencia cromosómica en o cerca de la secuencia a la que se dirige la proteína de modificación de ADN, pero que comprende al menos un cambio de nucleótido. De este modo, tras la integración o el intercambio con la secuencia nativa, la secuencia en la ubicación cromosómica diana comprende al menos un cambio de nucleótido. Por ejemplo, el cambio puede ser una inserción de uno o más nucleótidos, una delección de uno o más nucleótidos, una sustitución de uno o más nucleótidos o combinaciones de los mismos. Como consecuencia de la integración de la "corrección genética" de la secuencia modificada, la célula puede producir un producto genético modificado a partir de la secuencia cromosómica diana.

En otros aspectos, la secuencia donante del polinucleótido donante puede ser una secuencia exógena. Como se utiliza en el presente documento, una secuencia "exógena" se refiere a una secuencia que no es nativa de la célula, o una secuencia cuya ubicación nativa está en una ubicación diferente en el genoma de la célula. Por ejemplo, la secuencia exógena puede comprender una secuencia codificante de proteína, que puede estar vinculada operativamente a una secuencia de control de promotor exógeno de tal manera que, tras la integración en el genoma, la célula sea capaz de expresar la proteína codificada por la secuencia integrada. Alternativamente, la secuencia exógena puede integrarse en la secuencia cromosómica de tal manera que su expresión esté regulada por una secuencia de control del promotor endógeno. En otras iteraciones, la secuencia exógena puede ser una secuencia de control transcripcional, otra secuencia de control de expresión,

una secuencia codificante de ARN, etc. Como se señaló anteriormente, la integración de una secuencia exógena en una secuencia cromosómica se denomina "inserción".

Como podrán apreciar los expertos en la técnica, la longitud de la secuencia donante puede variar y variará. Por ejemplo, la secuencia donante puede variar en longitud desde unos pocos nucleótidos hasta cientos de nucleótidos o cientos de miles de nucleótidos.

Por lo general, la secuencia donante en el polinucleótido donante está flanqueada por una secuencia corriente arriba y una secuencia corriente abajo, que tienen una identidad de secuencia sustancial con las secuencias ubicadas corriente arriba y corriente abajo, respectivamente, de la secuencia a la que se dirige la proteína de modificación de ADN programable. Debido a estas similitudes de secuencia, las secuencias corriente arriba y corriente abajo del polinucleótido donante permiten la recombinación homóloga entre el polinucleótido donante y la secuencia cromosómica diana, de manera que la secuencia donante puede integrarse en (o intercambiarse con) la secuencia cromosómica.

La secuencia corriente arriba, como se utiliza en el presente documento, se refiere a una secuencia de ácido nucleico que comparte una identidad de secuencia sustancial con una secuencia cromosómica corriente arriba de la secuencia a la que se dirige la proteína de modificación de ADN programable. De manera similar, la secuencia corriente abajo se refiere a una secuencia de ácido nucleico que comparte una identidad de secuencia sustancial con una secuencia cromosómica corriente abajo de la secuencia a la que se dirige la proteína de modificación de ADN programable. Como se utiliza en el presente documento, la frase "identidad de secuencia sustancial" se refiere a secuencias que tienen al menos aproximadamente un 75 % de identidad de secuencia. De este modo, las secuencias corriente arriba y corriente abajo en el polinucleótido donante pueden tener aproximadamente un 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia con la secuencia corriente arriba o corriente abajo a la secuencia diana. En una realización de ejemplo, las secuencias corriente arriba y corriente abajo en el polinucleótido donante pueden tener aproximadamente un 95 % o un 100 % de identidad de secuencia con las secuencias cromosómicas corriente arriba o corriente abajo a la secuencia a la que se dirige la proteína de modificación de ADN programable.

En algunas realizaciones, la secuencia corriente arriba comparte una identidad de secuencia sustancial con una secuencia cromosómica ubicada inmediatamente corriente arriba de la secuencia a la que se dirige la proteína de modificación de ADN programable. En otras realizaciones, la secuencia corriente arriba comparte una identidad de secuencia sustancial con una secuencia cromosómica que se encuentra dentro de aproximadamente cien (100) nucleótidos corriente arriba de la secuencia diana. De este modo, por ejemplo, la secuencia corriente arriba puede compartir una identidad de secuencia sustancial con una secuencia cromosómica que se encuentra aproximadamente entre 1 y aproximadamente 20, aproximadamente entre 21 y aproximadamente 40, aproximadamente entre 41 y aproximadamente 60, aproximadamente entre 61 y aproximadamente 80, o aproximadamente entre 81 y aproximadamente 100 nucleótidos corriente arriba de la secuencia diana. En algunas realizaciones, la secuencia corriente abajo comparte una identidad de secuencia sustancial con una secuencia cromosómica ubicada inmediatamente corriente abajo de la secuencia a la que se dirige la proteína de modificación de ADN programable. En otras realizaciones, la secuencia corriente abajo comparte una identidad de secuencia sustancial con una secuencia cromosómica que se encuentra dentro de aproximadamente cien (100) nucleótidos corriente abajo de la secuencia diana. De este modo, por ejemplo, la secuencia corriente abajo puede compartir una identidad de secuencia sustancial con una secuencia cromosómica que se encuentra aproximadamente entre 1 y aproximadamente 20, aproximadamente entre 21 y aproximadamente 40, aproximadamente entre 41 y aproximadamente 60, aproximadamente entre 61 y aproximadamente 80, o aproximadamente entre 81 y aproximadamente 100 nucleótidos corriente abajo de la secuencia diana.

Cada secuencia corriente arriba o corriente abajo puede variar en longitud desde aproximadamente 20 nucleótidos hasta aproximadamente 5000 nucleótidos. En algunas realizaciones, las secuencias corriente arriba y corriente abajo pueden comprender aproximadamente 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2800, 3000, 3200, 3400, 3600, 3800, 4000, 4200, 4400, 4600, 4800 o 5000 nucleótidos. En realizaciones específicas, las secuencias corriente arriba y corriente abajo pueden variar en longitud desde aproximadamente 50 a aproximadamente 1500 nucleótidos.

(d) Tipos de células

Una variedad de células son adecuadas para su uso en los métodos divulgados en el presente documento. En general, la célula es una célula eucariota. Por ejemplo, la célula puede ser una célula humana, una célula de un mamífero no humano, una célula de un vertebrado no mamífero, una célula de un invertebrado, una célula de un insecto, una célula vegetal, una célula de levadura o un organismo eucariota unicelular. En algunas realizaciones, la célula también puede ser un embrión unicelular no humano. Por ejemplo, un embrión de mamífero no humano, incluidos embriones de rata, hámster, roedor, conejo, felino, canino, ovino, porcino, bovino, equino y primate. En otras realizaciones más, la célula puede ser una célula madre, tal como células

madre embrionarias, células madre tipo ES, células madre fetales, células madre adultas y similares. En una realización, la célula madre no es una célula madre embrionaria humana. Además, las células madre pueden incluir aquellas obtenidas mediante las técnicas divulgadas en los documentos WO2003/046141, o Chung et al. (Cell Stem Cell, 2008, 2:113-117). La célula puede ser *in vitro*. En realizaciones de ejemplo, la célula es una célula de mamífero o una línea celular de mamífero. En realizaciones particulares, la célula es una célula humana o una línea celular humana.

Ejemplos no limitantes de células o líneas celulares de mamíferos adecuadas incluyen células de riñón embrionario humano (HEK293, HEK293T); células de carcinoma cervical humano (HELA); células de pulmón humano (W138); células de hígado humano (Hep G2); células de osteosarcoma U2-OS humano, células A549 humanas, células A-431 humanas y células K562 humanas; células de ovario de hámster chino (CHO), células de riñón de hámster bebé (BHK); células de mieloma NS0 de ratón, células de fibroblastos embrionarios de ratón 3T3 (NIH3T3), células de linfoma B A20 de ratón; células de melanoma B16 de ratón; células de mioblastos C2C12 de ratón; células de mieloma SP2/0 de ratón; células mesenquimales embrionarias de ratón C3H-10T1/2; células de carcinoma CT26 de ratón, células DuCuP de próstata de ratón; células EMT6 de mama de ratón; células de hepatoma Hepa1c1c7 de ratón; células de mieloma de ratón J5582; células epiteliales MTD-1A de ratón; células MyEnd de miocardio de ratón; células RenCa renales de ratón; células RIN-5F pancreáticas de ratón; células X64 de melanoma de ratón; células YAC-1 de linfoma de ratón; células 9L de glioblastoma de rata; células RBL de linfoma B de rata; células B35 de neuroblastoma de rata; células de hepatoma de rata (HTC); células BRL 3A de hígado de rata búfalo; células de riñón canino (MDCK); células mamarias caninas (CMT); células D17 de osteosarcoma de rata; células DH82 de monocitos/macrófagos de rata; células de fibroblastos transformados SV-40 de riñón de mono (COS7); células CVI-76 de riñón de mono; células de riñón de mono verde africano (VERO-76). Se puede encontrar una lista extensa de líneas celulares de mamíferos en the American Type Culture Collection catalog (ATCC, Manassas, VA).

(VII) Métodos para detectar loci genómicos específicos

En realizaciones en las que la proteína de fusión comprende una modificación de ADN programable que tiene actividad no nucleasa o el complejo CRISPR comprende una proteína CRISPR catalíticamente inactiva que tiene actividad no nucleasa, dicha proteína de fusión o complejo CRISPR se puede utilizar en métodos para detectar o visualizar loci genómicos específicos en células eucariotas. En tales realizaciones, la proteína de fusión o proteína CRISPR del complejo comprende además al menos una etiqueta detectable, tal como un fluoróforo (por ejemplo, FAM, TMR, Cy3, Cy5, Texas Red, Oregon Green, Alexa Fluors, etiquetas Halo o colorante fluorescente adecuado), una etiqueta de detección (por ejemplo, biotina, digoxigenina y similares), puntos cuánticos o partículas de oro. Alternativamente, el ARN guía del complejo CRISPR puede comprender además una etiqueta detectable para detección *in situ* (por ejemplo, FISH o CISH). El al menos un dominio de proteína que interactúa con el nucleosoma de la proteína de fusión o complejo CRISPR aumenta el acceso de la proteína de modificación de ADN programable o proteína CRISPR que tiene actividad no nucleasa a la secuencia cromosómica diana, potenciando así la detección de loci genómicos específicos o secuencias cromosómicas diana.

El método comprende introducir en la célula eucariota la proteína de fusión marcada de forma detectable, el complejo CRISPR marcado de forma detectable o el ácido nucleico codificante, y detectar la proteína de modificación de ADN programable marcada o la proteína CRISPR marcada unida a la secuencia cromosómica diana. La detección puede realizarse mediante imágenes dinámicas de células vivas, microscopía fluorescente, microscopía confocal, inmunofluorescencia, inmunodetección, unión de ARN-proteína, unión proteína-proteína y similares. La etapa de detección se puede realizar en células vivas o células fijadas.

