

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 690 386**

51 Int. Cl.:

<b>C12N 15/11</b>	(2006.01)
<b>C12N 15/90</b>	(2006.01)
<b>C12N 9/22</b>	(2006.01)
<b>C12N 9/16</b>	(2006.01)
<b>C12N 15/10</b>	(2006.01)
<b>C12N 15/52</b>	(2006.01)
<b>C12N 15/63</b>	(2006.01)
<b>C12N 15/82</b>	(2006.01)

12

### TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **23.10.2013 PCT/KR2013/009488**
- 87 Fecha y número de publicación internacional: **01.05.2014 WO14065596**
- 96 Fecha de presentación y número de la solicitud europea: **23.10.2013 E 13849670 (8)**
- 97 Fecha y número de publicación de la concesión europea: **22.08.2018 EP 2912175**

54 Título: **Composición para escindir un ADN diana que comprende un ARN guía específico para el ADN diana y el ácido nucleico que codifica la proteína Cas o la propia proteína Cas, y sus usos**

30 Prioridad:

**23.10.2012 US 201261717324 P**  
**20.03.2013 US 201361803599 P**  
**20.06.2013 US 201361837481 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**20.11.2018**

73 Titular/es:

**TOOLGEN INCORPORATED (100.0%)**  
**1204, Byucksan Digital Valley 6-cha, 219 Gasan**  
**Digital 1-ro, Geumcheon-gu**  
**Seoul 08501, KR**

72 Inventor/es:

**KIM, JIN-SOO;**  
**CHO, SEUNG WOO;**  
**KIM, SOJUNG;**  
**KIM, JONG MIN y**  
**KIM, SEOKJOONG**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 690 386 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

Composición para escindir un ADN diana que comprende un ARN guía específico para el ADN diana y el ácido nucleico que codifica la proteína Cas o la propia proteína Cas, y sus usos

### Campo de la invención

5 Como se define en las reivindicaciones, la presente invención se refiere de forma general a la edición dirigida del genoma en células u organismos eucarióticos. Más particularmente, la presente invención se refiere de forma general a una composición para escindir un ADN diana en células eucariotas u organismos que comprenden un ARN guía específico para el ADN diana y el ácido nucleico que codifica la proteína Cas9 o la propia proteína Cas9, y sus usos.

10 Las CRISPR (por sus siglas en inglés, sistema de repeticiones palindrómicas cortas agrupadas y regularmente inter-espaciadas) son loci que contienen múltiples repeticiones directas cortas que se encuentran en los genomas de aproximadamente el 40% de las bacterias secuenciadas y el 90% de las arqueas secuenciadas. Las CRISPR funcionan como un sistema inmune procariótico, ya que confiere resistencia a elementos genéticos exógenos tales como plásmidos y fagos. El sistema CRISPR proporciona una forma de inmunidad adquirida. Los segmentos cortos de ADN ajeno, llamados espaciadores, se incorporan en el genoma entre repeticiones CRISPR y sirven como memoria de exposiciones pasadas. Los espaciadores CRISPR se usan luego para reconocer y silenciar elementos genéticos exógenos de una manera análoga al ARNi en organismos eucariotas.

20 Cas9, un componente proteico esencial en el sistema CRISPR/Cas de tipo II, forma una endonucleasa activa cuando se compleja con dos ARN denominados ARN CRISPR (ARNcr) y ARNcr trans-activador (ARNtracr), rebanando así elementos genéticos externos en fagos o plásmidos invasores para proteger las células huésped. El ARNcr se transcribe desde el elemento CRISPR en el genoma del huésped, que previamente se capturó de tales invasores externos. Recientemente, Jinek et al. (1) han demostrado que un ARN quimérico monocatenario producido fusionando una porción esencial de ARNcr y ARNtracr puede reemplazar los dos ARN en el complejo Cas9/ARN para formar una endonucleasa funcional.

25 Los sistemas CRISPR/Cas ofrecen una ventaja al dedo de zinc y a las proteínas de unión de ADN efectoras tipo activadoras de la transcripción, ya que la especificidad del sitio en las proteínas CRISPR-Cas de unión de nucleótidos está gobernada por una molécula de ARN, en lugar de la proteína de unión al ADN, que puede ser más complicada de diseñar y sintetizar.

30 Sin embargo, hasta ahora, no se ha desarrollado un método de edición del genoma que utilice la endonucleasa guiada por ARN (RGEN) basada en el sistema CRISPR/Cas.

Mientras tanto, el polimorfismo de longitud de fragmentos de restricción (RFLP) es uno de los métodos de genotipificación más antiguo, más conveniente y menos costoso que todavía se usa ampliamente en biología molecular y genética, pero a menudo está limitado por la falta de sitios apropiados reconocidos por las endonucleasas de restricción.

35 Las mutaciones inducidas por nucleasas modificadas se detectan mediante diversos métodos, que incluyen la T7 endonucleasa I sensible al malapareamiento (T7E1) o ensayos de nucleasa Surveyor, RFLP, electroforesis capilar de productos de PCR fluorescentes, secuenciación didesoxi y secuenciación profunda. Los ensayos T7E1 y Surveyor son ampliamente utilizados, pero son engorrosos. Además, estas enzimas tienden a subestimar las frecuencias de mutación debido a que las secuencias mutantes pueden formar homodúplex entre ellas y no pueden distinguir clones mutantes bi-alélicos homocigóticos de las células naturales. RFLP está libre de estas limitaciones y, por lo tanto, es un método de elección. De hecho, RFLP fue uno de los primeros métodos para detectar mutaciones modificadas por nucleasas en células y animales. Desafortunadamente, sin embargo, RFLP está limitado por la disponibilidad de sitios de restricción apropiados. Es posible que no haya sitios de restricción disponibles en el sitio de interés de destino.

### 45 Descripción de la invención

Hasta ahora, no se ha desarrollado un método de edición y genotipificación del genoma utilizando la endonucleasa guiada por ARN (RGEN) basada en el sistema CRISPR/Cas.

50 Bajo estas circunstancias, los presentes inventores han realizado muchos esfuerzos para desarrollar un método de edición del genoma basado en el sistema CRISPR/Cas y, finalmente, han establecido una endonucleasa guiada por ARN programable que escinde el ADN de manera dirigida en células y organismos eucarióticos.

Además, los presentes inventores han realizado muchos esfuerzos para desarrollar un nuevo método para usar endonucleasas guiadas por ARN (RGEN) en el análisis de RFLP. Han usado RGEN para genotipificar mutaciones recurrentes encontradas en el cáncer y aquellas inducidas en células y organismos por nucleasas diseñadas que incluyen las propias RGENs, completando así la presente invención.

### Sumario de la invención

Un objeto de la presente invención es proporcionar una composición para escindir ADN diana en células u organismos eucariotas que comprende un ARN específico guía para ADN diana o ADN que codifica el ARN guía, y el ácido nucleico que codifica la proteína Cas9 o la proteína Cas9, como se define en las reivindicaciones.

- 5 En un primer aspecto, la presente invención proporciona un sistema de repeticiones palindrómicas cortas agrupadas con regularidad entremezcladas Tipo II (CRISPR)/proteína asociada a CRISPR 9 (Cas9) para introducir roturas bicatenarias en una secuencia de ADN diana en una célula de mamífero, en donde el sistema Tipo II CRISPR/Cas9 comprende: a) una proteína Cas9 con una señal de localización nuclear (NLS), en la que el NLS está en el extremo C, o un ácido nucleico que codifica la proteína Cas9; y b) un ARN guía monocatenario que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción de trans-activación de ARNcr (ARNtracr).

10 En una realización, la proteína Cas9 es de *Streptococcus pyogenes*.

En otra realización, el ARN guía monocatenario es un ARN transcrito in vitro.

En otra realización, la secuencia de ADN diana es una secuencia genómica localizada en su sitio endógeno en el genoma de la célula de mamífero.

- 15 En otra realización, la célula de mamífero es una célula humana.

En otra realización, el ácido nucleico que codifica la proteína Cas9 está optimizada en el codón para la expresión en células humanas.

- 20 En otra realización, la secuencia de ADN diana consiste en 20 nucleótidos complementarios a la porción de ARNcr del ARN guía monocatenario y un motivo adyacente del motivo adyacente de protoespaciador del trinucleótido (PAM), y en donde el PAM consiste en el trinucleótido 5'-NGG-3'.

En un segundo aspecto, la invención proporciona un sistema CRISPR del tipo II/Cas9, como se define en el primer aspecto para su uso en la escisión de un ADN diana en una célula de mamífero.

En un tercer aspecto, la invención proporciona un sistema CRISPR del tipo II/Cas9, como se define en el primer aspecto para su uso en terapia génica.

- 25 En un cuarto aspecto, la invención proporciona un método in vitro para introducir una rotura bicatenaria específica del sitio en una secuencia de ADN diana en una célula de mamífero, comprendiendo el método introducir en la célula de mamífero el sistema CRISPR del tipo II/Cas9 del primer aspecto.

- 30 En un quinto aspecto, la invención proporciona un método in vitro para introducir una rotura bicatenaria específica del sitio en una secuencia de ADN diana en una célula de mamífero, comprendiendo el método introducir en la célula de mamífero: a) una proteína Cas9 con una señal de localización nuclear (NLS), en la que el NLS está en el extremo C-terminal, o un ácido nucleico que codifica la proteína Cas9; y b) un ARN guía monocatenario que comprende una porción de ARN CRISPR (ARNcr) fijada a una porción de ARNcr de trans-activación (ARNtracr).

En una realización, el ácido nucleico que codifica la proteína Cas9 se introduce en la célula de mamífero antes de introducir el ARN guía monocatenario en la célula de mamífero.

- 35 En otra realización, la célula de mamífero es una célula humana.

El sistema de la invención también se denomina en este documento "composición", "composición de RGEN" o "RGEN(s)".

Se debe entender que cualquier referencia en este documento a embriones excluye embriones humanos; y, del mismo modo, cualquier referencia a los animales debe entenderse que excluye a los seres humanos.

- 40 La composición de la invención puede usarse para inducir mutagénesis dirigida en células u organismos eucarióticos.

El sistema de la invención puede proporcionarse en un kit para escindir un ADN diana en células eucariotas u organismos.

- 45 El sistema de la invención puede proporcionarse en un kit para inducir mutagénesis dirigida en células u organismos eucarióticos.

El sistema de la invención puede usarse para preparar una célula u organismo eucariótico que comprende la proteína Cas9 y un ARN guía que comprende una etapa de co-transfección o transfección en serie de la célula u organismo eucariótico.

El sistema de la invención puede estar comprendido en una célula u organismo eucariótico.

El sistema de la invención puede usarse para escindir un ADN diana en células u organismos eucarióticos que comprende una etapa de transfección de las células u organismos eucarióticos.

El sistema de la invención puede usarse para inducir mutagénesis dirigida en una célula u organismo eucariótico.

5 El sistema de la invención puede usarse para proporcionar un embrión, un animal modificado genéticamente o una planta modificada genéticamente que comprende un genoma editado por el sistema.

El sistema de la invención puede usarse en para preparar un animal genéticamente modificado que comprende una etapa para introducir el sistema en un embrión de un animal; y una etapa para transferir el embrión a un oviducto de una madre adoptiva pseudopreñada para producir un animal genéticamente modificado.

10 La composición de la invención puede usarse para genotipar mutaciones o variaciones en una muestra biológica aislada.

La composición de endonucleasa guiada por ARN (RGEN) de la invención puede usarse para genotipar mutaciones inducidas por nucleasas modificadas genéticamente en células o mutaciones o variaciones de origen natural.

15 La composición de endonucleasa guiada por ARN (RGEN) de la invención puede proporcionarse en un kit para genotipar mutaciones inducidas por nucleasas modificadas genéticamente en células o mutaciones o variaciones de origen natural.

La composición de la invención puede usarse para escindir ADN diana en células eucariotas u organismos.

La composición de la invención puede usarse para inducir mutagénesis dirigida en células u organismos eucarióticos.

20 La composición de la invención puede usarse para genotipar secuencias de ácidos nucleicos en microorganismos patógenos en una muestra biológica aislada.

El sistema de la invención puede proporcionarse en un kit para genotipar mutaciones o variaciones en una muestra biológica aislada.

El sistema de la invención puede usarse para genotipar mutaciones o variaciones en una muestra biológica aislada.

Efectos ventajosos de la invención

25 El sistema, su uso y los métodos de la invención proporcionan nuevas herramientas de edición del genoma adecuadas. Además, debido a que pueden diseñarse RGEN personalizados para reconocer a cualquier secuencia de ADN, puede analizarse casi cualquier polimorfismo de nucleótido simple o inserción/delección (indel) pequeña a través de RFLP mediado por RGEN; por lo tanto, la composición y el método de la presente invención pueden usarse en la detección y eliminación de variaciones y mutaciones de origen natural.

30 **Breve descripción de los dibujos**

35 La Fig. 1 muestra la escisión catalizada por Cas9 de ADN de plásmido in vitro. (a) Representación esquemática de ADN diana y secuencias de ARN quimérico. Los triángulos rojos indican sitios de escisión. La secuencia PAM reconocida por Cas9 se muestra en negrita. Las secuencias en el ARN guía derivadas de ARNcr y ARNtracr se muestran en recuadro y se subrayan, respectivamente. (b) Eliminación in vitro del ADN del plásmido por Cas9. Se incubó un plásmido circular intacto o un plásmido digerido con ApaI con Cas9 y ARN guía.

40 La Fig. 2 muestra la mutagénesis inducida por Cas9 en un sitio diana episómico. (a) Resumen esquemático de ensayos basados en células usando un informador RFP-GFP. GFP no se expresa desde este reportero porque la secuencia GFP se fusiona con la secuencia RFP fuera del marco. La proteína de fusión RFP-GFP se expresa solo cuando el sitio diana entre las dos secuencias se escinde mediante una nucleasa específica de sitio. (b) Citometría de flujo de células transfectadas con Cas9. Se indica el porcentaje de células que expresan la proteína de fusión RFP-GFP.

45 La Fig. 3 muestra mutaciones impulsadas por RGEN en sitios cromosómicos endógenos. (a) Locus CCR5. (b) Locus C4BPB. (Arriba) El ensayo T7E1 se usó para detectar mutaciones inducidas por RGEN. Las flechas indican la posición esperada de las bandas de ADN escindidas por T7E1. Las frecuencias de mutación (Inserciones-delecciones (%)) se calcularon midiendo las intensidades de la banda. (Parte inferior) Secuencias de ADN de CCR5 y C4BPB naturales (WT) y clones mutantes. La región de la secuencia diana complementaria al ARN guía se muestra en boc. La secuencia PAM se muestra en negrita. Los triángulos indican el sitio de división. Las bases correspondientes a microhomologías están subrayadas. La columna a la derecha indica el número de bases insertadas o eliminadas.

50 La Fig. 4 muestra que las mutaciones fuera de la diana impulsadas por RGEN son indetectables. (a) Secuencias en la diana y potenciales fuera de la diana. El genoma humano se buscó in silico para posibles sitios fuera de la diana. Se identificaron cuatro sitios, cada uno de los cuales lleva apareamientos erróneos de 3 bases con el sitio CCR5 en

la diana. Las bases no coincidentes están subrayadas. (b) El ensayo T7E1 se usó para investigar si estos sitios estaban mutados en células transfectadas con el complejo Cas9/ARN. No se detectaron mutaciones en estos sitios. N/A (no aplicable), un sitio intergénico. (c) Cas9 no indujo deleciones cromosómicas asociadas a la diana. El RGEN y ZFN específicos de CCR5 se expresaron en células humanas. La PCR se usó para detectar la inducción de deleciones cromosómicas de 15 kb en estas células.

La Fig. 5 muestra el gen *Foxn1* inducido por RGEN dirigido a ratones. (a) Un diagrama esquemático que representa un ARNsg específico para el exón 2 del gen *Foxn1* de ratón. PAM en el exón 2 se muestra en rojo y la secuencia en el ARNsg que es complementaria al exón 2 está subrayada. Los triángulos indican sitios de escisión. (b) Ensayos representativos de T7E1 que demuestran las eficacias de reconocimiento génico del ARNm de Cas9 más el ARNsg específico de *Foxn1* que se administraron mediante inyección intracitoplasmática en embriones de ratón en fase de una célula. Los números indican ratones fundadores independientes generados a partir de la dosis más alta. Las flechas indican bandas escindidas por T7E1. (c) Secuencias de ADN de alelos mutantes observados en tres fundadores mutantes de *Foxn1* identificados en b. La cantidad de apariciones se muestra entre paréntesis. (d) genotipificación por PCR de progenies F1 derivadas del cruce del fundador de *Foxn1* n°. 108 y FVB/NTac natural. Téngase en cuenta la segregación de los alelos mutantes encontrados en el fundador *Foxn1* n°. 108 en las progenies.

La Fig. 6 muestra el reconocimiento del gen *Foxn1* en embriones de ratón mediante inyección intracitoplasmática de ARNm de Cas9 y *Foxn1*-ARNsg. (a) Un resultado representativo de un ensayo T7E1 que controla la tasa de mutación después de inyectar la dosis más alta. Las flechas indican bandas escindidas por T7E1. (b) Un resumen de los resultados del ensayo T7E1. Se indican las fracciones mutantes entre los embriones cultivados in vitro obtenidos después de la inyección intracitoplasmática de las dosis de RGEN indicadas. (c) Secuencias de ADN de alelos mutantes de *Foxn1* identificados a partir de un subconjunto de embriones mutantes positivos a T7E1. La secuencia diana del alelo natural se indica en recuadro.

La Fig. 7 muestra el reconocimiento del gen *Foxn1* en embriones de ratón usando la proteína Cas9 recombinante: complejo *Foxn1*-ARNsg. (a) y (b) son resultados de ensayos T7E1 representativos y sus resúmenes. Los embriones se cultivaron in vitro después de que se sometieron a una inyección pronuclear (a) o intracitoplasmática (b). Los números en rojo indican ratones fundadores mutantes T7E1-positivos. (c) Secuencias de ADN de alelos mutantes de *Foxn1* identificados a partir de los embriones cultivados in vitro que se obtuvieron mediante la inyección pronúcleo de la proteína Cas9 recombinante: complejo *Foxn1*-ARNsg en la dosis más alta. La secuencia diana del alelo natural se indica en recuadro.

La Fig. 8 muestra la transmisión en línea germinal de los alelos mutantes encontrados en el fundador del mutante *Foxn1* n°. 12. (a) análisis de fPCR. (b) genotipificación por PCR de FVB/NTac natural, el ratón fundador y sus progenies F1.

La Fig. 9 muestra genotipos de embriones generados cruzando los fundadores mutantes de *Prkdc*. Se cruzaron los mutantes *Prkdc* ♂25 y ♀15 y se aislaron los embriones E13.5. (a) análisis de fPCR natural, fundador ♂25 y fundador ♀15. Téngase en cuenta que, debido a las limitaciones técnicas del análisis de fPCR, estos resultados mostraron pequeñas diferencias de las secuencias precisas de los alelos mutantes; por ejemplo, a partir del análisis de secuencia, se identificaron  $\Delta 269/\Delta 61/WT$  y  $\Delta 5 + 1/+ 7/+ 12/WT$  en los fundadores ♂25 y ♀15, respectivamente. (b) Genotipos de los embriones generados.

La Fig. 10 muestra la mutación dirigida inducida por el complejo de proteína Cas9/ARNsg.

La Fig. 11 muestra mutaciones inducidas por la proteína Cas9 recombinante en protoplastos de *Arabidopsis*.

La Fig. 12 muestra secuencias mutantes inducidas por la proteína Cas9 recombinante en el gen *BR11* de *Arabidopsis*.

La Fig. 13 muestra el ensayo T7E1 que muestra la disrupción del gen *CCR5* endógeno en células 293 mediante el tratamiento de Cas9-mal-9R4L y el complejo ARNsg/C9R4LC.

La Fig. 14 (a, b) muestra las frecuencias de mutación en los sitios en la diana y fuera de la diana de los RGEN informados en Fu et al. (2013). Ensayos T7E1 que analizan ADN genómico de células K562 (R) transfectadas en serie con 20  $\mu$ g de plásmido que codifica Cas9 y con 60  $\mu$ g y 120  $\mu$ g en ARNcr y ARNtracr GX19 transcrito in vitro, respectivamente ( $1 \times 10^6$  células), o (D) cotransfectadas con 1  $\mu$ g de plásmido que codifica Cas9 y 1  $\mu$ g de plásmido de expresión de GX<sub>19</sub> ARNsg ( $2 \times 10^5$  células).

La Fig. 15 (a, b) muestra la comparación de la estructura del ARN guía. Las frecuencias de mutación de los RGEN informados en Fu et al. (2013) se midieron en sitios en la diana y fuera de la diana utilizando el ensayo T7E1. Las células K562 se cotransfectaron con el plásmido que codifica Cas9 y el plásmido que codifica el ARNsg de GX19 o el ARNsg de GGX20. Los sitios fuera de la diana (OT1-3, etc.) están marcados como en Fu et al. (2013).

La Fig. 16 muestra que la escisión del ADN in vitro mediante miasas Cas9. (a) Resumen esquemático de la nucleasa Cas9 y la nickasa Cas9 emparejada. Las secuencias de PAM y los sitios de escisión se muestran en recuadro. (b)

Sitios diana en el locus humano AAVS1. La posición de cada sitio diana se muestra en triángulo. (c) Visión general esquemática de las reacciones de escisión del ADN. Los colorantes FAM (que se muestran en la caja) se unieron a ambos extremos 5' del sustrato de ADN. (d) DSB y SSB analizados usando fluorescencia de electroforesis capilar. Los sustratos de ADN marcados fluorescentemente se incubaron con nucleasas Cas9 o nickasas antes de la electroforesis.

5 La Fig. 17 muestra la comparación de la nucleasa Cas9 y el comportamiento de la nickasa. (a) Frecuencias de mutación en la diana asociadas con nucleasas Cas9 (WT), nickasas (D10A) y nickasas emparejadas. Se indican nickasas emparejadas que producirían extremos 5' o extremos 3'. (b) Análisis de los efectos fuera de diana de las nucleasas Cas9 y las nickasas emparejadas. Se analizaron un total de siete posibles sitios fuera de la diana para tres ARNsg.

10 La Fig. 18 muestra las nickasas Cas9 emparejadas probadas en otros loci humanos endógenos. (a,c) Los sitios diana de ARNsg en los loci CCR5 y BRCA2 humanos. Las secuencias de PAM están indicadas en rojo. (b,d) Las actividades de edición del genoma en cada sitio diana se detectaron mediante el ensayo T7E1. La reparación de dos mellas que producirían extremos 5' condujo a la formación de inserciones-delecciones mucho más frecuentemente que aquellos que producían extremos de 3'.

15 La Fig. 19 muestra que las nickasas Cas9 emparejadas intervienen en la recombinación homóloga. (a) Estrategia para detectar la recombinación homóloga. El ADN del donante incluía un sitio de la enzima de restricción XbaI entre dos brazos de homología, mientras que el sitio diana endógeno carecía de este sitio. Se usó un ensayo de PCR para detectar secuencias que se habían sometido a recombinación homóloga. Para evitar la amplificación del ADN del donante contaminante, se usaron cebadores específicos para el ADN genómico. (b) Eficiencia de la recombinación homóloga. Solo los amplicones de una región en la que se había producido recombinación homóloga podían digerirse con XbaI; las intensidades de las bandas de división se usaron para medir la eficacia de este método.

20 La Fig. 20 muestra el corte y empalme de ADN inducido por las nickasas de Cas9 emparejadas. (a) Los sitios diana de las nickasas emparejadas en el locus AAVS1 humano. Se muestran las distancias entre el sitio AS2 y cada uno de los otros sitios. Las flechas indican cebadores de PCR. (b) Eliminaciones genómicas detectadas mediante PCR. Los asteriscos indican productos de PCR específicos de eliminación. (c) Secuencias de ADN de productos de PCR específicos de delección obtenidos utilizando ARNsg de AS2 y L1. Las secuencias de PAM del sitio diana se muestran en recuadro y las secuencias de coincidencia de ARNsg se muestran en letras mayúsculas. Las secuencias de emparejamiento de ARNsg intactas están subrayadas. (d) Un modelo esquemático de delecciones cromosómicas mediadas por la nickasa Cas9 emparejadas. Las cadenas de ADN recién sintetizadas se muestran en recuadro.

25 La figura 21 muestra que las nickasas de Cas9 emparejadas no inducen translocaciones. (a) Resumen esquemático de las translocaciones cromosómicas entre los sitios en la diana y fuera de la diana. (b) Amplificación por PCR para detectar translocaciones cromosómicas. (c) Translocaciones inducidas por nucleasas Cas9 pero no por el par de la nickasa.

30 La Fig. 22 muestra un diagrama conceptual de los ensayos T7E1 y RFLP. (a) Comparación de las reacciones de escisión del ensayo en cuatro escenarios posibles después del tratamiento con nucleasa modificada en una célula diploide: (A) natural, (B) una mutación monoalélica, (C) diferentes mutaciones bialélicas (hetero) y (D) idénticas mutaciones bialélicas (homo). Las líneas negras representan productos de PCR derivados de cada alelo; los recuadros con línea discontinua y punteados indican mutaciones de inserción/eliminación generadas por NHEJ. (b) Resultados esperados de la digestión de T7E1 y RGEN resueltos por electroforesis.

35 La Fig. 23 muestra el ensayo de eliminación in vitro de un plásmido linealizado que contiene el sitio diana C4BPB que lleva inserciones-delecciones. Secuencias de ADN de sustratos de plásmidos individuales (panel superior). La secuencia PAM está subrayada. Las bases insertadas se muestran en la caja. Las flechas (panel inferior) indican las posiciones esperadas de las bandas de ADN escindidas por el RGEN natural después de la electroforesis.

40 La Fig. 24 muestra la genotipificación de mutaciones inducidas por nucleasas modificadas genéticamente en células mediante RFLP mediado por RGEN. (a) Genotipo de clones de células K562 mutantes de C4BPB. (b) Comparación del ensayo de T7E1 sensible al desapareamiento con análisis de RFLP mediado por RGEN. Las flechas negras indican el producto de escisión por tratamiento de la enzima T7E1 o RGEN.

45 La Fig. 25 muestra la genotipificación de mutaciones inducidas por RGEN mediante la técnica RGEN-RFLP. (a) Análisis de clones modificados con C4BPB usando ensayos RGEN-RFLP y T7E1. Las flechas indican las posiciones esperadas de las bandas de ADN escindidas por RGEN o T7E1. (b) Comparación cuantitativa del análisis de RGEN-RFLP con ensayos de T7E1. Las muestras de ADN genómico de células K562 de tipo natural y modificadas con C4BPB se mezclaron en diversas proporciones y se sometieron a amplificación por PCR. (c) Genotipificación de mutaciones inducidas por RGEN en el gen HLA-B en células HeLa con análisis RFLP y T7E1.

50 La Fig. 26 muestra la genotipificación de mutaciones inducidas por nucleasas modificadas genéticamente en organismos a través de RFLP mediado por RGEN. (a) Genotipo de ratones fundadores del mutante Pibf1. (b)

Comparación del ensayo de T7E1 sensible al desapareamiento con análisis de RFLP mediado por RGEN. Las flechas negras indican el producto de escisión por el tratamiento de la enzima T7E1 o RGEN.

La Fig. 27 muestra la genotipificación mediada por RGEN de mutaciones inducidas por ZFN. El sitio diana de ZFN se muestra en el recuadro. Las flechas negras indican bandas de ADN escindidas por T7E1.

5 La Fig. 28 muestra sitios polimórficos en una región del gen HLA-B humano. La secuencia, que rodea el sitio diana de RGEN, es la de un amplicón de PCR de células HeLa. Las posiciones polimórficas se muestran en el recuadro. El sitio diana RGEN y la secuencia PAM se muestran en cuadro con líneas discontinuas y en negrita, respectivamente. Las secuencias del cebador están subrayadas.

10 La Fig. 29 muestra la genotipificación de mutaciones oncogénicas mediante análisis de RGEN-RFLP. (a) Una mutación recurrente (deleción c.133-135 de TCT) en el gen CTNNB1 humano en células HCT116 fue detectada por RGEN. Las células HeLa se usaron como control negativo. (b) Genotipificación de la mutación de sustitución de KRAS (c.34 G>A) en la línea celular de cáncer A549 con RGEN que contienen ARN guía no coincidentes. Los nucleótidos mal emparejados se muestran en el recuadro. Las células HeLa se usaron como control negativo. Las flechas indican bandas de ADN escindidas por RGEN. Se muestran secuencias de ADN confirmadas por  
15 secuenciación de Sanger.

La Fig. 30 muestra la genotipificación del alelo CCR5 delta32 en células HEK293T mediante análisis de RGEN-RFLP. (a) ensayos RGEN-RFLP de líneas celulares. Las células K562, SKBR3 y HeLa se usaron como controles naturales. Las flechas indican bandas de ADN escindidas por RGEN. (b) secuencia de ADN de alelos CCR5 de tipo natural y delta32. Ambos sitios de RGEN en la diana y fuera de la diana utilizados en el análisis de RFLP están  
20 subrayados. Un desapareamiento de nucleótido único entre los dos sitios se muestra en el recuadro. La secuencia PAM está subrayada. (c) Eliminación in vitro de plásmidos que albergan alelos WT o del32 CCR5 usando el RGEN natural. (d) Confirmación de la presencia de un sitio fuera de la diana del RGEN específico de CCR5-delta32 en el locus CCR5. Ensayos de eliminación in vitro de plásmidos que albergan secuencias en la diana o fuera de la diana usando diversas cantidades del RGEN específico del del32.

25 La Fig. 31 muestra la genotipificación de una mutación puntual de KRAS (c.34 G>A). (a) Análisis RGEN-RFLP de la mutación KRAS (c.34 G>A) en líneas celulares de cáncer. Los productos de PCR de células HeLa (utilizadas como control natural) o células A549, que son homocigóticas para la mutación puntual, se digirieron con RGEN con un ARNcr perfectamente emparejado específico para la secuencia natural o la secuencia mutante. Los genotipos de KRAS en estas células se confirmaron mediante la secuenciación de Sanger. (b) Los plásmidos que albergan las  
30 secuencias de KRAS naturales o mutante se digirieron usando RGENs con ARNcr perfectamente emparejados o ARNcrs emparejados por una sola base y atenuados. Los ARNcr atenuados que se eligieron para el genotipado se marcan en un recuadro encima de los geles.

La Fig. 32 muestra la genotipificación de una mutación puntual PIK3CA (c.3140 A>G). (a) Análisis de RGEN-RFLP de la mutación PIK3CA (c.3140 A>G) en líneas celulares de cáncer. Los productos de PCR de células HeLa (utilizadas como control natural) o células HCT116 que son heterocigóticas para la mutación puntual se digirieron  
35 con RGEN con un ARNcr perfectamente emparejado específico para la secuencia natural o la secuencia mutante. Los genotipos PIK3CA en estas células se confirmaron mediante la secuenciación de Sanger. (b) Los plásmidos que albergan las secuencias de PIK3CA natural o mutante se digirieron usando RGENs con ARNcrs perfectamente emparejados o ARNcrs mal emparejados de una sola base o atenuados. Los ARNcr atenuados que se eligieron para el genotipado se marcan en un recuadro encima de los geles.  
40

La Fig. 33 muestra la genotipificación de mutaciones puntuales recurrentes en líneas celulares cancerosas. (a) ensayos RGEN-RFLP de mutaciones puntuales oncogénicas recurrentes en IDH (c.394c> T), (b) PIK3CA (c.3140A> G), (c) NRAS (c.181C> A), (d) y Genes BRAF (c.1799T> A). Se muestran los genotipos de cada línea celular confirmada por secuenciación de Sanger. Los nucleótidos mal emparejados se muestran en el recuadro. Las flechas  
45 negras indican bandas de ADN escindidas por RGEN.

### Descripción detallada

En un aspecto, la presente invención proporciona un sistema o una composición para escindir el ADN diana en células eucariotas u organismos, como se define en las reivindicaciones. Además, la presente invención proporciona un uso del sistema o de la composición para escindir el ADN diana en células eucariotas u organismos.

50 En la presente invención, el sistema o la composición también se denomina composición de endonucleasa guiada por ARN (RGEN).

ZFNs y TALENs permiten la mutagénesis dirigida en células de mamíferos, organismos modelo, plantas y ganado, pero las frecuencias de mutación obtenidas con nucleasas individuales son muy diferentes entre sí. Además, algunos ZFN y TALEN no muestran las actividades de edición del genoma. La metilación del ADN puede limitar la  
55 unión de estas nucleasas modificadas a los sitios diana. Además, son técnicamente complicadas y lleva mucho tiempo fabricar nucleasas personalizadas.

Los presentes inventores han desarrollado una nueva composición de endonucleasa guiada por ARN basada en la proteína Cas9 para superar las desventajas de ZFN y TALEN.

5 Antes de la presente invención, se conocía una actividad endonucleasa de las proteínas Cas. Sin embargo, no se sabe si la actividad endonucleasa de la proteína Cas funcionaría en una célula eucariota debido a la complejidad del genoma eucariota. Además, hasta ahora, no se ha desarrollado ninguna composición que comprenda la proteína Cas o el ácido nucleico que codifique la proteína Cas y un ARN guía específico para que el ADN diana escinda un ADN diana en células u organismos eucariotas.

10 En comparación con los ZFN y TALEN, la presente composición de RGEN basada en la proteína Cas9 se puede personalizar más fácilmente porque solo se reemplaza el componente de ARN guía sintético para formar una nueva nucleasa de edición del genoma. No se involucran pasos de subclonación para hacer endonucleasas guiadas por ARN personalizadas. Además, el tamaño relativamente pequeño del gen Cas (por ejemplo, 4,2 kbp para Cas9) en comparación con un par de genes TALEN (~ 6 kbp) proporciona una ventaja para esta composición de endonucleasa guiada por ARN en algunas aplicaciones tales como la mediada por la liberación de genes mediada por virus. Además, esta endonucleasa guiada por ARN no tiene efectos fuera de la diana y, por lo tanto, no induce mutaciones, deleciones, inversiones y duplicaciones indeseadas. Estas características hacen que la presente composición de endonucleasa guiada por ARN sea una herramienta escalable, versátil y conveniente para la ingeniería del genoma en células y organismos eucarióticos. Además, RGEN puede diseñarse para reconocer cualquier secuencia de ADN, casi cualquier polimorfismo de nucleótido simple o inserción/delección (indel) pequeña puede analizarse a través de RFLP mediado por RGEN. La especificidad de los RGEN está determinada por el componente de ARN que se hibrida con una secuencia de ADN diana de hasta 20 pares de bases (pb) de longitud y por la proteína Cas9 que reconoce el motivo adyacente de protoespaciador (PAM). Los RGEN se reprograman fácilmente reemplazando el componente de ARN. Por lo tanto, los RGEN proporcionan una plataforma para usar análisis de RFLP simples y robustos para diversas variaciones de secuencia.

El ADN diana puede ser un ADN endógeno o ADN artificial, preferiblemente, ADN endógeno.

25 Como se usa en el presente documento, la expresión "proteína Cas9" se refiere a un componente proteico esencial en el sistema CRISPR/Cas9, forma una endonucleasa o una nickasa activa cuando se complejan con dos ARN denominados ARN CRISPR (ARNcr) y ARNcr transactivo (ARNtracr).

La información sobre el gen y la proteína de Cas está disponible en GenBank del Centro Nacional de Información Biotecnológica (NCBI), sin limitación.

30 Los genes (cas) asociados a CRISPR que codifican las proteínas Cas a menudo se asocian con matrices repetidoras espaciadoras CRISPR. Se han descrito más de cuarenta familias diferentes de proteínas Cas. De estas familias de proteínas, Cas1 parece ser omnipresente entre diferentes sistemas CRISPR/Cas. Hay tres tipos de sistema CRISPR-Cas. Entre ellos, el sistema Tipo II CRISPR/Cas que implica la proteína Cas9 y ARNcr y ARNtracr es representativo y es bien conocido. Se han usado combinaciones particulares de genes cas y estructuras repetidas para definir 8 subtipos CRISPR (Ecoli, Ypest, Nmeni, Dvulg, Tneap, Hmari, Apern y Mtube).

La proteína Cas puede estar unida a un dominio de transducción de proteínas. El dominio de transducción de proteínas puede ser poli-arginina o una proteína TAT derivada de VIH, pero no está limitada a esto.

La presente composición puede comprender el componente Cas9 en forma de una proteína o en forma de un ácido nucleico que codifica la proteína Cas9.

40 En la presente invención, la proteína Cas9 puede ser cualquier proteína Cas9 con la condición de que tenga una actividad endonucleasa o nickasa cuando forma un complejo con un ARN guía.

La expresión proteína Cas9 puede incluir variantes de la misma.

45 La variante de la proteína Cas9 puede ser una forma mutante de Cas9 en la que el residuo de aspartato catalítico se cambia a cualquier otro aminoácido. Preferiblemente, el otro aminoácido puede ser una alanina, pero no está limitado a esto.

Además, la proteína Cas9 puede ser la aislada de un organismo tal como *Streptococcus* sp., preferiblemente *Streptococcus pyogenes* o una proteína recombinante, pero no se limita a esto.

La proteína Cas9 derivada de *Streptococcus pyogenes* puede reconocer el trinucleótido NGG. La proteína Cas9 puede comprender una secuencia de aminoácidos de la SEQ ID NO: 109, pero no está limitada a la misma.

50 El término "recombinante", cuando se utiliza con referencia, por ejemplo, a una célula, ácido nucleico, proteína o vector, indica que la célula, el ácido nucleico, la proteína o el vector se han modificado mediante la introducción de un ácido nucleico o proteína heteróloga o la alteración de un ácido nucleico o proteína nativa, o que la célula se deriva de una célula así modificada. De este modo, por ejemplo, se puede generar una proteína Cas9 recombinante reconstituyendo la secuencia que codifica la proteína Cas9 usando la tabla de codones humanos.

- En cuanto a la presente invención, el ácido nucleico que codifica la proteína Cas9 puede ser una forma de vector, tal como un plásmido que comprende la secuencia que codifica Cas9 bajo un promotor tal como CMV o CAG. La secuencia que codifica Cas9 se puede derivar de *Streptococcus sp.*, y, preferiblemente, se deriva de *Streptococcus pyogenes*. Por ejemplo, el ácido nucleico que codifica Cas9 puede comprender la secuencia de nucleótidos de la SEQ ID NO: 1. Además, el ácido nucleico que codifica Cas9 puede comprender la secuencia de nucleótidos que tiene una homología de al menos 50% con la secuencia de SEQ ID NO: 1, preferiblemente al menos 60, 70, 80, 90, 95, 97, 98 o 99% con la SEQ ID NO: 1, pero no está limitado a esto. El ácido nucleico que codifica Cas9 puede comprender la secuencia de nucleótidos de las SEQ ID NOs. 108, 110, 106 ó 107.
- Como se usa en el presente documento, la expresión "ARN guía" se refiere a un ARN que es específico para el ADN diana y puede formar un complejo con la proteína Cas9 y llevar la proteína Cas9 al ADN diana.
- Como se describe en este documento, el ARN guía puede consistir en dos moléculas de ARN, es decir, ARN CRISPR (ARNcr) y ARNcr transactivador (ARNtracr). En la presente invención, el ARN guía es un ARN monocatenario (ARNsg) producido por fusión de una porción esencial de ARNcr y ARNtracr.
- El ARN guía puede ser un ARN doble que comprenda un ARNc y un ARNtracr.
- Si el ARN guía comprende la porción esencial de ARNcr y ARNtracr y una porción complementaria de una diana, se puede usar cualquier ARN guía en la presente invención.
- El ARNcr se puede hibridar con un DNA diana.
- El RGEN puede consistir en la proteína Cas9 y ARNsg (fusión de una porción esencial del ARNtracr invariable y ARNcr específico de la diana), y puede reprogramarse fácilmente reemplazando el ARNcr.
- El ARN guía comprende además uno o más nucleótidos adicionales en el extremo 5' del ARN guía monocatenario o el ARNcr del ARN dual.
- Preferiblemente, el ARN guía comprende adicionalmente 2 nucleótidos de guanina adicionales en el extremo 5' del ARN guía monocatenario o el ARNcr del ARN dual.
- El ARN guía puede transferirse a una célula o a un organismo en forma de ARN o ADN que codifica el ARN guía. El ARN guía puede estar en forma de un ARN aislado, ARN incorporado en un vector viral, o está codificado en un vector. Preferiblemente, el vector puede ser un vector viral, un vector plasmídico o un vector agrobacteriano, pero no está limitado a este.
- Un ADN que codifica el ARN guía puede ser un vector que comprenda una secuencia que codifique el ARN guía. Por ejemplo, el ARN guía puede transferirse a una célula u organismo transfecando la célula u organismo con el ARN guía o el ADN plasmídico aislado que comprende una secuencia que codifica el ARN guía y un promotor.
- Alternativamente, el ARN guía puede transferirse a una célula u organismo usando la administración génica mediada por virus.
- Cuando el ARN guía se transfecta en forma de un ARN aislado en una célula u organismo, el ARN guía se puede preparar mediante transcripción *in vitro* usando cualquier sistema de transcripción *in vitro* conocido en la técnica. El ARN guía preferiblemente se transfiere a una célula en forma de ARN aislado en lugar de en forma del plásmido que comprende una secuencia codificante para un ARN guía. Como se usa en el presente documento, la expresión "ARN aislado" puede ser intercambiable con "ARN desnudo". Esto ahorra costos y tiempo porque no requiere un paso de clonación. Sin embargo, no se excluye el uso de ADN plasmídico o liberación de genes mediada por virus para la transfección del ARN guía.
- La presente composición de RGEN que comprende proteína Cas9 o ácido nucleico que codifica la proteína Cas9 y un ARN guía puede escindir específicamente un ADN diana debido a una especificidad del ARN guía por una diana y una actividad de endonucleasa o nickasa de la proteína Cas9.
- Como se usa en este documento, el término "escisión" se refiere a la rotura de la cadena principal covalente de una molécula de nucleótido.
- En la presente invención, se puede preparar un ARN guía para que sea específico para cualquier diana que deba escindirse. Por lo tanto, la presente composición de RGEN puede escindir cualquier ADN diana manipulando o genotipificando la porción específica de la diana del ARN guía.
- El ARN guía y la proteína Cas9 pueden funcionar como un par. Tal como se usa en la presente memoria, la expresión "Cas9 nickasa emparejada" puede referirse al ARN guía y a la proteína Cas9 que funciona como un par. El par comprende dos ARN guía. El ARN guía y la proteína Cas9 pueden funcionar como un par e inducir dos mellas en diferentes cadenas de ADN. Las dos mellas pueden estar separadas por al menos 100 bps, pero no están limitadas a esto.

