

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 788 176**

51 Int. Cl.:

C12Q 1/68 (2008.01)

C07K 16/40 (2006.01)

C12Q 1/6841 (2008.01)

C12Q 1/6853 (2008.01)

C12Q 1/682 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.09.2016 PCT/US2016/053386**

87 Fecha y número de publicación internacional: **30.03.2017 WO17053762**

96 Fecha de presentación y número de la solicitud europea: **23.09.2016 E 16849727 (9)**

97 Fecha y número de publicación de la concesión europea: **18.03.2020 EP 3337909**

54 Título: **Métodos y reactivos para la detección de proximidad molecular usando proteínas de unión a ácido nucleico guiadas por ARN**

30 Prioridad:

24.09.2015 US 201562232023 P

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

20.10.2020

73 Titular/es:

**SIGMA ALDRICH CO. LLC (100.0%)
3050 Spruce Street
St. Louis, MO 63103, US**

72 Inventor/es:

**DAVIS, GREGORY D.;
PALHAN, VIKAS y
KREADER, CAROL A.**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 788 176 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

Métodos y reactivos para la detección de proximidad molecular usando proteínas de unión a ácido nucleico guiadas por ARN

Campo

5 La presente divulgación generalmente se refiere a sondas que comprenden proteínas de unión a ácido nucleico guiadas por ARN y al uso de complejos que comprenden dichas sondas para detectar ácidos nucleicos endógenos *in situ* por proximidad molecular.

Antecedentes

10 El ADN genómico nuclear eucariótico se empaqueta en cromatina y se dispone en el núcleo como cromosomas. La organización espacial de la cromatina y las posiciones relativas de regiones cromosómicas específicas están estrechamente asociadas a la regulación génica durante la función celular normal y la enfermedad. La hibridación fluorescente *in situ* (FISH), una técnica citogenética que usa sondas fluorescentes que se unen solo a aquellas partes del cromosoma con un alto grado de complementariedad de secuencia, se ha usado para visualizar dominios de cromatina, genes o ARN específicos, incluidos tanto ARN mensajeros (ARNm) como ARN no codificantes (ARNnc).
 15 Sin embargo, esta técnica requiere un tratamiento riguroso para desnaturalizar el ADN genómico bicatenario para la hibridación de la sonda, y dicho tratamiento afecta a la integridad de la estructura y/u organización cromosómica de la cromatina y a la estructura secundaria y estabilidad del ARN y sus interacciones con las proteínas y la cromatina. Por tanto, existe la necesidad de técnicas de imagenología de cromatina y ARN *in situ* que utilicen condiciones de reacción más suaves y proporcionen una amplificación de señal robusta. Wulan Deng et al.: "CASFISH: CRISP/Cas-9 mediated
 20 *in situ* labelling of genomic loci in fixed cells", Proceedings of National Academy of Sciences of the United States of America, vol. 112, nº. 38, agosto de 2015, páginas 11870-11875, XP055553074 divulga la ligadura de un oligonucleótido marcado a ARNs para generar una sonda sgMajSat. El documento WO 2012/152942 se relaciona con el desarrollo de sondas de proximidad y su uso en un ensayo de detección.

Compendio

25 Entre los diversos aspectos de la presente divulgación en la provisión de un complejo adecuado para el ensayo de ligadura de proximidad se comprende una primera sonda que comprende (i) una primera proteína de unión a ácido nucleico guiada por ARN que es un sistema asociado a repeticiones palindrómicas cortas agrupadas y regularmente espaciadas (CRISPR) (Cas) manipulado genéticamente de origen no natural que comprende una proteína CRISPR/Cas y un ARN guía y (ii) un primer oligonucleótido que está ligado directa o indirectamente al primer sistema
 30 CRISPR/Cas; el complejo comprende además una segunda sonda que comprende (a) un segundo sistema CRISPR/Cas y un segundo oligonucleótido que está ligado directa o indirectamente al segundo sistema CRISPR/Cas o (b) un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico y un segundo oligonucleótido que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico. En algunas realizaciones, la proteína CRISPR/Cas es una proteína Cas9. En algunos casos, la proteína Cas9 comprende dos dominios nucleasa funcionales, un dominio de nucleasa funcional o ningún dominio de nucleasa funcional. En otras realizaciones, la proteína Cas9 comprende al menos una señal de localización nuclear. En una realización, la proteína Cas9 no tiene dominio de nucleasa funcional y carece de toda actividad de nucleasa. En otra realización, la proteína Cas9 carece de toda actividad de nucleasa y comprende al menos una señal de localización nuclear. En algunas realizaciones, la primera sonda comprende el primer sistema CRISPR/Cas y el primer oligonucleótido está ligado a la proteína CRISPR/Cas o al ARN guía del sistema CRISPR/Cas mediante enlaces covalentes o no covalentes, y opcionalmente la segunda sonda comprende (i) el segundo sistema CRISPR/Cas y el segundo oligonucleótido está ligado a la proteína CRISPR/Cas del segundo sistema CRISPR/Cas mediante enlaces covalentes o no covalentes o al ARN guía del segundo sistema CRISPR/Cas mediante enlaces covalentes o no covalentes. En realizaciones adicionales, la primera sonda comprende el primer sistema CRISPR/Cas y el primer oligonucleótido está ligado a un anticuerpo primario dirigido contra la proteína CRISPR/Cas; y opcionalmente la segunda sonda comprende un anticuerpo dirigido contra una proteína asociada a ácido nucleico o modificación de ácido nucleico y el segundo oligonucleótido que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico. En realizaciones adicionales, la primera sonda comprende un sistema CRISPR/Cas, los primer y segundo oligonucleótidos que están ligados a anticuerpos secundarios antiespecie, y anticuerpos primarios dirigidos contra la proteína CRISPR/Cas. En algunas iteraciones, el complejo que comprende la sonda que contiene CRISPR/Cas que comprende el primer y el segundo anticuerpos ligados a anticuerpos secundarios antiespecie comprende además al menos un sistema CRISPR/Cas adicional. En ciertas realizaciones, los oligonucleótidos de los complejos detallados anteriormente son monocatenarios y/o comprenden estructuras secundarias de tallo, bucle y/u horquilla. En otras realizaciones, los oligonucleótidos de los complejos comprenden desoxirribonucleótidos, ribonucleótidos o combinaciones de los mismos, en los que los desoxirribonucleótidos o ribonucleótidos son nucleótidos estándares, análogos de nucleótidos o nucleótidos modificados con 2'-O-metilo.

Un aspecto adicional de la presente divulgación abarca kits que comprenden uno cualquiera de los complejos detallados anteriormente.

Todavía otro aspecto de la presente divulgación proporciona un método para detectar un ácido nucleico endógeno en una célula. El método comprende poner en contacto la célula con un complejo de sonda de detección de proximidad como se define en cualquiera de las realizaciones mencionadas anteriormente, en el que la proteína de unión a ácido nucleico guiada por ARN es guiada por ARN a un primer sitio en el ácido nucleico endógeno para la unión, formando así un complejo de sonda de detección de proximidad unido y visualizando el complejo de sonda de detección de proximidad unido a través de una reacción de amplificación dependiente de proximidad *in situ* para detectar el ácido nucleico endógeno. En algunas realizaciones, la segunda sonda comprende (a) el segundo sistema CRISPR/Cas que comprende un segundo ARN guía orientado a un segundo sitio en el ácido nucleico endógeno y el segundo oligonucleótido de detección de proximidad está ligado a la proteína CRISPR/Cas del segundo sistema CRISPR/Cas mediante enlaces covalentes o no covalentes; (b) el segundo sistema CRISPR/Cas comprende un segundo ARN guía orientado a un segundo sitio en el ácido nucleico endógeno y el segundo oligonucleótido de detección de proximidad está ligado al segundo ARN guía del segundo sistema CRISPR/Cas mediante enlaces covalentes o no covalentes; o (c) un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico y un segundo oligonucleótido de detección de proximidad que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico. En otras realizaciones, el primer oligonucleótido de detección de proximidad está ligado a un anticuerpo primario dirigido contra la proteína CRISPR/Cas, y opcionalmente la segunda sonda comprende un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico en el ácido nucleico endógeno y el segundo oligonucleótido de detección de proximidad está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico. En realizaciones adicionales, el complejo de sonda de detección de proximidad comprende la primera sonda que comprende el primer sistema CRISPR/Cas que comprende la proteína CRISPR/Cas y el ARN guía que está orientado al primer sitio en el ácido nucleico endógeno, y los primer y segundo oligonucleótidos que están ligados a anticuerpos secundarios antiespecie, y el complejo de sonda de detección de proximidad comprende además anticuerpos primarios dirigidos contra la proteína CRISPR/Cas. En algunas iteraciones de esta realización, la célula se pone en contacto con el segundo o adicional sistema CRISPR/Cas que comprende el segundo o adicional ARN guía dirigido al segundo o adicional sitio en el ácido nucleico endógeno. En diversas realizaciones, el ácido nucleico endógeno es ADN cromosómico nuclear, ARN mensajero o ARN no codificante. En ciertas realizaciones, la célula es una célula primaria, una célula de estirpe celular o dentro de una muestra de tejido o fluido obtenida de un sujeto. En otras realizaciones, la célula está viva, fijada o congelada. En realizaciones adicionales, la reacción de amplificación dependiente de proximidad *in situ* comprende un ensayo de ligadura de proximidad (PLA) o un inicio de reacción en cadena de hibridación dependiente de la proximidad (proxHCR).

Otros aspectos e iteraciones de la divulgación se detallan a continuación.

Breve descripción de las figuras

La FIG. 1 esquematiza la detección de un locus genómico a través de la unión genómica proximal de dos complejos dCas9-ARNg y amplificación de señal a través de un ensayo de ligadura de proximidad (PLA). La amplificación de la señal se inicia mediante moléculas de ácido nucleico monocatenario unidas covalentemente o no covalentemente a las proteínas dCas9.

La FIG. 2 ilustra la detección de un locus genómico y una proteína proximal X a través de la unión genómica proximal de un complejo dCas9-ARNg compatible con PLA y un anticuerpo compatible con PLA a la proteína X proximal, y una amplificación de señal basada en PLA.

La FIG. 3 esquematiza la detección de un locus genómico y una proteína proximal X a través de la unión genómica proximal de un complejo dCas9-ARNg y un anticuerpo compatible con PLA a la proteína proximal X. La amplificación basada en PLA utiliza un anticuerpo anti-Cas9 compatible con PLA.

La FIG. 4 ilustra la detección de un locus genómico a través de la unión genómica proximal de dos complejos cas9-ARNg compatibles con PLA en los que la señal de PLA se inicia por moléculas de ácido nucleico monocatenario unidas covalentemente o no covalentemente a las moléculas de ARNg.

La FIG. 5 esquematiza la detección de un locus genómico usando un complejo dCas9-ARNg y el inicio de la amplificación de la señal a través de PLA usando anticuerpos policlonales (de conejo) anti-Cas9 y sondas anti-PLA de conejo (+) y (-).

La FIG. 6 ilustra la detección de un locus genómico utilizando complejos dCas9-ARNg (ensamblados con ARNg que recubren el locus genómico) y el inicio de la amplificación de la señal a través de PLA usando anticuerpos policlonales (de conejo) anti-Cas9 y sondas anti-PLA de conejo (+) y (-).

La FIG. 7 presenta un diagrama esquemático de un cromosoma. La cromatina centromérica es única porque CENP-A (izquierda) reemplaza a la histona H3 (derecha).

Las FIG. 8A-D muestran imágenes de células U2OS nucleofectadas con complejos CRISPR RNP formados con dCas9, DIG-ARNtracr y ARNcr satélite menor o ARNcr de control negativo 1 (NC1). El ensayo PLA (DUOLINK®) se realizó usando anticuerpos anti-Cas9 y/o anti-CENP-A. Las imágenes superpuestas muestran la señal de Duolink en rojo (Cy5), el núcleo (ADN) en azul (DAPI) y la actina en verde (Phalloidin Atto488). La FIG. 8A muestra una célula en

contacto con ARNcr satélite menor y anticuerpos anti-Cas9 y anti-CENP-A. La FIG. **8B** presenta una célula en contacto con ARNcr de control negativo y anticuerpos anti-Cas9 y anti-CENP-A. La FIG. **8C** muestra una célula en contacto con ARNcr satélite menor y anticuerpos anti-CENP-A. La FIG. **8D** presenta una célula en contacto con ARNcr satélite menor y anticuerpos anti-Cas9.

5 La FIG. **9A** ilustra esquemáticamente un telómero con las proteínas TRF1 y TRF2 asociadas. La FIG. **9B** ilustra la unión de las proteínas TRF1 y TRF2 a la secuencia de repetición de los telómeros TTAGGG (subrayado) del ADN de los telómeros (5'-TTAGGGTTAGGGTTAGGGTTAGGG-3'; SEQ ID NO: 18), que está apareada por bases con el ADN (AATCCCAATCCC; SEQ ID NO: 19) o ARN de molde de telomerasa que comprende CAAUCCCAAUC (SEQ ID NO: 20).

10 Las FIG. **10A-D** muestran imágenes de células U2OS nucleofectadas con complejos CRISPR RNP formados con dCas9, DIG-ARNtracr y ARNcr de telómero o ARNcr de control negativo 1 (NC1). El ensayo PLA (DUOLINK®) se realizó usando anticuerpos anti-Cas9 y/o anti-TRF2. Las imágenes superpuestas muestran la señal de Duolink en rojo (Cy5) y el núcleo (ADN) en azul (DAPI). La FIG. **10A** muestra una célula en contacto con ARNcr de telómero y anticuerpos anti-Cas9 y anti-TRF2A. La FIG. **10B** presenta una célula en contacto con ARNcr de control negativo y anticuerpos anti-Cas9 y anti-TRF2. La FIG. **10C** muestra una célula en contacto con ARNcr de telómero y anticuerpos anti-Cas9. La FIG. **10D** presenta una célula en contacto con ARNcr de telómeros y anticuerpos anti-TRF2.

15 Las FIG. **11A-D** muestran imágenes de células U2OS fijadas incubadas con complejos CRISPR RNP formados con dCas9, DIG-ARNtracr y ARNcr satélite menor o ARNcr de control negativo 1 (NC1). El ensayo PLA (DUOLINK®) se realizó usando anticuerpos anti-Cas9 y/o anti-CENP-A. Las imágenes superpuestas muestran la señal de Duolink en rojo (Cy5) y el núcleo (ADN) en azul (DAPI). La FIG. **11A** muestra una célula en contacto con ARNcr satélite menor y anticuerpos anti-Cas9 y anti-CENP-A. La FIG. **11B** presenta una célula en contacto con ARNcr de control negativo y anticuerpos anti-Cas9 y anti-CENP-A. La FIG. **11C** muestra una célula en contacto con ARNcr satélite menor y anticuerpos anti-Cas9. La FIG. **11D** presenta una célula en contacto con ARNcr satélite menor y anticuerpos anti-CENP-A.

25 Descripción detallada

La presente divulgación proporciona reactivos y métodos para la detección de proximidad molecular de ácidos nucleicos endógenos específicos *in situ* usando proteínas de unión a ácido nucleico guiadas por ARN. Las proteínas de unión a ácido nucleico guiadas por ARN, tales como las codificadas por los sistemas asociados (CAS) a repeticiones palindrómicas cortas agrupadas y regularmente espaciadas de tipo II (CRISPR) (CRISPR/Cas) poseen propiedades de unión a ácido nucleico potentes y flexibles. El sistema CRISPR/Cas de *Streptococcus pyogenes* se ha estudiado con gran detalle bioquímico y estructural con un enfoque en su capacidad enigmática para desenrollar ADN bicatenario en ausencia de cualquier actividad de helicasa dependiente de energía (Sternberg et al., Nature, 2014, 506 (7490): 62-67) El sistema CRISPR/Cas se ha usado como una sonda enzimática altamente específica y eficiente para marcar secuencias repetidas de ADN en núcleos no alterados de células y tejidos fijados (Deng et al., Proc Natl Acad Sci U S A., 2015, 112 (37): E5123-32). En la presente memoria, la capacidad de detección del sistema CRISPR/Cas se incrementa dramáticamente por la detección de proximidad. En la presente memoria, se proporcionan complejos de sonda de detección de proximidad que comprenden al menos un sistema CRISPR/Cas, métodos para usar dichos complejos de sonda de detección de proximidad para detectar ácidos nucleicos endógenos *in situ* y kits que comprenden dichos complejos de sonda de detección de proximidad.

40 I. Complejos de sonda de detección de proximidad

Un aspecto de la divulgación proporciona un complejo de sonda de detección de proximidad que comprende una proteína de unión a ácido nucleico guiada por ARN. La proteína de unión a ácido nucleico guiada por ARN puede ser un sistema CRISPR/Cas que comprende una proteína CRISPR/Cas y un ARN guía (ARNg). El complejo de sonda de detección de proximidad divulgado en la presente memoria comprende al menos una sonda que comprende un sistema CRISPR/Cas, que dirige la unión de la sonda a un ácido nucleico diana, y un oligonucleótido de detección de proximidad, que facilita la detección de la sonda unida mediante un ensayo de detección de proximidad amplificable. En general, el oligonucleótido de detección de proximidad está ligado directa o indirectamente al sistema CRISPR/Cas. En algunas realizaciones, el oligonucleótido de detección de proximidad puede estar ligado directamente a la proteína CRISPR/Cas (véase la FIG. 1) o al ARN guía (véase la FIG. 4) de la sonda. En otras realizaciones, el oligonucleótido de detección de proximidad puede estar ligado a un anticuerpo dirigido contra la proteína CRISPR/Cas (véase la FIG. 3, sonda en el sitio genómico 1). En realizaciones adicionales, el o los oligonucleótidos de detección de proximidad pueden estar ligados a anticuerpos secundarios antiespecie, y el complejo de sonda de detección de proximidad comprende además anticuerpos primarios dirigidos contra la proteína CRISPR/Cas (véanse las FIG. 5 y 6). El complejo de sonda de detección de proximidad puede comprender además sondas adicionales como se detalla a continuación.

55 (a) Sondas que comprenden el sistema CRISPR/Cas

El complejo de sonda de detección de proximidad comprende al menos una sonda que comprende un resto de unión que comprende una proteína de unión a ácido nucleico guiada por ARN y al menos un oligonucleótido de detección de proximidad. La proteína de unión a ácido nucleico guiada por ARN puede ser un sistema CRISPR/Cas que

comprende una proteína CRISPR/Cas y un ARN guía

(i) Sistema CRISPR/Cas

El complejo de sonda de detección de proximidad divulgado en la presente memoria comprende al menos una sonda que comprende un sistema CRISPR/Cas. El sistema CRISPR/Cas puede ser un sistema CRISPR/Cas genomanipulado no natural. Un sistema CRISPR/Cas comprende una proteína CRISPR/Cas y un ARN guía.

