

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 927 463**

51 Int. Cl.:

C12N 15/09 (2006.01)
C12N 15/10 (2006.01)
C12N 15/11 (2006.01)
C12Q 1/68 (2008.01)
C12Q 1/6816 (2008.01)
C12Q 1/6811 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.12.2017** **PCT/US2017/065477**

87 Fecha y número de publicación internacional: **14.06.2018** **WO18107129**

96 Fecha de presentación y número de la solicitud europea: **08.12.2017** **E 17879489 (7)**

97 Fecha y número de publicación de la concesión europea: **29.06.2022** **EP 3551753**

54 Título: **Diagnósticos basados en el sistema efector CRISPR**

30 Prioridad:

09.12.2016 US 201662432553 P
08.02.2017 US 201762456645 P
15.03.2017 US 201762471930 P
12.04.2017 US 201762484869 P
04.10.2017 US 201762568268 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
07.11.2022

73 Titular/es:

THE BROAD INSTITUTE, INC. (33.3%)
415 Main Street
Cambridge, MA 02142, US;
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
(33.3%) y
PRESIDENT AND FELLOWS OF HARVARD
COLLEGE (33.3%)

72 Inventor/es:

ABUDAYYEH, OMAR;
COLLINS, JAMES JOSEPH;
GOOTENBERG, JONATHAN;
ZHANG, FENG y
LANDER, ERIC, S.

74 Agente/Representante:

ARIAS SANZ, Juan

Observaciones:

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

ES 2 927 463 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

Diagnósticos basados en el sistema efector CRISPR

Campo técnico

El objeto dado a conocer en el presente documento se refiere en general a diagnósticos rápidos relacionados con el uso de sistemas efectores CRISPR.

Antecedentes

Los ácidos nucleicos son una firma universal de información biológica. La capacidad de detectar rápidamente ácidos nucleicos con alta sensibilidad y especificidad de base única en una plataforma portátil tiene el potencial de revolucionar el diagnóstico y el seguimiento de muchas enfermedades, proporcionar información epidemiológica valiosa y servir como una herramienta científica generalizable. Aunque se han desarrollado muchos métodos para la detección de ácidos nucleicos (Du *et al.*, 2017; Green *et al.*, 2014; Kumar *et al.*, 2014; Pardee *et al.*, 2014; Pardee *et al.*, 2016; Urdea *et al.*, 2006), inevitablemente presentan compensaciones entre sensibilidad, especificidad, simplicidad y velocidad. Por ejemplo, los enfoques de qPCR son sensibles pero costosos y dependen de una instrumentación compleja, lo que limita la usabilidad a operadores altamente capacitados en entornos de laboratorio. Otros enfoques, tales como los nuevos métodos que combinan la amplificación isotérmica de ácidos nucleicos con plataformas portátiles (Du *et al.*, 2017; Pardee *et al.*, 2016), ofrecen una alta especificidad de detección en un entorno ambulatorio (POC), pero tienen aplicaciones algo limitadas debido a la baja sensibilidad. A medida que los diagnósticos de ácidos nucleicos se vuelven cada vez más relevantes para una variedad de aplicaciones de diagnóstico inmediato, las tecnologías de detección que brindan alta especificidad y sensibilidad a bajo coste serían de gran utilidad tanto en entornos de investigación clínica como básicos.

East-Seletsky *et al.* (2016), Nature, 538(7624): 270-273 (doi: 10.1038/nature19802) y Abudayyeh *et al.* (2016), Science, 353(6299), aaf5573-aaf5573 (doi: 10.1126/science.aaf5573) describen que las ARNasas guiadas por ARN, tales como C2c2, permanecen activas después de la escisión de su ARN diana, lo que lleva a la escisión "colateral" de los ARN no seleccionados como diana en las proximidades.

Sumario

En un aspecto, la invención proporciona un sistema de detección de ácidos nucleicos que comprende: un sistema CRISPR que comprende una proteína efectora y uno o más ARN guía diseñados para unirse a las moléculas diana correspondientes; un constructo de enmascaramiento basado en ARN; y reactivos de amplificación de ácidos nucleicos para amplificar moléculas de ARN diana en una muestra. También se describe en el presente documento, pero no según la invención, un sistema de detección de polipéptidos que comprende: un sistema CRISPR que comprende una proteína efectora y uno o más ARN guía diseñados para unirse a un ARN desencadenante, un constructo de enmascaramiento basado en ARN; y uno o más aptámeros de detección que comprenden un sitio de unión a promotor de la ARN polimerasa enmascarado o un sitio de unión de cebador enmascarado.

Según la invención, el sistema comprende reactivos de amplificación de ácidos nucleicos. Los reactivos de amplificación de ácidos nucleicos pueden comprender un cebador que comprende un promotor de la ARN polimerasa. En determinadas realizaciones, los ácidos nucleicos de muestra se amplifican para obtener un molde de ADN que comprende un promotor de la ARN polimerasa, mediante lo cual puede generarse una molécula de ARN diana mediante transcripción. El ácido nucleico puede ser ADN y amplificarse mediante cualquier método descrito en el presente documento. El ácido nucleico puede ser ARN y amplificarse mediante un método de transcripción inversa tal como se describe en el presente documento. La secuencia de aptámero puede amplificarse al desenmascarar el sitio de unión de cebadores, mediante lo cual se transcribe un ARN desencadenante a partir del producto de ADN amplificado. La molécula diana puede ser un ADN diana y el sistema puede comprender además un cebador que se une al ADN diana y comprende un promotor de la ARN polimerasa.

Según la invención, la proteína efectora del sistema CRISPR es una proteína efectora que selecciona como diana ARN. Los ejemplos de proteínas efectoras que seleccionan como diana ARN incluyen Cas13b y C2c2 (ahora conocidas como Cas13a). Se entenderá que el término "C2c2" en el presente documento se usa de manera intercambiable con "Cas13a". En otra realización de ejemplo, la proteína efectora que selecciona como diana ARN es C2c2. En otras realizaciones, la proteína efectora C2c2 es de un organismo de un género seleccionado del grupo que consiste en: *Leptotrichia*, *Listeria*, *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor*, *Mycoplasma*, *Campylobacter* y *Lachnospira*, o la proteína efectora C2c2 es un organismo seleccionado del grupo que consiste en: *Leptotrichia shahii*, *Leptotrichia wadei*, *Listeria seeligeri*, *Clostridium aminophilum*, *Carnobacterium gallinarum*, *Paludibacter propionigenes*, *Listeria weihenstephanensis* o la proteína efectora C2c2 es una proteína efectora C2c2 de *L. wadei* F0279 o *L. wadei* F0279 (Lw2). En otra realización, el uno o más ARN guía están diseñados para detectar un polimorfismo de un solo nucleótido, una variante de corte y empalme de un transcrito o una mutación de cambio de marco en un ARN o ADN diana.

En otras realizaciones, el uno o más ARN guía están diseñados para unirse a una o más moléculas diana que son de diagnóstico para un estado patológico. En todavía realizaciones adicionales, el estado patológico es una infección, una enfermedad de un órgano, una enfermedad de la sangre, una enfermedad del sistema inmunitario, un cáncer, una enfermedad del cerebro y del sistema nervioso, una enfermedad endocrina, una enfermedad relacionada con el embarazo o el parto, una enfermedad hereditaria o una enfermedad adquirida por el ambiente. En todavía realizaciones adicionales, el estado patológico es cáncer o una enfermedad autoinmunitaria o una infección.

En realizaciones adicionales, el uno o más ARN guía están diseñados para unirse a una o más moléculas diana que comprenden mutaciones somáticas específicas de cáncer. La mutación específica de cáncer puede conferir resistencia a los fármacos. La mutación de resistencia a los fármacos puede ser inducida por el tratamiento con ibrutinib, erlotinib, imatinib, gefitinib, crizotinib, trastuzumab, vemurafenib, RAF/MEK, terapia de bloqueo de puntos de control o terapia antiestrógeno. Las mutaciones específicas de cáncer pueden estar presentes en uno o más genes que codifican para una proteína seleccionada del grupo que consiste en el ligando 1 de muerte programada (PD-L1), el receptor de andrógenos (AR), la tirosina cinasa de Bruton (BTK), el receptor del factor de crecimiento epidérmico (EGFR), BCR-Abl, c-kit, PIK3CA, HER2, EML4-ALK, KRAS, ALK, ROS1, AKT1, BRAF, MEK1, MEK2, NRAS, RAC1 y ESR1. La mutación específica de cáncer puede ser una mutación en un gen seleccionado del grupo que consiste en CASP8, B2M, PIK3CA, SMC1A, ARID5B, TET2, ALPK2, COL5A1, TP53, DNER, NCOR1, MORC4, CIC, IRF6, MYOCD, ANKLE1, CNKSR1, NF1, SOS1, ARID2, CUL4B, DDX3X, FUBP1, TCP11L2, HLA-A, B o C, CSNK2A1, MET, ASXL1, PD-L1, PD-L2, IDO1, IDO2, ALOX12B y ALOX15B, o ganancia de número de copias, excluyendo acontecimientos de cromosomas completos, que afecten a cualquiera de las siguientes bandas cromosómicas: 6q16.1-q21, 6q22.31-q24.1, 6q25.1-q26, 7p11.2-q11.1, 8p23.1, 8p11.23- p11.21 (que contiene IDO1, IDO2), 9p24.2-p23 (que contiene PDL1, PDL2), 10p15.3, 10p15.1-p13, 11p14.1, 12p13.32-p13.2, 17p13.1 (que contiene ALOX12B, ALOX15B) y 22q11.1-q11.21.

En realizaciones adicionales, el uno o más ARN guía pueden diseñarse para unirse a una o más moléculas diana que comprenden marcadores de pérdida de heterocigosidad (LOH).

En realizaciones adicionales, el uno o más ARN guía pueden diseñarse para unirse a una o más moléculas diana que comprenden polimorfismos de un solo nucleótido (SNP). La enfermedad puede ser una enfermedad cardíaca y las moléculas diana pueden ser VKORC1, CYP2C9 y CYP2C19.

En realizaciones adicionales, el estado patológico puede ser una enfermedad relacionada con el embarazo o el parto o una enfermedad hereditaria. La muestra puede ser una muestra de sangre o una muestra mucosa. La enfermedad puede seleccionarse del grupo que consiste en trisomía 13, trisomía 16, trisomía 18, síndrome de Klinefelter (47, XXY), (47, XYY) y (47, XXX), síndrome de Turner, síndrome de Down (trisomía 21), fibrosis quística, enfermedad de Huntington, beta talasemia, distrofia miotónica, anemia de células falciformes, porfiria, síndrome del cromosoma X frágil, translocación robertsoniana, síndrome de Angelman, síndrome de DiGeorge y síndrome de Wolf-Hirschhorn.

En realizaciones adicionales, la infección está provocada por un virus, una bacteria o un hongo, o la infección es una infección viral. En realizaciones específicas, la infección viral está provocada por un virus de ARN bicatenario, un virus de ARN de sentido positivo, un virus de ARN de sentido negativo, un retrovirus o una combinación de los mismos, o la infección viral está provocada por un virus *Coronaviridae*, un virus *Picornaviridae*, un virus *Caliciviridae*, un virus *Flaviviridae*, un virus *Togaviridae*, un *Bornaviridae*, un *Filoviridae*, un *Paramyxoviridae*, un *Pneumoviridae*, un *Rhabdoviridae*, un *Arenaviridae*, un *Bunyaviridae*, un *Orthomyxoviridae* o un *Deltavirus*, o la infección viral está provocada por coronavirus, SARS, virus de la poliomielitis, rinovirus, hepatitis A, virus de Norwalk, virus de la fiebre amarilla, virus del Nilo Occidental, virus de la hepatitis C, virus de fiebre del dengue, virus de Zika, virus de la rubéola, virus del río Ross, virus Sindbis, virus de Chikungunya, virus de la enfermedad de Borna, virus del ébola, virus de Marburgo, virus del sarampión, virus de las paperas, virus de Nipah, virus de Hendra, virus de enfermedad de Newcastle, virus sincicial respiratorio humano, virus de la rabia, virus de Lassa, *Hantavirus*, virus de la fiebre hemorrágica de Crimea-Congo, gripe o virus de la hepatitis D.

En otras realizaciones de la invención, el constructo de enmascaramiento basado en ARN suprime la generación de una señal positiva detectable o el constructo de enmascaramiento basado en ARN suprime la generación de una señal positiva detectable enmascarando la señal positiva detectable, o generando en su lugar una señal negativa detectable, o el constructo de enmascaramiento basado en ARN comprende un ARN de silenciamiento que suprime la generación de un producto génico codificado por un constructo indicador, en el que el producto génico genera la señal positiva detectable cuando se expresa.

En realizaciones adicionales, el constructo de enmascaramiento basado en ARN es una ribozima que genera la señal detectable negativa, y en el que la señal detectable positiva se genera cuando la ribozima se desactiva, o la ribozima convierte un sustrato en un primer color y en el que el sustrato se convierte en un segundo color cuando la ribozima está desactivada.

En otras realizaciones, el agente de enmascaramiento basado en ARN es un aptámero de ARN, o el aptámero secuestra una enzima, en el que la enzima genera una señal detectable tras su liberación del aptámero actuando sobre un sustrato, o el aptámero secuestra un par de agentes que cuando se liberan de los aptámeros se combinan para generar una señal detectable.

En otra realización, el constructo de enmascaramiento basado en ARN comprende un oligonucleótido de ARN al que se unen un ligando detectable y un componente de enmascaramiento. En otra realización, el ligando detectable es un fluoróforo y el componente de enmascaramiento es una molécula extintora, o los reactivos para amplificar moléculas de ARN diana tales como, pero sin limitarse a, reactivos de NASBA o RPA.

- 5 En otro aspecto, la invención proporciona un dispositivo de diagnóstico que comprende uno o más volúmenes discretos individuales, comprendiendo cada volumen discreto individual una proteína efectora CRISPR, uno o más ARN guía diseñados para unirse a la molécula diana correspondiente, un constructo de enmascaramiento basado en ARN y reactivos de amplificación de ácidos nucleicos.

- 10 En otro aspecto, la invención proporciona un dispositivo de diagnóstico que comprende uno o más volúmenes discretos individuales, comprendiendo cada volumen discreto individual una proteína efectora CRISPR, uno o más ARN guía diseñados para unirse a un ARN desencadenante, uno o más aptámeros de detección que comprenden un sitio de unión a promotor de la polimerasa ARN enmascarado o un sitio de unión de cebador enmascarado, y reactivos de amplificación de ácidos nucleicos.

- 15 En algunas realizaciones, los volúmenes discretos individuales son gotas, o los volúmenes discretos individuales se definen sobre un sustrato sólido, o los volúmenes discretos individuales son micropocillos, o los volúmenes discretos individuales son puntos definidos sobre un sustrato, tal como un sustrato de papel.

- Según la invención, la proteína efectora que selecciona como diana ARN es una proteína efectora CRISPR que selecciona como diana ARN de tipo VI, tal como C2c2 o Cas13b. En determinadas realizaciones de ejemplo, la proteína efectora C2c2 es de un organismo seleccionado del grupo que consiste en: *Leptotrichia*, *Listeria*,
 20 *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor*, *Mycoplasma* y *Campylobacter*, o la proteína efectora C2c2 se selecciona del grupo que consiste en: *Leptotrichia shahii*, *L. wadei*, *Listeria seeligeri*, *Lachnospiraceae bacteria*,
 25 *Clostridium aminophilum*, *Carnobacterium gallinarum*, *Paludibacter propionigenes*, *Listeria weihenstephanensis*, *bacteria Listeriaceae* y *Rhodobacter capsulatus*, la proteína efectora C2c2 es una proteína efectora C2c2 de *L. wadei* F0279 o *L. wadei* F0279 (Lw2). En otra realización, el uno o más ARN guía están diseñados para unirse a una o más secuencias de ARN diana que son de diagnóstico para un estado patológico.

- En determinadas realizaciones de ejemplo, el constructo de enmascaramiento basado en ARN suprime la generación de una señal positiva detectable, o el constructo de enmascaramiento basado en ARN suprime la
 30 generación de una señal positiva detectable enmascarando la señal positiva detectable, o generando en su lugar una señal negativa detectable, o el constructo de enmascaramiento basado en ARN comprende un ARN de silenciamiento que suprime la generación de un producto génico codificado por un constructo indicador, en el que el producto génico genera la señal positiva detectable cuando se expresa.

- En otra realización de ejemplo, el constructo de enmascaramiento basado en ARN es una ribozima que genera una
 35 señal detectable negativa, y en el que la señal detectable positiva se genera cuando la ribozima se desactiva. En una realización de ejemplo, la ribozima convierte un sustrato en un primer color y en el que el sustrato se convierte en un segundo color cuando la ribozima se desactiva. En otra realización de ejemplo, el agente de enmascaramiento basado en ARN es un aptámero que secuestra una enzima, en el que la enzima genera una señal detectable tras su liberación del aptámero actuando sobre un sustrato, o el aptámero secuestra un par de agentes que cuando se
 40 liberan de los aptámeros se combinan para generar una señal detectable.

En otra realización de ejemplo, el constructo de enmascaramiento basado en ARN comprende un oligonucleótido de ARN al que se unen un oligonucleótido de ligando detectable y un componente de enmascaramiento. En determinadas realizaciones de ejemplo, el ligando detectable es un fluoróforo y el componente de enmascaramiento es una molécula extintora.

- 45 En otro aspecto, la invención proporciona un método para detectar ARN diana en muestras, que comprende: distribuir una muestra o conjunto de muestras en uno o más volúmenes discretos individuales, comprendiendo los volúmenes discretos individuales un sistema CRISPR que comprende una proteína efectora, uno o más ARN guía, un constructo de enmascaramiento basado en ARN; incubar la muestra o conjunto de muestras en condiciones suficientes para permitir la unión del uno o más ARN guía a una o más moléculas diana; activar la proteína efectora
 50 CRISPR mediante la unión del uno o más ARN guía a una o más moléculas diana, en el que la activación de la proteína efectora CRISPR da como resultado la modificación del constructo de enmascaramiento basado en ARN de manera que se produce una señal positiva detectable; y detectar la señal positiva detectable, en el que la detección de la señal positiva detectable indica la presencia de una o más moléculas diana en la muestra.

- También se describe en el presente documento, pero no según la invención, un método para detectar péptidos en
 55 muestras, que comprende: distribuir una muestra o conjunto de muestras en un conjunto de volúmenes discretos individuales, comprendiendo los volúmenes discretos individuales aptámeros de detección de péptidos, un sistema CRISPR que comprende una proteína efectora, uno o más ARN guía, un constructo de enmascaramiento basado en ARN, en el que los aptámeros de detección de péptidos comprenden un sitio de ARN polimerasa enmascarado y

están configurados para unirse a una o más moléculas diana; incubar la muestra o conjunto de muestras en condiciones suficientes para permitir la unión de los aptámeros de detección de péptidos a una o más moléculas diana, en el que la unión del aptámero a una molécula diana correspondiente expone el sitio de unión de la ARN polimerasa, lo que da como resultado la síntesis de ARN de un ARN desencadenante; activar la proteína efectora CRISPR mediante la unión de uno o más ARN guía al ARN desencadenante, en el que la activación de la proteína efectora CRISPR da como resultado la modificación del constructo de enmascaramiento basado en ARN de manera que se produce una señal positiva detectable; y detectar la señal positiva detectable, en el que la detección de la señal positiva detectable indica la presencia de una o más moléculas diana en una muestra.

En determinados ejemplos, dichos métodos comprenden además amplificar el ARN de muestra o el ARN desencadenante. En otros ejemplos, la amplificación del ARN comprende la amplificación mediante NASBA o RPA.

En determinadas realizaciones de ejemplo, la proteína efectora CRISPR es una proteína efectora CRISPR que selecciona como diana ARN de tipo VI, tal como C2c2 o Cas13b. En otras realizaciones de ejemplo, la proteína efectora C2c2 es de un organismo seleccionado del grupo que consiste en: *Leptotrichia*, *Listeria*, *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor*, *Mycoplasma* y *Campylobacter*, o la proteína efectora C2c2 se selecciona del grupo que consiste en: *Leptotrichia shahii*, *L. wadei*, *Listeria seeligeri*, *Lachnospiraceae bacteria*, *Clostridium aminophilum*, *Carnobacterium gallinarum*, *Paludibacter propionigenes*, *Listeria weihenstephanensis*, bacteria *Listeriaceae* y *Rhodobacter capsulatus*. En una realización específica, la proteína efectora C2c2 es una proteína efectora C2c2 de *L. wadei* F0279 o *L. wadei* F0279 (Lw2).

En determinadas realizaciones de ejemplo, el uno o más ARN guía están diseñados para unirse a una o más moléculas diana que son de diagnóstico para un estado patológico. En determinadas otras realizaciones de ejemplo, el estado patológico es una infección, una enfermedad de un órgano, una enfermedad de la sangre, una enfermedad del sistema inmunitario, un cáncer, una enfermedad del cerebro y del sistema nervioso, una enfermedad endocrina, una enfermedad relacionada con el embarazo o el parto, una enfermedad hereditaria o una enfermedad adquirida por el ambiente, cáncer, o una infección fúngica, una infección bacteriana, una infección parasitaria, o una infección viral.

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento basado en ARN suprime la generación de una señal positiva detectable, o el constructo de enmascaramiento basado en ARN suprime la generación de una señal positiva detectable enmascarando la señal positiva detectable, o generando en su lugar una señal negativa detectable, o el constructo de enmascaramiento basado en ARN comprende un ARN de silenciamiento que suprime la generación de un producto génico codificado por un constructo indicador, en el que el producto génico genera la señal positiva detectable cuando se expresa, o el constructo de enmascaramiento basado en ARN es una ribozima que genera la señal detectable negativa, y en el que la señal detectable positiva se genera cuando la ribozima se desactiva. En otras realizaciones de ejemplo, la ribozima convierte un sustrato en un primer estado y en el que el sustrato se convierte en un segundo estado cuando la ribozima se desactiva, o el agente de enmascaramiento basado en ARN es un aptámero, o el aptámero secuestra una enzima, en el que la enzima genera una señal detectable tras su liberación del aptámero actuando sobre un sustrato, o el aptámero secuestra un par de agentes que cuando se liberan de los aptámeros se combinan para generar una señal detectable. En todavía realizaciones adicionales, el constructo de enmascaramiento basado en ARN comprende un oligonucleótido de ARN con un ligando detectable en un primer extremo del oligonucleótido de ARN y un componente de enmascaramiento en un segundo extremo del oligonucleótido de ARN, o el ligando detectable es un fluoróforo y el enmascaramiento componente es una molécula extintora.

Estos y otros aspectos, objetos, características y ventajas de las realizaciones de ejemplo resultarán evidentes para los expertos en la técnica al considerar la siguiente descripción detallada de las realizaciones de ejemplo ilustradas.

Breve descripción de los dibujos

La figura 1 es un diagrama de un sistema efector CRISPR basado en C2c2 de ejemplo.

La figura 2 proporciona (A) un diagrama del locus de CRISPR/C2c2 de *Leptotrichia wadei*. Se muestran estructuras representativas de ARNcr de los sistemas LwC2c2 y LshC2c2. (SEQ. ID. NO. 142 y 143) (B) Diagrama del ensayo bacteriano *in vivo* para determinar la actividad C2c2. Se clona un protoespaciador en el sentido de 5' del gen de la betalactamasa en un plásmido resistente a la ampicilina, y este constructo se transforma en *E. coli* que expresa C2c2 junto con un espaciador o bien de direccionamiento o bien de no direccionamiento. Para cuantificar la actividad, se cuentan las transformantes exitosas. (C) Cuantificación de la actividad *in vivo* de LwC2c2 y LshC2c2. (n=2 réplicas biológicas; las barras representan la media \pm e.e.m.) (D) Filtración final en gel de exclusión molecular de LwC2c2. (E) Gel de acrilamida teñido con azul de Coomassie de la purificación por etapas de LwC2c2. (F) Actividad de LwC2c2 frente a diferentes dianas de PFS. LwC2c2 se dirigió contra ARN fluorescente con PFS variable en 3' flanqueando el espaciador, y los productos de la reacción se visualizaron en un gel desnaturalizante. LwC2c2 muestra una ligera preferencia contra SSP G.

Las figuras 3-6 muestran la detección de un constructo de enmascaramiento de ejemplo a diferentes diluciones usando 1 mg, 100 ng, 10 ng y 1 ng de diana con 4 cantidades diferentes de proteína/ARNcr (1:4, 1:16, 1:32, 1:64) con 2 agrupaciones de ARNcr, sin condición de ARNcr, duplicados técnicos, en $(96+48) \times 2 = 288$ reacciones, medidas en intervalos de 5 minutos durante 3 horas.

- 5 La figura 7 proporciona un diagrama de un esquema de detección de ejemplo usando un constructo de enmascaramiento y la proteína efectora CRISPR, según determinadas realizaciones de ejemplo.

La figura 8 proporciona un conjunto de gráficos que muestran cambios en la fluorescencia a lo largo del tiempo cuando se detecta una diana usando diferentes agrupaciones de ARN guía.

- 10 La figura 9 proporciona un gráfico que muestra la fluorescencia normalizada detectada a través de diferentes diluciones de ARN diana a concentraciones variables de la proteína efectora CRISPR.

La figura 10 es un diagrama que muestra las etapas generales de una reacción de amplificación NASBA.

La figura 11 proporciona un gráfico que muestra la detección del ácido nucleico diana ARNmc 1 amplificado por NASBA con tres conjuntos diferentes de cebadores y luego sometido a la detección colateral con C2c2 usando una sonda fluorescente extinguida. (n=2 réplicas técnicas; las barras representan la media \pm e.e.m.)

- 15 La figura 12 proporciona un gráfico que muestra que el efecto colateral puede usarse para detectar la presencia de un ARN diana lentiviral.

La figura 13 proporciona un gráfico que demuestra que el efecto colateral y NASBA pueden detectar especies a concentraciones aM.

- 20 La figura 14 proporciona un gráfico que demuestra que el efecto colateral y NASBA discriminan rápidamente las muestras de baja concentración.

La figura 15 muestra que la fluorescencia normalizada en puntos de tiempo particulares es predictiva de la concentración de entrada de la muestra. Las mediciones de fluorescencia de la detección con Cas13a sin amplificación están correlacionadas con la concentración de ARN de entrada. (n=2 réplicas biológicas; las barras representan la media \pm e.e.m.).

- 25 La figura 16 proporciona un diagrama de la reacción de RPA, mostrando los componentes participantes en la reacción.

La figura 17 diagrama de SHERLOCK; proporciona un diagrama que muestra la detección de las dianas tanto de ADN como de ARN mediante la incorporación de una etapa de RPA o RT-RPA debidamente. Tras el reconocimiento del ARN diana, el efecto colateral provoca que C2c2 corte el indicador de escisión, generando fluorescencia. Cantidades de moléculas individuales de ARN o ADN pueden amplificarse a ADN mediante la amplificación por recombinasa y polimerasa (RPA) y transcribirse para producir ARN que luego es detectado por C2c2.

- 30

La figura 18 proporciona un diagrama de la diana de ARNmc detectada mediante la detección colateral con C2c2 (SEQ. ID. NO. 144 y 145).

- 35 La figura 19 proporciona un conjunto de gráficos que demuestran la detección de moléculas individuales de ADN usando RPA (es decir, en el plazo de 15 minutos tras la adición de C2c2).

La figura 20 proporciona un conjunto de gráficos que demuestran que mezclar T7 polimerasa en una reacción de RPA afecta negativamente a la detección de ADN.

La figura 21 proporciona un conjunto de gráficos que demuestran que la mezcla de polimerasa en una reacción de RPA no afecta negativamente a la detección de ADN.

- 40 La figura 22 proporciona un gráfico que demuestra que las reacciones de detección con C2c2, transcripción con T7 y de RPA son compatibles y logran la detección de moléculas individuales cuando se incuban simultáneamente (n=2 réplicas técnicas; las barras representan la media \pm e.e.m.).

La figura 23 proporciona un conjunto de gráficos que demuestran la eficacia de las incubaciones rápidas de RPA-ARN.

- 45 La figura 24 proporciona un conjunto de gráficos que demuestran que aumentar la cantidad de polimerasa T7 refuerza la sensibilidad para RPA-ARN.

La figura 25 proporciona un conjunto de gráficos que muestran los resultados de un ensayo de detección de RPA-ADN usando una reacción de una sola etapa con enzimas 1,5x. La detección de moléculas individuales (2 aM) se consigue tan pronto como en 30 minutos.

- 50 La figura 26 proporciona un conjunto de gráficos que demuestran que una reacción de RPA-ADN de una sola etapa

demuestra una disminución cuantitativa de la fluorescencia en relación con la concentración de entrada. La curva ajustada revela la relación entre la concentración de entrada diana y la fluorescencia de salida.

La figura 27 proporciona un conjunto de gráficos que demuestran que (A) la detección de ARN con C2c2 sin amplificación puede detectar la diana de ARNmc a concentraciones de hasta 50 fM. (n=2 réplicas técnicas; las barras representan la media \pm e.e.m.), y que (B) la reacción de RPA-C2c2 puede detectar moléculas individuales de ADN (n=4 réplicas técnicas; las barras representan la media \pm e.e.m.).

La figura 28 proporciona un conjunto de gráficos que demuestran que una señal de C2c2 generada según determinadas realizaciones de ejemplo puede detectar una diana de 20 pM en un sustrato de papel.

La figura 29 proporciona un gráfico que muestra que un inhibidor de ARNasa específico puede eliminar la señal de fondo en el papel.

La figura 30 es un conjunto de gráficos que muestran la detección mediante sistemas según determinadas realizaciones de ejemplo en sustratos de fibra de vidrio.

La figura 31 proporciona un conjunto de gráficos que proporcionan (A) un diagrama de la detección de ARN de Zika según determinadas realizaciones de ejemplo. Se empaquetaron en el lentivirus el ARN de Zika o los fragmentos homólogos de ARN de dengue seleccionados como diana por la detección colateral con C2c2. Los medios se cosechan después de 48 horas y se someten a lisis térmica, detección con C2c2 y RT-RPA. (B) La detección con RT-RAP-C2c2 puede detectar con gran sensibilidad las partículas lentivirales de Zika (n=4 réplicas técnicas, prueba de la t de Student bilateral; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.) (C) Un diagrama de la detección de ARN de Zika usando C2c2 secado por congelación en papel, según determinadas realizaciones de ejemplo. (D) El ensayo basado en papel puede detectar con gran sensibilidad las partículas lentivirales de Zika (n=4 réplicas técnicas, prueba de la t de Student bilateral; ****, $p < 0,0001$; **, $p < 0,01$, las barras representan la media \pm e.e.m.).

La figura 32 proporciona un conjunto de gráficos que demuestran (A) Un diagrama para la detección con C2c2 de ARN de Zika aislado de suero humano. El ARN de Zika en el suero se somete a transcripción inversa, a degradación por ARNasa H del ARN, a RPA del ADNc y a detección con C2c2. (B) C2c2 puede detectar con gran sensibilidad muestras de suero humano de Zika. Las concentraciones de ARN de Zika mostradas se verificaron mediante qPCR (n=4 réplicas técnicas, prueba de la t de Student bilateral; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.).

La figura 33 proporciona un conjunto de gráficos que demuestran (A) que el C2c2 secado por congelación puede detectar sensiblemente el ARNmc 1 en el intervalo femtomolar bajo. C2c2 puede detectar rápidamente una diana de ARNmc 1 de 200 pM en papel en forma líquida (B) o secada por congelación (C). La reacción puede detectar con sensibilidad fragmentos de ARN de Zika sintetizados en disolución (D) (n=3) y en forma secada por congelación (E) (n=3). (F) Curva cuantitativa para la detección de ADNc de Zika humano que muestra una correlación significativa entre la concentración de entrada y la fluorescencia detectada. (G) Detección con C2c2 de ARNmc 1 realizada en presencia de cantidades variables de suero humano (n=2 réplicas técnicas, a menos que se indique lo contrario; las barras representan la media \pm e.e.m.).

La figura 34 proporciona (A) un diagrama de la detección con C2c2 del gen ARNr 16S de genomas bacterianos usando un conjunto de cebadores universales V3 de RPA, y (B) la capacidad de lograr una detección sensible y específica de ADNc de *E. coli* o *P. aeruginosa* usando un ensayo realizado según determinadas realizaciones de ejemplo (n=4 réplicas técnicas, prueba de la t de Student bilateral; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.). *Ec*, *Escherichia coli*; *Kp*, *Klebsiella pneumoniae*; *Pa*, *Pseudomonas aeruginosa*; *Mt*, *Mycobacterium tuberculosis*; *Sa*, *Staphylococcus aureus*.

La figura 35 proporciona un conjunto de gráficos que demuestran (A) la detección de dos genes diferentes de resistencia a los carbapenémicos (KPC y NDM-1) a partir de cuatro aislados clínicos diferentes de *Klebsiella pneumoniae*, y (B) la detección de genes de resistencia a los carbapenémicos (parte A) se normaliza como una relación de señal entre los ensayos de ARNcr de KPC y NDM-1 (n=2 repeticiones técnicas, prueba de la t de Student bilateral; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.).

La figura 36 proporciona un conjunto de gráficos que demuestran que (A) C2c2 no es sensible a apareamientos erróneos individuales, pero puede distinguir entre las diferencias de un solo nucleótido en la diana cuando se carga con ARNcr con apareamientos erróneos adicionales. Se detectaron las dianas de ARNmc 1-3 con 11 ARNcr, con 10 espaciadores que contenían apareamientos erróneos sintéticos en diversas posiciones en el ARNcr. Los espaciadores con apareamientos erróneos no mostraron una reducción de la escisión de la diana 1, pero mostraron una escisión inhibida de las dianas 2 y 3 con apareamiento erróneo (SEQ. ID. NO. 146 a 159). (B) Diagrama que muestra el proceso de diseño racional de espaciadores específicos de una sola base con apareamientos erróneos sintéticos. Los apareamientos erróneos sintéticos se colocan en la proximidad del SNP o de la base de interés (SEQ. ID. NO. 160 a 164). (C) La detección altamente específica de los SNP de la cepa permite diferenciar las dianas de ARN de Zika africano frente al americano que difieren en un solo nucleótido usando la detección con C2c2 con ARNcr truncado (23 nucleótidos) (n=2 réplicas técnicas, prueba de la t de Student unilateral; *, $p < 0,05$; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.).

La figura 37 proporciona un conjunto de gráficos que demuestran: (A) Diagrama de las regiones diana de las cepas de Zika y las secuencias de ARNcr usadas para la detección (SEQ. ID. NO. 165 a 170). Los SNP en la diana se resaltan en rojo o azul y los apareamientos erróneos en la secuencia guía se colorean de rojo. (B) La detección altamente específica de los SNP de las cepas permite diferenciar las dianas de ARN de Zika africano frente al americano usando SHERLOCK (n=2 réplicas técnicas, prueba de la t de Student bilateral; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.) (SEQ. ID. NO. 171 a 176). (C) Diagrama de las regiones diana de las cepas del dengue y las secuencias de ARNcr usadas para la detección. Los SNP en la diana se resaltan en rojo o azul y los apareamientos erróneos sintéticos en la secuencia guía se colorean en rojo. (D) La detección altamente específica de los SNP de las cepas permite diferenciar las dianas de ARN de la cepa 1 del dengue frente a la cepa 3 usando SHERLOCK (n=2 réplicas técnicas, prueba de la t de Student bilateral; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.).

La figura 38 proporciona un conjunto de gráficos que muestran (A) el diagrama Circos que muestra la ubicación de los SNP humanos detectados con C2c2. (B) El ensayo realizado según determinadas realizaciones de ejemplo puede distinguir entre SNP humanos. SHERLOCK puede genotipar correctamente a cuatro individuos diferentes en cuatro sitios diferentes de SNP en el genoma humano. Los genotipos de cada individuo y las identidades de los ARNcr que detectan alelos se anotan debajo de cada gráfico (n=4 réplicas técnicas; prueba de la t de Student bilateral; *, $p < 0,05$; **, $p < 0,01$; ***, $p < 0,001$; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.). (C) Un diagrama del proceso para la detección de ADNlc (tal como la detección de ADN libre de células de mutaciones de cáncer) según determinadas realizaciones de ejemplo. (D) Secuencias de ARNcr de ejemplo para la detección de EGFR L858R y BRAF V600E. (SEQ. ID. NO. 177 a 182). Secuencias de dos loci genómicos sometidos a ensayo para detectar mutaciones cancerígenas en ADN libre de células. Se muestra la secuencia genómica diana con el SNP resaltado en azul y las secuencias de ARNcr de detección mutantes/silvestres con apareamientos erróneos sintéticos coloreados en rojo.

La figura 39 proporciona un conjunto de gráficos que demuestran que C2c2 puede detectar el alelo menor mutante en muestras de ADN libre de células de imitación del alelo menor EGFR L858R (C) o BRAF V600E (B). (n=4 réplicas técnicas, prueba de la t de Student bilateral; *, $p < 0,05$; **, $p < 0,01$; ***, $p < 0,001$; ****, $p < 0,0001$; las barras representan \pm e.e.m.)

La figura 40 proporciona un conjunto de gráficos que demuestran que (A) el ensayo puede distinguir entre genotipos en rs5082 (n=4 réplicas técnicas; *, $p < 0,05$; **, $p < 0,01$; ***, $p < 0,001$; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.). (B) el ensayo puede distinguir entre genotipos en rs601338 en ADNg directamente de saliva centrifugada, desnaturalizada y hervida (n=3 réplicas técnicas; *, $p < 0,05$; las barras representan la media \pm e.e.m.).

La figura 41 proporciona (A) un diagrama de una realización de ejemplo realizada en ADNmc 1 en el fondo de una diana que difiere de ADNmc 1 en sólo un único apareamiento erróneo. (B) El ensayo logra la detección de la especificidad de un solo nucleótido del ADNmc 1 en presencia de un fondo de apareamiento erróneo (diana que difiere en sólo un único apareamiento erróneo del ADNmc). Se combinaron diversas concentraciones de ADN diana con un exceso de fondo de ADN con un apareamiento erróneo y se detectaron mediante el ensayo.

La figura 42 es un gráfico que muestra un constructo de enmascaramiento con un colorante diferente Cy5 que también permite una detección eficaz.

La figura 43 es un diagrama de un ensayo basado en nanopartículas de oro colorimétricas. Las AuNP se agregan usando una combinación de ligadores de ADN y un puente de ARN. Al añadir la actividad ARNasa, el puente de ARNmc se escinde y las AuNP se liberan, provocando un cambio de color característico hacia el rojo.

La figura 44 es un gráfico que muestra la capacidad de detectar el cambio de color de las nanopartículas dispersas a 520 nm. Las nanopartículas se basaron en la realización de ejemplo mostrada en la figura 43 y se dispersaron mediante la adición de ARNasa A a concentraciones variables.

La figura 45 es un gráfico que muestra que la prueba colorimétrica de la ARNasa es cuantitativa.

La figura 46 es una imagen de una placa de micropocillos que muestra que el cambio de color en la nanopartícula dispersa es detectable visualmente.

La figura 47 es una imagen que demuestra que el cambio colorimétrico es visible en un sustrato de papel. La prueba se realizó durante 10 minutos a 37°C en la fibra de vidrio 934-AH.

La figura 48 es un diagrama de aptámeros de cambio de conformación según determinadas realizaciones de ejemplo para la detección de proteínas o pequeñas moléculas. El producto ligado (B) se usa como una diana completa para el efector que selecciona como diana ARN, que no puede detectar el producto de entrada no ligado (SEQ. ID. NO. 202 y 424).

La figura 49 es una imagen de un gel que muestra que el ligamiento basado en aptámeros puede crear sustratos detectables por RPA. Los aptámeros se incubaron con diversos niveles de trombina y luego se ligaron con la sonda. Los constructos ligados se usaron como moldes para una reacción de RPA de 3 minutos. La trombina 500 nM tiene niveles significativamente más altos de diana amplificada que el fondo.

La figura 50 muestra la secuencia de aminoácidos de los dominios HEPN de los ortólogos de C2c2 seleccionados (SEQ. ID. NO. 204-233).

Figura 51. La detección con Cas13a de ARN con amplificación RPA (SHERLOCK) puede detectar la diana de ARNmc a concentraciones de hasta ~2 aM, más sensible que Cas13a solo (n=4 réplicas técnicas; las barras representan la media ± e.e.m.).

Figura 52. La detección con Cas13a puede usarse para detectar patógenos virales y bacterianos. (A) Diagrama de la detección con SHERLOCK de ARN de ZIKV aislado de muestras clínicas humanas. (B) SHERLOCK puede detectar con gran sensibilidad muestras de suero (S) u orina (U) positivas para el ZIKV humano. Las concentraciones aproximadas de ARN de ZIKV mostradas se determinaron mediante qPCR. (n=4 réplicas técnicas, prueba de la t de Student bilateral; ****, p < 0,0001; las barras representan la media ± e.e.m.; n.d., no detectado).

Figura 53. Comparación de la detección de ARNmc 1 mediante NASBA con el conjunto de cebadores 2 (de la figura 11) y SHERLOCK. (n=2 réplicas técnicas; las barras representan la media ± e.e.m.)

Figura 54. Amplificación de ácidos nucleicos con RPA y SHERLOCK de reacción única.

(A) Cuantificación por PCR digital en gotas de ARNmc 1 para las diluciones usadas en la figura 1C. Las concentraciones ajustadas para las diluciones basadas en los resultados de la ddPCR se muestran encima de los gráficos de barras.

(B) Cuantificación por PCR digital en gotas de ADNmc 1 para las diluciones usadas en la figura 1D. Las concentraciones ajustadas para las diluciones basadas en los resultados de la ddPCR se muestran encima de los gráficos de barras.

(C) Las reacciones de detección con Cas13a, transcripción con T7 y de RPA son compatibles y logran la detección de moléculas individuales de ADN 2 cuando se incuban simultáneamente (n=3 réplicas técnicas, prueba de la t de Student bilateral; n.s., no significativo; **, p < 0,01; ****, p < 0,0001; las barras representan la media ± e.e.m.).

Figura 55. Comparación de SHERLOCK con otras herramientas sensibles de detección de ácidos nucleicos. (A) Análisis de detección de series de dilución de ADNmc 1 con PCR digital en gotas (n=4 réplicas técnicas, prueba de la t de Student bilateral; n.s., no significativo; *, p < 0,05; **, p < 0,01; ****, p < 0,0001; las líneas rojas representan la media, las barras representan la media ± e.e.m. No se muestran las muestras con una copia medida/μl por debajo de 10-1). (B) Análisis de detección de series de dilución de ADNmc 1 con PCR cuantitativa (n=16 réplicas técnicas, prueba de la t de Student bilateral; n.s., no significativo; **, p < 0,01; ****, p < 0,0001; las líneas rojas representan la media, las barras representan la media ± e.e.m. No se muestran las muestras con señal relativa inferior a 10-10). (C) Análisis de detección de series de dilución de ADNmc 1 con RPA con SYBR Green II (n=4 réplicas técnicas, prueba de la t de Student bilateral; *, p < 0,05; **, p < 0,01; las líneas rojas representan la media, las barras representan la media ± e.e.m. No se muestran las muestras con señal relativa inferior a 100). (D) Análisis de detección de series de dilución de ADNmc 1 con SHERLOCK (n=4 réplicas técnicas, prueba de la t de Student bilateral; **, p < 0,01; ****, p < 0,0001; las líneas rojas representan la media, las barras representan la media ± e.e.m. No se muestran las muestras con señal relativa inferior a 100). (E) Coeficiente de variación porcentual para una serie de diluciones de ADNmc 1 para cuatro tipos de métodos de detección. (F) Coeficiente de variación porcentual medio para las diluciones de ADNmc 1 de 6e2, 6e1, 6e0 y 6e-1 para cuatro tipos de métodos de detección (las barras representan la media ± e.e.m.).

Figura 56. Detección de la resistencia a los carbapenémicos en aislados bacterianos clínicos. Detección de dos genes diferentes de resistencia a los carbapenémicos (KPC y NDM-1) en cinco aislados clínicos de *Klebsiella pneumoniae* y un control de *E. coli* (n=4 réplicas técnicas, prueba de la t de Student bilateral; ****, p < 0,0001; las barras representan la media ± e.e.m.; n.d., no detectado).

Figura 57. Caracterización de la sensibilidad de LwCas13a a los espaciadores truncados y a los apareamientos erróneos individuales en la secuencia diana. (A) Secuencias de ARNcr con espaciador truncado (SEQ. ID. NO. 425-436) usadas en (B)-(G). También se muestran las secuencias de los ARNmc 1 y 2, que tienen una diferencia de un solo par de bases resaltada en rojo. Los ARNcr que contienen apareamientos erróneos sintéticos se muestran con las posiciones de los apareamientos erróneos coloreadas en rojo. (B) Actividad de escisión colateral en los ARNmc 1 y 2 para ARNcr con espaciador de 28 nt con apareamientos erróneos sintéticos en las posiciones 1-7 (n=4 réplicas técnicas; las barras representan la media ± e.e.m.). (C) Razones de especificidad de los ARNcr sometidos a prueba en (B). Las razones de especificidad se calculan como la razón de la escisión colateral del ARN específico (ARNmc 1) con respecto a la escisión colateral del ARN inespecífico (ARNmc 2). (n=4 réplicas técnicas; las barras representan la media ± e.e.m.) (D) Actividad de escisión colateral en los ARNmc 1 y 2 para ARNcr con espaciador de 23 nt con apareamientos erróneos sintéticos en las posiciones 1-7 (n=4 réplicas técnicas; las barras representan la media ± e.e.m.). (E) Razones de especificidad de los ARNcr sometidos a prueba en (D). Los coeficientes de especificidad se calculan como la razón de la escisión colateral del ARN específico (ARNmc 1) con respecto a la escisión colateral del ARN inespecífico (ARNmc 2) (n=4 réplicas técnicas; las barras representan la media ± e.e.m.). (F) Actividad de escisión colateral en los ARNmc 1 y 2 para ARNcr con espaciadores de 20 nt con apareamientos

erróneos sintéticos en las posiciones 1-7 (n=4 réplicas técnicas; las barras representan la media \pm e.e.m.). (G) Razones de especificidad de los ARNcr sometidos a prueba en (F). Las razones de especificidad se calculan como la razón de la escisión colateral del ARN específico (ARNmc 1) con respecto a la escisión colateral del ARN inespecífico (ARNmc 2) (n=4 réplicas técnicas; las barras representan la media \pm e.e.m.).

Figura 58. Identificación de la posición ideal de apareamiento erróneo sintético en relación con las mutaciones en la secuencia diana. (A) Secuencias para la evaluación de la posición ideal de apareamiento erróneo sintético para detectar una mutación entre el ARNmc 1 y el ARNmc (SEQ. ID. NO. 437 - 462). En cada una de las dianas, se someten a prueba los ARNcr con apareamientos erróneos sintéticos en las ubicaciones coloreadas (en rojo). Cada conjunto de ARNcr con apareamientos erróneos sintéticos se diseña de manera que la ubicación de la mutación se desplaza en posición respecto a la secuencia de espaciador. Los espaciadores se diseñan de manera que la mutación se evalúa en las posiciones 3, 4, 5 y 6 dentro del espaciador. (B) Actividad de escisión colateral en los ARNmc 1 y 2 para ARNcr con apareamientos erróneos sintéticos en distintas posiciones. Hay cuatro conjuntos de ARNcr con la mutación en la posición 3, 4, 5 o 6 dentro de la región dúplex espaciador:diana (n=4 réplicas técnicas; las barras representan la media \pm e.e.m.). (C) Razones de especificidad del ARNcr sometidos a prueba en (B). Las razones de especificidad se calculan como la razón de la escisión colateral del ARN específico (ARNmc 1) con respecto a la escisión colateral del ARN inespecífico (ARNmc 2) (n=4 réplicas técnicas; las barras representan la media \pm e.e.m.).

Figura 59. Genotipado con SHERLOCK en un locus adicional y genotipado directo a partir de saliva hervida. SHERLOCK puede distinguir entre genotipos en el sitio de SNP rs601338 en ADN genómico directamente de saliva centrifugada, desnaturalizada y hervida (n=4 réplicas técnicas, prueba de la t de Student bilateral; **, p < 0,01; ****, p < 0,001; las barras representan la media \pm e.e.m.).

Figura 60. Desarrollo de patrones sintéticos de genotipado para genotipar con precisión SNP humanos. (A) Genotipado con SHERLOCK en el sitio de SNP rs601338 para cada uno de los cuatro individuos en comparación con los patrones de genotipo amplificados por PCR (n=4 réplicas técnicas; las barras representan la media \pm e.e.m.). (B) Genotipado con SHERLOCK en el sitio de SNP rs4363657 para cada uno de los cuatro individuos en comparación con patrones de genotipo amplificados por PCR (n=4 réplicas técnicas; las barras representan la media \pm e.e.m.). (C) Mapas de calor de los valores de p calculados entre los resultados de SHERLOCK para cada individuo y los patrones sintéticos en el sitio de SNP rs601338. Se muestra un mapa de calor para cada uno de los ARNcr que detectan alelos. El mapa de color del mapa de calor está escalado de manera que la insignificancia (p > 0,05) es roja y la significación (p < 0,05) es azul (n=4 réplicas técnicas, ANOVA unifactorial). (D) Mapas de calor de los valores de p calculados entre los resultados de SHERLOCK para cada individuo y los patrones sintéticos en el sitio de SNP rs4363657. Se muestra un mapa de calor para cada uno de los ARNcr que detectan alelos. El mapa de colores del mapa de calor está escalado de manera que la insignificancia (p > 0,05) es roja y la significación (p < 0,05) es azul (n=4 réplicas técnicas, ANOVA unifactorial). (E) Una guía para entender los resultados del mapa de calor del valor de p del genotipado SHERLOCK. El genotipado puede designarse fácilmente eligiendo el alelo que corresponde a un valor de p > 0,05 entre el individuo y los patrones sintéticos alélicos. Los bloques rojos corresponden a diferencias no significativas entre el patrón sintético y el resultado SHERLOCK del individuo y, por tanto, a un resultado de genotipo positivo. Los bloques azules corresponden a diferencias significativas entre el patrón sintético y el resultado SHERLOCK del individuo y, por tanto, a un resultado de genotipo negativo.

Figura 61. Detección de ADNmc 1 como una pequeña fracción del objetivo de fondo no coincidente. Detección de SHERLOCK de una serie de diluciones de ADNmc 1 sobre un fondo de ADN genómico humano. Obsérvese que no debe haber ninguna similitud de secuencia entre la diana de ADNmc 1 detectada y el ADN genómico de fondo (n=2 réplicas técnicas; las barras representan la media \pm e.e.m.).

Figura 62. Muestras de orina (A) o suero (B) de pacientes con el virus de Zika se inactivaron por calor durante 5 minutos a 95°C (orina) o 65°C (suero). Se usó un microlitro de orina o suero inactivado como entrada para una reacción de RPA de 2 horas seguida de una reacción de detección con C2c2/Cas13a de 3 horas, según una realización de ejemplo. Las barras de error indican 1 DE basada en n=4 réplicas técnicas para la reacción de detección.

Figura 63. Las muestras de orina de pacientes con el virus de Zika se inactivaron por calor durante 5 minutos a 95°C. Se usó un microlitro de orina inactivada como entrada para una reacción de RPA de 30 minutos seguida de una reacción de detección con C2c2/Cas13 de 3 horas (A) o de 1 hora (B), según las realizaciones de ejemplo. Las barras de error indican 1 DE basada en n=4 réplicas técnicas para la reacción de detección.

Figura 64. Las muestras de orina de pacientes con el virus de Zika se inactivaron por calor durante 5 minutos a 95°C. Se usó un microlitro de orina inactivada como entrada para una reacción de RPA de 20 minutos seguida de una reacción de detección con C2c2/Cas13a de 1 hora. Se usó orina de ser humano sano como control negativo. Las barras de error indican 1 DE basada en n=4 réplicas técnicas o en la reacción de detección.

Figura 65. Las muestras de orina de pacientes con el virus de Zika se inactivaron por calor durante 5 minutos a 95°C. Se usó un microlitro de orina inactivada como entrada para una reacción de RPA de 20 minutos seguida de una reacción de detección con C2c2/Cas13a de 1 hora en presencia o ausencia de ARN guía. Los datos se

normalizan sustrayendo los valores medios de fluorescencia de las reacciones de detección sin guías de las reacciones de detección con guías. La orina de ser humano sano se usó como control negativo. Las barras de error indican 1 DE basada en n=4 réplicas técnicas para la reacción de detección.

5 La figura 66 muestra la detección de dos dianas específicas de la malaria con cuatro diseños diferentes de ARN guía, según las realizaciones de ejemplo (SEQ. ID. NO. 463-474).

La figura 67 proporciona gráficos que muestran las preferencias de edición de los diferentes ortólogos de Cas13b. Véase la leyenda en la tabla 3.

10 La figura 68 proporciona A) un diagrama de un ensayo múltiplex usando diferentes ortólogos de Cas13b con diferentes preferencias de edición, y B) datos que demuestran la viabilidad de un ensayo de este tipo usando Cas13b10 y Cas13b5.

La figura 69 proporciona gráficos que muestran el multiplexado dual con los ortólogos de Cas13b5 (*Prevotella sp.* MA2106) y Cas13b9 (*Prevotella intermedia*). Tanto las proteínas efectoras como las secuencias guía estaban contenidas en la misma reacción, lo que permitió el multiplexado dual en la misma reacción usando diferentes lecturas fluorescentes (poli U a 530 nm y poli A a 485 nm).

15 La figura 70 proporciona lo mismo que la figura 69 pero en este caso usando ortólogos de Cas13a (*Leptorichia wadei* LwaCas13a) y de Cas13b (*Prevotella sp.* MA2016, Cas13b5).

La figura 71 proporciona un método para el embaldosado de secuencias diana con múltiples secuencias guía con el fin de determinar la robustez de la orientación, según determinadas realizaciones de ejemplo (SEQ. ID. NO. 475 y 476).

20 La figura 72 proporciona geles de reacción en cadena híbrida (HCR) que muestran que las proteínas efectoras Cas13 pueden usarse para desbloquear un iniciador, por ejemplo, un iniciador incorporado en un constructo de enmascaramiento tal como se describe en el presente documento, para activar una reacción en cadena de hibridación.

25 La figura 73 proporciona datos que muestran la capacidad de detectar *Pseudomonas aeruginosa* en lisados complejos.

La figura 74 proporciona datos que muestran las preferencias iónicas de determinados ortólogos de Cas13 según determinadas realizaciones de ejemplo. Todas las concentraciones diana fueron 20 nM de entrada con concentraciones de iones de (1 mM y 10 mM).

30 La figura 75 proporciona datos que muestran que Cas13b12 tiene una preferencia de sulfato de zinc 1 mM para la escisión.

La figura 76 proporciona datos que demuestran que la optimización del tampón puede reforzar la relación señal/ruido de Cas13b5 en un indicador de poliA. El antiguo tampón comprende Tris-HCl 40 mM, NaCl 60 mM, MgCl₂ 6 mM, pH 7,3. El nuevo tampón comprende HEPES 20 mM pH 6,8, MgCl₂ 6 mM y NaCl 60 mM.

35 La figura 77 proporciona un diagrama de los sistemas Crispr de tipo VI-A/C y de los sistemas de tipo VI-B1 y B2, así como un árbol filogenético de ortólogos representativos de Cas13b.

La figura 78 proporciona la actividad de escisión relativa en diferentes nucleótidos de diversos ortólogos de Cas13b y en relación con un LwCas13a.

La figura 79 proporciona un gráfico que muestra la sensibilidad relativa de diversos ejemplos de ortólogos de Cas13.

40 La figura 80 proporciona un gráfico que muestra la capacidad de lograr niveles zepto molares (zM) de detección usando una realización de ejemplo.

La figura 81 proporciona un diagrama de un ensayo múltiplex usando ortólogos de Cas13 con diferentes preferencias de edición y constructos de enmascaramiento basados en poli N.

FIG 82 proporciona datos que muestran los resultados de los experimentos de optimización de cebadores y la detección de *Pseudomonas* a lo largo de una variedad de condiciones.

45 La figura 83 ilustra la caracterización bioquímica de la familia Cas13b de enzimas que seleccionan como diana ARN guiadas por ARN y sensibilidad aumentada y el SHERLOCK cuantitativo. A) Diagrama de los loci de CRISPR-Cas13 y de la estructura de ARNcr. B) Un mapa de calor de la preferencia de bases de 15 ortólogos de Cas13b que seleccionan como diana ARNmc 1 con sondas de sensor que consisten en un homopolímero hexámero de bases de A, C, G o U. C) Diagrama del examen de descubrimiento de la preferencia de motivos de corte y motivos preferidos de dos bases para LwaCas13a y PsmCas13b. Los valores representados en el mapa de calor son los recuentos de
50 cada dos bases en todos los motivos agotados. Los motivos se consideran agotados si el valor en $-\log_2$

(específico/inespecífico) es superior a 1,0 en la condición de LwaCas13a o a 0,5 en la condición de PsmCas13b. En el valor en $-\log_2$ (específico/inespecífico), específico e inespecífico denotan la frecuencia de un motivo en las condiciones específicas e inespecíficas, respectivamente. D) Preferencias de bases ortogonales de PsmCas13b y LwaCas13a que seleccionan como diana ARNmc 1 con una sonda de sensor o bien U₆ o bien A₆. E) Detección de SHERLOCK de moléculas individuales con LwaCas13a y PsmCas13b que seleccionan como diana la diana ARNmc de dengue. F) Detección de SHERLOCK de moléculas individuales con LwaCas13a y PsmCas13b en grandes volúmenes de reacción para seleccionar como diana, con sensibilidad aumentada, la diana ARNmc 1. G) Cuantificación del ADN sintético de *P. aeruginosa* a diversas concentraciones de cebadores de RPA. H) Correlación de la concentración de ADN sintético de *P. aeruginosa* con la fluorescencia detectada.

La figura 84 ilustra la multiplexación en muestra de SHERLOCK con enzimas Cas13 ortogonales. A) Diagrama de la multiplexación en muestra usando enzimas Cas13 ortogonales. B) Detección multiplexada en muestra de 20 nM de ARN sintético de Zika y dengue con actividad colateral de LwaCas13a y PsmCas13b. C) Detección multiplexada en muestra de RPA y colateral a concentraciones decrecientes de dianas sintéticas de termonucleasa de *S. aureus* y aciltransferasa de *P. aeruginosa* con LwaCas13a y PsmCas13b. D) Genotipado multiplexado con muestras humanas en rs601338 con LwaCas13a y CcaCas13b. E) Diagrama de la línea de tiempo terapéutica para la detección de los alelos de la enfermedad, la corrección con REPAIR y la evaluación de la corrección con REPAIR. F) Detección multiplexada en muestra de alelos APC de muestras sanas y que simulan la enfermedad con LwaCas13a y PsmCas13b. G) Cuantificación de la eficacia de la edición con REPAIR en la mutación APC seleccionada como diana. H) Detección multiplexada en muestra de alelos APC de muestras de direccionamiento y de no direccionamiento a REPAIR con LwaCas13a y PsmCas13b.

La figura 85 proporciona un árbol de 15 ortólogos de Cas13b purificados y evaluados para la actividad colateral *in vitro*. Se muestran el gen Cas13b (azul), el gen Csx27/Csx28 (rojo/amarillo) y la matriz de CRISPR (gris).

La figura 86 ilustra la purificación de proteínas de los ortólogos de Cas13. A) Cromatogramas de cromatografía de exclusión molecular para Cas13b, LwCas13a y LbaCas13a usados en este estudio. Se muestra la absorbancia UV medida (mUA) frente al volumen de elución (ml). B) Gel SDS-PAGE de los ortólogos de Cas13b purificados. Catorce ortólogos de Cas13b se cargan de izquierda a derecha. A la izquierda se muestra un marcador de tamaño molecular de proteínas. C) Gel final de SDS-PAGE de las diluciones de LbaCas13a (derecha) y de la valoración convencional de BSA (izquierda). Se muestran cinco diluciones de BSA y dos de LbaCas13.

La figura 87 muestra gráficos que ilustran la preferencia de bases de la escisión colateral de los ortólogos de Cas13b. A) Actividad de escisión de catorce ortólogos de Cas13b que seleccionan como diana ARNmc 1 usando un sensor homopolimérico de adenina de seis nucleótidos de longitud. B) Actividad de escisión de catorce ortólogos de Cas13b que seleccionan como diana ARNmc 1 usando un sensor homopolimérico de uridina de seis nucleótidos de longitud. C) Actividad de escisión de catorce ortólogos de Cas13b que seleccionan como diana ARNmc 1 usando un sensor homopolimérico de guanina de seis nucleótidos de longitud. D) Actividad de escisión de catorce ortólogos de Cas13b que seleccionan como diana ARNmc 1 usando un sensor homopolimérico de citidina de seis nucleótidos de longitud.

La figura 88 muestra el análisis del tamaño de la biblioteca de motivos aleatorios después de la escisión colateral de Cas13. Trazos del bioanalizador para muestras de bibliotecas tratadas con LwaCas13a-, PsmCas13b-, CcaCas13b- y ARNasa A que muestran los cambios en el tamaño de la biblioteca tras la actividad ARNasa. Los ortólogos de Cas13 seleccionan como diana ARNmc de dengue y escinden la biblioteca de motivos aleatorios debido a la escisión colateral. Los patrones de marcadores se muestran en el primer carril.

La figura 89 muestra una representación de diversos motivos después de la escisión por ARNasas. A) Diagramas de cajas que muestran la distribución de motivos de las razones de específico con respecto a inespecífico para LwaCas13a, PsmCas13b, CcaCas13b y ARNasa A en puntos de tiempo de 5 y 60 minutos. Las razones de ARNasa A se compararon con el promedio de las tres condiciones inespecíficas de Cas13. Las razones son también un promedio de dos réplicas de la reacción de escisión. B) Número de motivos enriquecidos para LwaCas13a, PsmCas13b, CcaCas13b y ARNasa A en el punto de tiempo de 60 minutos. El motivo de enriquecimiento se calculó como motivos por encima de umbrales en $-\log_2$ (específico/inespecífico) de 1 (LwaCas13a, CcaCas13b, y ARNasa A) o 0,5 (PsmCas13b). Un umbral de 1 corresponde a un agotamiento de al menos el 50%, mientras que un umbral de 0,5 corresponde a un agotamiento de al menos el 30%. C) Logos de secuencias generados a partir de motivos enriquecidos para LwaCas13a, PsmCas13b y CcaCas13b. LwaCas13a y CcaCas13b muestran una fuerte preferencia por las bases U, como cabría esperar, mientras que PsmCas13b muestra una preferencia única por las bases A en todo el motivo, lo que es coherente con las preferencias de actividad colateral de los homopolímeros. D) Mapa de calor que muestra las preferencias de motivos ortogonales de LwaCas13a, PsmCas13b y CcaCas13b. Los valores representados en el mapa de calor son el valor en $-\log_2$ (específico/inespecífico) de cada motivo mostrado. En el valor en $-\log_2$ (específico/inespecífico), específico e inespecífico denotan la frecuencia de un motivo en las condiciones específicas e inespecíficas, respectivamente.

La figura 90 muestra las preferencias de una sola base y de dos bases de las ARNasas determinadas por el examen de la biblioteca de motivos aleatorios. A) Mapas de calor que muestran las preferencias de una sola base para LwaCas13a, PsmCas13b, CcaCas13b y ARNasa A en el punto de tiempo de 60 minutos tal como se determinó por

el examen de escisión de la biblioteca de motivos aleatorios. Los valores representados en el mapa de calor son los recuentos de cada base en todos los motivos agotados. Los motivos se consideran agotados si el valor en $-\log_2$ (específico/inespecífico) es superior a 1,0 en las condiciones de LwaCas13a, CcaCas13b y ARNasa A o a 0,5 en la condición de PsmCas13b. En el valor en $-\log_2$ (específico/inespecífico), específico e inespecífico denotan la frecuencia de un motivo en las condiciones específicas e inespecíficas, respectivamente. B) Mapas de calor que muestran la preferencia de dos bases para CcaCas13b tal como se determinó por el examen de escisión de la biblioteca de motivos aleatorios. Los valores representados en el mapa de calor son los recuentos de cada dos bases en todos los motivos agotados. Los motivos se consideran agotados si el valor en $-\log_2$ (específico/inespecífico) es superior a 1,0 en las condiciones de LwaCas13a, CcaCas13b y ARNasa A o a 0,5 en la condición de PsmCas13b. En el valor en $-\log_2$ (específico/inespecífico), específico e inespecífico denotan la frecuencia de un motivo en las condiciones específicas e inespecíficas, respectivamente. C) Mapas de calor que muestran la preferencia de dos bases para la ARNasa A tal como se determinó en el examen de la biblioteca de motivos aleatorios. Los valores representados en el mapa de calor son los recuentos de cada dos bases en todos los motivos agotados. Los motivos se consideran agotados si el valor en $-\log_2$ (específico/inespecífico) es superior a 1,0 en las condiciones de LwaCas13a, CcaCas13b y ARNasa A o a 0,5 en la condición de PsmCas13b. En el valor en $-\log_2$ (específico/inespecífico), específico e inespecífico denotan la frecuencia de un motivo en las condiciones específicas e inespecíficas, respectivamente.

La figura 91 ilustra las preferencias de tres bases de las ARNasas determinadas por el examen de la biblioteca de motivos aleatorios. Los mapas de calor muestran las preferencias de tres bases para LwaCas13a, PsmCas13b, CcaCas13b y ARNasa A en el punto de tiempo de 60 minutos, tal como se determinó en el examen de escisión de la biblioteca de motivos aleatorios. Los valores representados en el mapa de calor son los recuentos de cada 3 bases en todos los motivos agotados. Los motivos se consideran agotados si el valor en $-\log_2$ (específico/inespecífico) es superior a 1,0 en las condiciones de LwaCas13a, CcaCas13b y ARNasa A o a 0,5 en la condición de PsmCas13b. En el valor en $-\log_2$ (específico/inespecífico), específico e inespecífico denotan la frecuencia de un motivo en las condiciones específicas e inespecíficas, respectivamente.

La figura 92 ilustra las preferencias de cuatro bases de las ARNasas determinadas por el examen de la biblioteca de motivos aleatorios. Los mapas de calor muestran las preferencias de cuatro bases para LwaCas13a, PsmCas13b, CcaCas13b y ARNasa A en el punto de tiempo de 60 minutos, tal como se determinó por el examen de escisión de la biblioteca de motivos aleatorios. Los valores representados en el mapa de calor son los recuentos de cada 4 bases en todos los motivos agotados. Los motivos se consideran agotados si el valor en $-\log_2$ (específico/inespecífico) es superior a 1,0 en las condiciones de LwaCas13a, CcaCas13b y ARNasa A o a 0,5 en la condición de PsmCas13b. En el valor en $-\log_2$ (específico/inespecífico), específico e inespecífico denotan la frecuencia de un motivo en las condiciones específicas e inespecíficas, respectivamente.

La figura 93 muestra los resultados de las pruebas de las preferencias de escisión de bases de los ortólogos de Cas13 con la escisión *in vitro* de sustratos con poli-X. A) Escisión *in vitro* con LwaCas13a de dianas con poli-U, C, G y A incubadas con y sin ARNcr. B) Escisión *in vitro* con CcaCas13b de dianas con poli-U, C, G y A incubadas con y sin ARNcr. C) Escisión *in vitro* con PsmCas13b de dianas con poli-U, C, G y A incubadas con y sin ARNcr.

La figura 94 muestra los resultados de la optimización del tampón de la actividad de escisión de PsmCas13b. A) Se someten a prueba varios tampones para determinar su efecto sobre la actividad colateral de PsmCas13b tras seleccionar como diana ARNmc 1. B) El tampón optimizado se compara con el tampón original a diferentes concentraciones del complejo PsmCas13b-ARNcr.

La figura 95 ilustra la preferencia de iones de los ortólogos de Cas13 por la escisión colateral. A) Actividad de escisión de PsmCas13b con un sensor de poli U fluorescente para cationes divalentes Ca, Co, Cu, Mg, Mn, Ni y Zn. PsmCas13b se incubaba con un ARNcr que selecciona como diana un ARNmc sintético de dengue. B) Actividad de escisión de PsmCas13b con un sensor de poli A fluorescente para cationes divalentes Ca, Co, Cu, Mg, Mn, Ni y Zn. PsmCas13b se incubaba con un ARNcr que selecciona como diana un ARNmc sintético de dengue. C) Actividad de escisión de Pin2Cas13b con un sensor de poli U fluorescente para cationes divalentes Ca, Co, Cu, Mg, Mn, Ni y Zn. Pin2Cas13b se incubaba con un ARNcr que selecciona como diana un ARNmc sintético de dengue. D) Actividad de escisión de Pin2Cas13b con un sensor de poli A fluorescente para cationes divalentes Ca, Co, Cu, Mg, Mn, Ni y Zn. Pin2Cas13b se incubaba con un ARNcr que selecciona como diana un ARNmc sintético de dengue. E) Actividad de escisión de CcaCas13b con un sensor de poli U fluorescente para cationes divalentes Ca, Co, Cu, Mg, Mn, Ni y Zn. CcaCas13b se incubaba con un ARNcr que selecciona como diana un ARNmc sintético de dengue. F) Actividad de escisión de CcaCas13b con un sensor de poli A fluorescente para cationes divalentes Ca, Co, Cu, Mg, Mn, Ni y Zn. CcaCas13b se incubaba con un ARNcr que selecciona como diana un ARNmc sintético de dengue.

La figura 96 muestra la comparación de la actividad de escisión de los ortólogos de Cas13 con preferencia de escisión de adenina. A) Actividad de escisión de PsmCas13b y LbaCas13a incubadas con los respectivos ARNcr que seleccionan como diana una diana sintética de Zika a diferentes concentraciones ($n=4$ réplicas técnicas, prueba de la t de Student bilateral; n.s., no significativo; *, $p < 0,05$; **, $p < 0,01$; ***, $p < 0,001$; ****, $p < 0,0001$; las barras representan la media \pm e.e.m.). B) Actividad de escisión de PsmCas13b y LbaCas13a incubadas con los respectivos ARNcr que seleccionan como diana una diana sintética del dengue a diferentes concentraciones ($n=4$ réplicas técnicas, prueba de la t de Student bilateral; n.s., no significativo; *, $p < 0,05$; **, $p < 0,01$; ***, $p < 0,001$; ****, $p < 0,0001$).

$p < 0,0001$; las barras representan la media \pm e.e.m.).

La figura 97 ilustra la detección attomolar de la diana 4 del ARNmc de Zika con SHERLOCK con LwaCas13a y PsmCas13b. A) Detección con SHERLOCK del ARNmc de Zika a diferentes concentraciones con LwaCas13a y el sensor de poli U. B) Detección con SHERLOCK del ARNmc de Zika a diferentes concentraciones con PsmCas13b y el sensor de poli A.

La figura 98 ilustra la detección attomolar del ARNmc del dengue con SHERLOCK a diferentes concentraciones de CcaCas13b.

La figura 99. Pruebas de reprogramación del ortólogo de Cas13 con ARNcr que embaldosan el ARNmc 1. A) Actividad de escisión de LwaCas13a y CcaCas13b con ARNcr embaldosados a lo largo de de ARNmc1. B) Actividad de escisión de PsmCas13b con ARNcr embaldosados a lo largo de de ARNmc1.

La figura 100 muestra el efecto de la longitud del espaciador del ARNcr en la escisión del ortólogo de Cas13. A) Actividad de escisión de PsmCas13b con ARNcr que seleccionan como diana ARNmc1 de longitudes de espaciador variables. B) Actividad de escisión de CcaCas13b con ARNcr que seleccionan como diana ARNmc1 de longitudes de espaciador variables.

La figura 101 ilustra la optimización de la concentración del cebador para SHERLOCK cuantitativo. A) Curvas cinéticas de SHERLOCK de LwaCas13a incubada con dianas de ARN de Zika de diferente concentración y un ARNcr complementario a una concentración de cebador de RPA de 480 nM. B) Curvas cinéticas de SHERLOCK de LwaCas13a incubada con dianas de ARN de Zika de diferente concentración y un ARNcr complementario a una concentración de cebador de RPA de 240 nM. C) Curvas cinéticas de SHERLOCK de LwaCas13a incubada con dianas de ARN de Zika de diferente concentración y un ARNcr complementario a una concentración de cebador de RPA de 120 nM. D) Curvas cinéticas de SHERLOCK de LwaCas13a incubada con dianas de ARN de Zika de diferente concentración y un ARNcr complementario a una concentración de cebador de RPA de 24nM. E) Detección de SHERLOCK de ARN de Zika de diferentes concentraciones con cuatro concentraciones diferentes de cebadores de RPA: 480 nM, 240 nM, 120 nM, 60 nM y 24 nM. F) La correlación media R^2 entre la fluorescencia sustraída de fondo de SHERLOCK y la concentración de ARN diana de Zika a diferentes concentraciones de cebadores de RPA. G) Detección cuantitativa de SHERLOCK de las dianas de ARN de Zika a diferentes concentraciones en una serie de dilución de 10 veces (puntos negros) y una serie de dilución de 2 veces (puntos rojos). Se usó una concentración de cebador de RPA de 120 nM.

La figura 102 ilustra la detección multiplexada de dianas de Zika y dengue. A) Detección multiplexada en dos colores usando LwaCas13a que selecciona como diana una diana de ARNmc de Zika y PsmCas13b que selecciona como diana una diana de ARNmc del dengue. Ambas dianas están en 20 nM de entrada. Todos los datos mostrados representan un punto de tiempo de reacción de 180 minutos. B) Detección multiplexada en dos colores usando LwaCas13a que selecciona como diana una diana de ARNmc de Zika y PsmCas13b que selecciona como diana una diana de ARNmc del dengue. Ambas dianas están a 200 pM de entrada. C) Detección multiplexada en muestra de 20 pM de ARN sintético de Zika y dengue con actividad colateral de CcaCas13a y PsmCas13b.

La figura 103 ilustra la detección de ARN multiplexada en muestra de ARNmc de Zika y dengue. Detección multiplexada en muestra de RPA y colateral a concentraciones decrecientes de dianas sintéticas de Zika y dengue con PsmCas13b y CcaCas13b.

La figura 104 ilustra la detección teranóstica no multiplexada de mutaciones y la edición con REPAIR. A) Detección de alelos APC de muestras sanas y simuladas de enfermedad con LwaCas13a. B) Detección con LwaCas13a de la corrección de la edición en los alelos APC de muestras de direccionamiento y de no direccionamiento a REPAIR.

La figura 105 ilustra la detección colorimétrica de la actividad ARNasa con agregación de nanopartículas de oro. A) Diagrama de la lectura colorimétrica basada en nanopartículas de oro para la actividad ARNasa. En ausencia de actividad ARNasa, los ligadores de ARN agregan las nanopartículas de oro, lo que provoca la pérdida de color rojo. B) La escisión de los ligadores de ARN libera las nanopartículas y da como resultado un cambio de color rojo. C) Imagen de los indicadores colorimétricos después de 120 minutos de digestión con ARNasa a diversas unidades de ARNasa A. D) Cinética a 520 nM de absorbancia de los indicadores colorimétricos de AuNP con la digestión a diversas concentraciones unitarias de ARNasa A. E) Absorbancia a 520 nM de los indicadores colorimétricos de AuNP después de 120 minutos de digestión a diversas concentraciones unitarias de ARNasa A. F) Tiempo hasta A_{520} semimáxima de los indicadores colorimétricos de AuNP con la digestión a diversas concentraciones unitarias de ARNasa A.

La figura 106 ilustra la detección cuantitativa del gen CP4-EPSPS del ADN genómico de sojas. A) La correlación media R^2 de la fluorescencia sustraída de fondo SHERLOCK y el porcentaje de sojas de CP4-EPSPS en diferentes puntos de tiempo de la detección. El porcentaje de sojas representa la cantidad de sojas "round-up ready" en una mezcla de sojas "round-up ready" y silvestres. El gen CP4-EPSPS sólo está presente en las sojas round-up ready. B) Detección de SHERLOCK del gen de resistencia CP4-EPSPS en diferentes porcentajes de sojas que muestran la naturaleza cuantitativa de la detección de SHERLOCK a los 30 minutos de incubación. C) Detección de SHERLOCK del gen de la lectina en diferentes porcentajes de sojas. El porcentaje de sojas representa la cantidad de sojas

“round-up ready” en una mezcla de sojas “round-up ready” y silvestres. El gen de la lectina está presente en ambos tipos de sojas y, por tanto, no muestra ninguna correlación con el porcentaje de sojas round-up ready.

Descripción detallada de las realizaciones de ejemplo

Definiciones generales

- 5 A menos que se defina lo contrario, los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta divulgación. Pueden encontrarse definiciones de términos y técnicas comunes en biología molecular en Molecular Cloning: A Laboratory Manual, 2ª edición (1989) (Sambrook, Fritsch, y Maniatis); Molecular Cloning: A Laboratory Manual, 4ª edición (2012) (Green y Sambrook); Current Protocols in Molecular Biology (1987) (F.M. Ausubel *et al.* eds.); las series Methods in Enzymology (Academic Press, Inc.): PCR 2: A Practical Approach (1995) (M.J. MacPherson, B.D. Hames, y G.R. Taylor eds.); Antibodies, A Laboratory Manual (1988) (Harlow y Lane, eds.); Antibodies A Laboratory Manual, 2ª edición 2013 (E.A. Greenfield ed.); Animal Cell Culture (1987) (R.I. Freshney, ed.); Benjamin Lewin, Genes IX, publicado por Jones y Bartlet, 2008 (ISBN 0763752223); Kendrew *et al.* (eds.), The Encyclopedia of Molecular Biology, publicada por Blackwell Science Ltd., 1994 (ISBN 0632021829); Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, publicado por VCH Publishers, Inc., 1995 (ISBN 9780471185710); Singleton *et al.*, Dictionary of Microbiology and Molecular Biology 2ª ed., J. Wiley & Sons (New York, N.Y. 1994), March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4ª ed., John Wiley & Sons (New York, N.Y. 1992); y Marten H. Hofker y Jan van Deursen, Transgenic Mouse Methods and Protocols, 2ª edición (2011).

- 20 Tal como se usa en el presente documento, las formas singulares “un”, “una” y “el/la” incluyen los referentes tanto en singular como en plural a menos que el contexto dicte claramente lo contrario.

El término “opcional” u “opcionalmente” significa que el acontecimiento, circunstancia o sustituyente descrito posterior puede producirse o no, y que la descripción incluye casos en los que el acontecimiento o la circunstancia se produce y casos en los que no se produce.

- 25 La enumeración de intervalos numéricos por los puntos finales incluye todos los números y fracciones integrados dentro de los intervalos respectivos, así como los puntos finales enumerados

- Se pretende que los términos “aproximadamente” o “alrededor de” tal como se usa en el presente documento cuando se hace referencia a un valor medible tal como un parámetro, una cantidad, una duración temporal, y similares, abarquen variaciones de y a partir del valor especificado, tal como variaciones de +/-10% o menos, +/-5% o menos, +/-1% o menos, y +/-0,1% o menos de y a partir del valor especificado, siempre que tales variaciones sean apropiadas para realizar en la invención dada a conocer. Debe entenderse que el valor al que se refiere el modificador “aproximadamente” o “alrededor de” en sí mismo también se da a conocer específica y preferiblemente.

“C2c2” se denomina ahora “Cas13a”, y los términos se usan de manera intercambiable en el presente documento a menos que se indique lo contrario.

35 Resumen

- Los sistemas inmunitarios adaptativos microbianos de repeticiones palindrómicas cortas regularmente espaciadas (CRISPR) y asociados a CRISPR (CRISPR-Cas) contienen endonucleasas programables, tales como Cas9 y Cpf1 (Shmakov *et al.*, 2017; Zetsche *et al.*, 2015). Aunque tanto Cas9 como Cpf1 seleccionan como diana ADN, recientemente se han descubierto (Shmakov *et al.*, 2015) y caracterizado (Abudayyeh *et al.*, 2016; Smargon *et al.*, 2017) ARNasas guiadas por ARN efectoras únicas, incluyendo C2c2, lo que proporciona una plataforma para la detección específica de ARN. Las ARNasas guiadas por ARN pueden ser fácil y convenientemente reprogramadas usando ARN de CRISPR (ARNcr) para escindir ARN diana. A diferencia de las endonucleasas de ADN Cas9 y Cpf1, que sólo escinden su diana de ADN, las ARNasas guiadas por ARN, tales como C2c2, permanecen activas después de la escisión de su diana de ARN, lo que lleva a la escisión “colateral” de los ARN no seleccionados como diana en la proximidad (Abudayyeh *et al.*, 2016). Esta actividad de escisión colateral de ARN programada por ARNcr presenta la oportunidad de usar las ARNasas guiadas por ARN para detectar la presencia de un ARN específico desencadenando la muerte celular programada *in vivo* o la degradación de ARN inespecífica *in vitro* que puede servir como lectura (Abudayyeh *et al.*, 2016; East-Seletsky *et al.*, 2016).

- Las realizaciones dadas a conocer en el presente documento utilizaron efectores que seleccionan como diana ARN para proporcionar un diagnóstico robusto basado en CRISPR con sensibilidad atómolar. Las realizaciones dadas a conocer en el presente documento pueden detectar tanto el ADN como el ARN con niveles comparables de sensibilidad y pueden diferenciar las dianas de las no dianas basándose en diferencias de un solo par de bases. Además, las realizaciones dadas a conocer en el presente documento pueden prepararse en formato secado por congelación para una distribución conveniente y aplicaciones de diagnóstico inmediato (POC). Tales realizaciones son útiles en múltiples escenarios de la salud humana, incluyendo, por ejemplo, la detección viral, la tipificación de cepas bacterianas, el genotipado sensible y la detección de ADN libre de células asociado a la enfermedad. Para facilitar la referencia, las realizaciones dadas a conocer en el presente documento también pueden denominarse

SHERLOCK (desbloqueo del indicador enzimático específico de alta sensibilidad).

En un aspecto, las realizaciones dadas a conocer en el presente documento se refieren a un sistema de detección de ácidos nucleicos que comprende un sistema CRISPR, uno o más ARN guía diseñados para unirse a las moléculas diana correspondientes, un constructo de enmascaramiento y reactivos de amplificación para amplificar las moléculas de ácido nucleico diana en una muestra. En determinadas realizaciones de ejemplo, el sistema puede comprender además uno o más aptámeros de detección. Los uno o más aptámeros de detección pueden comprender un sitio de ARN polimerasa o un sitio de unión de cebadores. Los uno o más aptámeros de detección se unen específicamente a uno o más polipéptidos diana y están configurados de tal manera que el sitio de ARN polimerasa o el sitio de unión de cebadores se exponen sólo al unirse el aptámero de detección a un péptido diana. La exposición del sitio de ARN polimerasa facilita la generación de un oligonucleótido de ARN desencadenante usando la secuencia del aptámero como molde. Por consiguiente, en tales realizaciones, el uno o más ARN guía está configurado para unirse a un ARN desencadenante.

En otro aspecto, las realizaciones dadas a conocer en el presente documento se refieren a un dispositivo de diagnóstico que comprende una pluralidad de volúmenes discretos individuales. Cada volumen discreto individual comprende una proteína efectora CRISPR, uno o más ARN guía diseñados para unirse a una molécula diana correspondiente, y un constructo de enmascaramiento. En determinadas realizaciones de ejemplo, los reactivos de amplificación de ARN pueden cargarse previamente en los volúmenes discretos individuales o añadirse a los volúmenes discretos individuales simultáneamente o después de la adición de una muestra a cada volumen discreto individual. El dispositivo puede ser un dispositivo basado en microfluidos, un dispositivo portátil o un dispositivo que comprende un sustrato de material flexible en el que se definen los volúmenes discretos individuales.

En otro aspecto, las realizaciones dadas a conocer en el presente documento se refieren a un método para detectar ácidos nucleicos diana en una muestra que comprende la distribución de una muestra o conjunto de muestras en un conjunto de volúmenes discretos individuales, comprendiendo cada volumen discreto individual una proteína efectora CRISPR, uno o más ARN guía diseñados para unirse a uno de los oligonucleótidos diana, y un constructo de enmascaramiento. Luego el conjunto de muestras se mantiene en condiciones suficientes para permitir la unión de uno o más ARN guía a una o más moléculas diana. La unión de uno o más ARN guía a un ácido nucleico diana activa a su vez la proteína efectora CRISPR. Una vez activada, la proteína efectora CRISPR desactiva el constructo de enmascaramiento, por ejemplo, escindiendo el constructo de enmascaramiento de manera que se desenmascara, libera o genera una señal positiva detectable. La detección de la señal positiva detectable en un volumen individual discreto indica la presencia de las moléculas diana.

También se describe en el presente documento, pero no según la invención, un método para detectar polipéptidos. El método para detectar polipéptidos es similar al método para detectar ácidos nucleicos diana descrito anteriormente. Sin embargo, también se incluye un aptámero de detección de péptidos. Los aptámeros de detección de péptidos actúan tal como se describió anteriormente y facilitan la generación de un oligonucleótido desencadenante al unirse a un polipéptido diana. Los ARN guía están diseñados para reconocer los oligonucleótidos desencadenantes, activando así la proteína efectora CRISPR. La desactivación del constructo de enmascaramiento por parte de la proteína efectora CRISPR activada conduce al desenmascaramiento, la liberación o la generación de una señal positiva detectable.

Proteínas efectoras de CRISPR

En general, un sistema CRISPR-Cas o CRISPR, tal como se usa en el presente documento y en documentos tales como WO 2014/093622 (documento PCT/US2013/074667), se refiere en conjunto a los transcritos y otros elementos que intervienen en la expresión o dirigen la actividad de los genes asociados a CRISPR ("Cas"), incluyendo las secuencias que codifican para un gen Cas, una secuencia tracr (CRISPR transactivante) (por ejemplo, ARNtracr o un ARNtracr parcial activo), una secuencia tracr-mate (que abarca una "repetición directa" y una repetición directa parcial procesada por ARNtracr en el contexto de un sistema CRISPR endógeno), una secuencia guía (también denominada "espaciador" en el contexto de un sistema CRISPR endógeno), o "ARN" tal como se usa este término en el presente documento (por ejemplo, ARN para guiar a Cas, tal como Cas9, por ejemplo, ARN de CRISPR y ARN transactivante (tracr) o un ARN guía único (ARNgu) (ARN quimérico)) u otras secuencias y transcritos de un locus de CRISPR. En general, un sistema CRISPR se caracteriza por tener elementos que promueven la formación de un complejo de CRISPR en el sitio de una secuencia diana (también denominado protoespaciador en el contexto de un sistema CRISPR endógeno). Cuando la proteína CRISPR es una proteína C2c2, no se requiere un ARNtracr. C2c2 se ha descrito en Abudayyeh *et al.* (2016) "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector"; Science; DOI: 10.1126/science.aaf5573; y Shmakov *et al.* (2015) "Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems", Molecular Cell, DOI: dx.doi.org/10.1016/j.molcel.2015.10.008. Cas13b se ha descrito en Smargon *et al.* (2017) "Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNases Differentially Regulated by Accessory Proteins Csx27 and Csx28", Molecular Cell. 65, 1-13; dx.doi.org/10.1016/j.molcel.2016.12.023.

En determinadas realizaciones, un motivo adyacente al protoespaciador (PAM) o un motivo similar al PAM dirige la unión del complejo de la proteína efectora tal como se da a conocer en el presente documento al locus diana de interés. En algunas realizaciones, el PAM puede ser un PAM en 5' (es decir, situado en el sentido de 5' del extremo

5' del protoespaciador). En otras realizaciones, el PAM puede ser un PAM en 3' (es decir, situado en el sentido de 3' del extremo 5' del protoespaciador). El término "PAM" puede usarse de manera intercambiable con el término "PFS" o "sitio de flanqueo del protoespaciador" o "secuencia de flanqueo del protoespaciador".

En una realización preferida, la proteína efectora CRISPR puede reconocer un PAM en 3'. En determinadas realizaciones, la proteína efectora CRISPR puede reconocer un PAM en 3' que es H en 5', en el que H es A, C o U. En determinadas realizaciones, la proteína efectora puede ser C2c2p de *Leptotrichia shahii*, más preferiblemente C2c2 de *Leptotrichia shahii* DSM 19757, y el PAM en 3' es un H en 5'.

En el contexto de la formación de un complejo de CRISPR, "secuencia diana" se refiere a una secuencia con respecto a la cual una secuencia guía esté diseñada para tener complementariedad, donde la hibridación entre una secuencia diana y una secuencia guía promueve la formación de un complejo de CRISPR. Una secuencia diana puede comprender polinucleótidos de ARN. El término "ARN diana" se refiere a un polinucleótido de ARN que es o comprende la secuencia diana. En otras palabras, el ARN diana puede ser un polinucleótido de ARN o una parte de un polinucleótido de ARN con el que una parte del ARNg, es decir, la secuencia guía, está diseñada para tener complementariedad y al que debe dirigirse la función efectora mediada por el complejo que comprende la proteína efectora CRISPR y un ARNg. En algunas realizaciones, una secuencia diana se encuentra en el núcleo o en el citoplasma de una célula.

La molécula de ácido nucleico que codifica para una proteína efectora CRISPR, en particular C2c2, es ventajosamente una proteína efectora CRISPR con codones optimizados. Un ejemplo de secuencia con codones optimizados es, en este caso, una secuencia optimizada para la expresión en eucariotas, por ejemplo, en humanos (es decir, optimizada para la expresión en humanos), o para otro eucariota, animal o mamífero, tal como se comenta en el presente documento; véase, por ejemplo, la secuencia con codones optimizados de SaCas9 humana en el documento WO 2014/093622 (documento PCT/US2013/074667). Aunque se prefiere esto, se apreciará que son posibles otros ejemplos y se conoce la optimización de codones para una especie huésped distinta de la humana, o para la optimización de codones para órganos específicos. En algunas realizaciones, una secuencia de codificación de enzimas que codifica para una proteína efectora CRISPR es un codón optimizado para la expresión en células particulares, tales como células eucariotas. Las células eucariotas pueden ser aquellas de o derivadas de un organismo particular, tal como una planta o un mamífero, incluyendo pero no limitándose al humano, o al eucariota no humano o animal o mamífero tal como se comenta, por ejemplo, ratón, rata, conejo, perro, ganado, o mamífero o primate no humano. En algunas realizaciones, pueden excluirse los procesos para modificar la identidad genética de la línea germinal de los seres humanos y/o los procesos para modificar la identidad genética de los animales que puedan provocarles sufrimiento sin ningún beneficio médico sustancial para el hombre o el animal, y también los animales resultantes de tales procesos. En general, la optimización de codones se refiere a un proceso de modificación de una secuencia de ácido nucleico para mejorar su expresión en las células huésped de interés mediante la sustitución de al menos un codón (por ejemplo, alrededor de 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 o más codones) de la secuencia nativa por codones que se usan con mayor frecuencia o más frecuentemente en los genes de esa célula huésped, manteniendo al mismo tiempo la secuencia nativa de aminoácidos. Diversas especies muestran un sesgo particular para determinados codones de un aminoácido en particular. El sesgo de los codones (las diferencias en el uso de los codones entre los organismos) a menudo se correlaciona con la eficiencia de la traducción del ARN mensajero (ARNm), que a su vez se cree que depende, entre otras cosas, de las propiedades de los codones que se traducen y de la disponibilidad de determinadas moléculas de ARN de transferencia (ARNt). El predominio de determinados ARNt en una célula suele ser un reflejo de los codones usados más frecuentemente en la síntesis de péptidos. Por consiguiente, los genes pueden adaptarse para una expresión genética óptima en un organismo determinado, basándose en la optimización de codones. Las tablas de uso de codones están fácilmente disponibles, por ejemplo, en la "Base de datos de uso de codones" disponible en kazusa.or.jp/codon/ y estas tablas pueden adaptarse de varias maneras. Véase Nakamura, Y., *et al.* "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28:292 (2000). También se dispone de algoritmos informáticos para la optimización de codones de una secuencia particular para la expresión en una célula huésped particular, tal como Gene Forge (Aptagen; Jacobus, PA). En algunas realizaciones, uno o más codones (por ejemplo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, o más, o todos los codones) en una secuencia que codifica para una Cas corresponden al codón más frecuentemente usado para un aminoácido particular.

En determinadas realizaciones, los métodos tal como se describen en el presente documento pueden comprender la provisión de una célula transgénica con Cas, en particular una célula transgénica con C2c2, en la que se proporcionan o introducen uno o más ácidos nucleicos que codifican para uno o más ARN guía operablemente conectados en la célula con un elemento regulador que comprende un promotor de uno o más genes de interés. Tal como se usa en el presente documento, el término "célula transgénica con Cas" se refiere a una célula, tal como una célula eucariota, en la que se ha integrado genómicamente un gen Cas. La naturaleza, el tipo o el origen de la célula no son particularmente limitativos según la presente invención. Además, la forma en que se introduce el transgén de Cas en la célula puede variar y puede ser cualquier método conocido en la técnica. En determinadas realizaciones, la célula transgénica con Cas se obtiene introduciendo el transgén Cas en una célula aislada. En otras realizaciones, la célula transgénica con Cas se obtiene aislando células de un organismo transgénico con Cas. A modo de ejemplo, y sin limitación, la célula transgénica con Cas a la que se hace referencia en el presente documento puede derivarse de un eucariota transgénico con Cas, tal como un eucariota con inserción de Cas. Se hace referencia al documento WO 2014/093622 (documento PCT/US13/74667). Los métodos de la publicación de patente estadounidense n.º

20120017290 y 201165198 asignados a Sangamo BioSciences, Inc. que se refieren a la selección como diana del locus Rosa pueden modificarse para utilizar el sistema CRISPR Cas de la presente invención. Los métodos de la publicación de patente estadounidense n.º 20130236946 asignada a Collectis que se refieren a la selección como diana del locus Rosa también pueden modificarse para utilizar el sistema CRISPR Cas de la presente invención. A modo de ejemplo adicional se hace referencia a Platt *et al.* (Cell; 159(2):440-455 (2014)), que describen un ratón con inserción de Cas9. El transgén Cas puede incluir además un casete Lox-Stop-polyA-Lox(LSL) que hace que la expresión de Cas sea inducible por la recombinasa Cre. Alternativamente, la célula transgénica con Cas puede obtenerse introduciendo el transgén Cas en una célula aislada. Los sistemas de administración de transgenes son bien conocidos en la técnica. Por ejemplo, el transgén Cas puede administrarse, por ejemplo, en una célula eucariota mediante un vector (por ejemplo, AAV, adenovirus, lentivirus) y/o administración de partículas y/o nanopartículas, tal como también se describe en el presente documento.

El experto entenderá que la célula, tal como la célula transgénica con Cas, a la que se hace referencia en el presente documento, puede comprender otras alteraciones genómicas además de tener un gen Cas integrado o las mutaciones que surgen de la acción específica de la secuencia de Cas cuando se complejiza con el ARN que puede guiar a Cas a un locus diana.

En determinados aspectos, la invención implica vectores, por ejemplo, para administrar o introducir en una célula Cas y/o ARN que puede guiar Cas a un locus diana (es decir, ARN guía), pero también para propagar estos componentes (por ejemplo, en células procariotas). Un "vector", tal como se usa el presente documento, es una herramienta que permite o facilita la transferencia de una entidad de un entorno a otro. Es un replicón, tal como un plásmido, un fago o un cósmido, en el que puede insertarse otro segmento de ADN para provocar la replicación del segmento insertado. Generalmente, un vector puede replicarse cuando está asociado con los elementos de control adecuados. En general, el término "vector" se refiere a una molécula de ácido nucleico que puede transportar otro ácido nucleico al que se ha unido. Los vectores incluyen, pero no se limitan a, moléculas de ácido nucleico que son monocatenarias, bicatenarias o parcialmente bicatenarias; moléculas de ácido nucleico que comprenden uno o más extremos libres, sin extremos libres (por ejemplo, circulares); moléculas de ácido nucleico que comprenden ADN, ARN o ambos; y otras variedades de polinucleótidos conocidas en la técnica. Un tipo de vector es un "plásmido", que se refiere a un bucle circular de ADN bicatenario en el que pueden insertarse segmentos adicionales de ADN, por ejemplo, mediante técnicas convencionales de clonación molecular. Otro tipo de vector es un vector viral, en el que las secuencias de ADN o ARN derivadas del virus están presentes en el vector para su empaquetamiento en un virus (por ejemplo, retrovirus, retrovirus de replicación defectuosa, adenovirus, adenovirus de replicación defectuosa y virus adenoasociados (AAV)). Los vectores virales también incluyen polinucleótidos transportados por un virus para su transfección en una célula huésped. Algunos vectores pueden replicarse de manera autónoma en una célula huésped en la que se introducen (por ejemplo, vectores bacterianos que tienen un origen de replicación bacteriano y vectores episomales de mamíferos). Otros vectores (por ejemplo, los vectores de mamíferos no episomales) se integran en el genoma de una célula huésped al ser introducidos en la célula huésped, y por tanto se replican junto con el genoma del huésped. Además, determinados vectores pueden dirigir la expresión de los genes a los que están unidos operativamente. Tales vectores se denominan en el presente documento "vectores de expresión". Los vectores de expresión comunes de utilidad en las técnicas de ADN recombinante a menudo están en forma de plásmidos.

Los vectores de expresión recombinantes pueden comprender un ácido nucleico de la invención en una forma adecuada para la expresión del ácido nucleico en una célula huésped, lo que significa que los vectores de expresión recombinantes incluyen uno o más elementos reguladores, que pueden seleccionarse en función de las células huésped que van a usarse para la expresión, que está operativamente unido a la secuencia de ácido nucleico que va a expresarse. Dentro de un vector de expresión recombinante, por "operativamente unido" se entiende que la secuencia de nucleótidos de interés está unida al/a los elemento(s) regulador(es) de manera que permite la expresión de la secuencia de nucleótidos (por ejemplo, en un sistema de transcripción/traducción *in vitro* o en una célula huésped cuando el vector se introduce en la célula huésped). Con respecto a los métodos de recombinación y clonación, se menciona la solicitud de patente estadounidense 10/815.730, publicada el 2 de septiembre de 2004 como US 2004-0171156 A1. Por tanto, las realizaciones dadas a conocer en el presente documento pueden comprender también células transgénicas que comprenden el sistema efector CRISPR. En determinadas realizaciones de ejemplo, la célula transgénica puede funcionar como un volumen individual discreto. En otras palabras, las muestras que comprenden un constructo de enmascaramiento pueden administrarse a una célula, por ejemplo, en una vesícula de administración adecuada y si la diana está presente en la vesícula de administración, el efector CRISPR se activa y se genera una señal detectable.

El/los vector(es) puede(n) incluir el/los elemento(s) regulador(es), por ejemplo, el/los promotor(es). El/los vector(es) puede(n) comprender secuencias de codificación de Cas, y/o un solo, pero posiblemente también puede(n) comprender al menos 3 u 8 o 16 o 32 o 48 o 50, ARN guía (por ejemplo, ARN_{gu}) que codifican para secuencias, tales como 1-2, 1-3, 1-4 1-5, 3-6, 3-7, 3-8, 3-9, 3-10, 3-8, 3-16, 3-30, 3-32, 3-48, 3-50 ARN (por ejemplo, ARN_{gu}). En un único vector puede haber un promotor para cada ARN (por ejemplo, ARN_{gu}), ventajosamente cuando hay hasta aproximadamente 16 ARN; y, cuando un único vector proporciona más de 16 ARN, uno o más promotores pueden impulsar la expresión de más de uno de los ARN, por ejemplo cuando hay 32 ARN, cada promotor puede impulsar la expresión de dos ARN, y cuando hay 48 ARN, cada promotor puede impulsar la expresión de tres ARN. Por simple aritmética y protocolos de clonación bien establecidos y las enseñanzas de esta divulgación, un experto en la técnica

puede practicar fácilmente la invención en cuanto a los ARN para un vector a modo de ejemplo adecuado tal como AAV, y un promotor adecuado tal como el promotor U6. Por ejemplo, el límite de empaquetamiento de AAV es de ~4,7 kb. La longitud de un único ARNg-U6 (más los sitios de restricción para la clonación) es de 361 pb. Por tanto, el experto puede ajustar fácilmente aproximadamente 12-16, por ejemplo, 13 casetes de U6-ARNg en un único vector.

5 Esto puede ensamblarse por cualquier medio adecuado, tal como la estrategia Golden Gate utilizada para el ensamblaje de TALE (genome-engineering.org/talectors/). El experto también puede usar una estrategia de guía en tándem para aumentar el número de U6-ARNg en aproximadamente 1,5 veces, por ejemplo, para aumentar desde 12-16, por ejemplo, desde 13 hasta aproximadamente 18-24, por ejemplo, aproximadamente 19 U6-ARNg.

10 Por tanto, un experto en la técnica puede alcanzar fácilmente aproximadamente 18-24, por ejemplo, aproximadamente 19 promotor-ARN, por ejemplo, U6-ARNg en un único vector, por ejemplo, un vector AAV. Otro medio para aumentar el número de promotores y ARN en un vector es usar un único promotor (por ejemplo, U6) para expresar un conjunto de ARN separados por secuencias escindibles. Y un medio incluso adicional para aumentar el número de promotores-ARN en un vector es expresar un conjunto de promotores-ARN separados por secuencias escindibles en el intrón de una secuencia o gen codificante; y, en este caso, es ventajoso usar un

15 promotor de la polimerasa II, que puede tener una expresión aumentada y permitir la transcripción de ARN largos de manera específica del tejido. (véase, por ejemplo, nar.oxfordjournals.org/content/34/7/e53.short y nature.com/mt/journal/v16/n9/abs/mt2008144a.html). En una realización ventajosa, el AAV puede empaquetar ARNg en tándem U6 que selecciona como diana hasta aproximadamente 50 genes. Por consiguiente, a partir de los conocimientos de la técnica y de las enseñanzas de la presente divulgación, el experto puede elaborar y usar

20 fácilmente vector(es), por ejemplo, un único vector, que exprese múltiples ARN o guías bajo el control o unido operativa o funcionalmente a uno o más promotores, especialmente en lo que respecta al número de ARN o guías comentado en el presente documento, sin ninguna experimentación indebida.

Las secuencias codificadoras de ARN guía y/o las secuencias codificadoras de Cas pueden estar unidas funcional u operativamente a elemento(s) regulador(es) y, por tanto, el/los elemento(s) regulador(es) impulsa(n) la expresión.

25 El/los promotor(es) puede(n) ser constitutivo(s) y/o condicional(es) y/o inducible(s) y/o específico(s) de tejido. El promotor puede seleccionarse del grupo que consiste en las ARN polimerasas, pol I, pol II, pol III, T7, U6, HI, el promotor LTR del virus del sarcoma de Rous (RSV) retroviral, el promotor del citomegalovirus (CMV), el promotor del SV40, el promotor de la dihidrofolato reductasa, el promotor de la β -actina, el promotor de la fosfoglicerol cinasa (PGK) y el promotor de EF1 α . Un promotor ventajoso es el promotor U6.

30 En algunas realizaciones, uno o más elementos de un sistema de selección como diana de ácidos nucleicos se deriva de un organismo particular que comprende un sistema de selección como diana de ARN de CRISPR endógeno. En determinadas realizaciones de ejemplo, el sistema que selecciona como diana ARN de CRISPR de la proteína efectora comprende al menos un dominio HEPN, incluyendo pero sin limitarse a los dominios HEPN descritos en el presente documento, dominios HEPN conocidos en la técnica, y dominios reconocidos como

35 dominios HEPN por comparación con motivos de secuencia de consenso. En el presente documento, se proporcionan varios de estos dominios. En un ejemplo no limitativo, una secuencia de consenso puede derivarse de las secuencias de C2c2 u ortólogos de Cas13b proporcionados en el presente documento. En determinadas realizaciones de ejemplo, la proteína efectora comprende un único dominio HEPN. En otras determinadas realizaciones de ejemplo, la proteína efectora comprende dos dominios HEPN.

40 En una realización de ejemplo, la proteína efectora comprende uno o más dominios HEPN que comprenden una secuencia de motivo RxxxxH. La secuencia de motivo RxxxxH puede ser, sin limitación, de un dominio HEPN descrito en el presente documento o un dominio HEPN conocido en la técnica. Las secuencias de motivos RxxxxH incluyen además secuencias de motivos creadas mediante la combinación de porciones de dos o más dominios HEPN. Tal como se indica, las secuencias de consenso pueden derivarse de las secuencias de los ortólogos dados

45 a conocer en la solicitud de patente provisional estadounidense 62/432.240 titulada "Novel CRISPR Enzymes and Systems", la solicitud de patente provisional estadounidense 62/471,710 titulada "Novel Type VI CRISPR Orthologs and Systems" presentada el 15 de marzo de 2017, y la solicitud de patente provisional estadounidense titulada "Novel Type VI CRISPR Orthologs and Systems," clasificada con el n.º de expediente 47627-05-2133 y presentada el 12 de abril de 2017.

50 En una realización de la invención, un dominio HEPN comprende al menos un motivo RxxxxH que comprende la secuencia de R{N/H/K}X1X2X3H. En una realización de la invención, un dominio HEPN comprende un motivo RxxxxH que comprende la secuencia de R{N/H}X1X2X3H. En una realización de la invención, un dominio HEPN comprende la secuencia de R{N/K}X1X2X3H. En determinadas realizaciones, X1 es R, S, D, E, Q, N, G, Y, o H. En determinadas realizaciones, X2 es I, S, T, V, o L. En determinadas realizaciones, X3 es L, F, N, Y, V, I, S, D, E, o A.

55 Los efectores adicionales para su uso según la invención pueden identificarse por su proximidad a los genes cas1, por ejemplo, aunque sin limitarse a ello, dentro de la región a 20 kb del inicio del gen cas1 y a 20 kb del final del gen cas1. En determinadas realizaciones, la proteína efectora comprende al menos un dominio HEPN y al menos 500 aminoácidos, y en las que la proteína efectora C2c2 está presente de manera natural en un genoma procariota dentro de 20 kb en el sentido de 5' o en el sentido de 3' de un gen Cas o de un conjunto CRISPR. Los ejemplos no

60 limitativos de proteínas Cas incluyen Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocidas como Csn1 y Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX,

Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homólogos de las mismas o versiones modificadas de las mismas. En determinadas realizaciones de ejemplo, la proteína efectora C2c2 está presente de manera natural en un genoma procarionte dentro de 20 kb en el sentido de 5' o en el sentido de 3' de un gen Cas 1. Los términos "ortólogo" y "homólogo" son bien conocidos en la técnica. A modo de orientación adicional, un "homólogo" de una proteína, tal como se usa en el presente documento, es una proteína de la misma especie que realiza la misma o similar función que la proteína de la que es homólogo. Las proteínas homólogas pueden estar relacionadas estructuralmente, pero no es necesario que lo estén, o sólo parcialmente. Un "ortólogo" de una proteína, tal como se usa en el presente documento, es una proteína de una especie diferente que realiza una función igual o similar a la de la proteína de la que es ortólogo. Las proteínas ortólogas pueden estar relacionadas estructuralmente, pero no es necesario que lo estén, o sólo parcialmente.

En realizaciones particulares, la enzima Cas que selecciona como diana ARN de tipo VI es C2c2. En otras realizaciones de ejemplo, la enzima Cas que selecciona como diana ARN de tipo VI es Cas 13b. En realizaciones particulares, el homólogo u ortólogo de una proteína de tipo VI tal como C2c2, tal como se menciona en el presente documento, tiene una homología o identidad de secuencia de al menos el 30%, o al menos el 40%, o al menos el 50%, o al menos el 60%, o al menos el 70%, o al menos el 80%, más preferiblemente al menos el 85%, incluso más preferiblemente al menos el 90%, tal como por ejemplo al menos el 95% con una proteína de tipo VI tal como C2c2 (por ejemplo, basada en la secuencia silvestre de cualquiera de C2c2 de *Leptotrichia shahii*, C2c2 de *Lachnospiraceae bacterium* MA2020, C2c2 de *Lachnospiraceae bacterium* NK4A179, C2c2 de *Clostridium aminophilum* (DSM 10710), C2c2 de *Carnobacterium gallinarum* (DSM 4847), C2c2 de *Paludibacter propionigenes* (WB4), C2c2 de *Listeria weihenstephanensis* (FSL R9-0317), C2c2 de *Listeriaceae bacterium* (FSL M6-0635), C2c2 de *Listeria newyorkensis* (FSL M6-0635), C2c2 de *Leptotrichia wadei* (F0279), C2c2 de *Rhodobacter capsulatus* (SB 1003), C2c2 de *Rhodobacter capsulatus* (R121), C2c2 de *Rhodobacter capsulatus* (DE442), C2c2 de *Leptotrichia wadei* (Lw2), o C2c2 de *Listeria seeligeri*). En otras realizaciones, el homólogo u ortólogo de una proteína de tipo VI, tal como C2c2, tal como se menciona en el presente documento, tiene una identidad de secuencia de al menos el 30%, o al menos el 40%, o al menos el 50%, o al menos el 60%, o al menos el 70%, o al menos el 80%, más preferiblemente al menos el 85%, incluso más preferiblemente al menos el 90%, tal como por ejemplo al menos el 95% con la C2c2 silvestre (por ejemplo, basado en la secuencia silvestre de cualquiera de C2c2 de *Leptotrichia shahii*, C2c2 de *Lachnospiraceae bacterium* MA2020, C2c2 de *Lachnospiraceae bacterium* NK4A179, C2c2 de *Clostridium aminophilum* (DSM 10710), C2c2 de *Carnobacterium gallinarum* (DSM 4847), C2c2 de *Paludibacter propionigenes* (WB4), C2c2 de *Listeria weihenstephanensis* (FSL R9-0317), C2c2 de *Listeriaceae bacterium* (FSL M6-0635), C2c2 de *Listeria newyorkensis* (FSL M6-0635), C2c2 de *Leptotrichia wadei* (F0279), C2c2 de *Rhodobacter capsulatus* (SB 1003), C2c2 de *Rhodobacter capsulatus* (R121), C2c2 de *Rhodobacter capsulatus* (DE442), C2c2 de *Leptotrichia wadei* (Lw2), o C2c2 de *Listeria seeligeri*).

En otras determinadas realizaciones de ejemplo, el sistema CRISPR la proteína efectora es una nucleasa C2c2. La actividad de C2c2 puede depender de la presencia de dos dominios HEPN. Se ha demostrado que estos son dominios de ARNasa, es decir, una nucleasa (en particular una endonucleasa) que corta el ARN. HEPN de C2c2 también puede seleccionar como diana ADN, o potencialmente ADN y/o ARN. Basándose en que los dominios HEPN de C2c2 pueden al menos unirse a ARN y, en su forma silvestre, cortar el ARN, entonces se prefiere que la proteína efectora de C2c2 tenga función de ARNasa. En cuanto a los sistemas CRISPR C2c2, se hace referencia a la provisional estadounidense 62/351,662 presentada el 17 de junio de 2016 y la provisional estadounidense 62/376.377 presentada el 17 de agosto de 2016. También se hace referencia a la provisional estadounidense 62/351.803 presentada el 17 de junio de 2016. También se hace referencia a la provisional estadounidense titulada "Novel Crispr Enzymes and Systems" presentada el 8 de diciembre de 2016 con el número 10035.PA4 del Broad Institute y el número de expediente 47627.03.2133. Se hace referencia además a East-Seletsky *et al.* "Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection" Nature doi:10.1038/nature19802 y Abudayyeh *et al.* "C2c2 is a single-component programmable RNA-guided RNA targeting CRISPR effector" bioRxiv doi:10.1101/054742.

La función de las ARNasas en los sistemas CRISPR es conocida, por ejemplo, se ha informado de la selección como diana de ARNm en determinados sistemas CRISPR-Cas de tipo III (Hale *et al.*, 2014, Genes Dev, vol. 28, 2432-2443; Hale *et al.*, 2009, Cell, vol. 139, 945-956; Peng *et al.*, 2015, Nucleic acids research, vol. 43, 406-417) y proporciona importantes ventajas. En el sistema de tipo III-A de *Staphylococcus epidermis*, la transcripción a través de las dianas da lugar a la escisión del ADN diana y sus transcritos, mediada por sitios activos independientes dentro del complejo de proteínas efectoras de ribonucleoproteínas Cas10-Csm (véase, Samai *et al.*, 2015, Cell, vol. 151, 1164-1174). Por tanto, se proporciona un sistema, composición o método CRISPR-Cas que selecciona como diana ARN a través de las presentes proteínas efectoras.

En una realización, la proteína Cas puede ser un ortólogo de C2c2 de un organismo de un género que incluye pero no se limita a *Leptotrichia*, *Listeria*, *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor*, *Mycoplasma* y *Campylobacter*. Las especies de organismos de un género de este tipo pueden ser tal como se indica de otro modo en el presente documento.

Algunos métodos de identificación de ortólogos de las enzimas del sistema CRISPR-Cas pueden implicar la

identificación de secuencias tracr en los genomas de interés. La identificación de secuencias tracr puede referirse a las siguientes etapas: buscar las repeticiones directas o secuencias tracr apareadas en una base de datos para identificar una región CRISPR que comprenda una enzima CRISPR. Buscar secuencias homólogas en la región CRISPR que flanquean la enzima CRISPR tanto en la dirección sentido con antisentido. Buscar terminadores transcripcionales y estructuras secundarias. Identificar cualquier secuencia que no sea una repetición directa o una secuencia tracr apareada pero que tenga más del 50% de identidad con la repetición directa o la secuencia tracr apareada como una secuencia tracr potencial. Tomar la secuencia tracr potencial y analizar para detectar secuencias terminadoras transcripcionales asociadas con la misma.

Se apreciará que cualquiera de las funcionalidades descritas en el presente documento puede modificarse por ingeniería para dar enzimas CRISPR a partir de otros ortólogos, incluyendo enzimas quiméricas que comprenden fragmentos de múltiples ortólogos. Los ejemplos de tales ortólogos se describen en otra parte del presente documento. Por tanto, las enzimas quiméricas pueden comprender fragmentos de ortólogos de enzimas CRISPR de un organismo que incluye, pero no se limita a, *Leptotrichia*, *Listeria*, *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor*, *Mycoplasma* y *Campylobacter*. Una enzima quimérica puede comprender un primer fragmento y un segundo fragmento, y los fragmentos pueden ser de ortólogos de enzimas CRISPR de organismos de los géneros mencionados en el presente documento o de las especies mencionadas en el presente documento; ventajosamente los fragmentos son de ortólogos de enzimas CRISPR de diferentes especies.