- En los Ejemplos, los presentes inventores confirmaron que la Cas9 nickasa emparejada permite la mutagénesis dirigida y grandes deleciones de segmentos cromosómicos de hasta 1 kbp en células humanas. De manera importante, las nickasas emparejadas no indujeron inserciones-deleciones en sitios fuera de la diana en los que sus correspondientes nucleasas inducen mutaciones. Además, a diferencia de las nucleasas, las nickasas apareadas no promocionaron translocaciones indeseadas asociadas con las divisiones de ADN fuera de la diana. En principio, las nickasas emparejadas duplican la especificidad de la mutagénesis mediada por Cas9 y ampliarán la utilidad de las enzimas guiadas por ARN en aplicaciones que requieren la edición precisa del genoma, como la terapia génica y celular.
- En la presente invención, la composición puede usarse en la genotipificación de un genoma en células u organismos eucarióticos *in vitro*.
- En una realización específica, el ARN guía puede comprender la secuencia de nucleótidos de Seq ID. 1, en donde la porción de la posición de nucleótido 3 ~ 22 es una porción específica de la diana y, por lo tanto, la secuencia de esta porción puede cambiarse dependiendo de un diana.
- Como se usa en el presente documento, una célula u organismo eucariótico puede ser levadura, hongo, protozoos, planta, planta superior, e insectos, o células anfibias, o células de mamíferos tales como CHO, HeLa, HEK293 y COS-1, por ejemplo, células cultivadas (*in vitro*), células de injerto y cultivo de células primarias (*in vitro* y *ex vivo*), y células *in vivo*, y también células de mamífero incluyendo humanos, que se usan comúnmente en la técnica, sin limitación.
- En una realización específica, se descubrió que la proteína Cas9/ARN guía monocatenario podía generar cortes de doble cadena de ADN específicos del sitio *in vitro* y en células de mamífero, cuya reparación espontánea indujo mutaciones del genoma dirigido a altas frecuencias.
- Además, se descubrió que podían inducirse ratones con genes silenciados mediante la inyección de complejos de proteína Cas9/ARN guía o ARNm de Cas9/ARN guía en embriones en estadio de una célula y que podían ser generadas mutaciones transmisibles de la línea germinal por el sistema de Cas9/ARN guía.
- El uso de la proteína Cas9 en lugar de un ácido nucleico que codifica la proteína Cas9 para inducir una mutagénesis dirigida es ventajoso porque el ADN exógeno no se introduce en un organismo. Por lo tanto, la composición que comprende la proteína Cas9 y un ARN guía puede usarse para desarrollar productos terapéuticos o cultivos de valor añadido, ganado, aves de corral, peces, mascotas, etc.
- En otro aspecto, la presente descripción proporciona una composición para inducir mutagénesis dirigida en células u organismos eucarióticos. Además, la presente composición puede usarse para inducir mutagénesis dirigida en células u organismos eucarióticos.
- Un ARN guía, un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 son como se describe en lo anterior.
- El sistema de la invención puede proporcionarse en un kit para escindir un ADN diana o inducir mutagénesis dirigida en células u organismos eucarióticos.
- Un ARN guía, un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 son como se describe en lo anterior.
- El kit puede comprender un ARN guía y un ácido nucleico que codifica una proteína Cas9 o proteína Cas9 como componentes separados o como una composición.
- El presente kit puede comprender algunos componentes adicionales necesarios para transferir el ARN guía y el componente Cas9 a una célula o un organismo. Por ejemplo, el kit puede comprender un tampón de inyección tal como un tampón de inyección tratado con DEPC, y materiales necesarios para el análisis de la mutación de un ADN diana, pero no están limitados a los mismos.
- El sistema de la invención puede usarse en un método para preparar una célula u organismo eucariótico que comprende la proteína Cas9 y un ARN guía que comprende una etapa de cotransfección o transfección en serie de la célula u organismo eucariótico con un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9, y un ARN guía o ADN que codifica el ARN guía.
- Un ARN guía, un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 son como se describe en lo anterior.
- En la presente invención, un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 y un ARN o ADN guía que codifica el ARN guía pueden transferirse a una célula mediante diversos métodos conocidos en la técnica, tales como microinyección, electroporación, tratamiento con DEAE-dextrano, lipofección, transfección mediada por nanopartículas, transducción mediada por dominio de transducción de proteínas, administración génica mediada por virus, y transfección mediada por PEG en protoplastos, y similares, pero no se limitan a ellos. Además, una proteína

- 5 Cas9 que codifica un ácido nucleico o proteína Cas9 y un ARN guía pueden transferirse a un organismo por diversos métodos conocidos en la técnica para administrar un gen o una proteína tal como una inyección. Un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 se pueden transferir a una célula en forma de complejo con un ARN guía, o por separado. La proteína Cas9 fusionada a un dominio de transducción de proteínas tal como Tat también se puede administrar de manera eficiente en las células.
- Preferiblemente, la célula u organismos eucarióticos se cotransfectan o transfectan en serie con una proteína Cas9 y un ARN guía de la invención.
- 10 La transfección en serie se puede realizar por transfección con ácido nucleico que codifica la proteína Cas9 primero, seguido de una segunda transfección con ARN guía desnudo. Preferiblemente, la segunda transfección es después de 3, 6, 12, 18, 24 horas, pero no está limitada a esto.
- El sistema de la invención puede comprender una célula u organismo eucariótico.
- Las células u organismos eucariotas pueden prepararse transfiriendo la composición de la invención en la célula u organismo.
- 15 Las células eucariotas pueden ser levaduras, hongos, protozoos, plantas superiores, e insectos, o células anfibias, o células de mamíferos tales como CHO, HeLa, HEK293 y COS-1, por ejemplo, células cultivadas (in vitro), células de injerto y cultivo de células primarias (in vitro y ex vivo), y células in vivo, y también células de mamífero incluyendo humanas, que se usan comúnmente en la técnica, sin limitación. Además, el organismo puede ser una levadura, un hongo, protozoos, una planta, una planta superior, un insecto, anfibio o mamífero.
- 20 En otro aspecto, la presente invención proporciona un método in vitro para escindir un ADN diana o inducir mutagénesis dirigida en células u organismos eucarióticos, que comprende una etapa de tratar una célula u organismo que comprende un ADN diana con una composición de la invención.
- La etapa de tratar una célula u organismo con la composición puede realizarse transfiriendo la presente composición que comprende un ARN guía específico para el ADN o ADN diana que codifica el ARN guía, y un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 en la célula u organismo.
- 25 Como se describe en lo anterior, dicha transferencia puede realizarse por microinyección, transfección, electroporación, etc.
- El sistema de la invención puede usarse para proporcionar un embrión que comprenda un genoma editado por la presente composición de RGEN.
- 30 Se puede usar cualquier embrión no humano; el embrión puede ser un embrión de un ratón. El embrión puede producirse inyectando PMSG (gonadotropina de yegua gestante preñada) y hCG (gonadotropina coriónica humana) en un ratón hembra de 4 a 7 semanas y el ratón hembra superovulado puede acoplarse a machos, y los embriones fertilizados pueden recogerse de oviductos.
- La presente composición de RGEN introducida en un embrión puede escindir un ADN diana complementario al ARN guía mediante la acción de la proteína Cas9 y causar una mutación en el ADN diana. Por lo tanto, el embrión en el que se ha introducido la presente composición de RGEN tiene un genoma editado.
- 35 Se descubrió que la presente composición de RGEN podría causar una mutación en un embrión de ratón y que la mutación podía transmitirse a las crías.
- Un método para introducir la composición de RGEN en el embrión puede ser cualquier método conocido en la técnica, tal como microinyección, inserción de células madre, inserción de retrovirus, y así sucesivamente. Preferiblemente, se puede usar una técnica de microinyección.
- 40 El sistema de la invención puede usarse para proporcionar un animal genéticamente modificado obtenido transfiriendo el embrión que comprende un genoma editado por la presente composición de RGEN a los oviductos de un animal.
- 45 La expresión "animal modificado con genoma" se refiere a un animal cuyo genoma ha sido modificado en la etapa de embrión por la presente composición de RGEN y el tipo de animal no está limitado; sin embargo, la expresión no abarca a seres humanos.
- El animal modificado con genoma tiene mutaciones causadas por una mutagénesis dirigida basada en la presente composición de RGEN. Las mutaciones pueden ser cualquiera de delección, inserción, translocación, inversión. El sitio de la mutación depende de la secuencia del ARN guía de la composición de RGEN.
- 50 El animal modificado con genoma que tiene una mutación de un gen puede usarse para determinar la función del gen.

La composición RGEN de la invención puede usarse en un método para preparar un animal genéticamente modificado que comprende una etapa de introducción de la presente composición de RGEN en un embrión de un animal; y un paso de transferir el embrión a un oviducto de madre adoptiva pseudopreñada para producir un animal genéticamente modificado.

- 5 La etapa de introducir la presente composición de RGEN puede realizarse por cualquier método conocido en la técnica, tal como microinyección, inserción de células madre, inserción retroviral, y así sucesivamente.

La composición RGEN de la invención puede usarse para producir una planta regenerada a partir de protoplastos modificados con genoma preparados mediante el método para células eucariotas que comprende la composición de RGEN.

- 10 La composición de la invención puede usarse para genotipar mutaciones o variaciones en una muestra biológica aislada. La composición de la invención también puede usarse para genotipar secuencias de ácidos nucleicos en microorganismos patógenos en una muestra biológica aislada.

Un ARN guía, un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 son como se describe en lo anterior.

- 15 Como se usa en el presente documento, el término "genotipado" se refiere al "ensayo de polimorfismo de longitud de fragmentos de restricción (RFLP)".

- 20 El RFLP puede usarse en 1) la detección de inserciones-deleciones en células u organismos inducidos por las nucleasas modificadas genéticamente, 2) las mutaciones genotípicas que ocurren naturalmente o las variaciones en células u organismos, o 3) el genotipado del ADN de microorganismos patógenos infectados, incluyendo virus o bacterias, etc.

Las mutaciones o variaciones pueden ser inducidas por nucleasas modificadas en las células.

La nucleasa modificada por ingeniería genética puede ser una nucleasa de dedo de cinc (ZFN), nucleasas efectoras de tipo transcriptor de activador de la transcripción (TALEN) o RGEN, pero no está limitada a la misma.

- 25 Como se usa en este documento, la expresión "muestra biológica" incluye muestras para análisis, tales como tejidos, células, sangre completa, suero, plasma, saliva, esputo, fluido cerebroespinal u orina, pero no se limita a estos.

Las mutaciones o variaciones pueden ser mutaciones o variaciones naturales.

Las mutaciones o variaciones son inducidas por los microorganismos patógenos. A saber, las mutaciones o variaciones ocurren debido a la infección de microorganismos patógenos, cuando se detectan los microorganismos patógenos, la muestra biológica se identifica como infectada.

- 30 Los microorganismos patógenos pueden ser virus o bacterias, pero no están limitados a ellos.

- 35 Las mutaciones inducidas por nucleasas modificadas se detectan mediante diversos métodos, que incluyen ensayos de Surveyor o T7 endonucleasa I (T7E1) sensibles al desapareamiento, análisis de RFLP, PCR fluorescente, análisis de fusión de ADN y secuenciación profunda. Los ensayos T7E1 y Surveyor son ampliamente utilizados, pero a menudo subestiman las frecuencias de mutación porque los ensayos detectan hetero dúplex (formados por la hibridación de secuencias mutantes y naturales o dos secuencias mutantes diferentes); no pueden detectar homodúplex formados por la hibridación de dos secuencias mutantes idénticas. Por lo tanto, estos ensayos no pueden distinguir clones mutantes bialélicos homocigóticos de células naturales ni mutantes bialélicos heterocigotos de mutantes monogénicos heterocigotos (Fig. 22). Además, los polimorfismos de secuencia cerca del sitio diana de nucleasa pueden producir resultados confusos porque las enzimas pueden escindir hetero dúplex formados por la hibridación de estos alelos naturales diferentes. El análisis de RFLP está libre de estas limitaciones y, por lo tanto, es un método de elección. De hecho, el análisis de RFLP fue uno de los primeros métodos utilizados para detectar mutaciones mediadas por nucleasa modificadas genéticamente. Desafortunadamente, sin embargo, está limitado por la disponibilidad de sitios de restricción apropiados.
- 40

- 45 El sistema de la invención puede proporcionarse en un kit para genotipar mutaciones o variaciones en una muestra biológica aislada. Además, el sistema de la invención puede proporcionarse en un kit para genotipar secuencias de ácidos nucleicos en microorganismos patógenos en una muestra biológica aislada.

Un ARN guía, un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 son como se describe en lo anterior.

- 50 El sistema de la invención puede usarse en un método para genotipar mutaciones o variaciones en una muestra biológica aislada. Además, el sistema de la invención puede usarse en un método para genotipar secuencias de ácidos nucleicos en microorganismos patógenos en una muestra biológica aislada.

Un ARN guía, un ácido nucleico que codifica la proteína Cas9 o la proteína Cas9 son como se describe en lo anterior.

## Ejemplos

5 En lo sucesivo en este documento, la presente invención se describirá con más detalle con referencia a los Ejemplos. Sin embargo, estos ejemplos son solo para fines ilustrativos, y la invención no pretende estar limitada por estos ejemplos.

Ejemplo 1: Ensayo de edición de genoma

### 1-1. Actividad de escisión del ADN de la proteína Cas9

10 En primer lugar, se probó la actividad de escisión del ADN de Cas9 derivada de *Streptococcus pyogenes* en presencia o ausencia de un ARN guía quimérico in vitro.

15 Para este fin, se usó la proteína Cas9 recombinante que se expresó y se purificó a partir de *E. coli* para escindir un ADN de plásmido predigerido o circular que contenía la secuencia diana de CCR5 humano de 23 pares de bases (pb). Una secuencia diana Cas9 consiste en una secuencia de ADN de 20 pb complementaria al ARNcr o un ARN guía quimérico y el motivo adyacente de protoespaciador (PAM) del trinucleótido (5'-NGG-3') reconocido por la propia Cas9 (Fig. 1A).

20 Específicamente, la secuencia que codifica Cas9 (4.104 pb), derivada de la cepa M1 GAS de *Streptococcus pyogenes* (NC\_002737.1), se reconstituyó usando la tabla de uso de codones humanos y se sintetizó usando oligonucleótidos. En primer lugar, se ensamblaron segmentos de ADN de 1 kb usando oligonucleótidos solapantes de ~ 35 mers y polimerasa Phusion (New England Biolabs) y se clonaron en el vector T (SolGent). Se ensambló una secuencia de Cas9 de longitud completa usando cuatro segmentos de ADN de 1 kbp mediante PCR solapada. El segmento de ADN que codifica Cas9 se subclonó en p3s, que se derivó de pcDNA3.1 (Invitrogen). En este vector, se añadió un marcador de péptido (NH<sub>2</sub>-GGSGPPKKRKYYPYDVPDYA-COOH, SEQ ID NO: 2) que contiene el epítipo HA y una señal de localización nuclear (NLS) al extremo C-terminal de Cas9. La expresión y la localización nuclear de la proteína Cas9 en células HEK 293T se confirmaron mediante transferencia Western usando anticuerpo anti-HA (Santa Cruz).

30 Luego, el casete Cas9 se subclonó en pET28-b (+) y se transformó en BL21 (DE3). La expresión de Cas9 se indujo usando IPTG 0.5 mM durante 4 horas a 25°C. La proteína Cas9 que contiene el marcador His6 en el extremo C se purificó usando resina de agarosa Ni-NTA (Qiagen) y se dializó frente a HEPES 20 mM (pH 7,5), KCl 150 mM, DTT 1 mM y glicerol al 10% (1). Se incubó Cas9 purificado (50 nM) con ADN plásmido súper enrollado o predigerido (300 ng) y ARN quimérico (50 nM) en un volumen de reacción de 20 µl en tampón NEB 3 durante 1 hora a 37°C. El ADN digerido se analizó por electroforesis usando geles de agarosa al 0,8%.

Cas9 escindió el ADN del plásmido eficientemente en la posición esperada solo en presencia del ARN sintético y no escindió un plásmido de control que carecía de la secuencia diana (Fig. 1B).

### 1 -2. Escisión de ADN por Cas9/complejo de ARN guía en células humanas

35 Se usó un informador RFP-GFP para investigar si el complejo Cas9/RNA guía podía escindir la secuencia diana incorporada entre las secuencias RFP y GFP en células de mamífero.

40 En este informador, la secuencia GFP se fusiona con la secuencia RFP fuera de marco (2). La GFP activa se expresa solo cuando la secuencia diana se escinde mediante nucleasas específicas del sitio, lo que causa pequeñas inserciones o inserciones (inserciones-delecciones) en el marco de la secuencia diana mediante la reparación de la recombinación no homóloga o unión de extremos no homólogos (NHEJ) de la rotura del ADN bicatenario (DSB) (Fig 2).

Los plásmidos informadores RFP-GFP utilizados en este estudio se construyeron tal como se describió anteriormente (2). Los oligonucleótidos correspondientes a los sitios diana (Tabla 1) se sintetizaron (Macrogen) y se hibridaron. Los oligonucleótidos hibridados se ligaron en un vector informador digerido con EcoRI y BamHI.

45 Las células HEK 293T se cotransfectaron con un plásmido que codificaba Cas9 (0,8 µg) y el plásmido informador RFP-GFP (0,2 µg) en una placa de 24 pocillos utilizando Lipofectamina 2000 (Invitrogen).

50 Mientras tanto, el ARN quimérico transcrito in vitro se había preparado de la siguiente manera. El ARN se transcribió in vitro mediante reacciones de derrame utilizando el kit MEGAShortsript T7 (Ambion) de acuerdo con el manual del fabricante. Las plantillas para el ARN en la transcripción in vitro se generaron mediante la hibridación de dos ADN complementarios de cadena simple o mediante amplificación por PCR (Tabla 1). El ARN transcrito se resolvió en un gel de urea-PAGE desnaturalizante al 8%. La rebanada de gel que contenía ARN se cortó y se transfirió al tampón de elución de la sonda. El ARN se recuperó en agua libre de nucleasas seguido de extracción con fenol:cloroformo, extracción con cloroformo y precipitación con etanol. Los ARN purificados se cuantificaron por espectrometría.

## ES 2 690 386 T3

A las 12 h después de la transfección, el ARN quimérico (1 µg) preparado mediante transcripción in vitro se transfectó usando Lipofectamina 2000.

En la post - transfección 3d, las células transfectadas se sometieron a citometría de flujo y se contaron las células que expresaban tanto RFP como GFP.

- 5 Se encontró que las células que expresaban GFP se obtuvieron solo cuando las células se transfectaron primero con el plásmido Cas9 y luego con el ARN guía 12 h más tarde (Fig. 2), demostrando que los RGEN podían reconocer y escindir la secuencia de ADN diana en células humanas cultivadas. De este modo, las células que expresaban GFP se obtuvieron mediante transfección en serie del plásmido Cas9 y el ARN guía en lugar de la cotransfección.

10 Tabla 1

Gen		Secuencia (de 5' a 3')	SEQ ID NO.
Oligonucleótidos utilizados para la construcción del plásmido informador			
CCR5	F	AATTCATGACATCAATTATTATACATCGGAGGAG	3
	R	GATCCTCCTCCGATGTATAATAATTGATGTCATG	4
Cebadores utilizados en el ensayo T7E1			
CCR5	F1	CTCCATGGTGCTATAGAGCA	5
	F2	GAGCCAAGCTCTCCATCTAGT	6
	R	GCCCTGTCAAGAGTTGACAC	7
C4BPB	F1	TATTTGGCTGGTTGAAAGGG	8
	R1	AAAGTCATGAAATAAACACACCCA	9
	F2	CTGCATTGATATGGTAGTACCATG	10
	R2	GCTGTTCAATGCAATGGAATG	11
Cebadores utilizados para la amplificación de sitios fuera de la diana			
ADCY5	F1	GCTCCCACCTTAGTGCTCTG	12
	R1	GGTGGCAGGAACCTGTATGT	13
	F2	GTCATTGGCCAGAGATGTGGA	14
	R2	GTCCCATGACAGGCGTGTAT	15
KCNJ6	F	GCCTGGCCAAGTTTCAGTTA	16
	R1	TGGAGCCATTGGTTTGCATC	17
	R2	CCAGAACTAAGCCGTTTCTGAC	18
CNTNAP2	F1	ATCACCGACAACCAGTTTCC	19
	F2	TGCAGTGCAGACTCTTTCCA	20
	R	AAGGACACAGGGCAACTGAA	21
N/A Chr. 5	F1	TGTGGAACGAGTGGTGACAG	22
	R1	GCTGGATTAGGAGGCAGGATTC	23
	F2	GTGCTGAGAACGCTTCATAGAG	24
	R2	GGACCAAACCACATTCTTCTCAC	25
Cebadores utilizados para la detección de deleciones cromosómicas			

Gen		Secuencia (de 5' a 3')	SEQ ID NO.
Supresión	F	CCACATCTCGTTCTCGGTTT	26
	R	TCACAAGCCCACAGATATTT	27

### 1-3. Disrupción dirigida de genes endógenos en células de mamíferos por RGEN

Para probar si los RGEN podían usarse para la disrupción dirigida de genes endógenos en células de mamífero, se analizaron ADN genómico aislado de células transfectadas usando T7 endonucleasa I (T7E1), una endonucleasa sensible al desapareamiento que reconoce y escinde específicamente hetero dúplex formados por la hibridación natural y las secuencias de ADN mutante (3).

Para introducir DSB en células de mamífero usando RGEN, se transfectaron  $2 \times 10^6$  células K562 con 20  $\mu\text{g}$  de plásmido que codifica Cas9 usando 4D - Nucleofector, SF Cell Line 4D - Nucleofector X Kit, Programa FF - 120 (Lonza) según el protocolo del fabricante. Para este experimento, se cultivaron células K562 (ATCC, CCL-243) en RPMI-1640 con 10% de SFB y la mezcla de penicilina/estreptomicina (100 U/ml y 100  $\mu\text{g}/\text{ml}$ , respectivamente).

Después de 24 h, se nucleofectaron 10-40  $\mu\text{g}$  de ARN quimérico transcrito in vitro en  $1 \times 10^6$  células K562. El ARN quimérico transcrito in vitro se preparó como se describe en el Ejemplo 1-2.

Las células se recogieron dos días después de la transfección del ARN y se aisló el ADN genómico. La región que incluía el sitio diana se amplificó por PCR usando los cebadores descritos en la Tabla 1. Los amplicones se sometieron al ensayo T7E1 como se describió previamente (3). Para el análisis de secuenciación, los productos de PCR correspondientes a las modificaciones genómicas se purificaron y se clonaron en el vector T-Blunt usando el kit de clonación T-Blunt PCR (SolGent). Los productos clonados se secuenciaron usando el cebador M13.

Se descubrió que las mutaciones solo se inducían cuando las células se transfectaban en serie con plásmido que codificaba Cas9 y luego con ARN guía (Fig. 3). Las frecuencias de mutación (Inserciones-delecciones (%)) en la Fig. 3A) estimadas a partir de las intensidades relativas de la banda de ADN dependían de la dosis de ARN, variando de 1,3% a 5,1%. El análisis de secuenciación de ADN de los amplicones de PCR corroboró la inducción de mutaciones mediadas por RGEN en los sitios endógenos. Las inserciones-delecciones y microhomologías, características de NHEJ propenso a error, se observaron en el sitio diana. La frecuencia de mutación medida por secuenciación directa fue del 7,3% (= 7 clones mutantes/96 clones), en el par con los obtenidos con nucleasas de dedos de zinc (ZFN) o nucleasas efectoras de tipo transcriptor-activador (TALEN).

Se requirió la transfección en serie del plásmido Cas9 y el ARN guía para inducir mutaciones en las células. Pero en el caso de los plásmidos que codifican ARN guía, la transfección en serie era innecesaria y las células se cotransfectaron con el plásmido Cas9 y el plásmido que codifica el ARN guía.

Mientras tanto, tanto ZFN como TALEN se habían desarrollado con éxito para alterar el gen CCR5 humano (3-6), que codifica un receptor de quimiocina acoplado a proteína G, un co-receptor esencial de la infección por VIH. Una ZFN específica para CCR5 está ahora bajo investigación clínica en los EE.UU. para el tratamiento del SIDA (7). Estas ZFN y TALEN, sin embargo, tienen efectos fuera de diana, induciendo tanto mutaciones locales en sitios cuyas secuencias son homólogas a la secuencia en la diana (6, 8-10) y reordenamientos del genoma que surgen de la reparación de dos DSB simultáneos inducidos en sitios en la diana y fuera de la diana (11-12). Los sitios fuera de diana más llamativos asociados con estas nucleasas modificadas específicamente para CCR5 residen en el locus CCR2, un homólogo cercano de CCR5, localizado 15 kbp aguas arriba de CCR5. Para evitar mutaciones fuera de diana en el gen CCR2 y delecciones indeseadas, inversiones y duplicaciones del segmento cromosómico de 15 kbp entre los sitios CCR5 en la diana y CCR2 fuera de la diana, los presentes inventores eligieron intencionadamente el sitio diana del RGEN específico de CCR5 para reconocer una región dentro de la secuencia CCR5 que no tiene homología aparente con la secuencia CCR2.

Los presentes inventores investigaron si el RGEN específico de CCR5 tenía efectos fuera de la diana. Con este fin, se buscaron posibles sitios fuera de la diana en el genoma humano identificando los sitios que son más homólogos con la secuencia diana de 23 pb. Como se esperaba, no se encontraron tales sitios en el gen CCR2. En cambio, se encontraron cuatro sitios, cada uno de los cuales llevaba apareamientos erróneos de 3 bases con el sitio en la diana (Fig. 4A). Los ensayos T7E1 mostraron que las mutaciones no se detectaron en estos sitios (sensibilidad del ensayo, ~ 0,5%), lo que demuestra especificidades exquisitas de los RGEN (Fig. 4B). Además, se usó PCR para detectar la inducción de delecciones cromosómicas en células transfectadas por separado con plásmidos que codificaban ZFN y RGEN específicos para CCR5. Mientras que ZFN indujo delecciones, RGEN no lo hizo (Fig. 4C).

A continuación, los RGEN se reprogramaron reemplazando el ARN guía específico de CCR5 por un ARN recién sintetizado diseñado para reconocer el gen C4BPB humano, que codifica la cadena beta de proteína de unión a C4b, un factor de transcripción. Este RGEN indujo mutaciones en el sitio diana cromosómico en células K562 a altas frecuencias (Fig. 3B). Las frecuencias de mutación medidas por el ensayo T7E1 y por secuenciación directa fueron

14% y 8,3% (= 4 clones mutantes/48 clones), respectivamente. De cuatro secuencias mutantes, dos clones contenían una inserción de una sola base o dos bases precisamente en el sitio de corte, un patrón que también se observó en el sitio diana CCR5. Estos resultados indican que los RGEN escinden el ADN diana cromosómico en posiciones esperadas en las células.

5 Ejemplo 2: Edición de genoma mediado por RGEN proteínico

Los RGEN se pueden liberar en las células de muchas formas diferentes. Los RGEN consisten en la proteína Cas9, ARNcr, y ARNtracr. Los dos ARN se pueden fusionar para formar un ARN guía monocatenario (ARNsg). Un plásmido que codifica Cas9 bajo un promotor tal como CMV o CAG se puede transfectar en las células. El ARNcr, el ARNtracr o el ARNsg también pueden expresarse en células usando plásmidos que codifican estos RNA. El uso de plásmidos, sin embargo, a menudo da como resultado la integración de la totalidad o parte de los plásmidos en el genoma del huésped. Las secuencias bacterianas incorporadas en el ADN del plásmido pueden provocar una respuesta inmune no deseada in vivo. Las células transfectadas con plásmidos para terapia celular o animales y plantas derivadas de células transfectadas con ADN deben someterse a un procedimiento de regulación costoso y prolongado antes de la aprobación del mercado en la mayoría de los países desarrollados. Además, el ADN plasmídico puede persistir en las células durante varios días después de la transfección, lo que agrava los efectos fuera de la diana de los RGEN.

En este documento se utiliza la proteína recombinante Cas9 complejada con el ARN guía transcrito in vitro para inducir la interrupción dirigida de genes endógenos en células humanas. La proteína Cas9 recombinante fusionada con el marcador de hexahistidina se expresó y se purificó a partir de E. coli usando cromatografía de afinidad de Ni estándar y filtración en gel. La proteína Cas9 recombinante purificada se concentró en tampón de almacenamiento (HEPES 20 mM pH 7,5, KCl 150 mM, DTT 1 mM y glicerol al 10%). El complejo de proteína Cas9/ARNsg se introdujo directamente en las células K562 por nucleofección: 1x10<sup>6</sup> células K562 se transfectaron con 22,5-225 (1,4-14 µM) de proteína Cas9 mezclada con 100ug (29 µM) de ARNsg transcrito in vitro (o ARNcr 40 µg y ARNtracr 80 µg) en una solución de 100 µl usando el 4D-Nucleofector, SF Cell Line 4D-Nucleofector X Kit, Programa FF-120 (Lonza) de acuerdo con el protocolo del fabricante. Después de la nucleofección, las células se colocaron en medio de crecimiento en placas de 6 pocillos y se incubaron durante 48 horas. Cuando se transfectaron 2x10<sup>5</sup> células K562 con protocolo de escala 1/5, se usaron 4,5-45 µg de proteína Cas9 mezclada con 6-60 µg de ARNsg transcrito in vitro (o 8 µg de ARNcr y 16 µg de ARNtracr) y se nucleofectaron en una solución de 20 µl. Las células nucleofectadas se colocaron luego en medio de crecimiento en placas de 48 pocillos. Después de 48 horas, se recogieron las células y se aisló el ADN genómico. La región de ADN genómico que abarca el sitio diana se amplificó por PCR y se sometió al ensayo T7E1.

Como se muestra en la Fig. 10, el complejo proteína/ARNsg de Cas9 indujo la mutación dirigida en el locus CCR5 a frecuencias que oscilaron entre 4,8 y 38% en una manera dependiente de la dosis de ARNsg o proteína Cas9, en el par con la frecuencia obtenida con la transfección del plásmido Cas9 (45%). El complejo proteína Cas9/ARNcr/ARNtracr fue capaz de inducir mutaciones con una frecuencia del 9,4%. La proteína Cas9 sola no logró inducir mutaciones. Cuando se transfectaron 2x10<sup>5</sup> células con 1/5 de dosis reducidas de proteína Cas9 y ARNsg, las frecuencias de mutación en el locus CCR5 variaron de 2,7 a 57% de una manera dependiente de la dosis, mayor que la obtenida con la cotransfección del plásmido Cas9 y el plásmido ARNsg (32%).

También se probó el complejo de proteína Cas9/ARNsg que se dirige al gen ABCC11 y se descubrió que este complejo inducía inserciones-delecciones a una frecuencia del 35%, lo que demuestra la utilidad general de este método.

Tabla 2

Secuencias de ARN guía

Diana	Tipo de ARN	Secuencia de ARN (de 5' a 3')	Longitud	SEQ ID NO
CCR5	ARNsg	GGUGACAUCAAUUUUUAUACAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCA CCGAGUCGGUGCUUUUUUU	104 bp	28
	ARNcr	GGUGACAUCAAUUUUUAUACAUGUUUUAGAGCUAUGCUGUUUU	44 pb	29
	ARNtracr	GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU	86 pb	30

Ejemplo 3: Edición del genoma guiado por ARN en ratones

Para examinar el potencial de reconocimiento génico de los RGEN en embriones de ratón pronuclear (PN), se usaron el gen de la caja forkhead N1 (Foxn1), que es importante para el desarrollo del timo y la diferenciación de queratinocitos (Nehls et al., 1996) y la proteína quinasa, gen activado por ADN, polipéptido catalítico (Prkdc), que codifica una enzima crítica para la reparación y recombinación de ADN DSB (Taccioli et al., 1998).

Para evaluar la actividad de edición del genoma de Foxn1-RGEN, se inyectó ARNm de Cas9 (10 ng/μl de solución) con varias dosis de ARNsg (figura 5a) en el citoplasma de embriones de ratón en etapa PN, y se llevaron a cabo Ensayos con T7 endonucleasa I (T7E1) (Kim et al., 2009) usando ADN genómicos obtenidos de embriones cultivados in vitro (figura 6a).

Alternativamente, se inyectó directamente el RGEN en forma de proteína Cas9 recombinante (de 0,3 a 30 ng/μl) formando un complejo con el exceso molar doble del ARNsg específico de Foxn1 (de 0,14 a 14 ng/μl) en el citoplasma o pronúcleo de embriones de ratón de una célula, y las mutaciones se analizaron en el gen Foxn1 usando embriones cultivados in vitro (Fig. 7).

Específicamente, se sintetizaron ARNm y ARNsg de Cas9 en moldes de ADN lineal in usando el kit mMMESSAGE mMACHINE T7 Ultra (Ambion) y el kit MEGAscript T7 (Ambion), respectivamente, de acuerdo con las instrucciones del fabricante, y se diluyeron con cantidades apropiadas de pirocarbonato de dietilo (Tampón de inyección tratado con DEPC, Sigma) (EDTA 0,25 mM, Tris 10 mM, pH 7,4). Las plantillas para la síntesis de ARNsg se generaron usando los oligonucleótidos enumerados en la Tabla 3. La proteína Cas9 recombinante se obtuvo de ToolGen, Inc.

Tabla 3

Nombre de ARN	Dirección	Secuencia (de 5' a 3')	SEQ ID NO
Foxn1 n°. 1 ARNsg	F	GAAATTAATACGACTCACTATAGGCAGTCTGACGTCACACTTCCGTT TTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	31
Foxn1 n°. 2 ARNsg	F	GAAATTAATACGACTCACTATAGGACTTCCAGGCTCCACCCGACGT TTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	32
Foxn1 n°. 3 ARNsg	F	GAAATTAATACGACTCACTATAGGCCAGGCTCCACCCGACTGGAGT TTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	33
Foxn1 n°. 4 ARNsg	F	GAAATTAATACGACTCACTATAGGACTGGAGGGCGAACCCCAAGGT TTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	34
Foxn1 n°. 5 ARNsg	F	GAAATTAATACGACTCACTATAGGACCCCAAGGGGACCTCATGCGT TTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	35
PRKDC n°. 1 ARNsg	F	GAAATTAATACGACTCACTATAGGTTAGTTTTTCCAGAGACTTGTTT TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	36
Prkdc n°. 2 ARNsg	F	GAAATTAATACGACTCACTATAGGTTGGTTTGCTTGTTTATCGTTT TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	37
Prkdc n°. 3 ARNsg	F	GAAATTAATACGACTCACTATAGGCACAAGCAAACCAAAGTCTCGTT TTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	38
Prkdc n°. 4 ARNsg	F	GAAATTAATACGACTCACTATAGGCCTCAATGCTAAGCGACTTCGTT TTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG	39

Todos los experimentos con animales se realizaron de acuerdo con las directrices de la Administración de Alimentos y Medicamentos de Corea (KFDA). Los protocolos fueron revisados y aprobados por los Comités Institucionales de Cuidado y Uso Animal (IACUC) del Centro de Investigación de Animales de Laboratorio de la Universidad de Yonsei

(Número de Permiso: 2013-0099). Todos los ratones se mantuvieron en las instalaciones libres de patógenos específicos del Centro de investigación de animales de laboratorio de Yonsei. Las cepas de ratón FVB/NTac (Taconic) e ICR se usaron como donantes de embriones y madres adoptivas, respectivamente. Ratones hembra FVB/NTac (7-8 semanas de edad) fueron super-ovulados por inyecciones intraperitoneales de 5 UI de gonadotropina sérica de yegua preñada (PMSG, Sigma) y 5 UI de gonadotropina coriónica humana (hCG, Sigma) a intervalos de 48 horas. Las hembras superóvuladas se aparearon con machos FVB/NTac y se recolectaron embriones fertilizados de los oviductos.

Se inyectaron ARNm y ARNsgs de Cas9 en medio M2 (Sigma) en el citoplasma de huevos fertilizados con pronúcleos bien reconocidos utilizando un micromanipulador piezoeléctrico (Prime Tech).

En el caso de inyección de proteína Cas9 recombinante, el complejo de proteína Cas9 recombinante: Foxn1-ARNsg se diluyó con tampón de inyección tratado con DEPC (EDTA 0,25 mM, Tris 10 mM, pH 7,4) y se inyectó en pronúcleos masculinos utilizando un micromanipulador TransferMan NK2 y un microinyector FemtoJet (Eppendorf).

Los embriones manipulados se transfirieron a los oviductos de madres adoptivas pseudopreñadas para producir animales vivos, o se cultivaron in vitro para análisis adicionales.

Para cribar ratones F0 y embriones de ratón cultivados in vitro con mutaciones inducidas por RGEN, se realizaron ensayos T7E1 como se describió previamente usando muestras de ADN genómico de biopsias de cola y lisados de embriones enteros (Cho et al., 2013).

Brevemente, la región genómica que abarca el sitio diana RGEN se amplificó por PCR, se fundió y se re-hibridó para formar ADN heterodúplex, que se trató con T7 endonucleasa I (New England Biolabs), y luego se analizó mediante electroforesis en gel de agarosa. Los posibles sitios fuera de la diana se identificaron mediante la búsqueda con bowtie 0.12.9 y también se monitorizaron de forma similar mediante ensayos T7E1. Los pares de cebadores usados en estos ensayos se enumeraron en las Tablas 4 y 5.

Tabla 4

Cebadores usados en el ensayo T7E1

Gen	Dirección	Secuencia (de 5' a 3')	SEQ ID NO
Foxn1	F1	GTCTGTCTATCATCTCTTCCCTTCTCTCC	40
	F2	TCCCTAATCCGATGGCTAGCTCCAG	41
	R1	ACGAGCAGCTGAAGTTAGCATGC	42
	R2	CTACTCAATGCTCTTAGAGCTACCAGGCTTGC	43
Prkdc	F	GACTGTTGTGGGGAGGGCCG	44
	F2	GGGAGGGCCGAAAGTCTTATTTTG	45
	R1	CCTGAAGACTGAAGTTGGCAGAAGTGAG	46
	R2	CTTTAGGGCTTCTTCTCTACAATCACG	47

25

Tabla 5

Cebadores usados para la amplificación de sitios fuera de la diana

Gen	Anotación	Dirección	Secuencia (de 5' a 3')	SEQ ID NO
Foxn1	off 1	F	CTCGGTGTGTAGCCCTGAC	48
		R	AGACTGGCCTGGAACCTCACAG	49
	off 2	F	CACTAAAGCCTGTCAGGAAGCCG	50
		R	CTGTGGAGAGCACACAGCAGC	51
	off 3	F	GCTGCGACCTGAGACCATG	52
		R	CTTCAATGGCTTCTGCTTAGGCTAC	53

Gen	Anotación	Dirección	Secuencia (de 5' a 3')	SEQ ID NO
	off 4	F	GGTTCAGATGAGGCCATCCTTTC	54
		R	CCTGATCTGCAGGCTTAACCCTTG	55
Prkdc	off 1	F	CTCACCTGCACATCACATGTGG	56
		R	GGCATCCACCCTATGGGGTC	57
	off 2	F	GCCTTGACCTAGAGCTTAAAGAGCC	58
		R	GGTCTTGTTAGCAGGAAGGACACTG	59
	off 3	F	AAAACCTCTGCTTGATGGGATATGTGGG	60
		R	CTCTCACTGGTTATCTGTGCTCCTTC	61
	off 4	F	GGATCAATAGGTGGTGGGGGATG	62
		R	GTGAATGACACAATGTGACAGCTTCAG	63
	off 5	F	CACAAGACAGACCTCTCAACATTCAGTC	64
		R	GTGCATGCATATAATCCATTCTGAT TGCTCTC	65
	off 6	F1	GGGAGGCAGAGGCAGGT	66
		F2	GGATCTCTGTGAGTTTGAGGCCA	67
		R1	GCTCCAGAACTCACTCTTAGGCTC	68

5 Los fundadores mutantes identificados por el ensayo T7E1 se analizaron adicionalmente mediante fPCR. Las regiones apropiadas de ADN genómico se secuenciaron como se describió previamente (Sung et al., 2013). Para la genotipificación por PCR de rutina de progenies F1, se usaron los siguientes pares de cebadores tanto para alelos mutantes como naturales:

5'-CTACTCCCTCCGCAGTCTGA-3' (SEQ ID NO: 69) y

5'-CCAGGCCTAGGTTCCAGGTA-3' (SEQ ID NO: 70) para el gen Foxn1,

5'-CCCCAGCATTGCAGATTTCC-3' (SEQ ID NO: 71) y

5'-AGGGCTTCTTCTCTACAATCACG-3' (SEQ ID NO: 72) para el gen Prkdc.

10 En el caso de la inyección de ARNm de Cas9, las fracciones mutantes (el número de embriones mutantes/número de embriones totales) fueron dependientes de la dosis, variando desde 33% (1 ng/μl de ARNsg) hasta 91% (100 ng/μl) (Fig. 6b). El análisis de secuencia confirmó las mutaciones en el gen Foxn1; la mayoría de las mutaciones fueron pequeñas deleciones (Fig. 6c), que recuerdan a las inducidas por ZFN y TALEN (Kim et al., 2013).

15 En el caso de la inyección de la proteína Cas9, estas dosis y métodos de inyección afectaron mínimamente la supervivencia y el desarrollo de los embriones de ratón in vitro: más del 70% de los embriones inyectados con RGEN eclosionaron normalmente en ambos experimentos. De nuevo, las fracciones mutantes obtenidas con la inyección de proteína Cas9 fueron dependientes de la dosis y alcanzaron hasta el 88% en la dosis más alta mediante inyección de pronúcleo y al 71% mediante inyección intracitoplasmática (Figuras 7a y 7b). De forma similar a los patrones de mutación inducidos por ARNm de Cas9 más ARNsg (figura 6c), los inducidos por el complejo  
20 proteína-ARNsg Cas9 eran en su mayoría pequeñas deleciones (figura 7c). Estos resultados demuestran claramente que los RGEN tienen una alta actividad de selección de genes en embriones de ratón.

Alentados por las altas frecuencias mutantes y la baja citotoxicidad inducida por RGEN, se produjeron animales vivos transfiriendo los embriones de ratón a los oviductos de madres adoptivas pseudopreñadas.

25 En particular, las tasas de natalidad fueron muy altas, que oscilaron entre el 58% y el 73%, y no se vieron afectadas por las crecientes dosis de Foxn1-ARNsg (Tabla 6).

Tabla 6

Reconocimiento génico mediado por RGEN en ratones FVB/NTac

Gen diana	ARNm + ARNsg de Cas9 (ng/μl)	Embriones inyectados	Embriones transferidos (%)	Total de recién nacidos (%)	Recién nacidos vivos* (%)	Fundadores† (%)
Foxn1	10 + 1	76	62 (82)	45 (73)	31 (50)	12 (39)
	10 + 10	104	90 (87)	52 (58)	58 (64)	33 (57)
	10 + 100	100	90 (90)	62 (69)	58 (64)	54 (93)
	Total	280	242 (86)	159 (66)	147 (61)	99 (67)
Prkdc	50 + 50	73	58 (79)	35 (60)	33 (57)	11 (33)
	50 + 100	79	59 (75)	22 (37)	21 (36)	7 (33)
	50 + 250	94	73 (78)	37 (51)	37 (51)	21 (57)
	Total	246	190 (77)	94 (49)	91 (48)	39 (43)

- 5 De 147 recién nacidos, se obtuvieron 99 ratones mutantes fundadores. De acuerdo con los resultados observados en embriones cultivados (Fig. 6b), las fracciones mutantes fueron proporcionales a las dosis de Foxn1-ARNsg, y alcanzaron hasta 93% (100 ng/μl de Foxn1-ARNsg) (Tablas 6 y 7, Fig. 5b).

Tabla 7

- 10 Secuencias de ADN de los alelos mutantes de Foxn1 identificados a partir de un subconjunto de fundadores mutantes positivos con T7E1

ACTTCCAGGCTCCACCCGACTGGAGGGCGAACCCCAAGGGGACCTCATGCAGG	del + ins	n°.	Ratones fundadores
ACTTCCAGGC ----- AACCCCAAGGGGACCTCATGCAGG	Δ19	1	20
ACTTCCAGGC ----- GAACCCCAAGGGGACCTCATGCAGG	Δ18	1	115
ACTTCCAGGCTCC -----	Δ60	1	19
ACTTCCAGGCTCC -----	Δ44	1	108
ACTTCCAGGCTCC ----- CAAGGGGACCTCATGCAGG	Δ21	1	64
ACTTCCAGGCTCC ----- TTAGGAGGCGAACCCCAAGGGGACCTCA	Δ12 + 6	1	126
ACTTCCAGGCTCCACC ----- TCATGCAGG	Δ28	1	5
ACTTCCAGGCTCCACCC ----- CCAAGGGACCTCATG	Δ21 + 4	1	61
ACTTCCAGGCTCCACCC ----- AAGGGGACCTCATGCAGG	Δ18	2	95, 29
ACTTCCAGGCTCCACCC ----- CAAGGGGACCTCATGCAGG	Δ17	7	12, 14, 27, 66, 108, 114, 126
ACTTCCAGGCTCCACCC ----- ACCCAAGGGGACCTCATGCAG	Δ15 + 1	1	32
ACTTCCAGGCTCCACCC ----- CACCCCAAGGGGACCTCATGCA	Δ15 + 2	1	124

ES 2 690 386 T3

ACTTCCAGGCTCCACCCGACTGGAGGGCGAACCCCAAGGGGACCTCATGCAGG	del + ins	n°.	Ratones fundadores
ACTTCCAGGCTCCACCC ----- ACCCCAAGGGGACCTCATGCAGG	Δ13	1	32
ACTTCCAGGCTCCACCC-----GGCGAACCCCAAGGGGACCTCATGCAGG	Δ8	1	110
ACTTCCAGGCTCCACCCCT ----- GGGGACCTCATGCAGG	Δ20 + 1	1	29
ACTTCCAGGCTCCACCCG ----- AACCCCAAGGGGACCTCATGCAGG	Δ11	1	111
ACTTCCAGGCTCCACCCGA ----- ACCTCATGCAGG	Δ22	1	79
ACTTCCAGGCTCCACCCGA ----- GGGGACCTCATGCAGG	Δ18	2	13, 127
ACTTCCAGGCTCCACCCCA ----- AGGGGACCTCATGCAGG	Δ17	1	24
ACTTCCAGGCTCCACCCGA ----- ACCCCAAGGGGACCTCATGCAGG	Δ11	5	14, 53, 58, 69, 124
ACTTCCAGGCTCCACCCGA ----- GACCCCAAGGGGACCTCATGCAGG	Δ10	1	14
ACTTCCAGGCTCCACCCGA ----- GGCGAACCCCAAGGGGACCTCATGCAGG	Δ5	3	53, 79, 115
ACTTCCAGGCTCCACCCGAC ----- CTCATGCAGG	Δ23	1	108
ACTTCCAGGCTCCACCCGAC ----- CCCCAAGGGGACCTCATGCAGG	Δ11	1	3
ACTTCCAGGCTCCACCCGAC ----- GAAGGGCCCAAGGGGACCTCA	Δ11 + 6	1	66
ACTTCCAGGCTCCACCCGAC ----- GAACCCCAAGGGGACCTCATGCAGG	Δ8	2	3, 66
ACTTCCAGGCTCCACCCGAC ----- GGCGAACCCCAAGGGGACCTCATGCAGG	Δ5	1	27
ACTTCCAGGCTCCACCCGAC - GTGCTTGAGGGCGAACCCCAAGGGGACCTCA	Δ2 + 6	2	5
ACTTCCAGGCTCCACCCGACT ----- CACTATCTTCTGGGCTCCTCCATGTC	Δ6 + 25	2	21, 114
ACTTCCAGGCTCCACCCGACT ---- TGGCGAACCCCAAGGGGACCTCATGCAG	Δ4 + 1	1	53
ACTTCCAGGCTCCACCCGACT - TGCAGGGCGAACCCCAAGGGGACCTCATGC	Δ2 + 3	1	126
ACTTCCAGGCTCCACCCGACTTGGAGGGCGAACCCCAAGGGGACCTCATGCAG	+1	15	3, 5, 12, 19, 29, 55, 56, 61, 66, 68, 81, 108, 111, 124, 127
ACTTCCAGGCTCCACCCGACTTTGGAGGGCGAACCCCAAGGGGACCTCATGCA	+2	2	79, 120
ACTTCCAGGCTCCACCCGACTGTTGGAGGGCGAACCCCAAGGGGACCTCATGC	+3	1	55
ACTTCCAGGCTCCACCCGACTGGAG (+455) GGCGAACCCCAAGGGGACCTCC	+455	1	13

Para generar ratones dirigidos a Prkdc, se aplicó una concentración 5 veces mayor de ARNm de Cas9 (50 ng/μl) con dosis crecientes de Prkdc-ARNsg (50, 100 y 250 ng/μl). Una vez más, las tasas de natalidad fueron muy altas, variando del 51% al 60%, lo suficiente como para producir un número suficiente de recién nacidos para el análisis (Tabla 6). La fracción mutante fue del 57% (21 fundadores mutantes entre 37 recién nacidos) a la dosis máxima de Prkdc-ARNsg. Estas tasas de natalidad obtenidas con RGEN fueron aproximadamente de 2 a 10 veces más altas

5

## ES 2 690 386 T3

que aquellas con TALEN informadas en el estudio anterior de los inventores (Sung et al., 2013). Estos resultados demuestran que los RGEN son potentes reactivos dirigidos a genes con toxicidad mínima.