Proteínas CRISPR/Cas. La proteína CRISPR/Cas del complejo de sonda de detección de proximidad puede derivar de un sistema CRISPR de tipo I (concretamente, IA, IB, IC, ID, IE o IF), tipo II (concretamente, IIA, IIB o IIC), tipo III (concretamente, IIIA o IIIB) o tipo V, que están presentes en varias bacterias y arqueas. El sistema CRISPR/Cas puede ser de *Streptococcus sp.* (p.ej., *Streptococcus pyogenes*), *Campylobacter sp.* (p.ej., *Campylobacter jejuni*), *Francisella sp.* (p.ej., *Francisella novicida*), *Acaryochloris sp.*, *Acetohalobium sp.*, *Acidaminococcus sp.*, *Acidithiobacillus sp.*, *Alicyclobacillus sp.*, *Allochromatium sp.*, *Ammonifex sp.*, *Anabaena sp.*, *Arthrospira sp.*, *Bacillus sp.*, *Burkholderiales sp.*, *Caldicelulosiruptor sp.*, *Candidatus sp.*, *Clostridium sp.*, *Crocospaera sp.*, *Cyanothece sp.*, *Exiguobacterium sp.*, *Fingoldia sp.*, *Ktedonobacter 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.* o *Thermosiphon sp.*

Los ejemplos no limitantes de proteínas CRISPR/Cas adecuadas incluyen proteínas Cas, proteínas Cpf, proteínas Cmr, proteínas Csa, proteínas Csb, proteínas Csc, proteínas Cse, proteínas Csf, proteínas Csm, proteínas Csn, proteínas Csx, proteínas Csy, proteínas Csz, y derivados o variantes de las mismas. En realizaciones específicas, la proteína CRISPR/Cas puede ser una proteína Cas9 de tipo II, una proteína Cpf1 de tipo V o un derivado de la misma. En algunas realizaciones, la proteína CRISPR/Cas puede ser Cas9 de *Streptococcus pyogenes* (SpCas9) o Cas9 de *Streptococcus thermophilus* (StCas9). En otras realizaciones, la proteína CRISPR/Cas puede ser Cas9 de *Campylobacter jejuni* (CjCas9). En realizaciones alternativas, la proteína CRISPR/Cas puede ser Cas9 de *Francisella novicida* (FnCas9). En todavía otras realizaciones, la proteína CRISPR/Cas puede ser Cpf1 de *Francisella novicida* (FnCpf1).

La proteína CRISPR/Cas puede comprender al menos un dominio de reconocimiento de ARN y/o un dominio de unión a ARN, que interactúan con el ARN guía. La proteína CRISPR/Cas también puede comprender dominios de nucleasa (concretamente, dominios de ADNasa o ARNasa), dominios de unión a ADN, dominios de helicasa, dominios de ARNasa, dominios de interacción proteína-proteína, dominios de dimerización, así como otros dominios.

En algunas realizaciones, la proteína CRISPR/Cas puede derivar de una proteína Cas9 de tipo silvestre o fragmento de la misma. En general, las proteínas Cas9 comprenden al menos dos dominios de nucleasa (concretamente, ADNasa). Por ejemplo, una proteína Cas9 puede comprender un dominio de nucleasa similar a RuvC y un dominio de nucleasa similar a HNH. Los dominios RuvC y HNH trabajan juntos cortando cadenas individuales para hacer una ruptura bicatenaria en el ADN. (Jinek et al., Science, 2013, 337: 816-821) En otras realizaciones, la proteína CRISPR/Cas puede derivar de la proteína Cas9 modificada. Por ejemplo, una proteína Cas9 o un derivado de la misma puede modificarse para contener solo un dominio de nucleasa funcional (ya sea un dominio de nucleasa similar a RuvC o similar a HNH). En particular, la proteína derivada de Cas9 o derivado de la misma puede modificarse de modo que uno de los dominios de nucleasa se elimine o mute de modo que el dominio de nucleasa ya no sea funcional (concretamente, la actividad nucleasa está ausente). En realizaciones en las que uno de los dominios de nucleasa está inactivo, la proteína Cas9 es capaz de escindir una hebra de ADN bicatenario (concretamente, tal proteína se denomina "nicasa"). Una nicasa no puede crear una ruptura bicatenaria en el ADN. Por ejemplo, una conversión de aspartato a alanina (D10A) en un dominio similar a RuvC convierte la proteína Cas9 o derivado de la misma en una nicasa. Del mismo modo, una conversión de histidina a alanina (H840A o H839A) en un dominio HNH convierte la proteína Cas9 o derivado de la misma en una nicasa. En realizaciones adicionales, tanto el dominio de nucleasa similar a RuvC como el dominio de nucleasa similar a HNH pueden modificarse o eliminarse de modo que la proteína Cas9 o derivado de la misma no sea capaz de mellar o escindir ácido nucleico bicatenario (tal proteína se denomina Cas9 muerta o dCas9). Por ejemplo, la proteína Cas9 puede comprender mutaciones D10A y H840A (o D10A y H839A). En algunas realizaciones, todos los dominios de nucleasa de una proteína Cas9 o derivado de la misma pueden modificarse o eliminarse de modo que la proteína derivada carezca de toda actividad nucleasa.

En cualquiera de las realizaciones descritas anteriormente, cualquiera o todos los dominios de nucleasa pueden inactivarse mediante una o más mutaciones de delección, mutaciones de inserción y/o mutaciones de sustitución usando métodos bien conocidos, tales como mutagénesis dirigida al sitio, mutagénesis mediada por PCR y síntesis génica total, así como otros métodos conocidos en la técnica.

En algunas realizaciones, la proteína CRISPR/Cas puede comprender además al menos una señal de localización nuclear. En general, una NLS comprende un tramo de aminoácidos básicos. Las señales de localización nuclear son conocidas en la técnica (véase, por ejemplo, Lange y col., J. Biol. Chem., 2007, 282: 5101-5105) Por ejemplo, en una realización, la NLS puede ser una secuencia monopartita, tal como PKKKRKV (SEQ ID NO: 1) o PKKKRRV (SEQ ID NO: 2). En otra realización, la NLS puede ser una secuencia bipartita, tal como KRPAATKKAGQAKKKK (SEQ ID NO: 3). La NLS puede localizarse en el extremo N, el extremo C o en una localización interna de la proteína CRISPR/Cas.

En todavía otras realizaciones, la proteína CRISPR/Cas puede comprender además al menos un dominio de penetración celular. En una realización, el dominio de penetración celular puede ser una secuencia peptídica penetrante de células derivada de la proteína TAT del VIH-1. Como ejemplo, la secuencia de penetración celular TAT puede ser GRKKRRQRRRPPQPKKKRKY (SEQ ID NO: 4). En otra realización, el dominio de penetración celular puede ser TLM (PLSSIFSRIGDPPKKKRKY; SEQ ID NO: 5), una secuencia peptídica penetrante celular derivada del virus de la hepatitis B humana. En todavía otra realización, el dominio de penetración celular puede ser MPG (GALFLGWLGAAGSTMGAPKKKRKY; SEQ ID NO: 6 o GALFLGFLGAAGSTMGAWSPKKKRKY; SEQ ID NO: 7). En una realización adicional, el dominio de penetración celular puede ser Pep-1 (KETWWETWWTEWSQPKKKRKY; SEQ ID NO: 8), VP22, un péptido penetrante celular del virus del herpes simple o una secuencia peptídica de poliarginina. El dominio de penetración celular puede localizarse en el extremo N, el extremo C o en una localización interna de la proteína CRISPR/Cas.

En todavía otras realizaciones, la proteína CRISPR/Cas también puede comprender al menos un dominio marcador. Los ejemplos no limitantes de dominios marcadores incluyen proteínas fluorescentes, marcajes de purificación y marcajes de epítipo. En algunas realizaciones, el dominio marcador puede ser una proteína fluorescente. Los ejemplos no limitantes de proteínas fluorescentes adecuadas incluyen proteínas fluorescentes verdes (p. ej., GFP, GFP-2, tagGFP, turboGFP, EGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), proteínas fluorescentes amarillas (p.ej., YFP, EYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), proteínas fluorescentes azules (p.ej., EBFP, EBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-Sapphire), proteínas fluorescentes cian (p.ej., ECFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), proteínas fluorescentes rojas (p.ej., mKate, mKate2, mPlum, monómero DsRed, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred) y proteínas fluorescentes naranjas (p. ej., mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, mTangerine, tdTomato) o cualquier otra proteína fluorescente adecuada. En otras realizaciones, el dominio marcador puede ser un marcaje de purificación y/o un marcaje de epítipo. Los marcajes ejemplares incluyen, pero sin limitación, glutation-S-transferasa (GST), proteína de unión a quitina (CBP), proteína de unión a maltosa, tiorredoxina (TRX), poli(NANP), marcaje de purificación de afinidad en tándem (TAP), myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, proteína portadora de carboxilo de biotina (BCCP) y calmodulina. El dominio marcador puede localizarse en el extremo N o el extremo C de la proteína CRISPR/Cas.

La proteína CRISPR/Cas se puede obtener de fuentes comerciales. Como alternativa, el ADN que codifica la proteína CRISPR/Cas puede clonarse en un vector de expresión y la proteína CRISPR/Cas puede expresarse y purificarse a partir de células bacterianas o eucarióticas usando procedimientos estándares. En todavía otras realizaciones, la proteína CRISPR/Cas se puede proporcionar como un ácido nucleico codificante. Por ejemplo, el ácido nucleico codificante puede ser ARN mensajero. Como alternativa, el ácido nucleico codificante puede ser ADN (p.ej., puede ser parte de un vector), en el que el ADN codificante puede estar ligado operativamente a una secuencia de control promotora adecuada (p.ej., promotores eucarióticos/mamíferos tales como CMV, SV40, RSV, MMTV). El ácido nucleico codificante puede optimizarse de codón para expresión en la célula de interés.

ARN guía. El sistema CRISPR/Cas también comprende un ARN guía. El ARN guía interactúa con la proteína CRISPR/Cas para dirigir la proteína CRISPR/Cas a un sitio diana específico en una secuencia cromosómica o ácido nucleico endógeno, en el que el extremo 5' del ARN guía se empareja por bases con una secuencia protoespaciadora específica en la secuencia diana.

El sitio diana no tiene limitación de secuencia, excepto que la secuencia es seguida inmediatamente (en dirección 3') por una secuencia de consenso. Esta secuencia de consenso también se conoce como motivo adyacente protoespaciador (PAM en inglés) Por ejemplo, las secuencias de PAM para Cas9 incluyen 3'-NGG, 3'-NGGNG, 3'-NNAGAAW y 3'-ACAY y las secuencias de PAM para Cpf1 incluyen 5'-TTN (en la que N se define como cualquier nucleótido, W se define como A o T, e Y se define como C o T). En algunas realizaciones, el sitio diana puede estar en la región de codificación de un gen, en un intrón de un gen, en una región de control de un gen, en una región espaciadora entre genes, en una región no codificante, etc. El gen puede ser un gen codificante de proteína o un gen codificante de ARN. En otras realizaciones, el sitio diana puede estar en una molécula de ARN.

Cada ARN guía comprende tres regiones: una primera región en el extremo 5' que es complementaria del sitio diana en la secuencia de ácido nucleico, una segunda región interna que forma una estructura de bucle y tallo y una tercera región 3' que permanece esencialmente monocatenaria. La primera región (en el extremo 5') de cada ARN guía es diferente, de tal modo que cada ARN guía guía la proteína CRISPR/Cas a un sitio diana específico. La segunda y tercera regiones de cada ARN guía pueden ser las mismas en todos los ARN guía.

La primera región del ARN guía es complementaria del sitio diana en la secuencia de ácido nucleico de tal modo que la primera región del ARN guía puede emparejarse por bases con el sitio diana. La primera región de cada ARN guía comprende una secuencia que es complementaria de la secuencia diana. En diversas realizaciones, la primera región del ARN guía puede comprender de aproximadamente 10 nucleótidos a más de aproximadamente 25 nucleótidos. Por ejemplo, la región de emparejamiento de bases entre la primera región del ARN guía y el sitio diana en la secuencia de ácido nucleico puede ser de aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25 o más de 25 nucleótidos de longitud. En realizaciones específicas, la primera región del ARN guía tiene aproximadamente 20 nucleótidos de longitud. Como ejemplo, un ARNg de Cas9 puede comprender GN₁₇₋₂₀GG.

5 El ARN guía también comprende una segunda región que forma una estructura secundaria. En algunas realizaciones, la estructura secundaria comprende un tallo (u horquilla) y un bucle. La longitud del bucle y el tallo pueden variar. Por ejemplo, el bucle puede oscilar de aproximadamente 3 a aproximadamente 10 nucleótidos de longitud, y el tallo puede oscilar de aproximadamente 6 a aproximadamente 20 pares de bases de longitud. El tallo puede comprender una o más protuberancias de 1 a aproximadamente 10 nucleótidos. Por tanto, la longitud total de la segunda región puede oscilar de aproximadamente 16 a aproximadamente 60 nucleótidos de longitud. En realizaciones específicas, el bucle tiene aproximadamente 4 nucleótidos de longitud y el tallo comprende aproximadamente 12 pares de bases.

10 El ARN guía también comprende una tercera región en el extremo 3' que permanece esencialmente monocatenaria. Por tanto, la tercera región no tiene complementariedad con ninguna secuencia de ácido nucleico en la célula de interés y no tiene complementariedad con el resto del ARN guía. La longitud de la tercera región puede variar. En general, la tercera región tiene más de aproximadamente 4 nucleótidos de longitud. Por ejemplo, la longitud de la tercera región puede oscilar de aproximadamente 5 a aproximadamente 30 nucleótidos de longitud.

15 En algunas realizaciones, el ARN guía comprende una sola molécula. En otras realizaciones, el ARN guía puede comprender dos moléculas separadas. La primera molécula de ARN (también llamada ARNcr) puede comprender la primera región del ARN guía y la mitad del "tallo" de la segunda región del ARN guía. La segunda molécula de ARN (también llamada ARNtracr) puede comprender la otra mitad del "tallo" de la segunda región del ARN guía y la tercera región del ARN guía. Por tanto, en esta realización, la primera y segunda moléculas de ARN contienen cada una una secuencia de nucleótidos que son complementarias entre sí. Por ejemplo, en una realización, la primera y segunda moléculas de ARN comprenden cada una una secuencia (de aproximadamente 6 a aproximadamente 20 nucleótidos) que se empareja por bases con la otra secuencia.

20 El ARN guía puede comprender ribonucleótidos estándares o análogos de ribonucleótidos. Un análogo de nucleótido hace referencia a análogos conocidos de nucleótidos naturales, así como a nucleótidos que se modifican en los restos de base, azúcar y/o fosfato. En algunas realizaciones, el ARN guía puede comprender nucleótidos de 2'-O-metilo.

25 Los ARN guía pueden sintetizarse químicamente usando procedimientos de síntesis de oligonucleótidos estándares. Por ejemplo, los ARNcr pueden sintetizarse químicamente. Como alternativa, los ARN guía pueden sintetizarse a través de transcripción *in vitro*. Para esto, el ADN que codifica el ARN guía puede estar ligado operativamente a una secuencia promotora (p.ej., promotor T7, T3 o SP6) y se puede producir el ARN guía *in vitro* por una polimerasa (p.ej., ARN polimerasa T7, T3 o SP6). En algunas realizaciones, los ARNcr pueden sintetizarse químicamente y los ARNtracr pueden transcribirse *in vitro*. Como alternativa, los ARN guía pueden proporcionarse como moléculas de ADN codificantes para la expresión en la célula eucariótica de interés. Por ejemplo, el ADN que codifica el ARN guía puede estar ligado operativamente a una secuencia promotora que es reconocida por la ARN polimerasa III (Pol III). Los ejemplos de promotores de Pol III adecuados incluyen, pero sin limitación, los promotores de ARN de mamífero U6, U3, H1 y 7SL.

(ii) *Oligonucleótidos de detección de proximidad*

35 La sonda que contiene CRISPR/Cas del complejo de sonda de detección de proximidad también comprende un oligonucleótido de detección de proximidad. El oligonucleótido de detección de proximidad puede estar ligado directa o indirectamente al sistema CRISPR/Cas de la sonda.

40 En algunas realizaciones, el oligonucleótido de detección de proximidad puede ser un ácido nucleico monocatenario. En otras realizaciones, el oligonucleótido de detección de proximidad puede comprender regiones monocatenarias y bicatenarias. Es decir, el oligonucleótido de detección de proximidad puede comprender tallos, bucles y/o regiones de horquilla.

45 El oligonucleótido de detección de proximidad puede comprender desoxirribonucleótidos, ribonucleótidos o combinaciones de los mismos. Los desoxirribonucleótidos/ribonucleótidos pueden ser nucleótidos estándares o análogos de nucleótidos. Un análogo de nucleótido hace referencia a análogos conocidos de nucleótidos naturales (p. ej., inosina), así como a nucleótidos que se modifican en los restos de base, azúcar y/o fosfato (p.ej., fosforotioato). En algunas realizaciones, el extremo 3' del oligonucleótido de detección de proximidad puede comprender uno o más nucleótidos de bloqueo. Ejemplos no limitantes de nucleótidos de bloqueo adecuados incluyen 2'-O-metilnucleótidos, 2'-fluoronucleótidos, 3'-ONH₂-nucleótidos, didesoxinucleótidos y nucleótidos de propino. En realizaciones específicas, el extremo 3' terminal del oligonucleótido de detección de proximidad puede comprender tres 2'-O-metilribonucleótidos.

50 La longitud del oligonucleótido de detección de proximidad puede variar y variará. En general, el oligonucleótido de detección de proximidad puede oscilar de aproximadamente 15 nucleótidos a aproximadamente 200 nucleótidos de longitud. En algunas realizaciones, el oligonucleótido de detección de proximidad puede oscilar de aproximadamente 20 nucleótidos a aproximadamente 100 nucleótidos de longitud. En otras realizaciones, el oligonucleótido de detección de proximidad puede oscilar de aproximadamente 30 nucleótidos a aproximadamente 60 nucleótidos de longitud.

55 El oligonucleótido de detección de proximidad se puede obtener comercialmente. Como alternativa, el oligonucleótido de detección de proximidad puede sintetizarse usando procedimientos de síntesis de oligonucleótidos estándares.

En algunas realizaciones, el oligonucleótido de detección de proximidad puede estar ligado directamente a la proteína CRISPR/Cas del sistema CRISPR/Cas de la sonda (véase la FIG. 1). El ligamiento puede ser a través de enlaces covalentes o no covalentes. Por ejemplo, el oligonucleótido de detección de proximidad puede estar ligado a la proteína CRISPR-Cas mediante un enlace covalente. El enlace puede ser directo o a través de un ligador o una molécula adaptadora. Las técnicas para ligar covalentemente un oligonucleótido a una proteína son bien conocidas en la materia. Como alternativa, el oligonucleótido de detección de proximidad puede estar ligado a la proteína CRISPR-Cas mediante un enlace no covalente. Por ejemplo, el oligonucleótido de detección de proximidad podría unirse a la proteína CRISPR/Cas a través de enlaces de hidrógeno o iónicos.

