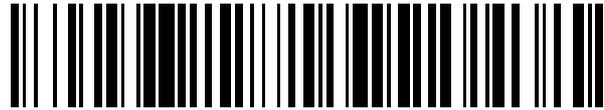


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 705 540**

51 Int. Cl.:

C12N 15/11 (2006.01)

C12N 15/113 (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.08.2016 PCT/US2016/045915**

87 Fecha y número de publicación internacional: **16.02.2017 WO17027423**

96 Fecha de presentación y número de la solicitud europea: **05.08.2016 E 16753538 (4)**

97 Fecha y número de publicación de la concesión europea: **17.10.2018 EP 3320092**

54 Título: **Composiciones de CRISPR-Cas9 modificadas por ingeniería genética y procedimientos de uso**

30 Prioridad:

07.08.2015 US 201562202715 P
24.08.2015 US 201562209334 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.03.2019

73 Titular/es:

CARIBOU BIOSCIENCES, INC. (100.0%)
2929 7th Street Suite 105
Berkeley, CA 94710, US

72 Inventor/es:

DONOHUE, PAUL DANIEL y
MAY, ANDREW PAUL

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 705 540 T3

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

DESCRIPCIÓN

Composiciones de CRISPR-Cas9 modificadas por ingeniería genética y procedimientos de uso

Campo técnico

5 La presente invención se refiere a polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética, sistemas, composiciones y procedimientos de uso.

Antecedentes de la invención

10 La modificación genómica incluye alterar el genoma mediante supresión, inserción, mutación o sustitución de secuencias de ácido nucleico específicas. La alteración puede ser específica de gen o localización. La modificación genómica puede usar nucleasas para cortar ADN, generando de este modo un sitio para alteración. En casos determinados, la escisión puede introducir roturas de doble cadena en el ADN diana. Las roturas de doble cadena pueden repararse, por ejemplo, mediante unión de extremos no homólogos (UENH) endógena o reparación dirigida por homología (RDH). La RDH se basa en la presencia de un molde para reparación. En algunos ejemplos de modificación genómica, un polinucleótido donante, o parte del mismo, puede insertarse en la rotura.

15 Las repeticiones palindrómicas cortas agrupadas y espaciadas de forma regular (CRISPR) y proteínas Cas asociadas constituyen el sistema CRISPR-Cas. Este sistema proporciona inmunidad adaptativa contra ADN extraño en bacterias (Barrangou, R., y col., "CRISPR provides acquired resistance against viruses in prokaryotes," *Science* 315, 1709-1712 (2007); Makarova, K. S., y col., "Evolution and classification of the CRISPR-Cas systems," *Nat Rev Microbiol* 9, 467-477 (2011); Garneau, J. E., y col., "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA," *Nature* 468, 67-71 (2010); Sapranasuskas, R., y col., "The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*," *Nucleic Acids Res* 39, 9275-9282 (2011)). La endonucleasa Cas9 guiada por ARN se dirige y escinde específicamente ADN de una manera dependiente de secuencia (Jinek, M., y col., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science* 337, 816-821 (2012); Sternberg, S. H., y col., "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9," *Nature* 507, 62 (2014)), y se ha usado ampliamente para edición de genoma programable en una diversidad de organismos y sistemas modelo (Publicación de Solicitud de Patente de los Estados Unidos n.º 2015-0284727, publicada el 8 de octubre de 2015; Cong, L., y col., "Multiplex genome engineering using CRISPR/Cas systems," *Science* 339, 819-823 (2013); Jiang, W., y col., "RNA-guided editing of bacterial genomes using CRISPR-Cas systems," *Nat. Biotechnol.* 31, 233-239 (2013); Sander, J. D. y Joung, J. K., "CRISPR-Cas systems for editing, regulating and targeting genomes," *Nature Biotechnol.* 32, 347-355. (2014)).

20 30 Jinek, M., y col., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science* 337(6096), 816-21 (2012) mostró que en un subconjunto de sistemas asociados con CRISPR (Cas) el CRISPR maduro (ARNcr) que forma pares de bases con el ARNcr de activación en trans (ARNcrtra) forma una estructura de ARN de dos partes que dirige la proteína asociada a CRISPR Cas9 para introducir roturas de doble cadena en ADN diana. En sitios complementarios de la secuencia guía de ARNcr (espaciador), el dominio de nucleasa HNH de Cas9 escinde la cadena complementaria y el dominio de tipo RuvC de Cas9 escinde la cadena no complementaria. En 35 Jinek, M., y col., se modificaron moléculas dobles de ARNcr/ARNcrtra a moléculas monocatenarias de ARNcr/ARNcrtra y estas ARNcr/ARNcrtra dirigieron la escisión de ADN bicatenario por Cas9 específica de secuencia diana.

40 Jinek, M., y col., diseñaron dos versiones de ARNcr/ARNcrtra monocatenario que contenían una secuencia de reconocimiento de diana (espaciador) en el extremo 5' seguida de una estructura en horquilla que conserva las interacciones de formación de pares de bases que normalmente aparecen entre el ARNcrtra y el ARNcr (véase FIG. 5B de Jinek, M., y col.). Para cada ARNcr/ARNcrtra monocatenario, el extremo 3' de ARNcr se unió covalentemente con el extremo 5' de ARNcrtra. En ensayos de escisión que usan ADN plasmídico, Jinek, M., y col., observaron que un ARNcr/ARNcrtra monocatenario truncado en 3' no escindía ADN diana tan eficazmente en el ensayo como un 45 ARNcr/ARNcrtra monocatenario más largo que no estaba truncado en el extremo 3' (véase FIG. 5B y FIG. S14 A, B y C de Jinek, M., y col.). Estos datos confirmaron que las "5 a 12 posiciones más allá de la interacción de formación de pares de bases de ARNcrtra:ARNcr son importantes para la unión y/o el reconocimiento de diana de Cas9 eficaces" (Jinek, M., y col., *Science* 337(6096), 820 (2012)).

50 Briner, A., y col., "Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality," *Molecular Cell* 56(2), 333-339 (2014) dilucidaron la base molecular de las interacciones de Cas9/ARN guía selectivas identificando y caracterizando módulos estructurales y de secuencia definidos en ARN guías que dirigen la actividad endonucleasa de Cas9 y definen la ortogonalidad. Establecieron seis módulos en dobles cadenas de ARNcr:ARNcrtra nativas y ARN guías individuales (ARNgi) a lo largo de cuarenta y un sistemas de tres familias de Cas9 distintas. Los seis módulos identificados son el espaciador, el tallo inferior, la protuberancia, el tallo superior, la horquilla de nexos y 55 horquillas 3'. Estos módulos se ilustran en referencia a un ARNgi en la FIG. 2.

Usando el sistema de ARNgi/Cas9 de *Streptococcus pyogenes*, Briner, A., y col., mostraron que una protuberancia en el ARNgi es estructuralmente necesaria para escisión de ADN tanto *in vitro* como *in vivo*, mientras que se toleran sustituciones de secuencias en otras regiones. Asimismo, pueden retirarse elementos prescindibles para generar

ARNgi en miniatura funcionales. También identificaron un módulo conservado "denominado el nexo; este elemento muestra elementos de secuencia y estructurales importantes para la escisión" (Briner, A., y col., página 2). Indicaron que este módulo, el nexo, es "necesario para la escisión de ADN" (Briner, A., y col., Sumario). La horquilla de nexo confiere actividad a su Cas9 afín. La localización de esta horquilla de nexo corresponde a las 5 a 12 posiciones más allá de la interacción de formación de pares de bases de ARNcrtra:ARNcr que son importantes para la unión y/o el reconocimiento de diana de Cas9 eficaces como identificaron Jinek, M., y col. (véase anteriormente).

Briner, A., y col., mostraron que la forma de horquilla de nexo general con un tallo rico en GC y un uracilo desplazado estaba compartida entre las dos familias de *Streptococcus*. En cambio, el característico tallo doble de la horquilla de nexo era único de, y ubicuo en, sistemas de *Lactobacillus*. Algunas bases en la horquilla de nexo estaban estrictamente conservadas incluso entre familias distintas, incluyendo A52 y C55, lo que destaca adicionalmente el papel importante de este módulo. En la estructura cristalina de SpyCas9 (es decir, Cas9 de *Streptococcus pyogenes*), A52 interacciona con la cadena principal de restos 1103-1107 cerca del extremo 5' de la cadena diana en la estructura cristalina de SpyCas9, lo que sugiere que la interacción de la horquilla de nexo con la cadena principal proteica puede ser necesaria para la unión de motivo adyacente a protoespaciador (MAP).

Wright, A. V., y col., "Rational design of a split-Cas9 enzyme complex," PNAS 112(10), 2984-2989 (2015) determinaron los determinantes moleculares de ARN de motivos de ARNgi que promueven la heterodimerización de los lóbulos α -helicoidal y de nucleasa para formar un complejo ternario. Las estructuras cristalinas de ARNgi/Cas9 unida a ADN mostraron que el espaciador y los motivos de tallo-bucle (es decir, los módulos del tallo inferior, la protuberancia y el tallo superior descritos por Briner, A., y col.) en el extremo 5' del ARNgi entran en contacto principalmente con el lóbulo α -helicoidal, mientras que las dos horquillas (es decir, el módulo de horquillas descrito por Briner, A., y col.) en el extremo 3' se unen con la cara externa del lóbulo de nucleasa. Indicaron que "el motivo de nexo, que se ha mostrado recientemente que es crítico para la actividad" (Wright, A. V., y col., página 2986, col. 1), ocupa una posición central entre los lóbulos y forma interacciones extensivas con la hélice de enlace. Basándose en este perfil de interacción, Wright, A. V. y col., generaron un ARNgi de longitud completa y dos construcciones de ARNgi más cortas que se truncaron selectivamente desde el extremo 5' o 3' (no se realizaron modificaciones en la horquilla de nexo crítica) y determinaron sus afinidades para Cas9 de tipo silvestre, los lóbulos α -helicoidal y de nucleasa individuales, y Cas9 dividida.

Al contrario que las enseñanzas anteriormente descritas de la técnica anterior, experimentos realizados en apoyo de la presente invención demostraron inesperadamente que las funciones de Cas9 (por ejemplo, unión y corte de ADN bicatenario) están apoyadas por ARN guías que tienen un nexo dividido, así como ARN guías que tienen modificaciones del nexo dividido.

Los resultados presentados en la presente memoria descriptiva abren nuevas vías de diseño y modificación para tecnologías de CRISPR y preparan el terreno para el desarrollo de tecnologías basadas en CRISPR de nueva generación.

Sumario de la invención

La invención se define en las reivindicaciones adjuntas.

Aspectos de la presente invención se refieren a polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética en los que son necesarios al menos dos polinucleótidos para formar un elemento de tallo de nexo.

Un aspecto de la presente invención se refiere a un polinucleótido de nexo dividido asociado con CRISPR-Cas9 de tipo II. La composición comprende tres polinucleótidos. Un primer polinucleótido de nexo dividido asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn1-casPN) comprende, en la dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y una secuencia de nucleótidos de elemento de tallo de nexo I. Un segundo polinucleótido de nexo dividido asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn2-casPN) comprende, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo de nexo II. La secuencia de nucleótidos de elemento de tallo de nexo I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo de nexo II del sn2-casPN son capaces de formar un elemento de tallo de nexo mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II. Un tercer polinucleótido de nexo dividido asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn3-casPN) comprende, en la dirección de 5' a 3', una secuencia de unión a diana de ácido nucleico y una primera secuencia de nucleótidos de elemento de tallo II. La primera secuencia de nucleótidos de elemento de tallo I del sn1-casPN y la primera secuencia de nucleótidos de elemento de tallo II del sn3-casPN son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II.

En algunas realizaciones, la primera secuencia de nucleótidos de elemento de tallo I del sn1-casPN comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I y una secuencia de nucleótidos de elemento de tallo inferior I. La primera

secuencia de nucleótidos de elemento de tallo II del sn3-casPN comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II. En esta realización, la secuencia de nucleótidos de elemento de tallo superior I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo superior II del sn3-casPN son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo inferior II del sn3-casPN son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

En algunas realizaciones, el sn2-casPN comprende uno o más polinucleótidos adjuntos. Por ejemplo, el sn2-casPN comprende, en la dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexa II y una segunda secuencia de nucleótidos de elemento de tallo I, y un primer polinucleótido adjunto que comprende un extremo 5' y un extremo 3' comprende una segunda secuencia de nucleótidos de elemento de tallo II. La segunda secuencia de nucleótidos de elemento de tallo I del sn2-casPN y la segunda secuencia de nucleótidos de elemento de tallo II del primer polinucleótido adjunto son capaces de formar un segundo elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II. En realizaciones adicionales, el primer polinucleótido adjunto comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la segunda secuencia de nucleótidos de elemento de tallo II, en la que el extremo 5' del primer polinucleótido adjunto se une covalentemente con el extremo 3' del sn2-casPN. En realizaciones adicionales, el primer polinucleótido adjunto comprende, en la dirección de 5' a 3', la secuencia de nucleótidos de elemento de bucle, la segunda secuencia de nucleótidos de elemento de tallo II y una tercera secuencia de nucleótidos de elemento de tallo I y un segundo polinucleótido adjunto que comprende un extremo 5' y un extremo 3' comprende una tercera secuencia de nucleótidos de elemento de tallo II. La tercera secuencia de nucleótidos de elemento de tallo I del primer polinucleótido adjunto y la tercera secuencia de nucleótidos de elemento de tallo II del segundo polinucleótido adjunto son capaces de formar un tercer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II. En realizaciones adicionales, el segundo polinucleótido adjunto comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la tercera secuencia de nucleótidos de elemento de tallo II, en la que el extremo 5' del segundo polinucleótido adjunto se une covalentemente con el extremo 3' del primer polinucleótido adjunto.

Asimismo, las realizaciones de la invención incluyen uno o más polinucleótidos adyuvantes con el sn1-casPN. Por ejemplo, el sn1-casPN comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I. En algunas realizaciones, el sn2-casPN comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II. En realizaciones adicionales, el sn1-casPN comprende un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I en la que el primer polinucleótido adyuvante comprende además una secuencia de nucleótidos de elemento de unión efector I. El sn2-casPN comprende un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II y el segundo polinucleótido adyuvante comprende además una secuencia de nucleótidos de elemento de unión efector II. La secuencia de nucleótidos de elemento de unión efector I del primer polinucleótido adyuvante y la secuencia de nucleótidos de elemento de unión efector II del segundo polinucleótido adyuvante son capaces de formar un elemento de unión efector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector II. Un ejemplo de dicho elemento de unión efector es un ARN bicatenario que comprende un sitio de unión a proteína Csy4. Las realizaciones relacionadas incluyen el primer polinucleótido adyuvante que comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento de unión efector, y el segundo polinucleótido adyuvante que comprende, en la dirección de 5' a 3', la secuencia de nucleótidos de elemento de unión efector II y una secuencia de nucleótidos de elemento conector II. La secuencia de nucleótidos de elemento conector I del primer polinucleótido adyuvante y la secuencia de nucleótidos de elemento conector II del segundo polinucleótido adyuvante son capaces de formar un elemento conector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento conector II. Las realizaciones adicionales incluyen, el primer polinucleótido adyuvante que comprende una horquilla (es decir, una estructura de bucle de tallo) y/o el segundo polinucleótido adyuvante que comprende una horquilla. Las realizaciones relacionadas incluyen el primer polinucleótido adyuvante que comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento conector I y la horquilla, y el segundo polinucleótido adyuvante que comprende además, en la dirección de 5' a 3', la horquilla y una secuencia de nucleótidos de elemento conector II. La secuencia de nucleótidos de elemento conector I del primer polinucleótido adyuvante y la secuencia de nucleótidos de elemento conector II del segundo polinucleótido adyuvante son capaces de formar un elemento conector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento conector II.

Otro aspecto de la invención incluye sistemas de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 que comprenden composiciones de polinucleótidos asociados con CRISPR-Cas9 de tipo II

como se describe en el presente documento y una proteína Cas9 o un polinucleótido (por ejemplo, secuencia de ADN) que codifica una proteína Cas9.

La presente invención también incluye composiciones farmacéuticas que comprenden composiciones de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 como se describe en el presente documento y un excipiente farmacéuticamente aceptable. En algunas realizaciones, las composiciones farmacéuticas comprenden además una proteína Cas9 o un polinucleótido (por ejemplo, secuencia de ADN) que codifica una proteína Cas9.

También se incluyen en la presente invención complejos nucleoproteicos de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 que comprenden composiciones de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II como se describe en el presente documento y una proteína Cas9, en los que el complejo nucleoproteico de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 tiene capacidad de unión dirigida a sitio con un ácido nucleico diana complementario de la unión a diana de ácido nucleico. Asimismo, la presente invención incluye kits que comprenden dichos complejos nucleoproteicos de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 y, por ejemplo, un tampón y/o instrucciones para su uso. Los procedimientos de la invención incluyen procedimientos de corte de un ácido nucleico diana y procedimientos de unión de un ácido nucleico diana, pero no incluyen procedimientos del tratamiento del cuerpo humano o animal por cirugía o terapia. En una realización, la invención se refiere a un procedimiento de corte de un ácido nucleico diana que comprende poner en contacto un ácido nucleico que comprende el ácido nucleico diana con los complejos nucleoproteicos de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 de la presente invención, facilitando de este modo la unión del complejo nucleoproteico de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 con el ácido nucleico diana y el corte del ácido nucleico diana por el complejo nucleoproteico de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9. En otra realización, la invención se refiere a un procedimiento de unión de un ácido nucleico diana que comprende poner en contacto un ácido nucleico que comprende el ácido nucleico diana con los complejos nucleoproteicos de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 de la presente invención facilitando de este modo la unión de los complejos nucleoproteicos de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 con el ácido nucleico diana.

La presente invención también incluye kits que comprenden polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II y, por ejemplo, un tampón y/o instrucciones para su uso. En algunas realizaciones, los kits comprenden además una proteína Cas9 o un polinucleótido (por ejemplo, secuencia de ADN) que codifica una proteína Cas9.

A la vista de la divulgación del presente documento, los expertos en la materia tendrán fácilmente acceso a estos aspectos y otras realizaciones de la presente invención que usan los polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de la presente invención.

Breve descripción de las figuras

La FIG. 1A y la FIG. 1B presentan ejemplos ilustrativos de ARN asociadas con CRISPR-Cas9 de tipo II guías dobles. La FIG. 1A muestra un sistema de CRISPR-Cas9 de tipo II de dos componentes de ARN que comprende un ARNcr (FIG. 1A, 101) y un ARNcrtra (FIG. 1A, 102). La FIG. 1B ilustra la formación de enlaces de hidrógeno de pares de bases entre el ARNcr y el ARNcrtra para formar una estructura secundaria (véase Solicitud de Patente Publicada de los Estados Unidos N.º 2014-0068797, publicada el 6 de marzo de 2014; véase también Jinek, M., y col., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science*, 337, 816-21 (2012)). La figura presenta una visión de conjunto de y nomenclatura para elementos estructurales secundarios del ARNcr y ARNcrtra de la Cas9 de *Streptococcus pyogenes* (SpyCas9) que incluyen los siguientes: un elemento espaciador (FIG. 1B, 103); un primer elemento de tallo que comprende un elemento de tallo inferior (FIG. 1B, 104), un elemento de protuberancia que comprende nucleótidos desapareados (FIG. 1B, 105) y un elemento de tallo superior (FIG. 1B, 106); un elemento de nexo (FIG. 1B, 107); un segundo elemento de horquilla que comprende un segundo elemento de tallo (FIG. 1B, 108); y un tercer elemento de horquilla que comprende un tercer elemento de tallo (FIG. 1B, 109). Las figuras no se presentan proporcionalmente ni están a escala. Las localizaciones de indicadores son aproximadas.

La FIG. 2 muestra otro ejemplo de un ARN asociado con CRISPR-Cas9. La figura ilustra un ARN guía individual (ARNGi) en el que el ARNcr se une covalentemente con el ARNcrtra y forma una estructura secundaria de polinucleótido de ARN mediante enlaces de hidrógeno de pares de bases (véase, por ejemplo, la solicitud publicada de patente de los Estados Unidos n.º 2014-0068797, publicada el 6 de marzo de 2014). La figura presenta una visión de conjunto de y nomenclatura para elementos estructurales secundarios de un ARNGi de la Cas9 de *Streptococcus pyogenes* que incluyen los siguientes: un elemento espaciador (FIG. 2, 201); un primer elemento de tallo que comprende un elemento de tallo inferior (FIG. 2, 202), un elemento de protuberancia que comprende nucleótidos desapareados (FIG. 2, 205) y un elemento de tallo superior (FIG. 2, 203); un elemento de bucle (FIG. 2, 204) que comprende nucleótidos desapareados; (un primer elemento de horquilla comprende el primer elemento de tallo y el elemento de bucle); un elemento de nexo (FIG. 2, 206); un segundo elemento de horquilla que comprende un segundo elemento de tallo (FIG. 2, 207); y un tercer elemento de horquilla que comprende un tercer elemento de tallo (FIG. 2, 208). (Véase, por ejemplo, Figuras 1 y 3 de Briner, A. E., y col.,

"Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality," Molecular Cell Volumen 56(2), 333-339 (2014)). La figura no se presenta proporcionalmente ni está a escala. Las localizaciones de indicadores son aproximadas.

5 La FIG. 3A, la FIG. 3B, la FIG. 3C, la FIG. 3D, la FIG. 3E, la FIG. 3F, la FIG. 3G y la FIG. 3H presentan una diversidad de polinucleótidos de los sistemas de polinucleótidos de nexo dividido asociado con CRISPR-Cas9 de tipo II de la presente invención. Los "polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II" (sn-casPN) de la presente invención comprenden dos o más polinucleótidos, en los que la cadena principal de polinucleótido se rompe dentro del elemento de nexo (es decir, un elemento de nexo dividido). Estas figuras presentan estructuras de sn-casPN ejemplares. Otras modificaciones de sn-casPN se describen en la presente memoria descriptiva. Las figuras no se presentan proporcionalmente ni están a escala. Los indicadores para localizaciones correspondientes a elementos son solamente ilustrativos para proporcionar puntos de referencia en los polinucleótidos ejemplares.

10 La Tabla 1 presenta una serie de indicadores usados uniformemente en la FIG. 3A, la FIG. 3B, la FIG. 3C, la FIG. 3D, la FIG. 3E, la FIG. 3F, la FIG. 3G y la FIG. 3H.

15

Tabla 1

Indicadores numéricos usados para ilustrar regiones de secuencia de nucleótidos asociada con ejemplos de sn-casPN	
Indicadores y elementos correspondientes	
Segundo polinucleótido (sn2-casPN)	
304 a 305	corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido II
305 a 306	corresponde a una segunda secuencia de nucleótidos conectiva
306 a 307	corresponde a una segunda secuencia de nucleótidos de elemento de tallo I
307 a 308	corresponde a una secuencia de nucleótidos de elemento de bucle
308 a 309	corresponde a una segunda secuencia de nucleótidos de elemento de tallo II
309 a 310	corresponde a una tercera secuencia de nucleótidos conectiva
310 a 311	corresponde a una tercera secuencia de nucleótidos de elemento de tallo I
311 a 312	corresponde a una secuencia de nucleótidos de elemento de bucle
312 a 313	corresponde a una tercera secuencia de nucleótidos de elemento de tallo II
313 a 314	corresponde a una secuencia de nucleótidos 3'
Primer polinucleótido (sn1-casPN)	
315 a 316	corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido I
316 a 317	corresponde a una primera secuencia de nucleótidos conectiva
317 a 320	corresponde a una primera secuencia de nucleótidos de elemento de tallo I
317 a 318	corresponde a una secuencia de nucleótidos de elemento de tallo inferior I
318 a 319	corresponde a una secuencia de nucleótidos de elemento de protuberancia I
319 a 320	corresponde a una secuencia de nucleótidos de elemento de tallo superior I
320 a 321	corresponde a una secuencia de nucleótidos de elemento de bucle
Polinucleótidos adicionales (sn3-casPN, sn4-casPN)	
321 a 324	corresponde a una primera secuencia de nucleótidos de elemento de tallo II
321 a 322	corresponde a una secuencia de nucleótidos de elemento de tallo superior II
322 a 323	corresponde a una secuencia de nucleótidos de elemento de protuberancia II

(continuación)

Indicadores numéricos usados para ilustrar regiones de secuencia de nucleótidos asociada con ejemplos de sn-casPN
Indicadores y elementos correspondientes
323 a 324 corresponde a una secuencia de nucleótidos de elemento de tallo inferior II
324 a 325 corresponde a una secuencia de unión a diana de ácido nucleico (un elemento espaciador)

La FIG. 3A ilustra un ejemplo de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de tres polinucleótidos. La FIG. 3A, 301 ilustra un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido. La FIG. 3A, 302 ilustra un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexo dividido. La FIG. 3A, 303 ilustra un tercer polinucleótido (sn3-casPN) que comprende un elemento espaciador. Los ejemplos de estructuras secundarias de polinucleótidos que se forman mediante enlaces de hidrógeno de pares de bases entre secuencias indicadas incluyen los siguientes: secuencia de nucleótidos de elemento de tallo de nexo dividido de sn1-casPN I/secuencia de nucleótidos de elemento de tallo de nexo dividido de sn2-casPN II forman un elemento de tallo de nexo dividido; segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN I/segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN II forman un segundo elemento de tallo; tercera secuencia de nucleótidos de elemento de tallo de sn2-casPN I/tercera secuencia de nucleótidos de elemento de tallo de sn2-casPN II forman un tercer elemento de tallo; primera secuencia de nucleótidos de elemento de tallo de sn1-casPN I/primer elemento de tallo, el primer elemento de tallo que comprende secuencia de nucleótidos de elemento de tallo inferior de sn1-casPN I/secuencia de nucleótidos de elemento de tallo inferior de sn3-casPN II forman un elemento de tallo inferior, y secuencia de nucleótidos de elemento de tallo superior de sn1-casPN I/secuencia de nucleótidos de elemento de tallo superior de sn3-casPN II forman un elemento de tallo superior.

La FIG. 3B ilustra un ejemplo de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos. La FIG. 3B, 326 ilustra un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido. La FIG. 3B, 302 ilustra un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexo dividido. Los ejemplos de estructuras secundarias de polinucleótidos que se forman mediante enlaces de hidrógeno de pares de bases entre secuencias indicadas incluyen los siguientes: secuencia de nucleótidos de elemento de tallo de nexo dividido de sn1-casPN I/secuencia de nucleótidos de elemento de tallo de nexo dividido de sn2-casPN II forman un elemento de tallo de nexo dividido; segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN I/segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN II forman un segundo elemento de tallo; tercera secuencia de nucleótidos de elemento de tallo de sn2-casPN I/tercera secuencia de nucleótidos de elemento de tallo de sn2-casPN II forman un tercer elemento de tallo; primera secuencia de nucleótidos de elemento de tallo de sn1-casPN I/primer elemento de tallo, el primer elemento de tallo que comprende secuencia de nucleótidos de elemento de tallo inferior de sn1-casPN I/secuencia de nucleótidos de elemento de tallo inferior de sn1-casPN II que forman un elemento de tallo inferior, y secuencia de nucleótidos de elemento de tallo superior de sn1-casPN I/secuencia de nucleótidos de elemento de tallo superior de sn1-casPN II que forman un elemento de tallo superior.

La FIG. 3C ilustra un ejemplo de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de tres polinucleótidos. La FIG. 3C, 327 ilustra un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido. La FIG. 3C, 302 ilustra un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexo dividido. La FIG. 3C, 328 ilustra un tercer polinucleótido (sn3-casPN) que comprende un elemento espaciador. Los ejemplos de estructuras secundarias de polinucleótidos que se forman mediante enlaces de hidrógeno de pares de bases entre secuencias indicadas incluyen los siguientes: secuencia de nucleótidos de elemento de tallo de nexo dividido de sn1-casPN I/secuencia de nucleótidos de elemento de tallo de nexo dividido de sn2-casPN II forman un elemento de tallo de nexo dividido; segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN I/segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN II forman un segundo elemento de tallo; tercera secuencia de nucleótidos de elemento de tallo de sn2-casPN I/tercera secuencia de nucleótidos de elemento de tallo de sn2-casPN II forman un tercer elemento de tallo; primera secuencia de nucleótidos de elemento de tallo de sn1-casPN I/primer elemento de tallo, el primer elemento de tallo que comprende secuencia de nucleótidos de elemento de tallo inferior de sn1-casPN I/secuencia de nucleótidos de elemento de tallo inferior de sn3-casPN II forman un primer elemento de tallo.

La FIG. 3D ilustra un ejemplo de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos. La FIG. 3D, 329 ilustra un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido. La FIG. 3D, 302 ilustra un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexo dividido. Los ejemplos de estructuras secundarias de polinucleótidos que se forman mediante enlaces de hidrógeno de pares de bases entre secuencias indicadas incluyen los siguientes: secuencia de nucleótidos de elemento de tallo de nexo dividido de sn1-casPN I/secuencia de nucleótidos de elemento de tallo de nexo dividido de sn2-casPN II forman un elemento de tallo de nexo dividido; segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN I/segunda secuencia de nucleótidos de elemento de tallo de sn2-

5 La FIG. 4A y la FIG. 4B presentan modificaciones de sn-PN. La FIG. 4A presenta modificaciones del polinucleótido 1 (sn1-casPN) y el polinucleótido 2 (sn2-casPN) descritos anteriormente en la FIG. 3A a la FIG. 3H. La FIG. 4B presenta ejemplos de modificaciones adicionales del polinucleótido 1 (sn1-casARN; descrito anteriormente en la FIG. 3A a la FIG. 3H) y el polinucleótido 3 (sn3-casARN; descrito anteriormente en la FIG. 3A, la FIG. C, la FIG. E y la FIG. 3G) descritos anteriormente en la FIG. 4A y la FIG. 4B presenta ejemplos de estructuras de sn1-casPN, sn2-casPN y sn3-casPN. Otras modificaciones de sn1-casPN, sn2-casPN y sn3-casPN se describen en la presente memoria descriptiva. Las figuras no se presentan proporcionalmente ni están a escala. Los indicadores para localizaciones correspondientes a elementos son solamente ilustrativos para proporcionar puntos de referencia en los polinucleótidos ejemplares. La Tabla 2 presenta una serie de indicadores usados uniformemente en la FIG. 4A y la FIG. 4B.

Tabla 2

Indicadores numéricos usados para ilustrar regiones de secuencias de nucleótidos asociadas con ejemplos de sn1-casPN, sn2-casPN y sn3-casPN
Indicadores y elementos correspondientes
Segundo polinucleótido (sn2-casPN; segundo polinucleótido adyuvante; primer polinucleótido adjunto; segundo polinucleótido adjunto)
405 a 406 corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido II
406 a 407 corresponde a una segunda secuencia de nucleótidos conectiva
407 a 408 corresponde a una segunda secuencia de nucleótidos de elemento de tallo I
408 a 409 corresponde a una secuencia de nucleótidos de elemento de bucle
409 a 410 corresponde a una segunda secuencia de nucleótidos de elemento de tallo II
410 a 411 corresponde a una tercera secuencia de nucleótidos conectiva
411 a 412 corresponde a una tercera secuencia de nucleótidos de elemento de tallo I
412 a 413 corresponde a una secuencia de nucleótidos de elemento de bucle
413 a 414 corresponde a una tercera secuencia de nucleótidos de elemento de tallo II
414 a 415 corresponde a una secuencia de nucleótidos 3'
405 a 418 corresponde a un segundo polinucleótido adyuvante
405 a 416 corresponde a una secuencia de nucleótidos de elemento conector II
416 a 417 corresponde a una secuencia de nucleótidos de afinidad II
417 a 418 corresponde a una secuencia de nucleótidos de elemento de unión efector II
Primer polinucleótido (sn1-casPN; polinucleótido adyuvante)
419 a 420 corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido I
420 a 421 corresponde a una primera secuencia de nucleótidos conectiva
419 a 424 corresponde a un primer polinucleótido adyuvante
419 a 422 corresponde a una secuencia de nucleótidos de elemento conector I
422 a 423 corresponde a una secuencia de nucleótidos de afinidad I
423 a 424 corresponde a una secuencia de nucleótidos de elemento de unión efector I
421 a 425 corresponde a una secuencia de nucleótidos de elemento de tallo inferior I
425 a 426 corresponde a una secuencia de nucleótidos de elemento de protuberancia I
426 a 427 corresponde a una secuencia de nucleótidos de elemento de tallo superior I

(continuación)

Indicadores numéricos usados para ilustrar regiones de secuencias de nucleótidos asociadas con ejemplos de sn1-casPN, sn2-casPN y sn3-casPN
Indicadores y elementos correspondientes
427 a 428 corresponde a un primer polinucleótido accesorio
Tercer polinucleótido (sn3-casPN; polinucleótido accesorio)
429 a 430 corresponde a un segundo polinucleótido accesorio
430 a 431 corresponde a una secuencia de nucleótidos de elemento de tallo superior II
431 a 432 corresponde a una secuencia de nucleótidos de elemento de protuberancia II
432 a 433 corresponde a una secuencia de nucleótidos de elemento de tallo inferior II
433 a 434 corresponde a una secuencia de unión a diana de ácido nucleico (un elemento espaciador)

La FIG. 4A, 401 ilustra un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido y un primer polinucleótido adyuvante opcional que se localiza 3' del elemento de nexo dividido. La FIG. 4A, 402 ilustra un ejemplo de un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexo dividido, una segunda secuencia conectiva opcional y un segundo polinucleótido adyuvante opcional que se localiza 5' del elemento de nexo dividido. La FIG. 4A, 402, 403 ilustra un sn2-casPN que comprende un primer polinucleótido adjunto. La FIG. 4A, 402, 403, 404 ilustra un sn2-casPN que comprende además un segundo polinucleótido adjunto. En la FIG. 4A, los tres puntos 5' representan una secuencia polinucleotídica adicional.

En algunas realizaciones, un sn2-casPN puede comprender uno o más de los siguientes: un primer polinucleótido adjunto, un segundo polinucleótido adjunto, un segundo polinucleótido adyuvante o combinaciones de los mismos. Un primer polinucleótido adjunto comprende uno o más de los siguientes: una secuencia de nucleótidos de elemento de bucle, una segunda secuencia de nucleótidos de elemento de tallo II, una tercera secuencia de nucleótidos conectiva, una tercera secuencia de nucleótidos de elemento de tallo I o combinaciones de las mismas. Un segundo polinucleótido adjunto comprende uno o más de los siguientes: una secuencia de nucleótidos de elemento de bucle, una tercera secuencia de nucleótidos de elemento de tallo II, una secuencia de nucleótidos 3' o combinaciones de las mismas.

En algunas realizaciones, ni sn1-casPN ni sn2-casPN comprenden un polinucleótido adyuvante. Las combinaciones de sn1-casPN y/o sn2-casPN que comprenden un polinucleótido adyuvante incluyen, pero sin limitación, las siguientes: sn1-casPN-primer polinucleótido adyuvante/sn2-casPN; sn1-casPN/sn2-casPN-segundo polinucleótido adyuvante; o sn1-casPN-primer polinucleótido adyuvante/sn2-casPN-segundo polinucleótido adyuvante. Asimismo, el primer polinucleótido adyuvante comprende uno o más de los siguientes: una secuencia de nucleótidos de elemento conector I, una secuencia de nucleótidos de afinidad I, una secuencia de nucleótidos de elemento de unión efector I o combinaciones de las mismas. Además, el segundo polinucleótido adyuvante comprende uno o más de los siguientes: la secuencia de nucleótidos de elemento conector II, una secuencia de nucleótidos de afinidad II, una secuencia de nucleótidos de elemento de unión efector II o combinaciones de las mismas.

Los ejemplos de estructuras secundarias de polinucleótidos que se pueden formar mediante enlaces de hidrógeno de pares de bases entre secuencias indicadas (cuando las secuencias están presentes) incluyen los siguientes: secuencia de nucleótidos de elemento de tallo de nexo dividido de sn1-casPN I/secuencia de nucleótidos de elemento de tallo de nexo dividido de sn2-casPN II forman un elemento de tallo de nexo dividido; segunda secuencia de nucleótidos de elemento de tallo de sn2-casPN I/secunda secuencia de nucleótidos de elemento de tallo de primer polinucleótido adjunto II forman un segundo elemento de tallo; y tercera secuencia de nucleótidos de elemento de tallo de primer polinucleótido adjunto I/tercera secuencia de nucleótidos de elemento de tallo de segundo polinucleótido adjunto II forman un tercer elemento de tallo.

Asimismo, en algunas realizaciones el primer polinucleótido adyuvante y el segundo polinucleótido adyuvante son capaces de formar una estructura secundaria mediante enlaces de hidrógeno de pares de bases entre secuencias indicadas, por ejemplo, que incluye uno o más de los siguientes: forma de primer polinucleótido adyuvante sn1-casPN/segundo polinucleótido adyuvante sn2-casPN; secuencia de nucleótidos de afinidad de sn1-casPN I/secuencia de nucleótidos de afinidad de sn2-casPN II; secuencia de nucleótidos de elemento de unión efector de sn1-casPN I/secuencia de nucleótidos de elemento de unión efector de sn2-casPN II; y secuencia de nucleótidos de elemento conector de sn1-casPN I/secuencia de nucleótidos de elemento conector de sn2-casPN II.

Sin embargo, en otras realizaciones no son necesarios enlaces de hidrógeno de pares de bases entre una o más de estas secuencias. Además, en algunas realizaciones la estructura secundaria se forma mediante enlaces de hidrógeno de pares de bases en una secuencia indicada, por ejemplo, el primer polinucleótido adyuvante sn1-casPN

puede comprender una horquilla y/o el segundo polinucleótido adyuvante sn2-casPN puede comprender una horquilla.

Las modificaciones adicionales de las variaciones de sn2-casPN descritas anteriormente en la FIG. 4A incluyen un segundo elemento de horquilla que comprende un segundo elemento de tallo y un elemento de bucle, un tercer elemento de horquilla que comprende un tercer elemento de tallo y un elemento de bucle y tanto el segundo elemento de horquilla como el tercer elemento de horquilla. Por ejemplo, conectando el extremo 3' de la segunda secuencia de nucleótidos de elemento de tallo I (FIG. 4A, 408) con el extremo 5' de la segunda secuencia de nucleótidos de elemento de tallo II (FIG. 4A, 409) se forma un segundo elemento de horquilla. De forma similar, conectando el extremo 3' de la tercera secuencia de nucleótidos de elemento de tallo I (FIG. 4A, 412) con el extremo 5' de la tercera secuencia de nucleótidos de elemento de tallo II (FIG. 4A, 413) se forma un tercer elemento de horquilla.

En algunas realizaciones, un sn2-casPN puede comprender uno o más de los siguientes: un primer polinucleótido adjunto, un segundo polinucleótido adjunto, un segundo polinucleótido adyuvante o combinaciones de los mismos. Un primer polinucleótido adjunto comprende uno o más de los siguientes: una secuencia de nucleótidos de elemento de bucle, una segunda secuencia de nucleótidos de elemento de tallo II, una tercera secuencia de nucleótidos conectiva, una tercera secuencia de nucleótidos de elemento de tallo I o combinaciones de las mismas. Un segundo polinucleótido adjunto comprende uno o más de los siguientes: una secuencia de nucleótidos de elemento de bucle, una tercera secuencia de nucleótidos de elemento de tallo II, una secuencia de nucleótidos 3' o combinaciones de las mismas.

La FIG. 4B, 401 ilustra un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido, un primer polinucleótido adyuvante opcional que se localiza 3' del elemento de nexo dividido y un primer polinucleótido accesorio opcional que se localiza 5' de la secuencia de nucleótidos de elemento de tallo superior I. La FIG. 4B, 435 ilustra un ejemplo de un tercer polinucleótido (sn3-casPN) que comprende un segundo polinucleótido accesorio opcional que se localiza 3' de la secuencia de nucleótidos de elemento de tallo superior II.

En algunas realizaciones, ni sn1-casPN ni sn3-casPN comprende un polinucleótido accesorio. Las combinaciones de sn1-casPN y/o sn3-casPN que comprenden un polinucleótido adyuvante incluyen, pero sin limitación, las siguientes: sn1-casPN-primer polinucleótido accesorio/sn3-casPN; sn1-casPN/sn3-casPN-segundo polinucleótido accesorio; o sn1-casPN-primer polinucleótido accesorio/sn3-casPN-segundo polinucleótido accesorio. Asimismo, el primer polinucleótido accesorio puede comprender uno o más de los siguientes: un elemento conector, una secuencia de afinidad (por ejemplo un ligando o resto de unión a ligando), un elemento de unión efector o combinaciones de los mismos. Además, el segundo polinucleótido adyuvante puede comprender uno o más de los siguientes: un elemento conector, una secuencia de afinidad (por ejemplo, un ligando o resto de unión a ligando), un elemento de unión efector o combinaciones de los mismos.

La FIG. 5A, la FIG. 5B y la FIG. 5C se refieren a información estructural para una realización de un complejo de sn1-casARN/sn2-casARN/proteína Cas9, en el que el sn1-casARN, sn2-casARN corresponden a sn1-casPN y sn2-casPN de la FIG. 3B. La FIG. 5A y la FIG. 5B proporcionan una vista ampliada, abierta de SpyCas9. La FIG. 5A presenta un modelo del lóbulo α -helicoidal de SpyCas9 (FIG. 5A, 501) en complejo con sn1-casARN (FIG. 5A, 502). La sección del sn1-casARN correspondiente al elemento espaciador (es decir, una secuencia de unión a diana de ácido nucleico) está indicada por un paréntesis (FIG. 5A, 503). El extremo 5' del sn1-casARN (FIG. 5A, 504) también se indica. El extremo 3' del sn1-casARN es la localización de la rotura en el elemento de nexo, que es el extremo 3' de la primera parte del nexo dividido (FIG. 5A, 505). La FIG. 5B presenta un modelo del lóbulo de nucleasa catalítica (FIG. 5B, 506) de SpyCas9 en complejo con sn2-casARN (FIG. 5B, 507). El extremo 5' del sn2-casARN es la localización de la rotura en el elemento de nexo, que es el extremo 5' de la segunda parte del nexo dividido (FIG. 5A, 508). El extremo 3' del sn2-casARN (FIG. 5AB, 509) también se indica. Las posiciones relativas del dominio RuvC (FIG. 5B, 510; dominio H de RNasa) y el dominio HNH (FIG. 5B, 511; dominio de nucleasa HNH) están indicadas. La FIG. 5C proporciona una vista de un complejo de sn1-casARN/sn2-casARN/proteína Cas9. Las localizaciones relativas de los siguiente elementos están indicadas: el lóbulo α -helicoidal de SpyCas9 (FIG. 5C, 501); el lóbulo de nucleasa catalítico (FIG. 5C, 506) de SpyCas9; el sn1-casARN (FIG. 5C, 502); el sn2-casARN (FIG. 5C, 507); el extremo 3' del sn2-casARN (FIG. 5C, 509); el extremo 5' del sn1-casARN (FIG. 5C, 504) también está indicado; la posición relativa del dominio RuvC (FIG. 5C, 510); y el área de los extremos 5' y 3' del elemento de nexo dividido (FIG. 5C 508, 505).

La FIG. 6A, la FIG. 6B y la FIG. 6C ilustran un ejemplo de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos. Este sistema corresponde a un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido (FIG. 3B, 326) y un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexo dividido (FIG. 3B, 302). Las figuras no se presentan proporcionalmente ni están a escala. Los indicadores para localizaciones correspondientes a elementos son solamente ilustrativos para proporcionar puntos de referencia en los polinucleótidos ejemplares. La Tabla 3 presenta una serie de indicadores usados en la FIG. 6A y la FIG. 6B.

Tabla 3

Indicadores numéricos usados para ilustrar regiones de secuencias de nucleótidos asociadas con ejemplos de sn1-casARN y sn2-casARN
Indicadores y elementos correspondientes
601 a 602 corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido I
602 a 604 corresponde a un primer polinucleótido adyuvante
602 a 603 corresponde a una secuencia de nucleótidos de elemento conector I
603 a 604 corresponde a una secuencia de nucleótidos de elemento de unión efector I
605 a 606 corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido II
606 a 608 corresponde a un segundo polinucleótido adyuvante
606 a 607 corresponde a una secuencia de nucleótidos de elemento conector II
607 a 608 corresponde a una secuencia de nucleótidos de elemento de unión efector II

La FIG. 6A ilustra un sn1-casARN que comprende un primer polinucleótido adyuvante (FIG. 6A, 602 a 604) y un sn2-casARN que comprende un segundo polinucleótido adyuvante (FIG. 6A, 606 a 608). La figura muestra el sn1-casARN y el sn2-casARN antes de la asociación y formación de pares de bases (pb) con enlaces de hidrógeno entre ellos. La FIG. 6B ilustra el sn1-casARN que comprende un primer polinucleótido adyuvante y el sn2-casARN que comprende un segundo polinucleótido adyuvante después de la formación de pares de bases con enlaces de hidrógeno entre ellos. Se forma un elemento conector entre la secuencia de nucleótidos de elemento conector I (FIG. 6B, 602 a 603) y la secuencia de nucleótidos de elemento conector II (FIG. 6B, 606 a 607). La caja de línea discontinua inferior (FIG. 6B, 609) muestra la formación de un elemento de tallo de nexo. La caja de línea discontinua superior (FIG. 6B, 610) muestra la formación de un elemento de unión efector, en este ejemplo un elemento de unión a ARN de Csy4. La FIG. 6C ilustra la asociación del sn2-casARN con el lóbulo de nucleasa catalítica (FIG. 6C, 613) de SpyCas9 y la asociación del sn1-casARN con el lóbulo α -helicoidal (FIG. 6C, 614) de SpyCas9. También se muestra una proteína efectora Csy4* (FIG. 6C, 615), que es una variante de Csy4 sin actividad endorribonucleasa. Asimismo, la primera parte del nexo dividido (FIG. 6C, 616), la segunda parte del nexo dividido (FIG. 6C, 617), el extremo 3' del sn2-casARN (FIG. 6C, 611), el extremo 5' del sn1-casARN (FIG. 6C, 612), el primer polinucleótido adyuvante (FIG. 6C, 602 a 604) y el segundo polinucleótido adyuvante (FIG. 6C, 606 a 608) están indicadas. La flecha gruesa que apunta hacia abajo indica el ensamblaje del sn2-casARN/lóbulo de nucleasa catalítica (FIG. 6C, 613) de SpyCas9, el sn1-casARN/lóbulo α -helicoidal (FIG. 6C, 614) de SpyCas9, y la proteína Csy4* (FIG. 6C, 615) en un complejo (FIG. 6C, 618). En el complejo (FIG. 6C, 618) el lóbulo de sn2-casARN/nucleasa catalítica (FIG. 6C, 613) de SpyCas9 y el sn1-casARN/lóbulo α -helicoidal (FIG. 6C, 614) de SpyCas9 se han ensamblado en un complejo de sn1-casARN/sn2-casARN/Cas9 (FIG. 6C, 619). La proteína Csy4* (FIG. 6C, 615) se ha unido al elemento de unión a ARN de Csy4 (FIG. 6C, 610). El elemento conector (620) también está indicado.

La FIG. 7A y la FIG. 7B ilustran un ejemplo de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos. Este sistema corresponde a un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexo dividido (FIG. 3B, 326) y un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexo dividido (FIG. 3B, 302). Las figuras no se presentan proporcionalmente ni están a escala. Los indicadores para localizaciones correspondientes a elementos son solamente ilustrativos para proporcionar puntos de referencia en los polinucleótidos ejemplares. La Tabla 4 presenta una serie de indicadores usados en la FIG. 7A y la FIG. 7B.

Tabla 4

Indicadores numéricos usados para ilustrar regiones de secuencias de nucleótidos asociadas con ejemplos de sn1-casARN y sn2-casARN
Indicadores y elementos correspondientes
701 a 702 corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido I
702 a 703 corresponde a un primer polinucleótido adyuvante
704 un elemento de horquilla formado por formación de pares de bases con enlace de hidrógeno entre bases en el primer polinucleótido adyuvante

(continuación)

Indicadores numéricos usados para ilustrar regiones de secuencias de nucleótidos asociadas con ejemplos de sn1-casARN y sn2-casARN
Indicadores y elementos correspondientes
705 a 706 corresponde a una secuencia de nucleótidos de elemento de tallo de nexo dividido II
706 a 707 corresponde a un segundo polinucleótido adyuvante
708 un elemento de horquilla formado por formación de pares de bases con enlace de hidrógeno entre bases en el segundo polinucleótido adyuvante

La FIG. 7A ilustra un sn1-casARN que comprende un primer polinucleótido adyuvante (FIG. 7A, 702 a 703) y un sn2-casARN que comprende un segundo polinucleótido adyuvante (FIG. 7A, 706 a 707). La figura muestra el sn1-casARN y el sn2-casARN antes de la asociación y formación de pares de bases con enlaces de hidrógeno entre ellos. La figura muestra un elemento de horquilla formado por formación de pares de bases con enlaces de hidrógeno entre bases en el primer polinucleótido adyuvante (FIG. 7A, 704) y un elemento de horquilla formado por formación de pares de bases con enlaces de hidrógeno en el segundo polinucleótido adyuvante (FIG. 7A, 708). La FIG. 7B ilustra el sn1-casARN que comprende un primer polinucleótido adyuvante y el sn2-casARN que comprende un segundo polinucleótido adyuvante ensamblados en un complejo de sn1-casARN/sn2-casARN/Cas9 activo. En la FIG. 7B la proteína Cas9 (FIG. 7B, 709), el primer polinucleótido adyuvante que comprende un elemento de horquilla (FIG. 7B, 704) y el segundo polinucleótido adyuvante que comprende un elemento de horquilla (FIG. 7B, 708) están indicados.

La FIG. 8 presenta los resultados del ensayo de escisión de Cas9 usando el ADN bicatenario diana de AAVS-1. En la figura, se muestran repeticiones de tres para cada combinación de sn-casARN^{EX}. En la parte superior de cada panel hay una representación gráfica de los sn-casARN^{EX} usados en el ensayo. La FIG. 8, Panel A muestra la actividad bioquímica de sn1-casARN^{EX}, sn2-casARN^{EX}, sn3-casARN^{EX}-AAVS1. La FIG. 8, Panel B muestra la actividad bioquímica de sn1-casARN^{EX} y sn2-casRNA^{EX}. La FIG. 8, Panel C muestra la actividad bioquímica de sn2-casARN^{EX} y sn3-casARN^{EX}-AAVS1, La FIG. 8, Panel D muestra la actividad bioquímica de sn1-casARN^{EX} y sn3-casARN^{EX}-AAVS1. El último carril de la FIG. 8, Panel D contiene patrones de peso molecular. Se muestran porcentajes de escisión en la parte inferior de cada carril. Para carriles indicados como LDD, cualquier actividad de escisión estuvo por debajo del límite de detección.

La FIG. 9 presenta los resultados del ensayo de escisión de Cas9 usando la proteína Csy4* para potenciar la actividad de escisión de los sn-casARN que comprenden una secuencia de unión a ARN de Csy4 adicional. Los ensayos de escisión usaron dos sistemas de dos polinucleótidos de nexo dividido diferentes que fueron variantes del sistema presente en la FIG. 3B. En el primer sistema, el sn1-casARN comprendía además un primer polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento de unión de Csy4 I y el sn2-casARN comprendía un segundo polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento de unión de Csy4 II, en el que el primer polinucleótido adyuvante y el segundo polinucleótido adyuvante se asocian para formar un elemento de unión a ARN de Csy4 (sn1-casARN^{EXCsy}-Csy/sn2-casARN^{EXCsy}-Csy). En el segundo sistema, el sn1-casARN comprendía además un primer polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento conector I y una secuencia de nucleótidos de elemento de unión de Csy4 I y el sn2-casARN comprendía un segundo polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento conector II y una secuencia de nucleótidos de elemento de unión de Csy4 II, en el que el primer polinucleótido adyuvante y el segundo polinucleótido adyuvante se asocian para formar un elemento conector y un elemento de unión a ARN de Csy4 (sn1-casARN^{EXCsy}-InkCsy/sn2-casARN^{EXCsy}-InkCsy). Cada uno de los dos sistemas se usó para escindir cuatro dianas diferentes, en las que los sn1-casARN comprendían cada uno un espaciador complementario de una de las cuatro dianas: AAVS-1, CD-34, CD-151 y JAK-1. En la figura, la actividad de escisión se muestra en la parte inferior de cada carril (excepto para los carriles 1 y 10, que son patrones de peso molecular). Para carriles indicados como LDD, cualquier actividad de escisión estuvo por debajo del límite de detección. Los sistemas usados en cada una de las reacciones de ensayo de escisión de Cas9 fueron como se muestra en la tabla 5.

Tabla 5

Componentes de polinucleótidos de nexo dividido usados en ensayos de escisión de Cas9		
Carril	sn-casARN ^{EXCsy}	¿Proteína Csy4* añadida?
1	No (patrón de peso molecular)	n/d
2	sn1-casARN ^{EXCsy} -Csy-AAVS1/sn2-casARN ^{EXCsy} -Csy	NO

(continuación)

Componentes de polinucleótidos de nexos divididos usados en ensayos de escisión de Cas9		
Carril	sn-casARN ^{EXCsy}	¿Proteína Csy4* añadida?
3	sn1-casARN ^{EXCsy} -lnkCsy-AAVS1/sn2-casARN ^{EXCsy} -lnkCsy	NO
4	sn1-casARN ^{EXCsy} -Csy-AAVS1/sn2-casARN ^{EXCsy} -Csy	SÍ
5	sn1-casARN ^{EXCsy} -lnkCsy-AAVS1/sn2-casARN ^{EXCsy} -lnkCsy	SÍ
6	sn1-casARN ^{EXCsy} -Csy-CD34/sn2-casARN ^{EXCsy} -Csy	NO
7	sn1-casARN ^{EXCsy} -lnkCsy-CD34/sn2-casARN ^{EXCsy} -lnkCsy	NO
8	sn1-casARN ^{EXCsy} -Csy-CD34/sn2-casARN ^{EXCsy} -Csy	SÍ
9	sn1-casARN ^{EXCsy} -lnkCsy-CD34/sn2-casARN ^{EXCsy} -lnkCsy	SÍ
10	No (patrón de peso molecular)	n/d
11	sn1-casARN ^{EXCsy} -Csy-CD151/sn2-casARN ^{EXCsy} -Csy	NO
12	sn1-casARN ^{EXCsy} -lnkCsy-CD151/sn2-casARN ^{EXCsy} -lnkCsy	NO
13	sn1-casARN ^{EXCsy} -Csy-CD151/sn2-casARN ^{EXCsy} -Csy	SÍ
14	sn1-casARN ^{EXCsy} -lnkCsy-CD151/sn2-casARN ^{EXCsy} -lnkCsy	SÍ
15	sn1-casARN ^{EXCsy} -Csy-JAK-1/sn2-casARN ^{EXCsy} -Csy	NO
16	sn1-casARN ^{EXCsy} -lnkCsy-JAK-1/sn2-casARN ^{EXCsy} -lnkCsy	NO
17	sn1-casARN ^{EXCsy} -Csy-JAK-1/sn2-casARN ^{EXCsy} -Csy	SÍ
18	sn1-casARN ^{EXCsy} -lnkCsy-JAK-1/sn2-casARN ^{EXCsy} -lnkCsy	SÍ

La FIG. 10 presenta el resultado del ensayo de escisión de Cas9 usando el sn1-casARN^{EX2} y sn2-casARN^{EX2}. Se muestran porcentajes de escisión en la parte inferior de cada carril excepto el carril 1, que es un patrón de peso molecular. La FIG. 10, carril 2, presenta resultados de escisión para un sistema de sn1-casARN^{EX2}-AAVS1 y sn2-casARN^{EX2}. La FIG. 10, carril 3, presenta resultados de escisión para un sistema de sn1-casARN^{EX2}-CD151 y sn2-casARN^{EX2}. La FIG. 10, carril 4, presenta los resultados para un sistema de sn1-casARN^{EX2}-JAK1 y sn2-casARN^{EX2}. En la parte superior de la figura hay una representación gráfica de los sn-casARN^{EX2} usados en el ensayo.