En realizaciones en las que el método comprende detectar la dinámica estructural de la cromatina en células vivas, la proteína de fusión marcada de forma detectable o el complejo CRISPR marcado de forma detectable se pueden introducir en la célula como proteínas o ácidos nucleicos. En realizaciones en las que el método comprende detectar la secuencia cromosómica diana en células fijadas, la proteína de fusión marcada de forma detectable o el complejo CRISPR marcado de forma detectable se pueden introducir en la célula como proteínas (o complejos proteína-ARN). Los medios para fijar y permeabilizar células son bien conocidos en la técnica. En algunas realizaciones, las células fijadas pueden someterse a procedimientos de desnaturalización química y/o térmica para convertir el ADN cromosómico de doble cadena en ADN monocatenario. En otras realizaciones, las células fijadas no están sujetas a procedimientos de desnaturalización química y/o térmica.

(VIII) Aplicaciones

Las composiciones y métodos divulgados en el presente documento se pueden utilizar en una variedad de aplicaciones terapéuticas, de diagnóstico, industriales y de investigación. En algunas realizaciones, la presente divulgación se puede utilizar para modificar cualquier secuencia cromosómica de interés en una célula, animal o planta con el fin de modelar y/o estudiar la función de los genes, estudiar afecciones genéticas o epigenéticas de interés o estudiar vías bioquímicas implicadas en diversas enfermedades o trastornos. Por ejemplo, se pueden crear organismos transgénicos que modelen enfermedades o trastornos, en los que se altera la expresión de una o más secuencias de ácidos nucleicos asociadas con una enfermedad o trastorno. El modelo

de enfermedad se puede utilizar para estudiar los efectos de las mutaciones en el organismo, estudiar el desarrollo y/o progresión de la enfermedad, estudiar el efecto de un compuesto farmacéuticamente activo en la enfermedad y/o evaluar la eficacia de una posible estrategia de terapia genética.

En otras realizaciones, las composiciones y los métodos se pueden utilizar para realizar exámenes genómicos funcionales eficientes y rentables, que se pueden utilizar para estudiar la función de los genes involucrados en un procedimiento biológico particular y cómo cualquier alteración en la expresión genética puede afectar el procedimiento biológico, o para realizar mutagénesis de saturación o de barrido profundo de loci genómicos junto con un fenotipo celular. La mutagénesis de saturación o de barrido profundo se puede utilizar para determinar características mínimas críticas y vulnerabilidades discretas de elementos funcionales necesarios para la expresión genética, la resistencia a fármacos y la reversión de enfermedades, por ejemplo.

En realizaciones adicionales, las composiciones y métodos divulgados en el presente documento se pueden utilizar para pruebas de diagnóstico para establecer la presencia de una enfermedad o trastorno y/o para su uso en la determinación de opciones de tratamiento. Entre los ejemplos de pruebas diagnósticas adecuadas se incluyen la detección de mutaciones específicas en células cancerosas (por ejemplo, mutación específica en EGFR, HER2 y similares), detección de mutaciones específicas asociadas con enfermedades particulares (por ejemplo, repeticiones de trinucleótidos, mutaciones en la β -globina asociadas con la enfermedad de células falciformes, SNP específicos, etc.), detección de hepatitis, detección de virus (por ejemplo, Zika), etc.

En realizaciones adicionales, las composiciones y métodos divulgados en el presente documento se pueden utilizar para corregir mutaciones genéticas asociadas con una enfermedad o trastorno particular, tales como, por ejemplo, corregir mutaciones del gen de la globina asociadas con la enfermedad de células falciformes o la talasemia, corregir mutaciones en el gen de la adenosina desaminasa asociado con la inmunodeficiencia combinada grave (SCID), reducir la expresión de HTT, el gen causante de la enfermedad de Huntington, o corregir mutaciones en el gen de la rodopsina para el tratamiento de la retinitis pigmentosa. Tales modificaciones pueden realizarse en células *ex vivo*.

En otras realizaciones más, las composiciones y métodos divulgados en el presente documento se pueden utilizar para generar plantas de cultivo con rasgos mejorados o mayor resistencia al estrés ambiental. La presente divulgación también se puede utilizar para generar animales de granja con rasgos mejorados o animales de producción. Por ejemplo, los cerdos tienen muchas características que los hacen atractivos como modelos biomédicos, especialmente en medicina regenerativa o xenotrasplante.

Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el significado comúnmente entendido por una persona experta en la técnica a la que pertenece esta invención. Las siguientes referencias proporcionan a un experto una definición general de muchos de los términos utilizados en esta invención: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd Ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991). Como se utilizan en el presente documento, los siguientes términos tienen los significados que se les atribuyen a menos que se especifique lo contrario.

Al introducir elementos de la presente divulgación o la(s) realización(es) preferidas, los artículos "un", "una", "el" y "dicho" pretenden significar que hay uno o más de los elementos. Los términos "que comprende", "incluyendo" y "que tiene" pretenden ser inclusivos y significan que puede haber elementos adicionales distintos de los enumerados.

El término "aproximadamente" cuando se utiliza en relación con un valor numérico, x, por ejemplo, significa $x \pm 5\%$.

Como se utilizan en el presente documento, los términos "complementario" o "complementariedad" se refieren a la asociación de ácidos nucleicos de doble cadena mediante emparejamiento de bases mediante enlaces de hidrógeno específicos. El emparejamiento de bases puede ser el emparejamiento de bases Watson-Crick estándar (por ejemplo, 5'-A G T C-3' se empareja con la secuencia complementaria 3'-T C A G-5'). El emparejamiento de bases también puede ser mediante enlaces de hidrógeno Hoogsteen o Hoogsteen invertido. La complementariedad se mide por lo general con respecto a una región dúplex y, de este modo, excluye los voladizos, por ejemplo. La complementariedad entre dos cadenas de la región dúplex puede ser parcial y expresarse como un porcentaje (por ejemplo, 70 %), si solo algunos (por ejemplo, el 70 %) de las bases son complementarias. Las bases que no son complementarias están "en emparejamiento erróneo". La complementariedad también puede ser completa (es decir, 100 %), si todas las bases en la región dúplex son complementarias.

Como se utiliza en el presente documento, el término "sistema CRISPR" se refiere a un complejo que comprende una proteína CRISPR (es decir, nucleasa, nickasa o proteína catalíticamente muerta) y un ARN guía.

El término "secuencia endógena", como se utiliza en el presente documento, se refiere a una secuencia cromosómica que es nativa de la célula.

Como se utiliza en el presente documento, el término "exógeno" se refiere a una secuencia que no es nativa de la célula, o una secuencia cromosómica cuya ubicación nativa en el genoma de la célula está en una ubicación cromosómica diferente.

Un "gen", como se utiliza en el presente documento, se refiere a una región de ADN (incluidos exones e intrones) que codifica un producto génico, así como a todas las regiones de ADN que regulan la producción del producto génico, independientemente de que dichas secuencias reguladoras sean adyacentes o no a las secuencias codificantes y/o transcritas. De acuerdo con lo anterior, un gen incluye, pero no se limita necesariamente a, secuencias promotoras, terminadores, secuencias reguladoras de la traducción tales como sitios de unión de ribosomas y sitios de entrada interna a ribosomas, potenciadores, silenciadores, aislantes, elementos límite, orígenes de replicación, sitios de unión a la matriz y regiones de control de locus.

El término "heterólogo" se refiere a una entidad que no es endógena ni nativa de la célula de interés. Por ejemplo, una proteína heteróloga se refiere a una proteína que se deriva o se derivó originalmente de una fuente exógena, tal como una secuencia de ácido nucleico introducida exógenamente. En algunos casos, la proteína heteróloga no es producida normalmente por la célula de interés.

El término "nickasa" se refiere a una enzima que escinde una cadena de una secuencia de ácido nucleico de doble cadena (es decir, corta una secuencia de doble cadena). Por ejemplo, una nucleasa con actividad de escisión de doble cadena puede modificarse mediante mutación y/o eliminación para funcionar como una nickasa y escindir solo una cadena de una secuencia de doble cadena.

El término "nucleasa", como se utiliza en el presente documento, se refiere a una enzima que escinde ambas cadenas de una secuencia de ácido nucleico de doble cadena.

Los términos "ácido nucleico" y "polinucleótido" se refieren a un polímero desoxirribonucleótido o ribonucleótido, en conformación lineal o circular, y en forma ya sea monocatenaria o bicatenaria. A los efectos de la presente divulgación, estos términos no deben interpretarse como limitantes con respecto a la longitud de un polímero. Los términos pueden abarcar análogos conocidos de nucleótidos naturales, así como nucleótidos que están modificados en las unidades estructurales de base, azúcar y/o fosfato (por ejemplo, cadenas principales de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases; es decir, un análogo de A se apareará con T.

El término "nucleótido" se refiere a desoxirribonucleótidos o ribonucleótidos. Los nucleótidos pueden ser nucleótidos estándar (es decir, adenosina, guanosina, citidina, timidina y uridina), isómeros de nucleótidos o análogos de nucleótidos. Un análogo de nucleótido se refiere a un nucleótido que tiene una base de purina o pirimidina modificada o una unidad estructural de ribosa modificada. Un análogo de nucleótido puede ser un nucleótido de origen natural (por ejemplo, inosina, pseudouridina, etc.) o un nucleótido que no se encuentra en la naturaleza. Los ejemplos no limitantes de modificaciones en las unidades estructurales de azúcar o base de un nucleótido incluyen la adición (o eliminación) de grupos acetilo, grupos amino, grupos carboxilo, grupos carboximetilo, grupos hidroxilo, grupos metilo, grupos fosforilo y grupos tiol, así como la sustitución de los átomos de carbono y nitrógeno de las bases con otros átomos (por ejemplo, 7-deaza purinas). Los análogos de nucleótidos también incluyen nucleótidos didesoxi, nucleótidos 2'-O-metil, ácidos nucleicos bloqueados (LNA), ácidos nucleicos peptídicos (PNA) y morfolidos.

Los términos "polipéptido" y "proteína" se utilizan indistintamente para referirse a un polímero de residuos de aminoácidos.

Como se utiliza en el presente documento, el término "proteína de modificación de ADN programable" se refiere a una proteína que está diseñada para unirse a una secuencia diana específica en el ADN cromosómico y que modifica el ADN o la(s) proteína(s) asociada(s) con el ADN en o cerca de la secuencia diana.

El término "identidad de secuencia", como se utiliza en el presente documento, indica una medida cuantitativa del grado de identidad entre dos secuencias de longitud sustancialmente igual. El porcentaje de identidad de dos secuencias, ya sean secuencias de ácidos nucleicos o de aminoácidos, es el número de coincidencias exactas entre dos secuencias alineadas dividido por la longitud de la secuencia más corta y multiplicado por 100. Una alineación aproximada de las secuencias de ácidos nucleicos se proporciona por el algoritmo de homología local de Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). Este algoritmo se puede aplicar a secuencias de aminoácidos utilizando la matriz de puntuación desarrollada por Dayhoff, *Atlas of Protein Sequences and Structure*, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, y normalizado por Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). El Genetics Computer Group (Madison, Wisconsin) proporciona una implementación de ejemplo de este algoritmo para determinar el porcentaje de identidad de una secuencia en la aplicación de utilidad "BestFit". En general, en la técnica se conocen otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias, por ejemplo, otro programa de alineamiento es BLAST, utilizado con parámetros predeterminados.