En otras realizaciones, el oligonucleótido de detección de proximidad puede estar ligado directamente al ARN guía del sistema CRISPR/Cas de la sonda (véase la FIG. 4). El ligamiento puede ser a través de enlaces covalentes o no covalentes. Por ejemplo, el oligonucleótido de detección de proximidad puede estar ligado al extremo 3', al extremo 5' o a un nucleótido interno del ARN guía mediante un enlace covalente. Como alternativa, el oligonucleótido de detección de proximidad puede estar ligado al ARN guía mediante un enlace no covalente. Por ejemplo, el oligonucleótido de detección de proximidad podría emparejarse por bases con una región en el extremo 3' del ARN guía.

En realizaciones adicionales, el oligonucleótido de detección de proximidad puede estar ligado indirectamente al sistema CRISPR/Cas de la sonda. Por ejemplo, el oligonucleótido de detección de proximidad puede estar ligado a un anticuerpo primario dirigido contra la proteína CRISPR/Cas (véase la FIG. 3, sonda en el sitio genómico 1). El anticuerpo anti proteína CRISPR/Cas puede ser un anticuerpo policlonal o un anticuerpo monoclonal. El ligamiento entre el oligonucleótido de detección de proximidad y el anticuerpo puede ser a través de enlaces covalentes o no covalentes. En algunas realizaciones, el oligonucleótido de detección de proximidad puede estar ligado al anticuerpo mediante un enlace covalente. El ligamiento puede ser directo o a través de un ligador o una molécula adaptadora. Las técnicas para ligar o conjugar oligonucleótidos con proteínas o anticuerpos son bien conocidas en la materia. En otras realizaciones, el oligonucleótido de detección de proximidad puede estar ligado al anticuerpo anti-CRISPR/Cas mediante un enlace no covalente. Por ejemplo, el oligonucleótido de detección de proximidad podría unirse al anticuerpo a través de enlaces de hidrógeno o iónicos.

En todavía otras realizaciones en las que el oligonucleótido de detección de proximidad está ligado indirectamente al sistema CRISPR/Cas, los primeros (y segundos) oligonucleótidos de detección de proximidad pueden ligarse a anticuerpos secundarios antiespecie (y el complejo de la sonda de detección de proximidad comprende además un anticuerpo primario dirigido contra la proteína CRISPR/Cas) (véanse las Figs. 5 y 6). El ligamiento entre el oligonucleótido de detección de proximidad y el anticuerpo puede ser mediante enlaces covalentes o no covalentes (como se detalla anteriormente). Los anticuerpos secundarios antiespecie se dirigen contra la especie (p.ej., conejo, ratón, etc.) en la que se generan los anticuerpos anti proteína CRISPR/Cas. En los casos en que el complejo comprende más de un oligonucleótido de detección de proximidad ligado a anticuerpos secundarios antiespecie, las secuencias y/o estructuras de los más de un oligonucleótido de detección de proximidad difieren. En realizaciones específicas, el complejo comprende los primero y segundo oligonucleótidos de detección de proximidad ligados a anticuerpos secundarios antiespecie, en los que la secuencia y/o estructura del primer oligonucleótido de detección de proximidad difiere de la del segundo oligonucleótido de detección de proximidad.

(iii) Primeras sondas específicas que comprenden proteína similar a CRISPR/Cas

En algunas realizaciones, la primera sonda del complejo de sonda de detección de proximidad comprende el sistema CRISPR/Cas, y un primer oligonucleótido que está ligado a la proteína CRISPR/Cas o al ARN guía del sistema CRISPR/Cas mediante enlaces covalentes o no covalentes. En realizaciones alternativas, la primera sonda del complejo de sonda de detección de proximidad comprende el sistema CRISPR/Cas, y un primer oligonucleótido que está ligado a un anticuerpo dirigido contra la proteína CRISPR/Cas. En todavía otras realizaciones, la primera sonda del complejo de sonda de detección de proximidad comprende el sistema CRISPR/Cas, y al menos un oligonucleótido que está ligado a anticuerpos secundarios antiespecie, y el complejo además comprende anticuerpos primarios dirigidos contra la proteína CRISPR/Cas. En algunas realizaciones, la proteína CRISPR/Cas del sistema CRISPR/Cas es una proteína Cas9. En diversas realizaciones, la proteína Cas9 tiene actividad nucleasa, tiene actividad nicasa o se modifica para carecer de toda actividad nucleasa. En algunas realizaciones, la proteína Cas9 se modifica para carecer de toda actividad nucleasa. En otras realizaciones, la proteína Cas9 comprende al menos una señal de localización nuclear.

(b) Sondas adicionales

El complejo de sonda de detección de proximidad comprende además una o más sondas adicionales. La segunda sonda (o adicional) comprende un resto de unión, que puede ser un sistema CRISPR/Cas que comprende un ARN guía diferente y un segundo (o adicional) oligonucleótido de detección de proximidad adicional cuya secuencia y/o estructura difiere de la del primer oligonucleótido de detección de proximidad. Por ejemplo, el primer o segundo oligonucleótido de sonda de detección de proximidad puede comprender uno o más nucleótidos de bloqueo en el extremo 3' terminal del oligonucleótido. Como alternativa, el primer o segundo oligonucleótidos de sonda de detección de proximidad pueden comprender diferentes secuencias y/o diferentes estructuras de horquilla. En algunos casos, el primer y segundo oligonucleótidos de detección de proximidad se pueden llamar oligonucleótidos (+) y (-).

En realizaciones en las que el primer oligonucleótido de detección de proximidad de la primera sonda está ligado directamente al sistema CRISPR/Cas (concretamente, a través de la proteína CRISPR/Cas o el ARN guía), el complejo de sonda de detección de proximidad puede comprender además una segunda sonda. La segunda sonda comprende un resto de unión y un segundo oligonucleótido de detección proximal, en el que el segundo oligonucleótido de detección proximal tiene una secuencia y/o estructura diferente de la del primer oligonucleótido de detección proximal de la primera sonda. En algunas realizaciones, el resto de unión de la segunda sonda puede ser otro sistema CRISPR/Cas, y el segundo oligonucleótido de detección de proximidad puede estar ligado a la proteína CRISPR/Cas o al ARN guía del segundo sistema CRISPR/Cas de la segunda sonda mediante enlaces covalentes o no covalentes (véanse la FIG. 1 y la FIG. 4) En otras realizaciones, el resto de unión de la segunda sonda puede ser un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico, y el segundo oligonucleótido de detección de proximidad puede estar ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico (véase la FIG. 2). La proteína asociada a ácido nucleico puede ser un factor de transcripción general, un factor de transcripción específico, una proteína reguladora de la transcripción, una proteína de unión a cromatina, una proteína o enzima de remodelación de cromatina, una enzima de modificación de cromatina (por ejemplo, metiltransferasa, acetiltransferasa, etc.), una proteína de unión a ADN, una proteína histona, una proteína histona modificada, una proteína/factor de corte y empalme, una proteína de modificación de ARN, una proteína de procesamiento de ARN, una proteína RISC, una proteína de unión a ARN, un factor de procesamiento de ARN no codificante, y similares. La modificación de ácido nucleico puede ser un desoxirribonucleótido o ribonucleótido modificado por metilación, hidroxilación, acetilación, formilación, acilación, carboxilación, tiolación, alquilación, aminación, esterificación, fosforilación o combinaciones de los mismos.

En realizaciones en las que el primer oligonucleótido de detección de proximidad está ligado indirectamente al sistema CRISPR/Cas a través de un anticuerpo anti-CRISPR/Cas, el complejo de sonda de detección de proximidad puede comprender además una segunda sonda. La segunda sonda comprende un resto de unión y un segundo oligonucleótido de detección proximal, en el que la secuencia y/o estructura del segundo oligonucleótido de detección proximal difiere de la del primer oligonucleótido de detección proximal de la primera sonda. En algunas realizaciones, el resto de unión de la segunda sonda puede ser un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico (como se detalla anteriormente), y el segundo oligonucleótido de detección de proximidad puede estar ligado a un anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico (véase la FIG. 3, sitio genómico 2).

En otras realizaciones en las que el primer oligonucleótido de detección de proximidad está ligado indirectamente al sistema CRISPR/Cas a través de anticuerpos secundarios antiespecie, el complejo de la sonda de detección de proximidad puede comprender además un segundo oligonucleótido de proximidad ligado a los anticuerpos antiespecie (véase la FIG. 5), en los que las secuencias y/o estructuras del primer y segundo oligonucleótidos de detección de proximidad difieren. En realizaciones adicionales en las que las sondas de detección de proximidad están ligadas a anticuerpos antiespecie, el complejo puede comprender además uno o más sistemas CRISPR/Cas, que pueden estar indirectamente ligados a oligonucleótidos de detección de proximidad a través de los anticuerpos antiespecie (véase la FIG. 6). Cada sistema CRISPR/Cas comprende un ARN guía diferente.

La Tabla A enumera diversas combinaciones de primera y segunda sondas.

TABLA A		
Complejos de ejemplo	Primera sonda	Segunda sonda
1	Sistema Cas con 1º oligo ligado a la proteína Cas	Sistema Cas con 2º oligo ligado a la proteína Cas
2	Sistema Cas con 1º oligo ligado a la proteína Cas	Sistema Cas con 2º oligo ligado a ARNg
3	Sistema Cas con 1º oligo ligado a la proteína Cas	Ab anticromosómico con 2º oligo ligado a Ab anticromosómico
4	Sistema Cas con 1º oligo ligado a ARNg	Sistema Cas con 2º oligo ligado a la proteína Cas
5	Sistema Cas con 1º oligo ligado a ARNg	Sistema Cas con 2º oligo ligado a ARNg
6	Sistema Cas con 1º oligo ligado a ARNg	Ab anticromosómico con 2º oligo ligado a Ab anticromosómico
7	Sistema Cas con 1º oligo ligado a Ab anti-Cas	Ab anticromosómico con 2º oligo ligado a Ab anticromosómico
8	Sistema Cas con 1º oligo ligado a Ab antiespecie (el complejo comprende Ab anti-Cas)	2º oligo ligado a Ab antiespecie
9	Sistema Cas con 1º y 2º oligos ligados a Ab antiespecie (el complejo comprende Ab anti-Cas)	Sistemas Cas adicionales

II. Métodos para detectar ácidos nucleicos endógenos

Otro aspecto de la presente divulgación abarca un método para detectar y visualizar un ácido nucleico endógeno *in situ* en una célula. El método comprende poner en contacto la célula con un complejo de sonda de detección de proximidad que comprende al menos una sonda que comprende una proteína de unión a ácido nucleico guiada por ARN, en el que la proteína de unión a ácido nucleico guiada por ARN es guiada por ARN al ácido nucleico endógeno para la unión, formando así un complejo de sonda de detección de proximidad unido y visualizando el complejo de sonda de detección de proximidad unido a través de una reacción de amplificación dependiente de proximidad *in situ* para detectar el ácido nucleico endógeno.

En general, el complejo de sonda de detección de proximidad comprende una primera sonda que comprende un sistema CRISPR/Cas dirigido a un primer sitio en el ácido nucleico endógeno y un primer oligonucleótido de detección de proximidad que está ligado directa o indirectamente al sistema CRISPR/Cas. En algunas realizaciones, la proteína CRISPR/Cas del sistema CRISPR/Cas es una proteína Cas9. En diversas realizaciones, la proteína Cas9 tiene actividad nucleasa, tiene actividad nicasa o se modifica para carecer de toda actividad nucleasa. En algunas realizaciones, la proteína Cas9 se modifica para carecer de toda actividad nucleasa. En otras realizaciones, la proteína Cas9 comprende al menos una señal de localización nuclear. El complejo comprende además una segunda sonda que comprende un resto de unión (que puede ser otro sistema CRISPR/Cas dirigido a un segundo sitio proximal en el ácido nucleico endógeno o un anticuerpo dirigido a una proteína o una modificación de ácido nucleico localizada proximalmente) y un segundo oligonucleótido de detección de proximidad que difiere del primer oligonucleótido de detección de proximidad y está ligado directa o indirectamente al resto de unión. Tras la unión, la primera y segunda sondas del complejo de sonda de detección de proximidad se localizan proximalmente y se pueden detectar y visualizar a través de una reacción de amplificación dependiente de proximidad *in situ*. La reacción de amplificación dependiente de proximidad *in situ* puede ser un ensayo de ligadura de proximidad (PLA, véase Söderberg, et al., Nature Methods, 2006, 2(12): 995-1000) o un inicio de reacción en cadena de hibridación dependiente de proximidad (proxHCR, véase Koos et al., Nature Communications, 2015, 6: 7294|DOI: 10.1038/ncomms8294).

(a) Puesta en contacto de la célula con un complejo de sonda de detección de proximidad

El método comprende poner en contacto la célula con un complejo de sonda de detección de proximidad que comprende al menos una sonda que comprende un sistema CRISPR/Cas. Los complejos de sonda de detección de proximidad se describen anteriormente en la sección (I).

En algunas realizaciones, la célula puede ponerse en contacto con un complejo de sonda de detección de proximidad que comprende (a) una primera sonda que comprende un primer sistema CRISPR/Cas orientado a un primer sitio en el ácido nucleico endógeno y un primer oligonucleótido de detección de proximidad que está ligado directamente a la proteína CRISPR/Cas o al ARN guía del primer sistema CRISPR/Cas y (b) una segunda sonda que comprende un segundo sistema CRISPR/Cas orientado a un segundo sitio en el ácido nucleico endógeno y un segundo oligonucleótido de detección de proximidad que está ligado directamente a la proteína CRISPR/Cas o al ARN guía del segundo sistema CRISPR/Cas (véanse los complejos 1, 2, 4 y 5 en la Tabla A y las FIG. 1 y 4).

En otras realizaciones, la célula puede ponerse en contacto con un complejo de sonda de detección de proximidad que comprende (a) una primera sonda que comprende un sistema CRISPR/Cas orientado a un primer sitio en el ácido nucleico endógeno y un primer oligonucleótido de detección de proximidad que está ligado directamente a la proteína CRISPR/Cas o al ARN guía del primer sistema CRISPR/Cas y (b) una segunda sonda que comprende un anticuerpo dirigido contra una proteína localizada proximalmente asociada a ácido nucleico endógeno o un anticuerpo dirigido contra una modificación de ácido nucleico en el ácido nucleico endógeno y un segundo oligonucleótido de detección de proximidad que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o al anticuerpo dirigido contra la modificación de ácido nucleico (véanse los complejos 3 y 6 en la Tabla A y la FIG. 2).

En otras realizaciones, la célula puede ponerse en contacto con un complejo de sonda de detección de proximidad que comprende (a) una primera sonda que comprende un sistema CRISPR/Cas orientado a un primer sitio en el ácido nucleico endógeno y un primer oligonucleótido de detección de proximidad que está ligado a un anticuerpo primario dirigido contra la proteína CRISPR/Cas y (b) una segunda sonda que comprende un anticuerpo dirigido contra una proteína localizada proximalmente asociada a el ácido nucleico endógeno o un anticuerpo dirigido contra una modificación de ácido nucleico en el ácido nucleico endógeno y un segundo oligonucleótido de detección de proximidad que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o al anticuerpo dirigido contra la modificación de ácido nucleico (véanse el complejo 7 en la Tabla A y la FIG. 3).

En realizaciones adicionales, la célula puede ponerse en contacto con un complejo de sonda de detección de proximidad que comprende una sonda que comprende un sistema CRISPR/Cas orientado a un primer sitio en el ácido nucleico endógeno, y el primero y segundo oligonucleótidos de detección de proximidad ligados a anticuerpos secundarios antiespecie, en el que el complejo de sonda de detección de proximidad comprende además anticuerpos primarios dirigidos contra la proteína CRISPR/Cas (véanse el complejo 8 en la Tabla A y la FIG. 5). En algunas iteraciones de esta realización, la célula puede ponerse en contacto con un segundo o adicionales sistemas

CRISPR/Cas dirigidos a un segundo o adicionales sitios en el ácido nucleico endógeno, en las que los sistemas CRISPR/Cas pueden estar indirectamente ligados al primer y segundo oligonucleótidos de detección de proximidad a través de los anticuerpos primarios antiproteína CRISPR/Cas y los anticuerpos secundarios antiespecie (véanse el complejo 9 en la Tabla A y la FIG. 6).

5 La puesta en contacto de la célula con el complejo de sonda de detección de proximidad puede ocurrir en una o varias etapas. Por ejemplo, la célula se puede poner en contacto con la primera sonda y luego la célula se puede poner en contacto con la segunda sonda en un momento posterior. De manera similar, la célula puede ponerse en contacto con una primera sonda que comprende un sistema CRISPR/Cas y luego ponerse en contacto con anticuerpos dirigidos contra la proteína CRISPR/Cas y/o anticuerpos dirigidos contra proteínas asociadas a ácidos nucleicos o modificaciones de ácidos nucleicos. En algunas realizaciones, la célula puede ponerse en contacto con una primera sonda que comprende un sistema CRISPR/Cas, luego ponerse en contacto con anticuerpos dirigidos contra la proteína CRISPR/Cas, y luego ponerse en contacto con anticuerpos secundarios antiespecie.

10 Como se detalla a continuación, la célula que se pone en contacto con el complejo de sonda de detección de proximidad puede estar viva, fijada o congelada. En general, la puesta en contacto puede realizarse a una temperatura que oscila de aproximadamente 20 °C a aproximadamente 40 °C.

15 En alguna realización, una célula viva puede ponerse en contacto con al menos una sonda que contiene el sistema CRISPR/Cas. Las sondas que contienen el sistema CRISPR/Cas pueden introducirse en la célula por diversos medios. Los medios de suministro adecuados incluyen microinyección, electroporación, sonoporación, biolística, transfección mediada por fosfato de calcio, transfección catiónica, transfección por liposomas, transfección por dendrímero, transfección por choque térmico, transfección por nucleofección, magnetofección, lipofección, impalefección, transfección óptica, captación de ácidos nucleicos potenciada por agente patentado y suministro a través de liposomas, inmunoliposomas, virosomas o viriones artificiales. En realizaciones específicas, las sondas que contienen el sistema CRISPR/Cas pueden introducirse en la célula mediante nucleofección. La sonda que contiene el sistema CRISPR/Cas se puede introducir como un complejo de ribonucleoproteína (RNP) que comprende una proteína CRISPR/Cas recombinante (p.ej., Cas9 o dCas9) complejada con un ARN guía (p.ej., ARNcr y ARNtracr sintetizados) o la proteína CRISPR/Cas y el ARN guía pueden introducirse en la célula individualmente. Como alternativa, el sistema CRISPR/Cas puede ensamblarse en la célula después de introducir en la célula el ácido nucleico (concretamente, ARNm o ADN) que codifica la proteína CRISPR/Cas y el ARN guía o de introducir en la célula el ácido nucleico (concretamente, ARNm o ADN) que codifica la proteína CRISPR/Cas y el ADN que codifica el ARN guía. Después de un período de tiempo adecuado, la célula se puede fijar (véase a continuación) y ponerse en contacto con (i) anticuerpos dirigidos contra la proteína CRISPR/Cas y/o anticuerpos dirigidos contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico o (ii) anticuerpos dirigidos contra la proteína CRISPR/Cas seguidos por anticuerpos secundarios antiespecie. La celda se puede lavar con tampones apropiados entre cualquiera de las etapas de puesta en contacto.