En las realizaciones, la proteína C2c2 a la que se hace referencia en el presente documento también abarca una variante funcional de C2c2 o un homólogo o un ortólogo de la misma. Una "variante funcional" de una proteína, tal como se usa en el presente documento, se refiere a una variante de tal proteína que conserva al menos parcialmente la actividad de esa proteína. Las variantes funcionales pueden incluir mutantes (que pueden ser mutantes de inserción, delección o sustitución), incluyendo polimorfos, etc. También se incluyen dentro de las variantes funcionales los productos de fusión de tal proteína con otro ácido nucleico, proteína, polipéptido o péptido, habitualmente no relacionado. Las variantes funcionales pueden producirse de manera natural o ser creadas por el hombre. Las realizaciones ventajosas pueden incluir proteínas efectoras que seleccionan como diana ARN de tipo VI obtenidas por ingeniería o que no se producen de manera natural.

En una realización, la(s) molécula(s) de ácido nucleico que codifica(n) para C2c2 o un ortólogo u homólogo de la misma, puede(n) tener codones optimizados para su expresión en una célula eucariota. Una célula eucariota puede ser tal como se comenta en el presente documento. La(s) molécula(s) de ácido nucleico puede(n) obtenerse por ingeniería o no producirse de manera natural.

En una realización, la C2c2 o un ortólogo u homólogo de la misma, puede comprender una o más mutaciones (y por tanto la(s) molécula(s) de ácido nucleico que lo codifica(n) puede(n) tener mutación/mutaciones). Las mutaciones pueden ser mutaciones introducidas artificialmente y pueden incluir, pero no se limitan a, una o más mutaciones en un dominio catalítico. Los ejemplos de dominios catalíticos con referencia a una enzima Cas9 pueden incluir, pero no se limitan a, los dominios RuvC I, RuvC II, RuvC III y HNH.

En una realización, la C2c2 o un ortólogo u homólogo de la misma, puede comprender una o más mutaciones. Las mutaciones pueden ser mutaciones introducidas artificialmente y pueden incluir, pero no se limitan a, una o más mutaciones en un dominio catalítico. Los ejemplos de dominios catalíticos con referencia a una enzima Cas pueden incluir, pero no se limitan, a los dominios HEPN.

En una realización, la C2c2 o un ortólogo u homólogo de la misma, puede usarse como una proteína genérica de unión a ácidos nucleicos con fusión a un dominio funcional o estando operativamente unida a él. Los dominios funcionales a modo de ejemplo pueden incluir, pero no se limitan a, un iniciador de la traducción, un activador de la traducción, un represor de la traducción, nucleasas, en particular ribonucleasas, un empalmosoma, perlas, un dominio inducible/controlable por luz o un dominio inducible/controlable químicamente.

En determinadas realizaciones de ejemplo, la proteína efectora C2c2 puede ser de un organismo seleccionado del grupo que consiste en; *Leptotrichia*, *Listeria*, *Corynebacter*, *Sutterella*, *Legionella*, *Treponema*, *Filifactor*, *Eubacterium*, *Streptococcus*, *Lactobacillus*, *Mycoplasma*, *Bacteroides*, *Flaviivola*, *Flavobacterium*, *Sphaerochaeta*, *Azospirillum*, *Gluconacetobacter*, *Neisseria*, *Roseburia*, *Parvibaculum*, *Staphylococcus*, *Nitratifactor*, *Mycoplasma* y *Campylobacter*.

En determinadas realizaciones, la proteína efectora puede ser una C2c2p de *Listeria sp.*, preferiblemente C2c2p de *Listeria seeligeri*, más preferiblemente C2c2p de *Listeria seeligeri* serovar 1/2b cepa SLCC3954 y la secuencia de ARNcr puede tener una longitud de 44 a 47 nucleótidos, con una repetición directa (DR) en 5' de 29 nt y un espaciador de 15 nt a 18 nt.

En determinadas realizaciones, la proteína efectora puede ser C2c2p de *Leptotrichia sp.*, preferiblemente C2c2p de *Leptotrichia shahii*, más preferiblemente C2c2p de *Leptotrichia shahii* DSM 19757 y la secuencia de ARNcr puede tener una longitud de 42 a 58 nucleótidos, con una repetición directa en 5' de al menos 24 nt, como una repetición

directa (DR) en 5' de 24-28 nt y un espaciador de al menos 14 nt, tal como un espaciador de 14 nt a 28 nt, o un espaciador de al menos 18 nt, tal como 19, 20, 21, 22, o más nt, tal como 18-28, 19-28, 20-28, 21-28, o 22-28 nt.

En determinadas realizaciones de ejemplo, la proteína efectora puede ser una *Leptotrichia* sp., *Leptotrichia wadei* F0279, o una *Listeria* sp., preferiblemente *Listeria newyorkensis* FSL M6-0635.

- 5 En determinadas realizaciones de ejemplo, las proteínas efectoras C2c2 de la invención incluyen, sin limitación, las siguientes 21 especies ortólogas (incluyendo múltiples loci de CRISPR: *Leptotrichia shahii*; *Leptotrichia wadei* (Lw2); *Listeria seeligeri*; *Lachnospiraceae bacterium* MA2020; *Lachnospiraceae bacterium* NK4A179; [*Clostridium*] *aminophilum* DSM 10710; *Carnobacterium gallinarum* DSM 4847; *Carnobacterium gallinarum* DSM 4847 (segundo loci de CRISPR); *Paludibacter propionici* WB4; *Listeria weihenstephanensis* FSL R9-0317; *Listeriaceae bacterium* FSL M6-0635; *Leptotrichia wadei* F0279; *Rhodobacter capsulatus* SB 1003; *Rhodobacter capsulatus* R121; *Rhodobacter capsulatus* DE442; *Leptotrichia buccalis* C-1013-b; *Herbinix hemicellulosilytica*; [*Eubacterium*] *rectale*; *Eubacteriaceae bacterium* CHKCI004; *Blautia* sp. Marseille-P2398; y *Leptotrichia* sp. taxón oral 879 cepa F0557. Otros doce (12) ejemplos no limitativos son: *Lachnospiraceae bacterium* NK4A144; *Chloroflexus aggregans*; *Demequina aurantiaca*; *Thalassospira* sp. TSL5-1; *Pseudobutyrvibrio* sp. OR37; *Butyrvibrio* sp. YAB3001; *Blautia* sp. Marseille-P2398; *Leptotrichia* sp. Marseille-P3007; *Bacteroides ihuae*; *Porphyromonadaceae bacterium* KH3CP3RA; *Listeria riparia*; e *Insolitispirillum peregrinum*.
- 10
- 15

- En determinadas realizaciones, la proteína C2c2 según la invención es o se deriva de uno de los ortólogos tal como se describe en la tabla a continuación, o es una proteína quimérica de dos o más de los ortólogos tal como se describe en la tabla a continuación, o es un mutante o variante de uno de los ortólogos tal como se describe en la tabla a continuación (o un mutante o variante quimérico), incluyendo C2c2 desactivada, C2c2 dividida, C2c2 desestabilizada, etc. tal como se define en el presente documento, con o sin fusión con un dominio heterólogo/funcional.
- 20

En determinadas realizaciones de ejemplo, la proteína efectora C2c2 se selecciona de la tabla 1 a continuación.

Tabla 1

Ortólogo de C2c2	Código	Múltiples letras
<i>Leptotrichia shahii</i>	C2-2	Lsh
<i>L. wadei</i> F0279 (Lw2)	C2-3	Lw2
<i>Listeria seeligeri</i>	C2-4	Lse
<i>Lachnospiraceae bacterium</i> MA2020	C2-5	LbM
<i>Lachnospiraceae bacterium</i> NK4A179	C2-6	LbNK179
<i>Clostridium aminophilum</i> DSM 10710	C2-7	Ca
<i>Carnobacterium gallinarum</i> DSM 4847	C2-8	Cg
<i>Carnobacterium gallinarum</i> DSM 4847	C2-9	Cg2
<i>Paludibacter propionigenes</i> WB4	C2-10	Pp
<i>Listeria weihenstephanensis</i> FSL R9-0317	C2-11	Lwei
<i>Listeriaceae bacterium</i> FSL M6-0635	C2-12	LbFSL
<i>Leptotrichia wadei</i> F0279	C2-13	Lw
<i>Rhodobacter capsulatus</i> SB 1003	C2-14	Rc
<i>Rhodobacter capsulatus</i> R121	C2-15	Rc
<i>Rhodobacter capsulatus</i> DE442	C2-16	Rc
<i>Leptotrichia buccalis</i> C-1013-b	C2-17	LbuC2c2
<i>Herbinix hemicellulosilytics</i>	C2-18	HheC2c2
<i>Eubacterium rectale</i>	C2-19	EreC2c2
<i>Eubacteriaceae bacterium</i> CHKC1004	C2-20	EbaC2c2
<i>Blautia</i> sp. Marseille-P2398	C2-21	BsmC2c2
<i>Leptotrichia</i> sp. taxón oral 879 cepa F0557	C2-22	LspC2c2
<i>Lachnospiraceae bacterium</i> NK4a144		
<i>Cloroflexus aggregans</i>		
<i>Demequina aurantiaca</i>		
<i>Thalassospira</i> sp. TSL5-1		
<i>Pseudobutyrvibrio</i> sp. OR37		
<i>Butyrvibrio</i> sp. YAB3001		

<i>Blautia</i> sp. Marseille-P2398		
<i>Leptotrichia</i> sp. Marseille-P300		
<i>Bacteroides ihuae</i>		
<i>Porphyromonadaceae bacterium</i> KH3CP3RA		
<i>Listeria riparia</i>		
<i>Insolitispirillum peregrinum</i>		

Las secuencias de las proteínas silvestres de las especies anteriores se enumeran en la tabla 2 a continuación. En determinadas realizaciones, se proporciona una secuencia de ácido nucleico que codifica para la proteína C2c2.

5 Tabla 2

C2c2-2		<i>L. shahii</i> (Lsh) (SEQ. ID. NO. 1)
C2c2-2		<i>L. shahii</i> (Lsh) (SEQ. ID. NO. 477) WP_018451595.1
c2c2-3		<i>L. wadei</i> (Lw2) (SEQ. ID. NO. 2)
c2c2-4		<i>Listeria seeligeri</i> (SEQ. ID. NO. 3)
c2c2-5	1	<i>Lachnospiraceae bacterium</i> MA2020 (SEQ. ID. NO. 4)
c2c2-6	2	<i>Lachnospiraceae bacterium</i> NK4A179 (SEQ. ID. NO. 5)
c2c2-7	3	<i>Clostridium aminophilum</i> DSM 10710 (SEQ. ID. NO. 6)
c2c2-8	5	<i>Carnobacterium gallinarum</i> DSM 4847 (SEQ. ID. NO. 7)
c2c2-9	6	<i>Carnobacterium gallinarum</i> DSM 4847 (SEQ. ID. NO. 8)
c2c2-10	7	<i>Paludibacter propionigenes</i> WB4 (SEQ. ID. NO. 9)
c2c2-11	9	<i>Listeria weihenstephanensis</i> FSL R9-0317 (SEQ. ID. NO. 10)
c2c2-12	10	<i>Listeriaceae bacterium</i> FSL M6-0635 = <i>Listeria newyorkensis</i> FSL M6-0635 (SEQ. ID. NO. 11)
c2c2-13	12	<i>Leptotrichia wadei</i> F0279 (SEQ. ID. NO. 12)
c2c2-14	15	<i>Rhodobacter capsulatus</i> SB 1003 (SEQ. ID. NO. 13)
c2c2-15	16	<i>Rhodobacter capsulatus</i> R121 (SEQ. ID. NO. 14)
c2c2-16	17	<i>Rhodobacter capsulatus</i> DE442 (SEQ. ID. NO. 15)
LbuC2c2 (C2-17)		<i>Leptotrichia buccalis</i> C-1013-b (SEQ ID NO: 309)
HheC2c2 (C2-18)		<i>Herbinix hemicellulosilytica</i> (SEQ ID NO:310)
EreC2c2 (C2-19)		<i>Eubacterium rectale</i> (SEQ ID NO: 311)
EbaC2C2 (C2-20)		<i>Eubacteriaceae bacterium</i> CHKCI004 (SEQ ID NO: 312)
C2c2 (C2-21)		<i>Blautia</i> sp. Marseille-P2398 (SEQ. I.D. No 319)
C2c2 (C2-22)		<i>Leptotrichia</i> sp. taxón oral 879 cepa F0557 (SEQ. ID. NO. 579)
C2c2 NK4A144 (C2-23)		<i>Lachnospiraceae bacterium</i> NK4A144 (SEQ. ID. NO. 313)
C2c2 Cloro_agg (C2-24)		Proteína S1 de unión a ARN de <i>Cloroflexus aggregans</i> (SEQ. ID. NO. 314)
C2c2 Dem_Aur (C2-25)		<i>Demequina aurantiaca</i> (SEQ. ID. NO. 315)
C2c2 Thal_Sp_TSL5 (C2-26)		<i>Thalassospira</i> sp. TSL5-1 (SEQ. I.D. No 316)
C2c2 Pseudo_sp (C2-27)		<i>Pseudobutyrvibrio</i> sp. OR37 (SEQ. ID. NO. 317)
C2c2_Buty_sp (C2-28)		<i>Butyrvibrio</i> sp. YAB3001 (SEQ. ID. NO. 318)
C2c2_Blautia_sp (C2-29)		<i>Blautia</i> sp. Marseille-P2398(SEQ. ID. NO. 478)
C2c2_Lepto_sp_Marseille (C2-30)		<i>Leptotrichia</i> sp. Marseille-P3007 (SEQ. ID No. 320)
C2c2_Bacteroides_ihuai (C2-31)		<i>Bacteroides ihuai</i> (SEQ. I.D. No 321)
C2c2_Porph_bacteria (C2-32)		<i>Porphyromonadaceae bacterium</i> KH3CP3RA(SEQ. ID. NO. 322)
C2c2_Listeria_riparia (C2-33)		<i>Listeria riparia</i> (SEQ. ID. NO. 323)
C2c2_insolitisperegrinum (C2-34)		<i>Insolitispirillum peregrinum</i> (SEQ. ID. NO. 324)

En una realización de la invención, se proporciona una proteína efectora que comprende una secuencia de aminoácidos que tiene al menos el 80% de homología de secuencia con la secuencia silvestre de cualquiera de C2c2 de *Leptotrichia shahii*, C2c2 de *Lachnospiraceae bacterium* MA2020, C2c2 de *Lachnospiraceae bacterium* NK4A179, C2c2 de *Clostridium aminophilum* (DSM 10710), C2c2 de *Carnobacterium gallinarum* (DSM4847), C2c2 de *Paludibacter propionigenes* (WB4), C2c2 de *Listeria weihenstephanensis* (FSL R9-0317), C2c2 de *Listeriaceae bacterium* (FSL M6-0635), C2c2 de *Listeria newyorkensis* (FSL M6-0635), C2c2 de *Leptotrichia wadei* (F0279), C2c2 de *Rhodobacter capsulatus* (SB 1003), C2c2 de *Rhodobacter capsulatus* (R121), C2c2 de *Rhodobacter capsulatus*

(DE442), C2c2 de *Leptotrichia wadei* (Lw2), o C2c2 de *Listeria seeligeri*.

En una realización de la invención, la proteína efectora comprende una secuencia de aminoácidos que tiene al menos el 80% de homología de secuencia con una secuencia consenso de proteína efectora de tipo VI que incluye, pero no se limita a, una secuencia consenso descrita en el presente documento.

- 5 Según la invención, una secuencia consenso puede generarse a partir de múltiples ortólogos de C2c2, que pueden asistir en la ubicación de residuos de aminoácido conservados, y motivos, que incluyen pero no se limitan a residuos catalíticos y motivos de HEPN en ortólogos de C2c2 que median en la función de C2c2. Una secuencia consenso de este tipo, generada a partir de los 33 ortólogos mencionados anteriormente usando la alineación Geneious es:

```

MKISKVXXXVXXKXXXGKLXKXVNERNRXAKRLSNXLBKYIXXIDKIXKKEXX
KKFXAXEEITLKLNQXXXBXLXKAXXDLRKDNXYSXJKKILHNEDINXEEEXELLI
NDXLEKLXKIESXKYSYQKXXXNYXMSVQEHSKKSIXRIXESAKRNKEALDKFL
KEYAXLDPRMEXLAKLRKLELYFYFKNDXIXXEEEXNVXXHKXLKENHPDFV
EXXXNKENAELNXYAIEXXKJLKYYFPXXAKNSNDKIFEKQELKKXWIHQJEN
AVERILLXXGKVXYKLQXGYLAELWKIRINEIFIKYIXVGKAVAXFALRNXXXKBE
NDILGGKIXKKLNGITSFYXEKIKAEELQREXAVEVAFAANXLYAXDLXXIRXSI
LQFFGGASNWDXFLFFHFATXSISDKKXNAELIXKKJGLVIREKLYSNNVAMFY
SKDDLEKLLNXLXXFXLRASQVPSFKKVYVRXBFPQNLLKKFNDEKDDEAYSAX
YYLLKEIYYNXFLPYFSANNXFFFXVKNLVLKANKDKFXXAFXDIREMNXGSPIE
YLXXTQXNXXNEGRKKKEEKEXDFIKFLLQIFXKGFDDYLNKNNXXFILKFIPEPTX
IEIXXELQAWYIVGKFLNARKXNLLGXFXSYLKLLDDIELRALRNENIKYQSSNX
EKEVLEXCLELIGLLSLDLNDYFBDEXDFAXYJGKXLD FEKKXMKDLAELXPYD
QNDGENPIVNRNIXLAKKYGTNLLEKJXDKVSEKEIKEYYELKKEIEEYXXKGE
ELHEEWXQXKNRVEXRDILEYXEELXGQIINYXNLXNKVLLYFQLGLHYLLLDIL
GRLVGTYGIWERDAXLYQIAAMYXNGLPEYIXXKKNDKYKDGQIVGXKINXFK
XDKKXLYNAGLELFENXNEHKNIXIRNYIAHFNYLSKAESSLLXYSENLRXLFSY
DRKLKNAVXKSLINILLRHGMVLKFKFGTDKKSXVIRSXXKIXHLKSIKKLYYP

```

- 10 EVXVSKEYCKLVKXLLKYK (SEQ ID NO: 325)

En otro ejemplo no limitativo, una herramienta de alineación de secuencia para asistir en la generación de una secuencia consenso e identificación de residuos conservados es la herramienta de alineación MUSCLE (www.ebi.ac.uk/Tools/msa/muscle/). Por ejemplo, usando MUSCLE, pueden identificarse las siguientes ubicaciones de aminoácido conservados entre ortólogos de C2c2 en C2c2 de *Leptotrichia wadei*: K2; K5; V6; E301; L331; I335; N341; G351; K352; E375; L392; L396; D403; F446; I466; I470; R474 (HEPN); H475; H479 (HEPN), E508; P556; L561; I595; Y596; F600; Y669; I673; F681; L685; Y761; L676; L779; Y782; L836; D847; Y863; L869; I872; K879; I933; L954; I958; R961; Y965; E970; R971; D972; R1046 (HEPN), H1051 (HEPN), Y1075; D1076; K1078; K1080; I1083; I1090.

- 20 En la figura 50 se muestra una alineación de secuencia de los dominios HEPN a modo de ejemplo que muestra los residuos altamente conservados.

- En determinadas realizaciones de ejemplo, la proteína efectora que selecciona como diana ARN es una proteína efectora de tipo VI-B, tal como Cas13b y las proteínas del grupo 29 o del grupo 30. En determinadas realizaciones de ejemplo, la proteína efectora que selecciona como diana ARN comprende uno o más dominios HEPN. En determinadas realizaciones de ejemplo, la proteína efectora que selecciona como diana ARN comprende un dominio HEPN C-terminal, un dominio HEPN N-terminal, o ambos. En cuanto a las proteínas efectoras de tipo VI-B de ejemplo que pueden usarse en el contexto de la presente invención, se hace referencia a la solicitud estadounidense n.º 15/331.792 titulada "Novel CRISPR Enzymes and Systems" y presentada el 21 de octubre de 2016, la solicitud internacional de patente n.º PCT/US2016/058302 titulada "Novel CRISPR Enzymes and Systems", y presentada el 21 de octubre de 2016, y Smargon *et al.* "Cas13b is a Type VI-B CRISPR-associated RNA-Guided RNase differentially regulated by accessory proteins Csx27 and Csx28" *Molecular Cell*, 65, 1-13 (2017); [dx.doi.org/10.1016/j.molcel.2016.12.023](https://doi.org/10.1016/j.molcel.2016.12.023), y la solicitud provisional estadounidense n.º aún por asignar, titulada "Novel

Cas13b Orthologues CRISPR Enzymes and System" presentada el 15 de marzo de 2017. En determinadas realizaciones, la enzima Cas13b se deriva de *Bergeyella zoohelcum*. En otras determinadas realizaciones de ejemplo, la proteína efectora es o comprende una secuencia de aminoácidos que tiene al menos el 80% de homología de secuencia con cualquiera de las secuencias enumeradas en la tabla 3.

5 Tabla 3

B-01	<i>Bergeyella zoohelcum</i>
B-02	<i>Prevotella intermedia</i>
B-03	<i>Prevotella buccae</i>
B-04	<i>Alistipes</i> sp. ZOR0009
B-05	<i>Prevotella</i> sp. MA2016
B-06	<i>Riemerella anatipestifer</i>
B-07	<i>Prevotella aurantiaca</i>
B-08	<i>Prevotella saccharolytica</i>
B-09	<i>Prevotella intermedia</i>
B-10	<i>Capnocytophaga canimorsus</i>
B-11	<i>Porphyromonas gulae</i>
B-12	<i>Prevotella</i> sp. P5-125
B-13	<i>Flavobacterium branchiophilum</i>
B-14	<i>Porphyromonas gingivalis</i>
B-15	<i>Prevotella intermedia</i>

En determinadas realizaciones de ejemplo, la secuencia silvestre del ortólogo de Cas13b se encuentra en la tabla 4a o 4b a continuación.

Tabla 4a

<i>Bergeyella zoohelcum</i> (SEQ. ID. NO. 326)	1
<i>Prevotella intermedia</i> (SEQ. ID. NO. 327)	2
<i>Prevotella buccae</i> (SEQ. ID. NO. 328)	3
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 329)	4
<i>Bacteroides pyogenes</i> (SEQ. ID. NO. 330)	5
<i>Alistipes</i> sp. ZOR0009 (SEQ. ID. NO. 331)	6
<i>Prevotella</i> sp. MA2016 (SEQ. ID. NO. 332)	7a
<i>Prevotella</i> sp. MA2016 (SEQ. ID. NO. 333)	7b

<i>Riemerella anatipestifer</i> (SEQ. ID. NO. 334)	8
<i>Prevotella aurantiaca</i> (SEQ. ID. NO. 335)	9
<i>Prevotella saccharolytica</i> (SEQ. ID. NO. 336)	10
HMPREF9712_03108 [<i>Myroides odoratimimus</i> CCUG 10230] (SEQ. ID. NO. 337)	11
<i>Prevotella intermedia</i> (SEQ. ID. NO. 338)	12
<i>Capnocytophaga canimorsus</i> (SEQ. ID. NO. 339)	13
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 340)	14
<i>Prevotella</i> sp. P5-125 (SEQ. ID. NO. 341)	15
<i>Flavobacterium branchiophilum</i> (SEQ. ID. NO. 342)	16
<i>Myroides odoratimimus</i> (SEQ. ID. NO. 343)	17
<i>Flavobacterium columnare</i> (SEQ. ID. NO. 344)	18
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 345)	19
<i>Porphyromonas</i> sp. COT-052 OH4946 (SEQ. ID. NO. 346)	20
<i>Prevotella intermedia</i> (SEQ. ID. NO. 347)	21
PIN17 0200 [<i>Prevotella intermedia</i> 17] (SEQ. ID. NO. 348)	AFJ07523
<i>Prevotella intermedia</i> (SEQ. ID. NO. 349)	BAU18623
HMPREF6485_0083 [<i>Prevotella buccae</i> ATCC 33574] (SEQ. ID. NO. 350)	EFU31981
HMPREF9144_1146 [<i>Prevotella pallens</i> ATCC 700821] (SEQ. ID. NO. 351)	EGQ18444
HMPREF9714_02132 [<i>Myroides odoratimimus</i> CCUG 12901] (SEQ. ID. NO. 352)	EH008761
HMPREF9711_00870 [<i>Myroides odoratimimus</i> CCUG 3837] (SEQ. ID. NO. 353)	EKB06014
HMPREF9699_02005 [<i>Bergeyella zoohelcum</i> ATCC 43767] (SEQ. ID. NO. 354)	EKB54193
HMPREF9151_01387 [<i>Prevotella saccharolytica</i> F0055] (SEQ. ID. NO. 355)	EKY00089
A343_1752 [<i>Porphyromonas gingivalis</i> JCVI SC001] (SEQ. ID. NO. 356)	EOA10535
HMPREF1981_03090 [<i>Bacteroides pyogenes</i> F0041] (SEQ. ID. NO. 357)	ERI81700
HMPREF1553_02065 [<i>Porphyromonas gingivalis</i> F0568] (SEQ. ID. NO. 358)	ERJ65637
HMPREF1988_01768 [<i>Porphyromonas gingivalis</i> F0185] (SEQ. ID. NO. 359)	ERJ81987
HMPREF1990 01800 [<i>Porphyromonas gingivalis</i> W4087] (SEQ. ID. NO. 360)	ERJ87335
M573_117042 [<i>Prevotella intermedia</i> ZT] (SEQ. ID. NO. 361)	KJJ86756
A2033_10205 [<i>Bacteroidetes bacterium</i> GWA2_31_9] (SEQ. ID. NO. 362)	OFX18020.1
SAMN05421542_0666 [<i>Chryseobacteria jejuense</i>] (SEQ. ID. NO. 363)	SDI27289.1
SAMN05444360_11366 [<i>Chryseobacteria carnipullorum</i>] (SEQ. ID. NO. 364)	SHM52812.1
SAMN05421786_1011119 [<i>Chryseobacteria ureilyticum</i>] (SEQ. ID. NO. 365)	SIS70481.1

<i>Prevotella buccae</i> (SEQ. ID. NO. 366)	WP_004343581
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 367)	WP_005873511
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 368)	WP_005874195
<i>Prevotella pallens</i> (SEQ. ID. NO. 369)	WP_006044833
<i>Myroides odoratimimus</i> (SEQ. ID. NO. 370)	WP_006261414
<i>Myroides odoratimimus</i> (SEQ. ID. NO. 371)	WP_006265509
<i>Prevotella</i> sp. MSX73 (SEQ. ID. NO. 372)	WP_007412163
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 373)	WP_012458414
<i>Paludibacter propionigenes</i> (SEQ. ID. NO. 374)	WP_013446107
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 375)	WP_013816155
<i>Flavobacterium columnare</i> (SEQ. ID. NO. 376)	WP_014165541
<i>Psychroflexus torquis</i> (SEQ. ID. NO. 377)	WP_015024765
<i>Riemerella anatipestifer</i> (SEQ. ID. NO. 378)	WP_015345620
<i>Prevotella pleuritidis</i> (SEQ. ID. NO. 379)	WP_021584635
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 380)	WP_021663197
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 381)	WP_021665475
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 382)	WP_021677657
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 383)	WP_021680012
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 384)	WP_023846767
<i>Prevotella falsenii</i> (SEQ. ID. NO. 385)	WP_036884929
<i>Prevotella pleuritidis</i> (SEQ. ID. NO. 386)	WP_036931485
[<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 387)	WP_039417390
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 388)	WP_039418912
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 389)	WP_039419792
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 390)	WP_039426176
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 391)	WP_039431778
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 392)	WP_039437199
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 393)	WP_039442171
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 394)	WP_039445055
<i>Capnocytophaga cynodegmi</i> (SEQ. ID. NO. 395)	WP_041989581
<i>Prevotella</i> sp. P5-119 (SEQ. ID. NO. 396)	WP_042518169
<i>Prevotella</i> sp. P4-76 (SEQ. ID. NO. 397)	WP_044072147
<i>Prevotella</i> sp. P5-60 (SEQ. ID. NO. 398)	WP_044074780
<i>Phaeodactilibacter xiamenensis</i> (SEQ. ID. NO. 399)	WP_044218239
<i>Flavobacterium</i> sp. 316 (SEQ. ID. NO. 400)	WP_045968377
<i>Porphyromonas gulae</i> (SEQ. ID. NO. 401)	WP_046201018
WP_047431796 (SEQ. ID. NO. 402)	<i>Chryseobacterium</i> sp. YR477
<i>Riemerella anatipestifer</i> (SEQ. ID. NO. 403)	WP_049354263
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 404)	WP_052912312
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 405)	WP_058019250
<i>Flavobacterium columnare</i> (SEQ. ID. NO. 406)	WP_060381855
<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 407)	WP_061156470

<i>Porphyromonas gingivalis</i> (SEQ. ID. NO. 408)	WP_061156637
<i>Riemerella anatipestifer</i> (SEQ. ID. NO. 409)	WP_061710138
<i>Flavobacterium columnare</i> (SEQ. ID. NO. 410)	WP_063744070
<i>Riemerella anatipestifer</i> (SEQ. ID. NO. 411)	WP_064970887
<i>Sinomicrobium oceani</i> (SEQ. ID. NO. 412)	WP_072319476.1
<i>Reichenbachiella agariperforans</i> (SEQ. ID. NO. 413)	WP_073124441.1

Tabla 4b

Nombre o n.º de registro
WP_015345620 (SEQ. ID. NO. 479)
WP_049354263 (SEQ. ID. NO. 480)
WP_061710138 (SEQ. ID. NO. 481)
6 (SEQ. ID. NO. 482) <i>Alistipes</i> sp. ZOR0009
SIS70481.1
15 <i>Prevotella</i> sp. (SEQ. ID. NO. 484)
WP_042518169 (SEQ. ID. NO. 485)
WP_044072147 (SEQ. ID. NO. 486)
WP_044074780 (SEQ. ID. NO. 487)
8_(modificado) (SEQ. ID. NO. 488)
WP_064970887 (SEQ. ID. NO. 489)
5 (SEQ. ID. NO. 490)
ERI81700 (SEQ. ID. NO. 491)
WP_036931485 (SEQ. ID. NO. 492)
19 (SEQ. ID. NO. 493)
WP_012458414 (SEQ. ID. NO. 494)
WP_013816155 (SEQ. ID. NO. 495)
WP_039417390 (SEQ. ID. NO. 496)
WP_039419792 (SEQ. ID. NO. 497)
WP_039426176 (SEQ. ID. NO. 498)
WP_039437199 (SEQ. ID. NO. 499)
WP_061156470 (SEQ. ID. NO. 500)
12 (SEQ. ID. NO. 501)
9 (SEQ. ID. NO. 502)
EGQ18444 (SEQ. ID. NO. 503)
KJJ86756 (SEQ. ID. NO. 504)
WP_006044833 (SEQ. ID. NO. 505)
2 (SEQ. ID. NO. 506)
3 (SEQ. ID. NO. 507)
EFU31981 (SEQ. ID. NO. 508)
WP_004343581 (SEQ. ID. NO. 509)
WP_007412163 (SEQ. ID. NO. 510)
WP_044218239 (SEQ. ID. NO. 511)
21 (SEQ. ID. NO. 512)

BAU18623 (SEQ. ID. NO. 513)
WP_036884929 (SEQ. ID. NO. 514)
WP_073124441.1 (SEQ. ID. NO. 515)
AFJ07523 (SEQ. ID. NO. 516)
4 (SEQ. ID. NO. 517)
ERJ65637 (SEQ. ID. NO. 518)
ERJ81987 (SEQ. ID. NO. 519)
ERJ87335 (SEQ. ID. NO. 520)
WP_005873511 (SEQ. ID. NO. 521)
WP_021663197 (SEQ. ID. NO. 522)
WP_021665475 (SEQ. ID. NO. 523)
WP_021677657 (SEQ. ID. NO. 524)
WP_021680012 (SEQ. ID. NO. 525)
WP_023846767 (SEQ. ID. NO. 526)
WP_039445055 (SEQ. ID. NO. 527)
WP_061156637 (SEQ. ID. NO. 528)
WP_021584635 (SEQ. ID. NO. 529)
WP_015024765 (SEQ. ID. NO. 530)
WP_047431796 (SEQ. ID. NO. 531)
WP_072319476.1 (SEQ. ID. NO. 532)
16 (SEQ. ID. NO. 533)
EKY00089 (SEQ. ID. NO. 534)
10 (SEQ. ID. NO. 535)
WP_013446107 (SEQ. ID. NO. 536)
WP_045968377 (SEQ. ID. NO. 537)
SHM52812.1 (SEQ. ID. NO. 538)
EHO08761 (SEQ. ID. NO. 539)
EKB06014 (SEQ. ID. NO. 540)
WP_006261414 (SEQ. ID. NO. 541)
WP_006265509 (SEQ. ID. NO. 542)
11 (SEQ. ID. NO. 543)
17 (SEQ. ID. NO. 544)
OFX18020.1 (SEQ. ID. NO. 545)
SDI27289.1 (SEQ. ID. NO. 546)
WP_039442171 (SEQ. ID. NO. 547)
14 (SEQ. ID. NO. 548)
20 (SEQ. ID. NO. 549)
EOA10535 (SEQ. ID. NO. 550)
WP_005874195 (SEQ. ID. NO. 551)
WP_039418912 (SEQ. ID. NO. 552)
WP_039431778 (SEQ. ID. NO. 553)
WP_046201018 (SEQ. ID. NO. 554)
WP_052912312 (SEQ. ID. NO. 555)

WP_058019250 (SEQ. ID. NO. 556)
WP_014165541 (SEQ. ID. NO. 557)
13 (SEQ. ID. NO. 558)
WP_060381855 (SEQ. ID. NO. 559)
WP_063744070 (SEQ. ID. NO. 560)
18 (SEQ. ID. NO. 561)
WP_041989581 (SEQ. ID. NO. 562)
1 (SEQ. ID. NO. 563)
EKB54193 (SEQ. ID. NO. 564)
7_(modificado) (SEQ. ID. NO. 565)
7_(modificado)_- sólo_residuos (SEQ. ID. NO. 566)

En determinadas realizaciones de ejemplo, la proteína efectora que selecciona como diana ARN es una proteína efectora Cas13c tal como se da a conocer en la solicitud de patente provisional estadounidense n.º 62/525.165 presentada el 26 de junio de 2017, y la solicitud PCT n.º US 2017/047193 presentada el 16 de agosto de 2017. Se proporcionan secuencias de ortólogos silvestres de Cas13c en la tabla 5 a continuación.

Tabla 5

Nombre
EHO19081 (SEQ. ID. NO. 567)
WP_094899336 (SEQ. ID. NO. 568)
WP_040490876 (SEQ. ID. NO. 569)
WP_047396607 (SEQ. ID. NO. 570)
WP_035935671 (SEQ. ID. NO. 571)
WP_035906563 (SEQ. ID. NO. 572)
WP_042678931 (SEQ. ID. NO. 573)
WP_062627846 (SEQ. ID. NO. 574)
WP_005959231 (SEQ. ID. NO. 575)
WP_027128616 (SEQ. ID. NO. 576)
WP_062624740 (SEQ. ID. NO. 577)
WP_096402050 (SEQ. ID. NO. 578)

En determinadas realizaciones de ejemplo, la proteína Cas13 puede seleccionarse de cualquiera de lo siguiente.

Tabla 6

ID	Especie	SEQ. ID. NO.:
Cas13a1	<i>Leptotrichia shahii</i>	580
Cas13a2	<i>Leptotrichia wadei</i> (Lw2)	581
Cas13a3	<i>Listeria seeligeri</i>	582
Cas13a4	<i>Lachnospiraceae bacterium</i> MA2020	583
Cas13a5	<i>Lachnospiraceae bacterium</i> NK4A179	584
Cas13a6	[<i>Clostridium</i>] <i>aminophilum</i> DSM 10710	585
Cas13a7	<i>Carnobacterium gallinarum</i> DSM 4847	586
Cas13a8	<i>Carnobacterium gallinarum</i> DSM 4847	587
Cas13a9	<i>Paludibacter propionisigenes</i> WB4	588

Cas13a10	<i>Listeria weihenstephanensis</i> FSL R9-0317	589
Cas13a11	<i>Listeriaceae bacterium</i> FSL M6-0635	590
Cas13a12	<i>Leptotrichia wadei</i> F0279	591
Cas13a13	<i>Rhodobacter capsulatus</i> SB 1003	592
Cas13a14	<i>Rhodobacter capsulatus</i> R121	593
Cas13a15	<i>Rhodobacter capsulatus</i> DE442	594
Cas13a16	<i>Leptotrichia buccalis</i> C-1013-b	595
Cas13a17	<i>Herbinix hemicellulosilytica</i>	596
Cas13a18	[<i>Eubacterium</i>] <i>rectale</i>	597
Cas13a19	<i>Eubacteriaceae bacterium</i> CHKCI004	598
Cas13a20	<i>Blautia</i> sp. Marseille-P2398	599
Cas13a21	<i>Leptotrichia</i> sp. taxón oral 879 cepa F0557	600
Cas13b1	<i>Bergeyella zoohelcum</i>	601
Cas13b2	<i>Prevotella intermedia</i>	602
Cas13b3	<i>Prevotella buccae</i>	603
Cas13b4	<i>Alistipes</i> sp. ZOR0009	604
Cas13b5	<i>Prevotella</i> sp. MA2016	605
Cas13b6	<i>Riemerella anatipestifer</i>	606
Cas13b7	<i>Prevotella aurantiaca</i>	607
Cas13b8	<i>Prevotella saccharolytica</i>	608
Cas13b9	<i>Prevotella intermedia</i>	609
Cas13b10	<i>Capnocytophaga canimorsus</i>	610
Cas13b11	<i>Porphyromonas gulae</i>	611
Cas13b12	<i>Prevotella</i> sp. P5-125	612
Cas13b13	<i>Flavobacterium branchiophilum</i>	613
Cas13b14	<i>Porphyromonas gingivalis</i>	614
Cas13b15	<i>Prevotella intermedia</i>	615
Cas13c1	<i>Fusobacterium necrophorum</i> subsp. <i>funduliforme</i> ATCC 51357 contig00003	616
Cas13c2	<i>Fusobacterium necrophorum</i> DJ-2 contig0065, secuencia indiscriminada de genoma completo	617
Cas13c3	<i>Fusobacterium necrophorum</i> BFTR-1 contig0068	618
Ca13c4	<i>Fusobacterium necrophorum</i> subsp. <i>funduliforme</i> 1_1_36S cont1.14	619
Cas13c5	<i>Fusobacterium perfoetens</i> ATCC 29250 T364DRAFT_scaffold00009.9_C	620
Cas13c6	<i>Fusobacterium ulcerans</i> ATCC 49185 cont2.38	621
Cas13c7	<i>Anaerosalibacter</i> sp. Ensamblaje de genoma ND1 <i>Anaerosalibacter massiliensis</i> ND 1	622

SECUENCIAS GUÍA

Tal como se usa en el presente documento, el término “secuencia guía”, “ARNcr”, “ARN guía” o “ARN de guía único” o “ARNg” se refiere a un polinucleótido que comprende cualquier secuencia de polinucleótido que tiene suficiente

complementariedad con una secuencia de ácido nucleico diana como para hibridarse con la secuencia de ácido nucleico diana y dirigir la unión específica de secuencia de un complejo que selecciona como diana un ARN que comprende la secuencia guía y una proteína efectora CRISPR a la secuencia de ácido nucleico diana. En algunas realizaciones de ejemplo, el grado de complementariedad, cuando se alinea óptimamente usando un algoritmo de alineación adecuado, es de aproximadamente o más de aproximadamente el 50%, el 60%, el 75%, el 80%, el 85%, el 90%, el 95%, el 97,5%, el 99%, o más. La alineación óptima puede determinarse con el uso de cualquier algoritmo adecuado para alinear secuencias, entre los que se incluyen, a modo de ejemplo no limitativo, el algoritmo de Smith-Waterman, el algoritmo de Needleman-Wunsch, los algoritmos basados en la transformada de Burrows-Wheeler (por ejemplo, el alineador de Burrows Wheeler), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; disponible en www.novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (disponible en soap.genomics.org.cn), y Maq (disponible en maq.sourceforge.net). La capacidad de una secuencia guía (dentro de un ARN guía que selecciona como diana un ácido nucleico) para dirigir la unión específica de secuencia de un complejo que selecciona como diana un ácido nucleico a una secuencia de ácido nucleico diana puede evaluarse mediante cualquier ensayo adecuado. Por ejemplo, los componentes de un sistema CRISPR que selecciona como diana un ácido nucleico suficientes para formar un complejo que selecciona como diana un ácido nucleico, incluyendo la secuencia guía que va a someterse a prueba, pueden proporcionarse a una célula huésped que tenga la secuencia de ácido nucleico diana correspondiente, tal como, por ejemplo, mediante transfección con vectores que codifican para los componentes del complejo que selecciona como diana un ácido nucleico, seguida de una evaluación del direccionamiento preferente (por ejemplo, escisión) dentro de la secuencia de ácido nucleico diana, tal como por ejemplo mediante el ensayo Surveyor tal como se describe en el presente documento. Del mismo modo, la escisión de una secuencia de ácido nucleico diana puede evaluarse en un tubo de ensayo proporcionando la secuencia de ácido nucleico diana, los componentes de un complejo que selecciona como diana un ácido nucleico, incluyendo la secuencia guía que va a someterse a prueba y una secuencia guía de control distinta de la secuencia guía de prueba, y comparando la unión o la tasa de escisión en la secuencia diana entre las reacciones de la secuencia guía de prueba y de control. Son posibles otros ensayos, y se les ocurrirán a los expertos en la técnica. Una secuencia guía, y por tanto una guía que selecciona como diana un ácido nucleico, puede seleccionarse para seleccionar como diana cualquier secuencia de ácido nucleico diana. La secuencia diana puede ser ADN. La secuencia diana puede ser cualquier secuencia de ARN. En algunas realizaciones, la secuencia diana puede ser una secuencia dentro de una molécula de ARN seleccionada del grupo que consiste en ARN mensajero (ARNm), pre-ARNm, ARN ribosómico (ARNr), ARN de transferencia (ARNt), micro-ARN (miARN), ARN interferente pequeño (ARNip), ARN nuclear pequeño (ARNnp), ARN nucleolar pequeño (ARNnop), ARN bicatenario (ARNbc), ARN no codificante (ARNnc), ARN no codificante largo (ARNncl) y ARN citoplasmático pequeño (ARNcp). En algunas realizaciones preferidas, la secuencia diana puede ser una secuencia dentro de una molécula de ARN seleccionada del grupo que consiste en ARNm y ARNncl. En algunas realizaciones más preferidas, la secuencia diana puede ser una secuencia dentro de una molécula de ARNm o una molécula de pre-ARNm.

En algunas realizaciones, una guía que selecciona como diana un ácido nucleico se selecciona para reducir el grado de estructura secundaria dentro de la guía que selecciona como diana un ácido nucleico. En algunas realizaciones, aproximadamente o menos de aproximadamente el 75%, el 50%, el 40%, el 30%, el 25%, el 20%, el 15%, el 10%, el 5%, el 1% o menos de los nucleótidos de la guía que selecciona como diana un ácido nucleico participan en el apareamiento de bases autocomplementarias cuando se pliegan de manera óptima. El plegado óptimo puede determinarse mediante cualquier algoritmo de plegado de polinucleótidos adecuado. Algunos programas se basan en el cálculo de la energía libre de Gibbs mínima. Un ejemplo de uno de estos algoritmos es mFold, tal como se describió Zuker y Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Otro ejemplo de algoritmo de plegado es el servidor web en línea RNAfold, desarrollado en el Instituto de Química Teórica de la Universidad de Viena, usando el algoritmo de predicción de estructura centroide (véase, por ejemplo, A.R. Gruber *et al.*, 2008, Cell 106(1): 23-24; y PA Carr y GM Church, 2009, Nature Biotechnology 27(12): 1151-62).

En determinadas realizaciones, un ARNcr o ARN guía puede comprender, consistir esencialmente en, o consistir en una secuencia de repetición directa (DR) y una secuencia guía o secuencia de espaciador. En determinadas realizaciones, el ARNcr o ARN guía puede comprender, consistir esencialmente en, o consistir en una secuencia de repetición directa fusionada o unida a una secuencia guía o secuencia de espaciador. En determinadas realizaciones, la secuencia de repetición directa puede estar situada en el sentido de 5' (es decir, 5') de la secuencia guía o secuencia de espaciador. En otras realizaciones, la secuencia de repetición directa puede estar situada en el sentido de 3' (es decir, 3') de la secuencia guía o secuencia de espaciador.

En determinadas realizaciones, el ARNcr comprende un tallo-bucle, preferiblemente un tallo-bucle único. En determinadas realizaciones, la secuencia de repetición directa forma un tallo-bucle, preferiblemente un único tallo-bucle.

En determinadas realizaciones, la longitud del espaciador del ARN guía es de desde 15 hasta 35 nt. En determinadas realizaciones, la longitud del espaciador del ARN guía es de al menos 15 nucleótidos. En determinadas realizaciones, la longitud del espaciador es de desde 15 hasta 17 nt, por ejemplo, 15, 16 o 17 nt, de desde 17 hasta 20 nt, por ejemplo, 17, 18, 19 o 20 nt, de desde 20 hasta 24 nt, por ejemplo, 20, 21, 22, 23 o 24 nt, de desde 23 hasta 25 nt, por ejemplo, 23, 24 o 25 nt, de desde 24 hasta 27 nt, por ejemplo, 24, 25, 26 o 27 nt, de

desde 27 hasta 30 nt, por ejemplo, 27, 28, 29 o 30 nt, de desde 30 hasta 35 nt, por ejemplo, 30, 31, 32, 33, 34 o 35 nt, o 35 nt o más.

En general, el sistema CRISPR-Cas, CRISPR-Cas9 o CRISPR puede ser tal como se usa en los documentos anteriores, tales como el documento WO 2014/093622 (documento PCT/US2013/074667) y se refiere en conjunto a los transcritos y otros elementos que intervienen en la expresión o dirigen la actividad de los genes asociados a CRISPR ("Cas"), incluyendo las secuencias que codifican para un gen Cas, en particular un gen Cas9 en el caso de CRISPR-Cas9, una secuencia *tracr* (CRISPR transactivante) (por ejemplo, *ARNtracr* o un *ARNtracr* parcial activo), una secuencia *tracr*-mate (que abarca una "repetición directa" y una repetición directa parcial procesada por *ARNtracr* en el contexto de un sistema CRISPR endógeno), una secuencia guía (también denominada "espaciador" en el contexto de un sistema CRISPR endógeno), o "ARN" tal como se usa ese término en el presente documento (por ejemplo, ARN para guiar a Cas9, por ejemplo, ARN de CRISPR y ARN transactivante (*tracr*) o ARN de guía único (ARNgu) (ARN quimérico)) u otras secuencias y transcritos de un locus de CRISPR. En general, un sistema CRISPR se caracteriza por elementos que promueven la formación de un complejo de CRISPR en el sitio de una secuencia diana (también denominado protoespaciador en el contexto de un sistema CRISPR endógeno). En el contexto de la formación de un complejo de CRISPR, "secuencia diana" se refiere a una secuencia con respecto a la cual una secuencia guía esté diseñada para tener complementariedad, donde la hibridación entre una secuencia diana y una secuencia guía promueve la formación de un complejo de CRISPR. La sección de la secuencia guía a través de la cual la complementariedad con la secuencia diana es importante para la actividad de escisión se denomina en el presente documento secuencia semilla. Una secuencia diana puede comprender cualquier polinucleótido, tal como polinucleótidos de ADN o ARN. En algunas realizaciones, una secuencia diana está ubicada en el núcleo o citoplasma de una célula y puede incluir ácidos nucleicos en o a partir de mitocondrias, orgánulos, vesículas, liposomas o partículas presentes dentro de la célula. En algunas realizaciones, especialmente para usos no nucleares, no se prefieren los NLS. En algunas realizaciones, un sistema CRISPR comprende una o más señales de exportación nuclear (NES). En algunas realizaciones, un sistema CRISPR comprende uno o más NLS y uno o más NES. En algunas realizaciones, pueden identificarse repeticiones directas *in silico* mediante la búsqueda de motivos repetitivos que cumplan alguno o todos los siguientes criterios: 1. encontrados en una ventana de 2 kb de secuencia genómica que flanquea el locus de CRISPR de tipo II; 2. van desde 20 a 50 pb; y 3. se separan unos de otros por de 20 a 50 pb. En algunas realizaciones, pueden usarse 2 de estos criterios, por ejemplo, 1 y 2, 2 y 3 o 1 y 3. En algunas realizaciones, pueden usarse los 3 criterios.

En realizaciones de la invención, los términos secuencia guía y ARN guía, es decir, ARN que puede guiar a Cas a un locus genómico diana, se usan manera intercambiable tal como en los documentos citados anteriormente, tales como el documento WO 2014/093622 (documento PCT/US2013/074667). En general, una secuencia guía es cualquier secuencia de polinucleótido que tenga suficiente complementariedad con una secuencia de polinucleótido diana para hibridarse con la secuencia diana y dirigir la unión específica de secuencia de un complejo de CRISPR a la secuencia diana. En algunas realizaciones, el grado de complementariedad entre una secuencia guía y su correspondiente secuencia diana, cuando se alinea de manera óptima usando un algoritmo de alineación adecuado, es de aproximadamente o superior a aproximadamente el 50%, el 60%, el 75%, el 80%, el 85%, el 90%, el 95%, el 97,5%, el 99% o más. La alineación óptima puede determinarse con el uso de cualquier algoritmo adecuado para alinear secuencias, cuyo ejemplo no limitativo incluye el algoritmo de Smith-Waterman, el algoritmo de Needleman-Wunsch, algoritmos basados en la transformada de Burrows-Wheeler (por ejemplo, el alineador de Burrows-Wheeler), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; disponible en www.novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (disponible en soap.genomics.org.cn) y Maq (disponible en maq.sourceforge.net). En algunas realizaciones, una secuencia guía tiene una longitud de aproximadamente o más de aproximadamente 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75 o más nucleótidos. En algunas realizaciones, una secuencia guía tiene una longitud de menos de aproximadamente 75, 50, 45, 40, 35, 30, 25, 20, 15, 12 nucleótidos o menos. Preferiblemente, la secuencia guía tiene una longitud de 10-30 nucleótidos. La capacidad de una secuencia guía para dirigir la unión específica de secuencia de un complejo de CRISPR a una secuencia diana puede evaluarse mediante cualquier ensayo adecuado. Por ejemplo, los componentes de un sistema CRISPR suficientes para formar un complejo de CRISPR, incluyendo la secuencia guía que va a someterse a prueba, pueden proporcionarse a una célula huésped que tenga la secuencia diana correspondiente, tal como mediante transfección con vectores que codifican para los componentes de la secuencia CRISPR, seguida de una evaluación de la escisión preferencial dentro de la secuencia diana, tal como mediante el ensayo Surveyor tal como se describe en el presente documento. Del mismo modo, la escisión de una secuencia de polinucleótido diana puede evaluarse en un tubo de ensayo proporcionando la secuencia diana, los componentes de un complejo de CRISPR, incluyendo la secuencia guía que va a someterse a prueba y una secuencia guía de control distinta de la secuencia guía de prueba, y comparando la unión o la tasa de escisión en la secuencia diana entre las reacciones de la secuencia guía de prueba y de control. Son posibles otros ensayos, y se les ocurrirán a los expertos en la técnica.

En algunas realizaciones de los sistemas CRISPR-Cas, el grado de complementariedad entre una secuencia guía y su correspondiente secuencia diana puede ser de aproximadamente o más de aproximadamente el 50%, el 60%, el 75%, el 80%, el 85%, el 90%, el 95%, el 97,5%, el 99% o el 100%; una guía o ARN o ARNgu puede tener una longitud de aproximadamente o más de aproximadamente 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75 o más nucleótidos; o una guía o ARN o ARNgu puede tener una

longitud de menos de aproximadamente 75, 50, 45, 40, 35, 30, 25, 20, 15, 12 nucleótidos o menos; y ventajosamente, el ARN tracr tiene una longitud de 30 o 50 nucleótidos. Sin embargo, un aspecto de la invención es reducir las interacciones inespecíficas, por ejemplo, reducir la interacción de la guía con una secuencia diana que tiene una baja complementariedad. De hecho, en los ejemplos, se muestra que la invención implica mutaciones que dan como resultado que el sistema CRISPR-Cas pueda distinguir entre secuencias diana y no diana que tienen una complementariedad de más del 80% a aproximadamente el 95%, por ejemplo, una complementariedad del 83%-84% o el 88-89% o el 94-95% (por ejemplo, distinguir entre una secuencia diana que tiene 18 nucleótidos de una secuencia no diana de 18 nucleótidos que tiene 1, 2 o 3 apareamientos erróneos). Por consiguiente, en el contexto de la presente invención el grado de complementariedad entre una secuencia guía y su correspondiente secuencia diana es de más del 94,5% o el 95% o el 95,5% o el 96% o el 96,5% o el 97% o el 97,5% o el 98% o el 98,5% o el 99% o el 99,5% o el 99,9% o el 100%. La secuencia no diana presenta una complementariedad de menos del 100% o el 99,9% o el 99,5% o el 99% o 99% o el 98,5% o el 98% o el 97,5% o el 97% o el 96,5% o el 96% o el 95,5% o el 95% o el 94,5% o el 94% o el 93% o el 92% o el 91% o el 90% o el 89% o el 88% o el 87% o el 86% o el 85% o el 84% o el 83% o el 82% o el 81% o el 80% entre la secuencia y la guía, con es ventajoso que la secuencia no diana tiene una complementariedad del 100% o el 99,9% o el 99,5% o el 99% o el 99% o el 98,5% o el 98% o el 97,5% o el 97% o el 96,5% o el 96% o el 95,5% o el 95% o el 94,5% entre el secuencia y la guía.

Modificaciones de la guía

En determinadas realizaciones, las guías de la invención comprenden ácidos nucleicos que no se producen de manera natural y/o nucleótidos y/o análogos de nucleótidos que no se producen de manera natural y/o modificaciones químicas. Los ácidos nucleicos que no se producen de manera natural pueden incluir, por ejemplo, mezclas de nucleótidos que se producen de manera natural y que no se producen de manera natural. Los nucleótidos y/o análogos de nucleótidos que no se producen de manera natural pueden modificarse en el resto de ribosa, fosfato y/o base. En una realización de la invención, un ácido nucleico guía comprende ribonucleótidos y no ribonucleótidos. En una realización de este tipo, una guía comprende uno o más ribonucleótidos y uno o más desoxirribonucleótidos. En una realización de la invención, la guía comprende uno o más nucleótidos o análogos de nucleótidos que no se producen de manera natural, tales como un nucleótido con enlace fosforotioato, enlace boranofosfato, nucleótidos de ácido nucleico bloqueado (LNA) que comprenden un puente de metileno entre los carbonos 2' y 4' del anillo de ribosa, o ácidos nucleicos en puente (BNA). Otros ejemplos de nucleótidos modificados incluyen análogos de 2'-O-metilo, análogos de 2'-desoxilo, análogos de 2-tiouridina, análogos de N6-metiladenosina o análogos de 2'-fluoro. Ejemplos adicionales de bases modificadas incluyen, pero no se limitan a, 2-aminopurina, 5-bromo-uridina, pseudouridina (ψ), N¹-metilpseudouridina ($\text{me}^1\psi$), 5-metoxiuridina (5moU), inosina, 7-metilguanosina. Los ejemplos de modificaciones químicas del ARN guía incluyen, sin limitación, la incorporación de 2'-O-metilo (M), 2'-O-metil-3'-fosforotioato (MS), fosforotioato (PS), etilo restringido en S (cEt), o 2'-O-metil-3'-tioPACE (MSP) en uno o más nucleótidos terminales. Tales guías modificadas químicamente pueden comprender una mayor estabilidad y una mayor actividad en comparación con las guías no modificadas, aunque la especificidad específica frente a la inespecífica no es predecible. (Véase, Hendel, 2015, Nat Biotechnol. 33(9):985-9, doi: 10.1038/nbt.3290, publicado en línea el 29 de junio de 2015; Ragdarm *et al.*, 0215, PNAS, E7110-E7111; Allerson *et al.*, J. Med. Chemistry 2005, 48:901-904; Bransen *et al.*, Front. Genet., 2012, 3:154; Deng *et al.*, PNAS, 2015, 112:11870-11875; Sharma *et al.*, MedChemComm., 2014, 5:1454-1471; Hendel *et al.*, Nat. Biotechnology. (2015) 33(9): 985-989; Li *et al.*, Nature Biomedical Engineering, 2017, 1, 0066 DOI:10.1038/s41551-017-0066). En algunas realizaciones, el extremo 5' y/o 3' de un ARN guía se modifica mediante una variedad de restos funcionales que incluyen colorantes fluorescentes, polietilenglicol, colesterol, proteínas o marcadores de detección. (Véase Kelly *et al.*, 2016, J. Biotech. 233:74-83). En determinadas realizaciones, una guía comprende ribonucleótidos en una región que se une a un ADN diana y uno o más desoxirribonucleótidos y/o análogos de nucleótidos en una región que se une a Cas9, Cpf1 o C2c1. En una realización de la invención, los desoxirribonucleótidos y/o los análogos de nucleótidos se incorporan en estructuras guía obtenidas por ingeniería, tales como, sin limitación, el extremo 5' y/o 3', las regiones de tallo-bucle y la región semilla. En determinadas realizaciones, la modificación no está en el asidero 5' de las regiones de tallo-bucle. La modificación química en el asidero 5' de la región del tallo-bucle de una guía puede abolir su función (véase Li, *et al.*, Nature Biomedical Engineering, 2017, 1:0066). En determinadas realizaciones, se modifican químicamente al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50 o 75 nucleótidos de una guía. En algunas realizaciones, se modifican químicamente 3-5 nucleótidos en el extremo 3' o 5' de una guía. En algunas realizaciones, sólo se introducen modificaciones menores en la región semilla, tales como modificaciones 2'-F. En algunas realizaciones, la modificación 2'-F se introduce en el extremo 3' de una guía. En determinadas realizaciones, se modifican químicamente de tres a cinco nucleótidos en el extremo 5' y/o 3' de la guía con 2'-O-metilo (M), 2'-O-metil-3'-fosforotioato (MS), S-etilo restringido (cEt), o 2'-O-metil-3'-tioPACE (MSP). Tal modificación puede mejorar la eficiencia de edición del genoma (véase Hendel *et al.*, Nat. Biotechnology. (2015) 33(9): 985-989). En determinadas realizaciones, todos los enlaces fosfodiéster de una guía se sustituyen con fosforotioatos (PS) para mejorar los niveles de alteración génica. En determinadas realizaciones, se modifican químicamente más de cinco nucleótidos en el extremo 5' y/o 3' de la guía con 2'-O-Me, 2'-F o etilo restringido en S (cEt). Tal guía modificada químicamente puede mediar en niveles mejorados de interrupción de genes (véase Ragdarm *et al.*, 0215, PNAS, E7110-E7111). En una realización de la invención, se modifica una guía para que comprenda un resto químico en su extremo 3' y/o 5'. Tales restos incluyen, pero no se limitan a, amina, azida, alquino, tio, dibenzociclooctino (DBCO) o rodamina. En determinada realización, el resto químico se conjuga con la guía mediante un ligador, tal como una cadena de alquilo. En determinadas realizaciones, el resto químico de la guía

modificada puede usarse para unir la guía a otra molécula, tal como ADN, ARN, proteína o nanopartículas. Tal guía modificada químicamente puede usarse para identificar o enriquecer células editadas genéricamente mediante un sistema CRISPR (véase Lee *et al.*, eLife, 2017, 6:e25312, DOI:10.7554).

En determinadas realizaciones, el sistema CRISPR tal como se proporciona en el presente documento puede hacer uso de ARNcr o polinucleótido análogo que comprende una secuencia guía, en el que el polinucleótido es un ARN, un ADN o una mezcla de ARN y ADN, y/o en el que el polinucleótido comprende uno o más análogos de nucleótidos. La secuencia puede comprender cualquier estructura incluyendo, pero sin limitarse a, una estructura de un ARNcr nativo, tal como una estructura de protuberancia, horquilla o tallo-bucle. En determinadas realizaciones, el polinucleótido que comprende la secuencia guía forma un dúplex con una segunda secuencia de polinucleótido que puede ser una secuencia de ARN o ADN.

En determinadas realizaciones, se hace uso de ARN guía modificados químicamente. Los ejemplos de modificaciones químicas del ARN guía incluyen, sin limitación, la incorporación de 2'-O-metilo (M), 2'-O-metil-3'fosforotioato (MS) o 2'-O-metil-3'tioPACE (MSP) en uno o más nucleótidos terminales. Tales ARN guía modificados químicamente pueden comprender una mayor estabilidad y una mayor actividad en comparación con los ARN guía no modificados, aunque la especificidad específica frente a inespecífica no es predecible. (Véase, Hendel, 2015, Nat Biotechnol. 33(9):985-9, doi: 10.1038/nbt.3290, publicado en línea el 29 de junio de 2015). Los ARN guía modificados químicamente incluyen además, sin limitación, ARN con enlaces fosforotioato y nucleótidos de ácido nucleico bloqueado (LNA) que comprenden un puente de metileno entre los carbonos 2' y 4' del anillo de ribosa.

En algunas realizaciones, una secuencia guía tiene una longitud de aproximadamente o más de aproximadamente 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75 o más nucleótidos. En algunas realizaciones, una secuencia guía tiene una longitud de menos de aproximadamente 75, 50, 45, 40, 35, 30, 25, 20, 15, 12 nucleótidos o menos. Preferiblemente, la secuencia guía tiene una longitud de 10 a 30 nucleótidos. La capacidad de una secuencia guía para dirigir la unión específica de secuencia de un complejo de CRISPR a una secuencia diana puede evaluarse mediante cualquier ensayo adecuado. Por ejemplo, los componentes de un sistema CRISPR suficientes para formar un complejo de CRISPR, incluyendo la secuencia guía que va a someterse a prueba, pueden proporcionarse a una célula huésped que tenga la secuencia diana correspondiente, tal como mediante transfección con vectores que codifican para los componentes de la secuencia CRISPR, seguida de una evaluación de la escisión preferencial dentro de la secuencia diana, tal como por ejemplo mediante el ensayo Surveyor. Del mismo modo, la escisión de un ARN diana puede evaluarse en un tubo de ensayo proporcionando la secuencia diana, los componentes de un complejo de CRISPR, incluyendo la secuencia guía que va a someterse a prueba y una secuencia guía de control distinta de la secuencia guía de prueba, y comparando la unión o la tasa de escisión en la secuencia diana entre las reacciones de la secuencia guía de prueba y de control. Son posibles otros ensayos, y se les ocurrirán a los expertos en la técnica.

En algunas realizaciones, la modificación de la guía es una modificación química, una inserción, una delección o una división. En algunas realizaciones, la modificación química incluye, pero no se limita a, la incorporación de análogos de 2'-O-metilo (M), análogos de 2'-desoxilo, análogos de 2-tiouridina, análogos de N6-metiladenosina, análogos de 2'-fluoro, 2-aminopurina, 5-bromo-uridina, pseudouridina (ψ), N¹-metilpseudouridina ($me^1\psi$), 5-metoxiuridina (5moU), inosina, 7-metilguanosina, 2'-O-metil-3'-fosforotioato (MS), etilo restringido en S (cEt), fosforotioato (PS) o 2'-O-metil-3'-tioPACE (MSP). En algunas realizaciones, la guía comprende una o más modificaciones de fosforotioato. En determinadas realizaciones, se modifican químicamente al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o 25 nucleótidos de la guía. En determinadas realizaciones, se modifican químicamente uno o más nucleótidos en la región semilla. En determinadas realizaciones, se modifican químicamente uno o más nucleótidos en el extremo 3'-terminal. En determinadas realizaciones, no se modifica químicamente ninguno de los nucleótidos en el asidero 5'. En algunas realizaciones, la modificación química en la región semilla es una modificación menor, tal como la incorporación de un análogo de 2'-fluoro. En una realización específica, un nucleótido de la región semilla se reemplaza con un análogo de 2'-fluoro. En algunas realizaciones, se modifican químicamente 5 o 10 nucleótidos en el extremo 3'-terminal. Tales modificaciones químicas en el extremo 3'-terminal del ARNcr de Cpf1 mejoran la eficiencia de corte de genes (véase Li, *et al.*, Nature Biomedical Engineering, 2017, 1:0066). En una realización específica, 5 nucleótidos en el extremo 3'-terminal se reemplazan con análogos de 2'-fluoro. En una realización específica, 10 nucleótidos en el extremo 3'-terminal se reemplazan con análogos de 2'-fluoro. En una realización específica, 5 nucleótidos en el extremo 3'-terminal se reemplazan con análogos de 2'-O-metilo (M).

En algunas realizaciones, se modifica el bucle del asidero 5' de la guía. En algunas realizaciones, el bucle del asidero 5' de la guía se modifica para tener una delección, una inserción, una división o modificaciones químicas. En determinadas realizaciones, el bucle comprende 3, 4 o 5 nucleótidos. En determinadas realizaciones, el bucle comprende la secuencia de UCUU, UUUU, UAUU o UGUU.

Puede seleccionarse una secuencia guía y, por tanto, un ARN guía que selecciona como diana un ácido nucleico para seleccionar como diana cualquier secuencia de ácido nucleico diana. En el contexto de la formación de un complejo de CRISPR, "secuencia diana" se refiere a una secuencia con respecto a la cual una secuencia guía esté diseñada para tener complementariedad, donde la hibridación entre una secuencia diana y una secuencia guía promueve la formación de un complejo de CRISPR. Una secuencia diana puede comprender polinucleótidos de ARN. El término "ARN diana" se refiere a un polinucleótido de ARN que es o comprende la secuencia diana. En

otras palabras, el ARN diana puede ser un polinucleótido de ARN o una parte de un polinucleótido de ARN con respecto al cual una parte del ARNg, es decir, la secuencia guía, esté diseñada para tener complementariedad y con respecto al cual va a dirigirse la función efectora mediada por el complejo que comprende la proteína efectora CRISPR y un ARNg. En algunas realizaciones, una secuencia diana está ubicada en el núcleo o citoplasma de una célula. La secuencia diana puede ser ADN. La secuencia diana puede ser cualquier secuencia de ARN. En algunas realizaciones, la secuencia diana puede ser una secuencia dentro de una molécula de ARN seleccionada del grupo que consiste en ARN mensajero (ARNm), pre-ARNm, ARN ribosómico (ARNr), ARN de transferencia (ARNt), micro-ARN (miARN), ARN interferente pequeño (ARNip), ARN nuclear pequeño (ARNnp), ARN nuclear pequeño (ARNnop), ARN bicatenario (ARNbc), ARN no codificante (ARNnc), ARN no codificante largo (ARNncl) y ARN citoplasmático pequeño (ARNcp). En algunas realizaciones preferidas, la secuencia diana puede ser una secuencia dentro de una molécula de ARN seleccionada del grupo que consiste en ARNm, pre-ARNm y ARNr. En algunas realizaciones preferidas, la secuencia diana puede ser una secuencia dentro de una molécula de ARN seleccionada del grupo que consiste en ARNnc y ARNncl. En algunas realizaciones más preferidas, la secuencia diana puede ser una secuencia dentro de una molécula de ARNm o una molécula de pre-ARNm.

En determinadas realizaciones, la longitud del espaciador del ARN guía es menos de 28 nucleótidos. En determinadas realizaciones, la longitud del espaciador del ARN guía es de al menos 18 nucleótidos y menos de 28 nucleótidos. En determinadas realizaciones, la longitud del espaciador del ARN guía está entre 19 y 28 nucleótidos. En determinadas realizaciones, la longitud del espaciador del ARN guía está entre 19 y 25 nucleótidos. En determinadas realizaciones, la longitud del espaciador del ARN guía es de 20 nucleótidos. En determinadas realizaciones, la longitud del espaciador del ARN guía es de 23 nucleótidos. En determinadas realizaciones, la longitud del espaciador del ARN guía es de 25 nucleótidos.

En determinadas realizaciones, las modulaciones de la eficiencia de escisión pueden explotarse mediante la introducción de apareamientos erróneos, por ejemplo, 1 o más apareamientos erróneos, tal como 1 o 2 apareamientos erróneos entre la secuencia de espaciador y la secuencia diana, incluyendo la posición del apareamiento erróneo a lo largo del espaciador/diana. Por ejemplo, cuanto más central (es decir, no en 3' o 5') un apareamiento erróneo doble, más se ve afectada la eficiencia de escisión. Por consiguiente, eligiendo una posición de apareamiento erróneo a lo largo del espaciador, puede modularse la eficiencia de escisión. A modo de ejemplo, si se desea menos del 100% de escisión de las dianas (por ejemplo, en una población celular), pueden introducirse 1 o más, tales como preferiblemente 2, apareamientos erróneos entre el espaciador y la secuencia diana en las secuencias de espaciador. Cuanto más central la posición de apareamiento erróneo a lo largo del espaciador, menor es el porcentaje de escisión.

En determinadas realizaciones de ejemplo, la eficiencia de escisión puede explotarse para diseñar guías únicas que pueden distinguir dos o más dianas que varían en un solo nucleótido, tal como un polimorfismo de un solo nucleótido (SNP), variación o mutación (puntual). El efector CRISPR puede tener una sensibilidad reducida a los SNP (u otras variaciones de un solo nucleótido) y continuar escindiendo las dianas de SNP con un determinado nivel de eficiencia. Por tanto, para dos dianas, o un conjunto de dianas, puede diseñarse un ARN guía con una secuencia de nucleótidos que sea complementaria a una de las dianas, es decir, el SNP específico. El ARN guía está diseñado además para tener un apareamiento erróneo sintético. Tal como se usa en el presente documento, un "apareamiento erróneo sintético" se refiere a un apareamiento erróneo que no se produce de manera natural que se introduce en el sentido de 5' o en el sentido de 3' del SNP que se produce de manera natural tal como, como máximo 5 nucleótidos en el sentido de 5' o en el sentido de 3', por ejemplo, 4, 3, 2 o 1 nucleótidos en el sentido de 5' o en el sentido de 3', preferiblemente como máximo 3 nucleótidos en el sentido de 5' o en el sentido de 3', más preferiblemente como máximo 2 nucleótidos en el sentido de 5' o en el sentido de 3', lo más preferiblemente 1 nucleótido en el sentido de 5' o en el sentido de 3' (es decir, adyacente al SNP). Cuando el efector CRISPR se une al SNP específico, sólo se formará un apareamiento erróneo único con el apareamiento erróneo sintético y el efector CRISPR continuará activado y se producirá una señal detectable. Cuando el ARN guía se hibrida con un SNP inespecífico, se formarán dos apareamientos erróneos, es decir, el apareamiento erróneo del SNP y el apareamiento erróneo sintético, y no se generará ninguna señal detectable. Por tanto, los sistemas descritos en el presente documento pueden diseñarse para distinguir los SNP dentro de una población. Por ejemplo, los sistemas pueden usarse para distinguir cepas patógenas que difieren en un solo SNP o detectar determinados SNP específicos de enfermedades, tales como, pero sin limitarse a, SNP asociados a enfermedades, tales como, sin limitación, SNP asociados a cáncer.

En determinadas realizaciones, el ARN guía está diseñado de manera que el SNP está ubicado en la posición 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 de la secuencia de espaciador (comenzando en el extremo 5'). En determinadas realizaciones, el ARN guía está diseñado de manera que el SNP está ubicado en la posición 1, 2, 3, 4, 5, 6, 7, 8 o 9 de la secuencia de espaciador (comenzando en el extremo 5'). En determinadas realizaciones, el ARN guía está diseñado de manera que el SNP está ubicado en la posición 2, 3, 4, 5, 6 o 7 de la secuencia de espaciador (comenzando en el extremo 5'). En determinadas realizaciones, el ARN guía está diseñado de manera que el SNP está ubicado en la posición 3, 4, 5 o 6 de la secuencia de espaciador (comenzando en el extremo 5'). En determinadas realizaciones, el ARN guía está diseñado de manera que el SNP está ubicado en la posición 3 de la secuencia de espaciador (comenzando en el extremo 5').

En determinadas realizaciones, el ARN guía está diseñado de tal manera que el apareamiento erróneo (por ejemplo,

el apareamiento erróneo sintético, es decir, una mutación adicional además de un SNP) está ubicado en la posición 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 de la secuencia de espaciador (comenzando en el extremo 5'). En determinadas realizaciones, el ARN guía está diseñado de manera que el apareamiento erróneo está ubicado en la posición 1, 2, 3, 4, 5, 6, 7, 8 o 9 de la secuencia de espaciador (comenzando en el extremo 5'). En determinadas realizaciones, el ARN guía está diseñado de manera que el apareamiento erróneo está ubicado en la posición 4, 5, 6 o 7 de la secuencia de espaciador (comenzando en el extremo 5'). En determinadas realizaciones, el ARN guía está diseñado de manera que el apareamiento erróneo está ubicado en la posición 5 de la secuencia de espaciador (comenzando en el extremo 5').

En determinadas realizaciones, el ARN guía está diseñado de manera que el apareamiento erróneo está ubicado 2 nucleótidos en el sentido de 5' del SNP (es decir, un nucleótido intermedio).

En determinadas realizaciones, el ARN guía está diseñado de manera que el apareamiento erróneo está ubicado 2 nucleótidos en el sentido de 3' del SNP (es decir, un nucleótido intermedio).

En determinadas realizaciones, el ARN guía está diseñado de manera que el apareamiento erróneo está ubicado en la posición 5 de la secuencia de espaciador (comenzando en el extremo 5') y el SNP está ubicado en la posición 3 de la secuencia de espaciador (comenzando en el extremo 5').

Las realizaciones descritas en el presente documento comprenden la inducción de una o más modificaciones de nucleótidos en una célula eucariota (*in vitro*, es decir, en una célula eucariota aislada) tal como se comenta en el presente documento, que comprende suministrar a la célula un vector tal como se comenta en el presente documento. La(s) mutación/mutaciones puede(n) incluir la introducción, delección o sustitución de uno o más nucleótidos en cada secuencia diana de la(s) célula(s) a través del/de los ARN guía(s). Las mutaciones pueden incluir la introducción, delección o sustitución de 1-75 nucleótidos en cada secuencia diana de dicha(s) célula(s) a través del/de los ARN guía(s). Las mutaciones pueden incluir la introducción, delección o sustitución de 1, 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50 o 75 nucleótidos en cada secuencia diana de dicha(s) célula(s) a través del/de los ARN guía(s). Las mutaciones pueden incluir la introducción, delección o sustitución de 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50 o 75 nucleótidos en cada secuencia diana de dicha(s) célula(s) a través del/de los ARN guía(s). Las mutaciones incluyen la introducción, delección o sustitución de 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50 o 75 nucleótidos en cada secuencia diana de dicha(s) célula(s) a través del/de los ARN guía(s). Las mutaciones pueden incluir la introducción, delección o sustitución de 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50 o 75 nucleótidos en cada secuencia diana de dicha(s) célula(s) a través del/de los ARN guía(s). Las mutaciones pueden incluir la introducción, delección o sustitución de 40, 45, 50, 75, 100, 200, 300, 400 o 500 nucleótidos en cada secuencia diana de dicha(s) célula(s) a través del/de los ARN guía(s).

Normalmente, en el contexto de un sistema CRISPR endógeno, la formación de un complejo de CRISPR (que comprende una secuencia guía hibridada con una secuencia diana y complejada con una o más proteínas Cas) da como resultado una escisión en o cerca de (por ejemplo, dentro de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 o más pares de bases de) la secuencia diana, pero puede depender, por ejemplo, de la estructura secundaria, en particular en el caso de dianas de ARN.

CONSTRUCTOS DE ENMASCARAMIENTO BASADOS EN ARN

Tal como se usa en el presente documento, un "constructo de enmascaramiento" se refiere a una molécula que puede escindirse o desactivarse de otro modo mediante una proteína efectora del sistema CRISPR activada descrita en el presente documento. El término "constructo de enmascaramiento" también puede denominarse alternativamente "constructo de detección". En determinadas realizaciones de ejemplo, el constructo de enmascaramiento es un constructo de enmascaramiento basado en ARN. El constructo de enmascaramiento basado en ARN comprende un elemento de ARN que puede escindirse mediante una proteína efectora CRISPR. La escisión del elemento de ARN libera agentes o produce cambios conformacionales que permiten que se produzca una señal detectable. A continuación se describen constructos de ejemplo que demuestran cómo puede usarse el elemento de ARN para prevenir o enmascarar la generación de señales detectables y las realizaciones de la invención comprenden variantes de los mismos. Antes de la escisión, o cuando el constructo de enmascaramiento está en un estado "activo", el constructo de enmascaramiento bloquea la generación o detección de una señal detectable positiva. Se entenderá que en determinadas realizaciones de ejemplo puede producirse una señal de fondo mínima en presencia de un constructo de enmascaramiento de ARN activo. Una señal detectable positiva puede ser cualquier señal que pueda detectarse usando métodos de detección ópticos, fluorescentes, quimioluminiscentes, electroquímicos u otros conocidos en la técnica. El término "señal detectable positiva" se usa para diferenciar de otras señales detectables que pueden ser detectables en presencia del constructo de enmascaramiento. Por ejemplo, en determinadas realizaciones, puede detectarse una primera señal cuando el agente de enmascaramiento está presente (es decir, una señal detectable negativa), que luego se convierte en una segunda señal (por ejemplo, la señal detectable positiva) tras la detección de las moléculas diana y la escisión o desactivación del agente de enmascaramiento por la proteína efectora CRISPR activada.

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede suprimir la generación de un

producto génico. El producto génico puede estar codificado por un constructo indicador que se añade a la muestra. El constructo de enmascaramiento puede ser un ARN de interferencia implicado en una ruta de interferencia de ARN, tal como un ARN de horquilla corta (ARNhc) o un ARN interferente pequeño (ARNip). El constructo de enmascaramiento también puede comprender microARN (miARN). Mientras está presente, el constructo de enmascaramiento suprime la expresión del producto génico. El producto génico puede ser una proteína fluorescente u otro transcrito de ARN o proteínas que de otro modo serían detectables por una sonda marcada, un aptámero o un anticuerpo de no ser por la presencia del constructo de enmascaramiento. Tras la activación de la proteína efectora, el constructo de enmascaramiento se escinde o se silencia de otro modo, lo que permite la expresión y detección del producto génico como señal detectable positiva.

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede secuestrar uno o más reactivos necesarios para generar una señal positiva detectable de manera que la liberación de uno o más reactivos del constructo de enmascaramiento da como resultado la generación de la señal positiva detectable. El uno o más reactivos pueden combinarse para producir una señal colorimétrica, una señal quimioluminiscente, una señal fluorescente o cualquier otra señal detectable y pueden comprender cualquier reactivo que se sepa que es adecuado para tales fines. En determinadas realizaciones de ejemplo, uno o más reactivos son secuestrados por aptámeros de ARN que se unen al uno o más reactivos. El uno o más reactivos se liberan cuando la proteína efectora se activa tras la detección de una molécula diana y los aptámeros de ARN se degradan.

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede inmovilizarse sobre un sustrato sólido en un volumen discreto individual (definido adicionalmente a continuación) y secuestrar un único reactivo. Por ejemplo, el reactivo puede ser una perla que comprende un colorante. Cuando son secuestradas por el reactivo inmovilizado, las perlas individuales son demasiado difusas como para generar una señal detectable, pero al liberarse del constructo de enmascaramiento pueden generar una señal detectable, por ejemplo, por agregación o por simple aumento en la concentración en disolución. En determinadas realizaciones de ejemplo, el agente de enmascaramiento inmovilizado es un aptámero basado en ARN que puede ser escindido por la proteína efectora activada tras la detección de una molécula diana.

En otras determinadas realizaciones de ejemplo, el constructo de enmascaramiento se une a un reactivo inmovilizado en disolución bloqueando así la capacidad del reactivo para unirse a una pareja de unión marcada separada que está libre en disolución. Por tanto, tras la aplicación de una etapa de lavado a una muestra, la pareja de unión marcada puede eliminarse de la muestra por lavado en ausencia de una molécula diana. Sin embargo, si se activa la proteína efectora, el constructo de enmascaramiento se escinde en un grado suficiente para interferir con la capacidad del constructo de enmascaramiento para unirse al reactivo, lo que permite que la pareja de unión marcada se una al reactivo inmovilizado. Por tanto, la pareja de unión marcada permanece después de la etapa de lavado, lo que indica la presencia de la molécula diana en la muestra. En determinados aspectos, el constructo de enmascaramiento que se une al reactivo inmovilizado es un aptámero de ARN. El reactivo inmovilizado puede ser una proteína y la pareja de unión marcada puede ser un anticuerpo marcado. Alternativamente, el reactivo inmovilizado puede ser estreptavidina y la pareja de unión marcada puede ser biotina marcada. El marcador en la pareja de unión usada en las realizaciones anteriores puede ser cualquier marcador detectable conocido en la técnica. Además, pueden usarse otras parejas de unión conocidas según el diseño general descrito en el presente documento.

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede comprender una ribozima. Las ribozimas son moléculas de ARN que tienen propiedades catalíticas. Las ribozimas, tanto las que se producen de manera natural como las obtenidas por ingeniería, comprenden o consisten en ARN que puede seleccionarse como diana por las proteínas efectoras dadas a conocer en el presente documento. La ribozima puede seleccionarse o obtenerse por ingeniería para catalizar una reacción que o bien genere una señal detectable negativa o bien impida la generación de una señal de control positiva. Tras la desactivación de la ribozima por la proteína efectora activada, la reacción que genera una señal de control negativa, o que impide la generación de una señal detectable positiva, se elimina permitiendo así que se genere una señal detectable positiva. En una realización de ejemplo, la ribozima puede catalizar una reacción colorimétrica que provoca que aparezca una disolución como primer color. Cuando se desactiva la ribozima, la disolución cambia a un segundo color, siendo el segundo color la señal positiva detectable. Un ejemplo de cómo pueden usarse las ribozimas para catalizar una reacción colorimétrica se describe en Zhao *et al.* "Signal amplification of glucosamine-6-phosphate based on ribozyme glmS", Biosens Bioelectron. 2014; 16:337-42, y proporcionar un ejemplo de cómo puede modificarse un sistema de este tipo para que funcione en el contexto de las realizaciones descritas en el presente documento. Alternativamente, las ribozimas, cuando están presentes, pueden generar productos de escisión de, por ejemplo, transcritos de ARN. Por tanto, la detección de una señal detectable positiva puede comprender la detección de transcritos de ARN no escindidos que sólo se generan en ausencia de la ribozima.

En determinadas realizaciones de ejemplo, el uno o más reactivos es una proteína, tal como una enzima, que puede facilitar la generación de una señal detectable, tal como una señal colorimétrica, quimioluminiscente o fluorescente, que se inhibe o se secuestra de tal manera que la proteína no puede generar la señal detectable por la unión de uno o más aptámeros de ARN a la proteína. Tras la activación de las proteínas efectoras dadas a conocer en el presente documento, los aptámeros de ARN se escinden o se degradan hasta el punto de que ya no inhiben la capacidad de la proteína para generar la señal detectable. En determinadas realizaciones de ejemplo, el aptámero es un aptámero

inhibidor de trombina. En determinadas realizaciones de ejemplo, el aptámero inhibidor de trombina tiene una secuencia de GGGAACAAAGCUGAAGUACUUACCC (SEQ ID NO: 414). Cuando se escinde este aptámero, la trombina se activará y escindirá un sustrato peptídico colorimétrico o fluorescente. En determinadas realizaciones de ejemplo, el sustrato colorimétrico es para-nitroanilida (pNA) unida covalentemente al sustrato peptídico para la trombina. Tras la escisión por la trombina, se libera pNA y se vuelve de color amarillo y fácilmente visible a simple vista. En determinadas realizaciones de ejemplo, el sustrato fluorescente es 7-amino-4-metilcumarina, un fluoróforo azul que puede detectarse usando un detector de fluorescencia. Los aptámeros inhibidores también pueden usarse para la peroxidasa de rábano picante (HRP), la beta-galactosidasa o la fosfatasa alcalina de ternera (CAP) y dentro de los principios generales establecidos anteriormente.

En determinadas realizaciones, la actividad ARNasa se detecta colorimétricamente a través de la escisión de aptámeros inhibidores de enzimas. Un modo potencial de convertir la actividad ARNasa en una señal colorimétrica es acoplar la escisión de un aptámero de ARN con la reactivación de una enzima que puede producir una salida colorimétrica. En ausencia de escisión de ARN, el aptámero intacto se unirá a la enzima diana e inhibirá su actividad. La ventaja de este sistema de lectura es que la enzima proporciona una etapa de amplificación adicional: una vez que se libera de un aptámero a través de la actividad colateral (por ejemplo, la actividad colateral de Cas13a), la enzima colorimétrica seguirá produciendo un producto colorimétrico, lo que dará lugar a una multiplicación de la señal.

En determinadas realizaciones, se usa un aptámero existente que inhibe una enzima con una lectura colorimétrica. Existen diversos pares de aptámeros/enzimas con lecturas colorimétricas, tales como trombina, proteína C, elastasa de neutrófilos y subtilisina. Estas proteasas tienen sustratos colorimétricos basados en pNA y están disponibles comercialmente. En determinadas realizaciones, se usa un nuevo aptámero que selecciona como diana una enzima colorimétrica común. Las enzimas comunes y robustas, tales como la beta-galactosidasa, la peroxidasa de rábano picante o la fosfatasa alcalina intestinal bovina, pueden seleccionarse como diana por aptámeros obtenidos por ingeniería y diseñados mediante estrategias de selección tales como SELEX. Tales estrategias permiten una selección rápida de aptámeros con eficiencias de unión nanomolar y pueden usarse para el desarrollo de pares de enzimas/aptámeros adicionales para la lectura colorimétrica.

En determinadas realizaciones, la actividad ARNasa se detecta colorimétricamente a través de la escisión de inhibidores anclados a ARN. Muchas enzimas colorimétricas comunes tienen inhibidores reversibles competitivos: por ejemplo, la galactosa puede inhibir la beta-galactosidasa. Muchos de estos inhibidores son débiles, pero su efecto puede aumentarse mediante aumentos de la concentración local. Al vincular la concentración local de inhibidores con la actividad ARNasa, los pares de enzimas colorimétricas e inhibidores pueden modificarse por ingeniería para dar sensores de ARNasa. El sensor colorimétrico de ARNasa basado en inhibidores de molécula pequeña consiste en tres componentes: la enzima colorimétrica, el inhibidor y un ARN en puente que está unido covalentemente tanto al inhibidor como a la enzima, anclando el inhibidor a la enzima. En la configuración no escindida, la enzima es inhibida por el aumento de la concentración local de la molécula pequeña; cuando se escinde el ARN (por ejemplo, mediante la escisión colateral de Cas13a), se liberará el inhibidor y se activará la enzima colorimétrica.

En determinadas realizaciones, la actividad ARNasa se detecta colorimétricamente a través de la formación y/o activación de cuádruples G. Los cuádruples G en el ADN pueden formar complejos con hemo (hierro (III)-protoporfirina IX) para formar una ADNzima con actividad peroxidasa. Cuando se suministra con un sustrato de peroxidasa (por ejemplo, ABTS: (sal diamónica del ácido 2,2'-azinobis[3-etilbenzotiazolin-6-ásulfónico]), el complejo cuádruple G-hemo en presencia de peróxido de hidrógeno provoca la oxidación del sustrato, que luego forma un color verde en disolución. Un ejemplo de secuencia de ADN que forma cuádruple G es: GGGTAGGGCGGGTTGGGA (SEQ. I.D. No. 415). Mediante la hibridación de una secuencia de ARN con este aptámero de ADN, se limitará la formación de la estructura cuádruple G. Tras la activación colateral de ARNasa (por ejemplo, activación colateral del complejo C2c2), el elemento básico de ARN se escindirá, lo que permitirá que se forme el cuádruple G y que se una el grupo hemo. Esta estrategia es particularmente atractiva porque la formación de color es enzimática, lo que significa que hay una amplificación adicional más allá de la activación de la ARNasa.

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede inmovilizarse sobre un sustrato sólido en un volumen discreto individual (definido adicionalmente a continuación) y secuestrar un único reactivo. Por ejemplo, el reactivo puede ser una perla que comprende un colorante. Cuando son secuestradas por el reactivo inmovilizado, las perlas individuales son demasiado difusas como para generar una señal detectable, pero al liberarse del constructo de enmascaramiento pueden generar una señal detectable, por ejemplo, por agregación o por simple aumento en la concentración en disolución. En determinadas realizaciones de ejemplo, el agente de enmascaramiento inmovilizado es un aptámero basado en ARN que puede ser escindido por la proteína efectora activada tras la detección de una molécula diana.

En una realización de ejemplo, el constructo de enmascaramiento comprende un agente de detección que cambia de color dependiendo de si el agente de detección se agrega o se dispersa en disolución. Por ejemplo, determinadas nanopartículas, tales como oro coloidal, experimentan un cambio de color visible de morado a rojo a medida que pasan de los agregados a las partículas dispersas. Por consiguiente, en determinadas realizaciones de ejemplo, tales agentes de detección pueden mantenerse en agregado por una o más moléculas puente. Véase, por ejemplo,

la figura 43. Al menos una porción de la molécula puente comprende ARN. Tras la activación de las proteínas efectoras descritas en el presente documento, la porción de ARN de la molécula puente se escinde permitiendo que el agente de detección se disperse y dando como resultado el cambio de color correspondiente. Véase, por ejemplo, la figura 46. En determinadas realizaciones de ejemplo, la molécula puente es una molécula de ARN. En determinadas realizaciones de ejemplo, el agente de detección es un metal coloidal. El material metálico coloidal puede incluir partículas metálicas insolubles en agua o compuestos metálicos dispersos en un líquido, un hidrosol o un sol metálico. El metal coloidal puede seleccionarse de los metales de los grupos IA, IB, IIB y IIIB de la tabla periódica, así como los metales de transición, especialmente los del grupo VIII. Los metales preferidos incluyen oro, plata, aluminio, rutenio, zinc, hierro, níquel y calcio. Otros metales adecuados también incluyen los siguientes en todos sus diversos estados de oxidación: litio, sodio, magnesio, potasio, escandio, titanio, vanadio, cromo, manganeso, cobalto, cobre, galio, estroncio, niobio, molibdeno, paladio, indio, estaño, tungsteno, renio, platino y gadolinio. Los metales se proporcionan preferiblemente en forma iónica, derivada de un compuesto metálico apropiado, por ejemplo, los iones Al^{3+} , Ru^{3+} , Zn^{2+} , Fe^{3+} , Ni^{2+} y Ca^{2+} .

Cuando el efector CRISPR activado corta el puente de ARN, se observa el cambio de color mencionado anteriormente. En determinadas realizaciones de ejemplo, las partículas son metales coloidales. En otras determinadas realizaciones de ejemplo, el metal coloidal es un oro coloidal. En determinadas realizaciones de ejemplo, las nanopartículas coloidales son nanopartículas de oro (AuNP) de 15 nm. Debido a las propiedades superficiales únicas de las nanopartículas de oro coloidal, la absorbancia máxima se observa a 520 nm cuando se dispersan por completo en disolución y aparecen de color rojo a simple vista. Tras la agregación de las AuNP, presentan un desplazamiento hacia el rojo en la absorbancia máxima y aparecen de color más oscuro, y finalmente se precipitan de la disolución como un agregado de color morado oscuro. En determinadas realizaciones de ejemplo, las nanopartículas se modifican para incluir ligadores de ADN que se extienden desde la superficie de la nanopartícula. Las partículas individuales están unidas entre sí por puentes de ARN monocatenario (ARNmc) que se hibridan en cada extremo del ARN con al menos una parte de los ligadores de ADN. Por tanto, las nanopartículas formarán una red de partículas unidas y se agregan, apareciendo como un precipitado oscuro. Tras la activación de los efectores CRISPR dados a conocer en el presente documento, el puente de ARNmc se escindirá, liberando las AuNP de la red unida y produciendo un color rojo visible. A continuación se enumeran ejemplos de ligadores de ADN y secuencias puente de ARN. Los ligadores de tiol en el extremo de los ligadores de ADN pueden usarse para la conjugación superficial con las AuNP. Pueden usarse otras formas de conjugación. En determinadas realizaciones de ejemplo, pueden generarse dos poblaciones de AuNP, una para cada ligador de ADN. Esto ayudará a facilitar la unión apropiada del puente de ARNmc con la orientación apropiada. En determinadas realizaciones de ejemplo, un primer ligador de ADN se conjuga por el extremo 3' mientras que un segundo ligador de ADN se conjuga por el extremo 5'.

ADN1 colorimétrico C2c2	TTATAACTATTCTAAAAA/3TioMC3 -D/ (SEQ. I.D. No. 183)
ADN2 colorimétrico C2c2	/5TioMC6- D/AAAAAACTCCCCTAATAACAAT (SEQ. I.D. No. 184)
Puente colorimétrico C2c2	GGGUAGGAAUAGUUAUAAUUUCCCUUCCCCA UUGUUAUUAGGGAG (SEQ. I.D. No. 185)

En otras determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede comprender un oligonucleótido de ARN al cual se unen un marcador detectable y un agente de enmascaramiento de ese marcador detectable. Un ejemplo de un par de marcador detectable/agente de enmascaramiento de este tipo es un fluoróforo y un extintor del fluoróforo. La extinción del fluoróforo puede producirse como resultado de la formación de un complejo no fluorescente entre el fluoróforo y otro fluoróforo o molécula no fluorescente. Este mecanismo se conoce como formación de complejos de estado fundamental, extinción estática o extinción por contacto. Por consiguiente, el oligonucleótido de ARN puede diseñarse de modo que el fluoróforo y el extintor estén lo suficientemente cerca uno del otro para que se produzca la extinción por contacto. Los fluoróforos y sus extintores afines son conocidos en la técnica y pueden seleccionarse para este fin por un experto habitual en la técnica. El par de fluoróforo/extintor particular no es crítico en el contexto de esta invención, sólo que la selección de los pares de fluoróforo/extintor asegura el enmascaramiento del fluoróforo. Tras la activación de las proteínas efectoras dadas a conocer en el presente documento, el oligonucleótido de ARN se escinde alterando así la proximidad entre el fluoróforo y el extintor necesaria para mantener el efecto de extinción por contacto. Por consiguiente, la detección del fluoróforo puede usarse para determinar la presencia de una molécula diana en una muestra.

En otras determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede comprender uno o más oligonucleótidos de ARN a los cuales se unen una o más nanopartículas metálicas, tales como nanopartículas de oro. En algunas realizaciones, el constructo de enmascaramiento comprende una pluralidad de nanopartículas metálicas reticuladas por una pluralidad de oligonucleótidos de ARN formando un bucle cerrado. En una realización, el constructo de enmascaramiento comprende tres nanopartículas de oro reticuladas por tres oligonucleótidos de ARN formando un bucle cerrado. En algunas realizaciones, la escisión de los oligonucleótidos de ARN por la proteína efectora CRISPR conduce a una señal detectable producida por las nanopartículas metálicas.

En otras determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede comprender uno o más oligonucleótidos de ARN a los cuales se unen uno o más puntos cuánticos. En algunas realizaciones, la escisión de los oligonucleótidos de ARN por la proteína efectora CRISPR conduce a una señal detectable producida por los puntos cuánticos.

En una realización de ejemplo, el constructo de enmascaramiento puede comprender un punto cuántico. El punto cuántico puede tener múltiples moléculas ligadoras unidas a la superficie. Al menos una parte de la molécula ligadora comprende ARN. La molécula ligadora está unida al punto cuántico en un extremo y a uno o más extintores a lo largo de la longitud de o en los extremos terminales del ligador, de manera que los extintores se mantengan lo suficientemente cerca para que se produzca la extinción del punto cuántico. El ligador puede estar ramificado. Como anteriormente, el par de punto cuántico/extintor no es crítico, sólo que la selección del par de punto cuántico/extintor asegura el enmascaramiento del fluoróforo. Los puntos cuánticos y sus extintores afines son conocidos en la técnica y pueden seleccionarse para este fin por un experto habitual en la técnica. Tras la activación de las proteínas efectoras dadas a conocer en el presente documento, la porción de ARN de la molécula ligadora se escinde alterando así la proximidad entre el punto cuántico y uno o más extintores necesarios para mantener el efecto de extinción. En determinadas realizaciones de ejemplo, el punto cuántico está conjugado con estreptavidina. Los ARN se unen a través de ligadores de biotina y reclutan moléculas de extinción con las secuencias */5Biosg/UCUCGUACGUUC/31AbRQSp/* (SEQ ID NO. 416) o */5Biosg/UCUCGUACGUUC/31AbRQSp/* (SEQ ID NO. 417), donde */5Biosg/* es una etiqueta de biotina y */31AbRQSp/* es un extintor negro de Iowa. Tras la escisión por los efectores activados descritos en el presente documento, el punto cuántico emitirá una fluorescencia visible.

Del mismo modo, la transferencia de energía de fluorescencia (FRET) puede usarse para generar una señal positiva detectable. FRET es un proceso no radiactivo mediante el cual un fotón de un fluoróforo excitado energéticamente (es decir, "fluoróforo donante") eleva el estado de energía de un electrón en otra molécula (es decir, "el aceptor") a niveles vibratorios más altos del estado de singlete excitado. El fluoróforo donante vuelve al estado fundamental sin emitir una fluorescencia característica de ese fluoróforo. El aceptor puede ser otro fluoróforo o una molécula no fluorescente. Si el aceptor es un fluoróforo, la energía transferida se emite como una fluorescencia característica de ese fluoróforo. Si el aceptor es una molécula no fluorescente, la energía absorbida se pierde en forma de calor. Por tanto, en el contexto de las realizaciones descritas en el presente documento, el par de fluoróforo/extintor se reemplaza con un par de fluoróforo donante/aceptor unido a la molécula de oligonucleótido. Cuando está intacto, el constructo de enmascaramiento genera una primera señal (señal detectable negativa) detectada por la fluorescencia o el calor emitido por el aceptor. Tras la activación de las proteínas efectoras dadas a conocer en el presente documento, el oligonucleótido de ARN se escinde y FRET se interrumpe de tal manera que ahora se detecta la fluorescencia del fluoróforo donante (señal detectable positiva).

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento comprende el uso de colorantes intercalados que cambian su absorbancia en respuesta a la escisión de ARN largos para dar nucleótidos cortos. Existen varios colorantes de este tipo. Por ejemplo, la pironina-Y formará un complejo con el ARN y formará un complejo que tiene una absorbancia de 572 nm. La escisión del ARN da como resultado una pérdida de absorbancia y un cambio de color. El azul de metileno puede usarse del mismo modo, con cambios en la absorbancia a 688 nm tras la escisión del ARN. Por consiguiente, en determinadas realizaciones de ejemplo, el constructo de enmascaramiento comprende un complejo de colorante intercalado y ARN que cambia la absorbancia tras la escisión del ARN por las proteínas efectoras dadas a conocer en el presente documento.

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede comprender un iniciador para una reacción de HCR. Véase, por ejemplo, Dirks y Pierce. PNAS 101, 15275-15278 (2004). Las reacciones de HCR utilizan la energía potencial en dos especies de horquilla. Cuando un iniciador monocatenario que tiene una porción complementaria a una región correspondiente en una de las horquillas se libera en la mezcla previamente estable, abre una horquilla de una especie. Este proceso, a su vez, expone una región monocatenaria que abre una horquilla de las otras especies. Este proceso, a su vez, expone una región monocatenaria idéntica al iniciador original. La reacción en cadena resultante puede conducir a la formación de una doble hélice cortada que crece hasta que se agota el suministro de horquilla. La detección de los productos resultantes puede realizarse en un gel o colorimétricamente. Los ejemplos de métodos de detección colorimétrica incluyen, por ejemplo, los dados a conocer en Lu *et al.* "Ultra-sensitive colorimetric assay system based on the hybridization chain reaction-triggered enzyme cascade amplification ACS Appl Mater Interfaces, 2017, 9(1):167-175, Wang *et al.* "An enzyme-free colorimetric assay using hybridization chain reaction amplification and split aptamers" Analyst 2015, 150, 7657-7662, y Song *et al.* "Non covalent fluorescent labelling of hairpin DNA probe coupled with hybridization chain reaction for sensitive DNA detection". Applied Spectroscopy, 70(4): 686-694 (2016).

En determinadas realizaciones de ejemplo, el constructo de enmascaramiento puede comprender una secuencia iniciadora de HCR y un elemento estructural escindible, tal como un bucle o una horquilla, que evita que el iniciador inicie la reacción de HCR. Tras la escisión del elemento estructural por una proteína efectora CRISPR activada, el iniciador se libera entonces para desencadenar la reacción de HCR, indicando su detección la presencia de una o más dianas en la muestra. En determinadas realizaciones de ejemplo, el constructo de enmascaramiento comprende una horquilla con un bucle de ARN. Cuando una proteína efectora CRISPR activada corta el bucle de ARN, el iniciador puede liberarse para desencadenar la reacción de HCR.

AMPLIFICACIÓN DE LA DIANA

En determinadas realizaciones de ejemplo, los ARN y/o ADN diana pueden amplificarse antes de activar la proteína efectora CRISPR. Puede usarse cualquier técnica de amplificación de ARN o ADN adecuada. En determinadas realizaciones de ejemplo, la amplificación de ARN o ADN es una amplificación isotérmica. En determinadas realizaciones de ejemplo, la amplificación isotérmica puede ser amplificación basada en secuencias de ácidos nucleicos (NASBA), amplificación por recombinasa y polimerasa (RPA), amplificación isotérmica mediada por bucle (LAMP), amplificación por desplazamiento de cadena (SDA), amplificación dependiente de helicasa (HDA), o reacción de amplificación de la enzima de corte (NEAR). En determinadas realizaciones de ejemplo, pueden usarse métodos de amplificación no isotérmicos que incluyen, pero no se limitan a, PCR, amplificación de desplazamiento múltiple (MDA), amplificación de círculo rodante (RCA), reacción en cadena de ligasa (LCR) o método de amplificación de ramificación (RAM).

En determinadas realizaciones de ejemplo, la amplificación de ARN o ADN es NASBA, que se inicia con la transcripción inversa del ARN diana mediante un cebador inverso específico de secuencia para crear un dúplex de ARN/ADN. La ARNasa H se usa luego para degradar el molde de ARN, lo que permite que un cebador directo que contiene un promotor, tal como el promotor T7, se una e inicie el alargamiento de la cadena complementaria, generando un producto de ADN bicatenario. La transcripción mediada por el promotor de la ARN polimerasa del molde de ADN crea entonces copias de la secuencia de ARN diana. Es importante destacar que cada uno de los nuevos ARN diana puede detectarse por los ARN guía, lo que mejora aún más la sensibilidad del ensayo. La unión de los ARN diana por los ARN guía conduce entonces a la activación de la proteína efectora CRISPR y los métodos proceden tal como se señaló anteriormente. La reacción de NASBA tiene la ventaja adicional de poder proceder en condiciones isotérmicas moderadas, por ejemplo, a aproximadamente 41°C, lo que la hace adecuada para sistemas y dispositivos implementados para la detección temprana y directa en el campo y lejos de los laboratorios clínicos.

En otras determinadas realizaciones de ejemplo, puede usarse una reacción de amplificación por recombinasa y polimerasa (RPA) para amplificar los ácidos nucleicos diana. Las reacciones de RPA emplean recombinasas que pueden emparejar cebadores específicos de secuencia con secuencia homóloga en ADN dúplex. Si el ADN diana está presente, se inicia la amplificación del ADN y no se requiere ninguna otra manipulación de la muestra, como termociclado o fusión química. Todo el sistema de amplificación de RPA es estable como formulación seca y puede transportarse de manera segura sin refrigeración. Las reacciones de RPA también pueden llevarse a cabo a temperaturas isotérmicas con una temperatura de reacción óptima de 37-42°C. Los cebadores específicos de secuencia están diseñados para amplificar una secuencia que comprende la secuencia de ácido nucleico diana que va a detectarse. En determinadas realizaciones de ejemplo, se añade un promotor de la ARN polimerasa, tal como un promotor T7, a uno de los cebadores. Esto da como resultado un producto de ADN bicatenario amplificado que comprende la secuencia diana y un promotor de la ARN polimerasa. Después o durante la reacción de RPA, se añade una ARN polimerasa que producirá ARN a partir de los moldes de ADN bicatenario. El ARN diana amplificado puede detectarse a su vez por el sistema efector CRISPR. De esta forma, el ADN diana puede detectarse usando las realizaciones descritas en el presente documento. Las reacciones de RPA también pueden usarse para amplificar el ARN diana. El ARN diana se convierte en primer lugar en ADNc usando una transcriptasa inversa, seguido de la síntesis de ADN de segunda hebra en cuyo punto la reacción de RPA avanza tal como se describió anteriormente.

Según la invención, los sistemas dados a conocer en el presente documento incluyen reactivos de amplificación. En el presente documento se describen diferentes componentes o reactivos útiles para la amplificación de ácidos nucleicos. Por ejemplo, un reactivo de amplificación tal como se describe en el presente documento puede incluir un tampón, tal como un tampón Tris. Puede usarse un tampón Tris en cualquier concentración apropiada para la aplicación o uso deseado, por ejemplo, pero sin limitarse a, una concentración de 1 mM, 2 mM, 3 mM, 4 mM, 5 mM, 6 mM, 7 mM, 8 mM, 9 mM, 10 mM, 11 mM, 12 mM, 13 mM, 14 mM, 15 mM, 25 mM, 50 mM, 75 mM, 1 M o similares. Un experto en la técnica podrá determinar una concentración apropiada de un tampón tal como Tris para su uso con la presente invención.

Una sal, tal como cloruro de magnesio (MgCl₂), cloruro de potasio (KCl) o cloruro de sodio (NaCl), pueden incluirse en una reacción de amplificación, tal como PCR, para mejorar la amplificación de fragmentos de ácido nucleico. Aunque la concentración de sal dependerá de la aplicación y reacción particular, en algunas realizaciones, los fragmentos de ácido nucleico de un tamaño particular pueden producir resultados óptimos en concentraciones de sal particulares. Los productos más grandes pueden requerir concentraciones de sal alteradas, normalmente menos sal, para producir los resultados deseados, mientras que la amplificación de productos más pequeños puede producir mejores resultados con concentraciones de sal más altas. Un experto en la técnica entenderá que la presencia y/o

concentración de una sal, junto con la alteración de las concentraciones de sal, puede alterar la rigurosidad de una reacción biológica o química y, por tanto, puede usarse cualquier sal que proporcione las condiciones apropiadas para una reacción de la presente invención y tal como se describe en el presente documento.

Otros componentes de una reacción biológica o química pueden incluir un componente de lisis celular para romper o lisar una célula para el análisis de los materiales en la misma. Un componente de lisis celular puede incluir, pero no se limita a, un detergente, una sal tal como se describió anteriormente, tal como NaCl, KCl, sulfato de amonio $[(\text{NH}_4)_2\text{SO}_4]$, u otras. Los detergentes que pueden ser apropiados para la invención pueden incluir Triton X-100, dodecilsulfato de sodio (SDS), CHAPS (3-[(3-colamidopropil)dimetilamonio]-1-propanosulfonato), bromuro de etiltrimetilamonio, nonilfenoxipolietoxietanol (NP-40). Las concentraciones de detergentes pueden depender de la aplicación particular y pueden ser específicas para la reacción en algunos casos. Las reacciones de amplificación pueden incluir dNTP y cebadores de ácido nucleico usados en cualquier concentración adecuada para la invención, tal como incluyendo, pero no se limita a, una concentración de 100 nM, 150 nM, 200 nM, 250 nM, 300 nM, 350 nM, 400 nM, 450 nM, 500 nM, 550 nM, 600 nM, 650 nM, 700 nM, 750 nM, 800 nM, 850 nM, 900 nM, 950 nM, 1 mM, 2 mM, 3 mM, 4 mM, 5 mM, 6 mM, 7 mM, 8 mM, 9 mM, 10 mM, 20 mM, 30 mM, 40 mM, 50 mM, 60 mM, 70 mM, 80 mM, 90 mM, 100 mM, 150 mM, 200 mM, 250 mM, 300 mM, 350 mM, 400 mM, 450 mM, 500 mM o similares. Asimismo, una polimerasa útil según la invención puede ser cualquier polimerasa específica o general conocida en la técnica y útil para la invención, incluyendo polimerasa Taq, polimerasa Q5 o similares.

En algunas realizaciones, los reactivos de amplificación tal como se describe en el presente documento pueden ser apropiados para su uso en la amplificación de inicio en caliente. La amplificación de inicio en caliente puede ser beneficiosa en algunas realizaciones para reducir o eliminar la dimerización de moléculas u oligos adaptadores, o para evitar de otro modo productos o artefactos de amplificación no deseados y obtener una amplificación óptima del producto deseado. Muchos componentes descritos en el presente documento para su uso en la amplificación también pueden usarse en la amplificación de inicio en caliente. En algunas realizaciones, pueden usarse reactivos o componentes apropiados para su uso con amplificación de inicio en caliente en lugar de uno o más de los componentes de la composición, según corresponda. Por ejemplo, puede usarse una polimerasa u otro reactivo que muestre una actividad deseada a una temperatura particular u otra condición de reacción. En algunas realizaciones, pueden usarse reactivos que están diseñados u optimizados para su uso en la amplificación de inicio en caliente, por ejemplo, puede activarse una polimerasa después de la transposición o después de alcanzar una temperatura particular. Tales polimerasas pueden estar basadas en anticuerpos o basadas en aptámeros. Las polimerasas tal como se describen en el presente documento son conocidas en la técnica. Los ejemplos de tales reactivos pueden incluir, pero no se limitan a, polimerasas de inicio en caliente, dNTP de inicio en caliente y dNTP fotoenjaulados. Tales reactivos son conocidos y están disponibles en la técnica. Un experto en la técnica podrá determinar las temperaturas óptimas, según corresponda, para los reactivos individuales.

La amplificación de ácidos nucleicos puede realizarse usando maquinaria o equipo de termociclado específico, y puede realizarse en reacciones únicas o en masa, de manera que puede realizarse cualquier número deseado de reacciones simultáneamente. En algunas realizaciones, la amplificación puede realizarse usando dispositivos microfluídicos o robóticos, o puede realizarse usando la alteración manual de las temperaturas para lograr la amplificación deseada. En algunas realizaciones, puede realizarse la optimización para obtener las condiciones de reacción óptimas para la aplicación o los materiales particulares. Un experto en la técnica comprenderá y podrá optimizar las condiciones de reacción para obtener una amplificación suficiente.

En determinadas realizaciones, la detección de ADN con los métodos o sistemas de la invención requiere la transcripción del ADN (amplificado) en ARN antes de la detección.

ENRIQUECIMIENTO DE ARN/ADN DIANA

En determinadas realizaciones de ejemplo, el ARN o ADN diana puede enriquecerse primero antes de la detección o amplificación del ARN o ADN diana. En determinadas realizaciones de ejemplo, este enriquecimiento puede lograrse mediante la unión de los ácidos nucleicos diana por un sistema efector CRISPR.

Los protocolos actuales de enriquecimiento específicos de diana requieren un ácido nucleico monocatenario antes de la hibridación con las sondas. Entre las diversas ventajas, las presentes realizaciones pueden omitir esta etapa y permitir el direccionamiento directo al ADN bicatenario (o bien parcial o bien completamente bicatenario). Además, las realizaciones dadas a conocer en el presente documento son métodos de direccionamiento impulsados por enzimas que ofrecen una cinética más rápida y un flujo de trabajo más sencillo, lo que permite el enriquecimiento isotérmico. En determinadas realizaciones de ejemplo, el enriquecimiento puede tener lugar a temperaturas tan bajas como 20-37°C. En determinadas realizaciones de ejemplo, se usa un conjunto de ARN guía para diferentes ácidos nucleicos diana en un solo ensayo, lo que permite la detección de múltiples dianas y/o múltiples variantes de una sola diana.

En determinadas realizaciones de ejemplo, una proteína efectora CRISPR desactivada puede unirse al ácido nucleico diana en disolución y luego aislarse posteriormente de dicha disolución. Por ejemplo, la proteína efectora CRISPR desactivada unida al ácido nucleico diana puede aislarse de la disolución usando un anticuerpo u otra molécula, tal como un aptámero, que se une específicamente a la proteína efectora CRISPR desactivada.