Para probar la transmisión de la línea germinal de los alelos mutantes, se cruzaron el fundador del mutante Foxn1 n°. 108, un mosaico con cuatro alelos diferentes (Fig. 5c y Tabla 8) con ratones naturales, y se monitorearon los genotipos de la descendencia F1.

5

Tabla 8

Genotipos de ratones mutantes Foxn1

Fundador n°.	ARNsg (ng/ml)	Resumen de genotipificación	Alelos detectados
58*	1	no determinado	$\Delta 11$
19	100	bi-alélico	$\Delta 60/+ 1$
20	100	bi-alélico	$\Delta 67/\Delta 19$
13	100	bi-alélico	$\Delta 18/+ 455$
32	10	bialélico (heterocigoto)	$\Delta 13/\Delta 15 + 1$
115	10	bialélico (heterocigoto)	$\Delta 18/\Delta 5$
111	10	bialélico (heterocigoto)	$\Delta 11/+ 1$
110	10	bialélico (homocigoto)	$\Delta 8/\Delta 8$
120	10	bi-alélico (homocigoto) + 2/+ 2	
81	100	heterocigoto	+ 1/WT
69	100	homocigoto	$\Delta 11/\Delta 11$
55	1	mosaico	$\Delta 18/\Delta 1/+ 1/+ 3$
56	1	mosaico	$\Delta 127/\Delta 41/\Delta 2/+ 1$
127	1	mosaico	$\Delta 18/+ 1/WT$
53	1	mosaico	$\Delta 11/\Delta 5/\Delta 4 + 1/WT$
27	10	mosaico	$\Delta 17/\Delta 5/WT$
29	10	mosaico	$\Delta 18/\Delta 20 + 1/+ 1$
95	10	mosaico	$\Delta 18/\Delta 14/\Delta 8/\Delta 4$
108	10	mosaico	+ 1/ $\Delta 17/\Delta 23/\Delta 44$
114	10	mosaico	$\Delta 17/\Delta 8/\Delta 6 + 25$
124	10	mosaico	$\Delta 11/\Delta 15 + 2/+ 1$
126	10	mosaico	$\Delta 17/\Delta 2 + 3/\Delta 12 + 6$
12	100	mosaico	$\Delta 30/\Delta 28/\Delta 17/+ 1$
5	100	mosaico	$\Delta 28/\Delta 11/\Delta 2 + 6/+ 1$
14	100	mosaico	$\Delta 17/\Delta 11/\Delta 10$
21	100	mosaico	$\Delta 127/\Delta 41/\Delta 2/\Delta 6 + 25$
24	100	mosaico	$\Delta 17/+ 1/WT$
64	100	mosaico	$\Delta 31/\Delta 21/+ 1/WT$
68	100	mosaico	$\Delta 17/\Delta 11/+ 1/WT$

Fundador nº.	ARNsg (ng/ml)	Resumen de genotipificación	Alelos detectados
79	100	mosaico	$\Delta 22/\Delta 5/+$ 2/WT
61	100	mosaico	$\Delta 21 + 4/\Delta 6/+1/+$ 9
66**	100	mosaico	$\Delta 17/\Delta 8/\Delta 11 + 6/+$ 1/WT
3	100	mosaico	$\Delta 11/\Delta 8/+$ 1

Los alelos subrayados fueron secuenciados.

Alelos en rojo, detectados por secuenciación, pero no por fPCR.

\*solo un clon secuenciado.

5 \*\*No determinado por fPCR.

Como se esperaba, todas las progenies fueron mutantes heterocigotos que poseían el alelo natural y uno de los alelos mutantes (Fig. 5d). También se confirmó la transmisión de la línea germinal en ratones fundadores independientes de Foxn1 (Fig. 8) y Prkdc (Fig. 9). Hasta donde se tiene conocimiento, estos resultados proporcionan la primera evidencia de que los alelos mutantes inducidos por RGEN se transmiten de manera estable a progenies F1 en animales.

Ejemplo 4: Edición del genoma guiado por ARN en plantas

#### 4-1. Producción de la proteína Cas9

La secuencia de codificación Cas9 (4104 bps), derivada de la cepa M1 GAS de *Streptococcus pyogenes* (NC\_002737.1), se clonó en el plásmido pET28-b (+). Se incluyó una secuencia de reconocimiento nuclear (NLS) en el extremo N de la proteína para asegurar la localización de la proteína en el núcleo. El plásmido pET28-b (+) que contenía Cas9 ORF se transformó en BL21(DE3). Entonces se indujo Cas9 usando IPTG 0,2 mM durante 16 horas a 18°C y se purificó usando perlas de agarosa Ni-NTA (Qiagen) siguiendo las instrucciones del fabricante. La proteína Cas9 purificada se concentró usando Ultracel - 100K (Millipore).

#### 4-2. Producción de ARN guía

La secuencia genómica del gen de *Arabidopsis* que codifica el BRI1 se cribó para detectar la presencia de un motivo NGG, el llamado protoespaciador de motivo adyacente (PAM), en un exón que se requería para el reconocimiento de Cas9. Para romper el gen BRI1 en *Arabidopsis*, se identificaron dos sitios diana de RGEN en un exón que contenían el motivo NGG. Se produjeron ARNsgs in vivo utilizando plantilla de ADN. Cada plantilla de ADN se generó por extensión con dos oligonucleótidos parcialmente superpuestos (Macrogen, Tabla X1) y Phusion polimerasa (Thermo Scientific) utilizando las siguientes condiciones: 98°C, 30 segundos (98°C, 10 s, 54°C, 20 s, 72°C, 2 min) x 20, 72°C 5 min.

Tabla 9

Oligonucleótidos para la producción del ADN plantilla para la transcripción in vitro

Oligonucleótidos	Secuencia (5'-3')	SEQ ID NO
BRI1 diana 1 (directa)	GAAATTAATACGACTCACTATAGGTTTGAAAGATGGAAGCGCGGGTTTTAGA GCTAGAAATAGCAAGTAAAATAAGGCTAGTCCG	73
BRI1 diana 2 (directa)	GAAATTAATACGACTCACTATAGGTGAACTAACTGGTCCACAGTTTTAGA GCTAGAAATAGCAAGTAAAATAAGGCTAGTCCG	74
Universal (inversa)	AAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTT ATTTAACTTGC	75

El ADN extendido se purificó y se usó como molde para la producción in vitro del ARN guía usando el kit MEGAshortscript T7 (Life Technologies). Los ARN guía se purificaron a continuación mediante extracción con fenol/cloroformo y precipitación con etanol. Para preparar complejos Cas9/ARNsg, 10 µl de proteína Cas9 purificada (12 µg/µl) y 4 µl de dos ARNsgs (11 µg/µl) se mezclaron en 20 µl de tampón NEB3 (New England Biolabs) y se incubaron durante 10 minutos a 37°C.

## 4-3. Transfección del complejo Cas9/ARNsg a protoplasto

Las hojas de plántulas de *Arabidopsis* de 4 semanas de edad cultivadas asépticamente en placas de Petri se digirieron en solución enzimática (1% de celulosa R10, 0,5% de macerozima R10, 450 mM de manitol, 20 mM de MES pH 5,7 y sal de CPW) durante 8 ~ 16 horas a 25°C con 40 rpm agitando en la oscuridad. Las soluciones de enzima/protoplasto se filtraron y centrifugaron a 100 X g durante 3 ~ 5 min. Los protoplastos se volvieron a suspender en una solución de CPW después de contar las células bajo el microscopio (X100) usando un hemacitómetro. Finalmente, los protoplastos se volvieron a suspender a  $1 \times 10^6$ /ml en solución de MMG (HEPES 4 mM, pH 5,7, manitol 400 mM y  $MgCl_2$  15 mM). Para transfectar los protoplastos con el complejo Cas9/ARNsg, 200  $\mu$ L (200.000 protoplastos) de la suspensión de protoplastos se mezclaron suavemente con 3,3 ó 10  $\mu$ L del complejo Cas9/ARNsg [proteína Cas9 (6  $\mu$ g/ $\mu$ L) y dos ARNsg (2,2  $\mu$ g/ $\mu$ L cada uno)] y 200  $\mu$ L de tampón de transfección de polietilenglicol al 40% (40% de PEG4000, 200 mM de manitol y 100 mM de  $CaCl_2$ ) en tubos de 2 ml. Después de 5 ~ 20 min de incubación a temperatura ambiente, la transfección se detuvo añadiendo tampón de lavado con solución de W5 (MES 2 mM, pH 5,7, NaCl 154 mM,  $CaCl_2$  125 mM y KCl 5 mM). Los protoplastos se recogieron después por centrifugación durante 5 minutos a 100 x g, se lavaron con 1 ml de solución de W5, se centrifugaron durante otros 5 minutos a 100 x g. La densidad de los protoplastos se ajustó a  $1 \times 10^5$ /ml y se cultivaron en medio líquido KM 8p modificado con glucosa 400 mM.

4-4. Detección de mutaciones en protoplastos y plantas de *Arabidopsis*

Después de 24 horas o 72 horas después de la transfección, se recogieron los protoplastos y se aisló el ADN genómico. La región de ADN genómica que abarca los dos sitios diana se amplificó por PCR y se sometió al ensayo T7E1. Como se muestra en la Figura 11, las inserciones-deleciones fueron inducidas por RGEN a altas frecuencias que oscilaban entre 50% y 70%. Sorprendentemente, las mutaciones se indujeron 24 horas después de la transfección. Aparentemente, la proteína Cas9 funciona inmediatamente después de la transfección. Los productos PCR se purificaron y se clonaron en el kit T-Blunt PCR Cloning (Solgent). Los plásmidos se purificaron y se sometieron a secuenciación de Sanger con cebador M13F. Una secuencia mutante tenía una delección de 7 pb en un sitio (Figura 12). Las otras tres secuencias mutantes tenían delecciones de segmentos de ADN de ~220 pb entre los dos sitios RGEN.

Ejemplo 5: Transducción de proteína Cas9 usando un péptido penetrante de células o dominio de transducción de proteínas

## 5-1. Construcción del plásmido que codifica His-Cas9

Cas9 con una cisteína en el extremo C-terminal se preparó por amplificación por PCR usando el plásmido Cas9 previamente descrito {Cho, 2013 n°. 166} como molde y se clonó en el vector pET28-(a) (Novagen, Merk Millipore, Alemania) que contenía el marcador His en el extremo N-terminal.

## 5-2. Cultivo de células

Se cultivaron 293T (línea celular de riñón embrionario humano) y HeLa (línea celular de cáncer de ovario humano) en DMEM (GIBCO-BRL Rockville) complementado con 10% de SFB y 1% de penicilina y estreptomicina.

## 5-3. Expresión y purificación de la proteína Cas9

Para expresar la proteína Cas9, las células BL21 de *E. coli* se transformaron con el vector pET28-(a) que codificaba Cas9 y se colocaron en placas en medio de agar Luria-Bertani (LB) que contenía 50  $\mu$ g/ml de kanamicina (Amresco, Solon, OH). Al día siguiente, se recogió una sola colonia y se cultivó en caldo LB que contenía 50  $\mu$ g/ml de kanamicina a 37°C durante la noche. El día siguiente, este cultivo iniciador a 0,1 OD600 se inoculó en caldo Luria que contenía 50  $\mu$ g/ml de kanamicina y se incubó durante 2 horas a 37°C hasta que OD600 alcanzó 0,6-0,8. Para inducir la expresión de la proteína Cas9, las células se cultivaron a 30°C durante la noche después de la adición de isopropil- $\beta$ -D-tiogalactopiranosido (IPTG) (Promega, Madison, WI) a la concentración final de 0,5 mM.

Las células se recogieron por centrifugación a 4000 rpm durante 15-20 minutos, se resuspendieron en un tampón de lisis (Tris-Cl 20 mM pH 8,0, NaCl 300 mM, imidazol 20 mM, cóctel inhibidor de proteasa 1X, lisozima 1 mg/ml) y se sometieron a lisis por sonicación (40% de trabajo, 10 segundos de pulso, 30 segundos de descanso, durante 10 minutos en hielo). La fracción soluble se separó como el sobrenadante después de la centrifugación a 15.000 rpm durante 20 minutos a 4°C. La proteína Cas9 se purificó a 4°C usando una columna que contenía resina de agarosa Ni-NTA (QIAGEN) y un instrumento AKTA prime (AKTA prime, GE Healthcare, RU). Durante esta etapa de cromatografía, las fracciones de proteína soluble se cargaron en una columna de resina de agarosa Ni-NTA (GE Healthcare, Reino Unido) a un caudal de 1 ml/min. La columna se lavó con un tampón de lavado (Tris-Cl 20 mM, pH 8,0, NaCl 300 mM, imidazol 20 mM, cóctel de inhibidor de proteasa 1X) y la proteína unida se eluyó a un caudal de 0,5 ml/min con un tampón de elución (Tris-Cl 20 mM, pH8,0, NaCl 300 mM, imidazol 250 mM, cóctel inhibidor de proteasa 1X). La fracción eluida reunida se concentró y se dializó frente a tampón de almacenamiento (Tris-HCl 50 mM, pH 8,0, KCl 200 mM, EDTA 0,1 mM, DTT 1 mM, PMSF 0,5 mM, glicerol al 20%). La concentración de proteína se cuantificó mediante el ensayo de Bradford (Biorad, Hercules, CA) y la pureza se analizó mediante SDS-PAGE usando albúmina de suero bovino como control.

5-4. Conjugación de Cas9 a 9R4L

1 mg de proteína Cas9 diluida en PBS a la concentración de 1 mg/ml y 50 µg de péptido maleimida-9R4L en 25 µl de DW (Peptron, Corea) se mezclaron suavemente usando un rotor a temperatura ambiente durante 2 horas y a 4°C durante la noche. Para eliminar la maleimida-9R4L no conjugada, las muestras se dializaron usando una membrana de corte de peso molecular de 50 kDa frente a DPBS (pH 7,4) a 4°C durante 24 horas. La proteína Cas9-9R4L se recogió de la membrana de diálisis y la cantidad de proteína se determinó usando el ensayo de Bradford.

5-5. Preparación de ARNsg-9R4L

Se añadió ARNsg (1 µg) suavemente a diversas cantidades de péptido C9R4LC (con una relación en peso de 1 a 40) en 100 µl de DPBS (pH 7,4). Esta mezcla se incubó a temperatura ambiente durante 30 minutos y se diluyó a 10 veces usando agua desionizada libre de ARNasa. El diámetro hidrodinámico y el potencial z de las nanopartículas formadas se midieron usando dispersión de luz dinámica (Zetasizer-nano analyser ZS, Malvern Instruments, Worcestershire, Reino Unido).

5-6. Tratamientos de la proteína Cas9 y ARNsg

Se trataron Cas9-9R4L y ARNsg-C9R4LC con las células de la siguiente manera: se añadieron 1 µg de ARNsg y 15 µg de péptido C9R4LC a 250 ml de medio OPTIMEM y se incubaron a temperatura ambiente durante 30 minutos. 24 horas después de la siembra, las células se lavaron con medio OPTIMEM y se trataron con el complejo ARNsg-C9R4LC durante 4 horas a 37°C. Las células se lavaron nuevamente con medio OPTIMEM y se trataron con Cas9-9R4L durante 2 horas a 37°C. Después del tratamiento, los medios de cultivo se reemplazaron por medio completo que contenía suero y se incubaron a 37°C durante 24 horas antes del siguiente tratamiento. Se siguió el mismo procedimiento para múltiples tratamientos de Cas9 y ARNsg durante tres días consecutivos.

5-7. Cas9-9R4L y ARNsg-9R4L pueden editar genes endógenos en células de mamíferos cultivadas sin el uso de herramientas de liberación adicionales

Para determinar si Cas9-9R4L y ARNsg-9R4L pueden editar genes endógenos en células de mamíferos cultivadas sin el uso de herramientas de administración adicionales, se trataron 293 células con Cas9-9R4L y ARNsg-9R4L dirigidas al gen CCR5 y se analizó el ADN genómico. El ensayo T7E1 mostró que el 9% del gen CCR5 se rompía en las células tratadas tanto con Cas9-9R4L como con ARNsg-9R4L y que la disrupción del gen CCR5 no se observaba en las células control, incluidas las no tratadas, las tratadas con Cas9-9R o ARNsg-9R4L, o las tratadas con Cas-9 y ARNsg no modificadas (Fig. 13), lo que sugiere que el tratamiento con la proteína Cas9-9R4L y el ARNsg conjugado con 9R4L, pero no con Cas9 y ARNsg no modificadas, puede conducir a una edición eficiente del genoma en células de mamífero.

Ejemplo 6: Control de la mutación fuera de la diana según la estructura del ARN guía

Recientemente, tres grupos informaron que los RGEN tenían efectos fuera de la diana en las células humanas. Para nuestra sorpresa, los RGEN indujeron mutaciones de manera eficiente en sitios fuera de la diana que diferían en de 3 a 5 nucleótidos de sitios en la diana. Se apreció, sin embargo, que había varias diferencias entre los RGEN de los inventores y los utilizados por otros. En primer lugar, se utilizó ARNdual, que es ARNcr más ARNtracr, en lugar de ARN guía sencillo (ARNsg) que se compone de porciones esenciales de ARNcr y ARNtracr. En segundo lugar, se transfectoron células K562 (pero no células HeLa) con ARNcr sintético en lugar de plásmidos que codificaban ARNcr. Las células HeLa se transfectoron con plásmidos que codificaban ARNcr. Otros grupos usaron plásmidos que codificaban ARNsg. En tercer lugar, el ARN guía de los inventores tenía dos nucleótidos de guanina adicionales en el extremo 5', que son necesarios para la transcripción eficaz por la polimerasa T7 in vitro. No se incluyeron tales nucleótidos adicionales en el ARNsg utilizado por otros. Por lo tanto, la secuencia de ARN del ARN guía de los inventores puede mostrarse como 5'-GGX<sub>20</sub>, mientras que 5'-GX<sub>19</sub>, en la que X<sub>20</sub> o GX<sub>19</sub> corresponde a la secuencia diana de 20 pb, representa la secuencia utilizada por otros. El primer nucleótido de guanina es requerido para la transcripción por la ARN polimerasa en las células. Para probar si los efectos RGEN fuera de la diana se pueden atribuir a estas diferencias, se eligieron cuatro RGEN que indujeron mutaciones fuera de la diana en las células humanas a altas frecuencias (13). En primer lugar, se compararon el método de los inventores de uso de ARNdual transcrito in vitro con el método de transfección de plásmidos que codificaban ARNsg en células K562 y las frecuencias de mutación medidas en los sitios en la diana y fuera de la diana a través del ensayo T7E1. Tres RGEN mostraron frecuencias de mutación comparables en sitios en la diana y fuera de la diana, independientemente de la composición del ARN guía. Curiosamente, un RGEN (VEFGA sitio 1) no indujo inserciones-delecciones en un sitio validado fuera de la diana, que difiere en tres nucleótidos del sitio en la diana (denominado OT1-11, Fig. 14), cuando se utilizó ARNdual sintético. Pero el ARNdual sintético no discriminó el otro sitio validado fuera de la diana (OT1-3), que difería en dos nucleótidos del sitio en la diana.

A continuación, se probó si la adición de dos nucleótidos de guanina en el extremo 5' de ARNsg podía hacer que los RGEN fueran más específicos al comparar ARNsg 5'-GGX<sub>20</sub> (o 5'-GGGX<sub>19</sub>) con ARNsg 5'-GX<sub>19</sub>. Cuatro ARNsg GX<sub>19</sub> formaron complejos con inserciones-delecciones inducidas por Cas9 igualmente eficientemente en sitios en la diana y fuera de la diana, tolerando hasta cuatro desapareamientos de nucleótidos. En marcado contraste, los ARNsg GGX<sub>20</sub> discriminaron los sitios fuera de la diana de manera efectiva. De hecho, el ensayo T7E1 apenas detectó

inserciones-delecciones inducidas por RGEN en seis de los siete sitios validados fuera de la diana cuando se usaron los cuatro ARNsg GGX<sub>20</sub> (Fig. 15). Sin embargo, se observó que dos ARNsg GGX<sub>20</sub> (VEGFA sitios 1 y 3) fueron menos activos en los sitios en la diana que los correspondientes ARNsg GX<sub>19</sub>. Estos resultados muestran que los nucleótidos extra en el extremo 5' pueden afectar las frecuencias de mutación en los sitios en la diana y fuera de la diana, tal vez alterando la estabilidad del ARN guía, la concentración o la estructura secundaria.

Estos resultados sugieren que tres factores: el uso de ARN guía sintético en lugar de los plásmidos que codifican RNA guía, ARNdual en lugar de ARNsg y ARNsg GGX<sub>20</sub> en lugar de ARNsg GX<sub>19</sub>, tienen efectos acumulativos sobre la discriminación de sitios fuera de la diana.

#### Ejemplo 7: Nickasas Cas9 emparejadas

En principio, las roturas monocatenarias (SSB) no pueden repararse mediante NHEJ proclive a errores, pero aún desencadenan la reparación dirigida por homología de alta fidelidad (HDR) o la reparación por escisión de base. Pero la mutagénesis dirigida inducida por nickasa vía HDR es mucho menos eficiente que la mutagénesis inducida por nucleasa. Pensamos que las nickasas Cas9 emparejadas producirían DSB compuestos, que desencadenarían la reparación del ADN a través de NHEJ o HDR, lo que llevaría a una mutagénesis eficiente (Fig. 16A). Además, las nickasas emparejadas duplicarían la especificidad de la edición del genoma basada en Cas9.

Primero se probaron varias nucleasas y nickasas Cas9 diseñadas para sitios diana en el locus AAVS1 (Fig. 16B) in vitro vía electroforesis capilar fluorescente. A diferencia de las nucleasas Cas9 que escindieron ambas cadenas de sustratos de ADN, las nickasas Cas9 compuestas de ARN guía y una forma mutante de Cas9 en la que un residuo catalítico de aspartato se cambia a una alanina (D10A Cas9) escindieron solo una cadena, produciendo mellas específicas del sitio (Fig. 16C,D). Curiosamente, sin embargo, algunas nickasas (AS1, AS2, AS3 y S6 en la Fig. 17A) indujeron inserciones-delecciones en sitios diana en células humanas, lo que sugiere que las mellas se pueden convertir a DSB, aunque de forma ineficiente, in vivo. Las nickasas Cas9 emparejadas, que producen dos mellas adyacentes en cadenas de ADN opuestas, produjeron inserciones-delecciones a frecuencias que variaron de 14% a 91%, comparable a los efectos de las nucleasas emparejadas (Fig. 17A). La reparación de dos mellas que producían proyecciones de 5' condujo a la formación de inserciones-delecciones mucho más frecuentemente que las que producían proyecciones de 3' en tres loci genómicos (Fig. 17A y Fig. 18). Además, las nickasas emparejadas permitieron la edición del genoma dirigida a través de la reparación dirigida por homología de manera más eficiente que lo hicieron las nickasas individuales (Fig. 19).

Seguidamente, se midieron las frecuencias de mutación de las nickasas y nucleasas emparejadas en los sitios fuera de la diana usando la secuenciación profunda. Las nucleasas Cas9 complejadas con tres ARNsg indujeron mutaciones fuera de la diana en seis sitios que diferían en uno o dos nucleótidos de sus sitios correspondientes en la diana con frecuencias que variaban de 0,5% a 10% (Fig. 17B). Por el contrario, las nickasas de Cas9 emparejadas no produjeron inserciones-delecciones por encima del límite de detección del 0,1% en ninguno de los seis sitios fuera de la diana. El sitio S2 Off-1 que se diferencia por un único nucleótido en la primera posición en el PAM (es decir, N en NGG) desde su sitio en la diana se puede considerar como otro sitio en la diana. Como era de esperar, la nucleasa Cas9 complejada con el ARNsg S2 fue igualmente eficiente en este sitio y en el sitio en la diana. En marcado contraste, D10A Cas9 complejado con ARNsg S2 y AS2 discriminó este sitio del sitio en la diana en un factor de 270 veces. Esta nickasa emparejada también discriminó los sitios AS2 fuera de la diana (Off-1 y Off-9 en la Fig. 17B) del sitio en la diana por factores de 160 veces y 990 veces, respectivamente.

#### Ejemplo 8: Corte y empalme de ADN cromosómico inducido por las nickasas Cas9 emparejadas

Dos DSB simultáneos producidos por nucleasas modificadas como ZFN y TALEN pueden promover grandes delecciones de los segmentos cromosómicos intermedios que se han reseñado. Se probó si dos SSB inducidas por las nickasas Cas9 emparejadas también podían producir delecciones en las células humanas. Se usó la PCR para detectar eventos de delección y se encontró que siete mellas emparejadas inducían delecciones de segmentos cromosómicos de hasta 1,1 kbp tan eficientemente como las nucleasas Cas9 emparejadas (Fig. 20A,B). Las secuencias de ADN de los productos de la PCR confirmaron los eventos de delección (Fig. 20C). Curiosamente, la secuencia de coincidencia de ARNsg permaneció intacta en dos de los siete amplicones de PCR específicos de delección (subrayados en la Fig. 20C). Por el contrario, los pares de nucleasa Cas9 no produjeron secuencias que contuvieran sitios diana intactos. Este hallazgo sugiere que las dos mellas distantes no se convirtieron en dos DSB separados para promover delecciones del segmento cromosómico intermedio. Además, es poco probable que dos mellas separadas por más de 100 bp puedan producir un DSB compuesto con grandes extremos en condiciones fisiológicas debido a que la temperatura de fusión es muy alta.

Se propone que dos mellas lejanas se reparan por desplazamiento de cadena en una dirección cabeza a cabeza, lo que resulta en la formación de un DSB en el medio, cuya reparación a través de NHEJ provoca pequeñas delecciones (Fig. 20D). Debido a que los dos sitios diana permanecen intactos durante este proceso, las nickasas pueden inducir SSB nuevamente, lo que activa el ciclo repetidamente hasta que se eliminan los sitios diana. Este mecanismo explica por qué dos mellas de compensación que producen proyecciones de 5' pero no las que producen proyecciones de 3' inducen inserciones-delecciones eficientemente en tres loci.

Luego se investigó si las nucleasas y las nickasas de Cas9 podían inducir translocaciones cromosómicas no deseadas que resultaran de la reparación de NHEJ de las divisiones de ADN en la diana y fuera de la diana (Fig. 21A). Se pudieron detectar translocaciones inducidas por nucleasas Cas9 usando PCR (Fig. 21B,C). Ninguno de dichos productos de PCR se amplificaron usando ADN genómico aislado de células transfectadas con los plásmidos que codificaban el par de nickasa Cas9 AS2 + S3. Este resultado está en línea con el hecho de que tanto las nickasas AS2 como las S3, a diferencia de sus correspondientes nucleasas, no produjeron inserciones-delecciones en sitios fuera de la diana (Fig. 17B).

Estos resultados sugieren que las mellas conjuntas Cas9 permiten la mutagénesis dirigida y grandes delecciones de segmentos cromosómicos de hasta 1 kbp en células humanas. De manera importante, las nickasas emparejadas no indujeron inserciones-delecciones en sitios fuera de la diana en los que sus correspondientes nucleasas inducen mutaciones. Además, a diferencia de las nucleasas, las mellas emparejadas no promovían translocaciones no deseadas asociadas con las divisiones de ADN fuera de la diana. En principio, las nickasas emparejadas duplican la especificidad de la mutagénesis mediada por Cas9 y ampliarán la utilidad de las enzimas guiadas por ARN en aplicaciones que requerirán la edición precisa del genoma, como la terapia génica y celular. Una advertencia para este enfoque es que se necesitan dos ARNsg altamente activos para formar un par de nickasa eficiente, lo que limita los sitios de destino. Como se muestra en este y otros estudios, no todos los ARNsg son igualmente activos. Cuando se utilizan clones individuales en lugar de poblaciones de células para estudios o aplicaciones adicionales, la elección de ARN guía que representa secuencias únicas en el genoma y el uso de ARN guía optimizados sería suficiente para evitar mutaciones fuera de la diana asociadas con nucleasas Cas9. Se propone que tanto las nucleasas Cas9 como las nickasas emparejadas son opciones potentes que facilitarán la edición precisa del genoma en células y organismos.

#### Ejemplo 9: Genotipado con endonucleasas guiadas por ARN derivadas de CRISPR/Cas

A continuación, se razonó que los RGEN se pueden utilizar en el análisis de polimorfismo de longitud de fragmentos de restricción (RFLP), en sustitución de las enzimas de restricción convencionales. Las nucleasas modificadas que incluyen RGEN inducen inserciones-delecciones en los sitios diana, cuando los DSB causados por las nucleasas son reparados por el sistema de unión al extremo no homólogo propenso a error (NHEJ). Los RGEN que están diseñados para reconocer las secuencias diana no pueden escindir las secuencias mutantes con inserciones-delecciones, sino que escindirán las secuencias diana naturales de manera eficiente.

##### 9-1. Componentes RGEN

Se prepararon ARNcr y ARNtracr mediante transcripción in vitro usando el kit MEGAscript T7 (Ambion) de acuerdo con las instrucciones del fabricante. Los ARN transcritos se resolvieron en un gel de urea-PAGE desnaturalizante al 8%. La rebanada de gel que contenía ARN se cortó y se transfirió a tampón de elución. El ARN se recuperó en agua sin nucleasas seguido de extracción con fenol:cloroformo, extracción con cloroformo y precipitación con etanol. El ARN purificado se cuantificó por espectrometría. Las plantillas para ARNcr se prepararon hibridando un oligonucleótido cuya secuencia se muestra como 5'-GAAATTAATACGACTCACTATAGGX<sub>20</sub>GTTTTAGAGCTATGCTGTTTTG-3' (SEQ ID NO: 76), en la que X<sub>20</sub> es la secuencia diana y su oligonucleótido complementario. La plantilla para el ARNtracr se sintetizó por extensión de los oligonucleótidos directo e inverso (5'-GAAATTAATACGCATAGACAGGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCG-3' (SEQ ID NO: 77) y 5'-AAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTAACTTGCTATG-3' (SEQ ID NO: 78) usando Phusion polimerasa (New England Biolabs).

##### 9-2. Purificación de la proteína Cas9 recombinante

La construcción de ADN Cas9 usada en el Ejemplo anterior de los inventores, que codifica Cas9 fusionado al marcador His6 en el extremo C, se insertó en el vector de expresión pET-28a. La proteína Cas9 recombinante se expresó en la cepa BL21 (DE3) de E. coli cultivada en medio LB a 25°C durante 4 horas después de la inducción con IPTG 1 mM. Las células se recogieron y se resuspendieron en tampón que contenía Tris 20mM pH 8,0 mM, NaCl 500 mM, imidazol 5 mM y PMSF 1 mM. Las células se congelaron en nitrógeno líquido, se descongelaron a 4°C y se sonicaron. Después de la centrifugación, la proteína Cas9 en el lisado se unió a resina de agarosa Ni-NTA (Qiagen), se lavó con tampón que contenía Tris 20 mM, pH 8,0, NaCl 500 mM e imidazol 20 mM, y se eluyó con tampón que contenía Tris 20 mM, pH 8,0, NaCl 500 mM e imidazol 250 mM. La proteína Cas9 purificada se dializó frente a HEPES 20 mM (pH 7,5), KCl 150 mM, DTT 1 mM y glicerol al 10% y se analizó mediante SDS-PAGE.

##### 9-3. Ensayo T7 endonucleasa I

El ensayo T7E1 se realizó de la siguiente manera. En resumen, los productos de PCR amplificados usando ADN genómico se desnaturalizaron a 95°C, se reasociaron a 16°C, y se incubaron con 5 unidades de T7 Endonucleasa I (New England BioLabs) durante 20 min a 37°C. Los productos de reacción se resolvieron usando electroforesis en gel de agarosa de 2 a 2,5%.

##### 9-4. Ensayo RGEN-RFLP

5 Los productos de la PCR (100-150 ng) se incubaron durante 60 min a 37°C con concentraciones optimizadas (Tabla 10) de la proteína Cas9, ARNtracr, ARNcr en 10 µl de tampón NEB 3 (1X). Después de la reacción de escisión, se añadió RNasa A (4 µg) y la mezcla de reacción se incubó durante 30 min a 37°C para eliminar el ARN. Las reacciones se detuvieron con tampón de solución de parada 6X que contenía 30% de glicerol, 1,2% de SDS y 100 mM de EDTA. Los productos se resolvieron con electroforesis en gel de agarosa al 1-2,5% y se visualizaron con tinción con EtBr.

Tabla 10

Concentración de componentes de RGEN en ensayos RFLP

Nombre de la diana	Cas9 (ng/µl)	ARNcr (ng/µl)	ARNtracr (ng/µl)
C4BPB	100	25	60
PIBF-NGG-RGEN	100	25	60
HLA-B	1,2	0,3	0,7
CCR5-ZFN	100	25	60
CTNNB1 específico natural	30	10	20
CTNNB1 mutante específico	30	10	20
CCR5 WT-específico	100	25	60
CCR5Δ32	10	2,5	6
KRASWT específico (wt)	30	10	20
KRAS mutante específico (m8)	30	10	20
KRAS WT específico (m6)	30	10	20
KRAS mutante específico (m6,8)	30	10	20
PIK3CA WT específico (wt)	100	25	60
PIK3CA mutante específico (m4)	30	10	20
PIK3CA WT específico (m7)	100	25	60
PIK3CA mutante específico (m4,7)	30	10	20
BRAF WT-específico	30	10	20
bRAF mutante específico	100	25	60
NRAS WT específico	100	25	60
NRAS mutante específico	30	10	20
IDH WT específico	30	10	20
IDH mutante específico	30	10	20
PIBF-NAG-RGEN	30	10	60

Tabla 11

Cebadores

Gen (sitio)	Dirección	Secuencia (de 5' a 3')	SEQ ID NO
CCR5 (RGEN)	F1	CTCCATGGTGCTATAGAGCA	79
	F2	GAGCCAAGCTCTCCATCTAGT	80
	R	GCCCTGTCAAGAGTTGACAC	81
CCR5 (ZFN)	F	GCACAGGGTGAACAAGATGGA	82
	R	GCCAGGTACCTATCGATTGTCAGG	83
CCR5 (del32)	F	GAGCCAAGCTCTCCATCTAGT	84
	R	ACTCTGACTGGGTCACCAGC	85
C4BPB	F1	TATTTGGCTGGTTGAAAGGG	86
	R1	AAAGTCATGAAATAAACACACCCA	87
	F2	CTGCATTGATATGGTAGTACCATG	88
	R2	GCTGTTCAATGCAATGGAATG	89
CTNNB1	F	ATGGAGTTGGACATGGCCATGG	90
	R	ACTCACTATCCACAGTTCAGCATTT ACC	91
KRAS	F	TGGAGATAGCTGTCAGCAACTTT	92
	R	CAACAAAGCAAAGGTAAAGTTGGT AATAG	93
PIK3CA	F	GGTTTCAGGAGATGTGTTACAAGG C	94
	R	GATTGTGCAATTCCTATGCAATCG GTC	95
NRAS	F	CACTGGGTACTTAATCTGTAGCCT C	96
	R	GGTTCCAAGTCATTCCCAGTAGC	97
IDH1	F	CATCACTGCAGTTGTAGGTTATAA CTATCC	98
	R	TTGAAAACACAGATCTGGTTGAA CC	99
BRAF	F	GGAGTGCCAAGAGAATATCTGG	100
	R	CTGAAACTGGTTTCAAATATTCGT TTTAAGG	101
PIBF	F	GCTCTGTATGCCCTGTAGTAGG	102
	R	TTTGCATCTGACCTTACCTTTG	103

9-5. Ensayo de escisión del plásmido

- 5 El plásmido linealizado tratado con enzima de restricción (100 ng) se incubó durante 60 minutos a 37°C con proteína Cas9 (0,1 µg), ARNtracr (60 ng) y ARNcr (25 ng) en 10 µl de tampón NEB3 (1X). Las reacciones se detuvieron con

una solución de parada 6X que contenía 30% de glicerol, 1,2% de SDS y 100 mM de EDTA. Los productos se resolvieron con electroforesis en gel de agarosa al 1% y se visualizaron con tinción con EtBr.

9-6. Estrategia de RFLP

5 Se pueden crear fácilmente nuevos RGEN con especificidades de ADN deseadas mediante la sustitución de ARNcr; no se requiere una purificación de novo de proteínas personalizadas una vez que esté disponible la proteína Cas9 recombinante. Las nucleasas modificadas genéticamente, que incluyen los RGEN, inducen pequeñas inserciones o deleciones (indels) en los sitios diana cuando los DSB causados por las nucleasas se reparan por unión final no homóloga propensa a error (NHEJ). Los RGEN que están diseñados para reconocer las secuencias diana escinden las secuencias naturales de manera eficiente pero no pueden escindir las secuencias mutantes con inserciones-  
10 deleciones (Fig. 22).

Primero se probó si los RGEN podían dividir diferencialmente plásmidos que contenían secuencias diana de C4BPB naturales o modificadas que albergaban inserciones-deleciones de 1 a 3 bases en el sitio de escisión. Ninguno de los seis plásmidos con estas inserciones-deleciones se escindió por un RGEN5 específico de C4BPB compuesto por ARNcr específico de diana, ARNtracr y proteína Cas9 recombinante (Fig. 23). En contraste, el plásmido con la  
15 secuencia diana intacta se escindió eficientemente por este RGEN.

9-7. Detección de mutaciones inducidas por los mismos RGEN usando RFLP mediado por RGEN

A continuación, para probar la viabilidad de RFLP mediado por RGEN para la detección de mutaciones inducidas por los mismos RGEN, se utilizaron clones de células de cáncer humano K562 modificadas genéticamente establecidas usando un gen C4BPB dirigido a RGEN (Tabla 12).

20 Tabla 12

Secuencia diana de RGEN usada en este estudio

Gen	Secuencia diana	SEQ ID NO
C4BPB humana	<u>AATGACCACTACATCCTCAAGGG</u>	104
Pibf1 de ratón	<u>AGATGATGTCTCATCATCAGAGG</u>	105

25 Los clones mutantes C4BPB utilizados en este estudio tienen diversas mutaciones que varían desde la delección de 94 pb a la inserción de 67 pb (Fig. 24A). De manera importante, todas las mutaciones ocurridas en los clones mutantes dieron como resultado la pérdida del sitio diana de RGEN. Entre los 6 clones C4BPB analizados, 4 clones tienen alelos mutantes y naturales (+/-) y 2 clones solo tienen alelos mutantes (-/-).

30 Los productos de PCR que abarcan el sitio diana de RGEN amplificado a partir de ADN genómico de K562 natural se digirieron completamente mediante el RGEN compuesto de ARNc específico de diana, ARNtracr y proteína Cas9 recombinante expresada en y purificada a partir de E. coli (Fig. 24B/Carril 1). Cuando los clones mutantes de C4BPB se sometieron a análisis de RFLP usando RGEN, los amplicones de PCR de clones +/- que contenían tanto alelos mutantes como naturales se digirieron parcialmente, y los de los clones -/- que no contenían el alelo natural no se digirieron del todo, no produciendo los productos de escisión correspondientes a la secuencia natural (Fig. 24B). Incluso una inserción de una sola base en el sitio diana bloqueó la digestión (clones n. ° 12 y n. ° 28) de los alelos mutantes amplificados por el C4BPB RGEN, lo que demuestra la alta especificidad del RFLP mediado por RGEN.  
35 Se sometieron los amplicones de PCR al ensayo T7E1 sensible al desapareamiento en paralelo (Fig. 24B). Notablemente, el ensayo T7E1 no fue capaz de distinguir los clones -/- de los clones +/- . Para empeorar las cosas, el ensayo T7E1 no puede distinguir clones mutantes homocigotos que contienen la misma secuencia mutante de los clones naturales, porque la hibridación de la misma secuencia mutante formará un homodúplex. Por lo tanto, el RFLP mediado por RGEN tiene una ventaja crítica sobre el ensayo de nucleasa sensible a desapareamientos  
40 convencional en el análisis de clones mutantes inducidos por nucleasas modificadas incluyendo ZFN, TALEN y RGEN.

9-8. Ensayo cuantitativo para el análisis RGEN-RFLP

45 También se investigó si el análisis RGEN-RFLP es un método cuantitativo. Las muestras de ADN genómico aisladas del clon nulo C4BPB y las células naturales se mezclaron en diversas proporciones y se usaron para amplificaciones por PCR. Los productos de PCR se sometieron a genotipado de RGEN y el ensayo de T7E1 en paralelo (figura 25b). Como se esperaba, la escisión del ADN por el RGEN fue proporcional a la relación natural a mutante. En contraste, los resultados del ensayo T7E1 se correlacionaron mal con las frecuencias de mutación inferidas a partir de las relaciones y fueron inexactos, especialmente en el alto % de mutante, una situación en la que las secuencias mutantes complementarias pueden hibridar entre sí para formar el homodúplex.

50 9-9. Análisis de los fundadores de ratones mutantes usando un genotipo RFLP mediado por RGEN

También se aplicó el genotipado de RFLP mediado por RGEN (genotipo RGEN en resumen) al análisis de los fundadores de ratones mutantes que se habían establecido mediante la inyección de TALEN en embriones de una célula de ratón (Fig. 26A). Se diseñó y se usó un RGEN que reconocía el sitio diana TALEN en el gen *Pibf1* (Tabla 10). El ADN genómico se aisló de un ratón natural y ratones mutantes y se sometió a genotipado de RGEN después de la amplificación por PCR. La genotipificación de RGEN detectó con éxito varias mutaciones, que variaron de las deleciones de una a 27 pb (Fig. 26B). A diferencia del ensayo T7E1, el genotipado de RGEN permitió la detección diferencial de los fundadores +/- y -/-.

#### 9-10. Detección de mutaciones inducidas en células humanas por un ZFN específico de CCR5 utilizando RGEN

Además, se usaron RGEN para detectar mutaciones inducidas en células humanas por un ZFN específico de CCR5, que representa otra clase más de nucleasas modificadas (Fig. 27). Estos resultados muestran que los RGEN pueden detectar mutaciones inducidas por nucleasas distintas de RGEN. De hecho, se espera que los RGEN puedan diseñarse para detectar mutaciones inducidas por la mayoría, sino todas, de las nucleasas modificadas genéticamente. La única limitación en el diseño de un ensayo de genotipado de RGEN es el requisito para el dinucleótido GG o AG (CC o CT en la cadena complementaria) en la secuencia PAM reconocida por la proteína Cas9, que ocurre una vez por 4 pb en promedio. Las inserciones-deleciones inducidas en cualquier lugar dentro de la región de la semilla de varias bases en el ARNcr y los nucleótidos PAM se espera que rompan la escisión del ADN catalizada por RGEN. De hecho, se identificó al menos un sitio RGEN en la mayoría (98%) de los sitios ZFN y TALEN.

#### 9-11. Detección de polimorfismos o variaciones usando RGEN

A continuación, se diseñó y probó un nuevo RGEN que reconocía un locus altamente polimórfico, HLA-B, que codificaba el antígeno leucocitario humano B (a.k.a. Proteína MHC clase I) (Fig. 28). Las células HeLa se transfectaron con plásmidos RGEN, y el ADN genómico se sometió a análisis T7E1 y RGEN-RFLP en paralelo. T7E1 produjo bandas positivas falsas que resultaron de polimorfismos de secuencia cerca del sitio diana (Fig. 25c). Como se esperaba, sin embargo, el mismo RGEN utilizado para la interrupción génica escindió productos de PCR de células naturales completamente, pero los de las células transfectadas con RGEN parcialmente, lo que indica la presencia de inserciones-deleciones inducidas por RGEN en el sitio diana. Este resultado muestra que el análisis RGEN-RFLP tiene una clara ventaja sobre el ensayo T7E1, especialmente cuando no se sabe si los genes diana tienen polimorfismos o variaciones en las células de interés.

#### 9-12. Detección de mutaciones recurrentes encontradas en el cáncer y polimorfismos naturales a través del análisis RGEN-RFLP

El análisis RGEN-RFLP tiene aplicaciones más allá del genotipado de mutaciones inducidas por nucleasa modificada genéticamente. Se intentó usar el genotipado de RGEN para detectar mutaciones recurrentes que se encuentran en el cáncer y polimorfismos naturales. Se prefirió la línea celular de cáncer colorrectal humano, HCT116, que lleva una deleción de 3 pb con ganancia de función en el gen oncogénico *CTNNB1* que codifica la beta-catenina. Los productos de PCR amplificados a partir del ADN genómico de HCT116 se escindieron parcialmente tanto por RGEN específicos naturales como por específicos de mutante, en línea con el genotipo heterocigótico en células HCT116 (figura 29a). En marcado contraste, los productos de PCR amplificados a partir de ADN de células HeLa que albergaban solo alelos naturales fueron digeridos completamente por el RGEN naturales y no se escindieron en absoluto por el RGEN específico de la mutación.