20 En otras realizaciones, una célula fijada puede ponerse en contacto con la al menos una sonda que contiene el sistema CRISPR/Cas. La celda se puede fijar usando cualquiera de una variedad de fijadores comúnmente usados. Los fijadores adecuados incluyen paraformaldehído, formaldehído, metanol, acetona, ácido acético, etanol, glutaraldehído, yodoformo, ácido láctico, ácido pícrico, cinc o combinaciones de los mismos. La concentración de fijador y la duración del proceso de fijación variarán dependiendo del tipo de célula o muestra. En una realización, la célula se puede fijar con una mezcla de metanol y ácido acético (1:1). En otra realización, la célula se puede fijar con paraformaldehído al 4 %. En general, la sonda que contiene el sistema CRISPR/Cas se introduce en la célula como un complejo de RNP (véase anteriormente).

25 En algunas realizaciones, la célula puede permeabilizarse por incubación con una solución que comprende al menos un tensioactivo y/o proteasa. Los ejemplos no limitantes de tensioactivos adecuados incluyen Tween-20, Tween-80, Triton X-100, alcohol cetílico, decilglucósido, digitonina, laurilglucósido, IGEPAL CA-630, leucoperm, NP-40, nonoxinol-9, octaetilenglicolmonododeciléter, *n*-octil- β -D-tioglucopiranosido, alcohol oleílico, octilglucósido, polisorbato 20, polisorbato 80, saponina, alcohol estearílico, o combinaciones de los mismos. Las proteasas adecuadas incluyen, sin límite, proteinasa K, caspasa, quimotripsina, papaína, pepsina y tripsina. En algunas realizaciones, la célula puede incubarse con una solución que comprende Tween-20 y/o Triton X-100. La concentración del tensioactivo o proteasa y la duración del período de incubación pueden variar y variarán dependiendo del tipo de célula o muestra.

30 En general, la célula no se somete a un proceso de desnaturalización química o térmica para convertir el ADN cromosómico bicatenario en ADN monocatenario. Sin embargo, en algunas realizaciones, la célula puede ponerse en contacto con una solución desnaturalizante para desnaturalizar el ADN cromosómico. La solución desnaturalizante puede ser ácida o alcalina. Una solución ácida comprende un ácido tal como ácido hidrocórico, y una solución alcalina comprende una base tal como un hidróxido de metal alcalino (p. ej., hidróxido de sodio o potasio). La concentración del ácido o la base en la solución desnaturalizante y la duración de la etapa de desnaturalización pueden variar y variarán dependiendo del tipo de célula o muestra. En otras realizaciones, la célula puede calentarse para desnaturalizar el ADN cromosómico. Por ejemplo, la celda se puede calentar a una temperatura de aproximadamente 70 °C a aproximadamente 80 °C en presencia de una solución que contiene formamida. La duración de la etapa de calentamiento puede variar y variará según el tipo de célula o muestra.

En realizaciones en las que el ácido nucleico endógeno es ARN, la puesta en contacto con la al menos una sonda que contiene el sistema CRISPR/Cas puede realizarse en presencia de un oligonucleótido presentador de PAM (concretamente, un PAMmero, véase O'Connell et al., Nature, 2014, 516: 263-266) Los PAMmeros pueden presentarse en *trans* al ARN guía. Los PAMmeros pueden comprender desoxirribonucleótidos estándares o modificados, ribonucleótidos o combinaciones de los mismos.

Después de la puesta en contacto con la primera y segunda sondas del complejo de sonda de detección de proximidad, la célula comprende al menos un complejo de sonda de detección de proximidad unido.

(b) Visualización del complejo de sonda de detección de proximidad unido

El método comprende además visualizar el complejo de sonda de detección de proximidad unido a través de una reacción de amplificación dependiente de proximidad *in situ* para detectar el ácido nucleico endógeno. La reacción de amplificación dependiente de proximidad *in situ* puede ser un ensayo de ligadura de proximidad (PLA), un inicio de reacción en cadena de hibridación dependiente de la proximidad (proxHCR) u otro método de amplificación que genera un producto insoluble o un producto que está enlazado al complejo de sonda de detección de proximidad.

PLA. En realizaciones en las que la reacción de amplificación dependiente de proximidad *in situ* comprende PLA, el método comprende además poner en contacto la célula con uno o más oligonucleótidos conectores (véase, Söderberg, et al., Nature Methods, 2006, 2(12): 995-1000). El uno o más oligonucleótidos conectores son generalmente ácidos nucleicos monocatenarios que tienen complementariedad de secuencia con el primer y segundo oligonucleótidos de detección de proximidad del complejo de sonda de detección de proximidad, y uno de los oligonucleótidos conectores comprende una secuencia única que no está presente en ácidos nucleicos endógenos de la célula. Los oligonucleótidos conectores pueden comprender desoxirribonucleótidos o ribonucleótidos estándares o modificados o combinaciones de los mismos. Los oligonucleótidos conectores pueden oscilar de longitud de aproximadamente 15 nucleótidos a aproximadamente 150 nucleótidos. En algunas realizaciones, los oligonucleótidos conectores pueden oscilar de longitud de aproximadamente 20 nucleótidos a aproximadamente 80 nucleótidos. En una realización específica, la célula se pone en contacto con dos desoxioligonucleótidos conectores monocatenarios. Al entrar en contacto con el oligonucleótido u oligonucleótidos conectores, el oligonucleótido u oligonucleótidos conectores se emparejan por bases o hibridan con el primer y segundo oligonucleótidos de detección de proximidad que están muy cercanos.

El método comprende además poner en contacto la célula con una ligasa, que liga el oligonucleótido u oligonucleótidos conectores hibridados para formar un producto de ligadura circularizado que comprende la secuencia única. La ligasa puede ser ADN ligasa de T4, ARN ligasa de T4 o ADN/ARN ligasa de App. La ligasa puede ser mesofílica o termoestable. En una realización específica, la ligasa es ADN ligasa de T4. La reacción de ligadura se lleva a cabo en presencia de un tampón de ligasa adecuado, y a una temperatura que oscila de aproximadamente 20 °C a aproximadamente 40 °C.

El método comprende además amplificar el producto de ligadura circularizado que comprende la secuencia única mediante amplificación de círculo rodante para formar un producto de amplificación de círculo rodante que comprende repeticiones de la secuencia única. Por tanto, el método comprende poner en contacto la célula con una polimerasa para replicación de círculo rodante. La polimerasa puede ser ADN polimerasa de phi29, ADN polimerasa de bst, ADN polimerasa de Taq, ADN polimerasa de T4 o ADN polimerasa de T7. La polimerasa puede ser mesofílica o termoestable. En realizaciones específicas, la polimerasa es la ADN polimerasa de phi29. La reacción de amplificación se lleva a cabo en presencia de una solución de amplificación que comprende un tampón de polimerasa adecuado, dNTP y otros reactivos. La reacción de amplificación se lleva a cabo a una temperatura que oscila de aproximadamente 20 °C a aproximadamente 40 °C.

El método comprende además poner en contacto la célula con oligonucleótidos marcados con fluorescencia que tienen complementariedad de secuencia con la secuencia única en el oligonucleótido conector. Por tanto, los oligonucleótidos marcados hibridan con la secuencia única repetida dentro del producto de amplificación de círculo rodante. El marcaje fluorescente del oligonucleótido marcado puede fluorecer en verde (p.ej., estar marcado con Cy2, Alexa 488 o fluoresceína y sus tales derivados como FAM, HEX, TET y TRITC), naranja (p.ej., estar marcado con Cy3 o Alexa 546); rojo (p.ej., estar marcado con Texas Red, Alexa 594 o rodamina y sus derivados tales como ROX), o rojo lejano (p.ej., estar marcado con Cy5). Los oligonucleótidos marcados pueden oscilar de longitud de aproximadamente 15 nucleótidos a aproximadamente 30 nucleótidos. En realizaciones específicas, los oligonucleótidos marcados pueden tener aproximadamente 18-22 nucleótidos de longitud. Los oligonucleótidos marcados con fluorescencia pueden ser parte de la solución de amplificación descrita anteriormente.

La etapa final de visualización del complejo de sonda de detección de proximidad unido comprende detectar la fluorescencia de los oligonucleótidos marcados con fluorescencia hibridados con el producto de replicación de círculo rodante usando microscopía de fluorescencia estándar y programas de software de análisis de imágenes. Un punto fluorescente discreto indica la localización del ácido nucleico endógeno. La célula puede comprender uno o más puntos fluorescentes discretos, dependiendo de la identidad y localización del ácido nucleico endógeno de interés.

proxHCR. En realizaciones en las que la reacción de amplificación dependiente de proximidad *in situ* comprende

proxHCR, el método comprende además poner en contacto la célula con uno o más oligonucleótidos adicionales (véase, Koos et al., Nature Communications, 2015, 6:7294|DOI: 10.1038/ncomms8294) Por ejemplo, la célula puede ponerse en contacto con un oligonucleótido activador, al menos un oligonucleótido de horquilla marcado con fluorescencia en 5', al menos un oligonucleótido de horquilla marcado con fluorescencia en 3' y un oligonucleótido iniciador. Los oligonucleótidos activador y en horquilla generalmente comprenden regiones monocatenarias, así como tallos, bucles, horquilla u otras estructuras secundarias. El oligonucleótido iniciador es generalmente monocatenario. Los oligonucleótidos pueden oscilar de longitud de aproximadamente 20 a aproximadamente 100 nucleótidos, y pueden comprender desoxirribonucleótidos o ribonucleótidos estándares o modificados o combinaciones de los mismos. El marcaje fluorescente puede ser como se describió anteriormente en la sección de PLA. El producto de amplificación marcado con fluorescencia se puede visualizar como se detalla anteriormente en la sección de PLA.

(c) Ácidos nucleicos endógenos

En algunas realizaciones, el ácido nucleico endógeno detectado por el método divulgado en la presente memoria puede ser ADN cromosómico nuclear. Por ejemplo, se pueden detectar loci cromosómicos o genómicos específicos en el ADN nuclear. El locus cromosómico o genómico específico puede estar dentro de la región de codificación de un gen, un intrón de un gen, una región de control de un gen, una isla de CpG, una región espaciadora entre genes, una región no codificante, una región centromérica, una región telomérica, una región que comprende repeticiones de trinucleótidos, y similares. El gen puede ser un gen codificante de proteína o un gen codificante de ARN. En algunas realizaciones, el locus cromosómico o genómico puede ser un cambio o modificación, por ejemplo, una delección, una inserción, una sustitución (p.ej., un SNP), una transversión, y similares, en un gen asociado a una enfermedad o trastorno. En general, el complejo de sonda de detección de proximidad se orienta a dos sitios distintos en el locus cromosómico o genómico que pueden estar separados por un máximo de aproximadamente 2 kb en *cis* en el mismo cromosoma

En otras realizaciones, el ácido nucleico endógeno detectado por el método divulgado en la presente memoria puede ser una molécula de ARN. La molécula de ARN puede ser un ARN mensajero (ARNm) o un fragmento del mismo. El ARNm puede estar poliadenilado o no poliadenilado. Como alternativa, la molécula de ARN puede ser un ARN no codificante (ARNnc). Por ejemplo, el ARNnc puede ser ARN largo no codificante (ARNlnc), ARN largo intergénico no codificante (ARNlinc), microARN (miARN), ARN interferente pequeño (ARNip), ARN que interactúa con Piwi (ARNpi), ARN de acción en trans (ARNrasi), ARN ribosómico (ARNr), ARN de transferencia (ARNt), ARNt mitocondrial (ARNt-MT), ARN nuclear pequeño (ARNnp), ARN nucleolar pequeño (ARNnup), ARN SmY, ARN Y, ARN líder cortado y empalmado (ARN SL), componente ARN de telomerasa (TERC), fragmentos de los mismos o combinaciones de los mismos.

En más realizaciones adicionales, el ácido nucleico endógeno detectado por el método divulgado en la presente memoria puede localizarse en genomas mitocondriales o plástidos.

En algunas realizaciones, el método para detectar ácidos nucleicos endógenos específicos se puede usar con fines de investigación. En otras realizaciones, los métodos para detectar ácidos nucleicos endógenos específicos pueden usarse con fines de diagnóstico

(d) Células

Se puede usar una variedad de células en el método divulgado en la presente memoria. En general, la célula es una célula eucariótica. En diversos aspectos, la célula puede ser una célula humana, una célula de mamífero no humano, una célula de vertebrado no mamífero, una célula de invertebrado, una célula de insecto, una célula vegetal, una célula de levadura o un organismo eucariótico monocelular. En aspectos ejemplares, la célula es una célula de mamífero. La célula puede ser una célula primaria o una célula de estirpe celular. La célula puede ser una célula adulta, una célula embrionaria o un citoblasto. La célula puede ser una célula normal, una célula enferma o una célula cancerosa.

En algunas realizaciones, la célula puede ser una célula de estirpe celular humana. Los ejemplos no limitantes de estirpes celulares adecuadas incluyen DU145 (cáncer metastásico), SW490 (cáncer de colon), DLD-1 (cáncer de colon), KM20L2 (cáncer de colon), COLO 205 (cáncer de colon), HCC-2998, (cáncer de colon) , HCT-116 (cáncer de colon), HCT-15 (cáncer de colon), HT29 (cáncer de colon), KM12 (cáncer de colon), SW-620 (cáncer de colon), SF-268 (SNC), SF-295 (SNC) , SF-539 (SNC), SNB-19 (SNC), SNB-75 (SNC), U251 (SNC), CCRF-CEM (leucemia), HL-60(TB) (leucemia), K-562 (leucemia) , MOLT-4 (leucemia), RPMI-8226 (leucemia), SR (leucemia), A549 (cáncer de pulmón no microcítico), EKVX (cáncer de pulmón no microcítico), HOP-62 (cáncer de pulmón no microcítico) , HOP-92 (cáncer de pulmón no microcítico), NCI-H226 (cáncer de pulmón no microcítico), NCI-H23 (cáncer de pulmón no microcítico), NCI-H322M (cáncer de pulmón no microcítico), NCI-H460 (cáncer de pulmón no microcítico), NCI-H522 (cáncer de pulmón no microcítico), LOX IMVI (melanoma), MALME-3M (melanoma), M14 (melanoma), MDA-MB-435 (melanoma) , SK-MEL-2 (melanoma), SK-MEL-28 (melanoma), SK-MEL-5 U (melanoma), ACC-257 (melanoma), UACC-62 (melanoma), IGR-OV1 (ovárico), OVCAR-3 (ovárico), OVCAR-4 OVCAR-5 (ovárico), OVCAR-8 (ovárico), SK-OV-3 (ovárico), 786-0 (renal), A498 (renal), ACHN (renal), CAKI-1 (renal), RXF 393 (renal), SN12C (renal), TK-10 (renal), UO-31 (renal), PC -3 (próstata), DU-145 (próstata), MCF7 (mama), MDA-MB-231 (mama), MDA-MB-468 (mama), HS 578T (mama), BT-549 (mama) y T-47D (mama).

En otras realizaciones, la célula puede ser una célula de estirpe celular de mamífero. Los ejemplos no limitantes de estirpes celulares de mamífero adecuadas incluyen células de ovario de hámster chino (CHO), células de riñón de hámster lactante (BHK); células de mieloma de ratón NS0, células de fibroblastos embrionarios de ratón 3T3 (NIH3T3), células de linfoma B de ratón A20; células de melanoma de ratón B16; células de mioblastos de ratón C2C12; células de mieloma de ratón SP2/0; células mesenquimatosas embrionarias de ratón C3H-10T1/2; células de carcinoma de ratón CT26, células de próstata de ratón DuCuP; células de mama de ratón EMT6; células de hepatoma de ratón Hepa1c1c7; células de mieloma de ratón J5582; células epiteliales de ratón MTD-1A; células de miocardio de ratón MyEnd; células renales de ratón RenCa; células pancreáticas de ratón RIN-5F; células de melanoma de ratón X64; células de linfoma de ratón YAC-1 ; células de glioblastoma de rata 9L; células de linfoma B de rata RBL; células de neuroblastoma de rata B35; células de hepatoma de rata (HTC); células de hígado de rata búfalo BRL 3A; células de riñón canino (MDCK); células mamarias caninas (CMT); células de osteosarcoma de rata D17; células de monocitos/macrófagos de rata DH82; células de fibroblastos transformados con SV-40 de riñón de mono (COS7); células de riñón de mono CVI-76; células de riñón de mono verde africano (VERO-76) y células de riñón embrionario humano (HEK293, HEK293T). Se puede encontrar una extensa lista de estirpes celulares de mamíferos en el catálogo de la Colección Americana de Cultivos Tipo (ATCC, Manassas, VA).

En todavía otras realizaciones, la célula puede estar dentro de una muestra de tejido o muestra de fluido obtenida de un sujeto. Por ejemplo, una muestra de tejido o una muestra de fluido puede extraerse mediante resección quirúrgica, biopsia por escisión, biopsia por incisión, biopsia central o biopsia por aspiración con aguja. El sujeto puede ser un mamífero humano, no humano (p. ej., roedores, gatos, perros, animales de ganado y similares) o un vertebrado no mamífero (por ejemplo, peces, aves y demás). La muestra de tejido se puede congelar o fijar usando un fijador como se detalla anteriormente. La muestra de tejido fijada puede incluirse en un medio de inclusión tal como parafina, Paraplast o medio de inclusión similar conocido en la materia.

III. Kits

Un aspecto adicional de la presente divulgación implica kits para detectar un ácido nucleico endógeno en una célula, en el que un kit comprende al menos una sonda que comprende un sistema CRISPR/Cas como se define anteriormente en la sección I(a).

En algunas realizaciones, el kit comprende (a) una primera sonda que comprende un primer sistema CRISPR/Cas y un primer oligonucleótido de detección de proximidad que está ligado directamente a la proteína CRISPR/Cas o al ARN guía del primer sistema CRISPR/Cas y (b) una segunda sonda que comprende un segundo sistema CRISPR/Cas y un segundo oligonucleótido de detección de proximidad que está ligado directamente a la proteína CRISPR/Cas o al ARN guía del segundo sistema CRISPR/Cas.

En otras realizaciones, el kit comprende (a) una primera sonda que comprende un sistema CRISPR/Cas y un primer oligonucleótido de detección de proximidad que está ligado directamente a la proteína CRISPR/Cas o al ARN guía del primer sistema CRISPR/Cas y (b) una segunda sonda que comprende un anticuerpo dirigido contra una proteína asociada a ácido nucleico endógeno o un anticuerpo dirigido contra una modificación de ácido nucleico en el ácido nucleico endógeno y un segundo oligonucleótido de detección de proximidad que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o al anticuerpo dirigido contra la modificación del ácido nucleico.

En otras realizaciones, el kit comprende (a) una primera sonda que comprende un sistema CRISPR/Cas y un primer oligonucleótido de detección de proximidad que está ligado a un anticuerpo primario dirigido contra la proteína CRISPR/Cas y (b) una segunda sonda que comprende un anticuerpo dirigido contra una proteína asociada a ácido nucleico endógeno o un anticuerpo dirigido contra una modificación de ácido nucleico en el ácido nucleico endógeno y un segundo oligonucleótido de detección de proximidad que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o el anticuerpo dirigido contra la modificación de ácido nucleico.