La FIG. 11 presenta los resultados de ensayos de escisión de Cas9. Los ensayos de escisión usaron dos sistemas diferentes de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos similares al sistema ilustrado en la FIG. 7A. En la figura, la actividad de escisión se muestra en la parte inferior de cada carril (excepto para los carriles 1 y 10, que son patrones de peso molecular). Para carriles indicados como LDD, cualquier actividad de escisión estuvo por debajo del límite de detección. Se ilustran representaciones de los sn-casARN usados en cada ensayo en la parte superior de la figura. Los sistemas usados en cada una de las reacciones de ensayo de escisión de Cas9 fueron como se muestra en la tabla 6.

Tabla 6

Componentes de polinucleótidos de nexos divididos usados en ensayos de escisión de Cas9		
Carril	sn-casARN ^{EX3Csy}	¿Proteína Csy4* añadida?
1	Ninguno (patrón de peso molecular)	n/d
2	sn1-casARN ^{EX3Csy} -Csy-AAVS1	NO
3	sn2-casARN ^{EX3Csy} -Csy	NO
4	sn1-casARN ^{EX3Csy} -Csy-AAVS1/sn2-casARN ^{EX3Csy} -Csy	
5	sn1-casARN ^{EX3Csy} -Csy-AAVS1/sn2-casARN ^{EX3Csy} -Csy	SÍ

(continuación)

Componentes de polinucleótidos de nexo dividido usados en ensayos de escisión de Cas9		
Carril	sn-casARN ^{EX3Csy}	¿Proteína Csy4* añadida?
6	sn1-casARN ^{EX3Csy} -InkCsy-AAVS1	
7	sn2-casARN ^{EX3Csy} -InkCsy	
8	sn1-casARN ^{EX3Csy} -InkCsy-AAVS1/sn2-casARN ^{EX3Csy} -InkCsy	
9	sn1-casARN ^{EX3Csy} -InkCsy-AAVS1/sn2-casARN ^{EX3Csy} -InkCsy	SI
10	Ninguno (patrón de peso molecular)	n/d

5 La FIG. 12 presenta ejemplos de disposiciones de polinucleótidos de nexo dividido potenciales basándose en secuencias de ARNcrtra conocidas de las especies bacterianas enumeradas en la Tabla 7. En la figura, la primera columna es un número de identificación para la especie bacteriana en relación con la columna de "ID" en la Tabla 7, la segunda columna de la figura es la secuencia del ARNcrtra de nexo dividido (cada uno de estos es un ejemplo de sn1-casARN/sn2-casARN), y la tercera columna es la SEQ ID NO del oligonucleótido. Todas las especies bacterianas enumeradas en la Tabla 7 tienen al menos un sistema de CRISPR-Cas9 de tipo II identificado.

Tabla 7

Especies bacterianas y secuencias de ARNcrtra de nexo dividido potenciales	
ID	Género/Especie
1	<i>Streptococcus pyogenes</i>
2	<i>Streptococcus thermophilus</i> CRISPR-I
3	<i>Listeria innocua</i>
4	<i>Neisseria meningitidis</i>
5	<i>Streptococcus gallolyticus</i>
6	<i>Staphylococcus aureus</i>
7	<i>Corynebacterium diphtheriae</i>
8	<i>Parvibaculum lavamentivorans</i>
9	<i>Campylobacter lari</i>
10	<i>Neisseria cinerea</i>
11	<i>Streptococcus pasteurianus</i>

10 La FIG. 13 es una tabla de oligonucleótidos que expone las secuencias de oligonucleótidos usados en los Ejemplos de la presente memoria descriptiva. La primera columna es una letra identificativa para el oligonucleótido, la segunda columna es la secuencia del oligonucleótido y la tercera columna es la SEQ ID NO del oligonucleótido.

Descripción detallada de la invención

15 Debe entenderse que la terminología usada en el presente documento es con el fin de describir solo realizaciones particulares y no se desea que sea limitante. Como se usa en la presente memoria descriptiva y en las reivindicaciones adjuntas, las formas en singular "un", "uno" y "el" incluyen referentes plurales salvo que el contexto indique claramente otra cosa. Por lo tanto, por ejemplo, la referencia a "un cebador" incluye uno o más cebadores, la referencia a "una célula recombinante" incluye una o más células recombinantes, la referencia a "un agente de reticulación" incluye uno o más agentes de reticulación, y similares.

20 Salvo que se definan de otra manera, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que un experto en la materia a la que la invención pertenece entiende habitualmente. Aunque pueden usarse otros procedimientos y materiales similares, o equivalentes, a los descritos en el presente documento en la práctica de la presente invención, en el presente documento se describen materiales y

procedimientos preferidos.

A la vista de las enseñanzas de la presente memoria descriptiva, un experto en la materia puede aplicar técnicas convencionales de inmunología, bioquímica, química, biología molecular, microbiología, biología celular, genómica y polinucleótidos recombinantes, como se enseña, por ejemplo, en los siguientes textos de referencia: Antibodies: A Laboratory Manual, Segunda edición, E. A. Greenfield, 2014, Cold Spring Harbor Laboratory Press, ISBN 978-1-936113-81-1; Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications, 6ª Edición, R. I. Freshney, 2010, Wiley-Blackwell, ISBN 978-0-470-52812-9; Transgenic Animal Technology, Tercera edición: A Laboratory Handbook, 2014, C. A. Pinkert, Elsevier, ISBN 978-0124104907; The Laboratory Mouse, Segunda edición, 2012, H. Hedrich, Academic Press, ISBN 978-0123820082; Manipulating the Mouse Embryo: A Laboratory Manual, 2013, R. Behringer, y col., Cold Spring Harbor Laboratory Press, ISBN 978-1936113019; PCR 2: A Practical Approach, 1995, M. J. McPherson, y col., IRL Press, ISBN 978-0199634248; Methods in Molecular Biology (Serie), J.M. Walker, ISSN 1064-3745, Humana Press; RNA: A Laboratory Manual, 2010, D. C. Rio, y col., Cold Spring Harbor Laboratory Press, ISBN 978-0879698911; Methods in Enzymology (Serie), Academic Press; Molecular Cloning: A Laboratory Manual (Cuarta edición), 2012, M. R. Green, y col., Cold Spring Harbor Laboratory Press, ISBN 978-1605500560; Bioconjugate Techniques, Tercera edición, 2013, G. T. Hermanson, Academic Press, ISBN 978-0123822390; Methods in Plant Biochemistry and Molecular Biology, 1997, W. V. Dashek, CRC Press, ISBN 978-0849394805; Plant Cell Culture Protocols (Methods in Molecular Biology), 2012, V. M. Loyola-Vargas, y col., Humana Press, ISBN 978-1617798177; Plant Transformation Technologies, 2011, C. N. Stewart, y col., Wiley-Blackwell, ISBN 978-0813821955; Recombinant Proteins from Plants (Methods in Biotechnology), 2010, C. Cunningham, y col., Humana Press, ISBN 978-1617370212; Plant Genomics: Methods and Protocols (Methods in Molecular Biology), 2009, D. J. Somers, y col., Humana Press, ISBN 978-1588299970; Plant Biotechnology: Methods in Tissue Culture and Gene Transfer, 2008, R. Keshavachandran, y col., Orient Blackswan, ISBN 978-8173716164.

Como se usa en el presente documento y se describe en detalle posteriormente, el término "sn-casPN" se refiere a polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II de la presente invención que son capaces de asociarse con proteína Cas9 para formar un complejo. La presente invención se refiere a los sn-casPN descritos en el presente documento y, en algunas realizaciones, los sn-casPN forman un complejo con una proteína Cas9. Una característica distintiva de los sn-casPN es que son necesarios al menos dos de los dos o más sn-casPN para formar un elemento de tallo de nexos. "Sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II" se refiere a dos o más sn-casPN.

Como se usa en el presente documento, la expresión "proteína Cas" y "proteína CRISPR-Cas" se refieren a proteínas asociadas con CRISPR incluyendo, pero sin limitación, proteínas Cas9, proteínas de tipo Cas9 codificadas por ortólogos de Cas9, proteínas sintéticas de tipo Cas9, proteínas Cpf1, proteínas codificadas por ortólogos de Cpf1, proteínas sintéticas de tipo Cpf1, proteínas C2c1, proteínas C2c2, proteínas C2c3, y variantes y modificaciones de las mismas. En una realización preferida, una proteína Cas es una proteína asociada con CRISPR de clase 2, por ejemplo una proteína asociada con CRISPR de tipo II de clase 2. Cada proteína CRISPR-Cas de tipo silvestre interactúa con uno o más polinucleótidos afines (más habitualmente ARN) para formar un complejo nucleoproteico (más habitualmente un complejo ribonucleoproteico).

La expresión "proteína Cas9" como se usa en el presente documento se refiere a una proteína Cas9 de tipo silvestre procedente de sistemas de CRISPR-Cas9 de tipo II, modificaciones de proteínas Cas9, variantes de proteínas Cas9, ortólogos de Cas9 y combinaciones de los mismos. El término "dCas9" como se usa en el presente documento se refiere a variantes de proteína Cas9 que son proteínas Cas9 desactivadas por nucleasa, también denominadas "proteína Cas9 catalíticamente inactiva", o "Cas9 enzimáticamente inactiva".

Como se usa en el presente documento un "sistema de sn-casPN/proteína Cas9" o un "sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9/proteína Cas9" se refiere a sistemas que comprenden al menos sn-casPN y proteína Cas9 o formas expresables de sn-casPN y proteína Cas9 o combinaciones de los sn-casPN y proteína Cas9 y formas expresables de los mismos. Los sistemas de sn-casPN/proteína Cas9 pueden comprender componentes de CRISPR Cas adicionales, tales como proteínas Cas adicionales.

Como se usa en el presente documento, el término "afín" se refiere habitualmente a una proteína Cas9 y sus uno o más polinucleótidos de Cas relacionados que son capaces de formar un complejo nucleoproteico con capacidad de unión dirigida a sitio con un ácido nucleico diana complementario de la secuencia de unión a ácido nucleico diana presente en uno de los polinucleótidos de Cas (por ejemplo, componentes ARN afín de una proteína Cas9).

Como se usa en el presente documento, las expresiones "tipo silvestre", "de origen natural" y "no modificado" se usan para indicar la forma, el aspecto, el fenotipo o la cepa típica (o más frecuente) existente en la naturaleza; por ejemplo, la forma típica de células, organismos, características, polinucleótidos, proteínas, complejos macromoleculares, genes, ARN, ADN o genomas como aparecen y pueden aislarse de una fuente en la naturaleza. La forma, el aspecto, el fenotipo o la cepa de tipo silvestre actúan como el precursor original antes de una modificación intencionada. Por lo tanto, las formas mutantes, variantes, modificadas por ingeniería genética, recombinantes y modificadas no son formas de tipo silvestre.

Como se usa en el presente documento, las expresiones "modificado por ingeniería genética", "modificado por

ingeniería genética", "recombinante", "modificado" y "de origen no natural" son intercambiables e indican manipulación humana intencionada.

Como se usa en el presente documento, las expresiones "ácido nucleico", "secuencia de nucleótidos", "oligonucleótido", y "polinucleótido" son intercambiables. Todas se refieren a una forma polimérica de nucleótidos. Los nucleótidos pueden ser desoxirribonucleótidos (ADN) o ribonucleótidos (ARN), o análogos de los mismos, y pueden ser de cualquier longitud. Los polinucleótidos pueden realizar cualquier función y pueden tener cualquier estructura secundaria y estructura tridimensional. Las expresiones abarcan análogos conocidos de nucleótidos naturales y nucleótidos que están modificados en los restos de base, azúcar y/o fosfato. Los análogos de un nucleótido particular tienen la misma especificidad de formación de pares de bases (por ejemplo, un análogo de base A se empareja con T). Un polinucleótido puede comprender un nucleótido modificado o múltiples nucleótidos modificados (por ejemplo, están disponibles muchos nucleótidos modificados de proveedores comerciales como TriLink (San Diego, CA) e Integrated DNA Technologies (Coralville, Iowa)). Los ejemplos de nucleótidos modificados incluyen nucleótidos metilados y análogos de nucleótidos. La estructura de nucleótidos puede modificarse antes o después de ensamblarse un polímero. Después de la polimerización, los polinucleótidos pueden modificarse adicionalmente mediante, por ejemplo, conjugación con un componente de marcaje o componente de unión a diana. Una secuencia de nucleótidos puede incorporar componentes no nucleotídicos. Las expresiones también abarcan ácidos nucleicos que comprenden restos o enlaces de cadena principal modificada, que (i) son sintéticos, de origen natural y de origen no natural, y (ii) tienen propiedades de unión similares a un polinucleótido de referencia (por ejemplo, ADN o ARN). Los ejemplos de dichos análogos incluyen, pero sin limitación, fosforotioatos, fosforamidatos, metilfosfonatos, metilfosfonatos quirales, 2-O-metilribonucleótidos, ácidos nucleicos peptídicos y estructuras de morfolino.

Los ácidos nucleicos peptídicos (PNA) son homólogos sintéticos de ácidos nucleicos en los que la cadena principal de fosfato-azúcar del polinucleótido se reemplaza con un polímero pseudopeptídico flexible. Las nucleobases se unen con el polímero. Los PNA tienen la capacidad de hibridar con alta afinidad y especificidad con secuencias complementarias de ARN y ADN.

En ácidos nucleicos de fosforotioato, el enlace de fosforotioato (PS) sustituye un oxígeno no enlazante por un átomo de azufre en la cadena principal de fosfato del polinucleótido. Esta modificación hace al enlace internucleotídico resistente a la degradación por nucleasa. En algunas realizaciones, se introducen enlace de fosforotioato entre los últimos 3-5 nucleótidos en el extremo 5' o 3' de una secuencia polinucleotídica para inhibir la degradación por exonucleasa. La colocación de enlaces de fosforotioato a lo largo de un oligonucleótido completo ayuda a reducir también la degradación por endonucleasas.

El ácido nucleico de treosa (TNA) es un polímero genético artificial. La estructura de cadena principal de TNA comprende repeticiones de azúcares treosa unidos por enlaces de fosfodiéster. Los polímeros de TNA son resistentes a la degradación por nucleasa. El TNA puede autoensamblarse por enlaces de hidrógeno de pares de bases en estructuras bicatenarias.

Pueden introducirse inversiones de enlaces en polinucleótidos mediante el uso de "fosforamiditas inversas" (véase, por ejemplo, www.ucalgary.ca/dnalab/synthesis/-modifications/linkages). Habitualmente dichos polinucleótidos tienen grupos fosforamidita en la posición 5'-OH y un grupo protector de dimetoxitritilo (DMT) en la posición 3'-OH. Normalmente, el grupo protector de DMT está en el 5'-OH y la fosforamidita está en el 3'-OH. El uso más frecuente de inversión de enlace es para añadir un enlace 3'-3' al extremo de un polinucleótido con una cadena principal de fosforotioato. El enlace 3'-3' estabiliza el polinucleótido frente a la degradación por exonucleasa creando un oligonucleótido que tiene dos extremos 5'-OH y ningún extremo 3'-OH.

Las secuencias polinucleotídicas se presentan en el presente documento en la orientación convencional de 5' a 3'.

Como se usa en el presente documento, el término "complementariedad" se refiere a la capacidad de una secuencia de ácido nucleico de formar enlace o enlaces de hidrógeno con otra secuencia de ácido nucleico (por ejemplo, mediante formación de pares de bases de Watson-Crick convencional). Un porcentaje de complementariedad indica el porcentaje de restos en una molécula de ácido nucleico que pueden formar enlaces de hidrógeno con una segunda secuencia de ácido nucleico. Cuando dos secuencias polinucleotídicas tienen 100 % de complementariedad, las dos secuencias son perfectamente complementarias, es decir, todos los restos contiguos de un enlace de hidrógeno de un primer polinucleótido con el mismo número de restos contiguos en un segundo polinucleótido.

Como se usa en el presente documento, la expresión "identidad de secuencia" se refiere en general al porcentaje de identidad de bases nucleotídicas o aminoácidos en la comparación de un primer polinucleótido o polipéptido con un segundo polinucleótido o polipéptido usando algoritmos que tienen diversos parámetros de ponderación. La identidad de secuencia entre dos polinucleótidos o dos polipéptidos puede determinarse usando alineamiento de secuencias por diversos procedimientos y programas informáticos (por ejemplo, BLAST, CS-BLAST, FASTA, HMMER, L-ALIGN, etc.), disponibles a través de internet en sitios que incluyen GENBANK (www.ncbi.nlm.nih.gov/genbank/) y EMBL-EBI (www.ebi.ac.uk). La identidad de secuencia entre dos polinucleótidos o dos secuencias polipeptídicas se calcula en general usando los parámetros por defecto convencionales de los

diversos procedimientos o programas informáticos. Un grado alto de identidad de secuencia, como se usa en el presente documento, entre dos polinucleótidos o dos polipéptidos es típicamente entre aproximadamente 90 % de identidad y 100 % de identidad, por ejemplo, aproximadamente 90 % de identidad o mayor, preferentemente aproximadamente 95 % de identidad o mayor, más preferentemente aproximadamente 98 % de identidad o mayor.

Un grado moderado de identidad de secuencia, como se usa en el presente documento, entre dos polinucleótidos o dos polipéptidos es típicamente entre aproximadamente 80 % de identidad y aproximadamente 85 % de identidad, por ejemplo, aproximadamente 80 % de identidad o mayor, preferentemente aproximadamente 85 % de identidad. Un grado bajo de identidad de secuencia, como se usa en el presente documento, entre dos polinucleótidos o dos polipéptidos es típicamente entre aproximadamente 50 % de identidad y 75 % de identidad, por ejemplo, aproximadamente 50 % de identidad, preferentemente aproximadamente 60 % de identidad, más preferentemente aproximadamente 75 % de identidad. Por ejemplo, una proteína Cas (por ejemplo, una Cas9 que comprende sustituciones de aminoácidos) puede tener un grado moderado de identidad de secuencia, o preferentemente un grado alto de identidad de secuencia, a lo largo de su longitud con una proteína Cas de referencia (por ejemplo, una Cas9 de tipo silvestre). Como otro ejemplo, un polinucleótido asociado con Cas9 (es decir, un polinucleótido asociado con Cas9 que es capaz de formar un complejo con una proteína Cas9) puede tener un grado moderado de identidad de secuencia, o preferentemente un grado alto de identidad de secuencia, a lo largo de su longitud con un polinucleótido de tipo silvestre de referencia que forma un complejo con la proteína Cas de referencia (por ejemplo, un ARNgi que forma complejo dirigido a sitio con Cas9).

Como se usa en el presente documento "hibridación" o "hibridar" o "hibridando" es el proceso de combinar dos moléculas de ADN o ARN monocatenarias complementarias y permitir que formen una única molécula bicatenaria (ADN/ADN, ADN/ARN, ARN/ARN) mediante formación de pares de bases por hidrógeno. La rigurosidad de hibridación se determina típicamente mediante la temperatura de hibridación y la concentración salina del tampón de hibridación, por ejemplo, alta temperatura y baja sal proporcionan condiciones de hibridación de alta rigurosidad. Son ejemplos de intervalos de concentración salina e intervalos de temperatura para diferentes condiciones de hibridación los siguientes: alta rigurosidad, sal de aproximadamente 0,01 M a aproximadamente 0,05 M, temperatura de hibridación de 5 °C a 10 °C por debajo de la T_m; rigurosidad moderada, sal de aproximadamente 0,16 M a aproximadamente 0,33 M, temperatura de hibridación de 20 °C a 29 °C por debajo de la T_m; baja rigurosidad, sal de aproximadamente 0,33 M a aproximadamente 0,82 M, temperatura de hibridación de 40 °C a 48 °C por debajo de la T_m. La T_m de ácidos nucleicos bicatenarios se calcula mediante procedimientos convencionales bien conocidos en la técnica (Maniatis, T., y col (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press: Nueva York; Casey, J., y col., (1977) *Nucleic Acids Res.*, 4: 1539; Bodkin, D.K., y col., (1985) *J. Virol. Methods*, 10: 45; Wallace, R.B., y col. (1979) *Nucleic Acids Res.* 6: 3545.) También están ampliamente disponibles herramientas de predicción de algoritmos para estimar la T_m. Las condiciones de alta rigurosidad para hibridación se refieren habitualmente a condiciones en las que un ácido nucleico que tiene complementariedad con una secuencia diana hibrida predominantemente con la secuencia diana y sustancialmente no hibrida con secuencias no diana. Habitualmente las condiciones de hibridación son de rigurosidad moderada, preferentemente alta rigurosidad.

Como se usa en el presente documento una "estructura de tallo-bucle" o un "elemento de tallo-bucle" se refiere a un polinucleótido que tiene una estructura secundaria que incluye una región de nucleótidos que se sabe o se predice que forman una región bicatenaria (el "elemento de tallo") que está unida en un lado por una región de nucleótidos predominantemente monocatenarios (el "elemento de bucle"). El término elemento de "horquilla" también se usa en el presente documento para referirse a estructuras de tallo-bucle. Dichas estructuras son bien conocidas en la técnica. La formación de pares de bases puede ser exacta. Sin embargo, como se sabe en la técnica, un elemento de tallo no requiere formación de pares de bases exacta. Por lo tanto, el elemento de tallo puede incluir uno o más desapareamientos de bases o bases no emparejadas.

Como se usa en el presente documento, el término "recombinación" se refiere a un proceso de intercambio de información genética entre dos polinucleótidos.

Como se usa en el presente documento, las expresiones "polinucleótido donante", "molde donante" y "oligonucleótido donante" se usan indistintamente y se refieren a un polinucleótido que proporciona una secuencia de ácido nucleico de la que se pretende que al menos una parte se integre en un sitio diana de ácido nucleico seleccionado. Habitualmente, un polinucleótido donante es un polinucleótido monocatenario o un polinucleótido bicatenario. Por ejemplo, un sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de la presente invención puede usarse en combinación con un molde de ADN donante para modificar una secuencia diana de ADN en un ADN genómico en el que el ADN genómico se modifica para comprender al menos una parte del molde de ADN donante en la secuencia diana de ADN. En algunas realizaciones, un vector comprende un polinucleótido donante (por ejemplo, un vector de dirección). En otras realizaciones, un polinucleótido donante es un oligonucleótido.

Como se usa en el presente documento, La expresión "reparación dirigida por homología (RDH)" se refiere a reparación de ADN que tiene lugar en células, por ejemplo, durante la reparación de roturas de doble cadena en ADN. La RDH requiere homología de secuencia de nucleótidos y usa un molde donante (por ejemplo, un molde de ADN donante) u oligonucleótido donante para reparar la secuencia en la que se produjo la rotura de doble cadena (por ejemplo, secuencia diana de ADN). Esto da como resultado la transferencia de información genética de, por ejemplo, el ADN molde donante a la secuencia diana de ADN. La RDH puede dar como resultado alteración de la

secuencia diana de ADN (por ejemplo, inserción, supresión, mutación) si la secuencia de ADN molde donante o secuencia oligonucleotídica difiere de la secuencia diana de ADN y parte de o todo el polinucleótido u oligonucleótido de ADN molde donante se incorpora en la secuencia diana de ADN. En algunas realizaciones, un polinucleótido de ADN molde donante completo, una parte del polinucleótido de ADN molde donante o una copia del polinucleótido donante se integra en el sitio de la secuencia diana de ADN.

Como se usa en el presente documento, la expresión "unión de extremos no homólogos (UENH)" se refiere a la reparación de roturas de doble cadena en ADN por ligamiento directo de un extremo de la rotura con el otro extremo de la rotura sin necesidad de un ADN molde donante. UENH en ausencia de un ADN molde donante da como resultado con frecuencia un número pequeño de nucleótidos insertados o suprimidos ("insup" o "insups") aleatoriamente en el sitio de la rotura de doble cadena.

Los términos "vector" y "plásmido" se usan indistintamente y como se usan en el presente documento se refieren a un vehículo polinucleotídico para introducir material genético en una célula. Los vectores pueden ser lineales o circulares. Los vectores pueden integrarse en un genoma diana de una célula huésped o replicarse de forma independiente en una célula huésped. Los vectores pueden comprender, por ejemplo, un origen de replicación, un sitio de clonación múltiple y/o un marcador seleccionable. Un vector de expresión comprende habitualmente un casete de expresión. Los vectores y plásmidos incluyen, pero sin limitación, vectores de integración, plásmidos procariotas, plásmidos eucariotas, cromosomas vegetales sintéticos, episomas, vectores víricos, cósmidos y cromosomas artificiales.

Como se usa en el presente documento la expresión "casete de expresión" es una construcción polinucleotídica, generada de forma recombinante o sintética, que comprende secuencias reguladoras unidas operativamente con un polinucleótido seleccionado para facilitar la expresión del polinucleótido seleccionado en una célula huésped. Por ejemplo, las secuencias reguladoras pueden facilitar la transcripción del polinucleótido seleccionado en una célula huésped o transcripción y traducción del polinucleótido seleccionado en una célula huésped. Un casete de expresión puede, por ejemplo, integrarse en el genoma de una célula huésped o estar presente en un vector de expresión.

Como se usa en el presente documento un "vector de dirección" es una construcción de ADN recombinante que comprende habitualmente ramas de ADN adaptadas homólogas de ADN genómico que flanquea elementos críticos de un gen diana o una secuencia diana. Cuando se introduce en una célula el vector de dirección se integra en el genoma celular mediante recombinación homóloga. Los elementos del gen diana pueden modificarse de varias maneras incluyendo supresiones y/o inserciones. Un gen diana defectuoso puede reemplazarse por un gen diana funcional o, como alternativa, puede inactivarse un gen funcional. Opcionalmente un vector de dirección comprende un casete de selección que comprende un marcador seleccionable que se introduce en el gen diana. Pueden usarse regiones de dirección adyacentes o en ocasiones dentro de un gen diana para afectar a la regulación de la expresión génica.

Como se usa en el presente documento, las expresiones "secuencias reguladoras", "elementos reguladores", y "elementos de control" son intercambiables y se refieren a secuencias polinucleotídicas que están cadena arriba (secuencias no codificantes 5'), dentro o cadena abajo (secuencias no traducidas 3') de una diana polinucleotídica para expresar. Las secuencias reguladoras influyen, por ejemplo, el momento de transcripción, la cantidad o el nivel de transcripción, el procesamiento o la estabilidad de ARN y/o la traducción de la secuencia de nucleótidos estructural relacionada. Las secuencias reguladoras pueden incluir secuencias de unión activadoras, potenciadores, intrones, secuencias de reconocimiento de poliadenilación, promotores, secuencias de unión a represores, estructuras de tallo-bucle, secuencias de inicio de la traducción, secuencias líderes de la traducción, secuencias de terminación de la transcripción, secuencias de terminación de la traducción, sitios de unión a cebadores y similares.

Como se usa en el presente documento la expresión "unido operativamente" se refiere a secuencias polinucleotídicas o secuencias de aminoácidos situadas en una relación funcional entre sí. Por ejemplo, un promotor o potenciador se une operativamente con una secuencia codificante si regula, o contribuye a la modulación de, la transcripción de la secuencia codificante. Las secuencias de ADN unidas operativamente que codifican secuencias reguladoras son típicamente contiguas con la secuencia codificante. Sin embargo, los potenciadores pueden actuar cuando están separados de un promotor por hasta varias kilobases o más. En consecuencia, algunos elementos polinucleotídicos pueden estar unidos operativamente pero no ser contiguos.

Como se usa en el presente documento, el término "expresión" se refiere a la transcripción de un polinucleótido a partir de un molde de ADN, que da como resultado, por ejemplo, un ARNm u otro transcrito de ARN (por ejemplo, ARN no codificantes, tales como estructurales o de almacén). El término se refiere además al proceso mediante el que el ARNm transcrito se traduce a péptidos, polipéptidos o proteínas. Los transcritos y polipéptidos codificados pueden denominarse colectivamente "producto génico". La expresión puede incluir corte y empalme del ARNm en una célula eucariota, si el polinucleótido procede de ADN genómico.

Como se usa en el presente documento, el término "gen" comprende una región de ADN que codifica un producto génico (por ejemplo, un ARN o una proteína), así como todas las regiones de ADN que regulan la producción del producto génico, tanto si dichas secuencias reguladoras están adyacentes a la región de ADN que codifica el producto génico como si no. Por ejemplo, además de la región de ADN que codifica el producto génico, un gen

puede incluir secuencias promotoras, secuencias de terminación, secuencias reguladoras de la traducción (por ejemplo, sitios de unión a ribosomas y sitios de entrada de ribosomas internos), potenciadores, silenciadores, aislantes, elementos limítrofes, orígenes de replicación, sitios de unión a matriz, regiones de control de locus y combinaciones de los mismos.

- 5 Como se usa en el presente documento el término "modular" se refiere a un cambio en la magnitud, el grado o la cantidad de una función. Por ejemplo, los sistemas de sn-casPN/proteína Cas9 desvelados en el presente documento pueden modular la actividad de una secuencia promotora mediante unión en o cerca del promotor. Dependiendo de la acción que se produzca después de la unión, los sistemas de sn-casPN/proteína Cas9 pueden inducir, potenciar, suprimir o inhibir la transcripción de un gen unido operativamente con la secuencia promotora. Por lo tanto, la "modulación" de la expresión génica incluye tanto activación génica como represión génica.

- 10 La modulación puede ensayarse determinando cualquier característica directa o indirectamente afectada por la expresión del gen diana. Dichas características incluyen, por ejemplo, cambios en los niveles de ARN o proteínas, actividad proteica, niveles de producto, expresión génica asociada o nivel de actividad de genes indicadores. En consecuencia, las expresiones "modulación de la expresión", "inhibición de la expresión" y "activación de la expresión" de un gen pueden referirse a la capacidad de un sistema de sn-casPN/proteína Cas9 para cambiar, activar o inhibir la transcripción de un gen.

Como se usa en el presente documento, el término "aminoácido" se refiere a aminoácidos naturales y sintéticos (no naturales), incluyendo análogos de aminoácidos, aminoácidos modificados, peptidomiméticos, glicina e isómeros ópticos D o L.

- 20 Como se usa en el presente documento, los términos "péptido", "polipéptido", y "proteína" se usan indistintamente y se refieren a polímeros de aminoácidos. Un polipéptido puede ser de cualquier longitud. Puede ser ramificado o lineal, puede estar interrumpido por moléculas distintas de aminoácidos y puede comprender aminoácidos modificados. Los términos pueden usarse para referirse a un polímero de aminoácidos que se ha modificado mediante, por ejemplo, acetilación, formación de enlaces disulfuro, glucosilación, lipidación, fosforilación, reticulación y/o conjugación (por ejemplo, con un componente de marcaje o ligando). Las secuencias polinucleotídicas se presentan en el presente documento en la orientación convencional de N-terminal a C-terminal.

Pueden prepararse polipéptidos y polinucleótidos usando técnicas rutinarias en el campo de la biología molecular (véase, por ejemplo, textos de referencia analizados anteriormente). Asimismo, esencialmente cualquier polipéptido o polinucleótido puede pedirse de forma personalizada de fuentes comerciales.

- 30 Como se usa en el presente documento, "no nativa" se refiere a una secuencia de ácido nucleico o secuencia polipeptídica que no se encuentra en la secuencia de ácido nucleico o secuencia polipeptídica nativa (o de tipo silvestre) correspondiente. No nativo también puede referirse a una secuencia de ácido nucleico o polipeptídica de origen natural que comprende mutaciones, inserciones, supresiones u otras modificaciones. Una secuencia de ácido nucleico o secuencia polipeptídica no nativa puede unirse a una secuencia de ácido nucleico o secuencia polipeptídica de origen natural por ingeniería genética para generar una secuencia de ácido nucleico o secuencia polipeptídica quimérica.

- 40 Como se usa en el presente documento, "fusión" se refiere a una secuencia polipeptídica ("polipéptido de fusión") y/o secuencia de ácido nucleico ("polinucleótido de fusión", "ácidos nucleicos de fusión") que comprende una o más secuencias no nativas. La fusión también puede referirse a la unión de un resto con una secuencia polipeptídica o secuencia de ácido nucleico, en la que el resto no es nativo de la secuencia de ácido nucleico o secuencia polipeptídica correspondiente (es decir, la secuencia de ácido nucleico o secuencia polipeptídica de tipo silvestre correspondiente no comprende el resto). Los ejemplos de secuencias y restos que pueden ser útiles en la generación de polipéptidos de fusión o polinucleótidos de fusión incluyen: una señal de localización subcelular o secuencias codificantes de la misma (por ejemplo, una señal de localización nuclear (SLN) para dirección al núcleo, una señal de localización mitocondrial para dirección a las mitocondrias, una señal de localización de cloroplastos para dirección a un cloroplasto, una señal de retención en el retículo endoplásmico (RE) y similares); una molécula pequeña tal como biotina o un colorante (por ejemplo, colorantes de alexa fluor, colorante Cyanine3, colorante Cyanine5); un marcador detectable, incluyendo un resto que puede proporcionar una señal detectable (por ejemplo, una enzima, un radioisótopo, un miembro de un par de unión específico; un fluoróforo; una proteína fluorescente; un punto cuántico; y similares); un miembro de un par de TERF (donante/aceptor) (por ejemplo, EDANS/fluoresceína, IAEDANS/fluoresceína, fluoresceína/tetrametilrodamina, fluoresceína/Cy5, EDANS/DABCYL, fluoresceína/QSY-7, fluoresceína/IC Red 640, fluoresceína/Cy 5.5 y fluoresceína/IC Red 705); un par de donante/aceptor de fluoróforo/punto cuántico; marcadores fluorescentes (por ejemplo, fluoresceína, rodamina, tetrametilrodamina, eosina, eritrosina, cumarina, metilcumarinas, pireno, verde de malaquita, estilbeno, Lucifer Yellow, Cascade Blue™, Texas Red, IAEDANS, EDANS, BODIPY® FL, LC Red 640, Cy 5, Cy 5.5, LC Red 705 y Oregon green); una enzima (peroxidasa de rábano picante, luciferasa, beta-galactosidasa, y similares); una proteína fluorescente (por ejemplo, una proteína verde fluorescente (GFP), una proteína roja fluorescente, una proteína amarilla fluorescente, cualquiera de una diversidad de proteínas fluorescentes y coloreadas); una nanopartícula (por ejemplo, nanopartículas fluorescentes o luminiscentes y nanopartículas magnéticas); puntos cuánticos (PC) (los PC pueden hacerse hidrosolubles aplicando capas de revestimiento que comprenden una diversidad de materiales diferentes.

Por ejemplo, los PC pueden solubilizarse usando polímeros anfífilos; los PC pueden conjugarse con un polipéptido mediante cualquiera de varios grupos funcionales o agentes de unión diferentes que pueden unirse directa o indirectamente a una capa de revestimiento); y radioisótopos.

5 El término "unión" como se usa en el presente documento se refiere a una interacción no covalente entre macromoléculas (por ejemplo, entre una proteína y un polinucleótido, entre un polinucleótido y un polinucleótido y entre una proteína y una proteína). Dicha interacción no covalente también se denomina "asociación" o "interacción" (por ejemplo, cuando una primera macromolécula interacciona con una segunda macromolécula, la primera macromolécula se une con la segunda macromolécula de una manera no covalente). Algunas partes de una interacción de unión pueden ser específicas de secuencia; sin embargo, no es necesario que todos los componentes de una interacción de unión sean específicos de secuencia, tales como los puntos de contacto de la proteína con restos de fosfato en una cadena principal de ADN. Las interacciones de unión pueden caracterizarse por una constante de disociación (Kd). La "afinidad" se refiere a la fuerza de unión. Una afinidad de unión aumentada se correlaciona con una Kd menor. Un ejemplo de unión no covalente es la formación de enlaces de hidrógeno entre pares de bases.

15 Como se usa en el presente documento, la expresión "proteína efectora" se refiere a cualquier polipéptido con un efecto funcional que se une de forma selectiva o específica con un elemento de unión a proteína efectora en un polinucleótido. Dichos elementos de unión de proteína efectora pueden ser polinucleótidos monocatenarios o bicatenarios. Por ejemplo, una proteína efectora puede comprender actividad enzimática, remodelar moléculas biológicas (por ejemplo, chaperonas de plegamiento) o ser una proteína de almacén. Además de unirse con un elemento de unión a proteína efectora correspondiente, una proteína efectora puede modificar un polinucleótido que comprende un elemento de unión efector correspondiente (por ejemplo, escisión, modificación enzimática, modificación transcripcional). Como alternativa, una proteína efectora puede unirse solamente con su elemento de unión a proteína efectora correspondiente. Las proteínas efectoras con actividad enzimática pueden modificarse para ser enzimáticamente inactivas, sin embargo, mantienen su capacidad para unirse con un elemento de unión a proteína efectora. Por ejemplo, Csy4 se une con un elemento de unión a ARN bicatenario de Csy4. Csy4 normalmente es una endorribonucleasa activa pero Csy4 tiene variantes en las que su actividad endonucleasa se ha eliminado (por ejemplo, Csy4*). Cas 7, Cas5 y Cas6 también son ejemplos de proteínas efectoras. Otros ejemplos de proteínas efectoras incluyen, pero sin limitación, proteínas de unión a ARN monocatenario (por ejemplo, proteína de unión a ARNip p19), proteínas de unión a ADN monocatenario (por ejemplo, DBP de adenovirus, proteína de unión a ADN monocatenario termoestable extrema), proteínas de unión a ARN bicatenario (por ejemplo, DICER), proteínas de unión a ADN bicatenario (por ejemplo, proteínas de dedos de cinc) e híbridos de ARN/ADN bicatenarios (por ejemplo, ribonucleasa H).

35 Como se usa en el presente documento, el término "aislado" puede referirse a un ácido nucleico o polipéptido que, por la mano del hombre, existe alejado de su ambiente nativo y no es por lo tanto un producto de la naturaleza. Aislado significa sustancialmente puro. Un ácido nucleico o polipéptido aislado puede existir en una forma purificada y/o puede existir en un ambiente no nativo tal como, por ejemplo, en una célula recombinante.

Como se usa en el presente documento, "organismo" se refiere a cualquier entidad biológica viva, tal como una bacteria, un protista, un hongo, una planta o un animal, compuesta de una o más células.

40 Como se usa en el presente documento, una "célula huésped" se refiere en general a una célula biológica. Una célula puede ser la unidad básica estructural, funcional y/o biológica de un organismo vivo. Una célula puede originarse de cualquier organismo que tenga una o más células. Los ejemplos de células huéspedes incluyen, pero sin limitación: una célula procariota, célula eucariota, una célula bacteriana, una célula de arquea, una célula de un organismo eucariota unicelular, una célula protozoaria, una célula de una planta (por ejemplo, células de plantas de cultivo (tales como soja, tomates, remolachas azucareras, calabaza, heno, cannabis, tabaco, plátanos, ñames, batatas, mandioca, patatas, trigo, sorgo, soja, arroz, trigo, maíz, *Brassica oleaginosa* (por ejemplo, colza y canola oleaginosas), algodón, caña de azúcar, girasol, mijo y alfalfa), frutas, hortalizas, cereales, semillas, angiospermas, coníferas, gimnospermas, helechos, licopodios, antocerotófitos, hepáticas, musgos), una célula de algas, (por ejemplo, *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Nannochloropsis gaditana*, *Chlorella pyrenoidosa*, *Sargassum patens* C. Agardh, y similares), algas marinas (por ejemplo, kelp), una célula fúngica (por ejemplo, una célula de levadura, una célula de un hongo), una célula animal, una célula de un animal invertebrado (por ejemplo, mosca de la fruta, cnidario, equinodermo, nematodo, etc.), una célula de un animal vertebrado (por ejemplo, pez, anfibio, reptil, ave, mamífero), una célula de un mamífero (por ejemplo, un cerdo, una vaca, una cabra, una oveja, un roedor, una rata, un ratón, un primate no humano, un ser humano, etc.). Asimismo, una célula puede ser una célula madre o célula progenitora.

55 Como se usa en el presente documento, la expresión "organismo transgénico" se refiere a un organismo que comprende un polinucleótido introducido de forma recombinante.

60 Como se usa en el presente documento, las expresiones "célula vegetal transgénica" y "planta transgénica" son intercambiables y se refieren a una célula vegetal o una planta que contiene un polinucleótido introducido de forma recombinante. En la expresión planta transgénica se incluye la descendencia (cualquier generación) de una planta transgénica o una semilla de modo que la descendencia o semilla comprenda una secuencia de ADN que codifique

un polinucleótido introducido de forma recombinante o un fragmento del mismo.

Como se usa en el presente documento, la frase "generar una célula vegetal o una planta transgénica" se refiere a usar procedimientos de ADN recombinante y técnicas para construir un vector para transformación de plantas para transformar la célula vegetal o la planta y para generar la célula vegetal transgénica o la planta transgénica.

5 El término "excipiente" como se usa en el presente documento se refiere habitualmente a cualquier sustancia farmacológicamente inactiva usada en la formulación o administración de composiciones farmacéuticas de la presente invención, por ejemplo, un transportador o vehículo. En el presente documento se describen ejemplos de excipientes útiles en la práctica de la presente invención.

10 La expresión "condiciones fisiológicas" como se usa en el presente documento se refiere a condiciones compatibles con células vivas, por ejemplo, condiciones predominantemente acuosas de una temperatura, pH, salinidad, etc.

15 Las expresiones "composición terapéutica", "composición farmacéutica", "preparación terapéutica" y "preparación farmacéutica" se usan indistintamente en el presente documento y abarcan composiciones de la presente invención adecuadas para aplicación o administración a un sujeto, típicamente un ser humano. En general dichas composiciones son seguras, estériles y preferentemente exentas de contaminantes que sean capaces de inducir respuestas indeseables en el sujeto (es decir, el compuesto o los compuestos que comprenden la composición son farmacéuticamente aceptables). Las composiciones pueden formularse para aplicación o administración a un sujeto que las necesite por varias vías de administración diferentes incluyendo oral (es decir, administradas por la boca o el canal alimentario) o parenteral (por ejemplo, bucal, rectal, transdérmica, transmucosa, subcutánea, intravenosa, intraperitoneal, intradérmica, intratraqueal, intratecal, pulmonar, y similares).

20 El término "sujeto" como se usa en el presente documento se refiere a cualquier miembro del subfilo cordados, incluyendo, sin limitación, seres humanos y otros primates, incluyendo primates no humanos tales como macaco rhesus, chimpancés y otras especies de simios y monos; animales de granja, tales como ganado vacuno, ovejas, cerdos, cabras y caballos; mamíferos domésticos, tales como perros y gatos; animales de laboratorio, incluyendo roedores, tales como ratones, ratas y cobayas; aves, incluyendo aves domésticas, silvestres y de caza, tales como pollos, pavos y otras aves gallináceas, patos, gansos; y similares. El término no denota una edad concreta. Por lo tanto, se pretende abarcar individuos adultos, jóvenes y neonatos.

30 Los sistemas de CRISPR-Cas se han reclasificado recientemente en dos clases que comprenden cinco tipos y dieciséis subtipos (Makarova, K., y col., Nature Reviews Microbiology, 13, 1-15 (2015)). Esta clasificación se basa en la identificación de todos los genes *cas* en un locus de CRISPR-Cas y después la determinación de los genes identificativos en cada locus de CRISPR-Cas, determinando en última instancia que los sistemas de CRISPR-Cas pueden situarse en la Clase 1 o Clase 2 basándose en los genes que codifican el módulo efector, es decir, las proteínas implicadas en la etapa de interferencia. Recientemente se ha identificado un sexto sistema de CRISPR-Cas (Abudayyeh O., y col. "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," Science, pii: aaf5573 [Epub] (2 de junio de 2016)).

35 Los sistemas de Clase 1 tienen un complejo multisubunitario de ARNcr-efector, mientras que los sistemas de Clase 2 tienen una única proteína, tal como Cas9, Cpf1, C2c1, C2c2, C2c3 o un complejo de ARNcr-efector. Los sistemas de Clase 1 comprenden sistemas de Tipo I, Tipo III y Tipo IV. Los sistemas de Clase 2 comprenden sistemas de Tipo II y Tipo V.

40 Todos los sistemas de Tipo I tienen una proteína Cas3 que tiene actividad helicasa y actividad de escisión. Los sistemas de Tipo I se dividen adicionalmente en siete subtipos (de I-A a I-F e I-U). Cada subtipo de tipo I tiene una combinación definida de genes distintivos y características definidas de organización de operones. Por ejemplo, los subtipos I-A e I-B parecen tener los genes *cas* organizados en dos o más operones, mientras que los subtipos I-C a I-F parecen tener los genes *cas* codificados por un único operón. Los sistemas de Tipo I tienen un complejo multiproteico de ARNcr-efector que está implicado en las etapas de procesamiento e interferencia del sistema inmunitario de CRISPR-Cas. Este complejo multiproteico se conoce como complejo asociado con CRISPR para defensa antivírica (Cascade). El subtipo I-A comprende *csa5* que codifica una proteína de subunidad pequeña y un gen *cas8* que se divide en dos, que codifica subunidades degradadas grandes y pequeñas y también tiene un gen *cas3* dividido. Un ejemplo de un organismo con un sistema de CRISPR-Cas de subtipo I-A es *Archaeoglobus fulgidus*.

50 El subtipo I-B tiene una disposición de genes *cas1-cas2-cas3-cas4-cas5-cas6-cas7-cas8* y carece de un gen *csa5*. Un ejemplo de un organismo con subtipo I-B es *Clostridium kluyveri*. El subtipo I-C no tiene un gen *cas6*. Un ejemplo de un organismo con subtipo I-C es *Bacillus halodurans*. El subtipo I-D tiene un CasIod en lugar de un *Cas8*. Un ejemplo de un organismo con subtipo I-D es *Cyanothece spp.* El subtipo I-E no tiene un *cas4*. Un ejemplo de un organismo con subtipo I-E es *Escherichia coli*. El subtipo I-F no tiene un *cas4* y tiene un *cas2* fusionado con un *cas3*.
55 Un ejemplo de un organismo con subtipo I-F es *Yersinia pseudotuberculosis*. Un ejemplo de un organismo con subtipo I-U es *Geobacter sulfurreducens*.

Todos los sistemas de tipo III poseen un gen *cas10*, que codifica una proteína multidominio que contiene un dominio Palm (una variante del motivo de reconocimiento de ARN (MRR)) que es homólogo del dominio de núcleo de

numerosas ácido nucleico polimerasas y ciclasas y que es la mayor subunidad de complejos de ARNcr-efector de tipo III. Todos los loci de tipo III codifican también la proteína de subunidad pequeña, una proteína Cas5 y habitualmente varias proteínas Cas7. El tipo III puede dividirse adicionalmente en cuatro subtipos, de III-A a III-D. El subtipo III-A tiene un gen *csm2* que codifica una subunidad pequeña y también tiene genes *cas1*, *cas2* y *cas6*. Un ejemplo de un organismo con subtipo III-A es *Staphylococcus epidermidis*. El subtipo III-B tiene un gen *cmr5* que codifica una subunidad pequeña y habitualmente también carece de genes *cas1*, *cas2* y *cas6*. Un ejemplo de un organismo con subtipo III-B es *Pyrococcus furiosus*. El subtipo III-C tiene una proteína Cas10 con un dominio de tipo ciclasa y carece de un gen *cas1* y *cas2*. Un ejemplo de un organismo con subtipo III-C es *Methanothermobacter thermautotrophicus*. El subtipo III-D tiene una proteína Cas10 que carece del dominio HD, carece de un gen *cas1* y *cas2* y tiene un gen de tipo *cas5* conocido como *csx10*. Un ejemplo de un organismo con subtipo III-D es *Roseiflexus spp.*

Los sistemas de tipo IV codifican un complejo multisubunitario mínimo de ARNcr-efector que comprende una subunidad grande parcialmente degradada, Csf1, Cas5, Cas7 y, en algunos casos, una subunidad pequeña potencial. Los sistemas de tipo IV carecen de genes *cas1* y *cas2*. Los sistemas de tipo IV no tienen subtipos, pero hay dos variantes definidas. Una variante de tipo IV tiene una helicasa de la familia DinG, mientras que una segunda variante de tipo IV carece de una helicasa de la familia DinG, pero tiene un gen que codifica una proteína alfa-helicoidal pequeña. Un ejemplo de un organismo con un sistema de tipo IV es *Acidithiobacillus ferrooxidans*.

Los sistemas de tipo II tienen genes *cas1*, *cas2* y *cas9*. *cas9* codifica una proteína multidominio que combina las funciones del complejo de ARNcr-efector con escisión de ADN diana. Los sistemas de tipo II también codifica un ARNcrtra. Los sistemas de tipo III se dividen adicionalmente en tres subtipos, subtipos II-A, II-B y II-C. El subtipo II-A contiene un gen adicional, *csn2*. Un ejemplo de un organismo con un sistema de subtipo II-A es *Streptococcus thermophilus*. El subtipo II-B carece de *csn2*, pero tiene *cas4*. Un ejemplo de un organismo con un sistema de subtipo II-B es *Legionella pneumophila*. El subtipo II-C es el sistema de tipo II más frecuente hallado en bacterias y tiene solamente tres proteínas, Cas1, Cas2 y Cas9. Un ejemplo de un organismo con un sistema de subtipo II-C es *Neisseria lactamica*.

Los sistemas de tipo V tienen un gen *cpf1* y genes *cas1* y *cas2*. El gen *cpf1* codifica una proteína, Cpf1, que tiene un dominio de nucleasa de tipo RuvC que es homólogo del dominio respectivo de Cas9, pero carece del dominio de nucleasa de HNH que está presente en proteínas Cas9. Se han identificado sistemas de tipo V en varias bacterias, incluyendo *Parcubacteria bacterium* GWC2011_GWC2_44_17 (PbCpf1), *Lachnospiraceae bacterium* MC2017 (Lb3Cpf1), *Butyrivibrio proteoclasticus* (BpCpf1), *Peregrinibacteria bacterium* GW2011_GWA_33_10 (PeCpf1), *Acidaminococcus spp.* BV3L6 (AsCpf1), *Porphyromonas macacae* (PmCpf1), *Lachnospiraceae bacterium* ND2006 (LbCpf1), *Porphyromonas crevioricanis* (PcCpf1), *Prevotella disiens* (PdCpf1), *Moraxella bovoculi* 237 (MbCpf1), *Smithella spp.* SC_K08D17 (SsCpf1), *Leptospira inadai* (LiCpf1), *Lachnospiraceae bacterium* MA2020 (Lb2Cpf1), *Franciscella novicida* U112 (FnCpf1), *Candidatus methanoplasma termitum* (CMtCpf1) y *Eubacterium eligens* (EeCpf1). Recientemente se ha demostrado que Cpf1 también tiene actividad RNasa y es responsable del procesamiento de pre-ARNcr (Fontara, I., y col., "The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA," Nature 28, 532(7600):517-521 (2016)).

En sistemas de clase 1, las etapas de expresión e interferencia implican complejos multisubunitarios de ARN de CRISPR (ARNcr)-efector. En sistemas de clase 2, las etapas de expresión e interferencia implican una única proteína grande, por ejemplo, Cas9, Cpf1, C2c1, C2c1 o C2c3.

En sistemas de clase 1, pre-ARNcr se une al complejo multisubunitario de ARNcr-efector y se procesa a un ARNcr maduro. En sistemas de tipo I y III esto implica una ARN endonucleasa, por ejemplo, Cas6. En sistemas de clase 2 de tipo II, pre-ARNcr se une a Cas9 y se procesa a un ARNcr maduro en una etapa que implica RNasa III y un ARNcrtra. Sin embargo, en al menos un sistema de CRISPR-Cas de tipo II, el de *Neisseria meningitidis*, se transcriben ARNcr con extremos 5' maduros directamente a partir de promotores internos, y no se produce procesamiento de ARNcr.

En sistemas de clase 1 el ARNcr se asocia con el complejo de ARNcr-efector y consigue interferencia combinando actividad nucleasa con dominios de unión a ARN y formación de pares de bases entre el ARNcr y un ácido nucleico diana.

En sistemas de tipo I, el ARNcr y unión a diana del complejo de ARNcr-efector implica Cas7, Cas5 y Cas8 fusionados con una proteína de subunidad pequeña. La escisión de ácido nucleico diana de sistemas de tipo I implica el dominio de nucleasa de HD, que está fusionado con la helicasa de la superfamilia 2 Cas3 o está codificado por un gen separado, *cas3*.

En sistemas de tipo III, el ARNcr y unión a diana del complejo de ARNcr-efector implica Cas7, Cas5, Cas10 y una proteína de subunidad pequeña. La escisión de ácido nucleico diana de sistemas de tipo III implica la acción combinada de las proteínas Cas7 y Cas10, con un dominio de nucleasa de HD definido fusionado con Cas10, que se cree que escinde ADN monocatenario durante la interferencia.

En sistemas de clase 2 el ARNcr se asocia con una única proteína y consigue interferencia combinando actividad

nucleasa con dominios de unión a ARN y formación de pares de bases entre el ARNcr y un ácido nucleico diana.

En sistemas de tipo II, el ARNcr y la unión a diana implica Cas9, como sucede con la escisión de ácido nucleico diana. En sistemas de tipo II, el dominio de nucleasa de tipo RuvC (pliegue H de RNasa) y el dominio de nucleasa de HNH (tipo McrA) de Cas9 escinden cada uno una de las cadenas del ácido nucleico diana. La actividad de escisión de Cas9 de sistemas de tipo II también requiere la hibridación de ARNcr con ARNcrtra para formar un dúplex que facilite la unión de ARNcr y diana por la Cas9.

En sistemas de tipo V, el ARNcr y la unión a diana implica CpfI, como sucede con la escisión de ácido nucleico diana. En sistemas de tipo V, el dominio de nucleasa de tipo RuvC de CpfI escinde una cadena del ácido nucleico diana y un dominio de nucleasa potencial escinde la otra cadena del ácido nucleico diana en una configuración escalonada, produciendo salientes 5', lo que se diferencia de los extremos romos generados por escisión mediante Cas9. Estos salientes 5' pueden facilitar la inserción de ADN mediante procedimientos de unión de extremos no homólogos.

La actividad de escisión de CpfI de sistemas de tipo V tampoco requiere hibridación de ARNcr a ARNcrtra para formar un dúplex, sino que el ARNcr de sistemas de tipo V usa un único ARNcr que tiene una estructura de tallo-bucle que forma un dúplex interno. CpfI se une con el ARNcr de una manera específica de secuencia y estructura, que reconoce el tallo-bucle y secuencias adyacentes al tallo-bucle, más notablemente, el nucleótido 5' de las secuencias espaciadoras que hibrida con el ácido nucleico diana. Esta estructura de tallo-bucle tiene habitualmente una longitud en el intervalo de 15 a 19 nucleótidos. Las sustituciones que alteran este dúplex de tallo-bucle anulan la actividad de escisión, mientras que otras sustituciones que no alteran el dúplex de tallo-bucle no anulan la actividad de escisión. En sistemas de tipo V, el ARNcr forma una estructura de tallo-bucle en el extremo 5' y la secuencia en el extremo 3' es complementaria de una secuencia en un ácido nucleico diana.

Otras proteínas asociadas con ARNcr de tipo V y unión y escisión de diana incluyen candidato 1 de clase 2 (C2c1) y candidato 3 de clase 2 (C2c3). Las proteínas C2c1 y C2c3 tienen una longitud similar a las proteínas Cas9 y CpfI, que varía de aproximadamente 1.100 aminoácidos a aproximadamente 1.500 aminoácidos. Las proteínas C2c1 y C2c3 también contienen dominios de nucleasa de tipo RuvC y tienen una arquitectura similar a CpfI. Las proteínas C2c1 son similares a proteínas Cas9 porque requieren un ARNcr y un ARNcrtra para unión y escisión de diana, pero tienen una temperatura de escisión óptima de 50 °C. Las proteínas C2c1 se dirigen a un MAP rico en AT, que, de forma similar a CpfI, está en dirección 5' de la secuencia diana (véase, por ejemplo, Shmakov, S., y col. *Molecular Cell* 60(3), 385-397 (2015)).