También se observó que las células HEK293 albergaban la deleción de 32 pb (del32) en el gen *CCR5*, que codifica un correceptor esencial de la infección por VIH: los portadores del32 *CCR5* homocigóticos son inmunes a la infección por VIH. Se diseñó un RGEN específico para el alelo del32 y el otro para el alelo natural. Como se esperaba, el RGEN natural escindió los productos de PCR obtenidos de las células K562, SKBR3 o HeLa (utilizados como controles naturales) por completo, pero los de las células HEK293 parcialmente (Fig. 30a), confirmando la presencia de los alelos del32 no escindibles en células HEK293. Inesperadamente, sin embargo, el RGEN específico de del32 escindió los productos de PCR de las células naturales tan eficientemente como los de las células HEK293. Curiosamente, este RGEN tenía un sitio fuera de la diana con una falta de coincidencia de una sola base inmediatamente cadena abajo del sitio en la diana (Fig. 30). Estos resultados sugieren que los RGEN pueden usarse para detectar inserciones-deleciones de origen natural pero no pueden distinguir secuencias con polimorfismos de un solo nucleótido o mutaciones puntuales debido a sus efectos fuera de la diana.

Para genotipificar las variaciones oncogénicas de un solo nucleótido usando RGEN, se atenuó la actividad de RGEN empleando un ARN guía mal emparejado de una única base en lugar de un ARN perfectamente combinado. Los RGEN que contenían el ARN guía perfectamente combinado específico para la secuencia natural o la secuencia mutante escindieron ambas secuencias (Figuras 31a y 32a). Por el contrario, los RGEN que contenían un ARN guía mal emparejado de una única base distinguieron las dos secuencias, permitiendo el genotipado de tres mutaciones puntuales oncogénicas recurrentes en los genes *KRAS*, *PIK3CA* e *IDH1* en líneas de células cancerosas humanas (Fig. 29b y Figs. 33a,b). Además, se pudieron detectar mutaciones puntuales en los genes *BRAF* y *NRAS* utilizando RGEN que reconocía la secuencia NAG PAM (figuras 33c,d). Se cree que se puede usar RGEN-RFLP para genotipar casi todas, si no todas, las mutaciones o polimorfismos en el ser humano y otros genomas.

Los datos anteriores proponen que RGEN proporciona una plataforma para usar análisis de RFLP simples y robustos para diversas variaciones de secuencia. Con una gran flexibilidad en la secuencia diana de reprogramación, los RGEN pueden usarse para detectar diversas variaciones genéticas (variaciones de nucleótidos únicos, pequeñas inserciones/deleciones, variaciones estructurales) tales como mutaciones recurrentes relacionadas con la enfermedad, genotipos relacionados con la respuesta del fármaco por un paciente y también mutaciones inducidas por nucleasas modificadas en las células. En esta invención, se utilizó el genotipado RGEN para detectar mutaciones inducidas por nucleasas modificadas en células y animales. En principio, también se podrían usar RGEN que detecten y rompan específicamente las variaciones y mutaciones que se producen de manera natural.

10

#### Referencias

1. M. Jinek et al., *Science* 337, 816 (17 de agosto de 2012).
2. H. Kim, E. Um, S.R. Cho, C. Jung, J.S. Kim, *Nat Methods* 8, 941 (noviembre de 2011).
3. H. J. Kim, H. J. Lee, H. Kim, S. W. Cho, J. S. Kim, *Genome Res* 19, 1279 (julio de 2009).
- 15 4. E. E. Perez et al., *Nat Biotechnol* 26, 808 (julio de 2008).
5. J. C. Miller et al., *Nat Biotechnol* 29, 143 (febrero de 2011).
6. C. Mussolino et al., *Nucleic Acids Res* 39, 9283 (noviembre de 2011).
7. J. Cohen, *Science* 332, 784 (13 de mayo de 2011).
8. V. Pattanayak, C. L. Ramirez, J. K. Joung, D. R. Liu, *Nat Methods* 8, 765 (Sep, 2011).
- 20 9. R. Gabriel et al., *Nat Biotechnol* 29, 816 (septiembre de 2011).
10. E. Kim et al., *Genome Res*, (20 de abril de 2012).
11. H. J. Lee, J. Kweon, E. Kim, S. Kim, J. S. Kim, *Genome Res* 22, 539 (marzo de 2012).
12. H. J. Lee, E. Kim, J. S. Kim, *Genome Res* 20, 81 (enero de 2010).
- 25 13. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Publicación en línea de *Nat Biotech advance* (2013).

# ES 2 690 386 T3

## LISTADO DE SECUENCIAS

<110> TOOLGEN INCORPORATED

<120> Composición para escindir un ADN diana que comprende un ARN guía específico para el ADN diana y el ácido nucleico que codifica la proteína Cas o la propia proteína Cas, y sus usos

5

<130> OPA13141PCT

<150> US 61/717,324

<151> 23-10-2012

10

<150> US 61/803,599

<151> 20-03-2013

15

<150> US 61/837,481

<151> 20-06-2013

<160> 111

20

<170> KopatentIn 2.0

<210> 1

<211> 4107

<212> ADN

<213> Secuencia Artificial

25

<220>

<223> Secuencia que codifica Cas9

<400> 1

atggacaaga agtacagcat cggcctggac atcggtagca acagcgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gccagcaag aagttcaagg tgctgggcaa caccgaccgc 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag 180

gccacccgcc tgaagcgcac cgcccgcgc cgctacacc gccgcaagaa ccgcatctgc 240

tacctgcagg agatcttcag caacgagatg gccaaaggtg acgacagctt cttccaccgc 300

ctggaggaga gcttctcgtt ggaggaggac aagaagcac agcgccacc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcgcaag 420

aagctggtgg acagcaccga caaggccgac ctgcgcctga tctacctggc cctggccacc 480

atgatcaagt tccgcggcca cttcctgatc gaggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg ccgcctgag caagagccgc 660

cgctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgcgc tgagcctggg cctgaccccc aacttaaga gcaacttcga cctggccgag 780

gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtagcg cgacctgttc ctggccgcca agaacctgag cgacgccatc 900

30

ctgctgagcg acatcctgcg cgtgaacacc gagatcacca aggccccct gagcgccagc 960

ES 2 690 386 T3

atgatcaagc gctacgacga gcaccaccag gacctgacct tgctgaaggc cctggtgctgc 1020  
 cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc 1080  
 ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg 1140  
 gagaagatgg acggcaccga ggagctgctg gtgaagetga accgcgagga cctgctgctgc 1200  
 aagcagcgca ccttcgacaa cggcagcatc cccaccaga tccacctggg cgagctgcac 1260  
 gccatcctgc gccgccagga ggacttctac cccttcctga aggacaaccg cgagaagatc 1320  
 gagaagatcc tgaccttccg catccctac tacgtgggcc ccctggcccg cggcaacagc 1380  
 cgcttcgctt ggatgacctg caagagcgag gagaccatca ccccctgaa cttcgaggag 1440  
 gtggtggaca agggcgccag cggccagagc ttcatcgagc gcatgaccaa cttcgacaag 1500  
 aacctgccc aagagaaggt gctgcccag cacagcctgc tgtacgagta cttcaccctg 1560  
 tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcgcaagcc cgccttcctg 1620  
 agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg caaggtgacc 1680  
 gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc 1740  
 agcggcgtgg aggaccgctt caacgccagc ctgggcacct accacgacct gctgaagatc 1800  
 atcaaggaca aggacttctt ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860  
 ctgacctga ccctgttca ggaccgagc atgatcgagg agcgcctgaa gacctacgcc 1920  
 cacctgttcg acgacaaggt gatgaagcag ctgaagcgcc gccgctacac cggctggggc 1980  
 cgcctgagcc gcaagcttat caacggcatc cgcgacaagc agagcggcaa gaccatcctg 2040  
 gacttctga agagcgacgg cttcgccaac cgcaacttca tgcagctgat ccacgacgac 2100  
 agcctgacct tcaaggagga catccagaag gccaggtga cgggccaggg cgacagcctg 2160  
 cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc 2220  
 gtgaaggtgg tggacgagct ggtgaaggtg atgggccgcc acaagcccga gaacatcgtg 2280  
 atcgagatgg cccgcgagaa ccagaccacc cagaagggcc agaagaacag ccgcgagcgc 2340  
 atgaagcgca tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcaccac 2400  
 gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgc 2460  
 gacatgtacg tggaccagga gctggacatc aaccgcctga gcgactacga cgtggaccac 2520  
 atcgtgcccc agagcttctt gaaggacgac agcatcgaca acaaggtgct gaccgcagc 2580  
 gacaagaacc gcggcaagag cgacaacgtg cccagcgagg aggtggtgaa gaagatgaag 2640  
 aactactgac gccagctgct gaacgccaa ctgatcacc agcgcgaagt cgacaacctg 2700  
 accaaggccg agcgcggcgg cctgagcgag ctggacaagg cggcttcat caagcgcag 2760  
 ctggtggaga cccgccagat caccaagcac gtggcccaga tcctggacag ccgcatgaac 2820  
 accaagtacg acgagaacga caagctgac cgcgaggtga aggtgatcac cctgaagagc 2880

ES 2 690 386 T3

aagctggtga gcgacttccg caaggacttc cagttctaca aggtgcgcgga gatcaacaac 2940  
 taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000  
 taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcgcaag 3060  
 atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120  
 aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccgcaagcgc 3180  
 cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgcgacttc 3240  
 gccaccgtgc gcaaggtgct gagcatgcc caggtgaaca tcgtgaagaa gaccgaggtg 3300  
 cagaccggcg gcttcagcaa ggagagcatc ctgccaagc gcaacagcga caagctgatc 3360  
 gcccgcaaga aggactggga cccaagaag tacggcggct tcgacagccc caccgtggcc 3420  
 tacagcgtgc tgggtggtggc caaggtggag aagggaaga gcaagaagct gaagagcgtg 3480  
 aaggagctgc tgggcatcac catcatggag cgcagcagct tcgagaagaa ccccatcgac 3540  
 ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgccaag 3600  
 tacagcctgt tcgagctgga gaacggccgc aagcgcagc tggccagcgc cggcgagctg 3660  
 cagaagggca acgagctggc cctgccagc aagtacgtga acttcctgta cctggccagc 3720  
 cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780  
 cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgcgtg 3840  
 atcctggccg acgccaacct ggacaaggtg ctgagcgcct acaacaagca ccgcgacaag 3900  
 cccatccgcg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960  
 cccgccgct tcaagtactt cgacaccacc atcgaccgca agcgctacac cagcaccaag 4020  
 gaggtgctgg acgccaccct gatccaccag agcatcaccg gtctgtacga gacccgcatc 4080  
 gacctgagcc agctgggfcg cgactaa 4107

5 <210> 2  
 <211> 21  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> marca de péptido

<400> 2  
 Gly Gly Ser Gly Pro Pro Lys Lys Lys Arg Lys Val Tyr Pro Tyr Asp  
 1 5 10 15

Val Pro Asp Tyr Ala  
 20

15 <210> 3  
 <211> 34  
 <212> ADN  
 <213> Secuencia Artificial

20 <220>  
 <223> Cebador F para CCR5

<400> 3  
 aatcatgac atcaattatt atacatcgga ggag 34

25 <210> 4  
 <211> 34

<212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> Cebador R para CCR5  
  
 <400> 4  
 gatcctcctc cgatgtataa taattgatgt catg 34  
  
 10 <210> 5  
 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
  
 15 <220>  
 <223> Cebador F1 para CCR5  
  
 <400> 5  
 ctccatggtg ctatagagca 20  
 20  
 <210> 6  
 <211> 21  
 <212> ADN  
 <213> Secuencia Artificial  
 25  
 <220>  
 <223> Cebador F2 para CCR5  
  
 <400> 6  
 30 gagccaagct ctccatctag t 21  
  
 <210> 7  
 <211> 20  
 <212> ADN  
 35 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador R para CCR5  
  
 <400> 7  
 40 gccctgtcaa gagttgacac 20  
  
 <210> 8  
 <211> 20  
 45 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador F1 para C4BPB  
 50  
 <400> 8  
 tatttgctg gttgaaagg 20  
  
 <210> 9  
 55 <211> 24  
 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 60 <223> Cebador R1 para C4BPB  
  
 <400> 9  
 aaagtcatga aataaacaca ccca 24  
  
 65 <210> 10  
 <211> 24

<212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> Cebador F2 para C4BPB  
  
 <400> 10  
 ctgcattgat atggtagtagtac catg 24  
  
 10 <210> 11  
 <211> 21  
 <212> ADN  
 <213> Secuencia Artificial  
  
 15 <220>  
 <223> Cebador R2 para C4BPB  
  
 <400> 11  
 gctgttcatt gcaatggaat g 21  
 20  
 <210> 12  
 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
 25  
 <220>  
 <223> Cebador F1 para ADCY5  
  
 <400> 12  
 30 gctcccacct tagtgctctg 20  
  
 <210> 13  
 <211> 20  
 <212> ADN  
 35 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador R1 para ADCY5  
  
 <400> 13  
 40 ggtggcagga acctgtatgt 20  
  
 <210> 14  
 <211> 21  
 45 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador F2 para ADCY5  
 50  
 <400> 14  
 gtcattggcc agagatgtgg a 21  
  
 <210> 15  
 55 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 60 <223> Cebador R2 para ADCY5  
  
 <400> 15  
 gtcccatgac aggcgtgtapt 20  
  
 65 <210> 16  
 <211> 20

<212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> Cebador F para KCNJ6  
  
 <400> 16  
 gcctggccaa gtttcagtta 20  
  
 10 <210> 17  
 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
  
 15 <220>  
 <223> Cebador R1 para KCNJ6  
  
 <400> 17  
 20 tggagccatt ggttgcatc 20  
  
 <210> 18  
 <211> 22  
 <212> ADN  
 <213> Secuencia Artificial  
 25  
 <220>  
 <223> Cebador R2 para KCNJ6  
  
 <400> 18  
 30 ccagaactaa gccgttctg ac 22  
  
 <210> 19  
 <211> 20  
 <212> ADN  
 35 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador F1 para CNTNAP2  
  
 <400> 19  
 40 atcaccgaca accagttcc 20  
  
 <210> 20  
 <211> 20  
 45 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador F2 para CNTNAP2  
 50  
 <400> 20  
 tgcagtgcag actctttcca 20  
  
 <210> 21  
 55 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 60 <223> Cebador R para CNTNAP2  
  
 <400> 21  
 aaggacacag ggcaactgaa 20  
  
 65 <210> 22  
 <211> 20

<212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> Cebador F1 para N/A Chr. 5  
  
 <400> 22  
 tgtggaacga gtggtgacag 20  
  
 10 <210> 23  
 <211> 22  
 <212> ADN  
 <213> Secuencia Artificial  
  
 15 <220>  
 <223> Cebador R1 para N/A Chr. 5  
  
 <400> 23  
 20 gctggattag gaggcaggat tc 22  
  
 <210> 24  
 <211> 22  
 <212> ADN  
 <213> Secuencia Artificial  
 25  
 <220>  
 <223> Cebador F2 para N/A Chr. 5  
  
 <400> 24  
 30 gtgctgagaa cgcttcatag ag 22  
  
 <210> 25  
 <211> 23  
 <212> ADN  
 35 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador R2 para N/A Chr. 5  
  
 <400> 25  
 40 ggaccaaacc acattctct cac 23  
  
 <210> 26  
 <211> 20  
 45 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador F para delección  
 50  
 <400> 26  
 ccacatctcg ttctcggtt 20  
  
 <210> 27  
 55 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 60 <223> Cebador R para delección  
  
 <400> 27  
 tcacaagccc acagatatt 20  
  
 65 <210> 28  
 <211> 105

ES 2 690 386 T3

<212> ARN  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> ARNsg para CCR5  
  
 <400> 28  
 ggugacauca auuauuauac auguuuuaga gcuagaaaua gcaaguuaaa auaaggcuag 60  
 uccguuauca acuugaaaaa guggcaccga gucggugcuu uuuuu 105  
  
 10 <210> 29  
 <211> 44  
 <212> ARN  
 <213> Secuencia Artificial  
  
 15 <220>  
 <223> ARNcr para CCR5  
  
 <400> 29  
 20 ggugacauca auuauuauac auguuuuaga gcuaugcugu uuug 44  
  
 <210> 30  
 <211> 86  
 <212> ARN  
 <213> Secuencia Artificial  
 25  
 <220>  
 <223> ARNtracr para CCR5  
  
 <400> 30  
 30 ggaaccuauuc aaaacagcau agcaaguuaa auaaggcuca guccguuauc aacuugaaaa 60  
 aguggcaccg agucggugcu uuuuuu 86  
  
 <210> 31  
 <211> 86  
 <212> ADN  
 35 <213> Secuencia Artificial  
  
 <220>  
 <223> Foxn1 n°1 ARNsg  
 40 <400> 31  
 gaaattaata cgactcacta taggcagtct gacgtcacac ttccgtttta gagctagaaa 60  
 tagcaagtta aaataaggct agtccg 86  
  
 45 <210> 32  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial  
  
 50 <220>  
 <223> Foxn1 n°2 ARNsg  
  
 <400> 32  
 gaaattaata cgactcacta taggacttcc aggctccacc cgacgtttta gagctagaaa 60  
 tagcaagtta aaataaggct agtccg 86  
 55  
 <210> 33  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial  
 60

ES 2 690 386 T3

<220>  
 <223> Foxn1 n°3 ARNsg

<400> 33  
 gaaattaata cgactcacta taggccaggc tccacccgac tggagtttta gagctagaaa 60

5 tagcaagtta aaataaggct agtccg 86

<210> 34  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial

10 <220>  
 <223> Foxn1 n°4 ARNsg

15 <400> 34  
 gaaattaata cgactcacta taggactgga gggcgaacc caaggtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 35  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial

20 <220>  
 <223> Foxn1 n°5 ARNsg

25 <400> 35  
 gaaattaata cgactcacta taggaccca aggggacctc atgcgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 36  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial

30 <220>  
 <223> Prkdc n°1 ARNsg

35 <400> 36  
 gaaattaata cgactcacta taggttagtt tttccagag acttgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

40 <210> 37  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial

45 <220>  
 <223> Prkdc n°2 ARNsg

<400> 37  
 gaaattaata cgactcacta taggttggtt tgcttggtt tatcgtttta gagctagaaa 60

50 tagcaagtta aaataaggct agtccg 86

<210> 38  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial

55 <220>  
 <223> Prkdc n°3 ARNsg

ES 2 690 386 T3

<400> 38  
 gaaattaata cgactcacta taggcacaag caaaccaaag tctcgtttta gagctagaaa 60  
 tagcaagtta aaataaggct agtccg 86

5 <210> 39  
 <211> 86  
 <212> ADN  
 <213> Secuencia Artificial

10 <220>  
 <223> Prkdc n°4 ARNsg  
 <400> 39  
 gaaattaata cgactcacta taggcctcaa tgctaagcga cttcgtttta gagctagaaa 60  
 tagcaagtta aaataaggct agtccg 86

15 <210> 40  
 <211> 29  
 <212> ADN  
 <213> Secuencia Artificial

20 <220>  
 <223> Cebador F1 para Foxn1

25 <400> 40  
 gtctgtctat catctctcc ctctctcc 29  
 <210> 41  
 <211> 25  
 <212> ADN  
 <213> Secuencia Artificial

30 <220>  
 <223> Cebador F2 para Foxn1

35 <400> 41  
 tccctaattcc gatggctagc tccag 25  
 <210> 42  
 <211> 23  
 <212> ADN  
 <213> Secuencia Artificial

40 <220>  
 <223> Cebador R1 para Foxn1

45 <400> 42  
 acgagcagct gaagtagca tgc 23  
 <210> 43  
 <211> 32  
 <212> ADN  
 <213> Secuencia Artificial

50 <220>  
 <223> Cebador R2 para Foxn1

55 <400> 43  
 ctactcaatg ctcttagagc taccaggctt gc 32

60 <210> 44  
 <211> 20  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador F para Prkdc  
 5  
 <400> 44  
 gactgtgtg gggaggccg 20  
 <210> 45  
 10 <211> 24  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador F2 para Prkdc  
 <400> 45  
 gggaggccg aaagtctat ttg 24  
 20 <210> 46  
 <211> 28  
 <212> ADN  
 <213> Secuencia Artificial  
 25 <220>  
 <223> Cebador R1 para Prkdc  
 <400> 46  
 30 cctgaagact gaagttgca gaagtgag 28  
 <210> 47  
 <211> 27  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador R2 para Prkdc  
 <400> 47  
 40 cttaggct tcttctac aatcacg 27  
 <210> 48  
 <211> 38  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador F para Foxn1  
 50 <400> 48  
 ctcggtgtg agccctgacc tcggtgtga gcctgac 38  
 <210> 49  
 <211> 21  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador R para Foxn1  
 60 <400> 49  
 agactggcct ggaactcaca g 21  
 <210> 50  
 65 <211> 23  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador F para Foxn1  
 5  
 <400> 50  
 cactaaagcc tgtcaggaag ccg 23  
 <210> 51  
 10 <211> 21  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador R para Foxn1  
 <400> 51  
 ctgtggagag cacacagcag c 21  
 <210> 52  
 20 <211> 19  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 25 <223> Cebador F para Foxn1  
 <400> 52  
 30 gctgcgacct gagaccatg 19  
 <210> 53  
 <211> 26  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador R para Foxn1  
 <400> 53  
 40 ctcaatggc ttctgctta ggctac 26  
 <210> 54  
 <211> 23  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador F para Foxn1  
 <400> 54  
 50 ggtcagatg aggccatcct ttc 23  
 <210> 55  
 <211> 24  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador R para Foxn1  
 60 <400> 55  
 cctgatctgc aggcttaacc cttg 24  
 <210> 56  
 65 <211> 22  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador F para Prkdc  
 5  
 <400> 56  
 ctcacctgca catcacatgt gg 22  
 <210> 57  
 10 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador R para Prkdc  
 <400> 57  
 ggcatccacc ctatggggtc 20  
 <210> 58  
 20 <211> 25  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 25 <223> Cebador F para Prkdc  
 <400> 58  
 30 gccttgacct agagcttaa gagcc 25  
 <210> 59  
 <211> 25  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador R para Prkdc  
 <400> 59  
 40 ggtcttgta gcaggaagga cactg 25  
 <210> 60  
 <211> 27  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador F para Prkdc  
 50 <400> 60  
 aaaactctgc ttgatgggat atgtggg 27  
 <210> 61  
 <211> 26  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador R para Prkdc  
 60 <400> 61  
 ctctcactgg ttatctgtgc tccttc 26  
 <210> 62  
 65 <211> 23  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador F para Prkdc  
 5 <400> 62  
 ggatcaatag gtggtggggg atg 23  
 <210> 63  
 10 <211> 27  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador R para Prkdc  
 <400> 63  
 gtgaatgaca caatgtgaca gcttcag 27  
 20 <210> 64  
 <211> 28  
 <212> ADN  
 <213> Secuencia Artificial  
 25 <220>  
 <223> Cebador F para Prkdc  
 <400> 64  
 30 cacaagacag acctctcaac attcagtc 28  
 <210> 65  
 <211> 32  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador R para Prkdc  
 <400> 65  
 40 gtgcatgcat ataatccatt ctgattgctc tc 32  
 <210> 66  
 <211> 17  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador F1 para Prkdc  
 50 <400> 66  
 gggaggcaga ggcaggt 17  
 <210> 67  
 <211> 23  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador F2 para Prkdc  
 60 <400> 67  
 ggatctctgt gagtttgagg cca 23  
 <210> 68  
 65 <211> 24  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador R1 para Prkdc  
 5 <400> 68  
 gctccagaac tcactcttag gctc 24  
 <210> 69  
 10 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador para Foxn1  
 <400> 69  
 ctactccctc cgcagtctga 20  
 <210> 70  
 20 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 25 <223> Cebador para Foxn1  
 <400> 70  
 30 ccaggcctag gttccagga 20  
 <210> 71  
 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador para Prkdc  
 <400> 71  
 40 cccagcatt gcagatttc 20  
 <210> 72  
 <211> 23  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador para Prkdc  
 <400> 72  
 50 agggcttctt ctctacaatc acg 23  
 <210> 73  
 <211> 86  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> BR11 diana 1  
 60 <400> 73  
 gaaattaata cgactcacta taggtttgaa agatggaagc gggggtttta gagctagaaa 60  
 tagcaagtta aaataaggct agtccg 86  
 <210> 74

ES 2 690 386 T3

<211> 86  
 <212> ADN  
 <213> Secuencia Artificial

5 <220>  
 <223> BRI1 diana 2

<400> 74  
 gaaattaata cgactcacta taggtgaaac taaactggtc cacagtttta gagctagaaa 60  
 tagcaagtta aaataaggct agtccg 86

10 <210> 75  
 <211> 64  
 <212> ADN  
 <213> Secuencia Artificial

15 <220>  
 <223> Universal

<400> 75  
 aaaaaagcac cgactcggcg ccactttttc aagttgataa cggactagcc ttattttaac 60

20 ttgc 64

<210> 76  
 <211> 65  
 <212> ADN  
 <213> Secuencia Artificial

25 <220>  
 <223> Plantillas para ARNcr

30 <400> 76  
 gaaattaata cgactcacta taggnnnnnn nnnnnnnnnn nnnngtttta gagctatgct 60  
 gtttt 65

<210> 77  
 <211> 67  
 <212> ADN  
 <213> Secuencia Artificial

35 <220>  
 <223> ARNtracr

40 <400> 77  
 gaaattaata cgactcacta taggaaccat tcaaaacagc atagcaagtt aaaataaggc 60  
 tagtccg 67

<210> 78  
 <211> 69  
 <212> ADN  
 <213> Secuencia Artificial

45 <220>  
 <223> ARNtracr

50 <400> 78  
 aaaaaaagca ccgactcggc gccacttttt caagttgata acggactagc cttattttaa 60  
 cttgctatg 69

55 <210> 79  
 <211> 20  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 5  
 <400> 79  
 ctccatggtg ctatagagca 20  
 <210> 80  
 10 <211> 21  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador  
 <400> 80  
 gagccaagct ctccatctag t 21  
 20 <210> 81  
 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
 25 <220>  
 <223> Cebador  
 <400> 81  
 30 gccctgtcaa gagtgacac 20  
 <210> 82  
 <211> 22  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador  
 <400> 82  
 40 gcacagggtg gaacaagatg ga 22  
 <210> 83  
 <211> 24  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 50 <400> 83  
 gccagggtacc tatcgattgt cagg 24  
 <210> 84  
 <211> 21  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 60 <400> 84  
 gagccaagct ctccatctag t 21  
 <210> 85  
 65 <211> 20  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 5  
 <400> 85  
 actctgactg ggtcaccagc 20  
 <210> 86  
 10 <211> 20  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador  
 <400> 86  
 tatttgctg gttgaaagg 20  
 20 <210> 87  
 <211> 24  
 <212> ADN  
 <213> Secuencia Artificial  
 25 <220>  
 <223> Cebador  
 <400> 87  
 30 aaagtcatga aataaacaca ccca 24  
 <210> 88  
 <211> 24  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador  
 <400> 88  
 40 ctgcattgat atggtagtac catg 24  
 <210> 89  
 <211> 21  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 50 <400> 89  
 gctgttcatt gcaatggaat g 21  
 <210> 90  
 <211> 22  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 60 <400> 90  
 atggagtgg acatggccat gg 22  
 <210> 91  
 65 <211> 28  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 5 <400> 91  
 actcactatc cacagttcag catttacc 28  
 <210> 92  
 10 <211> 23  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador  
 <400> 92  
 tggagatagc tgtcagcaac ttt 23  
 <210> 93  
 20 <211> 29  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 25 <223> Cebador  
 <400> 93  
 30 caacaaagca aaggtaaagt tgtaatatag 29  
 <210> 94  
 <211> 25  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador  
 <400> 94  
 40 ggttcagga gatgtgttac aaggc 25  
 <210> 95  
 <211> 27  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 <400> 95  
 50 gattgtgcaa tcctatgca atcggtc 27  
 <210> 96  
 <211> 25  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 60 <400> 96  
 cactgggtac ttaatctgta gcctc 25  
 <210> 97  
 65 <211> 23  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 5 <400> 97  
 ggtccaagt cattcccagt agc 23  
 <210> 98  
 10 <211> 30  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Cebador  
 <400> 98  
 catcactgca gttgtaggt ataactatcc 30  
 <210> 99  
 20 <211> 26  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 25 <223> Cebador  
 <400> 99  
 30 ttgaaaacca cagatctggt tgaacc 26  
 <210> 100  
 <211> 22  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Cebador  
 <400> 100  
 40 ggagtgcaa gagaatatct gg 22  
 <210> 101  
 <211> 32  
 <212> ADN  
 45 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 <400> 101  
 50 ctgaaactgg ttcaaaata ttcgtttaa gg 32  
 <210> 102  
 <211> 22  
 55 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 60 <400> 102  
 gctctgatg ccctgtagta gg 22  
 <210> 103  
 65 <211> 22  
 <212> ADN

# ES 2 690 386 T3

<213> Secuencia Artificial  
 <220>  
 <223> Cebador  
 5 <400> 103  
 tttgcatctg accttacctt tg 22  
 <210> 104  
 10 <211> 23  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Secuencia diana de RGEN  
 <400> 104  
 aatgaccact acatcctcaa ggg 23  
 <210> 105  
 20 <211> 23  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 25 <223> Secuencia diana de RGEN  
 <400> 105  
 30 agatgatgtc tcatcatcag agg 23  
 <210> 106  
 <211> 4170  
 <212> ADN  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Secuencia que codifica Cas9 en p3s-Cas9HC (humanizado, marcado en el extremo C, experimentos con células humanas)  
 40 <400> 106  
 atggacaaga agtacagcat cggcctggac atcggtagca acagcgtggg ctgggccgtg 60  
 atcaccgacg agtacaaggt gccagcaag aagttcaagg tgctgggcaa caccgaccgc 120

ES 2 690 386 T3

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgcccag 180  
gccacccgcc tgaagcgcac cgcccgcgcg cgctacaccc gccgcaagaa ccgcatctgc 240  
tacctgcagg agatcttcag caacgagatg gccaaagtgg acgacagctt cttccaccgc 300  
ctggaggaga gcttcctggt ggaggaggac aagaagcacg agcgccaccc catcttcggc 360  
aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcgcaag 420  
aagctggtgg acagcaccga caaggccgac ctgcgctga tctacctggc cctggcccac 480  
atgatcaagt tccgcggcca ettcctgate gaggcgacc tgaacccga caacagcgac 540  
gtggacaagc tgttcacca gctggtgcag acctacaacc agctgttcga ggagaacccc 600  
atcaacgccg gcgcgctgga cgccaaggcc atcctgagcg cccgcctgag caagagccgc 660  
cgcctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720  
ctgatcgccc tgagcctggg cctgaccccc aacttcaaga gcaacttcga cctggcccag 780  
gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggaaa cctgctggcc 840  
cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgag cgacgccatc 900  
ctgctgagcg acatcctgcg cgtgaacacc gagatcacca aggccccct gagcggcagc 960  
atgatcaagc gctacgacga gcaccaccag gacctgacc tgctgaaggc cctggtgcgc 1020  
cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc 1080  
ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg 1140  
gagaagatgg acggcaccga ggagctgctg gtgaagctga accgcgagga cctgctgcgc 1200  
aagcagcgca ccttcgaaa cggcagcatc cccaccaga tccacctggg cgagctgcac 1260  
gccatcctgc gcccccagga ggacttctac cccttcctga aggacaaccg cgagaagatc 1320  
gagaagatcc tgaccttcg catcccctac tacgtgggcc ccctggcccg cggcaacagc 1380  
cgcttcgctt ggatgacccg caagagcgag gagaccatca cccctggaa cttcgaggag 1440  
gtggtggaca agggcgccag cgcccagagc ttcctcagag ccatgaccaa cttcgacaag 1500  
aacctgccc aagagaaggt gctgcccagg cacagcctgc tgtacgagta cttcacctg 1560  
tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcgcaagcc cgccttcctg 1620  
agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg caaggtgacc 1680  
gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc 1740  
agcggcgtgg aggaccgctt caacgccagc ctgggcacct accacgacct gctgaagatc 1800  
atcaaggaca aggacttctt ggacaacgag gagaacgag acatcctgga ggacatcgtg 1860  
ctgaccctga ccctgttcga ggaccgag atgatcgagg agcgcctgaa gacctacgcc 1920  
cacctgttc acgacaaggt gatgaagcag ctgaagcgcc gccgctacac cggctggggc 1980  
cgcctgagcc gcaagcttat caacggcatc cgcgacaagc agagcggcaa gaccatcctg 2040

ES 2 690 386 T3

gacttctga agagcgacgg cttcgccaac cgcaacttca tgcagctgat ccacgacgac 2100  
agcctgacct tcaaggagga catccagaag gccaggtga gcggccaggg cgacagcctg 2160  
cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc 2220  
gtgaaggtgg tggacgagct ggtgaaggtg atgggccgcc acaagcccga gaacatcgtg 2280  
atcgagatgg cccgcgagaa ccagaccacc cagaagggcc agaagaacag ccgcgagcgc 2340  
atgaagcgca tcgaggagg catcaaggag ctgggcagcc agatcctgaa ggagcaccac 2400  
gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgc 2460  
gacatgtacg tggaccagga gctggacatc aaccgcctga gcgactacga cgtggaccac 2520  
atcgtgcccc agagcttctt gaaggacgac agcatcgaca acaaggtgct gacccgcagc 2580  
gacaagaacc gcggcaagag cgacaacgtg cccagcgagg agtggtgtaa gaagatgaag 2640  
aactactggc gccagctgct gaacgccaa ctgatcacc agcgcaagtt cgacaacctg 2700  
accaaggccg agcgcgccgg cctgagcgag ctggacaagg ccggcttcat caagcgccag 2760  
ctggtggaga cccgccagat caccaagcac gtggcccaga tcctggacag ccgcatgaac 2820  
accaagtacg acgagaacga caagctgac cgcgaggtga agtgatcac cctgaagagc 2880  
aagctggtga gcgacttccg caaggacttc cagttctaca agtgcgcgca gatcaacaac 2940  
taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000  
taccccaagc tggagagcga gttcgtgtac ggcgactaca agtggtacga cgtgcgcaag 3060  
atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120  
aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccgcaagcgc 3180  
cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgcgacttc 3240  
gccaccgtgc gcaaggtgct gagcatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300  
cagaccggcg gcttcagcaa ggagagcatc ctgccaaagc gcaacagcga caagctgac 3360  
gcccgcaaga aggactggga cccaagaag tacggcggct tcgacagccc caccgtggcc 3420  
tacagcgtgc tgggtggtgg caaggtggag aagggcaaga gcaagaagct gaagagcgtg 3480  
aaggagctgc tgggcatcac catcatggag cgcagcagct tcgagaagaa ccccatcgac 3540  
ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccagg 3600  
tacagcctgt tcgagctgga gaacggccgc aagcgcatgc tggccagcgc cggcgagctg 3660  
cagaagggca acgagctggc cctgccagc aagtacgtga acttctctgta cctggccagc 3720  
cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780  
cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgcgtg 3840  
atcctggccg acgccaacct ggacaaggtg ctgagcgcct acaacaagca ccgcgacaag 3900  
cccatccgcg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960  
cccgccgctt tcaagtactt cgacaccacc atcgaccgca agcgctacac cagcaccaag 4020  
gaggtgctgg acgccaccct gatccaccag agcatcaccg gtctgtacga gacccgcatc 4080  
gacctgagcc agctgggccc cgacggcggc tccggacctc caaagaaaa gagaaaagta 4140  
taccctacg acgtgcccga ctacgcctaa 4170

ES 2 690 386 T3

<211> 4194  
 <212> ADN  
 <213> Secuencia Artificial

5 <220>  
 <223> Secuencia que codifica Cas9 en p3s-Cas9HN (codón humanizado, marcado en el extremo N (subrayado), experimentos con células humanas)

<400> 107  
 atggtgtacc cctacgacgt gcccgactac gccgaattgc ctccaaaaa gaagagaaag 60  
 gtagggatcc gaattcccgg ggaaaaaccg gacaagaagt acagcatcgg cctggacatc 120  
 ggtaccaaca gcgtgggctg ggccgtgatc accgacgagt acaaggtgcc cagcaagaag 180  
 ttcaaggtgc tgggcaacac cgaccgccac agcatcaaga agaacctgat cggcgccctg 240  
 ctgttcgaca gcggcgagac cgccgaggcc acccgccctga agcgcaccgc ccgccgccgc 300  
 tacacccgcc gcaagaaccg catctgctac ctgcaggaga tcttcagcaa cgagatggcc 360  
 aaggtggacg acagcttctt ccaccgcctg gaggagagct tcctggtgga ggaggacaag 420  
 aagcacgagc gccaccccat cttcggcaac atcgtggacg aggtggccta ccacgagaag 480  
 taccaccacca tctaccacct gcgcaagaag ctggtggaca gcaccgacaa ggccgacctg 540  
 cgcctgatct acctggccct ggcccacatg atcaagttcc gcggccactt cctgatcgag 600  
 ggcgacctga accccgacaa cagcgacgtg gacaagctgt tcatccagct ggtgcagacc 660  
 tacaaccagc tgttcgagga gaaccccatc aacgccagcg gcgtggacgc caaggccatc 720  
 ctgagcgccc gcctgagcaa gagccgccgc ctggagaacc tgatcgcca gctgcccggc 780  
 gagaagaaga acggcctggt cggcaacctg atcgccctga gcctgggcct gaccccaac 840  
 ttcaagagca acttgcacct ggccgaggac gccaaagctgc agctgagcaa ggacacctac 900  
 gacgacgacc tggacaacct gctggcccag atcggcgacc agtacgccga cctgttcctg 960  
 gccgccaaga acctgagcga cgccatcctg ctgagcgaca tcctgcgcgt gaacaccgag 1020  
 atcaccaagg cccccctgag cgccagcatg atcaagcgct acgacgagca ccaccaggac 1080  
 ctgacctgc tgaaggccct ggtgcgccag cagctgcccg agaagtacaa ggagatcttc 1140  
 ttcgaccaga gcaagaaccg ctacgccggc tacatogacg gcggcgccag ccaggaggag 1200  
 10 ttctacaagt tcatcaagcc catcctggag aagatggacg gcaccgagga gctgctggtg 1260

ES 2 690 386 T3

aagctgaacc gcgaggacct gctgcgcaag cagcgcacct tcgacaacgg cagcatcccc 1320  
caccagatcc acctgggcga gctgcacgcc atcctgcgcc gccaggagga cttctacccc 1380  
ttcctgaagg acaaccgcga gaagatcgag aagatcctga ccttccgcat cccctactac 1440  
gtgggcccc tggcccgcgg caacagccgc ttcgcctgga tgaccgcaa gagcgaggag 1500  
accatcacc cctggaactt cgaggaggtg gtggacaagg gcgccagcgc ccagagcttc 1560  
atcgagcgca tgaccaactt cgacaagaac ctgcccacag agaaggtgct gcccaagcac 1620  
agcctgctgt acgagtactt caccgtgtac aacgagctga ccaaggtgaa gtacgtgacc 1680  
gagggcatgc gcaagcccgc cttcctgagc ggcgagcaga agaaggccat cgtggacctg 1740  
ctgttcaaga ccaaccgcaa ggtgaccgtg aagcagctga aggaggacta cttcaagaag 1800  
atcgagtgct tcgacagcgt ggagatcagc ggcgtggagg accgcttcaa cgccagcctg 1860  
ggcacctacc acgacctgct gaagatcatc aaggacaagg acttctgga caacgaggag 1920  
aacgaggaca tcctggagga catcgtgctg accctgacct tgttcgagga ccgcgagatg 1980  
atcgaggagc gcctgaagac ctacgcccac ctgttcgacg acaaggtgat gaagcagctg 2040  
aagcgccgcc gctacaccgg ctggggccgc ctgagccgca agcttatcaa cggcatccgc 2100  
gacaagcaga gcggcaagac catcctggac ttcctgaaga gcgacggctt cgccaaccgc 2160  
aacttcatgc agctgatcca cgacgacagc ctgacctca aggaggacat ccagaaggcc 2220  
caggtgagcg gccagggcga cagcctgcac gagcacatcg ccaacctggc cggcagcccc 2280  
gccatcaaga agggcatcct gcagaccgtg aaggtggtgg acgagctggt gaagtgatg 2340  
ggccgccaca agcccagaaa catcgtgatc gagatggccc gcgagaacca gaccaccag 2400  
aagggccaga agaacagccg cgagcgcagc aagcgcagc aggagggcat caaggagctg 2460  
ggcagccaga tcctgaagga gcaccccgtg gagaacaccc agctgcagaa cgagaagctg 2520  
tacctgtact acctgcagaa cggccgcgac atgtacgtgg accaggagct ggacatcaac 2580  
cgcctgagcg actacgacgt ggaccacatc gtgccccaga gcttctgaa ggacgacagc 2640  
atcgacaaca aggtgctgac ccgacgcgac aagaaccgcg gcaagagcga caacgtgccc 2700  
agcgaggagg tgggtaagaa gatgaagaac tactggcgcc agctgctgaa cgccaagctg 2760  
atcaccagc gcaagttcga caacctgacc aaggccgagc gcggcggcct gagcgagctg 2820  
gacaaggccg gcttcatcaa gcgccagctg gtggagacct gccagatcac caagcacgtg 2880  
gccagatcc tggacagccg catgaacacc aagtacgagc agaacgaaa gctgatccgc 2940  
gaggtgaagg tgatcacccct gaagagcaag ctggtgagcg acttccgcaa ggacttccag 3000  
ttctacaagg tgcgcgagat caacaactac caccacgccc acgacgccta cctgaacgcc 3060  
gtggtgggca ccgccctgat caagaagtac cccaagctgg agagcgagtt cgtgtacggc 3120

ES 2 690 386 T3

gactacaagg tgtacgacgt gcgcaagatg atcgccaaga gcgagcagga gatcggcaag 3180  
gccaccgcca agtactttctt ctacagcaac atcatgaact tcttcaagac cgagatcacc 3240  
ctggccaacg gcgagatccg caagcgcccc ctgatcgaga ccaacggcga gaccggcgag 3300  
atcgtgtggg acaagggccg cgacttcgcc accgtgcgca aggtgctgag catgccccag 3360  
gtgaacatcg tgaagaagac cgaggtgcag accgggggct tcagcaagga gagcatcctg 3420  
cccaagcgca acagcgacaa gctgatcgcc cgcaagaagg actgggaccc caagaagtac 3480  
ggcggcttcg acagccccac cgtggcctac agcgtgctgg tgggggcca ggtggagaag 3540  
ggcaagagca agaagctgaa gagcgtgaag gagctgctgg gcatcaccat catggagcgc 3600  
agcagcttcg agaagaacct catcgacttc ctggaggcca agggctacaa ggaggtgaag 3660  
aaggacctga tcatcaagct gcccaagtac agcctgttcg agctggagaa cggccgcaag 3720  
cgcatgctgg ccagcgccgg cgagctgcag aagggcaacg agctggccct gccagcaag 3780  
tacgtgaact tcctgtacct ggccagccac tacgagaagc tgaagggcag ccccgaggac 3840  
aacgagcaga agcagctggt cgtggagcag cacaagcact acctggacga gatcatcgag 3900  
cagatcagcg agttcagcaa gcgcgtgatc ctggccgacg ccaacctgga caaggtgctg 3960  
agcgcctaca acaagcaccg cgacaagccc atcccgagc aggccgagaa catcatccac 4020  
ctgttcaccc tgaccaacct gggcgcccc gccgccttca agtacttcga caccaccatc 4080  
gaccgcaagc gctacaccag caccaaggag gtgctggacg ccaccctgat ccaccagagc 4140  
atcaccggtc tgtacgagac ccgcatcgac ctgagccagc tgggcgcgga ctaa 4194

<210> 108  
<211> 4107  
5 <212> ADN  
<213> Secuencia Artificial

<220>  
10 <223> Secuencia que codifica Cas9 en Streptococcus pyogenes

<400> 108  
atggataaga aataactcaat aggcttagat atcggcacia atagcgtcgg atggcggtg 60  
atcactgatg aatataaggt tccgtctaaa aagttaagg ttctgggaaa tacagaccgc 120  
cacagtatca aaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa 180  
gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt 240  
tatctacagg agatthtttc aaatgagatg gcgaaagtag atgatagttt ctttcatcga 300  
cttgaagagt cttttttggt ggaagaagac aagaagcatg aacgtcatcc tatttttggg 360  
aatatagtag atgaagttgc ttatcatgag aatatccaa ctatctatca tctgcgaaaa 420  
aaattggtag attctactga taaagcggat ttgcgcttaa tctatttggc cttagcgcac 480