En realizaciones adicionales, la célula del kit comprende al menos un sistema CRISPR/Cas, el primero y segundo oligonucleótidos de detección de proximidad ligados a anticuerpos secundarios antiespecie y anticuerpos primarios dirigidos contra la proteína CRISPR/Cas. En algunas iteraciones, el kit comprende dos sistemas CRISPR/Cas, tres sistemas CRISPR/Cas o más de tres sistemas CRISPR/Cas.

La proteína CRISPR/Cas del sistema CRISPR/Cas puede ser una proteína Cas9. En algunas realizaciones, la proteína Cas9 tiene actividad nucleasa, tiene actividad nicasa o se modifica para carecer de toda actividad nucleasa. En realizaciones específicas, la proteína Cas9 comprende al menos una señal de localización nuclear y se modifica para carecer de toda actividad nucleasa.

En algunas realizaciones, la proteína CRISPR/Cas se puede proporcionar como una proteína purificada. En otras realizaciones, la proteína CRISPR/Cas se puede proporcionar como un ácido nucleico codificante (concretamente, ARNm o ADN). El ácido nucleico que codifica la proteína CRISPR/Cas puede ser de codón optimizado para expresión en la célula de interés. El ARNm que codifica la proteína CRISPR/Cas puede estar 5' protegido y/o 3' poliadenilado. El ADN que codifica la proteína CRISPR/Cas puede estar operativamente ligado a secuencias de control promotoras (véase a continuación), señales de poliadenilación (p.ej., señal poliA de SV40, señal poliA de hormona de crecimiento bovina, etc.), y/o secuencias de terminación transcripcional. El ADN que codifica la proteína CRISPR/Cas puede ser

parte de un constructo de ADN (p.ej., vector plasmídico, vector lentivírico, vector vírico adenoasociado, fagémido, cósmido, cromosoma artificial/minicromosoma y demás). En algunas realizaciones, el ADN puede estar ligado operativamente a una secuencia de control promotora para expresión en una célula de interés. Las secuencias de control promotoras de mamífero adecuadas incluyen el promotor temprano inmediato de citomegalovirus (CMV), el promotor de virus de simio (SV40), el promotor tardío principal de adenovirus, el promotor del virus de sarcoma de Rous (RSV), el promotor del virus de tumor mamario de ratón (MMTV), el promotor de fosfoglicerato cinasa (PGK), el promotor alfa de factor de elongación (ED1) y similares. En otras realizaciones, el ADN puede estar ligado operativamente a una secuencia de control promotora que es reconocida por una ARN polimerasa de fago para la síntesis de ARNm *in vitro*. Por ejemplo, la secuencia promotora puede ser una secuencia promotora de T7, T3 o SP6 o una variación de una secuencia promotora de T7, T3 o SP6. En realizaciones adicionales, el ADN puede estar ligado operativamente a una secuencia promotora para expresión *in vitro* en células bacterianas o eucarióticas. Los promotores bacterianos adecuados incluyen, sin límite, promotores de T7, promotores del operón *lac*, promotores de *trp*, promotores de *tac* (que son híbridos de promotores de *trp* y *lac*), variaciones de cualquiera de los anteriores, y combinaciones de los mismos de cualquiera de los anteriores. Los ejemplos no limitantes de promotores eucarióticos adecuados son bien conocidos en la materia e incluyen las secuencias promotoras de mamíferos enumeradas anteriormente.

En ciertas realizaciones, el kit puede comprender además al menos un sistema de vector para la transcripción *in vitro* de uno o más ARN guía. Por ejemplo, el sistema de vector puede comprender una secuencia promotora que es reconocida por una ARN polimerasa de fago para la síntesis de ARN *in vitro*. La secuencia del promotor de fago puede ser un promotor de T7, T3 o SP6 o una variante del mismo. En otras realizaciones, el kit puede comprender al menos un sistema de vector para la expresión *in vivo* de uno o más ARN guía. El ADN que codifica el o los ARN guía puede estar ligado operativamente a una secuencia promotora reconocida por la ARN polimerasa III (Pol III). Los ejemplos de promotores de Pol III adecuados incluyen, pero sin limitación, promotores de ARN de mamífero de U6, U3, H1 y 7SL. El sistema de vector puede comprender además una región de poliligador que contiene múltiples sitios de clonación para la inserción de ADN que codifica el ARN guía de interés. El sistema de vector puede comprender además un origen o replicación, un gen marcador de selección y/o sitios cebadores de amplificación o secuenciación.

En todavía otras realizaciones, el kit puede comprender además al menos un reactivo para una reacción de amplificación dependiente de proximidad *in situ*. En algunas realizaciones, el kit puede comprender al menos un reactivo para PLA. El reactivo de PLA puede ser uno o más oligonucleótidos conectores, una ligasa, un reactivo de ligadura, una polimerasa, un reactivo de amplificación, un oligonucleótido de detección marcado con fluorescencia o combinaciones de los mismos. En otras realizaciones, el kit puede comprender activador, detector marcado y/u oligonucleótidos iniciadores para proxHCR.

Definiciones

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

Cuando se introducen elementos de la presente divulgación o aspectos preferidos de la misma, los artículos "un", "una", "el/la" y "dicho/dicha" pretenden significar que hay uno o más de los elementos. Los términos "que comprende", "que incluye" y "que tiene" pretenden ser inclusivos y significan que puede haber elementos adicionales además de los elementos enumerados.

Como se usan en la presente memoria, los términos "complementario" o "complementariedad" hacen referencia a la asociación de ácidos nucleicos bicatenarios mediante emparejamiento de bases a través de enlaces de hidrógeno específicos. El emparejamiento de bases puede ser emparejamiento de bases de Watson-Crick estándar (por ejemplo, pares 5'-A G T C-3' con la secuencia complementaria 3'-T C A G-5'). El emparejamiento de bases también puede ser enlace de hidrógeno de Hoogsteen o Hoogsteen invertido. La complementariedad se mide típicamente con respecto a una región de dúplex y, por tanto, excluye los salientes, por ejemplo. La complementariedad entre dos hebras de la región dúplex puede ser parcial y expresarse como un porcentaje (por ejemplo, 70 %), si solo algunos de los pares de bases son complementarios. Las bases que no son complementarias son "desapareadas". La complementariedad también puede ser completa (concretamente, 100%), si todos los pares de bases de la región de dúplex son complementarios.

Como se usa en la presente memoria, el término "endógeno" hace referencia a un ácido nucleico que es nativo de la célula.

Un "gen", como se usa en la presente memoria, hace referencia a una región de ADN (que incluye 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 si tales secuencias reguladoras son adyacentes o no a las secuencias de codificación y/o

transcritas. Por consiguiente, un gen incluye, pero no se limita necesariamente a, secuencias promotoras, terminadores, secuencias reguladoras traduccionales tales como sitios de unión a ribosoma y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aisladores, elementos límite, orígenes de replicación, sitios de fijación a matriz y regiones de control de locus.

5 Los términos "ácido nucleico" y "polinucleótido" hacen referencia a un polímero de desoxirribonucleótido o ribonucleótido, en conformación lineal o circular, y en forma monocatenaria o bicatenaria. Para los fines 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 se modifican en los restos de base, azúcar y/o fosfato (p. ej., esqueletos de fosforotioato). En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases; concretamente, un análogo de A se emparejará por base con T.

10 El término "nucleótido" hace referencia a desoxirribonucleótidos o ribonucleótidos. Los nucleótidos pueden ser nucleótidos estándares (concretamente, adenosina, guanosina, citidina, timidina y uridina) o análogos de nucleótidos. Un análogo de nucleótido hace referencia a un nucleótido que tiene una base de purina o pirimidina modificada o un resto de ribosa modificado. Un análogo de nucleótido puede ser un nucleótido de origen natural (p. ej., inosina) o un nucleótido de origen no natural. Los ejemplos no limitantes de modificaciones en los restos 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, también como la sustitución de los átomos de carbono y nitrógeno de las bases por otros átomos (p. ej., 7-desazapurinas). Los análogos de nucleótidos también incluyen didesoxinucleótidos, 2'-O-metilnucleótidos, ácidos nucleicos bloqueados (LNA), ácidos nucleicos peptídicos (PNA) y morfolinos.

15 Las técnicas para determinar la identidad de secuencia de aminoácidos y ácidos nucleicos son conocidas en la materia. Típicamente, tales técnicas incluyen determinar la secuencia de nucleótidos del ARNm para un gen y/o determinar la secuencia de aminoácidos codificada por el mismo, y comparar estas secuencias con una segunda secuencia de nucleótidos o aminoácidos. Las secuencias genómicas también se pueden determinar y comparar de esta manera. En general, la identidad hace referencia 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. El porcentaje de identidad de dos secuencias, ya sea de ácido nucleico o de aminoácidos, es el número de coincidencias exactas entre dos secuencias alineadas dividido por la longitud de las secuencias más cortas y multiplicado por 100. Se proporciona un alineamiento aproximado de las secuencias de ácido nucleico el algoritmo de homología local de Smith y 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 Sequence and Structure*, M. O. Dayhoff ed., 5 supl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., EE. UU. y normalizada por Gribskov, *Nucl. Acids Res.* 14(6): 6745-6763 (1986). Se proporciona una implementación ejemplar de este algoritmo para determinar el porcentaje de identidad de una secuencia por el Genetics Computer Group (Madison, Wisconsin) en la aplicación de utilidad "BestFit". Otros programas adecuados para calcular el porcentaje de identidad o similitud entre secuencias son generalmente conocidos en la materia, por ejemplo, otro programa de alineamiento es BLAST, usado con parámetros predeterminados. Por ejemplo, BLASTN y BLASTP pueden usarse usando los siguientes parámetros predeterminados: código genético= estándar; filtro= ninguno; hebra= ambas; corte= 60; esperanza= 10; matriz= BLOSUM62; descripciones= 50 secuencias; ordenar por= PUNTUACIÓN ALTA; bases de datos= no redundantes, GenBank + EMBL + DDBJ + PDB + GenBank CDS Translations + Swiss protein + Spupdate + PIR. Los detalles de estos programas se pueden encontrar en el sitio web de GenBank.

Ejemplos

45 El siguiente ejemplo ilustra ciertos aspectos de la invención.

Ejemplo 1: Detección de centrómeros usando complejos de dCAS9-ARNg

50 El ADN satélite centromérico comprende grandes conjuntos de secuencias no codificantes que se repiten en tándem, y la cromatina centromérica es única en el sentido de que la histona H3 se reemplaza por CENP-A (véase la FIG. 7). Para determinar si los centrómeros podrían detectarse utilizando los métodos divulgados en la presente memoria, se construyó una sonda que comprende un sistema Cas9 orientado a una secuencia de repetición de satélite menor o una secuencia de repetición de satélite mayor para usar en combinación con un ensayo de ligadura de proximidad (PLA) (p. ej., ensayo DUOLINK®) y anticuerpos de CENP-A.

55 La sonda CRISPR/Cas se formó como complejos de ribonucleoproteína (RNP) mediante la combinación de dCas9 recombinante (concretamente, D10A/H840A doble mutante; 1 pmol) con DIG-ARNtracr sintetizado (2 pmol) y ARNcr (2 pmol) que se orientaron a la secuencia de repetición de satélite menor, a la secuencia de repetición de satélite mayor o a una secuencia de control 1 negativa (que no se orienta a ninguna secuencia humana conocida). Las secuencias de DIG-ARNtracr y ARNcr usadas eran como se indica en la Tabla 1. Se nucleofectaron células U2OS con los RNP usando reactivos y equipo estándares. Las células nucleofectadas se diluyeron 1:10 en medio completo (DMEM + FBS al 10 % sin antibióticos) y se sembraron en portaobjetos de vidrio con cámara de 8 pocillos. Para todas

las etapas posteriores en el procedimiento, se usaron 40 µl de reactivo/pocillo. Después de 6 horas, las células se fijaron con para-formaldehído al 4 % durante 15 minutos a TA, se inactivaron con 1/10 volumen de glicina 1,25 M durante 5 minutos a TA, se lavaron con PBS, se permeabilizaron durante 1 hora en tampón de permeabilización (0,75 % de Triton X100, 0,75 % de Tween 20) y se bloquearon con tampón de bloqueo de PLA durante 1 hora a TA.

Tabla 1. Secuencias de ARNcr y ARNtracr		
	Secuencia	SEQ ID NO:
sgSatélite mayor	5'-CCAUAUCCACGUCCUACAGUGUUUU AGAGCUAUGCUGUUUUG-3'	9
sgSatélite menor	5'-AUCUAAUAUGUUCUACAGUGUGUUUUA GAGCUAUGCUGUUUUG-3'	10
Control negativo 1	5'-CGCGAUAGCGCGAAUAUAUUGUUUUAG AGCUAUGCUGUUUUG-3'	11
ARNtracr	5'-DIG- AAACAGCAUAGCAAGUUAUUUUUAAGG CUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGCU-3'	12

5

El ensayo PLA se realizó con anticuerpos anti-Cas9 (ratón, Diagenode, dilución 1:3000) y anticuerpos anti-CENP-A (conejo, señalización celular, dilución 1:1000), que se incubaron durante la noche a 4 °C. Las células se lavaron con tampón de lavado de PLA A (2 x 5 minutos). Se añadieron soluciones de sonda de Ab-PLA secundaria de PLA diluidas [anti-conejo (+) y anti-ratón (-)] (dilución 1:5) y los portaobjetos se incubaron en una cámara de humedad precalentada durante 1 hora a 37 °C. Las células se lavaron con tampón de lavado de PLA A (2 x 5 min, TA) con agitación orbital suave. Se realizó la reacción de ligadura, se lavaron las células y se realizó la reacción de amplificación (kit Far red) esencialmente como se describe en la guía del usuario de PLA (DUOLINK® Fluorescent). Las células se tiñeron con DAPI (para ADN) y Phalloidin-Atto488 (para actina), se lavaron y cubrieron con PBS, se cubrieron con un portaobjetos de vidrio, las esquinas se sellaron con un endurecedor de esmalte de uñas transparente y se observaron en un microscopio fluorescente.

15

Los centrómeros se visualizaban como puntos rojos punteados solo cuando se usaban ARNcr orientados a las repeticiones de satélite menor para formar los RNP de CRISPR (FIG. 8A) pero no cuando se usaban ARNcr de control negativo (NC1) (FIG. 8B) o se usaban los principales ARNcr de satélite mayor (datos no mostrados). La amplificación de PLA se desencadenó usando ambos anticuerpos anti-Cas9 y anti-CENP-A (FIG. 8A). Usar anticuerpos anti-CENP-A solos (FIG. 8C) o anticuerpos anti-Cas9 solos (FIG. 8D) no daba una señal de PLA (DUOLINK®), lo que indica la especificidad dual y la naturaleza dependiente de la proximidad de la amplificación de PLA.

20

Ejemplo 2: Detección de telómeros utilizando complejos de dCAS9-ARNg

TRF1 y TRF2 son proteínas asociadas a repeticiones de telómeros en los extremos de los cromosomas (véanse las FIG. 9A, B). TRF2 y TRF1 se unen a la secuencia de repetición de telómero 5'-TTAGGG-3'. Los telómeros se detectaron en células U2OS usando un procedimiento similar al descrito anteriormente en el Ejemplo 1 con los siguientes cambios. (a) Los complejos de RNP se ensamblaron con ARNcr (5'-UAGGGUUAGGGUUAGGGUUAUGUUUUAGAGCUAUGCUGUUUUG-3'; SEQ ID NO: 13) orientados a la secuencia de repetición de los telómeros, y (b) el ensayo PLA se realizó con anticuerpos anti-Cas9 (ratón, Diagenode, dilución 1:3000) y anticuerpos anti-TRF2 (conejo, señalización celular, dilución 1:1000) (durante la noche a 4 °C).

25

Los telómeros se visualizaban como puntos rojos punteados solo cuando se usaban ARNcr orientados a la secuencia de repetición de telómeros para formar los RNP de CRISPR (FIG. 10A) pero no cuando se usaban ARNcr de control negativo 1 (NC1) (FIG. 10B). La amplificación de PLA se desencadenaba cuando se usaban ambos anticuerpos anti-Cas9 y anti-TRF2 (FIG. 10A). Usar cualquiera de los anticuerpos anti-Cas9 solos (FIG. 10C) o anticuerpos anti-TRF2 solos (FIG. 10D) no daba una señal de PLA (DUOLINK®), lo que indica la especificidad dual y la naturaleza dependiente de la proximidad de la amplificación de PLA.

30

35

Ejemplo 3: Detección de centrómeros en células fijadas usando complejos de dCAS9-ARNg

Se fijaron células U2OS.

Ejemplo 4: Detección de un locus genómico utilizando complejos de dCas9-ARNg

Puede complejarse dCas9 recombinante purificado (concretamente, D10A/H840A doble mutante) con ARNcr o agrupamientos de ARNcr (que se agrupan para orientarse a un locus genómico específico) y ARNtracr para formar complejos de dCas9-ARNg para uso para reconocer loci genómicos. Por ejemplo, uno o más ARNg pueden diseñarse dentro de una región de 2 kb que rodea sitios diana en AAVS1, EMX1 o Kras (concretamente, 1 kb en dirección 5' y 1 kb en dirección 3'). Las dianas de ARNg humano (y no dianas) se muestran a continuación.

1. AAVS1-ARNg (Sitio diana: GGGCCACTAGGGACAGGATTGG; SEQ ID NO: 14)
2. EMX1-ARNg (Sitio diana: AGTCCGAGCAGAAGAAGAA; SEQ ID NO: 15)
3. Kras-ARNg (Sitio diana: TAGTTGGAGCTGGTGGCGT; SEQ ID NO: 16)
4. No diana 1 (Sitio diana: CGCGATAGCGGAATATATT; SEQ ID NO: 17)

Los complejos de dCas9-ARNg pueden añadirse a células vivas o fijadas esencialmente como se describe anteriormente en los Ejemplos 1-3, y los complejos localizados pueden detectarse usando un ensayo PLA. Para esto, las células pueden incubarse con anticuerpos anti-Cas9 (conejo), luego incubarse con sondas de PLA anti-conejo con oligos (+) y (-) seguidos por los oligos conectores (que enlazarán los oligos (+) y (-) y tienen una secuencia de repetición interna no mamífera). A continuación, una reacción de ligadura formará el molde o moldes de ADN circular donde las sondas (+) y (-) se localizan dentro de una proximidad de aproximadamente 40 nm. Estas moléculas de ADN circular servirán como moldes para amplificación de círculo rodante usando la ADN polimerasa de phi29. El ADN amplificado localizado se puede visualizar usando oligosondas fluorescentes (p.ej., rojo lejano) que se hibridarán específicamente con la secuencia de repetición no mamífera del molde amplificado. Las señales rojas brillantes puntuales observadas usando un microscopio fluorescente con el filtro Cy5 indican una señal de PLA.

Listado de secuencias

<110> SIGMA-ALDRICH CO. LLC
 DAVIS, Gregory D.
 PALHAN, Vikas
 KREADER, Carol A.