El candidato 2 de clase 2 (C2c2) no comparte similitud de secuencia con otras proteínas efectoras de CRISPR y se ha identificado recientemente como un sistema de tipo VI (Abudayyeh O., y col., "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," *Science*, pii:aaf5573 [Epub] (2 de junio de 2016)). Las proteínas C2c2 tienen dos dominios HEPN y demuestran actividad de escisión de ARNmc. Las proteínas C2c2 son similares a proteínas CpfI porque requieren un ARNcr para unión y escisión de diana, sin requerir al mismo tiempo ARNcrtra. También como CpfI, el ARNcr para proteínas C2c2 forma una horquilla estable, o estructura de tallo-bucle, que ayuda en asociación con la proteína C2c2.

Con respecto a sistemas de CRISPR Cas de clase 2 de tipo II, se conoce en la técnica un gran número de ortólogos de Cas9 así como sus componentes polinucleotídicos asociados (ARNcrtra y ARNcr) (véase, por ejemplo, "Supplementary Table S2. List of bacterial strains with identified Cas9 orthologs," Fonfara, Ines, y col., "Phylogeny of Cas9 Determines Functional Exchangeability of Dual-RNA and Cas9 among Orthologous Type II CRISPR/Cas Systems," *Nucleic Acids Research* 42(4), 2577-2590 (2014), incluyendo todos los datos complementarios; Chylinski K., y col., "Classification and evolution of type II CRISPR-Cas systems," *Nucleic Acids Research*, 42(10), 6091-6105 (2014), incluyendo todos los datos complementarios).

Además, se conocen en la técnica proteínas sintéticas de tipo Cas9 (véase solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014). Un experto habitual en la materia puede practicar aspectos de la presente invención siguiendo la orientación de la memoria descriptiva para usar proteínas CRISPR Cas de tipo II y polinucleótidos que codifican proteína Cas, incluyendo, pero sin limitación, Cas9, tipo Cas9, proteínas codificadas por ortólogos de Cas9, proteínas sintéticas de tipo Cas9, y variantes y modificaciones de las mismas. Los componentes de ARN afines de estas proteínas Cas pueden manipularse y modificarse para su uso en la práctica de la presente invención por un experto habitual en la materia siguiendo la orientación de la presente memoria descriptiva.

Los sistemas de CRISPR-Cas9 de tipo II pueden subdividirse adicionalmente en II-A (contiene Csn2) y II-B (contiene Cas4) y tipo II-C (ni Csn2 ni Cas4, por ejemplo *N. meningitidis*). Se conoce en la técnica un gran número de ortólogos de Cas9 así como sus componentes de ARNcrtra y ARNcr asociados (véase, por ejemplo, "Supplementary Table S2. List of bacterial strains with identified Cas9 orthologs," Fonfara, Ines, y col., "Phylogeny of Cas9 Determines Functional Exchangeability of Dual-RNA and Cas9 among Orthologous Type II CRISPR/Cas Systems," *Nucleic Acids Research* 42(4), 2577-2590 (2014), incluyendo todos los datos complementarios; Chylinski K., y col., "Classification and evolution of type II CRISPR-Cas systems," *Nucleic Acids Research*, 42(10), 6091-6105 (2014), incluyendo todos los datos complementarios; Kevin M Esvelt, K. M., y col., "Orthogonal Cas9 proteins for RNA-

guided gene regulation and editing," *Nature Methods*, 10, 1116-1121 (2013)).

Además, se conocen en la técnica variantes y modificaciones de proteína Cas9. La solicitud publicada de patente de los Estados Unidos n.º 2014-0273226, publicada el 18 de septiembre de 2014, analiza el gen de Cas9 de *S. pyogenes*, proteína Cas9, variantes de la proteína Cas9 incluyendo secuencias codificantes de Cas9 con codones optimizados específicas de huésped (por ejemplo, 0129-0137, la solicitud publicada de patente de los Estados Unidos n.º 2014-0273226, publicada el 18 de septiembre de 2014) y proteína de fusión Cas9 (por ejemplo, 233-240, la solicitud publicada de patente de los Estados Unidos n.º 2014-0273226, publicada el 18 de septiembre de 2014). la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014, enseña un gran número de polipéptidos de Cas9 de tipo silvestre ejemplares (por ejemplo, SEQ ID NO: 1-256, SEQ ID NO: 795-1346, la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014, incluyendo la secuencia de Cas9 de *S. pyogenes* (SEQ ID NO: 8), la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014). También se analizan modificaciones y variantes de proteínas Cas9 (por ejemplo, 504-608, la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014).

Un experto habitual en la materia puede practicar aspectos de la presente invención siguiendo la orientación de la memoria descriptiva para usar proteínas CRISPR Cas de tipo II y polinucleótidos que codifican proteína Cas, incluyendo, pero sin limitación, Cas9, tipo Cas9, Cas1, Cas2, Cas3, Csn2, Cas4, proteínas codificadas por ortólogos de Cas9, proteínas sintéticas de tipo Cas9, y variantes y modificaciones de las mismas. Los componentes de ARN afines de estas proteínas Cas pueden manipularse y modificarse para su uso en la práctica de la presente invención por un experto habitual en la materia siguiendo la orientación de la presente memoria descriptiva.

Cas9 es una proteína CRISPR Cas ejemplar (Cas9 de *S. pyogenes*, SEQ ID NO: 127; CRISPR-1 Cas9 de *S. thermophilus*, SEQ ID NO: 128). Cas9 es una endonucleasa que puede programarse mediante el ARNcrtra/ARNcr para escindir, de forma específica de sitio, ADN diana usando dos dominios de endonucleasa definidos (dominios de HNH y tipo RuvC/RNasa H) (véase solicitud publicada de patente de los Estados Unidos n.º 2014-0068797, publicada el 6 de marzo de 2014; véase también Jinek M., y col., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science*, 337, 816-21 (2012)). Se ilustran dos componentes de ARN de un sistema de CRISPR-Cas9 de tipo II en la FIG. 1A. Típicamente cada sistema de CRISPR-Cas9 comprende un ARNcrtra y un ARNcr. Cas9 es la proteína distintiva característica de sistemas de CRISPR-Cas9 de tipo II.

El ARNcr tiene una región de complementariedad con una secuencia diana de ADN potencial y una segunda región que forma enlaces de hidrógeno de pares de bases con el ARNcrtra para formar una estructura secundaria, típicamente para formar al menos una estructura de tallo. La región de complementariedad con el ADN diana es el espaciador. El ARNcrtra y un ARNcr interactúan mediante varios enlaces de hidrógeno de pares de bases para formar estructuras secundarias de ARN, por ejemplo, como se ilustra en la FIG. 1B. La formación de complejo entre ARNcrtra/ARNcr y proteína Cas9 da como resultado un cambio conformacional de la proteína Cas9 que facilita la unión con ADN, actividades endonucleasa de la proteína Cas9 y escisión de ADN específica de sitio guiada por ARNcr mediante la endonucleasa. Para que un complejo ribonucleoproteico de proteína Cas9/ARNcrtra/ARNcr se una con o escinda una secuencia diana de ADN, la secuencia diana de ADN está adyacente a un motivo adyacente a protoespaciador (MAP) asociado con el complejo ribonucleoproteico de proteína Cas9/ARNcrtra/ARNcr.

El término ARNgi se refiere típicamente a un ARN guía individual (es decir, una única secuencia polinucleotídica contigua) que comprende esencialmente un ARNcr conectado en su extremo 3' con el extremo 5' de un ARNcrtra a través de una secuencia de "bucle" (véase, por ejemplo, la solicitud publicada de patente de los Estados Unidos n.º 2014-0068797, publicada el 6 de marzo de 2014). ARNgi interactúa con una proteína Cas9 afín esencialmente como se ha descrito para polinucleótidos de ARNcrtra/ARNcr, como se ha analizado anteriormente. De forma similar a ARNcr, el ARNgi tiene un elemento espaciador (FIG. 2, 201), una región de complementariedad con una secuencia diana de ADN, adyacente a una segunda región que forma enlaces de hidrógeno de pares de bases que forman una estructura secundaria, típicamente una estructura de tallo (por ejemplo, en la FIG. 2, 202, 203, 204, 205).

Usando un sistema de ARNgi/proteína Cas9, la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014, y posteriormente publicada Briner, A. E., y col., "Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality," *Molecular Cell* Volume, 56(2), 333-339 (2014), demostró que pueden retirarse elementos prescindibles para generar ARNgi en miniatura funcionales. Estas publicaciones analizan la importancia del "nexo", que se localiza en la parte de ARNgi que corresponde a ARNcrtra (no ARNcr), para conferir actividad de escisión a Cas9. El nexo confiere la capacidad de un ARNgi o un ARNcrtra para unirse con su proteína cas9 afín y conferir una transición conformacional de apoenzima a haloenzima.

El nexo se localiza inmediatamente cadena abajo del (es decir, se localiza en dirección 3' desde el) tallo inferior en sistemas de CRISPR-Cas9 de tipo II. Un ejemplo de la localización relativa del nexo se ilustra en el ARNgi mostrado en la FIG. 2 206. la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014, y Briner, A.E., y col., también desvela secuencias consenso y estructuras secundarias de ARNgi predichos para varias familias de ARNgi/Cas9. Estas referencias muestran que la disposición general de estructuras secundarias en los ARNgi predichos hasta e incluyendo el nexo corresponde a las mostradas en la FIG. 2 del presente documento, es decir, en una dirección de 5' a 3', un espaciador, un primer tallo y el nexo. La FIG. 2

presenta una visión de conjunto de y nomenclatura para elementos de un ARNgi de la Cas9 de *Streptococcus pyogenes*. En relación con la FIG. 2, hay variaciones en el número y la disposición de estructuras de tallo localizadas 3' del nexa en los ARNgi ilustrados en la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, y Briner, A.E., y col.

5 Fonfara, y col., "Phylogeny of Cas9 Determines Functional Exchangeability of Dual-RNA and Cas9 among Orthologous Type II CRISPR/Cas Systems," *Nucleic Acids Research*, 42(4), 2577-2590 (2014), incluyendo todos los datos complementarios, en particular la Figura complementaria S11, presentan las secuencias de ARNcr/ARNcrtra y estructuras secundarias de ocho sistemas de CRISPR-Cas9 de tipo II. Las estructuras secundarias de doble cadena de ARN se predijeron usando RNAcofold del paquete Vienna RNA (Bernhart, S.H., y col., "Partition function and base pairing probabilities of RNA heterodimers," *Algorithms Mol. Biol.*, 1,3 (2006); Hofacker, I.L., y col., "Secondary structure prediction for aligned RNA sequences," *J. Mol. Biol.*, 319, 1059-1066 (2002) y RNAhybrid (bibiserv.techfak.uni-bielefeld.de/rmahybrid/)). Las predicciones de estructura se visualizaron después usando VARNA (Darty, K., y col., "VARNA: Interactive drawing and editing of the RNA secondary structure," *Bioinformatics*, 25, 1974-1975 (2009)). Fonfara, I., y col., muestran que el complejo de ARNcr/ARNcrtra para *Campylobacter jejuni* no tiene la región de protuberancia ilustrada en la FIG. 1B, 105; sin embargo, conserva la disposición general de estructuras secundarias hasta e incluyendo el nexa correspondiente a las mostradas en la FIG. 1B del presente documento, es decir, en una dirección de 5' a 3', un espaciador, un primer tallo y el nexa. Ran, F.A., y col., "In vivo genome editing using *Staphylococcus aureus* Cas9," *Nature*, 9, 520(7546), 186-291 (2015), incluyendo todos los datos ampliados, presentan las secuencias de ARNcr/ARNcrtra y estructuras secundarias de ocho sistemas de CRISPR-Cas9 de tipo II (véase Figura 1 de datos ampliados de Ran, F.A., y col.). Las estructuras de ARNcrtra predichas se basaron en el modelo de plegamiento de ARN de generación restringida (Zuker, M., "Mfold web server for nucleic acid folding and hybridization prediction," *Nucleic Acids Res.*, 31, 3406-3415 (2003)). Las estructuras de ARNcr/ARNcrtra para las ocho especies bacterianas presentadas en la Figura 1 de Ran, F.A., y col., muestran que la disposición general de estructuras secundarias en los ARNcr/ARNcrtra predichos hasta e incluyendo el nexa corresponde a las mostradas en la FIG. 1B del presente documento, es decir, en una dirección de 5' a 3', un espaciador, un primer tallo y el nexa.

Como se ha analizado anteriormente y en los antecedentes de la presente memoria descriptiva, Jinek, M., y col., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science*, 337(6096), 816-21 (2012); Briner, A., y col., "Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality," *Molecular Cell* 56(2), 333-339 (2014); y Wright, A. V., y col., "Rational design of a split-Cas9 enzyme complex," *PNAS* 112(10), 2984-2989 (2015) indicaron todos la importancia de la horquilla de nexa para la actividad del complejo de ARN guía/enzima Cas9.

Sin embargo, al contrario que estas enseñanzas, los experimentos realizados en apoyo de la presente invención demostraron inesperadamente que la estructura de horquilla de nexa puede romperse y modificarse; proporcionando de este modo nuevas vías de diseño y genomodificación para tecnologías de CRISPR como se describe en el presente documento.

En un primer aspecto, la presente invención se refiere a un sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende dos o más polinucleótidos (sn-casPN) capaces de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con una primera secuencia de ADN que comprende una secuencia diana de ADN preferentemente en relación con una segunda secuencia de ADN sin la secuencia de unión a diana de ADN. En algunas realizaciones, el complejo corta la primera secuencia de ADN. En el sistema, son necesarios al menos dos de los dos o más polinucleótidos para formar un elemento de tallo de nexa. Además de la unión con la primera secuencia de ADN, el complejo de sn-casPN/Cas9 puede provocar que la proteína Cas9 se una con y escinda la primera secuencia de ADN. Una realización preferida comprende tres sn-casPN (sn1-casPN, sn2-casPN y sn3-casPN; se muestran dos ejemplos en la FIG. 3A, la FIG. 3C), en los que sn3-casPN comprende un elemento espaciador (es decir, una secuencia de unión a diana de ADN). Otra realización preferida comprende dos sn-casPN (sn1-casPN, sn2-casPN; se muestran dos ejemplos en la FIG. 3B, la FIG. 3D), en los que sn1-casPN comprende un elemento espaciador (es decir, una secuencia de unión a diana de ADN) y una primera parte del elemento de nexa. Se presentan dos variaciones de tres sn-casPN en la FIG. 3F, la FIG. 3H. Se presentan dos variaciones de cuatro sn-casPN en la FIG. 3E, la FIG. 3G).

En una realización del primer aspecto de la presente invención, los dos o más polinucleótidos comprenden un primer polinucleótido (por ejemplo, FIG. 3A, 301; FIG. 3C, 327; FIG. 3E, 301; FIG. 3G, 327) que comprende una primera secuencia de nucleótidos de elemento de tallo de nexa I y un segundo polinucleótido (por ejemplo, FIG. 3A, 302; FIG. 3C, 302; FIG. 3E, 302; FIG. 3G, 302) que comprende una primera secuencia de nucleótidos de elemento de tallo de nexa II, en los que (i) la primera secuencia de nucleótidos de elemento de tallo de nexa II y la primera secuencia de nucleótidos de elemento de tallo de nexa I son capaces de formar el elemento de tallo de nexa por enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo de nexa I y la segunda secuencia de nucleótidos de elemento de tallo de nexa II y (ii) el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

En algunas realizaciones del primer aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3A,

301; FIG. 3C, 327) comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexo I y un tercer polinucleótido (por ejemplo, FIG. 3A, 303; FIG. 3C, 328) comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II, en la que el tercer polinucleótido es un polinucleótido separado que comprende un extremo 5' y un extremo 3'.

En otras realizaciones del primer aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3A, 301) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I, la primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexo I, y el tercer polinucleótido (por ejemplo, FIG. 3A, 303) comprende, en una dirección de 5' a 3', la secuencia de unión a diana de ADN, la primera secuencia de nucleótidos de elemento de tallo II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II forman un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II forman el primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II.

En algunas realizaciones del primer aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3E, 301; FIG. 3G, 327) comprende además, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexo I, un tercer polinucleótido (por ejemplo, FIG. 3E, 330; FIG. 3G, 333) comprende una primera secuencia de nucleótidos de elemento de tallo II y un polinucleótido espaciador (por ejemplo, FIG. 3E, 331; FIG. 3G, 331) comprende una secuencia de unión a diana de ADN, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II forman un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II. En esta realización, el primer polinucleótido, el segundo polinucleótido, el tercer polinucleótido y el polinucleótido espaciador son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

En realizaciones adicionales del primer aspecto de la presente invención, el primer polinucleótido (FIG. 3E, 301) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo de nexo I, un tercer polinucleótido (FIG. 3E, 330) comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, y un polinucleótido espaciador (FIG. 3E, 331) comprende una secuencia de unión a diana de ADN, en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II forman un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II forman un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

Realizaciones adicionales resultarán evidentes para los expertos en la materia a la vista de las enseñanzas de la presente memoria descriptiva.

En un segundo aspecto de la presente invención, un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética comprende dos o más polinucleótidos. Los dos o más polinucleótidos comprenden un elemento de crtra que es capaz de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con secuencias de ADN diana que contienen secuencias de motivo adyacente a protoespaciador (MAP) preferentemente en relación con secuencias de ADN sin secuencias de MAP. En algunas realizaciones, el complejo se une preferentemente con y corta secuencias de ADN que contienen secuencias de MAP. El elemento de crtra comprende un primer polinucleótido (por ejemplo, FIG. 3A, 301; FIG. 3C, 327; FIG. 3E, 301; FIG. 3G, 327) que comprende una secuencia de nucleótidos de elemento de tallo de nexo I y un segundo polinucleótido (por ejemplo, FIG. 3A, 302; FIG. 3C, 302; FIG. 3E, 302; FIG. 3G, 302) que comprende una secuencia de nucleótidos de elemento de tallo de nexo II, en los que la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar un elemento de tallo de nexo por enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II y el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

En algunas realizaciones del segundo aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG.

3A, 301; FIG. 3C, 327) comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexa I y un tercer polinucleótido (por ejemplo, FIG. 3A, 303; FIG. 3C, 328) comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II, en la que el tercer polinucleótido es un polinucleótido separado que comprende un extremo 5' y un extremo 3'.

En otras realizaciones del segundo aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3A, 301) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I, la primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexa I, y el tercer polinucleótido (por ejemplo, FIG. 3A, 303) comprende, en una dirección de 5' a 3', la secuencia de unión a diana de ADN, la primera secuencia de nucleótidos de elemento de tallo II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II forman un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II forman el primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II.

En algunas realizaciones del segundo aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3E, 301; FIG. 3G, 327) comprende además, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexa I, un tercer polinucleótido (por ejemplo, FIG. 3E, 330; FIG. 3G, 333) comprende una primera secuencia de nucleótidos de elemento de tallo II y un polinucleótido espaciador (por ejemplo, FIG. 3E, 331; FIG. 3G, 331) comprende una secuencia de unión a diana de ADN, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II forman un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II. En esta realización, el primer polinucleótido, el segundo polinucleótido, el tercer polinucleótido y el polinucleótido espaciador son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

En realizaciones adicionales del segundo aspecto de la presente invención, el primer polinucleótido (FIG. 3E, 301) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo de nexa I, un tercer polinucleótido (FIG. 3E, 330) comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, y un polinucleótido espaciador (FIG. 3E, 331) comprende una secuencia de unión a diana de ADN, en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II forman un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II forman un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

Realizaciones adicionales resultarán evidentes para los expertos en la materia a la vista de las enseñanzas de la presente memoria descriptiva.

En referencia a la expresión "elemento de crtra", como se usa en el presente documento la expresión se refiere a dos o más sn-casPN capaces de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con secuencias de ADN que contienen secuencias de MAP preferentemente en relación con secuencias de ADN sin secuencias de MAP. Sternberg, S.H. y col., "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9," *Nature*, 507(7490), 62-67 (2014) enseñan procedimientos que usan cortinas de ADN con unión doble para examinar las localizaciones y tiempos de vida correspondientes de todos los acontecimientos de unión para ARNcrtra/ARNcr/Cas con ADN. Siguiendo la orientación de la presente memoria descriptiva, un experto habitual en la materia puede aplicar dichos procedimientos para evaluar la unión preferente (afinidad de unión mayor) de, por ejemplo, complejos de sn-casPN/Cas9 con secuencias de ADN que contienen secuencias de MAP frente a secuencias de ADN sin secuencias de MAP para confirmar la presencia de un elemento de crtra que comprende dos o más de los sn-casPN.

En referencia a los sn-casPN, un "espaciador" o "elemento espaciador" como se usa en el presente documento se refiere a una secuencia de unión a diana que puede hibridar específicamente con una secuencia de ácido nucleico diana complementaria y un "polinucleótido espaciador" se refiere a una secuencia polinucleotídica que comprende un elemento espaciador. El elemento espaciador interacciona con la secuencia de ácido nucleico diana mediante

enlaces de hidrógeno entre pares de bases complementarios (es decir, bases emparejadas). Habitualmente, un elemento espaciador (una secuencia de unión a diana de ADN) se une con una secuencia diana de ADN. El elemento espaciador determina la localización de la unión específica de sitio de proteína Cas9 y escisión endonucleolítica. Los elementos espaciadores varían de aproximadamente 17 a aproximadamente 84 nucleótidos de longitud, dependiendo de la proteína Cas9 con la que se asocian, y tienen una longitud promedio de 36 nucleótidos (Marraffini, L. A., y col., "CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea," *Nature Reviews Genetics*, 11(3), 181-190 (2010)). La variabilidad de la longitud funcional para un elemento espaciador se conoce en la técnica (véase, por ejemplo, la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014). Los polinucleótidos espaciadores en algunas realizaciones tienen secuencias polinucleotídicas además del elemento espaciador y dichas secuencias polinucleotídicas se localizan típicamente en el extremo 5' del elemento espaciador, el extremo 3' del elemento espaciador, interno para el elemento espaciador, o combinaciones de los mismos.

La creación de estructura secundaria entre dos polinucleótidos mediante enlaces de hidrógeno de pares de bases (por ejemplo, elementos de tallo y horquillas) puede determinarse mediante varios procedimientos conocidos por los expertos en la materia (por ejemplo, técnicas experimentales, incluyendo, pero sin limitación, cristalografía de rayos X, espectroscopia por resonancia magnética nuclear (RMN), criomicroscopia electrónica (crio-ME), exploración química/enzimática, desnaturalización térmica (estudios de fusión) y espectrometría de masas; técnicas predictivas, tales como predicción de estructura computacional; los procedimientos preferidos incluyen exploración química/enzimática, desnaturalización térmica (estudios de fusión)). Se conocen en la técnica procedimientos para predecir estructuras secundarias de secuencias de ARN o ADN monocatenarias, por ejemplo, el "servidor web RNAfold" (rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) predice estructuras secundarias de secuencias de ARN o ADN monocatenarias (véase, por ejemplo, Gruber, A.R., y col., "The Vienna RNA Websuite," *Nucleic Acids Res.* 1 de julio:36 (número de servidor web) (2008); Lorenz, R., y col., "ViennaRNA Package 2.0," *Algorithms for Molecular Biology*, 6, 26 (2011)). Un procedimiento preferido para evaluar la estructura secundaria de ARN es usar el procedimiento experimental y computacional combinado de SHAPE (Low J.T., y col., "SHAPE-Directed RNA Secondary Structure Prediction," *Methods*, 52(2), 150-158 (2010)).

En un tercer aspecto de la presente invención, un sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética comprende dos o más polinucleótidos que comprenden un primer polinucleótido y un segundo polinucleótido que tienen cada uno extremos 5' y 3'. El primer polinucleótido (por ejemplo, FIG. 3A, 301; FIG. 3C, 327; FIG. 3E, 301; FIG. 3G, 327) comprende una secuencia de nucleótidos de elemento de tallo de nexos I, comprendiendo la secuencia de nucleótidos de elemento de tallo de nexos I, en una dirección de 5' a 3', Nw-N1-N2-Nx, en el que Nw es una primera secuencia de nucleótidos conectiva en la que w es la longitud de la secuencia de nucleótidos conectiva y w es mayor de o igual a dos, N1 es un nucleótido, N2 es un nucleótido y Nx es un primer polinucleótido adyuvante en el que x es la longitud del primer polinucleótido adyuvante y x es mayor de o igual a cero. En algunas realizaciones, para Nw, w es mayor de o igual a cero, preferentemente w es mayor de o igual a 1, más preferentemente w es mayor de o igual a 2. El segundo polinucleótido (por ejemplo, FIG. 3A, 302; FIG. 3C, 302; FIG. 3E, 302; FIG. 3G, 302) comprende una secuencia de nucleótidos de elemento de tallo de nexos II, comprendiendo la secuencia de nucleótidos de elemento de tallo de nexos II, en una dirección de 5' a 3', Ny-Nc2-Nc1-Nz, en el que Ny es un segundo polinucleótido adyuvante en el que y es la longitud del segundo polinucleótido adyuvante e y es mayor de o igual a cero, Nc2 es un nucleótido que es complementario de N2, Nc1 es un nucleótido que es complementario de N1 y Nz es una segunda secuencia de nucleótidos conectiva en la que z es la longitud de la segunda secuencia de nucleótidos conectiva y z es mayor de o igual a cero. En algunas realizaciones, para Nz, z es mayor de o igual a 1, preferentemente z es mayor de o igual a 2. En este aspecto, la primera secuencia de nucleótidos de elemento de tallo de nexos y la segunda secuencia de nucleótidos de elemento de tallo de nexos son capaces de formar un elemento de tallo de nexos por enlaces de hidrógeno de pares de bases entre al menos N1/Nc1 y N2/Nc2 y el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados.

En algunas realizaciones del tercer aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3A, 301; FIG. 3C, 327) comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexos I y un tercer polinucleótido (por ejemplo, FIG. 3A, 303; FIG. 3C, 328) comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II, en la que el tercer polinucleótido es un polinucleótido separado que comprende un extremo 5' y un extremo 3'.

En otras realizaciones del tercer aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3A, 301) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I, la primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexos I, y el tercer polinucleótido (por ejemplo, FIG. 3A, 303) comprende, en una dirección de 5' a 3', la secuencia de unión a diana de ADN, la primera secuencia de nucleótidos de elemento de tallo II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, en el que la secuencia de nucleótidos de elemento de tallo superior I y la

secuencia de nucleótidos de elemento de tallo superior II forman un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II forman el primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II.

En algunas realizaciones del tercer aspecto de la presente invención, el primer polinucleótido (por ejemplo, FIG. 3E, 301; FIG. 3G, 327) comprende además, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexa I, un tercer polinucleótido (por ejemplo, FIG. 3E, 330; FIG. 3G, 333) comprende una primera secuencia de nucleótidos de elemento de tallo II y un polinucleótido espaciador (por ejemplo, FIG. 3E, 331; FIG. 3G, 331) comprende una secuencia de unión a diana de ADN, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II forman un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II. En esta realización, el primer polinucleótido, el segundo polinucleótido, el tercer polinucleótido y el polinucleótido espaciador son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

En realizaciones adicionales del tercer aspecto de la presente invención, el primer polinucleótido (FIG. 3E, 301) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo de nexa I, un tercer polinucleótido (FIG. 3E, 330) comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, y un polinucleótido espaciador (FIG. 3E, 331) comprende una secuencia de unión a diana de ADN, en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II forman un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II forman un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

Realizaciones adicionales resultarán evidentes para los expertos en la materia a la vista de las enseñanzas de la presente memoria descriptiva.

Un cuarto aspecto de la presente invención comprende una modificación del primer, segundo y tercer aspectos de la presente invención, en los que el extremo 5' del primer polinucleótido y el extremo 3' del tercer polinucleótido se unen por un elemento de bucle. En consecuencia, en el cuarto aspecto de la invención no hay "tercer polinucleótido" porque se ha incluido en "un primer polinucleótido que comprende una primera horquilla". En algunas realizaciones, este primer polinucleótido comprende, en una dirección de 5' a 3', una secuencia de unión a diana de ADN, un primer elemento de tallo y la secuencia de nucleótidos de elemento de tallo de nexa I, en el que el primer elemento de tallo comprende una primera horquilla (por ejemplo, FIG. 3D, 329). En realizaciones adicionales el primer elemento de tallo comprende además un elemento de tallo inferior, un elemento de protuberancia y un elemento de tallo superior, en el que el elemento de tallo inferior está adyacente al elemento de protuberancia, el elemento de protuberancia está adyacente al elemento de tallo superior, el elemento de protuberancia se interpone entre el elemento de tallo inferior y el elemento de tallo superior, y el elemento de tallo superior comprende la primera horquilla (por ejemplo, FIG. 3B, 326). En algunas realizaciones el elemento espaciador se separa del primer polinucleótido que comprende la primera horquilla (por ejemplo, FIG. 3F, 332; FIG. 3H, 334) y un polinucleótido espaciador (por ejemplo, FIG. 3F, 331; FIG. 3H, 331) comprende el elemento espaciador. Las variaciones de sn-casPN descritas posteriormente que usan un primer polinucleótido accesorio o un segundo polinucleótido accesorio no se aplican a los sn-casPN que comprenden un primer polinucleótido que comprende una primera horquilla.

Realizaciones adicionales resultarán evidentes para los expertos en la materia a la vista de las enseñanzas de la presente memoria descriptiva.

Se ilustran componentes de un sistema de sn1-casARN/sn2-casARN/proteína Cas9 en la FIG. 5A, la FIG. 5B y la FIG. 5C. Se presenta un ejemplo de sn1-casARN/sn2-casARN en la FIG. 3B, en el que el primer polinucleótido es sn1-casARN (FIG. 3B, 326) y el segundo polinucleótido es sn2-casARN (FIG. 3B, 302). La FIG. 5A presenta un modelo del lóbulo α -helicoidal de SpyCas9 (FIG. 5A, 501) en complejo con sn1-casARN (FIG. 5A, 502). La sección del sn1-casARN correspondiente al elemento espaciador (es decir, una secuencia de unión a diana de ácido nucleico) está indicada por un paréntesis (FIG. 5A, 503). La FIG. 5B presenta un modelo del lóbulo de nucleasa catalítica (FIG. 5B, 506) de SpyCas9 en complejo con sn2-casARN (FIG. 5B, 507). Las posiciones relativas del dominio RuvC (FIG. 5B, 510; dominio H de RNasa) y el dominio HNH (FIG. 5B, 511; dominio de nucleasa HNH) están indicadas. La FIG. 5C proporciona una vista de un complejo de sn1-casARN/sn2-casARN/proteína Cas9. Se indican las localizaciones relativas del extremo 3' del elemento de nexa dividido de sn1-casARN (FIG. 5C, 505) y extremo 3' del elemento de nexa dividido del sn2-casARN (FIG. 5C 508).

Un quinto aspecto de la presente invención comprende una modificación del primer, segundo y tercer aspectos de la presente invención, en los que la modificación es la adición de un polinucleótido accesorio opcional al primer polinucleótido, al tercer polinucleótido o tanto al primer polinucleótido como al tercer polinucleótido. En algunas realizaciones del quinto aspecto de la invención, el primer polinucleótido comprende además un primer polinucleótido accesorio localizado 5' de la primera secuencia de nucleótidos de elemento de tallo I. Cuando el primer elemento de tallo del primer polinucleótido comprende, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo superior I, la secuencia de elemento de protuberancia I, la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo de nexa I, entonces el primer polinucleótido accesorio se localiza 5' de la secuencia de nucleótidos de elemento de tallo superior I (por ejemplo, FIG. 4B, 401, 427 a 428).

En otras realizaciones del quinto aspecto de la presente invención, el tercer polinucleótido comprende además un segundo polinucleótido accesorio localizado 3' de la primera secuencia de nucleótidos de elemento de tallo II. Cuando el primer elemento de tallo del tercer polinucleótido comprende, en una dirección de 5' a 3', la secuencia de unión a diana de ADN, la primera secuencia de nucleótidos de elemento de tallo II, la secuencia de nucleótidos de elemento de protuberancia II y la secuencia de nucleótidos de elemento de tallo superior II, entonces el segundo polinucleótido accesorio se localiza 3' de la secuencia de nucleótidos de elemento de tallo superior II (por ejemplo, FIG. 4B, 435, 429 a 430).

El polinucleótido accesorio puede comprender una diversidad de restos incluyendo, pero sin limitación, un marcador de afinidad, un ligando, una secuencia de unión a ligando, una secuencia conectora, una horquilla, una secuencia de nucleótidos de afinidad, un elemento de unión efector, proteínas efectoras fusionadas, una señal de localización subcelular o secuencias codificantes de los mismos; una molécula pequeña, un marcador detectable, un miembro de un par de TERF, un par de donante/aceptor de fluoróforo/punto cuántico, marcadores fluorescentes, una enzima, una proteína fluorescente, una nanopartícula, un punto cuántico.

Un sexto aspecto de la presente invención se refiere a modificaciones del segundo polinucleótido de sn-casPN. En una realización, el segundo polinucleótido comprende, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexa II y un segundo elemento de tallo, en el que el segundo elemento de tallo comprende una horquilla. El segundo polinucleótido también puede comprender, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexa II, el segundo elemento de tallo y un tercer elemento de tallo, en el que el tercer elemento de tallo comprende una horquilla. Asimismo, el segundo polinucleótido comprende, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexa II y un segundo elemento de tallo, en el que el segundo elemento de tallo comprende una horquilla, y un tercer elemento de tallo, en el que el tercer elemento de tallo comprende una horquilla (por ejemplo, FIG. 3A, 302).

En otra realización del sexto aspecto de la presente invención, el segundo polinucleótido comprende además, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexa II, una segunda secuencia conectiva y una segunda secuencia de nucleótidos de elemento de tallo I (por ejemplo, FIG. 4A, 402, 406 a 407) y un primer polinucleótido adjunto que comprende una segunda secuencia de nucleótidos de elemento de tallo II (FIG. 4A, 403, 409 a 410), en el que la segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II forman un segundo elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II. En algunas realizaciones, el extremo 5' de la segunda secuencia de nucleótidos de elemento de tallo II y el extremo 3' de la segunda secuencia de nucleótidos de elemento de tallo I están conectados por un elemento de bucle para crear una segunda horquilla.

Asimismo, el primer polinucleótido adjunto puede comprender, en una dirección de 5' a 3', la segunda secuencia de nucleótidos de elemento de tallo I y una tercera secuencia de nucleótidos de elemento de tallo I (FIG. 4A, 403, 411 a 412), y un segundo polinucleótido adjunto (FIG. 4A, 404) comprende, en una dirección de 5' a 3', una tercera secuencia de nucleótidos de elemento de tallo II (FIG. 4A, 404, 413 a 414), en el que la tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II forman un tercer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II. En algunas realizaciones, el extremo 5' de la tercera secuencia de nucleótidos de elemento de tallo II y el extremo 3' de la tercera secuencia de nucleótidos de elemento de tallo I están conectados por un elemento de bucle para crear una tercera horquilla.

En otras realizaciones el extremo 3' del segundo polinucleótido comprende una secuencia terminal 3' que puede comprender una diversidad de restos incluyendo, pero sin limitación, un marcador de afinidad, un ligando, una secuencia de unión a ligando, una secuencia conectora, una horquilla, una secuencia de nucleótidos de afinidad, un elemento de unión efector, proteínas efectoras fusionadas, una señal de localización subcelular o secuencias codificantes de los mismos; una molécula pequeña, un marcador detectable, un miembro de un par de TERF, un par de donante/aceptor de fluoróforo/punto cuántico, marcadores fluorescentes, una enzima, una proteína fluorescente, una nanopartícula, un punto cuántico.

Un séptimo aspecto de la presente invención se refiere a la modificación del extremo 3' del nexa dividido del primer polinucleótido y el extremo 5' del nexa dividido del segundo polinucleótido en el que la modificación es la adición de

un polinucleótido adyuvante opcional al primer polinucleótido, al segundo polinucleótido o tanto al primer polinucleótido como al segundo polinucleótido. En una realización, el primer polinucleótido comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I. En otra realización, el segundo polinucleótido comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II. En otra realización más, el primer polinucleótido comprende un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I y el segundo polinucleótido comprende un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II. En algunas realizaciones un polinucleótido de elemento conector se interpone entre la secuencia de nucleótidos de elemento de nexa y el polinucleótido adyuvante. El primer polinucleótido adyuvante y/o segundo polinucleótido adyuvante pueden comprender un sitio de unión para una proteína de unión a polinucleótido monocatenario, tal como una proteína de unión a ARN monocatenario.

En una realización adicional del séptimo aspecto de la invención, el primer polinucleótido comprende un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I y el segundo polinucleótido comprende un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II, y el primer polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento de unión efector I, y el segundo polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento de unión efector II, en el que la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector II son capaces de formar un elemento de unión efector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector I. Por lo tanto, se proporciona un elemento de unión efector de polinucleótido bicatenario con el que se puede unir una proteína efectora. En algunas realizaciones el polinucleótido bicatenario es un ARN y la proteína efectora es una proteína de unión a ARN bicatenario capaz de unirse con el elemento de unión efector. Los ejemplos de proteínas efectoras de unión a ARN bicatenario incluyen Cas5, Cas6 y Csy4. En algunas realizaciones la proteína de unión efectora está catalíticamente inactiva (por ejemplo, Csy4*) pero aún se une con el elemento de unión efector.

En algunas realizaciones del séptimo aspecto de la invención, el primer polinucleótido adyuvante y/o segundo polinucleótido adyuvante comprenden además una o más primeras secuencias de nucleótidos de afinidad. Una secuencia de nucleótidos de afinidad puede unirse covalentemente con un polipéptido. Una secuencia de nucleótidos de afinidad puede comprender un ligando. En algunas realizaciones, una de las secuencias de nucleótidos de afinidad comprende un ligando y la otra secuencia de nucleótidos de afinidad comprende un resto de unión a ligando correspondiente.

La FIG. 4A ilustra un ejemplo de un primer polinucleótido (FIG. 4A, 401) que comprende un primer polinucleótido adyuvante (FIG. 4A, 401, 419 a 424), en el que el primer polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento conector I (FIG. 4A, 401, 419 a 422), una secuencia de nucleótidos de afinidad I (FIG. 4A, 401, 422 a 423), y una secuencia de nucleótidos de elemento de unión efector I (FIG. 4A, 401, 423 a 424), y un segundo polinucleótido (FIG. 4A, 402) que comprende un segundo polinucleótido adyuvante (FIG. 4A, 402, 405 a 418), en el que el segundo polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento conector II (FIG. 4A, 402, 405 a 416), una secuencia de nucleótidos de afinidad II (FIG. 4A, 402, 416 a 417), y una secuencia de nucleótidos de elemento de unión efector II (FIG. 4A, 402, 417 a 418).

Un ejemplo de uso de una proteína efectora es Csy4* con un elemento de unión de proteína efectora correspondiente puede proporcionarse con referencia a esta figura. La secuencia de nucleótidos de elemento de unión efector I (FIG. 4A, 401, 423 a 424), y secuencia de nucleótidos de elemento de unión efector II (FIG. 4A, 402, 417 a 418) forman una estructura de ARN bicatenario mediante enlaces de hidrógeno de pares de bases para formar un elemento de unión bicatenario de Csy4*. Después de la formación del elemento de unión a ARN bicatenario la proteína Csy4* se une con el elemento de unión y estabiliza la interacción del primer polinucleótido adyuvante y el segundo polinucleótido adyuvante. Csy4* y su elemento de unión correspondiente se usa de esta manera en los experimentos de escisión de Cas9 presentados en el Ejemplo 5 y el Ejemplo 7.

Se presenta un ejemplo relacionado de uso de una proteína efectora es Csy4* con un elemento de unión de proteína efectora correspondiente en la FIG. 6A, la FIG. 6B y la FIG. 6C para un sistema de dos polinucleótidos sn-casPN/proteína Cas9. Este sistema corresponde a un primer polinucleótido (sn1-casPN) que comprende una primera parte del elemento de nexa dividido (FIG. 3B, 326) y un segundo polinucleótido (sn2-casPN) que comprende una segunda parte del elemento de nexa dividido (FIG. 3B, 302).

La capacidad de Csy4* para facilitar la escisión por sn-casARN/proteína Cas9 de cuatro dianas de ADN bicatenario se demuestra en el Ejemplo 5. Los datos presentados en la FIG. 9 demuestran una escisión potenciada por proteína efectora (aquí Csy4*) de ADN bicatenario diana mediante sistemas de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II de la presente invención que comprenden polinucleótidos adyuvantes que tienen un elemento de unión efector (por ejemplo, la secuencia de unión a ARN de Csy4). Se presentan datos adicionales relacionados con la capacidad de Csy4* para facilitar la escisión por sn-casARN/proteína Cas9 de dianas de ADN bicatenario en el Ejemplo 7.

La FIG. 6A muestra el sn1-casARN y el sn2-casARN antes de la asociación y formación de pares de bases con

enlaces de hidrógeno entre ellos. La FIG. 6B ilustra el sn1-casARN que comprende un primer polinucleótido adyuvante y el sn2-casARN que comprende un segundo polinucleótido adyuvante después de la formación de pares de bases con enlaces de hidrógeno entre ellos para ilustrar la formación de un elemento de unión efector. La caja de línea discontinua superior (FIG. 6B, 610) muestra la formación de un elemento de unión efector, en este ejemplo un elemento de unión a ARN de Csy4*. La FIG. 6C ilustra la asociación del sn2-casARN con el lóbulo de nucleasa catalítico (FIG. 6C, 613) de SpyCas9 y la asociación del sn1-casARN con el lóbulo α -helicoidal (FIG. 6C, 614) de SpyCas9. También se muestra una proteína efectora Csy4* (FIG. 6C, 615), que es una variante de Csy4 sin actividad endorribonucleasa. La flecha gruesa que apunta hacia abajo indica el ensamblaje del sn2-casARN/lóbulo de nucleasa catalítica (FIG. 6C, 613) de SpyCas9, el sn1-casARN/lóbulo α -helicoidal (FIG. 6C, 614) de SpyCas9, y la proteína Csy4* (FIG. 6C, 615) en un complejo (FIG. 6C, 618). Este ejemplo ilustra el reclutamiento por sn1-casARN del lóbulo α -helicoidal y el reclutamiento por sn2-casARN del lóbulo de nucleasa catalítica en un complejo ternario estabilizado adicionalmente por la unión de la proteína Csy4* para recapitular la actividad de Cas9 para catalizar escisión de ADN específico de sitio.

En realizaciones adicionales del séptimo aspecto de la invención, la proteína efectora comprende al menos un dominio de dedo de cinc.

El primer polinucleótido adyuvante y/o segundo polinucleótido adyuvante también pueden comprender una o más horquillas. La FIG. 7A ilustra un sn1-casARN que comprende un primer polinucleótido adyuvante (FIG. 7A, 702 a 703) y un sn2-casARN que comprende un segundo polinucleótido adyuvante (FIG. 7A, 706 a 707). La figura muestra el sn1-casARN y el sn2-casARN antes de la asociación y formación de pares de bases con enlaces de hidrógeno entre ellos. La figura muestra un elemento de horquilla formado por formación de pares de bases con enlaces de hidrógeno entre bases en el primer polinucleótido adyuvante (FIG. 7A, 704) y un elemento de horquilla formado por formación de pares de bases con enlaces de hidrógeno en el segundo polinucleótido adyuvante (FIG. 7A, 708). La FIG. 7B ilustra el sn1-casARN/sn2-casARN ensamblado en un complejo activo con Cas9.

Asimismo, el primer polinucleótido adyuvante y/o segundo polinucleótido adyuvante también pueden comprender uno o más cadenas cuádruples G (también denominadas G4-ADN). Las cadenas cuádruples G son estructuras formadas en ácidos nucleicos por secuencias ricas en guanina. Habitualmente, cuatro bases de guanina se asocian mediante enlaces de hidrógeno de Hoogsteen para formar una estructura plana cuadrada (una tétrada de guanina). Cuando se apilan dos o más tétradas de guanina, forman una cadena cuádruple G. Una secuencia repetida ejemplar para la formación de estructuras de cadenas cuádruples G es $(GGN)_n$, en la que n es típicamente 4 o mayor. Pueden formarse una o más estructuras cuádruples intramolecularmente por secuencias dentro de, por ejemplo, un primer polinucleótido adyuvante y/o un segundo polinucleótido adyuvante. También se pueden formar cadenas cuádruples G mediante dos o más interacciones intermoleculares, por ejemplo, dos repeticiones en un primer polinucleótido adyuvante forman una cadena cuádruple mediante enlaces de hidrógeno de Hoogsteen con dos repeticiones diferentes en un segundo polinucleótido adyuvante. Además, pueden formarse una o más estructuras cuádruples entre secuencias de un primer polinucleótido adyuvante y un segundo polinucleótido adyuvante. Las estructuras cuádruples pueden estabilizarse adicionalmente por la presencia de un catión, por ejemplo, sodio o potasio. Dichos cationes generalmente ocupan un canal central entre cada par de tétradas. En algunas realizaciones, la formación de cadenas cuádruples G puede inducirse y/o estabilizarse mediante la introducción de una molécula que puede unir la estructura de cadena cuádruple G. Varias de dichas moléculas, tanto moléculas pequeñas como proteínas, son conocidas por los expertos habituales en la materia.

Todos los aspectos de la invención pueden comprender una proteína Cas9 (o según sea necesario secuencias de ácido nucleico que codifican una proteína Cas9) o una fusión de Cas9 (o según sea necesario secuencias de ácido nucleico que codifican una fusión de Cas9).

La expresión "marcador de afinidad" como se usa en el presente documento se refiere a uno o más restos que aumentan la afinidad de unión de un sn-casPN por otro sn-casPN y/o por una proteína Cas9. Algunas realizaciones de la presente invención usan una "secuencia de afinidad", que es una secuencia polinucleotídica que comprende uno o más marcadores de afinidad. Los ejemplos de secuencias de afinidad que pueden usarse para modificar un primer sn-casPN incluyen usar una secuencia de unión de MS2, secuencia de unión de U1A, secuencia de tallo-bucle, secuencia de unión de eIF4A, secuencia de unión de efector de tipo activador de la transcripción (TALE) (Valton, J., y col., "Overcoming Transcription Activator-like Effector (TALE) DNA Binding Domain Sensitivity to Cytosine Methylation," *J Biol Chem.*, 287(46), 38427-38432 (2012)) o secuencia de unión de dominio de dedo de cinc (Font, J., y col., "Beyond DNA: zinc finger domains as RNA-binding modules," *Methods Mol Biol.*, 649, 479-91 (2010); Isalan, M., y col., "A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter," *Nat Biotechnol.*, 19(7), 656-660 (2000)). Otros sn-casPN y/o la secuencia codificante de proteína Cas9 pueden modificarse para comprender un marcador de afinidad correspondiente: una secuencia codificante de MS2, secuencia codificante de U1A, secuencia codificante de proteína de unión a tallo-bucle, secuencia codificante de eIF4A, secuencia codificante de TALE o una secuencia codificante de dominio de dedo de cinc, respectivamente.

Se desvela una amplia diversidad de marcadores de afinidad en la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014.

Las expresiones "ligando" y "resto de unión a ligando" como se usan en el presente documento se refieren a restos

que facilitan la unión de un sn-casPN con otro sn-casPN o con una proteína Cas9. Los ligandos y restos de unión a ligando son marcadores de afinidad correspondientes.

Una realización de uso de un resto de ligando es construir un resto de unión a ligando en la proteína Cas9 o unir un resto de unión a ligando a un primer sn-casPN y modificar una secuencia polinucleotídica de un sn-casPN diferente para contener el ligando. Un ligando/resto de unión a ligando útil en la práctica de la presente invención es avidina o estreptavidina/biotina (véase, por ejemplo, Livnah, O., y col., "Three-dimensional structures of avidin and the avidin-biotin complex," PNAS, 90(11), 5076-5080 (1993); Airene, K.J., y col., "Recombinant avidin and avidin-fusion proteins," Biomol. Eng., 16(1-4), 87-92 (1999)). Un ejemplo de una proteína Cas9 con un resto de unión a ligando es una proteína Cas9 fusionada con un ligando avidina o estreptavidina diseñado para unirse con un sn-casPN biotinilado, en el que el sn-casPN comprende una secuencia polinucleotídica con la que se asocia la biotina. La biotina es un ligando de alta afinidad y alta especificidad para la proteína avidina o estreptavidina. Fusionando una cadena polipeptídica de avidina o estreptavidina con la proteína Cas9, la proteína Cas9 tiene una alta afinidad y especificidad por una sn-casPN-biotina biotinilada.

La biotilación es preferentemente en proximidad estrecha con los extremos 5' o 3' de un sn-casPN. La secuencia del sn-casPN y localización de la biotina se proporciona a fabricantes comerciales para síntesis de la sn-casPN-biotina. Se evalúan cambios del porcentaje de escisión y especificidad de un complejo modificado en unión a ligando de sn-casPN y proteína Cas9 como se describe, por ejemplo, en el Ejemplo 3, el Ejemplo 4 y/o el Ejemplo 9.

Los ejemplos de otros ligandos y restos de unión a ligando que pueden usarse de forma similar incluyen, pero sin limitación (ligando/resto de unión a ligando): estradiol/receptor de estrógenos (véase, por ejemplo, Zuo, J., y col., "Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants," Plant J., 24(2), 265-73 (2000)), rapamicina/FKBP12 y FK506/FKBP (véase, por ejemplo, B. Setscrew, y col., "A split-Cas9 architecture for inducible genome editing and transcription modulation," Nature Biotechnology, 33, 139-142 (2015); Chiu M.I., y col., "RAP1, a mammalian homolog of yeast Tor, interacts with the FKBP12/rapamycin complex," PNAS, 91(26), 12574-78 (1994)).

Otro ejemplo de un ligando y resto de unión a ligando es proporcionar uno o más aptámeros o aptámeros modificados en una secuencia polinucleotídica de un sn-casPN que tiene una alta afinidad y especificidad de unión por una región seleccionada de una proteína Cas9. En una realización, un resto de unión a ligando es un polinucleótido que comprende un aptámero (véase, por ejemplo, Navani, N.K., y col., "In vitro Selection of Protein-Binding DNA Aptamers as Ligands for Biosensing Applications," Biosensors and Biodetection, Methods in Molecular Biology, 504, 399-415 (2009); A. V. Kulbachinskiy, "Methods for Selection of Aptamers to Protein Targets," Biochemistry (Moscú), 72(13), 1505-18 (2007)). Los aptámeros son ácidos nucleicos funcionales monocatenarios que poseen capacidad de reconocimiento de ligando correspondiente. Habitualmente, el aptámero se localiza en el extremo 5' o 3' de un sn-casPN. En la práctica de la presente invención un ejemplo de un ligando es un complejo de casPN/Cas9.

En otra realización, un resto de unión a ligando comprende un polinucleótido modificado en el que un grupo funcional no nativo se introduce en posiciones orientadas en sentido contrario a la cara de enlace de hidrógeno de las bases del polinucleótido modificado, tales como la posición 5 de pirimidinas y la posición 8 de purinas (aptámeros modificados con tasa de disociación lenta o SOMAmeros, véase, por ejemplo, Rohloff, J.C., y col., "Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents," Molecular Therapy Nucleic Acids, 3, e201 (2014)). Un aptámero con alta especificidad y afinidad para proteínas Cas9 podría obtenerse mediante selección *in vitro* y exploración de una biblioteca de aptámeros.

En otra realización más, una secuencia de unión a aptámero establecida/aptámero se usa introduciendo la región de unión a aptámero en la proteína Cas9. Por ejemplo, un aptámero de unión a biotina puede introducirse en un sn-casPN y la proteína Cas9 puede biotinilarse de forma selectiva para formar un sitio de unión correspondiente para el aptámero de unión a biotina.

La creación de un sitio de unión de alta afinidad para un ligando seleccionado en una proteína Cas9 puede conseguirse usando procedimientos de ingeniería proteica conocidos por los expertos habituales en la materia a la vista de la orientación de la presente memoria descriptiva. Los ejemplos de dichos procedimientos de ingeniería proteica incluyen, diseño racional de proteínas, evolución dirigida usando diferentes procedimientos de selección y exploración para la biblioteca (por ejemplo, presentación en fagos), redistribución de ADN, procedimientos computacionales (por ejemplo ROSETTA, www.rosettacommons.org/software), o introducción de un ligando de alta afinidad conocido en Cas9. Pueden explorarse bibliotecas obtenidas por estos procedimientos para seleccionar agentes de unión de alta afinidad por proteína Cas9 usando, por ejemplo, un ensayo de presentación en fagos, un ensayo de supervivencia celular o un ensayo de unión.

En otro aspecto de la presente invención, al menos uno de los sn-casPN es un polinucleótido circular.

En otro aspecto más de la presente invención, al menos un sn-casPN lineal comprende una secuencia 5' terminal y/o una secuencia 3' terminal, y al menos una secuencia 5' terminal y/o secuencia 3' terminal comprende un resto de resistencia a exonucleasa asociado con la secuencia 5' terminal y/o secuencia 3' terminal. Los ejemplos de restos

resistentes a exonucleasa incluyen, pero sin limitación, una horquilla en la secuencia terminal, una secuencia de unión a polinucleótido monocatenario con la que se une un polinucleótido monocatenario, y una inversión de enlace.

Un aspecto de la invención se refiere a procedimientos de fabricación de los sn-casPN de la presente invención. En una realización, el procedimiento de fabricación comprende sintetizar químicamente uno o más de los sn-casPN. En algunas realizaciones, los sn-casPN comprenden bases de ARN, bases de ADN o una combinación de bases de ARN y bases de ADN. Asimismo, pueden sintetizarse cadenas principales de nucleobases distintas de o además de una cadena principal de fosfodiéster, por ejemplo, usando ácidos nucleicos, ácidos nucleicos peptídicos, ácido nucleico de treosa o combinaciones de los mismos. En algunas realizaciones, el procedimiento de fabricación comprende producir uno o más de los sn-casPN mediante transcripción *in vitro*.

En un aspecto, la presente invención se refiere a casetes de expresión que comprenden secuencias codificantes de polinucleótidos para dos o más sn-casPN y/o una proteína Cas9. Un casete de expresión de la presente invención comprende al menos un polinucleótido que codifica un sn-casPN de la presente invención. Los casetes de expresión útiles en la práctica de la presente invención pueden incluir además secuencias codificantes de proteína Cas9. En una realización, un casete de expresión comprende una secuencia codificante de sn-casPN. En otra realización, uno o más casetes de expresión comprenden secuencia codificante de sn-casPN y una secuencia codificante de proteína Cas9 afín. Los casetes de expresión habitualmente comprenden secuencias reguladoras que están implicadas en una o más de las siguientes: regulación de la transcripción, regulación postranscripcional y regulación de la traducción. Pueden introducirse casetes de expresión en una amplia diversidad de organismos incluyendo células bacterianas, células de levadura, células de insectos, células de mamíferos y células de plantas. Los casetes de expresión habitualmente comprenden secuencias reguladoras funcionales correspondientes a las células u organismo u organismos huéspedes en los que se introducen.

Un aspecto de la presente invención se refiere a vectores, incluyendo vectores de expresión, que comprenden secuencias codificantes de polinucleótidos para un sn-casPN y/o una proteína Cas9. Los vectores útiles para la práctica de la presente invención incluyen plásmidos, virus (incluyendo fagos) y fragmentos de ADN integrables (por ejemplo, fragmentos integrables en el genoma huésped mediante recombinación homóloga). Un vector se replica y actúa de forma independiente del genoma huésped, o puede, en algunos casos, integrarse en el genoma en sí mismo. Los vectores de replicación adecuados contendrán un replicón y secuencias de control procedentes de especies compatibles con la célula hospedadora de expresión pretendida. Un vector puede comprender uno o más casetes de expresión de secuencias codificantes de polinucleótidos para sn-casPN y/o una proteína Cas9. Los vectores incluyen, pero sin limitación, vectores bacterianos, vectores de levadura, vectores de algas, vectores de células de insectos, vectores de mamíferos y vectores víricos.