ES 2 690 386 T3

atgattaagt ttcgtggtca ttttttgatt gagggagatt taaatcctga taatagtgat	540
gtggacaaac tatttatcca gttggtacaa acctacaatc aattatttga agaaaaccct	600
attaacgcaa gtggagtaga tgctaaagcg attccttctg cacgattgag taaatcaaga	660
cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggctt atttggaat	720
ctcattgctt tgtcattggg tttgaccctt aattttaaat caaattttga tttggcagaa	780
gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg	840
caaattggag atcaatatgc tgatttgttt ttggcagcta agaatttata agatgctatt	900
ttactttcag atatcctaag agtaaatact gaaataacta aggctccccct atcagcttca	960
atgattaaac gctacgatga acatcatcaa gacttgactc ttttaaaagc tttagttcga	1020
caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatatgca	1080
ggttatattg atgggggagc tagccaagaa gaattttata aatttatcaa accaatttta	1140
gaaaaaatgg atggtactga ggaattattg gtgaaactaa atcgtgaaga tttgctgcgc	1200
aagcaacgga cctttgacaa cggctctatt ccccatcaaa ttcacttggg tgagctgcat	1260
gctattttga gaagacaaga agacttttat ccatttttaa aagacaatcg tgagaagatt	1320
gaaaaaatct tgacttttcg aattccttat tatgttggtc cattggcgcg tggcaatagt	1380
cgttttgcat ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa	1440
gttgctgata aagggtcttc agctcaatca tttattgaac gcatgacaaa ctttgataaa	1500
aatcttccaa atgaaaaagt actacaaaaa catagtttgc tttatgagta ttttacggtt	1560
tataacgaat tgacaaaggt caaatatggt actgaaggaa tgcgaaaacc agcatttctt	1620
tcaggtgaac agaagaaagc cattgttgat ttactcttca aaacaaatcg aaaagtaacc	1680
gttaagcaat taaaagaaga ttatttcaaa aaaatagaat gttttgatag tgttgaaatt	1740
tcaggagttg aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt	1800
attaaagata aagatttttt ggataatgaa gaaaatgaag atatcttaga ggatattggt	1860
ttaacattga ccttatttga agatagggag atgattgagg aaagacttaa aacatatgct	1920
cacctctttg atgataaggt gatgaaacag cttaaacgtc gccgttatac tggttgggga	1980
cgtttgtctc gaaaattgat taatggtatt agggataagc aatctggcaa aacaatatta	2040
gattttttga aatcagatgg ttttgccaat cgcaatttta tgcagctgat ccatgatgat	2100
agtttgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagttta	2160
catgaacata ttgcaaattt agctggtagc cctgctatta aaaaaggtat tttacagact	2220
gtaaaagttg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aaatatcgtt	2280
attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgagagcgt	2340
atgaaacgaa tcgaagaagg tatcaaagaa ttaggaagtc agattcttaa agagcatcct	2400

ES 2 690 386 T3

gttgaaaata ctcaattgca aaatgaaaag ctctatctct attatctcca aaatggaaga 2460  
gacatgtatg tggaccaaga attagatatt aatcgtttaa gtgattatga tgtcgcacac 2520  
attgttccac aaagtttcct taaagacgat tcaatagaca ataaggctct aacgcgttct 2580  
gataaaaatc gtggtaaatc ggataacggt ccaagtgaag aagtagtcaa aaagatgaaa 2640  
aactattgga gacaacttct aaacgccaaag ttaatcactc aacgtaagtt tgataattta 2700  
acgaaagctg aacgtggaggt tttgagtgaa cttgataaag ctgggttttat caaacgccaa 2760  
ttggttgaaa ctgcgcaaat cactaagcat gtggcacaaa ttttgatag tcgcatgaat 2820  
actaaatacg atgaaaatga taaacttatt cgagagggtta aagtgattac cttaaaaatct 2880  
aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat 2940  
taccatcatg cccatgatgc gtatctaaat gccgtcgttg gaactgcttt gattaagaaa 3000  
tatccaaaac ttgaatcggg gtttgtctat ggtgattata aagtttatga tgttcgtaaa 3060  
atgattgcta agtctgagca agaaataggc aaagcaaccg caaaatattt cttttactct 3120  
aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaacgc 3180  
cctctaatacg aaactaatgg ggaaactgga gaaattgtct gggataaagg gcgagatfff 3240  
gccacagtgc gcaaagtatt gtccatgccc caagtcaata ttgtcaagaa aacagaagta 3300  
cagacaggcg gattctccaa ggagtcaatt ttaccaaaaa gaaattcggg caagcttatt 3360  
gctcgtaaaa aagactggga tccaaaaaaa tatggtggtt ttgatagtcc aacggtagct 3420  
tattcagttc tagtggttgc taagggtgga aaagggaaat cgaagaagtt aaaatccggt 3480  
aaagagttac tagggatcac aattatggaa agaagttcct ttgaaaaaaa tccgattgac 3540  
tttttagaag ctaaaggata taagggaagtt aaaaaagact taatcattaa actacctaaa 3600  
tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaatta 3660  
caaaaaggaa atgagctggc tctgccaaagc aaatatgtga attttttata tttagctagt 3720  
cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780  
cagcataagc attatttaga tgagattatt gagcaaatca gtgaattttc taagcgtggt 3840  
attttagcag atgccaatff agataaagtt cttagtgcac ataacaaca tagagacaaa 3900  
ccaatacgtg aacaagcaga aaatattatt catttattta cgttgacgaa tcttgagct 3960  
cccgtgctt ttaaatattt tgatacaaca attgatcgta aacgatatac gtctacaaaa 4020  
gaagttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080  
gatttgagtc agctaggagg tgactaa 4107

<210> 109  
<211> 1368  
5 <212> PRT  
<213> Secuencia Artificial

<220>  
10 <223> Secuencia de aminoácidos de Cas9 de S.pyogenes

<400> 109

ES 2 690 386 T3

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val  
 1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe  
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile  
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu  
 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys  
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser  
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys  
 100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr  
 115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp  
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His  
 145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro  
 165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr  
 180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala  
 195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn  
 210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn  
 225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe  
 245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp  
 260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp  
 275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp  
 290 295 300

ES 2 690 386 T3

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser  
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys  
325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe  
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser  
355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp  
370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg  
385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu  
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe  
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile  
435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp  
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu  
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr  
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser  
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys  
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln  
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr  
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp  
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly  
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp  
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr  
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala  
625 630 635 640

ES 2 690 386 T3

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr  
 645 650 655  
 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp  
 660 665 670  
 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe  
 675 680 685  
 Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe  
 690 695 700  
 Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu  
 705 710 715 720  
 His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly  
 725 730 735  
 Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly  
 740 745 750  
 Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln  
 755 760 765  
 Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile  
 770 775 780  
 Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro  
 785 790 795 800  
 Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu  
 805 810 815  
 Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg  
 820 825 830  
 Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys  
 835 840 845  
 Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg  
 850 855 860  
 Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys  
 865 870 875 880  
 Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys  
 885 890 895  
 Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp  
 900 905 910  
 Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr  
 915 920 925  
 Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp  
 930 935 940  
 Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser  
 945 950 955 960  
 Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg  
 965 970 975

ES 2 690 386 T3

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val  
 980 985 990  
 Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe  
 995 1000 1005  
 Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys  
 1010 1015 1020  
 Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser  
 1025 1030 1035 1040  
 Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu  
 1045 1050 1055  
 Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile  
 1060 1065 1070  
 Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser  
 1075 1080 1085  
 Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly  
 1090 1095 1100  
 Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile  
 1105 1110 1115 1120  
 Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser  
 1125 1130 1135  
 Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly  
 1140 1145 1150  
 Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile  
 1155 1160 1165  
 Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala  
 1170 1175 1180  
 Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys  
 1185 1190 1195 1200  
 Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser  
 1205 1210 1215  
 Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr  
 1220 1225 1230  
 Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser  
 1235 1240 1245  
 Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His  
 1250 1255 1260  
 Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val  
 1265 1270 1275 1280  
 Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys  
 1285 1290 1295  
 His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu  
 1300 1305 1310  
 Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp  
 1315 1320 1325  
 Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp  
 1330 1335 1340  
 Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile  
 1345 1350 1355 1360  
 Asp Leu Ser Gln Leu Gly Gly Asp  
 1365

# ES 2 690 386 T3

<210> 110  
 <211> 4221  
 <212> ADN  
 <213> Secuencia Artificial

5  
 <220>  
 <223> Secuencia que codifica Cas9 en pET-Cas9N3T para la producción de proteína Cas9 recombinante en E. coli (codón humanizado; marca hexa-His y señal de localización nuclear en el extremo N)

10

<400> 110  
 atgggcagca gccatcatca tcatcatcat gtgtaccct acgacgtgcc cgactacgcc 60  
 gaattgcctc caaaaaagaa gagaaaggta gggatcgaga acctgtactt ccagggcgac 120  
 aagaagtaca gcatcggcct ggacatcggg accaacagcg tgggctgggc cgtgatcacc 180  
 gacgagtaca aggtgcccag caagaagttc aaggtgctgg gcaacaccga ccgccacagc 240  
 atcaagaaga acctgatcgg cgccctgctg ttcgacagcg gcgagaccgc cgaggccacc 300  
 cgctgaagc gcaccgcccg ccgccgctac acccgccgca agaaccgcat ctgctacctg 360  
 caggagatct tcagcaacga gatggccaag gtggacgaca gtttcttcca ccgcctggag 420  
 gagagcttcc tggaggagga ggacaagaag cagcagcgcc accccatctt cggcaacatc 480  
 gtggacgagg tggcctacca cgagaagtac cccaccatct accacctgcg caagaagctg 540  
 gtggacagca ccgacaaggc cgacctgcgc ctgatctacc tggccctggc ccacatgatc 600  
 aagttccgcg gccacttcct gatcgagggc gacctgaacc ccgacaacag cgacgtggac 660  
 aagctgttca tccagctggt gcagacctac aaccagctgt tcgaggagaa ccccatcaac 720  
 gccagcggcg tggacgcca ggccatcctg agcgcccgcc tgagcaagag ccgccgcctg 780  
 gagaacctga tcgccagct gcccgcgag aagaagaacg gcctgttcgg caacctgatc 840  
 gccctgagcc tgggectgac cccaacttc aagagcaact tcgacctggc cgaggacgcc 900  
 aagctgcagc tgagcaagga cacctacgac gacgacctgg acaacctgct ggcccagatc 960  
 ggcgaccagt acgccgacct gttcctggcc gccagaacc tgagcgacgc catcctgctg 1020  
 agcgacatcc tgcggtgaa caccgagatc accaaggccc ccctgagcgc cagcatgatc 1080  
 aagcgtacg acgagcacca ccaggacctg accctgctga aggcctggg gcgccagcag 1140

ES 2 690 386 T3

ctgcccgaga agtacaagga gatcttcttc gaccagagca agaacggcta cgccggctac	1200
atcgacggcg gcgccagcca ggaggagttc tacaagttca tcaagcccat cctggagaag	1260
atggacggca ccgaggagct gctggtgaag ctgaaccgcg aggacctgct gcgcaagcag	1320
cgcaccttcg acaacggcag catccccac cagatccacc tgggcgagct gcacgccatc	1380
ctgcgccgcc aggaggactt ctacccttc ctgaaggaca accgcgagaa gatcgagaag	1440
atcctgacct tccgcatccc ctactacgtg ggccccctgg cccgcggcaa cagccgcttc	1500
gcctggatga cccgcaagag cgaggagacc atcacccctt ggaacttcga ggagtggtg	1560
gacaaggcg ccagcggcca gagcttcac gagcgcata ccaacttcga caagaacctg	1620
cccaacgaga aggtgctgcc caagcacagc ctgctgtacg agtacttcac cgtgtacaac	1680
gagctgacca aggtgaagta cgtgaccgag ggcatgcga agcccgctt cctgagcggc	1740
gagcagaaga aggccatcgt ggacctgctg ttcaagacca accgcaaggt gacctgaag	1800
cagctgaagg aggactactt caagaagatc gagtgcttcg acagcgtgga gatcagcggc	1860
gtggaggacc gcttcaacgc cagcctgggc acctaccag acctgctgaa gatcatcaag	1920
gacaaggact tccctggaaa cgaggagaac gaggacatcc tggaggacat cgtgctgacc	1980
ctgacctgt tccgaggaccg cgagatgatc gaggagcgc tgaagacctg cggccacctg	2040
ttcgacgaca aggtgatgaa gcagctgaag cgccgccgct acaccggctg gggccgctg	2100
agccgcaagc ttatcaacgg catccgcgac aagcagagcg gcaagacct cctggacttc	2160
ctgaagagcg acggcttcgc caaccgcaac ttcatgcagc tgatccacga cgacagcctg	2220
accttcaagg aggacatcca gaaggcccag gtgagcggcc agggcgacag cctgcacgag	2280
cacatcgcca acctggccgg cagccccgcc atcaagaagg gcatcctgca gacctgaag	2340
gtggtggacg agctggtgaa ggtgatgggc cgccacaagc ccgagaacat cgtgatcgag	2400
atggcccgcg agaaccagac caccagaag ggccagaaga acagccgca gcgcatgaag	2460
cgcatcgagg agggcatcaa ggagctgggc agccagatcc tgaaggagca ccccgaggag	2520
aacaccagc tgcagaacga gaagctgtac ctgtactacc tgcagaacg cgcgacatg	2580
tacgtggacc aggagctgga catcaaccgc ctgagcgact acgacgtgga ccacatcgtg	2640
cccagagct tccctgaagga cgacagcatc gacaacaagg tgctgacctg cagcgacaag	2700
aaccgcggca agagcgaaa cgtgcccagc gaggaggtgg tgaagaagat gaagaactac	2760
tggcgcagc tgctgaacgc caagctgatc accagcgcga agttcgaaa cctgaccaag	2820
gccgagcgcg gcggcctgag cgagctggac aaggccggt tcatcaagcg ccagctggtg	2880
gagacccgcc agatcaccaa gcacgtggcc cagatcctgg acagccgcat gaacaccaag	2940
tacgacgaga acgacaagct gatccgcgag gtgaaggtga tcaccctgaa gagcaagctg	3000

ES 2 690 386 T3

gtgagcgact tccgaagga cttccagttc tacaaggtgc gcgagatcaa caactaccac 3060  
cacgcccacg acgcctacct gaacgccgtg gtgggcaccg ccctgatcaa gaagtacccc 3120  
aagctggaga gcgagttcgt gtacggcgac tacaaggtgt acgacgtgcg caagatgatc 3180  
gccaaagagc agcaggagat cggcaaggcc accgccaagt acttcttcta cagcaacatc 3240  
atgaacttct tcaagaccga gatcaccctg gccaaaggcg agatccgcaa gcgccccctg 3300  
atcgagacca acggcgagac cggcgagatc gtgtgggaca agggccgcga cttcgccacc 3360  
gtgcgcaagg tgctgagcat gccccaggtg aacatcgtga agaagaccga ggtgcagacc 3420  
ggcggcttca gcaaggagag catcctgccc aagcgaaca gcgacaagct gatcgcccgc 3480  
aagaaggact gggaccccaa gaagtacggc ggcttogaca gcccaccgt ggcctacagc 3540  
gtgctggtgg tggccaaggt ggagaaggc aagagcaaga agctgaagag cgtgaaggag 3600  
ctgctgggca tcaccatcat ggagcgcagc agcttogaga agaaccctat cgacttctg 3660  
gaggccaagg gctacaagga ggtgaagaag gacctgatca tcaagctgcc caagtacagc 3720  
ctgttcgagc tgagagaacg ccgcaagcgc atgtgggcca gcgccggcga gctgcagaag 3780  
ggcaacgagc tggccctgcc cagcaagtac gtgaacttcc tgtacctggc cagccactac 3840  
gagaagctga agggcagccc cgaggacaac gagcagaagc agctgttcgt ggagcagcac 3900  
aagcactacc tggacgagat catcgagcag atcagcgagt tcagcaagcg cgtgatcctg 3960  
gccgacgcca acctggacaa ggtgctgagc gcctacaaca agcaccgca caagcccatc 4020  
cgcgagcagg ccgagaacat catccacctg ttaccctga ccaacctggg cccccccgc 4080  
gccttcaagt acttogacac caccatcgac cgcaagcgct acaccagcac caaggaggtg 4140  
ctggacgcca ccctgatcca ccagagcatc accggtctgt acgagaccgc catcgacctg 4200  
agccagctgg gcggcgacta a 4221

<210> 111  
<211> 1406  
5 <212> PRT  
<213> Secuencia Artificial

<220>  
10 <223> Secuencia de aminoácidos de Cas9 (pET-Cas9N3T)

<400> 111  
Met Gly Ser Ser His His His His His His Val Tyr Pro Tyr Asp Val  
1 5 10 15  
Pro Asp Tyr Ala Glu Leu Pro Pro Lys Lys Lys Arg Lys Val Gly Ile  
20 25 30  
Glu Asn Leu Tyr Phe Gln Gly Asp Lys Lys Tyr Ser Ile Gly Leu Asp  
35 40 45  
Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys





ES 2 690 386 T3

				725						730					735				
Asp	Asp	Ser	Leu	Thr	Phe	Lys	Glu	Asp	Ile	Gln	Lys	Ala	Gln	Val	Ser				
			740					745						750					
Gly	Gln	Gly	Asp	Ser	Leu	His	Glu	His	Ile	Ala	Asn	Leu	Ala	Gly	Ser				
		755					760					765							
Pro	Ala	Ile	Lys	Lys	Gly	Ile	Leu	Gln	Thr	Val	Lys	Val	Val	Asp	Glu				
	770					775					780								
Leu	Val	Lys	Val	Met	Gly	Arg	His	Lys	Pro	Glu	Asn	Ile	Val	Ile	Glu				
785					790					795					800				
Met	Ala	Arg	Glu	Asn	Gln	Thr	Thr	Gln	Lys	Gly	Gln	Lys	Asn	Ser	Arg				
				805					810					815					
Glu	Arg	Met	Lys	Arg	Ile	Glu	Glu	Gly	Ile	Lys	Glu	Leu	Gly	Ser	Gln				
			820					825					830						
Ile	Leu	Lys	Glu	His	Pro	Val	Glu	Asn	Thr	Gln	Leu	Gln	Asn	Glu	Lys				
		835					840					845							
Leu	Tyr	Leu	Tyr	Tyr	Leu	Gln	Asn	Gly	Arg	Asp	Met	Tyr	Val	Asp	Gln				
	850					855					860								
Glu	Leu	Asp	Ile	Asn	Arg	Leu	Ser	Asp	Tyr	Asp	Val	Asp	His	Ile	Val				
865					870					875					880				
Pro	Gln	Ser	Phe	Leu	Lys	Asp	Asp	Ser	Ile	Asp	Asn	Lys	Val	Leu	Thr				
				885					890					895					
Arg	Ser	Asp	Lys	Asn	Arg	Gly	Lys	Ser	Asp	Asn	Val	Pro	Ser	Glu	Glu				
			900					905					910						
Val	Val	Lys	Lys	Met	Lys	Asn	Tyr	Trp	Arg	Gln	Leu	Leu	Asn	Ala	Lys				
		915					920					925							
Leu	Ile	Thr	Gln	Arg	Lys	Phe	Asp	Asn	Leu	Thr	Lys	Ala	Glu	Arg	Gly				
	930					935					940								
Gly	Leu	Ser	Glu	Leu	Asp	Lys	Ala	Gly	Phe	Ile	Lys	Arg	Gln	Leu	Val				
945					950					955					960				
Glu	Thr	Arg	Gln	Ile	Thr	Lys	His	Val	Ala	Gln	Ile	Leu	Asp	Ser	Arg				
				965					970					975					
Met	Asn	Thr	Lys	Tyr	Asp	Glu	Asn	Asp	Lys	Leu	Ile	Arg	Glu	Val	Lys				
			980					985					990						
Val	Ile	Thr	Leu	Lys	Ser	Lys	Leu	Val	Ser	Asp	Phe	Arg	Lys	Asp	Phe				
		995				1000					1005								
Gln	Phe	Tyr	Lys	Val	Arg	Glu	Ile	Asn	Asn	Tyr	His	His	Ala	His	Asp				
	1010					1015					1020								
Ala	Tyr	Leu	Asn	Ala	Val	Val	Gly	Thr	Ala	Leu	Ile	Lys	Lys	Tyr	Pro				
1025				1030						1035					1040				
Lys	Leu	Glu	Ser	Glu	Phe	Val	Tyr	Gly	Asp	Tyr	Lys	Val	Tyr	Asp	Val				
				1045					1050					1055					
Arg	Lys	Met	Ile	Ala	Lys	Ser	Glu	Gln	Glu	Ile	Gly	Lys	Ala	Thr	Ala				

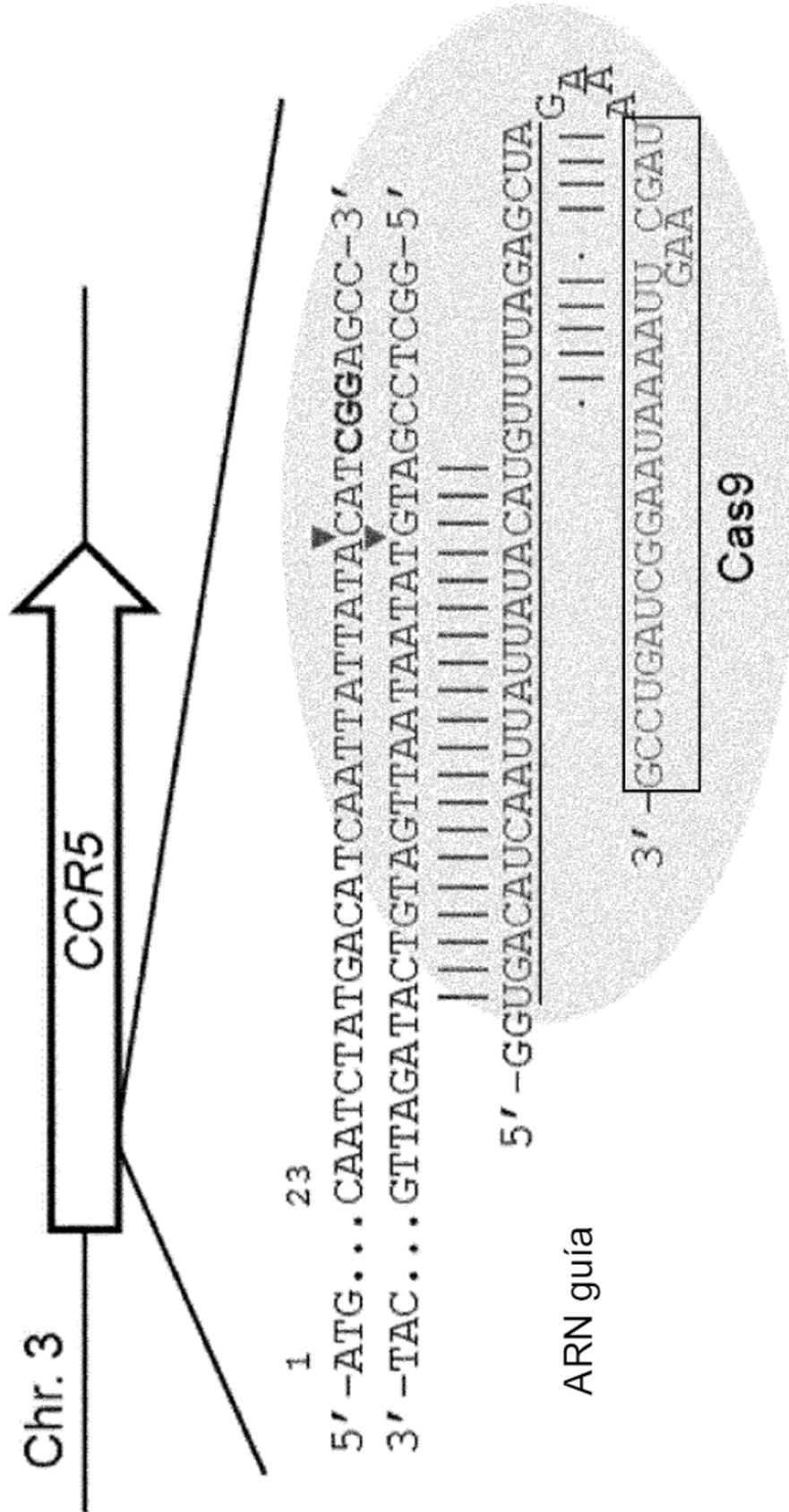
ES 2 690 386 T3

1060	1065	1070
Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile 1075 1080 1085		
Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn 1090 1095 1100		
Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr 1105 1110 1115 1120		
Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr 1125 1130 1135		
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg 1140 1145 1150		
Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys 1155 1160 1165		
Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val 1170 1175 1180		
Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu 1185 1190 1195 1200		
Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro 1205 1210 1215		
Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu 1220 1225 1230		
Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg 1235 1240 1245		
Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu 1250 1255 1260		
Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr 1265 1270 1275 1280		
Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe 1285 1290 1295		
Val Glu Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser 1300 1305 1310		
Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val 1315 1320 1325		
Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala 1330 1335 1340		
Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala 1345 1350 1355 1360		
Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser 1365 1370 1375		
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly 1380 1385 1390		
Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp 1395 1400 1405		

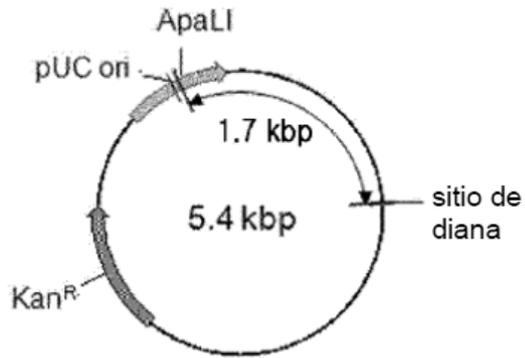
**REIVINDICACIONES**

1. Sistema de repeticiones palindrómicas cortas interespaciadas agrupadas regularmente tipo II (CRISPR)/proteína asociada CRISPR 9 (Cas9) para introducir roturas bicatenarias en una secuencia de ADN diana en una célula de mamífero, en la que el sistema CRISPR/Cas9 de tipo II comprende:
  - 5 a. una proteína Cas9 con una señal de localización nuclear (NLS), en la que el NLS está en el extremo C, o un ácido nucleico que codifica la proteína Cas9; y
  - b. un ARN guía monocatenario que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción ARNcr trans-activadora (ARNtracr).
2. El sistema CRISPR/Cas9 de tipo II de la reivindicación 1, en el que la proteína Cas9 es de *Streptococcus pyogenes*.
3. El sistema CRISPR/Cas9 de tipo II de la reivindicación 1 ó 2, en el que el ARN guía monocatenario es un ARN transcrito in vitro.
4. El sistema CRISPR/Cas9 de tipo II de cualquiera de las reivindicaciones 1-3, en el que la secuencia del ADN diana es una secuencia genómica localizada en su sitio endógeno en el genoma de la célula de mamífero.
- 15 5. El sistema CRISPR/Cas9 de tipo II de cualquiera de las reivindicaciones 1-4, en el que la célula de mamífero es una célula humana.
6. El sistema CRISPR/Cas9 de tipo II de una cualquiera de las reivindicaciones 1-5, en el que el ácido nucleico que codifica la proteína Cas9 está optimizado por codón para la expresión en células humanas.
- 20 7. El sistema CRISPR/Cas9 de tipo II de una cualquiera de las reivindicaciones 1-6, en el que la secuencia de ADN diana consiste en 20 nucleótidos complementarios a la porción ARNcr del ARN guía monocatenario y un trinucleótido motivo adyacente de protoespaciador (PAM), y en donde PAM consiste en el trinucleótido 5'-NGG-3'.
8. Un sistema CRISPR/Cas9 de tipo II como se define en una cualquiera de las reivindicaciones 1-7 para su uso en la escisión de un ADN diana en una célula de mamífero.
- 25 9. Un sistema CRISPR/Cas9 de tipo II como se define en una cualquiera de las reivindicaciones 1-7 para su uso en terapia génica.
10. Un método in vitro para introducir una rotura bicatenaria específica de sitio en una secuencia de ADN diana en una célula de mamífero, comprendiendo el método introducir en la célula de mamífero el sistema CRISPR/Cas9 Tipo II de una cualquiera de las reivindicaciones 1-7.
- 30 11. Un método in vitro para introducir una ruptura de cadena doble específica de sitio en una secuencia de ADN diana en una célula de mamífero, comprendiendo el método introducir en la célula de mamífero:
  - a. una proteína Cas9 con una señal de localización nuclear (NLS), en la que la NLS está en el extremo C, o un ácido nucleico que codifica la proteína Cas9; y
  - b. un ARN guía monocatenario que comprende una porción de ARN CRISPR (ARNcr) fusionada a una porción ARNcr trans-activadora (ARNtracr).
- 35 12. El método de la reivindicación 11, en el que el ácido nucleico que codifica la proteína Cas9 se introduce en la célula de mamífero antes de introducir el ARN guía monocatenario en la célula de mamífero.
13. El método de las reivindicaciones 10, 11 ó 12, en el que la célula de mamífero es una célula humana.

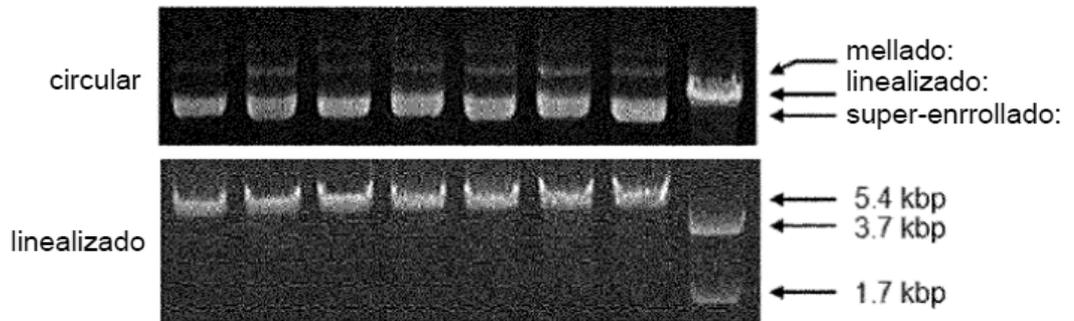
[Fig. 1a]



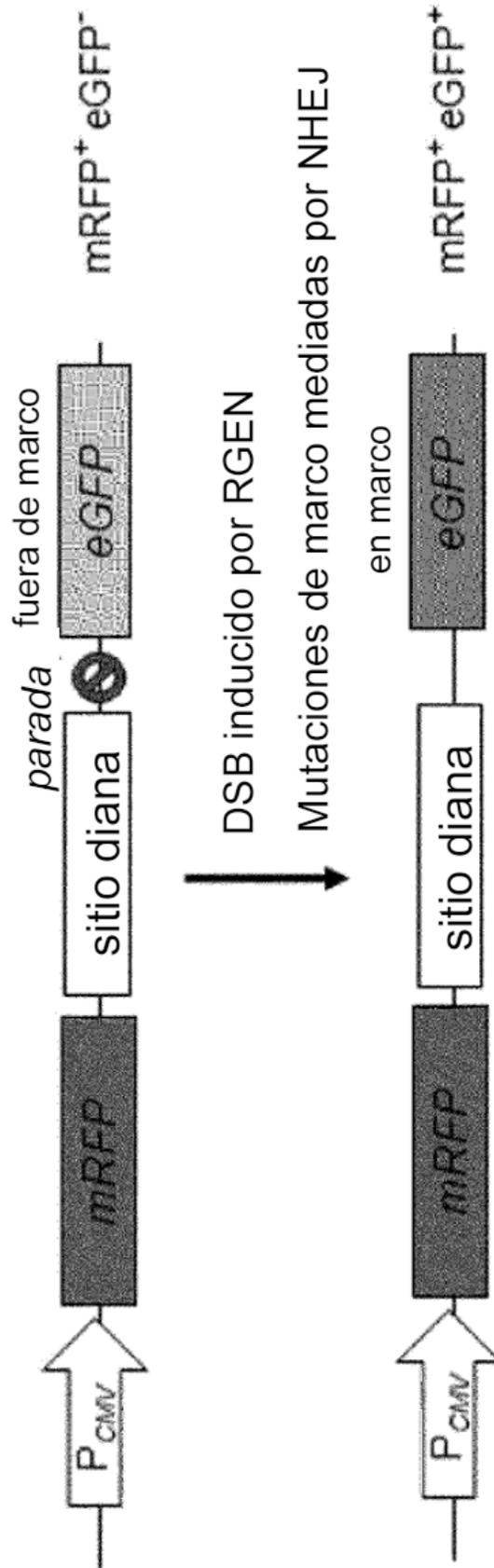
[Fig. 1b]



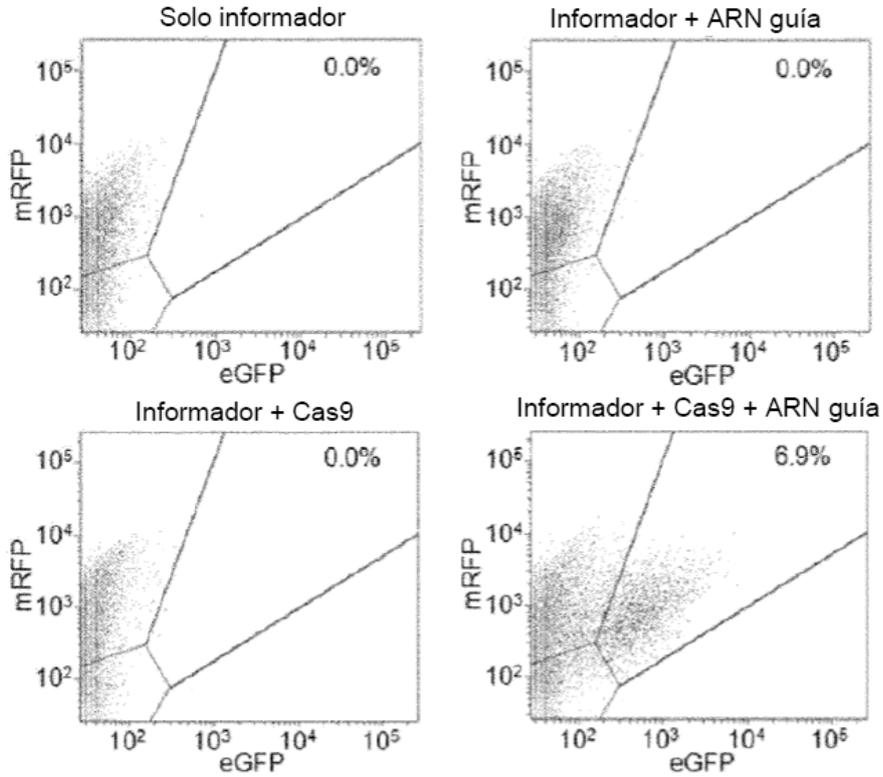
Cas9:	-	+	-	+	-	+	-	+
ARN guía:	-	-	+	+	-	-	+	+
secuencia diana:	-	-	-	-	+	+	+	+



[Fig. 2a]



[Fig. 2b]



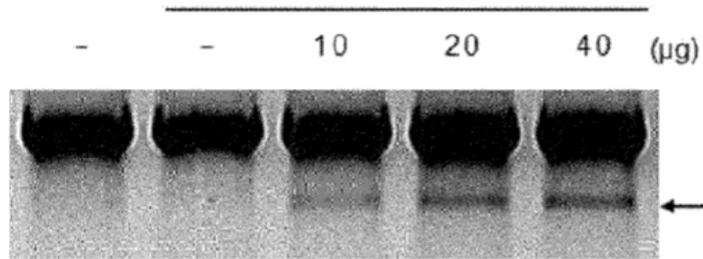
[Fig. 3a]

**CCR5**

ADN plasmídico: Vacío

Cas9

ARN guía



Inserciones-delecciones (%)

1.3 3.9 5.1

7,3% (7/96) mutado

```

CAATCTA TGACATCAATTATTATA-CAT CCGGAGCCCTGCCAAAAAATCAA WT
CAATCTA TGACATCAATTATTATA CAT CCGGAGCCCTGCCAAAAAATCAA +1
CAATCTA TGACATCAATTATTAT -----GCCAAAAAATCAA -13
CAATCTA TGACATC -----GGAGCCCTGCCAAAAAATCAA -14
CAATCTA TGACAT -----GCCCTGCCAAAAAATCAA -18
CAATCTA TGACATCAATTATTAT -----AAATCAA -19
CAATCTA TGACATC -----CAAAAAAATCAA -24
CAATCTA TGACA -----AAATCAA -30
    
```

[Fig. 3b]

**C4BPB**

ADN plasmídico: Vacío

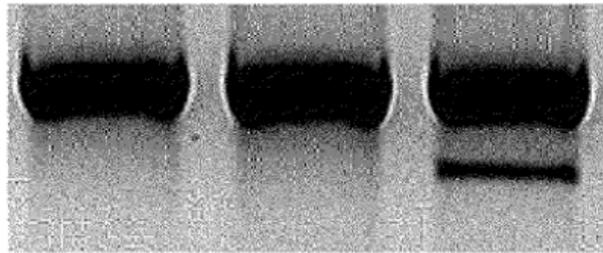
Cas9

ARN guía: -

-

-

+



Inserciones-delecciones (%)

14

8,3% (4/48) mutado

```

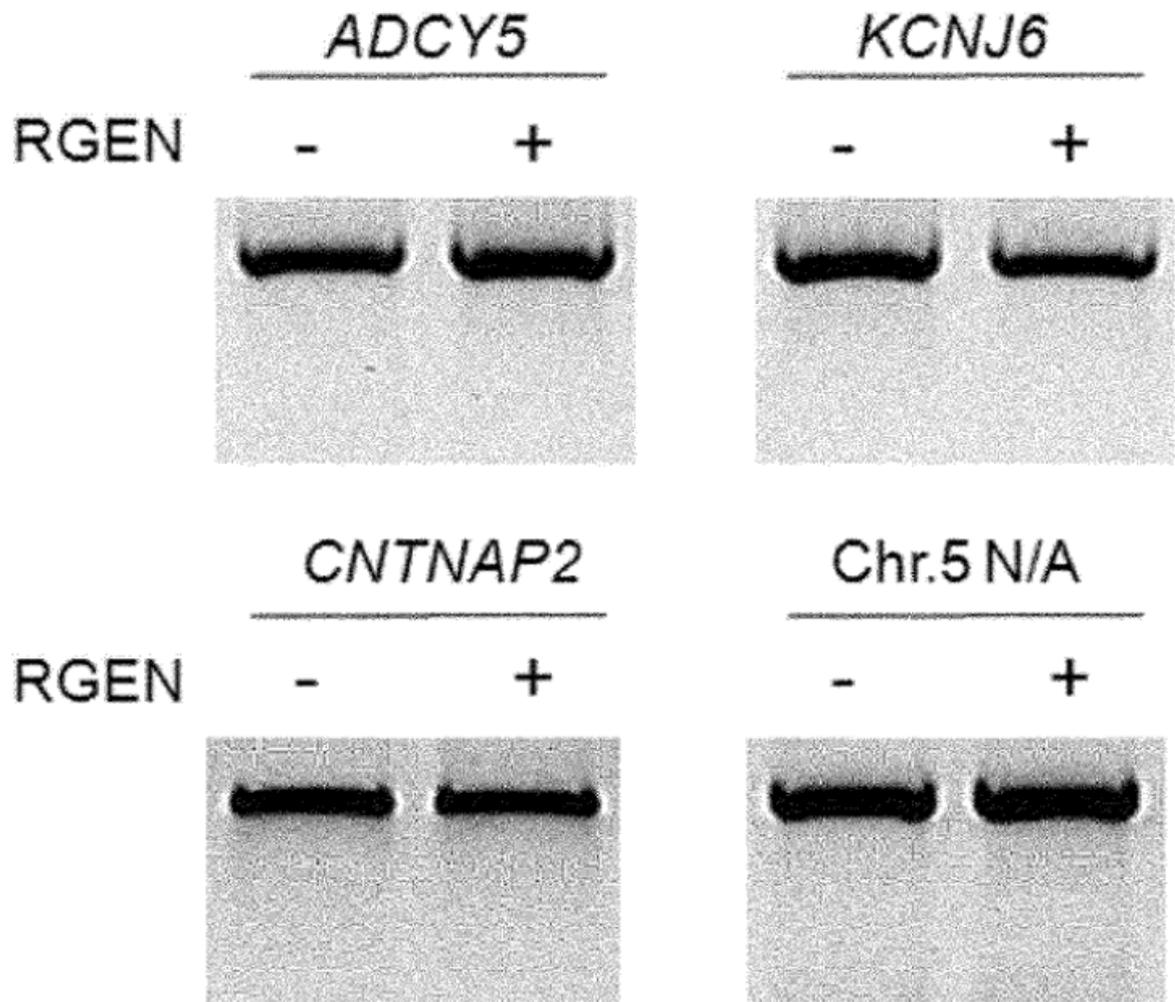
TATGTGCAATGACCACTACATCCT--CAAGGGGCAGCAATCGGAGCCAG  WT
TATGTGCAATGACCACTACATCCTT-CAAGGGGCAGCAATCGGAGCCAG  +1
TATGTGCAATGACCACTACATCCTCTCAAGGGGCAGCAATCGGAGCCAG  +2
TATGTGCAATGAG-----GAGCCAG  -30
TATGTGCAATGAC-----180 bp-----180
    
```

[Fig. 4a]

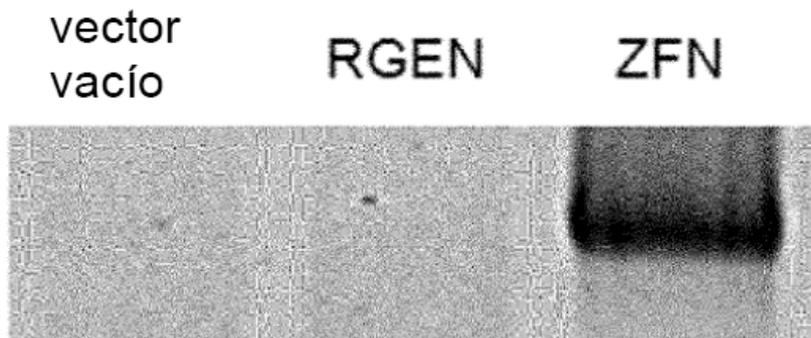
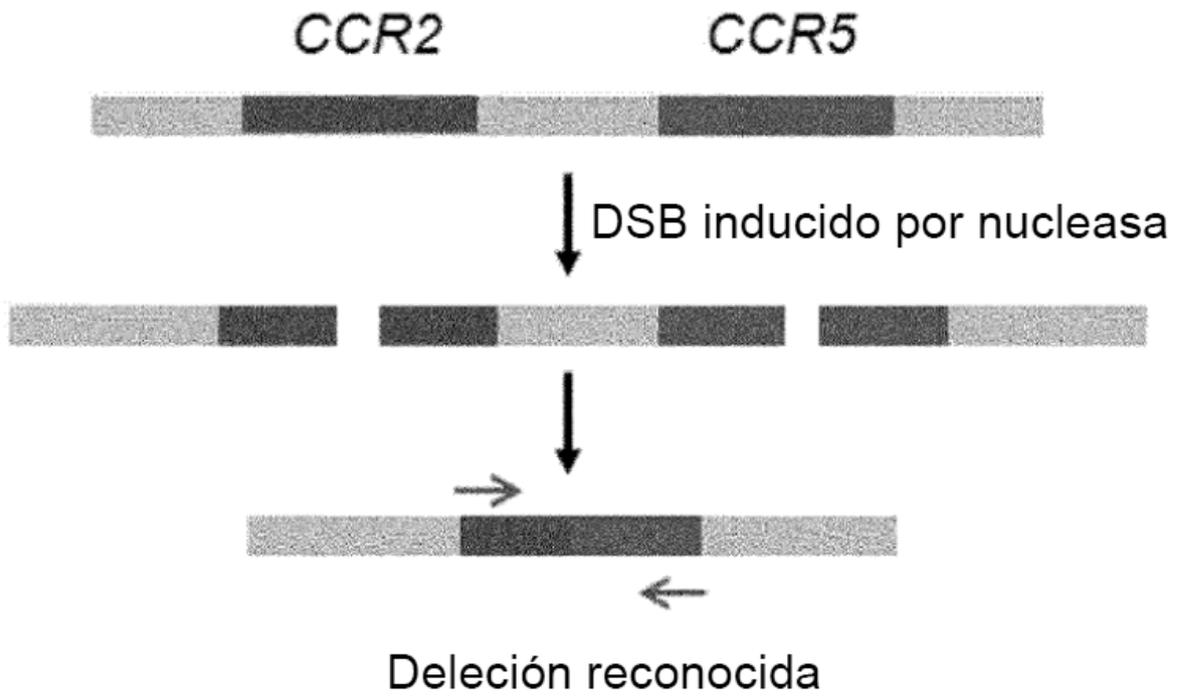
```

TGACATCAATTATTATACATCGG  CCR5
TGACATCAATTATTATAgATgGa  ADCY5
TGACATCAcTTATTATgCATgGG  KCNJ6
TGACATaAATTATTcTACATgGG  CNTNAP2
TGAaATCAATTATcATAgATCGG  Chr. 5 N/A
    
```

[Fig. 4b]