<120> MÉTODOS Y REACTIVOS PARA LA DETECCIÓN DE PROXIMIDAD MOLECULAR USANDO PROTEÍNAS DE UNIÓN A ÁCIDO NUCLEICO GUIADAS POR ARN

<130> 047497-550236

<150> US 62/232.023
 <151> 24-09-2015

<160> 20

<170> PatentIn versión 3.5

<210> 1
 <211> 7
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> SINTETIZADO

<400> 1
 Pro Lys Lys Lys Arg Lys Val
 1 5

<210> 2
 <211> 7
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> SINTETIZADO

<400> 2

Pro Lys Lys Lys Arg Arg Val
 1 5

<210> 3
 <211> 16
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> SINTETIZADO

10 <400> 3
Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
 1 5 10 15

15 <210> 4
 <211> 20
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> SINTETIZADO

<400> 4
Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Pro Lys Lys
 1 5 10 15

Lys Arg Lys Val
 20

25 <210> 5
 <211> 19
 <212> PRT
 <213> Secuencia artificial

30 <220>
 <223> SINTETIZADO

<400> 5
Pro Leu Ser Ser Ile Phe Ser Arg Ile Gly Asp Pro Pro Lys Lys Lys
 1 5 10 15

Arg Lys Val

35 <210> 6
 <211> 24
 <212> PRT
 <213> Secuencia artificial

40 <220>
 <223> SINTETIZADO

<400> 6
Gly Ala Leu Phe Leu Gly Trp Leu Gly Ala Ala Gly Ser Thr Met Gly
 1 5 10 15

Ala Pro Lys Lys Lys Arg Lys Val
 45 20

<210> 7
 <211> 27
 <212> PRT
 50 <213> Secuencia artificial

ES 2 788 176 T3

<220>
 <223> SINTETIZADO

<400> 7
 5 Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly
 1 5 10 15

Ala Trp Ser Gln Pro Lys Lys Lys Arg Lys Val
 20 25

<210> 8
 <211> 21
 10 <212> PRT
 <213> Secuencia artificial

<220>
 <223> SINTETIZADO

15 <400> 8
 Lys Glu Thr Trp Trp Glu Thr Trp Trp Thr Glu Trp Ser Gln Pro Lys
 1 5 10 15

Lys Lys Arg Lys Val
 20

<210> 9
 20 <211> 43
 <212> ARN
 <213> Secuencia artificial

<220>
 25 <223> SINTETIZADO

<400> 9
 ccuaauucca cguccuacag uguuuuagag cuaugcuguu uug 43

30 <210> 10
 <211> 43
 <212> ARN
 <213> Secuencia artificial

35 <220>
 <223> SINTETIZADO

<400> 10
 40 aucuaauaug uucuaacagug uguuuuagag cuaugcuguu uug 43

<210> 11
 <211> 42
 <212> ARN
 <213> Secuencia artificial

45 <220>
 <223> SINTETIZADO

<400> 11
 50 cgcgauagcg cgaauauauu guuuuagagc uaugcuguuu ug 42

<210> 12
 <211> 69
 <212> ARN
 55 <213> Secuencia artificial

<220>
 <223> SINTETIZADO

ES 2 788 176 T3

<400> 12
 aaacagcaua gcaaguuaaa auaaggcuag uccguuauca acuugaaaaa guggcaccga 60
 gucggugcu 69

5 <210> 13
 <211> 43
 <212> ARN
 <213> Secuencia artificial

10 <220>
 <223> SINTETIZADO

<400> 13
 uaggguuagg guuaggguaa uguuuuagag cuaugcuguu uug 43

15 <210> 14
 <211> 22
 <212> ADN
 <213> Homo sapiens

20 <400> 14
 gggccactag ggacaggatt gg 22

25 <210> 15
 <211> 19
 <212> ADN
 <213> Homo sapiens

30 <400> 15
 agtccgagca gaagaaga 19

35 <210> 16
 <211> 19
 <212> ADN
 <213> Homo sapiens

<400> 16
 tagttggagc tggtagcgt 19

40 <210> 17
 <211> 20
 <212> ADN
 <213> Homo sapiens

45 <400> 17
 cgcgatagcg cgaatatatt 20

50 <210> 18
 <211> 24
 <212> ADN
 <213> Homo sapiens

<400> 18
 ttagggtag ggtagggtt aggg 24

55 <210> 19
 <211> 12
 <212> ADN
 <213> Homo sapiens

60 <400> 19
 aatccaatc cc 12
 <210> 20

<211> 11
<212> ARN
<213> Homo sapiens

5 <400> 20
caauccaau c 11

REIVINDICACIONES

- 5 1. Un complejo de sonda de detección de proximidad adecuado para ensayo de ligadura de proximidad que comprende una primera sonda que comprende (i) una primera proteína de unión a ácido nucleico guiada por ARN que es una sistema asociado a repeticiones palindrómicas cortas agrupadas y regularmente espaciadas (CRISPR) (Cas) genomanipulado (CRISPR/Cas) de origen no natural que comprende una proteína CRISPR/Cas y un ARN guía y (ii) un primer oligonucleótido que está ligado directa o indirectamente al primer sistema CRISPR/Cas; el complejo comprende además una segunda sonda que comprende (a) un segundo sistema CRISPR/Cas y un segundo oligonucleótido que está ligado directa o indirectamente al segundo sistema CRISPR/Cas o (b) un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico y un segundo oligonucleótido que está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico.
- 10 2. El complejo de la reivindicación 1, en el que la proteína CRISPR/Cas es una proteína Cas9.
3. El complejo de la reivindicación 2, en el que la proteína Cas9 comprende dos dominios de nucleasa funcionales, un dominio de nucleasa funcional o ningún dominio de nucleasa funcional.
- 15 4. El complejo de una cualquiera de las reivindicaciones 1 a 3, en el que la primera sonda comprende (i) el primer sistema CRISPR/Cas y (ii) el primer oligonucleótido está ligado a la proteína CRISPR/Cas por enlaces covalentes o no covalentes.
5. El complejo de una cualquiera de las reivindicaciones 1 a 3, en el que la primera sonda comprende (i) el primer sistema CRISPR/Cas y (ii) el primer oligonucleótido está ligado al ARN guía por enlaces covalentes o no covalentes.
- 20 6. El complejo de las reivindicaciones 4 o 5, en el que la segunda sonda comprende el segundo sistema CRISPR/Cas y el segundo oligonucleótido está ligado a la proteína CRISPR/Cas del segundo sistema CRISPR/Cas por enlaces covalentes o no covalentes; o el segundo oligonucleótido está ligado al ARN guía del segundo sistema CRISPR/Cas por enlaces covalentes o no covalentes.
7. El complejo de una cualquiera de las reivindicaciones 1 a 3, en el que la primera sonda comprende (i) el primer sistema CRISPR/Cas y (ii) el primer oligonucleótido está ligado a un anticuerpo primario dirigido contra la proteína CRISPR/Cas.
- 25 8. El complejo de la reivindicación 7, en el que la segunda sonda comprende un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico, y el segundo oligonucleótido está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico.
- 30 9. El complejo de una cualquiera de las reivindicaciones 1 a 3, en el que la primera sonda comprende (i) el sistema CRISPR/Cas, (ii) el primer y segundo oligonucleótidos que están ligados a anticuerpos secundarios antiespecie, y (iii) anticuerpos primarios dirigidos contra la proteína CRISPR/Cas.
10. El complejo de la reivindicación 9 que comprende además al menos un sistema CRISPR/Cas adicional.
- 35 11. El complejo de una cualquiera de las reivindicaciones 1 a 10, en el que los oligonucleótidos son monocatenarios y/o comprenden estructuras secundarias de tallo, bucle y/u horquilla; los oligonucleótidos comprenden desoxirribonucleótidos, ribonucleótidos o combinaciones de los mismos; y los desoxirribonucleótidos o ribonucleótidos son nucleótidos estándares, análogos de nucleótidos o nucleótidos modificados con 2'-O-metilo.
- 40 12. Un método para detectar un ácido nucleico endógeno en una célula, comprendiendo el método poner en contacto la célula con el complejo de sonda de detección de proximidad de una cualquiera de las reivindicaciones 1 a 5, en el que la proteína de unión a ácido nucleico guiada por ARN es guiada por ARN a un primer sitio en el ácido nucleico endógeno para unión, formando así un complejo de sonda de detección de proximidad unido y visualizando el complejo de sonda de detección de proximidad unido a través de una reacción de amplificación dependiente de proximidad *in situ* para detectar el ácido nucleico endógeno.
- 45 13. El método de la reivindicación 12, en el que la segunda sonda comprende (a) el segundo sistema CRISPR/Cas que comprende un segundo ARN guía orientado a un segundo sitio en el ácido nucleico endógeno y el segundo oligonucleótido de detección de proximidad está ligado a la proteína CRISPR/Cas del segundo sistema CRISPR/Cas por enlaces covalentes o no covalentes; (b) el segundo sistema CRISPR/Cas que comprende un segundo ARN guía orientado a un segundo sitio en el ácido nucleico endógeno y el segundo oligonucleótido de detección de proximidad está ligado al segundo ARN guía del segundo sistema CRISPR/Cas por enlaces covalentes o no covalentes; o (c) un anticuerpo dirigido contra una proteína asociada a ácido nucleico o una modificación de ácido nucleico y el segundo oligonucleótido de detección de proximidad está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico.
- 50 14. El método de la reivindicación 12 o 13, en el que el primer oligonucleótido de detección de proximidad está ligado a un anticuerpo primario dirigido contra la proteína CRISPR/Cas.
15. El método de la reivindicación 14, en el que la segunda sonda comprende un anticuerpo dirigido contra una

proteína asociada a ácido nucleico o una modificación de ácido nucleico en el ácido nucleico endógeno, y el segundo oligonucleótido de detección de proximidad está ligado al anticuerpo dirigido contra la proteína asociada a ácido nucleico o la modificación de ácido nucleico.