Las células hospedadoras transformadas son células que se han transformado o transfectado con los vectores construidos usando técnicas de ADN recombinante.

Se conocen en la técnica procedimientos generales para la construcción de vectores de expresión. Están disponibles en el mercado vectores de expresión para la mayoría de células huéspedes. Existen varios productos de software comerciales diseñados para facilitar la selección de vectores apropiados y construcción de los mismos, tales como plásmidos bacterianos para transformación bacteriana y expresión génica en células bacterianas, plásmidos de levadura para transformación celular y expresión génica en levadura y otros hongos, sistemas de expresión de algas para su uso en células de algas, vectores de células de insectos para transformación de células de insecto y expresión génica en células de insectos, vectores de mamífero para transformación de células de mamífero y expresión génica en células de mamífero o mamíferos, vectores víricos (incluyendo vectores retrovíricos, lentivíricos y adenovíricos) para transformación celular y expresión génica y procedimientos para permitir fácilmente la clonación de dichos polinucleótidos. SnapGene™ (GSL Biotech LLC, Chicago, Ill.; snapgene.com/resources/plasmid_files/your_time_is_valuable/), por ejemplo, proporciona una lista exhaustiva de vectores, secuencias de vectores individuales y mapas de vectores, así como fuentes comerciales para muchos de los vectores.

Los vectores de expresión también pueden incluir polinucleótidos que codifican marcadores proteicos (por ejemplo, marcadores de poli-His, marcadores de hemaglutinina, marcadores de proteína fluorescente, marcadores bioluminiscentes). Las secuencias codificantes para dichos marcadores proteicos pueden fusionarse con una secuencia codificante de proteína Cas9 o pueden incluirse en un casete de expresión, por ejemplo, en un vector de dirección.

En algunas realizaciones, los polinucleótidos que codifican sn-casPN y/o proteína Cas9 se unen operativamente con un promotor inducible, un promotor reprimible o un promotor constitutivo.

Aspectos de la invención se refieren a sistemas de vectores que comprenden uno o más vectores para expresión de sn-casPN y proteínas Cas9 en células procariotas o eucariotas. Como alternativa, pueden transcribirse sn-casPN y proteínas Cas9 *in vitro*, por ejemplo usando secuencias reguladoras de promotor T7 y polimerasa T7. La traducción de proteínas Cas9 también puede llevarse a cabo *in vitro*.

Pueden introducirse vectores que comprenden sn-casPN (que incluyen opcionalmente secuencias codificantes de

proteína Cas9) y propagarse en un procarionta. Se conocen bien en la técnica vectores procariontas. Habitualmente un vector procarionta comprende un origen de replicación adecuado para la célula huésped diana (por ejemplo, oriC procedente de *E. coli*, pUC procedente de pBR322, pSC101 procedente de *Salmonella*), origen 15A (procedente de p15A) y cromosomas artificiales bacterianos). Los vectores pueden incluir un marcador seleccionable (por ejemplo, genes que codifican resistencia para ampicilina, cloranfenicol, gentamicina y kanamicina). Zeocin™ (Life Technologies, Grand Island, NY) puede usarse como una selección en bacterias, hongos (incluyendo levadura), plantas y líneas celulares de mamíferos. En consecuencia, pueden diseñarse vectores que portan solamente un gen de resistencia a fármacos para zeocin para trabajo de selección en varios organismos. Se conocen promotores útiles para expresión de proteínas en procariontas, por ejemplo, T5, T7, ramnosa (inducible), arabinosa (inducible) y PhoA (inducible). Asimismo, se usan ampliamente promotores T7 en vectores que también codifican la ARN polimerasa T7. Los vectores procariontas también pueden incluir sitios de unión a ribosomas de diversa fuerza y señales de secreción (por ejemplo, mal, sec, tat, ompC y pelB). Además, los vectores pueden comprender promotores de ARN polimerasa para la expresión de sn-casARN. También se conocen bien secuencias de terminación de la transcripción de ARN polimerasa procarionta (por ejemplo, secuencias de terminación de la transcripción de *S. pyogenes*).

También se conocen en la técnica vectores de integración para transformación estable de procariontas (véase, por ejemplo, Heap, J. T., y col., "Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker," *Nucleic Acids Res.*, 40(8), e59 (2012)).

La expresión de proteínas en procariontas se lleva a cabo habitualmente en *Escherichia coli* con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de proteínas de fusión o no de fusión.

Están disponibles una amplia diversidad de promotores de ARN polimerasa adecuados para la expresión de sn-casARN y proteínas Cas9 en procariontas (véase, por ejemplo, Jiang, Y., y col., "Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system," *Environ Microbiol.*, 81(7), 2506-14 (2015); Estrem, S. T., y col., "Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase alpha subunit," *Genes Dev.*, 15, 13(16), 2134-47 (1999)).

Los vectores de fusión añaden varios aminoácidos a una proteína codificada en los mismos (por ejemplo, al extremo amino de la proteína recombinante). Dichos vectores de fusión cumplen uno o más fines. Los ejemplos incluyen, pero sin limitación, los siguientes: (i) aumentar la expresión de proteína recombinante; (ii) aumentar la solubilidad de la proteína recombinante; y (iii) ayudar a la purificación de la proteína recombinante actuando como un ligando en la purificación por afinidad. En vectores de fusión-expresión, se introduce en ocasiones un sitio de escisión proteolítica en el punto de unión del resto de fusión y la proteína recombinante. Esto permite la separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión. Dichas enzimas, y sus sitios de escisión proteolítica correspondientes, incluyen Factor Xa, trombina y enteroquinasa. Los ejemplos de vectores de fusión-expresión incluyen, pero sin limitación, los siguientes: pGEX, pMAL y pRIT5 que fusionan glutatión S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente, con la proteína recombinante diana. Los ejemplos de vectores de expresión de *E. coli* no de fusión inducibles adecuados incluyen, pero sin limitación, pTrc y pET 11d.

En algunas realizaciones, un vector es un vector de expresión que comprende sn-casPN. Los vectores de expresión también pueden incluir secuencias codificantes de una proteína Cas9. Los ejemplos de vectores para expresión en levadura *Saccharomyces cerevisiae* incluyen, pero sin limitación, los siguientes: pYepSec1, pMFA, pJRY88, pYES2 y picZ. Se conocen en la técnica procedimientos para expresión génica en células de levadura (véase, por ejemplo, *Methods in Enzymology*, Volumen 194, "Guide to Yeast Genetics and Molecular and Cell Biology, Part A," Christine Guthrie y Gerald R. Fink (eds.), Elsevier Academic Press, San Diego, Calif. (2004)). Habitualmente, la expresión de genes que codifican proteínas en levadura requiere un promotor unido operativamente a una región codificante de interés más un terminador de la transcripción. Pueden usarse diversos promotores de levadura para construir casetes de expresión para expresión de genes en levadura. Los ejemplos de promotores incluyen, pero sin limitación, promotores de genes que codifican las siguientes proteínas de levadura: alcohol deshidrogenasa 1 (ADH1) o alcohol deshidrogenasa 2 (ADH2), fosfoglicerato quinasa (PGK), triosa fosfato isomerasa (TPI), gliceraldehído-3-fosfato deshidrogenasa (GAPDH; también conocida como TDH3 o triosa fosfato deshidrogenasa), galactosa-1-fosfato uridil-transferasa (GAL7), UDP-galactosa epimerasa (GAL10), citocromo c1 (CYC1) y fosfatasa ácida (PHO5). También pueden usarse promotores híbridos, tales como el ADH2/GAPDH, CYC1/GAL10 y el promotor ADH2/GAPDH (que es inducido a concentraciones celulares de glucosa bajas, por ejemplo, de aproximadamente 0,1 por ciento a aproximadamente 0,2 por ciento). En *S. pombe*, los promotores adecuados incluyen el promotor nmt1 reprimido por tiamina y el promotor de citomegalovirus constitutivo en pTL2M.

Se conocen en la técnica promotores de la ARN polimerasa III de levadura (por ejemplo, promotores de genes de 5S, U6 o RPR1) así como secuencias de terminación de polimerasa III (véase, por ejemplo, www.yeastgenome.org; Harismendy, O., y col., "Genome-wide location of yeast RNA polymerase III transcription machinery," *EMBO J.*, 22(18), 4738-4747 (2003)).

Un promotor de expresión de proteínas puede ser inducible o constitutivo. En algunas realizaciones, un promotor preferido es un promotor inducible estrechamente regulado, de modo que puede conseguirse un alto número de

copias en ausencia de la expresión. Los ejemplos incluyen, pero sin limitación, los promotores GAL1p y GAL1Op normalmente divergentes, que se suprimen estrechamente en medios de glucosa y son inducidos en gran medida por galactosa después de haberse aliviado la represión de catabolitos por crecimiento en una fuente de carbono no represiva tal como glicerol o lactato. Una fase abierta de lectura que codifica un polipéptido puede insertarse en un vector GAL1p (véase, por ejemplo, Cartwright, y col., "Use of β -lactamase as a secreted reporter of promoter function in yeast," *Yeast*, 10, 497 (1994); y Harley, y col., "Transmembrane Protein Insertion Orientation in Yeast Depends on the Charge Difference across Transmembrane Segments, Their Total Hydrophobicity, and Its Distribution," *J. Biol. Chem.*, 273, 24963 (1998)). Otros vectores y promotores que pueden usarse incluyen el promotor GAL1-CYCp híbrido en el vector de Yep URA3 leu2d pPAP1488 en la cepa PAP1502 (véase, por ejemplo, Pedersen, y col., "Expression in High Yield of Pig α 1 β 1 Na,K-ATPase and Inactive Mutants D369N and D807N in *Saccharomyces cerevisiae*," *J. Biol. Chem.*, 271, 2514-222 (1996)). Esta cepa tiene el plásmido pPAP1488 integrado en el locus Trp1. Esto proporciona una copia adicional del gen de GAL4 conducido por el promotor GAL10 y, cuando se induce expresión de GAL, se producen altos niveles del activador positivo de Gal4p.

En este sistema de vector, el crecimiento en ausencia de uracilo produce un número de copias del vector de 15 a 20, determinado por funciones de replicación de 2 micrómetros. El número de copias del vector puede aumentarse adicionalmente, en al menos 10 veces, cultivando las células de levadura en medios que carecen de leucina, debido al promotor muy débil asociado con el alelo leu2d defectuoso. Un aumento proporcional de la expresión conducida por GAL1p requiere los altos niveles inducidos por galactosa del activador de Gal4p proporcionado en la cepa PAP1502 por el plásmido PAP1488 integrado. Puede usarse cualquier otra cepa de *S. cerevisiae* ura3 leu2 Gal+ en la que se inserta este plásmido en lugar de PAP1502.

Otro promotor de levadura que puede usarse es el promotor del gen de glicerol-3-fosfato deshidrogenasa (GPD1). La expresión de polipéptidos usando el promotor de GPD1 puede estar regulada por la presencia (reprimida) o ausencia (desreprimida) de altos niveles de glucosa o sacarosa en un medio de fermentación. Como alternativa, una fuente de carbono no represora, tal como etanol o glicerol, puede añadirse al medio de fermentación (véase, por ejemplo, patente de los Estados Unidos n.º 5.667.986).

La regulación del número de copias de plásmidos puede proporcionar algún control sobre el nivel de productos de ARN expresados a partir de promotores de ARN polimerasa III. Asimismo, la transcripción de ARN polimerasa III puede regularse en levadura (Dingermann, T., y col., "RNA polymerase III catalysed transcription can be regulated in *Saccharomyces cerevisiae* by the bacterial tetracycline repressor-operator system," *EMBO J.*, 11(4), 1487-92 (1992)).

Además de un promotor, pueden usarse varias secuencias de activación cadena arriba (UAS), también denominadas potenciadores, para potenciar la expresión de polipéptidos. Las secuencias de activación cadena arriba ejemplares para expresión en levadura incluyen las UAS de genes que codifican estas proteínas: CYC1, ADH2, GAL1, GAL7, GAL10 y ADH2. Las secuencias de terminación de la transcripción ejemplares para expresión en levadura incluyen las secuencias de terminación de los genes de factor α , CYC1, GAPDH y PGK. Pueden usarse una o múltiples secuencias de terminación.

Pueden optimizarse los codones de cualquier región codificante de proteína expresada en células de levadura para expresión en la célula de levadura huésped específica para genomodificar, como es bien sabido en la técnica.

Pueden clonarse promotores, terminadores y regiones codificantes adecuados en vectores lanzadera de *E. coli*-levadura y transformarse en células de levadura. Estos vectores permiten la propagación de cepas en cepas tanto de levadura como de *E. coli*. Habitualmente, el vector contiene un marcador seleccionable y secuencias que permiten la replicación autónoma o integración cromosómica en cada huésped. Son ejemplos de plásmidos usados habitualmente en levadura los vectores lanzadera pRS423, pRS424, pRS425 y pRS426 (Colección Americana de Cultivos Tipo, Manassas, VA). Estos plásmidos contienen un origen de replicación de levadura de 2 micrómetros, un origen de replicación de *E. coli* (por ejemplo, pMB1) y un marcador seleccionable.

El ejemplo 15 presenta una ilustración de ingeniería genómica en *Saccharomyces cerevisiae* usando sn-casPN. Se usan un vector de Cas9 y dos pares de vectores de sn1-casARN/sn2-casARN para modificar el genoma de la levadura. Este protocolo proporciona datos para verificar que el sistema de sn1-casARN/sn2-casARN/proteína Cas9 proporciona actividad endonucleasa mediada por ARN específica en loci genómicos endógenos diana en levadura. Las construcciones también se usan en experimentos para verificar que el sistema de sn1-casARN/sn2-casARN/proteína Cas9 proporciona actividad endonucleasa mediada por ARN específica en loci genómicos endógenos diana en levadura y pueden estimular acontecimientos de recombinación homólogos en dichos loci usando ADN donante. Otros loci cromosómicos en *S. cerevisiae* pueden ser de forma similar diana para modificación por selección de secuencias espaciadoras y oligonucleótidos donantes apropiados. Pueden introducirse genes funcionales en el genoma de *S. cerevisiae* sin alteración de genes endógenos. Asimismo, puede usarse introducción de marcadores seleccionables en genes diana endógenos para proporcionar mutaciones de inactivación seleccionables de los genes diana.

También están ampliamente disponibles vectores de integración para transformación estable de levadura (Steams T., y col., "Manipulating yeast genome using plasmid vectors," *Methods Enzymol.*, 185, 280-97 (1990)).

Para el uso de sn-casPN en células de algas, se conocen bien en la técnica vectores y secuencias de control de la expresión adecuados (véase, por ejemplo, Hallmann, A., "Algal Transgenics and Biotechnology," *Transgenic Plant Journal*, 1(1), 81-98 (2007); Oey, M., y col., "Gateway-Assisted Vector Construction to Facilitate Expression of Foreign Proteins in the Chloroplast of Single Celled Algae," DOI: 10.1371/journal.pone.0086841 (2013)) incluyendo promotores de la ARN polimerasa III (véase, por ejemplo, Dieci, G., y col., "Eukaryotic snoRNAs: A paradigm for gene expression flexibility," *Genomics*, 94(2), 83-88 (2009)). Asimismo, están disponibles en el mercado sistemas de expresión de algas (Algae Expression and Engineering Products, ThermoFisher Scientific, Grand Island, NY). Los vectores de expresión también pueden incluir secuencias codificantes de una proteína Cas9.

Para el uso de sn-casPN en insectos o células de insectos, se conocen bien en la técnica secuencias de control de la expresión adecuadas. En algunas realizaciones, es deseable que la secuencia de control de la expresión comprenda un promotor constitutivo. Los ejemplos de promotores fuertes adecuados incluyen, pero sin limitación, los siguientes: los promotores de baculovirus para el piO, polihedrina (polh), p 6.9, cápside, UAS (contiene un sitio de unión a Gal4), Ac5, genes de tipo cathepsina, el promotor del gen de actina de *B. mori*; hsp70 de *Drosophila melanogaster*, actina, α -1-tubulina o promotores del gen de ubiquitina, promotores de VSR o VTMR, promotor de copia, promotor de gypsy y el promotor del gen de IE de citomegalovirus. Los ejemplos de promotores débiles que pueden usarse incluyen, pero sin limitación, los siguientes: los promotores de baculovirus para los genes de iel, ie2, ieO, etl, 39K (también conocido como pp31) y gp64. Si se desea aumentar la cantidad de expresión génica desde un promotor débil, pueden usarse elementos potenciadores, tales como el elemento potenciador de baculovirus, hr5, junto con el promotor. Los vectores de expresión también pueden incluir secuencias codificantes de una proteína Cas9.

En algunas realizaciones, la secuencia de control de la expresión comprende un promotor específico de órgano o específico de tejido. Muchas secuencias de control de la expresión, por ejemplo, promotores adecuados que dirigen la expresión en glándulas de seda de insectos incluyen el promotor p25 de *Bombyx mori*, que dirige la expresión específica de órgano en la glándula de seda posterior, y el promotor del gen de cadena pesada de fibroína de seda, que dirige la expresión específica de genes en la glándula de seda intermedia.

Los ejemplos de secuencias de control de la expresión regulables de insectos (por ejemplo, que comprenden un promotor inducible y/o elemento potenciador) incluyen, pero sin limitación, los siguientes: promotores de hsp70 de *Drosophila*, promotor de metalotioneína de *Drosophila*, un promotor regulado por ecdisona y otros sistemas inducibles bien conocidos. Puede usarse un interruptor molecular regulable por Tet junto con cualquier promotor constitutivo (por ejemplo, junto con el promotor de CMV-IE o promotores de baculovirus). Otro tipo de promotor inducible es un promotor tardío o muy tardío de baculovirus que se activa solamente después de infección por un baculovirus.

Para la expresión de sn-casPN en insectos, son conocidos en la técnica promotores de ARN polimerasa III, por ejemplo, el promotor U6. También se conocen elementos conservados de promotores de ARN polimerasa III en insectos (véase, por ejemplo, Hernandez, G., "Insect small nuclear RNA gene promoters evolve rapidly yet retain conserved features involved in determining promoter activity and RNA polymerase specificity," *Nucleic Acids Res.*, 35(1), 21-34 (2007)).

Son convencionales procedimientos para diseñar y preparar construcciones adecuadas para generar insectos transgénicos o vectores para infección de un insecto. También son convencionales procedimientos para transformación, cultivo y manipulación de líneas celulares de insectos. Los ejemplos de líneas celulares de insectos incluyen, pero sin limitación, los siguientes: células de *Antheraea*, Tn-368, células S2 de *Drosophila*, células High Five™ (Life Technologies, Grand Island, NY), células Sf21 y células Sf9. Están disponibles en la técnica líneas celulares de insectos, por ejemplo, de la Colección Americana de Cultivos tipo (Manassas, VA).

Una diversidad de líneas celulares de insectos lepidópteros inmortalizadas es adecuada para infección por vectores/construcciones que comprenden los sn-casPN/proteínas Cas9 de la presente invención. Los ejemplos de líneas celulares de insectos lepidópteros inmortalizadas que son adecuadas para infección por los vectores/construcciones de la invención incluyen, pero sin limitación, los siguientes: Sf9 y Tn 5B1-4.

En otra realización, el vector es un vector basado en transposones. Un vector basado en transposones es un vector vírico que comprende además repeticiones terminales invertidas de un transposón adecuado entre las que se clona el gen de interés. Uno o más genes, bajo el control de una secuencia de control de la expresión adecuada, se clonan en el vector basado en transposones. En algunos sistemas, el vector basado en transposones porta su propia transposasa. Sin embargo, típicamente el vector basado en transposones no codifica una transposasa adecuada. En este caso, el vector se coinfecta en un insecto (por ejemplo, una larva de insecto) con un virus o plásmido auxiliar que proporciona una transposasa. El vector recombinante, en general con un auxiliar, se introduce por procedimientos convencionales (por ejemplo, microinyección) en un óvulo o embrión temprano. Los transgenes se integran en un sitio de transposón (por ejemplo, secuencias correspondientes a la repetición terminal invertida del transposón) en el genoma de insecto. Los ejemplos de tipos adecuados de vectores basados en transposones incluyen, pero sin limitación, los siguientes: Mimos, mariner, Hermes, sleeping beauty y piggyBac.

Los elementos de repetición cortos, específicos de TTAA son un grupo de transposones (elementos móviles de clase

II) que tienen estructuras y propiedades de movimiento similares. Los vectores piggyBac son los más exhaustivamente estudiados de estos elementos de inserción. piggyBac es de 2,4 kb de longitud y termina en repeticiones invertidas perfectas de 13 pb, con repeticiones invertidas de 19 pb internas adicionales localizadas de forma asimétrica con respecto a los extremos. Un vector piggyBac puede codificar una transposasa de transacción que facilita su propio movimiento. Como alternativa, las secuencias codificantes de transposasa pueden suprimirse y esta función puede proporcionarse en un plásmido o virus auxiliar. Algunos vectores piggyBac tienen genes no esenciales suprimidos para facilitar la clonación de insertos grandes. Pueden clonarse insertos de hasta 15 kB en determinados vectores piggyBac. Por ejemplo, esto permite la inserción de aproximadamente seis o siete genes con sus secuencias de control de la expresión. Por ejemplo, puede introducirse una colección de sn-casPN juntos mediante un único vector de transposón en un único sitio en un genoma de insecto.

Se han desarrollado varios vectores piggyBac para transgénesis de insectos. Se desarrollaron dos construcciones por análisis de mutaciones de supresión dentro y fuera de los límites del transposón. Usando dichas construcciones, es posible aumentar la cantidad de material genético movilizado por la transposasa de piggyBac minimizando el tamaño del vector. Los requisitos mínimos para movimiento incluyen los dominios de repetición 5' y 3' terminales y secuencias diana TTAA asociadas. Se requiere típicamente un mínimo de 50 bases que separan los sitios diana TTAA del elemento para movilización eficaz.

El elemento piggyBac puede transponerse en células de insectos portando al mismo tiempo un gen marcador y puede producirse movimiento del elemento piggyBac en células de especies de lepidópteros relacionadas de forma distante con la especie de la que se originó. Por ejemplo, se ha mostrado que piggyBac transforma *D. melanogaster*, *Anastrepha suspensa* (mosca de la fruta oriental), *Bactrocera dorsalis*, *Bombyx mori*, *Pectinophora gossypiella*, *Tribolium castellani* y varias especies de mosquito. Al menos tres especies de lepidópteros, *P. gossypiella*, *T. ni* y *B. mori*, se han transformado con éxito usando el elemento piggyBac.

Habitualmente, un virus o plásmido auxiliar que expresa una transposasa se coinfecta con el vector basado en transposones. La expresión de la transposasa se determina por la elección de promotor para el sistema de insectos que se ensaya. Los ejemplos de construcciones auxiliares conducidas por promotor que son útiles para transformación de lepidópteros incluyen, pero sin limitación, los siguientes: hsp70 de *Drosophila*, promotor del baculovirus y promotor de actina 5C de *Drosophila*. Para orientación adicional sobre el uso de vectores basados en baculovirus, véase, por ejemplo, la publicación de solicitud de patente de PCT N.º WO/2005/042753, publicada el 12 de mayo de 2005.

Son convencionales procedimientos para generar insectos transgénicos. Por ejemplo, se colocan uno o más genes para introducir bajo el control de una secuencia de control de la expresión adecuada y se clonan en un vector (por ejemplo, un vector de baculovirus atenuado o un vector vírico no permisivo que no es infeccioso para el insecto diana). Las secuencias para introducir en el insecto están flanqueadas por secuencias genómicas del insecto. La construcción se introduce después en un óvulo de insecto (por ejemplo, mediante microinyección). Los transgenes se integran después por recombinación homóloga de las secuencias flanqueantes en secuencias complementarias en el genoma de insecto.

Se conocen procedimientos para introducir construcciones en un embrión para generar un insecto transgénico (por ejemplo, por microinyección). La supervivencia es típicamente alta (hasta 75 %) para embriones microinyectados. En general, se atraviesan óvulos preblastodérmicos con un capilar de vidrio fino que contiene una solución del ADN plasmídico y/o virus recombinante. Se exploran larvas GO eclosionadas de los óvulos inyectados con virus para expresión de los genes transfectados. La cría de insectos G1 transgénicos con insectos normales da como resultado herencia mendeliana.

Una vez que se ha integrado de forma estable un transgén en el genoma de un óvulo de insecto o embrión temprano, pueden usarse procedimientos convencionales para generar un insecto transgénico, en el que el transgén está presente en todas las células somáticas y germinales del insecto. Son convencionales procedimientos para producir insectos transgénicos homocigotos (por ejemplo, usando retrocruzamientos adecuados).

Seleccionando secuencias de control de la expresión apropiadas para cada uno de los genes, puede diseñarse un insecto transgénico múltiple que comprende copias integradas de forma genómica de sn-casPN y pueden diseñarse genes de proteínas Cas9 de modo que los genes se expresan a niveles adecuados, en el momento deseado durante el crecimiento de insectos.

En otro aspecto, los sn-casPN se incorporan en vectores de mamíferos para su uso en células de mamífero. Los vectores de expresión también pueden incluir secuencias codificantes de una proteína Cas9. Están disponibles en el mercado un gran número de vectores de mamíferos adecuados para su uso con los sn-casPN de la presente invención (por ejemplo, de Life Technologies, Grand Island, NY; NeoBiolab, Cambridge, MA; Promega, Madison, WI; DNA2.0, Menlo Park, CA; Addgene, Cambridge, MA).

Pueden usarse vectores procedentes de virus de mamíferos para expresar los sn-casPN y proteínas Cas9 de la presente invención en células de mamíferos. Estos incluyen vectores procedentes de virus tales como adenovirus, papovirus, herpesvirus, poliomavirus, citomegalovirus, lentivirus, retrovirus y virus de simio 40 (SV40) (véase, por

ejemplo, Kaufman, R. J., "Overview of vector design for mammalian gene expression," *Molecular Biotechnology*, 16(2), 151-160 (2000); Cooray S., y col., "Retrovirus and lentivirus vector design and methods of cell conditioning," *Methods Enzymol.*, 507, 29-57 (2012)). Las secuencias reguladoras unidas operativamente con los componentes de sn-casPN/Cas9 pueden incluir secuencias de unión a activador, potenciadores, intrones, secuencias de reconocimiento de poliadenilación, promotores, secuencias de unión a represores, estructuras de tallo-bucle, secuencias de inicio de la traducción, secuencias líderes de la traducción, secuencias de terminación de la transcripción, secuencias de terminación de la traducción, sitios de unión a cebadores y similares. Son promotores usados habitualmente los promotores de mamíferos constitutivos CMV, EF1a, SV40, PGK1 (de ratón o humano), Ubc, CAG, CaMKIIa y beta-Act, y otros conocidos en la técnica (véase, por ejemplo, Khan, K. H., "Gene Expression in Mammalian Cells and its Applications," *Advanced Pharmaceutical Bulletin*, 3(2), 257-263 (2013)). Asimismo, para expresión de los sn-casPN de la presente invención, se usan promotores de ARN polimerasa III de mamífero, incluyendo HI y U6.

El adenovirus es un miembro de la familia Adenoviridae. Los vectores de adenovirus proceden de adenovirus. El adenovirus es un virus icosaédrico sin envoltura, de tamaño mediano. Está compuesto de una nucleocápside y un genoma de ADN lineal bicatenario que puede usarse como un vector de clonación. Los conocimientos y datos extensivos sobre regulación de la transcripción de adenovirus favorecen la genomodificación de vectores de adenovirus modificados para la expresión de genes insertados. Para este fin, se suprimieron las regiones tempranas E1 y E3, haciendo de este modo al virus incapaz de replicar, lo que hace necesario que la célula huésped proporcione esta función en trans. Habitualmente se inserta un casete de expresión que comprende secuencias codificantes de proteínas (por ejemplo, una proteína Cas9) para reemplazar la región E1 suprimida. En el casete, se coloca un gen bajo el control de un promotor tardío principal adicional o bajo el control de un promotor exógeno, tal como citomegalovirus o promotor regulable seleccionado.

El genoma de adenovirus puede manipularse de tal modo que codifique y exprese un producto génico de interés inactivando al mismo tiempo la capacidad del adenovirus para replicar mediante un ciclo lítico normal. Algunos de dichos vectores adenovíricos incluyen los procedentes de la cepa de adenovirus Ad de tipo 5 dl324 u otras cepas de adenovirus (por ejemplo, Ad2, Ad3 y Ad7). En determinadas circunstancias, los adenovirus recombinantes pueden ser ventajosos porque no pueden infectar células que no están en división, y pueden usarse para infectar células epiteliales y una diversidad de otros tipos celulares. Además, la partícula vírica es relativamente estable y es susceptible de purificación y concentración. La capacidad de carga de un genoma adenovírico para ADN extraño es de hasta aproximadamente 8 kilobases, que es grande en comparación con otros vectores de suministro de genes. El adenovirus de ADN bicatenario no se integra en el genoma, lo que limita su uso a expresión episómica, transitoria. Debido a que no está integrado en el genoma de una célula huésped (a diferencia de ADN retrovívico) se evitan problemas potenciales tales como mutagénesis por inserción.

El virus adenoasociado (VAA), un miembro de ADN monocatenario de la familia Parvoviridae, es un virus deficiente en replicación de forma natural. Como el adenovirus, puede infectar células que no están en división; sin embargo, tiene la ventaja de competencia de integración. Los vectores de VAA están entre los vectores víricos más frecuentemente usados para terapia génica. Se conocen doce serotipos humanos de VAA (VAA de serotipo 1 [VAA-1] a VAA-12) y más de 100 serotipos no humanos. Varios factores han aumentado el potencial de VAA como un vehículo de suministro para aplicaciones de terapia génica, incluyendo la falta de patogenicidad del virus, la persistencia del virus y los muchos serotipos disponibles. VAA es un virus sin envoltura, pequeño (25 nm) que comprende un genoma de ADN monocatenario lineal. Habitualmente solo se produce infección productiva por VAA en presencia de un virus auxiliar, por ejemplo, adenovirus o herpesvirus. En ausencia de virus auxiliar, VAA (serotipo 2) puede hacerse latente integrándose en el cromosoma humano 19q13.4 (locus VAAS-1) (véase, por ejemplo, Daya, S., y col., "Gene Therapy Using Adeno-Associated Virus Vectors," *Clinical Microbiology Reviews*, 21(4), 583-593 (2008)).

El virus vaccinia es un miembro de la familia de poxvirus. Los vectores de vaccinia proceden de virus vaccinia. El genoma del virus vaccinia está comprendido por un ADN bicatenario de casi 200.000 pb. Se replica en el citoplasma de la célula huésped. Las células infectadas con el virus vaccinia producen hasta 5000 partículas víricas por célula, lo que conduce a niveles altos de expresión para productos génicos codificados. El sistema de vaccinia se ha usado eficazmente en cultivo a escala muy grande (1000 l) para producir varias proteínas, incluyendo rgp160 de VIH-1 y protrombina humana.

El retrovirus es un miembro de la familia Retroviridae. Los vectores retrovívicos proceden de retrovirus. Los retrovirus son virus de ARN que se replican mediante un intermedio de ARN bicatenario. Una ventaja de usar un retrovirus como vector es que la mayoría de retrovirus no matan al huésped, sino que en su lugar producen viriones descendientes durante un periodo de tiempo indefinido. Por lo tanto, los vectores retrovívicos (i) pueden usarse para fabricar líneas celulares transformadas de forma estable, (ii) proporcionar expresión génica vírica conducida por promotores fuertes, que pueden subvertirse para controlar la expresión de transgenes; y (iii) incluyen los procedentes de retrovirus que tienen una amplia serie de huéspedes (por ejemplo, cepas anfitriónicas de virus de leucemia murina (VLM)) permitiendo de este modo la transfección de muchos tipos celulares.

Los sistemas de expresión génica exógenos basados en el vector retrovívico también son un procedimiento para generar líneas celulares de mamíferos de alta expresión, estables.

El lentivirus es un miembro de la familia Retroviridae. Un virus de ARN monocatenario, puede infectar células tanto en división como no en división, así como proporcionar expresión estable mediante integración en el genoma. Para aumentar la seguridad del lentivirus, los componentes necesarios para producir un vector vírico se dividen entre múltiples plásmidos. Los vectores de transferencia son habitualmente incompetentes para replicación y pueden contener adicionalmente una supresión en la LTR 3', los que hace al virus "autoinactivante" (AIN) después de la integración. Habitualmente se usan plásmidos de empaquetamiento y envoltura en combinación con un vector de transferencia. Por ejemplo, un plásmido de empaquetamiento puede codificar los genes de Gag, Pol, Rev y Tat. Un plásmido de transferencia puede comprender LTR víricas y la señal de empaquetamiento psi. Generalmente uno o más promotores adecuados se unen operativamente a los genes para expresar (por ejemplo, sn-casPN y/o secuencias codificantes de proteína Cas9) debido a que la LTR 5' es un promotor débil y requiere la presencia de Tat para activar la expresión. El plásmido de envoltura comprende una proteína de envoltura (generalmente VSVG debido a su amplio intervalo de infecciosidad).

Los vectores lentivíricos basados en virus de la inmunodeficiencia humana de tipo 1 (VIH-1) tienen proteínas accesorias adicionales que permiten la integración en ausencia de división celular. Se han diseñado vectores de VIH-1 para abordar varias preocupaciones de seguridad. Estos incluyen expresión separada de los genes víricos en *trans* para prevenir acontecimientos de recombinación lo que conduce a la generación de virus competentes para replicación. Asimismo, el desarrollo de vectores autoinactivantes (AIN) reduce el potencial de transactivación de genes adyacentes y permite la incorporación de elementos reguladores para dirigir la expresión génica a tipos celulares particulares (véase, por ejemplo, Cooray, S., y col., "Retrovirus and lentivirus vector design and methods of cell conditioning," *Methods Enzymol.*, 507, 29-57 (2012)).

En algunas realizaciones, un vector de expresión de mamífero recombinante es capaz de dirigir preferentemente la expresión del ácido nucleico en un tipo celular particular (por ejemplo, usando elementos reguladores específicos de tejido para expresar un polinucleótido). Se conocen en la técnica elementos reguladores específicos de tejido e incluyen, pero sin limitación, el promotor de la albúmina, promotores específicos linfoides, promotores específicos de neuronas (por ejemplo, el promotor de neurofilamentos), promotores específicos del páncreas, promotores específicos de glándulas mamarias (por ejemplo, promotor de suero de leche) y, en particular, promotores de receptores de linfocitos T e inmunoglobulinas. También están abarcados promotores con desarrollo regulado, (por ejemplo, los promotores de *hox* murinos y el promotor de alfa-fetoproteína).

Están disponibles en el mercado varios vectores para su uso en células de mamíferos, por ejemplo: pcDNA3 (Life Technologies, Grand Island, NY); vectores de expresión adaptables, vectores transitorios, vectores estables y vectores lentivíricos (DNA 2.0, Menlo Park, CA); y vector pFNIOA (ACT) Flexi® (Promega, Madison, WI). Asimismo, los siguientes elementos pueden incorporarse en vectores para su uso en células de mamíferos: promotores de ARN polimerasa II unidos operativamente con secuencias codificantes de Cas9; promotores de ARN polimerasa III unidos operativamente con secuencias codificantes de sn-casARN; marcadores seleccionables (por ejemplo, G418, gentamicina, kanamicina y Zeocin™ (Life Technologies, Grand Island, NY)). También pueden añadirse secuencias de dirección nuclear, por ejemplo, secuencias codificantes de proteína Cas9.

Los elementos reguladores también pueden dirigir la expresión de una manera dependiente del tiempo, que también puede ser o no específico de tejido o tipo celular (por ejemplo, de una manera dependiente de ciclo celular o dependiente de estadio del desarrollo). En algunas realizaciones, los vectores comprenden uno o más promotores de ARN polimerasa III (por ejemplo, unido operativamente con secuencias codificantes de sn-casPN), uno o más promotores de ARN polimerasa II (por ejemplo, unidos operativamente con una secuencia codificante de proteína Cas9), uno o más promotores de ARN polimerasa I, o combinaciones de los mismos. Como se ha observado anteriormente, los ejemplos de promotores de ARN polimerasa III de mamífero incluyen, pero sin limitación, los siguientes: promotores U6 y H1. Se han analizado anteriormente ejemplos de promotores de ARN polimerasa II. Son bien conocidos en la técnica promotores de ARN polimerasa I.

Se han utilizado numerosas líneas celulares de mamíferos para la expresión de productos génicos que incluyen HEK 293 (riñón embrionario humano) y CHO (ovario de hámster chino). Estas líneas celulares pueden transfectarse mediante procedimientos convencionales (por ejemplo, usando fosfato de calcio o polietilimina (PEI), o electroporación). Otras líneas celulares de mamíferos típicas incluyen, pero sin limitación, HeLa, U2OS, 549, HT1080, CAD, P19, NIH 3T3, L929, N2a, células 293 de riñón embrionario humano, MCF-7, Y79, SO-Rb50, Hep G2, DUKX-X11, J558L y células de riñón de cría de hámster (BHK).

Los sistemas de sn-casPN/proteína Cas9 de la presente invención pueden usarse para manipular bioprocesos de células de mamífero para fabricación. Las células CHO y células de mieloma de ratón (incluyendo células Sp2/0 y NS0) son las células de mamíferos huéspedes más ampliamente usadas. Dos derivados de la línea celular CHO, CHO-K1 y CHO pro-3, han dado lugar a las dos líneas celulares más habitualmente usadas en el bioprocesamiento en la actualidad, DG44 y DUKX-X11 (ambas de estas líneas celulares se genomodificaron para ser deficientes en la actividad dihidrofolatorreductasa).

El Ejemplo 14 describe la modificación de células CHO para aplicaciones industriales. Este ejemplo describe uso de los polinucleótidos de nexos divididos (sn-casPN) de la presente invención para modificar el genoma de una célula CHO. También se describe un experimental para validación de secuencias y selección de células modificadas por

sn-casPN para usos futuros en aplicaciones industriales (por ejemplo, producción de anticuerpos). Los procedimientos proporcionan la modificación de loci cromosómicos en células CHO por selección de secuencias espaciadoras apropiadas para sn-casPN. La selección es específica de una diana génica específica y el procedimiento perfilado en el ejemplo es fácilmente modificable por un experto habitual en la materia para otras dianas génicas.

Se conocen en la técnica procedimientos de introducción de polinucleótidos (por ejemplo, un vector de expresión) en células huéspedes y se seleccionan típicamente basándose en el tipo de célula huésped. Dichos procedimientos incluyen, por ejemplo, infección vírica o de bacteriófago, transfección, conjugación, electroporación, precipitación con fosfato cálcico, transfección mediada por polietilimina, transfección mediada por DEAE-dextrano, fusión de protoplastos, lipofección, transfección mediada por liposomas, tecnología de pistola de partículas, microinyección directa y suministro mediado por nanopartículas.

En algunas realizaciones de la presente invención es útil expresar todos los componentes para complejos que comprenden sn-casPN y proteínas Cas9 en una célula huésped. Puede conseguirse expresión de secuencias que codifican sn-casARN y proteína Cas9 en una célula huésped mediante el uso de casetes de expresión como se ha descrito anteriormente. Sin embargo, no se realiza expresión de sn-casADN en una célula diana con el uso de vectores de clonación convencionales. Se han desarrollado vectores de expresión de ADN monocatenarios, que pueden generar intracelularmente moléculas de ADN monocatenario (Chen, Y., y col., "Intracellular production of DNA enzyme by a novel single-stranded DNA expression vector," *Gene Ther.*, 10(20), 1776-80 (2013); Miyata S., y col., "*In vivo* production of a stable single-stranded cDNA in *Saccharomyces cerevisiae* by means of a bacterial retron," *Proc Natl Acad Sci USA*, 89, 5735-39 (1992); Mirochnitchenko, O., y col., "Production of single-stranded DNA in mammalian cells by means of a bacterial retron," *J Biol Chem*, 269, 2380-2383 (1994); Mao J., y col., "Gene regulation by antisense DNA produced *in vivo*," *J Biol Chem*, 270, 19684-87 (1995)). Habitualmente estos vectores de expresión de ADN monocatenario se basan en la transcripción de una secuencia de ADN monocatenario seleccionada para formar un transcrito de ARN que es el sustrato para una transcriptasa inversa y RNasaH para generar el ADN monocatenario seleccionado en una célula huésped. Por ejemplo, los componentes de vectores de expresión de ADN monocatenario comprenden con frecuencia, una secuencia codificante de transcriptasa inversa (por ejemplo, un gen de transcriptasa inversa vírica de leucemia murina de Moloney), un sitio de unión a cebador (PBS) de transcriptasa inversa así como regiones del promotor que son esenciales para el inicio de transcripción inversa, la secuencia codificante de interés (por ejemplo, una secuencia codificante de sn-casADN), una estructura de tallo-bucle diseñada para la terminación de la reacción de transcripción inversa y un promotor de transcripción de ARN adecuado para su uso en una célula huésped (usada para crear un molde de ARNm que comprende los componentes previos). La transcriptasa inversa expresada en células usa ARNtpro endógeno como un cebador. Después de la transcripción inversa, se libera ADN monocatenario cuando el ARNm molde se degrada mediante RNasa H endógena o la actividad RNasa H de la transcriptasa inversa (Chen, Y., y col., "Expression of ssDNA in Mammalian Cells," *BioTechniques*, 34, 167-171 (2003)). Dichos vectores de expresión pueden emplearse para la expresión de un sn-casADN de la presente invención en una célula huésped.

La presente divulgación también abarca procedimientos de terapia génica para prevenir o tratar enfermedades, trastornos y afecciones usando los sistemas de sn-casPN/proteína Cas9 descritos en el presente documento. En un caso, un procedimiento de terapia génica usa la introducción de secuencias de ácido nucleico en un organismo o células de un organismo (por ejemplo, paciente) para conseguir la expresión de sn-casPN y componentes de proteína Cas9 de la presente invención para proporcionar modificación de una función diana. Por ejemplo, pueden genomodificarse células de un organismo, *ex vivo*, mediante (i) introducción de vectores que comprenden casetes de expresión que expresan los sn-casPN y la proteína Cas9, (ii) introducción directa de sn-casPN (por ejemplo, sn-casPN: polinucleótidos de ADN, polinucleótidos de ARN, polinucleótidos híbridos de ARN/ADN, nucleobases conectadas con cadenas principales alternativas, o combinaciones de los mismos) y proteína Cas9 o (iii) introducción de combinaciones de estos componentes. Las células modificadas por ingeniería genética se proporcionan a un organismo (por ejemplo, paciente) para tratar.

Se conocen en la materia ejemplos de terapia génica y técnicas de suministro para terapia (véase, por ejemplo, Kay, M.A., "State-of-the-art gene-based therapies: the road ahead," *Nature Reviews Genetics*, 12, 316-328 (2011); Wang, D., y col., "State-of-the-art human gene therapy: part I. Gene delivery technologies," *Discov Med.*, 18(97), 67-77 (2014); Wang, D., y col., "State-of-the-art human gene therapy: part II. Gene therapy strategies and clinical applications," *Discov Med.*, 18(98), 151-61 (2014); "The Clinibook: Clinical Gene Transfer State of the Art," Odile Cohen-Haguenaer (Editor), EDP Sciences, ISBN-10: 2842541715, (2012)).

El Ejemplo 11 ilustra el uso de sn-casARN de la presente invención para modificar dianas presentes en ADN genómico humano y medir el nivel de actividad de escisión en esos sitios. Los sitios diana se seleccionan en primer lugar de ADN genómico y después se diseñan sn-casARN para dirigirse a esas secuencias seleccionadas. Después se llevan a cabo mediciones para determinar el nivel de escisión diana que ha tenido lugar. Los datos de porcentaje de escisión y datos de especificidad proporcionan criterios sobre los que basar elecciones para una diversidad de aplicaciones. Por ejemplo, en algunas situaciones la actividad del sn-casARN puede ser el factor más importante. En otras situaciones, la especificidad del sitio de escisión puede ser relativamente más importante que el porcentaje de escisión.

En algunos aspectos, los componentes de la presente invención se suministran usando sistemas de suministro de nanoescala. Los componentes para suministrar incluyen, pero sin limitación, polinucleótidos que codifican sn-casPN y/o proteína Cas9, casetes de expresión que comprenden sn-casPN y/o proteínas Cas 9, sn-casPN, proteína Cas 9 y combinaciones de los mismos. Los componentes de la invención se pueden formular como nanopartículas. Están ampliamente disponibles bibliotecas extensivas de nanopartículas, compuestas de una mezcla de diferentes tamaños, formas y materiales, y con diversas propiedades químicas y superficiales. Los ejemplos de nanopartículas particularmente útiles en biotecnología y nanomedicina incluyen: fulerenos (por ejemplo, buckyballs y tubos de carbono); cristales líquidos; liposomas; nanopartículas basadas en sílice y silicio (por ejemplo, nanopartículas de sílice mesoporosas); nanosferas; nanobarras; nanopartículas metálicas y de óxidos metálicos (por ejemplo, ácidos nucleicos esféricos, polinucleótidos densamente empaquetados que rodean un núcleo de oro); policationes; y ciclodextrinas catiónicas.

Un ejemplo de formación de nanopartículas incluye el uso de ciclodextrinas catiónicas que pueden autoensamblarse en nanopartículas para formar partículas coloidales (Draz, M. S., y col., "Nanoparticle-Mediated Systemic Delivery of siRNA for Treatment of Cancers and Viral Infections," *Theranostics*, 4(9), 872-892 (2014)). El Ejemplo 18 describe la producción de componentes de proteína Cas9 y sn1-casARN/sn2-casARN/sn3-casARN. Estos componentes de sn-casARN y proteína Cas9 se forman en complejos ribonucleoproteicos y también se preparan como partículas con un vector de SC12CDCClickpropilamina. Se han descrito vectores de SC12CDCClickpropilamina para su uso con ARNip (véase, por ejemplo, O'Mahony, A. M., y col., "Cationic and PEGylated Amphiphilic Cyclodextrins: Co-Formulation Opportunities for Neuronal Sirna Delivery," *PLOS-ONE* 8(6), e66413 (2013)). La caracterización de las partículas de sn-casARN/proteína Cas9 del vector SC12CDCClickpropylamine se describe en el Ejemplo 18.

Las ciclodextrinas catiónicas incluyen, pero sin limitación, carboxietil- β -ciclodextrina, ciclodextrinas anfífilas (por ejemplo, heptakis[2-(ω -amino-oligo(etilenglicol))-6-desoxi-6-hexadeciltio]- β -ciclodextrina y heptakis[2-(ω -amino-oligo(etilenglicol))-6-desoxi-6-dodeciltio]- β -ciclodextrina); α -ciclodextrina multiramificada catiónica (a-CD):PEG polirotaxano.

Los liposomas son otro ejemplo de formación de nanopartículas. Los sn-casPN o complejos que comprenden sn-casPN y proteínas Cas9 de la presente invención pueden inmovilizarse en liposomas. Los liposomas para su uso con los sn-casPN comprenden habitualmente un lípido catiónico. Los ejemplos de los lípidos catiónicos incluyen DODAC (cloruro de dioctadecildimetilamonio), DOTMA (N-(2,3-dioleiloxi)propil-N,N,N-trimetilamonio), DDAB (bromuro de didodecilamonio), DOTAP (1,2-dioleiloxi-3-trimetilamonio propano), DC-Chol (3-beta-N-(N',N',-dimetil-aminoetano)-carbamol colesterol), DMRIE (1,2-dimiristoiloxipropil-3-dimetilhidroxietil amonio), DOSPA (trifluoroacetato de 2,3-dioleiloxi-N-[2(espermicarboxamido)etil]-N,N-dimetil-1-propanamino), DSTAP (1,2-Diestearoil-3-Trimetilamonio Propano), DODAP (dioleoil-3-dimetilamonio-propano), DOGS (dioctadecilamidoglicilcarboxiespermina), y similares. Puede usarse un único tipo de lípido catiónico solo, o puede usarse una combinación de dos o más tipos de lípidos catiónicos. Los lípidos catiónicos típicamente se combinan con otros lípidos (por ejemplo, fosfolípidos y colesterol) para formar liposomas.

Los ejemplos de fosfolípidos para formación de liposomas incluyen, pero sin limitación, los siguientes: fosfatidilcolina; L- α -fosfatidilcolina (fosfatidilcolina de huevo (EPC), o fosfatidilcolina de soja hidrogenada (HSPC)); 1,2-diestearoil-sn-glicero-3-fosfocolina (DSPC); 1,2-dipalmitoil-sn-glicero-3-fosfocolina (DPPC); fosfatidilserina (PS); fosfatidilinositol (PI); fosfatidilglicerol (PG); fosfatidiletanolamina (PE); dioleoil fosfatidilglicerol (DOPG); 1,2-Dioleoil-sn-glicero-3-fosfocolina (o dioleoil fosfatidilcolina) (DOPC); dioleoil fosfatidilserina (DOPS); 1,2-dioleoil-sn-glicero-3-fosfoetanolamina (DOPE); 1,2-Dioleoil-sn-glicero-3-fosfato (DOPA); 1-Miristoil-2-estearoil-sn-glicero-3-fosfocolina (MSPC); 1,2-Dipalmitoil-sn-glicero-3-fosfo-rac-(1-glicerol) (DPPG); 1,2-Dimiristoil-sn-glicero-3-fosfo-rac-(1-glicerol) (DMPG); 1,2-Dimiristoil-sn-glicero-3-fosfocolina (DMPC); 1,2-Diestearoil-sn-glicero-3-fosfo-rac-(1-glicerol) (DSPG); 1-Palmitoil-2-oleoil-sn-glicero-3-fosfocolina (POPC); diacilfosfatidilcolina; ácido diacilfosfatídico; N-dodecanoil fosfatidiletanolamina; N-succinil fosfatidiletanolamina; N-glutaril fosfatidiletanolamina; lisilfosfatidilglicerol; esfingolípidos (por ejemplo, esfingomielina); y mezclas de los mismos.

Puede usarse una diversidad de esteroides y derivados de los mismos (por ejemplo, colesterol) para estabilizar liposomas. El colesterol puede modificarse químicamente con un ligando diseñado para ser reconocido por un órgano o tipo celular particular tal como un ácido graso de cadena larga, un aminoácido, un derivado de aminoácido, una proteína, glucoproteína, un oligosacárido, una hormona, proteína modificada o similares. Los liposomas que contienen dichos esteroides modificados son adecuados para ser dirigidos a un órgano o tipo celular específico (véase, por ejemplo, patente de los Estados Unidos n.º 4.544.545).

Pueden usarse polímeros hidrófilos tales como polietilenglicol (PEG) y otros polímeros polietoxilados para proteger liposomas para potenciar la semivida en circulación del liposoma. Dichos polímeros hidrófilos pueden asociarse de forma no covalente con los liposomas o conjugarse o unirse de forma no covalente con un componente particular del liposoma (por ejemplo, lípidos derivatizados con PEG; tales como 1,2-distearoil-sn-glicero-3-fosfoetanolamina-N-[metoxi(polietilenglicol)-2000] (sal de amonio) (mPEG-DSPE) y PEG2000 estearilado). Los polímeros hidrófilos ejemplares adicionales incluyen, pero sin limitación, alcoholes polivinílicos, ácidos polilácticos, ácidos poliglicólicos, polivinilpirrolidonas, poliácridamidas, poligliceroles, polioxazolinas, poliaminoácidos (PAA) y mezclas de los mismos.

Los procedimientos para la preparación de las composiciones liposómicas incluyen liposomas formados por un

procedimiento de hidratación de película fina en el que la rehidratación usa una solución acuosa que comprende un sistema de sn-casPN/proteína Cas9 de la presente invención (véase, por ejemplo, Ejemplo 18).

5 El Ejemplo 18 describe complejos de sn-casARN/proteína Cas9 en vectores de suministro no víricos que comprenden moléculas catiónicas. En el ejemplo, se describe la producción de componentes de ARNm de Cas9 y sn1-casARN/sn2-casARN/sn3-casARN. Estos componentes se forman después en complejos ribonucleoproteicos así como partículas de ribonucleoproteína/SC12CDClickpropilamina. Los complejos y partículas se inmovilizan en liposomas. Estos liposomas se caracterizan usando varios criterios, incluyendo actividad *in vivo*. El ejemplo establece criterios para seleccionar composiciones liposómicas óptimas para encapsulación de complejos de sn-casARN/proteína Cas9 de la presente invención según sus ventajas y limitaciones.

10 En otras realizaciones, se forman liposomas por un procedimiento de inyección de solución de lípidos en el que se inyecta una solución lipídica en una solución acuosa que comprende componentes de los sistemas de sn-casPN/proteína Cas9 de la presente invención. Los lípidos se disuelven típicamente en un disolvente, por ejemplo, un disolvente orgánico (tal como un alcohol; por ejemplo, etanol), seguido de inyección en la solución acuosa que comprende el sistema de sn-casPN/proteína Cas9 agitando al mismo tiempo. Se forman vesículas liposómicas tras
15 inyección en la solución acuosa que inmovilizan cantidades pequeñas de solución acuosa en el compartimento o los compartimentos acuosos internos de las vesículas. Una ventaja de este procedimiento es que puede cambiarse de escala.

Los ejemplos de realizaciones que comprenden sn-casPN de la presente invención que pueden inmovilizarse en liposomas incluyen, pero sin limitación, polinucleótidos que codifican sn-casPN y/o proteína Cas9, casetes de
20 expresión que comprenden sn-casPN y/o proteínas Cas 9, sn-casPN, complejos de sn-casPN y proteína Cas9 y combinaciones de los mismos.

Aspectos de la presente invención incluyen, pero sin limitación, los siguientes: uno o más casetes de expresión que comprenden polinucleótidos que codifican sn-casPN y/o proteína Cas9; uno o más vectores, incluyendo vectores de
25 expresión, que comprenden polinucleótidos que codifican sn-casPN y/o proteína Cas9; procedimientos de fabricación de casetes de expresión que comprenden la producción de polinucleótidos que comprenden casetes de expresión que codifican sn-casPN y/o proteína Cas9; procedimientos de fabricación de vectores, incluyendo vectores de expresión, que comprenden la producción de vectores que codifican sn-casPN y/o proteína Cas9; procedimientos de introducción de uno o más casetes de expresión que comprenden introducir polinucleótidos que codifican sn-casPN y/o proteína Cas9 en una célula huésped seleccionada; procedimientos de
30 introducción de uno o más vectores, incluyendo vectores de expresión, que comprenden la introducción de vector o vectores que comprenden polinucleótidos que codifican sn-casPN y/o proteína Cas9 en una célula huésped seleccionada; células huésped que comprenden uno o más casetes de expresión que comprenden polinucleótidos que codifican sn-casPN y/o proteína Cas9 (células recombinantes); células huésped que comprenden uno o más vectores, incluyendo vectores de expresión, que comprenden polinucleótidos que codifican sn-casPN y/o proteína
35 Cas9 (células recombinantes); células huésped que comprenden uno o más polinucleótidos que codifican sn-casPN y/o proteína Cas9 (células recombinantes); células huésped que expresan los productos de uno o más polinucleótidos que codifican sn-casPN y/o proteína Cas9 (células recombinantes); procedimientos de fabricación de sn-casPN que comprenden producir sn-casPN mediante transcripción *in vitro* y/o producir proteína Cas9 mediante traducción *in vitro*; y procedimientos de fabricación de sn-casPN y/o proteína Cas9 que comprenden aislar los sn-casPN y/o proteína Cas9 de células huésped (células recombinantes) que expresan los productos de uno o más polinucleótidos que codifican sn-casPN y/o proteína Cas9.