5 Por ejemplo, BLASTN y BLASTP se pueden utilizar utilizando los siguientes parámetros predeterminados: código genético=estándar; filtro=ninguno; cadena=ambas; corte=60; expectativa=10; Matriz=BLOSUM62; Descripciones=50 secuencias; ordenar por=PUNTAJE ALTO; Bases de datos=no redundante, GenBank+EMB+DDBJ+PDB+GenBank CDS translations+Swiss protein+Spupdate+PIR. Los detalles de estos programas se pueden encontrar en el sitio web de GenBank. En general, las sustituciones son sustituciones conservadoras de aminoácidos: limitadas a intercambios dentro de los miembros del grupo 1: glicina, alanina, valina, leucina e isoleucina; grupo 2: serina, cisteína, treonina y metionina; grupo 3: prolina; grupo 4: fenilalanina, tirosina y triptófano; grupo 5: aspartato, glutamato, asparagina y glutamina.

10 Los términos "secuencia diana", "secuencia cromosómica diana" y "sitio diana" se utilizan indistintamente para referirse a la secuencia específica en el ADN cromosómico a la que se dirige la proteína de modificación de ADN programable, y al sitio en el que la proteína de modificación de ADN programable modifica el ADN o las proteínas asociadas con el ADN.

15 Se conocen en la técnica técnicas para determinar la identidad de la secuencia de ácidos nucleicos y aminoácidos. Por lo general, tales técnicas incluyen la determinación de la secuencia de nucleótidos del ARNm de un gen y/o la determinación de la secuencia de aminoácidos codificada por el mismo, y la comparación de estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también se pueden determinar y comparar de esta manera.

20 En general, la identidad se refiere a una correspondencia exacta de nucleótido a nucleótido o de aminoácido a aminoácido de dos polinucleótidos o secuencias de polipéptidos, respectivamente. Se pueden comparar dos o más secuencias (polinucleótidos o aminoácidos) determinando su porcentaje de identidad.

Se pretende que todo el material contenido en la descripción anterior y en los ejemplos que se dan a continuación se interprete como ilustrativo y no en sentido limitativo.

Ejemplos

25 Los siguientes ejemplos ilustran determinados aspectos de la divulgación. En la tabla 1 se enumeran las secuencias de péptidos de los dominios que interactúan con el nucleosoma y en la tabla 2 se presentan las secuencias cromosómicas diana utilizadas en los ejemplos 1-8 que se presentan a continuación.

Tabla 1. Secuencias peptídicas de dominios que interactúan con el nucleosoma		
Dominio de interacción del nucleosoma	Secuencia (NH ₂ -COOH)	SEQ ID NO:
Dominio de la caja A del gen HMGB1 humano (1-84 aa)	MGKGDPPKPRGKMSSYAFFVQTCREEHKKKHPDASVNF SEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMK TYIPPKGE	40
Proteína humana HMGN1	MPKRKVSSAEGAAKEPKRRSARLSAKPPAKVEAKPKKA AAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAE NGETKTEESPASDEAGEKEAKSD	41
Proteína humana HMGN2	MPKRKAEGDAKGDKAKVKDEPQRRSARLSAKPAPPKPE PKPKKAPAKKGEKVPKGKKGKADAGKEGNNPAENGDAK TDQAQKAEGAGDAK	42
Proteína humana HMGN3a	MPKRKSPENTEGKDGSKVTKQEPTRRSARLSAKPAPPK PEPKPRKTSAKKEPGAKISRGAKGKKEEKQEAGKEGTAP SENGETKAEAAQKTESVDNEGE	43

Proteína humana HMGN3b	MPKRKSPENTEGKDGSKVTKQEPTRRSARLSAKPAPPK PEPKPRKTSAKKEPGAKISRGAKGKKEEKQEAGKEGTEN	44
Dominio globular central de la histona H1 humana (22-101 aa)	STDHPKYSDMIVAAIQAEKNRAGSSRQSIQKYIKSHYKVG ENADSQIKLSIKRLVTTGVLKQTKGVGASGSFRLAKSDEP	45
Dominio de unión al ADN de la ATPasa ISW1 del complejo de remodelación de cromatina ISWI de levadura	LLNPTKRERKENYSIDNYYKDVLTNTGRSSTPSHPRMPKP HVFHSHQLQPPQLKVLYEKERMWTAKKTGYVPTMDDVK AAYGDISDEEEKKQKLELLKLSVNNSQPLTEEEKMKAD WESEGFTNWNKLEFRKFITVSGKYGRNSIQAIARELAPGK TLEEVRAKAFWSNIERIEDYKYLKIIENEEKIKRVKM QQEALRRKLSEYKNPFFDLKLKHPSSNNKRTYSEEDR FILLMLFKYGLDRDDVYELVRDEIRDCLFELDFYFRSRTP VELARRGNTLLQCLEKEFNAGIVLDDATKDRMKKEDENG KRIRREEFADQTANEKENVDGVESKKAKIEDTSNVGTEQLV AEKIPENETTH	46
Dominio de unión al ADN de la proteína 1 que contiene el dominio cromosómico (CHD1) de levadura	DMDSIGESEVRALYKAILKFGNLKEILDELIADGTLPVKSFE KYGETYDEMMEAACDCVHEEEKNRKEILEKLEKHATAYR AKLKSGEIKAENQPKDNPLTRLCLKREKKAVLFNFKGVK SLNAESLLSRVEDLKYLKNLINSNYKDDPLKFSLGNNTPK PVQNWSSNWTKEEDEKLLIGVFKYGYGSWTQIRDDPFL GITDKIFLNEVHNPAKKSASSSDTTPTPSKKGKGITGSSK KVPGAHLGRRVDYLLSFLRGGLNTKSPS	47

Tabla 2. Sitios diana cromosómicos			
Locus	Sitio	Secuencia (5'-3')	SEQ ID NO:
<i>Streptococcus pyogenes</i> Cas9 (SpCas9)			
POR	#1	AGCCGTGAGTGGAGGGAGCGTGG	48
POR	#2	AGAGGGAGGGGTTGGACTACAGG	49
POR	#3	CATTCGCCAGTACGAGCTTGTGG	50
CAR	#1	CTTTAATGCGCTGACTTGTGAGGG	51
EMX1	#1	GTGGCGCATTGCCACGAAGCAGG	52
EMX1	#2	TTCTTCTTCTGCTCGGACTCAGG	53
<i>Streptococcus pasteurianus</i> Cas9 (SpaCas9)			

POR	#1	TGCTGGAAAGGGGAGACCAAGGGTGA	54
POR	#2	AGAGCTAC GAGAAC CAGAAGC CGTGA	55
<i>Francisella novicida</i> Cpf1 (FnCpf1)			
POR	#1	TTCCGGCCTCACCCCTTGGTCTCCCC	56
POR	#2	TTGGTCTCCCCCTTCCAGCATTGCGC	57
POR	#3	TTCCAGCATTGCGCCAGTACGAGCTTG	58
<i>Campylobacter jejuni</i> Cas9 (CjCas9)			
POR	#1	GATCAACATGGGAGACTCCCACGTGGACAC	59
POR	#2	AGATACTTCTTCGCCACCGCTCGGACAC	60

Ejemplo 1. Mejora de la actividad de *Streptococcus pyogenes* Cas9 (SpCas9) utilizando el dominio A de la caja HMGB1 humano

Un dominio de caja A de HMGB1 humano (SEQ ID NO: 40) se fusionó con SpCas9 (+NLS) en el extremo carboxilo terminal de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cas9 y el dominio de caja A de HMGB1. Células humanas K562 (1×10^6) se transfectaron con ADN plasmídico que codifica la proteína de fusión o la proteína SpCas9 de tipo salvaje en cantidades equivalentes molares (5.2 y 5.0 μ g para la proteína de fusión y la proteína Cas9 de tipo salvaje, respectivamente) en combinación con 3 μ g de un plásmido ARNg para apuntar a un sitio genómico (#1) en el locus de la oxidoreductasa del citocromo p450 (POR) humano. La transfección se llevó a cabo mediante nucleofección en un nucleofector Amaxi. Tres días después de la transfección, las células se lisaron con una solución de extracción de ADN (QuickExtract™) y la región genómica diana se amplificó por PCR. Las actividades de escisión de la diana de la nucleasa Cas9 (% de indeles) se midieron utilizando ensayos Cel-I. Como se muestra en la tabla 3, la fusión del dominio de la caja A de HMGB1 humano con la nucleasa aumentó la eficiencia de escisión de SpCas9 en el sitio diana.

Tabla 3. Eficiencia de escisión		
Nucleasa	Sitio diana	Indel (%)
SpCas9 de tipo salvaje	POR/sitio #1	8.5
Fusión de la caja A de SpCas9-HMGB1	POR/sitio #1	21.3

Ejemplo 2. Mejora de la Actividad de *Streptococcus pyogenes* de Cas9 (SpCas9) utilizando HMGN1, HMGN2, HMGN3a y HMGN3b humanos

Los HMGN1, HMGN2, HMGN3a y HMGN3b humanos (SEQ ID NOS:41-44, respectivamente) se fusionaron cada uno con SpCas9 (+NLS) en el extremo carboxilo terminal de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cas9 y cada uno de los péptidos HMGN. Células humanas K562 (1×10^6) se transfectaron con ADN plasmídico que codifica cada una de las proteínas de fusión o la proteína SpCas9 de tipo salvaje en cantidades equivalentes molares (5.2 y 5.0 μ g para cada una de las proteínas de fusión y la proteína Cas9 de tipo salvaje, respectivamente) en combinación con 3 μ g de un plásmido ARNg para apuntar a un sitio genómico (#1) en el locus de la oxidoreductasa del citocromo p450 (POR) humano. La transfección se llevó a cabo mediante nucleofección en un nucleofector Amaxi. Tres días después de la transfección, las células se lisaron con una solución de extracción de ADN (QuickExtract™) y la región genómica diana se amplificó por PCR. Los resultados, resumidos en la tabla 4, muestran que la fusión de cada uno de los péptidos HMGN humanos con la nucleasa aumentó la eficiencia de escisión de SpCas9 en el sitio diana.

Tabla 4. Eficiencia de escisión

Nucleasa	Sitio diana	Indel (%)
SpCas9 de tipo salvaje	POR/sitio #1	8.5
Fusión SpCas9-HMGN1	POR/sitio #1	18.3
Fusión SpCas9-HMGN2	POR/sitio #1	13.3
Fusión SpCas9-HMGN3a	POR/sitio #1	13.5
Fusión SpCas9-HMGN3b	POR/sitio #1	14.4

Ejemplo 3. Mejora de actividad de *Streptococcus pyogenes* Cas9 (SpCas9) utilizando el dominio globular central de la histona H1 humano

Un dominio globular central de histona H1 humano (SEQ ID NO: 45) se fusionó con SpCas9 (+NLS) en el extremo carboxilo terminal de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cas9 y el dominio globular. Células humanas K562 (1×10^6) se transfectaron con ADN plasmídico que codifica la proteína de fusión o la proteína SpCas9 de tipo salvaje en cantidades equivalentes molares (5.2 y 5.0 μ g para la proteína de fusión y la proteína Cas9 de tipo salvaje, respectivamente) en combinación con 3 μ g de un plásmido ARN_gu para apuntar a un sitio genómico (#1) en el locus de la oxidorreductasa del citocromo p450 (POR) humano. La transfección se llevó a cabo mediante nucleofección en un nucleofector Amaxi. Tres días después de la transfección, las células se lisaron con una solución de extracción de ADN (QuickExtract™) y la región genómica diana se amplificó por PCR. Las actividades de escisión de la diana Cas9 (% de indeles) se midieron utilizando ensayos Ce-I. Los resultados se presentan en la tabla 5. La fusión del dominio globular central de la histona H1 humana con la nucleasa aumentó la eficiencia de escisión de SpCas9 en el sitio diana.