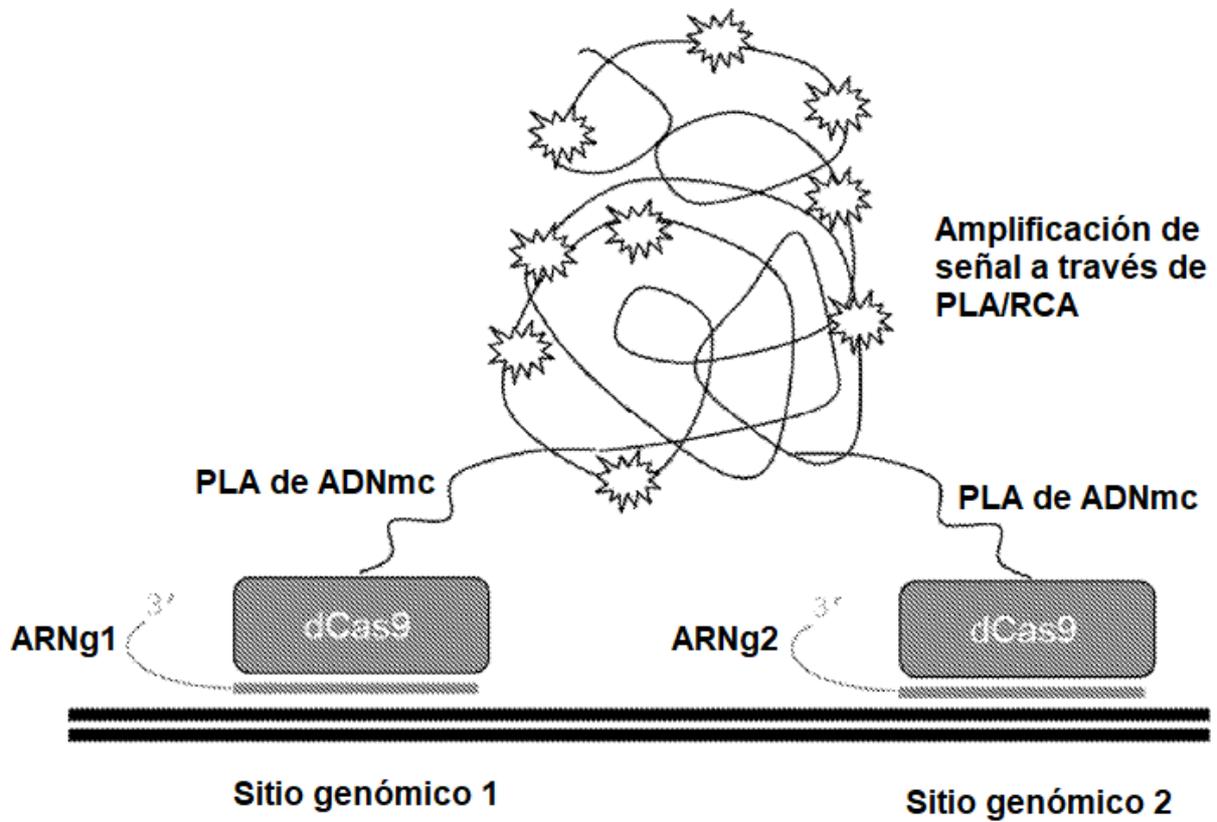


FIG. 1

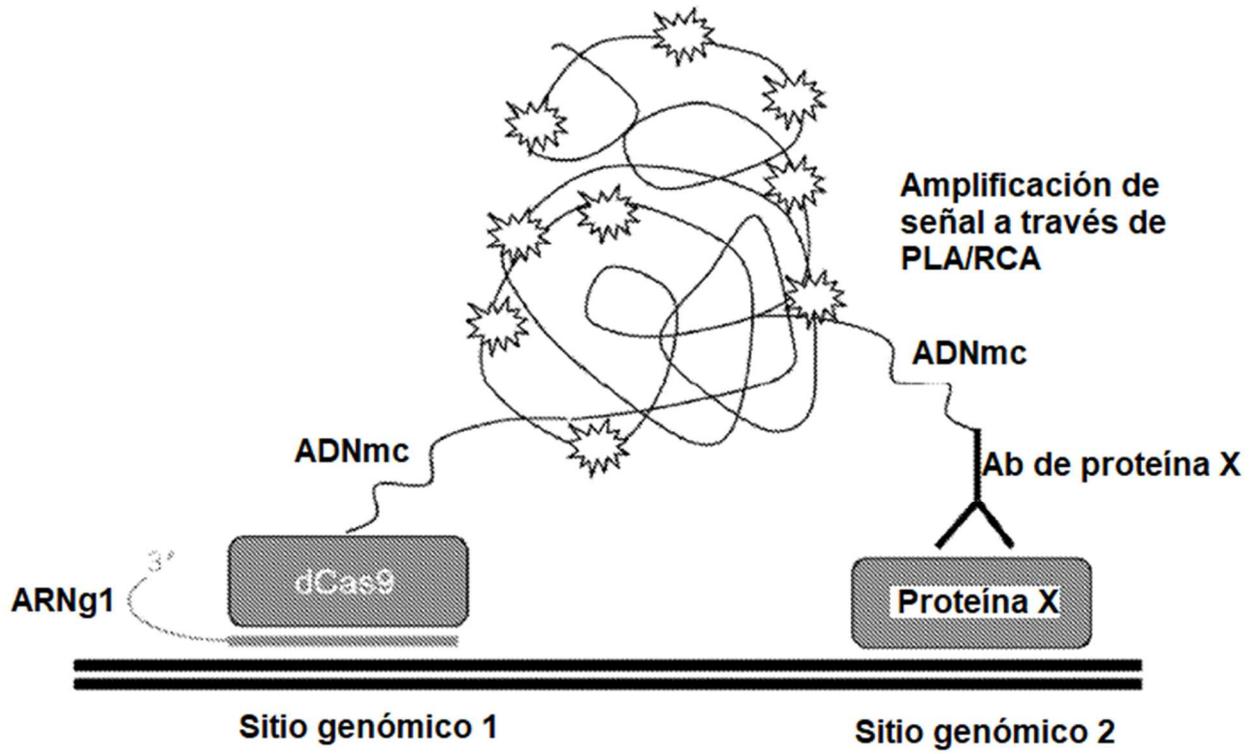


FIG. 2

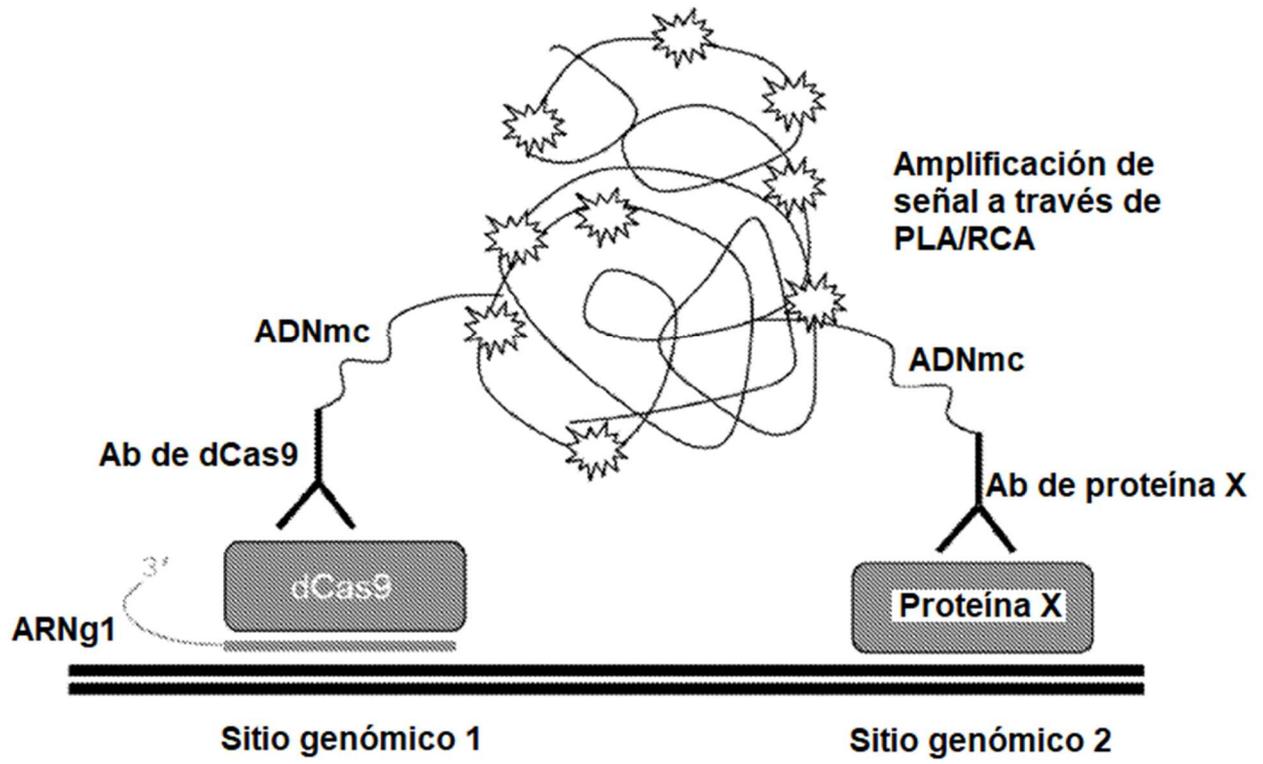


FIG. 3

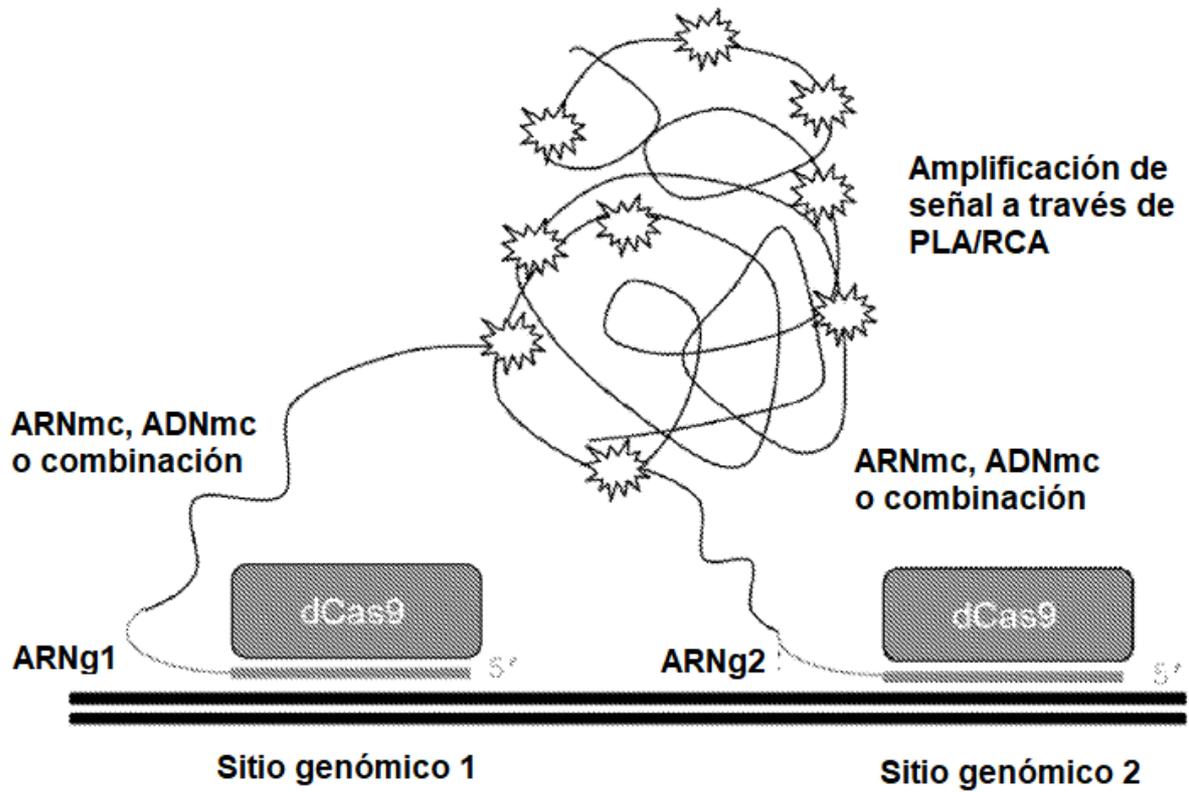


FIG. 4

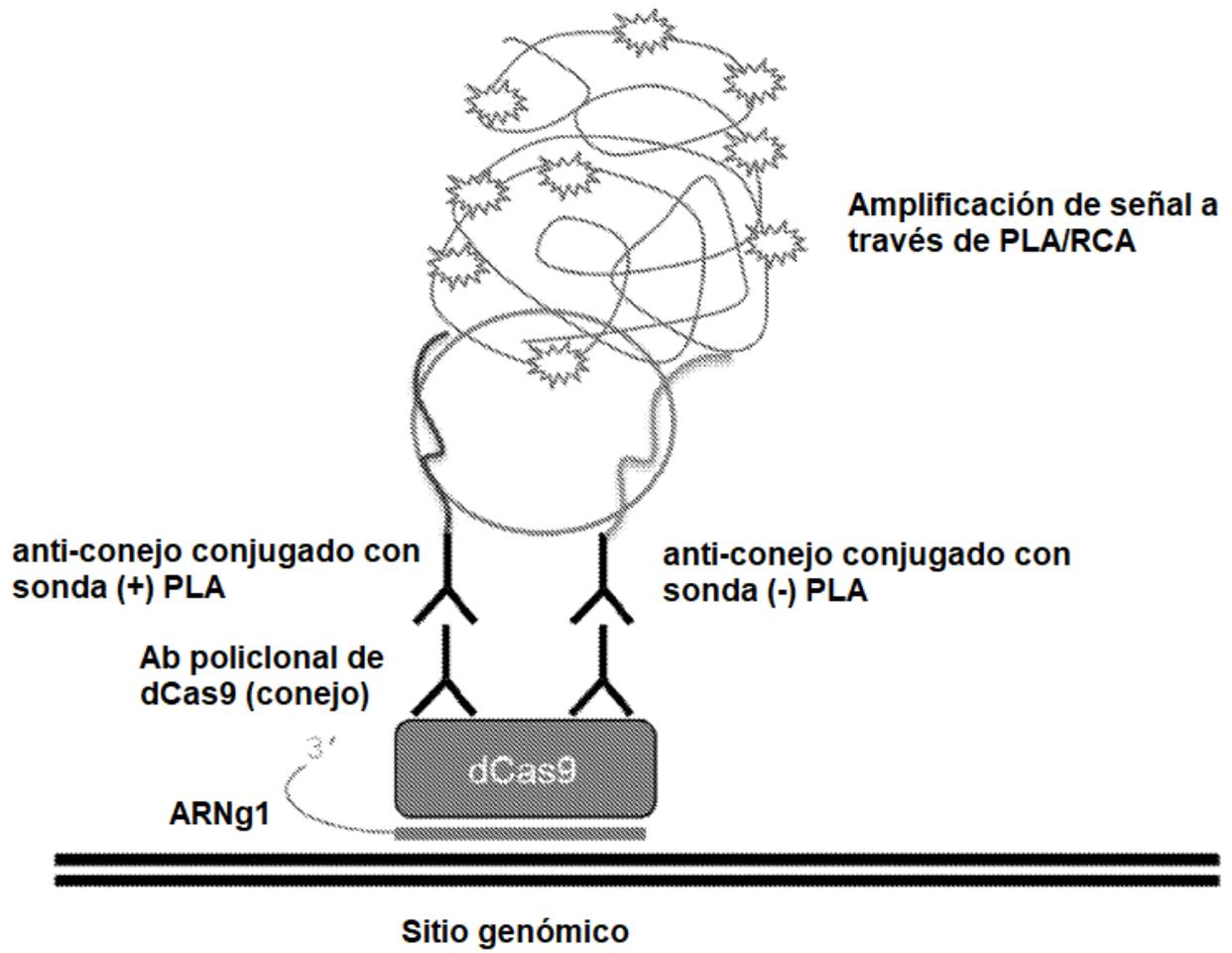


FIG. 5

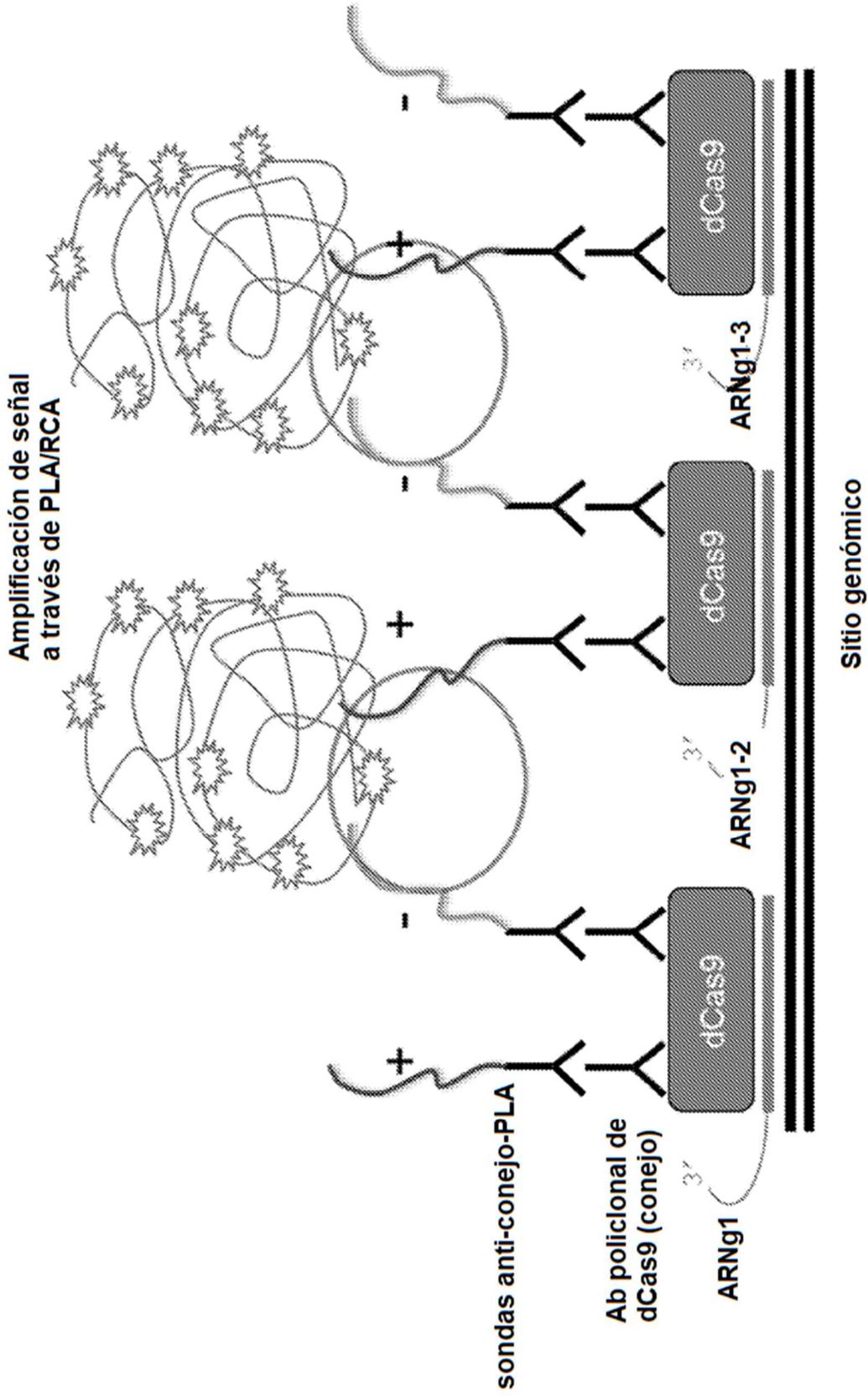


FIG. 6

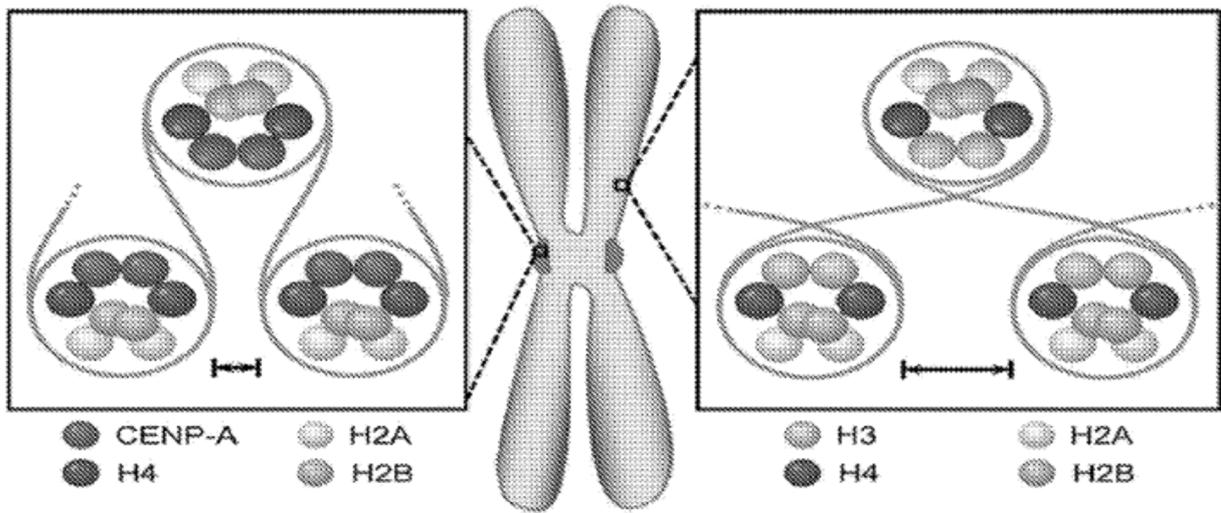


FIG. 7

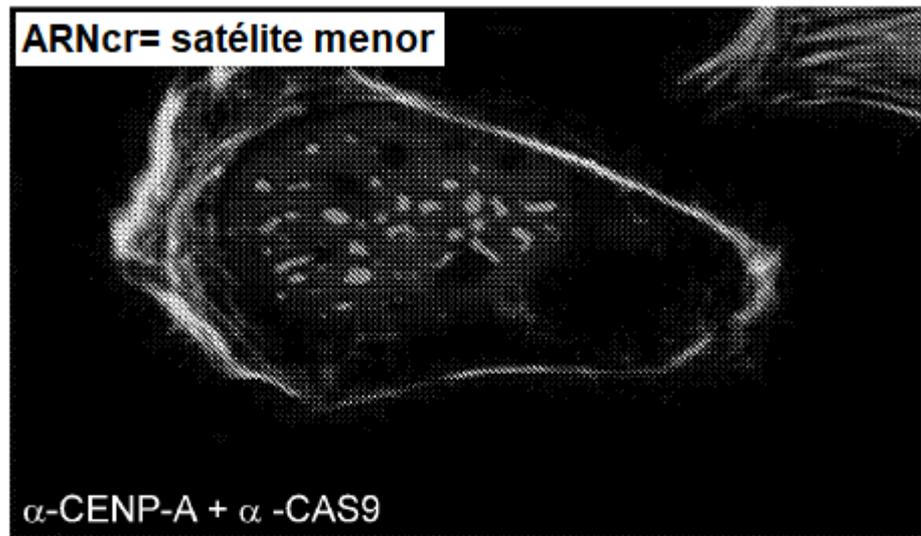


FIG. 8A

-  Duolink
-  DAPI
-  Actina

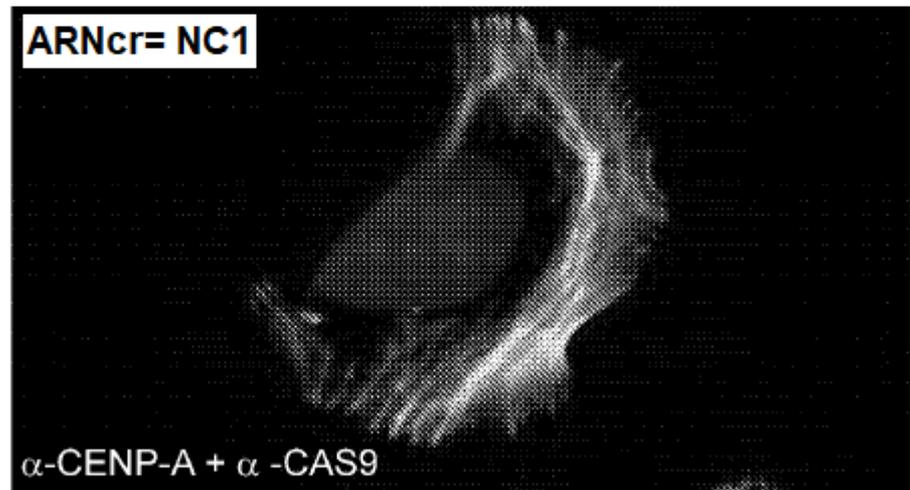


FIG. 8B

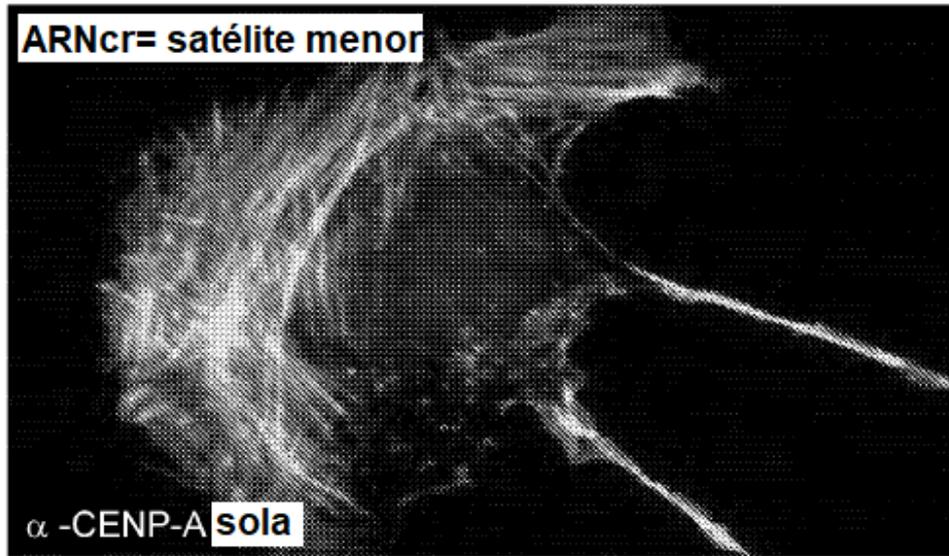


FIG. 8C

-  Duolink
-  DAPI
-  Actina