Otro aspecto de la presente invención se refiere a procedimientos para generar organismos modificados genéticamente no humanos. En general, en estos procedimientos se introducen casetes de expresión que comprenden secuencias polinucleotídicas de los sn-casPN y proteína Cas9, así como un vector de dirección en
45 células de cigotos para introducir de forma específica de sitio una secuencia polinucleotídica seleccionada en una secuencia diana de ADN en el genoma para generar una modificación del ADN genómico. La secuencia polinucleotídica seleccionada está presente en el vector de dirección. Las modificaciones del ADN genómico típicamente incluyen, inserción de una secuencia polinucleotídica, supresión de una secuencia polinucleotídica o mutación de una secuencia polinucleotídica, por ejemplo, corrección génica, reemplazo génico, marcaje génico, inserción de transgén, alteración génica, mutación génica, mutación de secuencias reguladoras génicas, y así sucesivamente. En una realización de procedimientos para generar organismos modificados genéticamente no humanos, el organismo es un ratón. Una realización de este aspecto de la invención es la generación de ratones modificados genéticamente.

La generación de ratones transgénicos implica cinco etapas básicas (Cho, A., y col., "Generation of Transgenic Mice," Current Protocols in Cell Biology; Capítulo.Unidad-19.11 (2009)). En primer lugar, Purificación de una construcción transgénica (por ejemplo, casetes de expresión que comprenden secuencias polinucleotídicas de los sn-casPN y proteína Cas9, así como un vector de dirección). En segundo lugar, recogida de cigotos donantes. En tercer lugar, microinyección de la construcción transgénica en el cigoto de ratón. En cuarto lugar, implantación de cigotos microinyectados en ratones receptores pseudoembrazados. En quinto lugar, realización de genotipado y
60 análisis de la modificación del ADN genómico establecido en ratones fundadores.

El Ejemplo 17 describe uso de los polinucleótidos de nexo dividido (sn-casPN) de la presente invención para crear modificaciones genómicas en animales no humanos. El ejemplo describe la generación de ratones transgénicos usando dos sn-casARN (sn1-casARN y sn2-casARN; véase, por ejemplo, FIG. 3B). Se describe la producción de ARNm de Cas9 y sn1-casARN/sn2-casARN. Los ARNm se usan para inyección en embrión de una célula. El ejemplo describe la creación de ratones mutantes de doble gen así como la evaluación de efectos fuera de diana *in vivo* del sistema de sn-casARN/proteína Cas9. Asimismo, el ejemplo incluye la evaluación de reparación génica *in vivo* usando un oligonucleótido donante con sn-casARN y proteínas Cas9. Los resultados de estos análisis son para demostrar que pueden generarse ratones con modificaciones de reparación genómica en múltiples genes usando complejos de los sn-casPN y proteínas Cas9 descritos en el presente documento.

En otra realización de procedimientos para generar organismos modificados genéticamente no humanos, el organismo es una planta. Los sistemas de sn-casPN/proteína Cas9 descritos en el presente documento se usan para efectuar edición y manipulación de genes rentable, eficaz, en células vegetales. Es en general de preferencia insertar un ADN recombinante funcional en un genoma vegetal en una localización inespecífica. Sin embargo, en determinados casos, puede ser útil usar integración específica de sitio para introducir una construcción de ADN recombinante en el genoma. Dicha introducción de ADN recombinante en plantas se facilita usando los sistemas de sn-casPN/proteína Cas9 de la presente invención. Se conocen en la técnica vectores recombinantes para su uso en planta. Los vectores pueden incluir, por ejemplo, regiones de unión a armazón (SAR), orígenes de replicación y/o marcadores seleccionables.

Para realizaciones en las que se usan polinucleótidos que codifican sn-casPN y/o proteína Cas9 para transformar una planta, se selecciona un promotor que demuestra la capacidad de conducir la expresión de la secuencia codificante en esa especie particular de planta. Son bien conocidos en la técnica promotores que pueden usarse eficazmente en diferentes especies vegetales, también. Pueden usarse promotores inducibles, víricos, sintéticos o constitutivos en plantas para expresión de polipéptidos. También pueden ser útiles promotores que están regulados espacialmente, regulados temporalmente y regulados espaciotemporales. Una lista de promotores preferidos incluye, pero sin limitación, el promotor FMV35S, los promotores potenciados CaMV35S, promotor CaMV 35S, promotores de opina, promotores de monocotiledóneas, promotor de ubiquitina vegetal (Ubi), promotor de actina de arroz 1 (Act-1), promotor de alcohol deshidrogenasa de maíz 1 (Adh-1).

Los factores que determinan qué secuencias reguladoras usar en una construcción recombinante, incluyen, pero sin limitación, el nivel de expresión deseado y la expresión preferente en células o tejidos, capacidad de inducción, eficacia y capacidad de selección. Un experto en la materia puede modular la expresión de una secuencia codificante seleccionando y situando secuencias reguladoras en relación con la secuencia codificante.

Las secuencias reguladoras adecuadas inician principalmente la transcripción o solamente la transcripción en determinados tipos celulares. Se conocen procedimientos para identificar y caracterizar secuencias reguladoras en ADN genómico vegetal. La publicación de solicitud de patente de los Estados Unidos N.º 2011-0177228, publicada el 21 de julio de 2011, describe un gran número de dichas secuencias reguladoras.

Los promotores activos en raíces confieren transcripción en tejido de raíz, por ejemplo, tejidos vasculares de raíz, epidermis de raíz o endodermis de raíz. Algunos promotores activos en raíz son promotores preferentes en raíz y confieren transcripción predominantemente en tejido de raíz. Los ejemplos de promotores preferentes en raíz incluyen, pero sin limitación, los siguientes: PT0625, PT0660, PT0683, PT0758, YP0128 e YP0275. Otros promotores preferentes en raíz incluyen los PT0613, PT0672, PT0688 y PT0837, que promueven la transcripción principalmente en tejido de raíz pero también en alguna medida en óvulos y/o semillas. Otros promotores preferentes en raíz incluyen los subdominios específicos de raíz del promotor CaMV 35S y el promotor RD2 del tabaco.

En algunas realizaciones, pueden usarse promotores específicamente activos en endospermo en maduración. La transcripción de un promotor de endospermo de maduración comienza en general después de fertilización y se produce principalmente en tejido de endospermo durante el desarrollo de semillas. La transcripción es habitualmente mayor durante la fase de celularización. Los ejemplos de promotores de endospermo en maduración que pueden usarse en construcciones de vectores de expresión incluyen, pero sin limitación, el promotor de napina, el promotor de inhibidor de tripsina de soja, la subunidad a' de soja del promotor de beta-conglicina, el promotor de Arcelina-5, el promotor de ACP, el promotor de faseolina, el promotor de esteroil-ACP desaturasa, el promotor de oleosina, los promotores de zeína (por ejemplo, promotores de zeína de 15 kD, 16 kD, 19 kD, 22 kD y 27 kD), el promotor de Osgt-1 del gen de glutelina-1 de arroz, el promotor de beta-amilasa y el promotor de hordeína de cebada. Otros promotores de endospermo en maduración incluyen los promotores PT0676, PT0708 e YP0092.

Los ejemplos de promotores activos en tejidos ováricos incluyen, pero sin limitación, los siguientes: el promotor de poligalacturonidasa, el promotor de TRX de plátano, el promotor de actina de melón, YP0396 y PT0623. Además, los ejemplos de promotores que están activos principalmente en óvulos incluyen YP0007, YP0008, YP0028, YP0039, YP0092, YP0103, YP0111, YP0115, YP0119, YP0120, YP0121 e YP0374.

Para conseguir expresión en saco embrionario/endospermo temprano, se usan secuencias reguladoras que están activas en núcleos polares y/o la célula central o en precursores de núcleos polares, pero no en óvulos o precursores de óvulos. También puede encontrarse un patrón de transcripción que se extiende desde núcleos polares a

- desarrollo de endospermo temprano con promotores preferentes en saco embrionario/endospermo temprano (aunque la transcripción se reduce típicamente de forma significativa en desarrollo de endospermo posterior durante y después de la fase de celularización). La expresión en el cigoto o embrión en desarrollo típicamente no está presente con promotores de saco embrionario/endospermo temprano. Los ejemplos de dichos promotores incluyen los procedentes de los siguientes genes, pero sin limitación, los siguientes: *Arabidopsis viviparous-1*, *atmycl* de *Arabidopsis*, FIE de *Arabidopsis*, MEA de *Arabidopsis*, FIS2 de *Arabidopsis*, FIE 1.1, MAC1 de maíz y Cat3 de maíz. Los promotores de *Arabidopsis* adicionales incluyen YP0039, YP0101, YP0102, YP0110, YP0117, YP0119, YP0137, DME, YP0285 e YP0212. Los ejemplos de promotores de arroz incluyen p530c10, pOsFIE2-2, pOsMEA, pOsYp102 y pOsYp285.
- 5 Las secuencias reguladoras que conducen preferentemente la transcripción en células cigóticas después de fertilización pueden proporcionar expresión preferente en embriones. Los ejemplos de promotores preferentes en embriones incluyen, pero sin limitación, los siguientes: el promotor de proteína de transferencia de lípidos de cebada (Ltpl), YP0088, YP0097, YP0107, YP0143, YP0156, PT0650, PT0695, PT0723, PT0740, PT0838 y PT0879.
- 10 Los promotores activos en tejido fotosintético confieren transcripción en tejidos verdes tales como hojas y tallos. Los ejemplos de promotores de tejidos fotosintéticos incluyen los promotores de ribulosa-1,5-bifosfato carboxilasa (RbcS) tales como el promotor de RbcS de alerce americano oriental (*Larix laricina*), el promotor de cab6 de pino, el promotor de Cab-1 de trigo, el promotor de CAB-1 de espinaca, el promotor de cab1R de arroz, el promotor de piruvato ortofosfato diquinasa (PPDK) de maíz, el promotor de Lhcb1*2 de tabaco, el promotor de simportador de SUC2 sacarosa-H⁺ de *Arabidopsis thaliana*, los promotores de proteína de membrana de tilacoide de espinaca (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab y rbcS) y los promotores PT0668, PT0886, YP0144, YP0380 y PT0585.
- 15 Los ejemplos de promotores que tienen actividad alta o preferente en haces vasculares incluyen YP0022, YP0080, YP0087, YP0093 y YP0108. Otros promotores preferentes en tejido vascular incluyen el promotor de GRP 1.8 de proteína de pared celular rica en glicina, el promotor de virus del moteado amarillo de la *Commelina* (CoYMV) y el promotor de virus baciliforme de tungro del arroz (RTBV).
- 20 Los promotores inducibles confieren transcripción en respuesta a estímulos externos tales como agentes químicos o estímulos ambientales. Por ejemplo, los promotores inducibles pueden conferir transcripción en respuesta a hormonas tales como ácido giberélico o etileno, o en respuesta a luz o sequía. Los ejemplos de promotores inducibles por sequía incluyen PD0901, PD1367, PT0710, PT0848, YP0286, YP0337, YP0356, YP0374, YP0377, YP0380, YP0381, YP0384, YP0385, YP0388, YP0396, PT0633 y PT0688. Los ejemplos de promotores inducibles por nitrógeno incluyen PT0863, PT0829, PT0665 y PT0886. Los ejemplos de promotores inducibles por sombra incluyen PRO924 y PT0678. Un ejemplo de un promotor inducido por sal es rd29A.
- 25 Un promotor de tallo puede ser específico de uno o más tejidos de tallo o específico de tallo y otras partes de plantas. Los promotores de tallo pueden tener actividad alta o preferente en, por ejemplo, epidermis y corteza, cámbium vascular, procámbium o xilema. Los ejemplos de promotores de tallo incluyen los siguientes: YP0018, CryIA(b) y CryIA(c).
- 30 Los ejemplos de otras clases de promotores incluyen promotores preferentes en brote, preferentes en callo, preferentes en células del tricoma, preferentes en células guarda tales como PT0678, preferentes en tubérculo, preferentes en células del parénquima y preferentes de senescencia. En algunas realizaciones, un promotor puede conducir preferentemente la expresión en tejidos reproductivos.
- 35 Una región no traducida (UTR) 5' puede incluirse en construcciones de vectores. Una UTR 5' se transcribe, pero no se traduce, y queda entre el sitio de inicio del transcrito y el codón de inicio de la traducción y puede incluir el nucleótido +1. Una UTR 3' puede situarse entre el codón de terminación de la traducción y el final del transcrito. Los UTR pueden tener funciones particulares tales como aumento de la estabilidad de ARNm o atenuación de la traducción. Los ejemplos de UTR 3' incluyen, pero sin limitación, los siguientes: las señales de poliadenilación y secuencias de terminación de la transcripción, (por ejemplo, una secuencia de terminación de nopalina sintasa).
- 40 Se describen secuencias reguladoras adicionales en la publicación de solicitud de patente de los Estados Unidos N.º 2011-0177228, publicada el 21 de julio de 2011.
- 45 Los promotores de ARN polimerasa III que puede usarse en vectores vegetales para la expresión de sn-casPN incluyen 7SL, U6 (por ejemplo, promotor de ARNpno de U6) y U3 (por ejemplo, promotor de ARNpno de U3).
- 50 En cualquier experimento de transformación, se introduce ADN solamente en un porcentaje pequeño de células diana. Los genes que codifican marcadores seleccionables son útiles y eficaces en la identificación de células que se transforman de forma estable cuando reciben e integran una construcción de ADN transgénico en sus genomas. Los genes marcadores preferidos proporcionan marcadores selectivos que confieren resistencia a un agente selectivo, tales como un antibiótico o herbicida. Cualquier herbicida al que pueden ser resistentes las plantas es un agente útil para un marcador selectivo.
- 55 Pueden usarse marcadores seleccionables para seleccionar plantas o células vegetales que contienen vectores que

comprenden los sn-casPN y/o proteína Cas9 de la presente invención. Un marcador seleccionable puede proporcionar un fenotipo seleccionable en una célula vegetal. Por ejemplo, un marcador puede proporcionar resistencia a un antibiótico (por ejemplo, kanamicina, G418, bleomicina o higromicina), a un herbicida (un gen de bar que codifica resistencia a bialafos; un gen de EPSP sintasa mutante que codifica resistencia a glifosato; un gen de nitrilasa que confiere resistencia a bromoxinilo; un gen de acetolactato sintasa (ALS) mutante que confiere resistencia a imidazolinona o sulfonilurea) o metotrexato (un gen de DHFR resistente a metotrexato). Los vectores de expresión también pueden incluir una secuencia marcadora diseñada para promover la detección o manipulación del polipéptido expresado. Las secuencias marcadoras pueden expresarse habitualmente como una fusión con el polipéptido codificado. Los ejemplos de secuencias marcadoras incluyen, pero sin limitación, los siguientes:

luciferasa, beta-glucuronidasa (GUS), proteína verde fluorescente (GFP), glutatión S-transferasa (GST), polihistidina, c-myc, hemaglutinina o epítipo (por ejemplo, un epítipo FLAG®, Sigma-Aldrich, St. Louis, MO). Dichos marcadores pueden insertarse en cualquier parte dentro del polipéptido, incluyendo en el extremo carboxilo o amino.

Se exponen células potencialmente transformadas al agente selectivo y, entre las células supervivientes habrá células en las que el gen que confiere resistencia se ha integrado y se expresa en niveles suficientes para la supervivencia celular. Las células pueden ensayarse adicionalmente para confirmar la integración estable del ADN exógeno.

Un marcador explorable, que puede usarse para supervisar la expresión, también puede incluirse en un vector recombinante o construcción de la presente invención. Los marcadores explorables incluyen, pero sin limitación, un gen de beta-glucuronidasa o uidA (GUS) que codifica una enzima para la que se conocen diversos sustratos cromogénicos; un gen de locus R, que codifica un producto que regula la producción de pigmentos de antocianina (color rojo) en tejidos vegetales; un gen de beta-lactamasa, un gen que codifica una enzima para la que se conocen diversos sustratos cromogénicos (por ejemplo, PADAC, una cefalosporina cromogénica); un gen de luciferasa; un gen de xylE que codifica una catecol dioxigenasa que puede convertir catecoles cromogénicos; un gen de β -amilasa; una tirosinasa que codifica una enzima capaz de oxidar tirosina a DOPA y dopaquinona, que a su vez se condensa en melanina; y una β -galactosidasa, que cataliza un sustrato de β -galactosa cromogénico.

Los polinucleótidos de la presente invención puede introducirse en una célula vegetal, bien de forma permanente o bien de forma transitoria, junto con otros elementos genéticos, por ejemplo, promotores, potenciadores, intrones y secuencias líder no traducidas.

Xing, H. L., y col., "A CRISPR/Cas9 toolkit for multiplex genome editing in plants," *BMC Plant Biology*, 14, 327 (2014) han descrito vectores modulares para la expresión de sistemas de CRISPR-Cas9 en dicotiledóneas y monocotiledóneas. Se utilizan vectores binarios con dos tipos de cadenas principales; el primero basado en vectores pGreen (Hellens, R.P., y col., "pGreen: a versatile and flexible binary Ti vector for agrobacterium-mediated plant transformation," *Plant Mol. Biol.*, 42, 819-832 (2000)); y el segundo basado en vectores pCAMBIA. Los vectores de tipo pGreen son relativamente pequeños, lo que permite usarlos para expresión transitoria de Cas9 y sn-casARN en protoplastos para ensayar la eficacia de sitios diana. Los vectores pueden usarse directamente para generar plantas transgénicas después de validación en protoplastos. En *Agrobacterium*, los vectores de tipo pGreen dependen de su origen de pSa para su propagación, y requieren un plásmido auxiliar para proporcionar proteína de replicación (RepA). Pueden usarse *Agrobacterium* que contienen plásmido auxiliar pSoup como hospedadores para vectores de tipo pGreen.

Entre los vectores binarios procedentes de pCAMBIA, los que tenían un gen de resistencia a higromicina como un marcador seleccionable se obtuvieron de pCAMBIA1300, mientras que los que tenían un gen de resistencia a kanamicina se obtuvieron de pCAMBIA2300 y los que tenían un gen de resistencia a Basta se obtuvieron de pCAMBIA3300. Los vectores pCAMBIA1300/2300/3300 (Curtis, M.D., y col., "A gateway cloning vector set for high-throughput functional analysis of genes in plants," *Plant Physiol.*, 133, 462-469 (2003); Lee, L. Y., Gelvin, S. B., "T-DNA binary vectors and systems," *Plant Physiol.*, 146, 325-332 (2008)) y sus derivados (incluyendo la serie pMDC compatible con Gateway) son algunos de los vectores binarios más ampliamente usados para una diversidad de especies vegetales y con varios protocolos de transformación de plantas optimizados específicamente basándose en estos vectores.

Dichos sistemas de vectores binarios pueden usarse con los casetes de expresión de la presente invención para proporcionar, por ejemplo, múltiples sn-casARN para edición de genoma múltiple. Por ejemplo, en un sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II de tres polinucleótidos, se colocan secuencias codificantes de ADN de sn1-casARN y sn2-casARN bajo el control de promotores de ARN polimerasa III en el primer vector. Se colocan múltiples sn3-casARN que comprenden cada uno una secuencia de dirección a ADN cada uno bajo el control de promotores de ARN polimerasa III y se clonan en el mismo vector. También se incluye en el vector una secuencia codificante de proteína Cas9 optimizada para expresión en la planta seleccionada.

Entre los vectores de transformación de plantas preferidos están los procedentes de un plásmido Ti de *Agrobacterium tumefaciens* (Lee, L.Y., y col., "T-DNA Binary Vectors and Systems," *Plant Physiol.*, 146(2), 325-332 (2008)). También son útiles y conocidos en la técnica plásmidos de *Agrobacterium rhizogenes*. Existen varios productos de software comerciales diseñados para facilitar la selección de plásmidos vegetales apropiados para facilitar la selección de plásmidos vegetales apropiados para transformación celular y expresión génica en plantas y

procedimientos para permitir fácilmente la clonación de dichos polinucleótidos. SnapGene™ (GSL Biotech LLC, Chicago, Ill.; www.snapgene.com/resources/plasmid_files/your_time_is_valuable/), por ejemplo, proporciona una lista exhaustiva de vectores vegetales que incluye secuencias de vectores individuales y mapas de vectores, así como fuentes comerciales para muchos de los vectores.

- 5 Los procedimientos y composiciones para transformar plantas introduciendo una construcción de ADN recombinante en un genoma vegetal incluye cualquiera de varios procedimientos conocidos en la técnica. Un procedimiento para construir plantas transformadas es bombardeo con microproyectiles. La transformación mediada por *Agrobacterium* es otro procedimiento para construir plantas transformadas. Como alternativa, pueden usarse otras especies distintas de *Agrobacterium* (por ejemplo, *Rhizobium*) y otras células procariotas que son capaces de infectar células vegetales e introducir secuencias de nucleótidos heterólogas en el genoma de una célula vegetal infectada. Los procedimientos de transformación adicionales incluyen electroporación, liposomas, transformación usando polen o virus, productos químicos que aumentan la captación de ADN libre o suministro de ADN libre por medio de bombardeo de microproyectiles. Las construcciones de ADN de la presente invención pueden introducirse en el genoma de un huésped vegetal usando técnicas de transformación convencionales que son bien conocidas por los expertos en la materia (véase, por ejemplo, Narusaka, Y., y col., "Methods to Transfer Foreign Genes to Plants," "Transgenic Plants - Advances and Limitations," ISBN 978-953-51-0181-9, 7 de marzo de 2012).

Habitualmente, una planta transgénica formada usando procedimientos de transformación con *Agrobacterium* contiene una secuencia de ADN recombinante sencilla insertada en un cromosoma; esto se denomina acontecimiento transgénico. Dichas plantas transgénicas son heterocigotas debido a la secuencia exógena insertada. Es posible formar una planta transgénica que es homocigota con respecto a un transgén apareando sexualmente (es decir, autofecundando) una planta transgénica segregante independiente que contiene una única secuencia génica exógena consigo misma, por ejemplo, una planta F0, para producir semillas F1. Las plantas formadas germinando semillas F1 pueden ensayarse para determinar la heterocigosidad. Los ensayos de cigosidad típicos incluyen, pero sin limitación, ensayos de polimorfismo de un único nucleótido y ensayos de amplificación térmica que distinguen entre homocigotos y heterocigotos.

Como una alternativa al uso de una construcción de ADN recombinante para la transformación directa de una planta, pueden formarse plantas transgénicas cruzando una primera planta que se han transformado con una construcción de ADN recombinante con una segunda planta que carece de la construcción. Como ejemplo, una primera línea vegetal en la que se ha introducido una construcción de ADN recombinante para supresión génica puede cruzarse con una segunda línea vegetal para introgresión del ADN recombinante en la segunda línea vegetal, formando de este modo una línea vegetal transgénica.

Los sn-casPN de la presente invención proporcionan a los obtentores de plantas una nueva herramienta para inducir mutaciones. En consecuencia, un experto en la materia puede analizar fuentes genómicas e identificar genes de interés que tienen rasgos o características deseados (por ejemplo, genes de resistencia a herbicidas) y usar los sn-casPN de la presente invención para introducir dichos genes en variedades de plantas que carecen de los genes; este resultado puede conseguirse con más precisión que usando agentes mutagénicos previos, acelerando y potenciando de este modo los programas de cultivo de plantas.

El Ejemplo 16 describe mutagénesis dirigida en *Zea mays* usando un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de sn-casARN de tipo II de tres partes (sn1-casARN, sn2-casARN y sn3-casARN) para crear modificaciones genómicas en plantas. Se dirigen tres secuencias diana genómicas de maíz diferentes para escisión. Se describen vectores que comprenden casetes de expresión de los sistemas de sn1-casARN/sn2-casARN/sn3-casARN/proteína Cas9. La generación de mutaciones en los sitios diana se usa para demostrar que los complejos de sn-casPN y proteínas Cas9 como se describe en el presente documento escinden ADN cromosómico de maíz y puede usarse para generar mutaciones genómicas.

Otro aspecto de la presente invención comprende procedimientos para modificar ADN usando sn-casPN y proteínas Cas9. En general, un procedimiento para modificar ADN implica poner en contacto un ADN diana con un complejo de sn-casPN/proteína Cas9 (un "complejo de dirección"). En algunos casos, el componente de proteína Cas9 muestra actividad nucleasa que escinde ambas cadenas de una diana de ADN bicatenario en un sitio en el ADN bicatenario que es complementario de una secuencia de unión a diana de ADN en los sn-casPN. Con proteínas Cas9 de tipo II con nucleasa activa, se produce escisión específica de sitio del ADN diana en sitios determinados por (i) complementariedad de pares de bases entre la secuencia de unión a diana de ADN en los sn-casPN y el ADN diana y (ii) un motivo adyacente a protoespaciador (MAP) presente en el ADN diana. La actividad nucleasa escinde en ADN diana para producir roturas de doble cadena. En células las roturas de doble cadena se reparan por mecanismos celulares incluyendo, pero sin limitación: unión de extremos no homólogos (UENH) y reparación dirigida por homología (RDH).

La reparación de roturas de doble cadena por UENH se produce por ligamiento directo de los extremos de la rotura entre sí. Habitualmente no se insertan nuevas secuencias polinucleotídicas en el sitio de la rotura de doble cadena; sin embargo, pueden producirse inserciones o supresiones cuando un número pequeño de nucleótidos se inserta o suprime aleatoriamente en el sitio de la rotura de doble cadena. Asimismo, dos sn-casPN diferentes que comprenden secuencias de unión a ADN que se dirigen a dos secuencias diana de ADN diferentes se usan para

proporcionar supresión de una secuencia de ADN intermedia (es decir, la secuencia de ADN entre las dos secuencias diana de ADN). Se produce supresión de la secuencia intermedia cuando la UENH vuelve a unir los extremos de las dos secuencias diana de ADN escindidas entre sí. De forma similar, puede usarse UENH para dirigir la inserción de ADN molde donante o parte del mismo usando ADN molde donante, por ejemplo, que contiene salientes compatibles. En consecuencia, una realización de la presente invención incluye procedimientos de modificación de ADN introduciendo inserciones y/o supresiones en un sitio de ADN diana.

La reparación de roturas de doble cadena mediante RDH usa un polinucleótido donante (ADN molde donante) u oligonucleótido que tiene homología con la secuencia de ADN diana escindida. El ADN u oligonucleótido molde donante se usa para reparación de la rotura de doble cadena en la secuencia de ADN diana que da como resultado la transferencia de información genética (es decir, secuencias polinucleotídicas) del ADN u oligonucleótido molde donante en el sitio de la rotura de doble cadena en el ADN. En consecuencia, puede insertarse o copiarse nueva información genética (es decir, secuencias polinucleotídicas) en un sitio de ADN diana.

Un aspecto de la presente invención se refiere a un procedimiento para modificar una secuencia de unión a diana de ácido nucleico (por ejemplo, ADN) que comprende, poner en contacto la secuencia de unión a diana de ácido nucleico (por ejemplo una secuencia diana de ADN en un polinucleótido de ADN) con un complejo que comprende sn-casPN de la presente invención (por ejemplo, un complejo de sn1-casPN/sn2-casPN/proteína Cas9 (tal como se muestra en la FIG. 3B, sn1-casPN, FIG. 3B, 326, y sn2-casPN, FIG. 3B, 302; o FIG. 3A sn1-casPN, FIG. 3A, 301, sn2-casPN, FIG. 3a, 302, y sn3-casPN, FIG. 3A, 303)), en el que el sn-casPN/proteína Cas9 forma un complejo que se une con y corta la secuencia diana de ácido nucleico (por ejemplo, una secuencias diana de ADN) dando como resultado una modificación del ácido nucleico diana (por ejemplo un polinucleótido de ADN que comprende la secuencia diana de ADN) en el que dicho procedimiento no comprende un procedimiento para el tratamiento del cuerpo humano o animal mediante cirugía o terapia. Este procedimiento puede llevarse a cabo *in vitro* o *in vivo*. El procedimiento puede, por ejemplo, usarse para modificar ADN procedente de una célula (por ejemplo, una célula eucariota) aislada de un organismo. Asimismo, en algunas realizaciones el procedimiento comprende poner en contacto una secuencia diana de ADN en ADN genómico con un molde de ADN donante en el que el ADN genómico se modifica porque comprende esa al menos una parte del molde de ADN donante integrada en la secuencia diana de ADN.

Se describen procedimientos para aproximar un polinucleótido donante al sitio de una rotura de doble cadena en un ácido nucleico diana en la solicitud publicada de patente de los Estados Unidos n.º 2014-0315985, publicada el 23 de octubre de 2014 (véase, por ejemplo, 0121, 0851-0860).

El Ejemplo 1 describe la producción de componentes de sn-casPN ejemplares de la presente invención. El Ejemplo 2 describe la producción de regiones diana de ADN bicatenario para su uso en ensayos de escisión de Cas9. El Ejemplo 3 y el Ejemplo 7 proporcionan ejemplos *in vitro* de un procedimiento para modificar ADN usando complejos de sn-casPN/proteínas Cas9 (sn1-casARN/sn2-casARN) de la presente invención. El Ejemplo 6 proporciona un ejemplo *in vitro* de un procedimiento para modificar ADN usando diferentes sn-casPN (sn1-casARN/sn2-casARN/sn3-casARN) para evaluar y comparar el porcentaje de escisión de complejos de sn1-casARN/sn2-casARN/proteína Cas9 en relación con secuencias diana de ADN bicatenario seleccionadas. Asimismo, los datos presentados en el Ejemplo 4 demuestran el uso de los complejos de sn-casPN/proteína Cas9 de la presente invención para análisis de secuenciación profunda para la detección de modificaciones diana en células eucariotas.

El Ejemplo 19 describe el uso de análisis de secuenciación profunda para la detección de modificaciones diana en células eucariotas usando un complejo de sn-casPN/proteína Cas9 de la presente invención (por ejemplo, un complejo de sn1-casPN/sn2-casPN/proteína Cas9). En el ejemplo se usaron dos polinucleótidos de nexo dividido (sn1-casARN y sn2-casARN) para modificar seis secuencias diana de ADN bicatenario seleccionadas en células humanas. Se usaron complejos de CRISPR-Cas ARNgi/Cas9 de *Streptococcus thermophilus* como controles. Los datos presentados en la Tabla 22 para actividad en la célula de un complejo nucleoproteico de sn1-casARN/sn2-casARN/proteína Cas se obtuvieron mediante análisis de datos de secuenciación profunda para detección de modificaciones diana en células eucariotas. Los datos demuestran que las construcciones de proteína Cas/sn1-casPN/sn2-casPN como se describen en el presente documento facilitan escisión específica de sitio mediada por Cas de ácidos nucleicos bicatenarios diana en células. Estos datos apoyan que los sistemas de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de la presente invención facilitan la escisión específica de sitio mediada por Cas en la célula de loci genómicos.

El Ejemplo 20 ilustra el uso de diseños de sn-casPN para evaluar y comparar el porcentaje de escisión *in vitro* de complejos de sn-casPN/proteína Cas9 en relación con secuencias diana de ácido nucleico seleccionadas. En este ejemplo, el elemento de nexo de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de sn1-casARN/sn2-casARN se dividió (es decir, se rompió la cadena principal de la secuencia de nucleótidos de elemento de nexo) en varias localizaciones diferentes. La Tabla 25 presenta los resultados de ensayos de escisión de Cas9 usando los diseños de sn1-casARN/sn2-casARN alternativos, en los que la división del nexo en cada sn1-casARN/sn2-casARN fue en una localización diferente en el elemento de nexo. Los datos demuestran que sn-casPN con roturas en diferentes localizaciones en la cadena principal del elemento de nexo facilitan la escisión específica de sitio mediada por proteína Cas9 *in vitro* de secuencias diana de ácido nucleico.

En algunos procedimientos de la presente invención, las células comprenden secuencias polinucleotídicas que codifican un sn-casPN y una proteína Cas9 que comprende dominios de nucleasa RuvC y HNH activos. La expresión de estas secuencias polinucleotídicas se coloca bajo el control de uno o más promotores inducibles. Cuando la secuencia de unión a ADN de un sn-casPN es complementaria de una diana de ADN en, por ejemplo, un promotor de un gen, tras inducir la expresión de los sn-casPN y proteína Cas9, la expresión del gen se inactiva (como resultado de la escisión de la secuencia promotora por el complejo de sn-casPN/proteína Cas9). Los polinucleótidos que codifican los sn-casPN y proteína Cas9 pueden estar integrados en el genoma celular, presentes en vectores, o combinaciones de los mismos.

En procedimientos de modificación de un ADN diana usando los complejos de sn-casPN/proteína Cas9 de la presente invención, la reparación de una rotura de doble cadena por UENH y/o RDH puede conducir a, por ejemplo, corrección génica, reemplazo génico, marcaje génico, alteración génica, mutación génica, inserción de transgén o supresión de nucleótidos. Pueden usarse procedimientos para modificar un ADN diana usando los complejos de sn-casPN/proteína Cas9 de la presente invención en combinación con un ADN molde donante para insertar o reemplazar secuencias polinucleotídicas en una secuencia diana de ADN, por ejemplo, para introducir un polinucleótido que codifica una proteína o ARN funcional (por ejemplo, ARNip), para introducir un marcador proteico, para modificar una secuencia reguladora de un gen o para introducir una secuencia reguladora para un gen (por ejemplo, un promotor, un potenciador, una secuencia de entrada de ribosoma interna, un codón de inicio, un codón de terminación, una señal de localización o señal de poliadenilación), para modificar una secuencia de ácido nucleico (por ejemplo, introducir una mutación) y similares.

En algunas realizaciones de los sistemas de sn-casPN/proteína Cas9 de la presente invención, se usa una forma mutada de la proteína Cas9. Las versiones modificadas de la proteína Cas9 pueden contener un único dominio catalítico inactivo (es decir, RuvC inactivo o HNH inactivo). Dichas proteínas Cas9 modificadas escinden solamente una cadena de un ADN diana creando de este modo una rotura de una única cadena. La proteína Cas9 modificada que tiene un único dominio catalítico inactivo puede unirse con ADN basándose en la especificidad conferida por sn-casPN; sin embargo, solamente cortará una de las cadenas de ADN bicatenarias. Como ejemplo, en la proteína Cas9 de *Streptococcus pyogenes* el dominio de RuvC puede inactivarse mediante una mutación D10A y el dominio de HNH puede inactivarse por una mutación H840A. Cuando se usa una proteína Cas9 modificada que tiene un único dominio catalítico inactivo en los complejos de sn-casPN/proteína Cas9 de la presente invención es menos probable que se produzca UENH en el sitio de rotura de una única cadena.

En otras versiones modificadas de la proteína Cas9 ambos dominios catalíticos son inactivos (es decir, RuvC inactivo y HNH inactivo; "dCas"). Dichas proteínas dCas9 no tienen ninguna actividad nucleasa sustancial; sin embargo, se pueden unir con ADN basándose en especificidad conferida por sn-casPN. Como ejemplo, en la proteína Cas9 de *Streptococcus pyogenes* una mutación D10A y una mutación H840A dan como resultado una proteína dCas 9 que no tiene ninguna actividad nucleasa sustancial.

La presente divulgación también incluye procedimientos de modulación de transcripción *in vitro* o *in vivo* usando complejos de sn-casPN/proteína Cas9 descritas en el presente documento. En un caso, un complejo de sn-casPN/proteína Cas9 puede reprimir expresión génica interfiriendo con la transcripción cuando un sn-casPN dirige la unión a diana de ADN del complejo de sn-casPN/proteína Cas9 con la región promotora de un gen. El uso de complejos de sn-casPN/proteína Cas9 para reducir la transcripción también incluye complejos en los que la proteína dCas9 se fusiona con un regulador negativo conocido de un gen diana (por ejemplo, un polipéptido represor). Por ejemplo, la expresión de un gen está bajo el control de secuencias reguladoras con las que puede unirse un polipéptido represor. Un sn-casPN puede dirigir la unión a diana de DNA de un complejo de sn-casPN/proteína Cas9-proteína represora con las secuencias de ADN que codifican las secuencias reguladoras o adyacentes a las secuencias reguladoras de modo que la unión del complejo de sn-casPN/proteína Cas9-proteína represora pone la proteína represora en contacto operativo con las secuencias reguladoras. Esto da como resultado la represión de expresión del gen diana. De forma similar, dCas9 se fusiona con un polipéptido activador para activar o aumentar la expresión de un gen bajo el control de secuencias reguladoras con las que puede unirse un polipéptido activador.

En un aspecto la presente divulgación se refiere a un procedimiento para modular la expresión de un gen que comprende elementos reguladores de la transcripción que comprenden, poner en contacto una secuencia diana de ADN en el gen con complejos de sn-casPN/proteína Cas9 de la presente invención, en los que los sn-casPN y la proteína Cas9 forman un complejo que se une con la secuencia diana de ADN dando como resultado modulación de la expresión del gen. En un caso, la proteína Cas9 es una Cas9 que es deficiente en nucleasa. En otros casos, el complejo de sn-casPN/Cas9 comprende además una proteína de fusión.

El Ejemplo 13 describe el uso de los sn-casPN de la presente invención para la represión o activación de genes endógenos en células humanas. Se usa la Cas9 de *S. pyogenes* deficiente en nucleasa (dCas9) con mutación D10A y H840A. La secuencia de sn1-casARN-CD71 comprende una secuencia espaciadora de 20 nucleótidos que dirige el complejo de sn-casARN/proteína Cas9 a la región no traducida cadena arriba del receptor de transferrina CD71. La activación de expresión de CD71 en muestras transfectadas con dCas9-VP64 se mide por el aumento de fluorescencia detectada en comparación con la fluorescencia medida de una población de control no transfectada de células HeLa detectada por separación FACS. La represión de la expresión de CD71 en muestras transfectadas con dCas9-KRAB se mide por la reducción de fluorescencia detectada en comparación con la fluorescencia medida de

una población de control no transfectada de células HeLa detectada por separación FACS. Este procedimiento proporciona datos para verificar que los sistemas de sn-casPN/proteína Cas9 de la presente invención pueden usarse en la activación o represión de genes endógenos.

5 En algunas realizaciones, una secuencia no nativa puede conferir nuevas funciones a una proteína de fusión. Los ejemplos de proteínas de fusión que incluyen una proteína Cas9 y otros dominios reguladores o funcionales incluyen, pero sin limitación, una nucleasa, una transposasa, una metilasa, un dominio represor o activador de factor de transcripción (por ejemplo, tal como KRAB y VP16), dominios correpresores y coactivadores, ADN metil transferasas, histona acetiltransferasas, histona desacetilasas y dominios de escisión de ADN (por ejemplo, un dominio de escisión de la endonucleasa FokI). Los ejemplos adicionales incluyen, pero sin limitación, los siguientes:

10 actividad metiltransferasa, actividad desmetilasa, actividad de desaminación, actividad de dismutasa, actividad de alquilación, actividad de despurinación, actividad de oxidación, actividad de formación de dímeros de pirimidina, actividad integrasa, actividad transposasa, actividad recombinasa, actividad polimerasa, actividad ligasa, actividad helicasa, actividad fotoliasa, actividad glucosilasa, actividad acetiltransferasa, actividad desacetilasa, actividad quinasa, actividad fosfatasa, actividad ubiquitina ligasa, actividad desubiquitinizante, actividad de adenilación,

15 actividad de desadenilación, actividad de sumoilación, actividad de desumoilación, actividad de ribosilación, actividad de desribosilación, actividad de miristoilación, actividad de remodelación, actividad proteasa, actividad oxidorreductasa, actividad transferasa, actividad hidrolasa, actividad liasa, actividad isomerasa, actividad sintasa, actividad sintetasa, actividad de desmiristoilación y cualquier combinación de las mismas.

En otro aspecto, los sn-casPN de la presente invención se usan en procedimientos para exploración genómica funcional de alto rendimiento. Las exploraciones genéticas directas son herramientas potentes para el descubrimiento y la anotación funcional de elementos genéticos (véase, por ejemplo, Gilbert, y col., "CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes," Cell 18, 154(2), 442-51 (2013)). Los complejos de sn-casPN/proteína Cas9 pueden usarse para generar bibliotecas de escala genómica de sn-casPN para exploración fenotípica, sin sesgos. Los enfoques para exploración a escala genómica incluyen enfoques de inactivación que inactivan loci genómicos y enfoques que modulan la actividad transcripcional. En exploración de inactivación, se generan mutaciones de pérdida de función mediadas por complejos de sn-casPN/proteína Cas9 mediante inducción de rotura de doble cadena y reparación mediada por UENH. Las exploraciones de inactivación son útiles para identificar funciones génicas esenciales, por ejemplo, funciones génicas relacionadas con sensibilidades a fármacos y toxinas. Se presenta un ejemplo de dicha exploración genómica funcional en el Ejemplo 12. En el ejemplo, se usa un sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II de sn-casARN de dos partes (sn1-casARN y sn2-casARN) para crear una biblioteca lentivírica de sn1-casARN. La biblioteca se usa en un procedimiento de inactivación para identificar genes candidatos importantes en la resistencia al tratamiento farmacológico. Este procedimiento proporciona datos para verificar que los complejos de sn-casPN/proteína Cas9 de la presente invención pueden usarse en la exploración funcional para consultar la función génica en una escala genómica.

20

25

30

35

Otro procedimiento de la presente invención es el uso de sistemas de sn-casPN/proteína dCas para aislar o purificar regiones de ADN genómico (ADNg). En una realización del procedimiento, una proteína dCas9 se fusiona con un epítipo (por ejemplo, un epítipo FLAG®, Sigma-Aldrich, St. Louis, MO) y un sn-casPN dirige la unión a diana de ADN de un complejo de sn-casPN/proteína dCas9-epítipo con secuencias de ADN en la región de ADN genómico para aislar o purificar.

40

Se usa un agente de afinidad para unir el epítipo y el ADNg asociado unido al complejo de sn-casPN/proteína dCas9-epítipo.

En un aspecto adicional, la presente invención incluye kits que comprenden sn-casPN o polinucleótidos que codifican sn-casPN e instrucciones. Los kits pueden comprender uno o más de los siguientes: sn-casPN y proteína Cas9 afín; polinucleótidos que codifican sn-casPN y/o proteína afín; células recombinantes que comprenden sn-casPN; células recombinantes que comprenden sn-casPN y proteína afín; y similares. Cualquier kit de la presente invención puede comprender además otros componentes tales como soluciones, tampones, sustratos, células, instrucciones, vectores (por ejemplo, vectores de dirección) y así sucesivamente.

45

También se desvela el uso de ensayos de T7E1 para evaluar y comparar el porcentaje de escisión *in vivo* de complejos de sn-casPN/proteína Cas9 en relación con secuencias diana de ADN bicatenario seleccionadas (Ejemplo 9).

50

También se desvelan procedimientos para identificación y exploración de ARN de CRISPR de activación en trans (Ejemplo 8), que puede modificarse para su uso en los sn-casPN y procedimientos de la presente invención. Asimismo, se desvelan procedimientos de la divulgación para generar y ensayar modificaciones de nexos dividido en ARNcrtra (Ejemplo 10), por ejemplo, basados en ARNcr/ARNcrtra conocidos en la técnica o identificados por procedimientos descritos en el Ejemplo 8.

55

La presente invención también incluye composiciones farmacéuticas que comprenden complejos de sn-casPN/proteína Cas9 o uno o más polinucleótidos que codifican sn-casPN y una proteína Cas9. Composiciones farmacéuticas, por ejemplo, las composiciones de nanopartículas que comprenden sn-casPN descritas

anteriormente, pueden comprender además excipientes farmacéuticamente aceptables.

Una composición farmacéutica puede comprender una combinación de cualquiera de los sn-casPN descritos en el presente documento con otros componentes, por ejemplo, excipientes (por ejemplo, vehículos, estabilizadores, diluyentes, agentes de suspensión, agentes espesantes y otros como se describe en el presente documento). Las composiciones facilitan la administración de, por ejemplo, complejos de sn-casPN/proteína Cas9 a un sujeto. Pueden administrarse composiciones farmacéuticas en cantidades terapéuticamente eficaces por diversas formas y vías incluyendo, por ejemplo, intravenosa, subcutánea o inhalación.

Los procedimientos para la preparación de composiciones farmacéuticas que comprenden los sn-casPN pueden incluir formularlas con uno o más excipientes farmacéuticamente aceptables, inertes. Por ejemplo, las composiciones farmacéuticas pueden ser soluciones o suspensiones líquidas. Los excipientes útiles habituales en la práctica de la presente invención incluyen, pero sin limitación, los siguientes: transportador o vehículo (por ejemplo, agua o soluciones acuosas tamponadas); sistemas de tampones (por ejemplo, que comprenden acetato, fosfato, citrato, borato, tartrato, histidina, succinato y mezclas de los mismos); antioxidantes (por ejemplo, tiosulfato sódico, ácido etilendiamintetraacético, ácido cítrico, cisteínas, tioglicerol, ácido tioglicólico, tiosorbitol, hidroxianisol butilado, hidroxitolueno butilado y galato de propilo, y mezclas de los mismos); agentes para mantener la isotonicidad (por ejemplo, cloruro sódico, azúcares, polioles (alcoholes de azúcares), ácido bórico, tartrato sódico, propilenglicol y mezclas de los mismos); uno o más azúcares (por ejemplo, trehalosa, maltosa, sacarosa, lactosa, manosa, dextrosa, fructosa, etc.) o alcohol de azúcar (por ejemplo, sorbitol, maltitol, lactitol, manitol, glicerol, etc.), alcohol (por ejemplo, etanol, t-butanol, etc.); y conservantes (alcoholes, ácido benzoico, ácido salicílico, fenol y sus derivados (por ejemplo, cresol, p-cresol, m-cresol y o-cresol), cetrimida, BHA (hidroxitolueno butilado), BHA (hidroxianisol butilado); y mezclas de los mismos).

Las ventajas de los sn-casPN de la presente invención incluyen, pero sin limitación, las siguientes. El uso de sistemas de sn-casPN/proteína Cas9 de múltiples partes permite mejor control de la actividad para sistemas *in vivo*. El control de la expresión de todas las partes de un sistema de sn-casPN/proteína Cas9 proporciona capas adicionales de regulación sobre el ensamblaje de los componentes específicos necesarios para constituir un sistema de sn-casPN/proteína Cas9 funcional, por ejemplo, en relación con un sistema de ARNgi/proteína Cas9.

El elemento de nexo dividido, polinucleótidos adyuvantes, auxiliares y adjuntos de los sn-casPN de la presente invención proporcionan sitios adicionales (en relación con complejos de ARNcr/ARNcrtra/Cas9 y ARNgi/Cas9) para añadir y/o unir restos funcionales (por ejemplo, polipéptidos, moléculas pequeñas, marcadores y similares).

En algunas realizaciones de la presente invención, por ejemplo, un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificado por ingeniería genética de tres polinucleótidos, la menor longitud de los sn-casPN (en relación con las mayores longitudes de ARNcr/ARNcrtra y ARNgi) permite síntesis química de mayor calidad y más rápida de los sn-casPN. Asimismo, la menor longitud de los sn-casPN facilita el empaquetamiento en vectores basados en virus.

Asimismo, puede usarse un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificado por ingeniería genética de tres polinucleótidos de la presente invención (por ejemplo, como se ilustra en la FIG. 3A) para proporcionar complejos de Cas9 parcialmente preformados en un sistema *in vivo* para permitir la activación rápida. Por ejemplo, sn1-casARN, sn3-casARN y proteína Cas9 se expresan en una célula. Estos componentes forman un complejo de sn1ARN/sn3-casARN/proteína Cas9, que no está activo para unión o escisión de una diana. Cuando el componente de sn2 se expresa o introduce en la célula, el complejo de sn1-casARN/sn2-casARN/sn3-casARN/proteína Cas9 se activa rápidamente, lo que permite el control temporal sobre la dirección específica de sitio.

Ventajas adicionales de la presente invención resultarán evidentes para los expertos en la materia a la vista de las enseñanzas de la presente memoria descriptiva.

En un aspecto, la presente invención se refiere a tres polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética capaces de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con una primera secuencia de ADN que comprende una secuencia diana de ADN preferentemente en relación con una segunda secuencia de ADN sin la secuencia de unión a diana de ADN. Son necesarios al menos dos de los tres polinucleótidos para formar un elemento de tallo de nexo. En algunas realizaciones, el sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética comprende además una proteína Cas9 o una secuencia de ADN que codifica una proteína Cas9. En realizaciones adicionales, la presente invención se refiere a los tres polinucleótidos en complejo con una proteína Cas9.

En una realización, un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de la presente invención comprende un primer polinucleótido, un segundo polinucleótido y un tercer polinucleótido que son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

El primer polinucleótido (sn1-casPN) comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos

de elemento de tallo I y una secuencia de nucleótidos de elemento de tallo de nexa I. El segundo polinucleótido (sn2-casPN) comprende una secuencia de nucleótidos de elemento de tallo de nexa II, en el que la secuencia de nucleótidos de elemento de tallo de nexa I y la secuencia de nucleótidos de elemento de tallo de nexa II son capaces de formar el elemento de tallo de nexa mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexa I y la secuencia de nucleótidos de elemento de tallo de nexa II. El tercer polinucleótido (sn3-casPN) comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II.

En otra realización, un sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de la presente invención comprende un primer polinucleótido, un segundo polinucleótido y un tercer polinucleótido que son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'. El primer polinucleótido (sn1-casPN) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo inferior I y una secuencia de nucleótidos de elemento de tallo de nexa I. El segundo polinucleótido (sn2-casPN) comprende una secuencia de nucleótidos de elemento de tallo de nexa II, en el que la secuencia de nucleótidos de elemento de tallo de nexa I y la secuencia de nucleótidos de elemento de tallo de nexa II son capaces de formar el elemento de tallo de nexa mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexa I y la secuencia de nucleótidos de elemento de tallo de nexa II. El tercer polinucleótido (sn3-casPN) comprende, en una dirección de 5' a 3', una secuencia de unión a diana de ADN, una secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

En realizaciones adicionales el segundo polinucleótido (sn2-casPN) comprende primer y/o segundo polinucleótidos adjuntos. El segundo polinucleótido puede comprender además, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexa II y una segunda secuencia de nucleótidos de elemento de tallo I, y un primer polinucleótido adjunto que comprende una segunda secuencia de nucleótidos de elemento de tallo II. La segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II son capaces de formar un segundo elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II. En algunas realizaciones, el primer polinucleótido adjunto comprende además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la segunda secuencia de nucleótidos de elemento de tallo II, en el que el extremo 5' de la secuencia de nucleótidos de elemento de bucle está unido covalentemente con el extremo 3' de la segunda secuencia de nucleótidos de elemento de tallo I, formando de este modo una horquilla. En realizaciones adicionales, el primer polinucleótido adjunto comprende, en una dirección de 5' a 3', la segunda secuencia de nucleótidos de elemento de tallo II y una tercera secuencia de nucleótidos de elemento de tallo I, y un segundo polinucleótido adjunto comprende, en una dirección de 5' a 3', una tercera secuencia de nucleótidos de elemento de tallo II. La tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II son capaces de formar un tercer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II. En algunas realizaciones el segundo polinucleótido adjunto comprende además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la tercera secuencia de nucleótidos de elemento de tallo II, en el que el extremo 5' de la secuencia de nucleótidos de elemento de bucle está unido covalentemente con el extremo 3' de la tercera secuencia de nucleótidos de elemento de tallo I.

Realizaciones adicionales de la presente invención incluyen el primer polinucleótido (sn1-casPN) que comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo I, el segundo polinucleótido (sn2-casPN) que comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II, o tanto el primer polinucleótido (sn1-casPN) como el segundo polinucleótido (sn2-casPN) que comprenden cada uno una secuencia adyuvante.

En algunas realizaciones de la presente invención, el primer polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento de unión efector I y el segundo polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento de unión efector II. La secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector II son capaces de formar un elemento de unión efector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector I. El elemento de unión efector puede ser, por ejemplo, un ARN bicatenario y la proteína efectora es una proteína de unión a ARN bicatenario capaz de unirse con el

elemento de unión efector. En realizaciones seleccionadas la proteína efectora es una variante catalíticamente inactiva de una proteína seleccionada del grupo que consiste en Cas5, Cas6 y Csy4.

5 En realizaciones adicionales, el primer polinucleótido adyuvante comprende además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento de unión efector, y el segundo polinucleótido adyuvante comprende, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de unión efector II y una secuencia de nucleótidos de elemento conector II. La secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento conector II son capaces de formar un elemento conector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector I.

10 En realizaciones adicionales, el primer polinucleótido adyuvante, el segundo polinucleótido adyuvante o tanto el primer polinucleótido adyuvante como el segundo polinucleótido adyuvante comprenden cada uno una horquilla. Asimismo, el primer polinucleótido adyuvante puede comprender además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento conector I y la horquilla, el segundo polinucleótido adyuvante comprende, en una dirección de 5' a 3', la horquilla y una secuencia de nucleótidos de elemento conector II o ambos, el primer polinucleótido adyuvante comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento conector I y la horquilla y el segundo polinucleótido adyuvante comprende, en una dirección de 5' a 3', la horquilla y una secuencia de nucleótidos de elemento conector II. La secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento conector II son capaces de formar un elemento conector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector I.

20 En otro aspecto más un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de la presente invención comprende tres polinucleótidos. Un primer polinucleótido (sn1-casPN) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo inferior I y una secuencia de nucleótidos de elemento de tallo de nexo I. Un segundo polinucleótido (sn2-casPN) comprende, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo de nexo II, un segundo elemento de tallo que comprende una horquilla, y un tercer elemento de tallo que comprende una horquilla. La secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar el elemento de tallo de nexo mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II. Un tercer polinucleótido (sn3-casPN) comprende, en una dirección de 5' a 3', una secuencia de unión a diana de ADN, una secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II. La secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética puede comprender además una proteína Cas9 o una secuencia de ADN que codifica una proteína Cas9.

35 En algunas realizaciones de este aspecto de la presente invención el primer polinucleótido (sn1-casPN) comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexo I y el segundo polinucleótido (sn2-casPN) comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexo II.

Sección experimental

45 Se ilustran adicionalmente aspectos de la presente invención en los siguientes ejemplos. Se ha intentado asegurar la precisión con respecto a los números usados (por ejemplo, cantidades, concentraciones, porcentajes de cambio, etc.) pero deberían tenerse en cuenta algunos errores y desviaciones experimentales. A menos que se indique otra cosa, la temperatura está en grados centígrados y la presión es atmosférica o cercana a la atmosférica. Debería entenderse que estos ejemplos, aunque indican algunas realizaciones de la invención, se proporcionan solamente como ilustración.

55 No se pretende que los siguientes ejemplos limiten el ámbito de lo que los inventores consideran diversos aspectos de la presente invención.

Materiales y procedimientos

Se proporcionaron secuencias oligonucleotídicas (por ejemplo, las secuencias de cebadores mostradas en la FIG. 13) a proveedores comerciales para síntesis (Integrated DNA Technologies, Coralville, IA; o Eurofins, Luxemburgo).

Se ensamblan sn-casPN mediante PCR usando cebadores solapantes 3' que contienen las secuencias de ADN correspondientes para los sn-casPN. Asimismo, pueden clonarse secuencias de ADN que codifican los sn-casPN en un vector adecuado para propagación y aislamiento posterior de secuencias de sn-casPN (por ejemplo, usando escisión por enzimas de restricción del vector para producir el sn-casPN).

5 **Ejemplo 1**

Producción de componentes de sn-casARN

Este ejemplo describe la producción de un sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II de tres polinucleótidos (por ejemplo, similar al sistema ilustrado en la FIG. 3A).

10 Se produjeron componentes de ARN mediante transcripción *in vitro* (por ejemplo, kit de síntesis de ARN de alto rendimiento rápido T7, New England Biolabs, Ipswich, MA) a partir de molde de ADN bicatenario que incorpora un promotor T7 en el extremo 5' de las secuencias de ADN.

15 El molde de ADN bicatenario para el componente de sn2-casARN específico, usado en los ejemplos (denominado en el presente documento sn2-casARN^{EX}), se ensambló mediante PCR usando cebadores solapantes 3' que contienen las secuencias de ADN correspondientes para el componente de sn2-casARN^{EX}. Los oligonucleótidos utilizados en el ensamblaje se presentan en la Tabla 8.