Tabla 5. Eficiencia de escisión

Nucleasa	Sitio diana	Indel (%)
SpCas9 de tipo salvaje	POR/sitio #1	8.5
Fusión del dominio globular central SpCas9-H1	POR/sitio #1	19.4

Ejemplo 4. Mejora de la actividad de *Streptococcus pyogenes* Cas9 (SpCas9) utilizando un dominio de unión al ADN de una proteína de remodelación de cromatina

SpCas9 (+NLS) se fusionó con el dominio de unión al ADN del complejo de remodelación de cromatina ISWI de levadura ATPasa ISW1 (SEQ ID NO:46) en el extremo amino de la nucleasa con el enlazante TGSG (SEQ ID NO:2) entre Cas9 y el dominio de unión al ADN. De forma independiente, el SpCas9 de tipo salvaje se fusionó con el dominio de unión al ADN de la proteína 1 que contiene el dominio cromosómico de levadura (CHD1) (SEQ ID NO: 47) en el extremo carboxilo terminal de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cas9 y el dominio de unión al ADN. Las células humanas K562 (1×10^6) se transfectaron con ADN plasmídico que codifica cada una de las proteínas de fusión o la proteína SpCas9 de tipo salvaje en cantidades equivalentes molares (6.0 y 5.0 μ g para cada una de las proteínas de fusión y la proteína Cas9 de tipo salvaje, respectivamente) en combinación con 3 μ g de un plásmido ARN_gu para apuntar a un sitio genómico (#1) en el locus de la citocromo p450 oxidorreductasa (POR) humana. La transfección se llevó a cabo mediante nucleofección en un nucleofector Amaxi. Tres días después de la transfección, las células se lisaron con una solución de extracción de ADN (QuickExtract™) y la región genómica diana se amplificó por PCR. Las actividades de escisión de la diana Cas9 (% de indeles) se midieron utilizando ensayos Ce-I. Los resultados, resumidos en la tabla 6, muestran que la fusión de cada uno de los dominios de unión del ADN con la nucleasa aumentó la eficiencia de escisión de SpCas9 en el sitio diana.

Tabla 6. Eficiencia de escisión		
Nucleasa	Sitio diana	Indel (%)
SpCas9 de tipo salvaje	POR/sitio #1	8.5
Fusión del dominio de unión al ADN ISW1 con SpCas9	POR/sitio #1	21.1
Fusión del dominio de unión al ADN SpCas9-CHD1	POR/sitio #1	20.8

Ejemplo 5. Mejora de la actividad de *Streptococcus pyogenes* Cas9 (SpCas9) utilizando combinaciones de dominios que interactúan con el nucleosoma

SpCas9 (+NLS) se fusionó con el HMGN1 humano (SEQ ID N:41) en el extremo amino de la nucleasa con el
 5 enlazante TGSG (SEQ ID NO:2) entre Cas9 y HMGN1 y con el dominio de caja A del HMGB1 humano (SEQ
 ID NO:40) o el dominio globular central de la histona H1 humana (SEQ ID NO: 45) o el dominio de unión al
 ADN de la proteína 1 que contiene el dominio cromosómico de levadura (CHD1) (SEQ ID NO: 47) en el extremo
 10 carboxilo terminal de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cas9 y cada uno de los
 dominios de la proteína. Las células humanas K562 (1×10^6) se transfectaron con ADN plasmídico que codifica
 cada una de las proteínas de fusión o la proteína SpCas9 de tipo salvaje en cantidades equivalentes molares
 (5.4 µg para las proteínas de fusión de dominio globular central de la caja A de HMGB1 y H1, 6.0 µg para la
 15 proteína de fusión del dominio de unión al ADN de CHD1 y 5.0 µg para la proteína Cas9 de tipo salvaje) en
 combinación con 3 µg de un plásmido ARN_g para apuntar a un sitio genómico (#1, #2, #3) en el locus de la
 oxidoreductasa del citocromo p450 (POR) humano, o un sitio genómico (#1) del locus del miembro 3 del grupo
 I de la subfamilia 1 del receptor nuclear humano (CAR), o un sitio genómico (#1, #2) del locus de la homeobox
 1 1 de los espiráculos vacíos humanos (EMX1). La transfección se llevó a cabo mediante nucleofección en un
 nucleofector Amaxi. Cinco días después de la transfección, las células se lisaron con una solución de extracción
 20 de ADN (QuickExtract™) y cada región genómica diana se amplificó mediante PCR. Las actividades de escisión
 de la diana Cas9 (% de indeles) se midieron utilizando ensayos Cel-I. Los resultados, resumidos en la tabla 7,
 muestran que la fusión combinatoria de estos dominios proteicos con la nucleasa aumentó la eficiencia de
 escisión de SpCas9 en los sitios diana.

Tabla 7. Eficiencia de escisión		
Nucleasa	Sitio diana	Indel (%)
SpCas9 de tipo salvaje	POR/sitio #1	3.4
	POR/sitio #2	1.3
	POR/sitio #3	22.2
	CAR/sitio #1	2.1
	EMX1/sitio #1	2.2
	EMX1/sitio #2	1.1
Fusión de caja A HMGN1-SpCas9-HMGB1	POR/sitio #1	28.2
	POR/sitio #2	8.3
	POR/sitio #3	42.7
	CAR/sitio #1	14.3
	EMX1/sitio #1	29.0

	EMX1/sitio #2	12.1
Fusión del dominio globular central HMGN1-SpCas9-H1	POR/sitio #1	24.3
	POR/sitio #2	6.5
	POR/sitio #3	44.2
	CAR/sitio #1	23.9
	EMX1/sitio #1	26.9
	EMX1/sitio #2	21.0
Fusión del dominio de unión al ADN HMGN1-SpCas9-CHD1	POR/sitio #1	21.5
	POR/sitio #2	3.6
	POR/sitio #3	39.8
	CAR/sitio #1	9.0
	EMX1/sitio #1	23.5
	EMX1/sitio #2	20.2

Ejemplo 6. Mejora de la actividad de *Streptococcus pasteurianus* Cas9 (SpaCas9) utilizando combinaciones de dominios que interactúan con el nucleosoma

5 *Streptococcus pasteurianus* Cas9 (SpaCas9) (+NLS) se fusionó con el HMGN1 humano (SEQ ID NO:41) en el extremo amino de la nucleasa con el enlazante TGSG (SEQ ID NO:2) entre Cas9 y HMGN1 y con el dominio de caja A HMGB1 humano (SEQ ID NO:41) o el dominio globular central de la histona H1 humana (SEQ ID NO:45) o el dominio de unión al ADN de la proteína 1 que contiene el dominio cromosómico de levadura (CHD1) (SEQ ID NO:47) en el extremo carboxilo de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cas9 y cada uno de los dominios de la proteína. Las células humanas K562 (1×10^6) se transfectoron con ADN plasmídico que codifica cada una de las proteínas de fusión o la proteína SpaCas9 de tipo salvaje en cantidades equivalentes molares (5.4 y 5.0 μg para cada una de las proteínas de fusión y la proteína Cas9 de tipo salvaje, respectivamente) en combinación con 3 μg de un plásmido ARN_g para apuntar a un sitio genómico (#1, #2) en el locus de la oxidoreductasa del citocromo p450 (POR) humano. La transfección se llevó a cabo mediante nucleofección en un nucleofector Amaxi. Tres días después de la transfección, las células se lisaron con una solución de extracción de ADN (QuickExtract™) y la región genómica diana se amplificó por PCR. Las actividades de escisión de la diana Cas9 (% de índices) se midieron utilizando ensayos Cel-I. Como se resume en la tabla 8, la fusión combinatoria de estos dominios de proteína con la nucleasa aumentó la eficiencia de escisión de SpaCas9 en los sitios diana.

Tabla 8. Eficiencia de escisión		
Nucleasa	Sitio diana	Indel (%)
Tipo salvaje SpaCas9	POR/sitio #1	16.6
	POR/sitio #2	12.9
Fusión de caja A HMGN1-SpCas9-HMGB1	POR/sitio #1	20.6
	POR/sitio #2	35.8

Fusión del dominio globular central HMGN1-SpaCas9-H1	POR/sitio #1	28.6
	POR/sitio #2	31.7
Fusión del dominio de unión al ADN HMGN1-SpaCas9-CHD1	POR/sitio #1	19.4
	POR/sitio #2	18.5

Ejemplo 7. Mejora de la actividad de *Francisella novicida* Cpf1 (FnCpf1) utilizando combinaciones de dominios que interactúan con el nucleosoma

5 *Francisella novicida* Cpf1 (FnCpf1) (+NLS) se fusionó con el HMGN1 humano (SEQ ID NO:41) en el extremo amino de la nucleasa con el enlazante TGSG (SEQ ID NO:2) entre Cpf1 y HMGN1 y con el dominio de caja A HMGB1 humano (SEQ ID NO:40) o el dominio globular central de la histona H1 humana (SEQ ID NO:45) o el dominio de unión al ADN de la proteína 1 que contiene el dominio cromo de levadura (CHD1) (SEQ ID NO:47) en el extremo carboxilo de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cpf1 y cada uno de los dominios de la proteína. Las células humanas K562 (1×10^6) se transfectaron con ADN plasmídico que

10 codifica cada una de las proteínas de fusión o la proteína FnCpf1 de tipo salvaje en cantidades equivalentes molares (5.4 y 5.0 μ g para cada una de las proteínas de fusión y la proteína Cas9 de tipo salvaje, respectivamente) en combinación con 3 μ g de un plásmido ARNg para apuntar a un sitio genómico (#1, #2, #3) en el locus de la oxidoreductasa del citocromo p450 (POR) humano. La transfección se llevó a cabo mediante nucleofección en un nucleofector Amaxi. Tres días después de la transfección, las células se lisaron

15 con una solución de extracción de ADN (QuickExtract™) y la región genómica diana se amplificó por PCR. Las actividades de escisión de la diana Cas9 (% de indeles) se midieron utilizando ensayos Cel-I. Los resultados, resumidos en la tabla 9, muestran que la fusión combinatoria de estos dominios proteicos con la nucleasa aumentó la eficiencia de escisión de FnCpf1 en los sitios diana.