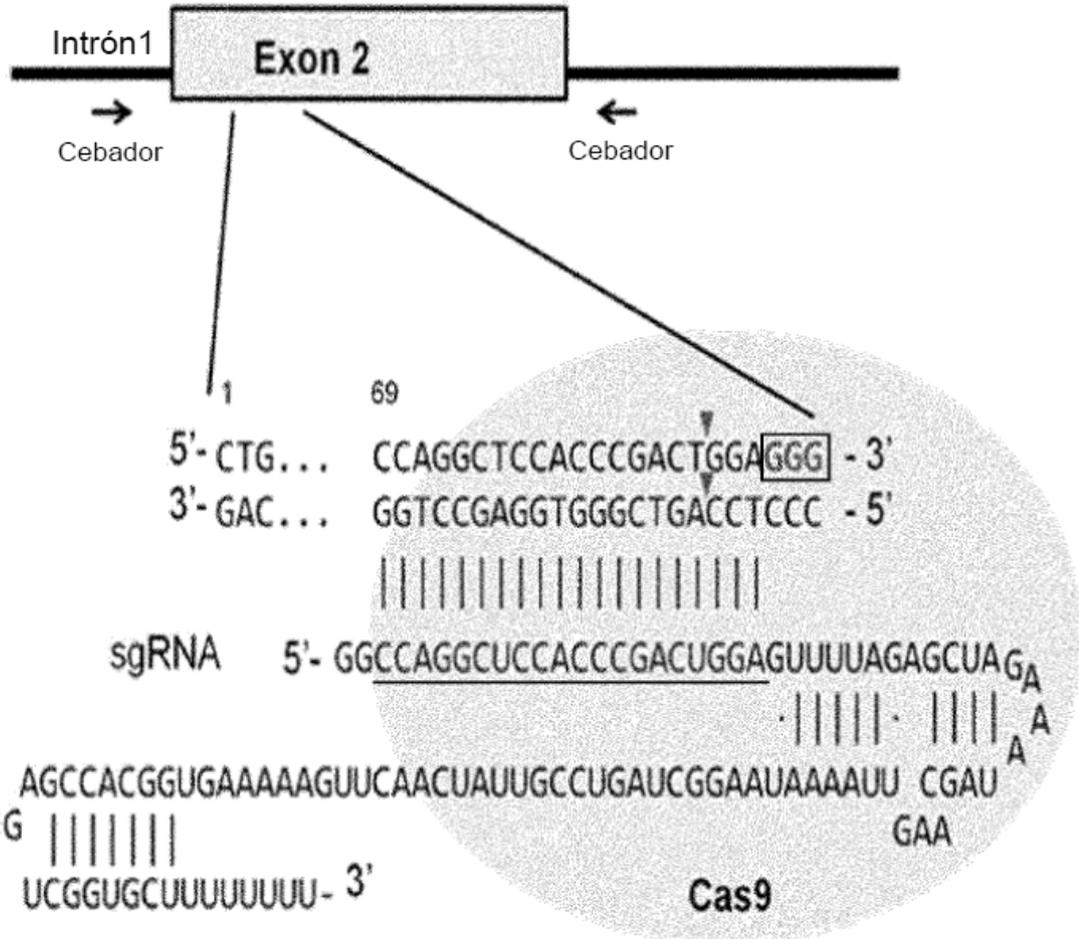


[Fig. 4c]

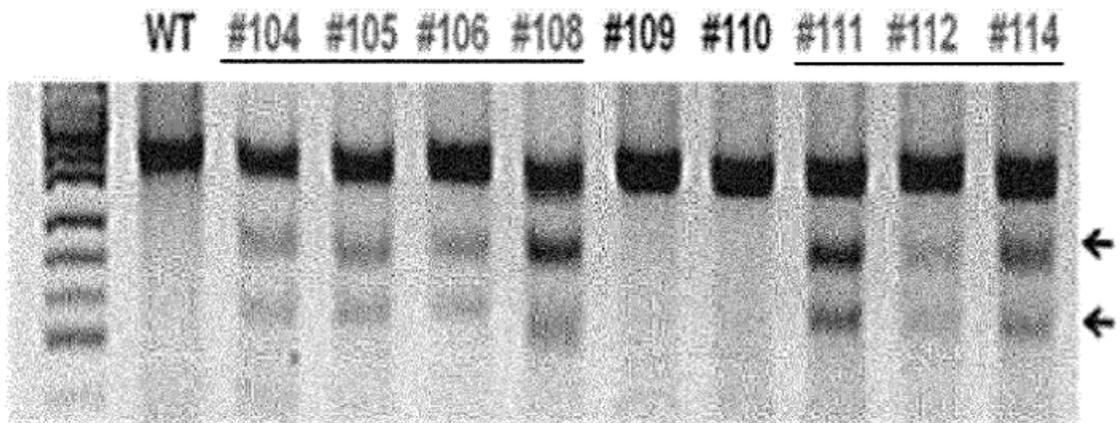


[Fig. 5a]

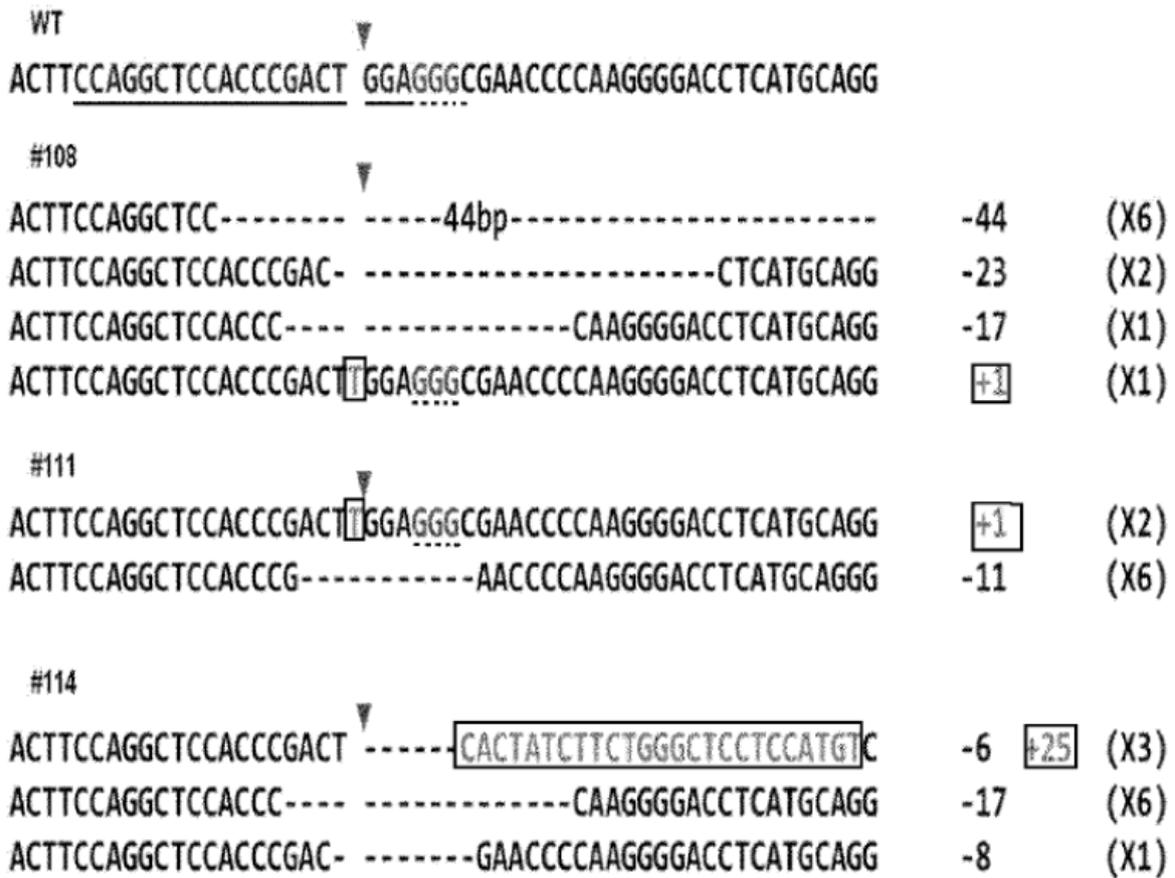
*Foxn1*



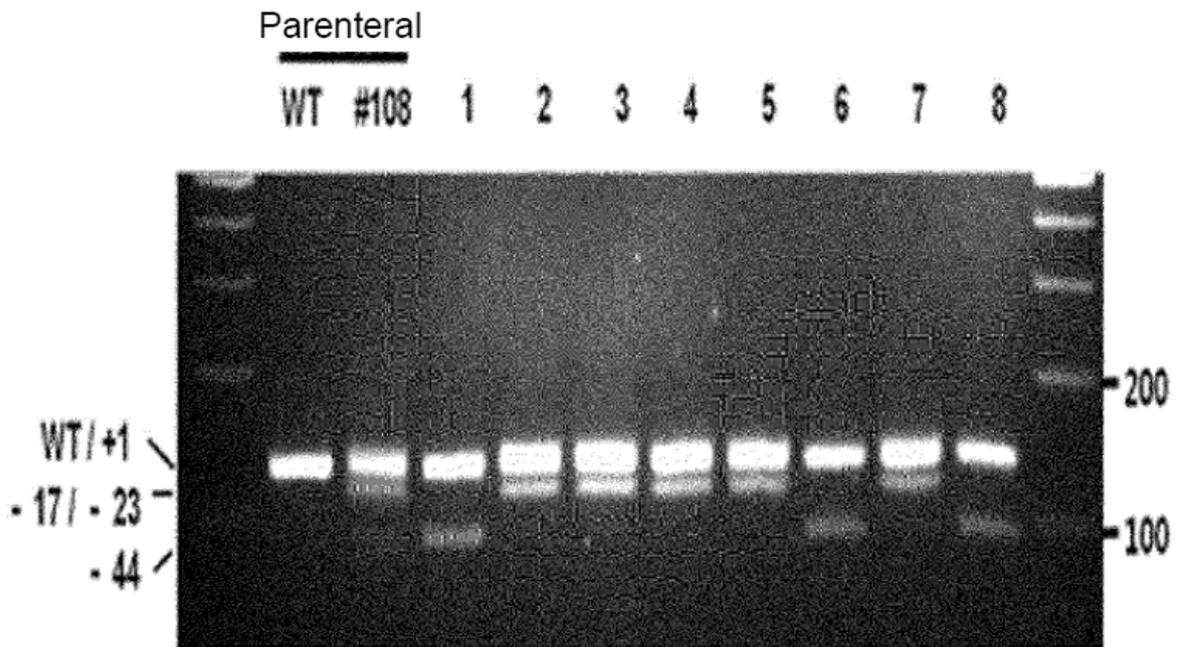
[Fig. 5b]



[Fig. 5c]



[Fig. 5d]





[Fig. 7a]

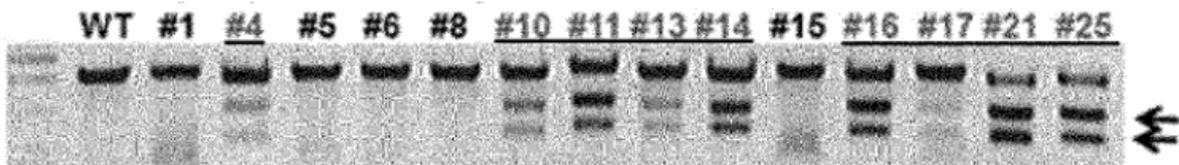
Inyección pronuclear



proteína Cas9 (nM)	ARNsg (nM)	Embriones analizados	Mutantes (%)
2	4	20	3 (15%)
20	40	15	5 (33%)
200	400	17	15 (88%)

[Fig. 7b]

Inyección intra-citoplásmica

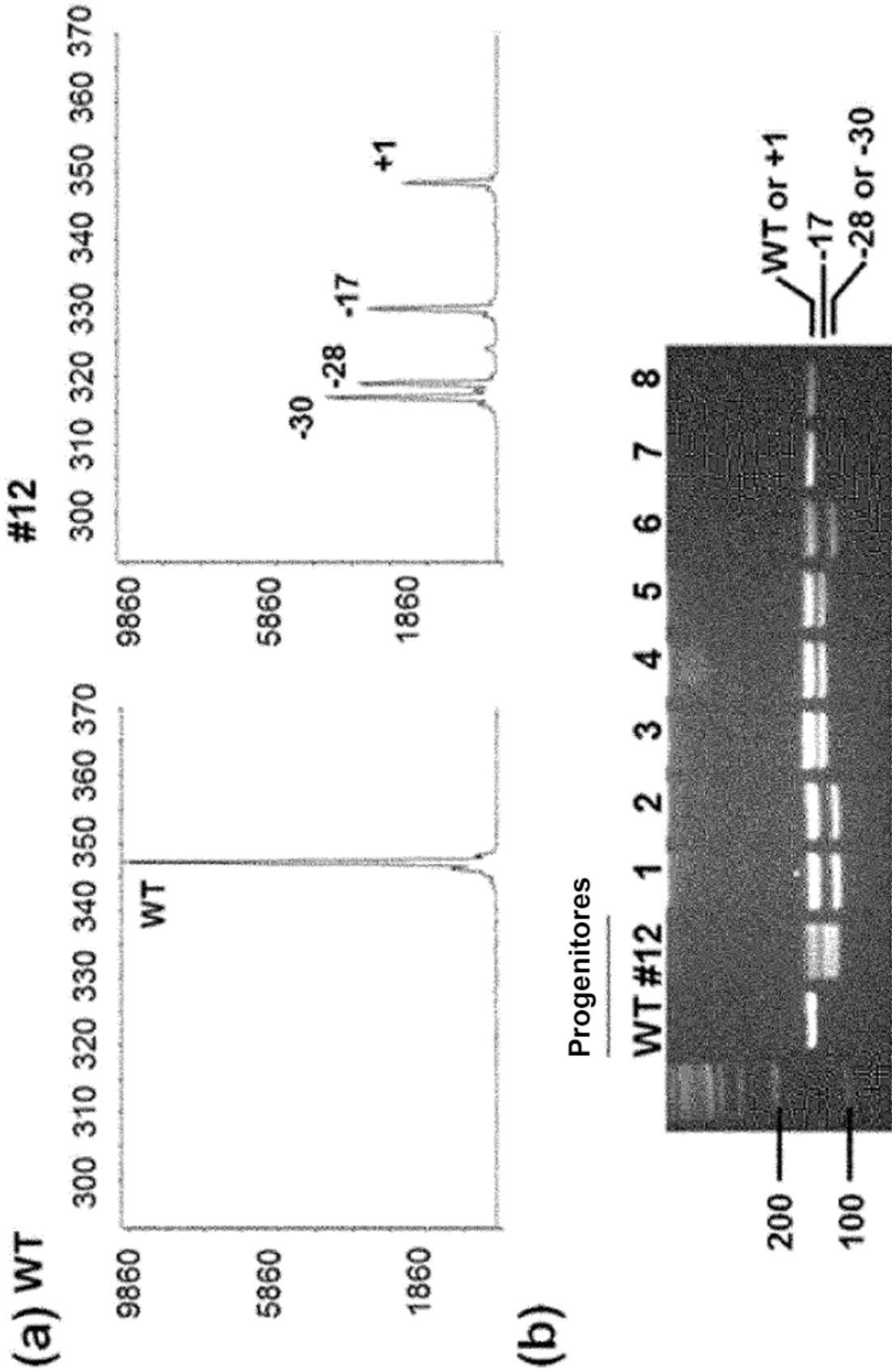


proteína Cas9 (nM)	ARNsg (nM)	Embriones analizados	Mutantes (%)
2	4	18	1 (6%)
20	40	19	6 (32%)
200	400	14	10 (71%)

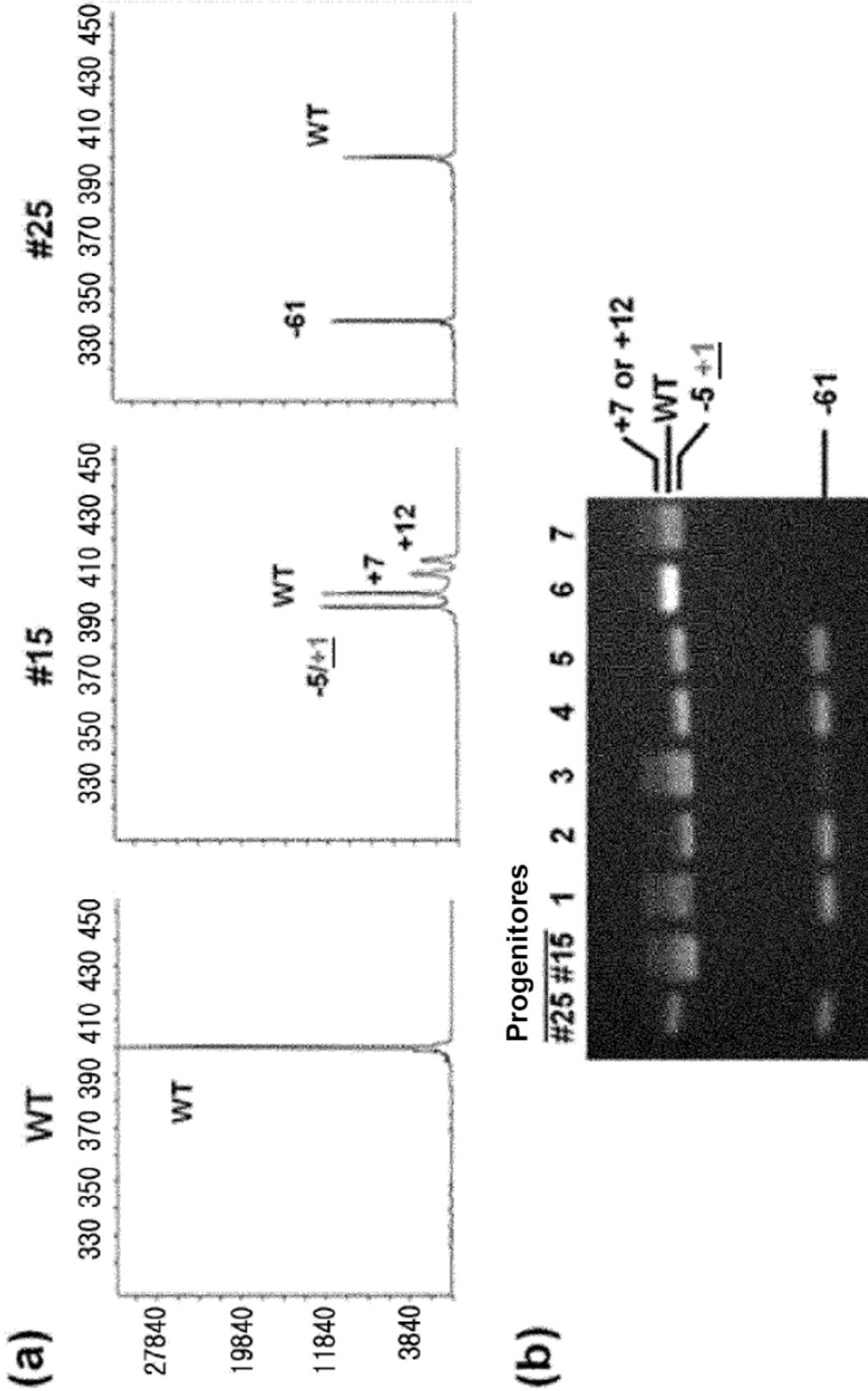
[Fig. 7c]

Secuencia	Inserciones- deleciones	Embrión nº.
<u>ACTTCCAGGCTCCACCCGACTGGAGGGCGAACCCCAAGGGAACTCATGCCAG</u> WT		
ACTTCCAGGCGAACCC-----AAGGGAACTCATGCCAG	Δ18	2
ACTTCCAGGCTCCAC-----AAGGGAACTCATGCCAG	Δ20	1
ACTTCCAGGCTCCACCC-----AAGGGAACTCATGCC	Δ19	1
ACTTCCAGGCTCCACCC-----CAAGGGAACTCATGCCAG	Δ17	1
ACTTCCAGGCTCCACCCGA-----ACCCCAAGGGAACTCATGCCAG	Δ11	3
ACTTCCAGGCTCCACCCGAA--GGAGGGCGAACCCCAAGGGAACTCATGCCA	Δ3+1	1
ACTTCCAGGCTCCACCCGACT--AGGGCGAACCCCAAGGGAACTCATGCCAG	Δ2	1
ACTTCCAGGCTCCACCCGACTGGAGGGCGAACCCCAAGGGAACTCATGCCA	+1	1
ACTTCCAGGCTCCACCCGACTTGGAGGGCGAACCCCAAGGGAACTCATGCCA	+1	10
ACTTCCAGGCTCCACCCGAA-----GGCGAACCCCAAGGGAACTCATGCCAG	Δ6	1
ACTTCCAGGCTCCACCCGAA-----GGCGAACCCCAAGGGAACTCATGCCAG	Δ5	2
ACTTCCAGGCTCCACC-----TCATGCCAG	Δ28	1
-----AGGGCGAACCCCAAGGGAACTCATGCCAG	Δ126	1
Total		26

[Fig. 8]

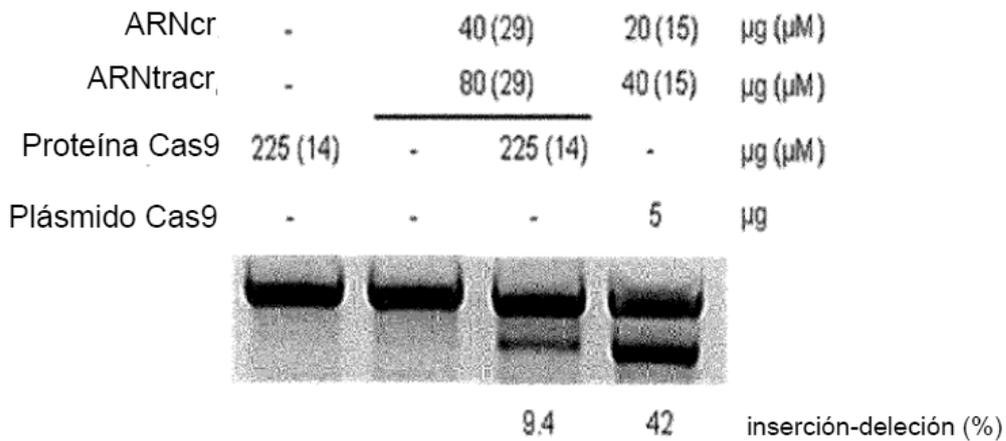
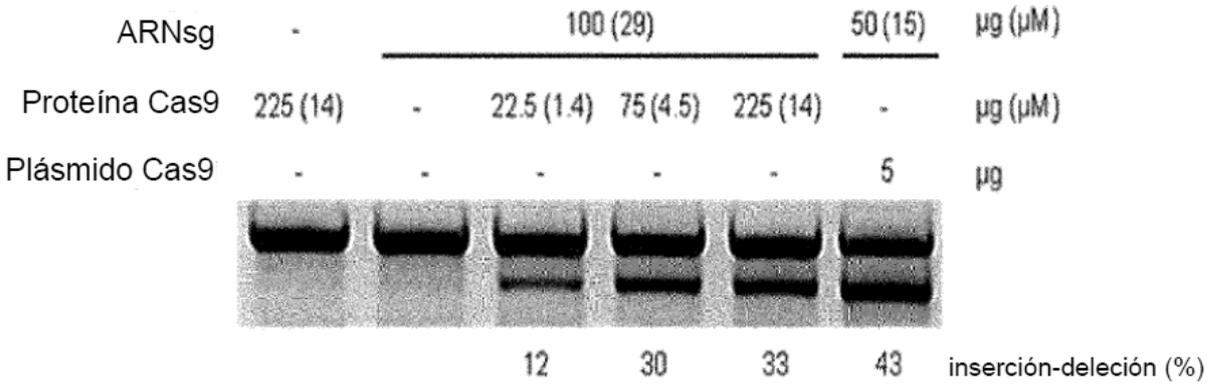
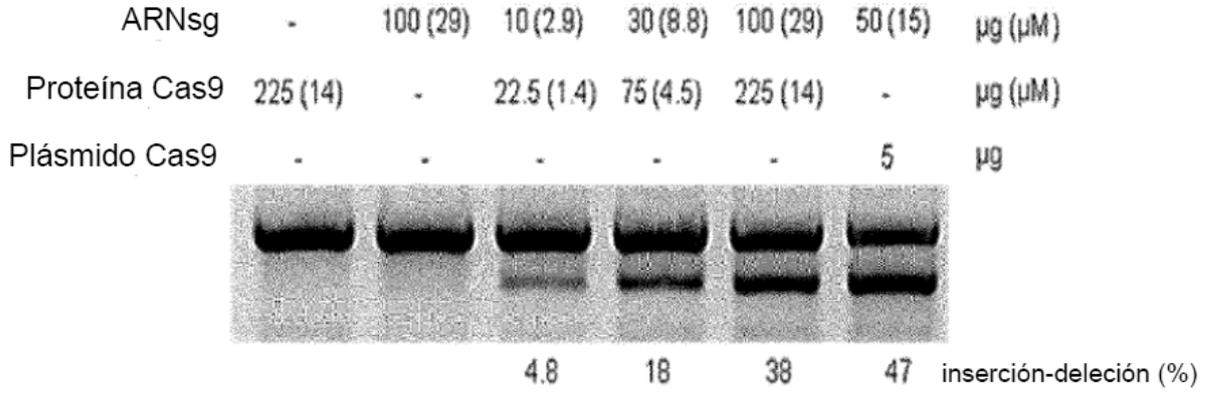


[Fig. 9]



[Fig. 10a]

**CCR5#4**





[Fig. 10e]

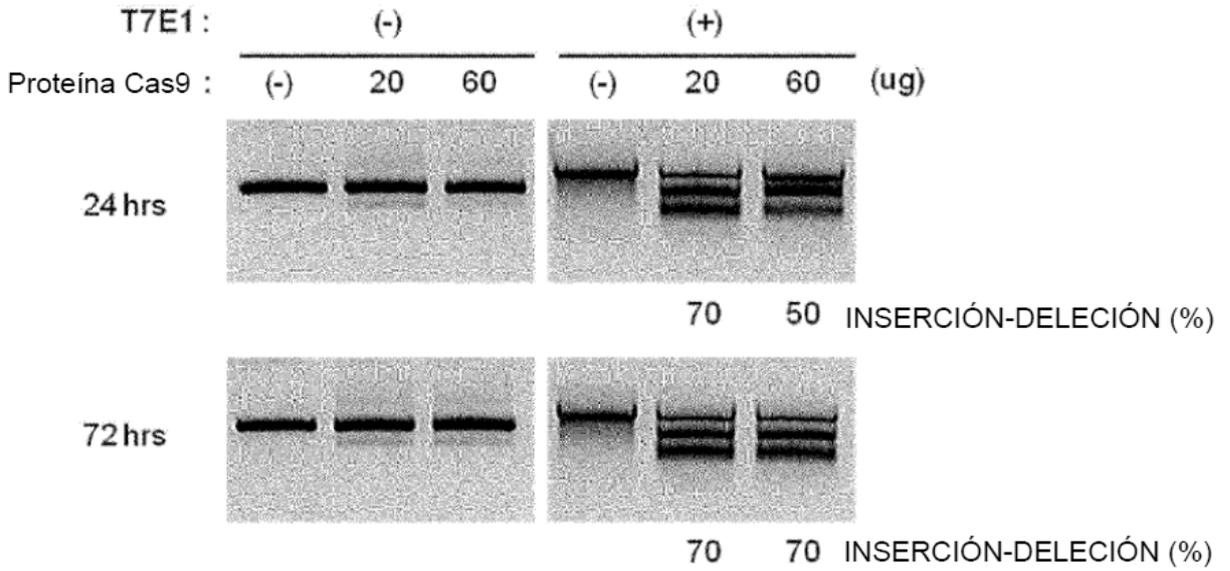
**ABCC11**

▼

```

TTCTCAAGGCAGCATCATACTTCCCCCACGGTGGGACAGCTGCCCTCCCTGG   WT
TTCTCAAGGCAGCATCATACTTCC-----CTGGGACAGCTGCCCTCCCTGG   -6
TTCTCAAGGCAGCATCATACTTC---CACGGTGGGACAGCTGCCCTCCCTGG   -3
TTCTCAAGGCAGC-----TGCCCTCCCTGG   -29
TTCTCAAGGCAGCATCATACTT-----CCCTCCCTGG   -20
TTCTCAAGGCAGCATCATACTT-----CCCTCCCTGG   -20
TTCTC----- -256
    
```

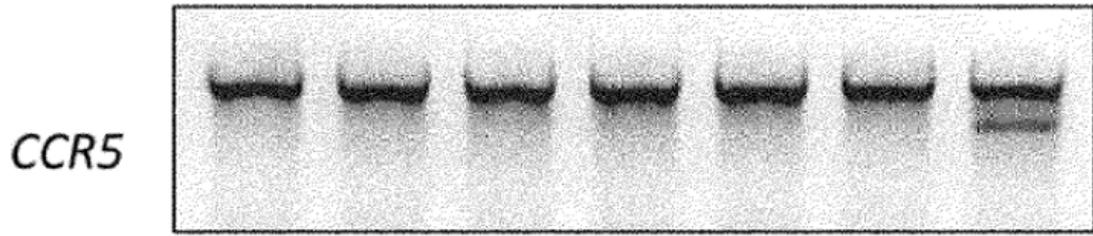
[Fig. 11]





[Fig. 13]

Cas9	-	+	-	-	-	+	-
Cas9-9R4L	-	-	+	-	-	-	+
ARNsg	-	-	-	+	-	+	-
sgRNA-9R4L	-	-	-	-	+	-	+

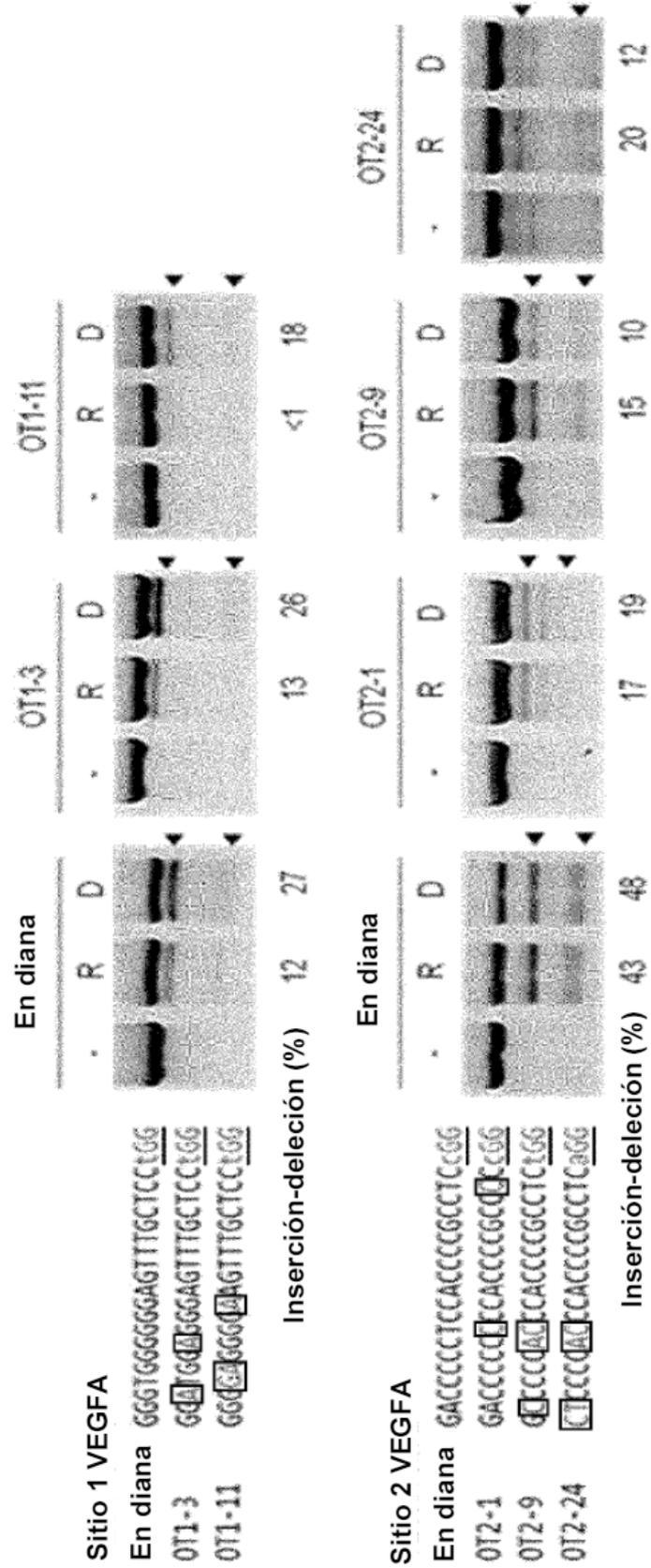


*CCR5*

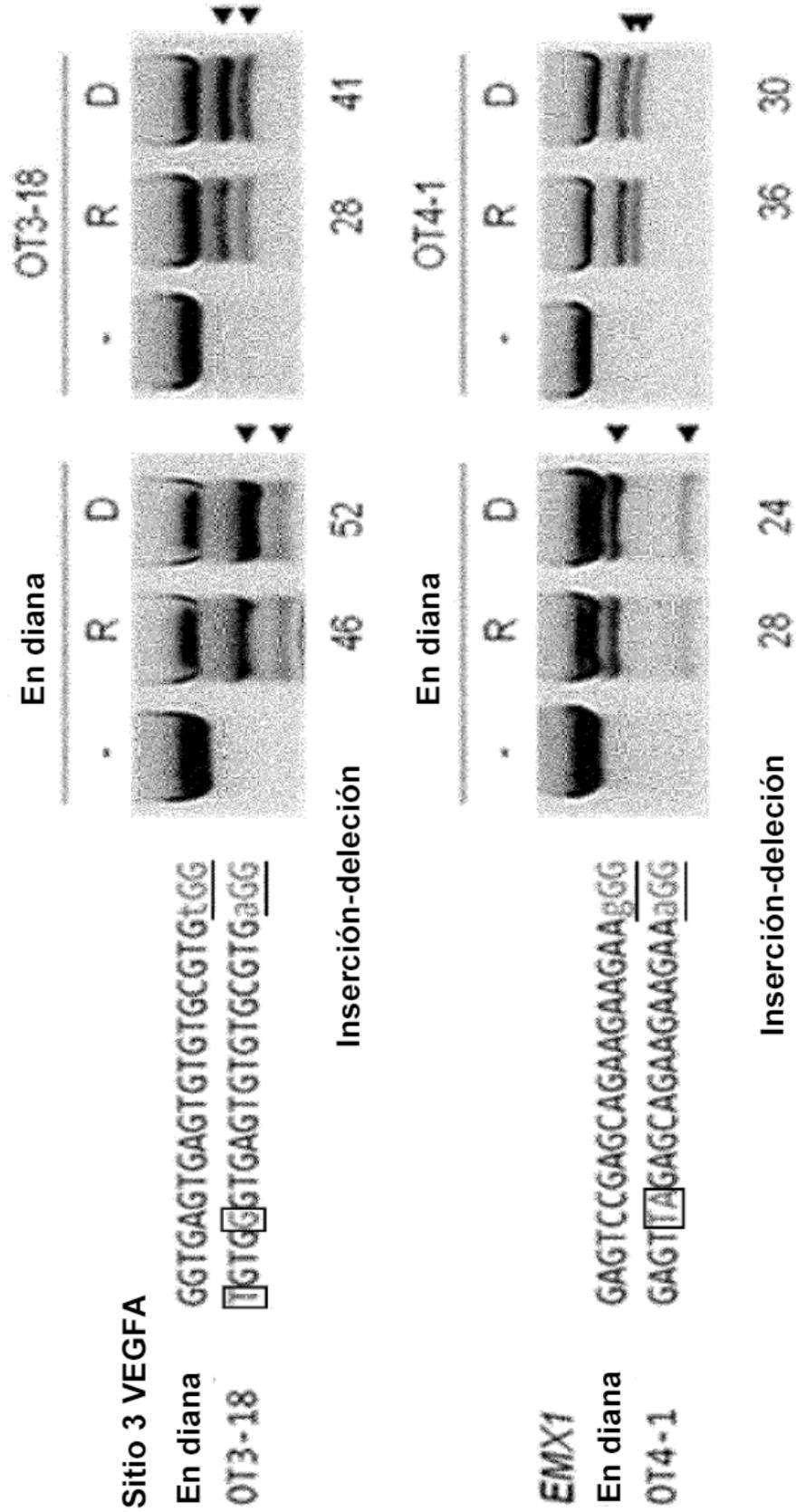
Inserción-Delección (%)

9

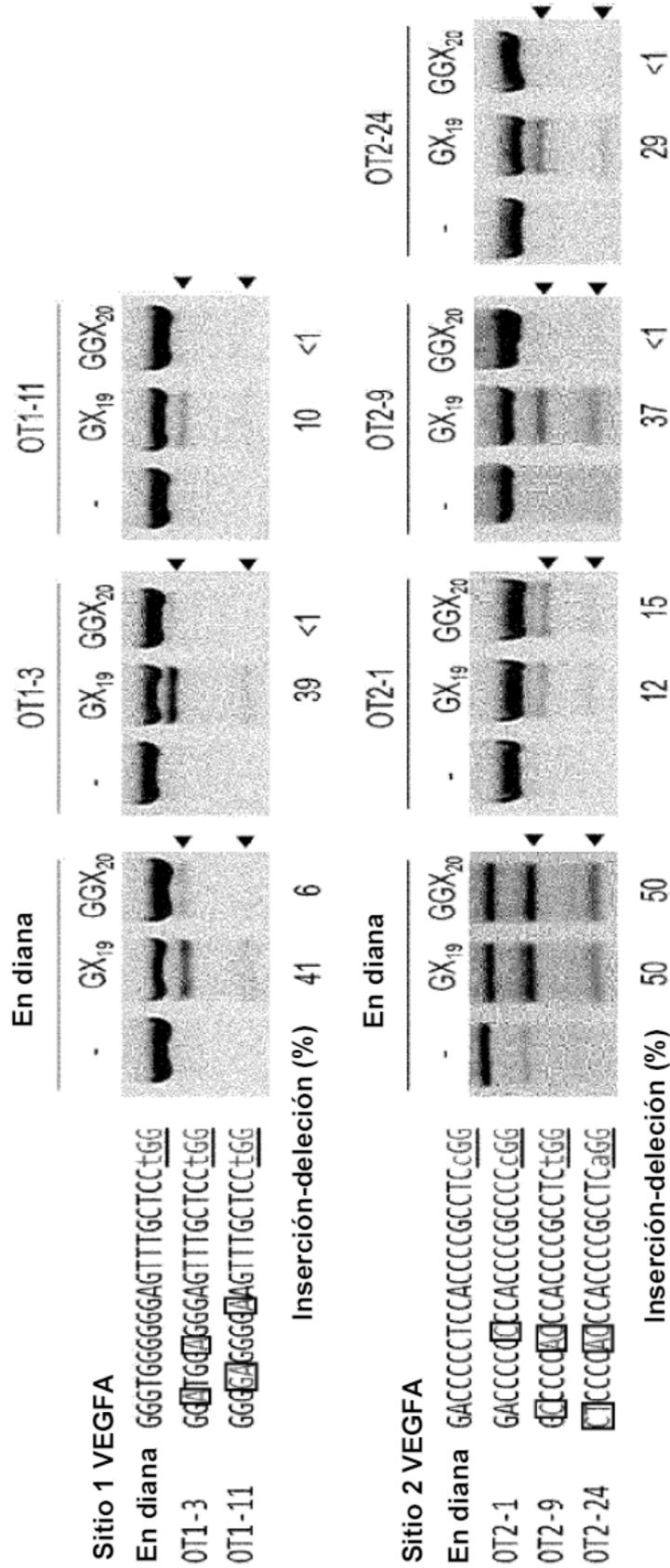
[Fig. 14a]



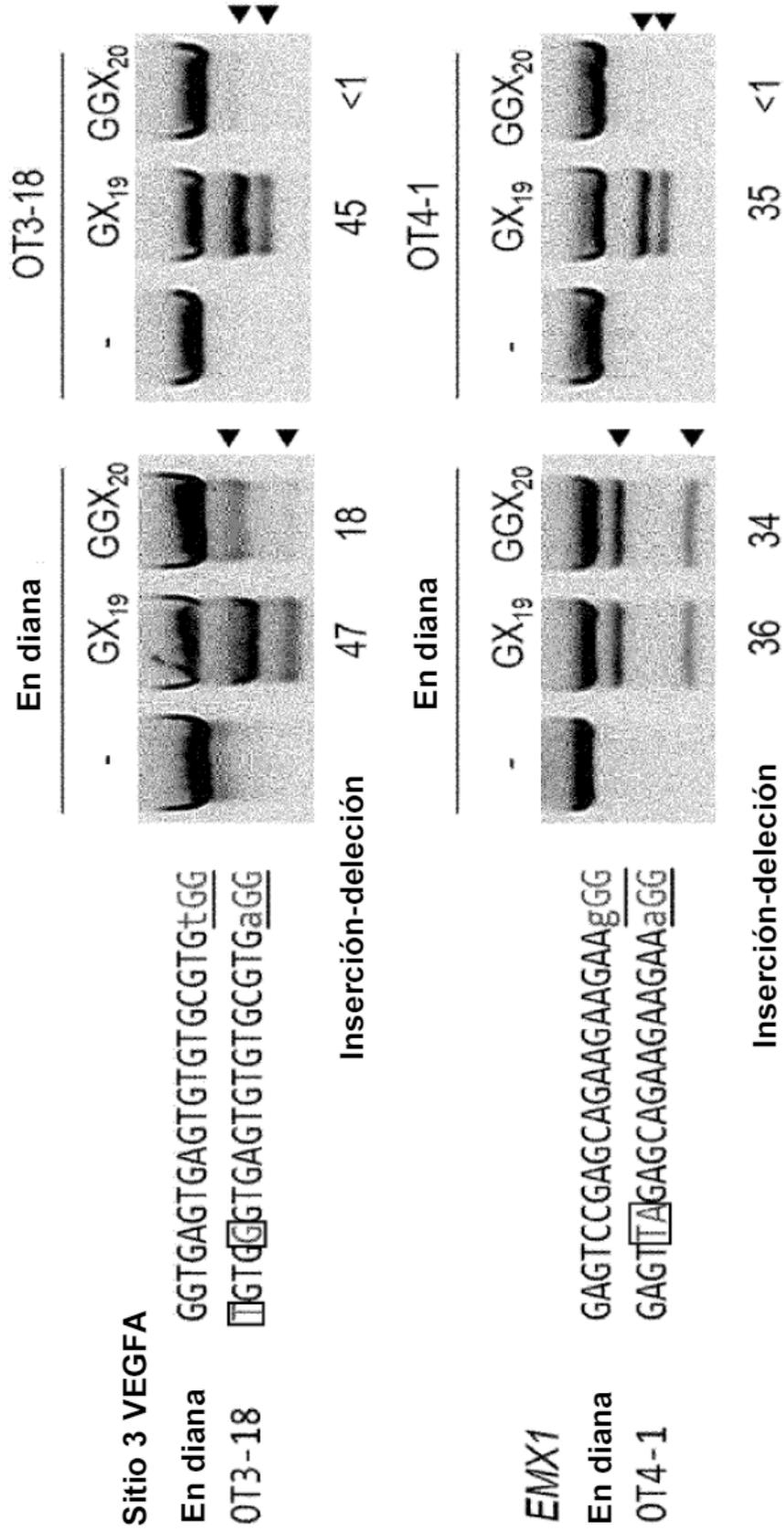
[Fig. 14b]



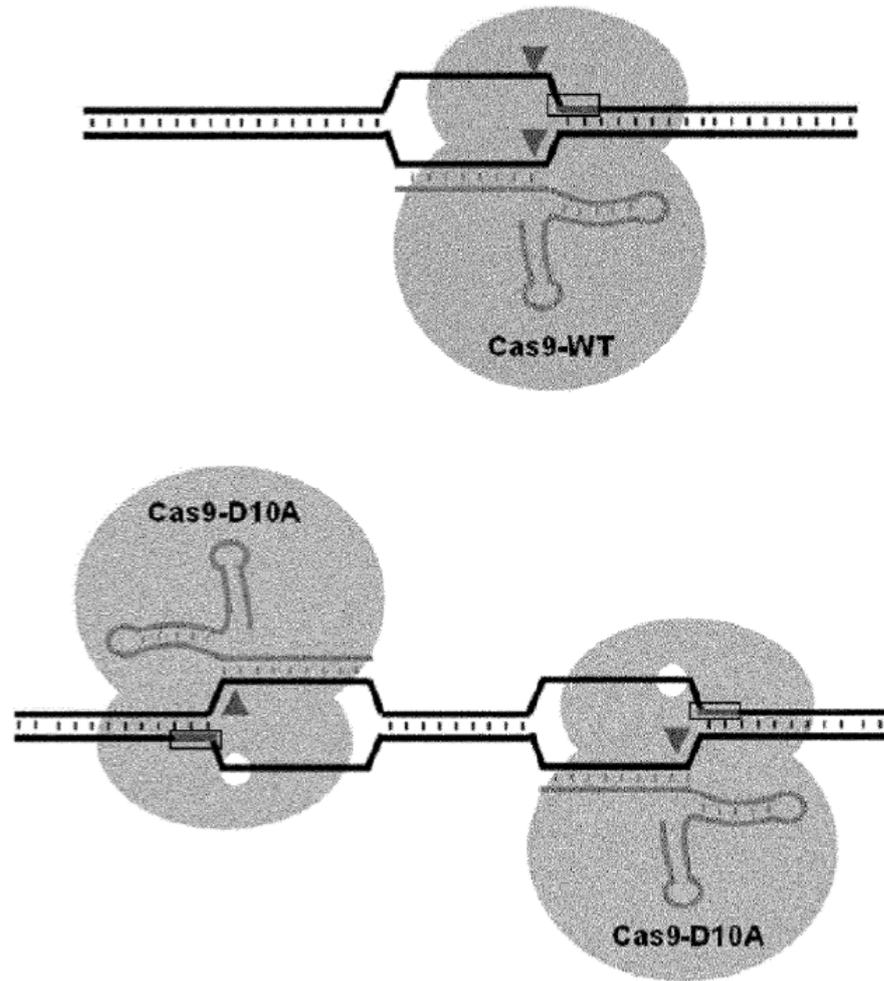
[Fig. 15a]



[Fig. 15b]

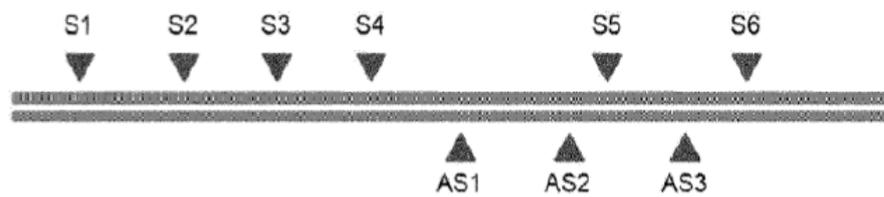


[Fig. 16a]