Tabla 8

Cebadores solapantes para generación de moldes de componentes de sn2-casARN		
Tipo de componente de sn-casPN	Diana para secuencia de unión a ADN	Oligonucleótidos*
sn2-casARN ^{EX}	n/d	A, B, C
*Se muestran secuencias oligonucleotídicas de cebadores de ADN en la FIG. 13		

20 Los cebadores de ADN estaban presentes a una concentración de 2 nM cada uno. Se usaron dos cebadores de ADN externos correspondientes al promotor T7 (cebador directo: Oligonucleótido A, Tabla 8) y el extremo 3' de la secuencia de ARN (cebadores inversos: Oligonucleótidos C, Tabla 8) a 640 nM para conducir la reacción de amplificación. Se realizaron reacciones de PCR usando mezcla maestra 2X de alta fidelidad de inicio en caliente Q5 (New England Biolabs, Ipswich, MA) siguiendo las instrucciones del fabricante. Se llevaron a cabo reacciones de ensamblaje de PCR usando las siguientes condiciones de termociclación: 98 °C durante 2 minutos, 35 ciclos de 15 segundos a 98 °C, 15 segundos a 62 °C, 15 segundos a 72 °C y una extensión final a 72 °C durante 2 min. La calidad del ADN se evaluó mediante electroforesis en gel de agarosa (SYBR® Safe, 1,5 %, Life Technologies, Grand Island, NY).

25 Los moldes de ADN bicatenario para los componentes de sn1-casARN y sn3-casARN específicos, usados en los ejemplos, se ensamblaron formando complejo con dos secuencias oligonucleotídicas complementarias (denominadas sn1-casARN^{EX} y sn3-casARN^{EX}). Los oligonucleótidos utilizados en el ensamblaje se presentan en la Tabla 9.

30

Tabla 9

Cebadores solapantes para generación de moldes de componentes de sn-casARN		
Tipo de componente de sn-casPN	Diana para secuencia de unión a ADN	Oligonucleótido ²
sn1-casARN ^{EX}	n/d	D, E
sn3-casARN ^{EX} -AAVS1	AAVS-1 ¹	F, G
1 - AAVS-1, sitio de integración de virus adenoasociado 1 - genoma humano; 2 - se muestran secuencias de cebadores de ADN en la FIG. 13.		

Los cebadores de ADN estaban presentes a una concentración de 10 µM cada uno, se mezclaron 10 µM de cada cebador entre sí y se incubaron durante 2 minutos a 95 °C, se retiraron del termociclador y se permitió que se equilibraran a temperatura ambiente.

35 Se transcribieron entre 0,25-0,5 µg del molde de ADN para cada componente de sn-casARN usando kit de síntesis de ARN de alto rendimiento T7 (New England Biolabs, Ipswich, MA) durante aproximadamente 16 horas a 37 °C. Las reacciones de transcripción se trataron con DNasa I (New England Biolabs, Ipswich, MA) y se purificaron usando el kit de limpieza y concentración de ARN GeneJet (Life Technologies, Grand Island, NY). Se cuantificó el

rendimiento de ARN usando el sistema Nanodrop™ 2000 (Thermo Scientific, Wilmington, DE). Se comprobó la calidad del ARN transcrito mediante electroforesis en gel de agarosa (SYBR® Safe, 2 %, Life Technologies, Grand Island, NY). Las secuencias de sn-casARN son como se muestra en la Tabla 10.

Tabla 10

Secuencias de sn-casARN		
sn2-casARN ^{EX}	5' – GUCCGUUAUC AACUUGAAAA AGUGGCACCG AGUCGGUGCU U– 3'	SEQ ID NO: 68
sn1-casARN ^{EX}	5' – GCAGGACAGC AUAGCAAGUU GAGUAAGGC UA– 3'	SEQ ID NO: 69
sn3-casARN ^{EX} -AAVS1	5' – GGGGCCACUA GGGACAGGAU GUCUCAGAGC UAUGCUGU– 3'	SEQ ID NO: 70

- 5 Este procedimiento para la producción de sn1-casARN^{EX}, sn2-casARN^{EX}, y sn3-casARN^{EX} puede aplicarse a la producción de otros sn-casARN como se describe en el presente documento.

Ejemplo 2

Producción de regiones diana de ADN bicatenario para su uso en ensayos de escisión de Cas9

- 10 Se produjo ADN bicatenario diana para su uso en ensayos de escisión de Cas9 *in vitro* usando amplificación por PCR de la región diana de ADN genómico.

- 15 Se amplificaron regiones diana de ADN bicatenario (por ejemplo, AAVS-1) para ensayos bioquímicos mediante PCR a partir de ADN genómico (ADNg) de la línea celular humana K562 preparado con fenolcloroformo (Colección Americana de Cultivos Tipo, Manassas, VA). Se llevaron a cabo reacciones de PCR con mezcla maestra 2X de alta fidelidad de inicio en caliente Q5 (New England Biolabs, Ipswich, MA) siguiendo las instrucciones del fabricante. Se usaron 20 ng/μl de ADNg en un volumen final de 25 μl para amplificar la región diana seleccionada en las siguientes condiciones: 98 °C durante 2 minutos, 35 ciclos de 20 s a 98 °C, 20 s a 60 °C, 20 s a 72 °C y una extensión final a 72 °C durante 2 min. Los productos de PCR se purificaron usando tubos de purificación de PCR Spin Smart™ (Denville Scientific, South Plainfield, NJ) y se cuantificaron usando espectrofotómetro Nanodrop™ 2000 UV-Vis (Thermo Scientific, Wilmington, DE).

- 20 Los cebadores directo e inverso usados para amplificación de secuencias diana seleccionadas a partir de ADNg fueron los siguientes: AAVS-1, oligonucleótidos H e I (FIG. 13). La diana de ADN bicatenario amplificada para AAVS-1 fue de 495 pb.

- 25 Se obtienen otras regiones diana de ADN bicatenario adecuadas usando esencialmente el mismo procedimiento. Para regiones diana no humanas, se usa ADN genómico del organismo seleccionado (por ejemplo, planta, bacterias, levadura, algas) en lugar de ADN procedente de células humanas. Asimismo, pueden usarse fuentes de polinucleótidos distintas de ADN genómico (por ejemplo, vectores y fragmentos de ADN aislados en gel).

Ejemplo 3

Ensayos de escisión de Cas9

- 30 Este ejemplo ilustra el uso de un sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II de tres polinucleótidos de la presente invención en ensayos de escisión de Cas9 *in vitro* para evaluar y comparar el porcentaje de escisión de complejos de sn-casARN/proteína Cas9 seleccionados en relación con secuencias diana de ADN bicatenario seleccionadas.

Se determinó la escisión de secuencias diana de ADN bicatenario para componentes de sn-casARN^{EX} del Ejemplo 2 contra una diana de ADN bicatenario (AAVS-1; Ejemplo 2).

- 35 Los tres componentes de sn-casARN^{EX} en cantidades equimolares se mezclaron en un tampón de hibridación (HEPES 1,25 mM, MgCl₂ 0,625 mM, KCl 9,375 mM a pH7,5), se incubaron durante 2 minutos a 95 °C, se retiraron del termociclador y se permitió que se equilibraran a temperatura ambiente. Se ensayaron combinaciones adicionales de dos de los tres sn-casARN^{EX} como se describe posteriormente en referencia a los datos presentados en la FIG. 8. Cuando se usaron solamente dos componentes del sistema de polinucleótidos de nexos dividido
40 asociados con CRISPR-Cas9 de tipo II se añadió agua en lugar del tercer componente de sn-casARN^{EX}.

Los sn-casARN^{EX} se añadieron a una mezcla de reacción de Cas9. La mezcla de reacción de Cas9 comprendía

proteína Cas9 diluida hasta una concentración final de 200 μM en tampón de reacción (HEPES 20 mM, KCl 100 mM, MgCl_2 5 mM, DTT 1 mM y glicerol 5 % a pH 7,4). En la mezcla de reacción la concentración final de cada sn-casARN^{EX} fue de 500 nM en cada mezcla de reacción. Cada mezcla de reacción se incubó a 37 °C durante 10 minutos. La reacción de escisión se inició mediante la adición de ADN diana hasta una concentración final de 15 nM.

5 Las muestras se mezclaron y centrifugaron brevemente antes de incubarse durante 15 minutos a 37 °C. Las reacciones de escisión se terminaron mediante la adición de proteinasa K (Denville Scientific, South Plainfield, NJ) a una concentración final de solución de RNasa A 0,2 $\mu\text{g}/\mu\text{l}$ y 0,44 $\mu\text{g}/\mu\text{l}$ (SigmaAldrich, St. Louis, MO).

Las muestras se incubaron durante 25 minutos a 37 °C y 25 minutos a 55 °C. Se evaluaron 12 μl de la reacción total para determinar la actividad de escisión mediante electroforesis en gel de agarosa (SYBR® Gold, 2 %, Life Technologies, Grand Island, NY). Para la diana de ADN bicatenario de AAVS-1, la aparición de bandas de ADN a aproximadamente 316 pb y aproximadamente 179 pb indicó que se había producido escisión del ADN diana. Se usó una escalera de ADN de 100 pb como un patrón de peso molecular (New Englan Biolabs, Ipswich, MA). Se calcularon los porcentajes de escisión usando valores de área bajo la curva calculados mediante FIJI (ImageJ; un programa de procesamiento de imágenes de Java de código abierto) para cada fragmento de escisión y el ADN diana, y dividiendo la suma de los fragmentos de escisión por la suma tanto de los fragmentos de escisión como del ADN diana.

10

15

La FIG. 8 presenta los resultados del ensayo de escisión de Cas9 usando el ADN bicatenario diana de AAVS-1. En la figura, se muestran repeticiones de tres para cada combinación de sn-casARN^{EX}. En la parte superior de cada panel hay una representación gráfica de los sn-casARN^{EX} usados en el ensayo. La FIG. 8, Panel A muestra la actividad bioquímica de sn1-casARN^{EX}, sn2-casARN^{EX}, sn3-casARN^{EX}-AAVS1. La FIG. 8, Panel B muestra la actividad bioquímica de sn1-casARN^{EX} y sn2-casARN^{EX}. La FIG. 8, Panel C muestra la actividad bioquímica de sn2-casARN^{EX} y sn3-casARN^{EX}-AAVS1, La FIG. 8, Panel D muestra la actividad bioquímica de sn1-casARN^{EX} y sn3-casARN^{EX}-AAVS1. El último carril de la FIG. 8, Panel D contiene patrones de peso molecular. Se muestran porcentajes de escisión en la parte inferior de cada carril. Como puede verse a partir de los datos en la FIG. 8, sn1-casARN^{EX}, sn2-casARN^{EX}, y sn3-casARN^{EX}-AAVS1 tuvieron un porcentaje promedio de escisión de 46,9 % (desviación típica de 0,3 %). Para todas las reacciones en las que solamente estaban presentes dos componentes de sn-casARN^{EX} (por ejemplo FIG. 8, Panel B, FIG. 8, Panel C, FIG. 8, Panel D) no se observó actividad de escisión (para los carriles indicados como LDD, cualquier actividad de escisión estuvo por debajo del límite de detección).

20

25

Los datos presentados en la FIG. 8 demuestran que los sistemas de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II de la presente invención facilitan la escisión específica de sitio mediada por Cas9 de ADN bicatenario diana. Los datos también muestran que los tres componentes de sn-casARN son necesarios para apoyar la actividad de escisión específica de sitio mediada por Cas9.

30

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el ensayo de escisión de Cas9 descrito en este ejemplo puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos divididos.

35

Ejemplo 4

Análisis de secuenciación profunda para detección de modificaciones diana en células eucariotas

Este ejemplo ilustra el uso de análisis de secuenciación profunda para evaluar y comparar el porcentaje de escisión *in vivo* de complejos de sn-casARN/proteína Cas9 seleccionados en relación con secuencias diana de ADN bicatenario seleccionadas.

40

A. Formación de complejos RNP de sn1-casARN^{EX}, sn2-casARN^{EX}, sn3-casARN^{EX}-AAVS1 y proteína Cas9

Se marcó Cas9 de *S. pyogenes* en el extremo C terminal con dos secuencias de localización nuclear (SLN) y se expresan de forma recombinante en *E. coli*. Se prepararon complejos ribonucleoproteicos (RNP) a dos concentraciones, Cas9 50 pmol:sn-casARN^{EX} 150 pmol y Cas9 200 pmol:sn-casARN^{EX} 600 pmol, por triplicado. Los tres componentes de sn-casARN^{EX} (sn1-casARN^{EX}, sn2-casARN^{EX}, sn3-casARN^{EX}-AAVS1) en cantidades equimolares se mezclaron en un tampón de hibridación (HEPES 1,25 mM, MgCl_2 0,625 mM, KCl 9,375 mM a pH 7,5) hasta la concentración deseada (150 pmol o 600 pmol) en un volumen final de 5 μl , se incubaron durante 2 minutos a 95 °C, se retiraron del termociclador y se permitió que se equilibraran a temperatura ambiente. La proteína Cas9 se diluyó hasta una concentración apropiada en tampón de unión (HEPES 20 mM, KCl 100 mM, MgCl_2 5 mM, DTT 1 mM y glicerol 5 % a pH 7,4) hasta un volumen final de 5 μl y se mezclaron con los 5 μl de sn-casARN^{EX} desnaturalizados por calor seguido de incubación a 37 °C durante 30 minutos.

45

50

B. Transfecciones celulares usando RNP de sn-casARN^{EX}/proteína Cas9

Los complejos RNP se transfectaron en células K562 (Colección Americana de Cultivos Tipo, Manassas, VA), usando el sistema de lanzadera de Nucleofector® de 96 pocillos (Lonza, Allendale, NJ) y el siguiente protocolo. Se distribuyeron complejos RNP en un volumen final de 10 μl en pocillos individuales de una placa de 96 pocillos. Se transfirieron células K562 suspendidas en medios de un matraz de cultivo a un tubo cónico de 50 ml. Las células se

55

sedimentaron por centrifugación durante 3 minutos a 200 x g, el medio de cultivo se aspiró y las células se lavaron una vez con PBS sin calcio y magnesio. Las células K562 se sedimentaron después por centrifugación durante 3 minutos a 200 x g, el PBS se aspiró y se resuspendió el sedimento celular en 10 ml de PBS sin calcio y magnesio.

5 Las células se contaron usando el contador de células automático Countess® II (Life Technologies, Grand Island, NY). Se transfirieron $2,2 \times 10^7$ células a un tubo de 50 ml y se sedimentaron. El PBS se aspiró y las células se resuspendieron en solución Nucleofector™ SF (Lonza, Allendale, NJ) hasta una densidad de 1×10^7 células/ml. Después se añaden 20 µl de la suspensión celular a pocillos individuales que contienen 10 µl de complejos RNP y se transfirió el volumen completo a los pocillos de una placa Nucleocuvette™ de 96 pocillos (Lonza, Allendale, NJ).
 10 La placa se cargó en la Shuttle™ de 96 pocillos Nucleofector™ (Lonza, Allendale, NJ) y las células se nucleofectaron usando el programa 96-FF-120 Nucleofector™ (Lonza, Allendale, NJ). Después de la nucleofección, se añadieron 70 µl de medio de Dulbecco modificado por Iscove (IMDM; Life Technologies, Grand Island, NY), complementado con FBS 10 % (Fisher Scientific, Pittsburgh, PA), penicilina y estreptomina (Life Technologies, Grand Island, NY), a cada pocillo y se transfirieron 50 µl de la suspensión celular a un cultivo celular de 96 pocillos que contiene 150 µl de medio de cultivo completo de IMDM precalentado. La placa se transfirió después a un
 15 incubador de cultivo tisular y se mantuvo a 37 °C en CO₂ 5 % durante 48 horas.

C. Generación de ADN bicatenario diana para secuenciación profunda

Se aisló ADNg de células K562 48 horas después de transfección de RNP usando 50 µl de solución de extracción de ADN QuickExtract (Epicentre, Madison, WI) por pocillo seguido de incubación a 37 °C durante 10 minutos, 65 °C durante 6 minutos y 95 °C durante 3 minutos para detener la reacción. El ADNg aislado se diluyó después con 50 µl
 20 de agua y se almacenaron muestras a -80 °C.

Usando el ADNg aislado, se realizó una primera PCR usando mezcla maestra 2X de alta fidelidad de inicio en caliente Q5 (New England Biolabs, Ipswich, MA) a una concentración 1x, cebadores a 0,5 µM cada uno (FIG. 13, oligonucleótidos H e I), 3,75 µl de ADNg en un volumen final de 10 l y se amplificaron a 98 °C durante 1 minuto, 35 ciclos de 10 s a 98 °C, 20 s a 60 °C, 30s a 72 °C y una extensión final a 72 °C durante 2 min. La reacción de PCR se
 25 diluyó 1:100 en agua.

Se preparó una PCR de "código de barras" usando cebadores únicos para cada muestra para facilitar la secuenciación múltiple. Los pares de cebadores se muestran en la Tabla 11.

Tabla 11

Cebadores de código de barras		
ID	Muestra	Cebadores*
Conjunto 1 de CEBADORES DE CÓDIGO DE BARRAS	Cas9 50 pmol: sn-casARN rep-1 150 pmol	L, M
Conjunto 2 de CEBADORES DE CÓDIGO DE BARRAS	Cas9 50 pmol: sn-casARN rep-2 150 pmol	L, N
Conjunto 3 de CEBADORES DE CÓDIGO DE BARRAS	Cas9 50 pmol: sn-casARN rep-3 150 pmol	L, O
Conjunto 4 de CEBADORES DE CÓDIGO DE BARRAS	Cas9 200 pmol: sn-casARN rep-1 600 pmol	L, P
Conjunto 5 de CEBADORES DE CÓDIGO DE BARRAS	Cas9 200 pmol: sn-casARN rep-2 600 pmol	L, Q
Conjunto 6 de CEBADORES DE CÓDIGO DE BARRAS	Cas9 200 pmol: sn-casARN rep-3 600 pmol	L, R
* Se muestran secuencias de cebadores en la FIG. 13		

30 La PCR de código de barras se realizó usando mezcla maestra 2X de alta fidelidad de inicio en caliente Q5 (New England Biolabs, Ipswich, MA) a una concentración 1x, cebadores a 0,5 µM cada uno (Tabla 11), 1 µl primera PCR diluida 1:100, en un volumen final de 10 µl y se amplificaron a 98 °C durante 1 minuto, 12 ciclos de 10 s a 98 °C, 20 s a 60 °C, 30 s a 72 °C y una extensión final a 72 °C durante 2 min.

D. Limpieza por SPRIselect

35 Se agruparon reacciones de PCR en un único tubo de microcentrífuga para limpieza basada en perlas SPRIselect (Beckman Coulter, Pasadena, CA) de amplicones para secuenciación.

A los amplicones agrupados, se añadieron volúmenes de 0,9x de perlas de SPRIselect y se mezclaron e incubaron a temperatura ambiente (TA) durante 10 minutos. El tubo de microcentrifuga se colocó en un soporte de tubo magnético (Beckman Coulter, Pasadena, CA) hasta que se hubo clarificado la solución. El sobrenadante se retiró y se descartó, y las perlas residuales se lavaron con 1 volumen de etanol al 85 % y se incubaron a TA durante 30 s. Después de la incubación se aspiró el etanol y las perlas se secaron al aire a TA durante 10 min. El tubo de microcentrifuga se retiró después del soporte magnético y se añadieron volúmenes de 0,25x de tampón Qiagen EB (Qiagen, Venlo, Países bajos) a las perlas, se mezcló vigorosamente y se incubó durante 2 min a temperatura ambiente. El tubo de microcentrifuga se devolvió al imán, se incubó hasta que se hubo eliminado la solución y el sobrenadante que contenía los amplicones purificados se distribuyó en un tubo de microcentrifuga limpio. Se cuantificó la biblioteca de amplicones purificados usando el sistema Nanodrop™ 2000 (Thermo Scientific, Wilmington, DE) y la calidad de la biblioteca se analizó usando el sistema Fragment Analyzer™ (Advanced Analytical Technologies, Inc., Ames, IA) y el kit de reactivo de ADN bicatenario DNF-910 (Advanced Analytical Technologies, Inc., Ames, IA).

E. Preparación de secuenciación profunda

La biblioteca de amplicones se normalizó a una concentración 4 nmolar como se calcula a partir de valores de Nanodrop y tamaño de los amplicones. La biblioteca se analizó en secuenciador MiSeq (Illumina, San Diego, CA) con kit de reactivo MiSeq v2 (Illumina, San Diego, CA) durante 300 ciclos con dos procesamientos de extremos emparejados de 151 ciclos más dos lecturas de índice de ocho ciclos.

F. Análisis de datos de secuenciación profunda

La identidad de productos en los datos de secuenciación se determinó basándose en las secuencias de código de barras de índice adaptadas a los amplicones en el ciclo de código de barras de PCR. Se usó un *script* informático para procesar los datos de MiSeq ejecutando las siguientes tareas:

- Las lecturas se alinearon con el genoma humano (versión GRCh38/38) usando software Bowtie (bowtie-bio.sourceforge.net/index.shtml).
- Se compararon lecturas alineadas con la secuencia de locus de AAVS-1 de tipo silvestre esperada, se descartaron las lecturas que no se alineaban con ninguna parte del locus de AAVS-1 (Tabla 12, "otras").
- Se contaron las lecturas que coincidían con la secuencia de AAVS-1 de tipo silvestre (Tabla 12, "TS").
- Las lecturas con insup (inserción o la supresión de bases) se clasificaron por tipo de insup y se contaron (Tabla 12, "insup").
- Las lecturas de insup totales se dividieron por la suma de lecturas de tipo silvestre y las lecturas de insup proporcionaron lecturas de porcentaje de mutación.

Los resultados de este análisis se presentan en la Tabla 12.

Tabla 12

Datos de secuenciación profunda					
Tipo de muestra ¹	Total ²	Alineadas ³	TS ⁴	insup ⁵	Otras ⁶
Cas9 50 pmol:sn-casARN ^{EX} rep1 150 pmol	33807	33680	18119	15561	3
Cas9 50 pmol:sn-casARN ^{EX} rep2 150 pmol	33070	32991	18225	14766	2
Cas9 50 pmol:sn-casARN ^{EX} rep3 150 pmol	33062	32986	18580	14406	5
Cas9 200 pmol:sn-casARN ^{EX} rep1 650 pmol	34089	33993	9321	24672	1
Cas9 200 pmol:sn-casARN ^{EX} rep2 650 pmol	28691	28600	7100	26893	2
Cas9 200 pmol:sn-casARN ^{EX} rep3 650 pmol	28573	28509	12184	16325	1

1 - Tipo de muestra; 2 - Lecturas de MiSeq totales; 3 - Las lecturas totales se alinean con el locus diana (AAVS-1); 4 - Lecturas ts totales (es decir secuencia no modificada); 5 - Lecturas mutadas (cas9 escindida); 6 - Lecturas que no se alinean con el locus de AAVS-1 locus.

Como se puede ver a partir de los insup medidos entre repeticiones en la Tabla 12, los sistemas de sn-casPN/proteína Cas9 son capaces de modificar *in vivo* un locus diana. Además, la mayor frecuencia de insup como resultado de aumento de la concentración de sn-casPN/Cas9 transfectado es indicativa de escisión mediada por sistema de sn-casPN/proteína Cas9 dependiente de dosis. Los datos presentados en la Tabla 12 demuestran que los sistemas de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II de la presente invención facilitan la escisión específica de sitio mediada por Cas9 *in vivo* de un locus genómico.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el análisis descrito en este ejemplo

puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexa dividido.

Ejemplo 5

5 Escisión de sn-casARN/proteína Cas9 facilitada por Csy4*

Este ejemplo ilustra el uso de sn-casARN de la presente invención y una proteína efectora, la proteína Csy4 de *P. aeruginosa* deficiente en nucleasa que posee la mutación H29A (Csy4*), para aumentar la asociación de dos sn-casARN aumentados con una secuencia de unión a ARN de Csy4.

A. Generación de componentes de sn-casARN

- 10 Los moldes de ADN bicatenario para los componentes de sn-casARN^{EXCsy} específicos que comprenden una secuencia de unión de Csy4 se ensamblaron mediante PCR usando cebadores solapantes 3' que contenían las secuencias de ADN correspondientes para los componentes de sn-casRNA^{EXCsy}. Los oligonucleótidos utilizados en el ensamblaje se presentan en la Tabla 13.

Tabla 13

Cebadores solapantes para la generación de sn-casARN ^{EXCsy} con secuencia de unión a ARN de Csy4		
Tipo de componente de ARN de Cas	Secuencia de unión a diana de ADN	Oligonucleótidos*
primer polinucleótido con secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -Csy)	AAVS-1	A, Y, S, T
primer polinucleótido con secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -Csy)	CD34	A, Z, S, T
primer polinucleótido con secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -Csy)	CD151	A, AA, S, T
primer polinucleótido con secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -Csy)	JAK-1	A, AB, S, T
Segundo polinucleótido con secuencia de unión a Csy4 3' del nexa dividido (sn2-casARN ^{EXCsy} -Csy)	n/d	A, U, AC
primer polinucleótido de AAVS con conector + secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -Csy)	AAVS-1	A, Y, V, W
primer polinucleótido de CD34 con conector + secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -InkCsy)	CD34	A, Z, V, W
primer polinucleótido de CD151 con conector + secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -InkCsy)	CD151	A, AA, V, W
primer polinucleótido de JAK-1 con conector + secuencia de unión a Csy4 3' del nexa dividido (sn1-casARN ^{EXCsy} -InkCsy)	JAK-1	A, AB, V, W
Segundo polinucleótido con conector + secuencia de unión a Csy4 3' del nexa dividido (sn2-casARN ^{EXCsy} -InkCsy)	n/d	A, X, AC

*se muestran secuencias de cebadores de ADN en la FIG. 13

- 15 Los cebadores de ADN estaban presentes a una concentración de 2 nM cada uno. Se usaron dos cebadores de ADN externos correspondientes al promotor T7 (cebador directo: Oligonucleótido A, Tabla 13) y el extremo 3' de la secuencia de ARN (cebadores inversos: Oligonucleótidos T, AC o W, Tabla 13) a 640 nM para conducir la reacción de amplificación. Se realizó PCR y transcripción como se ha descrito en el Ejemplo 1 descrito en la presente memoria descriptiva. Las secuencias de sn-casARN^{EXCsy} transcritas se muestran en la Tabla 14.

Tabla 14

Secuencias de sn-casARN ^{EXCsy}		
Tipo de componente de ARN de Cas	Secuencia de ARN	SEQ ID NO
sn1-casARN ^{EXCsy} -Csy-AAVS1	5' - GGGGCCACUA GGGACAGGAU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUACUG CC - 3'	SEQ ID NO: 71
sn1-casARN ^{EXCsy} -Csy-CD34	5' - GUUUGUGUUU CCAUAAACUG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUACUG CC - 3'	SEQ ID NO: 72
sn1-casARN ^{EXCsy} -Csy-CD151	5' - GCCCGCCACC ACCAGGAUGU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUACUG CC - 3'	SEQ ID NO: 73
sn1-casARN ^{EXCsy} -Csy-JAK-1	5' - GGCAGCCAGC AUGAUGAGAC GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUACUG CC - 3'	SEQ ID NO: 74
sn2-casARN ^{EXCsy} -Csy	5' - GGCAGGUCCG UUAUCAACUU GAAAAAGUGG CACCGAGUCG GUGCUU - 3'	SEQ ID NO: 75
sn1-casARN ^{EXCsy} -InkCsy-AAVS1	5' - GGGGCCACUA GGGACAGGAU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUAGUU CACUGCC - 3'	SEQ ID NO: 76
sn1-casARN ^{EXCsy} -InkCsy-CD34	5' - GUUUGUGUUU CCAUAAACUG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUAGUU CACUGCC - 3'	SEQ ID NO: 77
sn1-casARN ^{EXCsy} -InkCsy-CD151	5' - GCCCGCCACC ACCAGGAUGU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUAGUU CACUGCC - 3'	SEQ ID NO: 78

(continuación)

Secuencias de sn-casARN ^{EXCsy}		
Tipo de componente de ARN de Cas	Secuencia de ARN	SEQ ID NO
sn1-casARN ^{EXCsy} -InkCsy-JAK-1	5' - GGCAGCCAGC AUGAUGAGAC GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUAGUU CACUGCC - 3'	SEQ ID NO: 79
sn2-casARN ^{EXCsy} -InkCsy	5' - GGGCAGUGAA CUAGCCUUAU CUCAACUUGC UAUGCUGUCC UGUUUCCAGG ACAGCAUAGC UCUGAGAC - 3'	SEQ ID NO: 80

B. Generación de dianas de ADN bicatenario para ensayos bioquímicos

Se produjo ADN bicatenario diana para su uso en los ensayos de escisión de Cas9 *in vitro* usando amplificación por PCR como se ha descrito en el Ejemplo 2 en el presente documento. Los cebadores directo e inverso usados para amplificación a partir de ADNg fueron los siguientes: los oligonucleótidos de AAVS-1 fueron J y K (FIG. 13), la diana de ADN bicatenario amplificada para AAVS-1 fue de 288 pb; los oligonucleótidos de CD34 (antígeno de células progenitoras hematopoyéticas) fueron AD y AE (FIG. 13), la diana de ADN bicatenario amplificada para CD34 fue de 258 pb; los oligonucleótidos de CD151 (antígeno de tetraspanina de células endoteliales plaquetarias) fueron AF y AG (FIG. 13), la diana de ADN bicatenario de CD151 amplificada fue de 272 pb; y, los oligonucleótidos de JAK-1 (quinasa Janus 1) fueron AH y AI (FIG. 13), la diana de ADN bicatenario de JAK-1 amplificada fue de 298 pb.

C. Csy4* apoyó el ensayo bioquímico de escisión de Cas9

Se prepararon sn-casARN^{EXCsy} para su uso en el ensayo bioquímico como se ha descrito en el Ejemplo 3 en el presente documento. Con la modificación de que antes de la adición de Cas9, se añadió proteína Csy4* 250 nM a la reacción y se incubaron sn-casARN^{EXCsy} y Csy4* a 37 °C durante 5 min. Después de la incubación, se añadió Cas9 y se llevaron a cabo reacciones bioquímicas como se ha descrito en el Ejemplo 3. Se incluyó un control sin Csy4*.

La FIG. 9 presenta los resultados del ensayo de escisión de Cas9 usando la proteína Csy4* y los sn-casARN^{EXCsy}. Los ensayos de escisión usaron dos sistemas diferentes de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos que fueron variantes del sistema presente en la FIG. 3B. En el primer sistema los sn1-casARN^{EXCsy} comprendían además un primer polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento de unión de Csy4 I (sn1-casARN^{EXCsy}-Csy) y el sn2-casARN comprendía un segundo polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento de unión de Csy4 II (sn2-casARN^{EXCsy}-Csy), en el que el primer polinucleótido adyuvante y el segundo polinucleótido adyuvante se asocian para formar un elemento de unión a ARN de Csy4 (sn1-casARN/sn2-casARN/Csy4ARN). En el segundo sistema el sn1-casARN comprendía además un primer polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento conector I y una secuencia de nucleótidos de elemento de unión de Csy4 I (sn1-casARN^{EXCsy}-InkCsy) y el sn2-casARN comprendía un segundo polinucleótido adyuvante que comprende una secuencia de nucleótidos de elemento conector II y una secuencia de nucleótidos de elemento de unión de Csy4 II (sn2-casARN^{EXCsy}-InkCsy), en el que el primer polinucleótido adyuvante y el segundo polinucleótido adyuvante se asocian para formar un elemento conector y un elemento de unión a ARN de Csy4 (véase, por ejemplo, las representaciones generales en la FIG. 6A y la FIG. 6B). Cada uno de los dos sistemas se usó para dirigir la escisión a cuatro dianas diferentes, en las que los sn-casARN^{EXCsy} comprendían cada uno un espaciador complementario de una de las cuatro dianas: AAVS-1, CD-34, CD-151 y JAK-1 (véase Tabla 13 anterior). En la figura, la actividad de escisión se muestra en la parte inferior de cada carril (excepto para los carriles 1 y 10, que son patrones de peso molecular). Para carriles indicados como LDD, cualquier actividad de escisión estuvo por debajo del límite de detección. Los sistemas usados en cada una de las reacciones de ensayo de escisión de Cas9 fueron como se muestra en la Tabla 5/FIG. 9 (véase Breve descripción de las figuras, FIG. 9).

Para la diana de ADN bicatenario de AAVS-1, la aparición de bandas de ADN a aproximadamente 174 pb y aproximadamente 114 pb indicó que se había producido escisión del ADN diana. Para la diana de ADN bicatenario de CD34, la aparición de bandas de ADN a aproximadamente 105 pb y aproximadamente 153 pb indicó que se había producido escisión del ADN diana. Para la diana de ADN bicatenario de CD151, la aparición de bandas de ADN a aproximadamente 109 pb y aproximadamente 163 pb indicó que se había producido escisión del ADN diana. Para la diana de ADN bicatenario de JAK-1, la aparición de bandas de ADN a aproximadamente 204 pb y

aproximadamente 94 pb indicó que se había producido escisión del ADN diana. Se usó una escalera de ADN de 100 pb como un patrón de peso molecular (New Englan Biolabs, Ipswich, MA).

5 Como puede verse a partir de los datos en la FIG. 9, la adición de Csy4* potenció la actividad de escisión del sistema de sn-casARN^{EXCsy} para múltiples secuencias de ADN bicatenario: para AAVS-1 comparar los carriles 2/3 (sin proteína Csy4*) con los carriles 4/5, respectivamente; para CD-34 comparar los carriles 6/7 (sin proteína Csy4*) con los carriles 8/9; para CD-151 comparar los carriles 11/12 (sin proteína Csy4*) con los 13/14; y, para JAK-1 comparar los carriles 15/16 (sin proteína Csy4*) con los carriles 17/18.

10 Los datos presentados en la FIG. 9 demuestran que una proteína efectora (aquí Csy4*) potenció la escisión de ADN bicatenario diana mediante sistemas de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de la presente invención que comprenden polinucleótidos adyuvantes que tienen un elemento de unión efector (por ejemplo, la secuencia de unión a ARN de Csy4).

15 Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el aumento de la asociación de dos sn-casARN que comprendían una secuencia de unión a ARN de Csy4 con una proteína Csy4 de *P. aeruginosa* deficiente en nucleasa como se describe en este ejemplo puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexo dividido. Asimismo, a la vista de la orientación en la presente memoria descriptiva y los ejemplos un experto habitual en la materia puede usar otras combinaciones de proteína efectora/secuencia de unión a efector como se ejemplifica en el presente documento por la proteína Csy4*/secuencia de unión a ARN de Csy4.

Ejemplo 6

Actividad de escisión de sn1-CasARN/sn2-casARN/Cas9

25 Este ejemplo ilustra el uso de un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos de la presente invención en ensayos de escisión de Cas9 *in vitro* para evaluar y comparar el porcentaje de escisión de complejos de sn1-casARN/sn2-casARN/proteína Cas9 seleccionados en relación con secuencias diana de ADN bicatenario seleccionadas.

30 Los moldes de ADN bicatenario para los componentes de sn-casARN^{EX2} usados en este ejemplo se ensamblaron mediante PCR usando cebadores solapantes 3' que contenían las secuencias de ADN correspondientes para los componentes de sn-casARN^{EX2}. Se presenta una representación gráfica de los componentes de sn-casARN^{EX2} en la FIG. 10. Los oligonucleótidos utilizados en el ensamblaje se presentan en la Tabla 15.

Tabla 15

Cebadores solapantes para generación de sn1-casARN y sn2-casARN	
Tipo de componente de sn-casARN ^{EX2}	Oligonucleótidos*
AAVS-1 sn1-casARN	Y, AJ
CD151 sn1-casARN	AA, AJ
JAK-1 sn1-casARN	AB, AJ
sn2-casARN	A, C, B
*se muestran secuencias de cebadores de ADN en la FIG. 13	

Se realizó generación de molde de ADN bicatenario para transcripción de ARN como se ha descrito en el Ejemplo 1 en el presente documento. Las secuencias de sn-casARN^{EX2} transcritas se muestran en la Tabla 16.

Tabla 16

Secuencias de sn-casARN		
Tipo de componente de ARN de Cas	Secuencia de ARN	SEQ ID NO
sn1-casARN ^{EX2} -AAVS1	5' - GGGGCCACUA GGGACAGGAU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUA - 3'	SEQ ID NO: 81
sn1-casARN ^{EX2} -CD151	5' - GCCCGCCACC ACCAGGAUGU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUA - 3'	SEQ ID NO: 82
sn1-casARN ^{EX2} -JAK-1	5' - GGCAGCCAGC AUGAUGAGAC GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUA - 3'	SEQ ID NO: 83
sn2-casARN ^{EX2}	5' - GUCCGUUAUC AACUUGAAAA AGUGGCACCG AGUCGGUGCU U - 3'	SEQ ID NO: 84

Se produjo ADN bicatenario diana para su uso en los ensayos de escisión de Cas9 *in vitro* usando amplificación por PCR como se ha descrito en el Ejemplo 2 en el presente documento. Los cebadores directo e inverso usados para amplificación a partir de ADN_g fueron los siguientes: los oligonucleótidos de AAVS-1 fueron J y K (FIG. 13), la diana de ADN bicatenario amplificada para AAVS-1 fue de 288 pb; los oligonucleótidos de CD151 fueron AF y AG (FIG. 13), la diana de ADN bicatenario de CD151 amplificada fue de 272 pb; y, los oligonucleótidos de JAK-1 fueron AH y AI, la diana de ADN bicatenario de JAK-1 amplificada fue de 298 pb. Se realizó escisión *in vitro* como se ha descrito en el Ejemplo 3 en el presente documento.

La FIG. 10 presenta el resultado del ensayo de escisión de Cas9 usando el sn1-casARN^{EX2} y sn2-casARN^{EX2} descritos anteriormente. Se muestran porcentajes de escisión en la parte inferior de cada carril excepto el carril 1, que es un patrón de peso molecular. La FIG. 10, carril 2, presenta resultados de escisión para un sistema de sn1-casARN^{EX2}-AAVS1 y sn2-casARN^{EX2}, que demostraron una actividad de escisión de 97,6 %. La FIG. 10, carril 3, presenta resultados de escisión para un sistema de sn1-casARN^{EX2}-CD151 y sn2-casARN^{EX2}, que demostraron una actividad de escisión de 48,8 %. La FIG. 10, carril 4, presenta los resultados para un sistema de sn1-casARN^{EX2}-JAK1 y sn2-casARN^{EX2}, que demostraron una actividad de escisión de 60,0 %.

Los datos presentados en la FIG. 10 demostraron que las construcciones de sn1-casARN y sn2-casARN como se describe en el presente documento facilitan la escisión específica de sitio mediada por Cas9 *in vitro* de una diana de ADN bicatenario.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el ensayo de escisión de Cas9 descrito en este ejemplo puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexo dividido.

Ejemplo 7

Actividad de escisión de sn1-CasARN^{EX3Csy}/sn2-casARN^{EX3Csy}/Cas9

Este ejemplo ilustra el uso de dos sistemas diferentes de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II de dos polinucleótidos de la presente invención para evaluar y comparar su porcentaje de actividad de escisión en relación con secuencias diana de ADN bicatenario seleccionadas.

Los dos sistemas diferentes de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II fueron los siguientes: uno fue el sistema ilustrado en la FIG. 7A (sn1-casARN^{EX3Csy}-Csy-AAVS1/sn2-casARN^{EX3Csy}-Csy); y el segundo fue una variante del sistema presente en la FIG. 7A. En el segundo sistema el sn1-casARN^{EX3Csy}-lnkCsy-AAVS1 comprendía, de 5' a 3', una secuencia de nucleótidos de elemento de tallo de nexo dividido I, un primer polinucleótido adyuvante (que tiene una secuencia de nucleótidos de elemento conector I y un polinucleótido

5 formador de horquilla), y el sn2-casARN^{EX3Csy}-InkCsy-AAVS1 comprendía, de 5' a 3', un segundo polinucleótido adyuvante (que tiene un polinucleótido formador de horquilla y una secuencia de nucleótidos de elemento conector II) y una secuencia de nucleótidos de elemento de tallo de nexo dividido II. Cada uno de los dos sistemas se usó para dirigir la escisión de una diana de AAVS-1, en la que el sn1-casARN^{EX3Csy}-AAVS1 y el sn1-casARN^{EX3Csy}-InkCsy-AAVS1 comprendían cada uno un espaciador complementario del AAVS-1.

Los moldes de ADN bicatenario para componentes de sn-casARN^{EX3-Cys} usados en este ejemplo se ensamblaron mediante PCR usando cebadores solapantes 3' que contenían las secuencias de ADN correspondientes para los componentes de sn-casARN^{EX3-Cys}. Los oligonucleótidos utilizados en los ensamblajes se presentan en la Tabla 17.

Tabla 17

Cebadores solapantes para la generación de componentes de sn-casARN ^{EX3-Cys}	
Tipo de componente de sn-casARN ^{EX3-Cys}	Oligonucleótidos*
sn1-casARN ^{EX3Csy} -Csy-AAVS1	A, AK, AL, AM
sn2-casARN ^{EX3Csy} -Csy	A, AN, AO, AC
sn1-casARN ^{EX3Csy} -InkCsy-AAVS1	A, AK, AP, AQ
sn2-casARN ^{EX3Csy} -InkCsy	A, AR, AS, AC

*se muestran secuencias de cebadores de ADN en la FIG. 13

10 Se realizó generación de molde de ADN bicatenario para transcripción de ARN como se ha descrito en el Ejemplo 1. Las secuencias de sn-casARN^{EX3-Cys} transcritas se muestran en la Tabla 18.

Tabla 18

Secuencias de sn-casARN		
Tipo de componente de ARN de Cas9	Secuencia de ARN	SEQ ID NO
sn1-casARN ^{EX3Cs} -y-Csy-AAVS1	5' - GGGGCCACUA GGGACAGGAU GUCUCAGAGC UAUGCAGUCC UGGAAACAGG ACUGCAUAGC AAGUUGAGAU AAGGCUACUG CCGUAUAGGC AG - 3'	SEQ ID NO: 85
sn2-casARN ^{EX3Csy} -Csy	5' - CUGCCGUAUA GGCAGGUCCG UUAUCAACUU GAAAAAGUGG CACCGAGUCG GUGCUU - 3'	SEQ ID NO: 86
sn1-casARN ^{EX3Csy} -InkCsy-AAVS1	5' - GGGGCCACUA GGGACAGGAU GUCUCAGAGC UAUGCAGUCC UGGAAACAGG ACUGCAUAGC AAGUUGAGAU AAGGCUAGAC ACUGCCGUAU AGGCAG - 3'	SEQ ID NO: 87
sn2-casARN ^{EX3Csy} -InkCsy	5' - CUGCCGUAUA GGCAGAGACA GUCCGUUAUC AACUUGAAAA AGUGGCACCG AGUCGGUGCUU - 3'	SEQ ID NO: 88

15 Se produjo ADN bicatenario diana para su uso en los ensayos de escisión de Cas9 *in vitro* usando amplificación por PCR como se ha descrito en el Ejemplo 2. Los cebadores directo e inverso usados para amplificación a partir de ADN_g fueron los siguientes: AAVS-1, oligonucleótidos H e I (FIG. 13). La diana de ADN bicatenario amplificada para AAVS-1 fue de 495 pb. Se realizó escisión *in vitro* como se ha descrito en el Ejemplo 3.

La FIG. 11 presenta los resultados del ensayo de escisión de Cas9 usando los sn-casARN descritos anteriormente.

En la figura, la actividad de escisión se muestra en la parte inferior de cada carril (excepto para los carriles 1 y 10, que son patrones de peso molecular). Para carriles indicados como LDD, cualquier actividad de escisión estuvo por debajo del límite de detección. Los sistemas usados en cada una de las reacciones de ensayo de escisión de Cas9 fueron como se muestra en la Tabla 6 (véase Breve descripción de las figuras, FIG. 11).

- 5 Como puede verse a partir de los datos en la FIG. 11, tanto sn1-casARN^{EX3Csy}-Csy-AAVS1 como sn2-casARN^{EX3Csy}-Csy (FIG. 11, carriles 2 y 3) o sn1-casARN^{EX3Csy}-InkCsy-AAVS1 y sn2-casARN^{EX3Csy}-InkCsy son necesarios para actividad de escisión detectable (FIG. 11, carriles 6 y 7). Asimismo, la escisión potenciada fue detectable cuando se introdujo una secuencia de nucleótidos de elemento conector entre el elemento de nexo dividido (FIG. 11, carril 8 en comparación con el carril 4). Además, cuando se introduce proteína Csy4* se observa escisión potenciada con sn1-casARN^{EX3Csy}-InkCsy-AAVS1 y sn2-casARN^{EX3Csy}-InkCsy (FIG. 11, carril 9 en comparación con el carril 8), pero no en ausencia de las secuencias conectoras (sn1-casARN^{EX3Csy}-Csy-AAVS1 y sn2-casARN^{EX3Csy}-Csy; FIG. 11, carril 5 en comparación con el carril 4).

15 Los datos presentados en la FIG. 11 demuestran que las construcciones de sn1-casARN y sn2-casARN como se describe en el presente documento facilitan la escisión específica de sitio mediada por Cas9 *in vitro* de una diana de ADN bicatenario.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el ensayo de escisión de Cas9 descrito en este ejemplo puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexo dividido.

20 Ejemplo 8

Identificación y exploración de ARN de CRISPR de activación en trans

25 Este ejemplo ilustra el procedimiento mediante el que pueden identificarse ARN de CRISPR de activación en trans (ARNcrtra) de especies que tienen sistema de CRISPR-Cas9 de tipo II. El procedimiento presentado aquí está adaptado de Chylinski, y col., "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems," RNA Biol., 10(5), 726-37 (2013). No son necesarias todas las siguientes etapas para exploración ni es necesario que el orden de las etapas sea como se presenta.

A. Identificación de una especie que contiene un sistema de CRISPR-Cas9 de tipo II

30 Usando la herramienta de búsqueda de alineamiento local básica (BLAST, blast.ncbi.nlm.nih.gov/Blast.cgi), se realiza una búsqueda de los genomas de diversas especies para identificar proteínas Cas9 o de tipo Cas9. Los sistemas de CRISPR-Cas9 muestran una alta diversidad de secuencia entre especies, sin embargo, los ortólogos de Cas9 muestran arquitectura de dominio conservado de un dominio de endonucleasa de HNH central y un dominio de RuvC/RNasa H dividido. Los resultados de BLAST primarios se filtran para dominios identificados; se descartan secuencias incompletas o truncadas y se identifican ortólogos de Cas9.

35 Cuando se identifica un ortólogo de Cas9 en una especie, se exploran secuencias adyacentes a la secuencia codificante de ortólogo de Cas9 para determinar otras proteínas Cas y una matriz de repetición-espaciador asociada para identificar todas las secuencias que pertenecen al locus de CRISPR-Cas9. Esto puede realizarse mediante alineamiento con otros loci de CRISPR-Cas9 de tipo II ya conocidos en el dominio público, con el conocimiento de que especies estrechamente relacionadas muestran arquitectura de locus de CRISPR-Cas9 similar (es decir, composición de proteína Cas, tamaño, orientación, localización de la matriz, localización de ARNcrtra, etc.). El elemento de ARNcrtra está contenido habitualmente en el locus de CRISPR-Cas9 de tipo II y se identifica rápidamente por su complementariedad de secuencia con los elementos repetidos en la matriz de repetición-espaciador (secuencia de crtra anti repetición).

45 Una vez que se ha identificado la secuencia del locus de CRISPR-Cas9 para el ortólogo de Cas9 para la especie, se usa exploración predictiva por ordenador para extraer la secuencia anti repetición para identificar el ARNcrtra asociado. Se exploran anti repeticiones potenciales, por ejemplo, de la siguiente manera.

Si la secuencia repetida es de una especie conocida, se identifica en y se recupera de la base de datos CRISPRdb (crispr.u-psud.fr/crispr/). Si se desconoce que la secuencia repetida esté asociada con una especie, se predicen secuencias repetidas usando software CRISPRfinder (crispr.u-psud.fr/Server/) usando el locus de CRISPR-Cas9 de tipo II locus para la especie descrita anteriormente.

50 La secuencia repetida identificada para la especie se usa para explorar el locus de CRISPR-Cas9 para determinar la secuencia anti repetición (por ejemplo, usando el algoritmo de BLASTp o similares). La búsqueda se restringe habitualmente a regiones intergénicas del locus de CRISPR-Cas9.

Se valida una región anti repetición identificada para complementariedad con la secuencia repetida identificada.

Se explora una región anti repetición potencial tanto 5' como 3' de la anti repetición potencial para determinar un

terminador de la transcripción independiente de Rho (TransTerm HP, transterm.cbcb.umd.edu/).

Por lo tanto, se ha determinado que la secuencia identificada que comprende el elemento anti repetición y el terminador de la transcripción independiente de Rho es el ARNcrtra potencial de la especie dada.

B. Preparación de biblioteca de Sec-ARN

- 5 El ARNcrtra potencial que se identificó por ordenador se valida adicionalmente usando secuenciación de ARN (secARN).

Se obtienen células de especies de las que se identificó el ARNcrtra potencial de un depósito comercial (por ejemplo, Colección Americana de Cultivos Tipo, Manassas, VA; DSMZ, Braunschweig, Alemania).

- 10 Las células se cultivan hasta fase semilogarítmica y se prepara ARN total usando reactivo de Trizol (Sigma-Aldrich, St. Louis, MO) y se trata con DNAsal (Fermentas, Vilnius, Lituania).

Se tratan 10 ug del ARN total con kit de retirada de ARNr Ribo-Zero (Illumina, San Diego, Ca) y el ARN restante se purifica usando RNA Clean and Concentrators (Zymo Research, Irvine, CA).

- 15 Después se prepara una biblioteca usando el kit de preparación de biblioteca de ARN pequeños TruSeq (Illumina, San Diego, CA) siguiendo las instrucciones del fabricante, que da como resultado la presencia de secuencias adaptadoras asociadas con el ADNc.

La biblioteca de ADNc resultante se secuencia usando un secuenciador MiSeq (Illumina, San Diego, CA).

C. Procesamiento de datos de secuenciación

Se procesan lecturas de secuenciación de la biblioteca de ADNc usando el siguiente procedimiento.

- 20 Se retiran secuencias adaptadoras usando cutadapt 1.1 (pypi.python.org/pypi/cutadapt/1.1) y se recortan 15 nt del extremo 3' de la lectura para mejorar la calidad de lectura.

Las lecturas se alinean de nuevo con el genoma de especies respectivas (a partir del que se identificó ARNcrtra potencial) con una admisión de desapareamientos de 2 nucleótidos.

La cobertura de lectura se calcula usando BedTools (bedtools.readthedocs.org/en/latest/).

- 25 Se usa Integrative Genomics Viewer (IGV, www.broadinstitute.org/igv/) para mapear la posición de inicio (5') y terminación (3') de lecturas. Las lecturas totales recuperadas para el ARNcrtra se calculan a partir del archivo SAM de alineamientos.

- 30 Los datos de sec-ARN se usan para validar que un elemento de ARNcrtra potencial se transcriba de forma activa *in vivo*. Se validan aciertos confirmados del compuesto de las exploraciones por ordenador y sec-ARN para capacidad funcional de la secuencia de ARNcrtra identificada y su ARNcr afin para apoyar la escisión mediada por Cas9 de una diana de ADN bicatenario usando procedimientos esbozados en el presente documento (véase Ejemplos 1, 2 y 3).

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos en el presente documento, la identificación de secuencias de ARNcrtra novedosas puede ser practicada por un experto habitual en la materia.

Ejemplo 9

Ensayo de T7E1 para detección de modificaciones diana en células eucariotas

- 35 Este ejemplo ilustra el uso de ensayos de T7E1 para evaluar y comparar el porcentaje de escisión *in vivo* de sistemas de sn-casPN/proteína Cas9 en relación con secuencias diana de ADN bicatenario seleccionadas.

A. Transfecciones celulares usando componentes polinucleotídicos de Cas

- 40 Se transfectan sn-casPN en células HEK293 que expresan de forma constitutiva fusión de SpyCas9-GFP (HEK293-Cas9-GFP), usando el sistema de lanzadera de Nucleofector® de 96 pocillos (Lonza, Allendale, NJ) y el siguiente protocolo. Se preparan cantidades molares iguales de componentes polinucleotídicos de Cas en un tampón de hibridación (HEPES 1,25 mM, MgCl₂ 0,625 mM, KCl 9,375 mM a pH 7,5), se incuban durante 2 minutos a 95 °C, se retiran del termociclador, se permitió que se equilibraran a temperatura ambiente y se distribuyeron en un volumen final de 10 µl en una placa de 96 pocillos. Se aspira el medio de cultivo de células HEK293-Cas9-GFP, y las células se lavan una vez con PBS sin calcio y magnesio, y después se tripsiniza por la adición de TrypLE (Life Technologies, Grand Island, NY) seguido de incubación a 37 °C durante 3-5 minutos. Las células tripsinizadas se pipetea suavemente arriba y abajo para formar una suspensión de células individuales y se añaden a medio de cultivo completo DMEM compuesto de medio de cultivo DMEM (Life Technologies, Grand Island, NY) que contiene FBS 10 % (Fisher Scientific, Pittsburgh, PA) y se complementa con penicilina y estreptomina (Life Technologies, Grand Island, NY).
- 45

Las células se sedimentan después por centrifugación durante 3 minutos a 200 x g, el medio de cultivo se aspira y las células se resuspenden en PBS. Las células se cuentan usando el contador de células automático Countess® II (Life Technologies, Grand Island, NY). Se transfieren $2,2 \times 10^7$ células a un tubo de 50 ml y se sedimentan. El PBS se aspira y las células se resuspenden en solución Nucleofector™ SF (Lonza, Allendale, NJ) hasta una densidad de 1×10^7 células/ml. Después se añaden 20 µl de la suspensión celular a pocillos individuales que contienen 10 µl de componentes polinucleotídicos de Cas y se transfiere el volumen completo a los pocillos de una placa Nucleocuvette™ de 96 pocillos (Lonza, Allendale, NJ). La placa se carga en la Shuttle™ de 96 pocillos Nucleofector™ (Lonza, Allendale, NJ) y las células se nucleofectan usando el programa 96-CM-130 Nucleofector™ (Lonza, Allendale, NJ). Después de la nucleofección, se añaden 70 µl de medio de cultivo completo DMEM a cada pocillo y se transfieren 50 µl de la suspensión celular a una placa de cultivo celular de 96 pocillos recubierta con colágeno que contiene 150 µl de medio de cultivo completo DMEM precalentado. La placa se transfiere después a un incubador de cultivo tisular y se mantiene a 37 °C en CO₂ 5 % durante 48 horas.

B. Generación de ADN bicatenario diana para ensayo de T7E1

Se aísla ADNg de células HEK-293-SpyCas9 48 horas después de transfección de componentes polinucleotídicos de Cas usando 50 µl de solución de extracción de ADN QuickExtract (Epicentre, Madison, WI) por pocillo seguido de incubación a 37 °C durante 10 minutos, 65 °C durante 6 minutos y 95 °C durante 3 minutos para detener la reacción. Después se diluye ADNg con 150 µl de agua y las muestras se almacenan a -80 °C.