Tabla 9. Eficiencia de escisión		
Nucleasa	Sitio diana	Indel (%)
Tipo salvaje FnCpf1	POR/sitio #1	2.3
	POR/sitio #2	5.3
	POR/sitio #3	3.0
Fusión de caja A HMGN1-FnCpf1-HMGB1	POR/sitio #1	8.2
	POR/sitio #2	12.8
	POR/sitio #3	13.2
Fusión del dominio globular central HMGN1-FnCpf1-H1	POR/sitio #1	8.7
	POR/sitio #2	12.9
	POR/sitio #3	13.2
Fusión del dominio de unión al ADN HMGN1-FnCpf1-CHD1	POR/sitio #1	7.7
	POR/sitio #2	7.5
	POR/sitio #3	9.4

20 **Ejemplo 8. Mejora de eficiencia de la edición del gen *Campylobacter jejuni* Cas9 (CjCas9)**

Campylobacter jejuni Cas9 (CjCas9) (+NLS) se fusionó con el HMGN1 humano (SEQ ID NO:41) en el extremo amino de la nucleasa con el enlazante TGSG (SEQ ID NO:2) entre Cas9 y HMGN1 y con el dominio de caja A del HMGB1 humano (SEQ ID NO:40) o el dominio globular central de la histona H1 humana (SEQ ID NO:45) en el extremo carboxilo de la nucleasa con el enlazante LEGGGS (SEQ ID NO: 1) entre Cas9 y cada uno de los dominios de la proteína. El CjCas9 gRNA de tipo salvaje se modificó introduciendo una mutación de U a C en la región de repetición constante del crRNA y una mutación de A a G correspondiente en la región 5' de la secuencia del tracrRNA. La secuencia ARN_g modificada es: 5'-

La secuencia sgRNA modificada es: 5'-

NNNNNNNNNNNNNNNNNNNNNNNNNGUUCUAGUCCCUGAAAAGGGACUAGAAUAA
AGAGUUUGCGGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3',

donde los nucleótidos mutados en las unidades estructurales crRNA y tracrRNA están subrayados. Las secuencias guía dirigidas a dos sitios diferentes (#1, #2) en el gen de la oxidoreductasa del citocromo p450 humano (POR) se clonaron en el tipo salvaje y en el andamio CjCas9 ARN_g modificado, respectivamente. La expresión de los ARN_g estaba bajo el control de un promotor U6. Las células humanas K562 (1x 10⁶) se transfectaron con 4 µg de ADN plasmídico CjCas9 y 3 µg de un ADN plasmídico ARN_g. La transfección se llevó a cabo mediante nucleofección en un nucleofector Amaxi. Tres días después de la transfección, las células se lisaron con QuickExtract y las regiones genómicas dirigidas se amplificaron por PCR. Las actividades de escisión del ADN diana de CjCas9 (% de indeles) se midieron utilizando ensayos Cel-I. Los resultados se presentan en la figura 1 y muestran que las proteínas de fusión habían aumentado la eficiencia de escisión en los sitios diana, y que el andamiaje de CjCas9 ARN_g modificado aumentó efectivamente la eficiencia de escisión de CjCas9 en los sitios diana.

La tabla 10 presenta las secuencias de aminoácidos de las proteínas de fusión específicas. Los dominios de proteína que interactúan con el nucleosoma se muestran en **negrita**, los enlazantes se muestran en *cursiva* y el NLS está subrayado.

Tabla 10. Proteínas de fusión CRISPR

Fusión de la caja A de SpCas9-HMGB1 (SEQ ID NO: 61)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSSTDKADLRLLIYALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSLARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAMWTRKSEETITPWNFEFVVDKGASQSFIERMTNFDKNLPNEKVLPPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKITLDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQDSLHEHIANLAGSPAIIKKGILQTVKIVDELVKVMGHKPENIVIMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDYALNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRLIETNGTEGIWVDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYASHYEKLKGSPEDEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSIITGLYETRIDLSQLGGDPKKKRKVLEGGGSGMGKGDPKKPRGKMSSYAFFVQT

**CREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGFEDMAKADKARYEREMKTYIPPKG
E**

Fusión SpCas9-HMGN1 (SEQ ID NO: 62)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVK
LFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
DFYFPFLKDNREKIEKILTFRIPIYVVGPLARGNSRFAMTRKSEETITPWNFEVVVDKGASA
QSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNE
ENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTCWGRLSRKLINGIR
DKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
IKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
ILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNK
VLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELKAGF
IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVR
INNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKATAYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTE
VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGSKKLKS
KELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQ
KGNELALPSKYVNFYLYLASHYEKLKGSPEDEQQLFVEQHKHYLDEIIEQISEFSKRVIL
ADANLDKVL SAYNKH RD KPIREQAENIIHLETLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQSI TGLYETRIDLSQLGGDPKKRKVLEGGGGSMPKRKVSSAE GAKEE PKRRSA
RLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGQAEVANQETKEDLPAENGETKT
EESPASDEAGEKEAKSD

Fusión SpCas9-HMGN2 (SEQ ID NO: 63)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
 TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLVDSITDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVK
 LFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
 LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIPKILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
 DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASA
 QSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNE
 ENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
 DKQSGKTILDFLKSDFANRNFQMLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPA
 IKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ
 ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNK
 VLTRSDKNRGSNDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF
 IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVR
 INNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDKYVDVRKMIAKSEQEI GKATAYFF
 YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTE
 VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEGKSKKLKSV
 KELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQ
 KGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVIL
 ADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
 DATLIHQ SITGLYETRIDLSQLGGDPKKRKVL LEGGGGSMPKRKAEGDAKGD KAKVKDEPQ
 RRSARLSAKPAPPKPEPKPKKAPAKKGEKVPKGKKGKADAGKEGNNPAENGDAKTDQAQKA
 EGAGDAK

Fusión SpCas9-HMGN3a (SEQ ID NO: 64)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
 TRLKRTARRRYTRRNRI CYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLVDS TDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVK
 LFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
 LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
 DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVV DKGASA
 QSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNE
 ENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
 DKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
 IKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQ
 ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNK
 VLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF
 IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK VRE
 INNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVDVRKMIAKSEQEIGKAT AKYFF
 YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTE
 VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEGKSKKLKSV
 KELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ
 KGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEII EQISEFSKR VIL
 ADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
 DATLIHQSI TGLYETRIDLSQLGGDPKKRKVLEGGGGSMPKRKSPENTEGKDGSKVTQKE
 PTRRSARLSAKPAPPKPEPKPRKTS AKKEPGAKISRGAKGKKEEKQEAGKEGTAPSENGET
 KAEEAQKTESVDNEGE

Fusión SpCas9-HMGN3b (SEQ ID NO: 65)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
 TRLKRTARRRYTRRNRI CYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLVDS TDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVK
 LFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
 LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE
 DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVV DKGASA
 QSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
 DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNE
 ENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
 DKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
 IKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEG IKELGSQ
 ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNK
 VLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF
 IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK VRE
 INNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVDVRKMIAKSEQEIGKAT AKYFF

YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTE
VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSV
KELLGITIMERSSEKPNIDFLEAKGYKEVKKDIIKLPKYSLFELENGRKRMLASAGELQ
KGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVL
ADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQSI TGLYETRIDLSQLGGDPKKKRKVLEGGGGSM PKRKSPENTEGKD GSKVTKQE
PTRRSARLSAKPAPPKPEPKPRKTS AKKEPGA KISRGAKGKKEEKQEAGKEGTEN

Fusión globular SpCas9-Histona H1 (SEQ ID NO:66)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETABEA
TRLKRTARRRYTRRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNI
VDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVK
LFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI
LRVNTEITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAILRRQE
DFYPFLKDNREKIEKILTRIPYYVGPLARGNSRFWMTRKSEETITPWNFEVVVDKGASA
QSFIERTMNFNFKNL PNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIV
DLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNE
ENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR
DKQSGKTILDFLKS DGFANRNFQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
IKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQ
ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNK
VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF
IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRE
INNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVDVRKMIAKSEQEIGKATAKYFF
YSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTE
VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSV
KELLGITIMERSSEKPNIDFLEAKGYKEVKKDIIKLPKYSLFELENGRKRMLASAGELQ
KGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVL
ADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQSI TGLYETRIDLSQLGGDPKKKRKVLEGGGGSS TDHPKYS DMIVA AIQAEKNRA
GSSRQSIQKYIKSHYKVG ENADSQIKLSIKRLVTTGVLKQTKGVGASGSFRLAKSDEP

Fusión ISWI-SpCas9 (SEQ ID NO:67)

LLNPTKRERKENYSIDNYYKDVLTNTRGSSTPSHPRMPKPHVFHSHQLQPPQLKVLYEKERM
 WTAKKTGYVPTMDDVKAAYGDISDEEEKKQKLELLKLSVNNSQPLTEEEEKMKADWESEGF
 TNWNKLEFRKFITVSGKYGRNSIQAIARELAPGKTLEEVRAYAKAFWSNIERIEDYEKYLK
 I I E N E E E K I K R V K M Q Q E A L R R K L S E Y K N P F F D L K L K H P P S S N N K R T Y S E E E D R F I L L M L F K
 YGLDRDDVYELVRDEIRDCLFELDFYFRSRTFVELARRGNTLLQCLEKEFNAGIVLDDAT
 KDRMKKEDENGKRIRREEFADQTANEKENVDGVESSKKAKIEDTSNVTGTEQLVAEKIPENETT
 HTGSGMDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDG
 ETAEATRLKRTARRRYTRRNRI CYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHP
 IFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDN
 SDVDKLFILQVLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFG
 NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDIILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG
 YIDGGASQEEFYKFIPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAI
 LRRQEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAMMTRKSEETITPWNFEFVVD
 KGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQ
 KKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKLIKDKD

 FLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKL
 INGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
 AGSPAIAKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK
 ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDD
 SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
 DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF
 YKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKAT
 AKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNI
 VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSK
 KLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKDLIIKLPKYSLFELENGKRMLAS
 AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEQKQLFVEQHKHYLDEIIIEQISEFS
 KRVLADANLDKVL SAYNKHDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTS
 TKEVL DATLIHQSI TGLYETRIDLSQLGGDPKKKKRV

Fusión SpCas9-CHD1 (SEQ ID NO: 68)

MDKKYSIGLDIGTNSVGAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNFQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF IKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDPKKKRKVLEGGGGSDMDSIGESEVRALYKAILKFGNLKEIILDELIADGTLPVKSEFKYGETYDEMMEA AKDCVHEEEKNRKEILEKLEKHATAYRAKLKSGEIKAENQPKDNPLTRL SLKKREKKAVLFNFKGVKSLNAESLLSRVEDLKY LKNLINSNYKDDPLKFSLGNNTPKPVQNWSSNWTKEEDEKLLIGVFKYGYGSWTQIRDDPFLGITDKIFLNEVHN PVAKKSASSSDTTPTPSKKGKGITGSSKKVPGA IHLGRRVDYLLSFLRGGLNTKSPS

Fusión de la caja A HMGN1-SpCas9-HMGB1 (SEQ ID NO: 69)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSDTGSGMDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA

KLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILT
 FRIPYYVGPLARGNSRFAMWTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED
 YFKKIECFDSVEISGVEDRFNASLGTYHDLKLIKDKDFLDNEENEDILEDIVLTLTLFED
 REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKIVDELVKV
 MGHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSE
 EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQI
 LDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG
 ALIKKYPKLESEFVYGDKYVDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGE
 IRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSD
 KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKPNP
 IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
 SHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI
 TGLYETRIDLSQLGGDPKKKRKVLGGGGSGKGD PPKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEF
SKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGE

Fusión globular HMGN1-SpCas9-Histona H1 (SEQ ID NO: 70)

MPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAAGK
 QAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSDTGSMDKKYSIGLDIGTNSVGW
 AVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI
 CYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK
 KLV DSTDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPI
 NASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSDFDLAEDA
 KLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILT
 FRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED
 YFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTTLTFED
 REMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQCDLSLHEHIANLAGSPAIIKGILQTVKIVDELVKV
 MGHKPENIVIAMARENQTTQKGQKNSRERMKRIE EG I KELGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSE
 EVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQI
 LDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVG
 ALIKKYPKLESEFVYGDYKVYDVRKMIKSEQF IGKATAKYFFYSNIMNFFKTEITLANGE
 IRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSD
 KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKPNP
 IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
 SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRID
 LSQLGCDPKKKRKVLEGGGSSTDHPKYS DMIVAAIQAEKNRAGSSRQSIQKYIKSHYKVG
 ENADSQIKLSIKRLVTTGVLKQTKGVGASGSFRLAKSDEP

Fusión HMGN1-SpCas9-CDH1(SEQ ID NO: 71)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGK
 QAEVANQETKEDLPAENGETKTEES PASDEAGEKEAKSD TGSQMDKKYSIGLDIGTNSVGW
 AVITDEYKVPSSKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI
 CYLQEIFSNEMAKVDDSEFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK
 KLVDSTDKADLRILIYLAHAMIKFERGHFLIEGDLNPDNSDVDFKFIQLVQTYNQLFEEENPI
 NASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNEFLAEDA
 KLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILT
 FRIPYYVVGPLARGNSRFAMTRKSEETITPWNFEFVVDKGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED
 YFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTLTLFED
 REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQGDLSHEHIANLAGSPAIKKGILQTVKIVDELVKV
 MGHKPNIVIAMARENQTTQKGQKNSRERMKRIE EG I KELGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSE
 EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQI
 LDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVG
 ALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE
 IRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKESILPKRNSD
 KLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKPNP
 IDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLA
 SHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAYNKHARDK
 PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRID
 LSQLGGD PKKKRKVLEGGGGSDMDSIGESEVRALYKAILKFGNLKEILDELIADGTLPVKS
 FEKYGETYDEMMEA AKDCVHEEEKNRKEILEKLEKHATAYRAKLKSGE IKAENQPKDNPLT
 RLSLKKREKKAVLFNFKGKSLNAESLLSRVEDLKYLKNLINSNYKDDPLKFSLGNNTPKP
 VQNWSSNWTKEEDEKLLIGVFKYGYGSWTQIRDDPFLGITDKIFLNEVHNPAKKSASSSD
 TTTPPSKKGKGITGSSKKVPGA IHLGRRVDYLLSFLRGGLNTKSPS

Fusión de la caja A HMGN1-SpaCas9-HMGB1 (SEQ ID NO: 72)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGK
 QAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD TGSQMTNGKILGLDIGIASVGV
 GIIEAKTGKVVHANSRLFSAANAENNAERPGFRGSRRLNRRKKHRVKEVRDLFEKYGIVTD
 FRNLNLNPFYELPVKGLTEQLKNEELFAALPTISKRRGISYLDDEADDSTGSTDYAKSIDEN
 RRLKKNKTPGQIQLEKLEKYQQLRGNTVYDENGAEHRLINVFSTSDYEKEARKILETQAD
 YNKKITAEFIDDYVEILTQKRKYHGGPGNEKSRTDYGRFRDGTTLLENIFGILIGKCNFYF
 DEYRASKASYTAQEYNFLNDLNLKVSTETGKLSTEQKESLVEFAKNTATLGPALLKEIA
 KILDCKVDEIKGYPEDDKGKPDLTTFEPYRKLKFNLESINIDDLREVIDKLADILTLNTE
 REGIELAIKENLNPQNFTTEEQISEIIKVRKSQSTAFNKGWHSFSAKLMNELIPELYATSDEQ
 MTILTRLEKFKVNKKSSKNTKTIDEKEVTDEIYNPVVAKSVRQTIKIINAADVKKYGFDFKI
 VIEMPRDKNADDEKKFIDKRNKENKKEKDDALKRAAYLYNSSDKLPDEVFHNKQLETKIR
 LWYQOGERCLYSGKPISIQELVHNSNNFEIDHILPLSLSFDDSLANKVLVYAWTNQEKQK
 TPYQVIDSMDAAWSFREM KDYVLKQKGLGKKKRDYLLTTENIDKIEVKKKFIERNLVDTRY
 ASRVVLNSLQSALRELKGDTKVSVVRGQFTSQLRREWKIDKSRETYHHHAVDALIIAASSQ
 LKLWEKQDNPMFVDYGNQVVDKQTEGILSVSDDEYKELVFQPPYQGFVNTISSKGFEDFI
 LFSYQVDSKYNRKVS DATIYSTRKAKIGKDKKEETVVLGKIKDIYSQNGFDTFIKKYNKDK
 TQFLMYQKDSLWENVIEVILRDYPTTKKSEDGKNDVKCNPFEEYRRENGLICKYSKKGKG
 TPIKSLKYDYDKKLGNCIDITPEESRNKVILQSINPWKADVFNPETLKYELMGLKYSDSL
 EKGTGNYHISQEKYDAIKEKEGIGKKSEFKFTLYRNDLILIKDIASGEQEIYRFLSRTMPN
 VNHYVELKPYDKEKFDNVQELVEALGEADKVGRCIKGLNKPNI SIYKVRTDVLGNKYFVKK
 KGDKPKLDFKNNKKPKKKRKVLEGGGSGSGKDPKKPRGKMSSYAFFVQTCREEHKKKHPDA
 SVNFSFSSKKCSERWKTMSAKEKGFEDMAKADKARYEREMKTYIPPKGE

Fusión globular HMGN1-SpaCas9-Histona H1 (SEQ ID NO: 73)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGK
 QAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD TGSQMTNGKILGLDIGIASVGV
 GIIEAKTGKVVHANSRLFSAANAENNAERRGFRGSRRLNRRKKHRVKEVRDLFEKYGIVTD
 FRNLNLNPFYELRVKGLTEQLKNEELFAALRTISKRRGISYLDDEADDSTGSTDYAKSIDEN
 RRLKKNKTPGQIQLEKLEKYQQLRGNTVYDENGAEHRLINVFSTSDYEKEARKILETQAD
 YNKKITAEFIDDYVEILTQKRKYHGGPGNEKSRTDYGRFRDGTTLLENIFGILIGKCNFYF
 DEYRASKASYTAQEYNFLNDLNLKVSTETGKLSTEQKESLVEFAKNTATLGPALLKEIA
 KILDCKVDEIKGYREDDKGKPDLTTFEPYRKLKFNLESINIDDLREVIDKLADILTLNTE
 REGIEDAIKRNLPQNFTTEEQISEIIKVRKSQSTAFNKGWHSFSAKLMNELIPELYATSDEQ
 MTILTRLEKFKVNKKSSKNTKTIDEKEVTDEIYNPVVAKSVRQTIKIINAADVKKYGFDFKI
 VIEMPRDKNADDEKKFIDKRNKENKKEKDDALKRAAYLYNSSDKLPDEVFHNKQLETKIR
 LWYQOGERCLYSGKPISIQELVHNSNNFEIDHILPLSLSFDDSLANKVLVYAWTNQEKQK
 TPYQVIDSMDAAWSFREM KDYVLKQKGLGKKKRDYLLTTENIDKIEVKKKFIERNLVDTRY
 ASRVVLNSLQSALRELKGDTKVSVVRGQFTSQLRREWKIDKSRETYHHHAVDALIIAASSQ
 LKLWEKQDNPMFVDYGNQVVDKQTEGILSVSDDEYKELVFQPPYQGFVNTISSKGFEDFI
 LFSYQVDSKYNRKVS DATIYSTRKAKIGKDKKEETVVLGKIKDIYSQNGFDTFIKKYNKDK
 TQFLMYQKDSLWENVIEVILRDYPTTKKSEDGKNDVKCNPFEEYRRENGLICKYSKKGKG
 TPIKSLKYDYDKKLGNCIDITPEESRNKVILQSINPWKADVFNPETLKYELMGLKYSDSL
 EKGTGNYHISQEKYDAIKEKEGIGKKSEFKFTLYRNDLILIKDIASGEQEIYRFLSRTMPN
 VNHYVELKPYDKEKFDNVQELVEALGEADKVGRCIKGLNKPNI SIYKVRTDVLGNKYFVKK
 KGDKPKLDFKNNKKPKKKRKVLEGGGSSDTHPKYS DMIVAAIQAEKNRAGSSRSIQKYI
 KSHYKVGENADSIKLSIKRLVTTGVLKQTKGVGASGSFRLAKSDEP

Fusión HMGN1-SpaCas9-CHD1 (SEQ ID NO: 74)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGK
 QAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSDTGSGMTNGKILGLDIGIASVGV
 GIIEAKTGKVVHANSRLFSAANAENNAERRGFRGSRRLNRRKKHRVKRVRDLFEKYGIVTD
 FRNLNLPYELRVKGLTEQLKNEELFAALRTISKRRGISYLDDEADDSTGSTDYAKSIDEN
 RRLKKNKTPGQIQLEERLEKYGQLRGNFTVYDENGAEHRLINVFSTSDYEKEARKILETQAD
 YNKKITAEFIDDYVEILTQKRKYHGPCNEKSRTDYGRFRDGTTLNIFGILIGKCNFYF
 DEYRASKASYTAQEYNFLNDLNNLKVSTETGKLSTEQKESLVEFAKNTATLGPAKLLKEIA
 KILDCKVDEIKGYREDDKGKPDLTTFEPYRKLKFNLESINIDDLRSREVIDKLADILTLNTE
 REGIEDAIKRNLPNQFTTEEQISEIIKVRKSQSTAFNKGWHSFSAKLMNELIPELYATSDEQ
 MTILTRLEKFKVNKKSSKNTKTIDEKEVTDEIYNPVVAKSVRQTIKIINAAVKKYGDFDKI
 VIEMPRDKNADDEKKFIDKRNKENKKEKDDALKRAAYLYNSSDKLPDEVFVHGNKQLETKIR
 LWYQQGERCLYSGKPISIQELVHNSNNFEIDHILPLSLSFDDSLANKVLVYAWTNQEKGGK
 TPYQVIDSMDAAWSFREMKDYVLKQKGLGKKKRDYLLTTENIDKIEVKKKFIERNLVDTRY
 ASRVVLNSLQSAALRELKDKTKVSVVRGQFTSQLRRKWKIDKSRETYHHHHAVDALI IAASSQ
 LKLWEKQDNPMFVDYGNQVVDKQTGEILSVSDDEYKELVFQPPYQGFVNTISSKGFEDEI
 LFSYQVDSKYNRKVSDATIYSTRKAKIGKDKKEETYVLGKIKDIYSQNGFDTFIKKYNKDK
 TQFLMYQKDSL TWENVIEVILRDYPTTKKSEDGKNDVKCNPFEEYRRENGLICKYSKKGKG
 TPIKSLKYYDKKLGNCIDITPEESRNKVILQSINPWRADVYFNPETLKYELMGLKYSDSLSE
 EKG TGNYHISQEKYDAIKEKEGIGKKSEFKFTLYRNDLILIKDIASGEQEIYRFLSRTMPN
 VNHYVELKPYDKEKFDNVQELVEALGEADKVGRCIKGLNKPNI SIYKVRTDVLGNKYFVKK
 KGDKPKLDFKNNKPKKKRKLVEGGGSDMDSIGESEVRALYKAILKFGNLKEILDELIAD
 GTLPVKSEFEKYGE TYDEMEEAAKDCVHEEEKNRKEILEKLEKHATAYRAKLKSGEIKAEHQ
 PKDNPLTRLSLKKREKKAVLFNFKGVKSLNAESLLSRVEDLKYLKNLINSNYKODPLKFSL
 GNNTPKPVQNWSSNWTKEEDEKLLIGVFYGYGSWTQIRDDPFLGITDKIFLNEVHNPVAK
 KSASSSDTTPTPSKKGKGITGSSKKVPGAIHLGRRVDYLLSFLRGGLNTKSPS