En otras realizaciones de ejemplo, la proteína efectora CRISPR desactivada puede unirse a un sustrato sólido. Un sustrato fijo puede referirse a cualquier material que sea apropiado o que pueda modificarse para que sea apropiado para la unión de un polipéptido o un polinucleótido. Los posibles sustratos incluyen, pero no se limitan a, vidrio y vidrio funcionalizado modificado, plásticos (incluyendo acrílicos, poliestireno y copolímeros de estireno y otros materiales, polipropileno, polietileno, polibutileno, poliuretanos, Teflon™, etc.), polisacáridos, nailon o nitrocelulosa, cerámica, resinas, sílice o materiales basados en sílice, incluyendo silicio y silicio modificado, carbono, metales, vidrios inorgánicos, plásticos, haces de fibras ópticas y una variedad de otros polímeros. En algunas realizaciones, el soporte sólido comprende una superficie en patrones adecuada para la inmovilización de moléculas en un patrón ordenado. En determinadas realizaciones, una superficie en patrones se refiere a una disposición de diferentes regiones en o sobre una capa expuesta de un soporte sólido. En algunas realizaciones, el soporte sólido comprende una serie de pocillos o depresiones en una superficie. La composición y geometría del soporte sólido pueden variar con su uso. En algunas realizaciones, el soporte sólido es una estructura plana tal como un portaobjetos, un chip, un microchip y/o una matriz. Como tal, la superficie del sustrato puede tener la forma de una capa plana. En algunas realizaciones, el soporte sólido comprende una o más superficies de una celda de flujo. El término "celda de flujo", tal como se usa en el presente documento, se refiere a una cámara que comprende una superficie sólida a través de la cual pueden fluir uno o más reactivos fluidos. Las celdas de flujo de ejemplo y los sistemas fluidicos relacionados y las plataformas de detección que pueden usarse fácilmente en los métodos de la presente descripción se describen, por ejemplo, en Bentley *et al.* Nature 456: 53-59 (2008), los documentos WO 04/0918497, U.S. 7.057.026; WO 91/06678; WO 07/123744; US 7.329.492; US 7.211.414; US 7.315.019; U.S. 7.405.281, y US 2008/0108082. En algunas realizaciones, el soporte sólido o su superficie no es plano, tal como la superficie interna o externa de un tubo o recipiente. En algunas realizaciones, el soporte sólido comprende microesferas o perlas. Se pretende que "microesferas", "perla", "partículas" signifiquen, dentro del contexto de un sustrato sólido, pequeñas partículas discretas elaboradas de diversos materiales que incluyen, pero no se limitan a, plásticos, cerámica, vidrio y poliestireno. En determinadas realizaciones, las microesferas son microesferas o perlas magnéticas. Alternativa o adicionalmente, las perlas pueden ser porosas. Los tamaños de las perlas oscilan entre nanómetros, por ejemplo, 100 nm, y milímetros, por ejemplo 1 mm.

Una muestra que contiene, o se sospecha que contiene, los ácidos nucleicos diana puede entonces exponerse al sustrato para permitir la unión de los ácidos nucleicos diana a la proteína efectora CRISPR desactivada unida. Luego las moléculas no diana pueden eliminarse por lavado. En determinadas realizaciones de ejemplo, los ácidos nucleicos diana pueden liberarse entonces del complejo proteína efectora CRISPR/ARN guía para una detección adicional usando los métodos dados a conocer en el presente documento. En determinadas realizaciones de ejemplo, los ácidos nucleicos diana pueden amplificarse en primer lugar tal como se describe en el presente documento.

En determinadas realizaciones de ejemplo, el efector CRISPR puede marcarse con una etiqueta de unión. En determinadas realizaciones de ejemplo, el efector CRISPR puede marcarse químicamente. Por ejemplo, el efector CRISPR puede biotilarse químicamente. En otra realización de ejemplo, puede crearse una fusión agregando una secuencia adicional que codifica para una fusión al efector CRISPR. Un ejemplo de tal fusión es AviTag™, que emplea una conjugación enzimática altamente dirigida de una sola biotina en una etiqueta peptídica de 15 aminoácidos única. En determinadas realizaciones, el efector CRISPR puede marcarse con una etiqueta de captura tal como, pero sin limitarse a, GST, Myc, hemaglutinina (HA), proteína fluorescente verde (GFP), flag, etiqueta His, etiqueta TAP y etiqueta Fc. La etiqueta de unión, ya sea una fusión, una etiqueta química o una etiqueta de captura, puede usarse o bien para atraer el sistema efector CRISPR una vez que se ha unido a un ácido nucleico diana o bien para fijar el sistema efector CRISPR en el sustrato sólido.

En determinadas realizaciones de ejemplo, el ARN guía puede marcarse con una etiqueta de unión. En determinadas realizaciones de ejemplo, el ARN guía completo puede marcarse usando transcripción *in vitro* (IVT) incorporando uno o más nucleótidos biotinilados, tales como, uracilo biotinilado. En algunas realizaciones, la biotina puede añadirse química o enzimáticamente al ARN guía, tal como la adición de uno o más grupos de biotina al extremo 3' del ARN guía. La etiqueta de unión puede usarse para atraer el complejo ARN guía/ácido nucleico diana después de que se haya producido la unión, por ejemplo, exponiendo el ARN guía/ácido nucleico diana a un sustrato sólido recubierto de estreptavidina.

Por consiguiente, en determinadas realizaciones de ejemplo, puede usarse un efector CRISPR obtenido por ingeniería o que no se produce de manera natural para fines de enriquecimiento. En una realización, la modificación puede comprender la mutación de uno o más residuos de aminoácidos de la proteína efectora. La una o más mutaciones pueden estar en uno o más dominios catalíticamente activos de la proteína efectora. La proteína efectora puede tener actividad nucleasa reducida o abolida en comparación con una proteína efectora que carece de dicha una o más mutaciones. Es posible que la proteína efectora no dirija la escisión de la cadena de ARN en el locus diana de interés. En una realización preferida, la una o más mutaciones pueden comprender dos mutaciones. En una realización preferida, el uno o más residuos de aminoácidos se modifican en una proteína efectora C2c2, por ejemplo, una proteína efectora o C2c2 obtenida por ingeniería o que no se produce de manera natural. En realizaciones particulares, el uno o más modificados de residuos de aminoácidos mutados son uno o más de aquellos en C2c2 correspondientes a R597, H602, R1278 y H1283 (referidos a los aminoácidos Lsh C2c2), tales como las mutaciones R597A, H602A, R1278A y H1283A, o los residuos de aminoácidos correspondientes en los ortólogos de Lsh C2c2.

En realizaciones particulares, uno o más modificados de residuos de aminoácidos mutados son uno o más de aquellos en C2c2 correspondientes a K2, K39, V40, E479, L514, V518, N524, G534, K535, E580, L597, V602, D630, F676, L709, I713, R717 (HEPN), N718, H722 (HEPN), E773, P823, V828, I879, Y880, F884, Y997, L1001, F1009, L1013, Y1093, L1099, L1111, Y1114, L1203, D1222, Y1244, L1250, L1253, K1261, I1334, L1355, L1359, R1362, Y1366, E1371, R1372, D1373, R1509 (HEPN), H1514 (HEPN), Y1543, D1544, K1546, K1548, V1551, I1558, según la numeración de consenso de C2c2. En determinadas realizaciones, uno o más modificados de residuos de aminoácidos mutados son uno o más de aquellos en C2c2 correspondientes a R717 y R1509. En determinadas realizaciones, el uno o más modificados de residuos de aminoácidos mutados son uno o más de aquellos en C2c2 correspondientes a K2, K39, K535, K1261, R1362, R1372, K1546 y K1548. En determinadas realizaciones, dichas mutaciones dan como resultado una proteína que tiene una actividad alterada o modificada. En determinadas realizaciones, dichas mutaciones dan como resultado una proteína que tiene una actividad reducida, tal como una especificidad reducida. En determinadas realizaciones, dichas mutaciones dan como resultado una proteína que no tiene actividad catalítica (es decir, C2c2 "desactivada"). En una realización, dichos residuos de aminoácidos corresponden a residuos de aminoácidos Lsh C2c2, o los residuos de aminoácidos correspondientes de una proteína C2c2 de una especie diferente. Dispositivos que pueden facilitar estas etapas. En algunas realizaciones, para reducir el tamaño de una proteína de fusión del efector Cas13b y el uno o más dominios funcionales, el extremo C-terminal del efector Cas13b puede truncarse manteniendo al mismo tiempo su función de unión al ARN. Por ejemplo, al menos 20 aminoácidos, al menos 50 aminoácidos, al menos 80 aminoácidos, o al menos 100 aminoácidos, o al menos 150 aminoácidos, o al menos 200 aminoácidos, o al menos 250 aminoácidos, o al menos 300 aminoácidos, o al menos 350 aminoácidos, o hasta 120 aminoácidos, o hasta 140 aminoácidos, o hasta 160 aminoácidos, o hasta 180 aminoácidos, o hasta 200 aminoácidos, o hasta 250 aminoácidos, o hasta 300 aminoácidos, o hasta 350 aminoácidos, o hasta 400 aminoácidos, pueden truncarse en el extremo C-terminal del efector Cas13b. Los ejemplos específicos de truncamientos de Cas13b incluyen Δ 984-1090 C-terminal, Δ 1026-1090 C-terminal y Δ 1053-1090 C-terminal, Δ 934-1090 C-terminal, Δ 884-1090 C-terminal, Δ 834-1090 C-terminal, Δ 784-1090 C-terminal y Δ 734-1090 C-terminal, en los que las posiciones de aminoácidos corresponden a las posiciones de aminoácidos de la proteína P5-125 Cas13b de *Prevotella* sp.

Los sistemas de enriquecimiento anteriores también pueden usarse para agotar una muestra de determinados ácidos nucleicos. Por ejemplo, los ARN guía pueden diseñarse para unirse a ARN no diana para eliminar los ARN no diana de la muestra. En una realización de ejemplo, los ARN guía pueden diseñarse para unirse a ácidos nucleicos que portan una variación particular de ácido nucleico. Por ejemplo, en una muestra dada puede esperarse un mayor número de copias de ácidos nucleicos no variantes. Por consiguiente, las realizaciones descritas en el presente documento pueden usarse para eliminar los ácidos nucleicos no variantes de una muestra, para aumentar la eficacia con la que el sistema efector CRISPR de detección puede detectar las secuencias variantes diana en una muestra determinada.

AMPLIFICACIÓN Y/O POTENCIAMIENTO DE LA SEÑAL POSITIVA DETECTABLE

En determinadas realizaciones de ejemplo, pueden introducirse modificaciones adicionales que amplifiquen adicionalmente la señal positiva detectable. Por ejemplo, la activación colateral de la proteína efectora CRISPR activada puede usarse para generar una diana secundaria o una secuencia guía adicional, o ambas. En una realización de ejemplo, la disolución de reacción contendría una diana secundaria que se añade a una concentración alta. La diana secundaria puede ser distinta de la diana principal (es decir, la diana para la cual el ensayo está diseñada para detectar) y, en determinados casos, puede ser común en todos los volúmenes de reacción. Una secuencia guía secundaria para la diana secundaria puede protegerse, por ejemplo, por una característica estructural secundaria, tal como una horquilla con un bucle de ARN, y ser incapaz de unirse a la segunda diana o la proteína efectora CRISPR. La escisión del grupo protector por una proteína efectora CRISPR activada (es decir, después de la activación mediante la formación de un complejo con la(s) diana(s) principal(es) en disolución) y formación de un complejo con proteína efectora CRISPR libre en disolución y activación a partir de la diana secundaria añadida. En otras determinadas realizaciones de ejemplo, se usa un concepto similar con una segunda secuencia guía para una secuencia diana secundaria. La secuencia diana secundaria puede ser una característica estructural protegida o un grupo protector en la diana secundaria. La escisión de un grupo protector de la diana secundaria permite entonces que se forme un complejo adicional de proteína efectora CRISPR/segunda secuencia guía/diana secundaria. En aún otra realización de ejemplo, la activación de la proteína efectora CRISPR por las dianas principales puede usarse para escindir un cebador protegido o circularizado, que luego se libera para realizar una reacción de amplificación isotérmica, tal como las dadas a conocer en el presente documento, en un molde que codifica para una secuencia guía secundaria, una secuencia diana secundaria o ambas. La transcripción posterior de este molde amplificado produciría más secuencia guía secundaria y/o secuencia diana secundaria, seguida de una activación colateral adicional de la proteína efectora CRISPR.

DETECCIÓN DE PROTEÍNAS (NO SEGÚN LA INVENCION)

Los sistemas, dispositivos y métodos dados a conocer en el presente documento también pueden adaptarse para la detección de polipéptidos (u otras moléculas) además de la detección de ácidos nucleicos, mediante la incorporación de un aptámero de detección de polipéptido configurado específicamente. Los aptámeros de detección de polipéptido son distintos de los aptámeros de constructo de enmascaramiento comentados anteriormente. En primer lugar, los aptámeros están diseñados para unirse específicamente a una o más moléculas diana. En una realización

de ejemplo, la molécula diana es un polipéptido diana. En otro ejemplo de realización, la molécula diana es un compuesto químico diana, tal como una molécula terapéutica diana. Los métodos para diseñar y seleccionar aptámeros con especificidad para una diana dada, tal como SELEX, se conocen en la técnica. Además de la especificidad para una diana dada, los aptámeros se diseñan además para incorporar un sitio de unión al promotor de la ARN polimerasa. En determinadas realizaciones de ejemplo, el promotor de la ARN polimerasa es un promotor T7. Antes de unir el aptámero que se une a una diana, el sitio de ARN polimerasa no es accesible ni reconocible de otro modo por una ARN polimerasa. Sin embargo, el aptámero está configurado de modo que, tras la unión de una diana, la estructura del aptámero sufre un cambio conformacional de modo que el promotor de la ARN polimerasa queda expuesto. Una secuencia de aptámero en el sentido de 3' del promotor de la ARN polimerasa actúa como molde para la generación de un oligonucleótido de ARN desencadenante por una ARN polimerasa. Por tanto, la porción de molde del aptámero puede incorporar además un código de barras u otra secuencia de identificación que identifique un aptámero dado y su diana. Los ARN guía tal como se describieron anteriormente pueden diseñarse entonces para reconocer estas secuencias de oligonucleótidos desencadenantes específicas. La unión de los ARN guía a los oligonucleótidos desencadenantes activa las proteínas efectoras CRISPR que proceden a desactivar los constructos de enmascaramiento y generar una señal detectable positiva tal como se describió anteriormente.

Por consiguiente, en determinadas realizaciones de ejemplo, los métodos dados a conocer en el presente documento comprenden la etapa adicional de distribuir una muestra o conjunto de muestras en un conjunto de volúmenes discretos individuales, comprendiendo cada volumen discreto individual aptámeros de detección de péptidos, una proteína efectora CRISPR, uno o más ARN guía, un constructo de enmascaramiento, e incubar la muestra o conjunto de muestras en condiciones suficientes para permitir la unión de los aptámeros de detección a la una o más moléculas diana, en los que la unión del aptámero a una diana correspondiente da como resultado la exposición del sitio de unión al promotor de la ARN polimerasa de manera que la síntesis de un ARN desencadenante se inicia mediante la unión de una ARN polimerasa al sitio de unión al promotor de la ARN polimerasa.

En otra realización de ejemplo, la unión del aptámero puede exponer un sitio de unión del cebador tras la unión del aptámero a un polipéptido diana. Por ejemplo, el aptámero puede exponer un sitio de unión del cebador de RPA. Por tanto, la adición o inclusión del cebador se incorporará a una reacción de amplificación, tal como la reacción de RPA señalada anteriormente.

En determinadas realizaciones de ejemplo, el aptámero puede ser un aptámero de cambio de conformación, que al unirse a la diana de interés, puede cambiar la estructura secundaria y exponer nuevas regiones de ADN monocatenario. En determinadas realizaciones de ejemplo, estas nuevas regiones de ADN monocatenario pueden usarse como sustratos para el ligamiento, extendiendo los aptámeros y creando moléculas de ADNmc más largas que pueden detectarse específicamente usando las realizaciones descritas en el presente documento. El diseño de aptámero puede combinarse adicionalmente con complejos ternarios para la detección de dianas de bajo epítipo, tales como glucosa (Yang *et al.* 2015: <http://pubs.acs.org/doi/abs/10.1021/acs.analchem.5b01634>). A continuación se muestran aptámeros de cambio de conformación de ejemplo y los correspondientes ARN guía (ARNcr).

Aptámero de trombina	(SEQ. ID. NO. 186)
Sonda de ligamiento de trombina	(SEQ. ID. NO. 187)
Cebador 1 directo de RPA de trombina	(SEQ. ID. NO. 188)
Cebador 2 directo de RPA de trombina	(SEQ. ID. NO. 189)
Cebador 1 inverso de RPA de trombina	(SEQ. ID. NO. 190)
ARNcr 1 de trombina	(SEQ. ID. NO. 191)
ARNcr 2 de trombina	(SEQ. ID. NO. 192)
ARNcr 3 de trombina	(SEQ. ID. NO. 193)
Control de amplicón de longitud completa de PTK7	(SEQ. ID. NO. 194)
Aptámero de PTK7	(SEQ. ID. NO. 195)
Sonda de ligamiento de PTK7	(SEQ. ID. NO. 196)
Cebador 1 directo de RPA PTK7	(SEQ. ID. NO. 197)
Cebador 1 inverso de RPA PTK7	(SEQ. ID. NO. 198)
ARNcr 1 PTK7	(SEQ. ID. NO. 199)
ARNcr 2 PTK7	(SEQ. ID. NO. 200)
ARNcr 3 PTK7	(SEQ. ID. NO. 201)

DISPOSITIVOS

Los sistemas descritos en el presente documento pueden materializarse en dispositivos de diagnóstico. Pueden usarse diversos sustratos y configuraciones. Los dispositivos pueden definir múltiples volúmenes discretos individuales dentro del dispositivo. Tal como se usa en el presente documento, un "volumen discreto individual" se refiere a un espacio discreto, tal como un contenedor, receptáculo u otro volumen o espacio definido que puede definirse por propiedades que previenen y/o inhiben la migración de moléculas diana, por ejemplo, un volumen o espacio definido por propiedades físicas tales como paredes, por ejemplo, las paredes de un pocillo, un tubo o la superficie de una gota, que puede ser impermeable o semipermeable, o definirse por otros medios tales como químicos, velocidad de difusión limitada, electromagnética o iluminación de luz, o cualquier combinación de los mismos, que puede contener una muestra dentro de un espacio definido. Los volúmenes discretos individuales pueden identificarse mediante etiquetas moleculares, tal como códigos de barras de ácido nucleico. Por "tasa de difusión limitada" (por ejemplo, volúmenes definidos por difusión) se entiende espacios que sólo son accesibles para determinadas moléculas o reacciones porque las restricciones de difusión definen efectivamente un espacio o volumen, tal como sería el caso de dos corrientes laminares paralelas donde la difusión limitará la migración de una molécula diana de una corriente a la otra. Por volumen o espacio definido "químico" se entiende espacios en los que sólo determinadas moléculas diana pueden existir debido a sus propiedades químicas o moleculares, como el tamaño, donde, por ejemplo, las perlas de gel pueden impedir que determinadas especies entren en las perlas pero no otras, tal como por carga superficial, tamaño de la matriz u otra propiedad física de la perla que pueda permitir la selección de especies que puedan entrar en el interior de la perla. Por volumen o espacio definido "electromagnéticamente" se entiende espacios en los que las propiedades electromagnéticas de las moléculas diana o sus soportes, tales como la carga o las propiedades magnéticas, pueden usarse para definir determinadas regiones en un espacio, tal como capturar partículas magnéticas dentro de un campo magnético o directamente sobre imanes. Por volumen definido "ópticamente" se entiende cualquier región del espacio que pueda definirse iluminándola con luz visible, ultravioleta, infrarroja u otras longitudes de onda de luz de modo que sólo puedan marcarse moléculas diana dentro del espacio o volumen definido. Una ventaja del uso de volúmenes discretos sin paredes o semipermeables es que algunos reactivos, tales como tampones, activadores químicos u otros agentes, pueden pasar a través del volumen discreto, mientras que otros materiales, tales como las moléculas diana, pueden mantenerse en el volumen o espacio discreto. Normalmente, un volumen discreto incluirá un medio fluido (por ejemplo, una disolución acuosa, un aceite, un tampón y/o un medio que puede soportar el crecimiento celular) adecuado para marcar la molécula diana con el identificador de ácido nucleico indexable bajo condiciones que permitan el marcaje. Los ejemplos de volúmenes o espacios discretos útiles en los métodos dados a conocer incluyen gotas (por ejemplo, gotas microfluídicas y/o gotas de emulsión), perlas de hidrogel u otras estructuras poliméricas (por ejemplo, perlas de diacrilato de polietilenglicol o perlas de agarosa), portaobjetos de tejido (por ejemplo, portaobjetos de tejido embebido en parafina con formalina fija con regiones, volúmenes o espacios particulares definidos por medios químicos, ópticos o físicos), portaobjetos de microscopio con regiones definidas mediante el depósito de reactivos en matrices ordenadas o patrones aleatorios, tubos (tales como tubos de centrifuga, tubos de microcentrifuga, tubos de ensayo, cubetas, tubos cónicos y similares), botellas (tales como botellas de vidrio, botellas de plástico, botellas de cerámica, matraces Erlenmeyer, viales de centelleo y similares), pocillos (tales como pocillos en una placa), placas, pipetas o puntas de pipeta entre otros. En determinadas realizaciones, el compartimento es una gota acuosa en una emulsión de agua en aceite. En realizaciones específicas, cualquiera de las aplicaciones, métodos o sistemas descritos en el presente documento que requieran volúmenes exactos o uniformes pueden emplear el uso de un dispensador de líquido acústico.

En determinadas realizaciones de ejemplo, el dispositivo comprende un sustrato de material flexible sobre el que pueden definirse una serie de puntos. Los materiales de sustrato flexibles adecuados para su uso en diagnóstico y biodetección se conocen en la técnica. Los materiales de sustrato flexibles pueden estar elaborados de fibras derivadas de plantas, tales como fibras celulósicas, o pueden estar elaborados de polímeros flexibles, tales como películas de poliéster flexibles y otros tipos de polímeros. Dentro de cada punto definido, los reactivos del sistema descrito en el presente documento se aplican a los puntos individuales. Cada punto puede contener los mismos reactivos excepto por un ARN guía diferente o un conjunto de ARN guía diferente o, cuando corresponda, un aptámero de detección diferente para detectar múltiples dianas a la vez. Por tanto, los sistemas y dispositivos en el presente documento pueden examinar muestras de múltiples fuentes (por ejemplo, múltiples muestras clínicas de diferentes individuos) para detectar la presencia de la misma diana, o un número limitado de dianas, o alícuotas de una sola muestra (o múltiples muestras de la misma fuente) para detectar la presencia de múltiples dianas diferentes en la muestra. En determinadas realizaciones de ejemplo, los elementos de los sistemas descritos en el presente documento se secan por congelación sobre el sustrato de papel o tela. Los sustratos basados en materiales flexibles de ejemplo que pueden usarse en determinados dispositivos de ejemplo se describen en Pardee *et al.* Cell. 2016, 165(5):1255-66 y Pardee *et al.* Cell. 2014, 159(4):950-54. Los sustratos basados en materiales flexibles adecuados para su uso con líquidos biológicos, incluyendo la sangre, se dan a conocer en la publicación de solicitud de patente internacional n.º WO/2013/071301 titulada "Paper based diagnostic test" de Shevkoplyas *et al.* La publicación de solicitud de patente estadounidense n.º 2011/0111517 titulada "Paper-based microfluidic systems" de Siegel *et al.* y Shafiee *et al.* "Paper and Flexible Substrates as Materials for Biosensing Platforms to Detect Multiple Biotargets" Scientific Reports 5:8719 (2015). Materiales de base flexibles adicionales, incluyendo aquellos adecuados para su uso en dispositivos de diagnóstico portátiles, se dan a conocer en Wang *et al.* "Flexible Substrate-Based Devices for Point-of-Care Diagnostics" Cell 34(11):909-21 (2016). Los materiales de base flexible adicionales pueden incluir nitrocelulosa, policarbonato, metilcelulosa, poli(fluoruro de vinilideno) (PVDF), poliestireno o vidrio (véase, por ejemplo, el documento US20120238008). En determinadas realizaciones, los volúmenes discretos están separados

por una superficie hidrófoba tal como, pero sin limitarse a, cera, resina fotosensible o tinta sólida.

En algunas realizaciones, puede proporcionarse un dosímetro o placa que sirva como sensor o indicador de modo que se notifique al usuario de la exposición a determinados microbios u otros agentes. Por ejemplo, los sistemas descritos en el presente documento pueden usarse para detectar un patógeno particular. Del mismo modo, las realizaciones basadas en aptámeros dadas a conocer anteriormente pueden usarse para detectar tanto polipéptidos como otros agentes, tales como agentes químicos, a los que puede unirse un aptámero específico. Un dispositivo de este tipo puede ser útil para la vigilancia de soldados u otro personal militar, así como de médicos, investigadores, personal hospitalario y similares, con el fin de proporcionar información relacionada con la exposición a agentes potencialmente peligrosos lo más rápido posible, por ejemplo, para la detección biológica o de agentes de guerra química. En otras realizaciones, una placa de vigilancia de este tipo puede usarse para prevenir la exposición a microbios o patógenos peligrosos en pacientes inmunocomprometidos, pacientes quemados, pacientes sometidos a quimioterapia, niños o ancianos.

Las fuentes de muestras que pueden analizarse usando los sistemas y dispositivos descritos en el presente documento incluyen muestras biológicas de un sujeto o muestras ambientales. Las muestras ambientales pueden incluir superficies o líquidos. Las muestras biológicas pueden incluir, pero no se limitan a, saliva, sangre, plasma, sueros, heces, orina, esputo, moco, linfa, líquido sinovial, líquido cefalorraquídeo, un frotis de piel o una membrana mucosa, o una combinación de los mismos. En una realización de ejemplo, la muestra ambiental se toma de una superficie sólida, tal como una superficie usada en la preparación de alimentos u otras composiciones y materiales sensibles.

En otras realizaciones de ejemplo, los elementos de los sistemas descritos en el presente documento pueden colocarse en un sustrato de un solo uso, tal como un hisopo o un paño que se usa para tomar una muestra de una superficie o un fluido de muestra. Por ejemplo, el sistema puede usarse para analizar la presencia de un patógeno en un alimento tomando una muestra de la superficie de un producto alimenticio, tal como una fruta o verdura. Del mismo modo, el sustrato de un solo uso puede usarse para tomar unas muestras de otras superficies para la detección de determinados microbios o agentes, tal como para su uso en el control de seguridad. Los sustratos de un solo uso también pueden tener aplicaciones en medicina forense, donde los sistemas CRISPR están diseñados para detectar, por ejemplo, identificar SNP de ADN que pueden usarse para identificar a un sospechoso, o determinados marcadores de células o tejidos para determinar el tipo de materia biológica presente en una muestra. Asimismo, el sustrato de un solo uso puede usarse para recoger una muestra de un paciente, tal como una muestra de saliva de la boca, o un frotis de piel. En otras realizaciones, puede tomarse una muestra o un frotis de un producto cárnico con el fin de detectar la presencia o ausencia de contaminantes sobre o dentro del producto cárnico.

Se necesitan diagnósticos microbianos casi en tiempo real para entornos alimentarios, clínicos, industriales y otros entornos ambientales (véase, por ejemplo, Lu TK, Bowers J y Koeris MS., Trends Biotechnol. Junio de 2013; 31(6):325-7). En determinadas realizaciones, la presente invención se usa para la detección rápida de patógenos transmitidos por los alimentos usando ARN guía específicos para un patógeno (por ejemplo, *Campylobacter jejuni*, *Clostridium perfringens*, *Salmonella spp.*, *Escherichia coli*, *Bacillus cereus*, *Listeria monocytogenes*, *Shigella spp.*, *Staphylococcus aureus*, enteritis estafilocócica, *Streptococcus*, *Vibrio cholerae*, *Vibrio parahaemolyticus*, *Vibrio vulnificus*, *Yersinia enterocolitica* y *Yersinia pseudotuberculosis*, *Brucella spp.*, *Corynebacterium ulcerans*, *Coxiella burnetii* o *Plesiomonas shigelloides*).

En determinadas realizaciones, el dispositivo es o comprende una tira de flujo. Por ejemplo, una tira de flujo lateral permite la detección de ARNasa (por ejemplo, C2c2) por color. El indicador de ARN se modifica para tener una primera molécula (tal como, por ejemplo, FITC) unida al extremo 5' y una segunda molécula (tal como, por ejemplo, biotina) unida al extremo 3' (o viceversa). La tira de flujo lateral está diseñada para tener dos líneas de captura con anticuerpos anti-primer molécula (por ejemplo, anti-FITC) hibridados en la primera línea y anticuerpos anti-segunda molécula (por ejemplo, anti-biotina) en la segunda línea aguas abajo. A medida que la reacción fluye por la tira, el indicador no escindido se unirá a los anticuerpos anti-primer molécula en la primera línea de captura, mientras que los indicadores escindidos liberarán la segunda molécula y permitirán la unión de la segunda molécula en la segunda línea de captura. Los anticuerpos en sándwich de segunda molécula, por ejemplo, conjugados con nanopartículas, tales como nanopartículas de oro, se unirán a cualquier segunda molécula en la primera o segunda línea y darán como resultado una lectura/señal fuerte (por ejemplo, color). A medida que se escinde más indicador, se acumulará más señal en la segunda línea de captura y aparecerá menos señal en la primera línea. En determinados aspectos, la invención se refiere al uso de una tira de seguimiento tal como se describe en el presente documento para detectar ácidos nucleicos o polipéptidos. En determinados aspectos, la invención se refiere a un método para detectar ácidos nucleicos o polipéptidos con una tira de flujo tal como se define en el presente documento, por ejemplo, pruebas de flujo (lateral) o ensayos inmunocromatográficos de flujo (lateral).

En determinadas realizaciones de ejemplo, el dispositivo es un dispositivo microfluídico que genera y/o fusiona diferentes gotas (es decir, volúmenes discretos individuales). Por ejemplo, puede formarse un primer conjunto de gotas que contengan muestras que van a examinarse y un segundo conjunto de gotas formado que contenga los elementos de los sistemas descritos en el presente documento. Los conjuntos de gotas primero y segundo se fusionan luego y se llevan a cabo entonces métodos de diagnóstico tal como se describe en el presente documento

en el conjunto de gotas fusionadas. Los dispositivos microfluídicos dados a conocer en el presente documento pueden ser chips basados en silicona y pueden fabricarse usando una variedad de técnicas que incluyen, pero no se limitan en, estampado en caliente, moldeado de elastómeros, moldeado por inyección, LIGA, litografía blanda, fabricación de silicona y técnicas de procesamiento de película delgada relacionadas. Los materiales adecuados para fabricar los dispositivos microfluídicos incluyen, pero no se limitan a, copolímero de olefina cíclica (COC), policarbonato, poli(dimetilsiloxano) (PDMS) y poli(metilacrilato) (PMMA). En una realización, puede usarse litografía blanda en PDMS para preparar los dispositivos microfluídicos. Por ejemplo, puede elaborarse un molde usando fotolitografía que define la ubicación de los canales de flujo, válvulas y filtros dentro de un sustrato. El material del sustrato se vierte en un molde y se deja fraguar para crear un sello. Luego el sello se sella a un soporte sólido, tal como, sin limitación, vidrio. Debido a la naturaleza hidrófoba de algunos polímeros, tales como PDMS, que absorben algunas proteínas y pueden inhibir determinados procesos biológicos, puede ser necesario un agente pasivante (Schoffner *et al.* Nucleic Acids Research, 1996, 24:375-379). Los agentes pasivantes adecuados son conocidos en la técnica e incluyen, pero no se limitan a, silanos, parileno, n-dodecil-b-D-matósido (DDM), Pluronic, Tween-20, otros tensioactivos similares, polietilenglicol (PEG), albúmina, colágeno y otras proteínas y péptidos similares.

En determinadas realizaciones de ejemplo, el sistema y/o el dispositivo pueden adaptarse para la conversión a una lectura de citometría de flujo o permitir todas las mediciones sensibles y cuantitativas de millones de células en un solo experimento y mejorar los métodos basados en flujo existentes, tales como el ensayo PrimeFlow. En determinadas realizaciones de ejemplo, las células pueden moldearse en gotas que contienen monómero de gel sin polimerizar, que luego pueden moldearse en gotas de una sola célula adecuadas para el análisis por citometría de flujo. Un constructo de detección que comprende un marcador detectable fluorescente puede moldearse en la gota que comprende monómero de gel sin polimerizar. Tras la polimerización del monómero de gel para formar una perla dentro de una gota. Debido a que la polimerización del gel se produce mediante la formación de radicales libres, el indicador fluorescente se une covalentemente al gel. El constructo de detección puede modificarse adicionalmente para comprender un ligador, tal como una amina. Puede añadirse un extintor después de la formación del gel y se unirá a través del ligador al constructo informador. Por tanto, el extintor no está unido al gel y puede difundirse libremente cuando la proteína efectora CRISPR escinde el indicador. La amplificación de la señal en la gota puede lograrse acoplando el constructo de detección a una amplificación de reacción en cadena de hibridación (iniciador de HCR). Pueden incorporarse horquillas híbridas de ADN/ARN en el gel que puede comprender un bucle de horquilla que tiene un dominio sensible a la ARNasa. Al proteger un punto de apoyo (*toehold*) de desplazamiento de cadena dentro de un bucle en horquilla que tiene un dominio sensible a la ARNasa, los iniciadores de HCR pueden desprotegerse selectivamente después de la escisión del bucle en horquilla por la proteína efectora CRISPR. Después de la desprotección de los iniciadores de HCR a través del desplazamiento de la cadena mediado por el punto de apoyo, los monómeros de HCR fluorescentes pueden lavarse en el gel para permitir la amplificación de la señal cuando se desprotegen los iniciadores.

Un ejemplo de dispositivo microfluídico que puede usarse en el contexto de la invención se describe en Hour *et al.* "Direct Detection and Drug-resistance profiling of bacteremias using inertial microfluidics" Lap Chip. 15(10):2297-2307 (2016).

Los sistemas descritos en el presente documento pueden incorporarse adicionalmente en dispositivos médicos ponibles que evalúan muestras biológicas, tales como líquidos biológicos, de un sujeto fuera del entorno clínico e informan el resultado del ensayo de manera remota a un servidor central al que puede acceder un profesional de atención médica. El dispositivo puede incluir la capacidad de tomar muestras de sangre por sí mismo, tal como los dispositivos dados a conocer en publicación de solicitud de patente estadounidense n.º 2015/0342509 titulada "Needle-free Blood Draw" de Peeters *et al.*, la publicación de solicitud de patente estadounidense n.º 2015/0065821 titulada "Nanoparticle Phoresis" de Andrew Conrad.

En determinadas realizaciones de ejemplo, el dispositivo puede comprender pocillos individuales, tales como pocillos de microplaca. El tamaño de los pocillos de microplaca puede ser del tamaño de los pocillos de tamaño convencional 6, 24, 96, 384, 1536, 3456 o 9600. En determinadas realizaciones de ejemplo, los elementos de los sistemas descritos en el presente documento pueden secarse por congelación y aplicarse a la superficie del pocillo antes de su distribución y uso.

Los dispositivos dados a conocer en el presente documento pueden comprender además puertos o aberturas de entrada y salida, que a su vez pueden estar conectados a válvulas, tubos, canales, cámaras y jeringas y/o bombas para la introducción y extracción de fluidos hacia el interior y desde el dispositivo. Los dispositivos pueden estar conectados a actuadores de flujo de fluidos que permiten el movimiento direccional de fluidos dentro del dispositivo microfluídico. Los actuadores de ejemplo incluyen, pero no se limitan a, bombas de jeringa, bombas de recirculación accionadas mecánicamente, bombas electroosmóticas, bulbos, fuelles, diafragmas o burbujas destinados a forzar el movimiento de fluidos. En determinadas realizaciones de ejemplo, los dispositivos están conectados a controladores con válvulas programables que trabajan juntas para mover fluidos a través del dispositivo. En determinadas realizaciones de ejemplo, los dispositivos están conectados a los controladores que se analizan con más detalle a continuación. Los dispositivos pueden conectarse a actuadores de flujo, controladores y dispositivos de carga de muestras mediante un tubo que termina en pasadores metálicos para su inserción en los puertos de entrada del dispositivo.

Tal como se muestra en el presente documento, los elementos del sistema son estables cuando se secan por congelación, por tanto, también se contemplan realizaciones que no requieren un dispositivo de soporte, es decir, el sistema puede aplicarse a cualquier superficie o fluido que admita las reacciones descritas en el presente documento y permita la detección de una señal detectable positiva de esa superficie o disolución. Además de la liofilización, los sistemas también pueden almacenarse de manera estable y utilizarse en forma de gránulos. Los polímeros útiles para formar formas granuladas adecuadas son conocidos en la técnica.

En determinadas realizaciones, la proteína efectora CRISPR se une a cada volumen discreto en el dispositivo. Cada volumen discreto puede comprender un ARN guía diferente específico para una molécula diana diferente. En determinadas realizaciones, una muestra se expone a un sustrato sólido que comprende más de un volumen discreto, comprendiendo cada uno un ARN guía específico para una molécula diana. Sin estar sujeto a ninguna teoría, cada ARN guía capturará su molécula diana de la muestra y no es necesario dividir la muestra en ensayos separados. Por tanto, puede conservarse una muestra valiosa. La proteína efectora puede ser una proteína de fusión que comprenda una etiqueta de afinidad. Las etiquetas de afinidad son bien conocidas en la técnica (por ejemplo, etiqueta HA, etiqueta Myc, etiqueta Flag, etiqueta His, biotina). La proteína efectora puede unirse a una molécula de biotina y los volúmenes discretos pueden comprender estreptavidina. En otras realizaciones, un anticuerpo específico para la proteína efectora se une a la proteína efectora CRISPR. Los métodos para unir una enzima CRISPR se han descrito previamente (véase, por ejemplo, el documento US20140356867A1).

Los dispositivos dados a conocer en el presente documento también pueden incluir elementos de dispositivos de diagnóstico inmediato (POC) conocidos en la técnica para analizar muestras mediante otros métodos. Véase, por ejemplo St John y Price, "Existing and Emerging Technologies for Point-of-Care Testing" (Clin Biochem Rev. Agosto de 2014; 35(3): 155-167).

La presente invención puede usarse con un sistema de sensor de diagnóstico de laboratorio en chip (LOC) inalámbrico (véase, por ejemplo, la patente estadounidense n.º 9.470.699 "Diagnostic radio frequency identification sensors and applications thereof"). En determinadas realizaciones, la presente invención se realiza en un LOC controlado por un dispositivo inalámbrico (por ejemplo, un teléfono móvil, una agenda electrónica (PDA), una tablet) y los resultados se informan a dicho dispositivo.

Los sistemas de etiqueta de identificación por radiofrecuencia (RFID) incluyen una etiqueta RFID que transmite datos para su recepción por un lector RFID (también denominado interrogador). En un sistema RFID típico, los objetos individuales (por ejemplo, la mercancía de la tienda) están equipados con una etiqueta relativamente pequeña que contiene un transpondedor. El transpondedor tiene un chip de memoria que recibe un código de producto electrónico único. El lector RFID emite una señal que activa el transpondedor dentro de la etiqueta mediante el uso de un protocolo de comunicación. Por consiguiente, el lector RFID puede leer y escribir datos en la etiqueta. Además, el lector de etiqueta RFID procesa los datos según la aplicación del sistema de etiqueta RFID. Actualmente, existen etiquetas RFID de tipo pasivo y activo. La etiqueta RFID de tipo pasivo no contiene una fuente de alimentación interna, sino que se alimenta mediante las señales de radiofrecuencia recibidas del lector RFID. Alternativamente, la etiqueta RFID de tipo activo contiene una fuente de alimentación interna que permite que la etiqueta RFID de tipo activo posea mayores intervalos de transmisión y capacidad de memoria. El uso de una etiqueta pasiva frente a una activa depende de la aplicación particular.

La tecnología de laboratorio en chip está bien descrita en la bibliografía científica y consiste en múltiples canales microfluídicos, pocillos de entrada o químicos. Las reacciones en los pocillos pueden medirse usando la tecnología de etiqueta de identificación por radiofrecuencia (RFID), ya que los cables conductores del chip electrónico RFID pueden vincularse directamente a cada uno de los pocillos de prueba. Una antena puede imprimirse o montarse en otra capa del chip electrónico o directamente en la parte posterior del dispositivo. Además, los cables, la antena y el chip electrónico pueden incrustarse en el chip LOC, evitando así el cortocircuito de los electrodos o la electrónica. Dado que LOC permite análisis y separaciones de muestras complejas, esta tecnología permite que las pruebas de LOC se realicen independientemente de un lector complejo o costoso. Más bien puede usarse un dispositivo inalámbrico sencillo, tal como un teléfono móvil o una PDA. En una realización, el dispositivo inalámbrico también controla la separación y el control de los canales microfluídicos para análisis LOC más complejos. En una realización, se incluyen un LED y otros dispositivos electrónicos de medición o detección en el chip LOC-RFID. Sin estar sujeto a ninguna teoría, esta tecnología es desechable y permite realizar pruebas complejas que requieren separación y mezcla fuera de un laboratorio.

En realizaciones preferidas, el LOC puede ser un dispositivo microfluídico. El LOC puede ser un chip pasivo, en el que el chip se alimenta y controla a través de un dispositivo inalámbrico. En determinadas realizaciones, el LOC incluye un canal microfluídico para contener reactivos y un canal para introducir una muestra. En determinadas realizaciones, una señal del dispositivo inalámbrico alimenta el LOC y activa la mezcla de la muestra y los reactivos de ensayo. Específicamente, en el caso de la presente invención, el sistema puede incluir un agente de enmascaramiento, una proteína efectora CRISPR y ARN guía específicos para una molécula diana. Tras la activación del LOC, el dispositivo microfluídico puede mezclar la muestra y los reactivos de ensayo. Tras el mezclado, un sensor detecta una señal y transmite los resultados al dispositivo inalámbrico. En determinadas realizaciones, el agente de desenmascaramiento es una molécula de ARN conductora. La molécula de ARN conductora puede unirse al material conductor. Las moléculas conductoras pueden ser nanopartículas conductoras,

proteínas conductoras, partículas metálicas que se unen a la proteína o látex u otras perlas que son conductoras. En determinadas realizaciones, si se usa ADN o ARN, las moléculas conductoras pueden unirse directamente a las cadenas de ADN o ARN correspondientes. La liberación de las moléculas conductoras puede detectarse a través de un sensor. El ensayo puede ser un proceso de una sola etapa.

Dado que la conductividad eléctrica del área de la superficie puede medirse con precisión, es posible obtener resultados cuantitativos en los electroensayos RFID inalámbricos desechables. Además, el área de prueba puede ser muy pequeña, lo que permite realizar más pruebas en un área determinada y, por tanto, generar ahorros de costes. En determinadas realizaciones, se usan sensores separados, cada uno asociado con una proteína efectora CRISPR diferente y un ARN guía inmovilizado en un sensor, para detectar múltiples moléculas diana. Sin estar sujeto a una teoría, la activación de diferentes sensores puede distinguirse mediante el dispositivo inalámbrico.

Además de los métodos conductivos descritos en el presente documento, pueden usarse otros métodos que se basen en RFID o Bluetooth como la plataforma básica de comunicación y alimentación de bajo coste para un ensayo RFID desechable. Por ejemplo, pueden usarse medios ópticos para evaluar la presencia y el nivel de una molécula diana dada. En determinadas realizaciones, un sensor óptico detecta el desenmascaramiento de un agente de enmascaramiento fluorescente.

En determinadas realizaciones, el dispositivo de la presente invención puede incluir dispositivos de mano portátiles para la lectura diagnóstica de un ensayo (véase, por ejemplo, Vashist *et al.*, Commercial Smartphone-Based Devices and Smart Applications for Personalized Healthcare Monitoring and Management, Diagnostics 2014, 4(3), 104-128; mReader de Mobile Assay; y lector de pruebas de diagnóstico rápido Holomic).

Tal como se indica en el presente documento, determinadas realizaciones permiten la detección a través de cambios colorimétricos que tienen determinados beneficios concomitantes cuando las realizaciones se utilizan en situaciones de POC y/o en entornos con pocos recursos donde el acceso a equipos de detección más complejos para leer la señal puede ser limitado. Sin embargo, las realizaciones portátiles dadas a conocer en el presente documento también pueden acoplarse con espectrofotómetros de mano que permiten la detección de señales fuera del intervalo visible. Un ejemplo de un dispositivo de espectrofotómetro de mano que puede usarse en combinación con la presente invención se describe en Das *et al.* "Ultra-portable, wireless smartphone spectrophotometer for rapid, non-destructive testing of fruit ripeness". Nature Scientific Reports. 2016, 6:32504, DOI: 10.1038/srep32504. Finalmente, en determinadas realizaciones que utilizan constructos de enmascaramiento basados en puntos cuánticos, el uso de una luz UV de mano u otro dispositivo adecuado puede usarse con éxito para detectar una señal debido al rendimiento cuántico casi completo proporcionado por los puntos cuánticos.

EJEMPLOS DE MÉTODOS Y ENSAYOS

El bajo coste y la adaptabilidad de la plataforma de ensayo se presta a una serie de aplicaciones que incluyen (i) cuantificación general de ARN/ADN/proteína, (ii) detección rápida y multiplexada de expresión de proteínas y ARN/ADN, y (iii) detección sensible de los ácidos nucleicos, péptidos y proteínas diana en muestras clínicas y ambientales. Además, los sistemas dados a conocer en el presente documento pueden adaptarse para la detección de transcritos dentro de entornos biológicos, tales como las células. Dada la naturaleza altamente específica de los efectores CRISPR descritos en el presente documento, es posible monitorizar la expresión alélica específica de transcritos o mutaciones asociadas a enfermedades en células vivas.

En determinadas realizaciones de ejemplo, las secuencias de guía únicas específicas para una sola diana se colocan en volúmenes distintos. Cada volumen puede recibir entonces una muestra diferente o una alícuota de la misma muestra. En determinadas realizaciones de ejemplo, múltiples secuencias de guía, cada una para una diana distinta, pueden colocarse en un solo pocillo de modo que puedan examinarse múltiples dianas en un pocillo diferente. Con el fin de detectar múltiples ARN guía en un solo volumen, en determinadas realizaciones de ejemplo, pueden usarse múltiples proteínas efectoras con diferentes especificidades. Por ejemplo, pueden usarse diferentes ortólogos con diferentes especificidades de secuencia. Por ejemplo, un ortólogo puede cortar preferiblemente A, mientras que otros pueden cortar preferiblemente C, G, U/T. Por consiguiente, pueden generarse constructos de enmascaramiento que comprendan por completo, o estén compuestos por una porción sustancial, de un solo nucleótido, cada uno con un fluoróforo diferente que puede detectarse en diferentes longitudes de onda. De esta manera, pueden examinarse hasta cuatro dianas diferentes en un único volumen discreto individual. En determinadas realizaciones de ejemplo, pueden usarse diferentes ortólogos de una misma clase de proteína efectora CRISPR, como dos ortólogos de Cas13a, dos ortólogos de Cas13b o dos ortólogos de Cas13c. Las preferencias de nucleótidos de diversas proteínas Cas13 se muestran en la figura 67. En otras determinadas realizaciones de ejemplo, pueden usarse diferentes ortólogos con diferentes preferencias de edición de nucleótidos, tales como ortólogos de Cas13a y Cas13b, u ortólogos de Cas13a y Cas13c, u ortólogos de Cas13b y ortólogos de Cas13c, etc. En determinadas realizaciones de ejemplo, se usan la proteína Cas13 con una preferencia poliU y una proteína Cas13 con una preferencia poliA. En determinadas realizaciones de ejemplo, la proteína Cas13 con una preferencia poliU es una Cas13b de *Prevotella intermedia* y la proteína Cas13 con una preferencia poliA es una proteína Cas13b de *Prevotella sp.* MA2106 (PsmCas13b). En determinadas realizaciones de ejemplo, la proteína Cas13 con una preferencia poliU es una proteína Cas13a de *Leptotrichia wadei* (LwaCas13a) y la proteína Cas13 con una preferencia poliA es una proteína Cas13b de *Prevotella sp.* MA2106. En determinadas realizaciones de

ejemplo, la proteína Cas13 con una preferencia poliU es proteína Cas13b de *Capnocytophaga canimorsus* (CcaCas13b).

Además de las preferencias de edición de base única, pueden diseñarse constructos de detección adicionales basándose otras preferencias de corte de motivo de los ortólogos de Cas 13. Por ejemplo, los ortólogos de Cas13 pueden cortar preferiblemente una secuencia de dinucleótidos, una secuencia de trinucleótidos o motivos más complejos que comprenden 4, 5, 6, 7, 8, 9 o 10 motivos de nucleótidos. Por tanto, el límite superior para los ensayos múltiplex que usan las realizaciones dadas a conocer en el presente documento está limitado principalmente por el número de marcadores detectables distinguibles. Los métodos de ejemplo para identificar tales motivos se describen adicionalmente en los ejemplos de trabajo a continuación.

Tal como se demuestra en el presente documento, los sistemas efectores CRISPR pueden detectar concentraciones hasta attomolares de moléculas diana. Véase, por ejemplo, las figuras 13, 14, 19, 22 y los ejemplos de trabajo descritos a continuación. Los ensayos y aplicaciones de ejemplo se describen con mayor detalle a continuación.

APLICACIONES MICROBIANAS

En determinadas realizaciones de ejemplo, los sistemas, dispositivos y métodos dados a conocer en el presente documento se refieren a detectar la presencia de uno o más agentes microbianos en una muestra, tal como una muestra biológica obtenida de un sujeto. En determinadas realizaciones de ejemplo, el microbio puede ser una bacteria, un hongo, una levadura, un protozoo, un parásito o un virus. Por consiguiente, los métodos dados a conocer en el presente documento pueden adaptarse para su uso en otros métodos (o en combinación) con otros métodos que requieren una identificación rápida de especies de microbios, control de la presencia de proteínas microbianas (antígenos), anticuerpos, genes de anticuerpos, detección de determinados fenotipos (por ejemplo, resistencia bacteriana), seguimiento de la evolución y/o brote epidémico de la enfermedad y detección de antibióticos. Debido a las capacidades diagnósticas rápidas y sensibles de las realizaciones dadas a conocer en el presente documento, la detección del tipo de especie de microbio, hasta una diferencia de un solo nucleótido, y la capacidad de implementarse como un dispositivo de POC, las realizaciones dadas a conocer en el presente documento pueden usarse para guiar regímenes terapéuticos, tales como la selección del antibiótico o antiviral adecuado. Las realizaciones dadas a conocer en el presente documento también pueden usarse para examinar muestras ambientales (aire, agua, superficies, alimentos, etc.) para detectar la presencia de contaminación microbiana.

Se da a conocer un método para identificar especies microbianas, tales como especies bacterianas, virales, fúngicas, de levadura o parasitarias, o similares. Las realizaciones particulares dadas a conocer en el presente documento describen métodos y sistemas que identificarán y distinguirán especies microbianas dentro de una sola muestra, o entre múltiples muestras, lo que permitirá el reconocimiento de muchos microbios diferentes. Los presentes métodos permiten la detección de patógenos y la distinción entre dos o más especies de uno o más organismos, por ejemplo, bacterias, virus, levaduras, protozoos y hongos o una combinación de los mismos, en una muestra biológica o ambiental, detectando la presencia de una secuencia de ácido nucleico diana en la muestra. Una señal positiva obtenida de la muestra indica la presencia del microbio. Pueden identificarse múltiples microbios simultáneamente usando los métodos y sistemas de la invención, empleando el uso de más de una proteína efectora, en donde cada proteína efectora selecciona como diana una secuencia diana microbiana específica. De esta manera, puede realizarse un análisis multinivel para un sujeto en particular en el que puede detectarse cualquier cantidad de microbios a la vez. En algunas realizaciones, la detección simultánea de múltiples microbios puede realizarse usando un conjunto de sondas que pueden identificar una o más especies microbianas.

El análisis múltiplex de muestras permite la detección de muestras a gran escala, lo que reduce el tiempo y el coste de los análisis. Sin embargo, los análisis múltiplex a menudo están limitados por la disponibilidad de una muestra biológica. Según la invención, sin embargo, pueden realizarse alternativas al análisis múltiplex de modo que puedan añadirse múltiples proteínas efectoras a una sola muestra y cada constructo de enmascaramiento pueda combinarse con un colorante extintor separado. En este caso, las señales positivas pueden obtenerse de cada colorante extintor por separado para la detección múltiple en una sola muestra.

En el presente documento se dan a conocer métodos para distinguir entre dos o más especies de uno o más organismos en una muestra. Los métodos también pueden detectar una o más especies de uno o más organismos en una muestra.

Detección de microbios

En algunas realizaciones, se proporciona un método para detectar microbios en muestras que comprende distribuir una muestra o conjunto de muestras en uno o más volúmenes discretos individuales, comprendiendo los volúmenes discretos individuales un sistema CRISPR tal como se describe en el presente documento; incubar la muestra o conjunto de muestras en condiciones suficientes para permitir la unión de uno o más ARN guía a una o más dianas específicas de microbios; activar la proteína efectora CRISPR mediante la unión del uno o más ARN guía a la una o más moléculas diana, en el que la activación de la proteína efectora CRISPR da como resultado la modificación del constructo de enmascaramiento basado en ARN de manera que se genera una señal positiva detectable; y detectar

la señal positiva detectable, en el que la detección de la señal positiva detectable indica la presencia de una o más moléculas diana en la muestra. La una o más moléculas diana pueden ser ARNm, ADN_g (codificante o no codificante), ARNr o ARNr que comprende una secuencia de nucleótidos diana que puede usarse para distinguir dos o más especies/cepas microbianas entre sí. Los ARN guía pueden diseñarse para detectar secuencias diana.

Las realizaciones dadas a conocer en el presente documento también pueden utilizar determinadas etapas para mejorar la hibridación entre secuencias de ARN guía y ARN diana. Los métodos para mejorar la hibridación del ácido ribonucleico se dan a conocer en el documento WO 2015/085194, titulado "Enhanced Methods of Ribonucleic Acid Hybridization". La diana específica de microbios puede ser un ARN o ADN o una proteína. Si el método de ADN puede comprender además el uso de cebadores de ADN que introducen un promotor de ARN polimerasa tal como se describe en el presente documento. Si la diana es una proteína, entonces el método utilizará aptámeros y etapas específicas para la detección de proteínas descritos en el presente documento.

Detección de variantes de un solo nucleótido

En algunas realizaciones, una o más secuencias diana identificadas pueden detectarse usando ARN guía que son específicos para y se unen a la secuencia diana tal como se describe en el presente documento. Los sistemas y métodos de la presente invención pueden distinguir incluso entre polimorfismos de un solo nucleótido presentes entre diferentes especies microbianas y, por tanto, el uso de múltiples ARN guía según la invención puede ampliar o mejorar adicionalmente el número de secuencias diana que pueden usarse para distinguir entre especies. Por ejemplo, en algunas realizaciones, el uno o más ARN guía pueden distinguir entre microbios en la especie, género, familia, orden, clase, filo, reino o una combinación de los mismos.

Detección basada en secuencias de ARNr

En determinadas realizaciones de ejemplo, los dispositivos, sistemas y métodos descritos en el presente documento pueden usarse para distinguir múltiples especies microbianas en una muestra. En determinadas realizaciones de ejemplo, la identificación puede basarse en secuencias de ARN ribosómico, incluyendo las subunidades 16S, 23S y 5S. Los métodos para identificar secuencias de ARNr relevantes se dan a conocer en la publicación de solicitud de patente estadounidense n.º 2017/0029872. En determinadas realizaciones de ejemplo, puede diseñarse un conjunto de ARN guía para distinguir cada especie por una región variable que es única para cada especie o cepa. Los ARN guía también pueden diseñarse para seleccionar como diana genes de ARN que distinguen microbios en niveles de género, familia, orden, clase, filo, reino o una combinación de los mismos. En determinadas realizaciones de ejemplo en las que se usa la amplificación, puede diseñarse un conjunto de cebadores de amplificación para flanquear las regiones constantes de la secuencia de ARN ribosómico y puede diseñarse un ARN guía para distinguir cada especie por una región interna variable. En determinadas realizaciones de ejemplo, los cebadores y los ARN guía pueden diseñarse para regiones conservadas y variables en la subunidad 16S respectivamente. También pueden usarse otros genes o regiones genómicas que varían de manera única entre especies o un subconjunto de especies tales como la familia de genes RecA, la subunidad β de la ARN polimerasa. Otros marcadores filogenéticos adecuados y métodos para identificar los mismos se analizan, por ejemplo, en Wu *et al.* arXiv: 1307.8690 [q-bio.GN].

En determinadas realizaciones de ejemplo, se diseña un método o diagnóstico para examinar microbios a través de múltiples niveles filogenéticos y/o fenotípicos al mismo tiempo. Por ejemplo, el método o diagnóstico puede comprender el uso de múltiples sistemas CRISPR con diferentes ARN guía. Un primer conjunto de ARN guía puede distinguir, por ejemplo, entre micobacterias, bacterias Gram positivas y Gram negativas. Estas clases generales pueden subdividirse adicionalmente. Por ejemplo, los ARN guía pueden diseñarse y usarse en el método o diagnóstico que distinguen las bacterias entéricas y no entéricas dentro de las bacterias Gram negativas. Puede diseñarse un segundo conjunto de ARN guía para distinguir los microbios a nivel de género o especie. Por tanto, puede producirse una matriz que identifique todas las micobacterias, bacterias Gram positivas, bacterias Gram negativas (divididas además en entéricas y no entéricas) con cada género de especies de bacterias identificadas en una muestra dada que se encuentran dentro de una de esas clases. Lo anterior es sólo para fines de ejemplo. También se contemplan otros medios para clasificar otros tipos de microbios y seguirían la estructura general descrita anteriormente.

Examen para seleccionar la resistencia a los fármacos

En determinadas realizaciones de ejemplo, los dispositivos, sistemas y métodos dados a conocer en el presente documento pueden usarse para realizar un examen de selección de genes microbianos de interés, por ejemplo, genes de resistencia a antibióticos y/o antivirales. Los ARN guía pueden diseñarse para distinguir entre genes de interés conocidos. Luego las muestras, incluyendo las muestras clínicas, pueden examinarse usando las realizaciones dadas a conocer en el presente documento para la detección de tales genes. La capacidad de examinar la resistencia a los fármacos en el POC tendría un enorme beneficio en la selección de un régimen de tratamiento adecuado. En determinadas realizaciones de ejemplo, los genes de resistencia a antibióticos son carbapenemasas que incluyen KPC, NDM1, CTX-M15, OXA-48. Se conocen otros genes de resistencia a antibióticos y pueden encontrarse, por ejemplo, en la base de datos completa de resistencia a los antibióticos (Jia *et al.* "CARD 2017: expansion and model-centric curation of the Comprehensive Antibiotic Resistance Database". Nucleic Acids Research, 45, D566-573).

La ribavirina es un antiviral eficaz que ataca a varios virus de ARN. Varios virus clínicamente importantes han desarrollado resistencia a la ribavirina, incluyendo virus de la fiebre aftosa doi:10.1128/JVI.03594-13; virus de la poliomielitis (Pfeifer y Kirkegaard. PNAS, 100(12):7289-7294, 2003); y virus de la hepatitis C (Pfeiffer y Kirkegaard, J. Virol. 79(4):2346-2355, 2005). Varios de otros virus de ARN persistentes, tal como la hepatitis y el VIH, han desarrollado resistencia a fármacos antivirales existentes: virus de la hepatitis B (lamivudina, tenofovir, entecavir) doi:10.1002/hep.22900; virus de la hepatitis C (telaprevir, BILN2061, ITMN-191, SCH6, boceprevir, AG-021541, ACH-806) doi:10.1002/hep.22549; y VIH (muchas mutaciones de resistencia a los fármacos) hivb.stanford.edu. Las realizaciones dadas a conocer en el presente documento pueden usarse para detectar tales variantes, entre otras.

Aparte de la resistencia a los fármacos, hay una serie de mutaciones clínicamente relevantes que pueden detectarse con las realizaciones dadas a conocer en el presente documento, tales como infección persistente frente aguda en LCMV (doi: 10.1073/pnas.1019304108), y mayor infectividad de ébola (Diehl *et al.* Cell. 2016, 167(4):1088-1098).

Tal como se describe en otra parte en el presente documento, las especies microbianas estrechamente relacionadas (por ejemplo, que tienen una diferencia de un solo nucleótido en una secuencia diana determinada) pueden distinguirse mediante la introducción de un apareamiento erróneo sintético en el ARNg.

Enfoques de cobertura de conjuntos

En realizaciones particulares, se diseña un conjunto de ARN guía que pueden identificar, por ejemplo, todas las especies microbianas dentro de un conjunto definido de microbios. En determinadas realizaciones de ejemplo, los métodos para generar ARN guía tal como se describe en el presente documento pueden compararse con los métodos descritos en el documento WO 2017/040316. Tal como se describe en el documento WO 2017040316, una solución de cobertura de conjuntos puede identificar el número mínimo de sondas de secuencias diana o ARN guía necesarios para cubrir una secuencia diana completa o un conjunto de secuencias diana, por ejemplo, un conjunto de secuencias genómicas. Los enfoques de cobertura de conjuntos se han usados previamente para identificar cebadores y/o sondas de micromatrices, normalmente en el intervalo de 20 a 50 pares de bases. Véase, por ejemplo, Pearson *et al.*, cs.virginia.edu/~robins/papers/primers_dam11_final.pdf., Jabado *et al.* Nucleic Acids Res. 2006 34(22):6605-11, Jabado *et al.* Nucleic Acids Res. 2008, 36(1):e3 doi:10.1093/nar/gkm1106, Duitama *et al.* Nucleic Acids Res. 2009, 37(8):2483-2492, Phillippy *et al.* BMC Bioinformatics. 2009, 10:293 doi:10.1186/1471-2105-10-293. Sin embargo, tales enfoques generalmente implicaban tratar cada cebador/sonda como k-meros y buscar coincidencias exactas o permitir coincidencias inexactas usando matrices de sufijos. Además, los métodos generalmente adoptan un enfoque binario para detectar la hibridación seleccionando cebadores o sondas de modo que cada secuencia de entrada sólo necesita estar unida por un cebador o sonda y la posición de esta unión a lo largo de la secuencia es irrelevante. Los métodos alternativos pueden dividir un genoma diana en ventanas predefinidas y tratar de manera efectiva cada ventana como una secuencia de entrada separada bajo el enfoque binario: es decir, determinan si una sonda o guía de ARN dada se une dentro de cada ventana y requieren que todas las ventanas estén unidas por el estado de alguna sonda o guía de ARN. Efectivamente, estos enfoques tratan cada elemento del "universo" en el problema de cobertura de conjuntos como si fuera una secuencia de entrada completa o una ventana predefinida de una secuencia de entrada, y cada elemento se considera "cubierto" si el inicio de una sonda o ARN guía se une dentro del elemento. Estos enfoques limitan la fluidez con la cual se permite que diferentes diseños de sonda o ARN guía cubran una secuencia diana dada.

Por el contrario, las realizaciones dadas a conocer en el presente documento se refieren a detectar longitudes de ARN guía o de sonda más largas, por ejemplo, en el intervalo de 70 pb a 200 pb que son adecuadas para la secuenciación de selección híbrida. Además, los métodos dados a conocer por el documento WO 2017/040316 en el presente documento pueden aplicarse para adoptar un enfoque de secuencia pan-diana que puede definir una sonda o conjuntos de ARN guía que pueden identificar y facilitar la secuenciación de detección de todas las secuencias de especies y/o cepas en un conjunto de secuencias diana grande y/o variable. Por ejemplo, los métodos dados a conocer en el presente documento pueden usarse para identificar todas las variantes de un virus determinado o múltiples virus diferentes en un solo ensayo. Además, el método dado a conocer en el presente documento trata cada elemento del "universo" en el problema de la cobertura de conjuntos como si fuera un nucleótido de una secuencia diana, y cada elemento se considera "cubierto" siempre que una sonda o ARN guía se una a algún segmento de un genoma diana que incluye el elemento. Este tipo de métodos de cobertura de conjuntos pueden usarse en lugar del enfoque binario de los métodos anteriores, los métodos dados a conocer en el presente documento modelan mejor cómo una sonda o un ARN guía puede hibridarse con una secuencia diana. En lugar de sólo preguntar si una secuencia de ARN guía dada se une o no a una ventana determinada, estos enfoques pueden usarse para detectar un patrón de hibridación, es decir, donde una sonda o ARN guía dado se une a una secuencia diana o secuencias diana, y luego determinar a partir de esos patrones de hibridación el número mínimo de sondas o ARN guía necesarios para cubrir el conjunto de secuencias diana en un grado suficiente para permitir tanto el enriquecimiento de una muestra como la secuenciación de cualquiera y todas las secuencias diana. Estos patrones de hibridación pueden determinarse definiendo determinados parámetros que minimizan una función de pérdida, lo que permite la identificación de conjuntos mínimos de sonda o ARN guía de una manera que permite que los parámetros varíen para cada especie, por ejemplo, para reflejar la diversidad de cada especie, así como de una manera computacionalmente eficiente que no puede lograrse mediante una aplicación sencilla de una solución de cobertura de conjuntos, tales como las que se aplicaron anteriormente en el contexto de diseño de sonda o ARN guía.

La capacidad de detectar múltiples abundancias de transcritos puede permitir la generación de firmas microbianas únicas indicativas de un fenotipo particular. Pueden usarse diversas técnicas de aprendizaje automático para derivar las firmas génicas. Por consiguiente, los ARN guía de los sistemas CRISPR pueden usarse para identificar y/o cuantificar niveles relativos de biomarcadores definidos por la firma génica para detectar determinados fenotipos. En determinadas realizaciones de ejemplo, la firma génica indica susceptibilidad a un antibiótico, resistencia a un antibiótico o una combinación de las mismas.

En un aspecto de la invención, un método comprende detectar uno o más patógenos. De esta manera, puede obtenerse la diferenciación entre la infección de un sujeto por microbios individuales. En algunas realizaciones, tal diferenciación puede permitir la detección o el diagnóstico por parte de un médico de enfermedades específicas, por ejemplo, diferentes variantes de una enfermedad. Preferiblemente, la secuencia del patógeno es un genoma del patógeno o un fragmento del mismo. El método puede comprender además determinar la evolución del patógeno. Determinar la evolución del patógeno puede comprender la identificación de mutaciones del patógeno, por ejemplo, delección de nucleótidos, inserción de nucleótidos, sustitución de nucleótidos. Entre estos últimos, hay sustituciones no sinónimas, sinónimas y no codificantes. Las mutaciones son más frecuentemente no sinónimas durante un brote epidémico. El método puede comprender además determinar la tasa de sustitución entre dos secuencias de patógenos analizadas tal como se describió anteriormente. Sin embargo, si las mutaciones son perjudiciales o incluso adaptativas requeriría un análisis funcional, la tasa de mutaciones no sinónimas sugiere que la evolución continua de esta epidemia puede brindar una oportunidad para la adaptación del patógeno, lo que subraya la necesidad de una contención rápida. Por tanto, el método puede comprender además evaluar el riesgo de adaptación viral, en el que se determina el número de mutaciones no sinónimas. (Gire, *et al.*, Science 345, 1369, 2014).

Monitorización de brotes epidémicos de microbios

En algunas realizaciones, puede usarse un sistema CRISPR o métodos de uso del mismo tal como se describe en el presente documento para determinar la evolución de un brote epidémico de patógenos. El método puede comprender la detección de una o más secuencias diana de una pluralidad de muestras de uno o más sujetos, en el que la secuencia diana es una secuencia de un microbio que provoca los brotes epidémicos. Tal método puede comprender además determinar un patrón de transmisión del patógeno, o un mecanismo involucrado en un brote epidémico de enfermedad provocado por un patógeno.

El patrón de transmisión del patógeno puede comprender nuevas transmisiones continuas desde el reservorio natural del patógeno o transmisiones de sujeto a sujeto (por ejemplo, transmisión de persona a persona) después de una sola transmisión desde el reservorio natural o una combinación de ambas. En una realización, la transmisión del patógeno puede ser una transmisión viral o bacteriana, en tal caso, la secuencia diana es preferiblemente un genoma microbiano o fragmentos del mismo. En una realización, el patrón de transmisión del patógeno es el patrón temprano de la transmisión del patógeno, es decir, al comienzo del brote epidémico del patógeno. Determinar el patrón de transmisión del patógeno al comienzo del brote epidémico aumenta la probabilidad de detener el brote epidémico lo antes posible, reduciendo así la posibilidad de diseminación local e internacional.

Determinar el patrón de transmisión del patógeno puede comprender detectar una secuencia de patógenos según los métodos descritos en el presente documento. La determinación del patrón de transmisión del patógeno puede comprender además la detección de variaciones dentro del huésped compartidas de la secuencia patógena entre los sujetos y la determinación de si las variaciones dentro del huésped compartidas muestran patrones temporales. Los patrones en la variación dentro del huésped y entre huéspedes observada brindan información importante sobre la transmisión y la epidemiología (Gire, *et al.*, 2014).

La detección de variaciones dentro del huésped compartidas entre los sujetos que muestran patrones temporales es una indicación de vínculos de transmisión entre sujetos (en particular entre seres humanos) porque puede explicarse por infección del sujeto de múltiples fuentes (sobreinfección), contaminación de la muestra, mutaciones recurrentes (con o sin selección disruptiva para reforzar las mutaciones), o la cotransmisión de virus ligeramente divergentes que surgieron por mutación anterior en la cadena de transmisión (Park, *et al.*, Cell 161(7):1516-1526, 2015). La detección de variaciones dentro del huésped compartidas entre sujetos puede comprender la detección de variantes dentro del huésped ubicadas en posiciones comunes de polimorfismo de un solo nucleótido (SNP). La detección positiva de variantes dentro del huésped ubicadas en posiciones comunes (SNP) es indicativa de que la sobreinfección y la contaminación son explicaciones principales de las variantes dentro del huésped. La sobreinfección y la contaminación pueden distinguirse en función de la frecuencia de SNP que aparece como variantes entre huéspedes (Park, *et al.*, 2015). De lo contrario, puede descartarse la sobreinfección y la contaminación. En este último caso, la detección de variaciones dentro del huésped compartidas entre sujetos puede comprender además evaluar las frecuencias de variantes sinónimas y no sinónimas y comparar la frecuencia de variantes sinónimas y no sinónimas entre sí. Una mutación no sinónima es una mutación que altera el aminoácido de la proteína, lo que probablemente dé como resultado un cambio biológico en el microbio que está sujeto a la selección natural. La sustitución sinónima no altera una secuencia de aminoácidos. La misma frecuencia de variantes sinónimas y no sinónimas es indicativa de que las variantes dentro del huésped evolucionan de manera neutral. Si las frecuencias de las variantes sinónimas y no sinónimas son divergentes, es probable que las variantes dentro del huésped se mantengan mediante la selección disruptiva. Si las frecuencias de variantes sinónimas y no sinónimas son bajas,

esto es indicativo de una mutación recurrente. Si las frecuencias de variantes sinónimas y no sinónimas son altas, esto es indicativo de cotransmisión (Park, *et al.*, 2015).

Al igual que el virus del ébola, el virus de Lassa (LASV) puede provocar fiebre hemorrágica con altas tasas de letalidad. Andersen *et al.* generaron un catálogo genómico de casi 200 secuencias de LASV a partir de muestras clínicas y de reservorios de roedores (Andersen, *et al.*, Cell Volume 162, número 4, págs. 738-750, 13 de agosto de 2015). Andersen *et al.* muestran que mientras que la epidemia de EVE de 2013-2015 está impulsado por transmisiones de persona a persona, las infecciones por LASV se deben principalmente a infecciones de reservorio a persona. Andersen *et al.* dilucidaron la propagación de LASV en África occidental y muestran que esta migración estuvo acompañada de cambios en la abundancia del genoma de LASV, las tasas de mortalidad, la adaptación de codones y la eficiencia de traducción. El método puede comprender además comparar filogenéticamente una primera secuencia patógena con una segunda secuencia patógena y determinar si existe un vínculo filogenético entre las secuencias patógenas primera y segunda. La segunda secuencia patógena puede ser una secuencia de referencia anterior. Si existe un vínculo filogenético, el método puede comprender además enraizar la filogenia de la primera secuencia patógena en la segunda secuencia patógena. Por tanto, es posible construir el linaje de la primera secuencia patógena. (Park, *et al.*, 2015).

El método puede comprender además determinar si las mutaciones son perjudiciales o adaptativas. Las mutaciones perjudiciales son indicativas de virus con problemas de transmisión e infecciones sin salida, por lo que normalmente solo están presentes en un sujeto individual. Las mutaciones exclusivas de un sujeto individual son aquellas que se producen en las ramas externas del árbol filogenético, mientras que las mutaciones de ramas internas son aquellas presentes en múltiples muestras (es decir, en múltiples sujetos). Una tasa más alta de sustitución no sinónima es una característica de las ramas externas del árbol filogenético (Park, *et al.*, 2015).

En las ramas internas del árbol filogenético, la selección ha tenido más oportunidad de filtrar mutantes perjudiciales. Las ramas internas, por definición, han producido múltiples linajes descendientes y, por tanto, es menos probable que incluyan mutaciones con costes de adecuación biológica. Por tanto, una tasa más baja de sustitución no sinónima es indicativa de ramas internas (Park, *et al.*, 2015).

Las mutaciones sinónimas, que probablemente tienen menos impacto en la adecuación biológica, se produjeron a frecuencias más comparables en las ramas internas y externas (Park, *et al.*, 2015).

Al analizar la secuencia diana secuenciada, tal como los genomas virales, es posible descubrir los mecanismos responsables de la gravedad del episodio epidémico, tal como durante el brote epidémico de ébola de 2014. Por ejemplo, Gire *et al.* hicieron una comparación filogenética de los genomas del brote epidémico de 2014 con los 20 genomas de brotes epidémicos anteriores y sugirieron que el virus de África occidental de 2014 probablemente se propagó desde África central en la última década. Fue problemático enraizar la filogenia usando la divergencia de otros genomas de virus del Ébola (6, 13). Sin embargo, enraizar el árbol en el brote epidémico más antiguo reveló una fuerte correlación entre la fecha de la muestra y la distancia de la raíz a la punta, con una tasa de sustitución de 8×10^{-4} por sitio por año (13). Esto sugiere que los linajes de los tres brotes más recientes divergieron de un ancestro común aproximadamente al mismo tiempo, aproximadamente 2004, lo que respalda la hipótesis de que cada brote epidémico representa un acontecimiento zoonótico independiente de la misma población viral genéticamente diversa en su reservorio natural. También descubrieron que el brote epidémico de EBOV de 2014 puede ser provocado por una sola transmisión del reservorio natural, seguida de transmisión de persona a persona durante el brote epidémico. Sus resultados también sugirieron que el episodio epidémico en Sierra Leona puede deberse a la introducción de dos virus genéticamente distintos de Guinea aproximadamente al mismo tiempo (Gire, *et al.*, 2014).

También ha sido posible determinar cómo se propagó el virus de Lassa desde su punto de origen, en particular gracias a la transmisión de persona a persona, e incluso rastrear la historia de esta propagación hace 400 años (Andersen, *et al.*, Cell 162(4):738-50, 2015).

En relación con el trabajo necesario durante el brote epidémico de EBOV de 2013-2015 y las dificultades encontradas por el personal médico en el lugar del brote epidémico, y de manera más general, el método de la invención permite realizar una secuenciación usando menos sondas seleccionadas de manera que la secuenciación puede acelerarse, acortando así el tiempo necesario desde la toma de muestras hasta la obtención de resultados. Además, los kits y sistemas pueden diseñarse para que puedan usarse en el campo de manera que el diagnóstico de un paciente pueda realizarse fácilmente sin necesidad de enviar muestras a otra parte del país o del mundo.

En cualquier método descrito anteriormente, puede usarse cualquiera de los procesos de secuenciación descritos anteriormente para la secuenciación de la secuencia diana o fragmento de la misma. Además, la secuenciación de la secuencia diana o fragmento de la misma puede ser una secuenciación casi en tiempo real. La secuenciación de la secuencia diana o fragmento de la misma puede llevarse a cabo según los métodos descritos anteriormente (Procedimientos experimentales: Matranga *et al.*, 2014; y Gire, *et al.*, 2014). La secuenciación de la secuencia diana o fragmento de la misma puede comprender la secuenciación paralela de una pluralidad de secuencias diana. La secuenciación de la secuencia diana o fragmento de la misma puede comprender la secuenciación Illumina.

El análisis de la secuencia diana o fragmento de la misma que se hibrida con una o más de las sondas seleccionadas puede ser un análisis de identificación, en el que la hibridación de una sonda seleccionada con la secuencia diana o un fragmento de la misma indica la presencia de la secuencia diana dentro de la muestra.

Actualmente, los diagnósticos primarios se basan en los síntomas que tiene un paciente. Sin embargo, diversas enfermedades pueden compartir síntomas idénticos, por lo que los diagnósticos dependen mucho de las estadísticas. Por ejemplo, la malaria desencadena síntomas similares a los de la gripe: dolor de cabeza, fiebre, escalofríos, dolor en las articulaciones, vómitos, anemia hemolítica, ictericia, hemoglobina en la orina, daño en la retina y convulsiones. Estos síntomas también son comunes para la septicemia, la gastroenteritis y las enfermedades virales. Entre estos últimos, la fiebre hemorrágica del Ébola tiene los siguientes síntomas: fiebre, dolor de garganta, dolor muscular, dolores de cabeza, vómitos, diarrea, sarpullido, disminución de la función del hígado y los riñones, hemorragia interna y externa.

Cuando se presenta un paciente a una unidad médica, por ejemplo en África tropical, los diagnósticos básicos concluirán en malaria porque, estadísticamente, la malaria es la enfermedad más probable en esa región de África. Por consiguiente, el paciente recibe tratamiento para la malaria, aunque es posible que el paciente no haya contraído la enfermedad y el paciente termine sin recibir el tratamiento correcto. Esta falta de tratamiento correcto puede poner en riesgo la vida especialmente cuando la enfermedad que el paciente contrae presenta una evolución rápida. Puede ser demasiado tarde antes de que el personal médico se dé cuenta de que el tratamiento dado al paciente es ineficaz y llegue al diagnóstico correcto y administre el tratamiento adecuado al paciente.

El método de la invención proporciona una solución a esta situación. De hecho, debido a que el número de ARN guía puede reducirse drásticamente, esto hace posible proporcionar en un solo chip sondas seleccionadas divididas en grupos, siendo cada grupo específico para una enfermedad, de modo que una pluralidad de enfermedades, por ejemplo, infección viral, puede diagnosticarse al mismo tiempo. Gracias a la invención pueden diagnosticarse más de 3 enfermedades en un solo chip, preferiblemente más de 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 enfermedades al mismo tiempo, preferiblemente las enfermedades que se presentan con mayor frecuencia dentro de la población de un área geográfica determinada. Dado que cada grupo de sondas seleccionadas es específico para una de las enfermedades diagnosticadas, puede realizarse un diagnóstico más preciso, disminuyendo así el riesgo de administrar un tratamiento incorrecto al paciente.

En otros casos, una enfermedad como una infección viral puede ocurrir sin ningún síntoma, o haber causado síntomas pero estos desaparecieron antes de que el paciente fuera presentado al personal médico. En tales casos, o bien el paciente no busca asistencia médica o bien el diagnóstico se complica por la ausencia de síntomas el día de la presentación.

La presente invención también puede usarse junto con otros métodos de diagnóstico de enfermedades, identificación de patógenos y optimización del tratamiento basados en la detección de ácidos nucleicos, tales como ARNm en muestras en bruto no purificadas.

El método de la invención también proporciona una poderosa herramienta para abordar esta situación. En efecto, dado que dentro de un único diagnóstico están comprendidos una pluralidad de grupos de ARN guía seleccionados, siendo cada grupo específico de una de las enfermedades más comunes que se producen en la población de la zona determinada, el personal médico sólo necesita poner en contacto una muestra biológica tomada del paciente con el chip. La lectura del chip revela las enfermedades que ha contraído el paciente.

En algunos casos, el paciente se presenta al personal médico para el diagnóstico de síntomas particulares. El método de la invención permite no sólo identificar qué enfermedad provoca estos síntomas sino al mismo tiempo determinar si el paciente padece otra enfermedad que no conocía.

Esta información puede ser de suma importancia a la hora de buscar los mecanismos de un brote epidémico. De hecho, grupos de pacientes con virus idénticos también muestran patrones temporales que sugieren vínculos de transmisión de sujeto a sujeto.

Examen para detectar alteraciones genéticas microbianas

En determinadas realizaciones de ejemplo, los sistemas CRISPR dados a conocer en el presente documento pueden usarse para realizar un examen de detección de alteraciones genéticas microbianas. Tales métodos pueden ser útiles, por ejemplo, para trazar rutas microbianas y redes funcionales. Las células microbianas pueden modificarse genéticamente y luego examinarse en diferentes condiciones experimentales. Tal como se describió anteriormente, las realizaciones dadas a conocer en el presente documento pueden examinar múltiples moléculas diana en una única muestra, o una única diana en un único volumen discreto individual de manera múltiple. Los microbios modificados genéticamente pueden modificarse para incluir una secuencia de código de barras de ácido nucleico que identifique la modificación genética particular portada por una célula microbiana particular o población de células microbianas. Un código de barras es una secuencia corta de nucleótidos (por ejemplo, ADN, ARN o combinaciones de los mismos) que se usa como identificador. Un código de barras de ácido nucleico puede tener una longitud de 4 a 100 nucleótidos y ser monocatenario o bicatenario. Los métodos para identificar células con códigos de barras se conocen en la técnica. Por consiguiente, los ARN guía de los sistemas efectores CRISPR

descritos en el presente documento pueden usarse para detectar el código de barras. La detección de la señal detectable positiva indica la presencia de una modificación genética particular en la muestra. Los métodos dados a conocer en el presente documento pueden combinarse con otros métodos para detectar lecturas fenotípicas o genotípicas complementarias que indiquen el efecto de la modificación genética en las condiciones experimentales sometidas a prueba. Las modificaciones genéticas que se examinarán pueden incluir, pero no se limitan a, una inserción génica, una inactivación génica, inversiones, translocaciones, transposiciones o una o más inserciones, delecciones, sustituciones, mutaciones o adición de ácidos nucleicos que codifican para un epítipo con una consecuencia funcional tal como la alteración de la estabilidad o la detección de proteínas. Del mismo modo, los métodos descritos en el presente documento pueden usarse en aplicaciones de biología de síntesis para examinar la funcionalidad de disposiciones específicas de elementos reguladores de genes y módulos de expresión de genes.

En determinadas realizaciones de ejemplo, los métodos pueden usarse para examinar hipomorfos. La generación de hipomorfos y su uso en la identificación de genes funcionales bacterianos clave y la identificación de nuevos productos terapéuticos de antibióticos tal como se da a conocer en el documento PCT/US2016/060730 titulado "Multiplex High-Resolution Detection of Micro-organism Strains, Related Kits, Diagnostic Methods and Screening Assays" presentado el 4 de noviembre de 2016.

Las diferentes condiciones experimentales pueden comprender la exposición de las células microbianas a diferentes agentes químicos, combinaciones de agentes químicos, diferentes concentraciones de agentes químicos o combinaciones de agentes químicos, diferentes duraciones de exposición a agentes químicos o combinaciones de agentes químicos, diferentes parámetros físicos, o ambas cosas. En determinadas realizaciones de ejemplo, el agente químico es un antibiótico o antiviral. Los diferentes parámetros físicos a examinar pueden incluir diferentes temperaturas, presiones atmosféricas, diferentes concentraciones de gases atmosféricos y no atmosféricos, diferentes niveles de pH, diferentes composiciones de medios de cultivo o una combinación de los mismos.

Examen de muestras ambientales

Los métodos dados a conocer en el presente documento también pueden usarse para examinar muestras ambientales para detectar contaminantes mediante la detección de la presencia de polipéptidos o ácidos nucleicos diana. Por ejemplo, en algunas realizaciones, la invención proporciona un método para detectar microbios, que comprende: exponer un sistema CRISPR tal como se describe en el presente documento a una muestra; activar una proteína efectora de ARN mediante la unión de uno o más ARN guía a uno o más ARN diana específicos de microbios o uno o más ARN desencadenantes de manera que se produzca una señal positiva detectable. La señal positiva puede detectarse y es indicativa de la presencia de uno o más microbios en la muestra. En algunas realizaciones, el sistema CRISPR puede estar sobre un sustrato tal como se describe en el presente documento, y el sustrato puede exponerse a la muestra. En otras realizaciones, puede aplicarse el mismo sistema CRISPR y/o un sistema CRISPR diferente a múltiples ubicaciones discretas en el sustrato. En realizaciones adicionales, el sistema CRISPR diferente puede detectar un microbio diferente en cada ubicación. Tal como se describió con mayor detalle anteriormente, un sustrato puede ser un sustrato de materiales flexibles incluyendo, por ejemplo, pero no se limita a, un sustrato de papel, un sustrato de tela o un sustrato basado en polímero flexible.

Según la invención, el sustrato puede exponerse a la muestra de manera pasiva, sumergiendo temporalmente el sustrato en un fluido que va a muestrearse, aplicando un fluido que va a someterse a prueba al sustrato, o poniendo en contacto una superficie que va a someterse a prueba con el sustrato. Puede usarse cualquier medio de introducción de la muestra al sustrato según sea apropiado.

Tal como se describe en el presente documento, una muestra para su uso con la invención puede ser una muestra biológica o ambiental, tal como una muestra de comida (frutas o verduras frescas, carnes), una muestra de bebida, una superficie de papel, una superficie de tela, una superficie de metal, una superficie de madera, una superficie de plástico, una muestra de suelo, una muestra de agua dulce, una muestra de agua residual, una muestra de agua salada, exposición al aire atmosférico u otra muestra de gas, o una combinación de las mismas. Por ejemplo, puede tomarse una muestra de las superficies domésticas/comerciales/industriales elaboradas de cualquier material incluyendo, pero no se limitan a, metal, madera, plástico, caucho o similares, y analizarse para detectar contaminantes. Las muestras de suelo pueden someterse a prueba para detectar la presencia de bacterias o parásitos patógenos, u otros microbios, tanto con fines ambientales como para pruebas de enfermedades humanas, animales o vegetales. Las muestras de agua, tales como las muestras de agua dulce, las muestras de aguas residuales o las muestras de agua salada, pueden evaluarse en cuanto a limpieza y seguridad, y/o potabilidad, para detectar la presencia de, por ejemplo, *Cryptosporidium parvum*, *Giardia lamblia*, u otra contaminación microbiana. En realizaciones adicionales, una muestra biológica puede obtenerse de una fuente que incluye, pero no se limita a, una muestra de tejido, saliva, sangre, plasma, sueros, heces, orina, esputo, mucosa, linfa, líquido sinovial, líquido cefalorraquídeo, ascitis, derrame pleural, seroma, pus o frotis de piel o superficie de una membrana mucosa. En algunas realizaciones particulares, una muestra ambiental o muestras biológicas pueden ser muestras en bruto y/o una o más moléculas diana pueden no purificarse o amplificarse a partir de la muestra antes de la aplicación del método. La identificación de microbios puede ser útil y/o necesaria para cualquier número de aplicaciones y, por tanto, cualquier tipo de muestra de cualquier fuente que un experto en la técnica considere apropiada puede usarse según la invención.

En algunas realizaciones, comprobar la contaminación de los alimentos por bacterias, tales como *E. coli*, en restaurantes u otros proveedores de alimentos; superficies de alimentos; analizar el agua para detectar patógenos tales como *Salmonella*, *Campylobacter* o *E. coli*; también comprobar la calidad de los alimentos para que los fabricantes y los reguladores determinen la pureza de las fuentes de carne; identificar la contaminación del aire con patógenos tales como *Legionella*; comprobar si la cerveza está contaminada o estropeada por patógenos tales como *Pediococcus* y *Lactobacillus*; contaminación del queso pasteurizado o no pasteurizado por bacterias u hongos durante la fabricación.

Un microbio según la invención puede ser un microbio patógeno o un microbio que da como resultado el deterioro de alimentos o productos consumibles. Un microbio patógeno puede ser patógeno o de otro modo indeseable para seres humanos, animales o plantas. Para propósitos humanos o animales, un microbio puede causar una enfermedad o resultar en una enfermedad. Las aplicaciones animales o veterinarias de la presente invención pueden identificar animales infectados por un microbio. Por ejemplo, los métodos y sistemas de la invención pueden identificar animales de compañía con patógenos que incluyen, pero no se limita, tos de las perreras, virus de la rabia y *Dirofilaria immitis*. En otras realizaciones, los métodos y sistemas de la invención pueden usarse para pruebas de parentesco con fines de cría. Un microbio vegetal puede provocar daño o enfermedad a una planta, reducción en el rendimiento o alteración de rasgos tales como el color, el sabor, la consistencia, el olor, para fines de contaminación de alimentos o consumibles, un microbio puede afectar negativamente el sabor, el olor, el color, la consistencia u otras propiedades comerciales del alimento o producto consumible. En determinadas realizaciones de ejemplo, el microbio es una especie bacteriana. La bacteria puede ser un psicótrofo, un coliforme, una bacteria del ácido láctico o una bacteria formadora de esporas. En determinadas realizaciones de ejemplo, las bacterias pueden ser cualquier especie bacteriana que provoque una enfermedad o dolencia, o que de otro modo resulte en un producto o rasgo no deseado. Las bacterias según la invención pueden ser patógenas para seres humanos, animales o plantas.

Tipos de muestra

Las muestras apropiadas para su uso en los métodos dados a conocer en el presente documento incluyen cualquier muestra biológica convencional obtenida de un organismo o una parte del mismo, tal como una planta, un animal, una bacteria y similares. En realizaciones particulares, la muestra biológica se obtiene de un sujeto animal, tal como un sujeto humano. Una muestra biológica es cualquier muestra sólida o líquida obtenida de, excretada por o secretada por cualquier organismo vivo, incluyendo, sin limitación, organismos unicelulares, tales como bacterias, levaduras, protozoos y amebas, entre otros, organismos multicelulares (tales como plantas o animales, incluyendo muestras de un sujeto humano sano o aparentemente sano o de un paciente humano afectado por un estado o enfermedad que va a diagnosticarse o investigarse, tal como una infección por un microorganismo patógeno, tal como una bacteria o virus patógeno). Por ejemplo, una muestra biológica puede ser un líquido biológico obtenido de, por ejemplo, sangre, plasma, suero, orina, heces, esputo, moco, líquido linfático, líquido sinovial, bilis, ascitis, derrame pleural, seroma, saliva, líquido cefalorraquídeo, humor acuoso o vítreo, o cualquier secreción corporal, un trasudado, un exudado (por ejemplo, líquido obtenido de un absceso o cualquier otro sitio de infección o inflamación), o líquido obtenido de una articulación (por ejemplo, una articulación normal o una articulación afectada por una enfermedad, tal como artritis reumatoide, osteoartritis, gota o artritis séptica), o un frotis de piel o superficie de una membrana mucosa.

Una muestra también puede ser una muestra obtenida de cualquier órgano o tejido (incluyendo una biopsia o una muestra de autopsia, tal como una biopsia de tumor) o puede incluir una célula (ya sea una célula primaria o una célula cultivada) o un medio acondicionado por cualquier célula, tejido u órgano. Las muestras a modo de ejemplo incluyen, sin limitación, células, lisados de células, frotis de sangre, preparaciones de citocentrífuga, frotis de citología, líquidos corporales (por ejemplo, sangre, plasma, suero, saliva, esputo, orina, lavado broncoalveolar, semen, etc.), biopsias de tejido (por ejemplo, biopsias de tumores), aspirados con aguja fina y/o cortes de tejido (por ejemplo, cortes de tejido en criostato y/o cortes de tejido incrustados en parafina). En otros ejemplos, la muestra incluye células tumorales circulantes (que pueden identificarse mediante marcadores de superficie celular). En ejemplos particulares, las muestras se usan directamente (por ejemplo, recién preparadas o congeladas), o pueden manipularse antes de su uso, por ejemplo, por fijación (por ejemplo, usando formalina) y/o incrustación en cera (tal como muestras de tejido fijadas en formalina e incrustadas en parafina (FFPE)). Se apreciará que puede usarse cualquier método para obtener tejido de un sujeto, y que la selección del método usado dependerá de diversos factores tales como el tipo de tejido, la edad del sujeto o los procedimientos disponibles para el médico. Las técnicas convencionales para la adquisición de tales muestras están disponibles en la técnica. Véase, por ejemplo Schluger *et al.*, J. Exp. Med. 176:1327-33 (1992); Bigby *et al.*, Am. Rev. Respir. Dis. 133:515-18 (1986); Kovacs *et al.*, NEJM 318:589-93 (1988); y Ognibene *et al.*, Am. Rev. Respir. Dis. 129:929-32 (1984).

En otras realizaciones, una muestra puede ser una muestra ambiental, tal como agua, suelo o una superficie, tal como una superficie industrial o médica. En algunas realizaciones, los métodos tales como los dados a conocer en la publicación de patente estadounidense n.º 2013/0190196 pueden aplicarse para la detección de firmas de ácidos nucleicos, específicamente niveles de ARN, directamente a partir de muestras celulares en bruto con un alto grado de sensibilidad y especificidad. Las secuencias específicas de cada patógeno de interés pueden identificarse o seleccionarse comparando las secuencias codificantes del patógeno de interés con todas las secuencias codificantes en otros organismos mediante el software BLAST.

Varias realizaciones de la presente descripción implican el uso de procedimientos y enfoques que se conocen en la técnica que fraccionen de manera exitosa muestras clínicas de sangre. Véase, por ejemplo, el procedimiento descrito en Han Wei Hour *et al.*, Microfluidic Devices for Blood Fractionation, *Micromachines* 2011, 2, 319-343; Ali Asgar S. Bhagat *et al.*, Dean Flow Fractionation (DFF) Isolation of Circulating Tumor Cells (CTCs) from Blood, 15.^a Conferencia internacional sobre sistemas miniaturizados para química y ciencias biológicas, 2 al 6 de octubre de 2011, Seattle, WA; y la publicación de patente internacional n.º WO2011109762. Las muestras de sangre se expanden habitualmente en cultivo para aumentar el tamaño de muestra con fines de prueba. En algunas realizaciones de la presente invención, las muestras de sangre u otras muestras biológicas pueden usarse en métodos tal como se describen en el presente documento sin necesidad de expansión en cultivo.

Además, varias realizaciones de la presente divulgación implican el uso de procedimientos y enfoques que se conocen en la técnica que aislen de manera exitosa patógenos de la sangre completa usando microcanales en espiral, tal como se describe en Han Wei Hour *et al.*, Pathogen Isolation from Whole Blood Using Spiral Microchannel, n.º de caso 15995JR, Instituto de Tecnología de Massachusetts, manuscrito en preparación.

Debido a la mayor sensibilidad de las realizaciones descritas en el presente documento, en determinadas realizaciones de ejemplo, los ensayos y métodos pueden ejecutarse en muestras en bruto o muestras en las que las moléculas diana que van a detectarse no se fraccionan ni purifican adicionalmente de la muestra.

Microbios de ejemplo

La realización dada a conocer en el presente documento puede usarse para detectar varios microbios diferentes. El término microbio, tal como se usa en el presente documento, incluye bacterias, hongos, protozoos, parásitos y virus.

Bacterias

Lo siguiente proporciona una lista de ejemplo de los tipos de microbios que puede detectarse usando las realizaciones dadas a conocer en el presente documento. En determinadas realizaciones de ejemplo, el microbio es una bacteria. Los ejemplos de bacterias que pueden detectarse según los métodos dados a conocer incluyen sin limitación uno cualquiera o más de (o cualquier combinación de) *Acinetobacter baumannii*, *Actinobacillus* sp., *Actinomycetes*, *Actinomyces* sp. (tal como *Actinomyces israelii* y *Actinomyces naeslundii*), *Aeromonas* sp. (tal como *Aeromonas hydrophila*, *Aeromonas veronii* biovar *sobria* (*Aeromonas sobria*), y *Aeromonas caviae*), *Anaplasma phagocytophilum*, *Anaplasma marginale*, *Alcaligenes xilosoxidans*, *Acinetobacter baumannii*, *Actinobacillus actinomycetemcomitans*, *Bacillus* sp. (tal como *Bacillus anthracis*, *Bacillus cereus*, *Bacillus subtilis*, *Bacillus thuringiensis* y *Bacillus stearothermophilus*), *Bacteroides* sp. (tal como *Bacteroides fragilis*), *Bartonella* sp. (tal como *Bartonella bacilliformis* y *Bartonella henselae*, *Bifidobacterium* sp., *Bordetella* sp. (tal como *Bordetella pertussis*, *Bordetella parapertussis* y *Bordetella bronchiseptica*), *Borrelia* sp. (tal como *Borrelia recurrentis* y *Borrelia burgdorferi*), *Brucella* sp. (tal como *Brucella abortus*, *Brucella canis*, *Brucella melitensis* y *Brucella suis*), *Burkholderia* sp. (tal como *Burkholderia pseudomallei* y *Burkholderia cepacia*), *Campylobacter* sp. (tal como *Campylobacter jejuni*, *Campylobacter coli*, *Campylobacter lari* y *Campylobacter fetus*), *Capnocytophaga* sp., *Cardiobacterium hominis*, *Chlamydia trachomatis*, *Chlamydophila pneumoniae*, *Chlamydophila psittaci*, *Citrobacter* sp., *Coxiella burnetii*, *Corynebacterium* sp. (tal como *Corynebacterium diphtheriae*, *Corynebacterium jeikeum* y *Corynebacterium*), *Clostridium* sp. (tal como *Clostridium perfringens*, *Clostridium difficile*, *Clostridium botulinum* y *Clostridium tetani*), *Eikenella corrodens*, *Enterobacter* sp. (tal como *Enterobacter aerogenes*, *Enterobacter agglomerans*, *Enterobacter cloacae* y *Escherichia coli*, incluyendo *Escherichia coli* oportunista, tal como *E. coli* enterotoxigénica, *E. coli* enteroinvasiva, *E. coli* enteropatógena, *E. coli* enterohemorrágica, *E. coli* enteroagregativa y *E. coli* uropatógena) *Enterococcus* sp. (tal como *Enterococcus faecalis* y *Enterococcus faecium*) *Ehrlichia* sp. (tal como *Ehrlichia chaffeensis* y *Ehrlichia canis*), *Epidermophyton floccosum*, *Erysipelothrix rhusiopathiae*, *Eubacterium* sp., *Francisella tularensis*, *Fusobacterium nucleatum*, *Gardnerella vaginalis*, *Gemella morbillorum*, *Haemophilus* sp. (tal como *Haemophilus influenzae*, *Haemophilus ducreyi*, *Haemophilus aegyptius*, *Haemophilus parainfluenzae*, *Haemophilus haemolyticus* y *Haemophilus parahaemolyticus*, *Helicobacter* sp. (tal como *Helicobacter pylori*, *Helicobacter cinaedi* y *Helicobacter fennelliae*), *Kingella kingii*, *Klebsiella* sp. (tal como *Klebsiella pneumoniae*, *Klebsiella granulomatis* y *Klebsiella oxytoca*), *Lactobacillus* sp., *Listeria monocytogenes*, *Leptospira interrogans*, *Legionella pneumophila*, *Leptospira interrogans*, *Peptostreptococcus* sp., *Mannheimia hemolytica*, *Microsporum canis*, *Moraxella catarrhalis*, *Morganella* sp., *Mobiluncus* sp., *Micrococcus* sp., *Mycobacterium* sp. (tal como *Mycobacterium leprae*, *Mycobacterium tuberculosis*, *Mycobacterium paratuberculosis*, *Mycobacterium intracellulare*, *Mycobacterium avium*, *Mycobacterium bovis*, y *Mycobacterium marinum*), *Mycoplasma* sp. (tal como *Mycoplasma pneumoniae*, *Mycoplasma hominis*, y *Mycoplasma genitalium*), *Nocardia* sp. (tal como *Nocardia asteroides*, *Nocardia cyriacigeorgica* y *Nocardia brasiliensis*), *Neisseria* sp. (tal como *Neisseria gonorrhoeae* y *Neisseria meningitidis*), *Pasteurella multocida*, *Pityrosporum orbiculare* (*Malassezia furfur*), *Plesiomonas shigelloides*, *Prevotella* sp., *Porphyromonas* sp., *Prevotella melaninogenica*, *Proteus* sp. (tal como *Proteus vulgaris* y *Proteus mirabilis*), *Providencia* sp. (tal como *Providencia alcalifaciens*, *Providencia rettgeri* y *Providencia stuartii*), *Pseudomonas aeruginosa*, *Propionibacterium acnes*, *Rhodococcus equi*, *Rickettsia* sp. (tal como *Rickettsia rickettsii*, *Rickettsia akari* y *Rickettsia prowazekii*, *Orientia tsutsugamushi* (anteriormente: *Rickettsia tsutsugamushi*) y *Rickettsia typhi*), *Rhodococcus* sp., *Serratia marcescens*, *Stenotrophomonas maltophilia*, *Salmonella* sp. (tal como *Salmonella enterica*, *Salmonella typhi*, *Salmonella paratyphi*, *Salmonella enteritidis*, *Salmonella choleraesuis* y *Salmonella typhimurium*), *Serratia* sp. (tal como *Serratia marcescens* y *Serratia liquifaciens*), *Shigella* sp. (tal como *Shigella*

dysenteriae, *Shigella flexneri*, *Shigella boydii* y *Shigella sonnei*), *Staphylococcus* sp. (tal como *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus hemolyticus*, *Staphylococcus saprophyticus*), *Streptococcus* sp. (tal como *Streptococcus pneumoniae* (por ejemplo *Streptococcus pneumoniae* serotipo 4 resistente a cloranfenicol, *Streptococcus pneumoniae* serotipo 6B resistente a espectinomicina, *Streptococcus pneumoniae* serotipo 9V resistente a estreptomina, *Streptococcus pneumoniae* serotipo 14 resistente a eritromicina, *Streptococcus pneumoniae* serotipo 14 resistente a optoquina, *Streptococcus pneumoniae* serotipo 18C resistente a rifampicina, *Streptococcus pneumoniae* serotipo 19F resistente a tetraciclina, *Streptococcus pneumoniae* serotipo 19F resistente a penicilina, y *Streptococcus pneumoniae* serotipo 23F resistente a trimetoprim, *Streptococcus pneumoniae* serotipo 4 resistente a cloranfenicol, *Streptococcus pneumoniae* serotipo 6B resistente a espectinomicina, *Streptococcus pneumoniae* serotipo 9V resistente a estreptomina, *Streptococcus pneumoniae* serotipo 14 resistente a optoquina, *Streptococcus pneumoniae* serotipo 18C resistente a rifampicina, *Streptococcus pneumoniae* serotipo 19F resistente a penicilina o *Streptococcus pneumoniae* serotipo 23F resistente a trimetoprim), *Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus pyogenes*, estreptococos de grupo A, *Streptococcus pyogenes*, estreptococos de grupo B, *Streptococcus agalactiae*, estreptococos de grupo C, *Streptococcus anginosus*, *Streptococcus equismilis*, estreptococos de grupo D, *Streptococcus bovis*, estreptococos de grupo F, y *Streptococcus anginosus*, estreptococos de grupo G, *Spirillum minus*, *Streptobacillus moniliformi*, *Treponema* sp. (tal como *Treponema carateum*, *Treponema petenue*, *Treponema pallidum* y *Treponema endemicum*, *Trichophyton rubrum*, *T. mentagrophytes*, *Tropheryma whippelii*, *Ureaplasma urealyticum*, *Veillonella* sp., *Vibrio* sp. (tal como *Vibrio cholerae*, *Vibrio parahemolyticus*, *Vibrio vulnificus*, *Vibrio parahaemolyticus*, *Vibrio vulnificus*, *Vibrio alginolyticus*, *Vibrio mimicus*, *Vibrio hollisae*, *Vibrio fluvialis*, *Vibrio metchnikovii*, *Vibrio damsela* y *Vibrio furnisii*), *Yersinia* sp. (tal como *Yersinia enterocolitica*, *Yersinia pestis* y *Yersinia pseudotuberculosis*) y *Xanthomonas maltophilia* entre otras.

Hongos

En determinadas realizaciones de ejemplo, el microbio es un hongo o una especie fúngica. Los ejemplos de hongos que puede detectarse según los métodos dados a conocer incluyen, sin limitación, uno cualquiera o más de (o cualquier combinación de), *Aspergillus*, *Blastomyces*, *Candidiasis*, *Coccidioidomycosis*, *Cryptococcus neoformans*, *Cryptococcus gattii*, sp. *Histoplasma* sp. (tal como *Histoplasma capsulatum*), *Pneumocystis* sp. (tal como *Pneumocystis jirovecii*), *Stachybotrys* (tal como *Stachybotrys chartarum*), *Mucroymosis*, *Sporothrix*, tiña de infecciones oculares fúngicas, *Exserohilum*, *Cladosporium*.

En determinadas realizaciones de ejemplo, el hongo es una levadura. Los ejemplos de levadura que puede detectarse según los métodos dados a conocer incluyen, sin limitación, una o más de (o cualquier combinación de), especie de *Aspergillus* (tal como *Aspergillus fumigatus*, *Aspergillus flavus* y *Aspergillus clavatus*), *Cryptococcus* sp. (tal como *Cryptococcus neoformans*, *Cryptococcus gattii*, *Cryptococcus laurentii* y *Cryptococcus albidus*), una especie de *Geotrichum*, una especie de *Saccharomyces*, una especie de *Hansenula*, una especie de *Candida* (tal como *Candida albicans*), una especie de *Kluyveromyces*, una especie de *Debaryomyces*, una especie de *Pichia*, o combinación de las mismas. En determinadas realizaciones de ejemplo, el hongo es un moho. Los ejemplos de mohos incluyen, pero no se limitan a, una especie de *Penicillium*, una especie de *Cladosporium*, una especie de *Byssoschlamys*, o una combinación de los mismos.

Protozoos

En determinadas realizaciones de ejemplo, el microbio es un protozoo. Los ejemplos de protozoos que pueden detectarse según los métodos y dispositivos dados a conocer incluyen, sin limitación, uno cualquiera o más de (o cualquier combinación de), *Euglenozoa*, *Heterolobosea*, *Diplomonadida*, *Amoebozoa*, *Blastocystis* y *Apicomplexa*. Los *Euglenozoa* de ejemplo incluyen, pero no se limitan a, *Trypanosoma cruzi* (enfermedad de Chagas), *T. brucei gambiense*, *T. brucei rhodesiense*, *Leishmania braziliensis*, *L. infantum*, *L. mexicana*, *L. major*, *L. tropica* y *L. donovani*. Los *Heterolobosea* de ejemplo incluyen, pero no se limitan a, *Naegleria fowleri*. Los *Diplomonadids* de ejemplo incluyen, pero no se limitan a, *Giardia intestinalis* (*G. lamblia*, *G. duodenalis*). Los *Amoebozoa* de ejemplo incluyen, pero no se limitan a, *Acanthamoeba castellanii*, *Balamuthia mandrillaris*, *Entamoeba histolytica*. Los blastocitos de ejemplo incluyen, pero no se limitan a, *Blastocystis hominis*. Los *Apicomplexa* de ejemplo incluyen, pero no se limitan a, *Babesia microti*, *Cryptosporidium parvum*, *Ciclospora cayetanensis*, *Plasmodium falciparum*, *P. vivax*, *P. ovale*, *P. malariae* y *Toxoplasma gondii*.

Parásitos

En determinadas realizaciones de ejemplo, el microbio es un parásito. Los ejemplos de parásitos que pueden detectarse según métodos dados a conocer incluyen, sin limitación, uno o más de (o cualquier combinación de) una especie de *Onchocerca* y una especie de *Plasmodium*.

Virus

En determinadas realizaciones de ejemplo, los sistemas, dispositivos y métodos dados a conocer en el presente documento se refieren a detectar virus en una muestra. Las realizaciones dadas a conocer en el presente documento pueden usarse para detectar una infección viral (por ejemplo de un sujeto o planta) o determinar una

cepa viral, incluyendo cepas virales que difieren en un polimorfismo de un solo nucleótido. El virus puede ser un virus de ADN, un virus de ARN o un retrovirus. Los ejemplos no limitativos de virus útiles con la presente invención incluyen, pero no se limitan a, ébola, sarampión, SARS, Chikungunya, hepatitis, Marburgo, fiebre amarilla, MERS, dengue, Lassa, gripe, rabdovirus o VIH. Un virus de hepatitis puede incluir hepatitis A, hepatitis B o hepatitis C. Un virus de la gripe puede incluir, por ejemplo, gripe A o gripe B. Un VIH puede incluir VIH 1 o VIH 2. En determinadas realizaciones de ejemplo, la secuencia viral puede ser un virus respiratorio sincicial humano, virus del ébola de Sudán, virus de Bundibugyo, virus del ébola de Tai Forest, virus del ébola de Reston, Achimota, *Flavivirus* de Aedes, virus de Aguacate, virus de Akabane, reptarenavirus de Alethinophidia, *Mammarenavirus* de Allpahuayo, *Arenavirus* de Amapari, virus de los Andes, virus de Apoi, virus de Aravan, virus de Aroa, virus de Arumwot, paramixovirus del salmón del Atlántico, *Lyssavirus* del murciélago australiano, bornavirus aviar, metapneumovirus aviar, paramixovirus aviar, virus del pingüino o de las Islas Malvinas, poliomavirus BK, virus de Bagaza, virus de Banna, herpesvirus del murciélago, sapovirus del murciélago, *Mammarenavirus* de Bear Canon, virus de Beilong, *Betacoronavirus*, *Betapapillomavirus* 1-6, virus de Bhanja, *Lyssavirus* del murciélago Bokeloh, virus de la enfermedad de Borna, virus de Bourbon, *Hepacivirus* bovino, virus de la paragripe bovina 3, virus respiratorio sincicial bovino, virus de Brazoran, virus de Bunyamwera, virus *Caliciviridae*, virus de la encefalitis de California, virus *Candiru*, virus del moquillo canino, *Pneumovirus* canino, virus del cedro, virus de agente de fusión celular, *Morbillivirus* cetáceo, virus de Chandipura, virus de Chaoyang, *Mammarenavirus* de Chapare, virus de Chikungunya, *Papillomavirus* del mono colobo, virus de la fiebre por garrapatas de Colorado, virus de la viruela vacuna, virus de la fiebre hemorrágica de Crimea-Congo, *Flavivirus* de Culex, *Mammarenavirus* de Cupixi, virus del dengue, virus de Dobrava-Belgrado, virus de Donggang, virus de Dugbe, virus de Duvenhage, virus de la encefalitis equina occidental, virus del murciélago de Entebbe, *Enterovirus* A-D, *Lyssavirus* 1-2 del murciélago europeo, virus de Eyach, *Morbillivirus* felino, paramixovirus de Fer-de-Lance, virus del río Fitzroy, virus *Flaviviridae*, *Mammarenavirus* de Flexal, virus C de GB, virus de Gairo, *Gemycircularvirus*, paramixovirus SF02 de ganso, virus de Great Island, *Mammarenavirus* de Guanarito, virus de Hantaan, *Hantavirus* Z10, virus de Heartland, virus de Hendra, hepatitis A/B/C/E, virus de la hepatitis delta, bocavirus humano, coronavirus humano, retrovirus K endógeno humano, coronavirus entérico humano, virus-1 de ADN circular asociado a los genitales humanos, herpesvirus humano 1-8, virus de la inmunodeficiencia humana 1/2, *Mastadenovirus* humano A-G, *Papillomavirus* humano, virus de la paragripe humana 1-4, *Parechovirus* humano, picornavirus humano, smacovirus humano, *Lyssavirus* de Ikoma, virus de Ilheus, gripe A-C, *Mammarenavirus* de Ippy, virus de Irkut, virus J, *Polyomavirus* JC, virus de la encefalitis japonesa, *Mammarenavirus* de Junin, *Polyomavirus* KI, virus de Kadipiro, virus del río Kamiti, virus de Kedougou, virus de Khujand, virus de Kokobera, virus de la enfermedad de la selva de Kyasanur, virus del murciélago de Lagos, virus de Langat, *Mammarenavirus* de Lassa, *Mammarenavirus* Latino, virus de Leopards Hill, virus de Liao ning, virus de Ljungan, virus de Lloviu, virus de la enfermedad de Louping, *Mammarenavirus* Lujo, *Mammarenavirus* Luna, virus de Lunk, *Mammarenavirus* de la coriomeningitis linfocítica, *Lyssavirus* Ozernoe, virus MSS12.225, *Mammarenavirus* Machupo, *Mamastrovirus* 1, virus de Manzanilla, virus de Mapuera, virus de Marburgo, virus Mayaro, virus del sarampión, virus de Menangle, virus de Mercadeo, *Polyomavirus* de células de Merkel, coronavirus del síndrome respiratorio de Oriente Medio, *Mammarenavirus* de Mobala, virus de Modoc, virus de Moijang, virus de Mokolo, virus de la viruela de los simios, virus de la leucoencefalitis de *Myotis* de Montana, virus reagrupado 29 de Lassa de Mopeia, *Mammarenavirus* de Mopeia, virus de Morogoro, virus de Mossman, virus de las paperas, virus de la neumonía murina, virus de la encefalitis del valle de Murray, virus de Nariva, virus de la enfermedad de Newcastle, virus de Nipah, virus de Norwalk, *Hepacivirus* de la rata noruega, virus de Ntaya, virus de O'nyong-nyong, *Mammarenavirus* de Oliveros, virus de la fiebre hemorrágica de Omsk, virus de Oropouche, virus de la paragripe 5, *Mammarenavirus* de Paraná, virus del río Parramatta, virus de la peste de los pequeños rumiantes, *Mammarenavirus* de Pichande, virus *Picornaviridae*, *Mammarenavirus* de Pirital, *Piscihepevirus* A, virus 1 de la paragripe porcina, *Rubulavirus* porcino, virus de Powassan, virus 1-2 T-linfotrópico de primates, *Erythroparvovirus* 1 de primates, virus de Punta Toro, virus de Puumala, virus de Quang Binh, virus de la rabia, virus de Razdan, *Bornavirus* 1 de reptiles, *Rhinovirus* A-B, virus de la fiebre del valle del Rift, virus de la peste bovina, virus de Río Bravo, virus de Torque Teno de roedores, *Hepacivirus* de roedores, virus del río Ross, *Rotavirus* A-I, virus de Royal Farm, virus de la rubéola, *Mammarenavirus* de Sabia, virus de Salem, flebovirus de Nápoles, flebovirus de Sicilia, virus de Sapporo, virus de Sathuperi, anellovirus de las focas, virus del bosque Semliki, virus de Sendai, virus de Seúl, virus de Sepik, coronavirus relacionado con el síndrome respiratorio agudo grave, fiebre grave con virus síndrome de trombocitopenia, virus de Shamonda, virus del murciélago de Shimonivirus, virus de Shuni, virus de Simbu, Torque teno virus del simio, virus del simio 40-41, virus Sin Nombre, virus Sindbis, pequeño anellovirus, virus de Sosuga, virus de la encefalitis caprina española, virus de Spondweni, virus de la encefalitis de San Luis, virus Sunshine, mini virus similar a TTV, *Mammarenavirus* de Tacaribe, virus de Taila, virus del murciélago de Tamana, *Mammarenavirus* de Tamiami, virus de Tembusu, virus de Thogoto, virus de Thottapalayam, virus de la encefalitis transmitida por garrapatas, virus de Tioman, virus *Togaviridae*, Torque teno canis virus, Torque teno douroucouli virus, Torque teno feliz virus, Torque teno midi virus, Torque teno sus virus, Torque teno tamarin virus, Torque teno virus, Torque teno zalophus virus, virus de Tuhoko, virus de Tula, paramixovirus de Tupaia, virus de Usutu, virus de Uukuniemi, virus de la viruela vacuna, virus de la viruela, virus de la encefalitis equina venezolana, virus de la estomatitis vesicular de Indiana, *Polyomavirus* WU, virus de Wesselsbron, virus del murciélago del Cáucaso occidental, virus del Nilo Occidental, virus de la encefalitis equina occidental, virus de la encefalitis equina occidental, *Mammarenavirus* de Whitewater Arroyo, virus de la fiebre amarilla, virus de Yokose, virus de Yug Bogdanovac, virus del Ébola de Zaire, virus de Zika o secuencia viral del virus Z de *Zygosaccharomyces bailii*. Los ejemplos de virus de ARN que pueden detectarse incluyen uno o más (o cualquier combinación de) virus *Coronaviridae*, un virus *Picornaviridae*, un virus *Caliciviridae*, un virus *Flaviviridae*, un virus *Togaviridae*, un *Bornaviridae*, un *Filoviridae*, un *Paramyxoviridae*, un *Pneumoviridae*,

un *Rhabdoviridae*, un *Arenaviridae*, un *Bunyaviridae*, un *Orthomyxoviridae* o un *Deltavirus*. En determinadas realizaciones de ejemplo, el virus es coronavirus, SARS, poliovirus, rinovirus, hepatitis A, virus de Norwalk, virus de la fiebre amarilla, virus del Nilo Occidental, virus de la hepatitis C, virus de la fiebre del dengue, virus de Zika, virus de la rubéola, virus del río Ross, virus Sindbis, virus de Chikungunya, virus de la enfermedad de Borna, virus del Ébola, virus de Marburgo, virus del sarampión, virus de las paperas, virus de Nipah, virus de Hendra, virus de la enfermedad de Newcastle, virus sincicial respiratorio humano, virus de la rabia, virus de Lassa, hantavirus, virus de la fiebre hemorrágica de Crimea-Congo, gripe o virus de la hepatitis D.