Se genera ADN para T7E1 mediante amplificación por PCR de una secuencia de ADN bicatenario diana (por ejemplo, AAVS-1) de ADNg aislado. Se preparan reacciones de PCR usando 8 µl de ADNg como molde con polimerasa de inicio en caliente KAPA HiFi y que contiene 0,5 U de polimerasa, tampón de reacción 1x, dNTP 0,4 mM y cebadores directo e inverso 300 nM dirigidos al ADN bicatenario diana (por ejemplo, AAVS-1, oligonucleótidos K y L (FIG. 13)) en un volumen total de 25 µl. Se amplifica ADN diana usando las siguientes condiciones: 95 °C durante 5 minutos, 4 ciclos de 20 s a 98 °C, 20 s a 70 °C, menos 2 °C/ciclo, 30 s a 72 °C, seguido de 30 ciclos de 15 s a 98 °C, 20 s a 62 °C, 20 s a 72 °C y una extensión final a 72 °C durante 1 minuto.

C. Ensayo de T7E1

Se desnaturaliza ADN bicatenario diana amplificado por PCR para ensayos de T7E1 a 95 °C durante 10 minutos y después se permitió que se rehibridara enfriando hasta 25 °C a -0,5 °C/s en un termociclador. El ADN rehibridado se incubaba con 0,5 µl de endonucleasa T7 I en tampón NEBuffer 2 1x (New England Biolabs, Ipswich, MA) en un volumen total de 15 µl durante 25 minutos a 37 °C. Se analizan reacciones de T7E1 usando el sistema Fragment Analyzer™ (Advanced Analytical Technologies, Inc., Ames, IA) y el kit de reactivo de ADN bicatenario DNF-910 (Advanced Analytical Technologies, Inc., Ames, IA). El sistema Fragment Analyzer™ proporciona la concentración de cada fragmento de escisión y del ADN bicatenario diana que permanece después de la escisión.

Se calculan los porcentajes de escisión del ADN bicatenario diana a partir de la concentración de cada fragmento de escisión y el ADN bicatenario diana, que permanece después de haber tenido lugar la escisión, usando la siguiente fórmula:

$$\% \text{ de escisión} = \left(1 - \sqrt{1 - \frac{(\text{frag1} + \text{frag2})}{(\text{frag1} + \text{frag2} + \text{parental})}} \right) \text{ Ecuación 1}$$

En la Ecuación 1, las concentraciones de "frag1" y "frag2" corresponden a la concentración de fragmentos de escisión de Cas9 de la diana de ADN bicatenario y "parental" corresponde al ADN bicatenario diana que permanece después de haber tenido lugar la escisión.

El ensayo de T7E1 para detección de modificaciones diana en células eucariotas proporciona datos para demostrar que los sistemas de sn-casPN/proteína Cas9 como se describe en el presente documento facilitan la escisión *in vivo* específica de sitio mediada por Cas9 de ADN bicatenario diana. También pueden incluirse polinucleótidos de ARNgi y/o ARNcrtra/ARNcr que tienen la misma secuencia de unión a diana de ADN que los sn-casPN en el ensayo para comparar los porcentajes de escisión específica de sitio mediada por Cas9 entre las construcciones.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el ensayo de T7E1 descrito en este ejemplo puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos dividido.

Ejemplo 10

Ensayos de nexos dividido de ARNcrtra identificados

Este ejemplo describe la generación y el ensayo de modificaciones de nexos dividido en ARNcrtra, por ejemplo, basándose en ARNcr/ARNcrtra conocidos en la técnica o identificados por procedimientos descritos en el Ejemplo 8.

Se unen una secuencia de ARNcrtra y sus secuencias de ARNcr afines, con la secuencia de ARNcr colocada 5' de la secuencia de ARNcrtra que mantiene la polaridad de 5' a 3', con una secuencia conectora para generar un ARNgi. Una secuencia conectora adecuada es 5-GAAA-3.

5 El ARNgi se analiza para determinar motivos estructurales secundarios usando software de plegamiento de ARN disponible públicamente. Uno de dichos softwares es RNAstructure (rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html).

10 Las estructuras secundarias del ARNgi se analizan para determinar la estructura secundaria similar a ARNgi conocido que apoya la actividad de escisión dirigida por Cas9, que comprende tradicionalmente, en una dirección de 5' a 3', un primer elemento de tallo, un elemento de horquilla que comprende un segundo elemento de tallo (denominado en el presente documento un elemento de nexo) y cero, uno o dos, elementos de horquilla 3' del elemento de nexo.

15 El ARNgi se divide después en el elemento de nexo en al menos dos polinucleótidos: un primer polinucleótido (por ejemplo, un sn1-casPN, FIG. 3B) que comprende, en una dirección de 5' a 3', una secuencia de unión a diana de ADN seleccionada, el primer elemento de tallo y una primera parte del nexo (es decir, una secuencia de nucleótidos de elemento de tallo de nexo dividido I); y un segundo polinucleótido (por ejemplo, un sn2-casPN, FIG. 3B) que comprende, en una dirección de 5' a 3', una segunda parte del nexo (es decir, una secuencia de nucleótidos de elemento de tallo de nexo dividido II) y las cero, una o dos horquillas 3'.

20 Se construye una biblioteca de primeras secuencias polinucleotídicas y segundas secuencias polinucleotídicas, usando el procedimiento descrito en el Ejemplo 1 de la presente memoria descriptiva, en la que una división en el nexo del ARNgi se realiza en cada posición de nucleótido de la secuencia que comprende el nexo nativo.

La biblioteca se ensaya después para determinar la capacidad de cada primera secuencia polinucleotídica de nexo dividido y su segunda secuencia polinucleotídica de nexo dividido afin para apoyar la escisión mediada por Cas9 de una diana de ADN bicatenario seleccionada siguiendo los procedimientos descritos en el Ejemplo 2 al Ejemplo 4 de la presente memoria descriptiva.

25 Se muestran en la FIG. 12 disposiciones de nexo dividido potenciales de secuencias de ARNcrtra conocidas de diversas especies. En la figura, la primera columna es un número de identificación para la especie bacteriana (véase FIG. 12, Tabla 7, Breve descripción de las figuras), la segunda columna es la secuencia del sn1-casARN/sn2-casARN. Se muestra un nexo dividido de un sn1-casARN/sn2-casARN de *S. pyogenes* de la presente invención para referencia (FIG. 12, fila 1).

30 Se sabe que una única especie puede tener más de un locus de CRISPR del mismo tipo o más de un locus de CRISPR de diferentes tipos (por ejemplo, tipo I y tipo II). Habitualmente solamente son utilizables elementos repetidos de un locus de CRISPR para identificar el elemento anti repetición (y por lo tanto las secuencias de ARNcrtra) contenido en el mismo locus de CRISPR.

35 Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el ensayo descrito en este ejemplo puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexo dividido.

Ejemplo 11

Exploración de múltiples sn-casARN que comprenden secuencias de unión a diana de ADN

40 Este ejemplo ilustra el uso de sn-casARN de la presente invención para modificar dianas presentes en ADN genómico humano y medir el nivel de actividad de escisión en esos sitios. Los sitios diana se seleccionan en primer lugar de ADN genómico y después se diseñan sn-casARN para dirigirse a esas secuencias seleccionadas. Después se llevan a cabo mediciones para determinar el nivel de escisión diana que ha tenido lugar. No todas las siguientes etapas son necesarias para cada exploración ni es necesario que el orden de las etapas sea como se presenta, y la exploración puede acoplarse a otros experimentos, o formar parte de un mayor experimento.

A. Seleccionar una región diana de ADN de ADN genómico

Identificación de todas las secuencias de MAP (por ejemplo 'NGG') en la región genómica seleccionada.

Identificación y selección de una o más secuencias de 20 nucleótidos de longitud (secuencia de ADN diana) que es 5' adyacente a secuencias de MAP.

50 Los criterios de selección pueden incluir, pero sin limitación: homología con otras regiones en el genoma; porcentaje de contenido G-C; temperatura de fusión; presencias de homopolímero en el espaciador; y otros criterios conocidos por un experto en la materia.

Adjuntar una secuencia de sn-casARN apropiada (por ejemplo, un sn1-casARN, como se ilustra en la FIG. 3B, con la

secuencia espaciadora retirada) al extremo 3' de la secuencia de ADN diana identificada (sn-casARN-sudADN (secuencia de unión a diana de ADN)). Una construcción de sn-casARN-sudADN es sintetizada habitualmente por un fabricante comercial o es producida como se describe en el Ejemplo 1 mediante transcripción *in vitro*.

5 Un sn-casARN-sudADN como se describe en el presente documento se usa con sn-casARN afín o afines para completar un sistema de sn-casARN (por ejemplo, un sistema de nexo dividido de dos polinucleótidos de sn1-casARN-sudADN/sn2-casARN) para su uso con una proteína Cas9 afín.

B. Determinación de porcentajes de escisión y especificidad

Se comparan porcentajes de escisión y especificidad *in vitro* asociados con un sistema de sn-casARN-sudADN/sn-casARN, por ejemplo, usando los ensayos de escisión de Cas9 del Ejemplo 3, de la siguiente manera:

10 (a) Si solamente se identifica o selecciona una única secuencia de ADN diana, se determina el porcentaje de escisión y la especificidad para la región diana de ADN. Si se desea, el porcentaje de escisión y/o la especificidad se alteran en experimentos adicionales usando procedimientos de la presente invención incluyendo, pero sin limitación, modificación del ARN, introducción de proteínas efectoras/secuencias de unión a proteína efectora o ligando/restos de unión a ligando.

15 (b) Los datos de porcentaje de escisión y datos de especificidad de sitio obtenidos de los ensayos de escisión se comparan entre diferentes ADN que comprenden la secuencia de unión a diana para identificar las secuencias de ADN diana que tienen el mejor porcentaje de escisión y la mayor especificidad. Los datos de porcentaje de escisión y datos de especificidad proporcionan criterios sobre los que basar elecciones para una diversidad de aplicaciones. Por ejemplo, en algunas situaciones la actividad del sn-casARN puede ser el factor más importante. En otras situaciones, la especificidad del sitio de escisión puede ser relativamente más importante que el porcentaje de escisión. Si se desea, el porcentaje de escisión y/o la especificidad se alteran en experimentos adicionales usando procedimientos de la presente invención incluyendo, pero sin limitación, modificación del ARN, introducción de proteínas efectoras/secuencias de unión a proteína efectora o ligando/restos de unión a ligando.

25 Opcionalmente, o en lugar del análisis *in vitro*, se comparan porcentajes de escisión y especificidad *in vivo* asociados con un sistema de sn-casARN-sudADN/sn-casARN, por ejemplo, usando análisis de secuenciación profunda para detección de modificaciones diana en células eucariotas del Ejemplo 4, de la siguiente manera:

30 (a) Si solamente se identifica una secuencia de ADN diana se determina el porcentaje de escisión y la especificidad para la región diana de ADN. Si se desea, el porcentaje de escisión y/o la especificidad se alteran en experimentos adicionales usando procedimientos de la presente invención incluyendo, pero sin limitación, modificación del ARN, introducción de proteínas efectoras/secuencias de unión a proteína efectora o ligando/restos de unión a ligando.

35 (b) Los datos de porcentaje de escisión y datos de especificidad de sitios obtenidos de los ensayos de escisión se compara entre diferentes ADN diana para identificar las secuencias de sn-casARN que dan como resultado el mayor porcentaje de escisión o ADN diana y la mayor especificidad para el ADN diana. Los datos de porcentaje de escisión y datos de especificidad proporcionan criterios sobre los que basar elecciones para una diversidad de aplicaciones. Por ejemplo, en algunas situaciones la actividad del sn-casARN puede ser el factor más importante. En otras situaciones, la especificidad del sitio de escisión puede ser relativamente más importante que el porcentaje de escisión. Si se desea, el porcentaje de escisión y/o la especificidad se alteran en experimentos adicionales usando procedimientos de la presente invención incluyendo, pero sin limitación, modificación del ARN, introducción de proteínas efectoras/secuencias de unión a proteína efectora o ligando/restos de unión a ligando.

45 Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, la exploración descrita en este ejemplo puede ser practicada por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos dividido.

Ejemplo 12

Exploración genómica funcional

50 Este ejemplo describe el uso de los polinucleótidos de nexos-dividido (sn-casPN) de la presente invención para identificación del papel funcional de genes utilizando un procedimiento de exploración funcional y datos de secuencia.

55 Un sistema de sn-casARN de dos partes (sn1-casARN y sn2-casARN) (véase, por ejemplo, FIG. 3B) se usa en una modificación de los procedimientos descritos en Shalem, y col., "Genome-scale CRISPR-Cas9 knockout screening in human cells," *Science*, 3, 343(6166), 84-87 (2014) y Zhou, y col., "High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells," *Nature*, 509, 487-491 (2014), que usaron un único ARN guía que tenía una secuencia continua. La exploración descrita en el presente documento se diseña en torno a la vulnerabilidad de la línea celular de melanoma A375 al fármaco vemurafenib; cuando se trata con vemurafenib, se

detiene el crecimiento de las células. Se transducen células A375 con una biblioteca de sn1-casARN y estas células se tratan posteriormente con vemerafenib. La inactivación por sn1-casARN de genes importantes para sensibilidad de A375 a vemerafenib se enriquecerá en la población de células supervivientes y puede secuenciarse e identificarse.

- 5 Se describen ejemplos de vectores, medios, condiciones de cultivo adecuados, etc. Un experto habitual en la materia entenderá modificaciones de estos componentes y condiciones a la vista de las enseñanzas de la presente memoria descriptiva.

A. Biblioteca lentivírica y construcciones de Cas9

10 Se genera una biblioteca vírica de sn1-casARN sintetizando oligonucleótidos que contienen las secuencias espaciadoras diseñadas adjuntas a secuencias marcadoras universales para clonar en un plásmido de transferencia para producción de lentivirus (por ejemplo pD2107-CMV -DNA 2.0, Menlo Park, CA). Se sintetizan bibliotecas de oligonucleótidos en micromatrices programables y se escinden de la micromatriz por el fabricante de matrices (por ejemplo Agilent Technologies, Santa Clara, CA). Se amplifican oligonucleótidos de longitud completa por PCR usando polimerasa Q5 (NEB) y cebadores diseñados para amplificar ADN que contiene las secuencias de marcadores universales. Se lleva a cabo clonación en el vector de transferencia usando técnicas convencionales conocidas por los expertos en la materia. Un ejemplo incluye digerir el vector con una enzima de restricción de tipo II (por ejemplo BsbI) para revelar salientes monocatenarios, tratar con fosfatasa alcalina (Fermentas) y purificar el vector cortado de no cortado por purificación en gel. Se digieren bibliotecas de oligonucleótidos con una enzima de restricción para revelar extremos compatibles y se ligan en el vector usando ADN ligasa (Fermentas).

20 El vector de transferencia puede incluir un gen de Cas9 de *S. pyogenes* con codones optimizados para seres humanos marcado en el extremo N y el extremo C con una señal de localización nuclear de SV40 bajo el control del promotor corto de factor-1 α de elongación (EFS). Esta secuencia de NLS-Cas9-NLS se une con un péptido de autoescisión 2A y un marcador de selección adecuado para células de mamífero (es decir puromicina).

25 Como alternativa, puede suministrarse Cas9 a las células en un vector vírico separado o pueden generarse líneas celulares estables que expresan Cas9 de forma constitutiva. Después pueden usarse bibliotecas de sn1-casARN expresado en vector vírico para transducir las líneas celulares que expresan Cas9.

B. Producción y purificación de lentivirus

30 Se siembran células HEK293T a aproximadamente 40 % de confluencia 24 horas antes de la transfección en DMEM (Life Technologies, Grand Island, NY) complementado con suero bovino fetal 10 % (FBS). Se transfieren células a suero reducido OptiMEM (Life Technologies, Grand Island, NY) y se transfectan usando Lipofectamine 2000 y reactivo Plus según las instrucciones del fabricante. Para transfección, el vector de transferencia lentivírica se combina con plásmidos para empaquetamiento lentivírico tal como el sistema de empaquetamiento Lenti-X™ HTX (Takara Clontech, Mountain View, CA) según las instrucciones del fabricante.

35 Después de 60 horas, el medio se retira y centrifuga a 3000 rpm para retirar residuos celulares. El sobrenadante se filtra a través de una membrana de baja unión a proteínas de 0,45 μ m (por ejemplo Millipore Steriflip HV/PVDF). La biblioteca agrupada puede concentrarse por ultracentrifugación y después resuspenderse en DMEM complementado con FBS 10 % y BSA 1 % (Sigma-Aldrich, St. Louis, MO).

C. Cultivo celular

40 Se obtienen células A375 (ATCC CRL-1619) de la Colección Americana de Cultivos Tipo (Manassas, VA) y se cultivan en medio R8758 (Sigma-Aldrich, St. Louis, MO), complementado con FBS 10 % (Life Technologies, Grand Island, NY), Penicilina-Estreptomicina 1 % (Sigma-Aldrich, St. Louis, MO), HEPES 20 mM (Sigma-Aldrich, St. Louis, MO).

D. Transducción lentivírica

45 Se determina la multiplicidad de infección (MDI) para la biblioteca de vectores víricos usando procedimientos convencionales basándose en la transducción de células con volúmenes de virus predeterminados. Se siembran aproximadamente 3×10^6 células A375 en una placa de 12 pocillos en medio apropiado complementado con polibreno 8 μ g/ml (Sigma-Aldrich, St. Louis, MO). Las células se mezclan con el volumen de virus predeterminado para identificar una multiplicidad de infección (MDI) entre 0,3-0,5. Las células sembradas se centrifugan a 2.000 rpm durante 2 horas a 37 °C, después de lo cual el medio se aspira y se añade medio nuevo para cada tipo celular, sin polibreno. Las células se incuban durante 24 horas a 37 °C, CO₂ 5 %. Se incluye un control no transducido.

50 Después de 24 horas, las células se desprenden y se cuentan, se vuelven a sembrar aproximadamente $2,5 \times 10^6$ células tanto en un "pocillo de selección" como en un "pocillo no de selección". Los pocillos de selección se ponen bajo selección específica para la construcción de biblioteca lentivírica (es decir puromicina, Sigma-Aldrich, St. Louis, MO). Los pocillos no de selección no se tratan con puromicina. Las células se incuban hasta que no quedan células supervivientes colocadas bajo selección en el control no transducido. Las células se cuentan, y el número de células

en "pocillos de selección" dividido por el número de células en los "pocillos no de selección" correspondientes multiplicado por 100 produce la MDI, siendo una MDI cercana a 0,4 el valor ideal.

E. Exploración de resistencia a fármacos

5 Las células se siembran en placas en pocillos de 2×10^6 células por pocillo para cada condición para ensayar. Las células en cada pocillo se transducen con 10 μ l de la biblioteca para alcanzar una eficacia de transducción de 30 % (mínimo de 3-400 células por clon en la biblioteca). Se añade puromicina a los pocillos 24 horas después de la transducción y las células se mantienen durante 7 días. Las células se dividen en condiciones farmacológicas por duplicado con un mínimo de 2×10^7 células por pocillo repetido. Un pocillo se complementa con compuesto farmacológico 2 μ M (por ejemplo PLX4032, Thermo Fisher Scientific, South San Francisco, CA) y el otro con DMSO (Thermo Fisher Scientific, South San Francisco, CA). Las células se incuban a 37 °C, CO₂ 5 % durante 14 días, y se pasan cada 2-3 días a medio nuevo, complementado con PLX4032 o DMSO según sea apropiado. Después de 14 días, se prepara ADN genómico (ADNg) a partir de células usando la solución de extracción de ADN QuickExtract (Illumina, San Diego, CA) según las instrucciones del fabricante.

F. Secuenciación de ADNg

15 Se diseñan cebadores de PCR para amplificar secuencias diana de sn1-casPN lentivíricas de ADN genómico. Usando ADNg aislado, se realiza una primera PCR usando ADN polimerasa de fusión Herculase II (Agilent, Santa Clara, CA) con cebadores que comprenden una secuencia adaptadora y una secuencia específica del casete de sn1-casPN lentivírico. Se realiza una segunda PCR usando los amplicones del primer ciclo como molde a 1/20 del volumen del volumen de reacción de la segunda PCR. La segunda PCR usa un segundo conjunto de cebadores que comprende: secuencia complementaria de la secuencia adaptadora universal del primer par de cebadores, una secuencia de índice de código de barras única de cada muestra y una secuencia adaptadora de celda de flujo. Se agrupan reacciones de PCR para asegurar una cobertura de secuenciación 300x de cada muestra transducida. Se analizan reacciones de PCR agrupadas en un gel de TBE 2 %, las bandas de tamaños de amplicones esperados se purifican en gel usando el kit de extracción en gel QIAEX II (Qiagen, Venlo, Países Bajos). Las concentraciones de amplicones purificados se evalúan usando el kit de ensayo de ADNbc BR y sistema Qubit (Life Technologies, Grand Island, NY) y la calidad de la biblioteca se determina usando el DNA100Chip Agilent y sistema Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA). Se secuencian bibliotecas agrupadas en un MiSeq 2500 (Illumina, San Diego, CA).

G. Procesamiento y análisis de datos de secuenciación

30 Se procesan lecturas de secuenciación en bruto para contener solamente la secuencia del casete de sn1-casPN. Las lecturas de sn1-casPN se alinean con las secuencias diana contenidas en la biblioteca de exploración lentivírica y se cuenta el número de lecturas para cada secuencia diana única. Las lecturas contadas por secuencia diana se normalizan dividiendo las lecturas por diana por lecturas alineadas totales para todas las dianas en la muestra y multiplicando por 10^6 y añadiendo 1.

35 Las lecturas diana normalizadas identificadas en muestras tratadas con fármacos se comparan con lecturas de dianas normalizadas identificadas en las muestras tratadas con control de DMSO. Las dianas con alto recuento de lecturas presente en la muestra tratada con fármaco que están ausentes o reducidos en las muestras tratadas con control de DMSO pueden evaluarse adicionalmente como genes candidatos importantes en la resistencia a tratamiento farmacológico.

40 Pueden usarse otras exploraciones genómicas funcionales usando una biblioteca de sn1-casARN y el procedimiento de exploración perfilado en el presente documento para identificar genes candidatos importantes en esas exploraciones.

Este procedimiento proporciona datos para verificar que los sistemas de sn1-casARN/sn2-casARN/proteína Cas9 de la presente invención pueden usarse en la exploración funcional para consultar la función génica en una escala genómica.

45 Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, la exploración descrita en este ejemplo puede ser practicada por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos dividido.

Ejemplo 13

50 Represión/activación

Este ejemplo describe el uso de los polinucleótidos de nexos dividido (sn-casPN) de la presente invención para la represión o activación de genes endógenos en células humanas.

Un sistema de sn-casARN de dos partes (sn1-casARN y sn2-casARN) (véase, por ejemplo, FIG. 3B) se usa en una modificación de los procedimientos descritos en Gilbert y col., "CRISPR-Mediated Modular RNA-Guided Regulation

of Transcription in Eukaryotes," Cell, 154(2), 442-51, doi: 10.1016/j.cell.2013.06.044 (2013), que usaron un único ARN guía que tenía una secuencia continua.

5 Se describen ejemplos de vectores, medios, condiciones de cultivo adecuados, etc. Un experto habitual en la materia entenderá modificaciones de estos componentes y condiciones a la vista de las enseñanzas de la presente memoria descriptiva.

A. Construcciones de activador y represor de dCas9

10 El Cas9 de *S. pyogenes* deficiente en nucleasa (dCas9) con mutación D10A y H840A tiene codones optimizados para expresión en células de mamífero y marcadas en el extremo C con una señal de localización nuclear SV40 y el dominio de represión de caja asociada a Kruppel (KRAB) (dCas9-KRAB) o cuatro copias del activador de la transcripción VP16 (dCas9-VP64). Tanto el dCas9-KRAB como el dCas9-VP64 se insertan en un vector adyacente a un promotor de mamífero adecuado, tal como el promotor de citomegalovirus (CMV). Uno de dichos vectores, pJ607-03 (DNA2.0, Menlo Park, CA), está disponible en el mercado.

B. Construcción de sn-casPN

15 La secuencia de sn1-casARN-CD71 comprende una secuencia espaciadora de 20 nucleótidos que se dirige a la región no traducida (UTR) cadena arriba del receptor de transferrina CD71. La secuencia de sn1-casARN-CD71 se ensambla en un vector adecuado que también comprende la secuencia de sn2-casARN independiente. Cada secuencia está bajo el control independiente por un promotor U6 humano que dirige la transcripción por ARN polimerasa III. Una cadena principal de vector adecuado para la expresión de secuencias de sn1-casARN y sn2-casARN es el vector pRSFDuet-1 (Novagen, Merck, Darmstadt, Alemania).

20 C. Cultivo celular

Se obtienen células HeLa (ATCC CCL-2) de la Colección Americana de Cultivos Tipo (Manassas, VA) y se cultivan en medio de Eagle modificado por Dulbecco (DMEM, Life Technologies, Grand Island, NY), complementado con FBS 10 % (Life Technologies, Grand Island, NY), Penicilina-Estreptomicina 1 % (Sigma-Aldrich, St. Louis, MO), glutamina 2 mM (Life Technologies, Grand Island, NY) y se cultivan a 37 °C, CO2 5 %.

25 D. Transfección y separación FACS

30 Se transfectan de forma transitoria células HeLa con plásmido que contiene Cas9 y vector sn1-casARN-CD71 del mismo peso usando TransIT-LTI (Mirus, Madison, WI). Se incluye un control no transfectado. 72 horas después de la transfección las células se tripsinizan (Life Technologies, Grand Island, NY) y se disocian con EDTA-PBS 10 nM (Lonza, Allendale, NJ). Las células se incuban en los presentes de un anticuerpo específico de anti CD71 humano conjugado con un fluoróforo FITC (sBioscience, San Diego, CA) en tampón de tinción de citometría de flujo (eBioscience, San Diego, CA). Se realiza separación de células activadas por fluorescencia (FACS) de células transfectadas usando el láser azul (excitación a 488 nm) y el citómetro de flujo LSR II (BD Biosciences, Franklin Lakes, NJ) para detección del anticuerpo de CD71-FITC.

35 La activación de expresión de CD71 en muestras transfectadas con dCas9-VP64 se miden mediante el aumento de fluorescencia detectada (u.a. Log₁₀) en comparación con la fluorescencia medida de una población de control no transfectada de células HeLa como se detecta mediante separación FACS.

La represión de la expresión de CD71 en muestras transfectadas con dCas9-KRAB se mide por la reducción de fluorescencia detectada (u.a. Log₁₀) en comparación con la fluorescencia medida de una población de control no transfectada de células HeLa como se detecta mediante separación FACS.

40 Otros genes se activan o reprimen de forma similar usando el sn-casPN de la presente invención y los procedimientos perfilados aquí. Como resulta evidente para los expertos en la materia, pueden fusionarse otros dominios de activación y represión con un dCas9 para conseguir un resultado similar a los procedimientos descritos aquí.

45 Este procedimiento proporciona datos para verificar que los sistemas de sn1-casARN/sn2-casARN/proteína Cas9 de la presente invención pueden usarse en la activación o represión de genes endógenos.

50 Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, los ensayos de represión/activación descritos en este ejemplo pueden ser practicados por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos dividido.

Ejemplo 14

Modificación de células CHO para aplicación industrial

Este ejemplo describe uso de los polinucleótidos de nexos dividido (sn-casPN) de la presente invención para modificar el genoma de una célula CHO. Este ejemplo también contiene un perfil para la validación de secuencias y selección de células modificadas por sn-casPN para usos futuros en aplicaciones industriales (es decir, producción de anticuerpos).

- 5 Un sistema de sn-casARN de dos partes (sn1-casARN y sn2-casARN) (véase, por ejemplo, FIG. 3B) se usa en una modificación del procedimiento descrito en Ronda y col., "Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool," *Biotechnology and Bioengineering*, 111(8), 1604-1616, (2014), que usaron un único ARN guía que tenía una secuencia continua.

- 10 Se describen ejemplos de vectores, medios, condiciones de cultivo adecuados, etc. Un experto habitual en la materia entenderá modificaciones de estos componentes y condiciones a la vista de las enseñanzas de la presente memoria descriptiva.

A. Construcción de plásmidos

- 15 Se optimizan los codones de la secuencia de Cas9 de *S. pyogenes* para su expresión en células CHO y se marca en el extremo C con una señal de localización nuclear de SV40 y se inserta en un vector adyacente a un promotor de mamífero adecuado, tal como el promotor de citomegalovirus (CMV). Uno de dichos vectores, pJ607-03 (DNA2.0, Menlo Park, CA), está disponible en el mercado.

- 20 La secuencia de sn1-casARN-FUT8 comprende una secuencia espaciadora de FUT8 de 20 nucleótidos. La secuencia de sn1-casARN-FUT8 se ensambla en un vector adecuado que también comprende la secuencia de sn2-casARN independiente. Cada secuencia está bajo el control independiente por un promotor U6 que dirige la transcripción por ARN polimerasa III. Una cadena principal de vector adecuado para la expresión de secuencias de sn1-casARN y sn2-casARN es el vector pRSFDuet-1 (Novagen, Merck, Darmstadt, Alemania).

B. Cultivo celular

- 25 Se obtienen células CHO-K1 de la Colección Americana de Cultivos Tipo (Manassas, VA) y se cultivan en medio CHO-K1 F-12K (Colección Americana de Cultivos Tipo, Manassas, VA), FBS 10 % (Life Technologies, Grand Island, NY) y Penicilina-Estreptomicina 1 % (Sigma-Aldrich, St. Louis, MO). Se transfectan células CHO-K1 con plásmido que contiene Cas9 y vector que comprende sn1-casARN-FUT8/sn2-casARN del mismo peso usando el dispositivo Nucleofector 2b (Lonza, Allendale, NJ) y el kit de Nucleofector V de línea celular Amaxa (Lonza, Allendale, NJ) según las recomendaciones del fabricante. Las células se incuban a 30 °C en CO2 5 % durante las primeras 24 horas y después se movieron a 37 °C, CO2 5 % durante otro periodo de 24 horas.

- 30 C. Selección de células con inactivación de FUT8

- 35 Se seleccionan células con inactivación de FUT8 mediante la adición de aglutinina de *Lens culinaris* 50 µg/ml (LCA, Vector Laboratories, Burlingame, CA) cinco días después de la transfección del vector Cas9 y el sn1-casARN-FUT8. Las células se someten a 7 días de selección en LCA, las células se pasan y se añade medio nuevo, con LCA, cada 2-3 días o según sea necesario. Solamente células que tengan alteraciones en el gen de Fut8, provocadas por el sistema de sn-casPN, tendrán resistencia a la LCA.

Para confirmar la inactivación de FUT8, se vuelven a sembrar células seleccionadas en medio completo sin LCA y se incuban durante 48 horas. Después de volver a sembrar, se prepara ADN genómico (ADNg) usando la solución de extracción de ADN QuickExtract (Illumina, San Diego, CA) según las instrucciones del fabricante.

D. Validación de secuencia de modificación de Cas9 y construcción de biblioteca de Miseq

- 40 Se diseñan amplicones de secuenciación de entre 150 pb-200 pb para abarcar el sitio diana de sn1-casARN-FUT8. Usando ADNg aislado previamente, se realiza una primera PCR usando ADN polimerasa de fusión Herculase II (Agilent, Santa Clara, CA) con cebadores que comprenden una secuencia adaptadora y una secuencia específica de la región que flanquea el sitio diana de FUT8. Se realiza una segunda PCR usando los amplicones del primer ciclo de PCR como molde a 1/20 del volumen del volumen de reacción de PCR. La segunda PCR usa un segundo conjunto de cebadores que comprende: secuencia complementaria de la secuencia adaptadora del primer par de cebadores, una secuencia de índice de código de barras única de cada muestra y una secuencia adaptadora de celda de flujo. Se agrupan y analizan amplicones en un gel de TBE 2 %, las bandas de tamaños de amplicones esperados se purifican en gel usando el kit de extracción en gel QIAEX II (Qiagen, Venlo, Países Bajos). Las concentraciones de amplicones purificados se evalúan usando el kit de ensayo de ADN bicatenario BR y sistema Qubit (Life Technologies, Grand Island, NY) y la calidad de la biblioteca se determina usando el DNA100Chip Agilent y sistema Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA). Después de validación de la calidad de la biblioteca, la biblioteca se secuencia en un secuenciador de sobremesa MiSeq (Illumina, San Diego, CA) con el kit de reactivo MiSeq v2 (300 ciclos, Illumina, San Diego, CA) según las instrucciones del fabricante para lecturas de extremos emparejados de 151 pb.

- 55 E. Análisis de datos de secuenciación profunda

La identidad de productos en los datos de secuenciación se analiza basándose en la secuencia de código de barras de índice adaptada al amplicón en el segundo ciclo de PCR. Se usa un *script* informático para procesar los datos de MiSeq ejecutando las siguientes tareas:

- 5 1. Unión de lecturas de extremos emparejados con la ayuda de fastq-join (Aronesty 2011: code.google.com/p/ea-utils)
2. Validación de las lecturas de secuencias para secuencias de cebadores apropiadas que están presentes en los extremos tanto 5' como 3' de la secuencia de lectura usando fastx_barcode_splitter (hannonlab.cshl.edu/fastx_toolkit/index.html). Se descartan lecturas que carecen de secuencias de cebadores correctas en ambos extremos.
- 10 3. Comparación de secuencias de lectura con secuencia de FUT8 de tipo silvestre esperada, las secuencias de lecturas idénticas se clasifican como poseedoras de la misma modificación de insup.

Otros loci cromosómicos en células CHO se modifican de forma similar por selección de secuencia espaciadora apropiada para el sn1-casARN. La selección es específica de una diana génica específica y el procedimiento perfilado en este ejemplo es fácilmente modificable por un experto habitual en la materia para otras dianas génicas.

- 15 Este procedimiento proporciona datos para verificar los sistemas de sn1-casARN/sn2-casARN/proteína Cas9 de la presente invención proporciona actividad endonucleasa dirigida a ARN específica de secuencia en loci diana en células CHO y perfila los procedimientos para selección de dichas células CHO modificadas para su uso continuado.

20 Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, el ensayo descrito en este ejemplo puede ser practicado por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexa dividido.

Ejemplo 15

Ingeniería genómica en *Saccharomyces cerevisiae*

- 25 Este ejemplo describe uso de los polinucleótidos de nexa dividido (sn-casPN) de la presente invención para modificar el genoma de la levadura *S. cerevisiae*.

Un sistema de sn-casARN de dos partes (sn1-casARN y sn2-casARN) (véase, por ejemplo, FIG. 3B) se usa en una modificación del procedimiento de DiCarlo, y col., "Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems," *Nucleic Acids Res.*, 41(7), 4336-43 (2013), que usaron un único ARN guía que tenía una secuencia continua.

- 30 Se describen ejemplos de vectores, medios, condiciones de cultivo adecuados, etc. Un experto habitual en la materia entenderá modificaciones de estos componentes y condiciones a la vista de las enseñanzas de la presente memoria descriptiva.

A. Mutaciones genómicas específicas de sitio

- 35 Un gen de Cas9 de *Streptococcus pyogenes*, con codones optimizados para su expresión en células de levadura, se marca en el extremo C con una señal de localización nuclear de SV40 y se inserta en un vector de bajo número de copias adyacente a un promotor inducible, por ejemplo, secuencias promotoras de Gall. El vector también contiene un marcador seleccionable, tal como un marcador seleccionable de URA3. Uno de dichos vectores, p415-Gall-Cas9-CYC1t (Addgene, Cambridge, MA), está disponible en el mercado. La expresión del gen de Cas9 está bajo el control inducible del promotor de Gall.

- 40 La secuencia de sn1-casARN-CAN1.Y comprende una secuencia espaciadora de CAN1.Y de 20 nucleótidos. El sn1-casARN-CAN1.Z comprende una secuencia espaciadora de CAN1.Z de 20 nucleótidos. Se ensamblan casetes de expresión que comprenden cada uno sn1-casARN, el promotor SNR52 y secuencia flanqueante de SUP4 30. Cada casete de expresión se ensambla en un vector que comprende un origen de replicación de 2 micrómetros y un marcador seleccionable, por ejemplo, p426. Se insertan secuencias de ADN que codifican casete de expresión de sn1-casARN en un vector que contiene un marcador seleccionable de HIS3. Se insertan secuencias de ADN que codifican casete de expresión de sn2-casARN en un vector que contiene un marcador seleccionable de LEU2. Un vector de cadena principal adecuado para las secuencias codificantes de sn-casARN es p426 GPD (Colección Americana de Cultivos Tipo, Manassas, VA), en el que las secuencias codificantes de URA3 se mutan o suprimen y se inserta el marcador seleccionable apropiado. La expresión de las secuencias de sn-casARN está bajo el control constitutivo del promotor SNR52 que dirige la transcripción por ARN polimerasa III.
- 50

- El vector de Cas9 y cada par de vectores de sn1-casARN/sn2-casARN se transforman usando procedimientos convencionales en ATCC 200895 (MATa his3delta200 leu2delta0 met15delta0 trpldelta63 ura3delta0) (Colección Americana de Cultivos Tipo, Manassas, VA) y la presencia de los vectores se selecciona para usar los medios de retirada de SC sin uracilo, histidina o leucina. También se construyen cepas de levadura de control negativo mediante transformación de los vectores individuales que comprenden sn1-casARN-CAN1.Y, sn1-casARN-CAN1.Z,
- 55

sn2-casARN y Cas9 en ATCC 200895. Se usan medios de selección apropiados para cada vector.

5 Se cultivan células que comprenden Cas9 y sn1-casARN/sn2-casARN en medio de retirada de SC sin uracilo, histidina y leucina, y que contiene galactosa 2 % y rafinosa 1 %. Las células se cultivan durante aproximadamente 16 horas, se sedimentan y se siembran en placas YPAD, de SC-uracilo-histidina-leucina que contienen L-canavanina 60 mg/ml (Sigma-Aldrich, St. Louis, MO) y lisina-SC que contiene tialisina 100 mg/ml (S-2-aminoetil-1-cisteína, Sigma-Aldrich, St. Louis, MO). Se siembran aproximadamente 10^7 - 10^8 células en medio que contiene canavanina y tialisina, y las células se diluyen apropiadamente para siembra en placas en medio rico.

10 La relación del recuento de colonias en placas de canavanina o tialisina dividida por el recuento de colonias en placas de medios ricos (YPAD) para cada cultivo se usa como una medida de la frecuencia de mutación. Se cultivan y siembran en placas de forma similar cepas de control negativo.

Para controlar un fenotipo mutante de todo el genoma potencial, la frecuencia de mutación del gen de LYP1 endógeno no diana, una lisina permeasa, se supervisa seleccionando mutantes de *lyp1* usando un análogo de lisina tóxico, tialisina.

15 Los genes de LYP1 y CAN1 están en cromosomas separados. En consecuencia, la frecuencia de mutación local en cada locus debería ser independiente en ausencia de un mutador de todo el genoma.

El sn 1-casARN-CAN1.Y/sn2-casARN dirige la actividad de endonucleasa Cas9 a un sitio diana localizado 207 pb cadena abajo del codón de inicio del gen de CAN1. El sn1-casARN-CAN1.Z/sn2-casARN dirige la actividad endonucleasa de Cas9 a un sitio diana localizado 58 pb cadena abajo del codón de inicio de ATG del gen de CAN1.

20 Cuando la expresión de Cas9 es inducida por galactosa, una reducción de la viabilidad celular en placas de SC-uracilo-histidina-leucina que contienen L-canavanina 60 mg/ml frente a medio YPAD indica una mayor frecuencia de mutación en el gen de CAN1. La tasa de mutación en el gen de LYP1 proporciona un indicio de la tasa de mutación de fondo. Cuando la tasa de mutación del gen de LYP1 permanece constante entre todas las cepas sugiere que el sistema de sn1-casARN/sn2-casARN y Cas9 no induce mutaciones aleatorias en todo el genoma. Para validar adicionalmente que las mutaciones están provocadas por el sistema de sn1-casARN/sn2-casARN/proteína Cas9, el gen de CAN1 puede aislarse y secuenciarse a partir de poblaciones resistentes a canavanina. Las secuencias se alinean después para identificar la localización y los tipos de mutaciones en el gen de CAN1 en relación con la secuencia de unión a diana (es decir, la secuencia espaciadora) presente en el sn1-casARN.

Otros loci cromosómicos en *S. cerevisiae* son de forma similar diana para modificación por selección de secuencias espaciadoras para sn1-casARN.

30 Este análisis proporciona datos para verificar que los sistemas de sn1-casARN/sn2-casARN/proteína Cas9 de la presente invención proporcionan actividad endonucleasa dirigida por ARN específica en loci genómicos endógenos diana en levadura.

B. Recombinación homóloga específica de sitio con ADN donante

35 Se amplifica por PCR una secuencia oligonucleotídica de KanMX con ramas de homología de 50 pb para el locus de CAN1 a partir del plásmido pFA6a-KanMX6, que se usa habitualmente para la creación de inactivaciones génicas en levadura. El KanMX se usa como un ADN donante. El oligonucleótido de KanMX confiere resistencia a G418 y se diseña para alterar la secuencia de MAP asociada con CAN1.Y. Tras la integración, este ADN donante da como resultado resistencia a canavanina y resistencia a G418.

40 Se cultivan células que contienen los vectores de expresión de sn1-casARN-CAN1.Y, sn2-casARN, Cas9 hasta saturación en medio con retirada de SC sin uracilo, histidina y leucina. Este cultivo se usa para inocular medio de SC líquido sin uracilo, histidina y leucina y el cultivo se deja crecer hasta aproximadamente DO600 de 1,8. Las células se recogen mediante centrifugación y se transforman oligonucleótidos donantes en las células por electroporación. Se transfieren células electroporadas a medio SC-ura-his-leu que contiene galactosa 2 % y rafinosa 1 % y se cultivan durante aproximadamente 12 horas. Las cepas de control negativo se tratan de forma similar pero no se proporciona oligonucleótido donante.

45 Se siembran aproximadamente 10^6 - 10^7 células en placas en medio selectivo y las células se diluyen apropiadamente en medio rico. Se cultivan y siembran en placas de forma similar cepas de control negativo.

50 Las colonias que contienen los plásmidos se siembran en repeticiones en placas en medio de canavanina así como medio rico con G418 para seleccionar el acontecimiento de integración de KanMX. La relación de recuento de colonias en placas selectivas (es decir, colonias que son resistentes tanto a canavanina como a G418) con respecto a recuento de colonias en placas ricas se usa como una medida de la frecuencia de corrección que sugiere recombinación homóloga de las secuencias de KanMX en el sitio de escisión dirigida por sn1-casARN/sn2-casARN. Para validar adicionalmente que los acontecimientos de integración están dirigidos por el sistema de sn1-casARN/sn2-casARN/proteína Cas9, el gen de CAN1 que incluye las secuencias de KanMX puede aislarse y secuenciarse a partir de poblaciones resistentes a canavanina/G418. Las secuencias se alinean después para

55

identificar la localización y los tipos de inserciones en el gen de CAN1 en relación con la secuencia de unión a diana (es decir, la secuencia espaciadora) presente en el sn1-casARN.

Otros loci cromosómicos en *S. cerevisiae* son de forma similar diana para modificación por selección de secuencias espaciadoras apropiadas para sn1-casARN y oligonucleótidos donantes. Pueden introducirse genes funcionales en el genoma de *S. cerevisiae* sin alteración de genes endógenos. Además, puede usarse introducción de marcadores seleccionables en genes diana endógenos para proporcionar mutaciones de inactivación seleccionables de los genes diana.

Este análisis puede proporcionar datos para verificar que los sistemas de sn1-casARN/sn2-casARN/proteína Cas9 de la presente invención proporcionan actividad endonucleasa dirigida por ARN específica en loci genómicos endógenos diana en levadura y pueden estimular acontecimientos de recombinación homólogos en dichos loci usando ADN donante.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, los procedimientos descritos en este ejemplo pueden ser practicados por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos dividido.

Ejemplo 16

Mutagénesis diana en *Zea mays*

Este ejemplo describe uso de los polinucleótidos de nexos dividido (sn-casPN) de la presente invención para crear modificaciones genómicas en plantas. Aunque se describe un sistema polinucleotídico de sn-casARN de dos componentes, también pueden usarse otras realizaciones de la presente invención (por ejemplo, un sistema polinucleotídico de sn-casARN de tres componentes).

Se usa un sistema de sn-casARN (sn1-casARN, sn-2-casARN y sn3-casARN) (véase, por ejemplo, FIG. 3A) en una modificación del procedimiento de Cigan, A. M., y col., "Genome modification using guide polynucleotide/cas endonuclease systems and methods of use," publicación de solicitud de patente de los Estados Unidos N.º 2015-0059010, publicada el 26 de febrero de 2015, que usó ARN guía que tenían cada uno una secuencia continua.

Se describen ejemplos de vectores, medios, condiciones de cultivo adecuados, etc. Un experto habitual en la materia entenderá modificaciones de estos componentes y condiciones a la vista de las enseñanzas de la presente memoria descriptiva.

A. Casetes de expresión

Se optimizan los codones del gen de Cas9 *Streptococcus pyogenes* M1 GAS (SF370) para maíz según técnicas convencionales conocidas en este campo. El intrón de ST-LS1 de patata se introduce para eliminar su expresión en *E. coli* y *Agrobacterium*. La localización nuclear de la proteína Cas9 en células de maíz es facilitada por señales de localización nuclear de endonucleasa de límite de ADN-T de VirD2 de virus de simio 40 (SV40) monopartito y *Agrobacterium tumefaciens* bipartito incorporadas en los extremos amino y carboxilo de la fase abierta de lectura de Cas9, respectivamente. El gen de Cas9 se unió operativamente a un promotor de maíz constitutivo (por ejemplo un promotor de ubiquitina vegetal) o regulado por técnicas biológicas moleculares convencionales.

Los casetes de expresión para la expresión del sn1-casARN, sn2-casARN y sn3-casARN utilizan el promotor de la polimerasa III U6 del maíz (5' de cada secuencia codificante de sn-casARN) y terminador de la polimerasa III U6 del maíz (3' de cada secuencia codificante de sn-casARN) unidos operativamente con secuencias codificantes de ADN de sn-casARN usando técnicas de biología molecular convencionales para crear casetes de expresión de sn-casARN. Como se muestra en la FIG. 3A, sn3-casARN comprende una complementariedad de 20 regiones espaciadoras con la diana de ADN (dominio VT). Se selecciona una región diana cadena arriba de una secuencia de MAP para reconocimiento y escisión de sitio diana.

Los casetes de expresión para la proteína Cas9 y sn-casARN pueden colocarse en vectores de cadena principal adecuados (por ejemplo, como se describe en Belhaj, K., y col., "Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system," *Plant Methods*, 9(1), 39 (2013); Weber E., y col., "A Modular Cloning System for Standardized Assembly of Multigene Constructs," *PLoS ONE* 6(2), e16765 (2011)) usando técnicas de biología molecular convencionales.

B. Generación de mutaciones

Se dirigen tres secuencias diana genómicas de maíz diferentes para escisión usando el sistema de sn1-casARN/sn2-casARN/sn3-casARN/proteína Cas9. Las tres secuencias diana se localizan en el locus de LIG (aproximadamente 600 pb cadena arriba del codón de inicio del gen de Liguleless 1) y se examinan mediante secuenciación profunda para determinar la presencia de mutaciones. Las secuencias espaciadoras para cada sitio diana (LIGCas-1, LIGCas-2 y LIGCas3) son como se describen en la publicación de solicitud de patente de los

Estados Unidos N.º 2015-0059010, publicada el 26 de febrero de 2015 (véase dominios VT complementarios de la cadena antisentido de las secuencias diana genómicas del maíz enumeradas en la Tabla 1 de la publicación de solicitud de patente de los Estados Unidos N.º 2015-0059010, publicada el 26 de febrero de 2015). Los sn-casARN resultantes son los siguientes: sn1-casARN/sn2-casARN/sn3-casARN-LIGCas-1; sn1-casARN/sn2-casARN/sn3-casARN-LIGCas-2; y sn1-casARN/sn2-casARN/sn3-casARN-LIGCas-3.

Se suministran conjuntamente casetes de expresión que comprenden los sistemas de sn-casARN de tres componentes y casetes de expresión de Cas9 a 60-90 embriones de maíz inmaduros Hi-II mediante suministro mediado por partículas. Se transforman embriones de maíz Hi-II con los casetes de expresión de ARN guía largo y Cas9 (como se describe en la publicación de solicitud de patente de los Estados Unidos N.º 2015-0082478, publicada el 19 de marzo de 2015) que se dirigen al sitio diana genómico de LIGCas-3 para escisión para proporcionar un control positivo. Los embriones de maíz Hi-II transformados solamente con el casete de expresión de Cas9 proporciona un control negativo.

Las mazorcas de maíz se descascarillan y la superficie se esteriliza y se aclara dos veces con agua estéril. Los embriones inmaduros se aíslan y se colocan con el eje del embrión hacia abajo (escutelo hacia arriba), 25 embriones por placa, en medio 560Y durante 4 horas y después se alinean en la zona diana de 2,5 cm en preparación para el bombardeo.

Los vectores que comprenden los sistemas de sn1-casARN/sn2-casARN/sn3-casARN/proteína Cas9 se bombardean conjuntamente con vectores que contienen los genes del desarrollo ODP2 (proteína de desarrollo de óvulos 2, un factor de transcripción de dominio AP2; véase, por ejemplo, la solicitud publicada de patente de los Estados Unidos n.º 2009-0328252, publicada el 31 de diciembre de 2009; la solicitud publicada de patente de los Estados Unidos n.º 2011-0167516, publicada el 7 de julio de 2011).

Para cada sistema de sn1-casARN/sn2-casARN/sn3-casARN/proteína Cas9, los vectores correspondientes se precipitan en microgránulos de oro de 0,6 µm (diámetro promedio) usando un reactivo de transfección de lípidos catiónicos hidrosolubles. Se prepara solución de ADN en hielo usando vectores de sn1-casARN/sn2-casARN/sn3-casARN/Cas9 y plásmidos que contienen los genes del desarrollo ODP2 y Wushel. Se añaden partículas de oro preparadas al ADN premezclado. El reactivo de transfección de lípidos catiónicos hidrosolubles se añade en agua y se mezcla cuidadosamente. Las partículas de oro se sedimentan en una microcentrífuga y se retira el sobrenadante. El sedimento resultante se aclara cuidadosamente con etanol (EtOH) sin resuspender el sedimento y el aclarado de EtOH se retira cuidadosamente. Se añade EtOH al 100 % y las partículas se resuspenden por aplicación breve de ultrasonidos. Después, la mezcla se aplica puntualmente al centro de cada macrovehículo y se deja secar aproximadamente 2 minutos antes del bombardeo (Kikkert J.R., y col., "Stable transformation of plant cells by particle bombardment/biolistics," *Methods Mol Biol.*, 286, 61-78 (2005)).

Se bombardean placas con los embriones a nivel n.º 4 con un sistema de pistola génica Helios® (Bio-Rad, Hercules CA). Todas las muestras reciben un único disparo a 68,93 MPA (450 PSI) de partículas preparadas/ADN. Después del bombardeo, los embriones se incuban en medio de mantenimiento durante 12 a 48 horas a temperaturas que varían de 26 °C a 37 °C, y se colocan después a 26 °C.

Después de 7 días, se agrupan aproximadamente 30 de los embriones más uniformemente transformados de cada tratamiento y se extrae el ADN genómico total. La región que rodea el sitio diana pretendido se amplifica por PCR con mezcla maestra de PCR de alta fidelidad Phusion® (New England Biolabs, Ipswich, MA). La amplificación por PCR también se usa para añadir códigos de barras específicos de amplicón y cebadores de secuenciación de Illumina (Illumina, San Diego, CA). Los productos de amplificación por PCR resultantes se purifican con una columna de centrifugación de purificación de PCR (Qiagen, Venlo, Países Bajos), la concentración se mide con un ensayo fluorométrico basado en colorante de Hoechst, se combina en una relación equimolar y se realizó secuenciación profunda de 100 nucleótidos de longitud de una única lectura en un secuenciador personal MiSeq (Illumina, San Diego, CA).

Se determinan las frecuencias de mutaciones de UENH recuperadas por secuenciación profunda para los sistemas de sn1-casARN/sn2-casARN/sn3-casARN/proteína Cas9 que se dirigen a las tres dianas de LIGCas en comparación con el sistema de ARN guía largo individual/endonucleasa Cas9 que se dirige al locus correspondiente. Estos datos son para demostrar que los sistemas de sn1-casARN/sn2-casARN/sn3-casARN/proteína Cas9 como se describe en el presente documento escinden ADN cromosómico de maíz y genera mutaciones mediadas por UENH.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, los procedimientos descritos en este ejemplo pueden ser practicados por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos dividido.

Ejemplo 17

Generación de ratones transgénicos

Este ejemplo describe uso de los polinucleótidos de nexos dividido (sn-casPN) de la presente invención para crear

modificaciones genómicas en animales.

Un sistema de sn-casARN de dos partes (sn1-casARN y sn2-casARN) (véase, por ejemplo, FIG. 3B) se usa en una modificación del procedimiento de Wang, y col., "One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering," *Cell*, 153(4), 910-918 (2013), que usaron un único ARN guía que tenía una secuencia continua.

Se describen ejemplos de vectores, medios, condiciones de cultivo adecuados, etc. Un experto habitual en la materia entenderá modificaciones de estos componentes y condiciones a la vista de las enseñanzas de la presente memoria descriptiva.

A. Producción de ARNm de Cas9 y sn1-casARN/sn2-casARN

Se añade un promotor T7 a la región codificante de Cas9 optimizada para expresión en mamífero (por ejemplo, la secuencia codificante de Cas9 puede amplificarse por PCR a partir de pX330-U6-BB_quimérico-CBh-hSpCas9; Addgene, Cambridge, MA). El promotor T7 se añade por amplificación por PCR. El producto de PCR de T7-Cas9 se purifica en gel y se usa como el molde para transcripción *in vitro* usando el kit mMACHINE T7 ULTRA (Life Technologies, Grand Island, NY). El ARNm de Cas9 se purifica usando el kit MEGAclean (Life Technologies, Grand Island, NY) y se eluye en agua sin RNasa.

Se sintetizan químicamente secuencias de ADN que codifican los sn1-casARN y sn2-casARN (véase, por ejemplo, la FIG. 3B). Las secuencias espaciadoras de 20 nucleótidos para los sn1-casARN son las siguientes: sn1-casARN-Tet1, GGCTGCTGTC AGGGAGCTCA (SEQ ID NO: 89); y sn1-casARN-Tet 2, GAAAGTGCCA ACAGATATCC (SEQ ID NO: 90) (véase, Figura 1A de Wang, y col., *Cell*, 153(4), 910-918 (2013)). El promotor T7 se añade a cada uno de los moldes de sn1-casARN y sn2-casARN mediante amplificación por PCR. Los productos de PCR de T7-sn-casARN se purifican en gel y se usan como el molde para transcripción *in vitro* usando el kit MEGAshortscript T7 (Life Technologies, Grand Island, NY). Los sn-casARN se purifican usando el kit MEGAclean (Life Technologies, Grand Island, NY) y se eluyen en agua sin RNasa.

B. Inyección en embrión de una célula

Todos los procedimientos animales se realizan según las directrices de los Institutos Nacionales de Salud (NIH). Se usan ratones hembra B6D2F1 (C57BL/6 X DBA2) como donantes de embriones. Se usan cepas de ratón ICR como madres adoptivas. Se aparean ratones B6D2F1 hembras de siete a ocho semanas de edad, superovuladas con machos B6D2F1. Se recogen embriones fertilizados de los oviductos. Se inyectan ARNm de Cas9 (administrados a embriones individuales sobre un intervalo de aproximadamente 20 ng/ml a aproximadamente 200 ng/ml), sn1-casARN/sn2-casARN (administrados a embriones individuales sobre un intervalo de 20 ng/ml a 50 ng/ml) en el citoplasma de embriones fertilizados (que tienen pronúcleos bien reconocidos) en medio M2 (Sigma-Aldrich, St. Louis, MO).