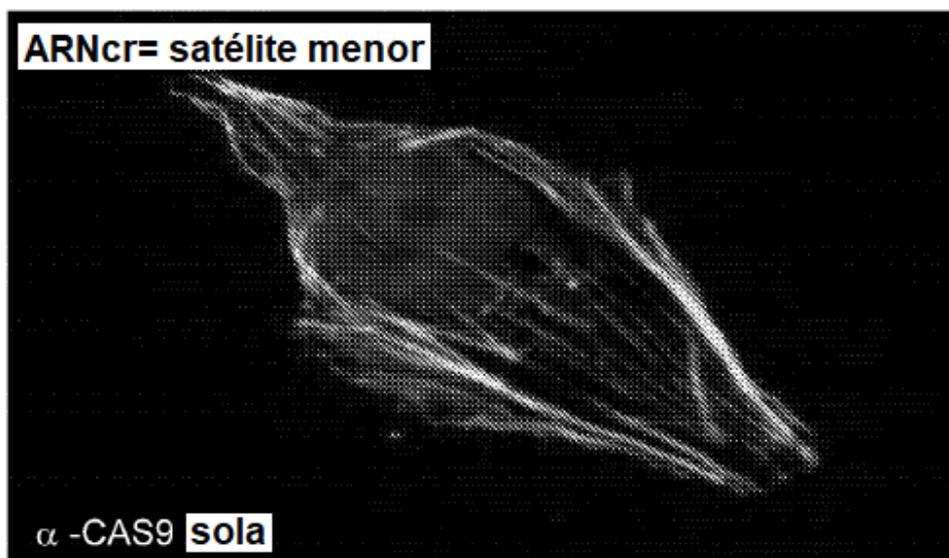
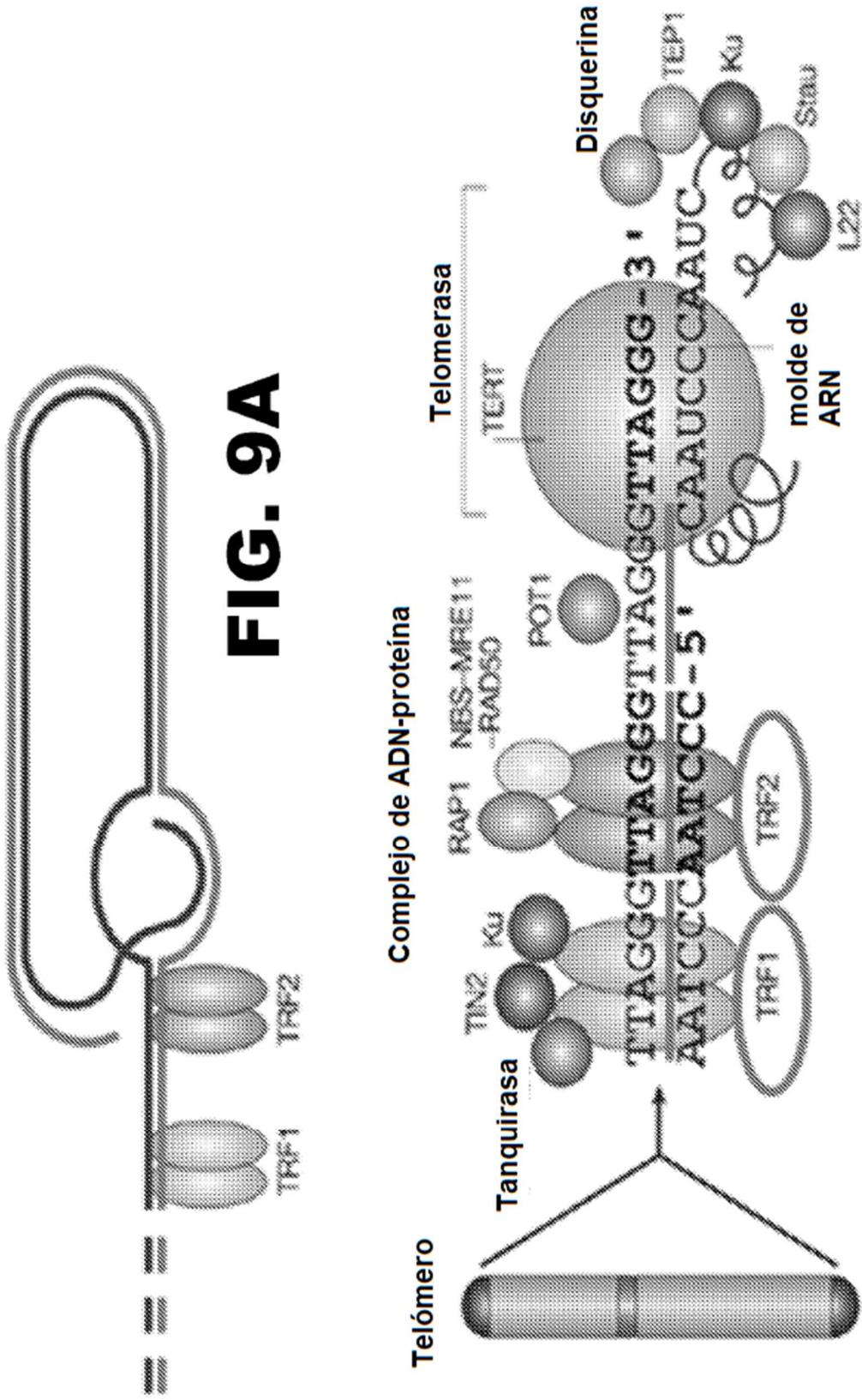
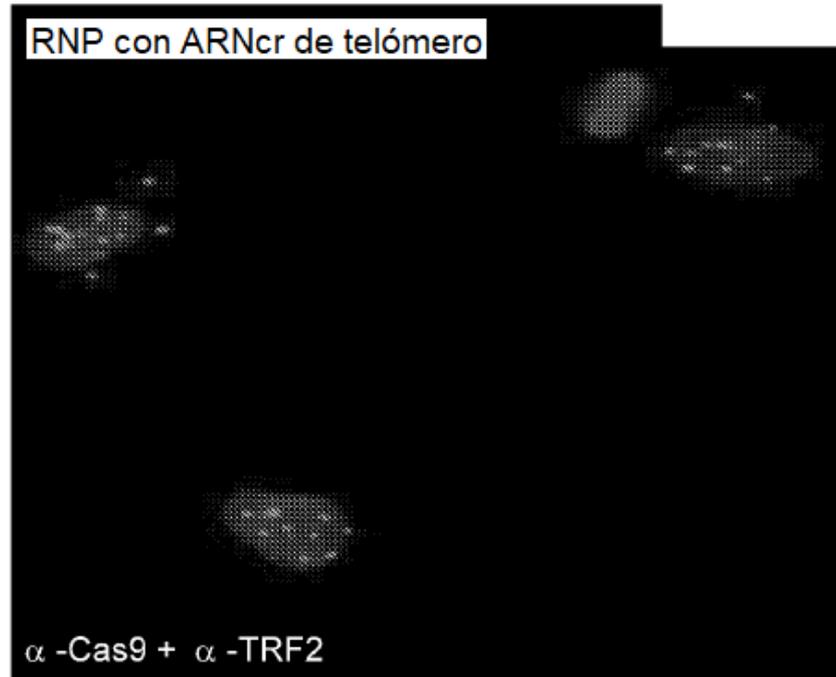


FIG. 8D





■ Duolink
■ DAPI

FIG. 10A

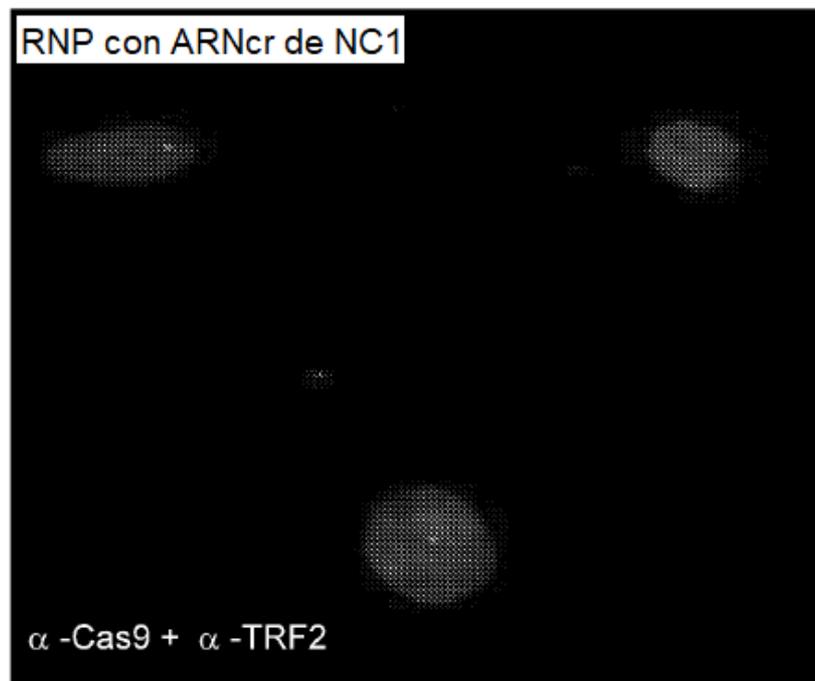


FIG. 10B

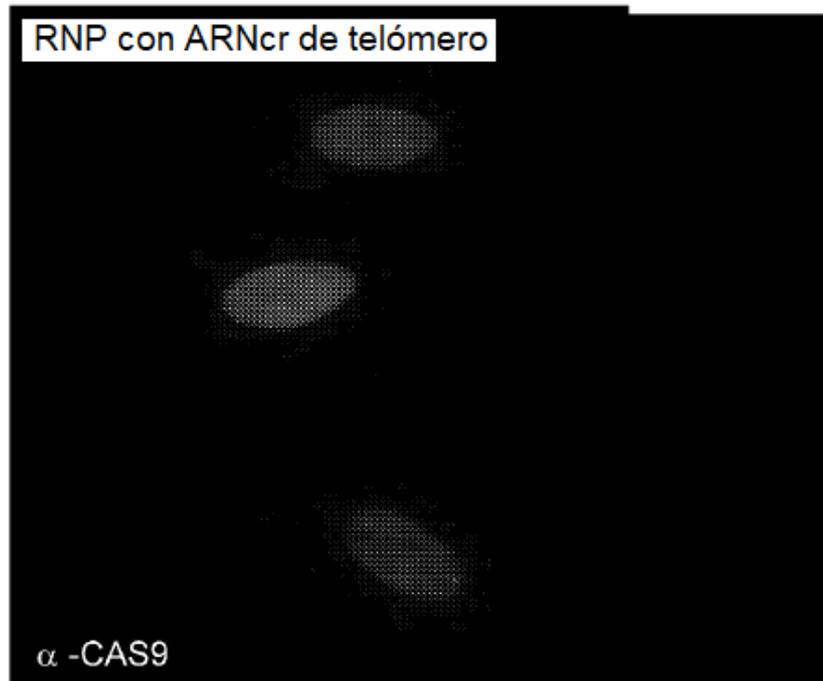


FIG. 10C

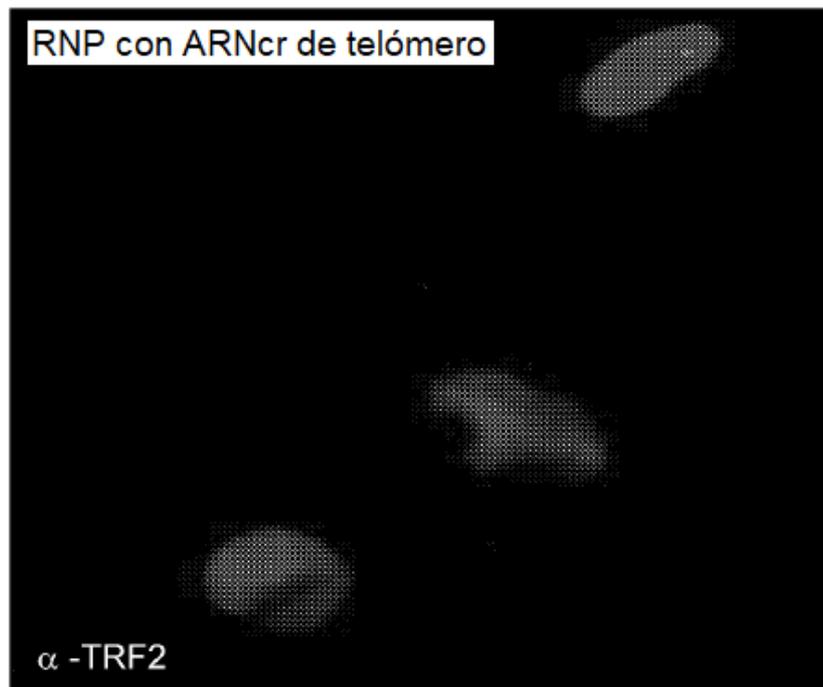


FIG. 10D

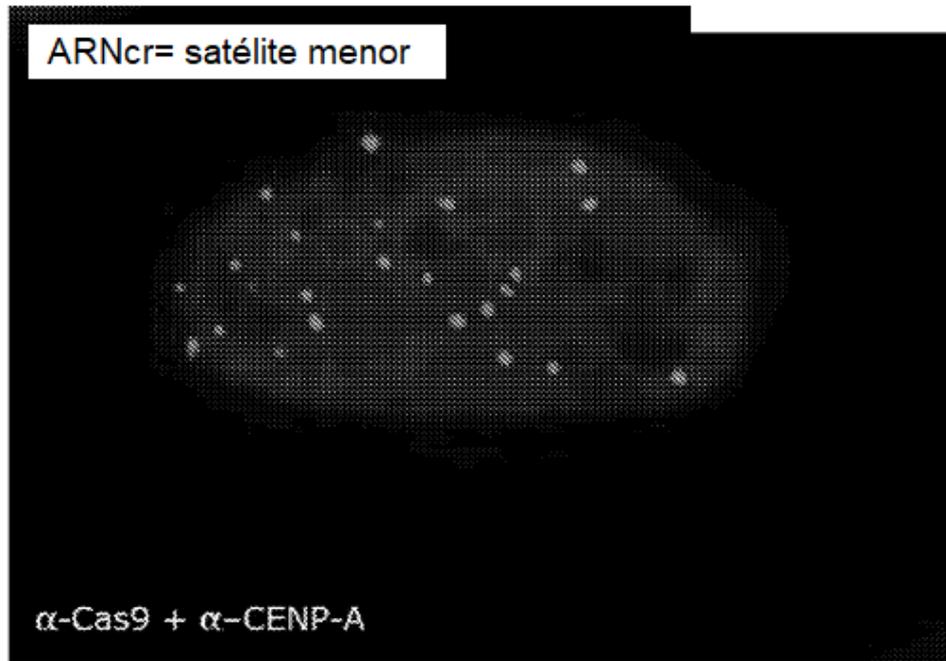


FIG. 11A

-  Duolink
-  DAPI

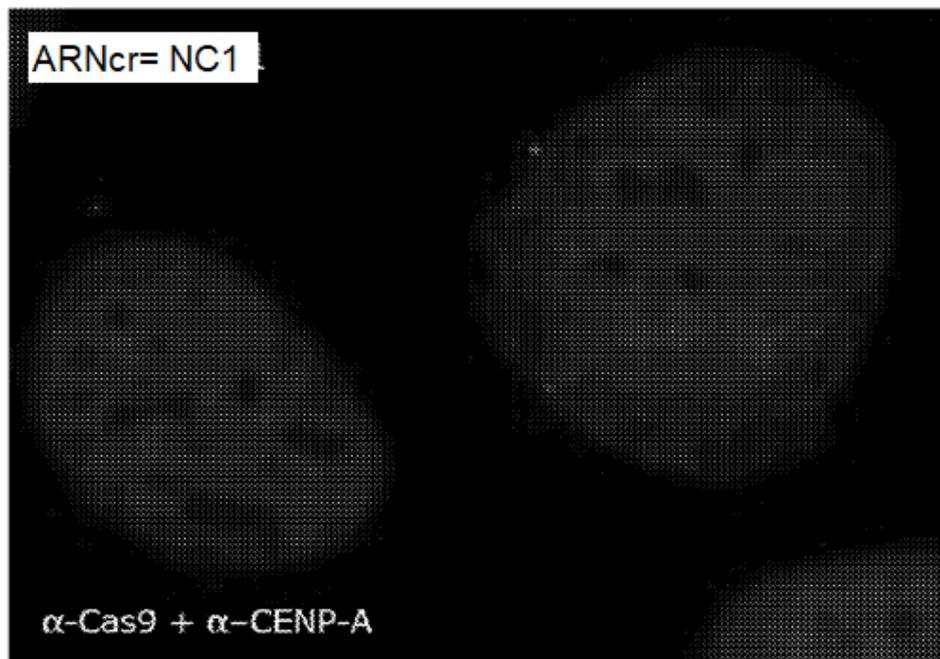


FIG. 11B

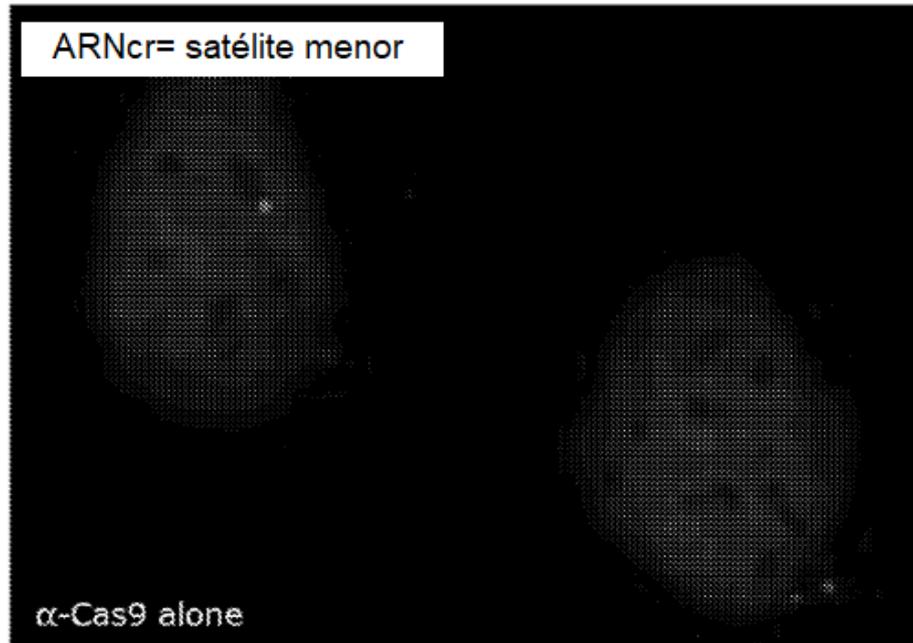


FIG. 11C

■ Duolink
■ DAPI

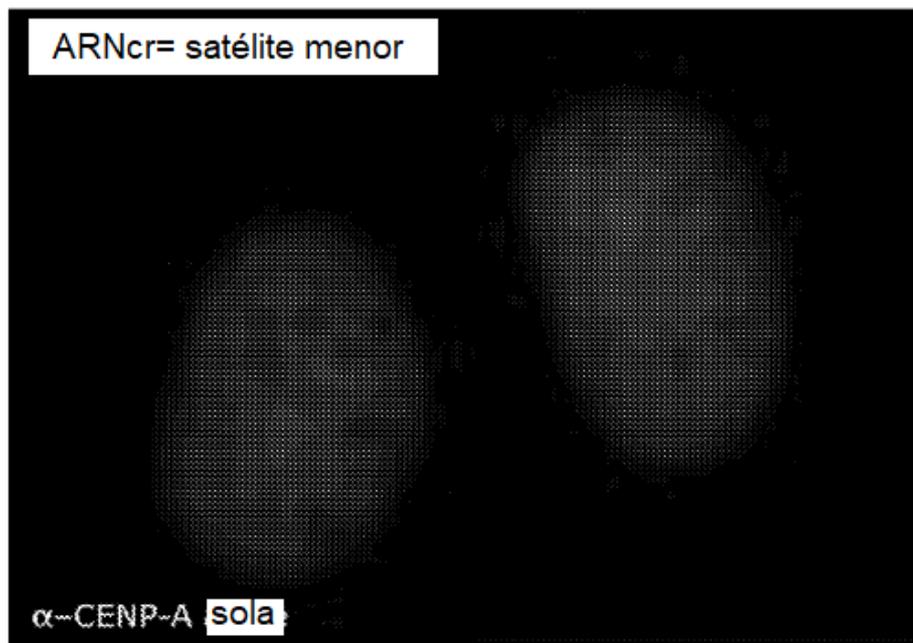


FIG. 11D