En determinadas realizaciones de ejemplo, el virus puede ser un virus vegetal seleccionado del grupo que comprende el virus del mosaico del tabaco (TMV), el virus del marchitamiento del tomate (TSWV), el virus del mosaico del pepino (CMV), virus Y de la patata (PVY), el virus del mosaico de la coliflor del virus RT (CaMV), virus de la viruela del ciruelo (PPV), virus del mosaico de Brome (BMV), virus X de la patata (PVX), virus de la tristeza de los cítricos (CTV), virus del enanismo amarillo de la cebada (BYDV), virus del enrollamiento de la hoja de la patata (PLRV), virus del enanismo arbustivo del tomate (TBSV), virus esférico del tungro del arroz (RTSV), virus del moteado amarillo del arroz (RYMV), virus de la hoja blanca del arroz (RHBV), virus del rayado fino del maíz (MRFV), virus del mosaico del enanismo del maíz (MDMV), virus del mosaico de la caña de azúcar (SCMV), virus del moteado plumoso de la batata (SPFMV), closterovirus de la vena hundida de la batata (SPSVV), virus de la hoja de abanico de la vid (GFLV), virus de la vid A (GVA), virus de la vid B (GVB), virus de la mancha de la vid (GFKV), virus-1, -2 y -3 asociado con el enrollamiento de la hoja de la vid, (GLRaV-1, -2 y -3), virus del mosaico Arabis (ArMV) o virus asociado con la picadura del tallo de *Rupestris* (RSPAV). En una realización preferida, la molécula de ARN diana forma parte de dicho patógeno o se transcribe a partir de una molécula de ADN de dicho patógeno. Por ejemplo, la secuencia diana puede estar comprendida en el genoma de un virus de ARN. Se prefiere además que la proteína efectora CRISPR hidrolice dicha molécula de ARN diana de dicho patógeno en dicha planta si dicho patógeno infecta o ha infectado dicha planta. Por tanto, se prefiere que el sistema CRISPR pueda escindir la molécula de ARN diana del patógeno de la planta cuando el sistema CRISPR (o las partes necesarias para completarlo) se aplica tanto terapéuticamente, es decir, después de que se haya producido ocurrido la infección, como profilácticamente, es decir, antes de que se produzca la infección.

En determinadas realizaciones de ejemplo, el virus puede ser un retrovirus. Los retrovirus de ejemplo que pueden detectarse usando las realizaciones dadas a conocer en el presente documento incluyen uno o más de o cualquier combinación de virus del género *Alpharetrovirus*, *Betaretrovirus*, *Gammaretrovirus*, *Deltaretrovirus*, *Epsilonretrovirus*, *Lentivirus*, *Spumavirus* o la familia *Metaviridae*, *Pseudoviridae* y *Retroviridae* (incluyendo el VIH), *Hepadnaviridae* (incluyendo el virus de la hepatitis B) y *Caulimoviridae* (incluyendo el virus del mosaico de la coliflor).

En determinadas realizaciones de ejemplo, el virus es un virus de ADN. Los virus de ADN de ejemplo que pueden detectarse usando las realizaciones dadas a conocer en el presente documento incluyen uno o más de (o cualquier combinación de) virus de la familia *Myoviridae*, *Podoviridae*, *Siphoviridae*, *Alloherpesviridae*, *Herpesviridae* (incluyendo virus del herpes humano y virus de la varicela zóster), *Malcoherpesviridae*, *Lipothrixviridae*, *Rudiviridae*, *Adenoviridae*, *Ampullaviridae*, *Ascoviridae*, *Asfarviridae* (incluyendo virus de la fiebre porcina africana), *Baculoviridae*, *Cicadaviridae*, *Clavaviridae*, *Corticoviridae*, *Fuselloviridae*, *Globuloviridae*, *Guttaviridae*, *Hytrosaviridae*, *Iridoviridae*, *Marseilleviridae*, *Mimiviridae*, *Nudiviridae*, *Nimaviridae*, *Pandoraviridae*, *Papillomaviridae*, *Phycodnaviridae*, *Plasmaviridae*, polidnavirus, *Polyomaviridae* (incluyendo virus del simio 40, virus JC, virus BK), *Poxviridae* (incluyendo viruela vacuna y viruela), *Sphaerolipoviridae*, *Tectiviridae*, *Turriviridae*, *Dinodnavirus*, *Salterprovirus*, *Rhizidovirus*, entre otros. En algunas realizaciones, un método para diagnosticar una infección bacteriana específica de especie en un sujeto que se sospecha que tiene una infección bacteriana se describe como obtener una muestra que comprende ácido ribonucleico ribosómico bacteriano del sujeto; poner en contacto la muestra con una o más de las sondas descritas, y detectar la hibridación entre la secuencia de ácido ribonucleico ribosómico bacteriano presente en la muestra y la sonda, en el que la detección de hibridación indica que el sujeto está infectado con *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Acinetobacter baumannii*, *Candida albicans*, *Enterobacter cloacae*, *Enterococcus faecalis*, *Enterococcus faecium*, *Proteus mirabilis*, *Staphylococcus agalactiae* o *Staphylococcus maltophilia* o una combinación de los mismos.

Detección y monitorización de la malaria

La malaria es una patología transmitida por mosquitos provocada por los parásitos *Plasmodium*. Los parásitos se transmiten a las personas a través de las picaduras de mosquitos *Anopheles* hembra infectados. Cinco especies de *Plasmodium* provocan la malaria en seres humanos: *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* y *Plasmodium knowlesi*. Entre ellas, según la Organización Mundial de la Salud (OMS), *Plasmodium falciparum* y *Plasmodium vivax* son las responsables de la mayor amenaza. *P. falciparum* es el parásito de la malaria más prevalente en el continente africano y es responsable de la mayoría de las muertes relacionadas con la malaria a nivel mundial. *P. vivax* es el parásito de la malaria dominante en la mayoría de los países fuera de África subsahariana.

En 2015, 91 países y regiones tenían una transmisión continua de la malaria. Según las últimas estimaciones de la OMS, hubo 212 millones de casos de malaria en 2015 y 429.000 muertes. En regiones con alta transmisión de la malaria, los niños menores de 5 años son particularmente susceptibles a la infección, a la enfermedad y a la muerte; más de dos tercios (el 70%) de todas las muertes por malaria se producen en este grupo de edades. Entre 2010 y

2015, la tasa de mortalidad por malaria en menores de 5 años se redujo en un 29% a nivel mundial. Sin embargo, la malaria sigue siendo una de las principales causas de muerte de niños menores de cinco años, acabando con la vida de un niño cada dos minutos.

5 Tal como describe la OMS, la malaria es una enfermedad febril aguda. En un individuo no inmunizado, los síntomas aparecen a los 7 días o más después de la picadura del mosquito infeccioso. Los primeros síntomas, fiebre, cefalea, escalofríos y vómitos, pueden ser leves y difíciles de reconocer como malaria, sin embargo, si no se tratan en el plazo de 24 horas, la malaria por *P. falciparum* puede evolucionar a enfermedad grave, que a menudo conduce a la muerte.

10 Con frecuencia, los niños con malaria grave desarrollan uno o más de los siguientes síntomas: anemia grave, insuficiencia respiratoria en relación con acidosis metabólica, o malaria cerebral. En adultos, también es frecuente la implicación multiorgánica. En regiones endémicas de la malaria, las personas pueden desarrollar una inmunidad parcial, lo que permite que se produzcan infecciones asintomáticas.

15 El desarrollo de pruebas de diagnóstico rápido y eficiente es de gran relevancia para la salud pública. De hecho, el diagnóstico y el tratamiento tempranos de la malaria no sólo reduce la enfermedad y previene las muertes sino que también contribuye a reducir la transmisión de la malaria. Según las recomendaciones de la OMS, todos los casos de malaria sospechosa deben confirmarse usando pruebas de diagnóstico basadas en parásitos (en particular, usando una prueba de diagnóstico rápido) antes de administrar el tratamiento (véase “WHO Guidelines for the treatment of malaria”, tercera edición, publicado en abril de 2015).

20 La resistencia a las terapias contra la malaria representa un problema de salud crítico que reduce drásticamente las estrategias terapéuticas. De hecho, tal como se notifica en el sitio web de la OMS, la resistencia de *P. falciparum* a las generaciones previas de medicamentos, tales como la cloroquina y la sulfadoxina/pirimetamina (SP), se generalizó en las décadas de 1950 y 1960, socavando los esfuerzos para el control de la malaria y revirtiendo los avances en la supervivencia infantil. Por tanto, la OMS recomienda la monitorización rutinaria de la farmacoresistencia a la malaria. De hecho, un diagnóstico preciso puede evitar tratamientos inapropiados y limitar la extensión de la resistencia a medicamentos contra la malaria.

En este contexto, la estrategia técnica global de la OMS para la Malaria 2016-2030, adoptada por la Asamblea Mundial de la Salud en mayo de 2015, proporciona un marco técnico para todos los países endémicos de la malaria. Se pretende guiar y respaldar los programas regionales y nacionales en su labor hacia el control y la eliminación de la malaria. La estrategia establece metas globales ambiciosas pero alcanzables, incluyendo:

- 30 • Reducir la incidencia de casos de malaria en al menos un 90% para el 2030.
- Reducir las tasas de mortalidad por malaria en al menos un 90% para el 2030.
- Eliminar la malaria en al menos 35 países para el 2030.
- Prevenir un resurgimiento de la malaria en todos los países libres de malaria.

35 Esta estrategia fue el resultado de un extenso procedimiento consultivo que duró 2 años e implicó la participación de más de 400 expertos técnicos de los 70 estados miembros. Se basa en 3 ejes clave:

- garantizar el acceso universal a la prevención, el diagnóstico y el tratamiento de la malaria;
- acelerar los esfuerzos hacia la eliminación y el logro de un estado libre de malaria; y
- transformar la vigilancia de la malaria en una intervención básica.

40 El tratamiento contra *Plasmodium* incluye aril-aminoalcoholes tales como quinina o derivados de quinina tales como cloroquina, amodiaquina, mefloquina, piperaquina, lumefantrina, primaquina; análogo lipófilo de hidroxinaftoquinona, tal como atovaquona; fármacos antifolato, tales como los fármacos sulfatados sulfadoxina, dapsona y pirimetamina; proguanilo; la combinación de atovaquona/proguanilo; fármacos artemisininas; y combinaciones de los mismos.

45 Las secuencias diana que son de diagnóstico para la presencia de un patógeno transmitido por mosquitos incluyen una secuencia que diagnostica la presencia de *Plasmodium*, en particular las especies de *Plasmodium* que afectan a seres humanos tales como *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* y *Plasmodium knowlesi*, incluyendo las secuencias de los genomas de los mismos.

Las secuencias diana que son de diagnóstico para la monitorización de la farmacoresistencia al tratamiento contra *Plasmodium*, en particular las especies de *Plasmodium* que afectan a seres humanos tales como *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* y *Plasmodium knowlesi*.

50 Las secuencias diana adicionales incluyen secuencias que incluyen moléculas diana/moléculas de ácido nucleico que codifican para proteínas implicadas en el proceso biológico esencial para el parásito *Plasmodium* y en particular

proteínas transportadoras, tales como la proteína de la familia de transportadores de fármacos/metabolitos, la proteína del casete de unión a ATP (ABC) implicada en la translocación del sustrato, tal como la subfamilia C del transportador ABC o el intercambiador de Na^+/H^+ , glutatión S-transferasa de membrana; proteínas implicadas en la ruta del folato, tales como la dihidropteroato sintasa, la actividad dihidrofolato reductasa o la dihidrofolato reductasa-timidilato sintasa; y proteínas implicadas en la translocación de protones a través de la membrana mitocondrial interna y en particular el complejo de citocromo b. Una diana adicional también puede incluir el/los gen(es) que codifica(n) para la hemo polimerasa.

Las secuencias diana adicionales que incluyen moléculas diana/moléculas de ácido nucleico que codifican para proteínas implicadas en el proceso biológico esencial pueden seleccionarse del gen transportador de resistencia a la cloroquina de *P. falciparum* (*pfcr*), el gen transportador de multirresistencia a fármacos 1 de *P. falciparum* (*pfmdr1*), el gen de la proteína asociada a multirresistencia a fármacos de *P. falciparum* (*Pfmrp*), el gen intercambiador de Na^+/H^+ de *P. falciparum* (*pfhhe*), el gen que codifica para la proteína 1 exportada de *P. falciparum*, la ATPasa 6 de transporte de Ca^{2+} de *P. falciparum* (*pfatp6*); la dihidropteroato sintasa de *P. falciparum* (*pfdhps*), los genes de actividad dihidrofolato reductasa (*pfdhpr*) y de dihidrofolato reductasa-timidilato sintasa (*pfdhfr*), el gen del citocromo b, la GTP ciclohrolasa y el gen *Kelch13* (*K13*), así como sus genes heterólogos funcionales en otras especies de *Plasmodium*.

Se han identificado varias mutaciones, en particular mutaciones puntuales únicas, en las proteínas que son las dianas de los tratamientos actuales y que están asociadas con fenotipos de resistencia específicos. Por consiguiente, la invención permite la detección de diversos fenotipos de resistencia de los parásitos transmitidos por mosquitos, tales como *Plasmodium*.

La invención permite detectar una o más mutaciones y en particular uno o más polimorfismos de un solo nucleótido en ácidos nucleicos/moléculas diana. Por consiguiente, una cualquiera de las mutaciones a continuación, o una combinación de las mismas, puede usarse como marcador de resistencia a fármacos y puede detectarse según la invención.

Las mutaciones puntuales únicas en *P. falciparum* K13 incluyen las siguientes mutaciones puntuales únicas en las posiciones 252, 441, 446, 449, 458, 493, 539, 543, 553, 561, 568, 574, 578, 580, 675, 476, 469, 481, 522, 537, 538, 579, 584 y 719 y en particular las mutaciones E252Q, P441L, F446I, G449A, N458Y, Y493H, R539T, I543T, P553L, R561H, V568G, P574L, A578S, C580Y, A675V, M476I, C469Y, A481V, S522C, N537I, N537D, G538V, M579I, D584V y H719N. Estas mutaciones están asociadas generalmente con fenotipos de farmacoresistencia a artemisininas (resistencia a la artemisinina y a la a la terapia de combinación basada en artemisinina, abril de 2016 WHO/HTM/GMP/2016.5).

En la dihidrofolato reductasa de *P. falciparum* (DHFR) (*PfDHFR*-TS, PFD0830w), los polimorfismos importantes incluyen mutaciones en las posiciones 108, 51, 59 y 164, en particular 108D, 164L, 51I y 59R que modulan la resistencia a la pirimetamina. Otros polimorfismos también incluyen 437G, 581G, 540E, 436A y 613S que están asociados con la resistencia a la sulfadoxina. Las mutaciones observadas adicionales incluyen Ser108Asn, Asn51Ile, Cys59Arg, Ile164Leu, Cys50Arg, Ile164Leu, Asn188Lys, Ser189Arg y Val213Ala, Ser108Thr y Ala16Val. Las mutaciones Ser108Asn, Asn51Ile, Cys59Arg, Ile164Leu, Cys50Arg, Ile164Leu están asociadas, en particular, con las resistencias a la terapia basada en pirimetamina y/o a la terapia de combinación de cloroguanina-dapsona. La resistencia al cicloguanilo parece estar asociada con las mutaciones dobles Ser108Thr y Ala16Val. La amplificación de dhfr también puede ser de alta relevancia para la resistencia a la terapia, en particular la resistencia a la pirimetamina.

En la dihidropteroato sintasa de *P. falciparum* (DHPS) (*PfDHPS*, PF08_0095), los polimorfismos importantes incluyen mutaciones en las posiciones 436, 437, 581 y 613, Ser436Ala/Phē, Ala437Gly, Lys540Glu, Ala581Gly y Ala613Thr/Ser. El polimorfismo en la posición 581 y/o 613 también se ha asociado con la resistencia a las terapias basadas en sulfadoxina-pirimetamina.

En el transportador de resistencia a la cloroquina de *P. falciparum* (*PfCRT*), el polimorfismo en la posición 76, en particular la mutación Lys76Thr, está asociado con la resistencia a la cloroquina. Los polimorfismos adicionales incluyen Cys72Ser, Met74Ile, Asn75Glu, Ala220Ser, Gln271Glu, Asn326Ser, Ile356Thr y Arg371Ile que pueden estar asociados con la resistencia a la cloroquina. PfCRT también está fosforilado en los residuos S33, S411 y T416, que pueden regular la actividad de transporte o la especificidad de la proteína.

En el transportador, de multirresistencia a fármacos de *P. falciparum* (*PfMDR1*) (PFE1150w), se ha identificado y notificado que los polimorfismos en las posiciones 86, 184, 1034, 1042, en particular Asn86Tyr, Tyr184Phe, Ser1034Cys, Asn1042Asp y Asp1246Tyr influyen en las susceptibilidades a lumefantrina, artemisinina, quinina, mefloquina, halofantrina y cloroquina. Además, la amplificación de PfNMR1 está asociada con una susceptibilidad reducida a lumefantrina, artemisinina, quinina, mefloquina y halofantrina, y la desamplificación de PfMDR1 conduce a un aumento en la resistencia a la cloroquina. También puede detectarse la amplificación de pfmdr1. El estado de fosforilación de PfMDR1 también es de alta relevancia.

En la proteína asociada con la multirresistencia a fármacos de *P. falciparum* (*PfMRP*) (referencia de gen

PFA0590w), se han identificado los polimorfismos en las posiciones 191 y/o 437, tales como Y191H y A437S, y se han asociado con los fenotipos de resistencia a la cloroquina.

En el intercambiador de NA⁺/H⁺ de *P. falciparum* (P₁NHE) (ref. PF13_0019), la repetición aumentada de DNNND en el microsatélite ms4670 puede ser un marcador para la resistencia a la quinina.

5 Las mutaciones que alteran el sitio de unión a ubiquinol de la proteína citocromo b codificada por el gen citocromo b (cytb, mal_mito_3) están asociadas con la resistencia a la atovaquona. Las mutaciones en las posiciones 26, 268, 276, 133 y 280, y en particular Tyr26Asn, Tyr268Ser, M133I y G280D, pueden estar asociadas con la resistencia a la atovaquona.

10 Por ejemplo, en *P. vivax*, las mutaciones en PvMDR1, el homólogo de PfMDR1, se han asociado con la resistencia a la cloroquina, en particular el polimorfismo en la posición 976 tal como la mutación Y976F.

Las mutaciones anteriores se definen en términos de secuencias de proteína. Sin embargo, el experto puede determinar las mutaciones correspondientes, incluyendo los SNP, que van a identificarse como secuencia diana de ácido nucleico.

15 En la técnica se conocen otros marcadores de multiresistencia a fármacos identificados, por ejemplo, tal como se describe en "Susceptibility of Plasmodium falciparum to antimalarial drugs (1996-2004)"; OMS; Artemisinin and artemisinin-based combination therapy resistance (abril de 2016, WHO/HTM/GMP/2016.5); "Drug-resistant malaria: molecular mechanisms and implications for public health" FEBS Lett. 6 de junio de 2011;585(11):1551-62. doi:10.1016/j.febslet.2011.04.042. Epub del 23 de abril de 2011. Review. PubMed PMID: 21530510.

20 En cuanto a polipéptidos que pueden detectarse según la presente invención, pueden usarse como dianas los productos génicos de todos los genes mencionados en el presente documento. De manera correspondiente, se contempla que tales polipéptidos pueden usarse para la identificación de especies, el tipado y/o la detección de farmacoresistencia.

25 En determinadas realizaciones de ejemplo, los sistemas, dispositivos y métodos dados a conocer en el presente documento se refieren a detectar la presencia de uno o más parásitos transmitidos por mosquitos en una muestra, tal como una muestra biológica obtenida de un sujeto. En determinadas realizaciones de ejemplo, el parásito puede seleccionarse de las especies *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* o *Plasmodium knowlesi*. Por consiguiente, los métodos dados a conocer en el presente documento pueden adaptarse para su uso en otros métodos o (en combinación) con otros métodos que requieren la rápida identificación de la especie de parásito, la monitorización de la presencia de parásitos y formas de parásitos (por ejemplo, 30 correspondientes a diversas fases de infección y ciclo de vida del parásito, tal como el ciclo exoeritrocítico, el ciclo eritrocítico, el ciclo esporogónico; las formas de parásitos incluyen merozoitos, esporozoitos, esquizontes, gametocitos); la detección de determinados fenotipos (por ejemplo, farmacoresistencia del patógeno), la monitorización de la evolución y/o el brote epidémico de la enfermedad, y el examen del tratamiento (fármaco). Además, en el caso de la malaria, puede transcurrir un tiempo prolongado tras la picadura infecciosa, concretamente 35 un periodo de incubación prolongado, durante el cual el paciente no muestra síntomas. De manera similar, los tratamientos profilácticos pueden retrasar la aparición de síntomas, y también pueden observarse periodos asintomáticos prolongados antes de una recaída. Tales retrasos pueden provocar fácilmente un diagnóstico erróneo o un diagnóstico retrasado y, por tanto, alterar la eficacia del tratamiento.

40 Debido a las capacidades de diagnóstico rápido y sensible de las realizaciones dadas a conocer en el presente documento, la detección del tipo de parásito, hasta una diferencia de un solo nucleótido, y a la capacidad de implementarse como dispositivo de POC, las realizaciones dadas a conocer en el presente documento pueden usarse para guiar regímenes terapéuticos, tales como la selección del transcurso de tratamiento apropiado. Las realizaciones dadas a conocer en el presente documento también pueden usarse para examinar muestras medioambientales (población de mosquitos, etc.) para determinar la presencia y el tipado del parásito. Las 45 realizaciones también pueden modificarse para detectar parásitos transmitidos por mosquitos y otros patógenos transmitidos por mosquitos simultáneamente. En algunos casos, la malaria y otros patógenos transmitidos por mosquitos pueden presentarse inicialmente con síntomas similares. Por tanto, la capacidad para distinguir rápidamente el tipo de infección puede guiar las decisiones importantes del tratamiento. Otros patógenos transmitidos por mosquitos que pueden detectarse junto con la malaria incluyen dengue, virus del Nilo Occidental, 50 chikungunya, fiebre amarilla, filariasis, encefalitis japonesa, encefalitis de Saint Louis, encefalitis equina occidental, encefalitis equina oriental, encefalitis equina venezolana, encefalitis de La Crosse y Zika.

En determinadas realizaciones de ejemplo, los dispositivos, sistemas y métodos dados a conocer en el presente documento pueden usarse para distinguir múltiples especies de parásitos transmitidos por mosquitos en una muestra. En determinadas realizaciones de ejemplo, la identificación puede basarse en secuencias de ARN 55 ribosómico, incluyendo las subunidades 18S, 16S, 23S y 5S. En determinadas realizaciones de ejemplo, la identificación puede basarse en secuencias de genes que están presentes en múltiples copias en el genoma, tales como genes mitocondriales como CYTB. En determinadas realizaciones de ejemplo, la identificación puede basarse en secuencias de genes que están altamente expresadas y/o altamente conservadas, tales como GAPDH, histona

H2B, enolasa o LDH. Se dan a conocer métodos para identificar secuencias de ARNr relevantes en la publicación de solicitud de patente estadounidense n.º 2017/0029872. En determinadas realizaciones de ejemplo, puede diseñarse un conjunto de ARN guía para distinguir cada especie mediante una región variable que es única para cada especie o cepa. Los ARN guía también pueden diseñarse para seleccionar como diana genes de ARN que distinguen microbios a nivel de género, familia, orden, clase, filo, reino, o una combinación de los mismos. En determinadas realizaciones de ejemplo en las que se usa amplificación, puede diseñarse un conjunto de cebadores de amplificación para flanquear regiones constantes de la secuencia de ARN ribosómico y un ARN guía diseñado para distinguir cada especie mediante una región interna variable. En determinadas realizaciones de ejemplo, los cebadores y los ARN guía pueden diseñarse para regiones conservadas y variables en la subunidad 16S respetuosamente. También pueden usarse genes o regiones genómicas que varían de manera única entre especies o un subconjunto de especies tales como la familia de genes RecA, la subunidad β de ARN polimerasa. Se comentan otros marcadores filogenéticos adecuados y métodos para identificar los mismos, por ejemplo, en Wu *et al.* arXiv: 1307.8690 [q-bio.GN].

En determinadas realizaciones de ejemplo, la identificación de especies puede realizarse basándose en los genes que están presentes en múltiples copias en el genoma, tales como genes mitocondriales como CYTB. En determinadas realizaciones de ejemplo, la identificación de especies puede realizarse basándose en genes altamente expresados y/o altamente conservados tales como GAPDH, histona H2B, enolasa o LDH.

En determinadas realizaciones de ejemplo, un método o diagnóstico se diseña para examinar los parásitos transmitidos por mosquitos a lo largo de múltiples niveles filogenéticos y/o fenotípicos al mismo tiempo. Por ejemplo, el método o diagnóstico puede comprender el uso de múltiples sistemas CRISPR con diferentes ARN guías. Puede distinguirse un primer conjunto de ARN guías, por ejemplo, entre *Plasmodium falciparum* o *Plasmodium vivax*. Estas clases generales pueden incluso subdividirse adicionalmente. Por ejemplo, los ARN guías pueden diseñarse y usarse en el método o diagnóstico que distingue cepas farmacorresistentes, en general o con respecto a un fármaco específico o una combinación de fármacos. Puede diseñarse un segundo conjunto de ARN guías para distinguir microbios a nivel de especie. Por tanto, puede producirse una matriz que identifica todas las especies o subespecies de parásitos transmitidos por mosquitos, divididas adicionalmente según la farmacorresistencia. Lo anterior es a modo de ejemplo únicamente. También se contemplan otros medios para clasificar otros tipos de parásitos transmitidos por mosquitos y seguirían la estructura general descrita anteriormente.

En determinadas realizaciones de ejemplo, los dispositivos, sistemas y métodos dados a conocer en el presente documento pueden usarse para examinar en busca de genes de parásitos transmitidos por mosquitos de interés, por ejemplo, genes de farmacorresistencia. Pueden diseñarse ARN guías para distinguir entre genes conocidos de interés. Luego pueden examinarse las muestras, incluyendo muestras clínicas, usando las realizaciones dadas a conocer en el presente documento para la detección de uno o más de tales genes. La capacidad para examinar en busca de farmacorresistencia en POC tendría un tremendo beneficio en la seleccionar de un régimen de tratamiento apropiado. En determinadas realizaciones de ejemplo, los genes de farmacorresistencia son genes que codifican para proteínas tales como proteínas transportadoras, tales como la proteína de la familia de transportadores de fármacos/metabolitos, la proteína del casete de unión a ATP (ABC) implicada en la translocación del sustrato, tal como la subfamilia C del transportador ABC o el intercambiador de Na^+/H^+ ; proteínas implicadas en la ruta del folato, tales como la dihidropteroato sintasa, la actividad dihidrofolato reductasa o la dihidrofolato reductasa-timidilato sintasa; y proteínas implicadas en la translocación de protones a través de la membrana mitocondrial interna y en particular el complejo de citocromo b. Las dianas adicionales también pueden incluir el/los gen(es) que codifica(n) para la hemo polimerasa. En determinadas realizaciones de ejemplo, los genes de farmacorresistencia se seleccionan del gen transportador de resistencia a la cloroquina de *P. falciparum* (*pfcr1*), el gen transportador de multiresistencia a fármacos 1 de *P. falciparum* (*pfmrd1*), el gen de la proteína asociada a multiresistencia a fármacos de *P. falciparum* (*Pfmrp*), el gen intercambiador de Na^+/H^+ de *P. falciparum* (*pfmhe*), la ATPasa 6 de transporte de Ca^{2+} de *P. falciparum* (*pfatp6*), la dihidropteroato sintasa de *P. falciparum* (*pfdhps*), los genes de actividad dihidrofolato reductasa (*pfdhpr*) y de dihidrofolato reductasa-timidilato sintasa (*pfdhfr*), el gen del citocromo b, la GTP ciclohrolasa y el gen *Kelch13* (*K13*), así como sus genes heterólogos funcionales en otras especies de *Plasmodium*. En la técnica se conocen otros marcadores de farmacorresistencia identificados, por ejemplo, tal como se describe en "Susceptibility of Plasmodium falciparum to antimalarial drugs (1996-2004)"; OMS; Artemisinin and artemisinin-based combination therapy resistance (abril de 2016 WHO/HTM/GMP/2016,5); "Drug-resistant malaria: molecular mechanisms and implications for public health" FEBS Lett. 6 de junio de 2011;585(11):1551-62. doi:10.1016/j.febslet.2011.04.042. Epub del 23 de abril de 2011. Review. PubMed PMID: 21530510.

En algunas realizaciones, puede usarse un sistema CRISPR, un sistema de detección o métodos de uso de los mismos tal como se describe en el presente documento para determinar la evolución del brote epidémico de parásitos transmitidos por mosquitos. El método puede comprender detectar una o más secuencias diana a partir de una pluralidad de muestras de uno o más sujetos, en el que la secuencia diana es una secuencia de un parásito transmitido por mosquitos que se propaga o provoca los brotes epidémicos. Un método de este tipo puede comprender además determinar un patrón de la transmisión de parásitos transmitidos por mosquitos, o un mecanismo implicado en un brote epidémico de enfermedad provocado por un parásito transmitido por mosquitos. Las muestras pueden derivarse de uno o más seres humanos y/o derivarse de uno o más mosquitos.

El patrón de la transmisión de patógenos puede comprender transmisiones nuevas continuas a partir de un

reservorio natural del parásito transmitido por mosquitos u otras transmisiones (por ejemplo, a través de mosquitos) tras una única transmisión a partir del reservorio natural, o una mezcla de ambas. En una realización, la secuencia diana es preferiblemente una secuencia dentro del genoma del parásito transmitido por mosquitos o fragmentos de los mismos. En una realización, el patrón de la transmisión de parásitos transmitidos por mosquitos es el patrón temprano de la transmisión de parásitos transmitidos por mosquitos, es decir, al comienzo del brote epidémico de parásitos transmitidos por mosquitos. La determinación del patrón de la transmisión de parásitos transmitidos por mosquitos al comienzo del brote epidémico aumenta la probabilidad de detener el brote epidémico lo más pronto posible, reduciendo de ese modo la posibilidad de diseminación local e internacional.

La determinación del patrón de la transmisión de parásitos transmitidos por mosquitos puede comprender detectar una secuencia del parásito transmitido por mosquitos según los métodos descritos en el presente documento. La determinación del patrón de la transmisión de patógenos puede comprender además detectar variaciones dentro del huésped compartidas de la secuencia del parásito transmitido por mosquitos entre los sujetos y determinar si las variaciones dentro del huésped compartidas muestran patrones temporales. Los patrones en una variación dentro del huésped y entre huéspedes observada proporcionan información importante sobre la transmisión y la epidemiología (Gire, *et al.*, 2014).

Además de otros tipos de muestras dados a conocer en el presente documento, la muestra puede derivarse de uno o más mosquitos, por ejemplo, la muestra puede comprender saliva de mosquito.

DETECCIÓN DE BIOMARCADORES

En determinadas realizaciones de ejemplo, los sistemas, dispositivos y métodos dados a conocer en el presente documento pueden usarse para la detección de biomarcadores. Por ejemplo, los sistemas, dispositivos y métodos dados a conocer en el presente documento pueden usarse para la detección y/o el genotipado de SNP. Los sistemas, dispositivos y métodos dados a conocer en el presente documento también pueden usarse para la detección de cualquier estado patológico o trastorno caracterizado por una expresión génica aberrante. La expresión génica aberrante incluye la aberración en el gen expresado, en la ubicación de la expresión y en el nivel de expresión. Pueden detectarse múltiples transcritos o marcadores de proteínas relacionados con trastornos cardiovasculares, trastornos inmunitarios y cáncer, entre otras enfermedades. En determinadas realizaciones de ejemplo, las realizaciones dadas a conocer en el presente documento pueden usarse para la detección de ADN libre de células de enfermedades que implican lisis, tales como fibrosis hepática y enfermedad pulmonar restrictiva/obstruktiva. En determinadas realizaciones de ejemplo, las realizaciones pueden utilizarse para una detección más rápida y más portátil para pruebas prenatales de ADN libre de células. Las realizaciones dadas a conocer en el presente documento pueden usarse para el examen de paneles de diferentes SNP asociados con, entre otros, la salud cardiovascular, las firmas lipídicas/metabólicas, la identificación de la etnia, la coincidencia de paternidad, la identificación humana (por ejemplo, coincidencia de un sospechoso con una base de datos criminal de firmas de SNP). Las realizaciones dadas a conocer en el presente documento también pueden usarse para la detección de ADN libre de células de mutaciones relacionadas con y liberadas a partir de tumores cancerosos. Las realizaciones dadas a conocer en el presente documento también pueden usarse para la detección de la calidad de la carne, por ejemplo, proporcionando una detección rápida de diferentes fuentes animales en un producto cárnico dado. Las realizaciones dadas a conocer en el presente documento también pueden usarse para la detección de los GMO o de la edición génica relacionada con el ADN. Tal como se describe en otras partes en el presente documento, pueden distinguirse genotipos/alelos o biomarcadores estrechamente relacionados (por ejemplo, que tienen sólo una diferencia de un solo nucleótido en una secuencia diana dada) mediante la introducción de un apareamiento erróneo sintético en el ARNg.

En un aspecto, la invención se refiere a un método para detectar ácidos nucleicos diana en muestras, que comprende:

- a. distribuir una muestra o un conjunto de muestras en uno o más volúmenes discretos individuales, comprendiendo los volúmenes discretos individuales un sistema CRISPR según la invención tal como se describe en el presente documento;
- b. incubar la muestra o el conjunto de muestras en condiciones suficientes para permitir la unión del uno o más ARN guías a una o más moléculas diana;
- c. activar la proteína efectora CRISPR mediante la unión del uno o más ARN guías a la una o más moléculas diana, en el que activar la proteína efectora CRISPR da como resultado la modificación del constructo de enmascaramiento basado en ARN de modo que se genera una señal positiva detectable; y
- d. detectar la señal positiva detectable, en el que la detección de la señal positiva detectable indica la presencia de una o más moléculas diana en la muestra.

Tipos de muestras de biomarcadores

La sensibilidad de los ensayos descritos en el presente documento es muy adecuada para la detección de ácidos nucleicos diana en una amplia variedad de tipos de muestras biológicas, incluyendo tipos de muestras en los que el

ácido nucleico diana está diluido o para los que el material de muestra es limitado. El examen de biomarcadores puede llevarse a cabo sobre varios tipos de muestras incluyendo, pero sin limitarse a, saliva, orina, sangre, heces, esputo y líquido cefalorraquídeo. Las realizaciones dadas a conocer en el presente documento también pueden usarse para detectar la regulación por incremento y/o por disminución de genes. Por ejemplo, una muestra puede diluirse en serie de tal manera que sólo permanezcan los genes sobreexpresados por encima del umbral de límite de detección del ensayo.

En determinadas realizaciones, la presente invención proporciona etapas de obtención de una muestra de líquido biológico (por ejemplo, orina, plasma o suero sanguíneo, esputo, líquido cefalorraquídeo) y extracción del ADN. La secuencia de nucleótidos mutante que va a detectarse puede ser una fracción de una molécula más grande o puede estar presente inicialmente como molécula discreta.

En determinadas realizaciones, el ADN se aísla a partir de plasma/suero de un paciente con cáncer. A modo de comparación, pueden aislarse muestras de ADN aisladas a partir de tejido neoplásico y una segunda muestra a partir de tejido no neoplásico del mismo paciente (control), por ejemplo, linfocitos. El tejido no neoplásico puede ser del mismo tipo que el tejido neoplásico o de una fuente orgánica diferente. En determinadas realizaciones, se extraen muestras de sangre e inmediatamente se separa el plasma de las células sanguíneas mediante centrifugación. El suero puede filtrarse y almacenarse congelado hasta la extracción del ADN.

En determinadas realizaciones de ejemplo, los ácidos nucleicos diana se detectan directamente a partir de una muestra en bruto o sin procesar, tal como sangre, suero, saliva, líquido cefalorraquídeo, esputo u orina. En determinadas realizaciones de ejemplo, el ácido nucleico diana es ADN libre de células.

Células tumorales circulantes

En una realización, pueden someterse a ensayo células circulantes (por ejemplo, células tumorales circulantes (CTC)) con la presente invención. Puede realizarse el aislamiento de células tumorales circulantes (CTC) para su uso en cualquiera de los métodos descritos en el presente documento. Se han descrito tecnologías a modo de ejemplo que logran una detección y una captura específicas y sensibles de células circulantes que pueden usarse en la presente invención (Mostert B, *et al.*, Circulating tumor cells (CTCs): detection methods and their clinical relevance in breast cancer. *Cancer Treat Rev.* 2009;35:463-474; y Talasz AH, *et al.*, Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device. *Proc Natl Acad Sci USA.* 2009; 106:3970-3975). Puede hallarse tan solo una CTC en el fondo de 105-106 células mononucleares de sangre periférica (Ross A peripheral, *et al.*, Detection and viability of tumor cells in peripheral blood stem cell collections from breast cancer patients using immunocytochemical and clonogenic assay techniques. *Blood.* 1993;82:2605-2610). La plataforma CellSearch® usa perlas inmunomagnéticas recubiertas con anticuerpos contra la molécula de adhesión de células epiteliales (EpCAM) para enriquecer las células epiteliales que expresan EPCAM, seguido de inmunotinción para confirmar la presencia de tinción de citoqueratina y la ausencia del marcador leucocitario CD45 para confirmar que las células capturadas son células tumorales epiteliales (Momburg F, *et al.*, Immunohistochemical study of the expresión of a Mr 34,000 human epithelium-specific surface glycoprotein in normal and malignant tissues. *Cancer Res.* 1987;47:2883-2891; y Allard WJ, *et al.*, Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin Cancer Res.* 2004; 10:6897-6904). Se ha demostrado prospectivamente que el número de células capturadas tiene significación pronóstica para pacientes con cáncer de mama, colorrectal y de próstata con enfermedad avanzada (Cohen SJ, *et al.*, *J Clin Oncol.* 2008;26:3213-3221; Cristofanilli M, *et al.* *N Engl J Med.* 2004;351:781-791; Cristofanilli M, *et al.*, *J Clin Oncol.* 2005;23: 1420-1430; y de Bono JS, *et al.* *Clin Cancer Res.* 2008; 14:6302-6309).

La presente invención también proporciona el aislamiento de CTC con la tecnología CTC-Chip. CTC-Chip es un dispositivo de captura de CTC basado en microfluidos en el que la sangre fluye a través de una cámara que contiene miles de micropostes recubiertos con anticuerpos anti-EpCAM a los que se unen las CTC (Nagrath S, *et al.* Isolation of rare circulating tumor cells in cancer patients by microchip technology. *Nature.* 2007;450: 1235-1239). CTC-Chip proporciona un aumento significativo en los recuentos de CTC y en la pureza en comparación con el sistema CellSearch® (Maheswaran S, *et al.* Detection of mutations in EGFR in circulating lung-cancer cells, *N Engl J Med.* 2008;359:366-377), pudiéndose usar ambas plataformas para el posterior análisis molecular.

Cromatina libre de células

En determinadas realizaciones, se aíslan y analizan fragmentos de cromatina libre de células según la presente invención. Los nucleosomas pueden detectarse en el suero de individuos sanos (Stroun *et al.*, *Annals of the New York Academy of Sciences* 906: 161-168 (2000)) así como de individuos aquejados de un estado patológico. Además, la concentración sérica de nucleosomas es considerablemente mayor en pacientes que padecen enfermedades benignas y malignas, tales como cáncer y enfermedad autoinmunitaria (Holdenrieder *et al.* (2001) *Int J Cancer* 95, 1 14-120, Trejo-Becerril *et al.* (2003) *Int J Cancer* 104, 663-668; Kuroi *et al.* 1999 *Breast Cancer* 6, 361-364; Kuroi *et al.* (2001) *Int J Oncology* 19, 143-148; Amoura *et al.* (1997) *Arth Rheum* 40, 2217-2225; Williams *et al.* (2001) *J Rheumatol* 28, 81-94). Sin estar restringidos a una teoría, la alta concentración de nucleosomas en pacientes que portan tumores deriva de la apoptosis, que se produce de manera espontánea en tumores en proliferación. Los nucleosomas que circulan en la sangre contienen histonas modificadas de manera única. Por

ejemplo, la publicación de patente estadounidense n.º 2005/0069931 (31 de marzo de 2005) se refiere al uso de anticuerpos dirigidos contra modificaciones N-terminales de histonas específicas como indicadores de diagnóstico de la enfermedad, empleando tales anticuerpos específicos de histonas para aislar los nucleosomas a partir de una muestra de sangre o suero de un paciente para facilitar la purificación y el análisis del ADN acompañante con propósitos de diagnóstico/examen. Por consiguiente, la presente invención puede usar ADN unido a cromatina para detectar y monitorizar, por ejemplo, mutaciones tumorales. La identificación del ADN asociado con histonas modificadas puede servir como marcadores de diagnóstico de la enfermedad y de defectos congénitos.

Por tanto, en otra realización, los fragmentos de cromatina aislados se derivan de cromatina circulante, preferiblemente de mono y oligonucleosomas circulantes. Los fragmentos de cromatina aislados pueden derivarse de una muestra biológica. La muestra biológica puede ser de un sujeto o un paciente que lo necesita. La muestra biológica puede ser suero, plasma, linfa, sangre, fracciones sanguíneas, orina, líquido sinovial, líquido espinal, saliva, células tumorales circulantes o mucosa.

ADN libre de células (ADNlc)

En determinadas realizaciones, la presente invención puede usarse para detectar ADN libre de células (ADNlc). El ADN libre de células en plasma o suero puede usarse como herramienta de diagnóstico no invasiva. Por ejemplo, se ha estudiado y optimizado el ADN fetal libre de células para someter a prueba factores de RhD compatibles, determinar el sexo para trastornos genéticos ligados al cromosoma X, someter a prueba trastornos genéticos individuales, identificar la preeclampsia. Por ejemplo, secuenciar la fracción celular fetal del ADNlc en plasma materno es un enfoque fiable para detectar cambios en el número de copias asociados con la aneuploidía cromosómica fetal. Como otro ejemplo, se ha usado ADNlc aislado a partir de pacientes con cáncer para detectar mutaciones en genes clave relevantes para las decisiones de tratamiento.

En determinadas realizaciones de ejemplo, la presente divulgación proporciona la detección directa de ADNlc a partir de una muestra del paciente. En otra realización de ejemplo determinada, la presente divulgación proporciona el enriquecimiento de ADNlc usando las realizaciones de enriquecimiento dadas a conocer anteriormente y antes de detectar el ADNlc diana.

Exosomas

En una realización, pueden someterse a ensayo exosomas con la presente invención. Los exosomas son pequeñas vesículas extracelulares que se ha demostrado que contienen ARN. En la técnica se conoce el aislamiento de exosomas mediante ultracentrifugación, filtración, precipitación química, cromatografía de exclusión molecular y microfluidos. En una realización, los exosomas se purifican usando un biomarcador de exosomas. El aislamiento y la purificación de exosomas a partir de muestras biológicas pueden realizarse mediante cualquier método conocido (véase, por ejemplo, el documento WO2016172598A1).

Detección y genotipado de SNP

En determinadas realizaciones, la presente invención puede usarse para detectar la presencia de polimorfismos de un solo nucleótido (SNP) en una muestra biológica. Los SNP pueden estar relacionados con las pruebas de maternidad (por ejemplo, determinación del sexo, defectos fetales). Pueden estar relacionados con una investigación criminal. En una realización, puede identificarse a un sospechoso en una investigación criminal mediante la presente invención. Sin estar restringidos a una teoría, las evidencias forenses basadas en ácidos nucleicos pueden requerir el ensayo más sensible disponible para detectar el material genético de un sospechoso o una víctima porque las muestras sometidas a pruebas pueden ser limitativas.

En otras realizaciones, la presente invención abarca los SNP asociados con una enfermedad. Los SNP asociados con enfermedades se conocen bien en la técnica, y un experto en la técnica puede aplicar los métodos de la presente invención para diseñar ARN guías adecuados (véase, por ejemplo, www.ncbi.nlm.nih.gov/clinvar?term=human%5Borgn%5D).

En un aspecto, la invención se refiere a un método para genotipar, tal como genotipar SNP, que comprende:

a) distribuir una muestra o un conjunto de muestras en uno o más volúmenes discretos individuales, comprendiendo los volúmenes discretos individuales un sistema CRISPR según la invención tal como se describe en el presente documento;

b) incubar la muestra o el conjunto de muestras en condiciones suficientes para permitir la unión del uno o más ARN guías a una o más moléculas diana;

c) activar la proteína efectora CRISPR mediante la unión del uno o más ARN guías a la una o más moléculas diana, en el que activar la proteína efectora CRISPR da como resultado la modificación del constructo de enmascaramiento basado en ARN de modo que se genera una señal positiva detectable; y

d) detectar la señal positiva detectable, en el que la detección de la señal positiva detectable indica la presencia

de una o más moléculas diana características para un genotipo particular en la muestra.

En determinadas realizaciones, la señal detectable se compara con (por ejemplo, mediante la comparación de la intensidad de señal) una o más señales patrón, preferiblemente una señal patrón sintética, tal como se ilustra, por ejemplo, en una realización de ejemplo en la figura 60. En determinadas realizaciones, el patrón es o corresponde a un genotipo particular. En determinadas realizaciones, el patrón comprende un SNP particular u otra variación nucleotídica (individual). En determinadas realizaciones, el patrón es un patrón de genotipo (amplificado por PCR). En determinadas realizaciones, el patrón es o comprende ADN. En determinadas realizaciones, el patrón es o comprende ARN. En determinadas realizaciones, el patrón es o comprende ARN que se transcribe a partir de ADN. En determinadas realizaciones, el patrón es o comprende ADN que se somete a transcripción inversa a partir de ARN. En determinadas realizaciones, la señal detectable se compara con uno o más patrones, cada uno de los cuales corresponde a un genotipo conocido, tal como un SNP u otra variación nucleotídica (individual). En determinadas realizaciones, la señal detectable se compara con una o más señales patrón y la comparación comprende un análisis estadístico, tal como mediante análisis estadístico paramétrico o no paramétrico, tal como mediante ANOVA unifactorial o bifactorial, etc. En determinadas realizaciones, la señal detectable se compara con una o más señales patrón y, cuando la señal detectable no se desvía significativamente (de manera estadística) del patrón, se determina el genotipo como el genotipo correspondiente a dicho patrón.

En otras realizaciones, la presente invención permite el genotipado rápido para la farmacogenómica de emergencias. En una realización, puede usarse un ensayo inmediato individual para genotipar a un paciente llevado a la sala de urgencias. El paciente puede ser sospechoso de tener un coágulo sanguíneo, y es necesario que un médico de urgencias decida la dosificación de antitrombótico que va a administrar. En realizaciones a modo de ejemplo, la presente invención puede proporcionar unas pautas para la administración de antitrombóticos durante el tratamiento de infarto de miocardio o accidente cerebrovascular basándose en el genotipado de marcadores tales como VKORC1, CYP2C9 y CYP2C19. En una realización, el antitrombótico es el anticoagulante warfarina (Holford, NH (diciembre de 1986). "Clinical Pharmacokinetics and Pharmacodynamics of Warfarin Understanding the Dose-Effect Relationship". Clinical Pharmacokinetics. Springer International Publishing. 11 (6): 483-504). En la técnica se conocen los genes asociados con la coagulación sanguínea (véanse, por ejemplo, el documento US20060166239A1; Litin SC, Gastineau DA (1995) "Current concepts in anticoagulant therapy". Mayo Clin. Proc. 70 (3): 266-72; y Rusdiana *et al.*, Responsiveness to low-dose warfarin associated with genetic variants of VKORC1, CYP2C9, CYP2C19, and CYP4F2 in an Indonesian population. Eur J Clin Pharmacol. Marzo de 2013;69(3):395-405). Específicamente, en el polimorfismo de un solo nucleótido VKORC1 1639 (o 3673), el alelo G ("silvestre") habitual se reemplaza por el alelo A. Las personas con un alelo A (o el "haplotipo A") producen menos VKORC1 que los que tienen el alelo G (o el "haplotipo distinto de A"). La prevalencia de estas variantes también varía según la raza, portando el alelo A el 37% de las personas de raza blanca y el 14% de los africanos. El resultado final es un menor número de factores de coagulación y, por tanto, una menor capacidad de coagulación.

En determinadas realizaciones de ejemplo, la disponibilidad de material genético para detectar un SNP en un paciente permite detectar los SNP sin la amplificación de una muestra de ADN o ARN. En el caso del genotipado, la muestra biológica sometida a prueba se obtiene fácilmente. En determinadas realizaciones de ejemplo, puede acortarse el tiempo de incubación de la presente invención. El ensayo puede realizarse en un periodo de tiempo requerido para que se produzca una reacción enzimática. Un experto en la técnica puede realizar reacciones bioquímicas en 5 minutos (por ejemplo, ligamiento en 5 minutos). La presente invención puede usar un dispositivo de extracción de ADN automático para obtener ADN a partir de la sangre. Luego puede añadirse el ADN a una reacción que genera una molécula diana para la proteína efectora. Inmediatamente después de generar la molécula diana, puede cortarse el agente de enmascaramiento y detectarse una señal. En realizaciones a modo de ejemplo, la presente invención permite un diagnóstico rápido de POC para determinar un genotipo antes de administrar un fármaco (por ejemplo, un antitrombótico). En el caso en el que se usa una etapa de amplificación, todas las reacciones se producen en la misma reacción en un procedimiento de una sola etapa. En realizaciones preferidas, el ensayo de POC puede realizarse en menos de una hora, preferiblemente 10 minutos, 20 minutos, 30 minutos, 40 minutos o 50 minutos.

En determinadas realizaciones, los sistemas, dispositivos y métodos dados a conocer en el presente documento pueden usarse para detectar la presencia o el nivel de expresión de ARN no codificantes largos (ARNncl). La expresión de determinados ARNncl está asociada con el estado patológico y/o la farmacorresistencia. En particular, determinados ARNncl (por ejemplo, TCONS_00011252, NR_034078, TCONS_00010506, TCONS_00026344, TCONS_00015940, TCONS_00028298, TCONS_00026380, TCONS_0009861, TCONS_00026521, TCONS_00016127, NR_125939, NR_033834, TCONS_00021026, TCONS_00006579, NR_109890 y NR_026873) están asociados con la resistencia al tratamiento del cáncer, tal como la resistencia a uno o más inhibidores de BRAF (por ejemplo, vemurafenib, dabrafenib, sorafenib, GDC-0879, PLX-4720 y LGX818) para tratar el melanoma (por ejemplo, melanoma nodular, lentigo maligno, melanoma sobre lentigo maligno, melanoma lentiginoso acro, melanoma de extensión superficial, melanoma mucoso, melanoma polipoide, melanoma desmoplásico, melanoma amelanótico y melanoma de tejidos blandos). La detección de ARNncl usando las diversas realizaciones descritas en el presente documento puede facilitar el diagnóstico de la enfermedad y/o la selección de las opciones de tratamiento.

En una realización, la presente invención puede guiar terapias dirigidas con ADN o ARN (por ejemplo, CRISPR,

TALE, proteínas de los dedos de Zinc, iARN), particularmente en entornos en los que es importante la rápida administración de la terapia para los resultados del tratamiento.

Detección de LOH

Las células cancerosas experimentan una pérdida de material genético (ADN) en comparación con las células normales. Esta delección de material genético que experimentan casi todos, si no todos, los cánceres se denominan "pérdida de heterocigosidad" (LOH). La pérdida de heterocigosidad (LOH) es un acontecimiento cromosómico macroscópico que da como resultado la pérdida de todo el gen y de la región cromosómica circundante. La pérdida de heterocigosidad es una aparición habitual en el cáncer, donde puede indicar la ausencia de un gen supresor tumoral funcional en la región perdida. Sin embargo, una pérdida puede ser silenciosa debido a que todavía queda un gen funcional en el otro cromosoma del par de cromosomas. La copia restante del gen supresor tumoral puede inactivarse mediante una mutación puntual, lo que conduce a la pérdida de un gen supresor tumoral. La pérdida de material genético a partir de las células cancerosas puede dar como resultado la pérdida selectiva de uno de los dos o más alelos de un gen vital para la viabilidad celular o el crecimiento celular en un locus particular en el cromosoma.

Un "marcador de LOH" es ADN de un locus de microsatélite, una delección, alteración o amplificación que, cuando se compara con células normales, está asociada con el cáncer u otras enfermedades. Con frecuencia, un marcador de LOH se asocia con la pérdida de un gen supresor tumoral u otro gen habitualmente relacionado con el tumor.

El término "microsatélites" se refiere a secuencias repetitivas cortas de ADN que se distribuyen ampliamente en el genoma humano. Un microsatélite es un tramo de motivos de ADN repetidos en tándem (es decir, adyacentes) que varían en longitud desde dos hasta cinco nucleótidos y que normalmente se repiten 5-50 veces. Por ejemplo, la secuencia TATATATATA (SEQ. ID. NO. 418) es un microsatélite dinucleotídico y GTCGTCGTCGTCGTC (SEQ. ID. NO. 419) es un microsatélite trinucleotídico (siendo A adenina, G guanina, C citosina y T timina). Se ha demostrado que las alteraciones somáticas en la longitud de repetición de tales microsatélites representan un rasgo característico de tumores. Pueden diseñarse ARN guías para detectar tales microsatélites. Además, la presente invención puede usarse para detectar alteraciones en la longitud de repetición, así como amplificaciones y delecciones basadas en la cuantificación de la señal detectable. Determinados microsatélites se ubican en las regiones intrónicas o flanqueantes reguladoras de los genes, o directamente en los codones de los genes. Las mutaciones de microsatélites en tales casos conducen a enfermedades y cambios fenotípicos, especialmente en enfermedades por expansión de tripletes tales como el síndrome del cromosoma X frágil y la enfermedad de Huntington.

Se ha notificado la pérdida de heterocigosidad (LOH) frecuente en regiones cromosómicas específicas en muchas clases de neoplasias malignas. Las pérdidas alélicas en regiones cromosómicas específicas son las alteraciones genéticas más habituales observadas en una variedad de neoplasias malignas, por lo que se ha aplicado el análisis de microsatélites para detectar el ADN de células cancerosas en muestras de líquidos corporales, tales como esputo para el cáncer de pulmón y orina para el cáncer de vejiga (Rouleau, *et al.* Nature 363, 515-521 (1993); y Latif, *et al.* Science 260, 1317-1320 (1993)). Además, se ha establecido que están presentes concentraciones marcadamente aumentadas de ADN soluble en plasma de individuos con cáncer y algunas otras enfermedades, lo que indica que puede usarse suero o plasma libre de células para detectar ADN canceroso con anomalías en los microsatélites. (Kamp, *et al.* Science 264, 436-440 (1994); y Steck, *et al.* Nat Genet. 15(4), 356-362 (1997)). Dos grupos han notificado alteraciones de microsatélites en plasma o suero de un número limitado de pacientes con cáncer microcítico de pulmón o cáncer de cabeza y cuello (Hahn, *et al.* Science 271, 350-353 (1996); y Miozzo, *et al.* Cancer Res. 56, 2285-2288 (1996)). La detección de la pérdida de heterocigosidad en tumores y suero de pacientes con melanoma también se ha demostrado previamente (véase, por ejemplo, la patente estadounidense n.º US6465177B1).

Por tanto, es ventajoso detectar los marcadores de LOH en un sujeto que padece o tiene riesgo de padecer cáncer. La presente invención puede usarse para detectar la LOH en células tumorales. En una realización, pueden usarse células tumorales circulantes como muestra biológica. En realizaciones preferidas, se usa ADN libre de células obtenido a partir de suero o plasma para detectar y/o monitorizar de manera no invasiva la LOH. En otras realizaciones, la muestra biológica puede ser cualquier muestra descrita en el presente documento (por ejemplo, una muestra de orina para el cáncer de vejiga). Sin estar restringidos a una teoría, la presente invención puede usarse para detectar marcadores de LOH con sensibilidad mejorada en comparación con cualquier otro método previo, proporcionando de ese modo la detección temprana de acontecimientos mutacionales. En una realización, la LOH se detecta en líquidos biológicos, en los que la presencia de LOH está asociada con la aparición del cáncer. El método y los sistemas descritos en el presente documento representan un avance significativo con respecto a las técnicas previas, tales como la PCR o la biopsia tisular, al proporcionar un método no invasivo, rápido y preciso para detectar la LOH de alelos específicos asociados con el cáncer. Por tanto, la presente invención proporciona métodos y sistemas que pueden usarse para examinar poblaciones de alto riesgo y monitorizar pacientes de alto riesgo que se someten a quimioprevención, quimioterapia, inmunoterapia u otros tratamientos.

Debido a que el método de la presente invención requiere sólo la extracción de ADN a partir de un líquido corporal tal como la sangre, puede realizarse en cualquier momento y de manera repetida sobre un paciente individual. La sangre puede extraerse y monitorizarse en busca de la LOH antes o después de la cirugía; antes, durante y después

del tratamiento, tal como quimioterapia, radioterapia, terapia génica o inmunoterapia; o durante el examen de seguimiento después del tratamiento para determinar la evolución, la estabilidad o la recaída de la enfermedad. Sin estar restringidos a una teoría, el método de la presente invención también puede usarse para detectar la presencia o recaída de una enfermedad subclínica con un marcador de LOH específico para ese paciente, puesto que los marcadores de LOH son específicos para el tumor de un paciente individual. El método también puede detectar si pueden estar presentes múltiples metástasis usando marcadores de LOH tumorales específicos.

Detección de modificaciones epigenéticas

Las variantes de histonas, las modificaciones de ADN y las modificaciones de histonas indicativas del cáncer o de la evolución del cáncer pueden usarse en la presente invención. Por ejemplo, la publicación de patente estadounidense 20140206014 describe que las muestras cancerosas tenían niveles elevados de nucleosoma H2AZ, macroH2A1.1, 5-metilcitosina, P-H2AX(Ser139) en comparación con sujetos sanos. La presencia de células cancerosas en un individuo puede generar un mayor nivel de nucleosomas libres de células en la sangre como resultado de la apoptosis aumentada de las células cancerosas. En una realización, puede usarse un anticuerpo dirigido contra marcas asociadas con la apoptosis, tales como H2B Ser 14(P), para identificar nucleosomas individuales que se han liberado a partir de las células neoplásicas apoptóticas. Por tanto, el ADN que surge de las células tumorales puede analizarse ventajosamente según la presente invención con alta sensibilidad y precisión.

Examen prenatal

En determinadas realizaciones, el método y los sistemas de la presente invención pueden usarse en el examen prenatal. En determinadas realizaciones, se usa ADN libre de células en un método de examen prenatal. En determinadas realizaciones, puede detectarse ADN asociado con nucleosomas u oligonucleosomas individuales con la presente invención. En realizaciones preferidas, se usa la detección de ADN asociado con nucleosomas u oligonucleosomas individuales para el examen prenatal. En determinadas realizaciones, se usan fragmentos de cromatina libres de células en un método de examen prenatal.

El diagnóstico prenatal o examen prenatal se refiere a pruebas para determinar enfermedades o estados en un feto o embrión antes de su nacimiento. El objetivo es detectar defectos congénitos tales como defectos del tubo neural, síndrome de Down, anomalías cromosómicas, trastornos genéticos y otros estados, tales como espina bífida, fisura palatina, enfermedad de Tay-Sachs, anemia drepanocítica, talasemia, fibrosis quística, distrofia muscular y síndrome del cromosoma X frágil. El examen también puede usarse para el discernimiento prenatal del sexo. Los procedimientos de pruebas habituales incluyen amniocentesis, ecografía incluyendo ecografía de translucidez nuchal, pruebas de marcadores séricos o examen genético. En algunos casos, las pruebas se administran para determinar si se abortará el feto, aunque los médicos y pacientes también lo encuentran útil para diagnosticar embarazos de alto riesgo de manera temprana, de modo que pueda programarse el parto en un hospital de atención terciaria en el que el recién nacido puede recibir una atención apropiada.

Se ha descubierto que hay células fetales que están presentes en la sangre de la madre, y que estas células presentan una posible fuente de cromosomas fetales para el diagnóstico prenatal basado en ADN. Además, el ADN fetal ADN oscila entre aproximadamente el 2-10% del ADN total en la sangre materna. Las pruebas genéticas prenatales actualmente disponibles implican habitualmente procedimientos invasivos. Por ejemplo, la biopsia de vellosidades coriónicas (CVS) realizada a una mujer embarazada alrededor de las 10-12 semanas de embarazo y la amniocentesis realizada alrededor de las 14-16 semanas contienen ambos procedimientos invasivos para obtener la muestra para las pruebas de anomalías cromosómicas en un feto. Habitualmente las células fetales obtenidas a través de estos procedimientos de muestreo se someten a prueba para determinar anomalías cromosómicas usando análisis citogenéticos o de hibridación *in situ* por fluorescencia (FISH). Se ha demostrado que existe ADN fetal libre de células en plasma y suero de mujeres embarazadas desde la sexta semana de gestación, aumentando las concentraciones durante el embarazo y alcanzando su máximo antes del parto. Debido a que estas células aparecen muy pronto en el embarazo, pueden constituir la base de una prueba precisa y no invasiva del primer trimestre. Sin estar restringidos a una teoría, la presente invención proporciona una sensibilidad sin precedentes en la detección de bajas cantidades de ADN fetal. Sin estar restringidos a una teoría, simultáneamente se recuperan grandes cantidades de ADN materno junto con el ADN fetal de interés, disminuyendo de ese modo la sensibilidad en la cuantificación del ADN fetal y la detección de mutaciones. La presente invención supera tales problemas mediante la sensibilidad inesperadamente alta del ensayo.

La clase H3 de histonas consiste en cuatro tipos de proteínas diferentes: los tipos principales, H3.1 y H3.2; el tipo de reemplazo, H3.3; y la variante específica de testículos, H3t. Aunque H3.1 y H3.2 están estrechamente relacionados, difiriendo sólo en Ser96, H3.1 difiere de H3.3 en al menos 5 posiciones de aminoácidos. Además, H3.1 está altamente enriquecido en hígado fetal, en comparación con su presencia en tejidos adultos incluyendo hígado, riñón y corazón. En tejido humano adulto, la variante H3.3 es más abundante que la variante H3.1, mientras que en el hígado fetal ocurre lo contrario. La presente invención puede usar estas diferencias para detectar nucleosomas fetales y ácido nucleico fetal en una muestra biológica materna que comprende células tanto fetales como maternas y/o ácido nucleico fetal.

En una realización, los nucleosomas fetales pueden obtenerse a partir de la sangre. En otras realizaciones, los

nucleosomas fetales se obtienen a partir de una muestra de moco cervical. En determinadas realizaciones, una muestra de moco cervical se obtiene mediante hisopado o lavado a partir de una mujer embarazada al principio del segundo trimestre o al final del primer trimestre del embarazo. La muestra puede colocarse en una incubadora para liberar el ADN atrapado en el moco. La incubadora puede ajustarse a 37°C. La muestra puede agitarse durante de aproximadamente 15 a 30 minutos. El moco puede disolverse adicionalmente con una mucinasa con el propósito de liberar el ADN. La muestra también puede someterse a condiciones tales como tratamiento químico y similares, tal como se conoce bien en la técnica, para inducir la apoptosis para liberar los nucleosomas fetales. Por tanto, una muestra de moco cervical puede tratarse con un agente que induce la apoptosis, mediante lo cual se liberan los nucleosomas fetales. Con respecto al enriquecimiento de ADN fetal circulante, se hace referencia a las publicaciones de patente estadounidense n.ºs 20070243549 y 20100240054. La presente invención es especialmente ventajosa cuando se aplican los métodos y sistemas al examen prenatal en el que sólo una pequeña fracción de nucleosomas o ADN puede ser de origen fetal.

El examen prenatal según la presente invención puede ser para una enfermedad incluyendo, pero sin limitarse a, trisomía del cromosoma 13, trisomía del cromosoma 16, trisomía del cromosoma 18, síndrome de Klinefelter (47, XXY), (47, XYY) y (47, XXX), síndrome de Turner, síndrome de Down (trisomía del cromosoma 21), fibrosis quística, enfermedad de Huntington, beta-talasemia, distrofia miotónica, anemia drepanocítica, porfiria, síndrome del cromosoma X frágil, translocación robertsoniana, síndrome de Angelman, síndrome de DiGeorge y síndrome de Wolf-Hirschhorn.

Varios aspectos adicionales de la invención se refieren a diagnosticar, pronosticar y/o tratar defectos asociados con una amplia gama de enfermedades genéticas que se describen adicionalmente en el sitio web de los Institutos Nacionales de Salud en la subsección temática Trastornos genéticos (sitio web en [health.nih.gov/topic/Genetic Disorders](http://health.nih.gov/topic/GeneticDisorders)).

Detección de cáncer y de farmacoresistencia contra el cáncer

En determinadas realizaciones, la presente invención puede usarse para detectar genes y mutaciones asociadas con el cáncer. En determinadas realizaciones, se detectan mutaciones asociadas con la resistencia. La amplificación de células tumorales resistentes o la aparición de mutaciones resistentes en poblaciones clonales de células tumorales puede surgir durante el tratamiento (véase, por ejemplo, Burger JA, *et al.*, Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat Commun.* 20 de mayo de 2016;7:11589; Landau DA, *et al.*, Mutations driving CLL and their evolution in progression and relapse. *Nature.* 22 de octubre de 2015;526(7574):525-30; Landau DA, *et al.*, Clonal evolution in hematological malignancies and therapeutic implications. *Leukemia.* Enero de 2014;28(1):34-43; y Landau DA, *et al.*, Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell.* 14 de febrero de 2013;152(4):714-26). Por consiguiente, la detección de tales mutaciones requiere ensayos altamente específicos y la monitorización requiere biopsias repetidas. Las biopsias repetidas son inconvenientes, invasivas y costosas. Las mutaciones resistentes pueden ser difíciles de detectar en una muestra de sangre u otra muestra biológica recogida de manera no invasiva (por ejemplo, sangre, saliva, orina) usando los métodos previos conocidos en la técnica. Las mutaciones resistentes pueden referirse a mutaciones asociadas con la resistencia a una quimioterapia, terapia dirigida o inmunoterapia.

En determinadas realizaciones, las mutaciones se producen en cánceres individuales que pueden usarse para detectar la evolución del cáncer. En una realización, se han caracterizado las mutaciones relacionadas con la actividad citolítica de células T contra tumores y pueden detectarse mediante la presente invención (véase, por ejemplo, Rooney *et al.*, Molecular and genetic properties of tumors associated with local immune cytolytic activity, *Cell.* 15 de enero de 2015; 160(1-2): 48-61). Pueden desarrollarse terapias personalizadas para un paciente basándose en la detección de estas mutaciones (véase, por ejemplo, el documento WO2016100975A1). En determinadas realizaciones, las mutaciones específicas del cáncer asociadas con la actividad citolítica pueden ser una mutación en un gen seleccionado del grupo que consiste en CASP8, B2M, PIK3CA, SMC1A, ARID5B, TET2, ALPK2, COL5A1, TP53, DNER, NCOR1, MORC4, CIC, IRF6, MYOCD, ANKLE1, CNKSR1, NF1, SOS1, ARID2, CUL4B, DDX3X, FUBP1, TCP11L2, HLA-A, B o C, CSNK2A1, MET, ASXL1, PD-L1, PD-L2, IDO1, IDO2, ALOX12B y ALOX15B, o un aumento en el número de copias, excluyendo acontecimiento de cromosomas completos, que afectan a cualquiera de las siguientes bandas cromosómicas: 6q16.1-q21, 6q22.31-q24.1, 6q25.1-q26, 7p11.2-q11.1, 8p23.1, 8p11.23-p11.21 (que contiene IDO1, IDO2), 9p24.2-p23 (que contiene PDL1, PDL2), 10p15.3, 10p15.1-p13, 11p14.1, 12p13.32-p13.2, 17p13.1 (que contiene ALOX12B, ALOX15B) y 22q11.1-q11.21.

En determinadas realizaciones, la presente invención se usa para detectar una mutación cancerosa (por ejemplo, mutación de resistencia) durante el transcurso de un tratamiento y después de completar el tratamiento. La sensibilidad de la presente invención puede permitir la detección no invasiva de mutaciones clonales que surgen durante el tratamiento y puede usarse para detectar una recaída en la enfermedad.

En determinadas realizaciones de ejemplo, puede usarse la detección de microARN (miARN) y/o firmas de miARN de miARN expresado de manera diferente para detectar o monitorizar la evolución de un cáncer y/o detectar la farmacoresistencia de una terapia contra el cáncer. Como ejemplo, Nadal *et al.* (*Nature Scientific Reports*, (2015) doi:10.1038/srep12464) describen firmas de ARNm que pueden usarse para detectar cáncer no microcítico de pulmón (NSCLC).

En determinadas realizaciones de ejemplo, puede usarse la presencia de mutaciones de resistencia en subpoblaciones clonales de células para determinar un régimen de tratamiento. En otras realizaciones, pueden administrarse terapias personalizadas para tratar a un paciente basándose en mutaciones tumorales habituales. En determinadas realizaciones, las mutaciones habituales surgen en respuesta al tratamiento y conducen a la farmacoresistencia. En determinadas realizaciones, la presente invención puede usarse para monitorizar a pacientes en busca de células que adquieren una mutación o amplificación de células que albergan tales mutaciones farmacoresistentes.

Con frecuencia, el tratamiento con diversos agentes quimioterápicos, particularmente con terapias dirigidas tales como inhibidores de tirosina cinasas, conduce a nuevas mutaciones en las moléculas diana que resisten la actividad del agente terapéutico. Están evaluándose múltiples estrategias para superar esta resistencia, incluyendo el desarrollo de terapias de segunda generación que no se ven afectadas por estas mutaciones y el tratamiento con múltiples agentes incluyendo aquellos que actúan en el sentido de 3' de la mutación de resistencia. En una realización a modo de ejemplo, una mutación habitual del ibrutinib, una molécula que selecciona como diana la tirosina cinasa de Bruton (BTK) y se usa para la LLC y determinados linfomas, es un cambio de cisteína a serina en la posición 481 (BTK/C481S). El erlotinib, que selecciona como diana el dominio tirosina cinasa del receptor del factor de crecimiento epidérmico (EGFR), se usa habitualmente en el tratamiento de cáncer de pulmón y de tumores resistentes que se desarrollan de manera invariable tras la terapia. Una mutación habitual hallada en clones resistentes es una mutación de treonina a metionina en la posición 790.

En la técnica se conocen mutaciones no silenciosas compartidas entre poblaciones de pacientes con cáncer y mutaciones resistentes habituales que pueden detectarse con la presente invención (véase, por ejemplo, el documento WO/2016/187508). En determinadas realizaciones, pueden inducirse mutaciones de farmacoresistencia mediante el tratamiento con ibrutinib, erlotinib, imatinib, gefitinib, crizotinib, trastuzumab, vemurafenib, RAF/MEK, terapia de bloqueo del punto de control o terapia con antiestrógenos. En determinadas realizaciones, las mutaciones específicas del cáncer están presentes en uno o más genes que codifican para una proteína seleccionada del grupo que consiste en ligando 1 de muerte programada (PD-L1), receptor de andrógenos (AR), tirosina cinasa de Bruton (BTK), receptor del factor de crecimiento epidérmico (EGFR), BCR-Abl, c-kit, PIK3CA, HER2, EML4-ALK, KRAS, ALK, ROS1, AKT1, BRAF, MEK1, MEK2, NRAS, RAC1 y ESR1.

Los puntos de control inmunitario son rutas inhibitoras que ralentizan o detienen las reacciones inmunitarias y previenen el daño tisular en exceso de una actividad incontrolada de las células inmunitarias. En determinadas realizaciones, el punto de control inmunitario seleccionado como diana es el gen 1 de muerte programada (PD-1 o CD279) (PDCD1). En otras realizaciones, el punto de control inmunitario seleccionado como diana es el antígeno asociado a linfocitos T citotóxicos (CTLA-4). En realizaciones adicionales, el punto de control inmunitario seleccionado como diana es otro miembro de la superfamilia de Ig CD28 y CTLA4 tal como BTLA, LAG3, ICOS, PDL1 o KIR. En aún más realizaciones adicionales, el punto de control inmunitario seleccionado como diana es un miembro de la superfamilia de TNFR, tal como CD40, OX40, CD137, GITR, CD27 o TIM-3.

Recientemente, se ha caracterizado la expresión génica en tumores y sus microentornos a nivel de célula individual (véase, por ejemplo, Tirosh, *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single cell RNA-seq. *Science* 352, 189-196, doi:10.1126/science.aad0501 (2016)); Tirosh *et al.*, Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature*. 10 de noviembre de 2016;539(7628):309-313. doi: 10.1038/nature20123. Epub del 2 de noviembre de 2016; y la publicación de patente internacional con número de serie WO 2017004153 A1). En determinadas realizaciones, pueden detectarse las firmas génicas usando la presente invención. En una realización, se monitorizan o detectan genes complementarios en un microentorno tumoral. En una realización, se monitorizan o detectan programas de MITF y AXL. En una realización, se detecta una firma de célula progenitora o célula madre tumoral específica. Tales firmas indican el estado de una respuesta inmunitaria y el estado de un tumor. En determinadas realizaciones, puede detectarse el estado de un tumor en términos de proliferación, resistencia al tratamiento y abundancia de células inmunitarias.

Por tanto, en determinadas realizaciones, la invención proporciona paneles de detección de cáncer rentables, rápidos y multiplexados para ADN circulante, tal como ADN tumoral, particularmente para monitorizar la recaída de la enfermedad o el desarrollo de mutaciones de resistencia habituales.

Aplicaciones de inmunoterapia

Las realizaciones dadas a conocer en el presente documento también pueden ser útiles en contextos de inmunoterapia adicionales. Por ejemplo, en algunas realizaciones, los métodos de diagnóstico, pronóstico y/o estadificación de una respuesta inmunitaria en un sujeto comprenden detectar un primer nivel de expresión, actividad y/o función de uno o más biomarcadores y comparar el nivel detectado con un nivel de control, en los que una diferencia en el nivel detectado y el nivel de control indica la presencia de una respuesta inmunitaria en el sujeto.

En determinadas realizaciones, la presente invención puede usarse para determinar una disfunción o activación de linfocitos infiltrantes tumorales (TIL). Los TIL pueden aislarse a partir de un tumor usando métodos conocidos. Los TIL pueden analizarse para determinar si deben usarse en terapias adoptivas de transferencia celular. Además,

pueden analizarse células T con receptor de antígeno quimérico (células T CAR) para determinar una firma de disfunción o activación antes de administrarlas a un sujeto. Se han descrito las firmas a modo de ejemplo para células T disfuncionales y activadas (véase, por ejemplo, Singer M, *et al.*, A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells. Cell. 8 de septiembre de 2016;166(6):1500-1511.e9. doi: 10.1016/j.cell.2016.08.052).

En algunas realizaciones, se usa C2c2 para evaluar el estado de las células inmunitarias, tales como células T (por ejemplo, células T CD8+ y/o CD4+). En particular, la activación y/o disfunción de células T puede determinarse, por ejemplo, basándose en genes o firmas génicas asociados con uno o más de los estados de las células T. De esta manera, puede usarse C2c2 para determinar la presencia de una o más subpoblaciones de células T.

En algunas realizaciones, puede usarse C2c2 en un ensayo de diagnóstico o puede usarse como método para determinar si un paciente es adecuado para que se le administre una inmunoterapia u otro tipo de terapia. Por ejemplo, la detección de firmas génicas o biomarcadores puede realizarse a través de C2c2 para determinar si un paciente está respondiendo a un tratamiento dado o, si el paciente no está respondiendo, si esto puede deberse a la disfunción de las células T. Tal detección es informativa con respecto a los tipos de terapia que son los más adecuados para que reciba el paciente. Por ejemplo, si el paciente debe recibir inmunoterapia.

En algunas realizaciones, los sistemas y ensayos dados a conocer en el presente documento pueden permitir a los médicos identificar si la respuesta de un paciente a una terapia (por ejemplo, una terapia adoptiva de transferencia de células (ACT)) se debe a la disfunción celular, y si es así, los niveles de regulación por incremento y regulación por disminución a lo largo de la firma de biomarcador permitirán que se aborden los problemas. Por ejemplo, si un paciente que recibe ACT no responde al tratamiento, las células administradas como parte de la ACT pueden someterse a ensayo mediante un ensayo dado a conocer en el presente documento para determinar el nivel de expresión relativo de una firma de biomarcador que se sabe que está asociada con los estados de activación y/o disfunción celular. Si una molécula o un receptor inhibidor particular está regulado por incremento en las células ACT, el paciente puede tratarse con un inhibidor de esa molécula o ese receptor. Si una molécula o un receptor estimulador particular está regulado por disminución en las células ACT, el paciente puede tratarse con un agonista de esa molécula o ese receptor.

En determinadas realizaciones de ejemplo, los sistemas, métodos y dispositivos descritos en el presente documento pueden usarse para examinar firmas génicas que identifican un tipo de célula particular, el fenotipo de la célula o el estado de la célula. Del mismo modo, mediante el uso de métodos tales como detección comprimida, las realizaciones dadas a conocer en el presente documento pueden usarse para detectar transcriptomas. Los datos de expresión génica son altamente estructurados, de tal manera que el nivel de expresión de algunos genes es predictivo del nivel de expresión de otros. El conocimiento de que los datos de expresión génica son altamente estructurados permite la suposición de que el número de grados de libertad en el sistema es pequeño, lo que permite suponer que la base para la computación de las abundancias génicas relativas es escasa. Es posible realizar varias suposiciones biológicamente motivadas que permitan a los solicitantes recuperar los términos de interacción no lineal mientras realizan un muestreo insuficiente sin tener ningún conocimiento específico de qué genes es probable que interaccionen. En particular, si los solicitantes suponen que las interacciones genéticas tienen un rango bajo, son escasas, o una combinación de los mismos, entonces el verdadero número de grados de libertad es pequeño en relación con la expansión combinatoria completa, lo que permite a los solicitantes inferir el panorama no lineal completo con un número de perturbaciones relativamente pequeño. El trabajo alrededor de estas suposiciones, las teorías analíticas de completitud de matriz y la detección comprimida pueden usarse para diseñar experimentos de perturbación combinatoria de muestreo insuficiente. Además, pueden usarse un marco de aprendizaje de núcleos para emplear el muestreo insuficiente mediante la construcción de funciones predictivas de perturbaciones combinatorias sin aprender directamente ningún coeficiente de interacción individual. La detección comprimida proporciona una manera de identificar el mínimo número de transcritos diana que van a detectarse con el fin de obtener un perfil de expresión génica exhaustivo. Se dan a conocer métodos para la detección comprimida en el documento PCT/US2016/059230 "Systems and Methods for Determining Relative Abundances of Biomolecules" presentado el 27 de octubre de 2016. Habiendo usado métodos como la detección comprimida para identificar un conjunto mínimo de dianas de transcritos, entonces puede diseñarse un conjunto de ARN guías correspondientes para detectar dichos transcritos. Por consiguiente, en determinadas realizaciones de ejemplo, un método para obtener un perfil de expresión génica de una célula comprende detectar, usando las realizaciones dadas a conocer en el presente documento, un conjunto mínimo de transcritos que proporcione un perfil de expresión génica de una célula o población de células.

DETECCIÓN DE EDICIONES GÉNICAS Y/O EFECTOS INESPECÍFICOS

Las realizaciones dadas a conocer en el presente documento pueden usarse en combinación con otras herramientas de edición génica para confirmar que una edición o ediciones génicas deseadas tuvieron éxito y/o para detectar la presencia de cualquier efecto inespecífico. Las células que se han editado pueden examinarse usando una o más guías para uno o más loci diana. Dado que las realizaciones dadas a conocer en el presente documento utilizan sistemas CRISPR, también se prevén aplicaciones terapéuticas. Por ejemplo, las realizaciones de genotipado dadas a conocer en el presente documento pueden usarse para seleccionar loci diana apropiados o identificar células o poblaciones de células que necesitan la edición diana. Luego puede usarse el mismo sistema, o uno

independiente, para determinar la eficiencia de edición. Tal como se describe en los ejemplos de trabajo a continuación, las realizaciones dadas a conocer en el presente documento pueden usarse para diseñar canalizaciones tratanósticas funcionales en tan solo una semana.

DETECCIÓN DE ELEMENTOS ETIQUETADOS CON ÁCIDOS NUCLEICOS

5 Alternativamente, las realizaciones descritas en el presente documento pueden usarse para detectar identificadores de ácidos nucleicos. Los identificadores de ácidos nucleicos son ácidos nucleicos no codificantes que pueden usarse para identificar un elemento particular. Se describen identificadores de ácidos nucleicos de ejemplo, tales como marcas de agua de ADN, en Heider y Barnekow. "DNA watermarks: A proof of concept" BMC Molecular Biology 9:40 (2008). Los identificadores de ácidos nucleicos también pueden ser un código de barras de ácido nucleico. Un
10 código de barras basado en ácido nucleico es una secuencia de nucleótidos corta (por ejemplo, ADN, ARN o combinaciones de los mismos) que se usa como identificador para una molécula asociada, tal como una molécula diana y/o un ácido nucleico diana. Un código de barras de ácido nucleico puede tener una longitud de al menos, por ejemplo, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45,
15 50, 60, 70, 80, 90 o 100 nucleótidos, y puede estar en forma monocatenaria o bicatenaria. Uno o más códigos de barras de ácido nucleico pueden unirse a, o "etiquetarse con", una molécula diana y/o un ácido nucleico diana. Esta unión puede ser directa (por ejemplo, unión covalente o no covalente del código de barras a la molécula diana) o indirecta (por ejemplo, a través de una molécula adicional, por ejemplo, un agente de unión específico, tal como un anticuerpo (u otra proteína) o un adaptador que recibe el código de barras (u otra molécula de ácido nucleico). La molécula diana y/o los ácidos nucleicos diana pueden marcarse con múltiples códigos de barras de ácido nucleico
20 de manera combinatoria, tal como un concatémero de código de barras de ácido nucleico. Normalmente, se usa un código de barras de ácido nucleico para identificar que las moléculas diana y/o los ácidos nucleicos diana pertenecen a un compartimento particular (por ejemplo, un volumen discreto), tienen una propiedad física particular (por ejemplo, afinidad, longitud, secuencia, etc.) o se han sometido a determinadas condiciones de tratamiento. La molécula diana y/o el ácido nucleico diana puede asociarse con múltiples códigos de barras de ácido nucleico para
25 proporcionar información sobre todas estas características (y más). Se dan a conocer métodos para generar códigos de barras de ácido nucleico, por ejemplo, en la publicación de solicitud de patente internacional n.º WO/2014/047561.

Enzimas

La solicitud proporciona además ortólogos de C2c2 que demuestran una actividad robusta, lo que les hace
30 particularmente adecuados para diferentes aplicaciones de escisión y detección de ARN. Estas aplicaciones incluyen, pero no se limitan a, las descritas en el presente documento. Más particularmente, un ortólogo que se demuestra que tiene una actividad más fuerte que los demás sometidos a prueba es el ortólogo de C2c2 identificado a partir del organismo *Leptotrichia wadei* (LwC2c2). Por tanto, la solicitud proporciona métodos para modificar un locus diana de interés, que comprende administrar a dicho locus una composición que no se produce de manera
35 natural o modificada por ingeniería que comprende una proteína efectora C2c2, más particularmente una proteína efectora C2c2 con actividad aumentada tal como se describe en el presente documento y uno o más componentes de ácido nucleico, en los que al menos el uno o más componentes de ácido nucleico está modificado por ingeniería, el uno o más componentes de ácido nucleico dirige el complejo a la diana de interés y la proteína efectora forma un complejo con el uno o más componentes de ácido nucleico y el complejo se une al locus diana de interés. En
40 realizaciones particulares, el locus diana de interés comprende ARN. La solicitud proporciona además el uso de las proteínas efectoras C2c2 con actividad aumentada en interferencia específica de secuencia del ARN, regulación génica específica de secuencia del ARN, examen de ARN o productos de ARN o ARNlinc o ARN no codificante, o ARN nuclear, o ARNm, mutagénesis, hibridación *in situ* por fluorescencia, o reproducción.

La invención se describe adicionalmente en los siguientes ejemplos, que no limitan el alcance de la invención
45 descrita en las reivindicaciones.

Ejemplos de trabajo

EJEMPLO 1 - Protocolos generales

Se proporcionan dos modos para realizar una prueba de diagnóstico C2c2 para ADN y ARN. Este protocolo también puede usarse con variantes de detección de proteínas después de la administración de los aptámeros de detección.
50 La primera es una reacción de dos etapas en la que se realizan por separado la amplificación y la detección con C2c2. La segunda es donde se combina todo en una reacción y esta se denomina reacción de dos etapas. Es importante tener en cuenta que la amplificación podría no ser necesaria para muestras de mayor concentración, por lo que es bueno disponer de un protocolo de C2c2 independiente que no incorpore amplificación.

Sólo efector de CRISPR - sin amplificación:

Componente	Volumen (μl)
Proteína (44 nM final)	2
ARNcr (12 nM final)	1
Diana de fondo (100 ng en total)	1
ARN diana (variable)	1
Sonda de sensor de ARN (125 nM)	4
MgCl ₂ (6 mM final)	2
Tampón de reacción 10x	2
Inhibidores de ARNasa (murina de NEB)	2
H ₂ O	5
total	20

El tampón de reacción es: Tris-HCl 40 mM, NaCl 60 mM, pH 7,3

Realizar esta reacción durante 20 min-3 h a 37°C. Lectura con excitación: 485 nm/20 nm, emisión: 528 nm/20 nm. Puede detectarse una señal para sensibilidad de molécula única comenzando a los 20 min pero, por supuesto, la sensibilidad es mayor para tiempos de reacción más prolongados.

5

Reacción de dos etapas:

Mezcla de amplificación RPA

Componente	Volumen (μl)
Cebador A (100 μM)	0,48
Cebador B (100 μM)	0,48
Tampón de RPA	59
MgAc	5
Diana (concentración variable)	5
ATP (100 μM del kit NEB)	2
GTP (100 μM del kit NEB)	2
UTP (100 μM del kit NEB)	2
CTP (100 μM del kit NEB)	2
Polimerasa T7 (del kit NEB)	2
H ₂ O	25
total	104,96

Mezclar esta reacción conjuntamente y luego resuspender de dos a tres tubos de mezcla de enzimas secada por congelación. Añadir 5 μl de MgAc 280 mM a la mezcla para comenzar la reacción. Realizar la reacción durante 10-20 min. Cada reacción es de 20 μl, por lo que esto es suficiente para hasta cinco reacciones.

10

Mezcla de detección con C2c2

Componente	Volumen (μl)
Proteína (44 nM final)	2
ARNcr (12 nM final)	1
Diana de fondo (100 ng en total)	1
Reacción de RPA	1
Sonda de sensor de ARN (125 nM)	4
MgCl ₂ (6 mM final)	2
Tampón de reacción 10x	2
Inhibidores de ARNasa (murina de NEB)	2
H ₂ O	5
total	20

El tampón de reacción es: Tris-HCl 40 mM, NaCl 60 mM, pH 7,3

Realizar esto durante 20 min – 3 horas. El tiempo mínimo de detección es de aproximadamente 20 min para observar sensibilidad de molécula única. Realizar la reacción durante más tiempo sólo mejora la sensibilidad.

5 Reacción en un solo recipiente:

Componente	Volumen (μl)
Cebador A (100 μM)	0,48
Cebador B (100 μM)	0,48
Tampón de RPA	59
MgAc	5
Lw2C2c2 (44 nM final)	2
ARNcr (12 nM final)	2
ARN de fondo (desde 250 ng/μl)	2
Sustr. RNase Alert (después de resuspender en 20 μl)	5
Inhib. de ARNasa murina de NEB	10
Diana (concentración variable)	5
ATP (100 μM del kit NEB)	2
GTP (100 μM del kit NEB)	2
UTP (100 μM del kit NEB)	2
CTP (100 μM del kit NEB)	2
Polimerasa T7 (del kit NEB)	2
H ₂ O	4
total	104,96

El kit NEB al que se hace referencia es el kit HighScribe T7 High Yield. Para resuspender el tampón, usar una concentración 1,5x: resuspender tres tubos de sustrato secado por congelación en 59 μl de tampón y usar en la mezcla anterior. Cada reacción es de 20 μl, por lo que esto es suficiente para 5 reacciones. La sensibilidad de molécula única con esta reacción se ha observado en tan solo 30-40 min.

10

EJEMPLO 2 - C2c2 DE *LEPTOTRICHIA WADEI* MEDIA EN LA DETECCIÓN ALTAMENTE SENSIBLE Y

ESPECÍFICA DE ADN Y ARN

La detección rápida, rentable y sensible de ácidos nucleicos puede ayudar en la detección de patógenos de diagnóstico inmediato, en el genotipado y en la monitorización de la enfermedad. El efector CRISPR que selecciona como diana ARN guiado por ARN Cas13a (previamente conocido como C2c2) presenta un “efecto colateral” de actividad ARNasa promiscua tras el reconocimiento de la diana. El solicitante combinó el efecto colateral de Cas13a con la amplificación isotérmica para establecer un diagnóstico basado en CRISPR (CRISPR-Dx), que proporciona una detección rápida de ADN o ARN con sensibilidad attomolar y especificidad de apareamiento erróneo de base única. El solicitante usó esta plataforma de detección molecular basada en Cas13a, denominada SHERLOCK (desbloqueo del indicador enzimático específico de alta sensibilidad), para detectar cepas específicas de virus de Zika y del dengue, distinguir bacterias patógenas, genotipar ADN humano e identificar mutaciones de ADN tumoral libre de células. Además, los reactivos de reacción de SHERLOCK pueden liofilizarse para independencia de la cadena de frío y almacenamiento a larga plazo, y reconstituirse fácilmente sobre papel para aplicaciones de campo.

La capacidad para detectar rápidamente ácidos nucleicos con alta sensibilidad y especificidad de base única en una plataforma portátil puede ayudar en el diagnóstico y la monitorización de enfermedades, en la epidemiología y en tareas generales de laboratorio. Aunque existen métodos para detectar ácidos nucleicos (1-6), tienen equilibrios entre sensibilidad, especificidad, sencillez, coste y velocidad. Los sistemas inmunitarios adaptativos microbianos de repeticiones palindrómicas cortas regularmente espaciadas (CRISPR) de agrupamientos microbianos y asociados a CRISPR (CRISPR-Cas) contienen endonucleasas programables que pueden aprovecharse para diagnósticos basados en CRISPR (CRISPR-Dx). Aunque algunas enzimas Cas seleccionan como diana el ADN (7, 8), las ARNasas guiadas por ARN efectoras únicas, tales como Cas13a (previamente conocida como C2c2) (8), pueden reprogramarse con ARN de CRISPR (ARNcr) (9-11) para proporcionar una plataforma para la detección específica de ARN. Tras el reconocimiento de su diana de ARN, la Cas13a activada participa en la escisión “colateral” de los ARN no seleccionados como diana cercanos (10). Esta actividad de escisión colateral programada por ARNcr permite que Cas13a detecte la presencia de un ARN específico *in vivo* desencadenando la muerte celular programada (10) o *in vitro* mediante la degradación inespecífica de ARN marcado (10, 12). En este caso, el solicitante describe SHERLOCK (desbloqueo del indicador enzimático específico de alta sensibilidad), una plataforma de detección de ácidos nucleicos *in vitro* con sensibilidad attomolar basada en la amplificación de ácidos nucleicos y 3 escisiones colaterales mediadas por Cas13a de un ARN indicador comercial (12), lo que permite la detección en tiempo real de la diana (figura 17).

Métodos

Clonación de loci de C2c2 y proteínas para la expresión

Para el ensayo de eficiencia *in vivo* en bacterias, se ordenaron proteínas C2c2 de *Leptotrichia wadei* F0279 y *Leptotrichia shahii* como genes con codones optimizados para la expresión en mamíferos (Genscript, Jiangsu, China) y se clonaron en estructuras principales de pACYC184 junto con las repeticiones directas correspondientes que flanquean o bien un espaciador que selecciona como diana beta-lactamasa o bien un espaciador que no selecciona como diana beta-lactamasa. La expresión del espaciador se impulsó mediante un promotor J23119.

Para la purificación de proteínas, se clonaron proteínas C2c2 con codones optimizados de mamífero en vector de expresión bacteriano para la purificación de proteínas (6x His/Twin Strep SUMO, un vector de expresión basado en pET recibido como un obsequio de Ilya Finkelstein).

Ensayo de eficiencia de C2c2 *in vivo* en bacterias

Se transformaron conjuntamente plásmidos de eficiencia *in vivo* LwC2c2 y LshC2c2 y un plásmido de beta-lactamasa previamente descrito (Abudayyeh 2016) en células competentes NovaBlue Singles (Millipore) a 90 ng y 25 ng, respectivamente. Después de la transformación, se sembraron diluciones de células en una placa de LB-agar con ampicilina y cloranfenicol y se incubaron durante la noche a 37°C. Se contaron las colonias al día siguiente.

Diana de ácido nucleico y preparación de ARNcr

Se amplificaron por PCR las dianas de ácido nucleico con KAPA Hifi Hot Start (Kapa Biosystems), se extrajeron en gel y se purificaron usando el kit de extracción en gel MinElute (Qiagen). Se incubó el ADNbc purificado con polimerasa T7 durante la noche a 30°C usando el kit de síntesis de ARN HiScribe T7 Quick High Yield (New England Biolabs) y se purificó el ARN con el kit de limpieza de transcripción MEGAclear (Thermo Fisher).

Para la preparación de ARNcr, se ordenaron constructos como ADN (Integrated DNA Technologies) con una secuencia promotora T7 unida. Se hibridó el ADN de ARNcr con un cebador T7 corto (concentraciones finales de 10 uM) y se incubó con polimerasa T7 durante la noche a 37°C usando el kit de síntesis de ARN HiScribe T7 Quick High Yield (New England Biolabs). Se purificó el ARNcr usando perlas de limpieza RNAXP (Beckman Coulter) a una razón 2x de perlas con respecto al volumen de reacción, con una complementación 1,8x adicional de isopropanol (Sigma).

Amplificación isotérmica NASBA

Se describen detalles de la reacción de NASBA en [Pardee 2016]. Para un volumen de reacción total de 20 µl, se combinaron 6,7 µl de tampón de reacción (Life Sciences, NECB-24), 3,3 µl de mezcla de nucleótidos (Life Sciences, NECN-24), 0,5 µl de agua libre de nucleasas, 0,4 µl de cebadores de NASBA 12,5 µM, 0,1 µl de inhibidor de ARNasa (Roche, 03335402001) y 4 µl de amplicón de ARN (o agua para el control negativo) a 4°C y se incubaron a 65°C durante 2 min y luego a 41°C durante 10 min. Se añadieron 5 µl de mezcla de enzimas (Life Sciences, NEC-1-24) a cada reacción, y se incubó la mezcla de reacción a 41°C durante 2 h. Los cebadores de NASBA usados fueron 5'-AATTCTAATACGACTCACTATAGGGGGATCCTCTAGAAATATGGATT-3' (SEQ ID NO. 16) y 5'-CTCGTATGTTGTGTGGAATTGT-3' (SEQ ID NO. 17), y la parte subrayada indica la secuencia promotora T7.

Amplificación por recombinasa y polimerasa

Se diseñaron cebadores para RPA usando Primer blast del NCBI (Ye *et al.*, BMC Bioinformatics 13, 134 (2012) usando los parámetros por defecto, con la excepción del tamaño de amplicón (entre 100 y 140 nt), las temperaturas de fusión del cebador (entre 54°C y 67°C) y el tamaño del cebador (entre 30 y 35 nt). Luego se ordenaron los cebadores como ADN (Integrated DNA Technologies).

Se ejecutaron las reacciones de RPA y RT-RPA según las instrucciones con TwistAmp® Basic o TwistAmp® Basic RT (TwistDx), respectivamente, con la excepción de que se añadió MgAc 280 mM antes del molde de entrada. Se ejecutaron las reacciones con 1 µl de entrada durante 2 h a 37°C, a menos que se describa lo contrario.

Purificación de proteína LwC2c2

Se transformaron vectores de expresión bacteriana de C2c2 en células competentes individuales Rosetta™ 2(DE3) pLysS (Millipore). Se hizo crecer un cultivo iniciador de 16 ml en medio de crecimiento Terrific Broth 4 (triptona 12 g/l, extracto de levadura 24 g/l, K₂HPO₄ 9,4 g/l, KH₂PO₄ 2,2 g/l, Sigma) (TB) que se usó para inocular 4 l de TB, que se incubó a 37°C, 300 RPM hasta una DO600 de 0,6. En ese momento, se indujo la expresión de proteínas mediante la complementación con IPTG (Sigma) hasta una concentración final de 500 µM, y se enfriaron las células hasta 18°C durante 16 h para la expresión de proteínas. Luego se centrifugaron las células a 5200 g, 15 min, 4°C. Se recogió el sedimento celular y se almacenó a -80°C para su posterior purificación.

Todas las etapas posteriores de la purificación de proteínas se realizaron a 4°C. se trituró el sedimento celular y se resuspendió en tampón de lisis (Tris-HCl 20 mM, NaCl 500 mM, DTT 1 mM, pH 8,0) complementado con inhibidores de proteasas (comprimidos libres de EDTA Complete Ultra), lisozima y benzonasa seguido de sonicación (Sonifier 450, Branson, Danbury, CT) con las siguientes condiciones: amplitud de 100 durante 1 segundo encendido y 2 segundos apagado con un tiempo de sonicación total de 10 minutos. Se aclaró el lisado mediante centrifugación durante 1 hora a 4°C a 10.000 g y se filtró el sobrenadante a través de un filtro Stericup de 0,22 micrómetros (EMD Millipore). Se aplicó el sobrenadante filtrado a StrepTactin Sepharose (GE) y se incubó con rotación durante 1 hora seguido del lavado de la resina StrepTactin unida a proteína tres veces en tampón de lisis. Se resuspendió la resina en tampón de digestión SUMO (Tris-HCl 30 mM, NaCl 500 mM, DTT 1 mM, Igepal al 0,15% (NP-40), pH 8,0) junto con 250 unidades de proteasa SUMO (ThermoFisher) y se incubó durante la noche a 4°C con rotación. Se confirmó la digestión mediante SDS-PAGE y tinción con azul de Coomassie y se aisló el eluido de proteína mediante la centrifugación de la resina. Se cargó la proteína en una columna de intercambio catiónico HiTrap SP HP de 5 ml (GE Healthcare Life Sciences) a través de FPLC (AKTA PURE, GE Healthcare Life Sciences) y se eluyó a lo largo de un gradiente de sales desde 130 mM hasta 2 M de NaCl en tampón de elución (Tris-HCl 20 mM, DTT 1 mM, glicerol al 5%, pH 8,0). Se sometieron a prueba las fracciones resultantes para determinar la presencia de LwC2c2 mediante SDS-PAGE y se agruparon las fracciones que contenían la proteína y se concentraron a través de una unidad de filtración centrífuga hasta 1 ml en tampón S200 (HEPES 10 mM, NaCl 1 M, MgCl₂ 5 mM, DTT 2 mM, pH 7,0). Se cargó la proteína concentrada en una columna de filtración en gel (Superdex® 200 Increase 10/300 GL, GE Healthcare Life Sciences) a través de FPLC. Se analizaron las fracciones resultantes de la filtración en gel mediante SDS-PAGE y se agruparon las fracciones que contenían LwC2c2 y se sometieron a intercambio de tampón a tampón de almacenamiento (NaCl 600 mM, Tris-HCl 50 mM pH 7,5, glicerol al 5%, DTT 2 mM) y se congelaron a -80°C para su almacenamiento.

Detección colateral de LwC2c2

Se realizaron ensayos de detección con LwC2c2 purificada 45 nM, ARNcr 22,5 nM, indicador de sustrato 125 nM (Thermo Scientific RNase Alert v2), 2 µl de inhibidores de ARNasa murina, 100 ng de ARN total de fondo y cantidades variables de diana de ácido nucleico de entrada, a menos que se indique lo contrario, en tampón de ensayo de nucleasas (Tris-HCl 40 mM, NaCl 60 mM, MgCl₂ 6 mM, pH 7,3). Si la entrada era ADN amplificado que incluía un promotor T7 de una reacción de RPA, se modificó la reacción de C2c2 anterior para incluir ATP 1 mM, GTP 1 mM, UTP 1 mM, CTP 1 mM y 0,6 µl de mezcla de polimerasa T7 (NEB). Se permitió que las reacciones avanzaran durante 1-3 horas a 37°C (a menos que se indique lo contrario) en un lector de placas de fluorescencia (BioTek), midiéndose la cinética de fluorescencia cada 5 minutos.

Se realizó la reacción en un solo recipiente que combinaba amplificación RPA de ADN, conversión por polimerasa T7 de ADN en ARN y detección con C2c2 mediante la integración de las condiciones de reacción anteriores con la mezcla de amplificación RPA. En resumen, en un ensayo en un solo recipiente de 50 µl que consistía en cebador

directo 0,48 μ M, cebador inverso 0,48 μ M, tampón de rehidratación de RPA 1x, cantidades variables de entrada de ADN, proteína recombinante Lwc2c2 45 nM, ARNcr 22,5 nM, 250 ng de ARN total de fondo, indicador de sustrato 200 nM (RNase Alert v2), 4 μ l de inhibidor de ARNasa, ATP 2 mM, GTP 2 mM, UTP 2 mM, CTP 2 mM, 1 μ l de mezcla de polimerasa T7, $MgCl_2$ 5 mM y $MgAc$ 14 mM.

5 *Análisis por PCR cuantitativa (qPCR) con sondas TaqMan*

Para comparar la cuantificación por SHERLOCK con otros métodos establecidos, se realizó qPCR sobre una serie de diluciones de ADNmc 1. Se diseñaron una sonda TaqMan y un conjunto de cebadores (secuencias a continuación) contra ADNmc 1 y se sintetizaron con IDT. Se realizaron los ensayos usando la mezcla maestra TaqMan Fast Advanced (Thermo Fisher) y se midieron en un dispositivo Roche LightCycler 480.

10 *Tabla de secuencias de cebadores/sondas de qPCR.*

Nombre	Secuencia
Cebador directo	GTG GAA TTG TGA GCG GAT AAA C (SEQ ID NO: 420)
Cebador inverso	AAC AGC AAT CTA CTC GAC CTG (SEQ ID NO: 421)
Sonda TaqMan	/56-FAM/AGGAAACAG/ZEN/CTATGACCATGATTACGCC/31ABkFQ/ (SEQ ID NO: 422 y 423)

RPA en tiempo real con SYBR Green II

Para comparar la cuantificación por SHERLOCK con otros métodos establecidos, el solicitante realizó RPA sobre una serie de diluciones de ADNmc 1. Para cuantificar la acumulación de ADN en tiempo real, el solicitante añadió 1x SYBR Green II (Thermo Fisher) a la mezcla de reacción de RPA típica descrita anteriormente, que proporciona una señal fluorescente que se correlaciona con la cantidad de ácido nucleico. Se permitió que las reacciones avanzaran durante 1 h a 37°C en un lector de placas de fluorescencia (BioTek), midiéndose la cinética de fluorescencia cada 5 min.

Preparación y procesamiento de lentivirus

La preparación y el procesamiento de lentivirus se basó en los métodos previamente conocidos. En resumen, se transfectaron 10 μ g de derivados de pSB700 que incluían un fragmento de ARN de Zika o de dengue, 7,5 μ g de psPAX2 y 2,5 μ g de pMD2.G con células HEK293FT (Life Technologies, R7007) usando el método de HeBS- $CaCl_2$. 28 h después de cambiar el medio, DMEM suplementado con FBS al 10%, penicilina-estreptomicina al 1% y GlutaMAX 4 mM (ThermoFisher Scientific), se filtró el sobrenadante usando un filtro de jeringa de 0,45 μ m. Se usaron el kit de purificación de lentivirus ViralBind (Cell Biolabs, VPK-104) y Lenti-X Concentrator (Clontech, 631231) para purificar y preparar lentivirus a partir del sobrenadante. Se cuantificó la concentración viral usando kit de lentivirus QuickTiter (Cell Biolabs, VPK-112). Se realizaron adiciones conocidas de muestras virales a suero humano al 7% (Sigma, H4522), se calentaron hasta 95°C durante 2 min y se usaron como entrada para la RPA.

Aislamiento y purificación de ADNc de muestras de suero humano de Zika

Se inactivaron muestras de orina o suero humano positivas para Zika sospechosas con tampón AVL (Qiagen) y se logró el aislamiento de ARN con el minikit de ARN viral QIAamp (Qiagen). Se convirtió el ARN aislado en ADNc mezclando cebadores aleatorios, dNTP y ARN de muestra seguido de desnaturalización por calor durante 7 minutos a 70°C. Luego se sometió a transcripción inversa el ARN desnaturalizado con Superscript III (Invitrogen) incubando a 22-25°C durante 10 minutos, 50°C durante 45 minutos, 55°C durante 15 minutos y 80°C durante 10 minutos. Luego se incubó el ADNc durante 20 minutos a 37°C con ARNasa H (New England Biolabs) para destruir el ARN en los híbridos ARN:ADNc.

Extracción de ADN genómico a partir de saliva humana

Se recogieron 2 ml de saliva de los voluntarios, a quienes se les restringió el consumo de alimento o bebida 30 minutos antes de la recogida. Luego se procesaron las muestras usando el minikit sanguíneo de ADN QIAamp® (Qiagen) según lo recomendado por el protocolo del kit. Para muestras de saliva hervidas, se añadieron 400 μ l de solución salina tamponada con fosfato (Sigma) a 100 μ l de saliva de voluntario y se centrifugaron durante 5 min a 1800 g. Se decantó el sobrenadante y se resuspendió el sedimento en solución salina tamponada con fosfato con Triton X-100 al 0,2% (Sigma) antes de la incubación a 95°C durante 5 min. Se usó 1 μ l de muestra como entrada directa a las reacciones de RPA.

45 *Secado por congelación y deposición sobre papel*

Se sometió a esterilización en autoclave un papel de filtro de fibra de vidrio (Whatman, 1827-021) durante 90 min (Consolidated Stills and Sterilizers, MKII) y se bloqueó en BSA libre de nucleasas al 5% (EMD Millipore, 126609-10GM) durante la noche. Después de enjuagar los papeles una vez con agua libre de nucleasas (Life technologies,

AM9932), se incubaron con RNAsure™ al 4% (Life technologies, AM7006) a 60°C durante 20 min y se enjuagaron tres veces más con agua libre de nucleasas. Se secaron los papeles tratados durante 20 min a 80°C en una placa caliente (Cole-Parmer, IKA C-Mag HS7) antes de su uso. Se colocaron 1,8 µl de mezcla de reacción de C2c2 tal como se indicó anteriormente sobre el disco (2 mm) que se colocó en una placa negra de 384 pocillos de fondo transparente (Corning, 3544). Para la prueba de secado por congelación, se ultracongeló en nitrógeno líquido la placa que contenía los discos con mezcla de reacción y se secó por congelación durante la noche tal como se describe en Pardee *et al.* (2). Se diluyeron 1:10 las muestras de RPA en agua libre de nucleasas, y se cargaron 1,8 µl de la mezcla en los discos de papel y se incubaron a 37°C usando un lector de placas (BioTek Neo).

Extracción de ADN genómico bacteriano

Para los experimentos que implican detección de CRE, se hicieron crecer cultivos bacterianos en caldo de lisogenia (LB) hasta la fase logarítmica media, luego se sedimentaron y se sometieron a extracción y purificación de ADN_g usando el kit sanguíneo y tisular Qiagen DNeasy, usando el protocolo del fabricante para bacterias o bien Gram negativas o bien Gram positivas, según fuese apropiado. Se cuantificó el ADN_g mediante el ensayo de ADN_{bc} Quant-It en un fluorímetro Qubit y se evaluó su calidad a través del espectro de absorbancia a 200-300 nm en un espectrofotómetro Nanodrop.

Para los experimentos que discriminan entre *E. coli* y *P. aeruginosa*, se hicieron crecer cultivos bacterianos hasta la fase estacionaria temprana en caldo de Luria-Bertani (LB). Se procesó 1,0 ml tanto de *E. coli* como de *P. aeruginosa* usando el kit de extracción de ADN_g bacteriano PureLyse portátil (Claremont BioSolutions). Se añadió tampón de unión 1X al cultivo bacteriano antes de realizar los pasos a través del cartucho de lisis que funciona con pilas durante tres minutos. Se usó tampón de unión 0,5X en agua como disolución de lavado antes de eluir con 150 µl de agua.

Cuantificación por PCR digital en gotas

Para confirmar la concentración de las diluciones patrón de ADN_{mc} 1 y ARN_{mc} 1 usadas en la figura 1C, el solicitante realizó PCR digital en gotas (ddPCR). Para la cuantificación de ADN, se prepararon gotas usando la supermezcla de ddPCR para sondas (sin dUTP), diseñándose los ensayos de sondas/cebadores de qPCR PrimeTime para seleccionar como diana la secuencia de ADN_{mc} 1. Para la cuantificación de ARN, se prepararon gotas usando el kit de RT-ddPCR de una sola etapa para sondas, diseñándose los ensayos de sondas/cebadores de qPCR PrimeTime para seleccionar como diana la secuencia de ARN_{mc} 1. En cualquier caso, se generaron las gotas usando el generador de gotas QX200 (BioRad) y se transfirieron a una placa de PCR. Se realizó amplificación basada en gotas en un termociclador tal como se describe en el protocolo del kit y posteriormente se determinaron las concentraciones de ácido nucleico a través de la medición en un lector de gotas QX200.

Patrones sintéticos para el genotipado humano

Para crear patrones para la llamada precisa de genotipos de muestras humanas, el solicitante diseñó cebadores alrededor de la diana de SNP para amplificar regiones de ~200 pb de ADN genómico humano que representan cada uno de los dos genotipos homocigóticos. Luego se preparó el patrón heterocigótico mezclando los patrones homocigóticos en una razón 1:1. Luego se diluyeron estos patrones hasta concentraciones de genoma equivalentes (~0,56 fg/µl) y se usaron como entrada para SHERLOCK junto con muestras humanas reales.

Detección de ADN libre de células (ADN_{lc}) mutante tumoral

Se adquirieron patrones de ADN_{lc} simulado que simulan muestras de ADN_{lc} de pacientes reales de un proveedor comercial (Horizon Discovery Group). Se proporcionaron estos patrones como cuatro fracciones alélicas (el 100% de WT y el 0,1%, el 1% y el 5% de mutante) para los mutantes BRAF V600E y EGFR L858R. Se proporcionaron 3 µl de estos patrones como entrada a SHERLOCK.

Análisis de los datos de fluorescencia

Para calcular los datos de fluorescencia sustraída del fondo, se sustrajo la fluorescencia inicial de las muestras para permitir comparaciones entre condiciones diferentes. Se sustrajo la fluorescencia para las condiciones de fondo (o bien sin entrada o bien sin condiciones de ARN_{cr}) de las muestras para generar la fluorescencia sustraída del fondo.