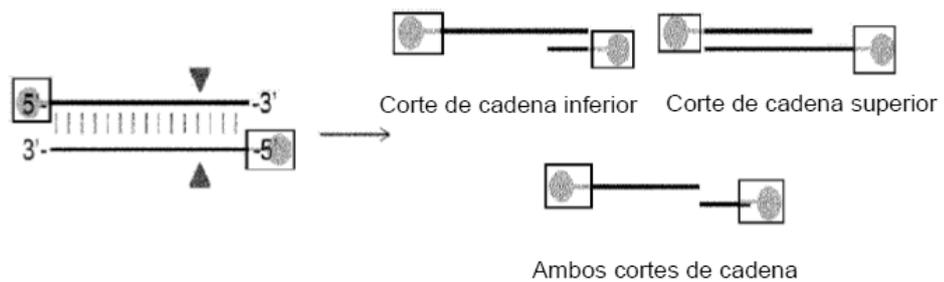


[Fig. 16b]

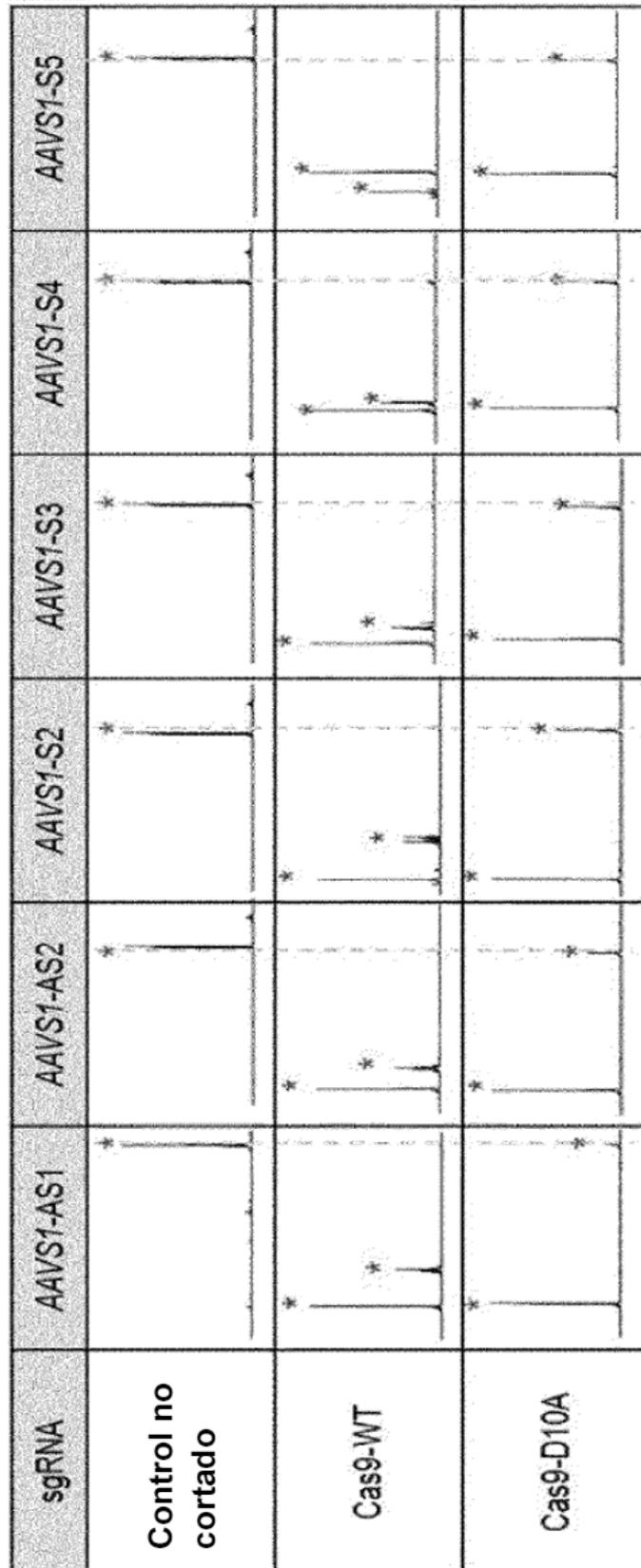
AAVS1 locus



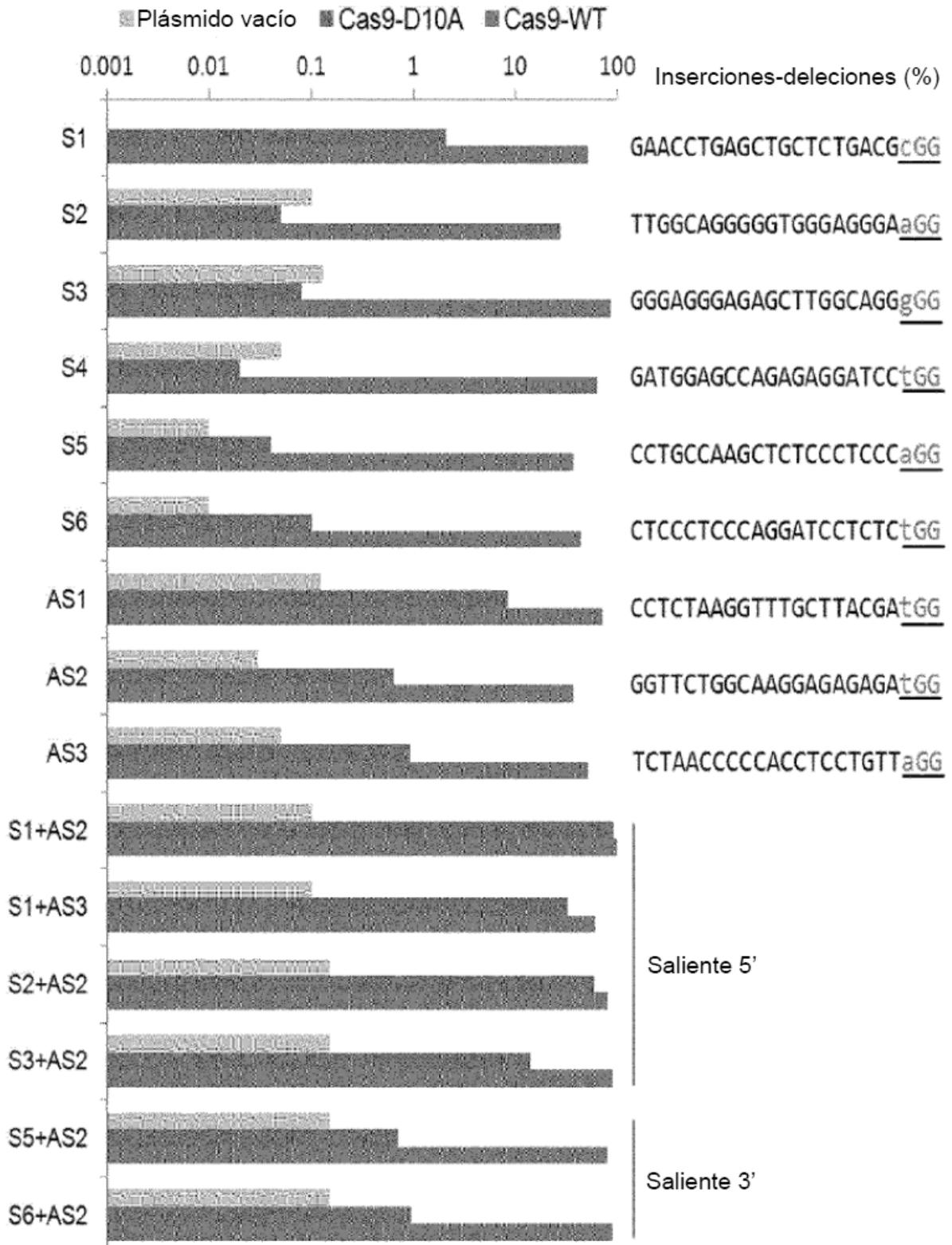
[Fig. 16c]



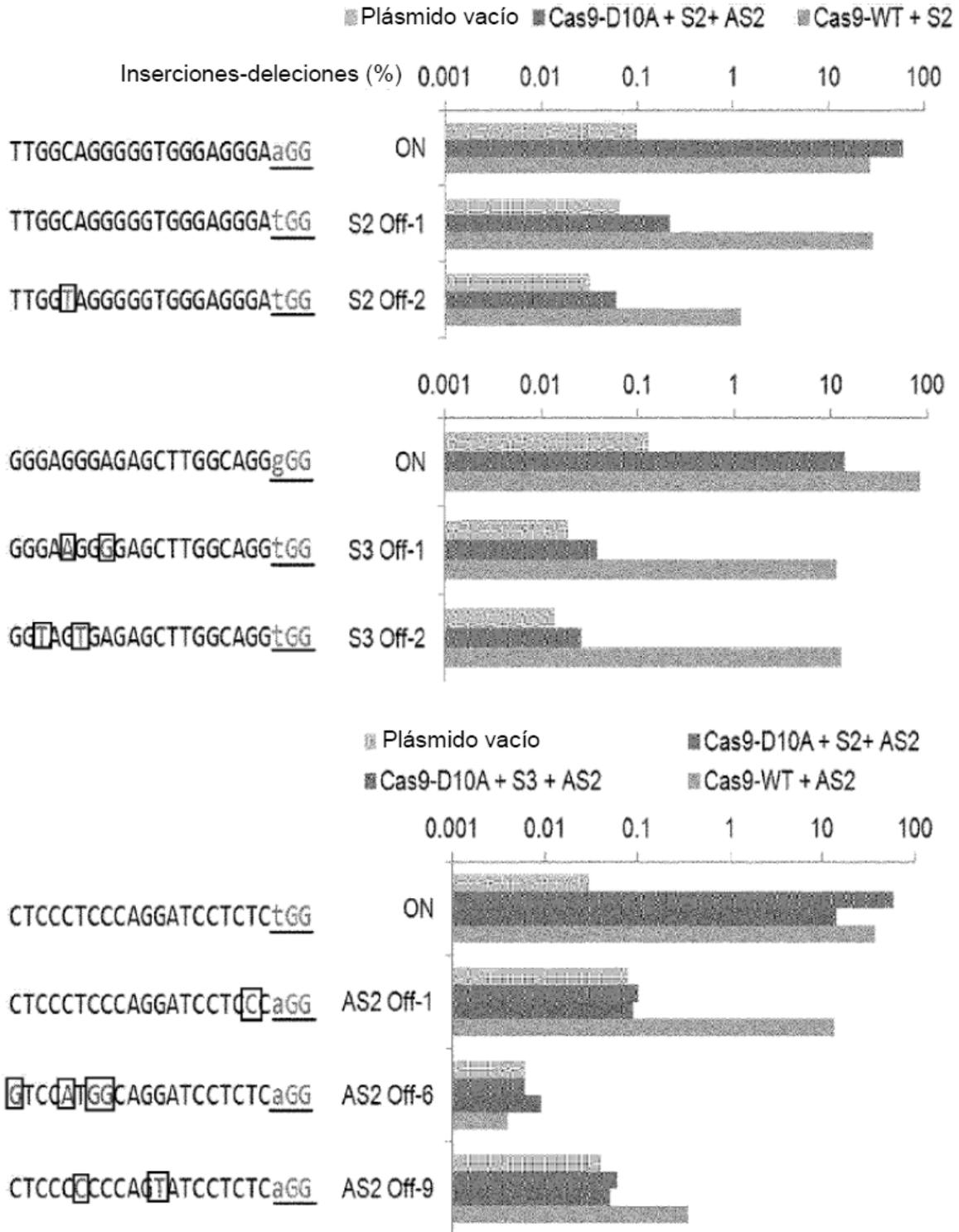
[Fig. 16d]



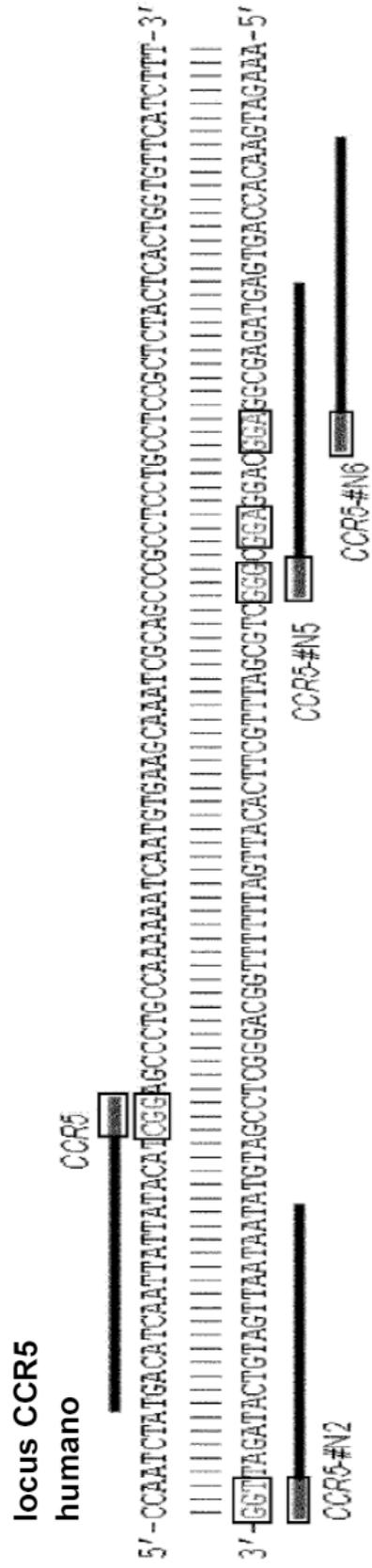
[Fig. 17a]



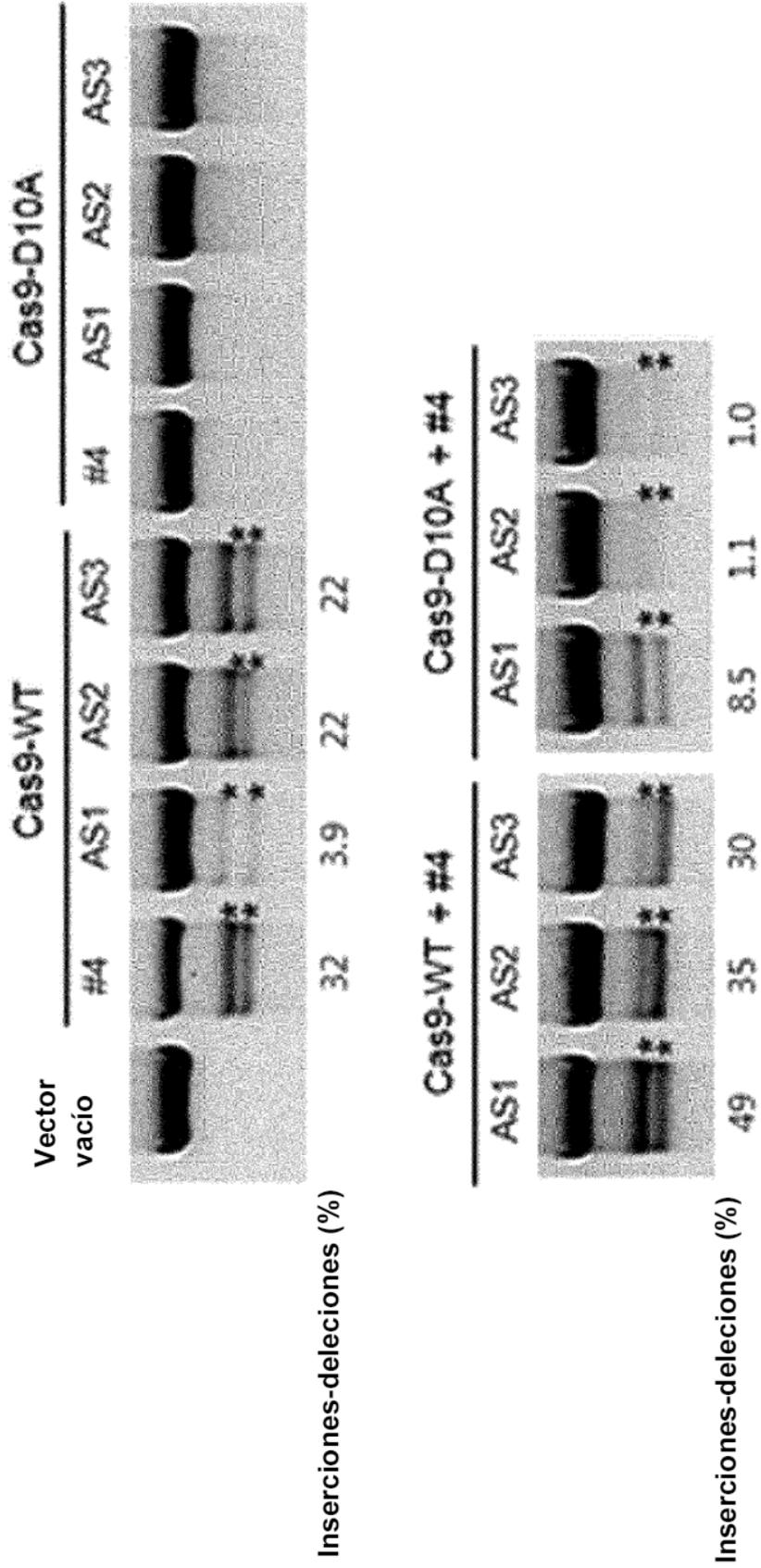
[Fig. 17b]



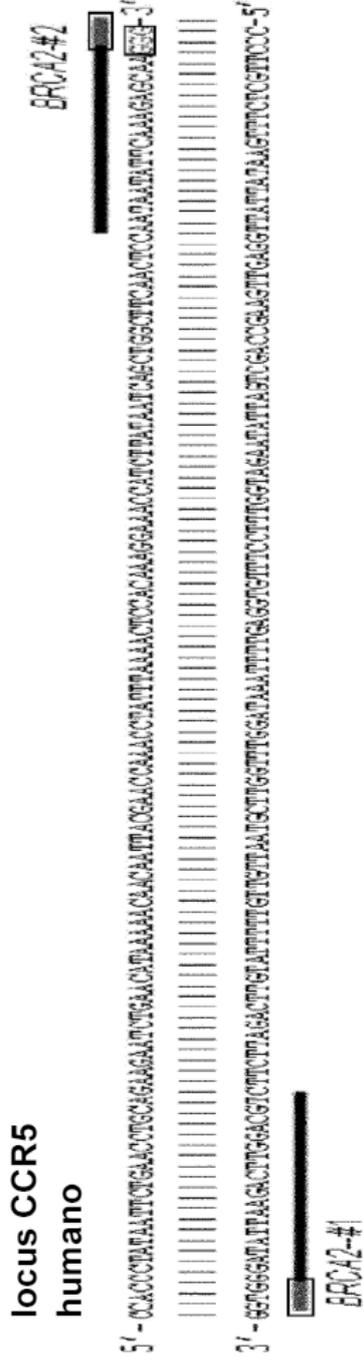
[Fig. 18a]



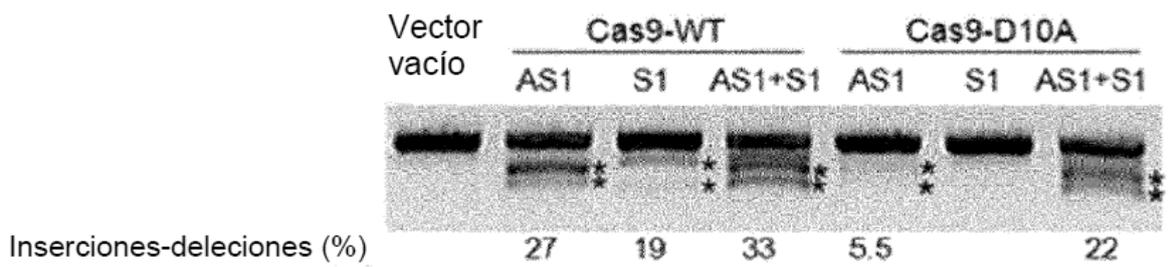
[Fig. 18b]



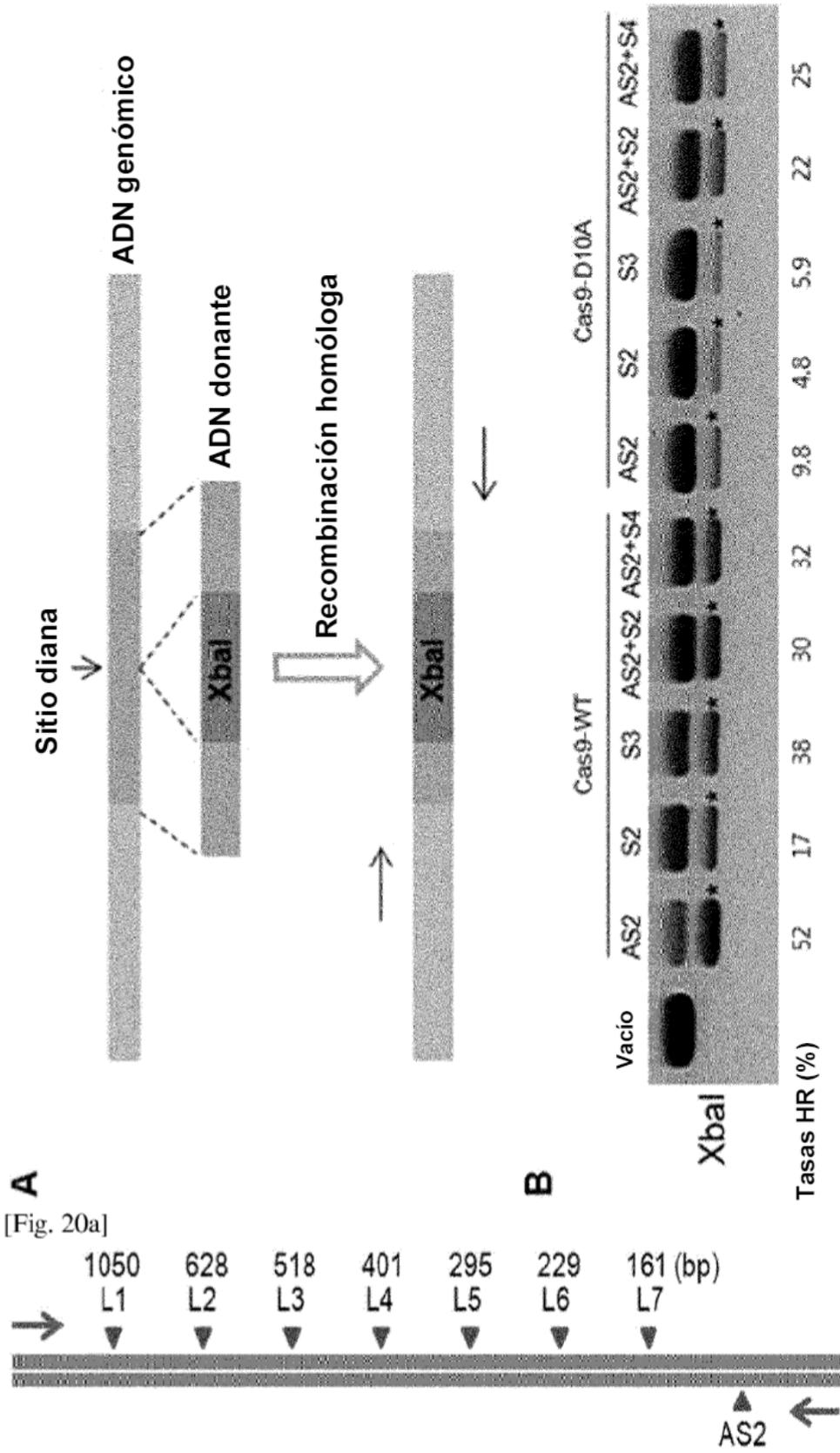
[Fig. 18c]



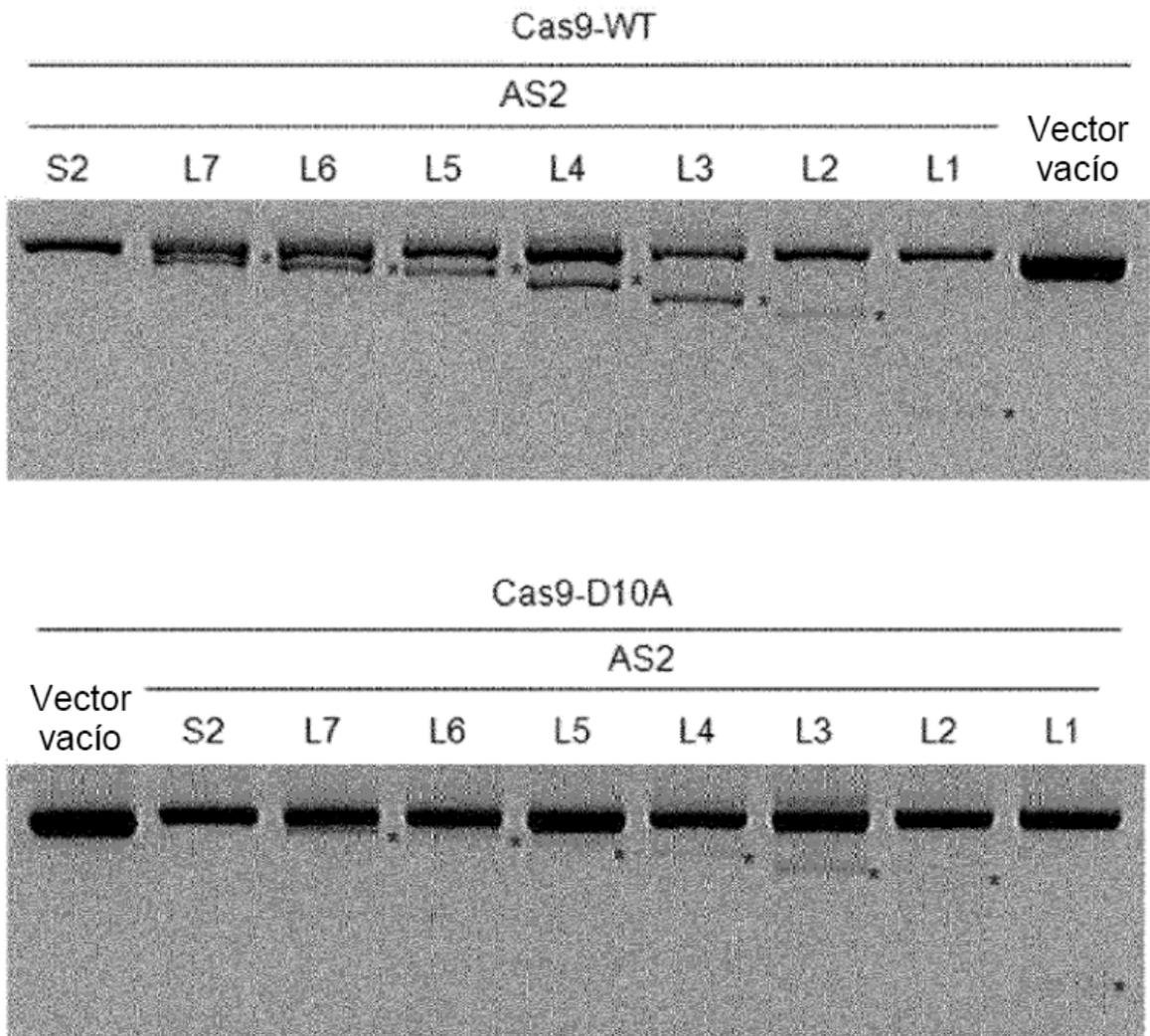
[Fig. 18d]



[Fig. 19]



[Fig. 20b]



[Fig. 20c]

**CAS9-WT AS2 + L1 (delección ~1050 bp)**

```

ggccggggaat caagagtgca CCGAGAGACAGTGACCAACCATCcctgttt...//...agctctccctcccaaggatcctctctcctcctcgttaagcaaaccttagagggtctcggccaaggagagagatg WT
ggccggggaat caagagtgca CCGAG-...--TGACCAACCATCctt-...--gtaagcaaaccttagagggtctcggccaaggagagagatg
ggccggggaat caagagtgca CCGAG-...--gaa
ggccggggaat caagagtgca CCGAGA-...--CCTCTTGGctccatcgttaagcaaaccttagagggtctcggccaaggagagagatg
ggccggggaat caagagtgca CCG-...--taacag
ggccggggaat caaga-...--cgTGGttccatcgttaagcaaaccttagagggtctcggccaaggagagagatg
ggc-...--tccatcgttaagcaaaccttagagggtctcggccaaggagagagatg
ggccggggaat caagagtgca CCGAGA-...--CTCTTGGctccatcgttaagcaaaccttagagggtctcggccaaggagagagatg

```

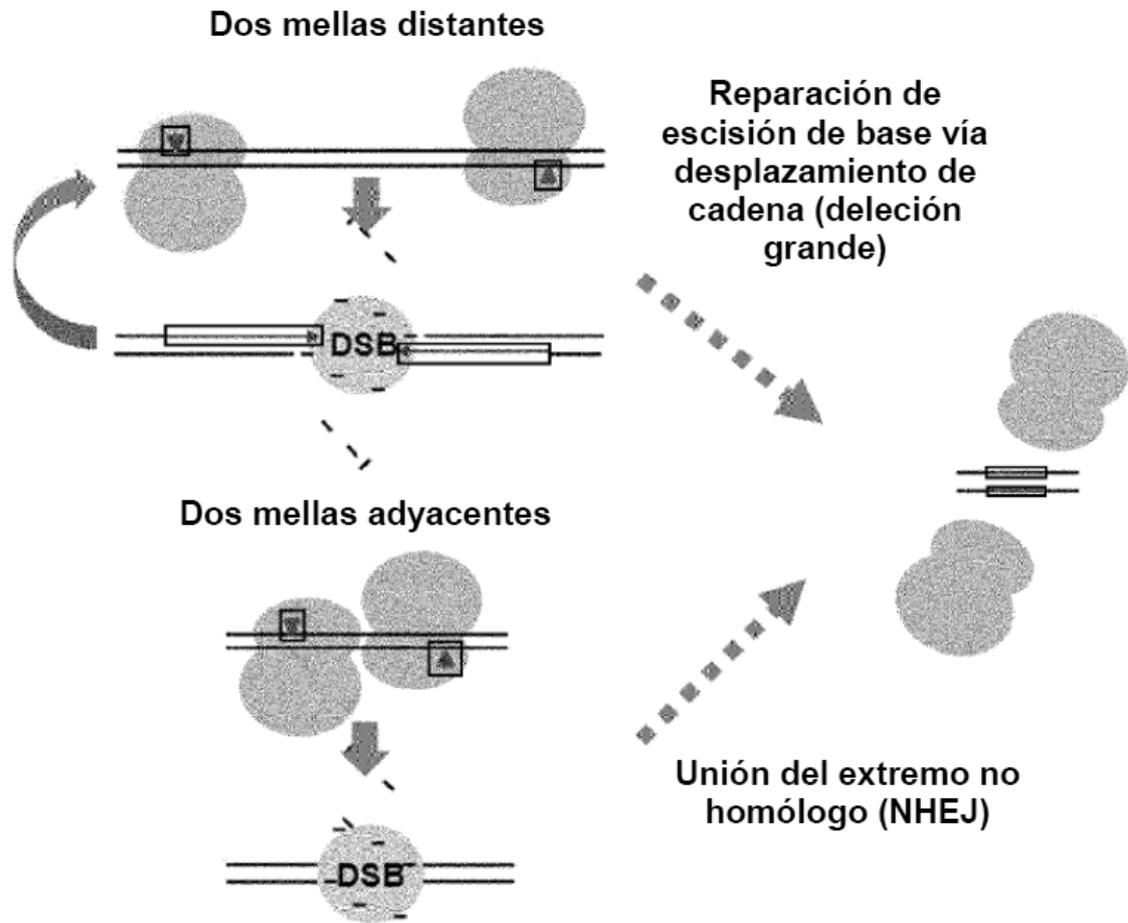
**CAS9-D10 AS2 + L1 (delección ~1050 bp)**

```

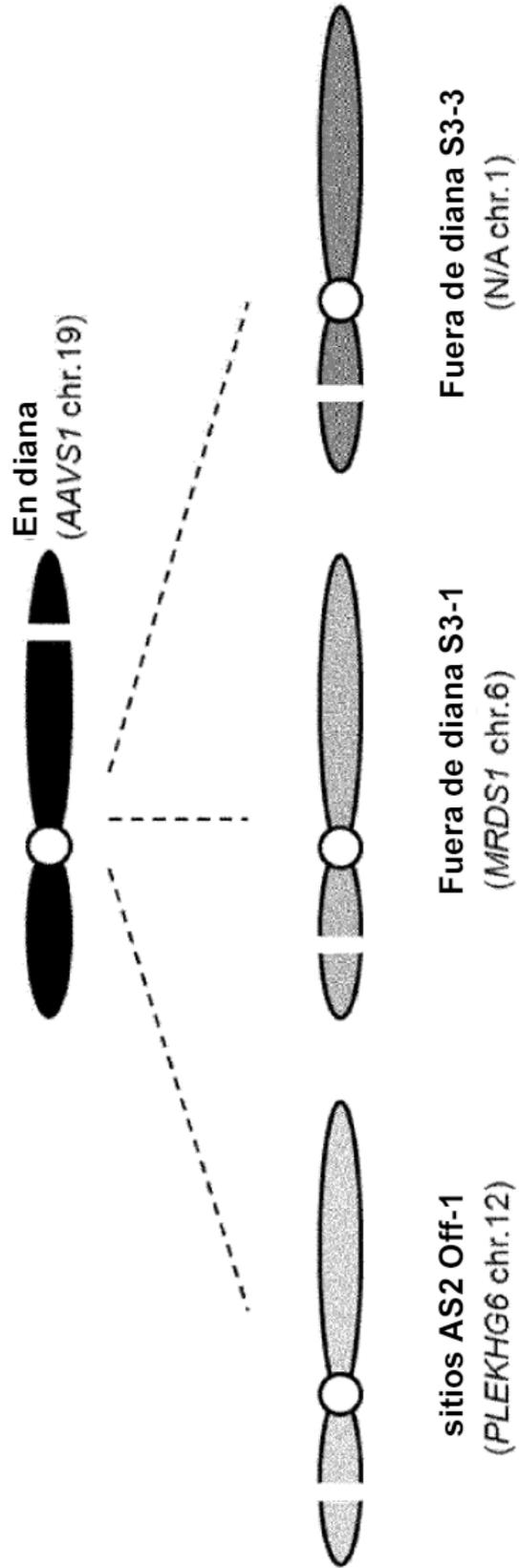
ggccggggaat caagagtgca CCGAGAGACAGTGACCAACCATCcctgttt...//...agctctccctcccaaggatcctctctcctcctcgttaagcaaaccttagagggtctcggccaaggagagagatg WT
ggccggggaat caagagtgca CCGAGAGACAGTGACCAACCATC-...--gtaagcaaaccttagagggtctcggccaaggagagagatg
gg-...--tccatcgttaagcaaaccttagagggtctcggccaaggagagagatg
ggccggggaat caagagtgca CCGA-...--tctccatcgttaagcaaaccttagagggtctcggccaaggagagagatg
ggccggggaat caagagtgca CCGAGA-...--CTTGGctccatcgttaagcaaaccttagagggtctcggccaaggagagagatg
gg-...--tccatcgttaagcaaaccttagagggtctcggccaaggagagagatg
ggccggggaat caagagtgca CCGAGAGACAGTGACCAACCATCcc-...--atcca
ggccggggaat caagagtgca-...--tgtaagcaaaccttagagggtctcggccaaggagagagatg x2

```

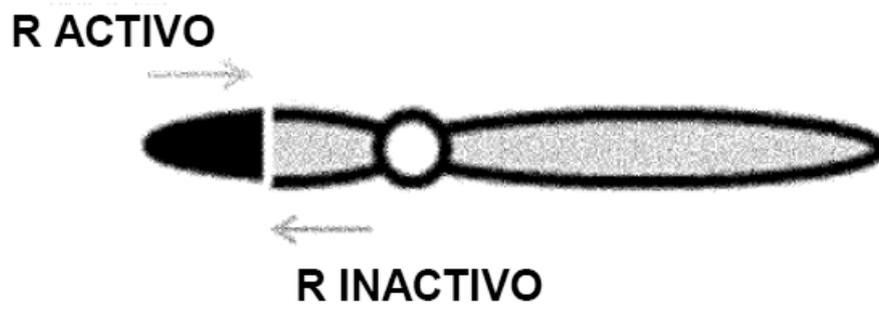
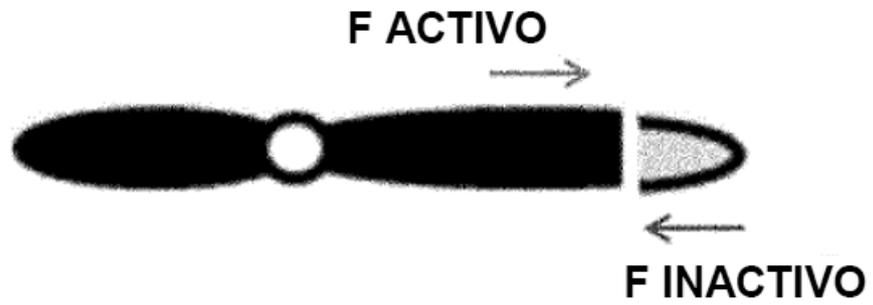
[Fig. 20d]



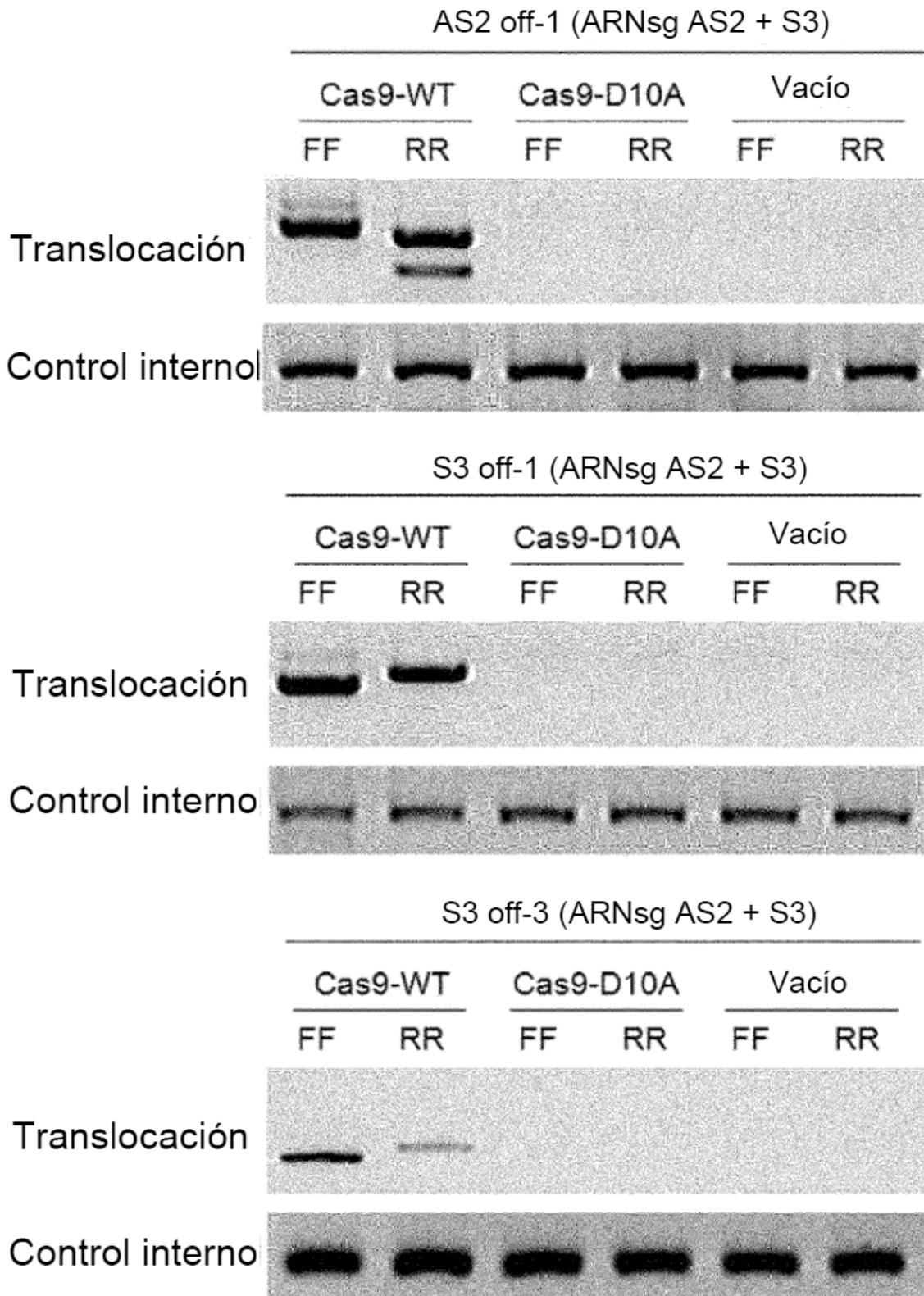
[Fig. 21a]



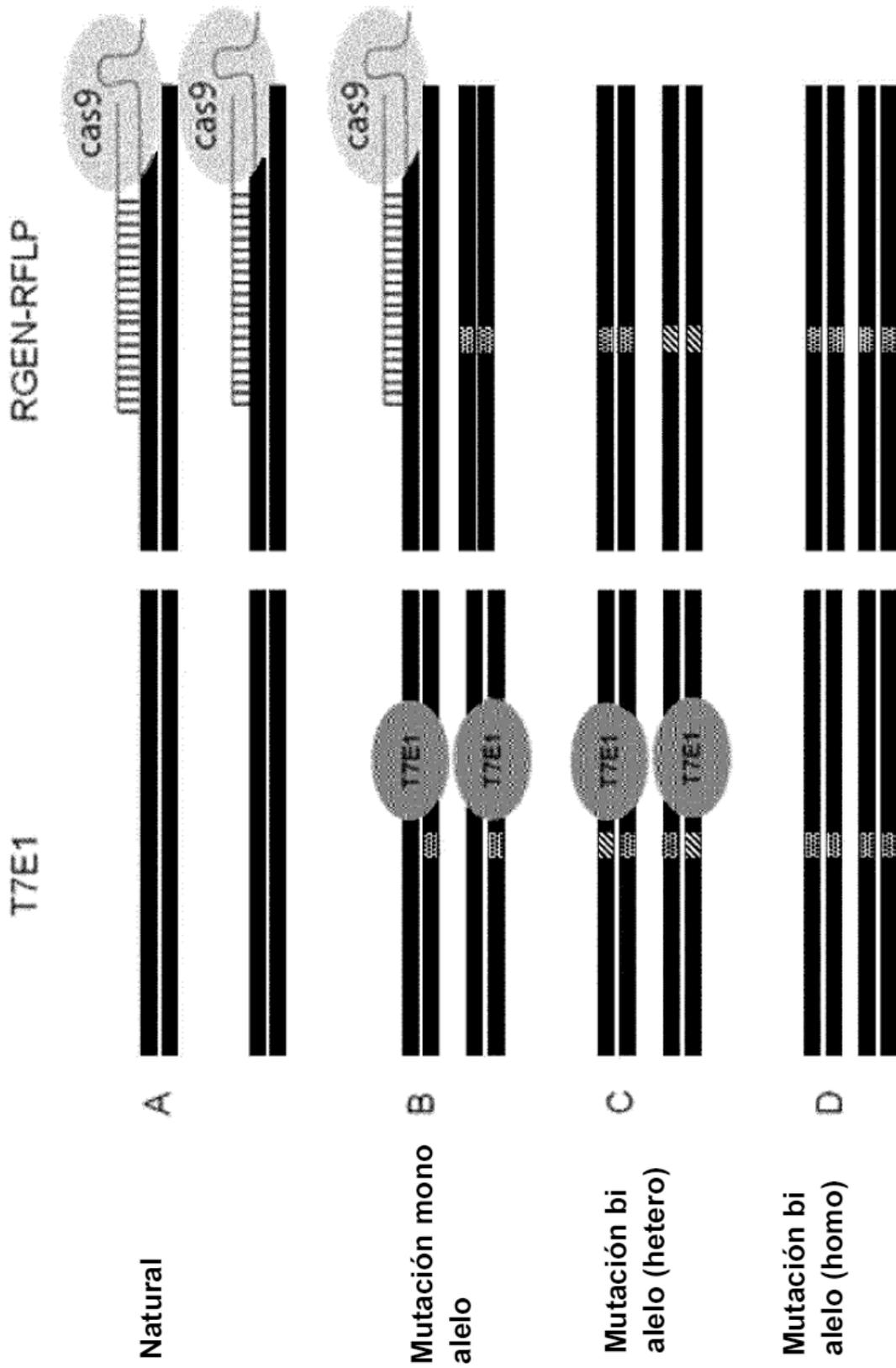
[Fig. 21b]