Fusión HMGN1-FnCpf1-HNGB1 (SEQ ID NO: 75)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDCKVQTKGKRGAKGK
 QAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD TGSMSIYQEFVNKYSLSKTLR
 FELIPQKGTLENIKARGLILDDEKRAKDYKKAQI IDKYHQFFIEEILSSVCISEDLLQNY
 SDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLIL
 WLKQSKDNGIELFKANSDITDIDEALEI IKSFKGWTTYFKGFHENRKNVYSSNDIPTSIY
 RIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDE
 VFEIANFNNYLNQSGITKENTIIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSV
 LFKQILSDTESKSFVIDKLEDDSDVVTMQSFYEQIAAFKTVEEKS IKETLSLFLDLDLKAQ
 KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKT
 EKAKYLSLETIKLALAEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKY
 QNQGKDLLQASAEDDVKAIKDLLDQTNLLHLKLIKFIHISQSEDKANILDKDEHFYLVFEE
 CYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKY
 YLGVMNKKNNKI FDDKA IENKGEYKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDIL
 RIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSI
 DEFYREVENQGYKLT FENISESYIDSVVNQGLYLFQIYNKDFSAYS KGRPNLHTLYWKAL
 FDERNLQDVVYKLNGEAELFYRKQSI PKKITHPAKEAIANKNDNPKKESVFEYDLIKDKR
 FTEDKFFHFHPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKG
 NIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAK
 LVIEYNAIVVFEDLNFQFKRGRFKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAY
 QLTAPFETFKKMGKQTGI IYVPAGFTSKI CPVTGFVNQLYPKYESVSKSQEFFSKFDKIC
 YNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVPYPTKELEKL
 LKDYSIEYGHGECIKAAICGESDKKFFAKLT SVLNTILQMRNSKTGTELDYLISPVADVNG
 NFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRKKNQEGKKNLVIKNEEYFEFVQNRN
 NPKKKRKVLEGGGSGKGD PPKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSE
 RWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGE

Fusión globular HMGN1-FnCpf1-Histona H1 (SEQ ID NO: 76)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDCKVQTKGKRGAKGK
 QAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD TGSMSIYQEFVNKYSLSKTLR
 FELIPQKGTLENIKARGLILDDEKRAKDYKKAQI IDKYHQFFIEEILSSVCISEDLLQNY
 SDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLIL
 WLKQSKDNGIELFKANSDITDIDEALEI IKSFKGWTTYFKGFHENRKNVYSSNDIPTSIY
 RIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDE
 VFEIANFNNYLNQSGITKENTIIGGKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSV
 LFKQILSDTESKSFVIDKLEDDSDVVTMQSFYEQIAAFKTVEEKS IKETLSLFLDLDLKAQ
 KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKT
 EKAKYLSLETIKLALAEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKY
 QNQGKDLLQASAEDDVKAIKDLLDQTNLLHLKLIKFIHISQSEDKANILDKDEHFYLVFEE
 CYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKY
 YLGVMNKKNNKI FDDKA IENKGEYKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDIL
 RIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSI
 DEFYREVENQGYKLT FENISESYIDSVVNQGLYLFQIYNKDFSAYS KGRPNLHTLYWKAL
 FDERNLQDVVYKLNGEAELFYRKQSI PKKITHPAKEAIANKNDNPKKESVFEYDLIKDKR
 FTEDKFFHFHPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKG
 NIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAK
 LVIEYNAIVVFEDLNFQFKRGRFKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAY
 QLTAPFETFKKMGKQTGI IYVPAGFTSKI CPVTGFVNQLYPKYESVSKSQEFFSKFDKIC

YNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKL
 LKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNG
 NFFDSRQAPKNMPQDADANGAYHIGLKGLMLLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRN
 NPKKKRKVLEGGGSSTDHPKYSDMIVAATQAEKNRAGSSRQSIQKYIKSHYKVGENADSQ
 IKLSIKRLVTTGVLKQTKGVGASGSFRLAKSDEP

Fusión HMGN1-FnCpf1-CHD1 (SEQ ID NO: 77)

MPKRKVSSAEGAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGK
 QAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSDTSGMSIYQEFVNKYSLSKTLR
 FELIPOGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCISEDLLQNY
 SDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFNLFNQNLIDAKKGQESDLIL
 WLKQSKDNGIELFKANSDITDIDEALEI IKSFKGWTTFYFKGFHENRKNVYSSNDIPTSIY
 RIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDE
 VFEIANFNMYNLNQSGITKFNTIIIGGKFVNGENTKRKGINEYINLYSQINDKTLKKYKMSV
 LFKQILSDTESKS FVIDKLEDDSDVVTMQSFYEQIAAFKTVEEKS IKETLSLLFDDLKAQ
 KLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKT
 EKAKYLSLETIKLAL EEFNKHRDIDKQCRFEEILANFAAIPMI FDEIAQNKDNLAQISIKY
 QNQGKKDLLQASAEDDVKAIKDLLDQTNLLHKLKIFHISQSEDKANILDKDEHYLVFEE
 CYFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKY
 YLGVMNKKNNKI FDDKAIKENKGEYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDIL
 RIRNHSTHTKNGSPQKGYEKFENIEDCRKFIDFYKQSIKHPPEWKDFGFRFSDTQRYNSI
 DEFYREVENQGYKLT FENISESYIDSVVNQGKLYLFQIYNKDFSAYS KGRPNLHTLYWKAL
 FDERNLQDVVYKLNGEAELFYRKQSI PKKI THPAKEAIANKNDNPKKESVFEYDLIKDKR
 FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKG
 NIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAK
 LVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNLYLVFKDNEFDKTGGVLRAY
 QLTAPFETFKKMGKQTGIIYYVPAGFTSKI CPVTGFVNQLYPKYESVSKSQEFFSKFDKIC
 YNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKL
 LKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNG
 NFFDSRQAPKNMPQDADANGAYHIGLKGLMLLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRN
 NPKKKRKVLEGGGSSDMSIGESEVRALYKAILKFGNLKEILDELIADGTLPVKSFEKYGE
 TYDEMMEAAKDCVHEEEKNRKEILEKLEKHATAYRAKLKSGEIKAEQPKDNPLTRLKSLKK
 REKKAVLFNFKGKSLNAESLLSRVEDLKYLNKLNINSNYKDDPLKFSLGNNTPKPVQNWSS
 NWTKEEDEKLLIGVFKYGYGSWTQIRDDPFLGITDKIFLNEVHNPAKKSASSSDTTPTPS
 KKGKGITGSSKKVPGAHLGRRVDYLLSFLRGGLNTKSPS

Fusión de la caja A HMGN1-CjCas9-HMGB1 (SEQ ID NO: 78)

MPKRKVSSAE GAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDCKVQTKGKRGAKGK
 QAEVANQE TKEDLPAENGETKTEESPASDEAGEKEAKSD TGSGMARILAFDIGISSIGWAF
 SENDELKDCGVRI FTKVENPKTGESLALPRRLARSARKRLARRKARLNHLKHLIANEFKLN
 YEDYQSFDES LAKAYKGLISPYELRFRALNELLSKQDFARVILHIAKRRGYDDIKNSDDK
 EKGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENSKEFTNVRNKKESYERCTAQSF
 DELKLI FKKQREFGFSFSKKFEEVLSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLA
 FMFVALTRI INLLNNLKNTEGILYTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFKG
 EKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDLNQNQIDS
 LSKLEFKDHLNISFKALKLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKD
 EVTNPVVLRAIKEYRKVLNALLKKYGVHVKINIELAREVGKNHSQRAKIEKEQNENYKAKK
 DAELECEKLGLKINSKNILKRLRFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSF
 DDSYMNKVLVFTKQNEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKE
 QKNFKDRNLNDTRYIARLVNLNTKDYLDLPLSDDENTKLNDTQKGSKVHVEAKSGMLTSA
 LRHTWGFSAKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYK
 NKRKFFEPFSGFRQKVLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQSYGKEGVLKA
 LELGKIRKVNGKIVKNGDMFRVDIFKHKKTNKIFYAVPIYTMDFALKVLPNKAVARSKKGEI
 KDWILMDENYEFCSFLYKDSLILIQTKDMQEPEFVYNAFTSSTVSLIVSKHDNKFETLSK
 NQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKKPKKKRKVLEG
 GGGSGKGD PKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKG
 KFEDMAKADKARYEREMKTYIPPKGE

Fusión globular HMGN1-CjCas9-Histona H1 (SEQ ID NO: 79)

MPKRKVSSAE GAAKEE PKRRSARLSAKPPAKVEAKPKKAAAKDKSSDCKVQTKGKRGAKGK
 QAEVANQE TKEDLPAENGETKTEESPASDEAGEKEAKSD TGSGMARILAFDIGISSIGWAF
 SENDELKDCGVRI FTKVENPKTGESLALPRRLARSARKRLARRKARLNHLKHLIANEFKLN
 YEDYQSFDES LAKAYKGLISPYELRFRALNELLSKQDFARVILHIAKRRGYDDIKNSDDK
 EKGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENSKEFTNVRNKKESYERCTAQSF
 DELKLI FKKQREFGFSFSKKFEEVLSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLA
 FMFVALTRI INLLNNLKNTEGILYTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFKG
 EKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDLNQNQIDS
 LSKLEFKDHLNISFKALKLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKD
 EVTNPVVLRAIKEYRKVLNALLKKYGVHVKINIELAREVGKNHSQRAKIEKEQNENYKAKK
 DAELECEKLGLKINSKNILKRLRFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSF
 DDSYMNKVLVFTKQNEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKE
 QKNFKDRNLNDTRYIARLVNLNTKDYLDLPLSDDENTKLNDTQKGSKVHVEAKSGMLTSA
 LRHTWGFSAKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYK
 NKRKFFEPFSGFRQKVLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQSYGKEGVLKA
 LELGKIRKVNGKIVKNGDMFRVDIFKHKKTNKIFYAVPIYTMDFALKVLPNKAVARSKKGEI
 KDWILMDENYEFCSFLYKDSLILIQTKDMQEPEFVYNAFTSSTVSLIVSKHDNKFETLSK
 NQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKKPKKKRKVLEG
 GGGSSSTDHPKYSDMIVAAIQAEKNRAGSSRSIQKYIKSHYKVGENADSIKLSIKRLVTT
 GVLKQTKGVGASGSFRLAKSDEP