Se calcularon las razones de guía para la discriminación de cepa o SNP dividiendo cada guía entre la suma de valores de guía, para ajustar la variación global de muestra a muestra. Se calcularon las razones de ARN_{cr} para la discriminación de cepa o SNP para ajustar la variación global de muestra a muestra de la siguiente manera:

$$\text{razón de } A_i \text{ de ARN}_{cr} = \frac{(m+n)A_i}{\sum_{i=1}^m A_i + \sum_{i=1}^n B_i}$$

en la que A_i y B_i se refieren a los valores de intensidad de SHERLOCK para la réplica técnica i de los ARN_{cr} que detectan el alelo A o el alelo B, respectivamente, para un individuo dado. Puesto que un ensayo tiene normalmente

cuatro réplicas técnicas por ARNcr, m y n son iguales a 4 y el denominador es equivalente a la suma de los ocho valores de intensidad de SHERLOCK del ARNcr para un locus de SNP e individuo dados. Debido a que hay dos ARNcr, el promedio de razones de ARNcr en cada uno de los ARNcr para un individuo siempre sumará dos. Por tanto, en el caso ideal de homocigosidad, la razón media de ARNcr para el ARNcr de alelo positivo será dos y la razón media de ARNcr para el ARNcr de alelo negativo será cero. En el caso ideal de heterocigosidad, la razón media de ARNcr para cada uno de los dos ARNcr será uno.

Caracterización de los requisitos de escisión de LwCas13a.

El sitio de flanqueo del protoespaciador (PFS) es un motivo específico presente cerca del sitio diana requerido para una actividad ribonucleasa robusta por parte de Cas13a. El PFS se ubica en el extremo 3' del sitio diana, y se caracterizó previamente para LshCas13a por el presente grupo como H (sin G) (1). Aunque este motivo es similar a un motivo adyacente al protoespaciador (PAM), una restricción de secuencia para los sistemas de clase 2 de direccionamiento de ADN, es funcionalmente diferente ya que no está implicado en prevenir el autodireccionamiento de los loci de CRISPR en sistemas endógenos. Estudios estructurales futuros de Cas13a probablemente dilucidarán la importancia del PFS para la actividad de escisión y la formación de complejos diana Cas13a:ARNcr.

El solicitante purificó la proteína LwCas13a recombinante a partir de *E. coli* (figura 2D-E) y sometió a ensayo su capacidad para escindir un ARNmc de 173 nt con cada nucleótido de sitio de flanqueo del protoespaciador (PFS) posible (A, U, C o G) (figura 2F). De manera similar a LshCas13a, LwCas13a puede escindir con robustez una diana con A, U o C PFS, con menos actividad en el ARNmc con un G PFS. Aunque se observó una actividad más débil contra ARNmc 1 con un G PFS, aun así el solicitante observó una detección robusta para los dos sitios diana con motivos G PFS (tabla 3; ARNcr de rs601338 y ARNcr 2 que selecciona como diana Zika). Es probable que no se requiera el H PFS en todas las circunstancias y que, en muchos casos, pueda lograrse una fuerte escisión o actividad colateral con un G PFS.

Discusión de la amplificación por recombinasa y polimerasa (RPA) y otras estrategias de amplificación isotérmica.

La amplificación por recombinasa y polimerasa (RPA) es una técnica de amplificación isotérmica que consiste en tres enzimas esenciales: una recombinasa, proteínas de unión a ADN monocatenario (SSB) y una polimerasa de desplazamiento de cadena. La RPA supera muchas de las dificultades técnicas presentes en otras estrategias de amplificación, particularmente la reacción en cadena de la polimerasa (PCR), al no requerir regulación de temperatura ya que las enzimas funcionan todas ellas a una temperatura constante de aproximadamente 37°C. La RPA reemplaza los ciclos de temperatura por la fusión global del molde bicatenario y la hibridación de cebadores por un enfoque enzimático inspirado por la replicación y reparación de ADN *in vivo*. Los complejos recombinasa-cebador exploran el ADN bicatenario y facilitan el intercambio de cadenas en sitios complementarios. Las SSB estabilizan el intercambio de cadenas, permitiendo que el cebador permanezca unido. El desensamblaje espontáneo de la recombinasa se produce en su estado unido a ADP, lo que permite que una polimerasa de desplazamiento de cadena invada y extienda el cebador, permitiendo la amplificación sin instrumentación compleja no disponible en entornos hospitalarios y de campo. La repetición cíclica de este procedimiento en un intervalo de temperaturas de 37-42°C da como resultado una amplificación exponencial del ADN. La formulación original publicada usa Pol I de *Bacillus subtilis* (Bsu) como polimerasa de desplazamiento de cadena, T4 uvsX como recombinasa y T4 gp32 como proteína de unión a ADN monocatenario (2), aunque no está claro qué componentes se encuentran en la formulación actual vendida por TwistDx usada en este estudio.

Además, la RPA presenta varias limitaciones:

1) Aunque la detección con Cas13a es cuantitativa (figura 15), la cuantificación por RPA en tiempo real puede ser difícil debido a su rápida saturación cuando la recombinasa usa todos los ATP disponibles. Aunque la PCR en tiempo real es cuantitativa debido a su capacidad de amplificación cíclica, la RPA no tiene ningún mecanismo para controlar estrictamente la tasa de amplificación. Pueden realizarse determinados ajustes para reducir la velocidad de amplificación, tales como reducir las concentraciones de magnesio o cebador disponibles, disminuir la temperatura de reacción o diseñar cebadores ineficientes. Aunque se observan algunos casos de SHERLOCK cuantitativo, tales como en la figura 31, 32 y 52, no siempre es el caso y puede depender del molde.

2) La eficiencia de la RPA puede ser sensible al diseño del cebador. El fabricante recomienda normalmente diseñar cebadores más largos para garantizar la unión eficiente de la recombinasa con un contenido de GC promedio (40-60%) y examinar hasta 100 pares de cebadores para hallar pares de cebadores altamente sensibles. El solicitante ha hallado con SHERLOCK que sólo tienen que diseñarse dos pares de cebadores para lograr una prueba attomolar con sensibilidad de una sola molécula. Esta robustez probablemente se deba a la amplificación adicional de la señal por la actividad colateral de Cas13a constitutivamente activa que compensa cualquier ineficiencia en la amplificación del amplicón. Esta cualidad es particularmente importante para la presente identificación de patógenos bacterianos en la figura 34. Se experimentaron problemas con la amplificación de regiones altamente estructuradas tales como los sitios del gen ARNr 16S en genomas bacterianos porque no existe ninguna etapa de fusión implicada en la RPA. Por tanto, la estructura secundaria en los cebadores se convierte en un problema, limitando la eficiencia de la amplificación y, por tanto, la sensibilidad. Se cree que las realizaciones dadas a conocer en el presente documento tuvieron éxito a pesar de estos problemas específicos de la RPA debido a la amplificación adicional de la señal de

Cas13a.

3) La longitud de secuencia de amplificación debe ser corta (100-200 pb) para una RPA eficiente. Para la mayoría de las aplicaciones, este no es un problema significativo y quizás es incluso ventajoso (por ejemplo, detección de ADNlc donde el tamaño de fragmento promedio es de 160 pb). A veces son importantes grandes longitudes de amplicón, tal como cuando se desean cebadores universales para la detección bacteriana y los SNP para la discriminación se extienden a lo largo de un gran área.

La modularidad de SHERLOCK permite el uso de cualquier técnica de amplificación, incluso enfoques no isotérmicos, antes de la transcripción con T7 y la detección con Cas13a. Se permite esta modularidad por la compatibilidad de las etapas de T7 y Cas13a en una única reacción que permite realizar la detección sobre cualquier entrada de ADN amplificado que tenga un promotor T7. Antes de usar la RPA, se intentó la amplificación basada en secuencias de ácido nucleico (NASBA) (3, 4) para el presente ensayo de detección (figura 10). Sin embargo, la NASBA no mejoró drásticamente la sensibilidad de Cas13a (figuras 11 y 53). Otras técnicas de amplificación que podrían emplearse antes de la detección incluyen PCR, amplificación isotérmica mediada por bucle (LAMP) (5), amplificación por desplazamiento de cadena (SDA) (6), amplificación dependiente de helicasa (HDA) (7) y reacción de amplificación por enzima de corte (NEAR) (8). La capacidad de intercambiar cualquier técnica isotérmica permite a SHERLOCK superar las limitaciones específicas de cualquier técnica de amplificación.

Diseño de apareamientos erróneos modificados por ingeniería.

El solicitante demuestra que se redujo la escisión de la diana LshCas13a cuando había dos o más apareamientos erróneos en el dúplex diana:ARNcr, pero no se vio afectada relativamente por apareamientos erróneos individuales, una observación que confirmó el solicitante para la escisión colateral de LwCas13a (figura 36^a). El solicitante planteó la hipótesis de que al introducir una mutación adicional en la secuencia espaciadora de ARNcr, el solicitante desestabilizaría la escisión colateral contra una diana con un apareamiento erróneo adicional (dos apareamientos erróneos en total) al tiempo que se conservaría la escisión colateral específica, ya que sólo habría un apareamiento erróneo individual. Para someter a prueba la posibilidad de diseñar por ingeniería una especificidad aumentada, el solicitante diseñó múltiples ARNcr que seleccionan como diana ARNmc 1 e incluyó apareamientos erróneos a lo largo de la longitud de los ARNcr (figura 36^a) para optimizar la escisión colateral específica y minimizar la escisión colateral de una diana que difiere en un apareamiento erróneo individual. El solicitante observó que estos apareamientos erróneos no reducían la escisión colateral de ARNmc 1, pero disminuían significativamente la señal para una diana que incluía un apareamiento erróneo adicional (ARNmc 2). El ARNcr diseñado que distinguía mejor entre ARNmc 1 y 2 incluía apareamientos erróneos sintéticos cercanos al apareamiento erróneo ARNmc 2, creando de hecho una "burbuja" o distorsión en el ARN hibridado. La pérdida de sensibilidad provocada por la coordinación de un apareamiento erróneo sintético y un apareamiento erróneo adicional presente en la diana (es decir, un apareamiento erróneo doble) concuerda con la sensibilidad de LshCas13a y LwCas13a a apareamientos erróneos dobles consecutivos o cercanos y presenta una base para el diseño racional de ARNcr que permiten la distinción de un solo nucleótido (figura 36B).

Para la detección de apareamientos erróneos de las cepas de ZIKV y DENV, el presente ARNcr de longitud completa contenía dos apareamientos erróneos (figura 37^a, B). Debido a la alta divergencia de secuencias entre las cepas, el solicitante no pudo hallar un tramo continuo de 28 nt con una única diferencia de un solo nucleótido entre los dos genomas. Sin embargo, el solicitante predijo que ARNcr más cortos seguirían siendo funcionales, y diseñó ARNcr de 23 nt más cortos contra dianas en las dos cepas de ZIKV que incluían un apareamiento erróneo sintético en la secuencia espaciadora y un único apareamiento erróneo en la secuencia diana. Estos ARNcr todavía podían distinguir las cepas africana y americana de ZIKV (figura 36C). Las pruebas posteriores de ARNcr de 23 nt y 20 nt muestran que las reducciones de la longitud del espaciador reducen la actividad pero mantienen o potencian la capacidad de discriminar apareamientos erróneos individuales (figura 57^a-G). Para entender mejor cómo pueden introducirse apareamientos erróneos sintéticos para facilitar la discriminación de mutaciones de un solo nucleótido, el solicitante dispuso el apareamiento erróneo sintético a lo largo de las primeras siete posiciones del espaciador a tres longitudes de espaciador diferentes: 28, 23 y 20 nt (figura 57^a). En una diana con una mutación en la tercera posición, LwCas13a muestra una especificidad máxima cuando el apareamiento erróneo sintético está en la posición 5 del espaciador, con una especificidad mejorada a longitudes del espaciador más cortas, aunque con niveles más bajos de actividad específica (figura 57B-G). El solicitante también desplazó la mutación diana a lo largo de las posiciones 3-6 y dispuso apareamientos erróneos sintéticos en el espaciador alrededor de la mutación (figura 58).

Genotipado con SHERLOCK usando patrones sintéticos.

La evaluación de patrones sintéticos creados a partir de amplificación por PCR de los loci de SNP permite la identificación precisa de genotipos (figura 60^a, B). Computando todas las comparaciones (ANOVA) entre los resultados de SHERLOCK de una muestra de individuo y los patrones sintéticos, puede identificarse cada genotipo de individuo hallando el patrón sintético que tiene la intensidad de detección por SHERLOCK más similar (figura 60C, D). Este enfoque de genotipado por SHERLOCK puede generalizarse a cualquier locus de SNP (figura 60E).

SHERLOCK es una plataforma de CRISPR-Dx asequible y adaptable.

Para el análisis de costes de SHERLOCK, se omitieron los reactivos que se determinó que tenían un coste insignificante, incluyendo los moldes de ADN para la síntesis de ARNcr, los cebadores usados en la RPA, los tampones habituales (MgCl₂, Tris HCl, glicerol, NaCl, DTT), el papel de filtro de microfibras de vidrio y el reactivo RNAsecure. Para los moldes de ADN, la síntesis de ultrámeros a partir de IDT proporciona un material para 40 reacciones de transcripción *in vitro* (siendo cada una suficiente para ~10.000 reacciones) por ~70 USD, lo que añade un coste insignificante a la síntesis de ARNcr. Para los cebadores de RPA, puede adquirirse una síntesis de IDT de 25 nmol de un cebador de ADN de 30 nt por ~10 USD, proporcionando un material adecuado para 5000 reacciones de SHERLOCK. El papel de microfibras de vidrio está disponible por 0,50 USD/lámina, que es suficiente para varios cientos de reacciones de SHERLOCK. El reactivo RNAsecure al 4% cuesta 7,20 USD/ml, que es suficiente para 500 pruebas.

Además, para todos los experimentos, excepto para los ensayos basados en el papel, se usaron placas de 384 pocillos (Corning 3544), al coste de 0,036 USD/reacción. Debido al coste insignificante, este no se incluyó en el análisis del coste global. Además, SHERLOCK-POC no requiere el uso de un recipiente de plástico, ya que puede realizarse fácilmente sobre papel. El método de lectura para SHERLOCK usado en el presente documento fue un lector de placas equipado o bien con un conjunto de filtros o bien con un monocromador. Como inversión de capital, no se incluyó el coste del lector en el cálculo, ya que el coste disminuye precipitadamente a medida que se ejecutan más reacciones en el instrumento y es insignificante. Para aplicaciones de POC, pueden usarse alternativas más baratas y portátiles, tales como espectrofotómetros portátiles (9) o lectores electrónicos portátiles (4), que reducen el coste de instrumentación hasta <200 USD. Si bien estas soluciones más portátiles reducirán la velocidad y la facilidad de la lectura en comparación con instrumentos más voluminosos, permiten un uso más amplio.

Resultados

El ensayo y los sistemas descritos en el presente documento pueden comprender generalmente un procedimiento de dos etapas de amplificación y detección. Durante la primera etapa, se amplifica la muestra de ácido nucleico, ya sea ARN o ADN, por ejemplo, mediante amplificación isotérmica. Durante la segunda etapa, se transcribe el ADN amplificado en ARN y posteriormente se incubaba con un efector CRISPR, tal como C2c2, y un ARNcr programado para detectar la presencia de la secuencia de ácido nucleico diana. Para permitir la detección, se añade a la reacción un ARN indicador que se ha marcado con un fluoróforo extinguido. La escisión colateral del ARN indicador da como resultado la no extinción del fluoróforo y permite la detección en tiempo real de la diana de ácido nucleico (figura 17^a).

Para lograr una detección robusta de la señal, se identificó un ortólogo de C2c2 a partir del organismo *Leptotrichia wadei* (LwC2c2) y se evaluó. Se evaluó la actividad de la proteína LwC2c2 expresándola junto con una matriz de CRISPR sintética en *E. coli* y programándola para escindir un sitio diana dentro del ARNm de beta-lactamasa, lo que conduce a la muerte de las bacterias bajo selección con ampicilina (figura 2B). Se observaron menos colonias de *E. coli* supervivientes con el locus de LwC2c2 que con el locus de LshC2c2, lo que demuestra una mayor actividad de escisión del ortólogo LwC2c2 (figura 2C). Luego se purificó la proteína LwC2c2 con codones optimizados humana a partir de *E. coli* (figura 2D-E) y se sometió a ensayo su capacidad para escindir un ARNmc de 173 nt con diferentes nucleótidos de sitio de flanqueo del protoespaciador (PFS) (figura 2F). LwC2c2 pudo escindir cada una de las cuatro dianas de PFS posibles, con una actividad ligeramente menor en el ARNmc con un G PFS.

La medición en tiempo real de la actividad colateral ARNasa de LwC2c2 se midió usando un lector de placas de fluorescencia de ARN disponible comercialmente (figura 17^a). Para determinar la sensibilidad inicial de la actividad de LwC2c2, se incubó LwC2c2 con diana de ARNmc 1 (ARNmc 1) y un ARNcr que es complementario a un sitio dentro de la diana de ARNmc, junto con la sonda de sensor de ARN (figura 18). Esto produjo una sensibilidad de ~50 fM (figura 27^a) que, aunque es más sensible que otras tecnologías de detección de ácidos nucleicos recientes (Pardee *et al.*, 2014), no es lo suficiente sensible para muchas aplicaciones de diagnóstico que requieren un rendimiento de detección por debajo de femtomolar (Barletta *et al.*, 2004; Emmadi *et al.*, 2011; Rissin *et al.*, 2010; Song *et al.*, 2013).

Para aumentar la sensibilidad, se añadió una etapa de amplificación isotérmica antes de la incubación con LwC2c2. El acoplamiento de detección mediada por LwC2c2 con enfoques de amplificación isotérmica previamente usados tales como amplificación basada en secuencias de ácido nucleico (NASBA) (Compton, 1991; Pardee *et al.*, 2016) mejoró la sensibilidad hasta un determinado grado (figura 11). Se sometió a prueba un enfoque de amplificación isotérmica alternativo, la amplificación por recombinasa y polimerasa (RPA) (Piepenburg *et al.*, 2006), que puede usarse para amplificar el ADN exponencialmente en menos de dos horas. Mediante la adición de un promotor de ARN polimerasa T7 a los cebadores de RPA, puede convertirse el ADN amplificado en ARN para su posterior detección mediante LwC2c2 (figura 17). Por tanto, en determinadas realizaciones de ejemplo, el ensayo comprende la combinación de amplificación mediante RPA, conversión con ARN polimerasa T7 del ADN en ARN y la posterior detección del ARN mediante desbloqueo con C2c2 de la fluorescencia a partir de un indicador extinguido.

Usando el método de ejemplo en una versión de ADN sintetizado de ARNmc 1, fue posible lograr una sensibilidad attomolar en el intervalo de 1-10 moléculas por reacción (figura 27B, parte izquierda). Con el fin de verificar la precisión de la detección, se calificó la concentración de ADN de entrada mediante PCR digital en gotas y se confirmó que la concentración de diana detectable más baja (2 aM) era a una concentración de una única molécula

por microlitro. Con la adición de una etapa de transcripción inversa, la RPA también puede amplificar el ARN en una forma de ADNbc, lo que permite lograr una sensibilidad attomolar en ARNmc 1 (27B, parte derecha). De manera similar, se confirmaron las concentraciones de dianas de ARN mediante PCR digital en gotas. Para evaluar la viabilidad del método de ejemplo para funcionar como prueba de diagnóstico de POC, se sometió a prueba la capacidad de todos los componentes, la RPA, la amplificación con polimerasa T7 y la detección de LwC2c2, para funcionar en una sola reacción y se halló una sensibilidad attomolar con una versión en un solo recipiente del ensayo (figura 22).

El ensayo puede realizar una detección viral sensible en líquido o sobre papel

A continuación se determinó si el ensayo sería eficaz en aplicaciones de enfermedades infecciosas que requieren una alta sensibilidad y podría beneficiarse de un diagnóstico portátil. Para someter a prueba la detección en un sistema modelo, se produjeron lentivirus que albergaban fragmentos de ARN del genoma del virus de Zika y del dengue de flavivirus relacionados (Dejnirattisai *et al.*, 2016), y se cuantificó el número de partículas virales (figura 31^a). Se detectaron los niveles de virus simulado hasta 2 aM. Al mismo tiempo, también fue posible mostrar una clara discriminación entre estos virus proxy que contiene fragmentos de ARN de Zika y dengue (figura 31B). Para determinar si el ensayo sería compatible con el secado por congelación para eliminar la dependencia de las cadenas de frío para la distribución, se secaron por congelación los componentes de reacción. Después de usar la muestra para rehidratar los componentes liofilizados, se detectaron 20 fM de ARNmc 1 (figura 33^a). Debido a que entornos de escasos recursos y de POC se beneficiarían de una prueba en papel para facilitar su uso, también se evaluó la actividad de detección con C2c2 sobre papel de fibra de vidrio y se halló que una reacción de C2c2 con manchas sobre papel podía realizar la detección de la diana (figura 33B). En combinación, el secado por congelación y el manchado sobre papel de la reacción de detección con C2c2 dio como resultado una detección sensible de ARNmc 1 (figura 33C). También se observaron niveles similares de sensibilidad para la detección de un fragmento de ARN viral de Zika sintético entre LwC2c2 en disolución y LwC2c2 secada por congelación, lo que demuestra la robustez de SHERLOCK secado por congelación y la posibilidad de un diagnóstico de virus de Zika de POC rápido (figura 33D-E). Para este fin, se sometió a prueba la capacidad de la variante de POC del ensayo para determinar la capacidad para discriminar el ARN de Zika del ARN de dengue (figura 31C). Si bien el manchado sobre papel y la liofilización redujeron ligeramente la señal absoluta de la lectura, el ensayo todavía detectó significativamente el virus de Zika simulado a concentraciones de tan sólo 20 aM (figura 31D), en comparación con la detección de virus simulado con la secuencia de control del dengue.

Se ha notificado que los niveles de ARN viral de Zika en seres humanos son de tan sólo 3×10^6 copias/ml (4,9 fM) en saliva del paciente y de $7,2 \times 10^5$ copias/ml (1,2 fM) en suero del paciente (Barzon *et al.*, 2016; Gourinat *et al.*, 2015; Lanciotti *et al.*, 2008). A partir de las muestras de pacientes obtenidas, se observaron concentraciones de tan sólo $1,25 \times 10^3$ copias/ml (2,1 aM). Para evaluar si el ensayo puede detectar el virus de Zika de aislados clínicos de bajo título, se extrajo ARN viral de los pacientes y se sometió a transcripción inversa, y se usó el ADNc resultante como entrada para el ensayo (figura 32^a). Se observó una detección significativa para las muestras de suero humano de Zika a concentraciones por debajo de 1,25 copias/ul (2,1 aM) (figura 32B). Además, la señal de las muestras de pacientes fue predictiva del número de copias de ARN viral de Zika y pudo usarse para predecir la carga viral (figura 31F). Para someter a prueba la amplia aplicabilidad para situaciones patológicas en las que no está disponible la purificación de ácidos nucleicos, se sometió a prueba la detección de ARNmc 1 añadido a suero humano, y se determinó que el ensayo se activaba a niveles séricos por debajo del 2% (figura 33G).

Distinción de patógenos bacterianos y distinción de genes

Para determinar si podía usarse el ensayo para distinguir patógenos bacterianos, se seleccionó la región 16S V3 como diana inicial, ya que las regiones flanqueantes conservadas permiten el uso de cebadores de RPA universales entre especies bacterianas, y la región interna variable que permite la diferenciación de especies. Se diseñaron un panel de 5 ARNcr de direccionamiento posibles para cepas patógenas y ADNg de *E. coli* y *Pseudomonas aeruginosa* aislado (figura 34^a). El ensayo pudo distinguir ADNg de *E. coli* o de *P. aeruginosa* y mostró una baja señal de fondo para el ARNcr de otras especies (figura 34^a, B).

El ensayo también puede adaptarse para detectar y distinguir rápidamente genes bacterianos de interés, tales como genes de resistencia a antibióticos. Las enterobacterias resistentes a los carbapenémicos (CRE) son un desafío emergente y significativo para la salud pública (Gupta *et al.*, 2011). Se evaluó la capacidad del ensayo para detectar genes de resistencia a los carbapenémicos y si la prueba podía distinguir entre diferentes genes de resistencia a los carbapenémicos. Se obtuvo *Klebsiella pneumoniae* de aislados clínicos que portaban los genes de resistencia a o bien carbapenemasa de *Klebsiella pneumoniae* (KPC) o bien metalo-beta-lactamasa 1 de Nueva Delhi (NDM-1) y ARNcr diseñados para distinguir entre los genes. Todas las CRE tenían una señal significativa con respecto a las bacterias que carecían de estos genes de resistencia (figura 35^a) y que pudo distinguirse significativamente entre las cepas de resistencia KPC y NDM-1 (figura 35B).

Especificidad de apareamientos erróneos de una sola base de ARNasas guiadas por ARN de CRISPR

Se ha demostrado que determinados ortólogos de ARNasas guiadas por ARN de CRISPR, tales como LshC2c2, no distinguen fácilmente apareamientos erróneos de una sola base (Abudayyeh *et al.*, 2016). Tal como se demuestra

en el presente documento, LwC2c2 también comparte esta característica (figura 37^a). Para aumentar la especificidad de la escisión de LwC2c, se desarrolló un sistema para introducir apareamientos erróneos sintéticos en el dúplex ARNcr:diana que aumenta la sensibilidad total a los apareamientos erróneos y permite una sensibilidad a apareamientos erróneos de una sola base. Se diseñaron múltiples ARNcr para la diana 1 y se incluyeron apareamientos erróneos a lo largo de la longitud del ARNcr (figura 37^a) para optimizar la escisión específica y minimizar la escisión de una diana que difieren en un apareamiento erróneo individual. Estos apareamientos erróneos no redujeron la eficiencia de escisión de la diana de ARNmc 1, pero disminuyeron significativamente la señal para una diana que incluía un apareamiento erróneo adicional (diana de ARNmc 2). El ARNcr diseñado que mejor distinguía entre las dianas 1 y 2 incluía apareamientos erróneos sintéticos cercanos al apareamiento erróneo de la diana 2, creando de hecho una “burbuja”. La pérdida de sensibilidad provocada por la coordinación de un apareamiento erróneo sintético y un apareamiento erróneo adicional presente en la diana (es decir, un apareamiento erróneo doble) concuerda con la sensibilidad de LshC2c2 a apareamientos erróneos dobles consecutivos o cercanos (Abudayyeh *et al.*, 2016) y presenta un formato para el diseño racional de ARNcr que permiten la distinción de un solo nucleótido (figura 37B).

Habiendo demostrado que C2c2 modificarse por ingeniería para reconocer apareamientos erróneos de una sola base, se determinó si esta especificidad modificada por ingeniería podría usarse para distinguir entre patógenos virales estrechamente relacionados. Se diseñaron múltiples ARNcr para detectar la cepa o bien africana o bien americana del virus de Zika (figura 37^a) y la cepa 1 o 3 del virus del dengue (figura 37C). Estos ARNcr incluían un apareamiento erróneo sintético en la secuencia espaciadora, provocando la formación de una única burbuja cuando se formó un dúplex con la cepa específica debido al apareamiento erróneo sintético. Sin embargo, cuando el espaciador de apareamiento erróneo sintético forma un dúplex con la cepa inespecífica, se forman dos burbujas debido al apareamiento erróneo sintético y al apareamiento erróneo de SNP. Los ARNcr de apareamiento erróneo sintético detectaron sus cepas correspondientes con una señal significativamente más alta que la cepa inespecífica, lo que permite una distinción robusta de las cepas (figura 37B, 37D). Debido a la similitud de secuencia significativa entre cepas, no fue posible hallar un tramo continuo de 28 nt con una única diferencia de un solo nucleótido entre los dos genomas con el fin de demostrar una verdadera distinción de cepas de un solo nucleótido. Sin embargo, se predijo que los ARNcr más cortos seguirían siendo funcionales, como lo son con LshC2c2 (Abudayyeh *et al.*, 2016), y, por consiguiente, se diseñaron ARNcr de 23 nt más cortos contra dianas en las dos cepas de Zika que incluían un apareamiento erróneo sintético en la secuencia espaciadora y un único apareamiento erróneo en la secuencia diana. Estos ARNcr todavía podían distinguir las cepas africana y americana de Zika con alta sensibilidad (figura 36C).

Genotipado rápido usando ADN purificado a partir de saliva

El genotipado rápido a partir de saliva humana podría ser útil en situaciones farmacogenómicas de emergencia o para diagnósticos domiciliarios. Para demostrar el potencial de las realizaciones dadas a conocer en el presente documento para el genotipado, se eligieron cinco loci para la detección distintiva de C2c2 usando datos de genotipado de 23andMe como patrón de referencia (Eriksson *et al.*, 2010) (figura 38^a). Los cinco loci abarcan una amplia gama de asociaciones funcionales, incluyendo la sensibilidad a fármacos tales como estatinas o acetaminofeno, la susceptibilidad a norovirus y el riesgo de cardiopatía (tabla 7).

Tabla 7: Variantes de SNP sometidas a prueba

ID	Gen	Categoría
rs5082	APOA2	Consumo de grasas saturadas y aumento de peso
rs1467558	CD44	Metabolismo del acetaminofeno
rs2952768	Cercano a CREB1	Dependencia de la morfina
rs4363657	SLCO1B1	Aumento de 4,5 veces en el riesgo de miopatía para usuarios de estatinas
rs601338	FUT2	Resistencia a norovirus

Se recogió saliva de cuatro sujetos humanos y se purificó el ADN genómico usando un kit comercial sencillo en menos de una hora. Los cuatro sujetos tenían un conjunto diverso de genotipos a lo largo de los cinco loci, lo que proporciona un espacio muestral suficientemente amplio con el que comparar el ensayo de genotipado. Para cada uno de los cinco loci de SNP, se amplificó el ADN genómico de un sujeto usando RPA con los cebadores apropiados seguido de la detección con LwC2c2 y pares de ARNcr diseñados para detectar específicamente uno de los dos alelos posible (figura 38B). El ensayo era lo suficientemente específico como para distinguir los alelos con alta significación e inferir genotipos tanto homocigóticos como heterocigóticos. Debido a que el protocolo de extracción de ADN se realizó sobre la saliva antes de la detección, se sometió a prueba el ensayo para determinar si puede hacerse incluso más susceptible para el genotipado de POC mediante el uso de saliva calentada hasta 95°C durante 5 minutos sin ninguna extracción adicional. El ensayo pudo genotipar correctamente dos pacientes cuya saliva se sometió sólo a calentamiento durante 5 minutos y luego a la posterior amplificación y detección con C2c2 (figura 40B).

Detección de mutaciones cancerosas en ADNlc en fracciones alélicas bajas

Debido a que el ensayo es altamente específico para diferencias de un solo nucleótido en las dianas, se ideó una prueba para determinar si el ensayo era lo suficientemente sensible como para detectar mutaciones cancerosas en ADN libre de células (ADNlc). Los fragmentos de ADNlc tienen un pequeño porcentaje (del 0,1% al 5%) de fragmentos de ADNlc silvestre (Bettegowda *et al.*, 2014; Newman *et al.*, 2014; Olmedillas Lopez *et al.*, 2016; Qin *et al.*, 2016). Un desafío significativo en el campo del ADNlc es detectar estas mutaciones porque normalmente son difíciles de descubrir dados los altos niveles de ADN no mutado hallado en el fondo en sangre (Bettegowda *et al.*, 2014; Newman *et al.*, 2014; Qin *et al.*, 2016). Una prueba para el cáncer de ADNlc de POC también sería útil para el examen regular de la presencia de cáncer, especialmente para pacientes en riesgo de remisión.

Se determinó la capacidad del ensayo para detectar ADN mutante en el fondo silvestre mediante la dilución de diana de ADNbc 1 en un fondo de ADNmc 1 con una única mutación en el sitio de diana de ARNcr (figura 41^a-B). LwC2c2 puede detectar ADNbc 1 a niveles de tan sólo el 0,1% del ADNbc de fondo y dentro de concentraciones attomolares de ADNbc 1. Este resultado muestra que la escisión de LwC2c2 de ADNbc 1 mutante de fondo es lo suficientemente baja como para permitir una detección robusta del ADNbc específico en la fracción alélica del 0,1%. A niveles inferiores al 0,1%, la actividad de fondo es probablemente un problema, lo que impide cualquier detección significativa adicional de la diana correcta.

Debido a que el ensayo podía detectar dianas sintéticas con fracciones alélicas en una gama clínicamente relevante, se evaluó si el ensayo podía detectar mutaciones cancerosas en ADNlc. Se diseñaron cebadores de RPA para dos mutaciones cancerosas diferentes, EGFR L858R y BRAF V600E, y se usaron patrones de ADNlc comerciales con fracciones alélicas del 5%, del 1% y del 0,1% que se asemejan a muestras de ADNlc humano reales para la prueba. Usando un par de ARNcr que podían distinguir el alelo mutante del alelo silvestre (figura 38C), se logró la detección de la fracción alélica del 0,1% para ambos loci mutantes (figura 39 A-B).

Discusión

Mediante la combinación de las propiedades naturales de C2c2 con la amplificación isotérmica y una sonda fluorescente extinguida, se ha demostrado que el ensayo y los sistemas dados a conocer en el presente documento son un método versátil y robusto para detectar ARN y ADN y adecuados para una variedad de diagnósticos rápidos incluyendo aplicaciones de enfermedades infecciosas y genotipado rápido. La ventaja principal de los ensayos y sistemas dados a conocer en el presente documento es que puede rediseñarse y sintetizarse una nueva prueba de POC en cuestión de días por tan solo 0,6 USD/prueba.

Debido a que muchas aplicaciones de enfermedades humanas requieren la capacidad para detectar apareamientos erróneos individuales, se desarrolló un enfoque racional para modificar por ingeniería ARNcr para que fueran altamente específicos para un apareamiento erróneo individual en la secuencia diana mediante la introducción de un apareamiento erróneo sintético en la secuencia espaciadora del ARNcr. Otros enfoques para lograr especificidad con efectores CRISPR se basan en métodos basados en examen de docenas de diseños de guías (Chavez *et al.*, 2016). Usando ARNcr diseñados con apareamientos erróneos, se demostró la discriminación de cepas virales de Zika y dengue en sitios que difieren en un apareamiento erróneo individual, el genotipado rápido de SNP a partir de ADNg de saliva humana y la detección de mutaciones cancerosas en muestras de ADNlc.

El bajo coste y la adaptabilidad de la plataforma de ensayo se presta a otras aplicaciones incluyendo (i) la experiencia de cuantificación de ARN/ADN general en sustitución de ensayos de qPCR específicos, tales como Taqman, (ii) la detección de la expresión de ARN rápida y multiplexada que se asemeja a los microalineamientos y (iii) otras aplicaciones de detección sensibles, tales como la detección de la contaminación de ácidos nucleicos a partir de otras fuentes en alimentos. Además, C2c2 puede usarse potencialmente para la detección de transcritos dentro de entornos biológicos, tal como en células, y dada la naturaleza altamente específica de la detección con C2c2, puede ser posible rastrear la expresión específica alélica de transcritos o mutaciones asociadas a enfermedad en células vivas. Con la amplia disponibilidad de aptámeros, también puede ser posible detectar proteínas mediante el acoplamiento de la detección de la proteína mediante un aptámero con la revelación de un sitio de amplificación críptico para RPA seguido de la detección con C2c2.

Detección de ácidos nucleicos con CRISPR-Cas13a/C2c2: sensibilidad attomolar y especificidad de un solo nucleótido

Para lograr una detección robusta de señales, el solicitante identificó un ortólogo de Cas13a de *Leptotrichia wadei* (LwCas13a), que presenta una mayor actividad ARNasa guiada por ARN en relación con Cas13a de *Leptotrichia shahii* (LshCas13a) (10) (figura 2, véase también "Caracterización de los requisitos de escisión de LwCas13a" anteriormente). LwCas13a incubada con diana de ARNmc 1 (ARNmc 1), ARNcr e indicador (ARN fluorescente extinguido) (figura 18) (13) produjo una sensibilidad de detección de ~50 fM (figura 51, 15), que no es lo suficientemente sensible para muchas aplicaciones de diagnóstico (12, 14-16). Por tanto, el solicitante exploró combinar la detección basada en Cas13a con diferentes etapas de amplificación isotérmica (figura 10, 11, 53, 16) (17, 18). De los métodos explorados, la amplificación por recombinasa y polimerasa (RPA) (18) proporcionó la mayor sensibilidad y puede acoplarse con la transcripción con T7 para convertir ADN amplificado en ARN para su posterior

detección mediante LwCas13a (véase también “Discusión de la amplificación por recombinasa y polimerasa (RPA) y otras estrategias de amplificación isotérmica” anteriormente). El solicitante se refiere a esta combinación de amplificación mediante RPA, transcripción con ARN polimerasa T7 de ADN amplificado en ARN y detección de ARN diana mediante liberación de señal indicadora mediada por la escisión de ARN colateral de Cas13a como SHERLOCK.

En primer lugar, el solicitante determinó la sensibilidad de SHERLOCK para la detección de dianas de ARN (cuando se acopla con transcripción inversa) o de ADN. El solicitante logró una sensibilidad de una sola molécula tanto para ARN como para ADN, tal como se verifica mediante PCR digital en gotas (ddPCR) (figura 27, 51, 54^a, B). Se mantuvo una sensibilidad attomolar cuando se combinaron todos los componentes de SHERLOCK en una única reacción, lo que demuestra la viabilidad de esta plataforma como diagnóstico inmediato (POC) (figura 54C). SHERLOCK tiene niveles de sensibilidad similares a la ddPCR y a la PCR cuantitativa (qPCR), dos enfoques de detección de ácidos nucleicos sensibles establecidos, mientras que la RPA sola no era lo suficientemente sensible como para detectar bajos niveles de diana (figura 55^a-D). Además, SHERLOCK muestra menos variación que ddPCR, qPCR y RPA, tal como se mide mediante el coeficiente de variación a lo largo de las réplicas (figura 55E-F).

A continuación, el solicitante examinó si SHERLOCK sería eficaz en aplicaciones de enfermedades infecciosas que requieren una alta sensibilidad. El solicitante produjo lentivirus que albergaban fragmentos de genoma o bien de virus de Zika (ZIKV) o bien del dengue de flavivirus relacionados (DENV) (19) (figura 31^a). SHERLOCK detectó partículas virales hasta 2 aM y pudo discriminar entre ZIKV y DENV (figura 31B). Para explorar el posible uso de SHERLOCK en el campo, el solicitante demostró en primer lugar que los complejos Cas13a-ARNcr liofilizados y posteriormente rehidratados (20) podían detectar 20 fM de ARNmc 1 no amplificado (figura 33^a) y que también fue posible la detección de la diana en papel de fibra de vidrio (figura 33B). Los otros componentes de SHERLOCK también son susceptibles de secado por congelación: RPA se proporciona como reactivo liofilizado a temperatura ambiental, y el solicitante demostró previamente que la polimerasa T7 tolera la liofilización (2). En combinación, el secado por congelación y el manchado sobre papel de la reacción de detección con Cas13a dieron como resultado niveles comparables de detección sensible de ARNmc 1 a las reacciones acuosas (figura 33C-E). Aunque el manchado sobre papel y la liofilización redujeron ligeramente la señal absoluta de la lectura, SHERLOCK (figura 31C) pudo detectar fácilmente virus ZIKV simulado a concentraciones de tan sólo 20 aM (figura 31D). SHERLOCK también puede detectar ZIKV en aislados clínicos (suero, orina o saliva) en los que los títulos pueden ser de tan sólo 2×10^3 copias/ml (3,2 aM) (21). Pudo detectarse el ARN de ZIKV extraído de muestras de suero u orina de paciente y sometido a transcripción inversa en ADNc (figuras 32E y 52^a) a concentraciones inferiores a $1,25 \times 10^3$ copias/ml (2,1 aM), tal como se verifica mediante qPCR (figuras 32F y 52B). Además, la señal de las muestras de paciente fue predictiva del número de copias de ARN de ZIKV y pudo usarse para predecir la carga viral (figura 33F). Para simular la detección de muestras sin purificación de ácidos nucleicos, el solicitante midió la detección de ARNmc 1 añadido a suero humano y halló que Cas13a podía detectar el ARN en reacciones que contenían hasta el 2% de suero (figura 33G). Otra aplicación epidemiológica importante para las realizaciones dadas a conocer en el presente documento es la identificación de patógenos bacterianos y la detección de genes bacterianos específicos. El solicitante seleccionó como diana la región V3 del gen ARNr 16S, en la que las regiones flanqueantes conservadas permiten el uso de cebadores de RPA universales entre especies bacterianas y la región interna variable permite la diferenciación de especies. En un panel de cinco ARNcr de direccionamiento posibles para diferentes cepas patógenas y ADNg aislados de *E. coli* y *Pseudomonas aeruginosa* (figura 34^a), SHERLOCK genotipó correctamente las cepas y mostró una baja reactividad cruzada (figura 34B). Además, el solicitante pudo usar SHERLOCK para distinguir entre aislados clínicos de *Klebsiella pneumoniae* con dos genes de resistencia diferentes: carbapenemasa de *Klebsiella pneumoniae* (KPC) y metalo-beta-lactamasa 1 de Nueva Delhi (NDM-1) (22) (figura 56).

Para aumentar la especificidad de SHERLOCK, el solicitante introdujo apareamientos erróneos sintéticos en el dúplex ARNcr:diana que permiten que LwCas13a discrimine entre dianas que difieren en un apareamiento erróneo de una sola base (figura 36^a,B; véase también “Diseño de apareamientos erróneos modificados por ingeniería” anteriormente). El solicitante diseñó múltiples ARNcr con apareamientos erróneos sintéticos en las secuencias espaciadoras para detectar la cepa o bien africana o bien americana de ZIKV (la figura 37^a) y la cepa 1 o 3 de DENV (figura 37C). El ARNcr con apareamientos erróneos sintéticos detectó sus cepas correspondientes con una señal significativamente mayor (prueba de la t de Student bidireccional; $p < 0,01$) que la cepa inespecífica, lo que permite una discriminación robusta de cepas basándose en apareamientos erróneos individuales (figura 37B, D, 36C). La caracterización adicional reveló que la detección con Cas13a logra una especificidad máxima al tiempo que mantiene una sensibilidad específica cuando una mutación está en la posición 3 del espaciador y el apareamiento erróneo sintético está en la posición 5 (figuras 57 y 58). La capacidad para detectar diferencias de una sola base abre la oportunidad de usar SHERLOCK para el genotipado humano rápido. El solicitante eligió cinco loci que abarcaban una gama de polimorfismos de un solo nucleótido (SNP) relacionados con la salud (tabla 1) y comparó la detección con SHERLOCK usando datos de genotipado de 23andMe como patrón de referencia en estos SNP (23) (figura 38^a). El solicitante recogió saliva de cuatro sujetos humanos con diversos genotipos a lo largo de los loci de interés y extrajo ADN genómico o bien a través de purificación en columna comercial o bien mediante calentamiento directo durante cinco minutos (20). SHERLOCK distinguió los alelos con alta significación y con especificidad suficiente para inferir genotipos tanto homocigóticos como heterocigóticos (figuras 38B, 40, 59, 60; véase también “Genotipado con SHERLOCK usando patrones sintéticos” anteriormente). Finalmente, el solicitante buscó determinar si SHERLOCK podía detectar mutaciones cancerosas de baja frecuencia en fragmentos de ADN libre de

células (1c), lo que supone un desafío debido a los altos niveles de ADN silvestre en sangre de paciente (24-26). El solicitante halló por primera vez que SHERLOCK podía detectar ADNmc 1 a concentraciones attomolares diluidas en un fondo de ADN genómico (figura 61). A continuación, el solicitante halló que SHERLOCK también podía detectar alelos que contienen polimorfismos de un solo nucleótido (SNP) (figura 41^a, B) a niveles de tan sólo el 0,1% de ADN de fondo, que se encuentra en el intervalo clínicamente relevante. El solicitante demostró entonces que SHERLOCK podía detectar dos mutaciones cancerosas diferentes, EGFR L858R y BRAF V600E, en muestras de ADNlc simulado con fracciones alélicas de tan sólo el 0,1% (figuras 38, 39) (20).

La plataforma SHERLOCK se presta a otras aplicaciones incluyendo (i) la cuantificación de ARN/ADN general en lugar de ensayos de qPCR específicos, tales como TaqMan, (ii) la detección de la expresión de ARN rápida y multiplexada y (iii) otras aplicaciones de detección sensibles, tales como detección de la contaminación de ácidos nucleicos. Además, Cas13a podría detectar posiblemente transcritos dentro de entornos biológicos y rastrear la expresión específica de alelo de transcritos o mutaciones asociadas a enfermedad en células vivas. SHERLOCK es un método versátil y robusto para detectar ARN y ADN, adecuado para diagnósticos rápidos incluyendo aplicaciones de enfermedades infecciosa y genotipado sensible. Una prueba en papel de SHERLOCK puede rediseñarse y sintetizarse en cuestión de días por tan solo 0,61 USD/prueba (véase también "SHERLOCK es una plataforma de CRISPR-Dx asequible y adaptable" anteriormente) con confianza, ya que casi todos los ARNcr sometidos a prueba dieron como resultado una sensibilidad y una especificidad altas. Estas cualidades destacan el poder de CRISPR-Dx y abren nuevas vías para la detección rápida, robusta y sensible de moléculas biológicas.

Tabla 8: Cebadores de RPA usados

Nombre	Secuencia	1ª figura
RP0683 - ADNmc/ARNmc 1 de RPA F	(SEQ. ID. NO. 18)	Figura 27B
RP0684 - ADNmc/ARNmc 1 de RPA R	(SEQ. ID. NO. 19)	Figura 27B
AMPL-25 Zika 8B long-rpa3-f	(SEQ. ID. NO. 20)	Figura 31B
AMPL-26 Zika 8B long-rpa3-r	(SEQ. ID. NO. 21)	Figura 31B
RP819 - región 8 de Zika F	(SEQ. ID. NO. 22)	Figura 31C
RP821 - región 8 de Zika R	(SEQ. ID. NO. 23)	Figura 31C
V3 bacteriana de 517 F	(SEQ. ID. NO. 24)	Figura 34B
V3 bacteriana de RP758 R	(SEQ. ID. NO. 25)	Figura 34B
wR0074 A2 rs5082 F	(SEQ. ID. NO. 26)	Figura 38B
wR0074 E2 rs5082 R	AA (SEQ. ID. NO. 27)	Figura 38B
wR0074 A4 rs1467558 F	(SEQ. ID. NO. 28)	Figura 38B
wR0074 E4 rs1467558 R	(SEQ. ID. NO. 29)	Figura 38B
wR0074 A5 rs2952768 F	(SEQ. ID. NO. 30)	Figura 38B
wR0074 E5 rs2952768 R	(SEQ. ID. NO. 31)	Figura 38B
wR0074 A9 rs4363657 F	(SEQ. ID. NO. 32)	Figura 38B
wR0074 E9 rs4363657 R	(SEQ. ID. NO. 33)	Figura 38B
wR0074 A11 rs601338 F	(SEQ. ID. NO. 34)	Figura 38B
wR0074 E11 rs601338 R	(SEQ. ID. NO. 35)	Figura 38B
ADNlc de RP824 BRAFV600E F	(SEQ. ID. NO. 36)	Figura 39A
ADNlc de RP769 BRAFV600E R	(SEQ. ID. NO. 37)	Figura 39A
ADNlc de RP826 EGFR858R F	(SEQ. ID. NO. 38)	Figura 39B
ADNlc de RP804 EGFR858R R	(SEQ. ID. NO. 39)	Figura 39B
AMPL-31 T1-nasba1-f	(SEQ. ID. NO. 40)	Figura 11
AMPL-32 T1-nasba1-r	(SEQ. ID. NO. 41)	Figura 11
AMPL-33 T1-nasba2-f	(SEQ. ID. NO. 42)	Figura 11
AMPL-34 T1-nasba2-r	(SEQ. ID. NO. 43)	Figura 11
AMPL-35 T1-nasba3-f	(SEQ. ID. NO. 44)	Figura 11
AMPL-36 T1-nasba3-r	(SEQ. ID. NO. 45)	Figura 11
wR0075 A1 KPC F	(SEQ. ID. NO. 46)	Figura 35A
wR0075 B1 KPC R	(SEQ. ID. NO. 47)	Figura 35A
wR0075 A3 NDM F	(SEQ. ID. NO. 48)	Figura 35A
wR0075 B3 NDMR	(SEQ. ID. NO. 49)	Figura 35A

Tabla 9: Secuencias de ARNcr usadas

Nombre	Secuencia de ARNcr completa	Secuencia espaciadora	1ª figura	PFS
ARNcr de diana 1	(SEQ. ID. NO. 50)	(SEQ. ID. NO. 51)	Figura 2F	C
ARNcr que selecciona como diana Zika 1	(SEQ. ID. NO. 52)	(SEQ. ID. NO. 53)	Figura 31A	U
ARNcr que selecciona como diana Zika 2	(SEQ. ID. NO. 54)	(SEQ. ID. NO. 55)	Figura 33D	G
ARNcr de detección de <i>E. coli</i>	(SEQ. ID. NO. 56)	(SEQ. ID. NO. 57)	Figura 22B	U
ARNcr de detección de <i>K. pneumoniae</i>	(SEQ. ID. NO. 58)	(SEQ. ID. NO. 59)	Figura 34B	U
ARNcr de detección de <i>P. aeruginosa</i>	(SEQ. ID. NO. 60)	(SEQ. ID. NO. 61)	Figura 34B	U
ARNcr de detección de <i>M. tuberculosis</i>	(SEQ. ID. NO. 62)	(SEQ. ID. NO. 63)	Figura 34B	U
ARNcr de detección de <i>S. aureus</i>	(SEQ. ID. NO. 64)	(SEQ. ID. NO. 65)	Figura 34B	G
ARNcr de KPC	(SEQ. ID. NO. 66)	(SEQ. ID. NO. 67)	Figura 35A	U
ARNcr de NDM	(SEQ. ID. NO. 68)	(SEQ. ID. NO. 69)	Figura 35A	C
ARNcr de apareamiento erróneo 1	(SEQ. ID. NO. 70)	(SEQ. ID. NO. 71)	Figura 36A	C
ARNcr de apareamiento erróneo 2	(SEQ. ID. NO. 72)	(SEQ. ID. NO. 73)	Figura 36A	C
ARNcr de apareamiento erróneo 3	(SEQ. ID. NO. 74)	(SEQ. ID. NO. 75)	Figura 36A	C
ARNcr de apareamiento erróneo 4	(SEQ. ID. NO. 76)	(SEQ. ID. NO. 77)	Figura 36A	C
ARNcr de apareamiento erróneo 5	(SEQ. ID. NO. 78)	(SEQ. ID. NO. 79)	Figura 36A	C
ARNcr de apareamiento erróneo 6	(SEQ. ID. NO. 80)	(SEQ. ID. NO. 81)	Figura 36A	C
ARNcr de apareamiento erróneo 7	(SEQ. ID. NO. 82)	(SEQ. ID. NO. 83)	Figura 36A	C
ARNcr de apareamiento erróneo 8	(SEQ. ID. NO. 84)	(SEQ. ID. NO. 85)	Figura 36A	C
ARNcr de apareamiento erróneo 9	(SEQ. ID. NO. 86)	(SEQ. ID. NO. 87)	Figura 36A	C
ARNcr de apareamiento erróneo 10	(SEQ. ID. NO. 88)	(SEQ. ID. NO. 89)	Figura 36A	C
ARNcr africano 1	(SEQ. ID. NO. 90)	(SEQ. ID. NO. 91)	Figura 38A	C
ARNcr africano 2	(SEQ. ID. NO. 92)	(SEQ. ID. NO. 93)	Figura 38A	C
ARNcr americano 1	(SEQ. ID. NO. 94)	(SEQ. ID. NO. 95)	Figura 38A	U
ARNcr americano 2	(SEQ. ID. NO. 96)	(SEQ. ID. NO. 97)	Figura 38A	U
ARNcr de la cepa 3 del dengue 1	(SEQ. ID. NO. 98)	(SEQ. ID. NO. 99)	Figura 38C	A
ARNcr de la cepa 3 del dengue 2	(SEQ. ID. NO. 100)	(SEQ. ID. NO. 101)	Figura 38C	A
ARNcr de la cepa 1 del dengue 1	(SEQ. ID. NO. 102)	(SEQ. ID. NO. 103)	Figura 38C	A
ARNcr de la cepa 1 del dengue 2	(SEQ. ID. NO. 104)	(SEQ. ID. NO. 105)	Figura 38C	A
ARNcr africano más corto 1	(SEQ. ID. NO. 106)	(SEQ. ID. NO. 107)	Figura 36C	C
ARNcr africano más corto 2	(SEQ. ID. NO. 108)	(SEQ. ID. NO. 109)	Figura 36C	C
ARNcr americano más corto 1	(SEQ. ID. NO. 110)	(SEQ. ID. NO. 111)	Figura 36C	U
ARNcr americano más corto 2	(SEQ. ID. NO. 112)	(SEQ. ID. NO. 113)	Figura 36C	U
ARNcr de rs1467558 C	(SEQ. ID. NO. 114)	(SEQ. ID. NO. 115)	Figura 38B	C
ARNcr de rs1467558 T	(SEQ. ID. NO. 116)	(SEQ. ID. NO. 117)	Figura 38B	C
ARNcr de rs2952768 C	(SEQ. ID. NO. 118)	(SEQ. ID. NO. 119)	Figura 38B	A
ARNcr de rs2952768 T	(SEQ. ID. NO. 120)	(SEQ. ID. NO. 121)	Figura 38B	A
ARNcr de rs4363657 C	(SEQ. ID. NO. 122)	(SEQ. ID. NO. 123)	Figura 38B	A
ARNcr de rs4363657 T	(SEQ. ID. NO. 124)	(SEQ. ID. NO. 125)	Figura 38B	A
ARNcr de rs601338 A	(SEQ. ID. NO. 126)	(SEQ. ID. NO. 127)	Figura 38B	G

ARNcr de rs601338 G	(SEQ. ID. NO. 128)	(SEQ. ID. NO. 129)	Figura 38B	G
ARNcr de rs5082 G	(SEQ. ID. NO. 130)	(SEQ. ID. NO. 131)	Figura 40A	A
ARNcr de rs5082 A	(SEQ. ID. NO. 132)			A
ARNcr silvestre de EGFR L858R	(SEQ. ID. NO. 134)	(SEQ. ID. NO. 135)	Figura 38C	C
ARNcr mutante de EGFR L858R	(SEQ. ID. NO. 136)	(SEQ. ID. NO. 137)	Figura 38C	C
ARNcr silvestre de BRAF V600E	(SEQ. ID. NO. 138)	(SEQ. ID. NO. 139)	Figura 38C	A
ARNcr mutante de BRAF V600E	(SEQ. ID. NO. 140)	(SEQ. ID. NO. 141)	Figura 38C	A
ARNcr de apareamiento erróneo de 23 nt 1	(SEQ. ID. NO. 303)	(SEQ. ID. NO. 304)	Figura 57D	C
ARNcr de apareamiento erróneo de 23 nt 2	(SEQ. ID. NO. 305)	(SEQ. ID. NO. 306)	Figura 57D	C
ARNcr de apareamiento erróneo de 23 nt 4	(SEQ. ID. NO. 307)	(SEQ. ID. NO. 308)	Figura 57D	C
ARNcr de apareamiento erróneo de 23 nt 5	(SEQ. ID. NO. 234)	(SEQ. ID. NO. 235)	Figura 57D	C
ARNcr de apareamiento erróneo de 23 nt 6	(SEQ. ID. NO. 236)	(SEQ. ID. NO. 237)	Figura 57D	C
ARNcr de apareamiento erróneo de 23 nt 7	(SEQ. ID. NO. 238)	(SEQ. ID. NO. 239)	Figura 57D	C
ARNcr de apareamiento erróneo de 20 nt 1	(SEQ. ID. NO. 240)	(SEQ. ID. NO. 241)	Figura 57F	C
ARNcr de apareamiento erróneo de 20 nt 2	(SEQ. ID. NO. 242)	(SEQ. ID. NO. 243)	Figura 57F	C
ARNcr de apareamiento erróneo de 20 nt 4	(SEQ. ID. NO. 244)	(SEQ. ID. NO. 245)	Figura 57F	C
ARNcr de apareamiento erróneo de 20 nt 5	(SEQ. ID. NO. 246)	(SEQ. ID. NO. 247)	Figura 57F	C
ARNcr de apareamiento erróneo de 20 nt 6	(SEQ. ID. NO. 248)	(SEQ. ID. NO. 249)	Figura 57F	C
ARNcr de apareamiento erróneo de 20 nt 7	(SEQ. ID. NO. 250)	(SEQ. ID. NO. 251)	Figura 57F	C
ARNcr de apareamiento erróneo 1 de diana con apareamiento erróneo 4	(SEQ. ID. NO. 252)	(SEQ. ID. NO. 253)	Figura 58B	C
ARNcr de apareamiento erróneo 2 de diana con apareamiento erróneo 4	(SEQ. ID. NO. 254)	(SEQ. ID. NO. 255)	Figura 58B	C
ARNcr de apareamiento erróneo 3 de diana con apareamiento erróneo 4	(SEQ. ID. NO. 256)	(SEQ. ID. NO. 257)	Figura 58B	C
ARNcr de apareamiento erróneo 5 de diana con apareamiento erróneo 4	(SEQ. ID. NO. 258)	(SEQ. ID. NO. 259)	Figura 58B	C
ARNcr de apareamiento erróneo 6 de diana con apareamiento erróneo 4	(SEQ. ID. NO. 260)	(SEQ. ID. NO. 261)	Figura 58B	C
ARNcr de apareamiento erróneo 7 de diana con apareamiento erróneo 4	(SEQ. ID. NO. 262)	(SEQ. ID. NO. 263)	Figura 58B	C
ARNcr de apareamiento erróneo 2 de diana con apareamiento erróneo 5	(SEQ. ID. NO. 264)	(SEQ. ID. NO. 265)	Figura 58B	C
ARNcr de apareamiento erróneo 3 de diana con apareamiento erróneo 5	(SEQ. ID. NO. 266)	(SEQ. ID. NO. 267)	Figura 58B	C
ARNcr de apareamiento erróneo 4 de diana con apareamiento erróneo 5	(SEQ. ID. NO. 268)	(SEQ. ID. NO. 269)	Figura 58B	C

ARNcr de apareamiento erróneo 6 de diana con apareamiento erróneo 5	(SEQ. ID. NO. 270)	(SEQ. ID. NO. 271)	Figura 58B	C
ARNcr de apareamiento erróneo 7 de diana con apareamiento erróneo 5	(SEQ. ID. NO. 272)	(SEQ. ID. NO. 273)	Figura 58B	C
ARNcr de apareamiento erróneo 8 de diana con apareamiento erróneo 5	(SEQ. ID. NO. 274)	(SEQ. ID. NO. 275)	Figura 58B	C
ARNcr de apareamiento erróneo 3 de diana con apareamiento erróneo 6	(SEQ. ID. NO. 276)	(SEQ. ID. NO. 277)	Figura 58B	C
ARNcr de apareamiento erróneo 4 de diana con apareamiento erróneo 6	(SEQ. ID. NO. 278)	(SEQ. ID. NO. 279)	Figura 58B	C
ARNcr de apareamiento erróneo 5 de diana con apareamiento erróneo 6	(SEQ. ID. NO. 280)	(SEQ. ID. NO. 281)	Figura 58B	C
ARNcr de apareamiento erróneo 7 de diana con apareamiento erróneo 6	(SEQ. ID. NO. 282)	(SEQ. ID. NO. 283)	Figura 58B	C
ARNcr de apareamiento erróneo 8 de diana con apareamiento erróneo 6	(SEQ. ID. NO. 284)	(SEQ. ID. NO. 285)	Figura 58B	C
ARNcr de apareamiento erróneo 9 de diana con apareamiento erróneo 6	(SEQ. ID. NO. 286)	(SEQ. ID. NO. 287)	Figura 58B	C

Tabla 10: Dianas de ARN y ADN usadas en este ejemplo

Nombre	Secuencia	1ª figura
ARNmc 1 (C PFS)	(SEQ. ID. NO. 288)	Figura 2F
ARNmc 1 (G PFS)	(SEQ. ID. NO. 289)	Figura 2F
ARNmc 1 (A PFS)	(SEQ. ID. NO. 290)	Figura 2F
ARNmc 1 (U PFS)	(SEQ. ID. NO. 291)	Figura 2F
ADNmc 1	(SEQ. ID. NO. 292)	Figura 27
ADN 2	(SEQ. ID. NO. 293)	Figura 54B
ZIKV en lentivirus	(SEQ. ID. NO. 294)	Figura 31B
DENV en lentivirus	(SEQ. ID. NO. 295)	Figura 31B
Diana de ZIKV sintética	(SEQ. ID. NO. 296)	Figura 33D
Diana de ZIKV africana sintética	(SEQ. ID. NO. 297)	Figura 37A
Diana de ZIKV americana sintética	(SEQ. ID. NO. 298)	Figura 37A
Diana de cepa 1 del dengue sintética	(SEQ. ID. NO. 299)	Figura 37C
Diana de cepa 3 del dengue sintética	(SEQ. ID. NO. 300)	Figura 37C
ARNmc 2	(SEQ. ID. NO. 301)	Figura 36A
ARNmc 3	(SEQ. ID. NO. 302)	Figura 36A

Tabla 11: Plásmidos usados en este ejemplo

Nombre de plásmido	Descripción	Enlace a mapa de plásmidos
pC004	Diana de examen de beta-lactamasa	https://benchling.com/s/IPJ1cCwR
pC009	Locus de LshCas13a en pACYC184 con espaciador de direccionamiento	https://benchling.com/s/seqylkMuglYmiG4A3VhShZg
pC010	Locus de LshCas13a en pACYC184 con espaciador de no direccionamiento	https://benchling.com/s/seq-2WApFr3zni1GOACyQY8a

pC011	Locus de LwCas13a en pACYC184 con espaciador de direccionamiento	https://benchling.com/s/seq-Vyk8qK2fyhzegfNgLJHM
pC012	Locus de LwCas13a en pACYC184 con espaciador de no direccionamiento	https://benchling.com/s/seq-RxZAgPBzBUGQThkxR2Kx
pC013	Twinstrep-SUMO-huLwCas13a para expresión bacteriana	https://benchling.com/s/seq-66CfLwu7sLMQMbcXe7lh

EJEMPLO 3 - CARACTERIZACIÓN DE ORTÓLOGOS DE Cas13b CON PREFERENCIAS DE BASES ORTOGONALES

El solicitante caracterizó bioquímicamente catorce ortólogos de la familia CRISPR-Cas13b de tipo VI recientemente definida de enzimas que selecciona como diana ARN guiadas por ARN para hallar nuevos candidatos para mejorar la tecnología de detección SHERLOCK (figuras 83A y 85). El solicitante pudo expresar de manera heteróloga catorce ortólogos de Cas13b en *E. coli* y purificar las proteínas para un ensayo de actividad ARNasa *in vitro* (figura 86). Debido a que ortólogos de Cas13 diferentes pueden tener preferencias de bases variables para una actividad de escisión óptima, el solicitante generó sensores homopoliméricos de ARNasa fluorescentes que consistían en 5 o bien A, G, C o bien U para evaluar las preferencias de escisión ortogonal. El solicitante incubó cada ortólogo con su ARNcr relacionado que selecciona como diana un ARNmc 1 de 173 nt sintético y midió la actividad de escisión colateral usando los sensores homopoliméricos fluorescentes (figuras 83B y 87).

EJEMPLO 4 - EXAMEN DE DESCUBRIMIENTO DE MOTIVOS CON BIBLIOTECA

Para explorar adicionalmente la diversidad de preferencias de escisión de los diversos ortólogos de Cas13a y Cas13b, el solicitante desarrolló un enfoque basado en biblioteca para caracterizar los motivos preferidos para la actividad endonucleasa en respuesta a la actividad colateral. El solicitante usó un indicador de ARN de 6 meros degenerado flanqueado por asideros de ADN constantes, lo que permitió la amplificación y la lectura de secuencias no escindidas (figura 83C). La incubación de esta biblioteca con enzimas Cas13 activadas colaterales dio como resultado una escisión detectable y dependía de la adición de ARN diana (figura 88). La secuenciación de motivos empobrecidos reveló un aumento en el sesgo de la biblioteca a lo largo del tiempo de digestión (figura 89A), indicativo de la preferencia de bases, y la selección de secuencias por encima de una razón de umbral produjo secuencias enriquecidas en número que se correspondían con la escisión de las enzimas (figura 89B). Los logotipos de secuencia de motivos enriquecidos reprodujeron la preferencia por U observada para LwaCas13a y CcaCas13b y la preferencia por A de PsmCas13b (figura 89C). El solicitante también determinó múltiples secuencias que mostraron escisión para un único ortólogo, pero no para otros, para permitir una lectura independiente (figura 89D).

Para entender los submotivos específicos de la preferencia de enzimas, el solicitante analizó los motivos empobrecidos para determinar las preferencias por una sola base (figura 90A), lo que coincidía con los motivos homopoliméricos sometidos a prueba así como con los motivos de dos bases (figuras 83C y 90B). Estos motivos de dos bases revelan una preferencia más compleja, especialmente para LwaCas13a y PsmCas13b, que prefiere secuencias de dibases TA, GA y AT. Motivos con órdenes mayores también revelaron preferencias adicionales (figuras 91 y 92).

El solicitante confirmó las preferencias colaterales de LwaCas13a, PsmCas13b y CcaCas13b con la digestión *in vitro* de las dianas (figura 93). Con el fin de mejorar la digestión débil de PsmCas13b, el solicitante optimizó la composición del tampón y la concentración de enzima (figura 94A, B). Otros dicaciones sometidos a prueba en ortólogos de PsmCas13b y Cas13b no tenían grandes efectos (figura 95A-F). El solicitante también comparó PsmCas13b con un miembro de la familia Cas13 con preferencia por A previamente caracterizado para dos dianas de ARN, y halló una sensibilidad comparable o mejorada (figura 96A, B). A partir de estos resultados, el solicitante comparó la cinética de LwaCas13a y PsmCas13b, en reacciones separadas con indicadores independientes, y halló bajos niveles de interferencia entre los dos canales (figura 83D).

EJEMPLO 5 - DETECCIÓN DE UNA SOLA MOLÉCULA CON LwaCas13a, PsmCas13b Y CcaCas13b

Una característica clave de la tecnología SHERLOCK es que permite la detección de una sola molécula (2 aM o 1 molécula/ μ l) mediante la actividad ARNasa colateral de LwaCas13a. Para caracterizar la sensibilidad de las enzimas Cas13b, el solicitante realizó SHERLOCK con PsmCas13b y CcaCas13b, otra enzima Cas13b altamente activa con preferencia por uridina (figura 83E). El solicitante halló que LwaCas13a, PsmCas13b y CcaCas13b podían lograr una detección de 2 aM de dos dianas de ARN diferentes, ARNmc 1 y un ARNmc de Zika sintético (figuras 83E; 97 y 98). Para investigar la robustez de direccionamiento con estas tres enzimas, el solicitante diseñó once ARNcr diferentes uniformemente espaciados a lo largo de ARNmc 1 y halló que LwaCas13a logró la detección de señales de manera más consistente, mientras que CcaCas13b y Psmcas13b mostraron ambas mucha más variabilidad en la detección de ARNcr a ARNcr (figura 99). Para identificar el ARNcr óptimo para la detección, el

solicitante varió la longitud de espaciador de PsmCas13b y CcaCas13b desde 34 hasta 12 nt y halló que PsmCas13b tenía una sensibilidad máxima a una longitud de espaciador de 30 mientras que CcaCas13b tenía una sensibilidad equivalente a longitudes de espaciador superiores a 28 nt (figura 100). El solicitante también sometió a prueba si el límite de detección podía aumentarse por encima de 2 aM, permitiendo entradas de volumen de muestra más grandes en SHERLOCK. Al aumentar de escala la etapa de preamplificación RPA, el solicitante halló que tanto LwaCas13a como PsmCas13b podían proporcionar señales de detección significativas para muestras de entrada de 200, 20 y 2 zM y permitir entradas de volumen de 250 μ l y 540 μ l.

EJEMPLO 6 - SHERLOCK CUANTITATIVO CON RPA

Dado que SHERLOCK se basa en una amplificación exponencial, la cuantificación precisa de ácidos nucleicos puede ser difícil. El solicitante planteó la hipótesis de que reducir la eficiencia de la etapa de RPA podría mejorar la correlación entre la cantidad de entrada y la señal de la reacción de SHERLOCK. El solicitante observó que la cinética de la detección por SHERLOCK era muy sensible a la concentración de cebador a lo largo de un intervalo de concentraciones de muestra (figura 101A-D). El solicitante diluyó las concentraciones de cebador, lo que aumentó tanto la señal como la precisión cuantitativa (figuras 83G y 101E). Esta observación puede deberse a una disminución en la formación de dímeros de cebadores, permitiendo una amplificación más eficaz al tiempo que impide la saturación. Concentraciones de cebador de 120 nM presentaron la correlación más alta entre señal y entrada (figura 101F). Esta precisión era sostenible a lo largo de un gran intervalo de concentraciones hasta el intervalo attomolar (figuras 83H y 101G).

EJEMPLO 7 - MULTIPLEXADO DE DOS COLORES CON ORTÓLOGOS DE Cas13 ORTOGONALES

Una característica ventajosa del diagnóstico de ácidos nucleicos es la capacidad para detectar simultáneamente múltiples entradas de muestra, lo que permite paneles de detección multiplexados o controles en la muestra. Las preferencias de bases ortogonales de las enzimas Cas13 ofrecen la oportunidad de tener SHERLOCK multiplexado. El solicitante puede someter a ensayo la actividad colateral de diferentes enzimas Cas13 en la misma reacción a través de sensores homopoliméricos fluorescentes de diferentes identidades de base y colores fluoróforos, permitiendo la medición simultánea de múltiples dianas (figura 84A). Para demostrar este concepto, el solicitante diseñó un ARNcr de LwaCas13a contra el ARNmc del virus de Zika y un ARNcr de PsmCas13b contra el ARNmc del virus del dengue. El solicitante halló que este ensayo con ambos conjuntos de complejos Cas13-ARNcr en la misma reacción podía identificar si el ARN de Zika o dengue, o ambos, estaban presentes en la reacción (figura 84B). El solicitante también halló que, debido a las preferencias ortogonales entre CcaCas13b y PsmCas13b, estas dos enzimas también podían aprovecharse para la detección multiplexada de dianas de Zika y dengue (figura 102). El solicitante pudo extender con éxito este concepto a toda la reacción de SHERLOCK, que contiene tanto cebadores de RPA multiplexados como complejos Cas13-ARNcr. El solicitante diseñó un ARNcr de LwaCas13a contra *P. aeruginosa* y un ARNcr de PsmCas13b contra *S. aureus* y pudieron detectar dianas de ADN hasta por debajo del intervalo attomolar (figura 84C). De manera similar, usando tanto PsmCas13b como LwaCas13a, el solicitante pudo lograr una detección multiplexada attomolar de ARN de Zika y dengue usando SHERLOCK (figura 103).

El solicitante ha demostrado que LwaCas13a permitió la detección de variantes de un solo nucleótido y que esto podía aplicarse para el genotipado rápido a partir de saliva humana, pero la detección requería dos reacciones independientes: una para cada ARNcr de detección de alelos. Para permitir un genotipado por SHERLOCK de reacción única, el solicitante diseñó un ARNcr de LwaCas13a contra el alelo G y un ARNcr de PsmCas13b contra el alelo A del SNP rs601338, una variante en el gen alfa(1,2)-fucosiltransferasa FUT2 que se asocia con la resistencia a norovirus. Usando este enfoque multiplexado de muestra única, el solicitante pudo genotipar con éxito cuatro sujetos humanos diferentes usando su saliva e identificar con precisión si eran homocigóticos o heterocigóticos.

Para demostrar adicionalmente la versatilidad de la familia de enzimas Cas13, el solicitante simuló un enfoque terapéutico que implica a Cas13 tanto como diagnóstico acompañante como terapia en sí. El solicitante desarrolló recientemente PspCas13b para la edición de transcritos con ARN programable, que puede usarse para la corrección de mutaciones en enfermedades genéticas, usando un sistema denominado edición con ARN para el reemplazo de A por I programable (REPAIR). Debido a que los diagnósticos pueden ser muy útiles cuando se combinan con terapias para guiar las decisiones de tratamiento o monitorizar el resultado de un tratamiento, el solicitante pensó que SHERLOCK podría usarse para el genotipado para guiar el tratamiento con REPAIR y también como lectura en el ARN editado para rastrear la eficiencia de edición de la terapia (figura 84E). El solicitante eligió para demostrar este concepto tratar un trastorno corregir una mutación APC (APC:c.1262G>A) en poliposis adenomatosa familiar 1, un trastorno hereditario que implica cáncer en el intestino grueso y en el recto. El solicitante diseñó ADNc sanos y mutantes del gen APC y los transfeció en células HEK293FT. El solicitante pudo recoger el ADN de estas células y genotipar con éxito las muestras correctas usando SHERLOCK multiplexado de muestra única con LwaCas13a y PsmCas13b (figura 84F). Simultáneamente, el solicitante diseñó y clonó ARN guías para el sistema de REPAIR y transfeció células que tenían el genotipo de enfermedad con el ARN guía y el sistema de REPAIR dPspCas13b-ADAR2_{dd}(E488Q). Después de 48 horas, el solicitante recogió el ARN, que el solicitante dividió para la entrada en SHERLOCK para detectar el resultado de edición y para el análisis por secuenciación de segunda generación (NGS) para confirmar la tasa de edición. La secuenciación reveló que el solicitante logró una edición del 43% con el sistema de REPAIR (figura 84G) y pudo detectarla con SHERLOCK ya que el ARNcr de detección sana mostró una mayor señal que la condición de control guía de no direccionamiento y el ARNcr de detección de enfermedad mostró

una disminución en la señal (figuras 84H y 104). En conjunto, el diseño y la síntesis de reactivos para este ensayo duraron 3 días, el genotipado duró 1 día y la corrección con REPAIR y la detección de la tasa de edición duraron 3 días, produciendo una canalización tratanóstica total que dura sólo 7 días.

El solicitante ha demostrado la detección altamente sensible y específica de ácidos nucleicos usando el ortólogo CRISPR-Cas13a que selecciona como diana ARN guiado por ARN de tipo VI de *Leptotrichia wadei*. El solicitante ha demostrado además que la familia de enzimas Cas13b son bioquímicamente activas y tienen propiedades únicas que las hacen susceptibles de la detección multiplexada de ácidos nucleicos mediante SHERLOCK. Mediante la caracterización de las preferencias de bases ortogonales de las enzimas Cas13b, el solicitante halló secuencias específicas de sensores de ARN fluorescentes que son reconocidos por PsmCas13b que LwaCas13a no reconoce. El solicitante puede aprovecharse de estas preferencias de bases para posibilitar la detección multiplexada en la muestra de dos dianas diferentes y mostrar la utilidad de esta característica para distinguir cepas virales y genotipar a los individuos. Además, a través de la modificación por ingeniería de la etapa de preamplificación, SHERLOCK puede hacerse cuantitativo, lo que permite la aproximación de la cuantificación o concentración de ácido nucleico de entrada. El solicitante ha demostrado además que PsmCas13b ortogonal es capaz de la detección de una sola molécula y que, a través del aumento de escala del volumen, el solicitante puede realizar la detección de muestras de hasta ~0,5 ml y a concentraciones por debajo de 2 zM.

La detección multiplexada con SHERLOCK es posible mediante la realización espacial de múltiples reacciones, pero el multiplexado en la muestra a través de preferencias de bases ortogonales permite la detección de muchas dianas a escala y un menor coste. Aunque el solicitante ha demostrado en este caso un multiplexado de dos entradas, los exámenes de motivos de escisión permiten el diseño de sensores de escisión ortogonal adicionales (figura 90). LwaCas13a y CcaCas13b, que escinden ambas el mismo homopolímero de uridina y, por tanto, no son ortogonales tal como se mide mediante sensores homopoliméricos (figura 83B), mostraron preferencias de escisión muy únicas mediante los exámenes de motivos (figura 90). Mediante el examen de ortólogos de Cas13a, Cas13b y Cas13c adicionales, es probable que muchos ortólogos revelen preferencias de motivos de 6 meros únicas, lo que en teoría podría permitir SHERLOCK altamente multiplexado limitado sólo por el número de sensores fluorescentes espectralmente únicos. SHERLOCK altamente multiplexado permite muchas aplicaciones tecnológicas, especialmente las que implican detección de entrada compleja y computación lógica.

Estos refinamientos adicionales de la detección basada en Cas13 para lecturas visuales, más sensibles y multiplexadas permiten aumentar las aplicaciones para la detección de ácidos nucleicos, especialmente en entornos en los que son necesarios análisis portátiles y sin instrumentación. El genotipado multiplexado rápido pueden informar sobre decisiones farmacogenómicas, someter a prueba múltiples rasgos de cultivo en el campo o evaluar para determinar la presencia de patógenos concurrentes. La lectura isotérmica rápida aumenta la accesibilidad de esta detección para entornos en los que no están disponibles lectores eléctricos o portátiles, incluso para especies raras como el ADN circulante. Las pruebas de ácidos nucleicos basadas en CRISPR mejoradas facilitan el entendimiento de la presencia de ácidos nucleicos en la agricultura, la detección de patógenos y las enfermedades crónicas.

REIVINDICACIONES

1. Un sistema de detección de ácidos nucleicos que comprende:
 reactivos para la amplificación de un ácido nucleico diana;
 un sistema CRISPR de detección que comprende una proteína efectora y uno o más ARN guía diseñados para unirse a los correspondientes ácidos nucleicos diana; y
 un constructo de enmascaramiento basado en ARN,
 en el que dicha proteína efectora es una proteína efectora CRISPR que selecciona como diana ARN de tipo VI.
2. El sistema según la reivindicación 1, en el que el ácido nucleico diana es un ADN diana y el sistema comprende además un cebador que se une al ADN diana y comprende un promotor de la ARN polimerasa.
3. El sistema según una cualquiera de las reivindicaciones 1 a 2, en el que el constructo de enmascaramiento basado en ARN suprime la generación de una señal positiva detectable, opcionalmente
 en el que el constructo de enmascaramiento basado en ARN suprime la generación de una señal positiva detectable enmascarando la señal positiva detectable, o generando en su lugar una señal negativa detectable;
 en el que el constructo de enmascaramiento basado en ARN comprende un ARN de silenciamiento que suprime la generación de un producto génico codificado por un constructo indicador, en el que el producto génico genera la señal positiva detectable cuando se expresa;
 en el que el constructo de enmascaramiento basado en ARN es una ribozima que genera la señal detectable negativa, y en el que la señal detectable positiva se genera cuando la ribozima se desactiva, opcionalmente en el que la ribozima convierte un sustrato en un primer color y en el que el sustrato se convierte en un segundo color cuando la ribozima se desactiva;
 en el que el agente de enmascaramiento basado en ARN es un aptámero de ARN y/o comprende un inhibidor anclado a ARN, opcionalmente en el que el aptámero o inhibidor anclado a ARN secuestra una enzima, en el que la enzima genera una señal detectable tras su liberación del aptámero o inhibidor anclado a ARN actuando sobre un sustrato, o en el que el aptámero es un aptámero inhibidor que inhibe una enzima y evita que la enzima catalice la generación de una señal detectable a partir de un sustrato o en el que el inhibidor anclado a ARN inhibe una enzima y evita que la enzima catalice la generación de una señal detectable a partir de un sustrato, opcionalmente en el que la enzima es trombina, proteína C, elastasa de neutrófilos, subtilisina, peroxidasa del rábano, beta-galactosidasa, o fosfatasa alcalina bovina, opcionalmente en el que la enzima es trombina y el sustrato es para-nitroanilida unida covalentemente a un sustrato peptídico para trombina, o 7-amino-4-metilcumarina unida covalentemente a un sustrato peptídico para trombina, o en el que el aptámero secuestra un par de agentes que cuando se liberan de los aptámeros se combinan para generar una señal detectable; opcionalmente en el que el constructo de enmascaramiento basado en ARN comprende un oligonucleótido de ARN al que se unen un ligando detectable y un componente de enmascaramiento, opcionalmente en el que el ligando detectable es un fluoróforo y el componente de enmascaramiento es una molécula extintora;
 en el que el constructo de enmascaramiento basado en ARN comprende una nanopartícula mantenida en agregado por moléculas puente, en el que al menos una porción de las moléculas puente comprende ARN, y en el que la disolución se somete a un cambio de color cuando la nanopartícula se dispersa en disolución, opcionalmente en el que la nanopartícula es un metal coloidal, opcionalmente en el que el metal coloidal es oro coloidal;
 en el que el constructo de enmascaramiento basado en ARN que comprende un punto cuántico unido a una o más moléculas extintoras mediante una molécula de unión, en el que al menos una porción de la molécula de unión comprende ARN;
- en el que el constructo de enmascaramiento basado en ARN comprende ARN en complejo con un agente intercalante, en el que el agente intercalante cambia la absorbancia tras la escisión del ARN, opcionalmente en el que el agente intercalante es pironina-Y o azul de metileno; o
 en el que dicho constructo de enmascaramiento comprende un oligonucleótido de ARN diseñado para unirse a una secuencia formadora de cuádruple G, en el que una estructura de cuádruple G se forma mediante la secuencia formadora de cuádruple G tras la escisión del constructo de enmascaramiento, y en el que la estructura de cuádruple G genera una señal positiva detectable.
4. El sistema según cualquiera de las reivindicaciones 1 a 3, en el que el uno o más ARN guía diseñados para

unirse a los correspondientes ácidos nucleicos diana comprenden un apareamiento erróneo (sintético), opcionalmente en el que dicho apareamiento erróneo está en el sentido de 5' o en el sentido de 3' de un SNP u otra variación de un solo nucleótido en dicho ácido nucleico diana, opcionalmente en el que el apareamiento erróneo sintético en dicho ARN guía está en la posición 3, 4, 5 o 6 del espaciador, preferiblemente en la posición 3, en el que dicho apareamiento erróneo en dicho ARN guía está en la posición 1, 2, 3, 4, 5, 6, 7, 8 o 9 del espaciador, preferiblemente en la posición 5, o en el que dicho apareamiento erróneo es de 1, 2, 3, 4 o 5 nucleótidos en el sentido de 5' o en el sentido de 3', preferiblemente 2 nucleótidos, preferiblemente en el sentido de 3' de dicho SNP u otra variación de un único nucleótido en dicho ARN guía.

5. El sistema según una cualquiera de las reivindicaciones 1 a 4, en el que el uno o más ARN guía se diseñan para detectar un polimorfismo de un solo nucleótido en un ARN o ADN diana, o una variante de corte y empalme de un transcrito de ARN, en el que el uno o más ARN guía se diseñan para unirse a uno o más ácidos nucleicos diana que son de diagnóstico para un estado patológico, o en el que dicho ácido nucleico diana es un gen o transcrito de resistencia o susceptibilidad a antibióticos o fármacos, opcionalmente en el que el estado patológico es cáncer, una enfermedad autoinmunitaria, una infección, o en el que dicho estado patológico se caracteriza por la presencia o ausencia de un gen o transcrito o polipéptido de resistencia o susceptibilidad a antibióticos o fármacos, preferiblemente en un patógeno o una célula, opcionalmente en el que la infección está provocada por un virus, una bacteria, un hongo, un protozoo o un parásito, opcionalmente en el que la infección viral está provocada por un virus de ADN, opcionalmente un *Myoviridae*, *Podoviridae*, *Siphoviridae*, *Alloherpesviridae*, *Herpesviridae* (incluyendo virus del herpes humano y virus de la varicela zóster), *Malcoherpesviridae*, *Lipothrixviridae*, *Rudiviridae*, *Adenoviridae*, *Ampullaviridae*, *Ascoviridae*, *Asfarviridae* (incluyendo virus de la fiebre porcina africana), *Baculoviridae*, *Cicaudaviridae*, *Clavaviridae*, *Corticoviridae*, *Fuselloviridae*, *Globuloviridae*, *Guttaviridae*, *Hytrosaviridae*, *Iridoviridae*, *Maseilleviridae*, *Mimiviridae*, *Nudiviridae*, *Nimaviridae*, *Pandoraviridae*, *Papillomaviridae*, *Phycodnaviridae*, *Plasmaviridae*, polidnavirus, *Polyomaviridae* (incluyendo virus del simio 40, virus JC, virus BK), *Poxviridae* (incluyendo viruela vacuna y viruela), *Sphaerolipoviridae*, *Tectiviridae*, *Turriviridae*, *Dinodnavirus*, *Salterprovirus*, *Rhizidovirus*, o en el que la infección viral está provocada por un virus de ARN bicatenario, un virus de ARN de sentido positivo, un virus de ARN de sentido negativo, un retrovirus, o una combinación de los mismos, opcionalmente un virus *Coronaviridae*, un virus *Picornaviridae*, un virus *Caliciviridae*, un virus *Flaviviridae*, un virus *Togaviridae*, un *Bornaviridae*, un *Filoviridae*, un *Paramyxoviridae*, un *Pneumoviridae*, un *Rhabdoviridae*, un *Arenaviridae*, un *Bunyaviridae*, un *Orthomyxoviridae* o un *Deltavirus*, opcionalmente coronavirus, SARS, virus de la poliometitis, rinovirus, hepatitis A, virus de Norwalk, virus de la fiebre amarilla, virus del Nilo Occidental, virus de la hepatitis C, virus de fiebre del dengue, virus de Zika, virus de la rubéola, virus del río Ross, virus Sindbis, virus de Chikungunya, virus de la enfermedad de Borna, virus del ébola, virus de Marburgo, virus del sarampión, virus de las paperas, virus de Nipah, virus de Hendra, virus de enfermedad de Newcastle, virus sincicial respiratorio humano, virus de la rabia, virus de Lassa, *Hantavirus*, virus de la fiebre hemorrágica de Crimea-Congo, gripe o virus de la hepatitis D, opcionalmente en el que la bacteria que provoca la infección bacteriana es una especie de *Acinetobacter*, una especie de *Actinobacillus*, una especie de *Actinomycetes*, una especie de *Actinomyces*, una especie de *Aerococcus*, una especie de *Aeromonas*, una especie de *Anaplasma*, una especie de *Alcaligenes*, una especie de *Bacillus*, una especie de *Bacteriodes*, una especie de *Bartonella*, una especie de *Bifidobacterium*, una especie de *Bordetella*, una especie de *Borrelia*, una especie de *Brucella*, una especie de *Burkholderia*, una especie de *Campylobacter*, una especie de *Capnocytophaga*, una especie de *Chlamydia*, una especie de *Citrobacter*, una especie de *Coxiella*, una especie de *Corynebacterium*, una especie de *Clostridium*, una especie de *Eikenella*, una especie de *Enterobacter*, una especie de *Escherichia*, una especie de *Enterococcus*, una especie de *Ehlichia*, una especie de *Epidermophyton*, una especie de *Erysipelothrix*, una especie de *Eubacterium*, una especie de *Francisella*, una especie de *Fusobacterium*, una especie de *Gardnerella*, una especie de *Gemella*, una especie de *Haemophilus*, una especie de *Helicobacter*, una especie de *Kingella*, una especie de *Klebsiella*, una especie de *Lactobacillus*, una especie de *Lactococcus*, una especie de *Listeria*, una especie de *Leptospira*, una especie de *Legionella*, una especie de *Leptospira*, una especie de *Leuconostoc*, una especie de *Mannheimia*, una especie de *Microsporum*, una especie de *Micrococcus*, una especie de *Moraxella*, una especie de *Morganell*, una especie de *Mobiluncus*, una especie de *Micrococcus*, una especie de *Mycobacterium*, una especie de *Mycoplasma*, una especie de *Nocardia*, una especie de *Neisseria*, una especie de *Pasteurella*, una especie de *Pediococcus*, una especie de *Peptostreptococcus*, una especie de *Pityrosporum*, una especie de *Plesiomonas*, una especie de *Prevotella*, una especie de *Porphyromonas*, una especie de *Proteus*, una especie de *Providencia*, una especie de *Pseudomonas*, una especie de *Propionibacterium*, una especie de *Rhodococcus*, una especie de *Rickettsia*, una especie de *Rhodococcus*, una especie de *Serratia*, una especie de *Stenotrophomonas*, una especie de *Salmonella*, una especie de *Serratia*, una especie de *Shigella*, una especie de *Staphylococcus*, una especie de *Streptococcus*, una especie de *Spirillum*, una especie de *Streptobacillus*, una especie de *Treponema*, una especie de *Tropheryma*, una especie de *Trichophyton*, una especie de *Ureaplasma*, una especie de *Veillonella*, una especie de *Vibrio*, una especie de *Yersinia*, una especie de *Xanthomonas*, o combinación de las mismas, opcionalmente en el hongo es *Aspergillus*, *Blastomyces*, *Candidiasis*, *Coccidioidomycosis*, *Cryptococcus neoformans*, *Cryptococcus gatti*, sp. *Histoplasma* sp. (tal como *Histoplasma capsulatum*),

Pneumocystis sp. (tal como *Pneumocystis jirovecii*), *Stachybotrys* (tal como *Stachybotrys chartarum*), *Mucroymcosis*, *Sporothrix*, tiña de infecciones oculares fúngicas, *Exserohilum*, *Cladosporium*, *Geotrichum*, *Saccharomyces*, una especie de *Hansenula*, una especie de *Candida*, una especie de *Kluyveromyces*, una especie de *Debaryomyces*, una especie de *Pichia*, una especie de *Penicillium*, una especie de *Cladosporium*, una especie de *Byssoschlamys* o una combinación de las mismas, opcionalmente en el que el protozoo es *Euglenozoa*, *Heterolobosea*, *Diplomonadida*, *Amoebozoa*, *Blastocystis*, *Apicomplexa*, o combinación de los mismos, opcionalmente en el que el parásito es *Trypanosoma cruzi* (enfermedad de Chagas), *T. brucei gambiense*, *T. brucei rhodesiense*, *Leishmania braziliensis*, *L. infantum*, *L. mexicana*, *L. major*, *L. tropica*, *L. donovani*, *Naegleria fowleri*, *Giardia intestinalis* (*G. lamblia*, *G. duodenalis*), *canthamoeba castellanii*, *Balamuthia mandrillaris*, *Entamoeba histolytica*, *Blastocystis hominis*, *Babesia microti*, *Cryptosporidium parvum*, *Ciclospora cayetanensis*, *Plasmodium falciparum*, *P. vivax*, *P. ovale*, *P. malariae*, y *Toxoplasma gondii*, o combinación de los mismos.

6. El sistema según una cualquiera de las reivindicaciones 1 a 5, en el que los reactivos para amplificar el ácido nucleico diana comprenden amplificación basada en la secuencia de ácido nucleico (NASBA), amplificación por recombinasa y polimerasa (RPA), amplificación isotérmica mediada por bucle (LAMP), amplificación por desplazamiento de cadena (SDA), amplificación dependiente de helicasa (HDA), reacción de amplificación por enzima de corte (NEAR), PCR, amplificación de múltiples desplazamientos (MDA), amplificación de círculo rodante (RCA), reacción en cadena de la ligasa (LCR), o método de amplificación por ramificación (RAM).

7. El sistema según una cualquiera de las reivindicaciones 1 a 6, que comprende además un sistema CRISPR de enriquecimiento, en el que el sistema CRISPR de enriquecimiento está diseñado para unirse al ácido nucleico diana antes de la detección mediante el sistema CRISPR de detección, opcionalmente en el que el sistema CRISPR de enriquecimiento comprende una proteína efectora CRISPR catalíticamente inactiva, opcionalmente en el que la proteína efectora CRISPR de enriquecimiento comprende además un marcador, en el que el marcador se usa para atraer el sistema efector CRISPR de enriquecimiento, o para unir el sistema CRISPR de enriquecimiento a un sustrato sólido, opcionalmente en el que el sustrato sólido es una celda de flujo.

8. Un dispositivo de diagnóstico que comprende uno o más volúmenes discretos individuales, comprendiendo cada volumen discreto individual un sistema CRISPR según una cualquiera de las reivindicaciones 1 a 7, opcionalmente en el que cada volumen discreto individual comprende además uno o más aptámeros de detección que comprenden un sitio de unión a promotor de la ARN polimerasa enmascarado o un sitio de unión de cebador enmascarado, en el que cada volumen discreto individual comprende además reactivos de amplificación de ácidos nucleicos, en el que la molécula diana es un ADN diana y los volúmenes discretos individuales comprenden además un cebador que se une al ADN diana y comprende un promotor de la ARN polimerasa, en el que los volúmenes discretos individuales son gotas, en el que los volúmenes discretos individuales son micropocillos, y/o en el que los volúmenes discretos individuales son puntos definidos sobre un sustrato, opcionalmente en el que el sustrato es un sustrato de materiales flexibles, opcionalmente en el que el sustrato de materiales flexibles es un sustrato de papel o un sustrato basado en polímeros flexibles.

9. Un método para detectar ácidos nucleicos diana en muestras que comprende:

poner en contacto una o más muestras con el sistema de detección de ácidos nucleicos según cualquiera de las reivindicaciones 1 a 7 o el dispositivo de diagnóstico según la reivindicación 8, y

detectar una señal de la escisión del constructo de enmascaramiento basado en ARN, detectando de ese modo el uno o más ácidos nucleicos diana en la muestra.

10. El método según la reivindicación 9, que comprende:

distribuir una muestra o conjunto de muestras en uno o más volúmenes discretos individuales, comprendiendo los volúmenes discretos individuales un sistema CRISPR según una cualquiera de las reivindicaciones 1 a 7;

incubar la muestra o conjunto de muestras en condiciones suficientes para permitir la unión del uno o más ARN guía a una o más moléculas diana;

activar la proteína efectora CRISPR mediante la unión del uno o más ARN guía a la una o más moléculas diana, en el que activar la proteína efectora CRISPR da como resultado la modificación del constructo de enmascaramiento basado en ARN de modo que se genera una señal positiva detectable; y

detectar la señal positiva detectable, en el que detección de la señal positiva detectable indica la presencia de una o más moléculas diana en la muestra

opcionalmente en el que amplificar ARN comprende amplificación mediante NASBA o RPA.

11. El método según una cualquiera de las reivindicaciones 9 a 10, en el que la muestra es una muestra biológica o una muestra ambiental, opcionalmente en el que la muestra biológica es sangre, plasma, suero, orina, heces, esputo, moco, líquido linfático, líquido sinovial, bilis, ascitis, efusión pleural, seroma, saliva, líquido cefalorraquídeo, humor acuoso o vítreo, o cualquier secreción corporal, un trasudado, un exudado (por ejemplo, líquido obtenido de cualquier absceso o cualquier otro sitio de infección o inflamación), o líquido obtenido de una articulación (por ejemplo, una articulación normal o una articulación afectada por una enfermedad, tal como artritis reumatoide, osteoartritis, gota o artritis séptica), o una muestra de la superficie de la piel o de la membrana mucosa, opcionalmente en el que la muestra ambiental se obtiene de una muestra alimentaria, una superficie de papel, un material textil, una superficie de metal, una superficie de madera, una superficie de plástico, una muestra de suelo, una muestra de agua dulce, una muestra de aguas residuales, una muestra de agua salina, o una combinación de los mismos.
12. El método según cualquiera de las reivindicaciones 9 a 11, que comprende además comparar la señal con una señal de patrón (sintético).
13. El método según cualquiera de las reivindicaciones 9 a 12, que comprende aplicar dicha muestra puesta en contacto a un ensayo inmunocromatográfico de flujo lateral.
14. El método según la reivindicación 13, en el que dicho sistema de detección de ácidos nucleicos comprende un constructo de enmascaramiento basado en ARN que comprende unas moléculas primera y segunda, y en el que dicho ensayo inmunocromatográfico de flujo lateral comprende detectar dichas moléculas primera y segunda, preferiblemente en sitios de detección discretos sobre una tira de flujo lateral, opcionalmente en el que dicha primera molécula y dicha segunda molécula se detectan mediante la unión a un anticuerpo que reconoce dicha primera o segunda molécula y detectar dicha molécula unida, preferiblemente con anticuerpos en sándwich, opcionalmente en el que dicha tira de flujo lateral comprende un primer anticuerpo en el sentido de 5' dirigido contra dicha primera molécula, y un segundo anticuerpo en el sentido de 3' dirigido contra dicha segunda molécula, y en el que dicho primer anticuerpo se une al constructo de enmascaramiento basado en ARN no escindido si el ácido nucleico diana no está presente en dicha muestra, y en el que tanto dicho primer anticuerpo como dicho segundo anticuerpo se unen al constructo de enmascaramiento basado en ARN escindido si el ácido nucleico diana está presente en dicha muestra.

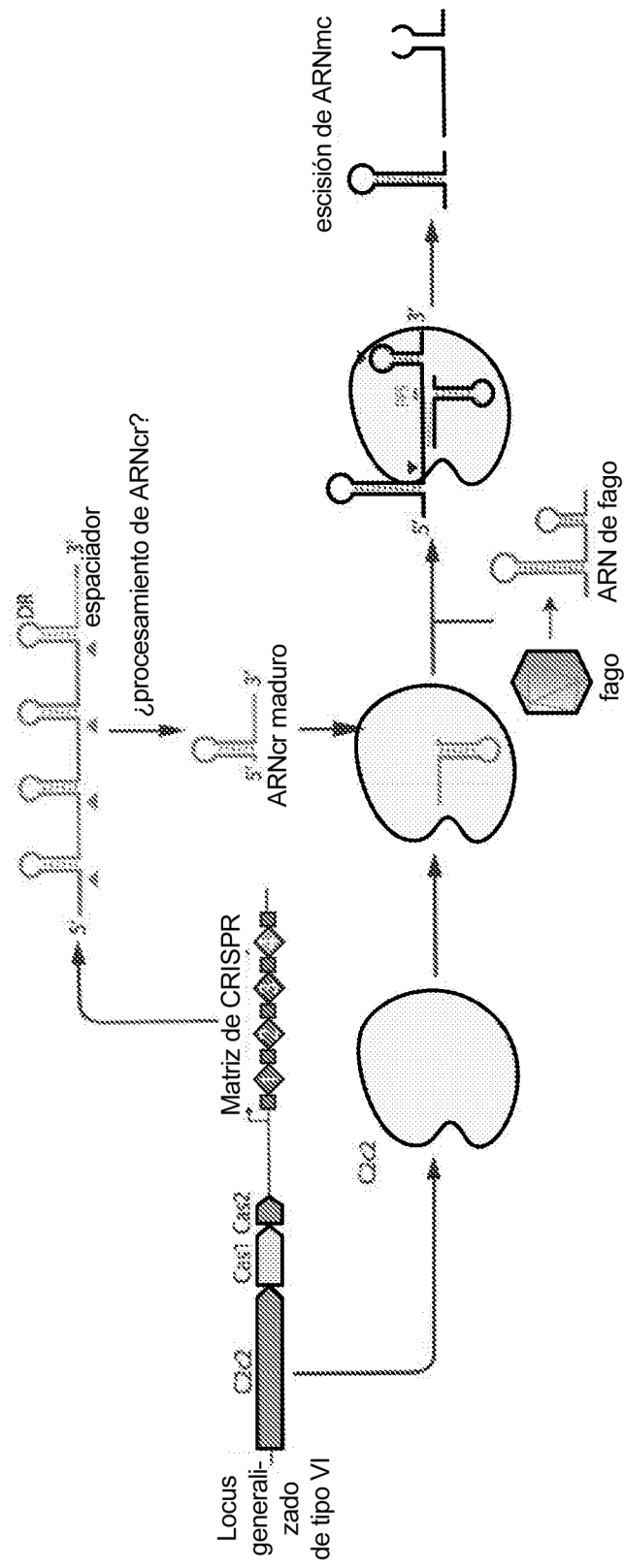


FIG. 1

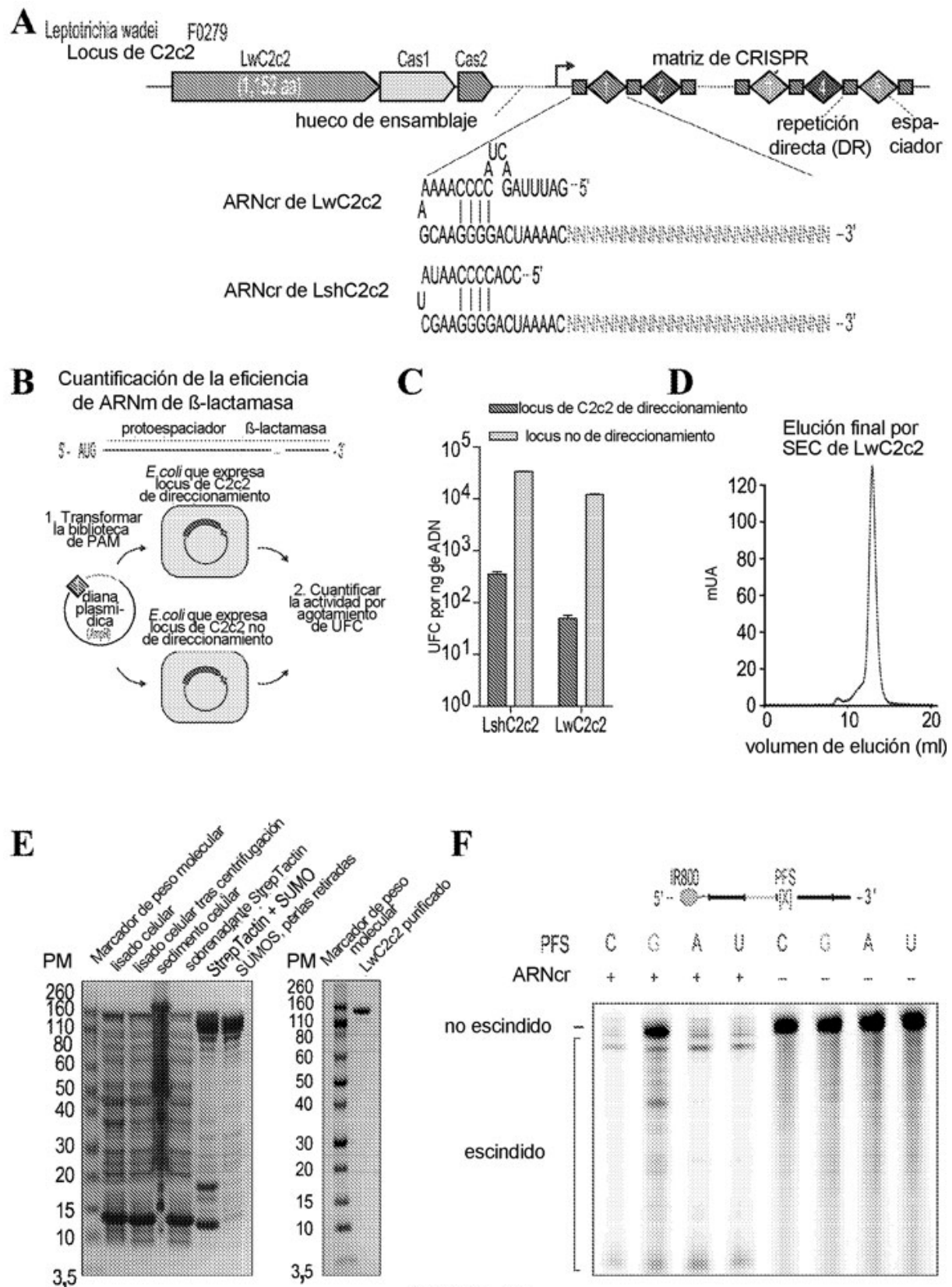


FIG. 2

Matiz Escalares

48

Datos:

[Carpas Lectura 1.485.528]

Editar la matriz

N° de lectura: Mostrar

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A																								
B																								
C																								
D																								
E																								
F																								
G																								
H																								
I																								
J																								
K																								
L																								
M																								
N																								
O																								
P																								

Editar

Enmascarar

Ayuda

FIG. 3

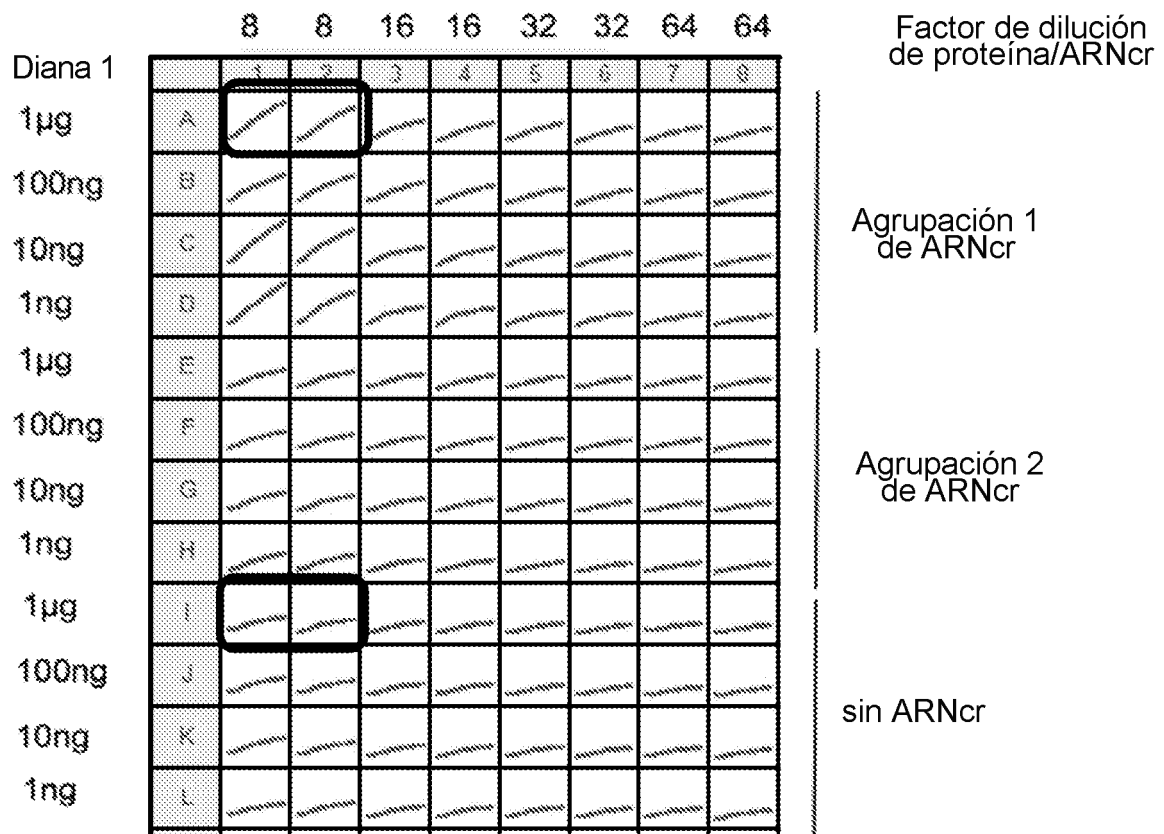


FIG. 4

Gráfico: Lectura 1:485,528 Editar los gráficos Ver: ☒ Gráfico ☐ Datos +

Pocillos: A1, A2, I1, I2 Seleccionar los pocillos

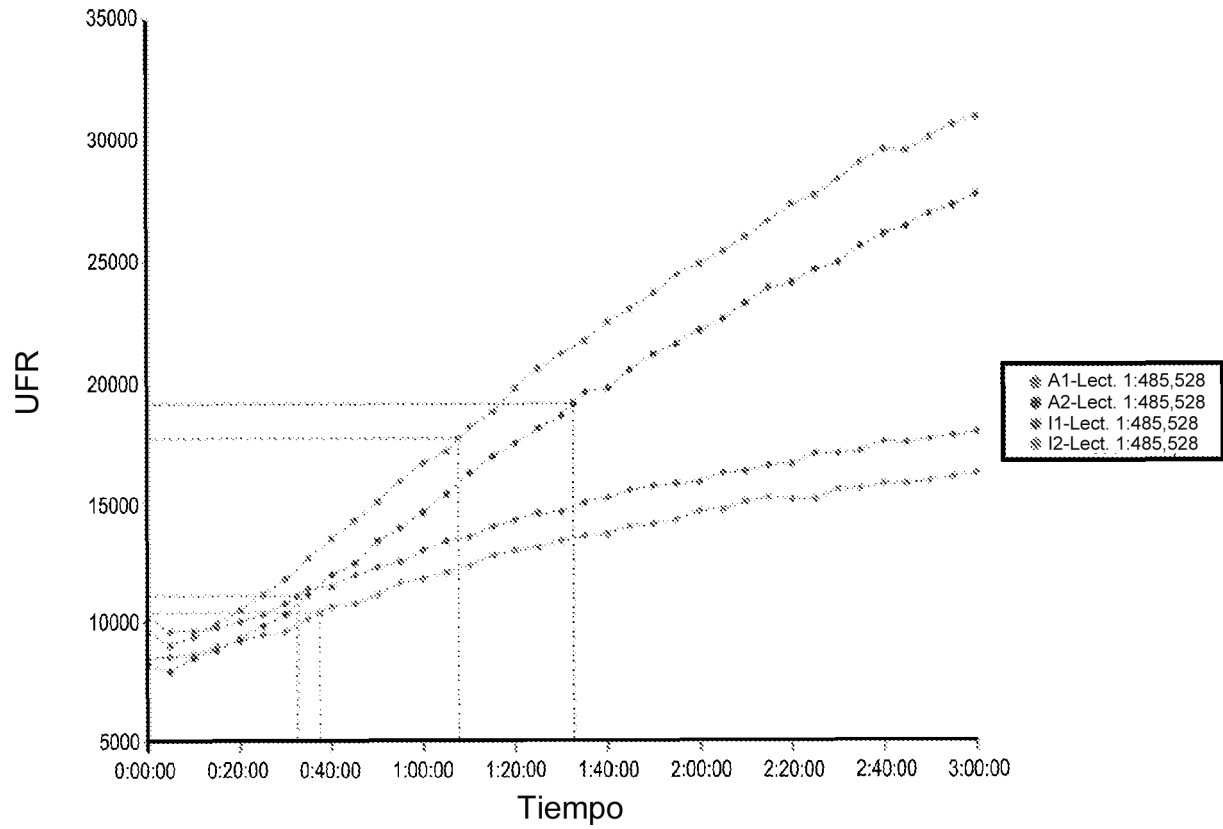


FIG. 5

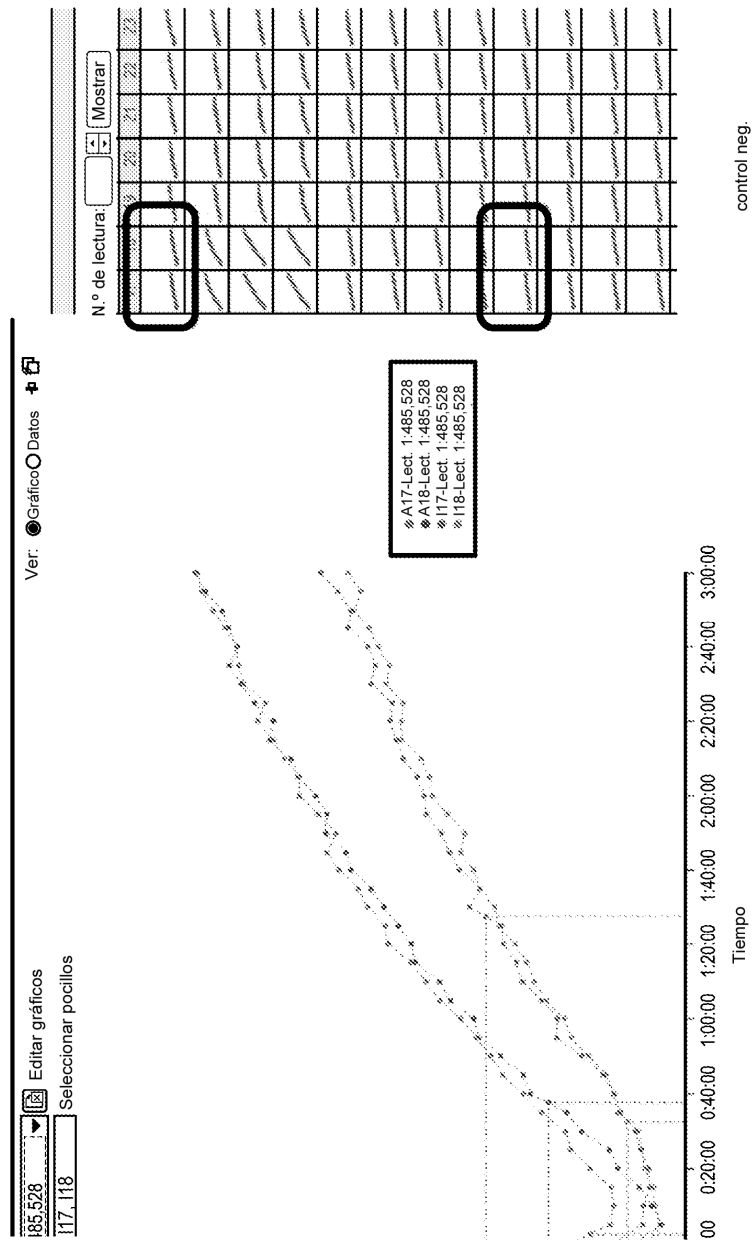


FIG. 6

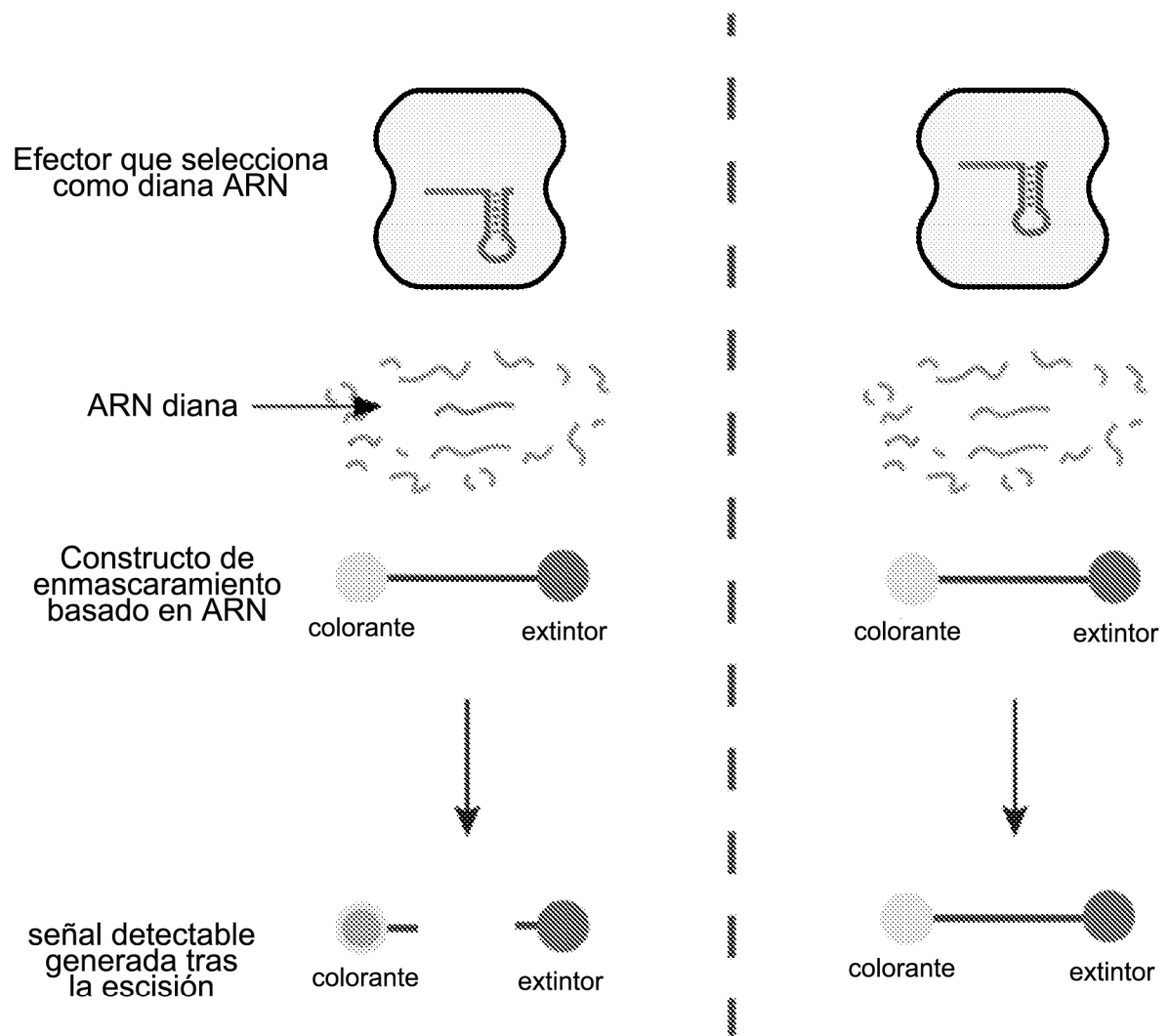


FIG. 7

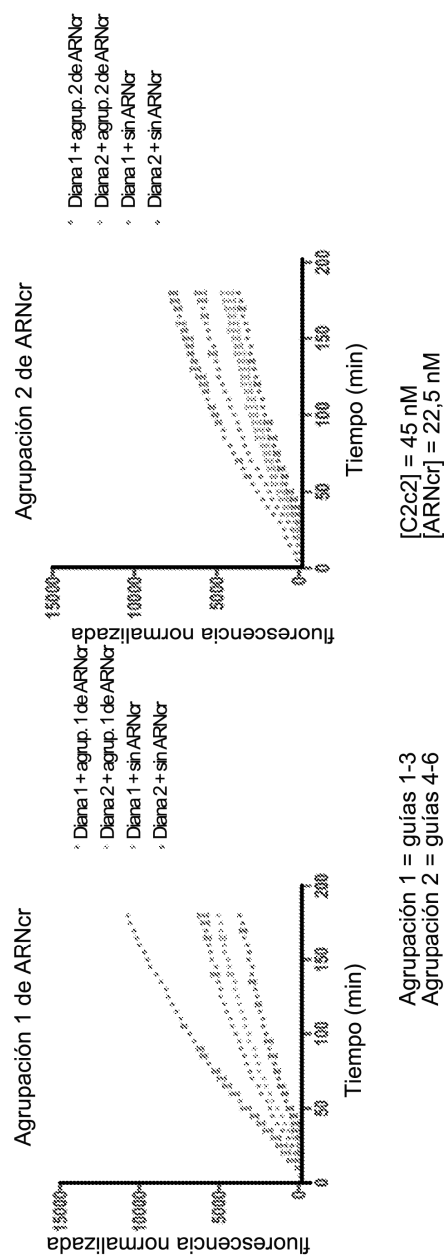
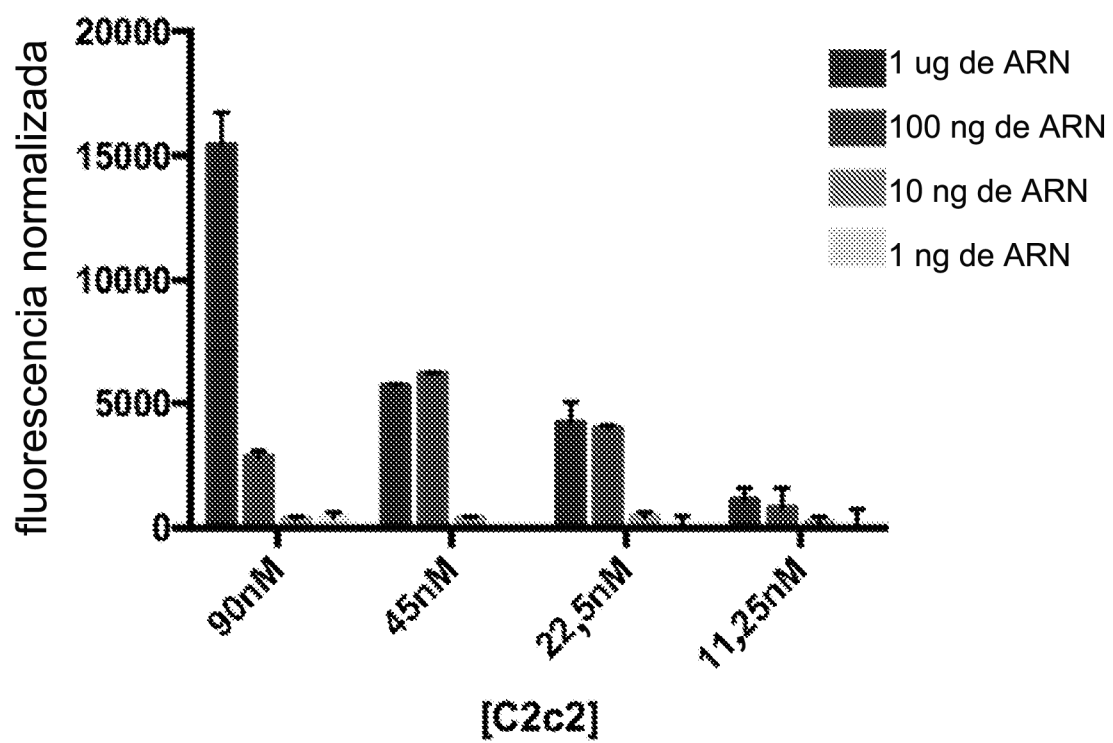