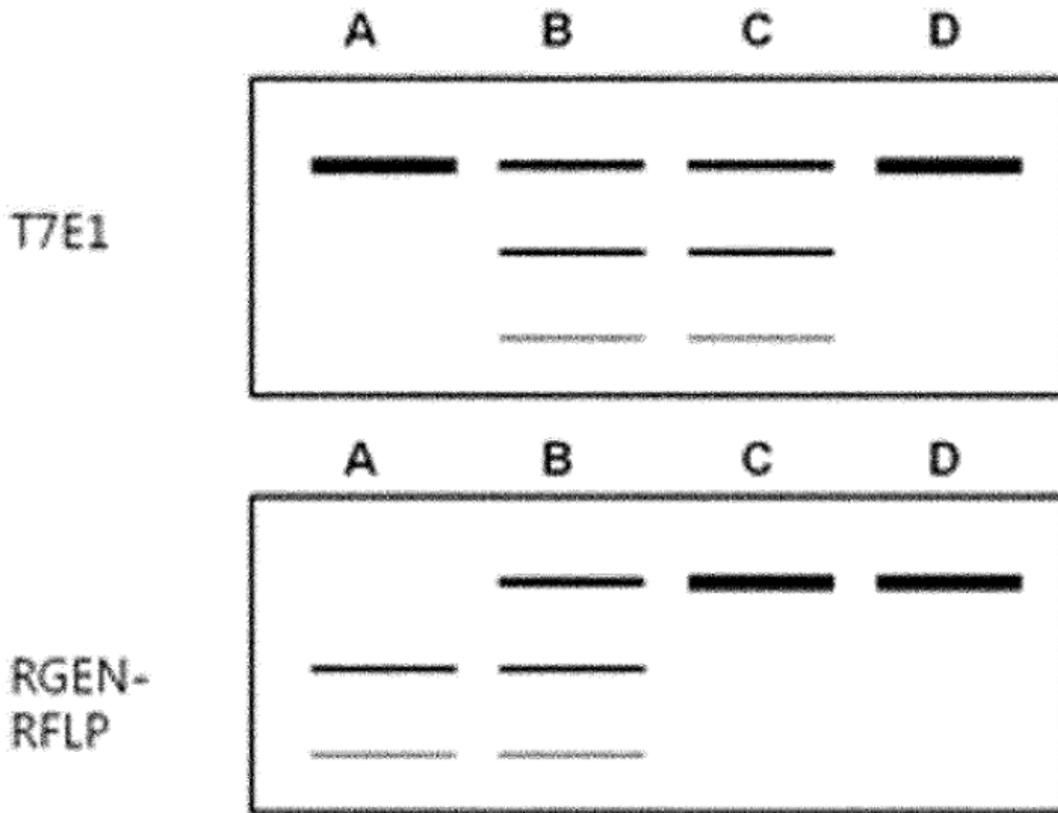
[Fig. 21c]



[Fig. 22a]



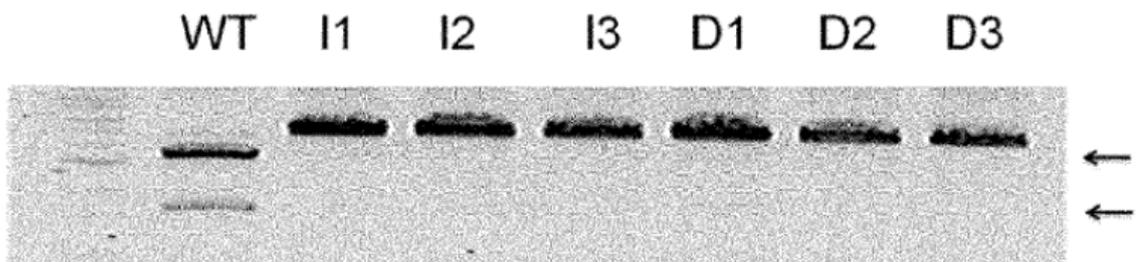
[Fig. 22b]



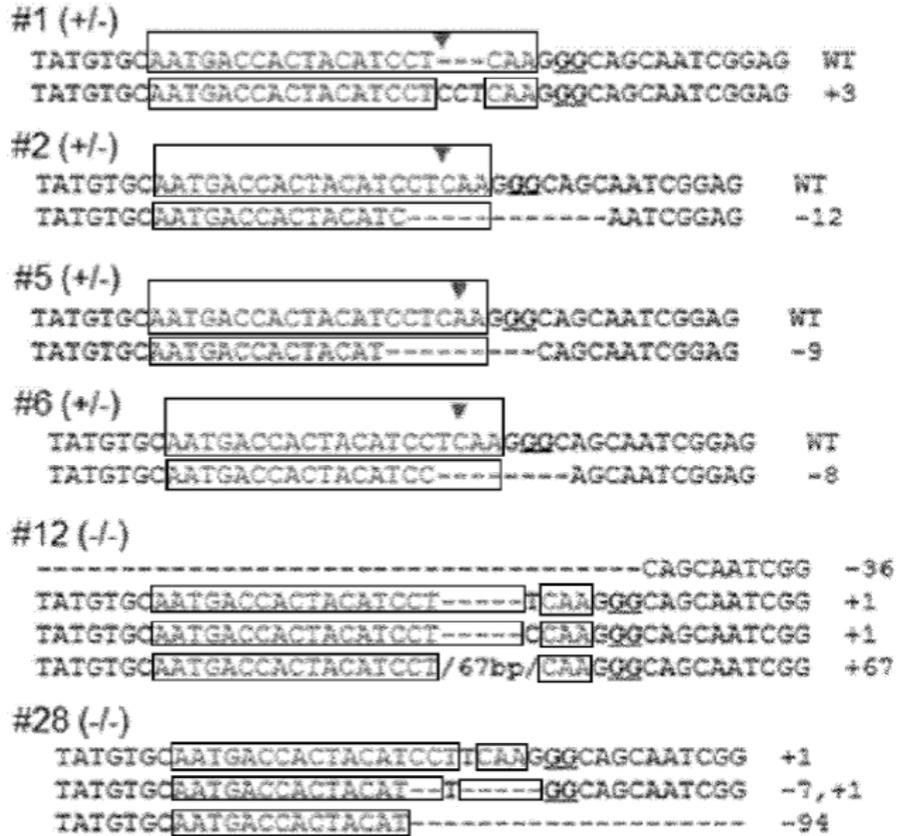
[Fig. 23]

Secuencia diana del plásmido

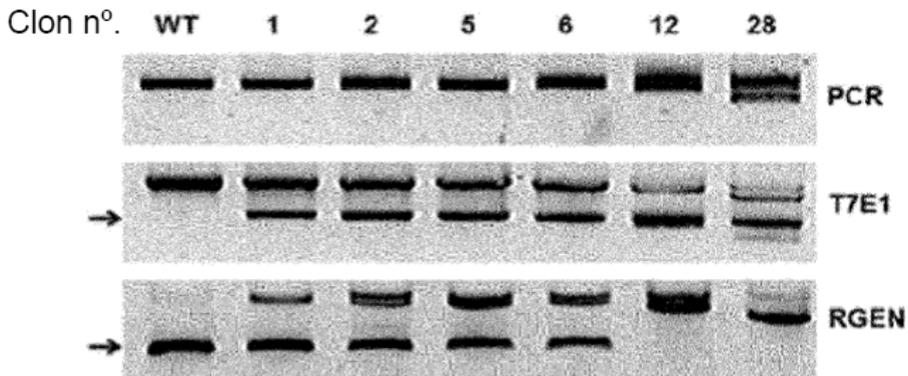
AATGACCACTACATCCT---	CAAGGG	WT
AATGACCACTACATCCT	<u>T</u> ---	I1
AATGACCACTACATCCT	<u>TT</u> ---	I2
AATGACCACTACATCCT	<u>TTT</u>	I3
AATGACCACTACATCCT----	AAGGG	D1
AATGACCACTACATCCT-----	AGGG	D2
AATGACCACTACATCCT-----	GGG	D3



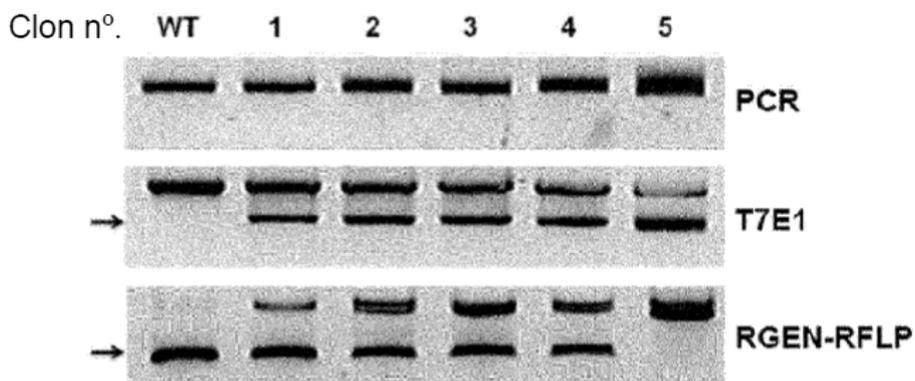
[Fig. 24a]



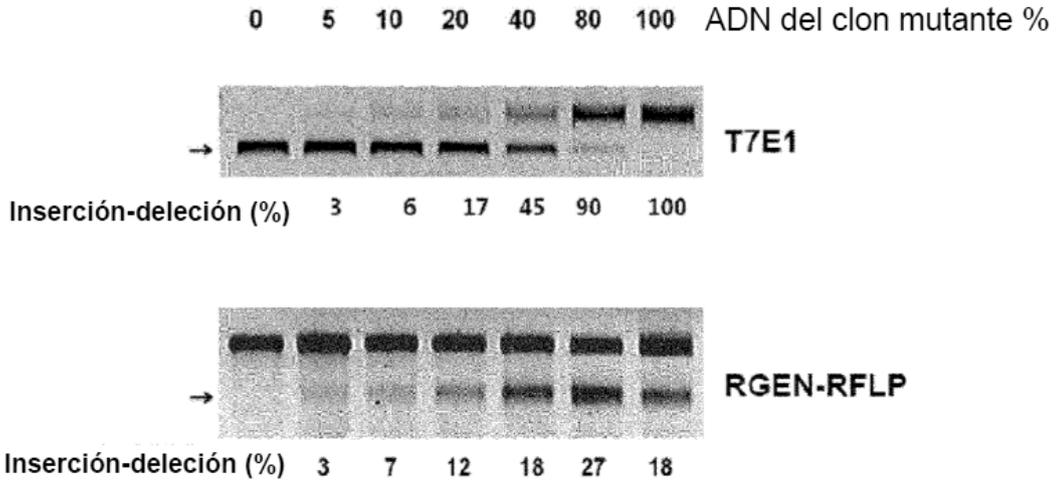
[Fig. 24b]



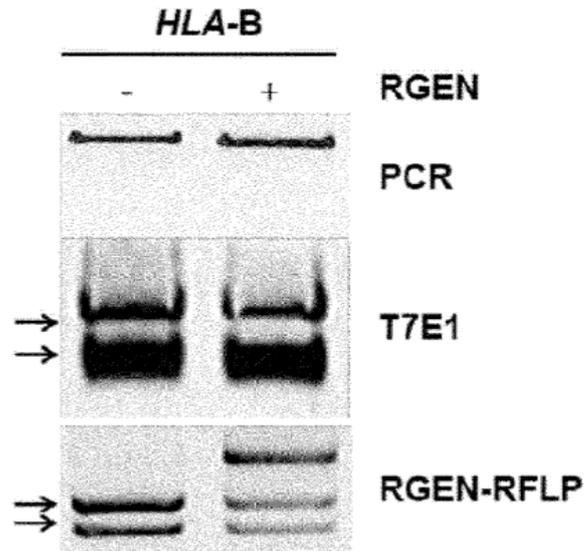
[Fig. 25a]



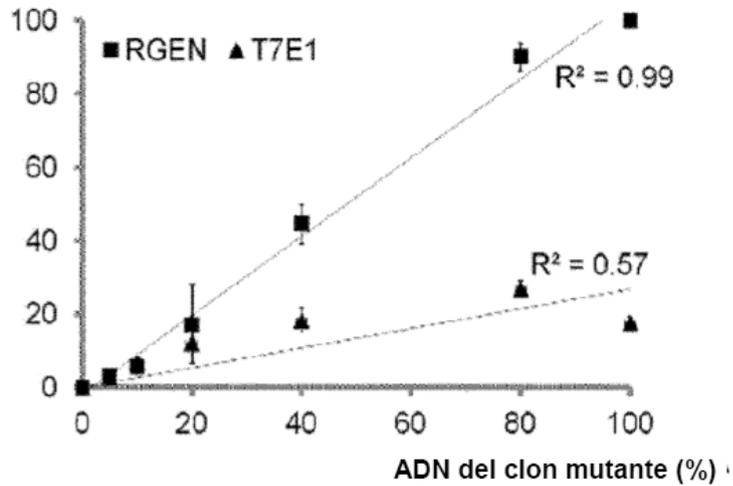
[Fig. 25b]



[Fig. 25c]



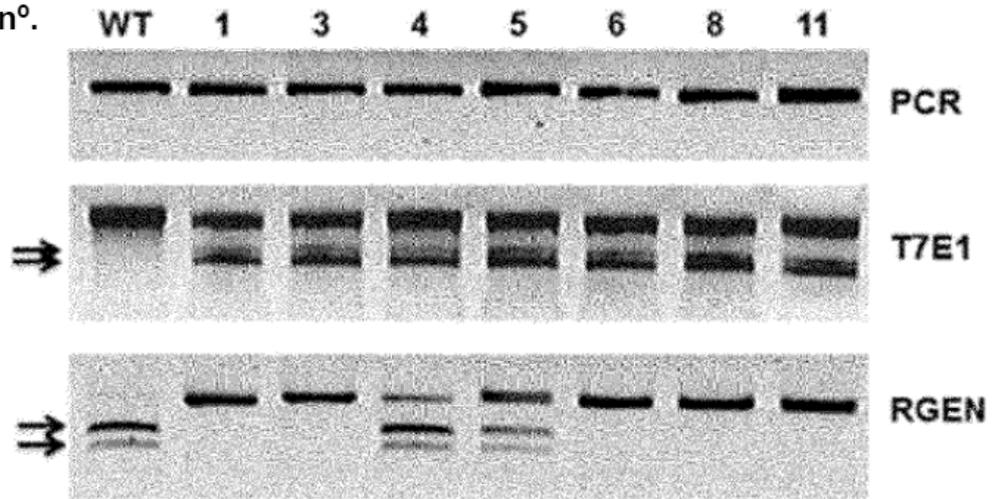
Inserción-delección (%)



[Fig. 26a]

	Natural																			
	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TCA	GAG	GAG	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
#1	T	CAT	ACA	GAT	GAT	AGC	---	---	---	---	---	---	--A	GGT	AAA	GTC	AAA	ATC	A	
	T	CAT	ACA	GGT	GAT	G	---	---	---	---	---	---	-AA	GGT	AAA	GTC	AAA	ATC	A	
#3	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TCA	GAG	---	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TCA	G	---	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
#4	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TCA	G-G	GAG	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TCA	GAG	GAG	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
#5	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TC-	---	-AG	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TCA	GAG	GAG	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
#6	T	CAT	ACA	GAT	GAT	GTC	TC-	---	---	---	---	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A	
	T	CAT	ACA	GAT	GAT	G	---	---	---	---	---	---	-AA	GGT	AAA	GTC	AAA	ATC	A	
#8	T	CAT	ACA	GAT	GA-	---	---	---	---	---	---	---	--A	GGT	AAA	GTC	AAA	ATC	A	
	T	CAT	ACA	GAT	GAT	GTC	TAG	AGA	↓	---	---	---	GAA	GGT	AAA	GTC	AAA	ATC	A	
#11	T	CAT	ACA	GAT	GAT	GTC	TCA	TCA	TCA	G	---	GAG	CGA	GAA	GGT	AAA	GTC	AAA	ATC	A
	T	CAT	ACA	GAT	GA-	---	---	---	---	---	---	---	--A	GGT	AAA	GTC	AAA	ATC	A	

[Fig. 26b]  
Fundador nº.

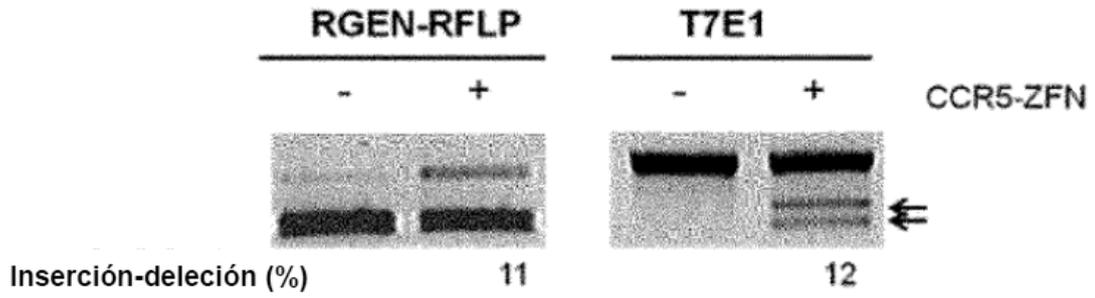


[Fig. 27]

CCR5-ZFN

GTCATCCTCATCCTGATAAACTGCAAAAGGCTGA

Sitio diana-RGEN-RFLP



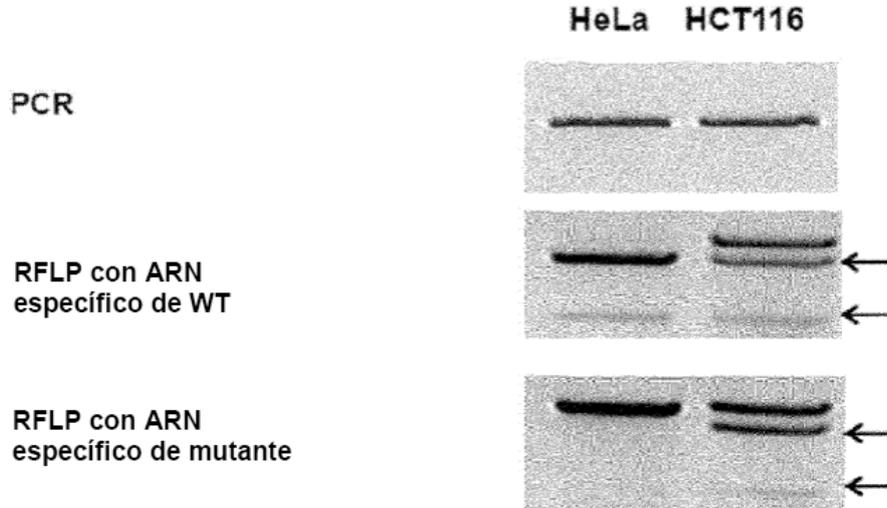
[Fig. 28]

**Cebador directo** →

GCTGGTGTCTGGGTTCTGTGCCCCCTTCCCCACCCAGCCACCCAGGTGTCTGTCCATTCTCAGGCTG  
GTCACATGGGTGGTCCTAGGGTGTCCCATGAGAGATGCAAGCGCCTGAATTTTCTGACTCTTCCCATCA  
GACCCCAAAGACAATGTGACCCACCACCCATCTTGACCATGAGGCCACCCTGAGGTGCTGGGCC  
TGGGTTCTACCCTGGGAGATCACACTGACCTGGCAGCGGGATGGCAGGACCAAACTCAGGACACGA  
GCTTGTGGAGACCAGACCAGGAGATAGAACCTTCCAGAAGTGGCAGCTGTGGTGGTCCTCTGGA  
GAGAGCAGAGATACATGCATGCCATGAGGGCTGCCGAAGCCCTCACCCTGAGATGGGTA  
AGGAGGGGGATGAGGGGTCATATCTGTCATATCTGTTCAGGGAAGCAGGAGCCTTCTGGAGCCCTTCTGGAGCCCTTCT  
TCAGCAGGGTCAGGCCCTCATCTTCCCCTCCTTCCCAGAGCCATCTCCCAGTCCCATCCCCATC  
GTGGGCATTGTTGCTGGCCTGGCTGTCCTAGCAGTTGTGGTCATCG

← **Cebador inverso**

[Fig. 29a]



**HeLa**

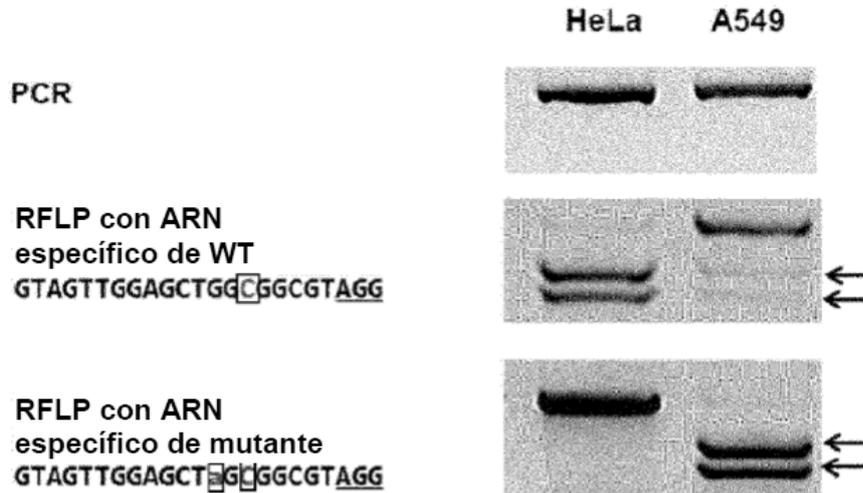
ACTACCACAGCTCCTTCTCTGAGTGG tipo natural

**HCT116**

ACTACCACAGCTCCTTCTCTGAGTGG tipo natural

ACTACCACAGCTCCT---CTGAGTGG c.133-135 del TCT

[Fig. 29b]



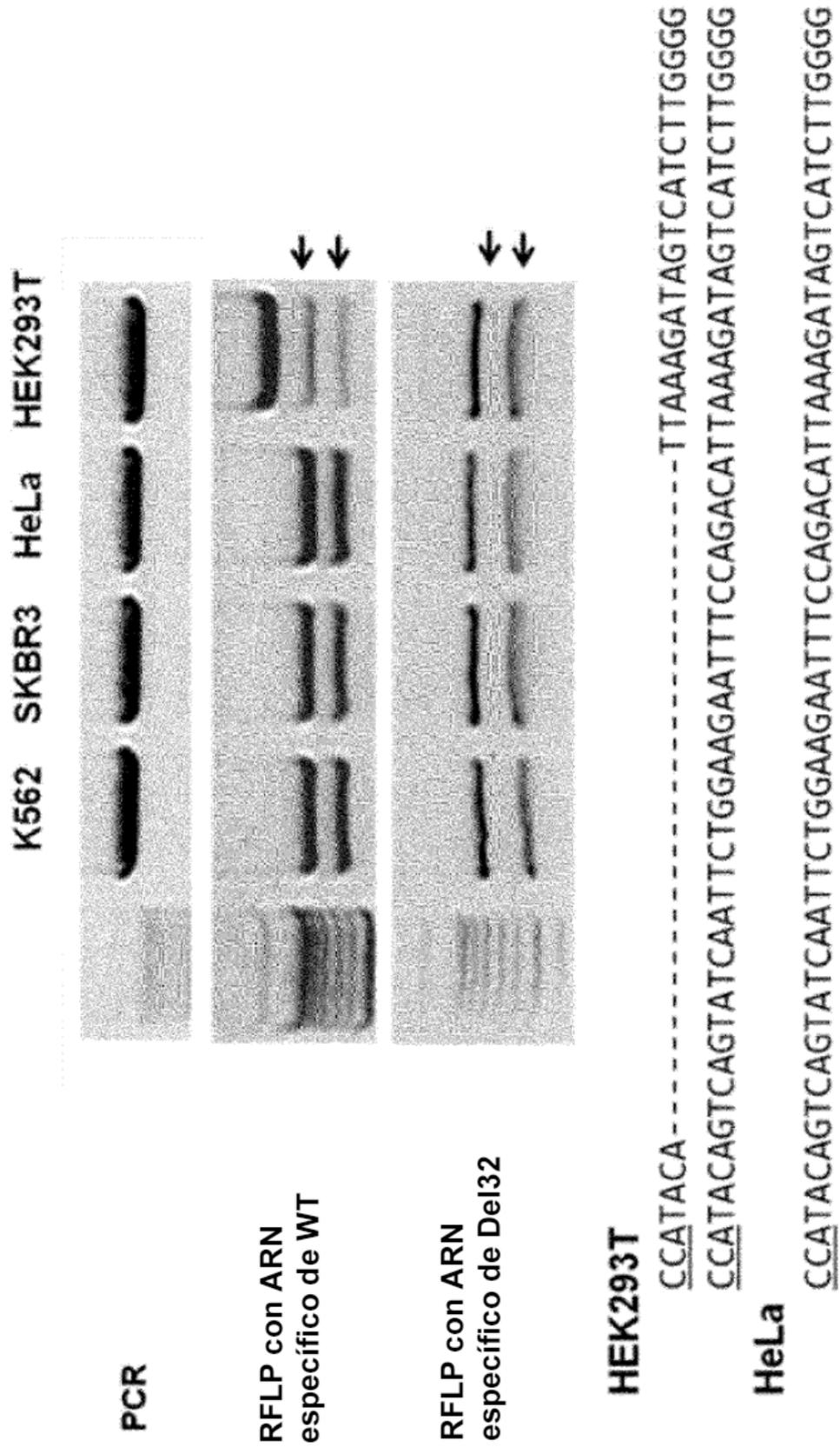
**HeLa**

GTAGTTGGAGCTGGTGGCGTAGG wild-type

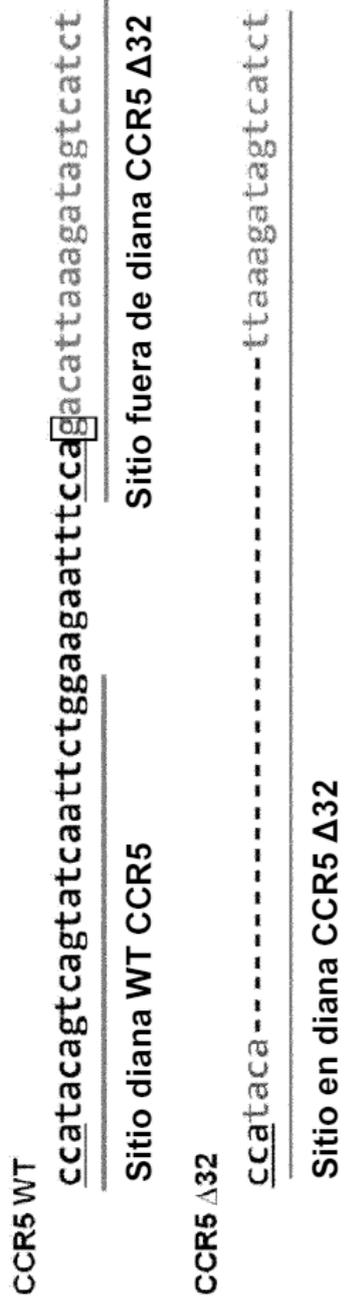
**A549**

GTAGTTGGAGCTG GTGGCGTAGG c.34G>A

[Fig. 30a]



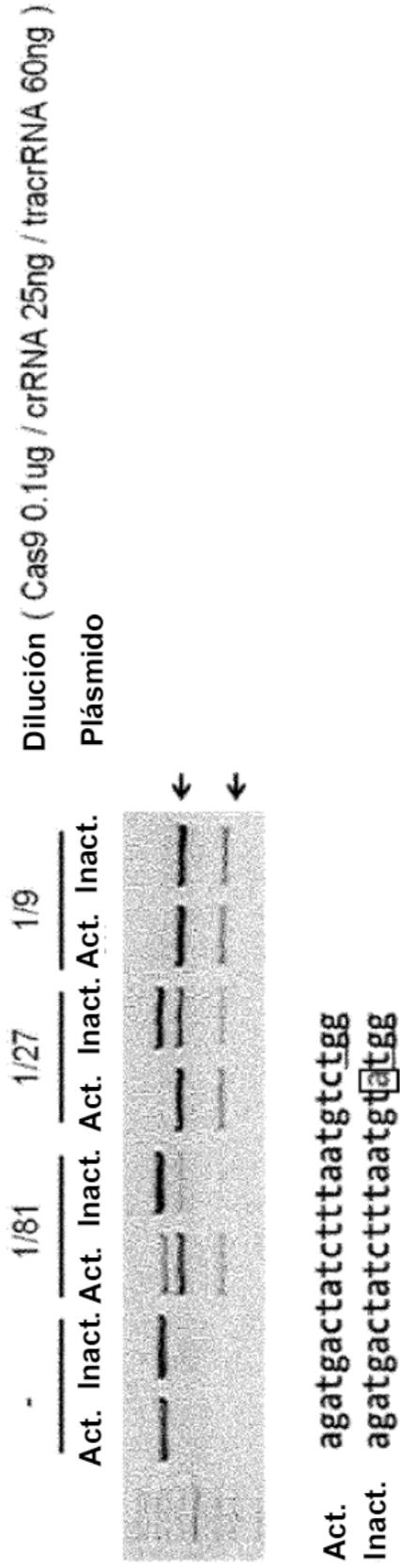
[Fig. 30b]



[Fig. 30c]



[Fig. 30d]



[Fig. 31a]

**KRAS**

**HeLa**

GTAGTTGGAGCTGGTGGCGTAGG Tipo natural

**A549**

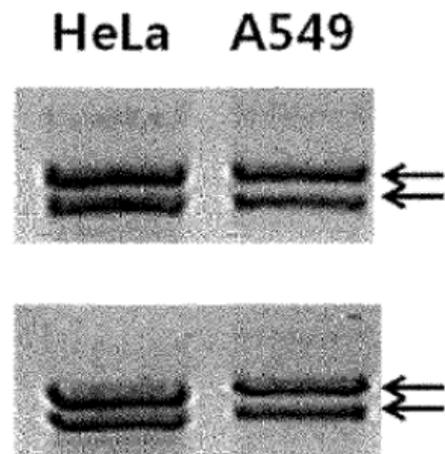
GTAGTTGGAGCT<sup>a</sup>GTGGCGTAGG c.34G>A

**RFLP con ARN  
específico de WT**

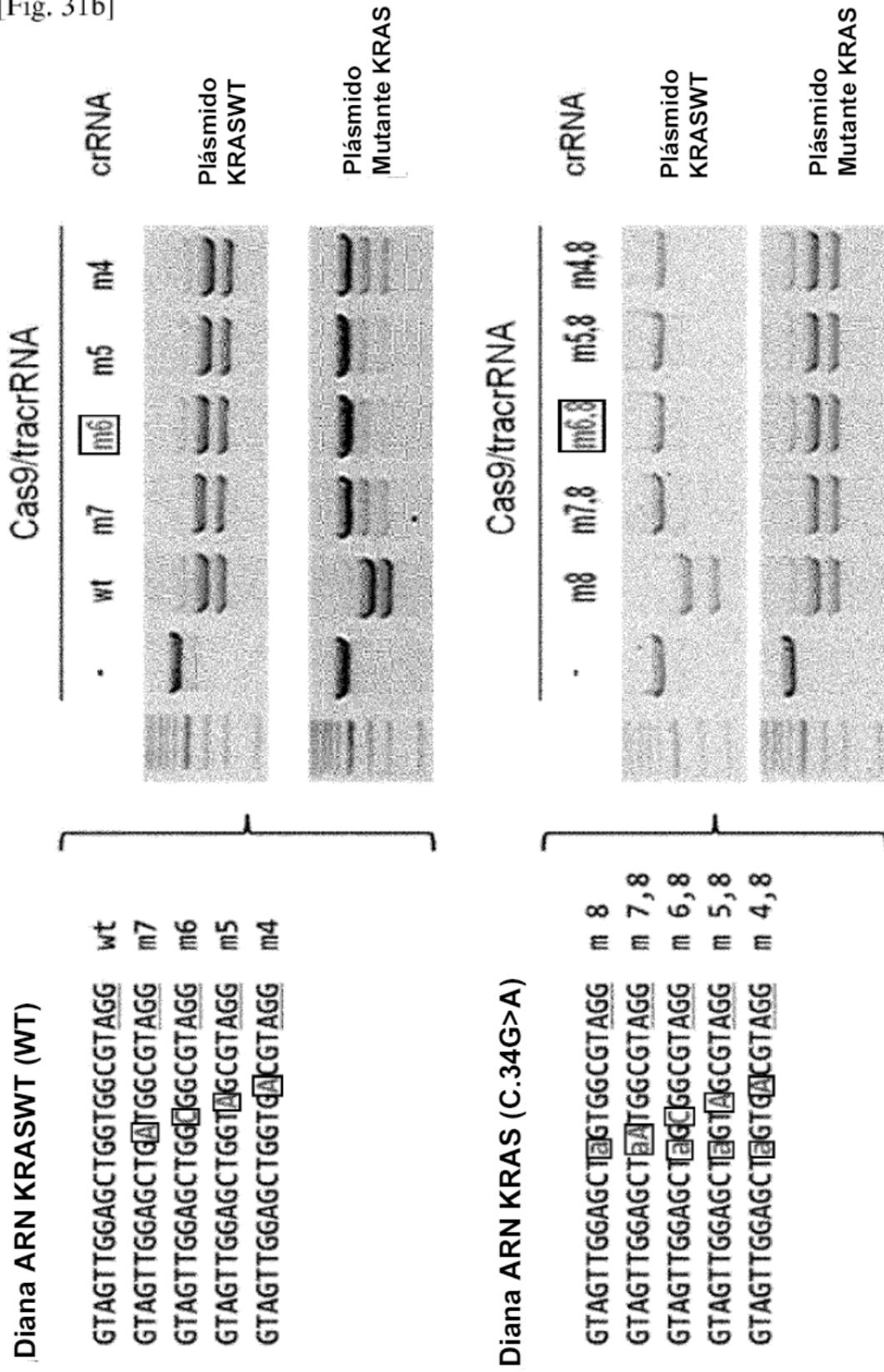
GTAGTTGGAGCTGGTGGCGTAGG

**RFLP con ARN  
específico de mutante**

GTAGTTGGAGCT<sup>a</sup>GTGGCGTAGG



[Fig. 31b]



[Fig. 32a]

**PIK3CA**

**HeLa**

CAAATGAATGATGCACATCATGG **Tipo natural**

**HCT116**

CAAATGAATGATGCACATCATGG **Tipo natural**

CAAATGAATGATGCACGTCATGG **C.3140A>G**

**RFLP con ARN  
específico de WT**

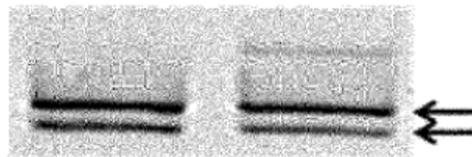
CAAATGAATGATGCACATCATGG

**HeLa      HCT116**

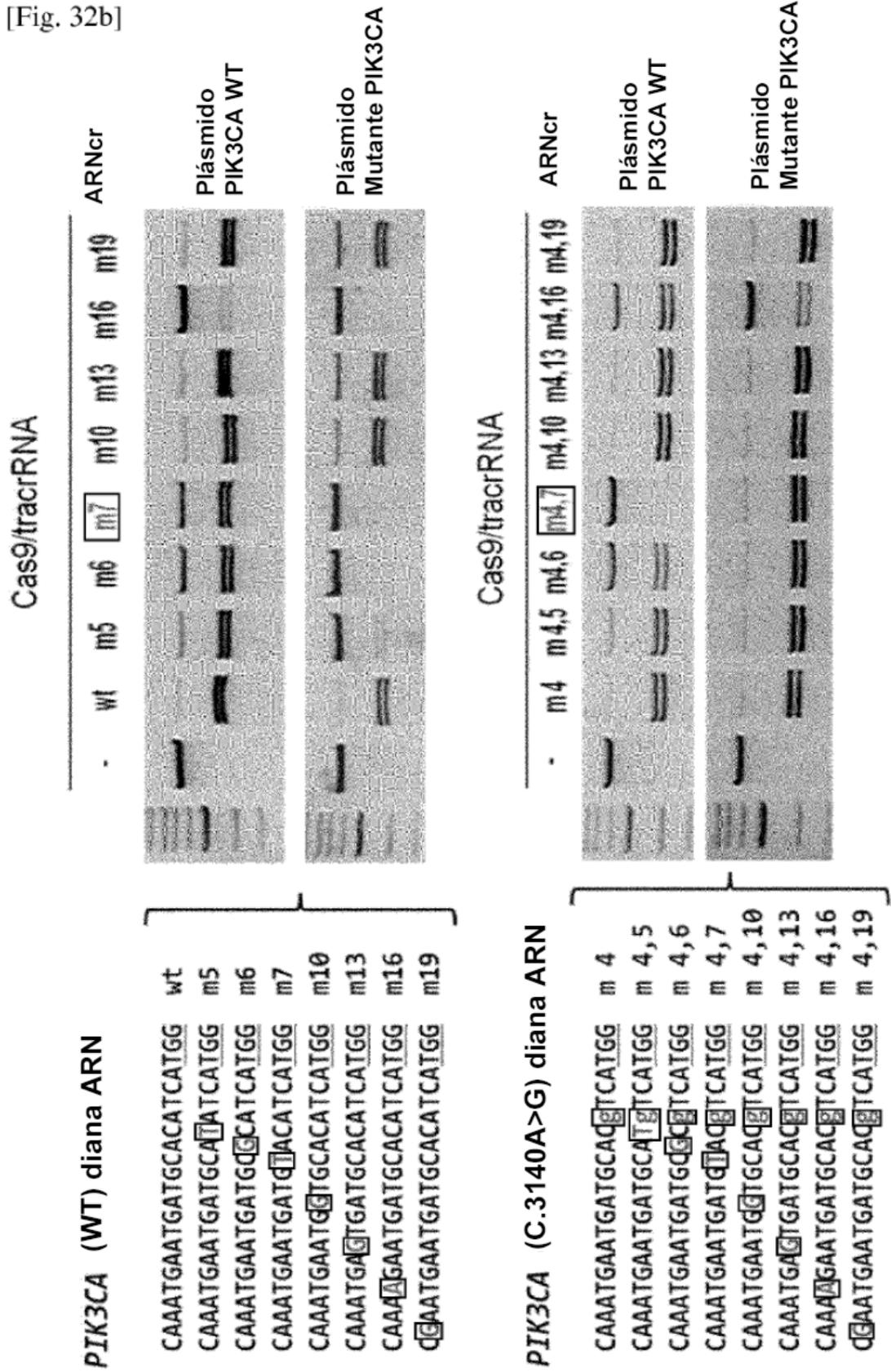


**RFLP con ARN  
específico de mutante**

CAAATGAATGATGCACGTCATGG



[Fig. 32b]



[Fig. 33a]

**IDH1**

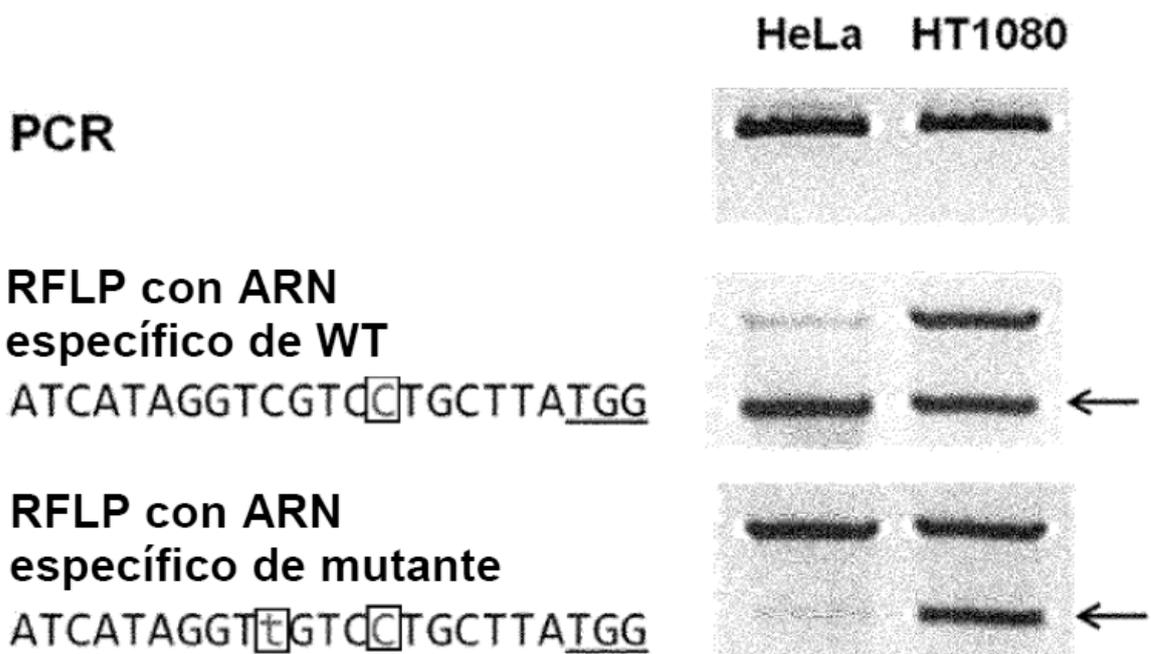
**HeLa**

ATCATAGGTCGTCATGCTTATGG **Tipo natural**

**HT1080**

ATCATAGGTCGTCATGCTTATGG **Tipo natural**

ATCATAGGT $\boxed{t}$ GTCATGCTTATGG **c.394C>T**



[Fig. 33b]

**PIK3CA**

**HeLa**

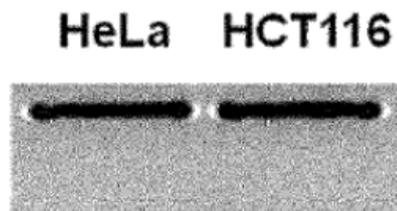
CAAATGAATGATGCACATCATGG **Tipo natural**

**HCT116**

CAAATGAATGATGCACATCATGG **Tipo natural**

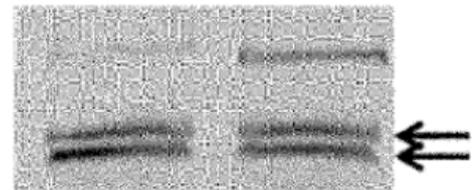
CAAATGAATGATGCACAGTCATGG **C.3140A>G**

**PCR**



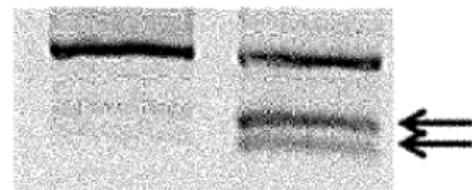
**RFLP con ARN  
específico de WT**

CAAATGAATGATGTACATCATGG



**RFLP con ARN  
específico de mutante**

CAAATGAATGATGTACAGTCATGG



[Fig. 33c]

**NRAS**

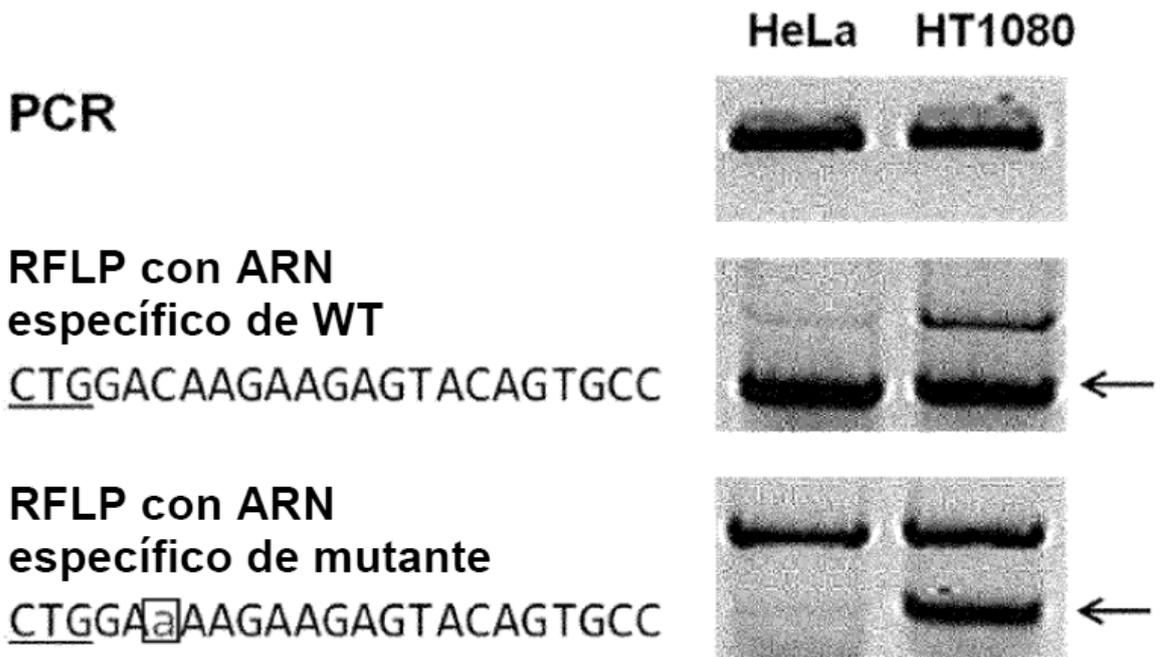
**HeLa**

CTGGACAAGAAGAGTACAGTGCC Tipo natural

**HT1080**

CTGGACAAGAAGAGTACAGTGCC Tipo natural

CTGGGA<sup>a</sup>AAGAAGAGTACAGTGCC c.181C>A



[Fig. 33d]

**BRAF**

**HeLa**

ACTCCATCGAGATTTCACTGTAG Tipo natural

**HT29**

ACTCCATCGAGATTTCACTGTAG Tipo natural

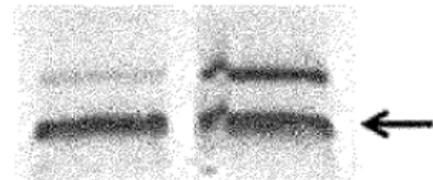
ACTCCATCGAGATTTCTTCTGTAG (c.1799T>A)

**PCR**



**RFLP con ARN específico de WT**

ACTCCATCGAGATTTCACTGTAG



**RFLP con ARN específico de mutante**

ACTCCATCGAGATTTCTTCTGTAG