REIVINDICACIONES

1. Una proteína de fusión que comprende una proteína de repeticiones palindrómicas cortas agrupadas y regularmente intercaladas (CRISPR) unida a al menos un dominio de proteína que interactúa con el nucleosoma, en la que
- 5 la proteína CRISPR es una nucleasa o nickasa CRISPR/Cas9 de tipo II, o la proteína CRISPR es una nucleasa o nickasa CRISPR/Cpf1 de tipo V, y en la que
la proteína de fusión comprende además al menos una señal de localización nuclear, y en la que
el al menos un dominio de proteína que interactúa con el nucleosoma es un dominio de unión a ADN de caja de grupo de alta movilidad (HMG) (HMGB), una proteína de unión a nucleosomas HMG (HMGN), un dominio
- 10 globular central de una variante de histona H1, un dominio de unión a ADN de una proteína del complejo de remodelación de cromatina, o una combinación de los mismos.
2. La proteína de fusión de la reivindicación 1, que comprende además al menos un dominio de penetración celular, al menos un dominio marcador o una combinación de los mismos.
3. La proteína de fusión de la reivindicación 1 o la reivindicación 2, en la que la proteína CRISPR es una proteína
- 15 CRISPR/Cas9 de tipo II modificada para carecer de toda actividad de nucleasa y unida a un dominio no nucleasa, o una proteína CRISPR/Cpf1 de tipo V modificada para carecer de toda actividad de nucleasa y unida a un dominio no nucleasa.
4. La proteína de fusión de la reivindicación 3, en la que el dominio no nucleasa tiene actividad de citosina desaminasa, actividad de histona acetiltransferasa, actividad de activación transcripcional o actividad represora
- 20 transcripcional.
5. La proteína de fusión de cualquier reivindicación anterior, en la que al menos un dominio de proteína que interactúa con el nucleosoma es el dominio de caja A de HMGB1, la proteína HMGN1, la proteína HMGN2, la proteína HMGN3a, la proteína HMGN3b, el dominio globular central de la histona H1, el dominio de unión al
- 25 ADN de la proteína de cambio de imitación (ISWI), el dominio de unión al ADN de la proteína 1 de cromodominio-helicasa-ADN (CHD1), o una combinación de los mismos.
6. La proteína de fusión de cualquier reivindicación anterior, en la que el al menos un dominio de proteína que interactúa con el nucleosoma está unido a la proteína CRISPR directamente mediante un enlace químico, indirectamente mediante un enlazante o una combinación de los mismos.
7. La proteína de fusión de cualquier reivindicación anterior, en la que el al menos un dominio de proteína que
- 30 interactúa con el nucleosoma está unido a la proteína CRISPR en su extremo-N, extremo-C, una ubicación interna o una combinación de los mismos.
8. La proteína de fusión de cualquier reivindicación anterior, en la que la proteína CRISPR es *Streptococcus pyogenes* Cas9 (SpCas9), *Streptococcus thermophilus* Cas9 (StCas9), *Streptococcus pasteurianus* (SpaCas9),
- 35 *Campylobacter jejuni* Cas9 (CjCas9), *Staphylococcus aureus* (SaCas9), *Francisella novicida* Cas9 (FnCas9), *Neisseria cinerea* Cas9 (NcCas9), *Neisseria meningitidis* Cas9 (NmCas9), *Francisella novicida* Cpf1 (FnCpf1), *Acidaminococcus* sp. Cpf1 (AsCpf1), o bacteria *Lachnospiraceae* ND2006 Cpf1 (LbCpf1).
9. La proteína de fusión de cualquier reivindicación anterior, en la que la proteína de fusión tiene una secuencia de aminoácidos que tiene al menos aproximadamente un 90 % de identidad de secuencia con SEQ ID NO:61,
- 40 SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, o SEQ ID NO:79.
10. La proteína de fusión de cualquier reivindicación anterior, en la que la proteína de fusión tiene una secuencia de aminoácidos como se establece en SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID
- 45 NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, o SEQ ID NO:79.
11. Un ácido nucleico que codifica la proteína de fusión de cualquier reivindicación anterior.
12. El ácido nucleico de la reivindicación 11, que está optimizado en codones para su traducción en una célula eucariota.
- 50 13. El ácido nucleico de la reivindicación 11 o 12, que es parte de un vector viral, un vector plasmídico o un ARN autorreplicante.

14. Un complejo que comprende al menos una proteína de fusión de cualquiera de las reivindicaciones 1-10 y al menos un ARN guía.
15. Un kit que comprende la proteína de fusión según una cualquiera de las reivindicaciones 1-10, el ácido nucleico según una cualquiera de las reivindicaciones 11-13 o el complejo según la reivindicación 14.
- 5 16. Una célula eucariota que comprende la proteína de fusión según una cualquiera de las reivindicaciones 1-10, el ácido nucleico según una cualquiera de las reivindicaciones 11-13 o el complejo según la reivindicación 14.
17. Un método para aumentar la eficiencia de la modificación genómica o epigenética dirigida en una célula eucariota, comprendiendo el método la introducción en la célula eucariota:
- 10 (a) al menos una proteína de fusión o un ácido nucleico que codifica al menos una proteína de fusión, comprendiendo cada proteína de fusión una proteína CRISPR unida a al menos un dominio de proteína que interactúa con el nucleosoma, en la que la proteína CRISPR (i) tiene actividad de nucleasa o nickasa o (ii) está modificada para carecer de toda actividad de nucleasa y está unida a un dominio no nucleasa, en la que la proteína CRISPR es una proteína CRISPR/Cas9 de tipo II o una proteína CRISPR/Cpf1 de tipo V, y en la que
- 15 la proteína de fusión comprende además al menos una señal de localización nuclear;
- (b) al menos un ARN guía o un ácido nucleico que codifica al menos un ARN guía;
- en el que la proteína CRISPR de la al menos una proteína de fusión está dirigida a una secuencia cromosómica diana y el al menos un dominio de proteína que interactúa con el nucleosoma de la al menos una proteína de fusión altera la estructura nucleosómica o de cromatina de tal manera que la al menos una proteína de fusión
- 20 tiene un mayor acceso a la secuencia cromosómica diana, aumentando así la eficiencia de la modificación genómica o epigenética dirigida, y
- en el que el al menos un dominio de proteína que interactúa con el nucleosoma es un dominio de unión a ADN de caja de grupo de alta movilidad (HMG) (HMGB), una proteína de unión a nucleosomas HMG (HMGN), un dominio globular central de una variante de histona H1, un dominio de unión a ADN de una proteína del
- 25 complejo de remodelación de cromatina, o una combinación de los mismos,
- en el que el método no comprende un método para el tratamiento del cuerpo humano o animal mediante cirugía o terapia.
18. El método de la reivindicación 17, en el que la proteína de fusión comprende además al menos un dominio de penetración celular, al menos un dominio marcador o una combinación de los mismos.
- 30 19. El método de la reivindicación 17 o la reivindicación 18, en el que el dominio no nucleasa tiene actividad de citosina desaminasa, actividad de histona acetiltransferasa, actividad de activación transcripcional o actividad represora transcripcional.
20. El método de cualquiera de las reivindicaciones 17-19, en el que el al menos un dominio de proteína que interactúa con el nucleosoma está unido a la proteína CRISPR directamente mediante un enlace químico,
- 35 indirectamente mediante un enlazante, o una combinación de los mismos; y/o
- en el que el al menos un dominio de proteína que interactúa con el nucleosoma está unido a la proteína CRISPR en su extremo-N, extremo-C, una ubicación interna o una combinación de los mismos.
21. El método de cualquiera de las reivindicaciones 17-20, en el que el ácido nucleico que codifica la al menos una proteína de fusión está optimizado en codones para la traducción en la célula eucariota, y/o en el que el
- 40 ácido nucleico que codifica la al menos una proteína de fusión es parte de un vector viral, un vector plasmídico o un ARN autorreplicante.
22. El método de cualquiera de las reivindicaciones 17-21, en el que el método comprende además introducir en la célula eucariota al menos un polinucleótido donante, comprendiendo el polinucleótido donante al menos una secuencia donante.
- 45 23. El método de cualquiera de las reivindicaciones 17-22, en el que la célula eucariota es una célula de mamífero o una célula humana.
24. El método de cualquiera de las reivindicaciones 17-22, en el que la célula es una célula vegetal.
25. La proteína de fusión de cualquiera de las reivindicaciones 1-10 o método de cualquiera de las
- 50 reivindicaciones 17-24, en el que la proteína CRISPR está unida a dos dominios de proteína que interactúan con el nucleosoma.

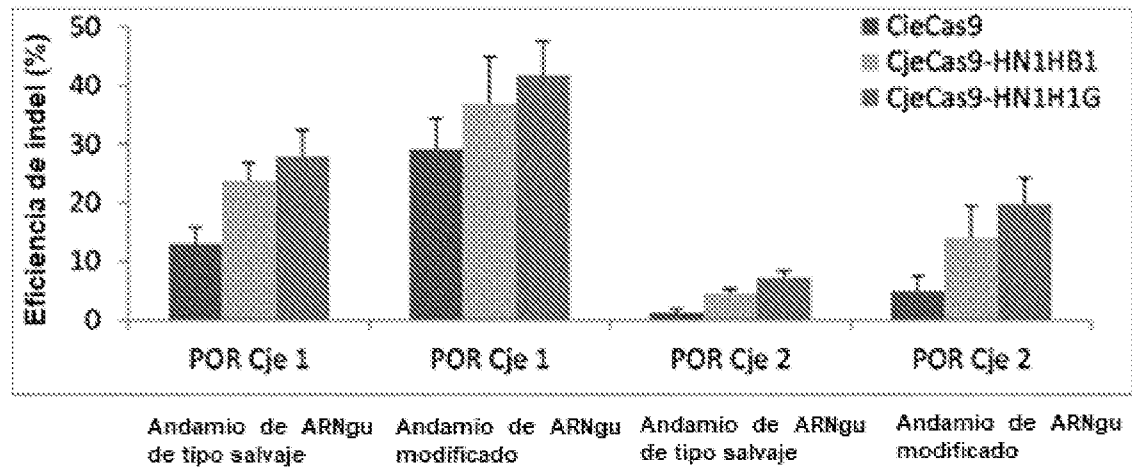


FIG. 1