FIG. 8

**FIG. 9**

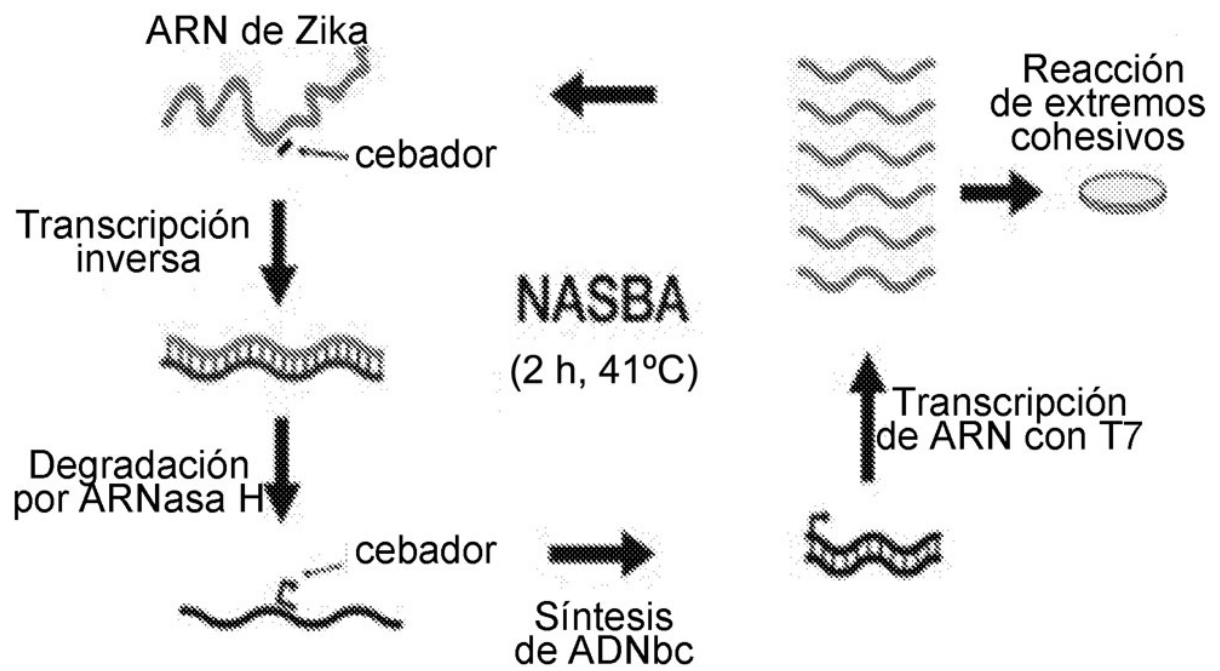
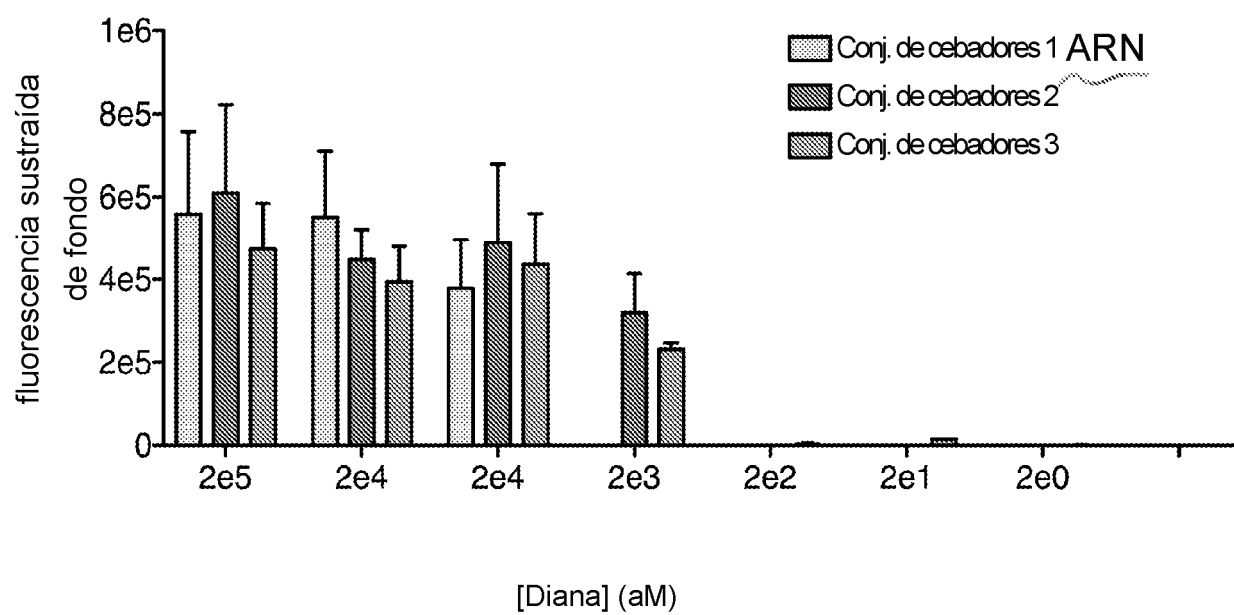


FIG. 10

**FIG. 11**

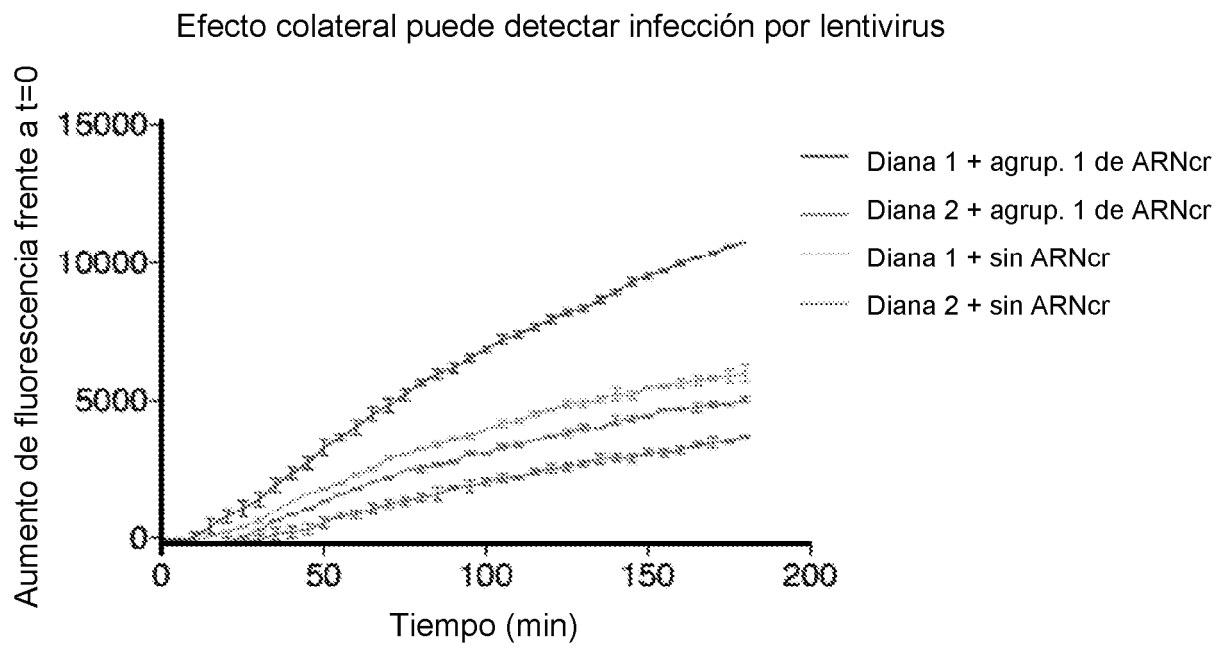


FIG. 12

Efecto colateral y NASBA pueden detectar especies a concentraciones aM

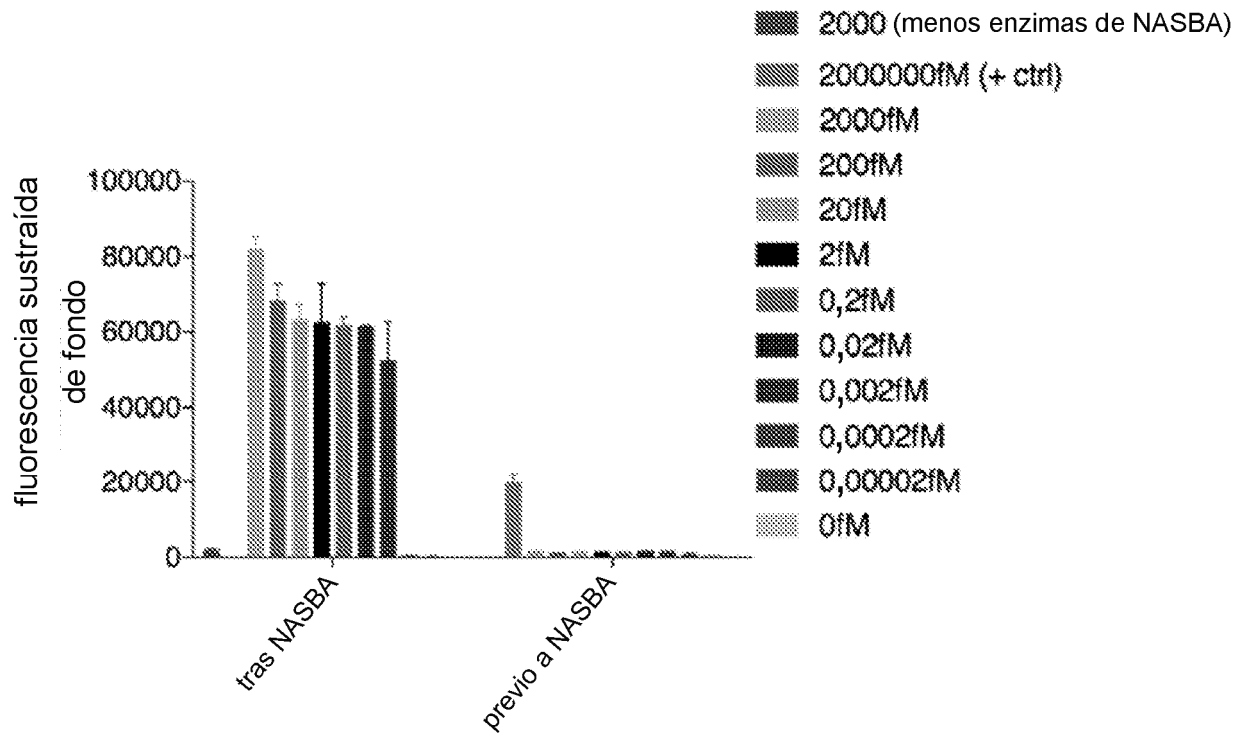


FIG. 13

Efecto colateral y NASBA pueden discriminar rápidamente muestras de baja concentración

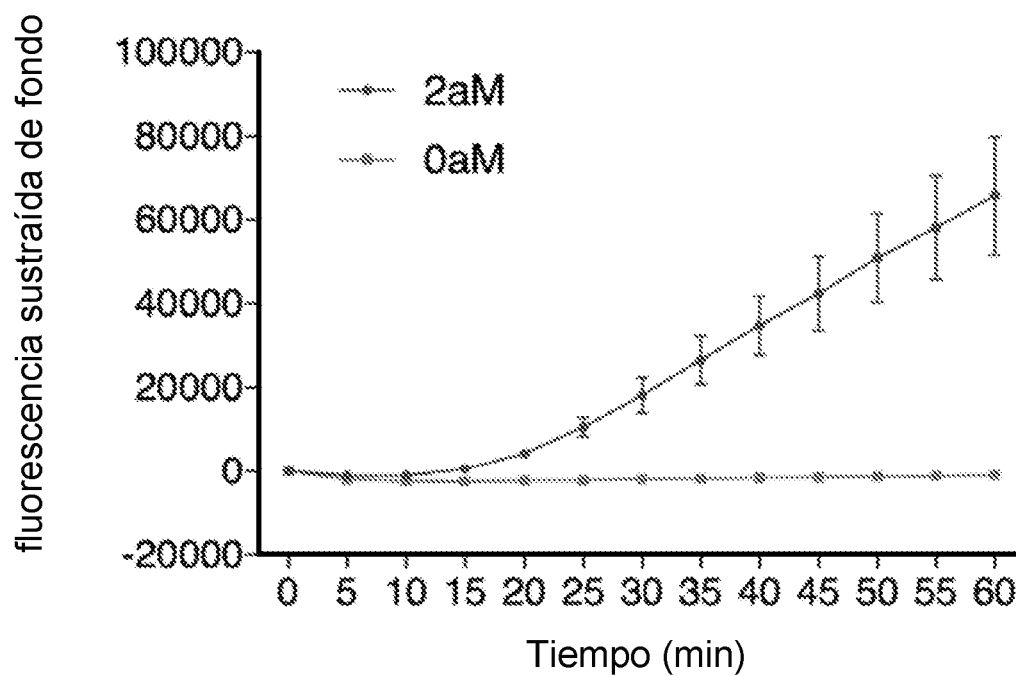


FIG. 14

Fluorescencia normalizada en puntos de tiempo
es predictiva de la concentración de muestra

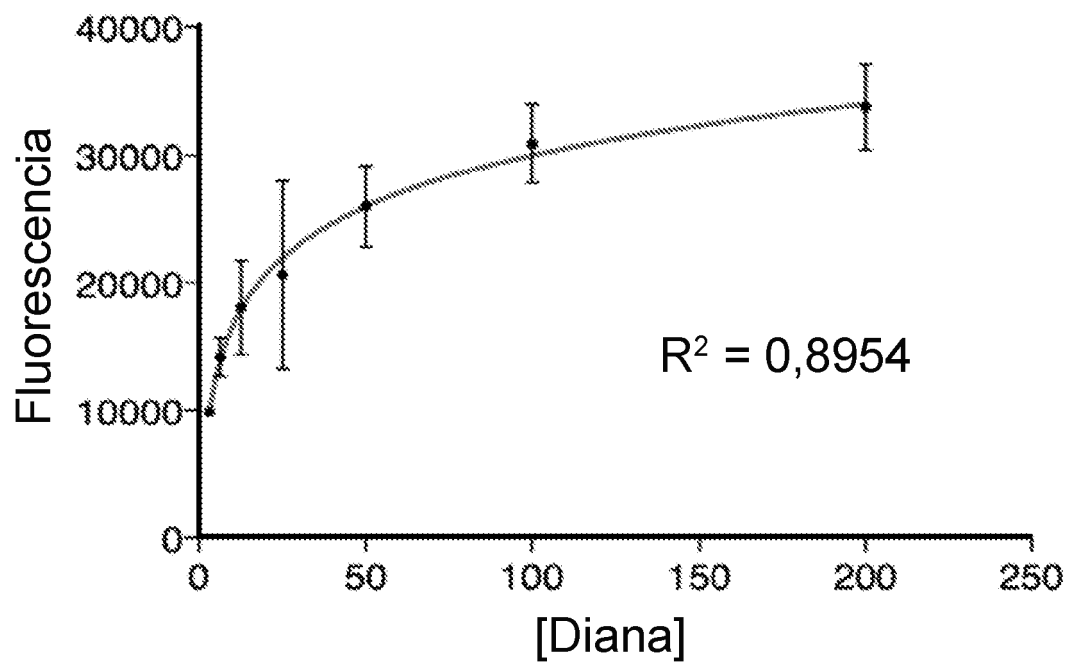


FIG. 15

A

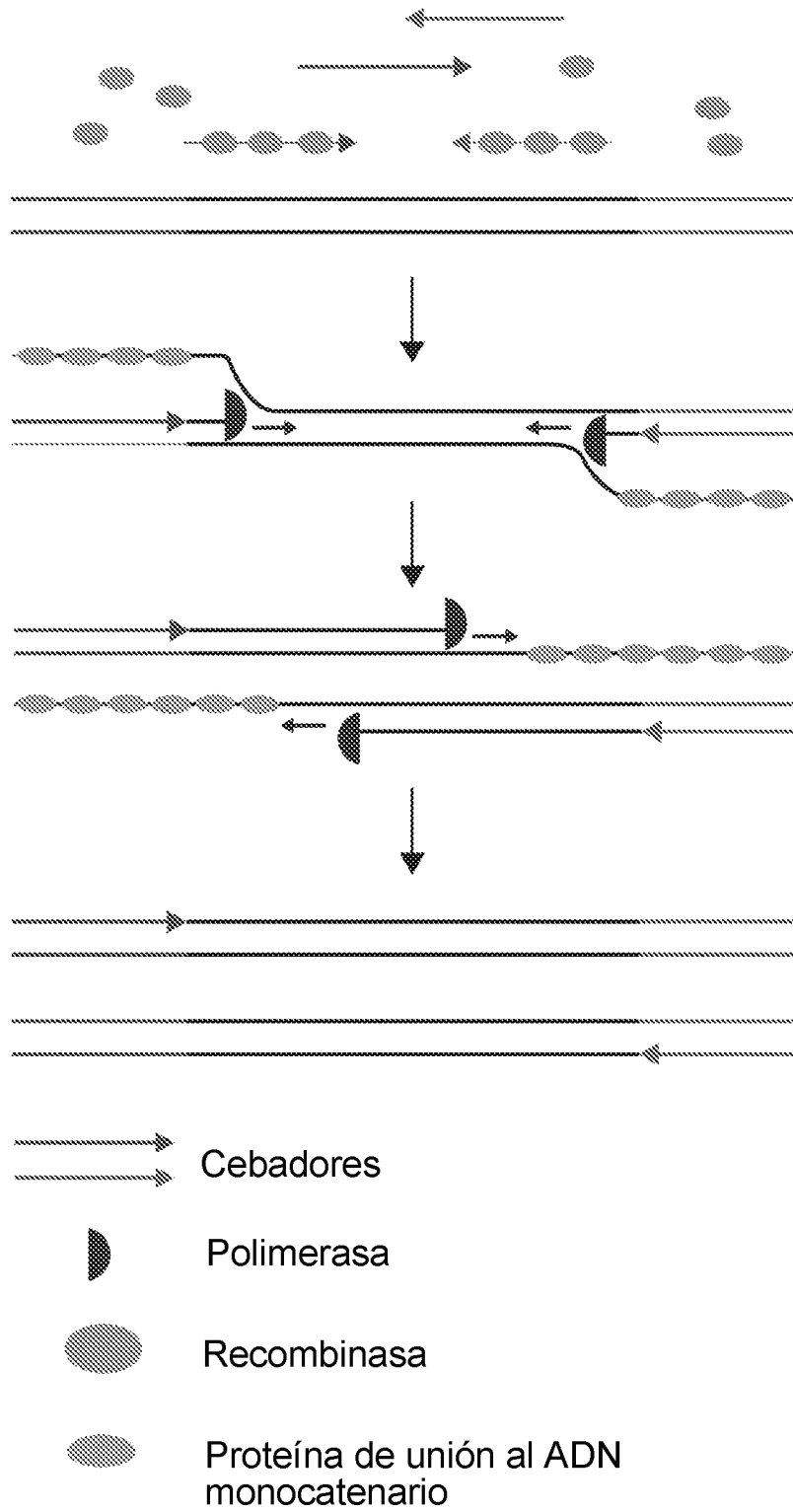


FIG. 16

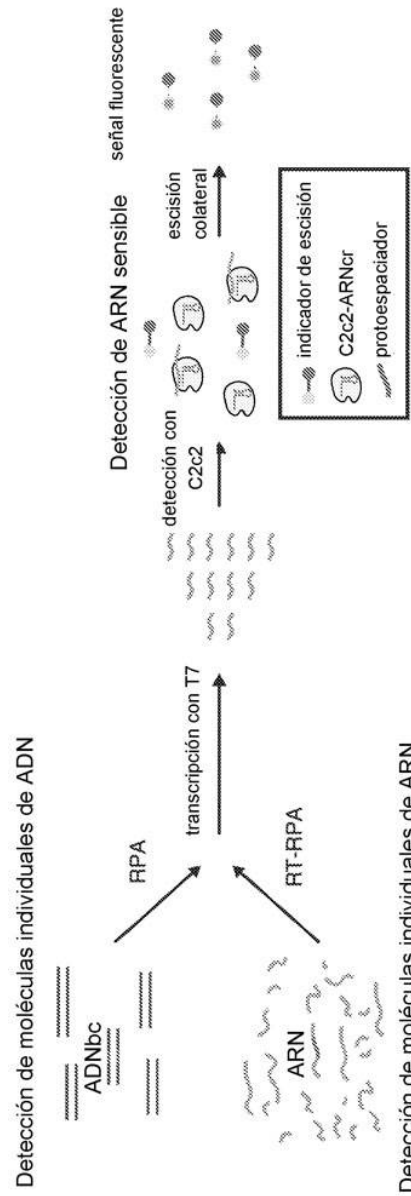


FIG. 17

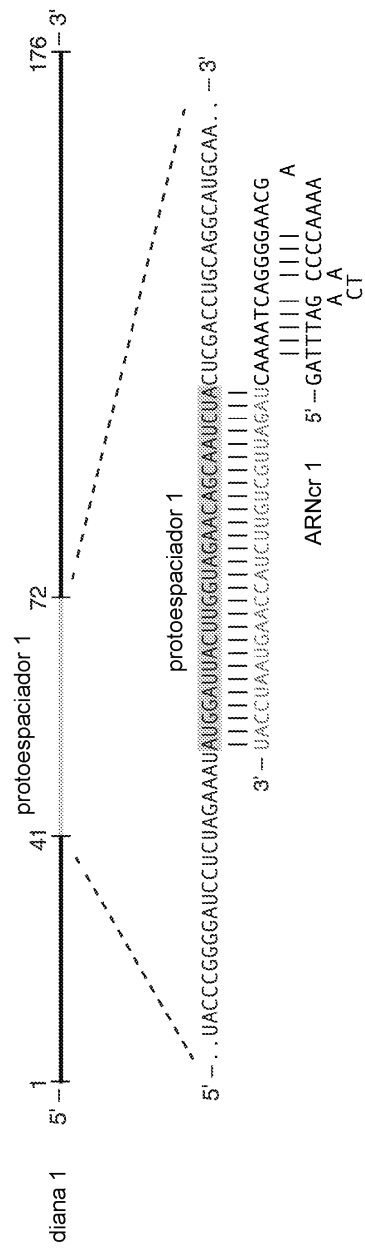
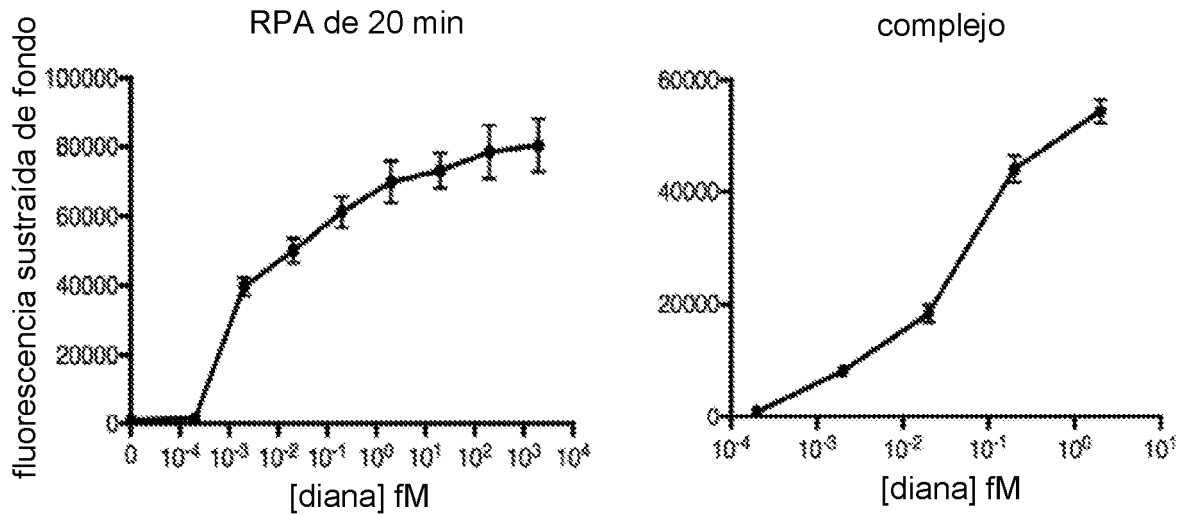


FIG. 18

Detección de moléculas individuales de ADN usando RPA

Nota: esto es un RPA de 20 min y C2c2 de 15 min (muy rápido)

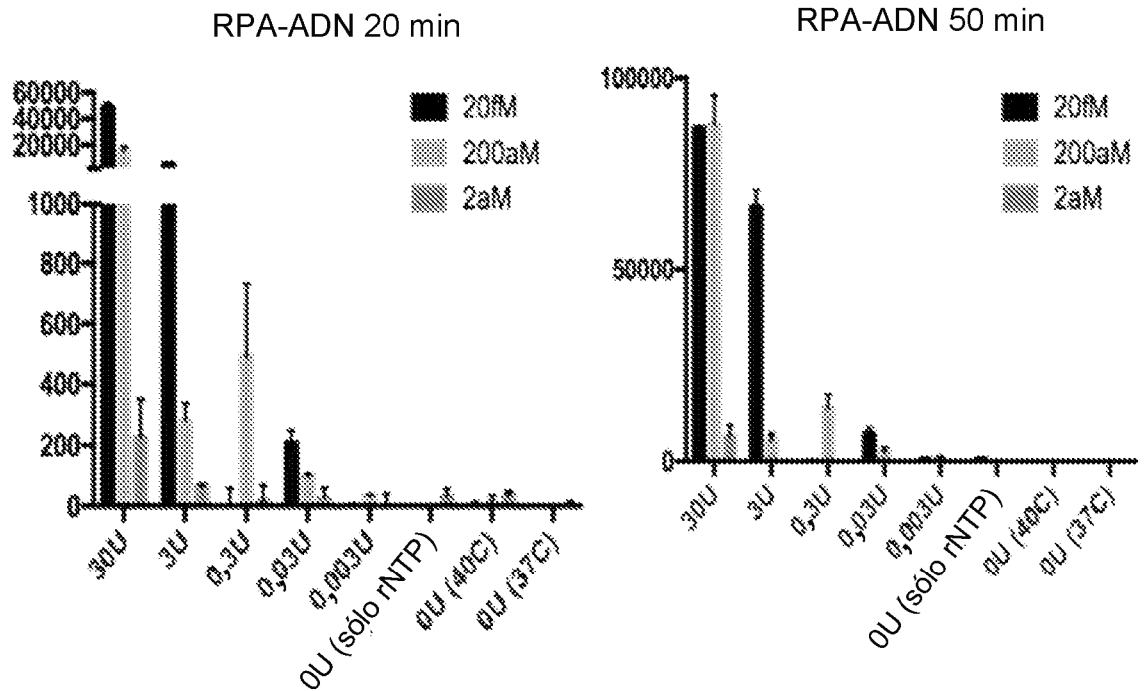


Detección a 2 aM es posible
(que es una molécula individual en 1 ul)

se diluye la diana en una mezcla compleja
para simular la muestra real

FIG. 19

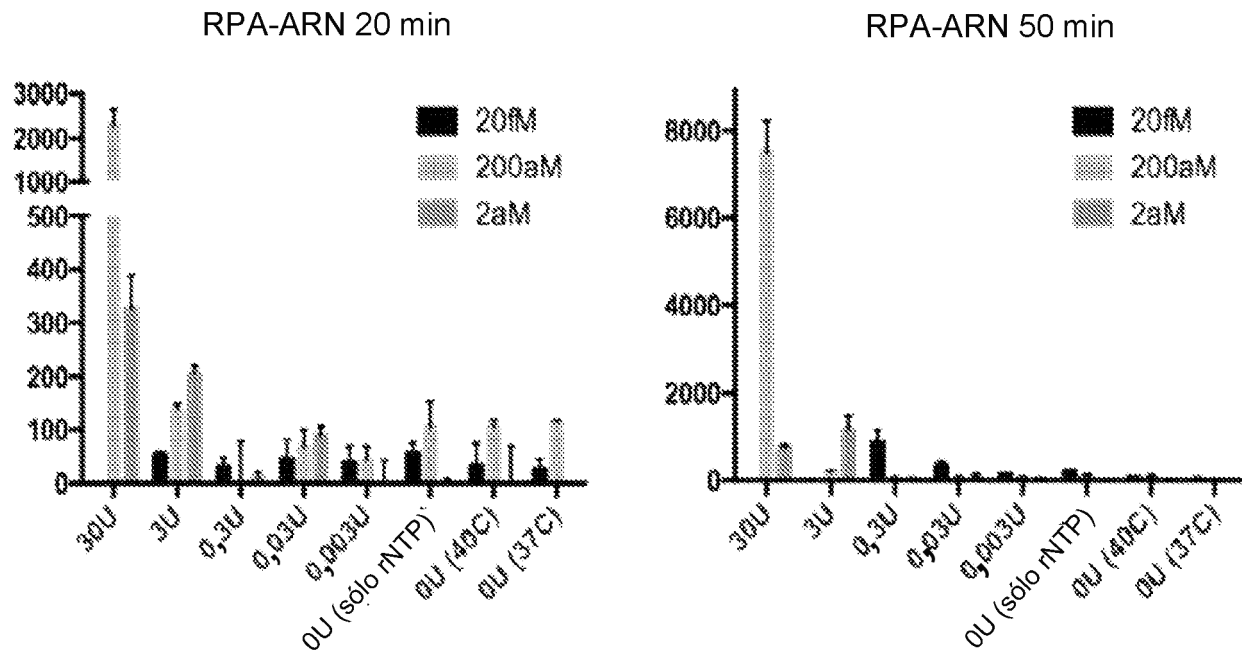
La mezcla de polimerasa T7 en una reacción sigue funcionando para la detección de ADN



Detección de moléculas individuales de ADN a partir de 20 min de incubación con C2c2. Más polimerasa T7 es mejor (el eje x son unidades de T7)

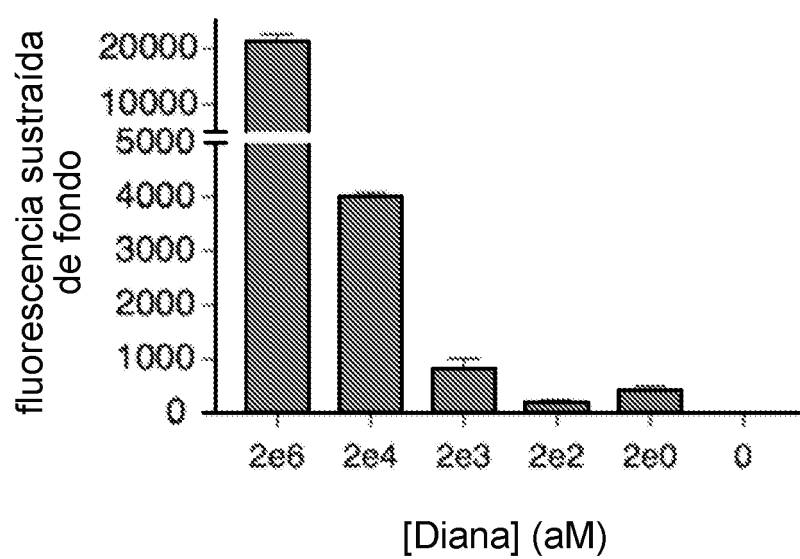
FIG. 20

La mezcla de polimerasa T7 en una reacción sigue funcionando para la detección de ARN



Detección de moléculas individuales de ARN a partir de 20 min de incubación con C2c2

FIG. 21

**FIG. 22**

Incubaciones rápidas de RPA-ARN en el tiempo siguen funcionando

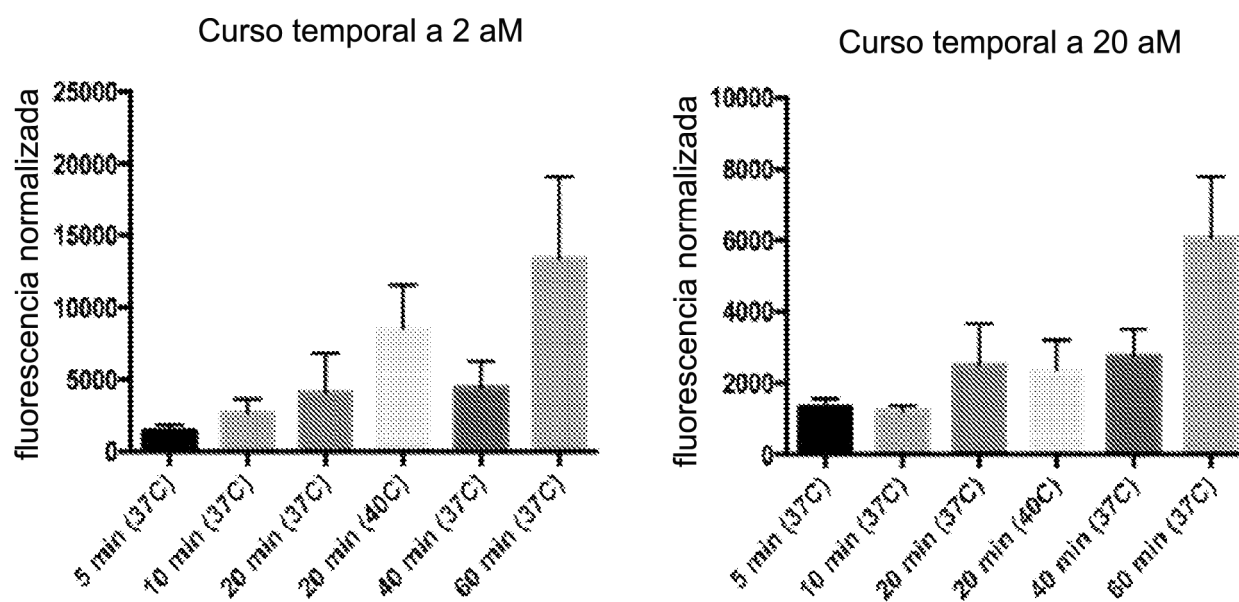


FIG. 23

El aumento de la cantidad de polimerasa T7
también refuerza la sensibilidad para el RPA-ARN

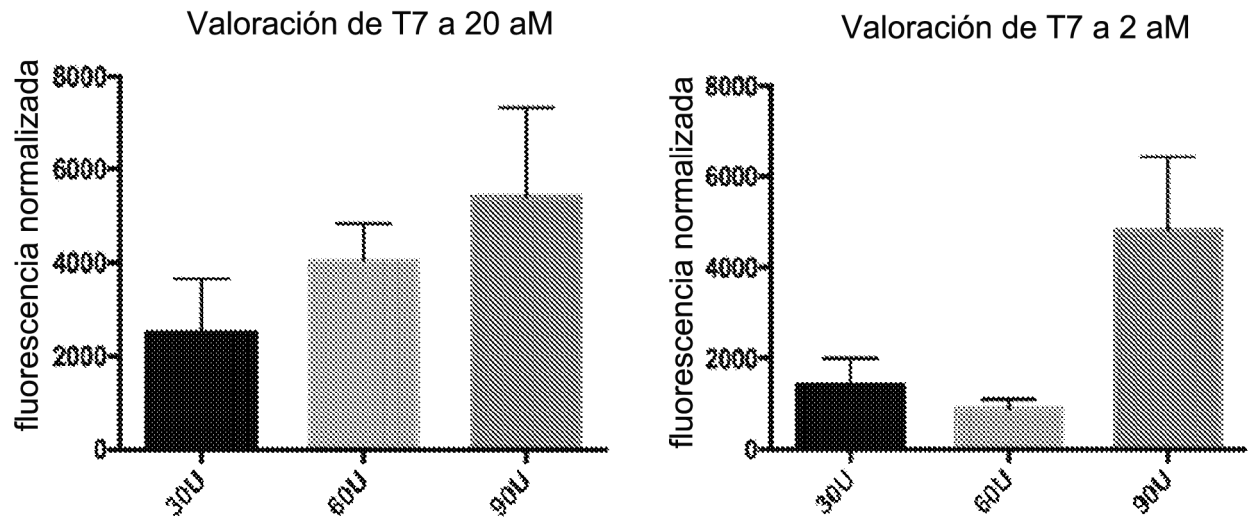


FIG. 24

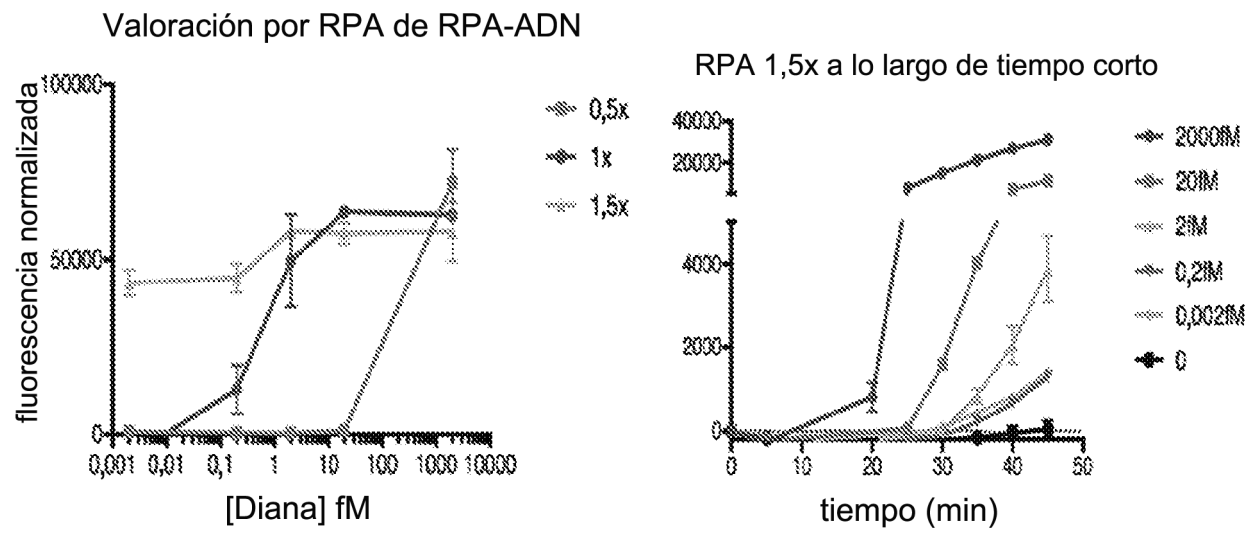


FIG. 25

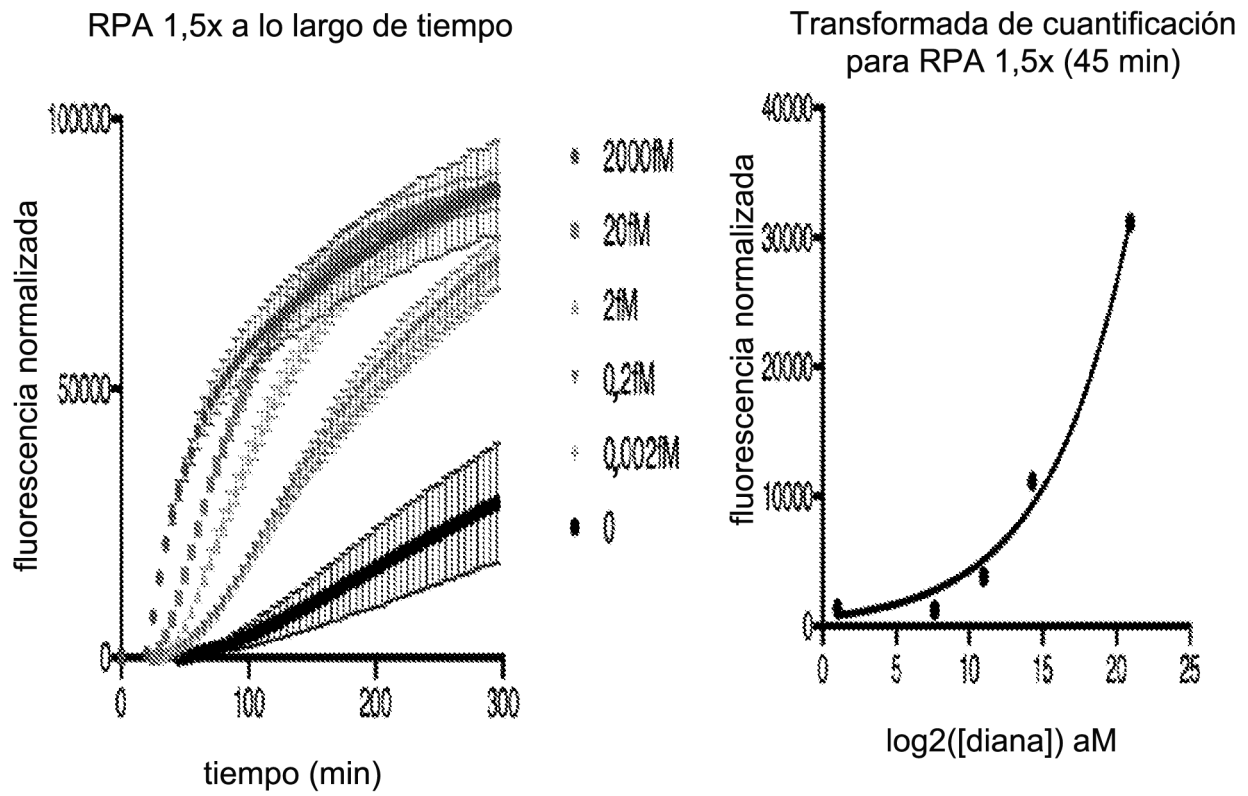
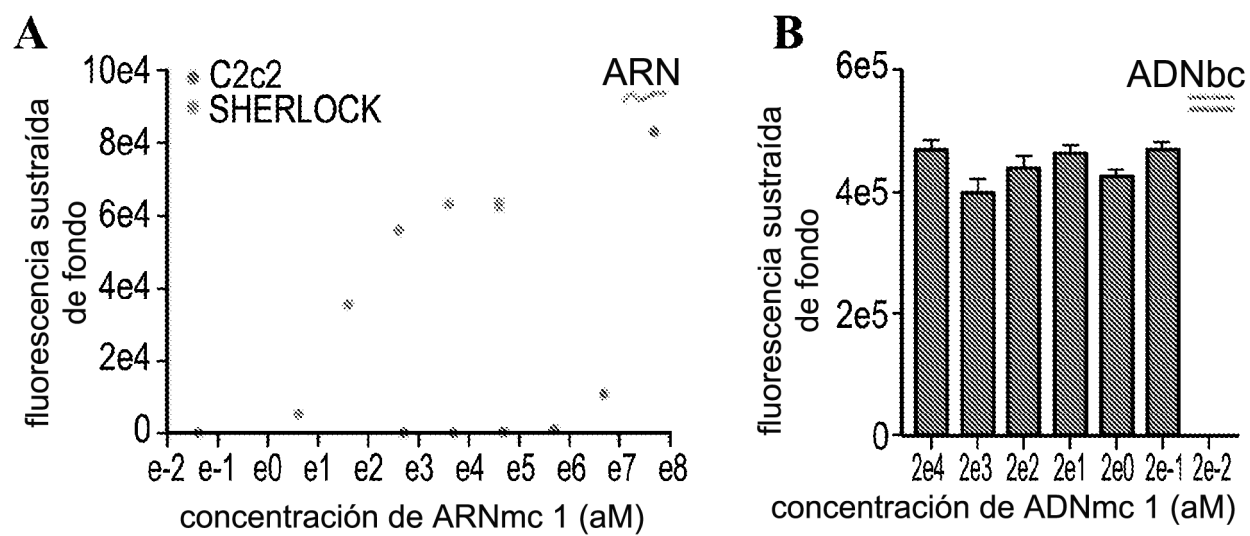
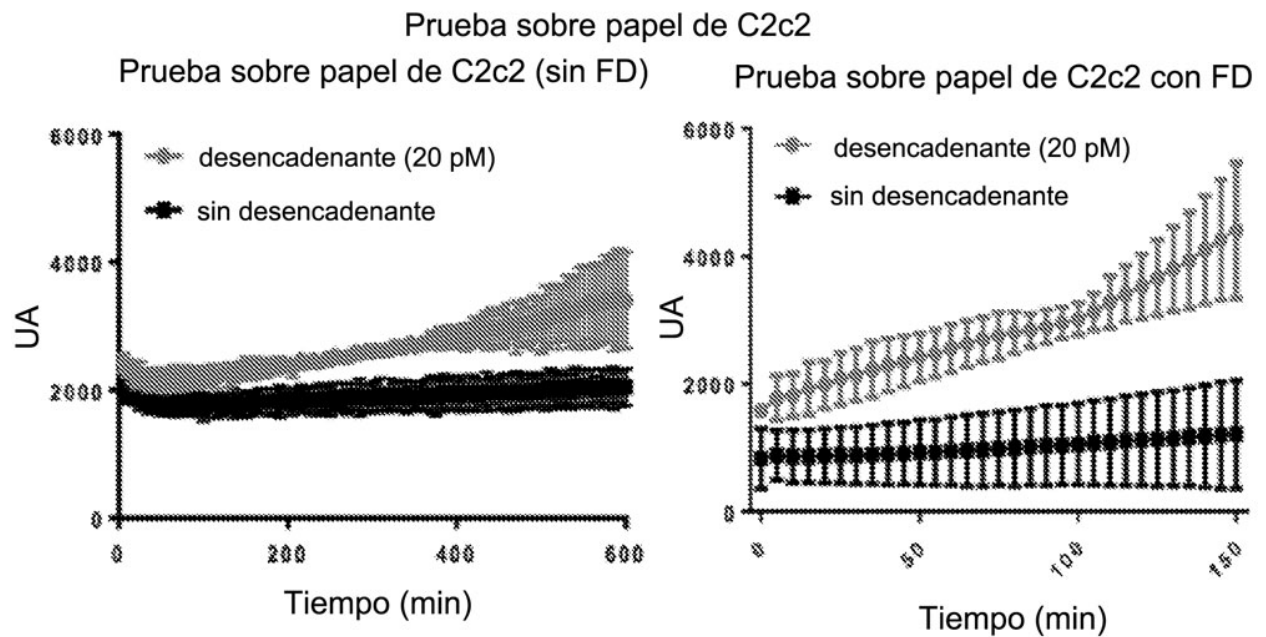


FIG. 26

**FIG. 27**



Esto es una señal de C2c2 sobre papel para una diana de 20 pM.
Conviene observar que, en realidad, el secado por congelación anterior
refuerza la señal (el gráfico de la derecha)

FIG. 28

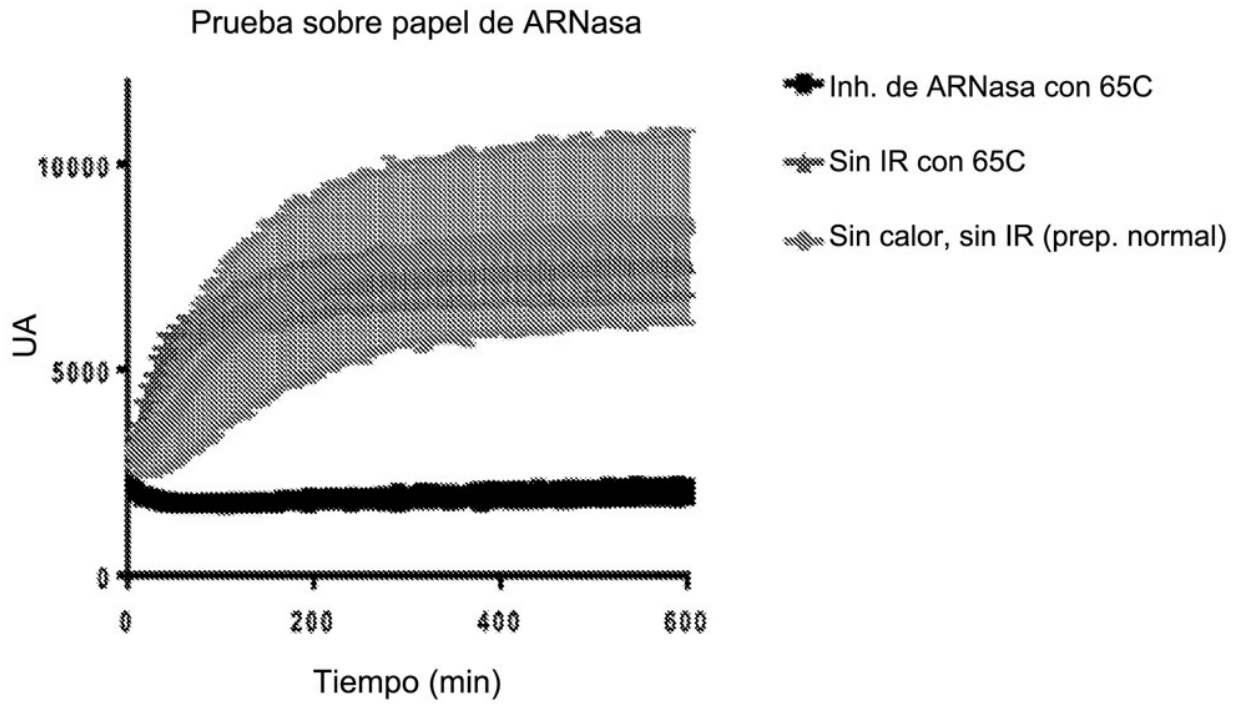
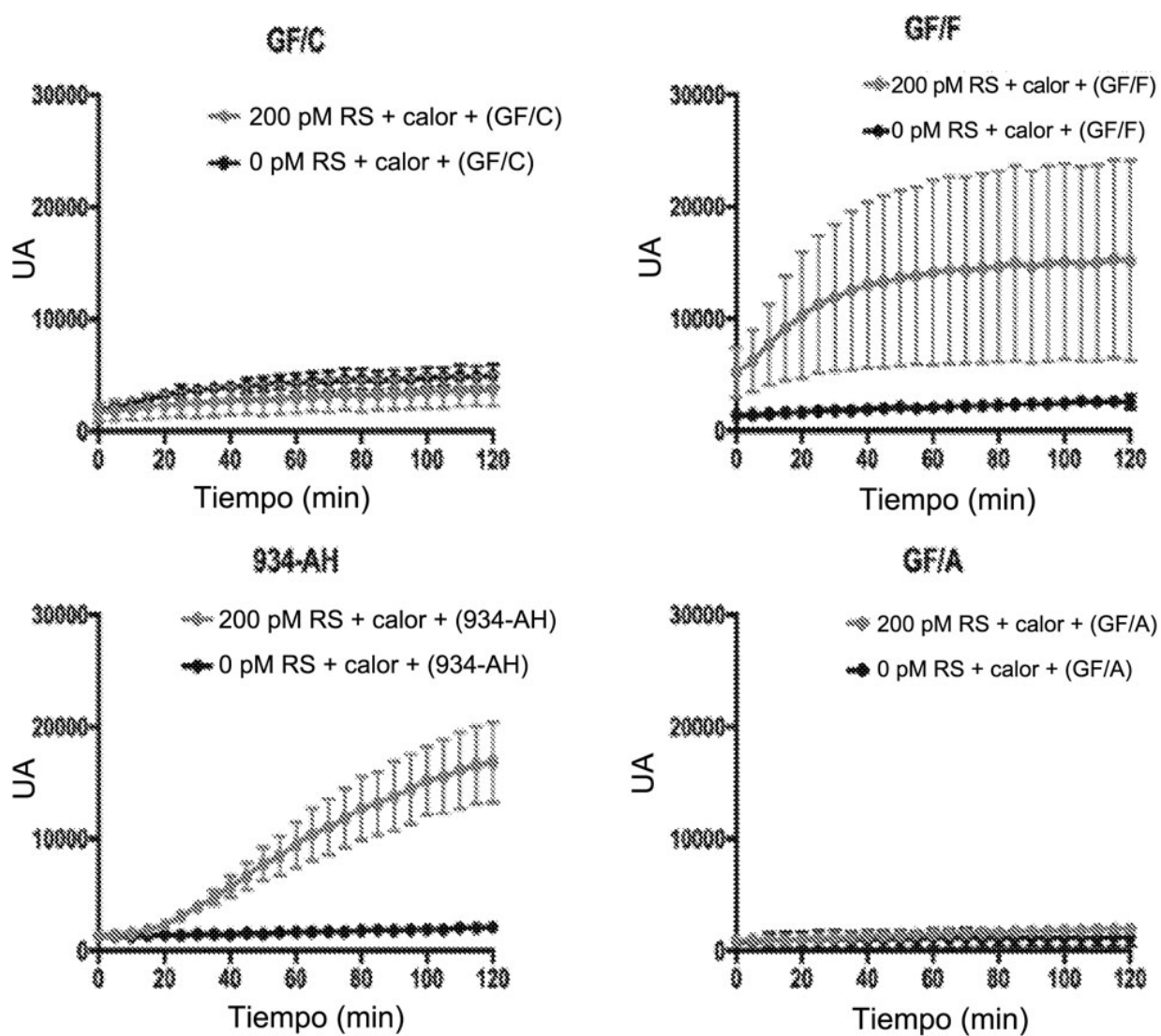
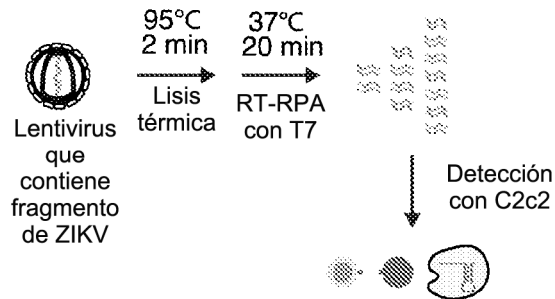
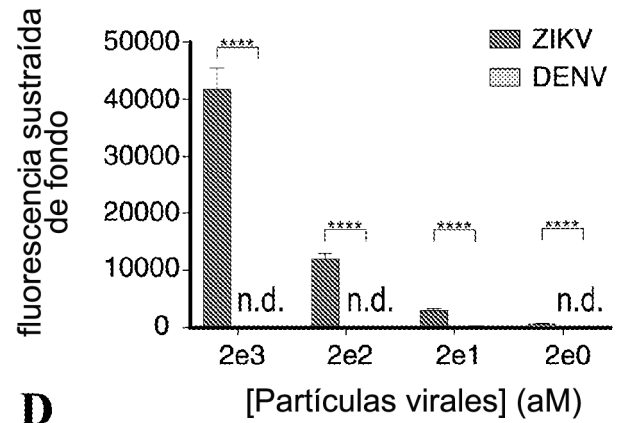
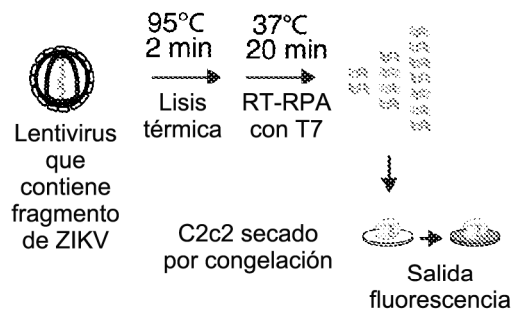
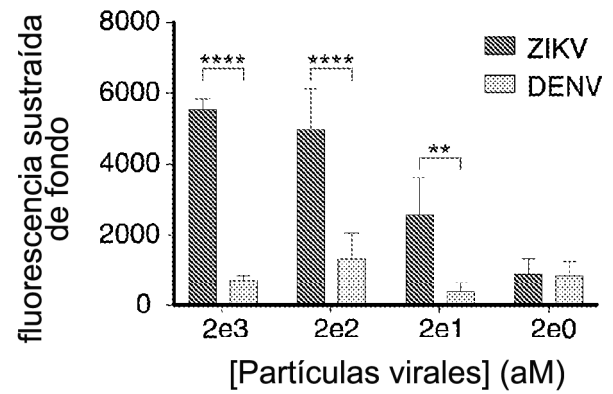


FIG. 29

Prueba sobre fibra de vidrio de C2c2

**FIG. 30**

A**B****C****D****FIG. 31**

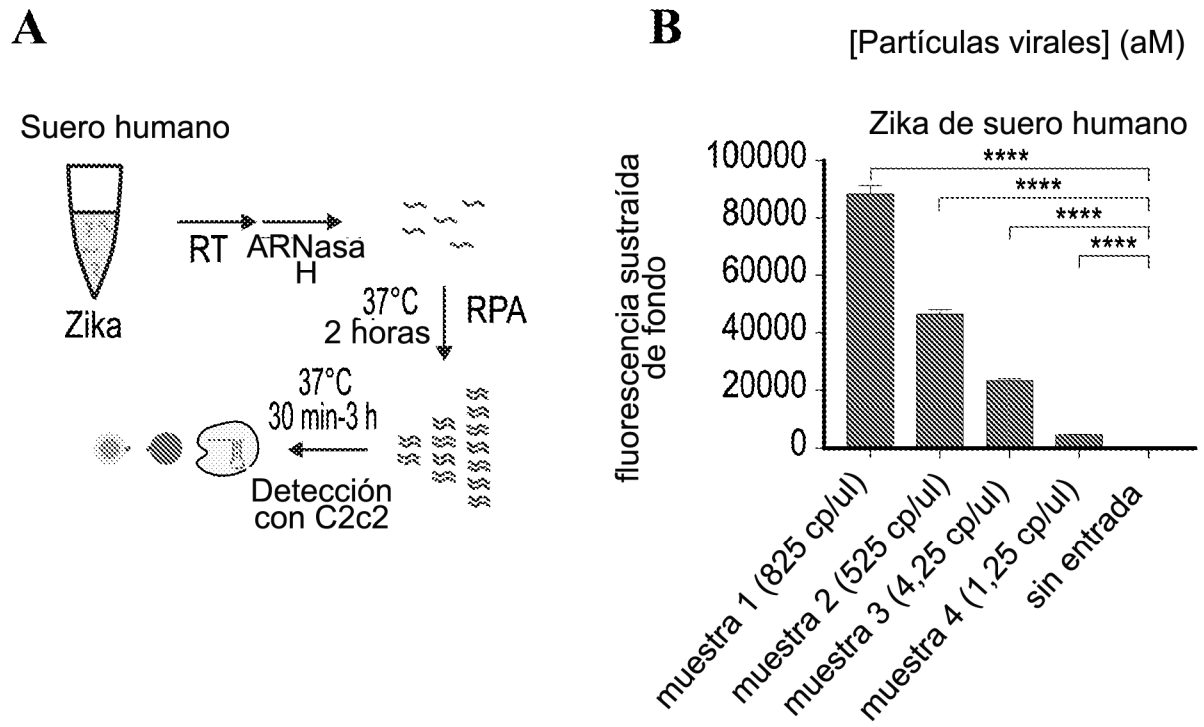


FIG. 32

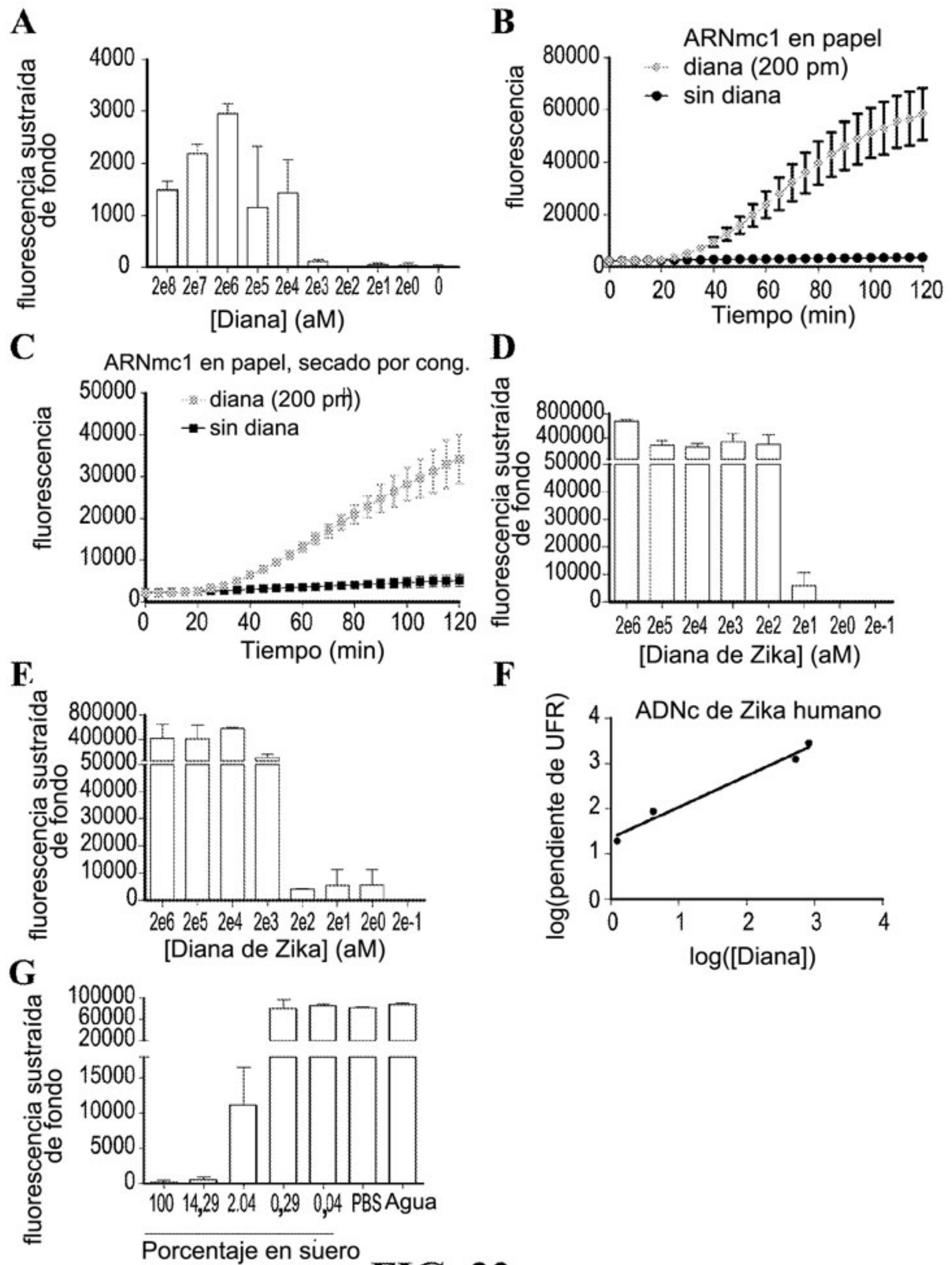


FIG. 33

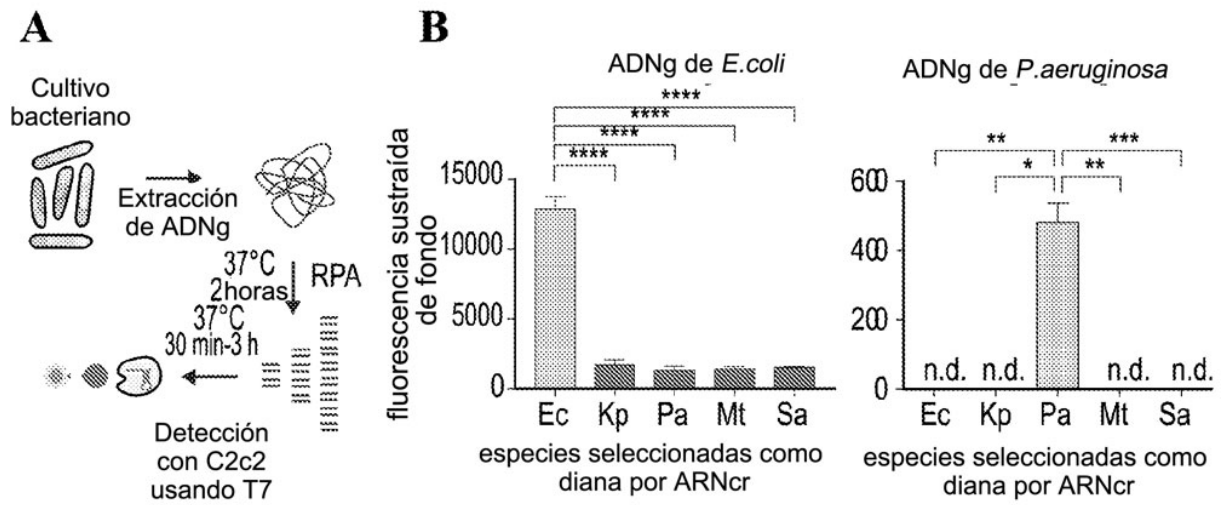


FIG. 34

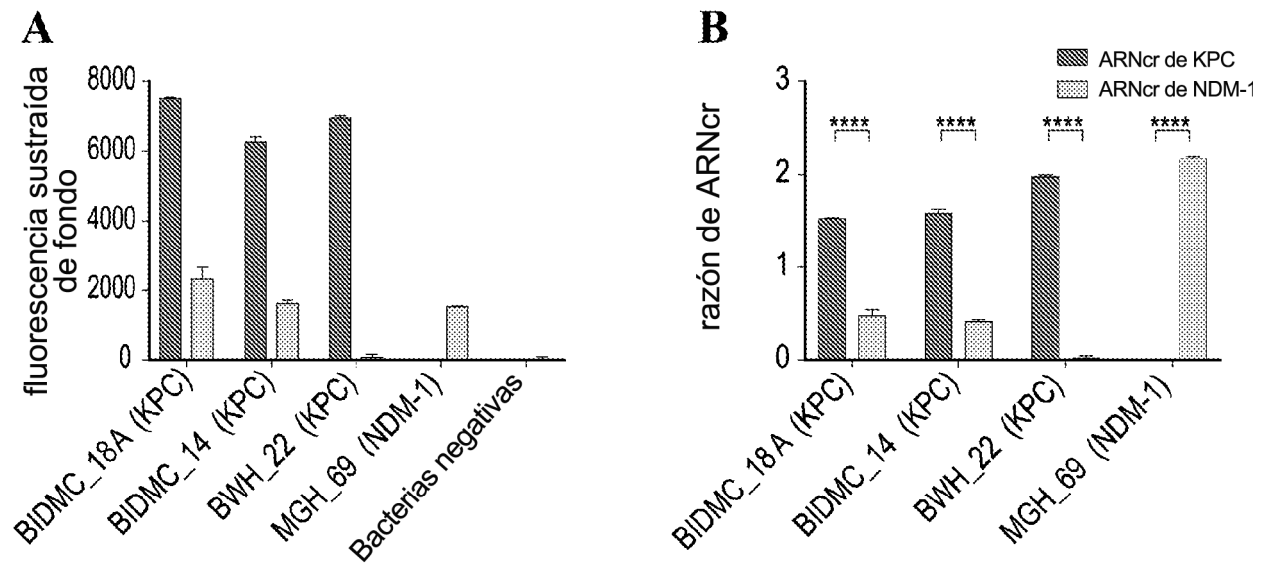


FIG. 35

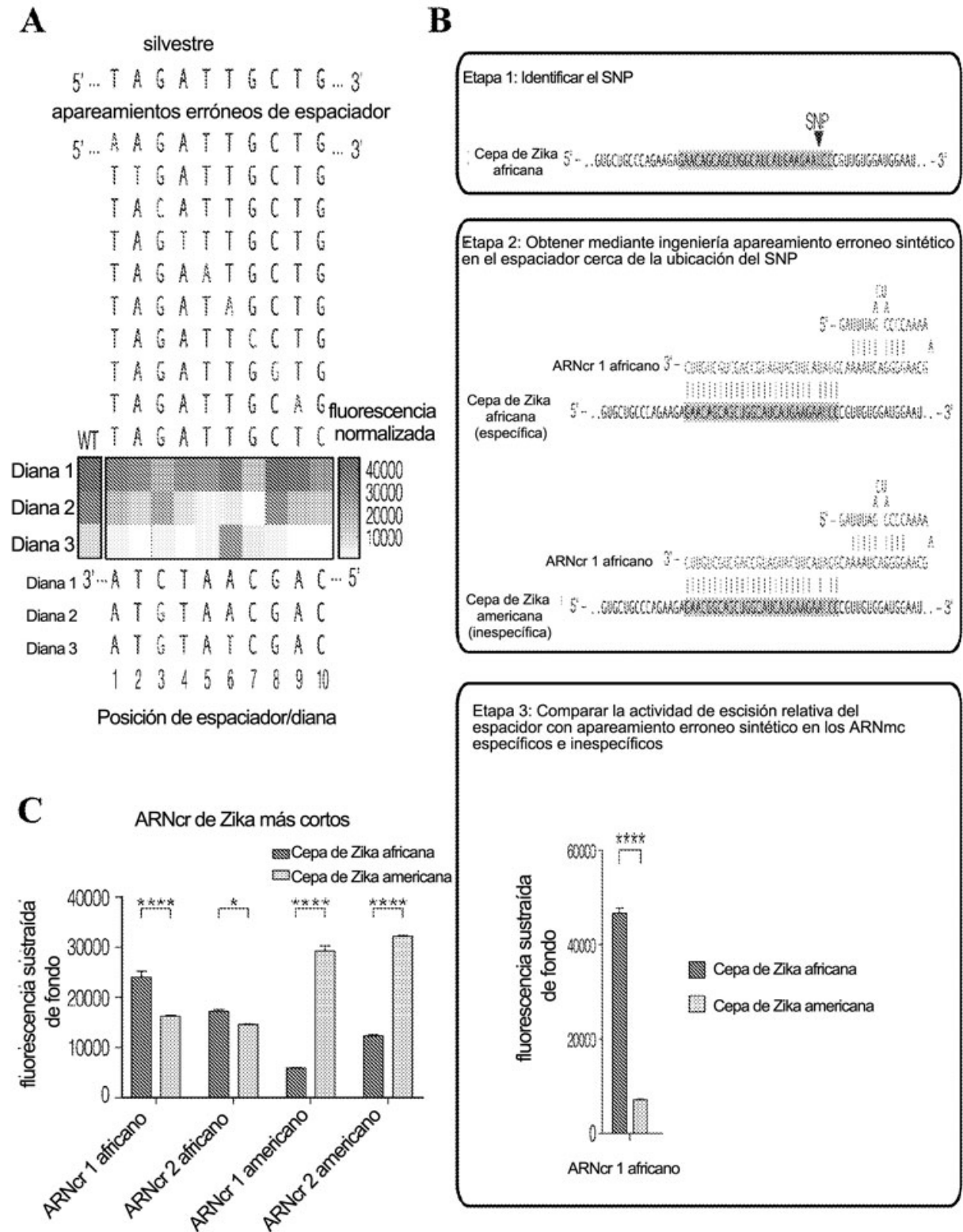
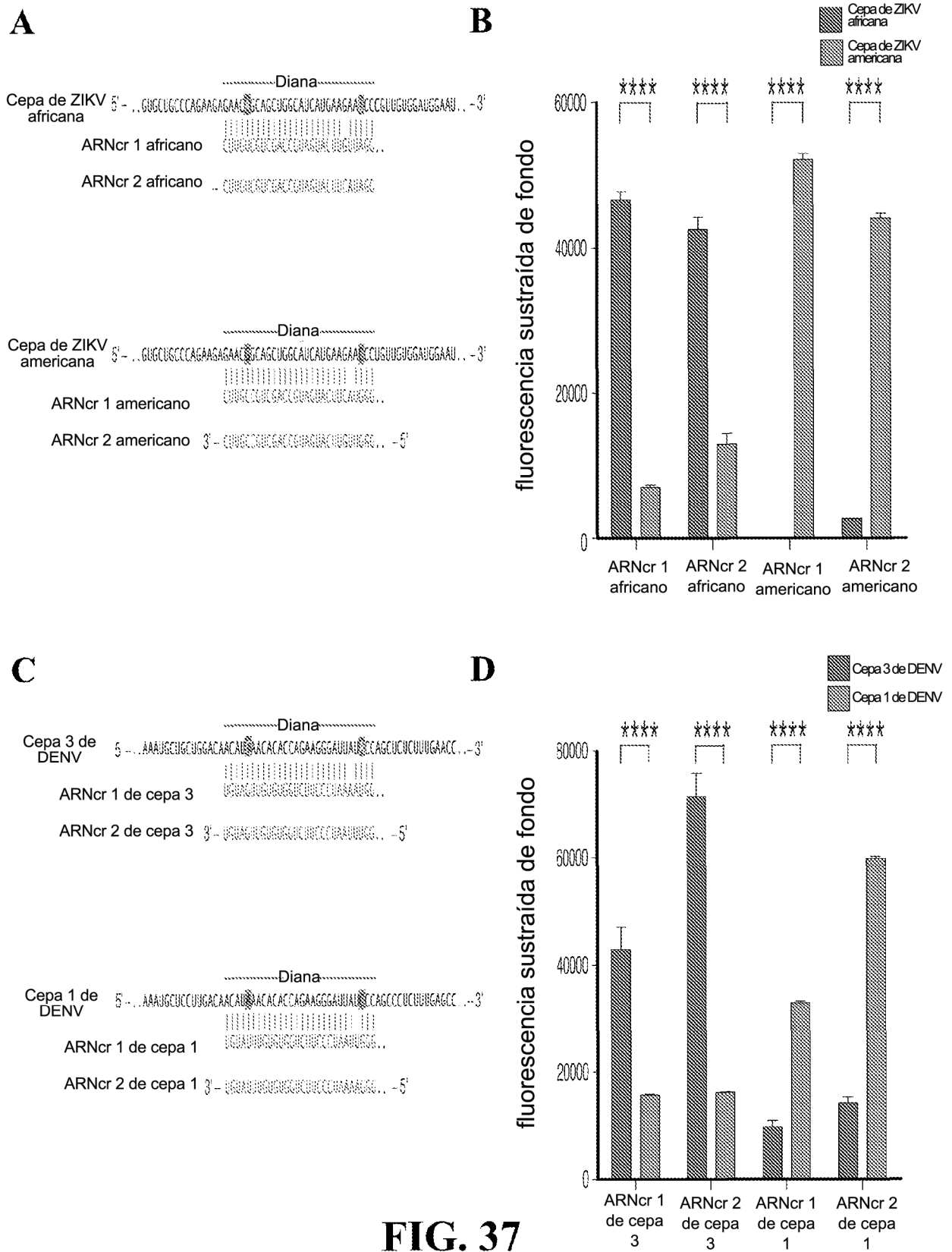


FIG. 36



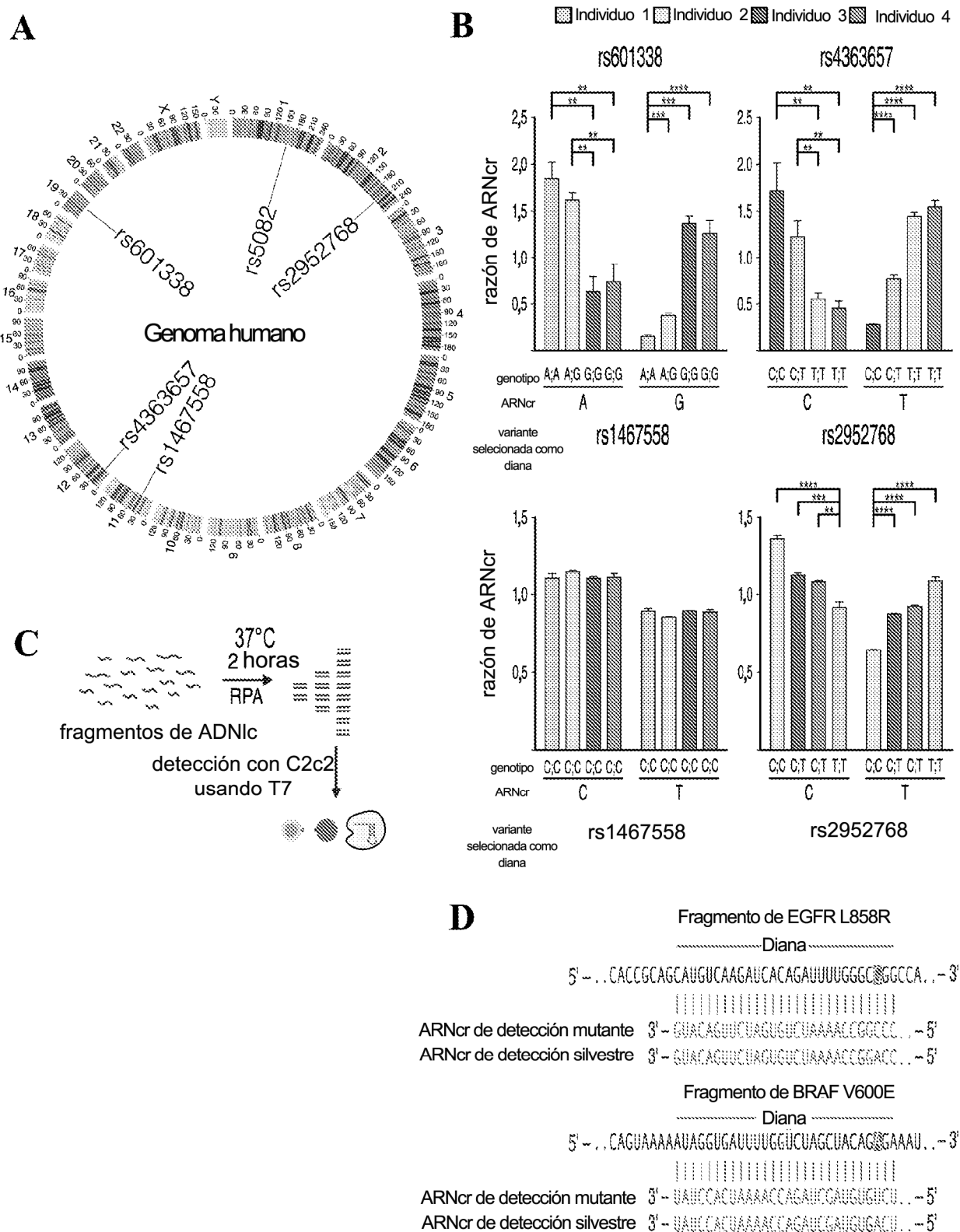
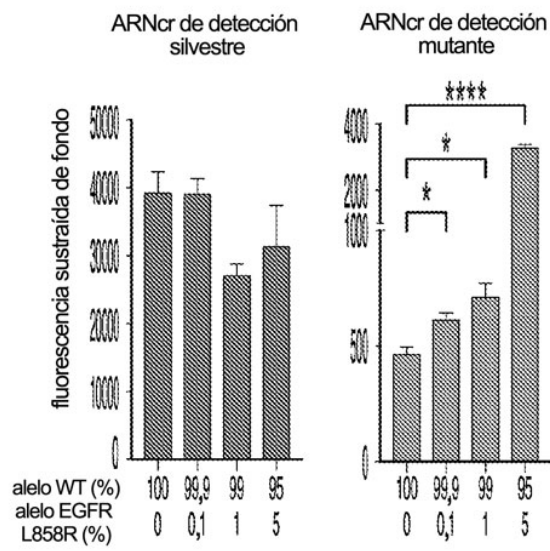


FIG. 38

A



B

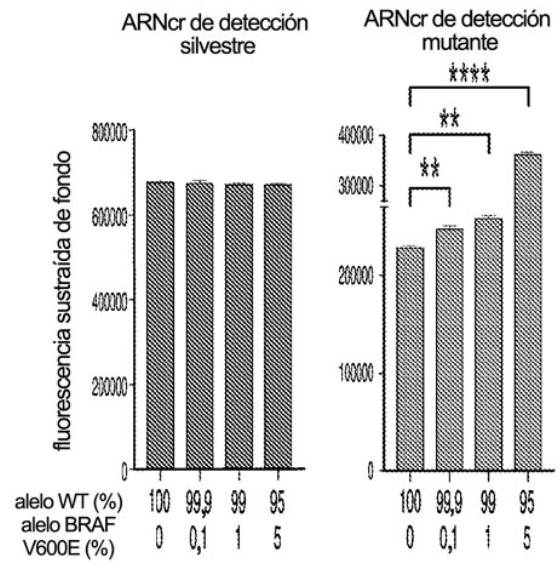
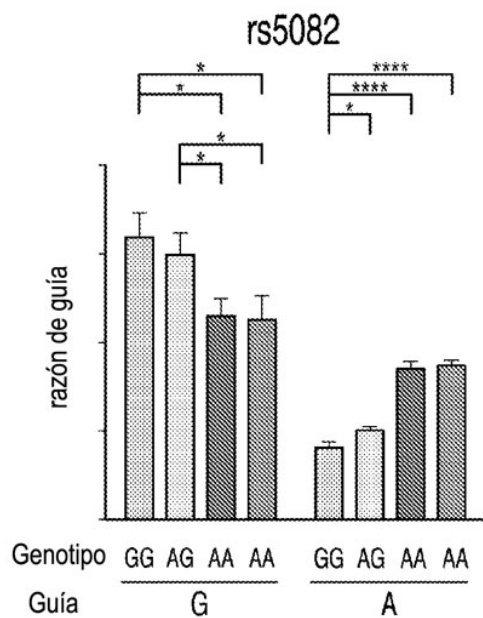


FIG. 39

A

Individuo 1 Individuo 2 Individuo 3 Individuo 4



B

Individuo 1 Individuo 4

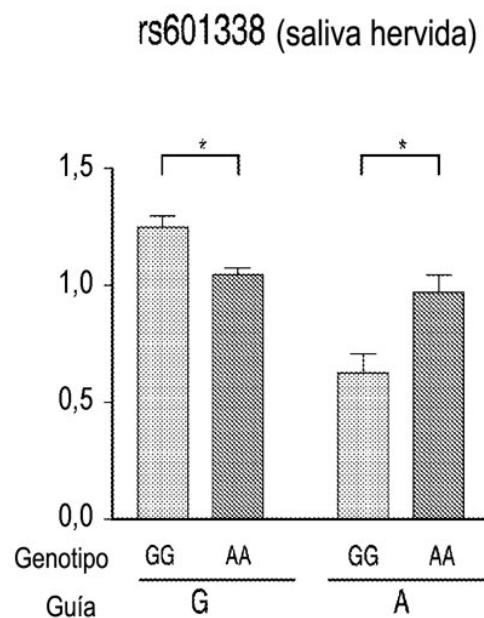
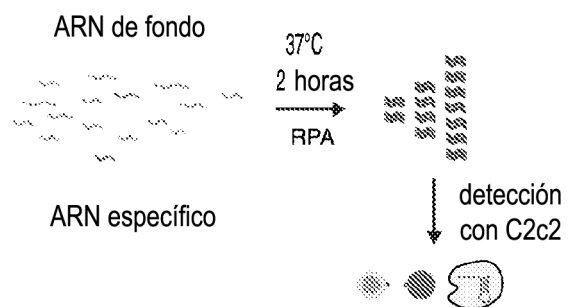


FIG. 40

A



B

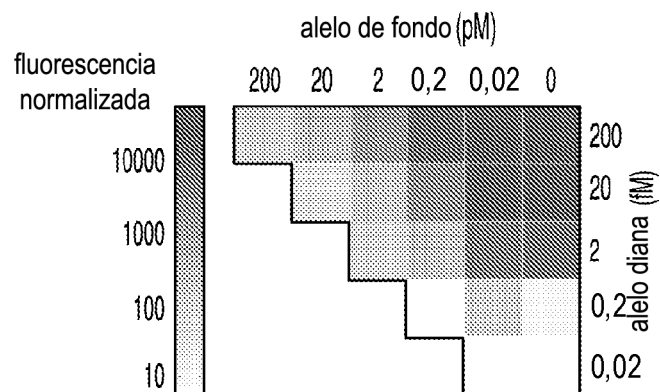


FIG. 41

Una sonda de sensor de ARN sintetizada con un colorante Cy5
diferente también funciona

Cy5 Ex/Em 638/668nm

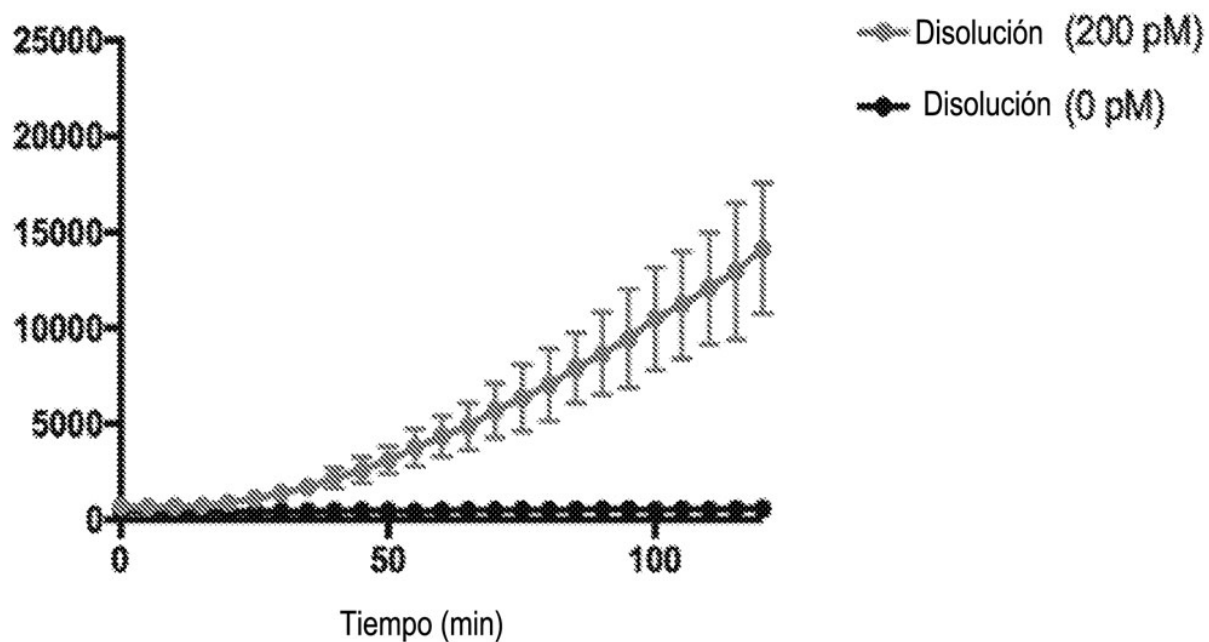


FIG. 42

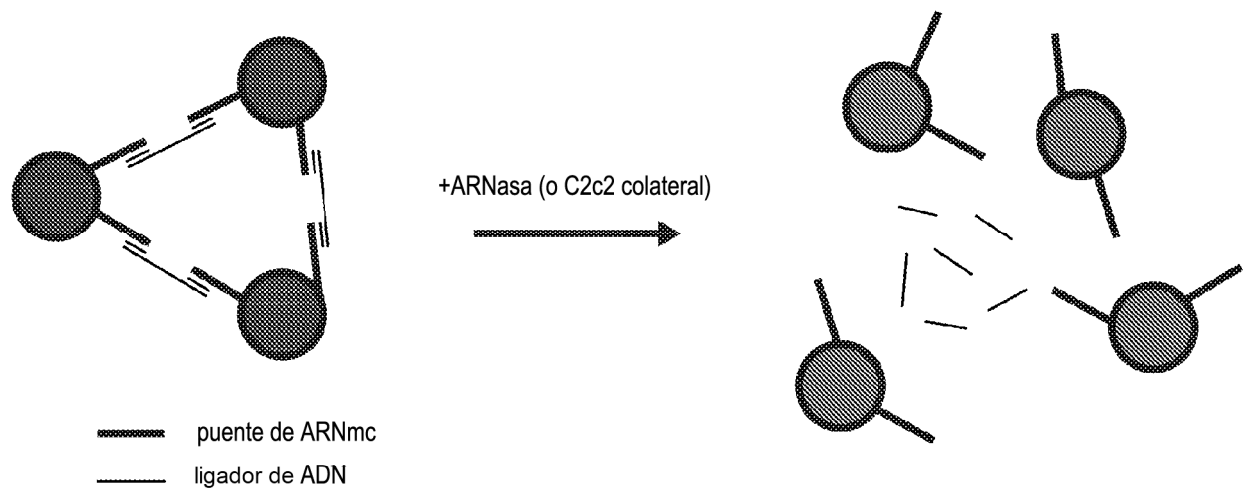
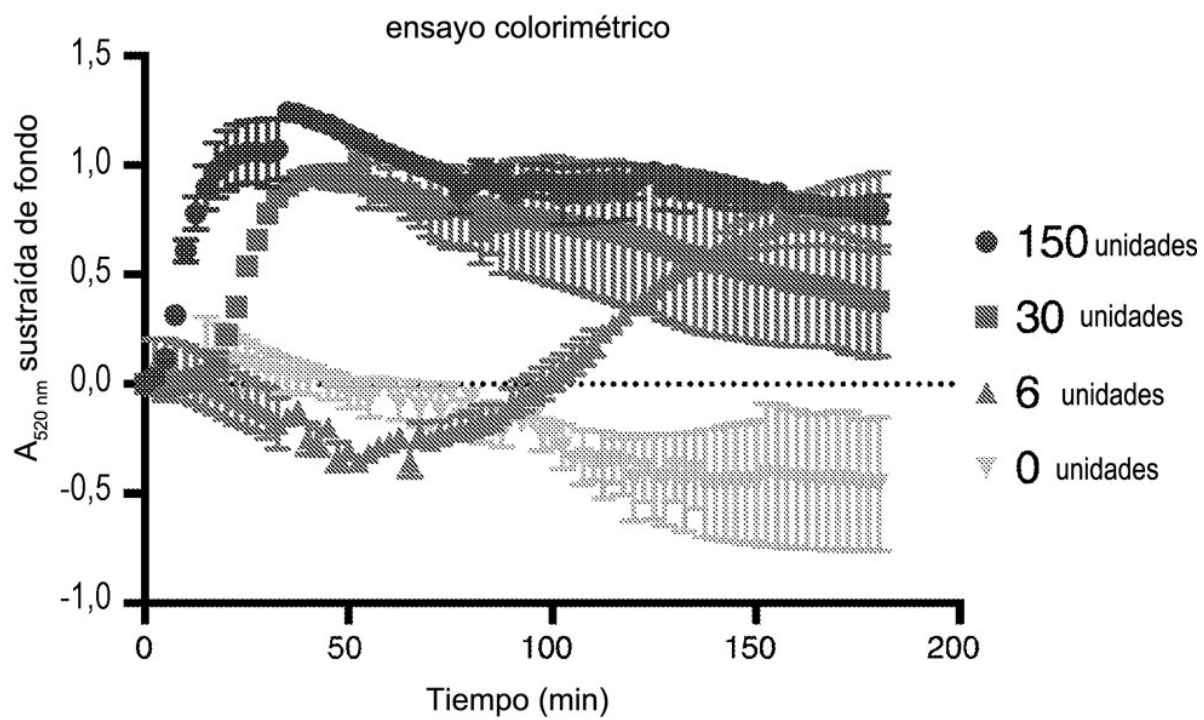


FIG. 43

**FIG. 44**

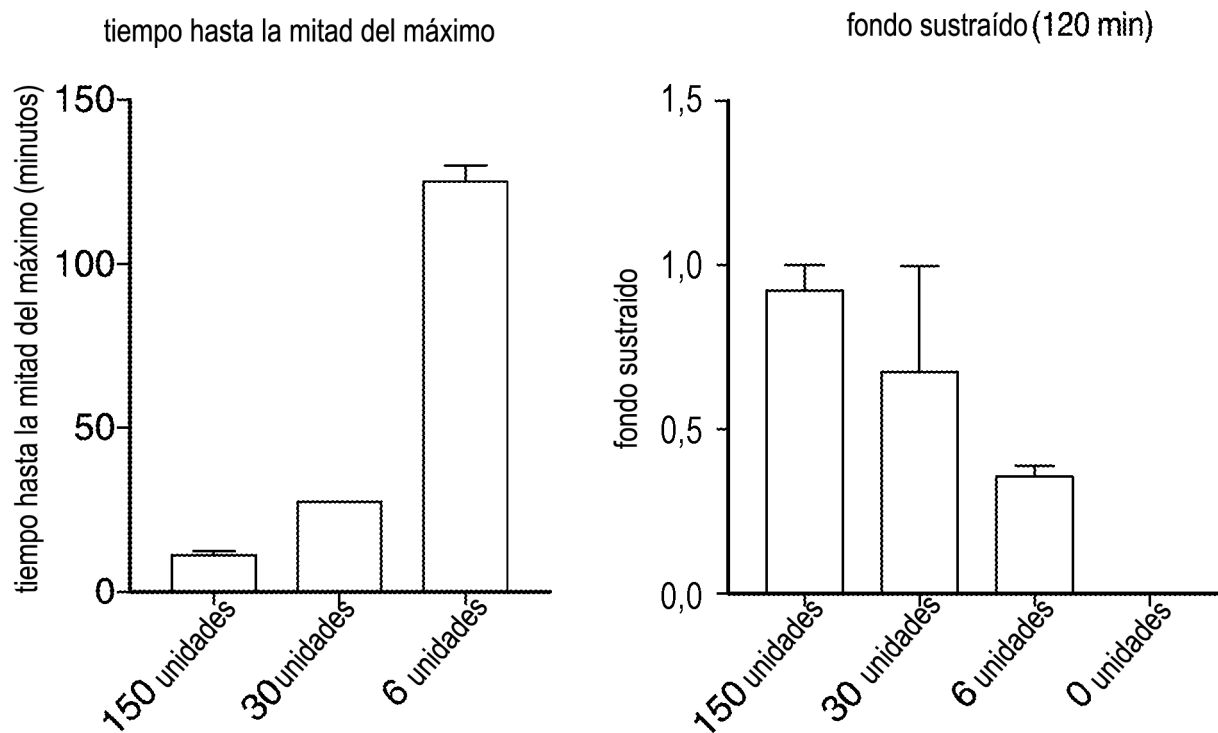


FIG. 45

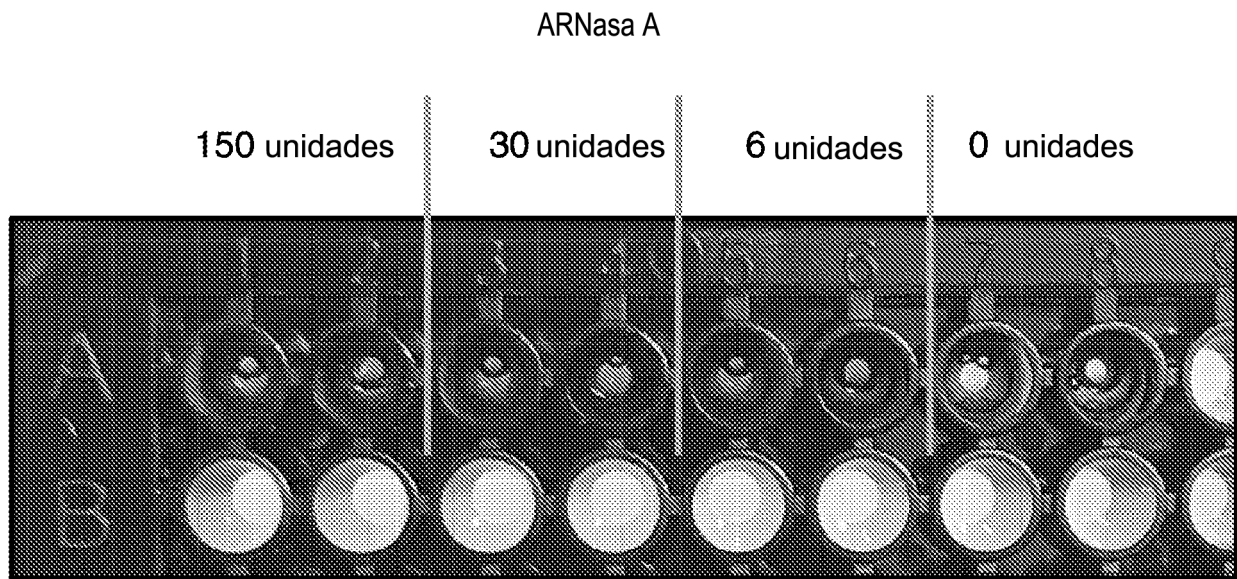
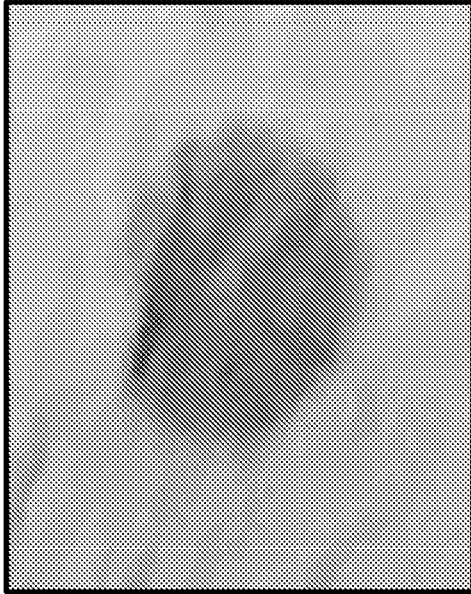
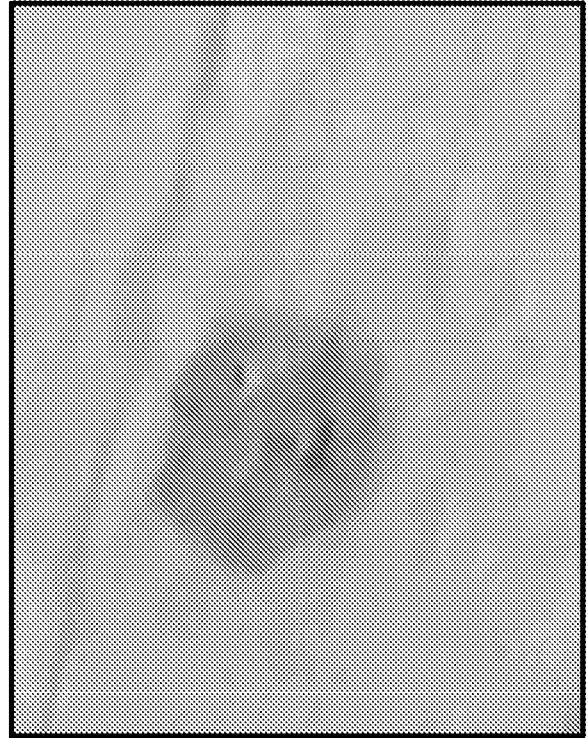


FIG. 46



30 unidades de ARNasa A



Sin ARNasa A

FIG. 47

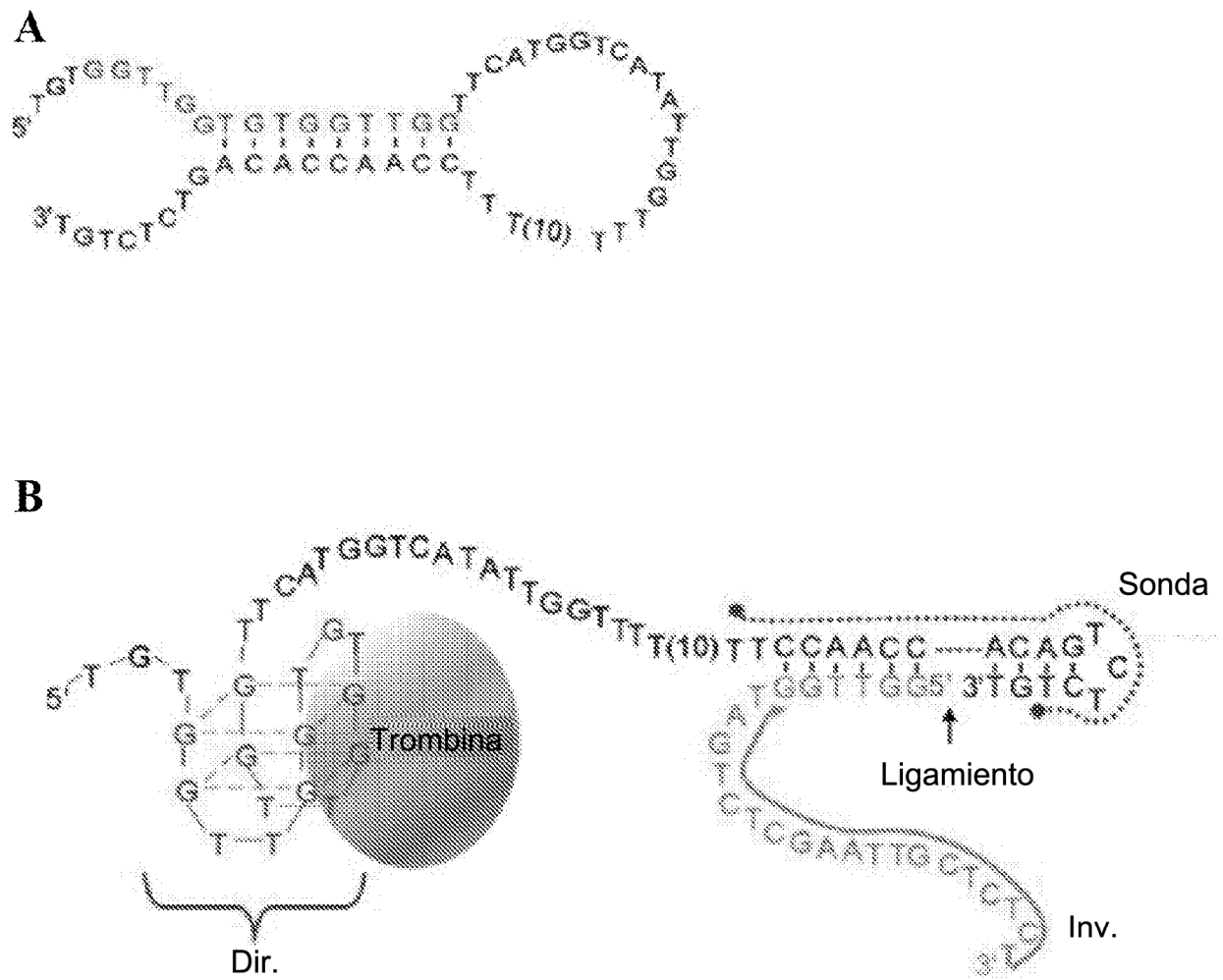


FIG. 48

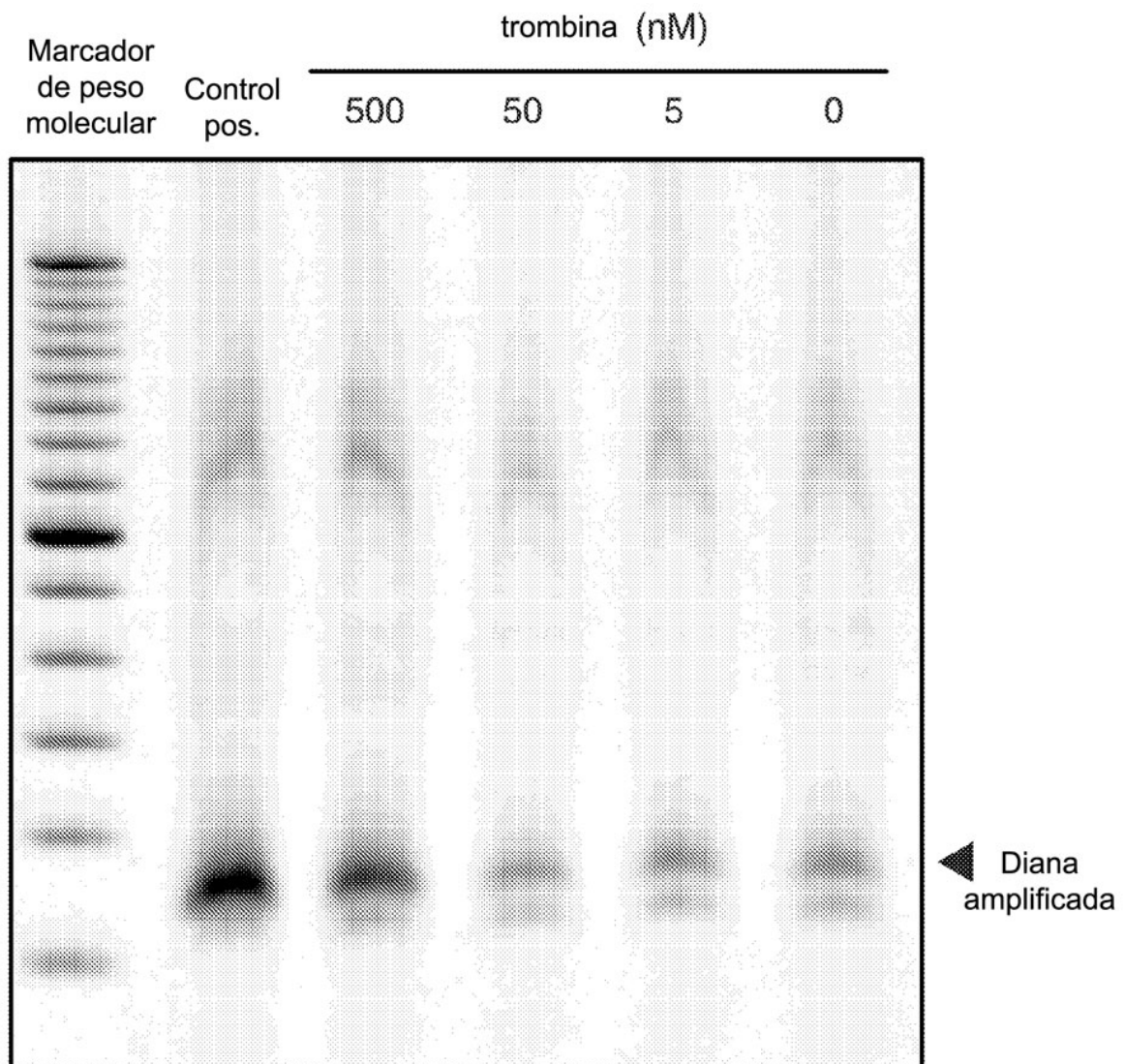
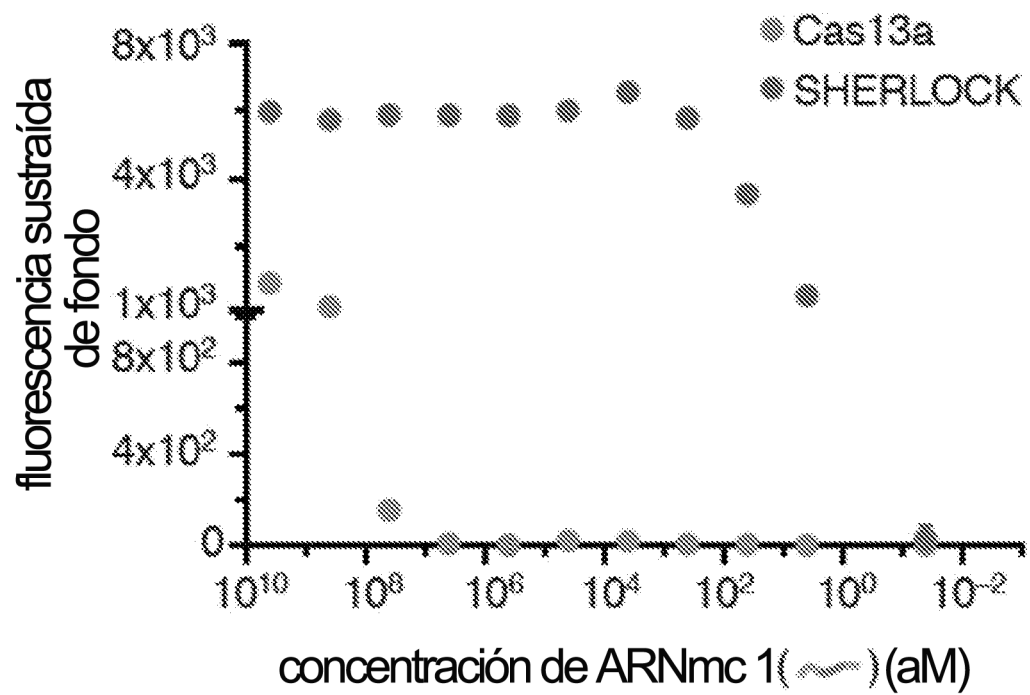


FIG. 49

	586	603	1276	1290
c2c2_Leptotnchia shahii	IRKFTKIGTN ERNRILHA		SIRNYISHFYIVRNP	
c2c2-5 Lachnospiraceae bacterium MA2020	LYSLKSNLYS MRNSSFHF		IFRNEIDHFFHYFYDR	
c2c2-6 Lachnospiraceae	LTDLKDVLYS MRNDSFHY		ELRNYIEHFYRYSSF	
c2c2-7 [Clostridium] ammophilum DSM 10710	ADDLRKAIYS LRNETFHF		DVRKYVDHFKYYATS	
c2c2-8 Camobacterium gallinarum DSM 4847	IWALRGSVQQ IRNEIFHS		KIRNQTAHLSVLQLE	
c2c2-9 Camobacterium gallinarum DSM 4847	LWAIRGAVQR VRNQIFHQ		EIRNNIAHLHVLKND	
c2c2-10 Paludibacter propionicigenes WB4	LWGIRGAVQQ IRNNVNHY		DIRNHIAHFNYLTkd	
c2c2-11 Listena weihenstephanensis FSL R9-0317	IWAIRGSIQQ IRNEVYHC		NARNHIAHLNLYSLK	
c2c2-12 Listenaceae bacterium FSL M6-0635	IWAIRGSIQQ IRNEVYHC		NARNHIAHLNLYSLK	
c2c2-13 Leptotrichia wadei F0279	FANIDEAISS IRHGIVHF		YIRNYIAHFNYIPHA	
c2c2-14 Rhodobacter capsulatus SB 1003	VFALLRYLRG CRNQTFHL		QTRKDLAHFNVLDRA	
c2c2-15 Rhodobacter capsulatus R121	VFALLRYLRG CRNQTFHL		QTRKDLAHFNVLDRA	
c2c2-16 Rhodobacter capsulatus DE442	VFALLRYLRG CRNQTFHL		QTRKDLAHFNVLDRA	
c2-3 L wadei (Lw2)	FANIDEAISS IRHGIVHF		YIRNYIAHFNYIPHA	
c2-4 Listeria seeligeri	SWGLRGAIAP IRNEIHL		EKRNNISHFNYLNQK	
		↑↑ ↑	↑↑ ↑	

FIG. 50

**FIG. 51**

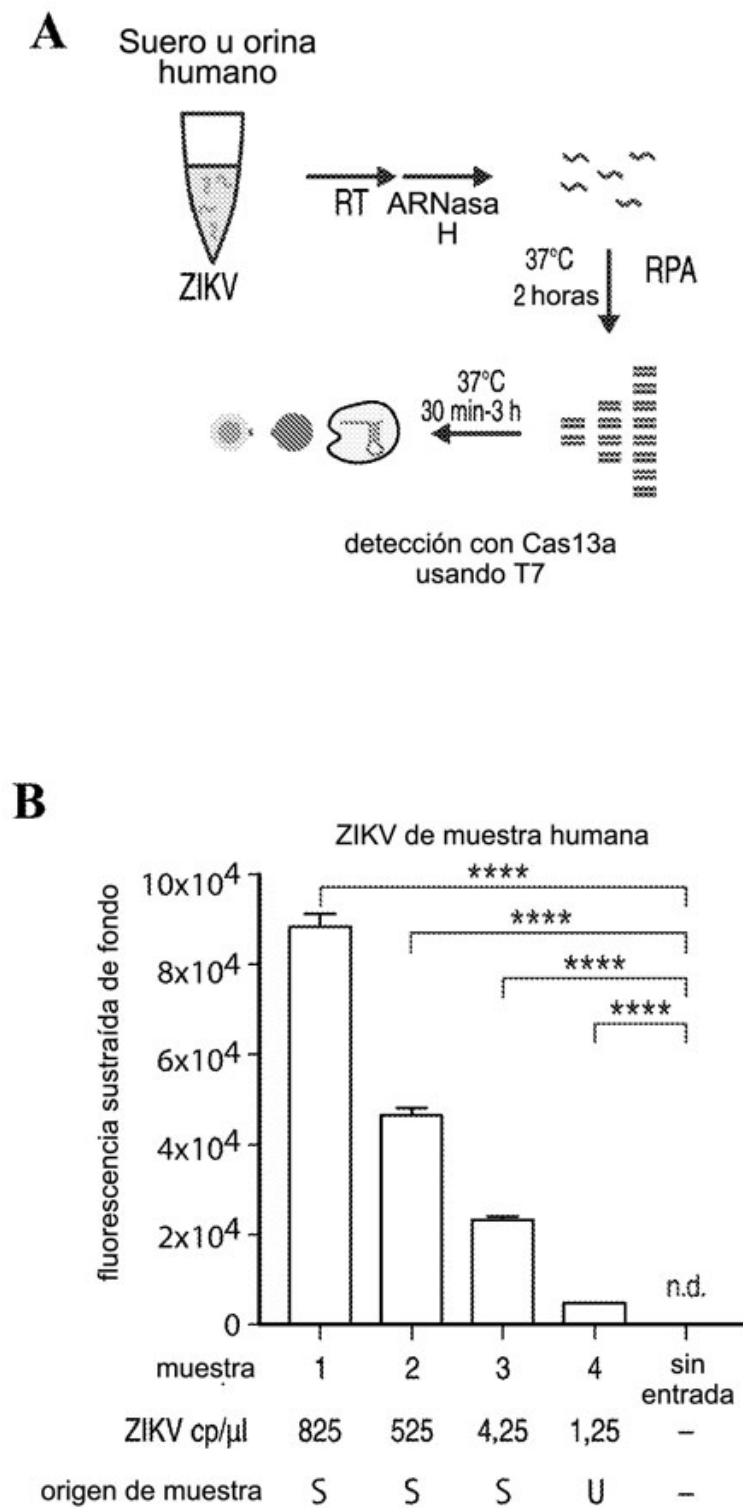


FIG. 52

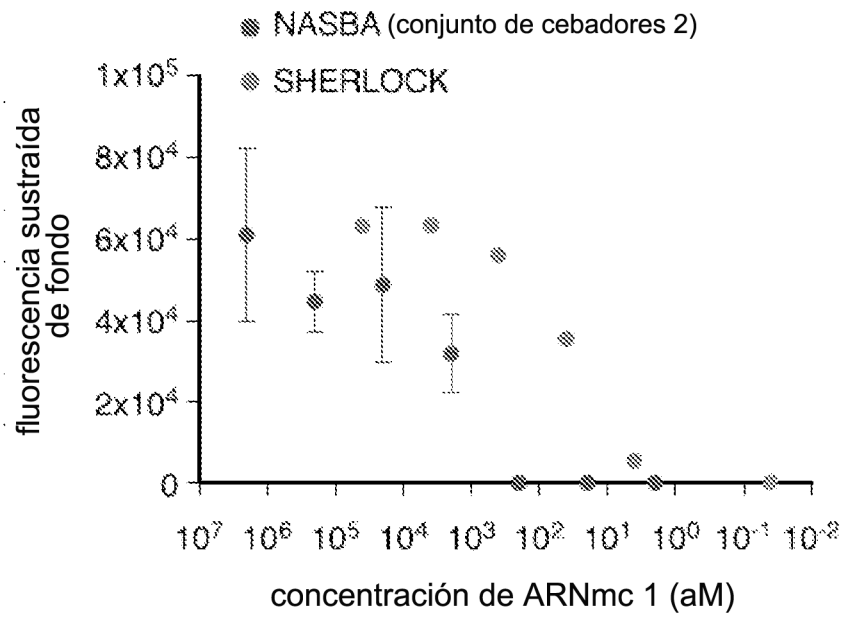


FIG. 53

A

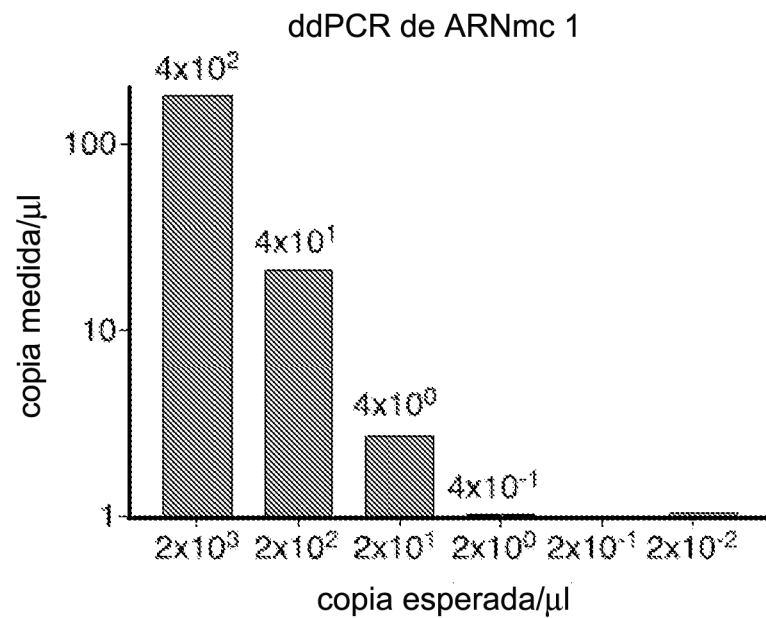
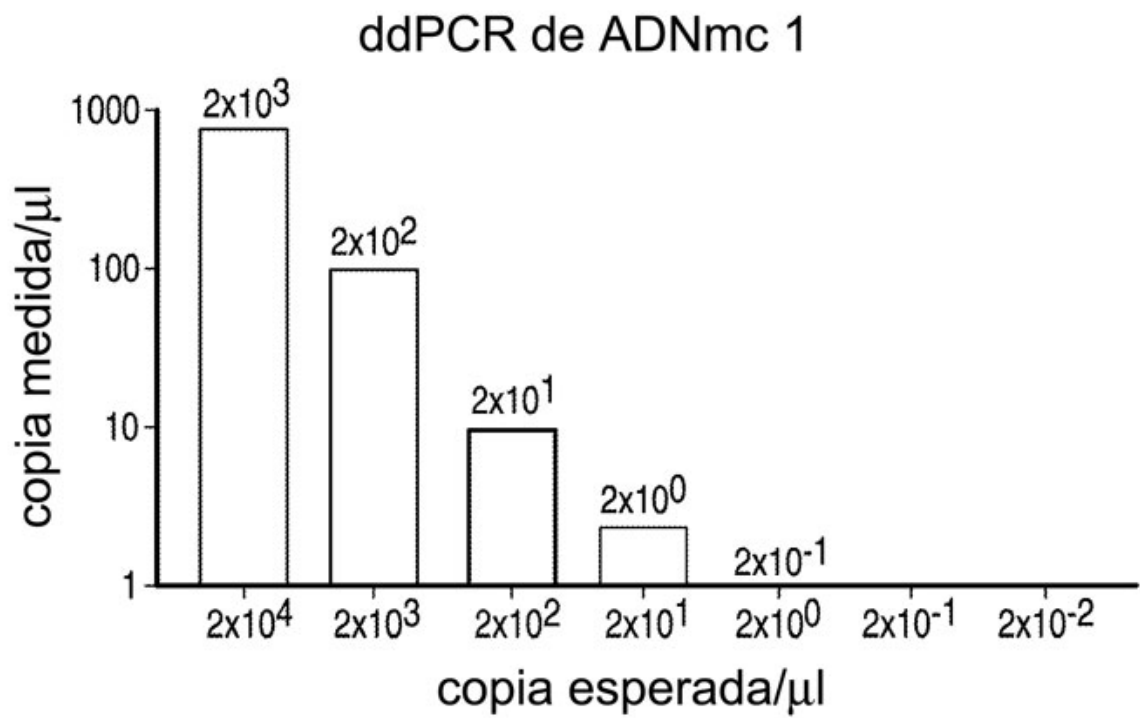
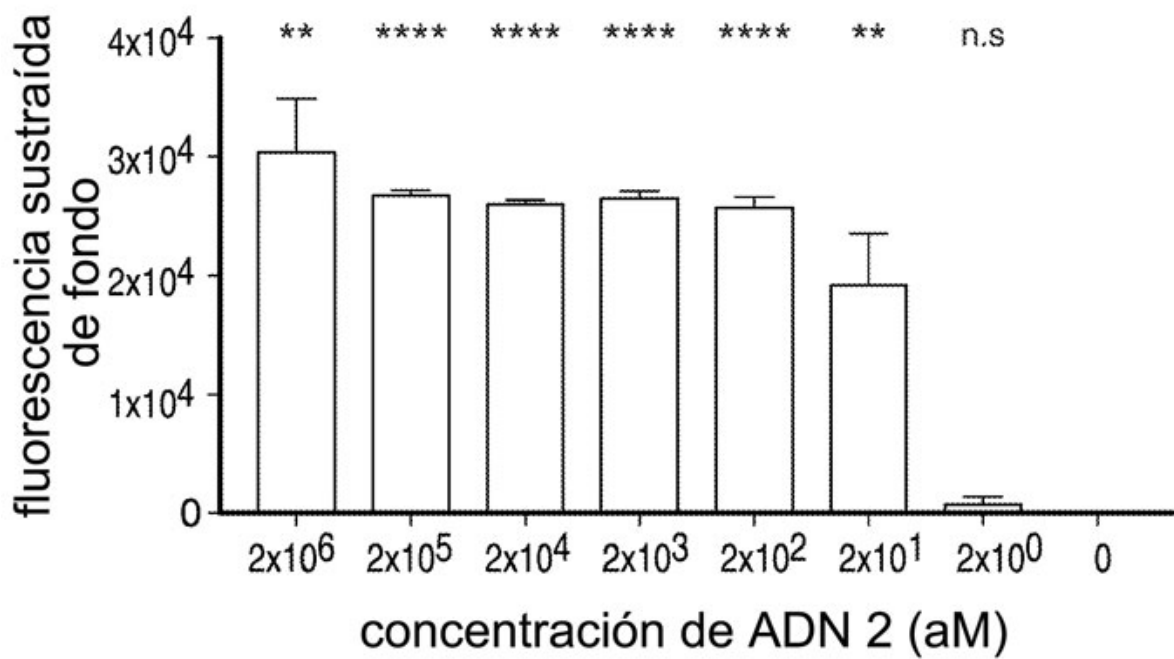


FIG. 54

B**C****FIG. 54 (continuación)**

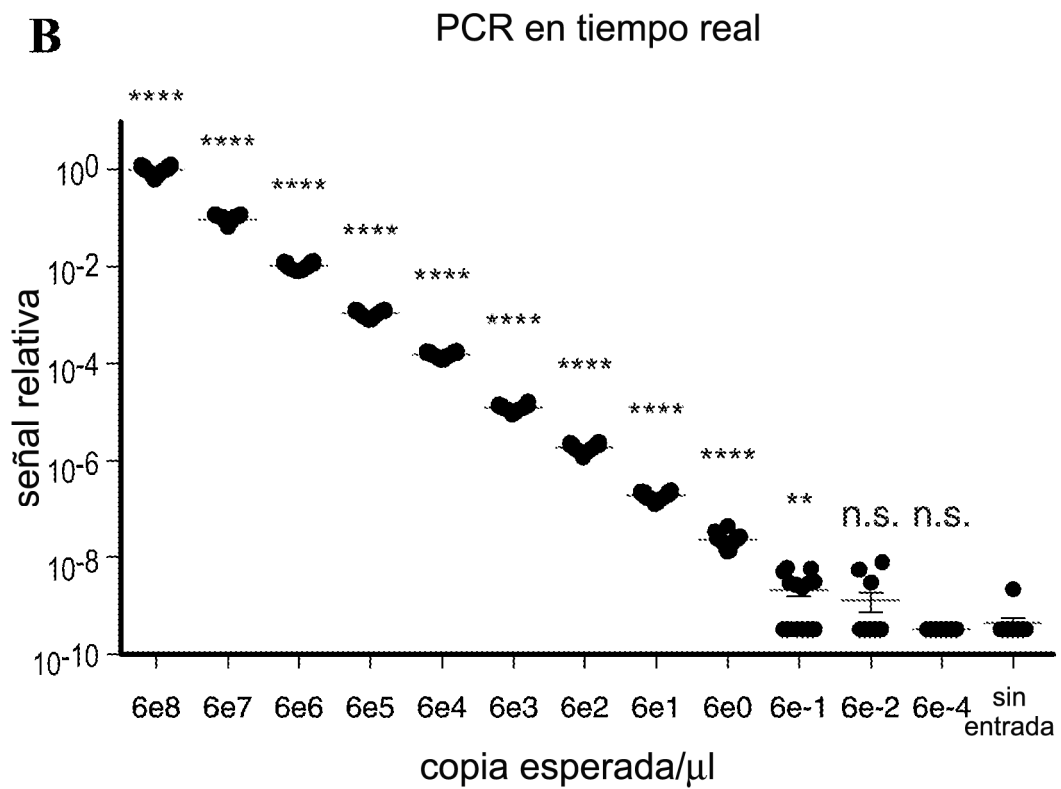
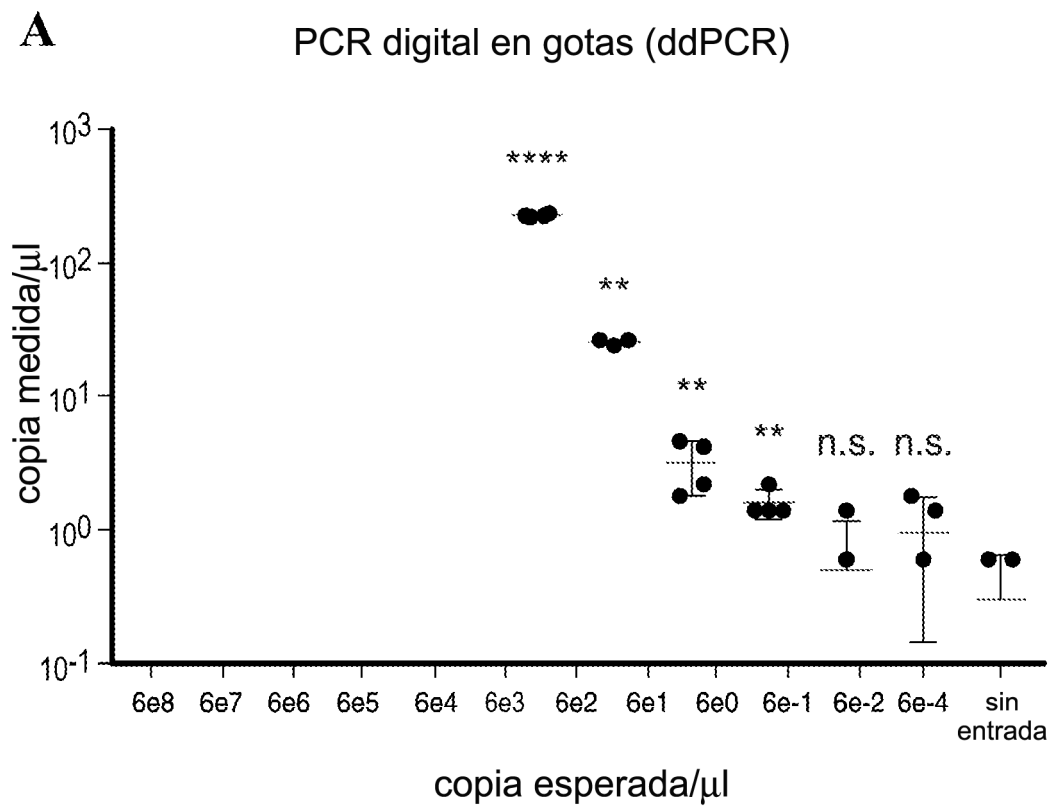


FIG. 55

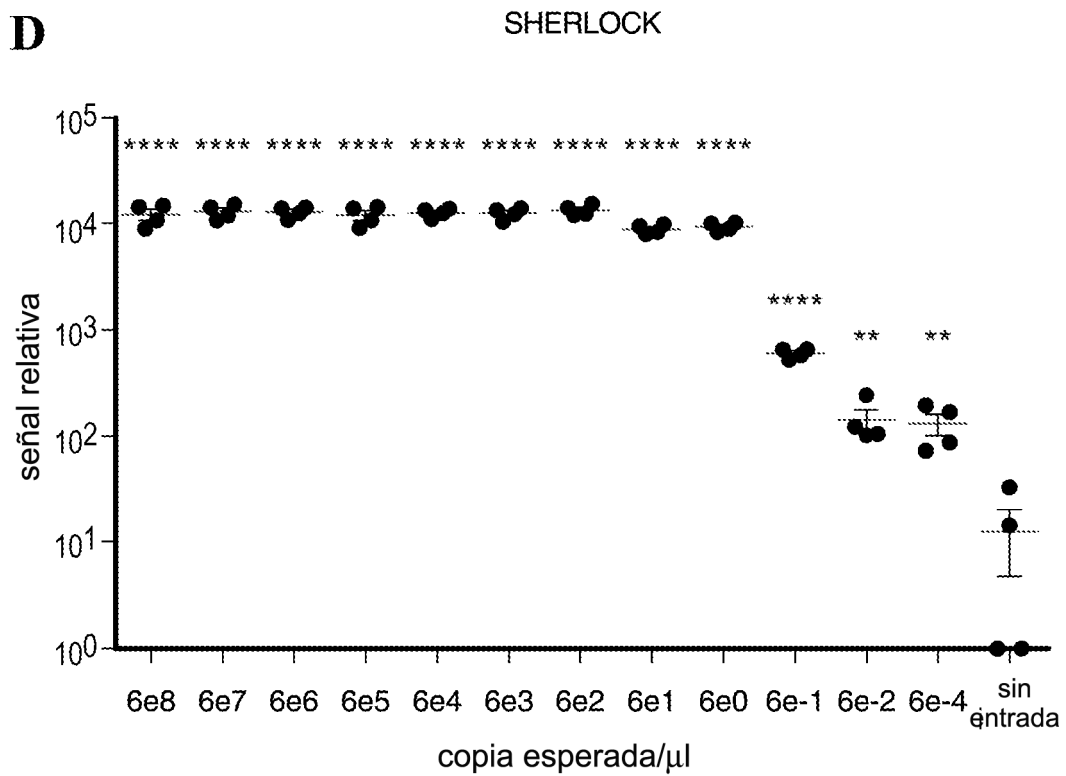
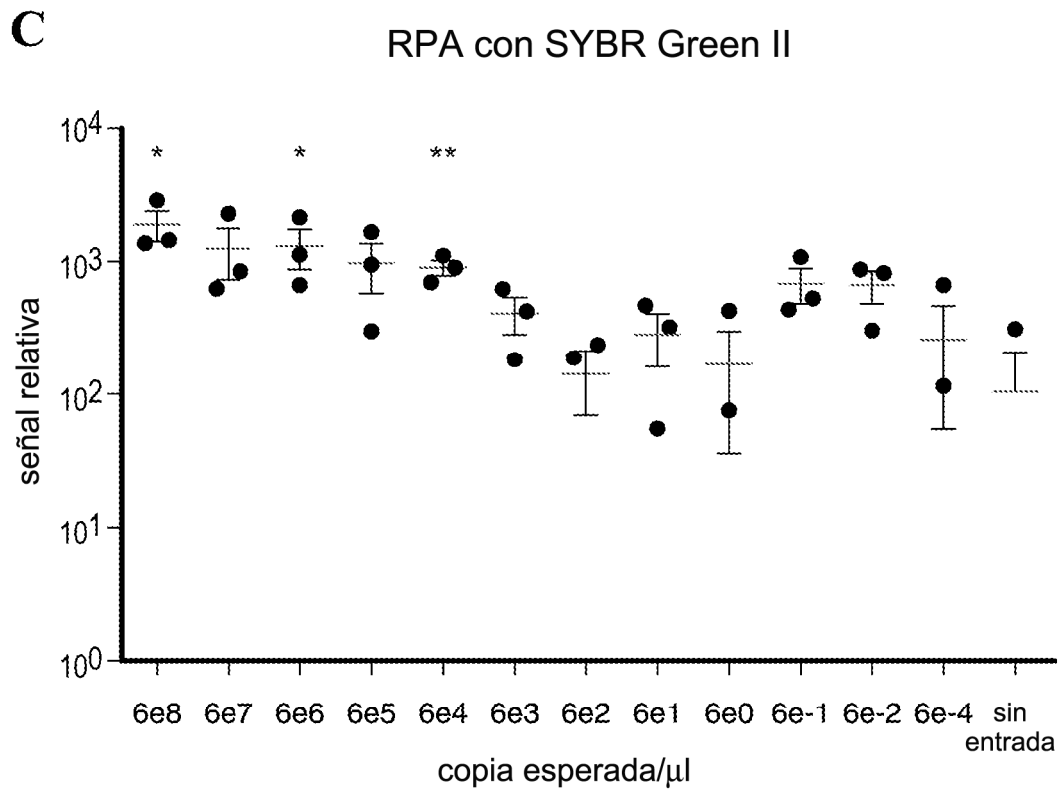
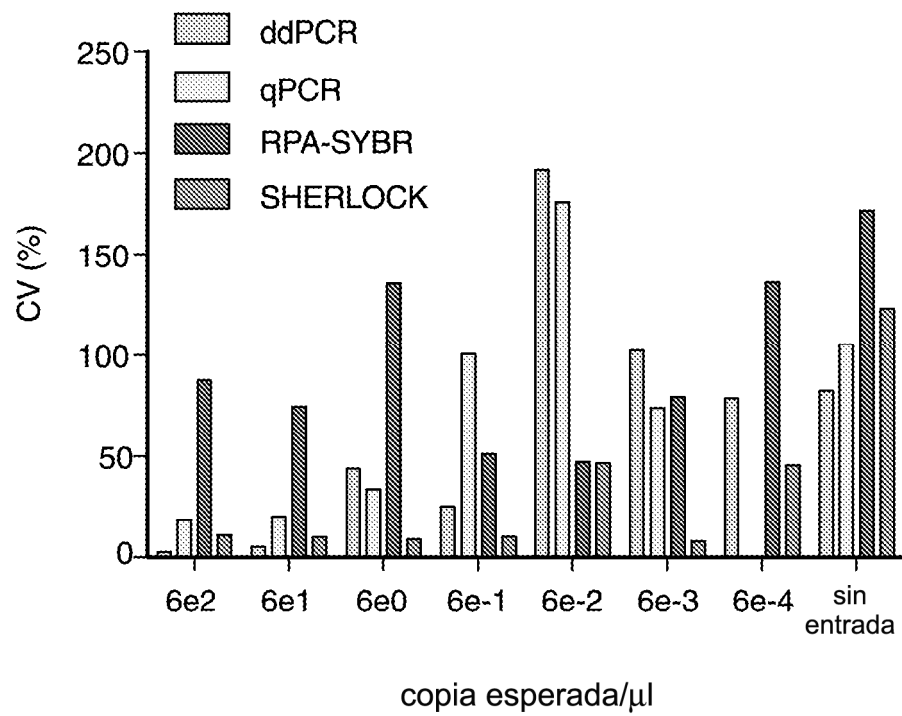


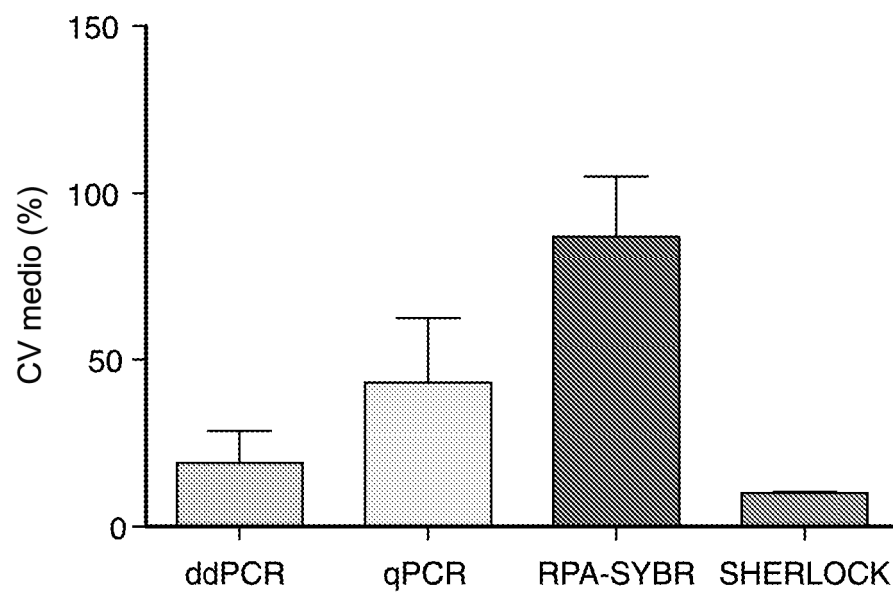
FIG. 55 (continuación)

E

Coeficiente de variación

**F**

análisis de CV medio

**FIG. 55 (continuación)**

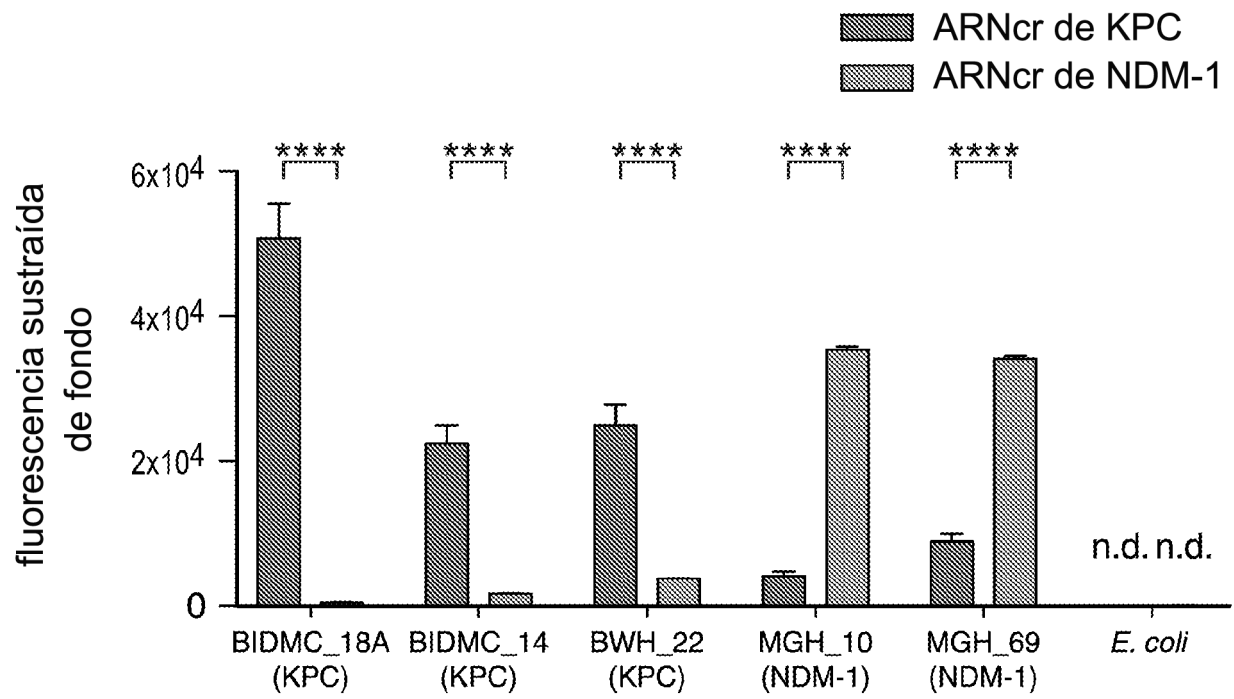


FIG. 56

A

```

      UC
      A A
AAAACCCC GAUUUAG -5'
A   |||| ||||
GCAAGGGGACUAAAACUAGAUUGCUGUUCUACCAAGUAAUCCAU -3' 28nt
                                UAGAUUGCUGUUCUACCAAGUAA -3' 23nt
                                UAGAUUGCUGUUCUACCAAG -3' 20nt

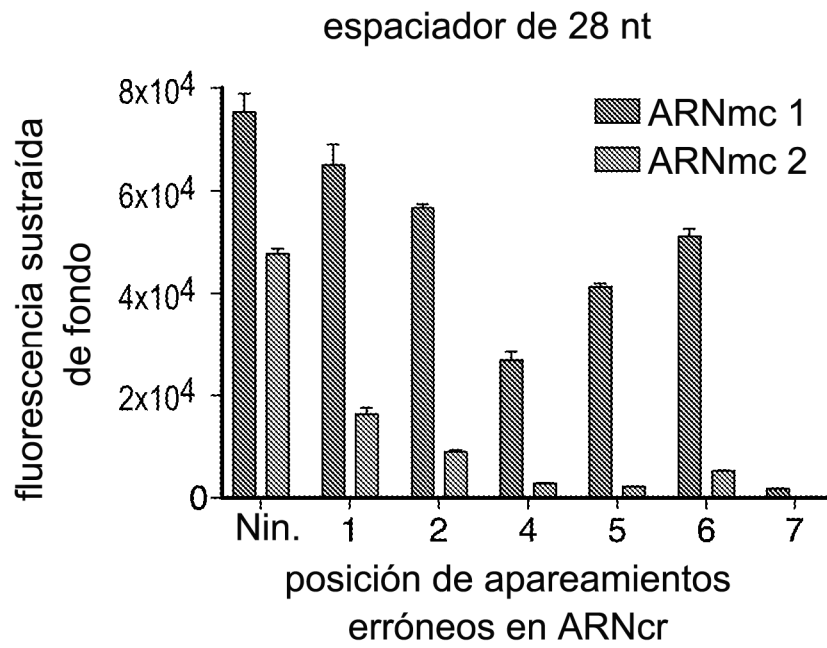
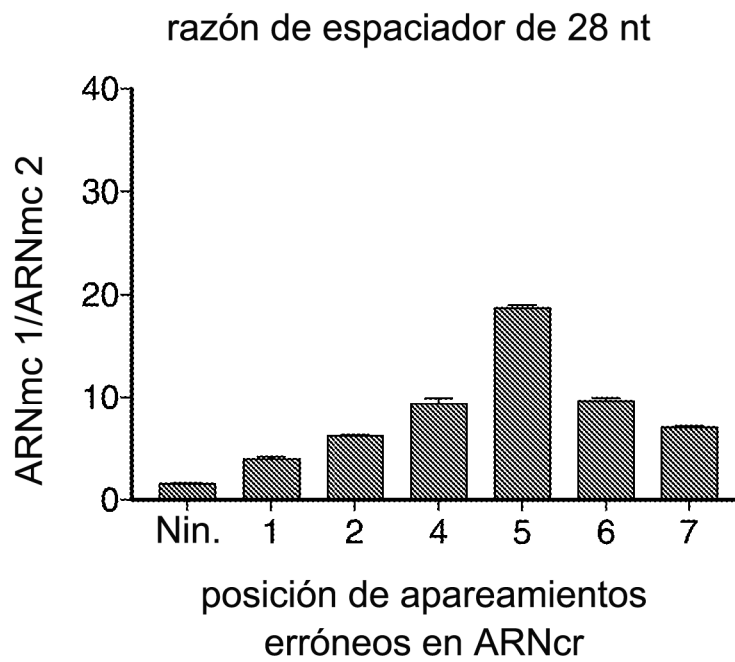
```

3' - . . AUCUAACGACAAGAUGGUUCAUUAGGUA . . - 5' ARNmc 1

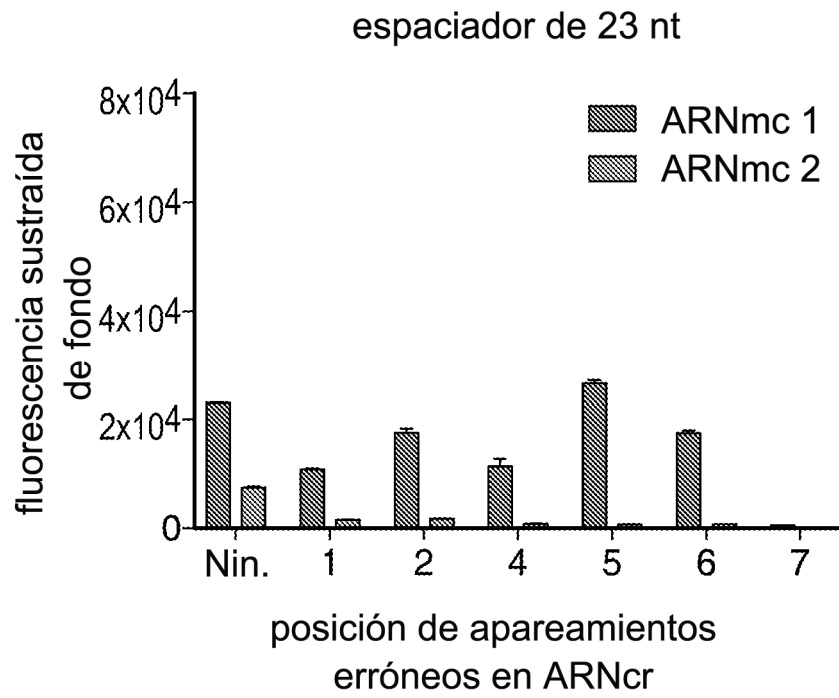
3' - . . AUUAACGACAAGAUGGUUCAUUAGGUA . . - 5' ARNmc 2

UAGAUUGCUGUUCUACC . .	Ninguna	
AAGAUUGCUGUUCUACC . .	1	
UUGAUUGCUGUUCUACC . .	2	posición de
UAGUUUGCUGUUCUACC . .	4	apareamientos
UAGAAUGCUGUUCUACC . .	5	erróneos
UAGAUAGCUGUUCUACC . .	6	en ARNcr
UAGAUUCCUGUUCUACC . .	7	

FIG. 57

B**C****FIG. 57 (continuación)**

D



E

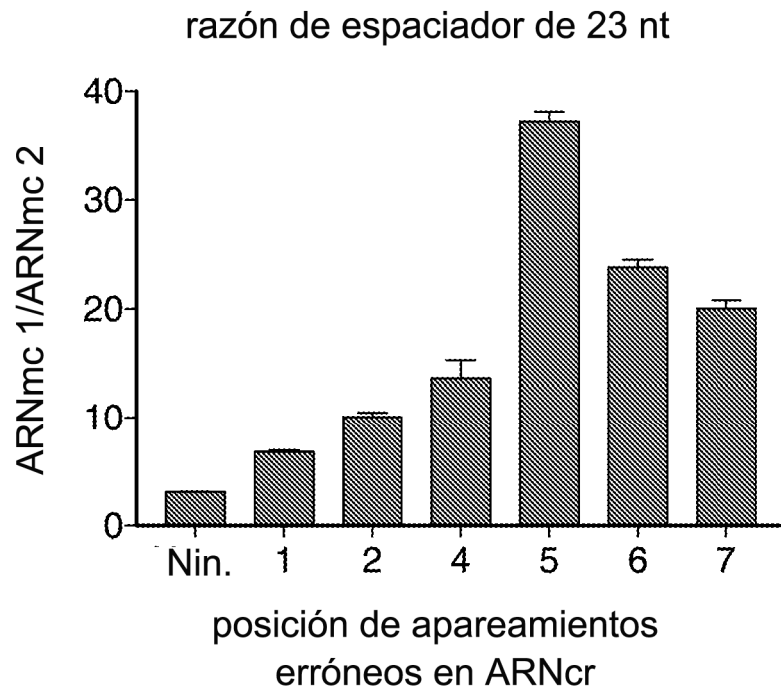
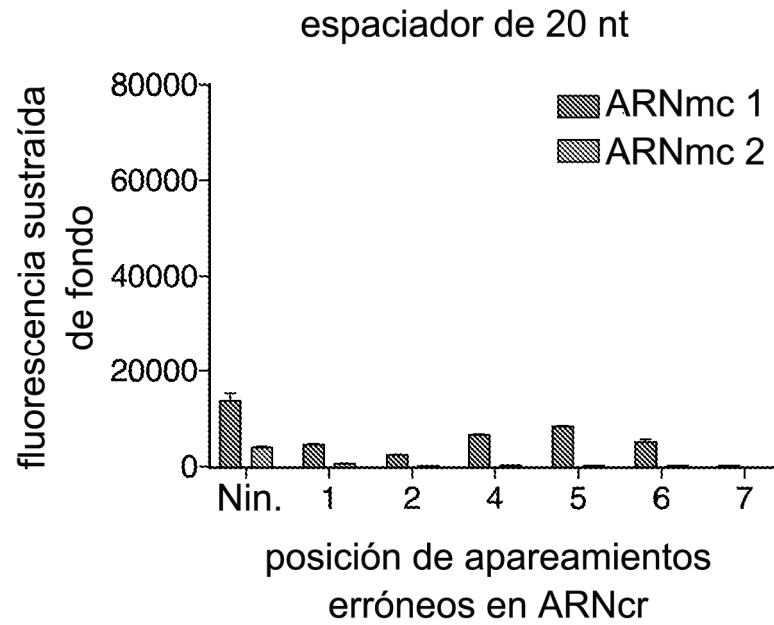
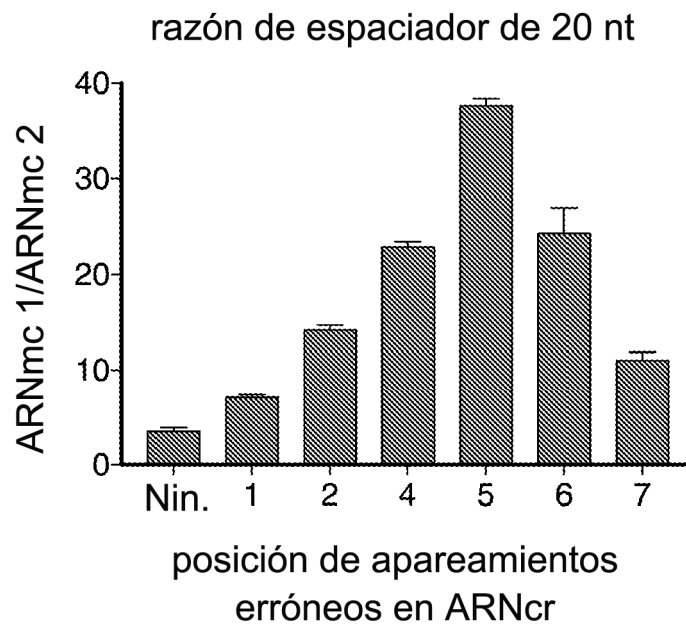


FIG. 57 (continuación)

F**G****FIG. 57** (continuación)

A

3' - . . CUCAUCUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 1

3' - . . CUCAUUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 2

posición de apareamientos erróneos en diana 3	. . AAGAUUGCUGUUCUACCAAGUAAUCCAU	1	posición de apareamientos erróneos en ARNcr
	. . UUGAUUGCUGUUCUACCAAGUAAUCCAU	2	
	. . UAGUUGCUGUUCUACCAAGUAAUCCAU	4	
	. . UAGAAUGCUGUUCUACCAAGUAAUCCAU	5	
	. . UAGAUAGCUGUUCUACCAAGUAAUCCAU	6	
	. . UAGAUUCCUGUUCUACCAAGUAAUCCAU	7	

3' - . . CUCAUCUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 1

3' - . . CUCAUUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 2

posición de apareamientos erróneos en diana 4	. . CUAGAUUGCUGUUCUACCAAGUAAUCCA	1	posición de apareamientos erróneos en ARNcr
	. . GAAGAUUGCUGUUCUACCAAGUAAUCCA	2	
	. . GUUGAUUGCUGUUCUACCAAGUAAUCCA	3	
	. . GUAGUUGCUGUUCUACCAAGUAAUCCA	5	
	. . GUAGAAUGCUGUUCUACCAAGUAAUCCA	6	
	. . GUAGAUAGCUGUUCUACCAAGUAAUCCA	7	

3' - . . CUCAUCUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 1

3' - . . CUCAUUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 2

posición de apareamientos erróneos en diana 5	. . ACUAGAUUGCUGUUCUACCAAGUAAUCC	2	posición de apareamientos erróneos en ARNcr
	. . AGAAGAUUGCUGUUCUACCAAGUAAUCC	3	
	. . AGUUGAUUGCUGUUCUACCAAGUAAUCC	4	
	. . AGUAGUUGCUGUUCUACCAAGUAAUCC	6	
	. . AGUAGAAUGCUGUUCUACCAAGUAAUCC	7	
	. . AGUAGAUAGCUGUUCUACCAAGUAAUCC	8	

3' - . . CUCAUCUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 1

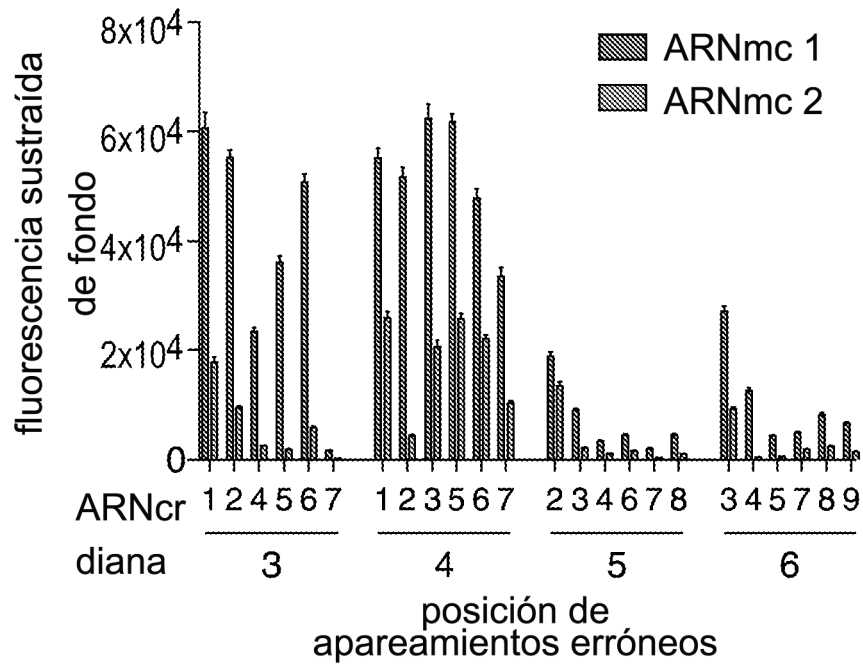
3' - . . CUCAUUAACGACAAGAUGGUUCAUUAGGUA. . - 5' ARNmc 2

posición de apareamientos erróneos en diana 6	. . GACUAGAUUGCUGUUCUACCAAGUAAUCC	3	posición de apareamientos erróneos en ARNcr
	. . GAGAAGAUUGCUGUUCUACCAAGUAAUCC	4	
	. . GAGUUGAUUGCUGUUCUACCAAGUAAUCC	5	
	. . GAGUAGUUGCUGUUCUACCAAGUAAUCC	7	
	. . GAGUAGAAUGCUGUUCUACCAAGUAAUCC	8	
	. . GAGUAGAUAGCUGUUCUACCAAGUAAUCC	9	

FIG. 58

B

Desplazamiento de la ubicación de ARNcr



C

Razón de desplazamiento de la ubicación de ARNcr

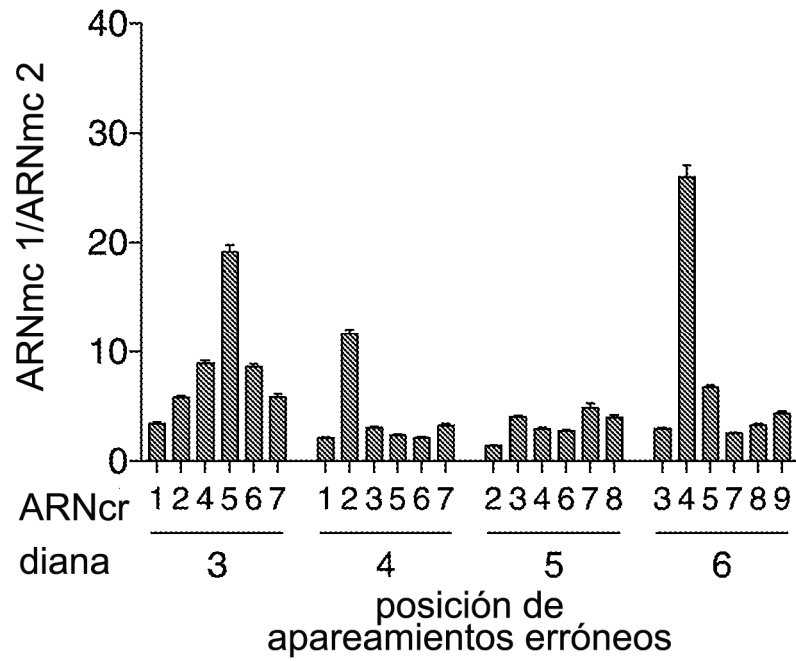


FIG. 58 (continuación)

B

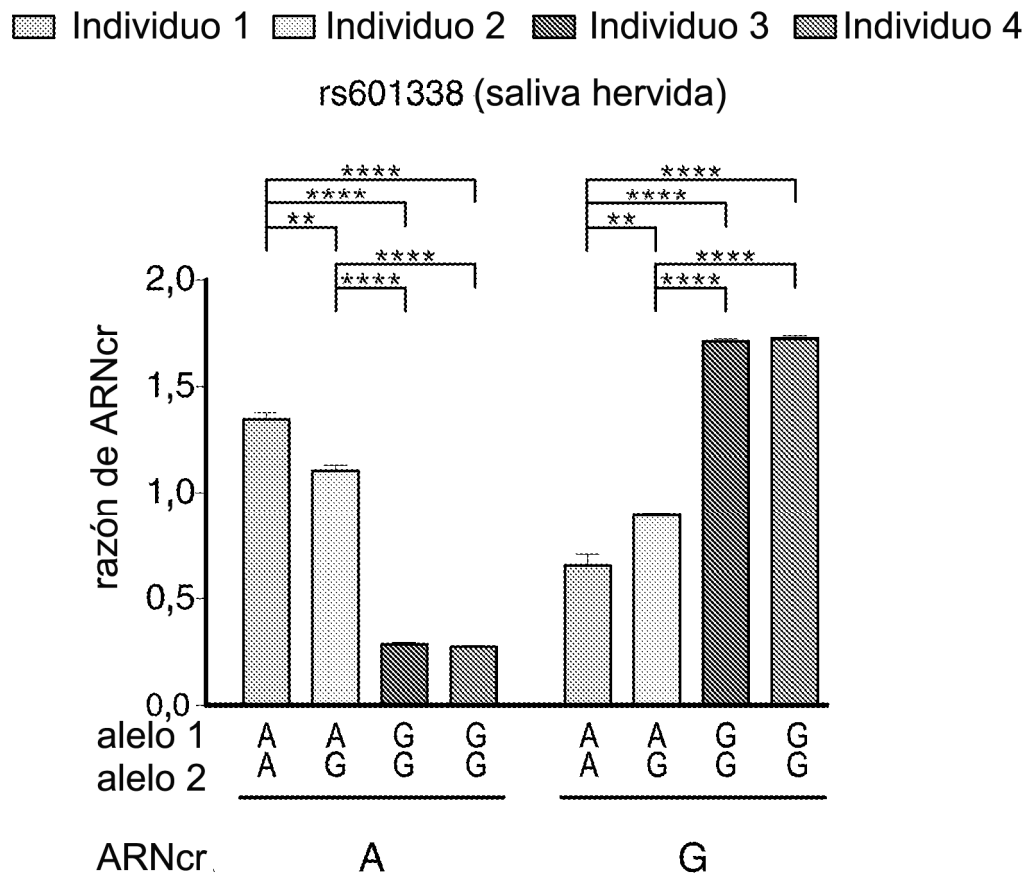
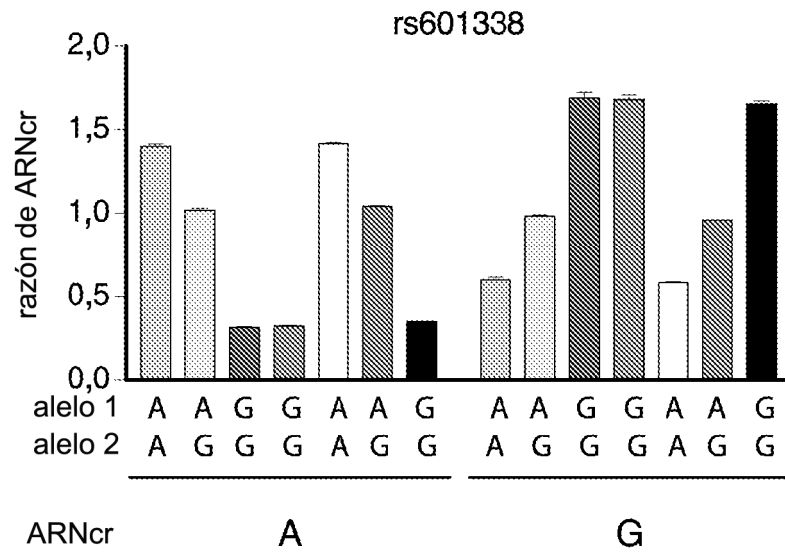


FIG. 59

A

Individuo 1
 Individuo 2
 Individuo 3
 Individuo 4
 patrón de A;A
 patrón de A;G
 patrón de G;G



B

Individuo 1
 Individuo 2
 Individuo 3
 Individuo 4
 patrón de C;C
 patrón de C;T
 patrón de T;T

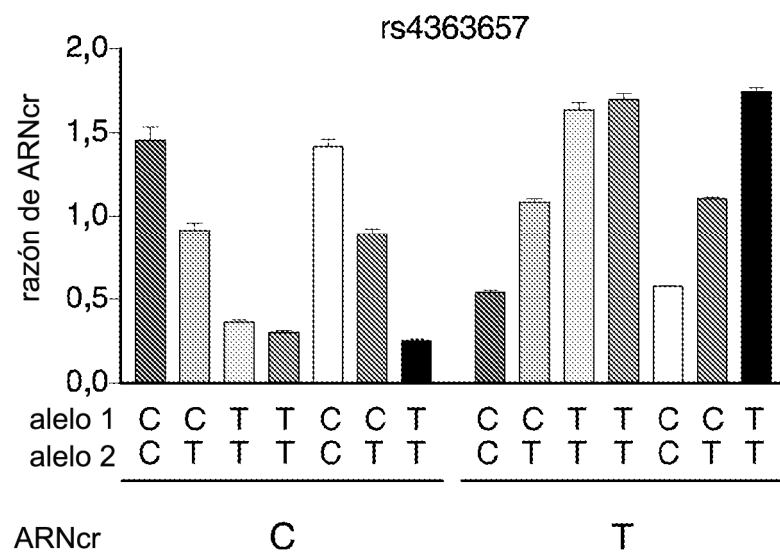


FIG. 60

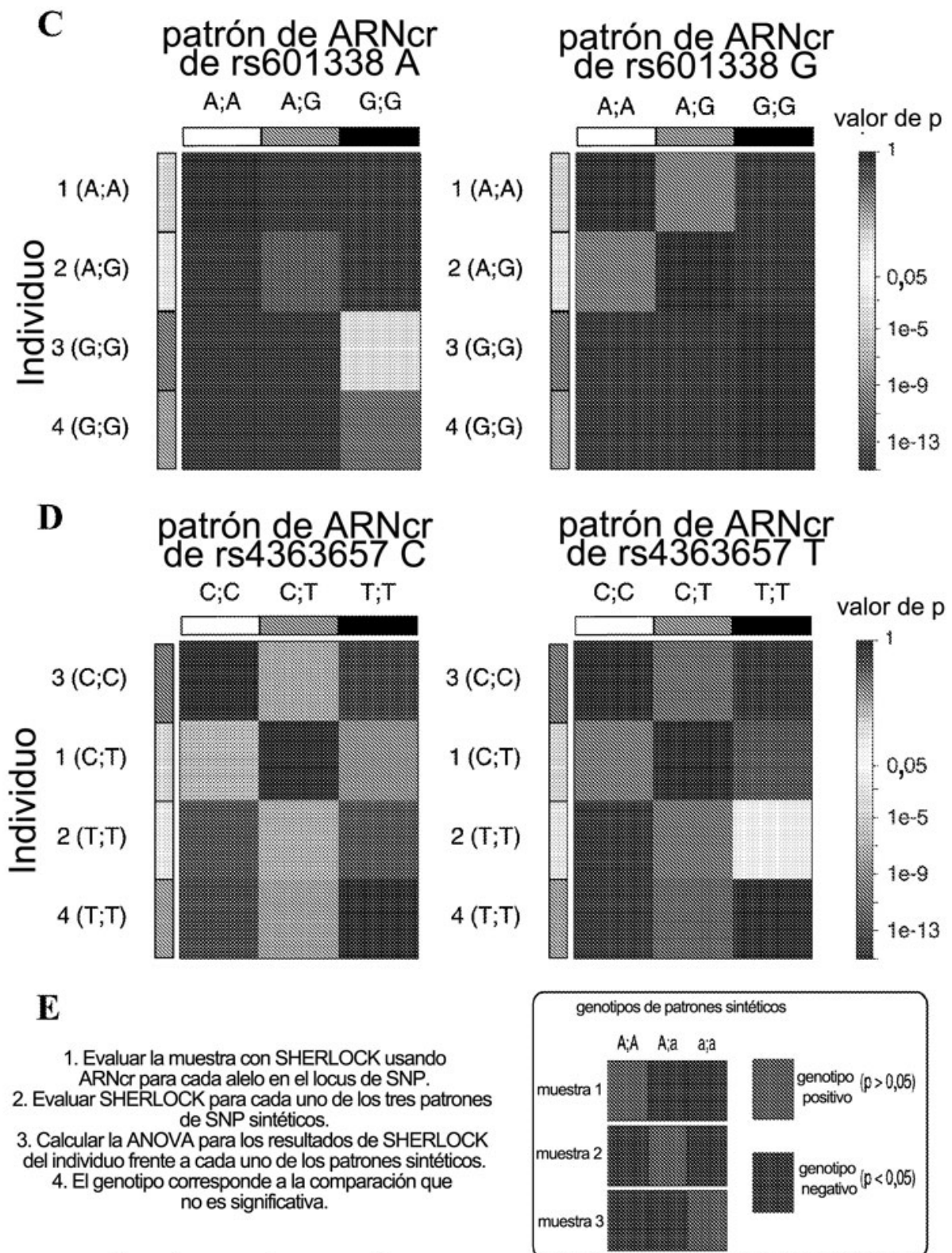
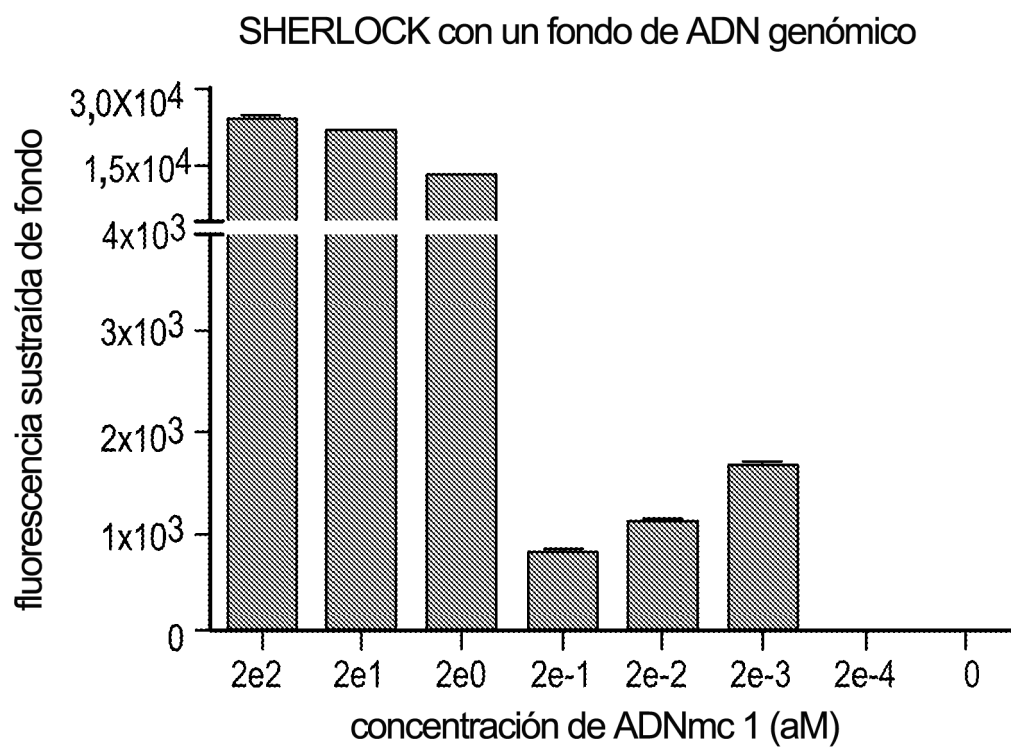
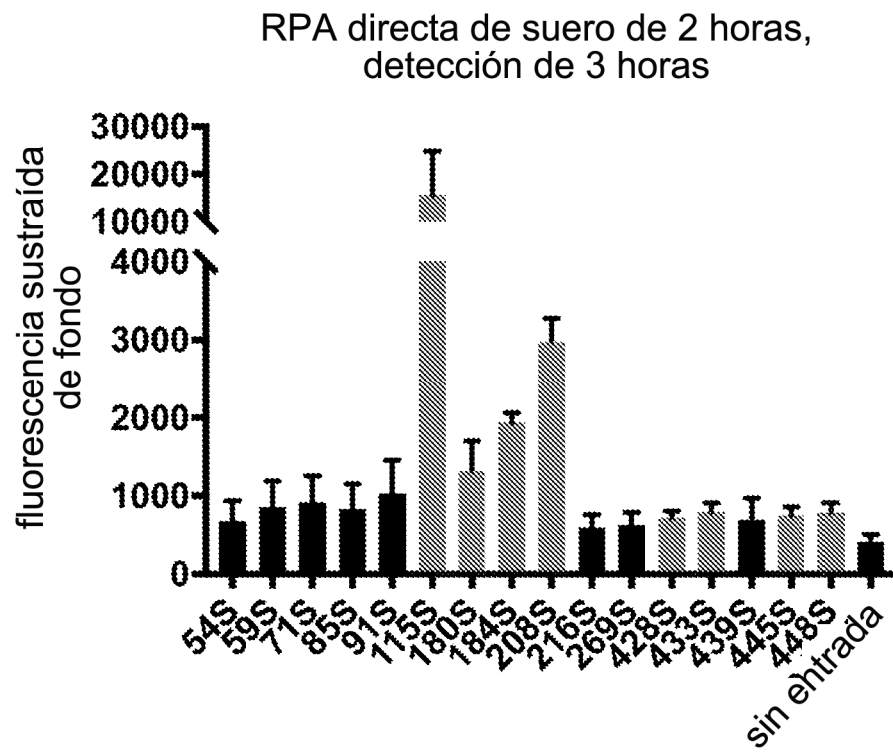
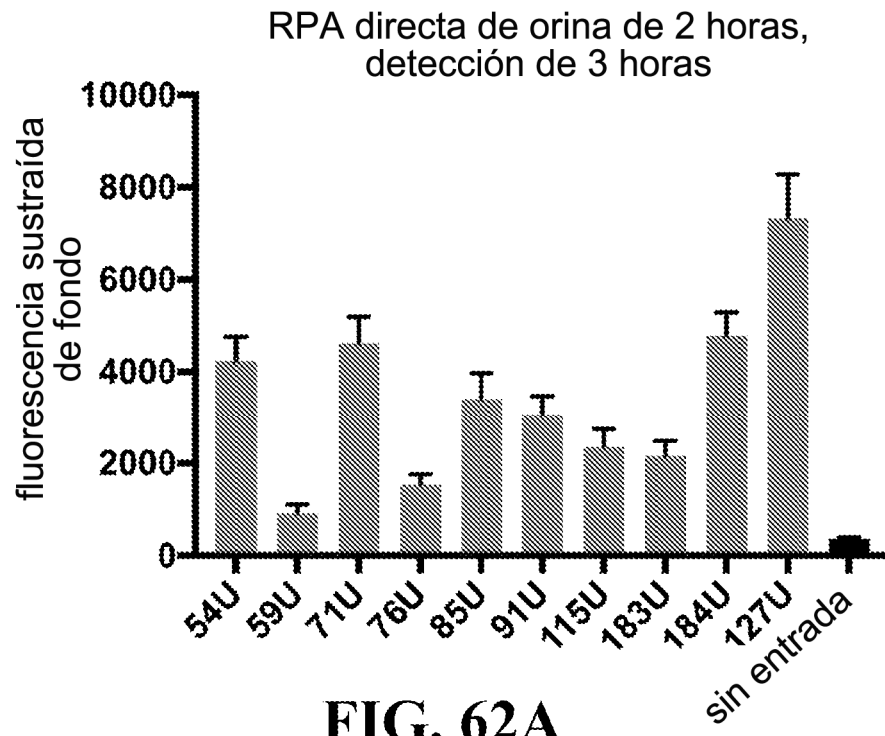


FIG. 60 (continuación)

**FIG. 61**



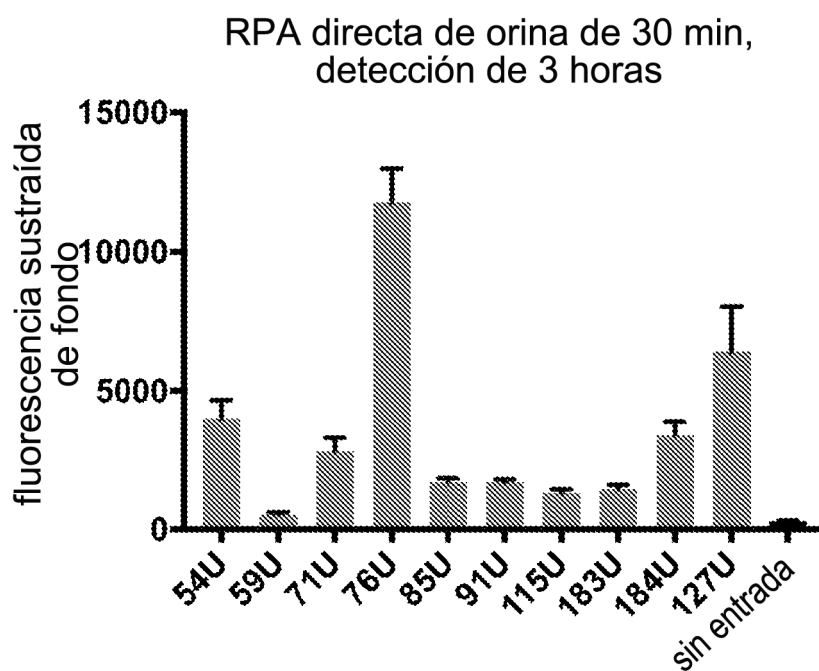


FIG. 63A

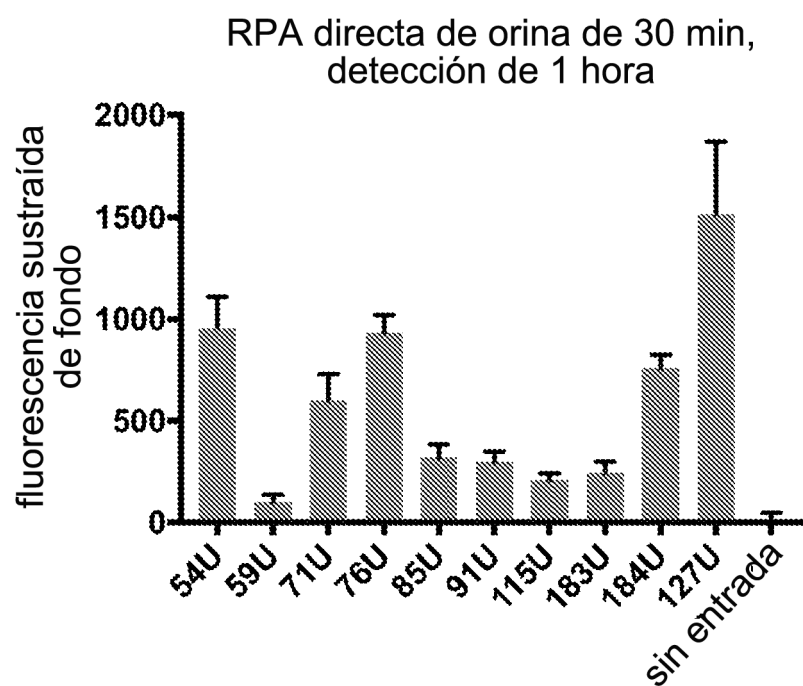


FIG. 63B

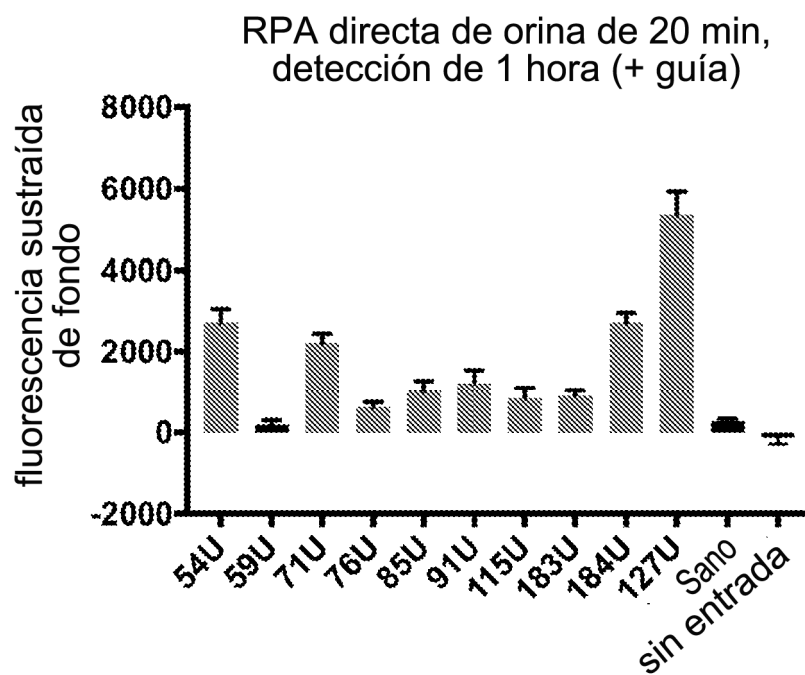


FIG. 64

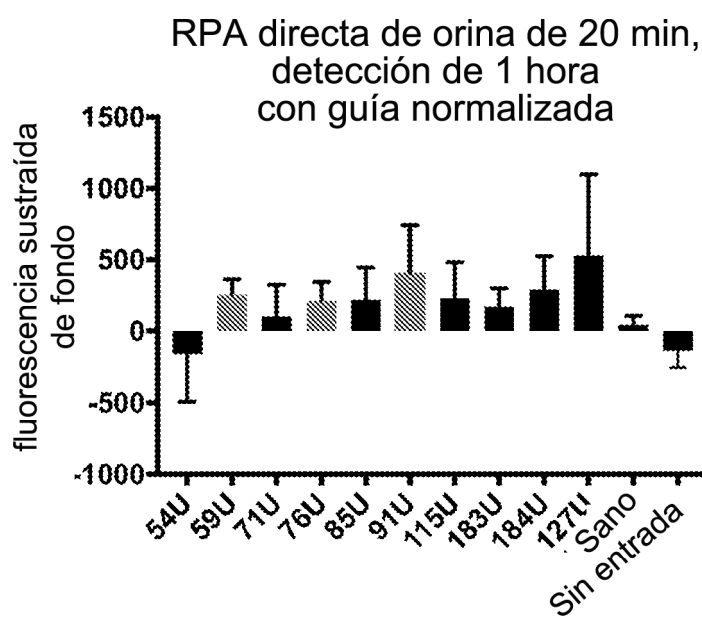


FIG. 65A

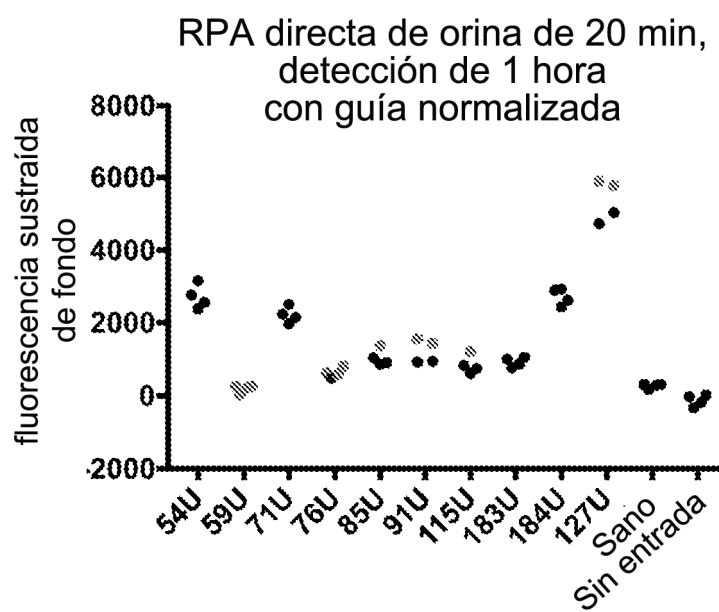
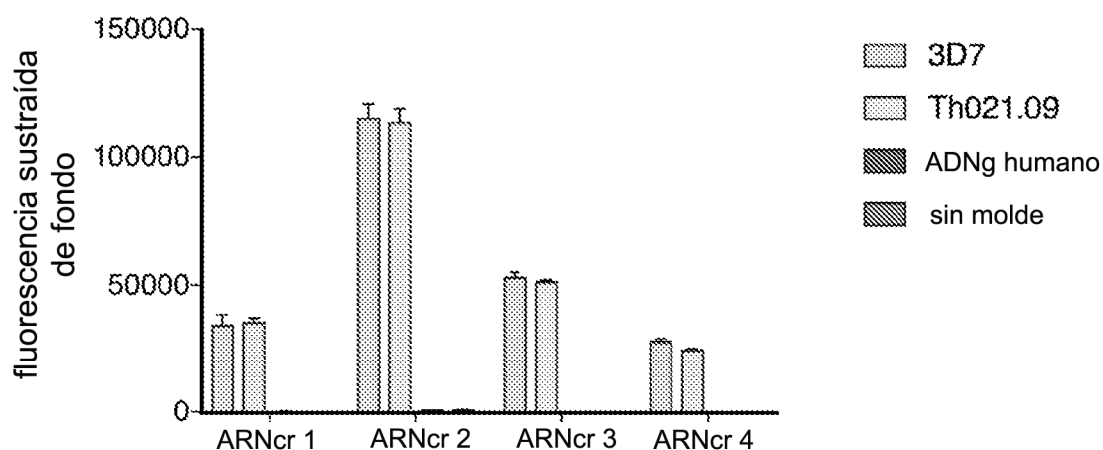


FIG. 65B



gaaattaatacgactcactatagggTTATTGCAATTATTAATCTTG AACGAGGAATG	conjunto de cebadores 1 F de diana 1 de <i>P. falciparum</i>
gaaattaatacgactcactatagggGATTGACAGATTAATAGCTCT TTCTTGATTTG	conjunto de cebadores 2 F de diana 1 de <i>P. falciparum</i>
ATTTTCTGTGTCACAAACATTCATCATATCTT	conjunto de cebadores 1 R de diana 1 de <i>P. falciparum</i>
TTCAATTTCAAATAAGATATAGTGACTCGC	conjunto de cebadores 2 R de diana 1 de <i>P. falciparum</i>
gaaattaatacgactcactatagggTTCTTATTAGCAGAACAAAGA AGTTTAACAAC	conjunto de cebadores 1 F de diana 2 de <i>P. falciparum</i>
gaaattaatacgactcactatagggATTTTATGCAATGTTAAAAAC TGTTCCAAGTA	conjunto de cebadores 2 F de diana 2 de <i>P. falciparum</i>
TAATTGACATCCAATCCATAATAAGCATAGA	conjunto de cebadores 1 R de diana 2 de <i>P. falciparum</i>
GAATTATAGTTGTAAACTTCTTTGTTCTGCT	conjunto de cebadores 2 R de diana 2 de <i>P. falciparum</i>
cctagtaagcatgattcatcagattgtggttttagtccccttcggt tttgggtagtctaaatcccctatagtgagtcgtattaatttc	ARNcr de conjunto de cebadores 1 de diana 1 de <i>P. falciparum</i>
ggatgggtgatgcatggccgttttttagttgttttagtccccttcggt tttgggtagtctaaatcccctatagtgagtcgtattaatttc	ARNcr de conjunto de cebadores 2 de diana 1 de <i>P. falciparum</i>
caatttaaaatgatttttgggtciagaggttttagtccccttcggt tttgggtagtctaaatcccctatagtgagtcgtattaatttc	ARNcr de conjunto de cebadores 1 de diana 2 de <i>P. falciparum</i>
gctggttttagtaattgtattattatcatgttttagtccccttcggt tttgggtagtctaaatcccctatagtgagtcgtattaatttc	ARNcr de conjunto de cebadores 2 de diana 2 de <i>P. falciparum</i>

FIG. 66

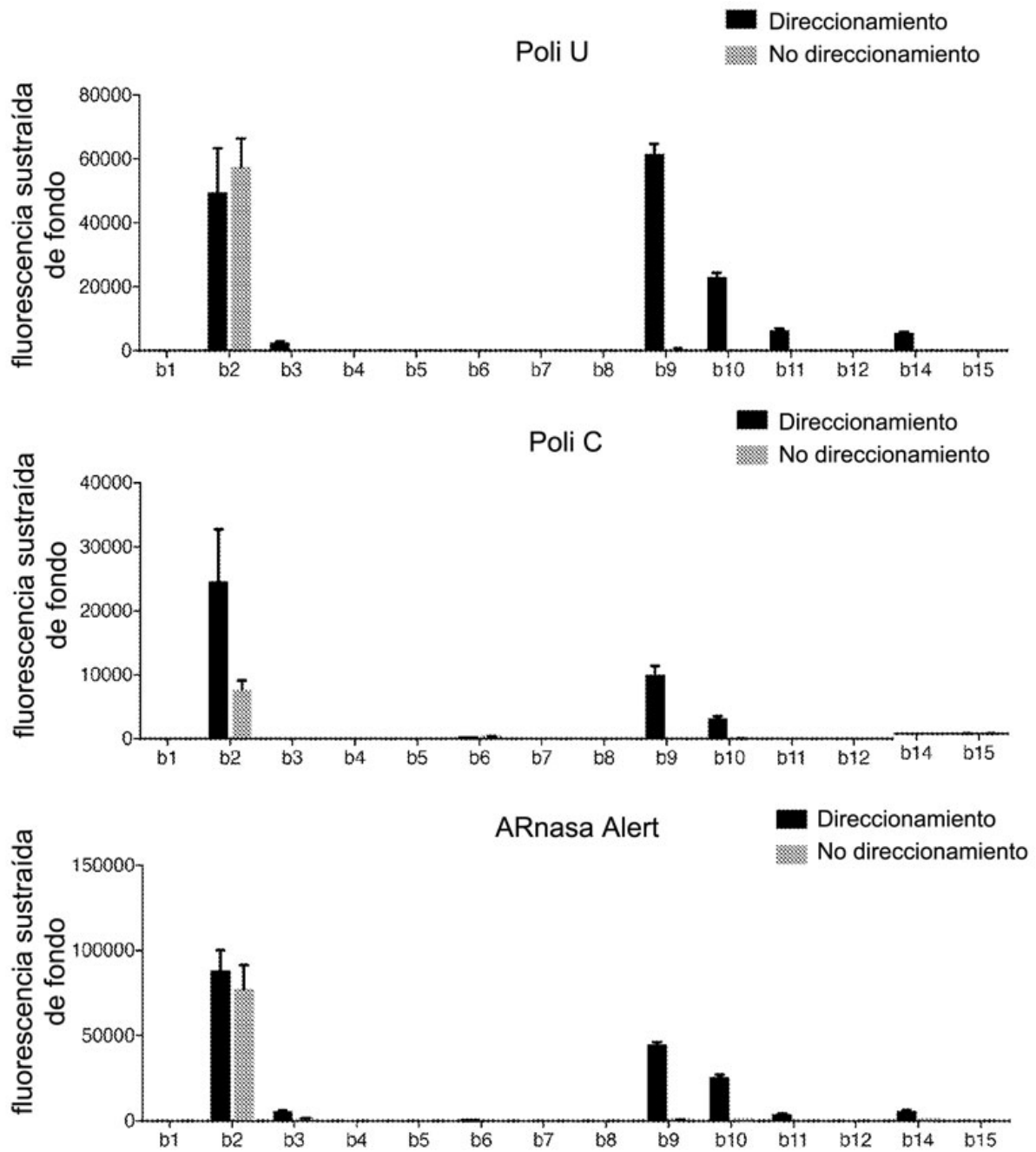


FIG. 67

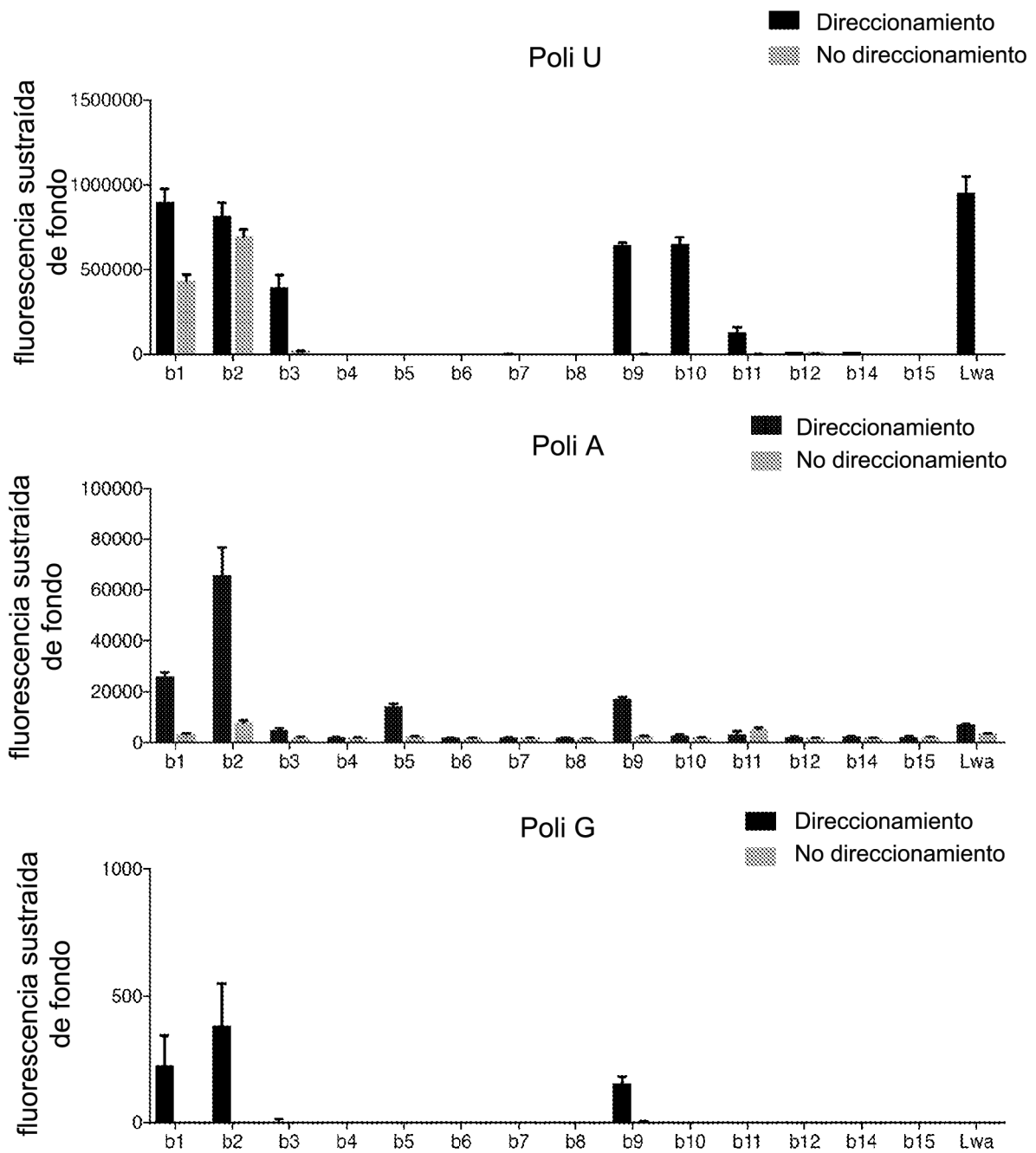


FIG. 67 (continuación)

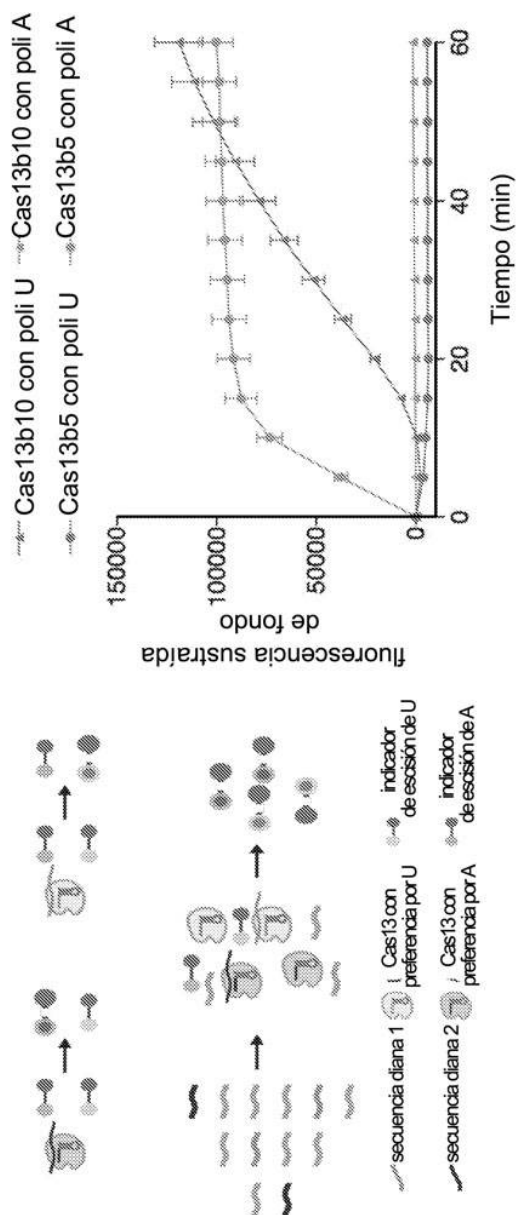


FIG. 68

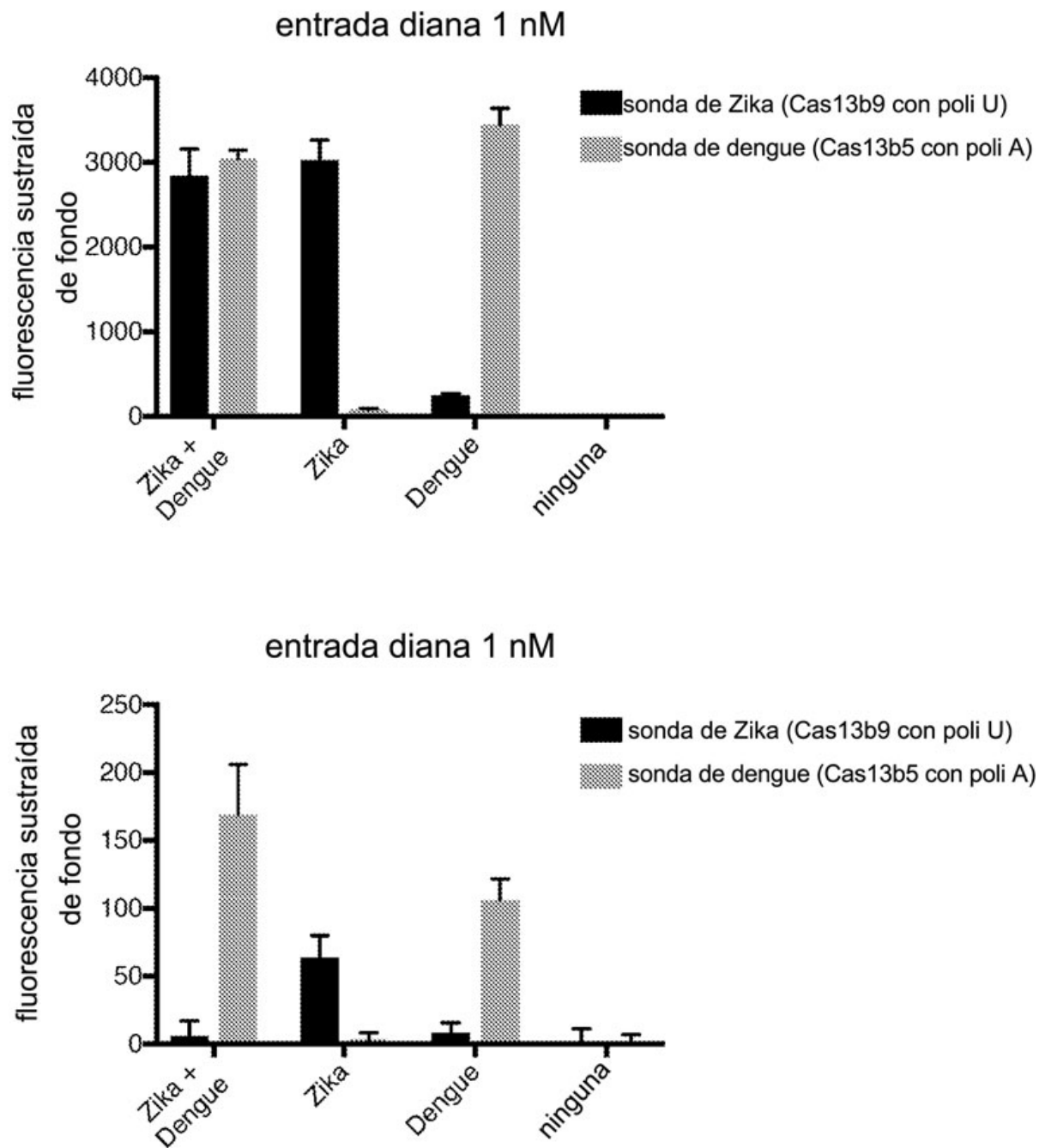


FIG. 69

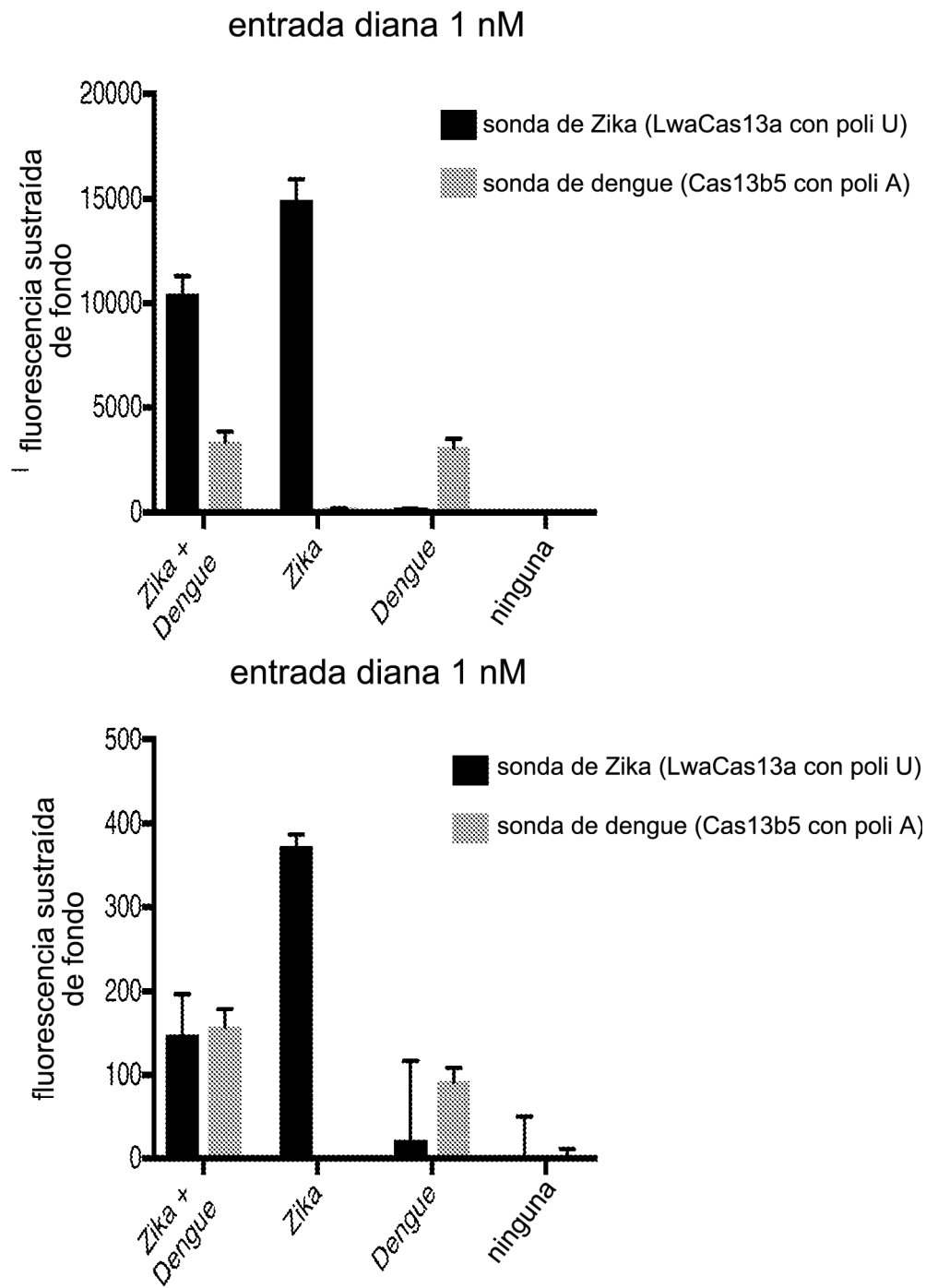


FIG. 70

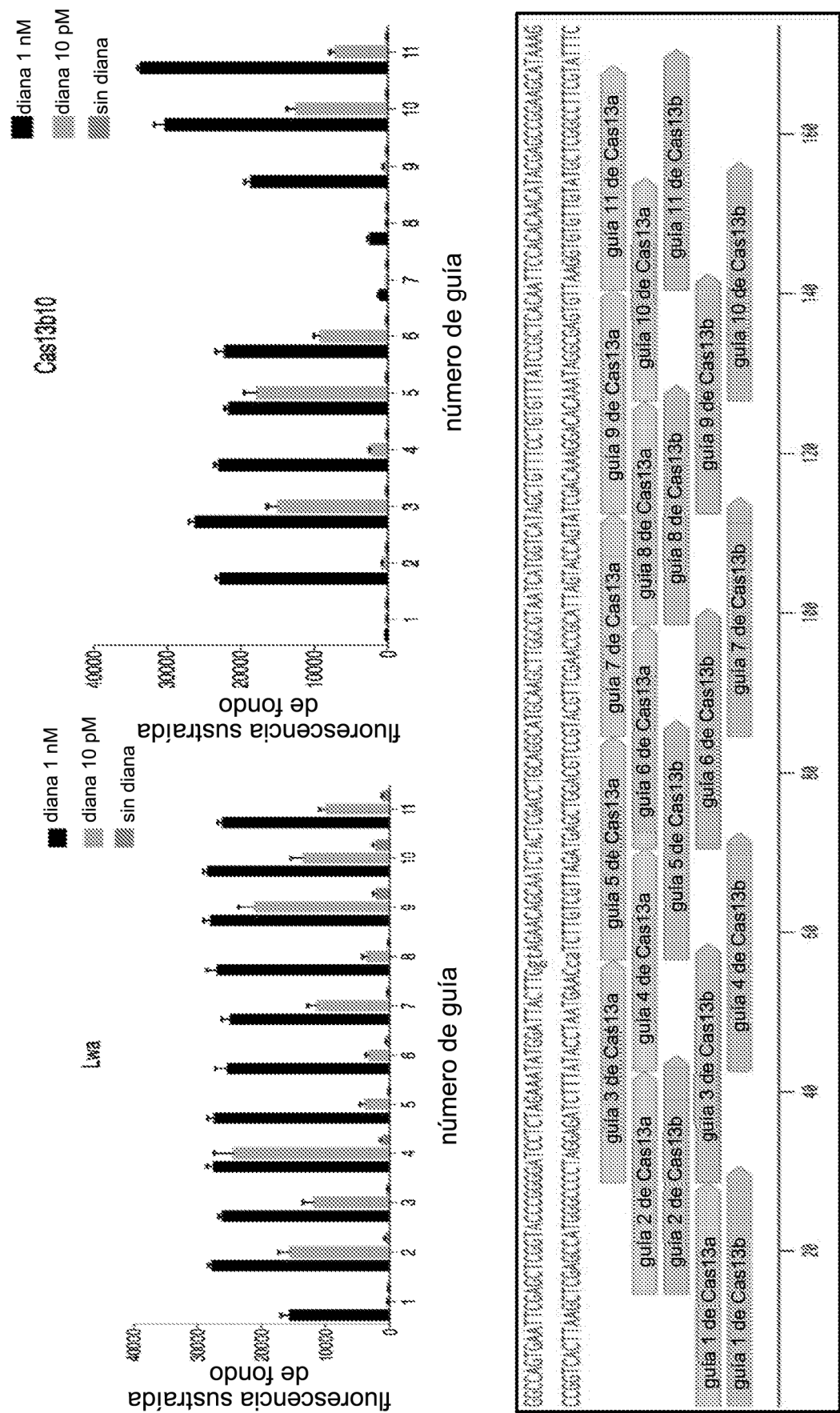


FIG. 71

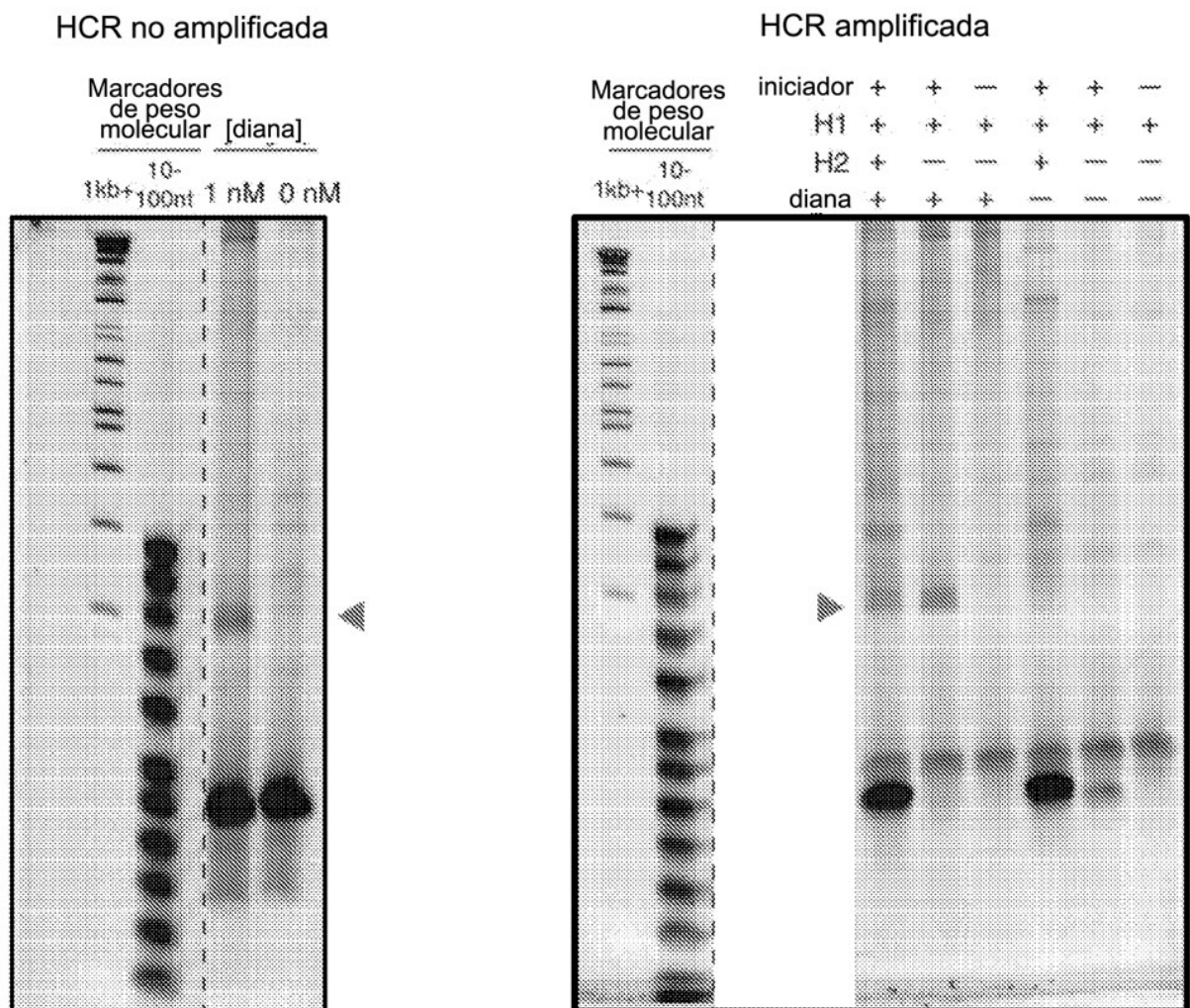
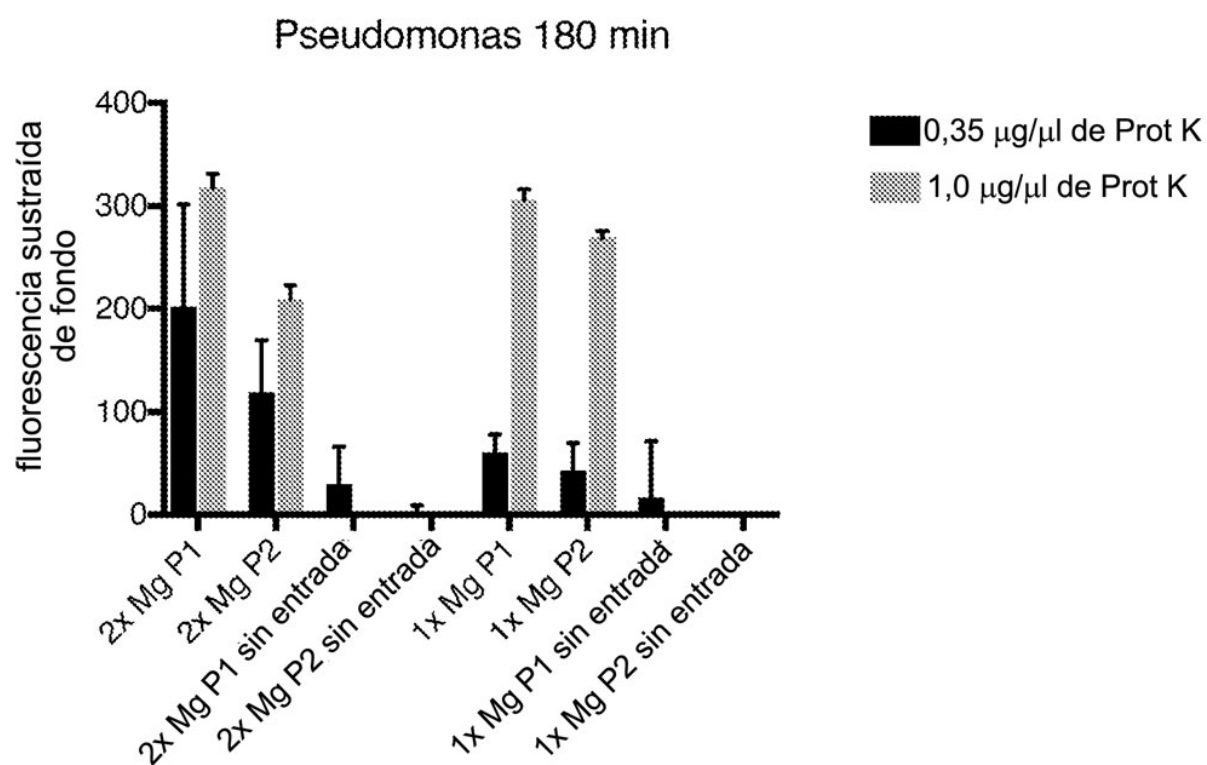


FIG. 72

**FIG. 73**

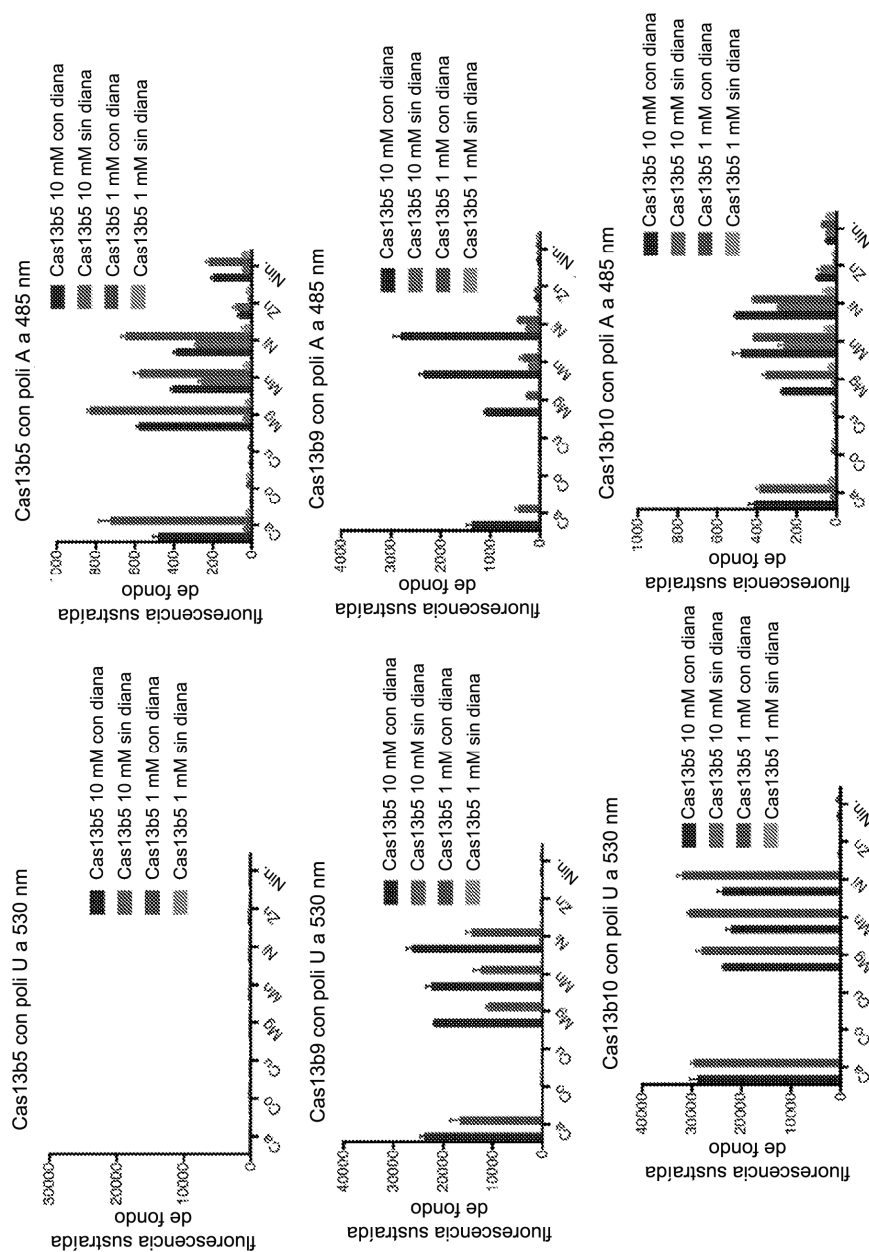


FIG. 74

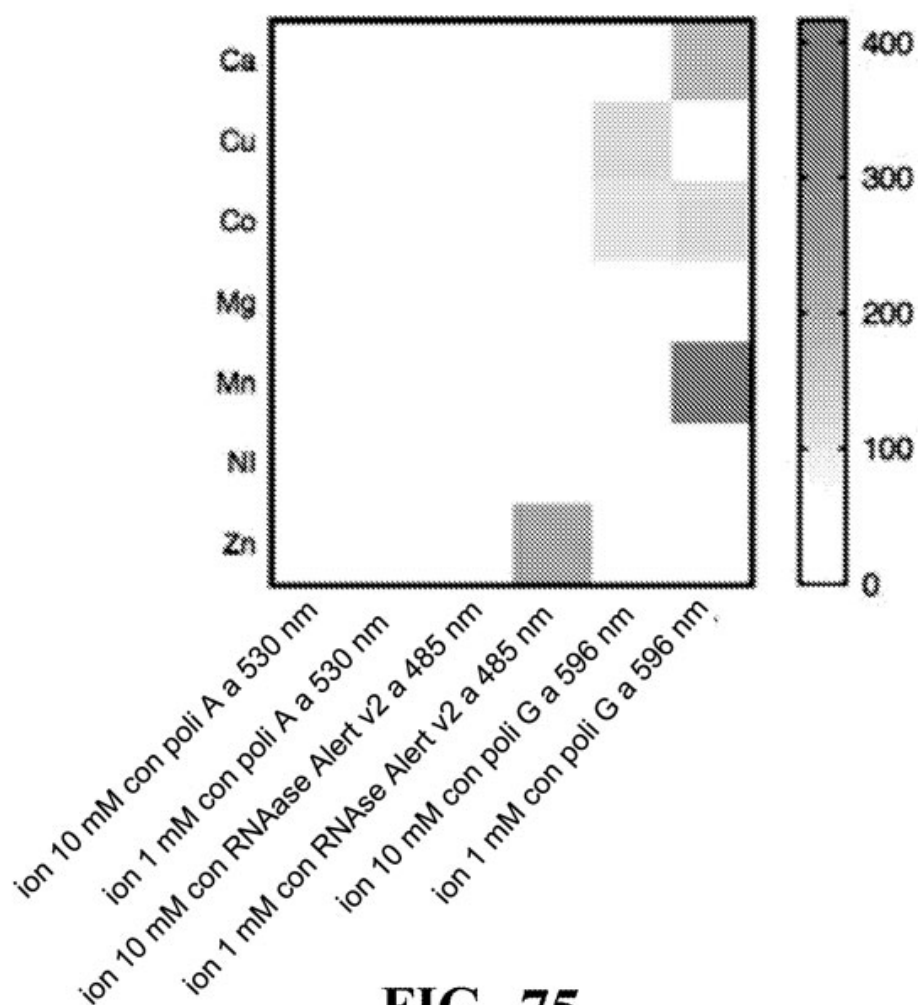
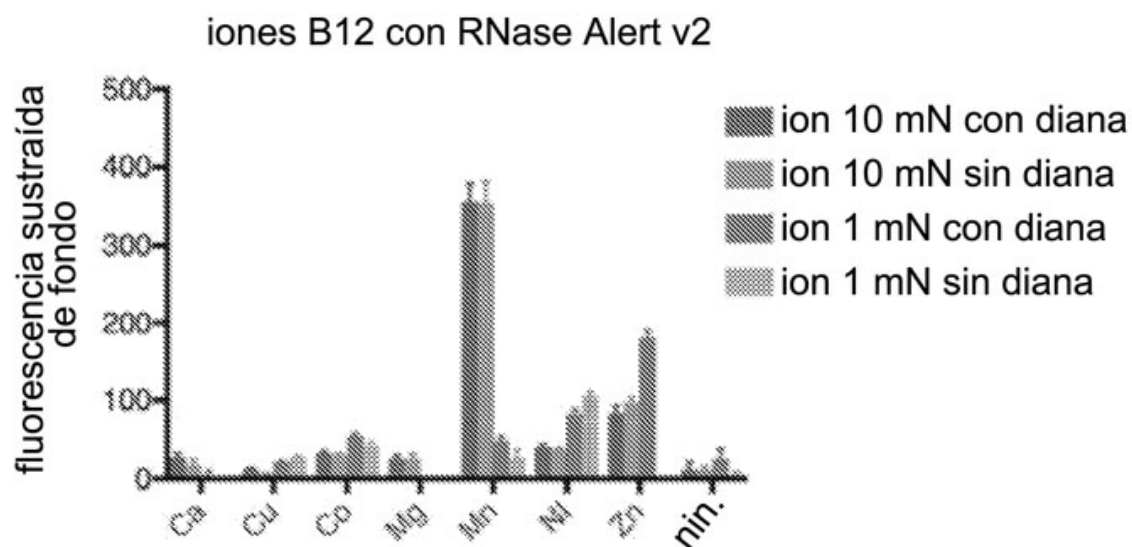


FIG. 75

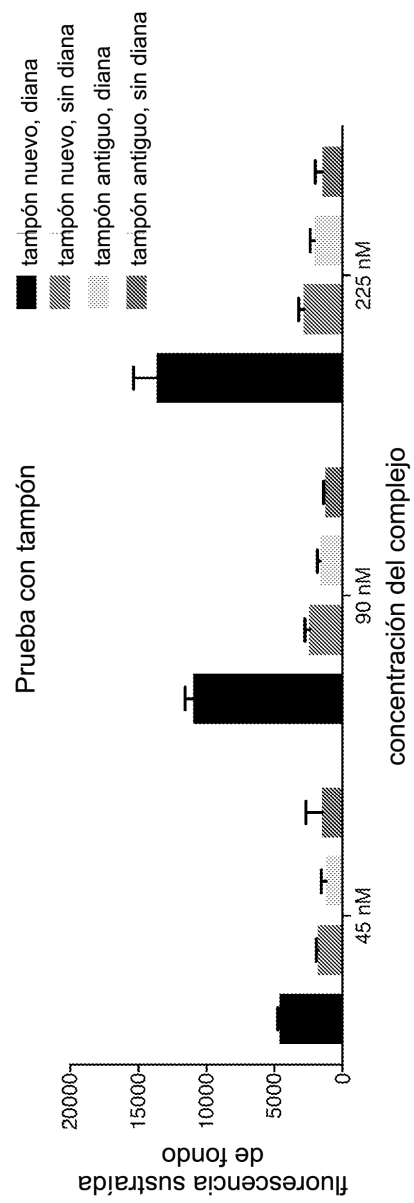


FIG. 76

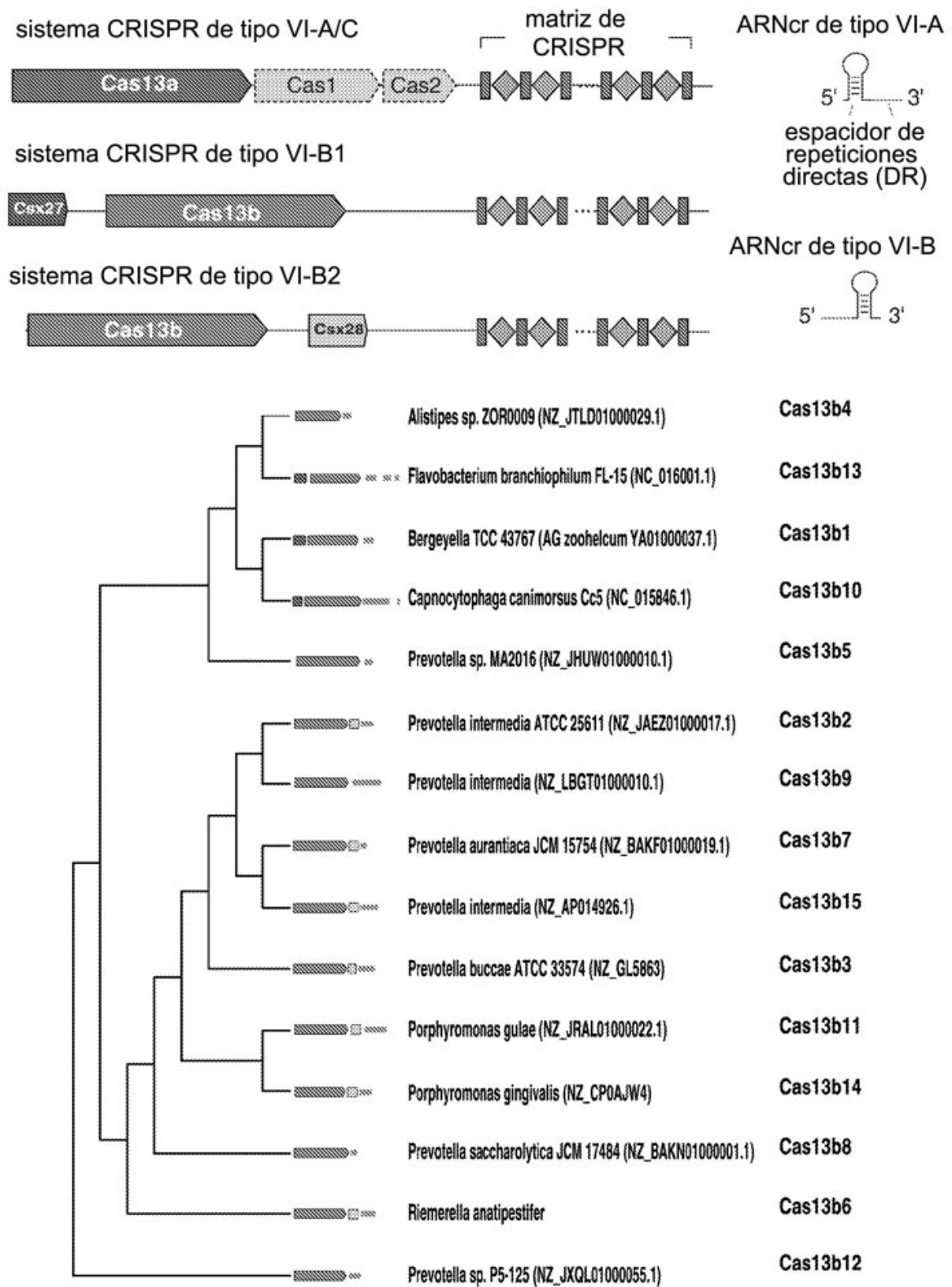


FIG. 77

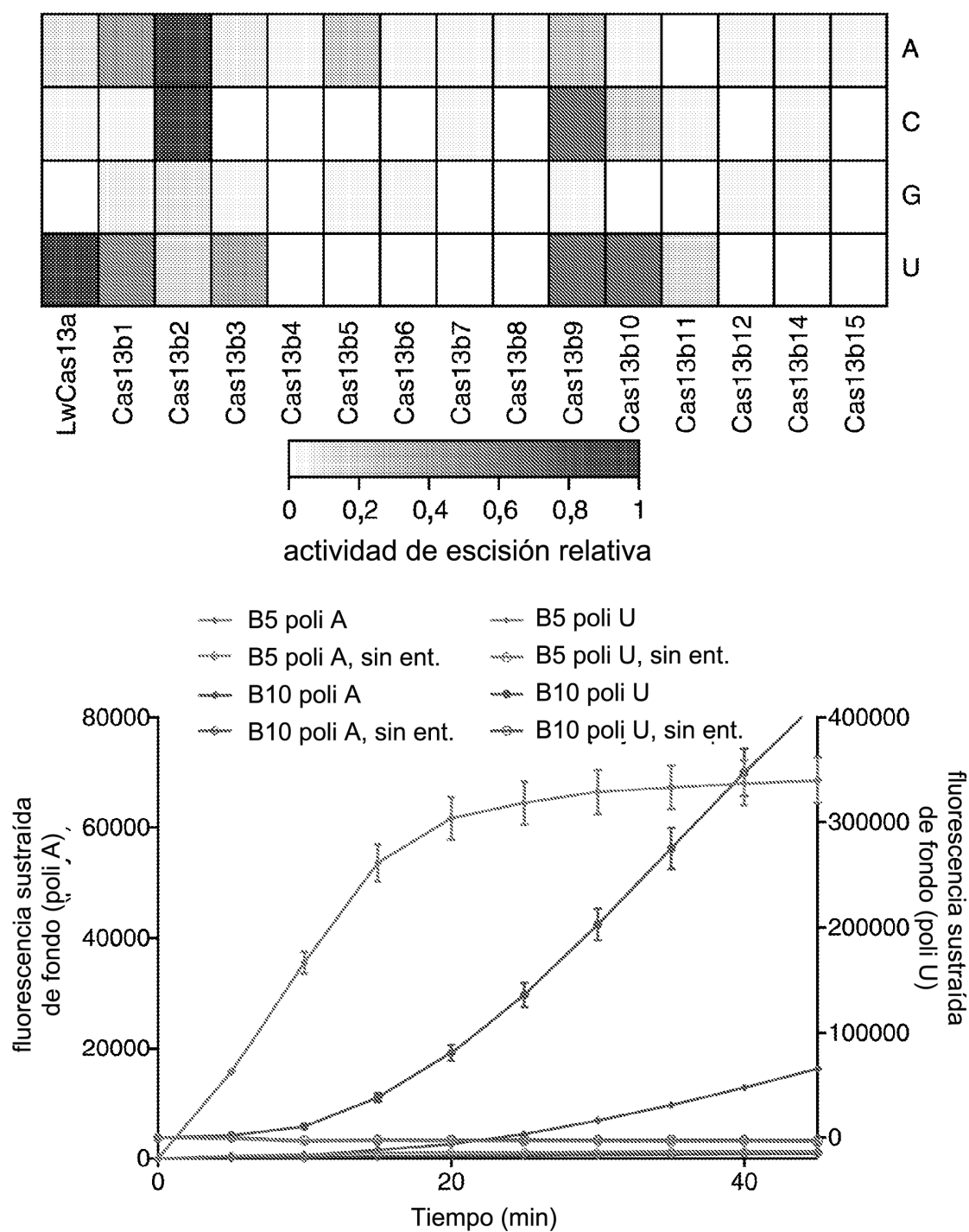


FIG. 78

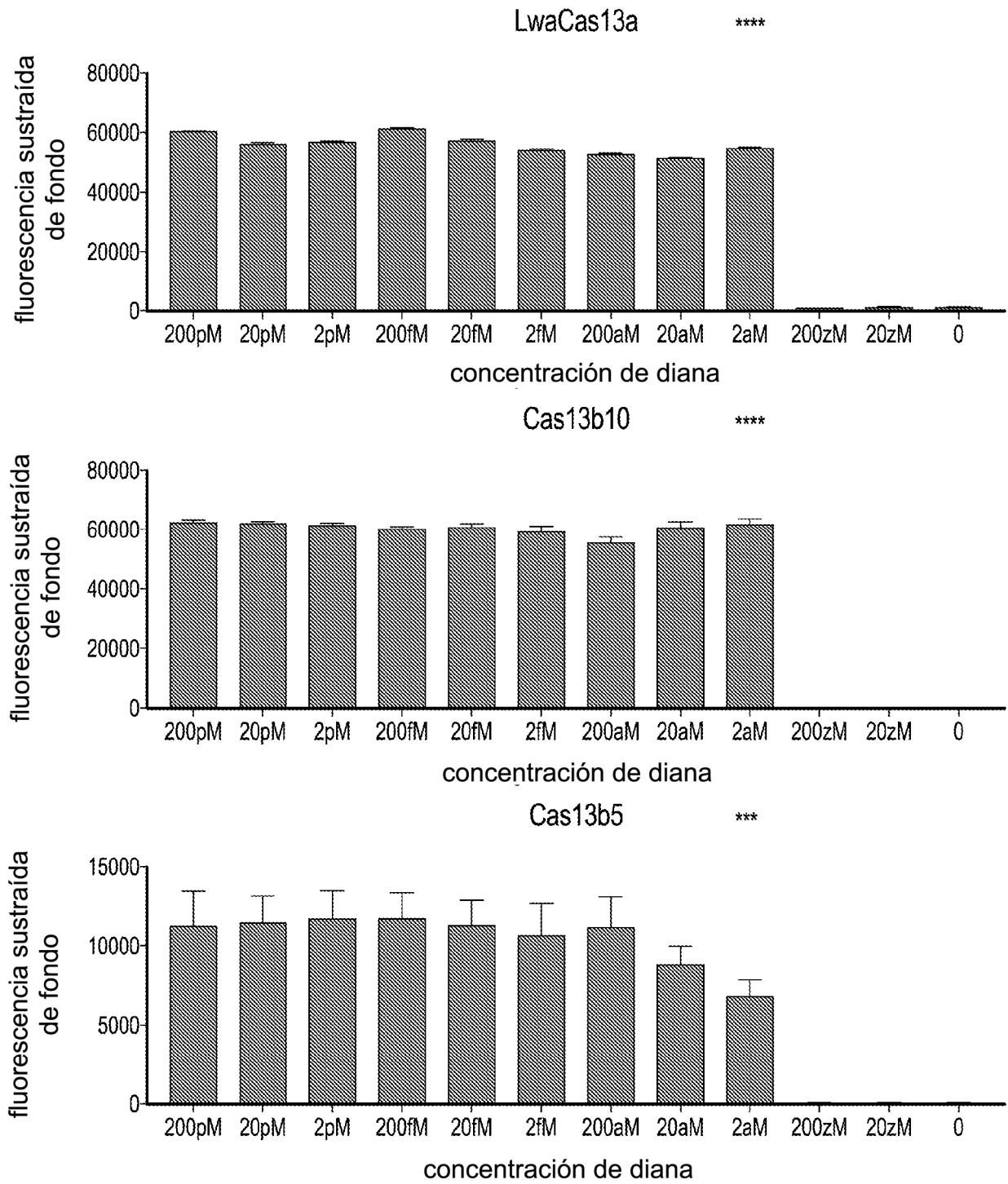


FIG. 79

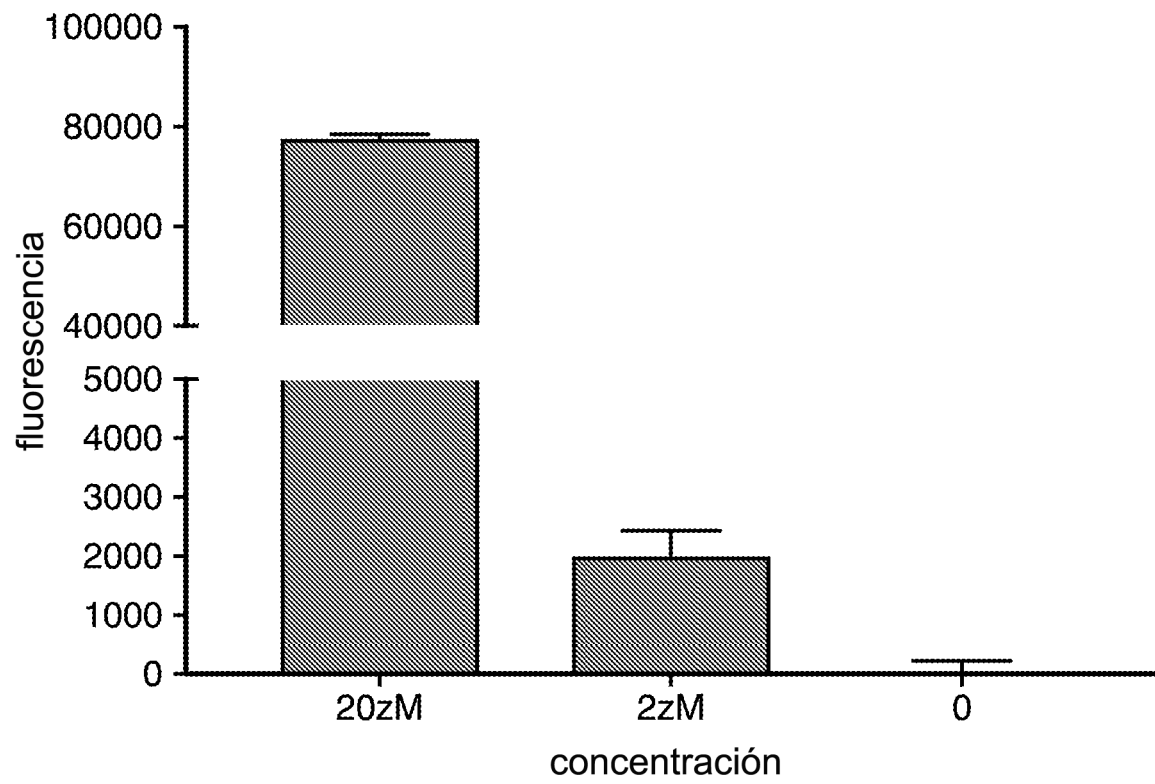


FIG. 80

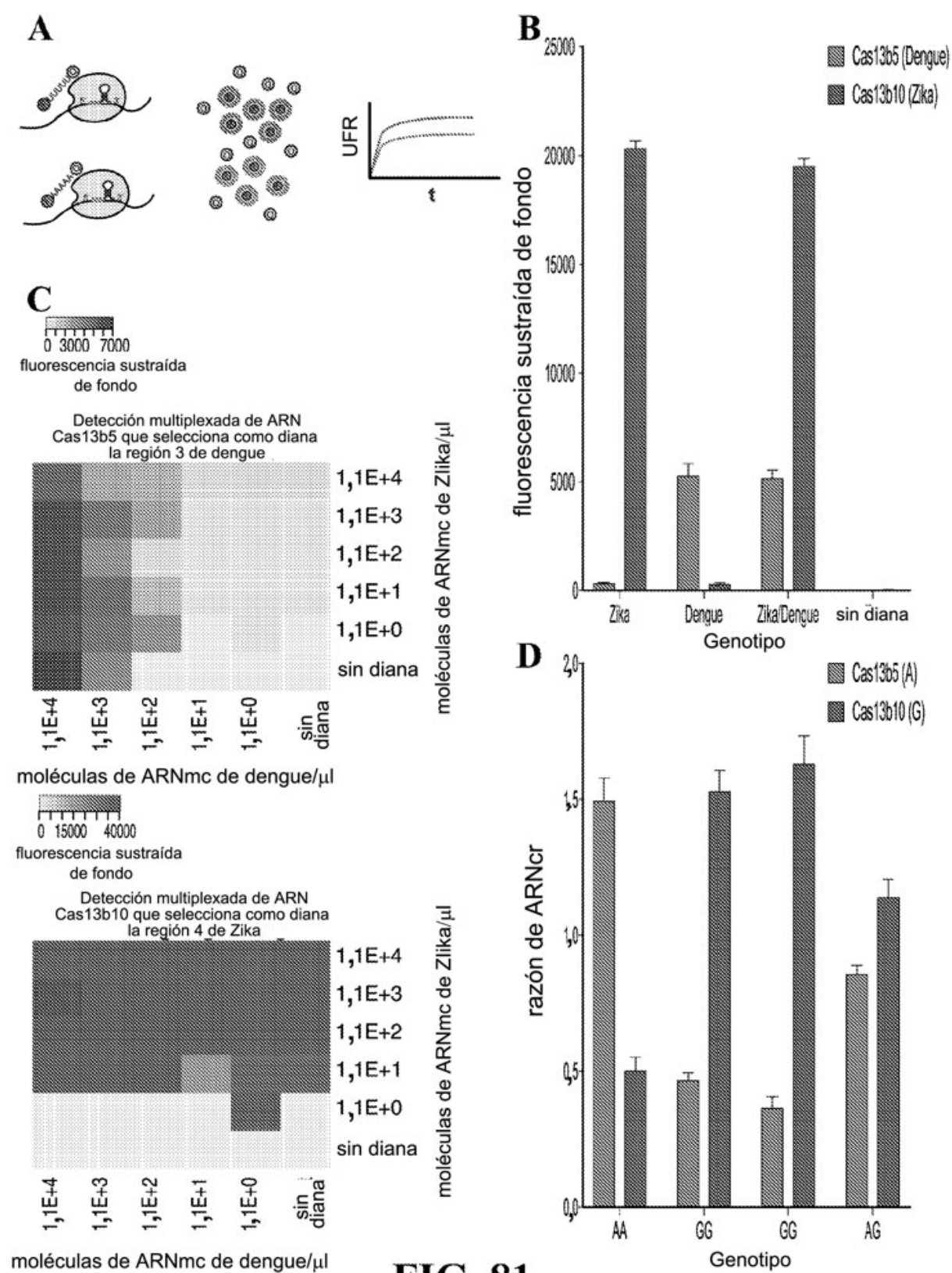


FIG. 81

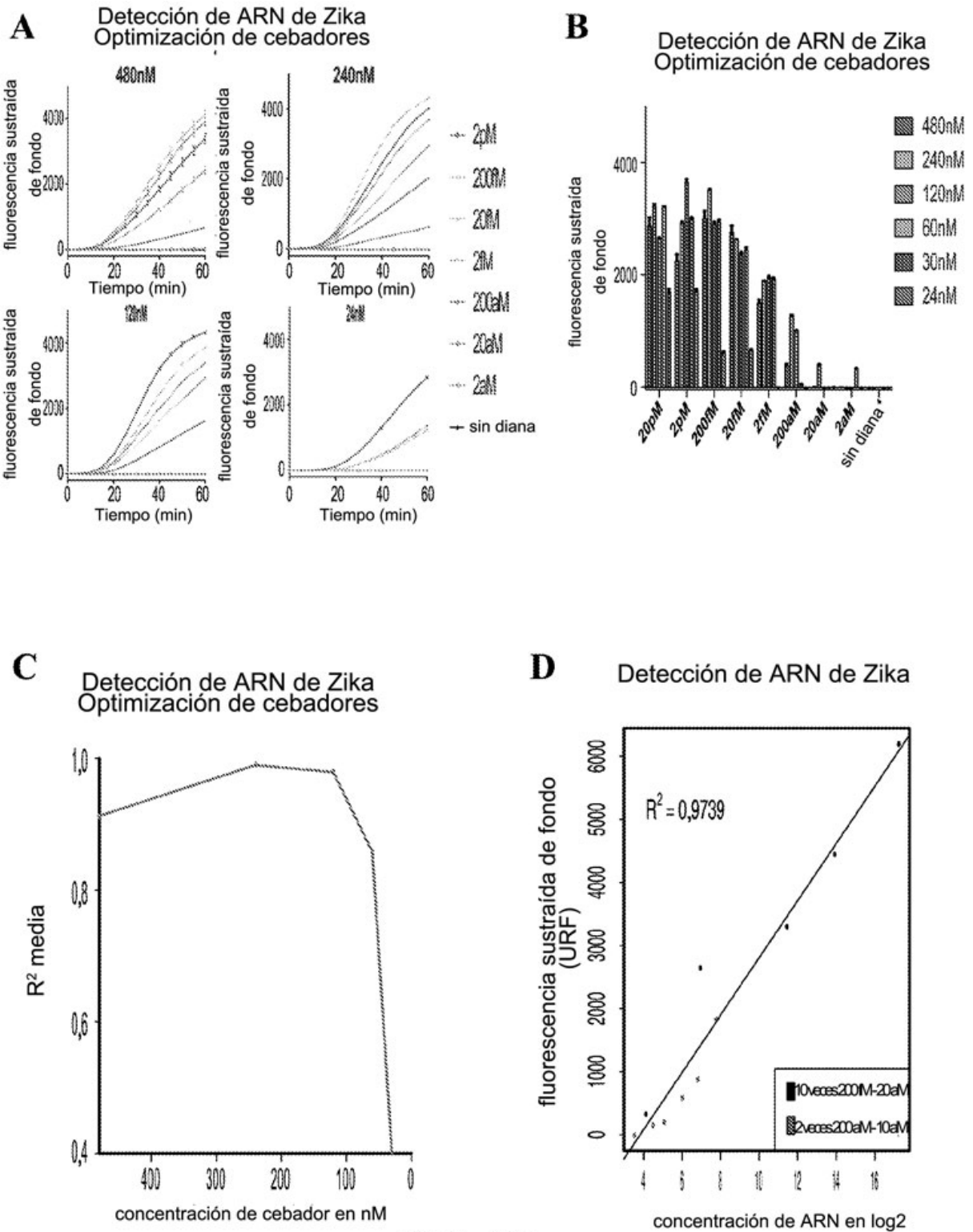
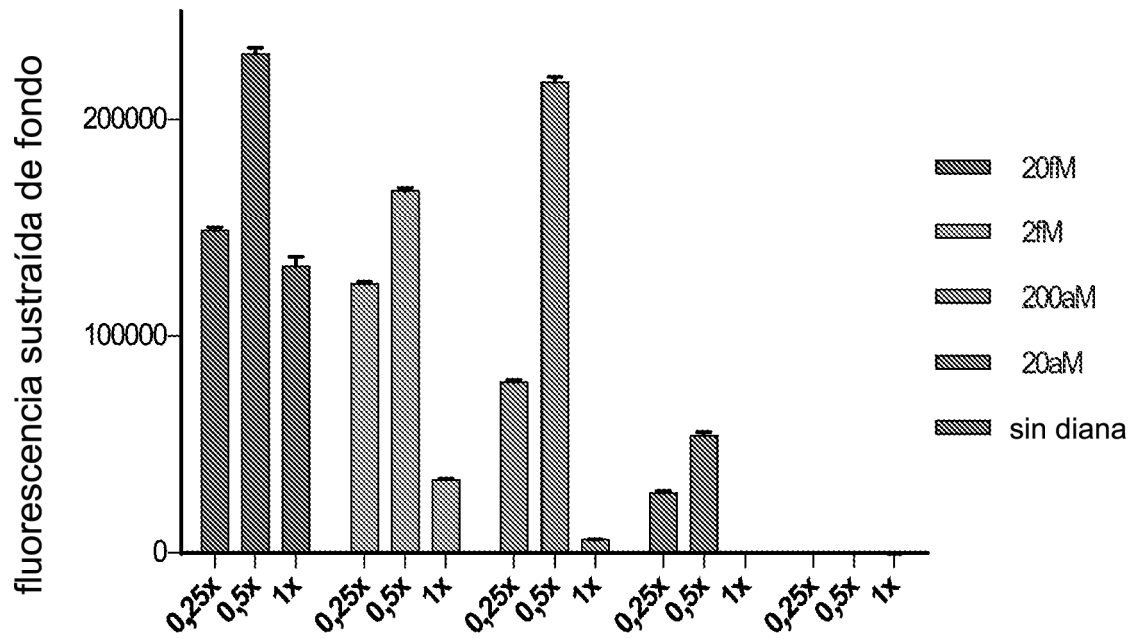


FIG. 82

E

Pseudomonas



F

Pseudomonas

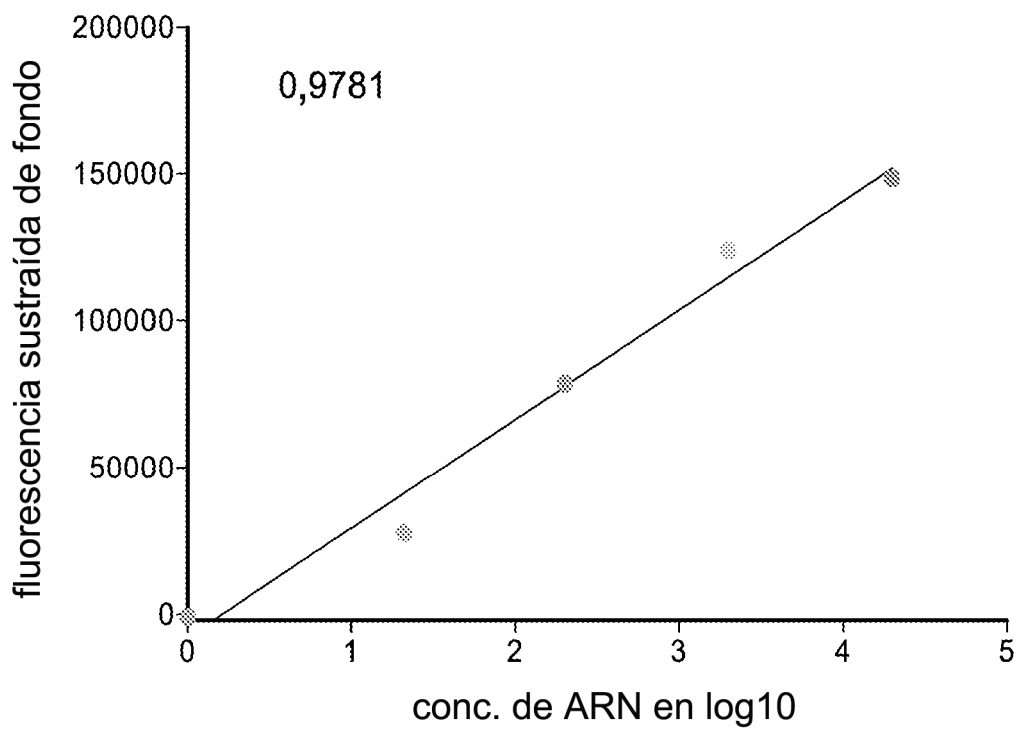


FIG. 82 (continuación)

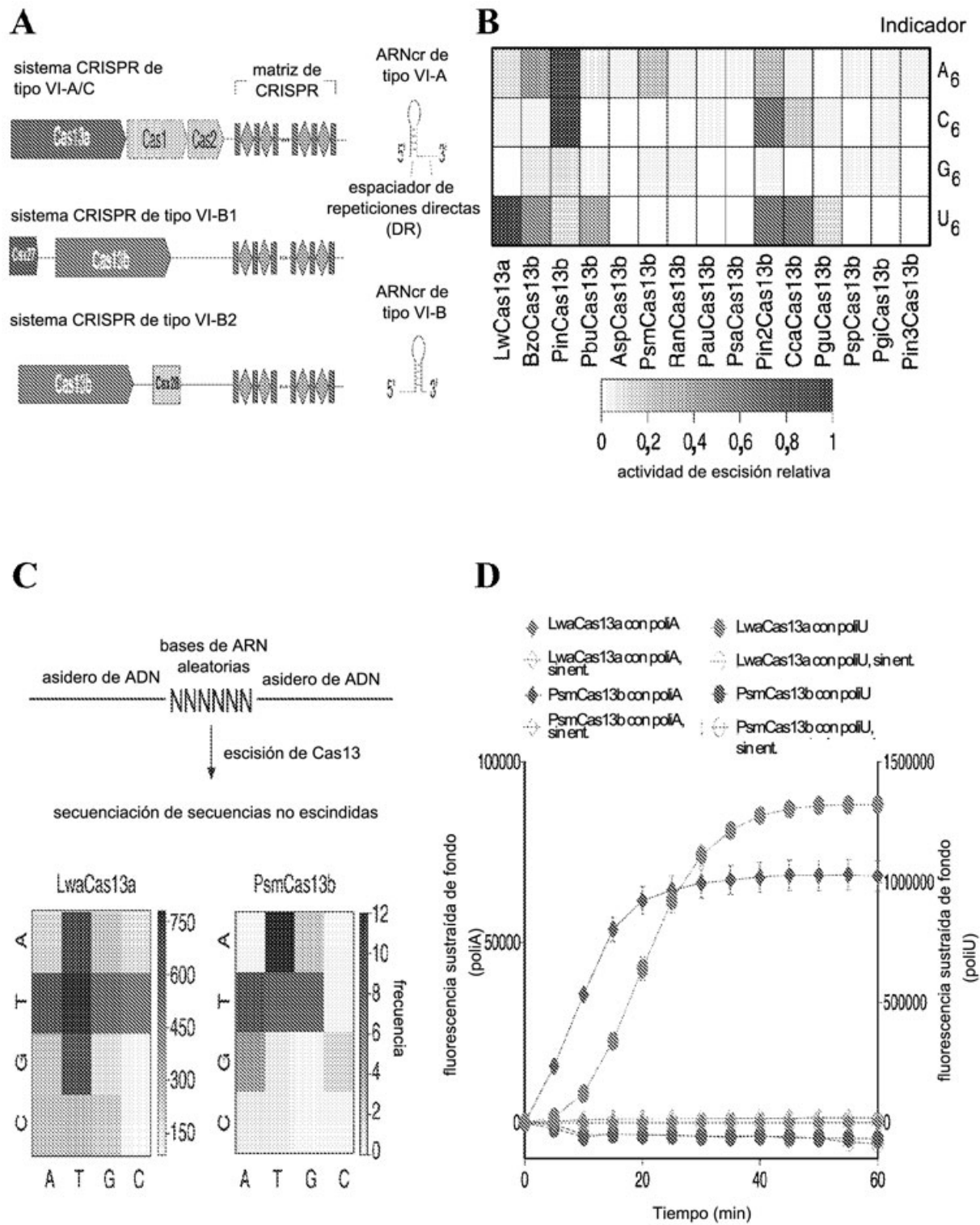


FIG. 83

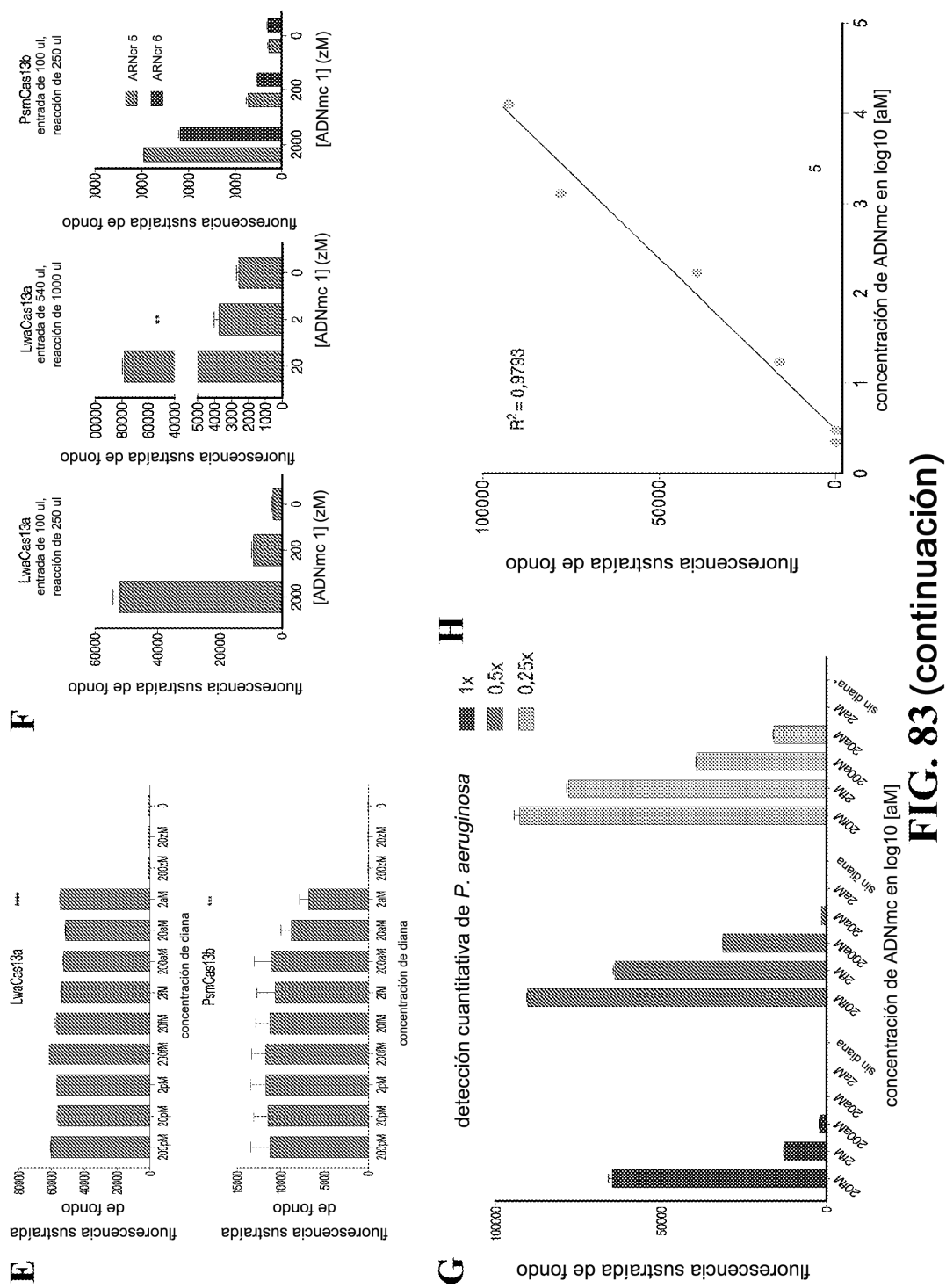


FIG. 83 (continuación)

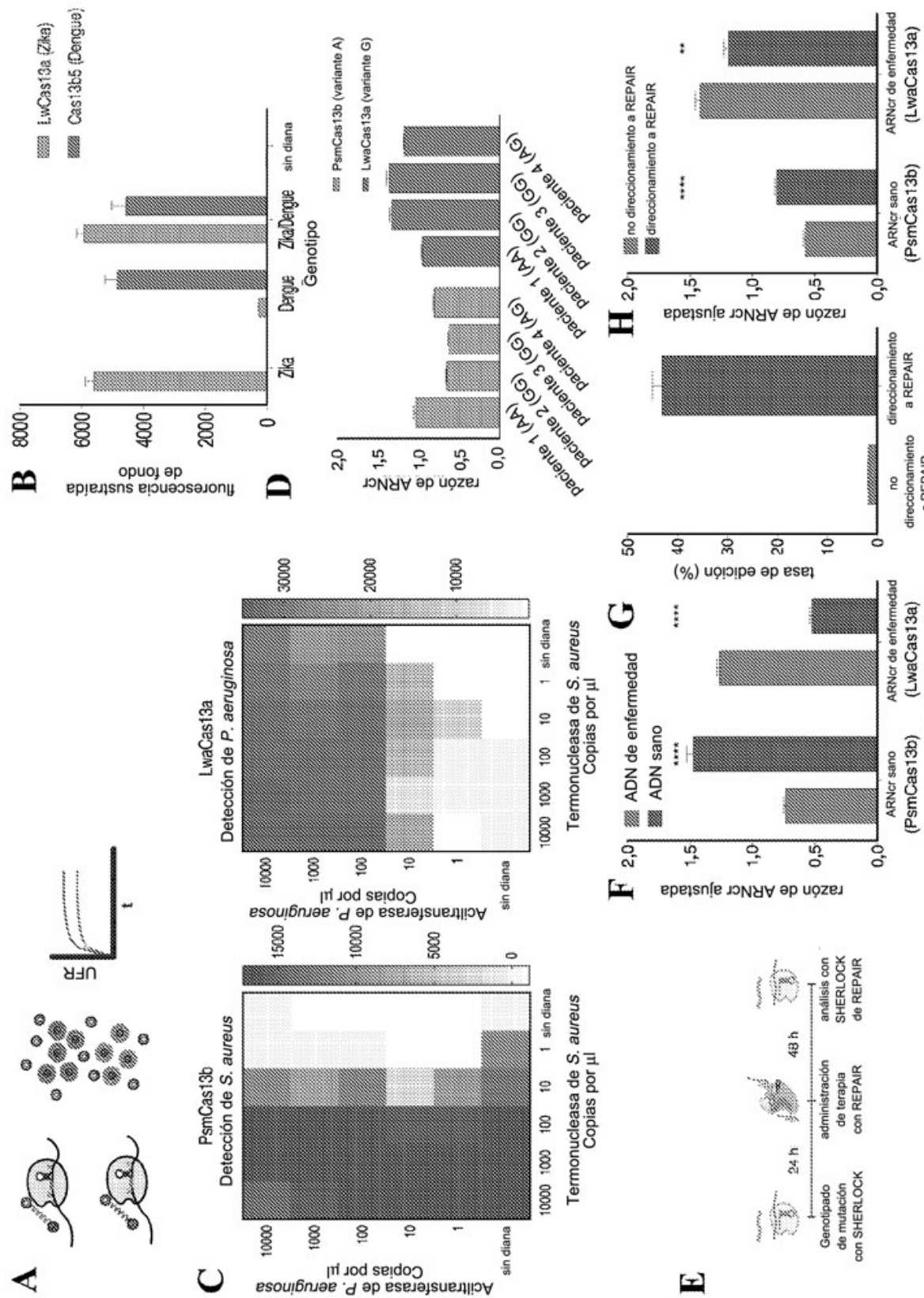
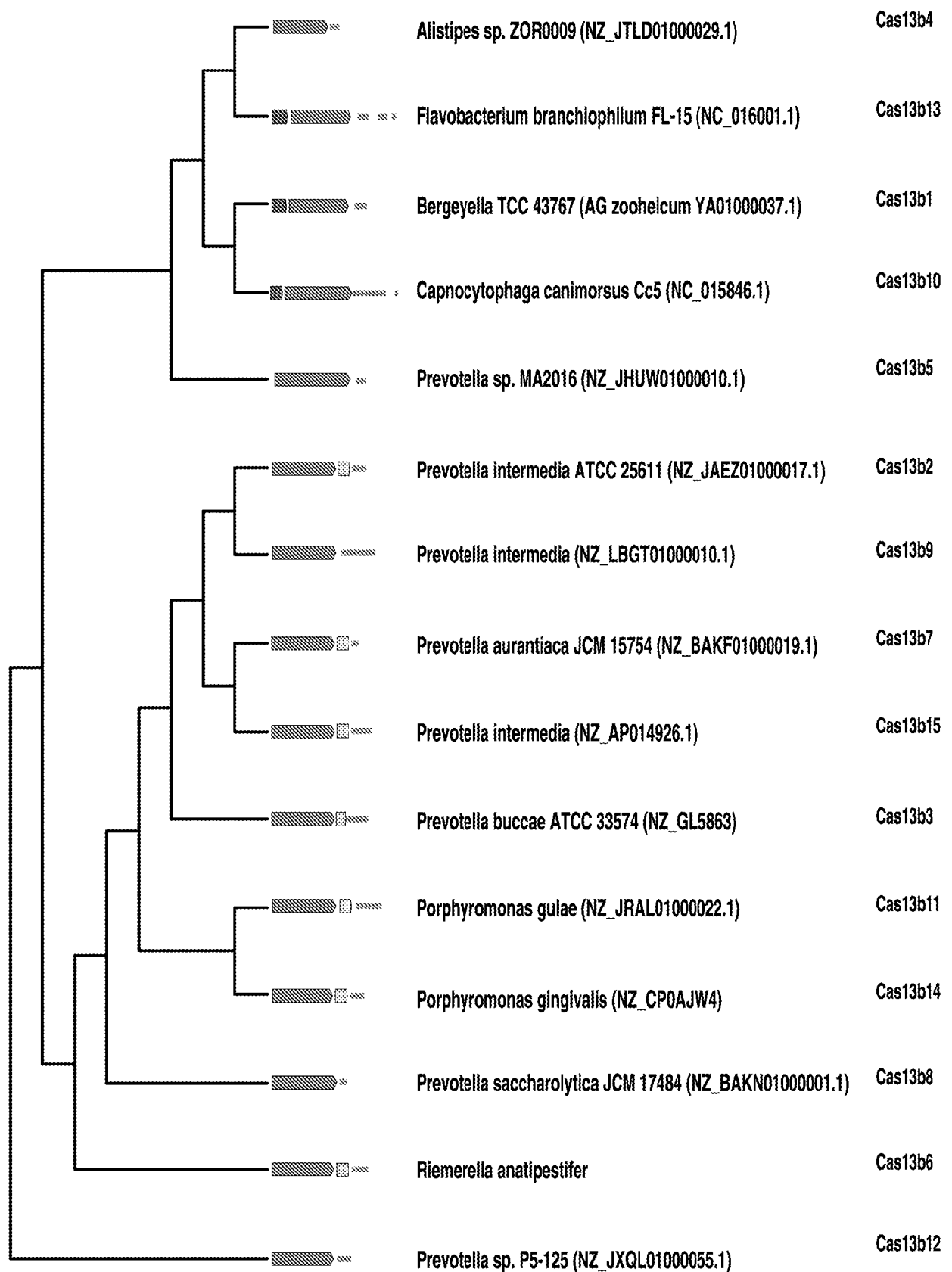


FIG. 84

**FIG. 85**

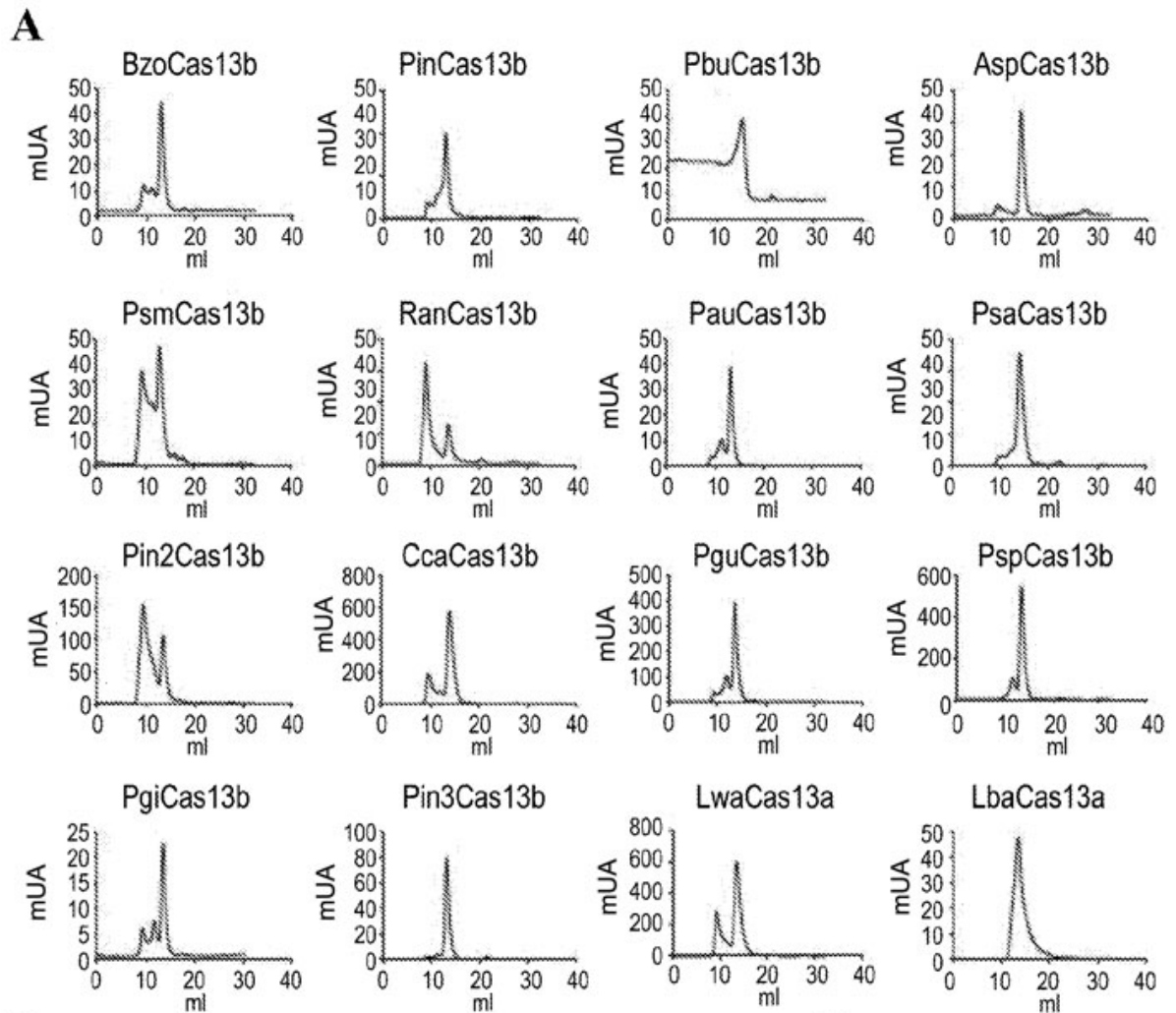


FIG. 86

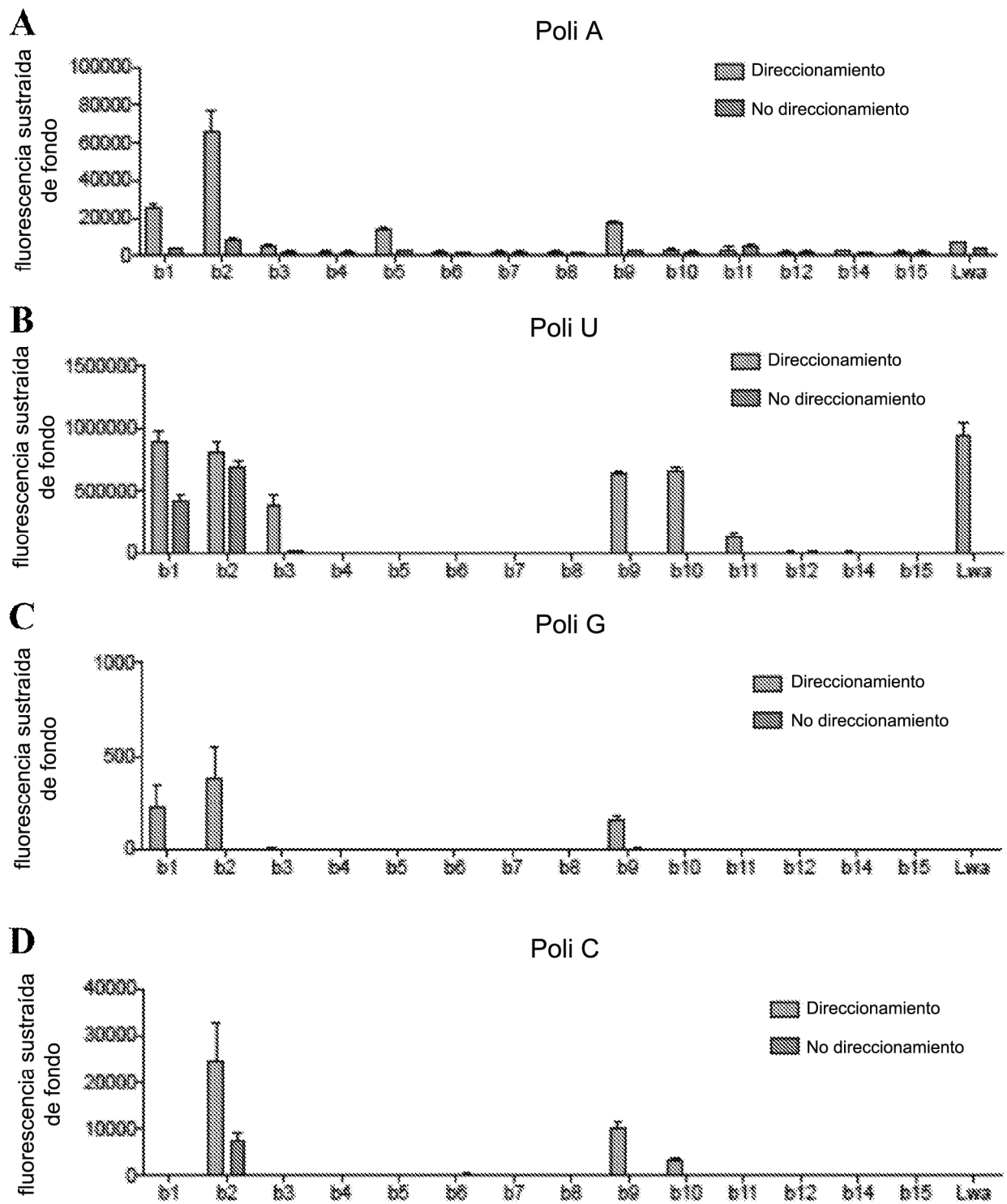


FIG. 87

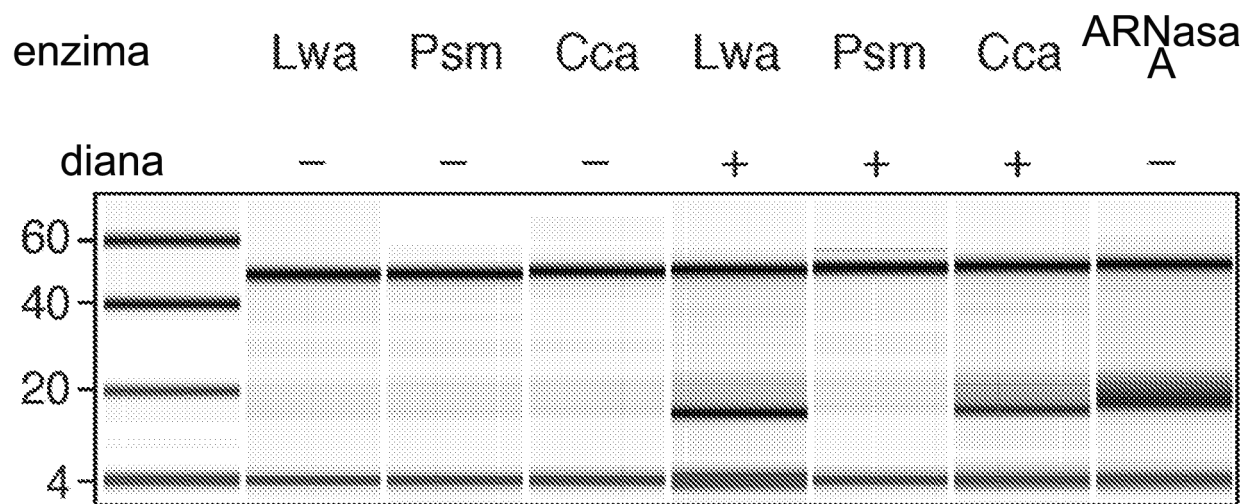


FIG. 88

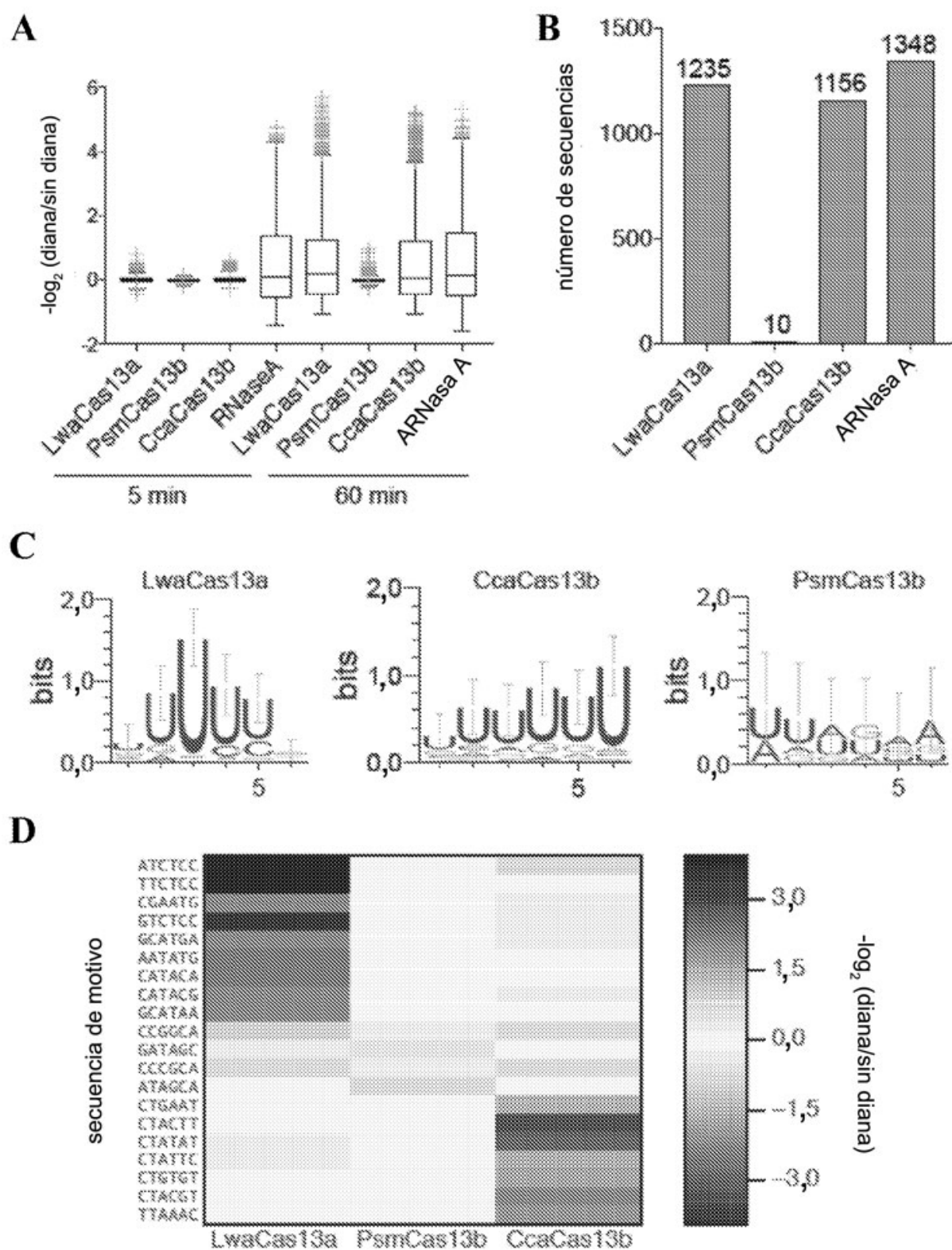


FIG. 89

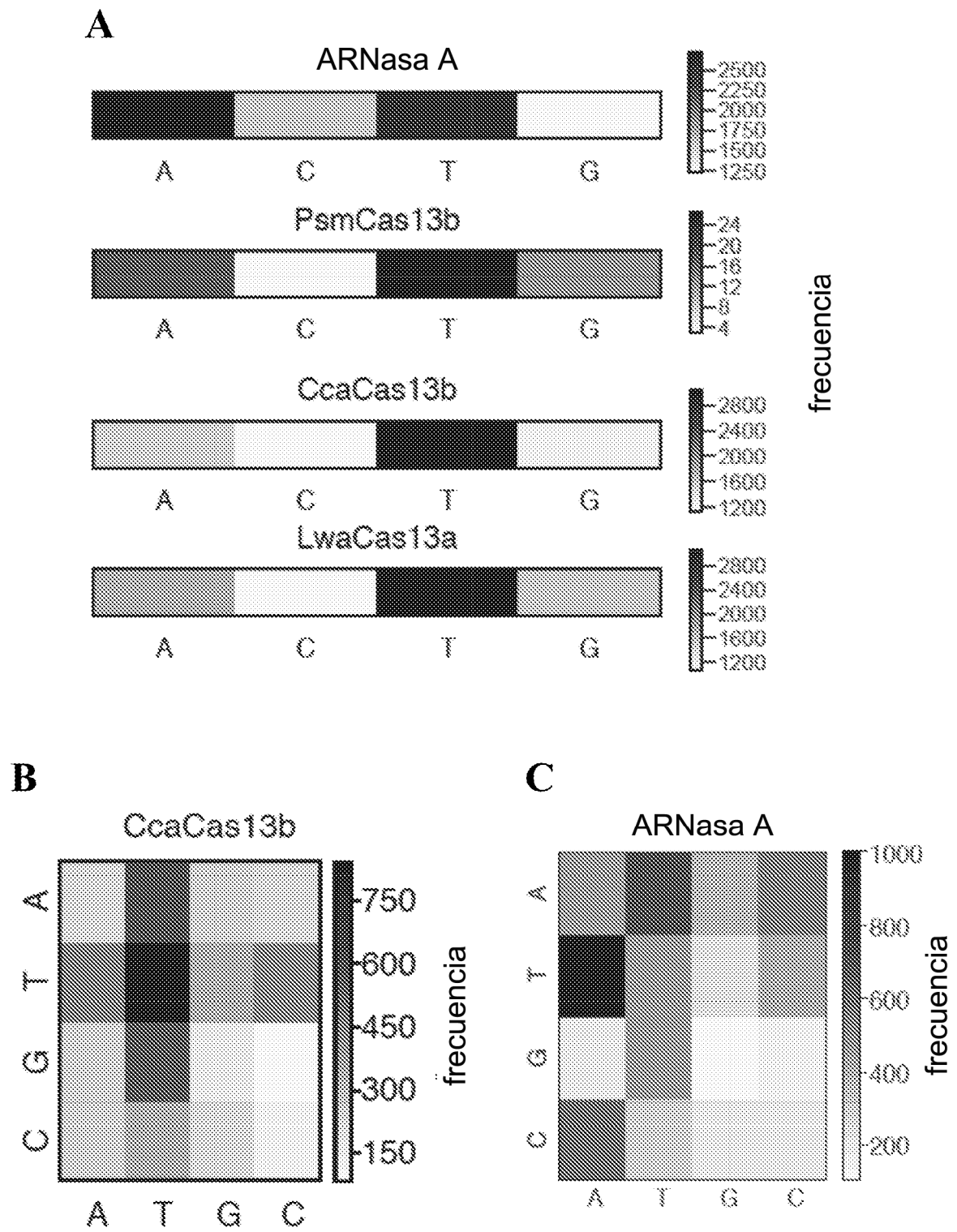


FIG. 90

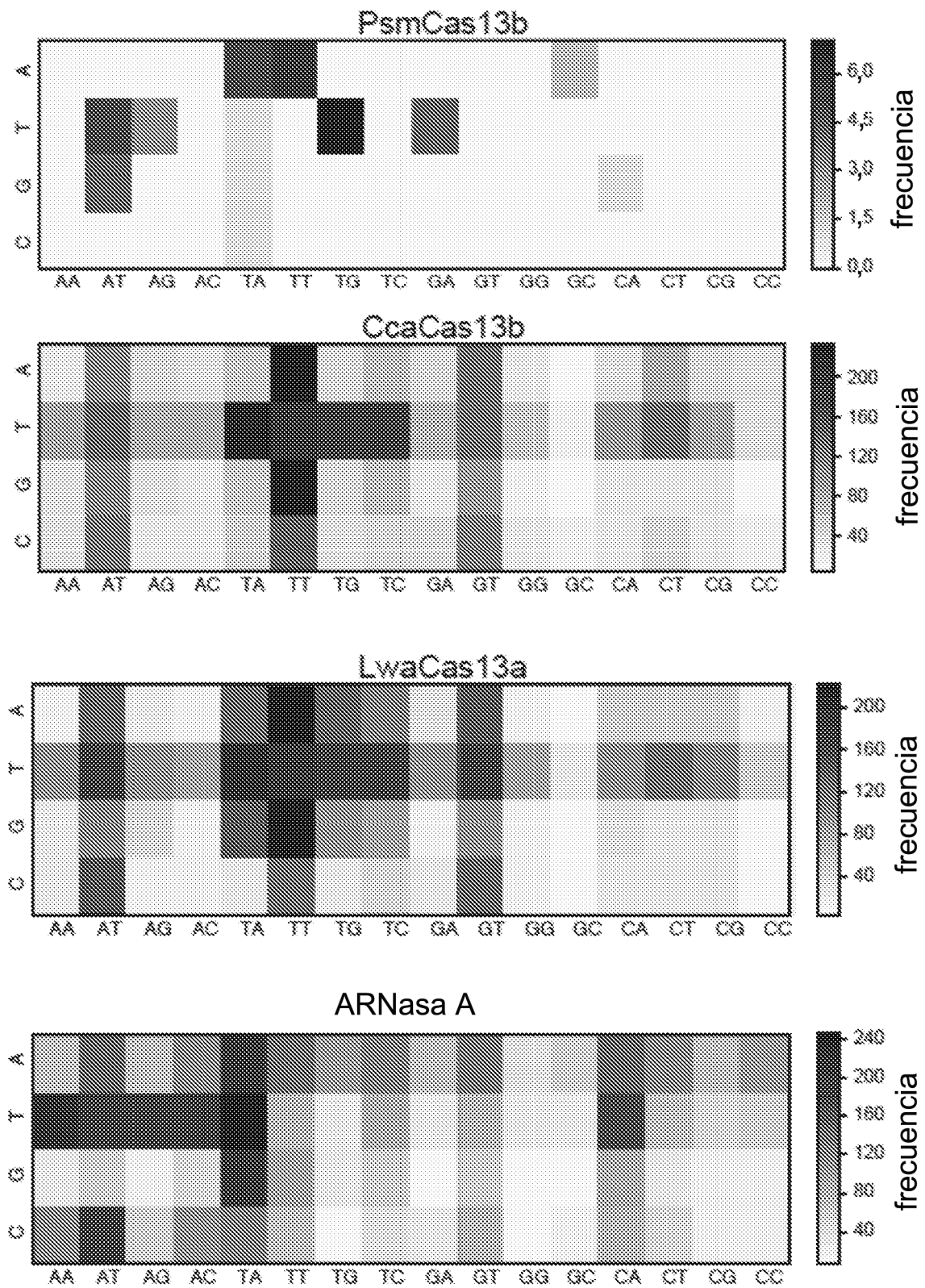


FIG. 91

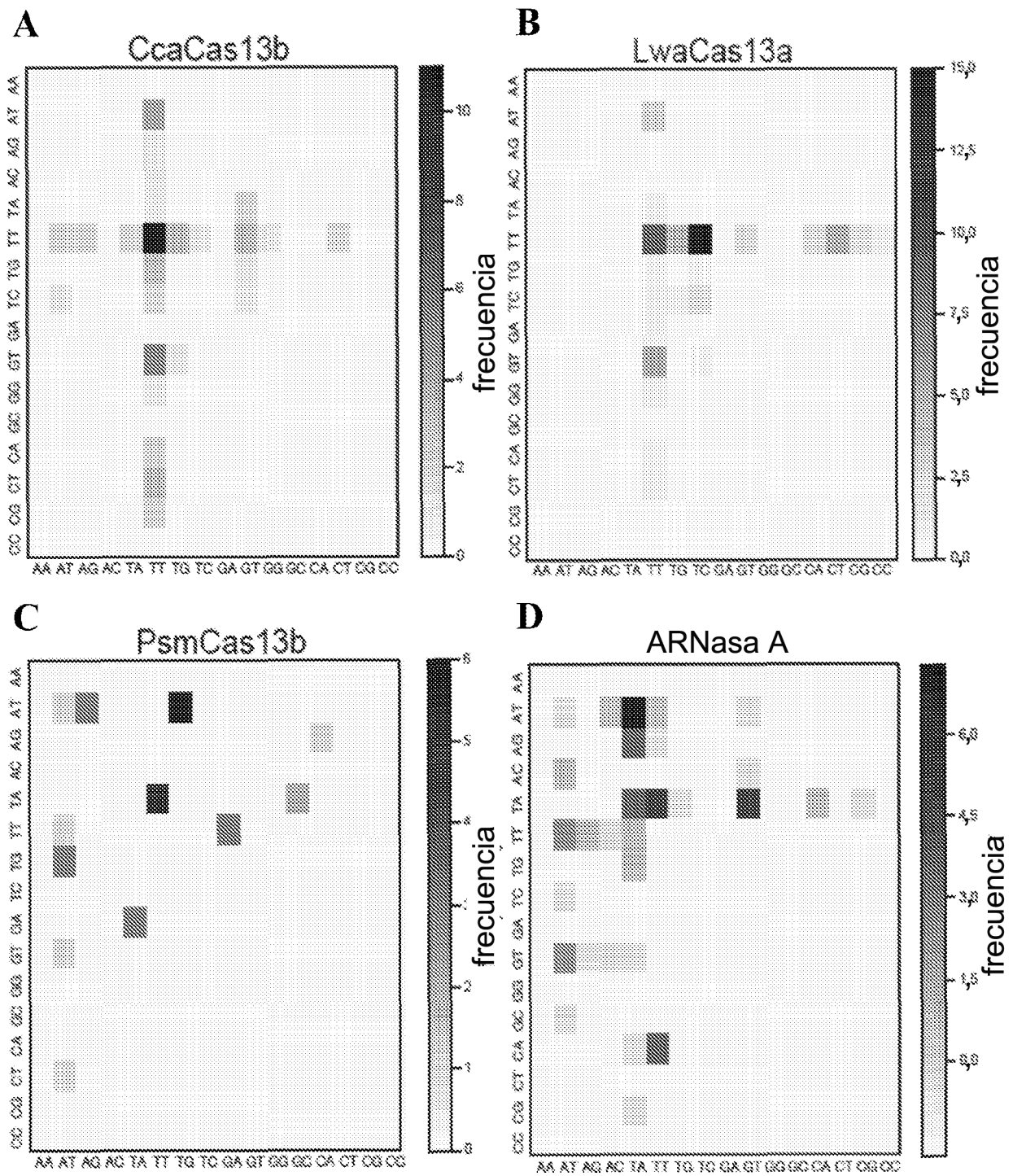


FIG. 92

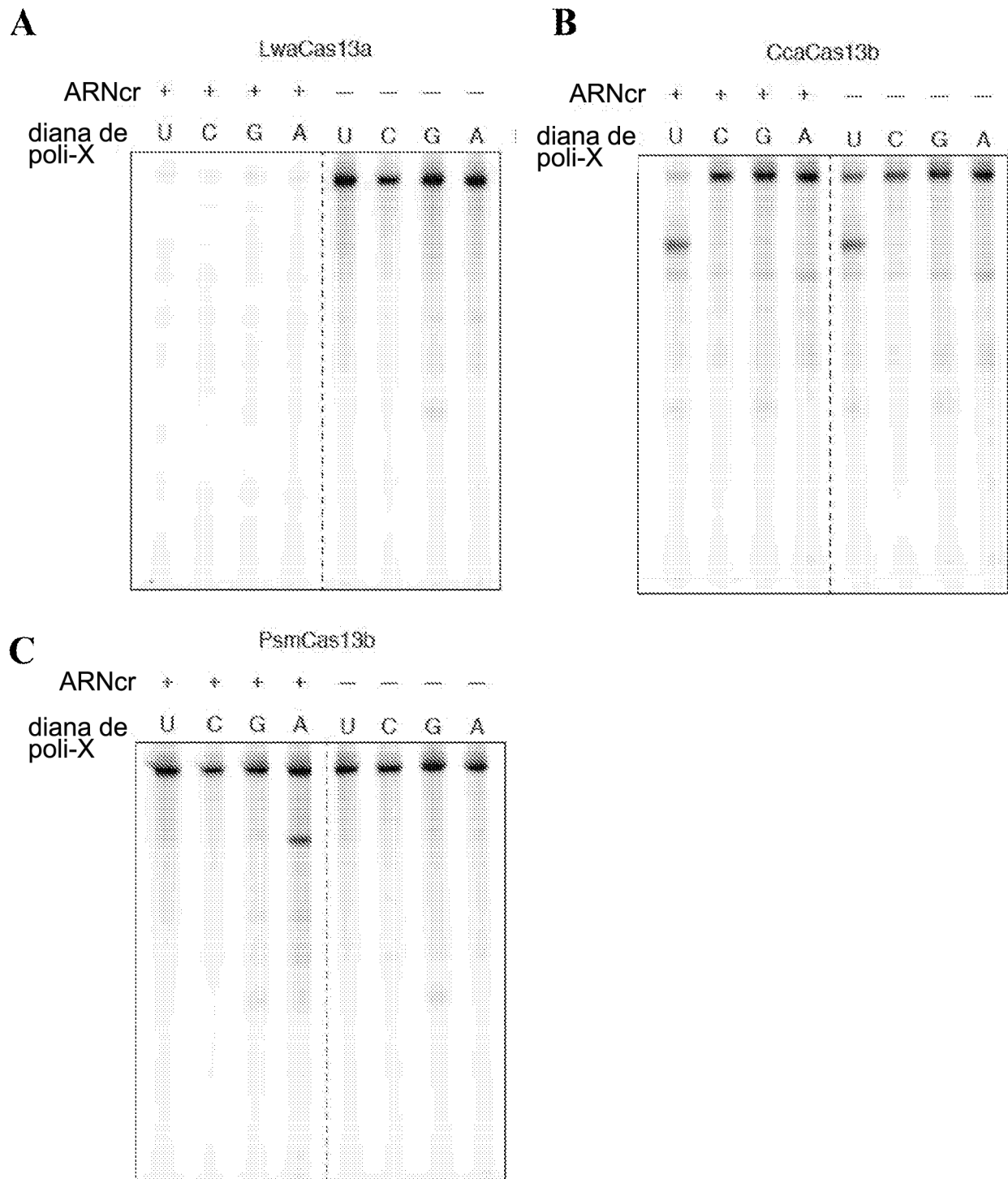


FIG. 93

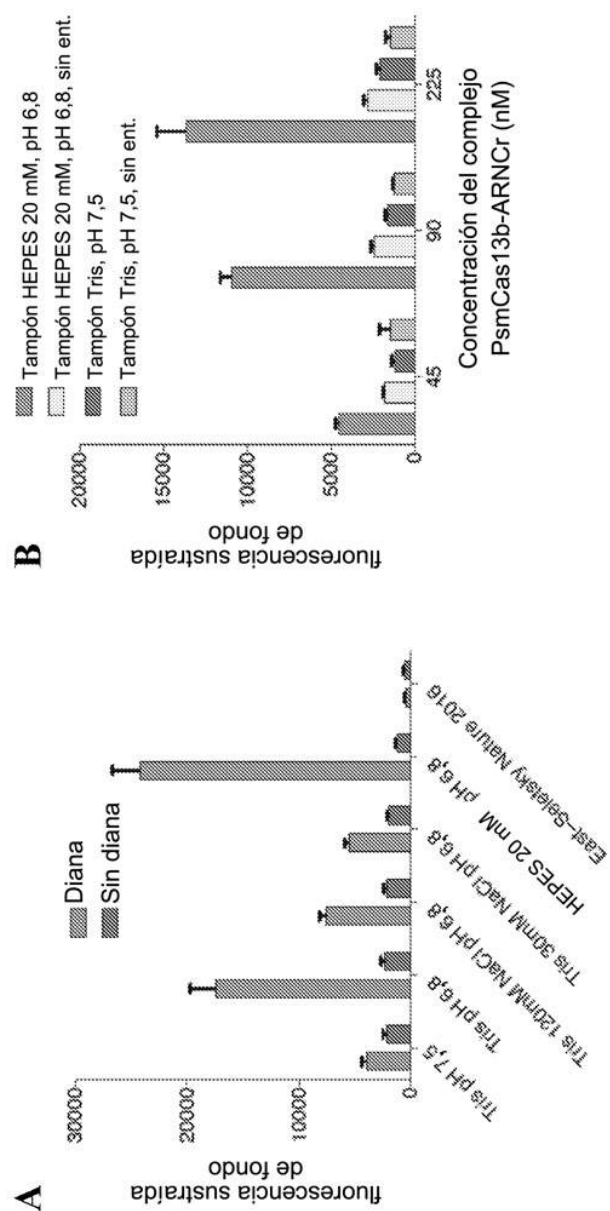


FIG. 94

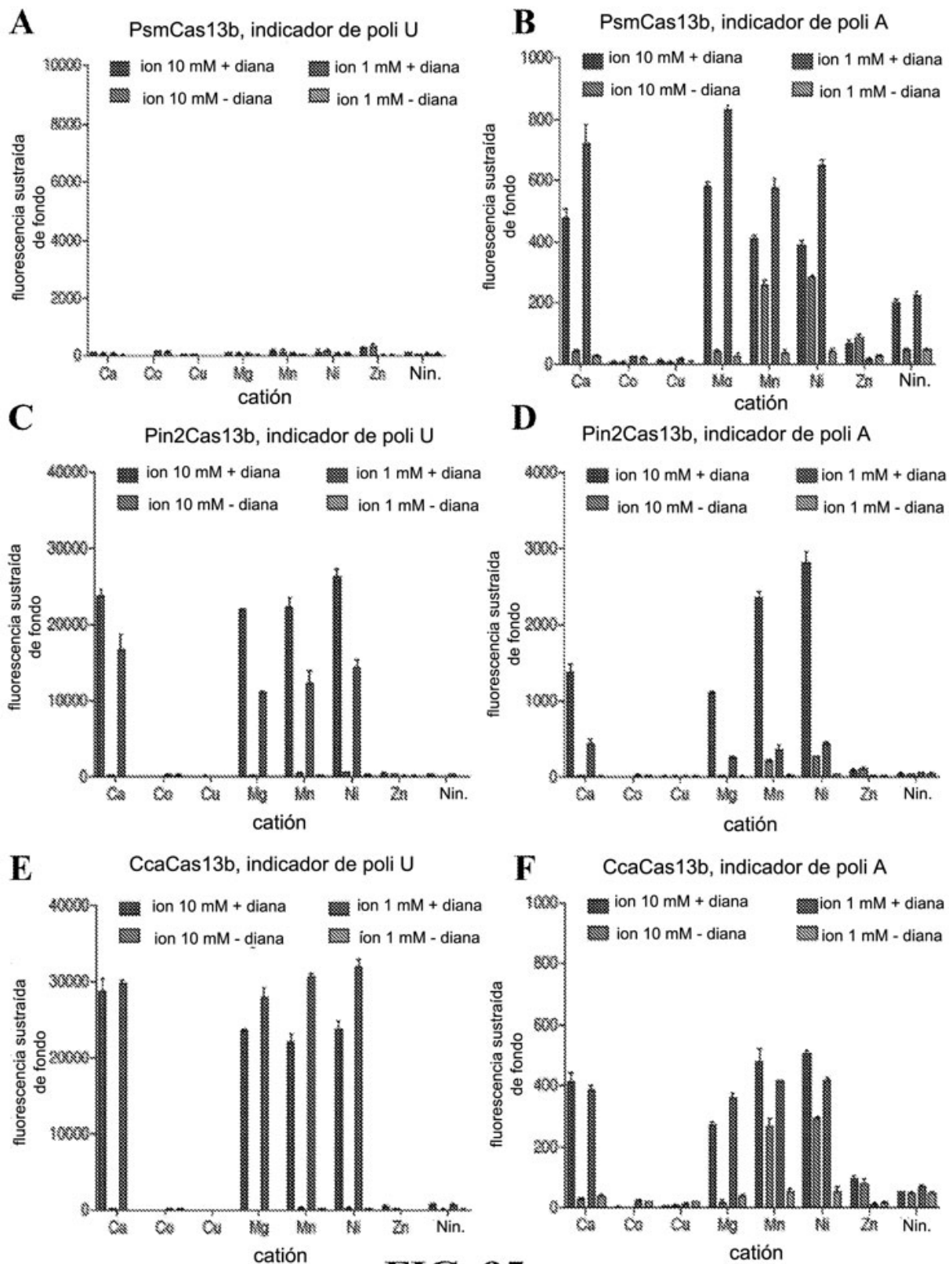


FIG. 95

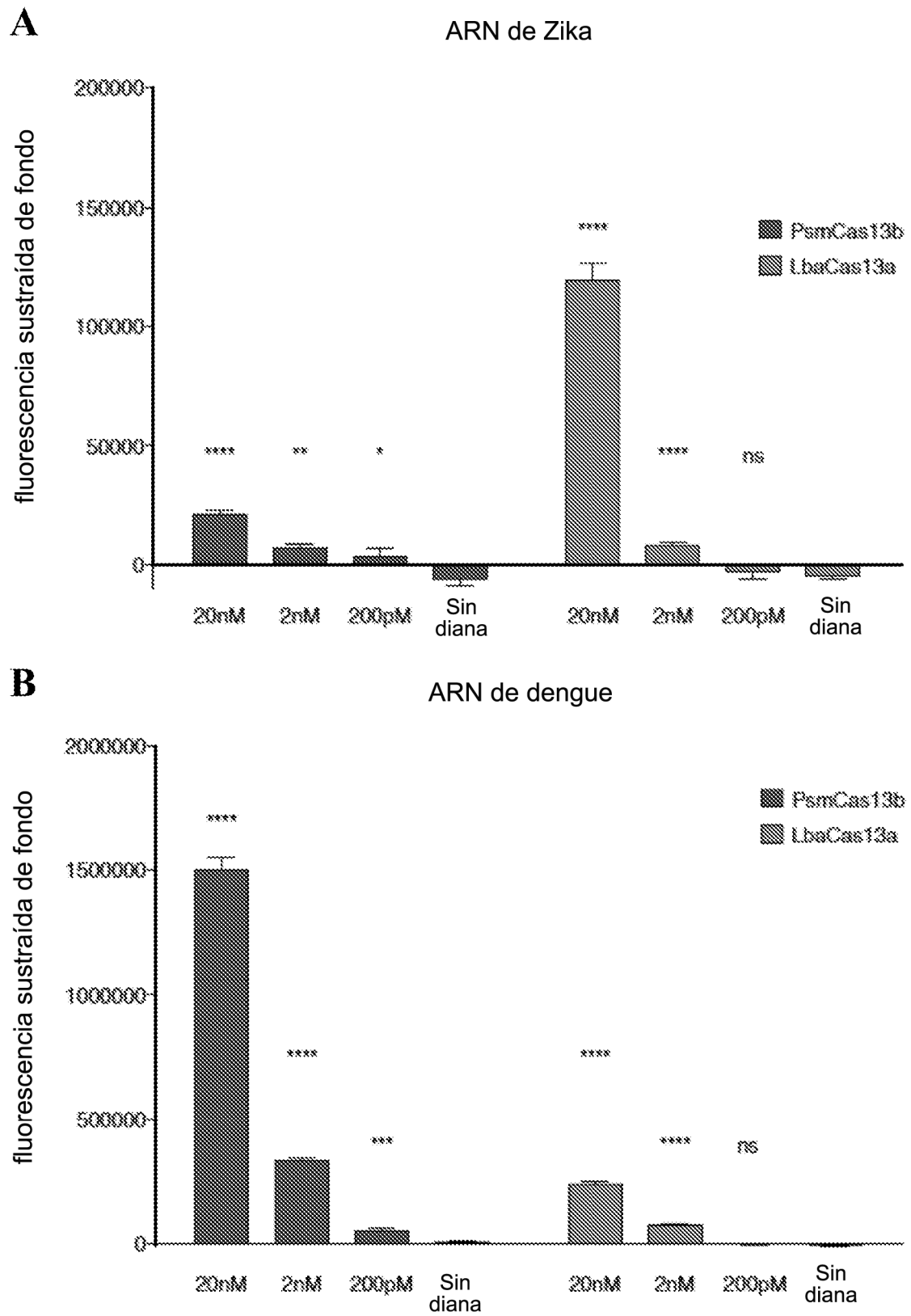
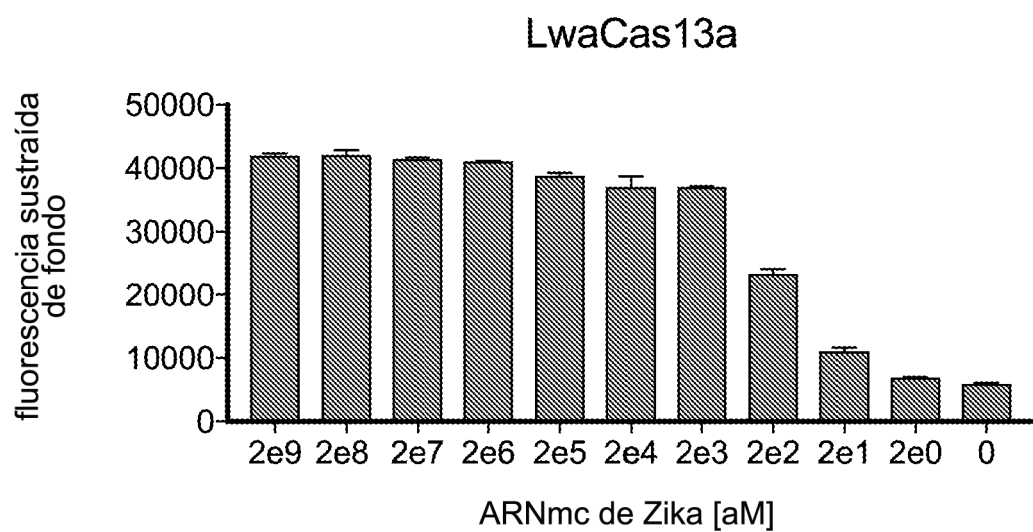
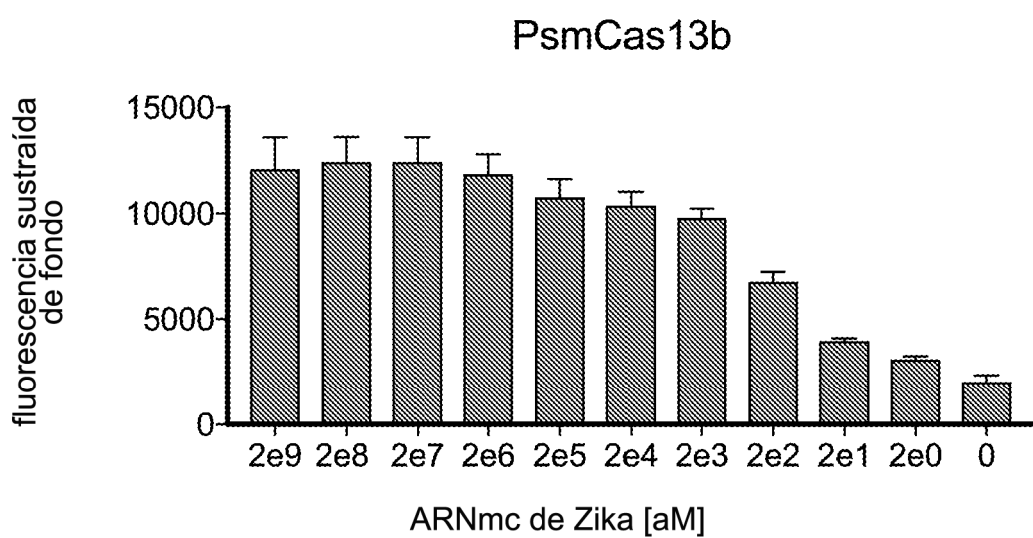


FIG. 96

A**B****FIG. 97**

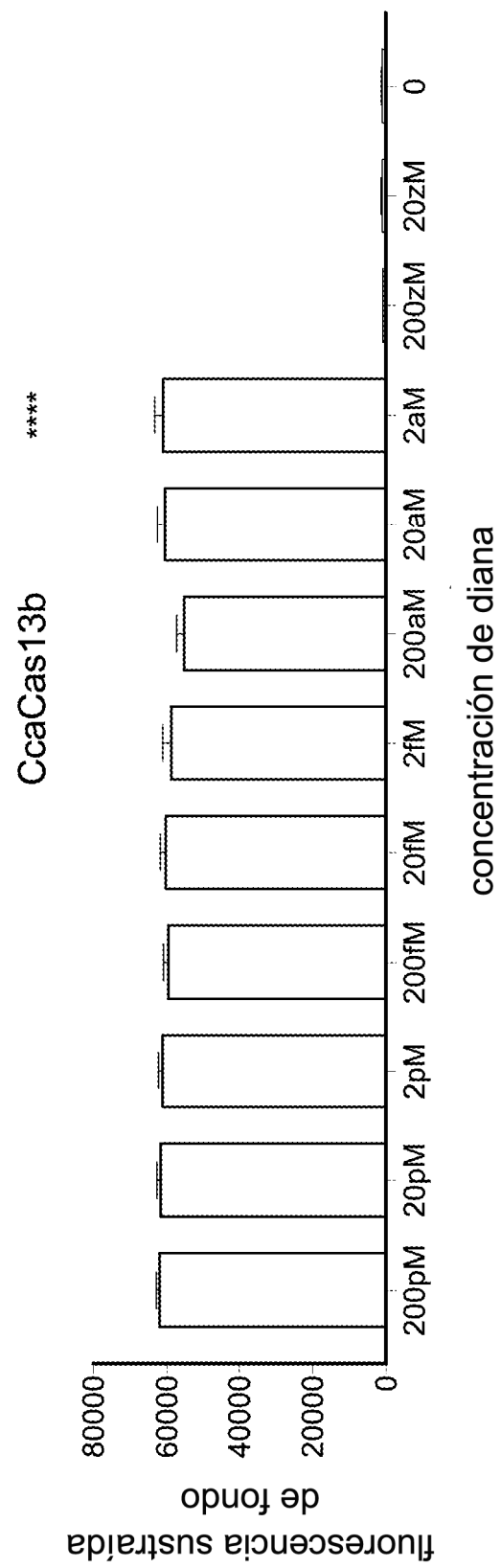


FIG. 98

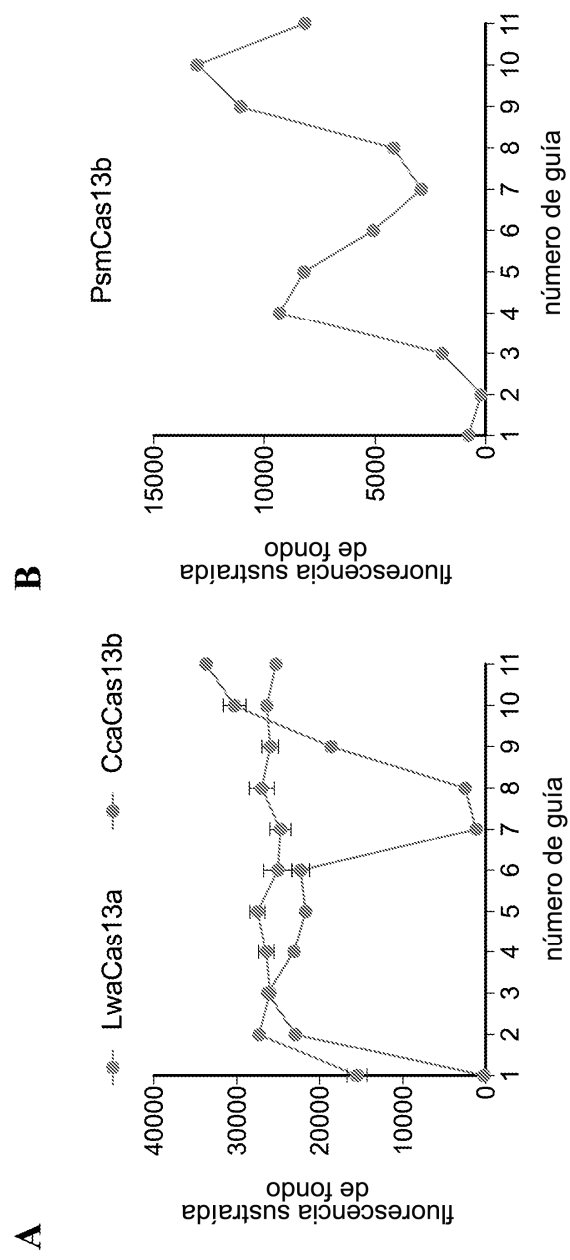


FIG. 99

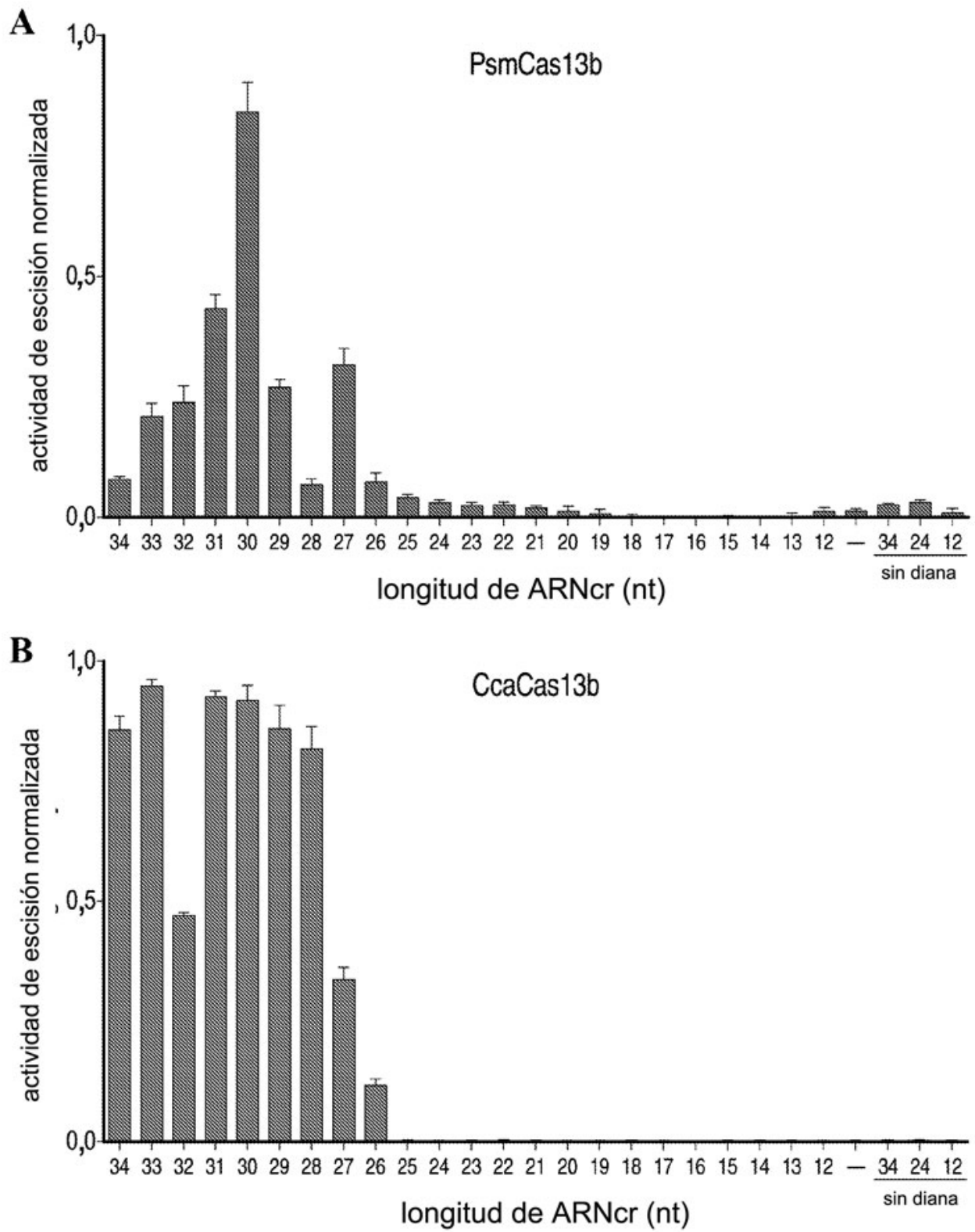


FIG. 100

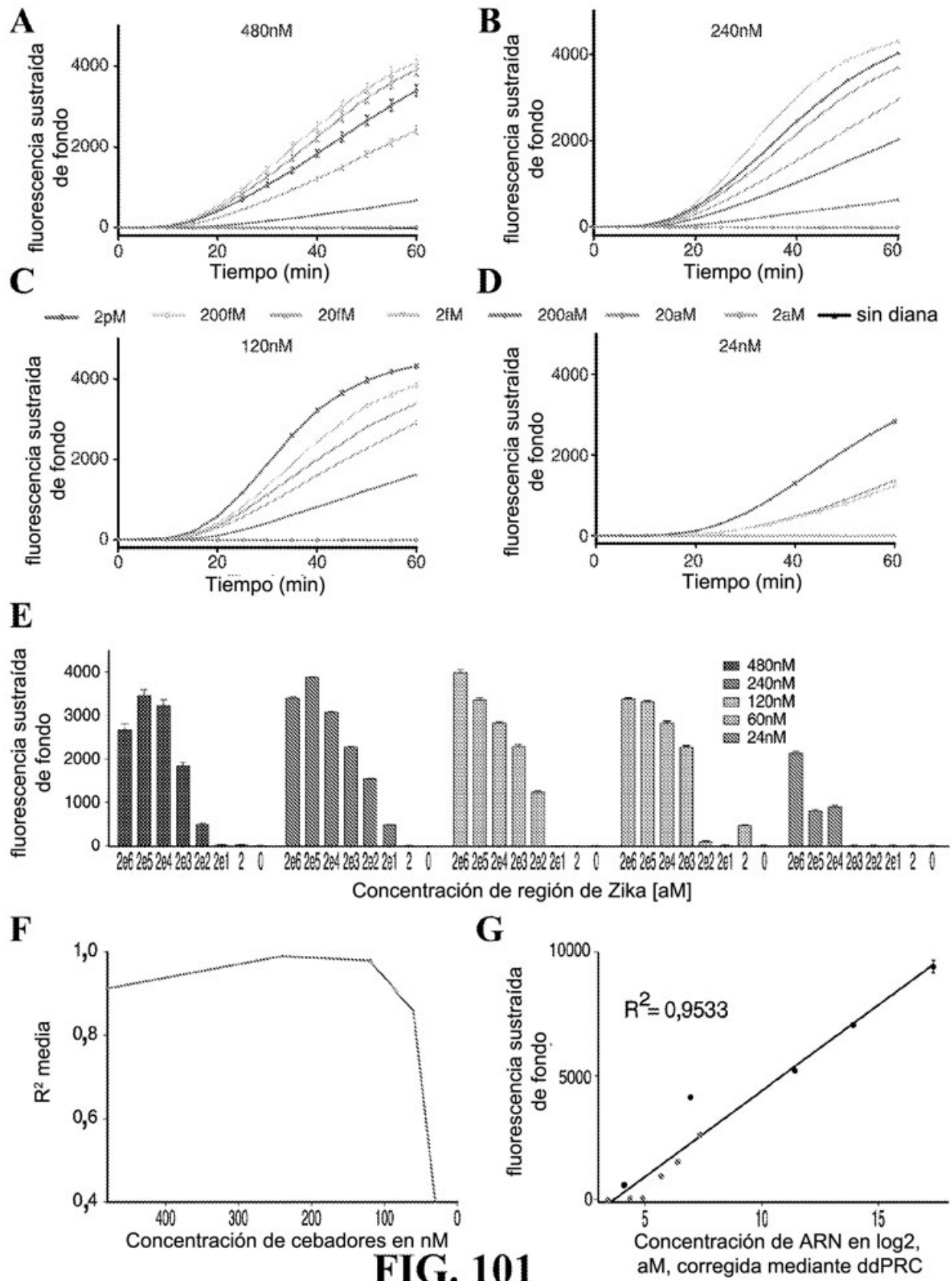
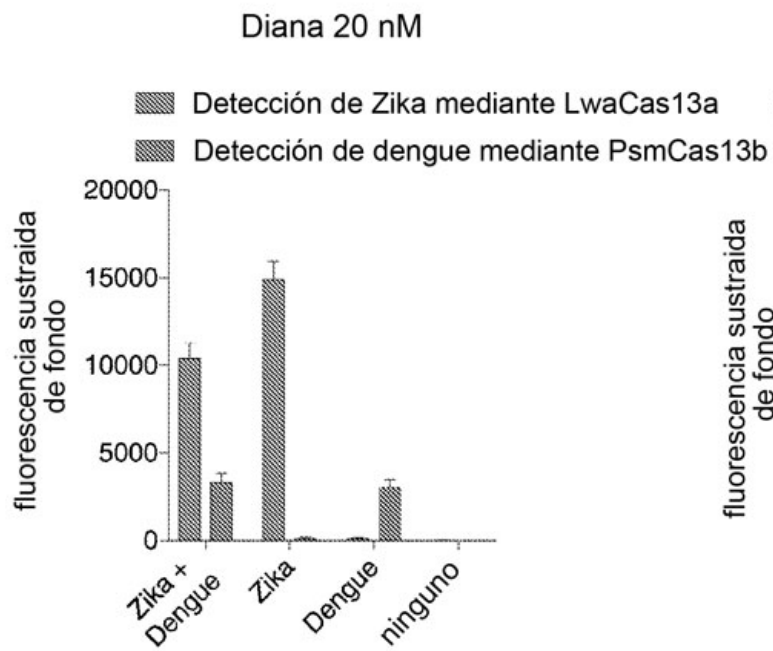
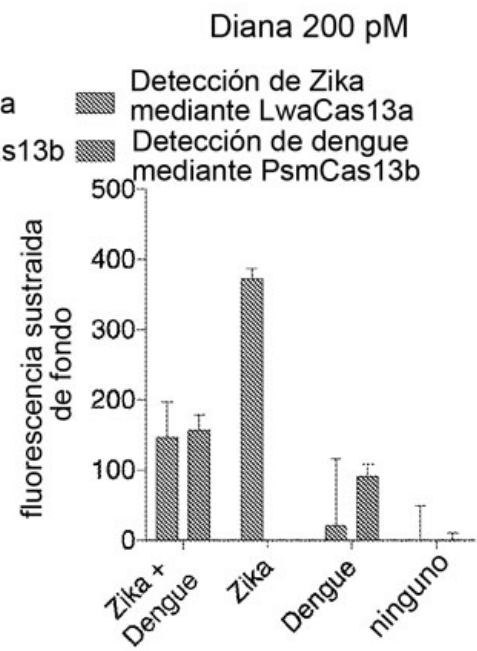
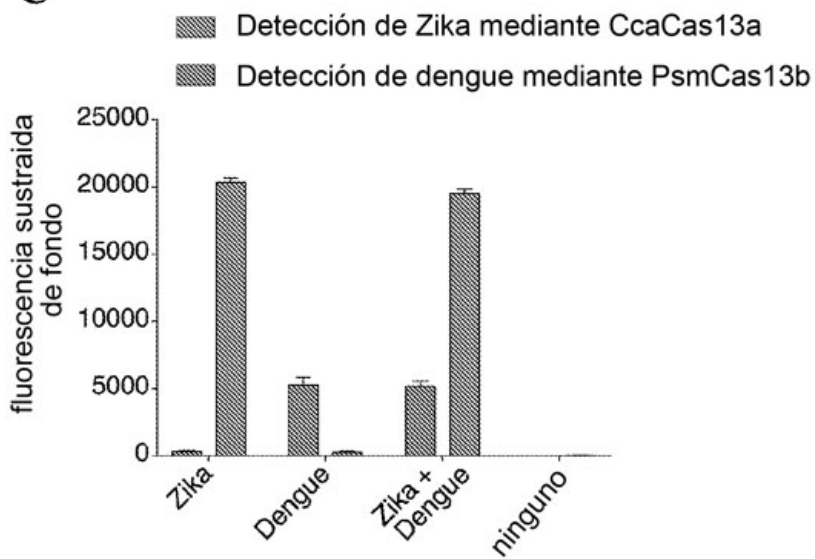


FIG. 101

A**B****C****FIG. 102**

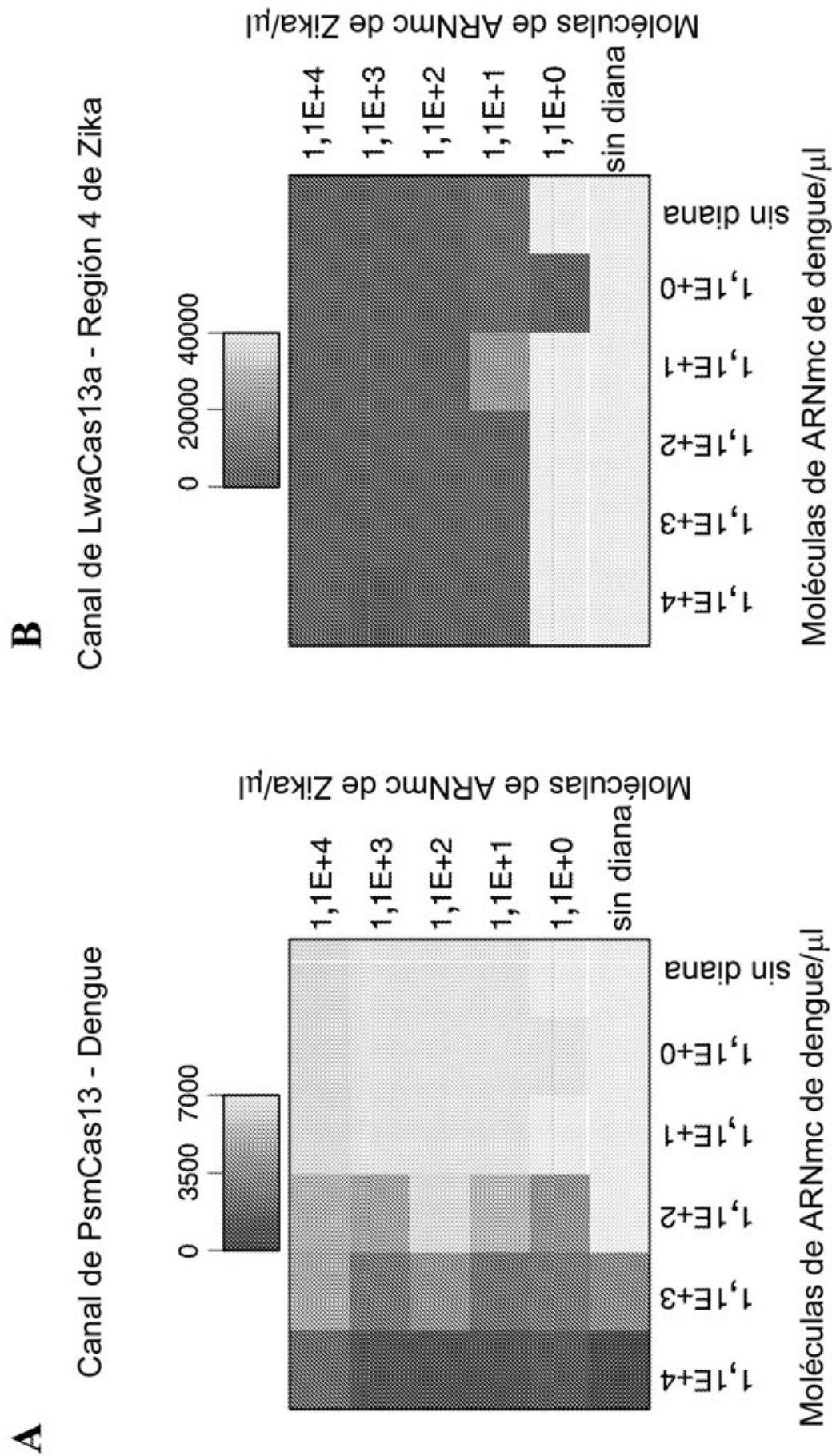


FIG. 103

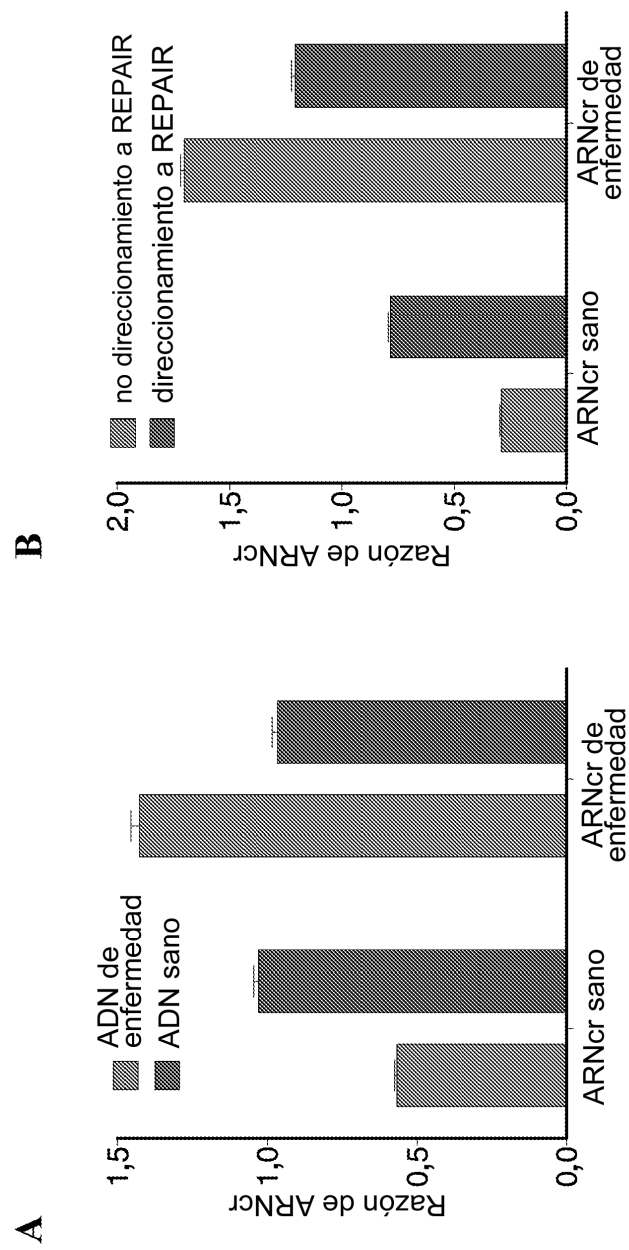


FIG. 104

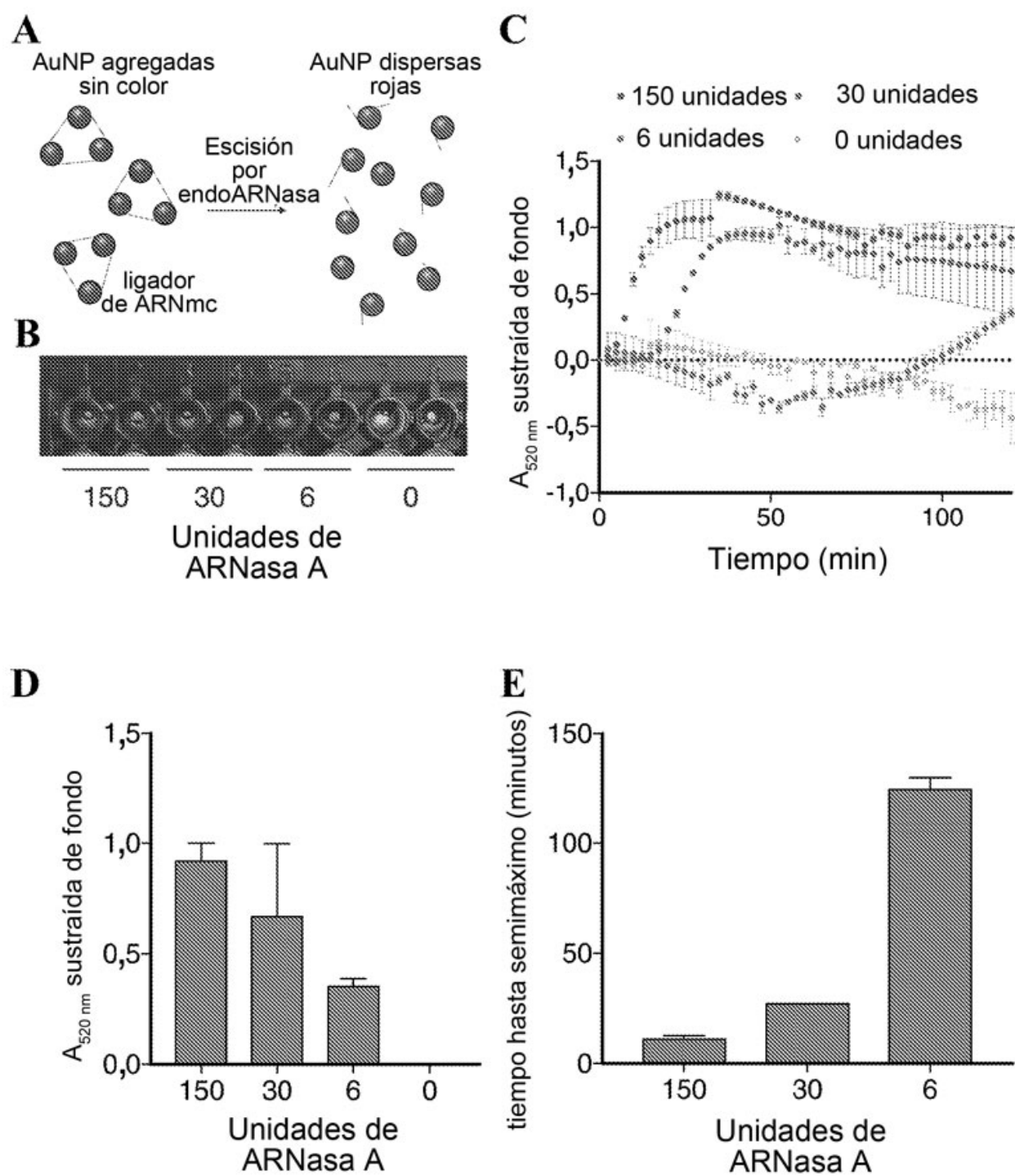
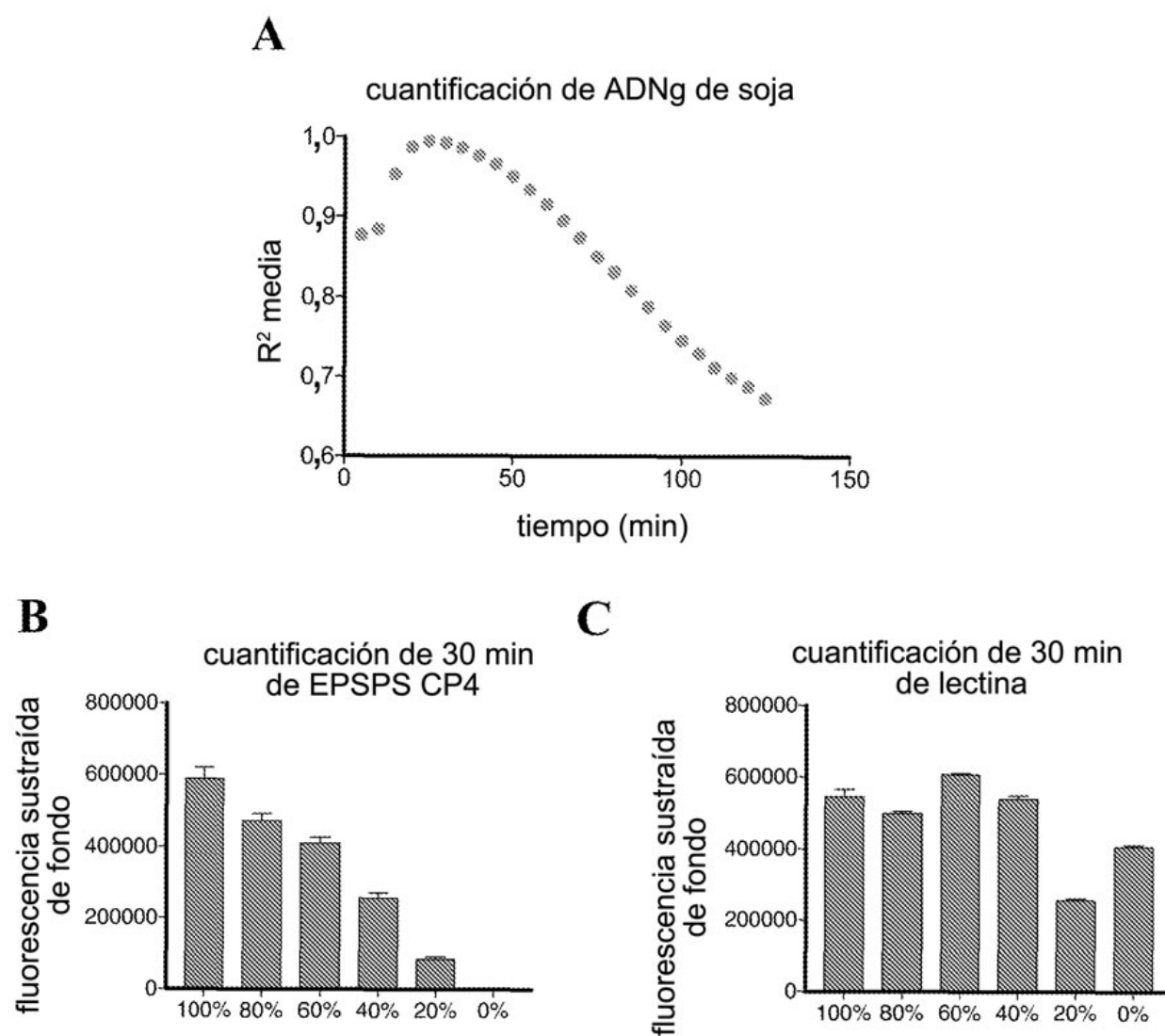


FIG. 105

**FIG. 106**