Cuando también se inyecta un oligonucleótido donante las concentraciones de los componentes del sistema de sn-casARN/proteína Cas9 son las siguientes: ARNm de Cas9 (aproximadamente 100 ng/ml), sn1-casARN/sn2-casARN (50 ng/ml) y oligonucleótido donantes (100 ng/ml). Los componentes se mezclan y se inyectan en cigotos en el estadio de pronúcleo. Los cigotos inyectados se cultivan en medio de cultivo de embriones KSOM PrimeQ™, con aminoácidos y rojo fenol (MTI-GlobalStem, Gaithersburg, MD) a 37 °C con CO₂ 5 % al aire durante aproximadamente 3,5 días (hasta el estadio de blastocisto). Se transfieren 15-25 blastocistos a los úteros de hembras ICR pseudoembarazadas a aproximadamente 2,5 días después del coito.

C. Ratones mutantes de doble gen

sn1-casARN-Tet1/sn2-casARN y sn1-casARN-Tet2/sn2-casARN se inyectan juntos como se ha descrito anteriormente en cigotos. El ADN genómico de las crías se evalúa mediante RFLP (análisis de polimorfismo de longitud de fragmentos de restricción), análisis de transferencia de Southern y análisis de secuenciación para identificar ratones que portan mutaciones diana en los cuatro alelos de los genes de Tet1 y Tet2. Los resultados de estos análisis son para demostrar que pueden generarse eficazmente ratones postnatales que portan mutaciones bialélicas en dos genes diferentes (es decir, los genes de Tet1 y Tet2).

También se evalúan los efectos fuera de diana *in vivo*. El trabajo previo *in vitro*, en bacterias, y en células humanas cultivadas sugiere que la secuencia de motivo adyacente a protoespaciador NGG y la "secuencia de siembra" de 8 a 12 bases de la secuencia espaciadora es importante para determinar la especificidad de escisión de ADN (Cong, L., y col., "Multiplex genome engineering using CRISPR/Cas systems," *Science* 339, 819-82 (2013); Jiang, W., y col., "RNA-guided editing of bacterial genomes using CRISPR-Cas systems," *Nat. Biotechnol.*, 31,233-239 (2013); y Jinek, M., y col., "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science*, 337, 816-821 (2012)). Usando esta norma, Wang, y col., identificaron que solamente tres sitios fuera de diana potenciales de Tet1 y cuatro de Tet2 existen en el genoma de ratón. Los efectos fuera de diana se evalúan usando el ensayo Surveyor (Guschin, D.Y., y col., "A rapid and general assay for monitoring endogenous gene modification," *Methods Mol. Biol.*, 649, 247-256 (2010)). El número de efectos fuera de diana proporciona una estimación de precisión de dirección *in vivo* del complejo de sn1-casARN/sn2-casARN/proteína Cas9.

D. Modificación de reparación de genes *in vivo*

Para evaluar la reparación de genes *in vivo* usando el sistema de sn1-casARN y sn2-casARN, un oligonucleótido donante se usa para dirigirse a Tet1 para cambiar dos pares de bases de un sitio de restricción de SmaI para crear un sitio de EcoRI (oligonucleótido de Tet1; 126 pb, para la secuencia véase la Figura 3A de Wang, y col.). Un segundo oligonucleótido donante se usa para dirigirse a Tet2 para cambiar dos pares de bases de un sitio de EcoRV en un sitio de EcoRI (oligonucleótido de Tet2; 126 pb, para la secuencia véase la Figura 3A de Wang, y col.). Los blastocistos proceden de cigotos a los que se ha inyectado ARNm de Cas9, sn1-casARN-Tet1/sn2-casARN y oligonucleótido de Tet1, ARNm de Cas9, sn1-casARN-Tet2/sn2-casARN y oligonucleótido de Tet2, y ARNm de Cas9, sn1-casARN-Tet1/sn2-casARN, oligonucleótido de Tet1, sn1-casARN-Tet2/sn2-casARN y oligonucleótido de Tet2.

Se aísla ADN de los blastocistos con ARNm de Cas9, sn1-casARN-Tet1/sn2-casARN y oligonucleótido de Tet1, se amplifica y se digiere con EcoRI para detectar acontecimientos de reparación de genes mediada por oligonucleótidos. Se aísla ADN de los blastocistos con ARNm de Cas9, sn1-casARN-Tet2/sn2-casARN y oligonucleótido de Tet2, se amplifica y se digiere con EcoRI para detectar acontecimientos de reparación de genes mediada por oligonucleótidos. Se aísla ADN de los blastocistos con ARNm de Cas9, sn1-casARN-Tet1/sn2-casARN, oligonucleótido de Tet1, sn1-casARN-Tet2/sn2-casARN y oligonucleótido de Tet2, se amplifica y se digiere con EcoRI para detectar acontecimientos de reparación de genes mediada por oligonucleótidos. El ADN genómico de los blastocistos se evalúa mediante RFLP, análisis de transferencia de Southern y análisis de secuenciación para identificar blastocistos que portan sitios de restricción modificados de los genes de Tet1 y Tet2. Los resultados de estos análisis son para demostrar que la reparación *in vivo* de genes de ratón (es decir, los genes de Tet1 y Tet2) pueden llevarse a cabo eficazmente.

Se usan análisis de RFLP usando escisión de SmaI y EcoRV para evaluar los loci de Tet1 y Tet2, respectivamente, para demostrar que los alelos que no son diana de un ARNm de Cas9, sn1-casARN-Tet1 o Tet2/sn2-casARN y oligonucleótido seleccionado (en cada una de las combinaciones enumeradas anteriormente) no se ven afectados.

Asimismo, se implantan blastocistos con dobles inyecciones de oligonucleótidos en madres adoptivas. El ADN genómico de las crías resultantes se evalúa mediante RFLP, análisis de transferencia de Southern y análisis de secuenciación para identificar blastocistos que portan sitios de restricción modificados de los genes de Tet1 y Tet2. Los resultados de estos análisis son para demostrar que pueden generarse ratones con modificaciones de reparación genómica en múltiples genes.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, los procedimientos descritos en este ejemplo pueden ser practicados por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexa dividido.

Ejemplo 18Complejos de sn-casARN/proteína Cas9 en vectores de administración que comprenden moléculas catiónicas

A. Producción de ARNm de Cas9 y sn1-casARN/sn2-casARN/sn3-casARN

Se añade un promotor T7 a la región codificante de Cas9 optimizada para expresión de mamífero y se marca en el extremo C con dos secuencias de localización nuclear (NLS). El promotor T7 se añade por amplificación por PCR. El producto de PCR de T7-Cas9 se purifica en gel y se clona en un vector para expresión de proteínas sin células (por ejemplo, vector pT7CFE1-NFtag para expresión de proteínas sin células de mamíferos, Life Technologies, Grand Island, NY). Se expresa y aísla proteína Cas9 usando un sistema de expresión de proteínas sin células (por ejemplo, kit de IVT de alto rendimiento de CHO de 1 etapa, Life Technologies, Grand Island, NY) y se suspende en agua sin RNasa.

Se preparan secuencias de ADN que codifican los sn1-casARN, sn2-casARN y sn3-casARN-AAVS-1 como se ha descrito en el Ejemplo 1. Los productos de PCR de T7-sn-casARN se purifican en gel y se usan como el molde para transcripción *in vitro* usando el kit de síntesis de ARN de alto rendimiento T7 (New England Biolabs, Ipswich, MA). Los sn-casARN se purifican usando el kit de limpieza y concentración de ARN GeneJet (Life Technologies, Grand Island, NY) y se eluyen en agua sin RNasa.

B. Formación de complejos ribonucleoproteicos

Se preparan complejos ribonucleoproteicos (RNP) a dos concentraciones, Cas9 50 pmol:sn-casARN 150 pmol y Cas9 200 pmol:sn-casARN 600 pmol. Los tres componentes de sn-casARN en cantidades equimolares se mezclan en un tampón de hibridación (HEPES 1,25 mM, MgCl₂ 0,625 mM, KCl 9,375 mM a pH 7,5) hasta la concentración deseada (150 pmol o 600 pmol) en un volumen final de 5 µl, se incuban durante 2 minutos a 95 °C, se retiran del termociclador y se permite que se equilibren a temperatura ambiente. La proteína Cas9 se diluye hasta una concentración apropiada en tampón de unión (HEPES 20 mM, KCl 100 mM, MgCl₂ 5 mM, DTT 1 mM y glicerol 5 % a pH 7,4) en un volumen final de 5 µl y se mezcla con los 5 µl de sn-casARN desnaturalizados por calor seguido de

incubación a 37 °C durante 30 minutos para formar los complejos de sn-casARN/proteína Cas9.

C. Preparación de vector de SC12CDClickpropilamina con complejos de sn-casARN/proteína Cas9

5 SC12CDClickpropilamina (una betaciclodextrina catiónica; O'Mahony A.M., y col., "Cationic and PEGylated Amphiphilic Cyclodextrins: Co-Formulation Opportunities for Neuronal Sirna Delivery," PLoS ONE 8(6), e66413 (2013)) se pesa y se disuelve en cloroformo (aproximadamente 1 mg/ml), después se mezcla en volúmenes apropiados para proporcionar relaciones molares de ciclodextrina catiónica con respecto a PEGilada (publicación de Solicitud de Patente de los Estados Unidos n.º 2014-0079770, publicada el 20 de marzo de 2014). El disolvente se retira bajo una corriente de nitrógeno para proporcionar una composición de ciclodextrina (CD) seca.

10 La composición de CD se rehidrata con tampón de unión (concentración final de aproximadamente 1 mg/ml) y se somete a ultrasonidos durante una hora a temperatura ambiente seguido de la adición inmediata de los complejos de sn-casARN/proteína Cas9 en tampón de unión. Los complejos de sn-casARN/proteína Cas9 tampón de unión se añaden en un volumen igual. La solución se mezcla y se incuba durante 20-30 minutos a temperatura ambiente para producir una composición de CD que comprende los complejos de sn-casARN/proteína Cas9 (CD-sn-casARN/proteína Cas9).

15 D. Preparación de complejos de sn-casARN/proteína Cas9 inmovilizados en liposomas

Se forman liposomas sin los complejos de casARN/proteína Cas9 para proporcionar controles negativos (liposomas vacíos).

En un matraz de fondo redondo de tamaño adecuado los componentes del liposoma se añaden y se solubilizan en un disolvente o mezcla de disolvente adecuado. Los siguientes son componentes de liposoma ejemplares:

20 Liposoma 1: Se preparan EPC (solución de EtOH) y Colesterol (solución de EtOH) en una relación molar de 70/30.

Liposoma 1-PEG: Se añade PEG2000 estearilado (solución de EtOH) para ser 5 % mol con respecto a la cantidad de lípido total del liposoma 1 (EPC+Colesterol).

25 Liposoma 2: DOTMA (solución de EtOH), Se preparan colesterol (solución de EtOH) y EPC (solución de EtOH) en una relación molar de 30/40/30.

Liposoma 2-PEG: Se añade PEG2000 estearilado (solución de EtOH) para ser 5 % mol con respecto a la cantidad de lípido total del liposoma 2 (DOTMA+Colesterol+EPC).

Liposoma 3: Se añadieron DODAP (solución de EtOH), colesterol (solución de EtOH) y EPC (solución de EtOH) en una relación molar de 30/40/30.

30 Liposoma 3-PEG: Se añade PEG2000 estearilado (solución de EtOH) para ser 5 % mol con respecto a la cantidad de lípido total del liposoma 3 (DODAP+Colesterol+EPC).

35 Se añade una cantidad de EtOH para solubilizar todos los componentes. El matraz se une a un evaporador rotatorio que gira a 50-100 rpm y se sumerge en un baño de agua ajustado a la mayor temperatura de transición de fase cristalina de gel-líquido (Tc) de los lípidos usados. Se permite que el matraz gire en el baño de agua durante aproximadamente 1 minutos para equilibrar. Se aplica un vacío lento, hasta < 1333 Pa (10 Torr), para obtener una película seca delgada en las paredes del matraz sin precipitación. Para retirar cualquier disolvente residual, el matraz se somete a alto vacío a temperatura ambiente durante varias horas o durante una noche.

40 Se añade una solución de complejos de sn-casARN/proteína Cas9 o CD-sn-casARN/proteína Cas9 para obtener una concentración final de lípido de entre aproximadamente 2 mM y aproximadamente 0,5 mM. Se realiza rehidratación de lípidos a temperatura ambiente durante 15 minutos o más. Se preparan liposomas por aplicación de ultrasonidos durante aproximadamente 1 minuto.

Los procedimientos anteriores producen las composiciones de partículas y composiciones de liposomas mostradas en la Tabla 19.

Tabla 19

Composiciones de partículas y liposomas	
Partícula	complejo de sn-casARN/proteína Cas9
Partícula	complejo de CD-sn-casARN/proteína Cas9
Liposoma 1	complejo de sn-casARN/proteína Cas9
Liposoma 1	complejo de CD-sn-casARN/proteína Cas9

45

(continuación)

Composiciones de partículas y liposomas	
Liposoma 1-PEG	complejo de sn-casARN/proteína Cas9
Liposoma 1-PEG	complejo de CD-sn-casARN/proteína Cas9
Liposoma 2	complejo de sn-casARN/proteína Cas9
Liposoma 2	complejo de CD-sn-casARN/proteína Cas9
Liposoma 2-PEG	complejo de sn-casARN/proteína Cas9
Liposoma 2-PEG	complejo de CD-sn-casARN/proteína Cas9
Liposoma 3	complejo de sn-casARN/proteína Cas9
Liposoma 3	complejo de CD-sn-casARN/proteína Cas9
Liposoma 3-PEG	complejo de sn-casARN/proteína Cas9
Liposoma 3-PEG	complejo de CD-sn-casARN/proteína Cas9

E. Caracterización de composiciones de partículas y composiciones de liposomas

Las composiciones de partículas y composiciones de liposomas descritas anteriormente se caracterizan (véase, por ejemplo, Laouini, A., y col., "Preparation, Characterization and Applications of Liposomes: State of the Art," Journal of Colloid Science and Biotechnology, 1, 147-168 (2012)) usando procedimientos convencionales de la siguiente manera.

(i) Análisis de tamaño

Los tamaños de las partículas y los liposomas se evalúan por técnicas convencionales. Están disponibles varias técnicas para evaluar el tamaño de los liposomas y la distribución de tamaños incluyendo técnicas de microscopía, cromatografía de exclusión por tamaño (SEC), fraccionamiento de flujo de campo y dispersión de luz estática o dinámica. Asimismo, los tamaños de partículas pueden evaluarse usando geles de agarosa no desnaturalizantes (por ejemplo, geles de agarosa al 1,5 %, SYBR® Safe, Life Technologies, Grand Island, NY). Diferentes tamaños de partículas y liposomas son útiles para diferentes aplicaciones, por ejemplo, transfección celular en cultivo o administración terapéutica a un animal.

(ii) Mediciones de carga

El tamaño promedio y la carga de las partículas y los liposomas se miden con un Zetasizer Nano ZS (Malvern, Westborough, MA) usando dispersión de luz dinámica (DLD). Si todas las partículas en una suspensión tienen un potencial zeta negativo grande o positivo grande estas tienden a repelerse entre sí. De este modo se reduce o elimina la tendencia a la agregación. Sin embargo, las partículas con valor de potencial zeta bajo no tienen fuerza para prevenir que las partículas floculen.

(iii) Morfología

Las morfologías de las composiciones se evalúan usando microscopía electrónica de transmisión (MET), por ejemplo, un microscopio electrónico de transmisión JEOL 2000 FXII (Jeol Ltd., Tokio, Japón). En general, son más deseables composiciones que tengan morfología de partículas o liposomas uniforme.

(iv) Estudios de agregación

Los efectos de medio que contiene sal y suero en la agregación de las partículas y los liposomas se evalúa incubando complejos en medio de transfección Opti-MEM® (Life Technologies, Grand Island, NY) o suero bovino fetal durante 24 horas a 37 °C. Se llevan a cabo mediciones de tamaños usando el Zetasizer Nano ZS. La ausencia de agregación es con frecuencia una cualidad deseable. Sin embargo, en algunas aplicaciones (por ejemplo, experimentos de transfección) puede ser deseable algo de agregación.

(v) Eficacia de encapsulación

Las preparaciones de liposomas son una mezcla de fracciones de CD-sn-casARN/proteína Cas9 o sn-casARN/proteína Cas9 encapsuladas o no encapsuladas. Se conocen varias técnicas para la determinación de la eficacia de encapsulación incluyendo HPLC y fraccionamiento de flujo de campo. Habitualmente, el porcentaje de encapsulación se expresa como la relación del área de pico no encapsulada con respecto a la de un patrón de referencia a la misma concentración inicial. Este procedimiento puede aplicarse si los liposomas no experimentan

ninguna purificación después de la preparación. En general un alto grado de eficacia de encapsulación es un parámetro importante para liposomas en aplicaciones terapéuticas. La eficacia de encapsulación baja necesita la incorporación de una etapa de separación postencapsulación (tal como diálisis, cromatografía de exclusión por tamaño o ultrafiltración) para retirar complejos no encapsulados. Habitualmente es deseable una eficacia de encapsulación de más del 45 %, más preferentemente más del 80 % y lo más preferentemente más del 95 %.

(v) Actividad *in vivo*

Las composiciones de partículas y composiciones de liposomas que comprenden el sn1-casARN, sn2-casARN y sn3-casARN-AAVS-1 se evalúan con respecto a sus capacidades relativas para suministrar los complejos de sn-casARN/proteína Cas9 a células para reparación génica. Para este experimento se incluye una molécula de ADN donante en la preparación de cada una de las composiciones de partículas y composiciones de liposomas. El ADN donante es un fragmento de EGFP para su uso con la línea celular indicadora EGIP 293T de control positivo de AAVS1 (System Biosciences, Mountain View, CA). Otros sistemas indicadores son adecuados para su uso en este análisis. Las composiciones de liposomas se concentran según sea necesario usando técnicas convencionales.

Las composiciones de partículas y composiciones de liposomas se transfectan en células. Un día antes de la transfección, las células se siembran en placas en medio de cultivo sin antibióticos. Las células deberían estar en confluencia en el momento de la transfección. Las composiciones de partículas y composiciones de liposomas se diluyen en medio Opti-MEM® I (Life Technologies, Grand Island, NY) sin suero para proporcionar un intervalo de concentración para los complejos de sn-casARN/proteína Cas9. Habitualmente, se aplica un volumen de aproximadamente 200 µl de estas suspensiones en placas multipocillo. Las células y suspensiones se mezclan suavemente sacudiendo las placas adelante y atrás. Las células se incuban a 37 °C en un incubador de CO₂ durante 5-24 horas. Al día siguiente, se añade medio completo a las células. Las células son células incubadas a 37 °C en un incubador de CO₂ durante 24-48 horas antes de ensayar.

La reparación génica se evalúa usando la línea celular EGIP 293T de control positivo de AAVS1 para supervisar la eficacia de HDR de ADN donante de EGFP. La secuencia de ARN de sn3-casARN-AAVS1 dirige el complejo de sn-casARN/proteína Cas9 a la diana y escinde una secuencia de AAVS1 de 53 pb integrada en una línea celular indicadora de proteína verde fluorescente inhibida potenciada (EGIP). La EGIP comprende un codón de terminación en el medio para inactivar la expresión de la proteína verde fluorescente potenciada (EGFP). En presencia de un complejo de sn-casARN/proteína Cas9 activo, el ADN donante de EGFP se recombina en el sitio diana y restaura la lectura abierta del gen de EGFP mediante recombinación homóloga. Se mide la eficacia de escisión del complejo de sn-casARN/proteína Cas9 que se dirige al locus de AAVS1 usando el ensayo Surveyor. La eficacia de restauración de la expresión de EGFP se supervisa usando microscopia de fluorescencia. Los resultados se expresan como un porcentaje de expresión del gen de EGFP en relación con la composición de partículas (sn-casARN/proteína Cas9 y CD-sn-casARN/proteína Cas9) y los controles de liposomas vacíos. Los resultados de los estudios de actividad *in vivo* proporcionan orientación para la selección de componentes y composiciones óptimos de partículas y/o liposomas.

En conjunto, estos datos proporcionan información que permite establecer criterios para seleccionar composiciones liposómicas óptimas para encapsulación de complejos de sn-casARN/proteína Cas9 de la presente invención según sus ventajas y limitaciones.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, los procedimientos descritos en este ejemplo pueden ser practicados por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexos dividido.

Ejemplo 19

Análisis de secuenciación profunda adicional para detección de modificaciones diana en células eucariotas

Este ejemplo describe el uso de un ARNgi de CRISPR-Cas9 de *Streptococcus thermophilus* y dos polinucleótidos de nexos dividido (sn1-casARN y sn2-casARN de la presente divulgación de invención) para modificar seis secuencias diana de ADN bicatenario seleccionadas en células humanas.

A. Selección de diana y generación de reactivo de sn-casARN

Se seleccionaron seis secuencias diana a partir del genoma humano para edición. Los sitios diana fueron de 20 nucleótidos de longitud y aparecieron 5' de la secuencia de MAP de CRISPR1/Cas9 de *S. thermophilus* (5'-NNAGAAW-3'), en la que 'N' es cualquier nucleótido y 'W' es una base de adenina, timina o uracilo. Las dianas y coordenadas del genoma humano seleccionadas (en relación con "Genome Reference Consortium," Centro Nacional para la Información Biotecnológica, ensamblaje de referencia GRCh37/hg19) se presentan en la Tabla 20.

Tabla 20

Secuencias diana de CRISPR1/Cas9 de <i>S. thermophilus</i>			
Nombre de la diana	Localización de la diana (hg19)	Secuencia diana con MAP*	SEQ ID NO
NMNAT2_tgt_3	cr1: 183220522-183220548	TGTAAGTCCAGGATGAACT <u>AGGAGAAA</u>	SEQ ID NO: 91
chr13_tgt_4	cr13: 108863533-108863559	TGTTCAACTTTAGAACGAA <u>TACAGAAA</u>	SEQ ID NO: 92
STAT3_tgt_2	cr17: 40466501-40466527	AGGGTGGGGCAGGAAGGA <u>AGCCAGAAT</u>	SEQ ID NO: 93
STAT3_tgt_5	cr17: 40481464-40481490	CTCCCTCAGGGTCTGTAAG <u>AAAAGAAA</u>	SEQ ID NO: 94
STAT1_tgt_1	cr2: 191835142-191835168	ATACACCCAAAGTATCAGG <u>ACGAGAAT</u>	SEQ ID NO: 95
XRCC5_tgt_3	cr2: 216974165-216974191	GCGGTCTGGGGAATAAGGTA <u>TAAAGAAA</u>	SEQ ID NO: 96
*la secuencia de MAP está subrayada			

Se usaron secuencias diana para diseñar elementos espaciadores para ARNgi de *S. thermophilus* (ARNgi^{S.th}), sn1-casARN de *S. thermophilus* sn1-casARN (sn1-casARN^{S.th}) y el sn2-casARN de *S. thermophilus* adjunto (sn2-casARN^{S.th}). Los moldes de ADN bicatenario y componentes de ARN se produjeron esencialmente como se ha descrito previamente (Ejemplo 1). Se enumeran secuencias de ARNgi y sn-casARN en la Tabla 21.

5

Tabla 21

Secuencias de ARNgi ^{S.th} , sn1-casARN ^{S.th} y sn2-casARN ^{S.th}		
Tipo de componente de ARN de Cas	Secuencia de ARN	SEQ ID NO
ARNgi ^{S.th} -NMNAT2_tgt_3	5' - UGUAAGUCCA GGAUGAACUA GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUCAU GCCGAAAUCA ACACCCUGUC AUUUUAUGGC AGGGUGUUU - 3'	SEQ ID NO: 97
ARNgi ^{S.th} -chr13_tgt_4	5' - UGUUCAACUU UAGAACGAAU GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUCAU GCCGAAAUCA ACACCCUGUC AUUUUAUGGC AGGGUGUUU - 3'	SEQ ID NO: 98

ES 2 705 540 T3

(continuación)

Secuencias de ARNgi ^{S.th} , sn1-casARN ^{S.th} y sn2-casARN ^{S.th}		
Tipo de componente de ARN de Cas	Secuencia de ARN	SEQ ID NO
ARNgi ^{S.th} -STAT3_tgt_2	5' - AGGGUGGGGC AGGAAGGAAG GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUCAU GCCGAAAUCA ACACCCUGUC AUUUUAUGGC AGGGUGUUU - 3'	SEQ ID NO: 99
ARNgi ^{S.th} -STAT3_tgt_5	5' - CUCCCUCAGG GUCUGUAAGA GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUCAU GCCGAAAUCA ACACCCUGUC AUUUUAUGGC AGGGUGUUU - 3'	SEQ ID NO: 100
ARNgi ^{S.th} -STAT1_tgt_1	5' - AUACACCCAA AGUAUCAGGA GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUCAU GCCGAAAUCA ACACCCUGUC AUUUUAUGGC AGGGUGUUU - 3'	SEQ ID NO: 101
ARNgi ^{S.th} -XRCC5_tgt_3	5' - GCGGUCGGGG AAUAAGGUAU GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUCAU GCCGAAAUCA ACACCCUGUC AUUUUAUGGC AGGGUGUUU - 3'	SEQ ID NO: 102
sn1-casARN ^{S.th} -NMNAT2_tgt_3	5' - UGUAAGUCCA GGAUGAACUA GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUC - 3'	SEQ ID NO: 103
sn1-casARN ^{S.th} -chr13_tgt_4	5' - UGUUCAACUU UAGAACGAU GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUC - 3'	SEQ ID NO: 104
sn1-casARN ^{S.th} -STAT3_tgt_2	5' - AGGGUGGGGC AGGAAGGAAG GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUC - 3'	SEQ ID NO: 105
sn1-casARN ^{S.th} -STAT3_tgt_5	5' - CUCCCUCAGG GUCUGUAAGA GUUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUC - 3'	SEQ ID NO: 106

(continuación)

Secuencias de ARNg ^{S.th} , sn1-casARN ^{S.th} y sn2-casARN ^{S.th}		
Tipo de componente de ARN de Cas	Secuencia de ARN	SEQ ID NO
sn1-casARN ^{S.th} -STAT1_tgt_1	5' - AUACACCCAA AGUAUCAGGA GUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUC - 3'	SEQ ID NO: 107
sn1-casARN ^{S.th} -XRCC5_tgt_3	5' - GCGGUCGGGG AAUAAGGUAU GUUUUGUAC UCGAAAGAAG CUACAAAGAU AAGGCUUC - 3'	SEQ ID NO: 108
sn2-casARN ^{S.th}	5' - AUGCCGAAAU CAACACCCUG UCAUUUAUG GCAGGGUGUU U - 3'	SEQ ID NO: 109

C. Formación de complejos de sn1-casARN/sn2-casARN/proteína Cas

Una proteína Cas9 de *Streptococcus thermophilus* marcada con NLS se expresó a partir de un vector de expresión bacteriano en *E. coli* (BL21 (DE3)) y se purifica usando cromatografía de intercambio iónico de afinidad y exclusión por tamaño según procedimientos descritos en Jinek, M., y col., (Science, 337, 816-21 (2012)). La secuencia codificante para la proteína Cas9 se diseñó para incluir dos secuencias de localización nuclear (NLS) en el extremo C. Los complejos se ensamblaron, por triplicado a una concentración de 40 pmol de Cas9 y 120 pmol del ARNg^{S.th} o el sn1-casARN^{S.th}/sn2-casARN^{S.th}. Los componentes del sn1-cas-ARN^{S.th} y el sn2-casARN^{S.th} se mezclaron en cantidades equimolares hasta la concentración deseada (120 pmol) en un volumen final de 2,5 µl, se incubaron durante 2 minutos a 95 °C, se retiraron del termociclador y se permitió que se equilibraran a temperatura ambiente. Se diluyeron ARNg^{S.th} de control hasta la concentración final (120 pmol) en un volumen final de 2,5 µl, se incubaron durante 2 minutos a 95 °C, se retiraron del termociclador y se permitió que se equilibraran a temperatura ambiente. La proteína Cas9 se diluyó hasta una concentración apropiada en tampón de unión (HEPES 20 mM, KCl 100 mM, MgCl₂ 5 mM, DTT 1 mM y glicerol 5 % a pH 7,4) hasta un volumen final de 2,5 µl y se mezcló con cada 2,5 µE del ARNg^{S.th} desnaturalizado por calor o cada sn1-casARN^{S.th}/sn2-casARN^{S.th} seguido de incubación a 37 °C durante 30 minutos.

D. Transfecciones celulares y secuenciación profunda usando complejos nucleoproteicos de ARNg^{S.th}/proteína Cas9 y sn1-casARN^{S.th}/sn2-casARN^{S.th}/proteína Cas9.

Se transfectaron complejos nucleoproteicos de ARNg^{S.th}/proteína Cas9 o el sn1-casRNA^{S.th}/sn2-casRNA^{S.th}/proteína Cas9 en células HEK293 (Colección Americana de Cultivos Tipo, Manassas, VA), usando el sistema de lanzadera de Nucleofector® de 96 pocillos (Lonza, Allendale, NJ). Se realizaron nucleofección y secuenciación profunda para la detección de modificaciones diana esencialmente como se ha descrito previamente (Ejemplo 4).

Se muestran insup detectados para edición de ARNg^{S.th}/Cas9 y sn1-casARN^{S.th}/sn2-casARN^{S.th}/Cas9 para los seis sitios diana en la Tabla 22. Los valores son el promedio de tres repeticiones técnicas.

Tabla 22

Insup detectados para edición en células				
Diana	ARNg ^{S.th}		sn1-casARN ^{S.th} /sn2-casARN ^{S.th}	
	Porcentaje de insup detectados	Desv. típica	Porcentaje de insup detectados	Desv. típica
NMNAT2_tgt_3	61,3 %	0,2 %	80,4 %	3,5 %
chr13_tgt_4	11,6 %	1,2 %	9,0 %	0,5 %
STAT3_tgt_2	78,1 %	0,7 %	71,0 %	1,3 %
STAT3_tgt_5	76,5 %	3,9 %	33,7 %	2,6 %

(continuación)

Insup detectados para edición en células				
Diana	ARNGi ^{S.tn}		sn1-casARN ^{S.tn} /sn2-casARN ^{S.tn}	
	Porcentaje de insup detectados	Desv. típica	Porcentaje de insup detectados	Desv. típica
STAT1_tgt_1	77,1 %	2,4 %	12,6 %	2,3 %
XRCC5_tgt_3	89,3 %	1,2 %	87,7 %	0,8 %

Los datos presentados en la Tabla 22 para actividad en la célula de un complejo nucleoproteico de sn1-casARN/sn2-casARN/proteína Cas evaluado por análisis de datos de secuenciación profunda para detección de modificaciones diana en células eucariotas proporciona datos para demostrar que las construcciones de proteína Cas/sn1-casPN/sn2-casPN como se describe en el presente documento facilitan escisión específica de sitio mediado por Cas de ácidos nucleicos bicatenarios diana en células. Estos datos apoyan que los sistemas de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II de la presente invención facilitan la escisión específica de sitio mediada por Cas en la célula de loci genómicos.

- 5 Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, los procedimientos descritos en este ejemplo pueden ser practicados por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexa dividido.

Ejemplo 20

15 Diseños de sn-casPN alternativos

Este ejemplo ilustra el uso de diseños de sn-casPN alternativos para evaluar y comparar el porcentaje de escisión *in vitro* de sistemas de sn-casPN/proteína Cas9 de *S. pyogenes* en relación con secuencias diana de ácido nucleico seleccionadas. Se realizaron tres sn-casPN con una única rotura en la cadena principal de nexa en localizaciones diferentes en el elemento de nexa para cada uno de los diseños de sn-casPN alternativos. Los tres diseños de sn-casPN se ensayaron con respecto a su capacidad para dirigir la actividad de escisión bioquímica contra tres dianas de ADN genómico humanas.

- 20 Se produjo ADN bicatenario diana como se ha descrito previamente (Ejemplo 2). Se amplificó una única región genómica que comprende las tres secuencias diana a partir de ADN genómico (ADNg) humano. Los cebadores oligonucleotídicos directo e inverso usados para amplificación de la secuencia diana de 670 pb a partir de ADNg se muestran en la Tabla 23.

Tabla 23

Secuencias de cebadores oligonucleotídicos de ADN para generar una secuencia diana de ADN bicatenario		
Nombre de la diana	Secuencias oligonucleotídicas de ADN	SEQ ID NO
tgt-directo	5' - TGTAAGGAAG CTGCAGCACC AGG – 3'	SEQ ID NO: 110
tgt-inverso	5' - TCTTTGCCTG GACACCCCGT TC – 3'	SEQ ID NO: 111

- 30 Se produjeron moldes de ADN bicatenario para componentes de ARN como se ha descrito previamente (Ejemplo 1). La escisión de las secuencias diana de 670 pb produjo tamaños de fragmentos de 375 pb y 295 pb para diana 1, 312 pb y 359 pb para diana 2, y 215 pb y 457 pb para diana 3. Para este ejemplo, los diseños de sn-casARN se etiquetan como v2.2, v2.3 y v2.4. **Las dianas** se indican como tgt1 (diana 1), tgt2 (diana 2) y tgt3 (diana 3). Se enumeran secuencias de ARNGi y sn-casARN en la Tabla 24.

Tabla 24

Secuencias para ARNgi y diseños de sn-casARN alternativos		
Tipo de componente de ARN	Secuencia de ARN	SEQ ID NO.
ARNgi-tgt1	5' – AGACCCAAUA UCAGGAGACU GUCUCAG AGCUAUGCU GUCCUGGAA ACAGGACAG CAUAGCAAG UUGAGAUGA GGCUAGUCC GUUAUCAAC UUGAAAAAG UGGCACCGA GUCGGUGCU U - 3'	SEQ ID NO: 112
ARNgi-tgt2	5' – UGUCCCUAGU GGCCCCACUG GUCUCAG AGCUAUGCU GUCCUGGAA ACAGGACAG CAUAGCAAG UUGAGAUGA GGCUAGUCC GUUAUCAAC UUGAAAAAG UGGCACCGA GUCGGUGCU U - 3'	SEQ ID NO: 113
ARNgi-tgt3	5' – CGGGGACACA GGAUCCCUUG GUCUCAG AGCUAUGCU GUCCUGGAA ACAGGACAG CAUAGCAAG UUGAGAUGA GGCUAGUCC GUUAUCAAC UUGAAAAAG UGGCACCGA GUCGGUGCU U - 3'	SEQ ID NO: 114
sn1-casARN ^{v2.2} -tgt1	5' - AGACCCAAUA UCAGGAGACU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAG - 3'	SEQ ID NO: 115
sn1-casARN ^{v2.2} -tgt2	5' - UGUCCCUAGU GGCCCCACUG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAG - 3'	SEQ ID NO: 116
sn1-casARN ^{v2.2} -tgt3	5' - CGGGGACACA GGAUCCCUUG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAG - 3'	SEQ ID NO: 117
sn2-casARN ^{v2.2}	5' - GCUAGUCCGU UAUCAACUUG AAAAAGUGGC ACCGAGUCGG UGCUU - 3'	SEQ ID NO: 118
sn1-casARN ^{v2.3} -tgt1	5' - AGACCCAAUA UCAGGAGACU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCU - 3'	SEQ ID NO: 119
sn1-casARN ^{v2.3} -tgt2	5' - UGUCCCUAGU GGCCCCACUG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCU - 3'	SEQ ID NO: 120
sn1-casARN ^{v2.3} -tgt3	5' - CGGGGACACA GGAUCCCUUG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCU - 3'	SEQ ID NO: 121
sn2-casARN ^{v2.3}	5' - GAGUCCGUUA UCAACUUGAA AAAGUGGCAC CGAGUCGGUG CUU - 3'	SEQ ID NO: 122

(continuación)

Secuencias para ARNgi y diseños de sn-casARN alternativos		
Tipo de componente de ARN	Secuencia de ARN	SEQ ID NO.
sn1-casARN ^{v2.4} -tgt1	5' - AGACCCAAUA UCAGGAGACU GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUAG - 3'	SEQ ID NO: 123
sn1-casARN ^{v2.4} -tgt2	5' - UGUCCCUAGU GGCCCCACUG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUAG - 3'	SEQ ID NO: 124
sn1-casARN ^{v2.4} -tgt3	5' - CGGGGACACA GGAUCCUGG GUCUCAGAGC UAUGCUGUCC UGGAAACAGG ACAGCAUAGC AAGUUGAGAU AAGGCUAG - 3'	SEQ ID NO: 125
sn2-casARN ^{v2.4}	5' - GUCCGUUAUC AACUUGAAA AGUGGCACCG AGUCGGUGCU U - 3'	SEQ ID NO: 126

Se mezclaron componentes de sn1-casARN y sn2-casARN emparejados (Tabla 24) para cada diana en cantidades equimolares de tampón de hibridación (HEPES 1,25 mM, MgCl₂ 0,625 mM, KCl 9,375 mM a pH7,5), se incubaron durante 2 minutos a 95 °C, se retiraron del termociclador y se permitió que se equilibraran a temperatura ambiente. Se diluyeron ARNgi de control positivo hasta concentración apropiada en tampón de hibridación. Se diluyeron componentes solamente de sn1-casARN de control negativo hasta la concentración apropiada en tampón de hibridación.

Se realizaron reacciones de escisión de Cas9 y se analizaron esencialmente como se ha descrito previamente (Ejemplo 3). Los componentes de sn1-casARN, sn2-casARN y ARNgi estuvieron a una concentración final de 500 nM en la reacción.

La Tabla 25 presenta los resultados del ensayo de escisión de Cas9 usando diseños de sn1-casARN/sn2-casARN alternativos, en los que la rotura en la cadena principal del nexo se realizó en diferentes localizaciones en el elemento de nexo. Para todas las reacciones de control negativo en las que estaba presente solamente un componente de sn1-casARN no se observó actividad de escisión (es decir, por debajo del límite de detección).

Tabla 25

Actividad de escisión bioquímica			
Tipo de componente de ARN	tgt-1	tgt-2	tgt-3
ARNgi	90 %	54 %	29 %
sn1-casARN ^{v2.2} /sn2-casARN ^{v2.2}	20 %	LDD*	3 %
sn1-casARN ^{v2.3} /sn2-casARN ^{v2.3}	76 %	25 %	26 %
sn1-casARN ^{v2.4} /sn2-casARN ^{v2.4}	45 %	3 %	14 %
*LDD: por debajo del límite de detección			

Los datos presentados en la Tabla 25 demuestran que diferentes pares de sn1-casARN/sn2-casARN, en los que cada par representa una localización diferente de la división en el elemento de nexo, facilitan la escisión específica de sitio mediada por proteína Cas9 *in vitro* de una diana de ADN bicatenario.

Siguiendo la orientación de la presente memoria descriptiva y los ejemplos, los procedimientos descritos en este ejemplo pueden ser practicados por un experto habitual en la materia con otras proteínas Cas9 de CRISPR de tipo II

incluyendo, pero sin limitación, Cas9 y fusiones de Cas9 combinadas con sus componentes polinucleotídicos afines modificados como se describe en el presente documento para comprender un elemento de nexo dividido.

Los artículos de la presente divulgación incluyen, pero sin limitación, los siguientes:

- 5 1. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:
 - dos o más polinucleótidos capaces de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con una primera secuencia de ADN que comprende una secuencia diana de ADN preferentemente en relación con una segunda secuencia de ADN sin la secuencia de unión a diana de ADN, en el que son necesarios al menos dos de los dos o más polinucleótidos para formar un elemento de tallo de nexo.
- 10 2. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:
 - dos o más polinucleótidos capaces de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 corte una primera secuencia de ADN que comprende una secuencia diana de ADN preferentemente en relación con una segunda secuencia de ADN sin la secuencia de unión a diana de ADN, en el que son necesarios al menos dos de los dos o más polinucleótidos para formar un elemento de tallo de nexo.
- 15 3. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 1 o 2, los dos o más polinucleótidos que comprenden:
 - un primer polinucleótido que comprende una primera secuencia de nucleótidos de elemento de tallo de nexo I y un segundo polinucleótido que comprende una primera secuencia de nucleótidos de elemento de tallo de nexo II, en el que (i) la primera secuencia de nucleótidos de elemento de tallo de nexo I y la primera secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar el elemento de tallo de nexo por enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo de nexo I y la segunda secuencia de nucleótidos de elemento de tallo de nexo II y (ii) el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.
- 20 4. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:
 - tres polinucleótidos capaces de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con una primera secuencia de ADN que comprende una secuencia diana de ADN preferentemente en relación con una segunda secuencia de ADN sin la secuencia de unión a diana de ADN, en el que son necesarios al menos dos de los tres polinucleótidos para formar un elemento de tallo de nexo.
- 25 5. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:
 - tres polinucleótidos capaces de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 corte una primera secuencia de ADN que comprende una secuencia diana de ADN preferentemente en relación con una segunda secuencia de ADN sin la secuencia de unión a diana de ADN, en el que son necesarios al menos dos de los dos o más polinucleótidos para formar un elemento de tallo de nexo.
- 30 6. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 4 o 5, que comprende un primer polinucleótido que comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y una secuencia de nucleótidos de elemento de tallo de nexo I, y
 - un segundo polinucleótido que comprende una secuencia de nucleótidos de elemento de tallo de nexo II, en el que la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar el elemento de tallo de nexo mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II; y
 - un tercer polinucleótido que comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II; en el que el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.
- 35 7. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende: dos o más polinucleótidos, comprendiendo los dos o más polinucleótidos un elemento de crtra que es capaz de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con secuencias de ADN que contienen secuencias de motivo adyacente a protoespaciador (MAP) preferentemente en relación con secuencias de ADN sin secuencias de MAP, comprendiendo el elemento de crtra,
 - (i) un primer polinucleótido que comprende una secuencia de nucleótidos de elemento de tallo de nexo I y un segundo polinucleótido que comprende una secuencia de nucleótidos de elemento de tallo de nexo II, en el que la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar un elemento de tallo de nexo por enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de
- 40
- 45
- 50
- 55
- 60

elemento de tallo de nexo II y (ii) el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

8. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:

5 dos o más polinucleótidos, comprendiendo los dos o más polinucleótidos un elemento de crtra que es capaz de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con y corte secuencias de ADN que contienen secuencias de motivo adyacente a protoespaciador (MAP) preferentemente en relación con secuencias de ADN sin secuencias de MAP, comprendiendo el elemento de crtra,

10 (i) un primer polinucleótido que comprende una secuencia de nucleótidos de elemento de tallo de nexo I y un segundo polinucleótido que comprende una secuencia de nucleótidos de elemento de tallo de nexo II, en el que la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar un elemento de tallo de nexo por enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II y (ii) el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

9. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:

20 un elemento de crtra que es capaz de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con secuencias de ADN que contienen secuencias de motivo adyacente a protoespaciador (MAP) preferentemente en relación con secuencias de ADN sin secuencias de MAP, comprendiendo el elemento de crtra,

un primer polinucleótido que comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y una secuencia de nucleótidos de elemento de tallo de nexo I, y

25 un segundo polinucleótido que comprende una secuencia de nucleótidos de elemento de tallo de nexo II, en el que la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar un elemento de tallo de nexo mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II; y

30 un tercer polinucleótido que comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II; en el que el

35 primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

10. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:

40 un elemento de crtra que es capaz de formar un complejo con una proteína Cas9 para provocar que la proteína Cas9 se una con y corte secuencias de ADN que contienen secuencias de motivo adyacente a protoespaciador (MAP) preferentemente en relación con secuencias de ADN sin secuencias de MAP, comprendiendo el elemento de crtra,

un primer polinucleótido que comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y una secuencia de nucleótidos de elemento de tallo de nexo I, y

45 un segundo polinucleótido que comprende una secuencia de nucleótidos de elemento de tallo de nexo II, en el que la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II son capaces de formar un elemento de tallo de nexo mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexo I y la secuencia de nucleótidos de elemento de tallo de nexo II; y

50 un tercer polinucleótido que comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II; en el que el

55 primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

11. Un sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende: dos o más polinucleótidos que comprenden

un primer polinucleótido, que comprende un extremo 5' y un extremo 3', comprendiendo el primer

polinucleótido una secuencia de nucleótidos de elemento de tallo de nexa I, comprendiendo la secuencia de nucleótidos de elemento de tallo de nexa I, en una dirección de 5' a 3', Nw-N1-N2-Nx, en el que Nw es una primera secuencia de nucleótidos conectiva en la que w es la longitud de la secuencia de nucleótidos conectiva y w es mayor de o igual a dos, N1 es un nucleótido, N2 es un nucleótido y Nx es un primer polinucleótido adyuvante en el que x es la longitud del primer polinucleótido adyuvante y x es mayor de o igual a cero;

un segundo polinucleótido, que comprende un extremo 5' y un extremo 3', comprendiendo el segundo polinucleótido una secuencia de nucleótidos de elemento de tallo de nexa II, comprendiendo la secuencia de nucleótidos de elemento de tallo de nexa II, en una dirección de 5' a 3', Ny-Nc2-Nc1-Nz, en el que Ny es un segundo polinucleótido adyuvante en el que y es la longitud del segundo polinucleótido adyuvante e y es mayor de o igual a cero, Nc2 es un nucleótido que es complementario de N2, Nc1 es un nucleótido que es complementario de N1 y Nz es una segunda secuencia de nucleótidos conectiva en la que z es la longitud de la segunda secuencia de nucleótidos conectiva y z es mayor de o igual a cero; en el que (i) la primera secuencia de nucleótidos de elemento de tallo de nexa I y la segunda secuencia de nucleótidos de elemento de tallo de nexa son capaces de formar un elemento de tallo de nexa por enlaces de hidrógeno de pares de bases entre al menos N1/Nc1 y N2/Nc2 y (ii) el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados.

12. Un sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética que comprende:

un primer polinucleótido, que comprende un extremo 5' y un extremo 3', comprendiendo el primer polinucleótido, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I, una secuencia de nucleótidos de elemento de tallo de nexa I, comprendiendo la secuencia de nucleótidos de elemento de tallo de nexa I, en una dirección de 5' a 3', Nw-N1-N2-Nx, en el que Nw es una primera secuencia de nucleótidos conectiva en la que w es la longitud de la secuencia de nucleótidos conectiva y w es mayor de o igual a dos, N1 es un nucleótido, N2 es un nucleótido y Nx es un primer polinucleótido adyuvante en el que x es la longitud del primer polinucleótido adyuvante y x es mayor de o igual a cero;

un segundo polinucleótido, que comprende un extremo 5' y un extremo 3', comprendiendo el segundo polinucleótido una secuencia de nucleótidos de elemento de tallo de nexa II, comprendiendo la secuencia de nucleótidos de elemento de tallo de nexa II, en una dirección de 5' a 3', Ny-Nc2-Nc1-Nz, en el que Ny es un segundo polinucleótido adyuvante en el que y es la longitud del segundo polinucleótido adyuvante e y es mayor de o igual a cero, Nc2 es un nucleótido que es complementario de N2, Nc1 es un nucleótido que es complementario de N1 y Nz es una segunda secuencia de nucleótidos conectiva en la que z es la longitud de la segunda secuencia de nucleótidos conectiva y z es mayor de o igual a cero; y un tercer polinucleótido, que comprende un extremo 5' y un extremo 3', que comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II;

en el que (i) la primera secuencia de nucleótidos de elemento de tallo de nexa I y la segunda secuencia de nucleótidos de elemento de tallo de nexa II son capaces de formar un elemento de tallo de nexa por enlaces de hidrógeno de pares de bases entre al menos N1/Nc1 y N2/Nc2, (ii) la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II y (iii) el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido son polinucleótidos separados.

13. El sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8 u 11, el primer polinucleótido comprende además, en una dirección de 5' a 3', una secuencia de unión a diana de ADN, un primer elemento de tallo y la secuencia de nucleótidos de elemento de tallo de nexa I, en el que el primer elemento de tallo comprende una primera horquilla.

14. El sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 13, el primer elemento de tallo comprende además un elemento de tallo inferior, un elemento de protuberancia y un elemento de tallo superior, en el que el elemento de tallo inferior está adyacente al elemento de protuberancia, el elemento de protuberancia está adyacente al elemento de tallo superior, el elemento de protuberancia se interpone entre el elemento de tallo inferior y el elemento de tallo superior, y el elemento de tallo superior comprende la primera horquilla.

15. El sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8 u 11, comprendiendo los dos o más polinucleótidos además

un tercer polinucleótido que comprende, en una dirección de 5' a 3', una secuencia de unión a diana de DNA y una primera secuencia de nucleótidos de elemento de tallo II, y

el primer polinucleótido comprende además, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexa I,

en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares

de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II, y

el tercer polinucleótido, el primer polinucleótido y el segundo polinucleótido son polinucleótidos separados, que comprenden cada uno un extremo 5' y un extremo 3'.

5 16. El sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8 u 11, comprendiendo los dos o más polinucleótidos además

un tercer polinucleótido que comprende, en una dirección de 5' a 3', una secuencia de unión a diana de ADN, una secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, y

10 el primer polinucleótido que comprende además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo de nexos I,

15 en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

20 17. El sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10 o 12,

25 comprendiendo el tercer polinucleótido, en una dirección de 5' a 3', la secuencia de unión a diana de ADN; la primera secuencia de nucleótidos de elemento de tallo II que comprende una secuencia de nucleótidos de elemento de tallo inferior II; una secuencia de nucleótidos de elemento de protuberancia II; y una secuencia de nucleótidos de elemento de tallo superior II; y

30 comprendiendo el primer polinucleótido, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I; una secuencia de nucleótidos de elemento de protuberancia I; la primera secuencia de nucleótidos de elemento de tallo I que comprende una secuencia de nucleótidos de elemento de tallo inferior I; y la secuencia de nucleótidos de elemento de tallo de nexos I;

35 en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

40 18. El sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8 u 11, comprendiendo los dos o más polinucleótidos además

un polinucleótido espaciador que comprende una secuencia de unión a diana de ADN, y

45 comprendiendo el primer polinucleótido además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo inferior I, una secuencia de nucleótidos de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de bucle, una secuencia de nucleótidos de elemento de tallo superior II, una secuencia de nucleótidos de elemento de protuberancia II, una secuencia de nucleótidos de elemento de tallo inferior II y la primera secuencia de nucleótidos de elemento de tallo de nexos,

50 en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

55 19. El sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8 u 11, comprendiendo los dos o más polinucleótidos además

un polinucleótido espaciador que comprende una secuencia de unión a diana de ADN, un tercer polinucleótido que comprende una primera secuencia de nucleótidos de elemento de tallo II, y

60 el primer polinucleótido comprende además, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y la secuencia de nucleótidos de elemento de tallo de nexos I,

65 en el que la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II.

20. El sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II modificados por

ingeniería genética de uno cualquiera de los artículos 3, 7, 8 u 11, comprendiendo los dos o más polinucleótidos además

un polinucleótido espaciador que comprende una secuencia de unión a diana de ADN,

un tercer polinucleótido que comprende, en una dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo II que comprende una secuencia de nucleótidos de elemento de tallo inferior II; una secuencia de nucleótidos de elemento de protuberancia II; y una secuencia de nucleótidos de elemento de tallo superior II; comprendiendo el primer polinucleótido, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I; una secuencia de elemento de protuberancia I; la primera secuencia de nucleótidos de elemento de tallo I que comprende una secuencia de nucleótidos de elemento de tallo inferior I; y la secuencia de nucleótidos de elemento de tallo de nexo I,

en el que la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

21. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15 o 19, comprendiendo el primer polinucleótido además un primer polinucleótido accesorio localizado 5' de la primera secuencia de nucleótidos de elemento de tallo I.

22. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 21, comprendiendo el tercer polinucleótido además un segundo polinucleótido accesorio localizado 3' de la primera secuencia de nucleótidos de elemento de tallo II.

23. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15 o 19, comprendiendo el tercer polinucleótido además un segundo polinucleótido accesorio localizado 3' de la primera secuencia de nucleótidos de elemento de tallo II.

24. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 16, 17 o 20, comprendiendo el primer polinucleótido además un primer polinucleótido accesorio localizado 5' de la secuencia de nucleótidos de elemento de tallo superior I.

25. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 24, comprendiendo el tercer polinucleótido además un segundo polinucleótido accesorio localizado 3' de la secuencia de nucleótidos de elemento de tallo superior II.

26. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 16, 17 o 20, comprendiendo el tercer polinucleótido además un segundo polinucleótido accesorio localizado 3' de la secuencia de nucleótidos de elemento de tallo superior II.

27. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3 o 6 a 26, comprendiendo el segundo polinucleótido, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexo II y un segundo elemento de tallo, en el que el segundo elemento de tallo comprende una horquilla.

28. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 27, comprendiendo el segundo polinucleótido, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexo II, el segundo elemento de tallo y un tercer elemento de tallo, en el que el tercer elemento de tallo comprende una horquilla.

29. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3 o 6 a 26, que comprende además el segundo polinucleótido que comprende además, en una dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexo I y una segunda secuencia de nucleótidos de elemento de tallo I, y

un primer polinucleótido adjunto que comprende una segunda secuencia de nucleótidos de elemento de tallo II, en el que la segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II son capaces de formar un segundo elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II.

30. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 29, comprendiendo el primer polinucleótido adjunto además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la segunda secuencia de nucleótidos de elemento de tallo II, en el que el extremo 5' de la secuencia de nucleótidos de elemento de bucle está unido covalentemente con el extremo 3' de la segunda secuencia de nucleótidos de elemento de tallo I.

31. El sistema de polinucleótidos de nexo dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 29 o 30, comprendiendo el primer polinucleótido adjunto, en una dirección de 5' a 3', la segunda secuencia de nucleótidos de elemento de tallo II y una tercera secuencia de nucleótidos de elemento de tallo I, y

un segundo polinucleótido adjunto que comprende, en una dirección de 5' a 3', una tercera secuencia de nucleótidos de elemento de tallo II, en el que la tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II son capaces de formar un tercer elemento de tallo

mediante enlaces de hidrógeno de pares de bases entre la tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II.

32. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 31, comprendiendo el segundo polinucleótido adjunto además, en una dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la tercera secuencia de nucleótidos de elemento de tallo II, en el que el extremo 5' de la secuencia de nucleótidos de elemento de bucle está unido covalentemente con el extremo 3' de la tercera secuencia de nucleótidos de elemento de tallo I.

33. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de cualquiera de los artículos 3 o 6 a 32, comprendiendo el primer polinucleótido además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexos I.

34. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 33, comprendiendo el segundo polinucleótido además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexos II.

35. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de cualquiera de los artículos 3 o 6 a 32, comprendiendo el segundo polinucleótido además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexos II.

36. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 33, en el que el primer polinucleótido adyuvante comprende un sitio de unión para una proteína de unión a polinucleótido monocatenario.

37. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 36, en el que la proteína de unión a polinucleótido monocatenario es una proteína de unión a ARN monocatenario.

38. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 34 o 35, en el que el segundo polinucleótido adyuvante comprende un sitio de unión para una proteína de unión a polinucleótido monocatenario.

39. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 38, en el que la proteína de unión a polinucleótido monocatenario es una proteína de unión a ARN monocatenario.

40. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 34, en el que el primer polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento de unión efector I y el segundo polinucleótido adyuvante comprende una secuencia de nucleótidos de elemento de unión efector II y la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector II son capaces de formar un elemento de unión efector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector I.

41. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 40, en el que el elemento de unión efector es un polinucleótido bicatenario.

42. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 41, en el que una proteína efectora es capaz de unirse con el elemento de unión efector.

43. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 42, en el que el elemento de unión efector es un ARN bicatenario y la proteína efectora es una proteína de unión a ARN bicatenario capaz de unirse con el elemento de unión efector.

44. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 43, en el que la proteína efectora es una variante catalíticamente inactiva de una proteína seleccionada del grupo que consiste en Cas5, Cas6 y Csy4.

45. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 43, en el que la proteína efectora comprende al menos un dominio de dedo de cinc.

46. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de cualquiera de los artículos 3 o 6 a 32, en el que el primer polinucleótido adyuvante comprende una horquilla.

47. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 46, en el que el segundo polinucleótido adyuvante comprende una horquilla.

48. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de cualquiera de los artículos 3 o 6 a 32, en el que el segundo polinucleótido adyuvante comprende una horquilla.

49. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 33 o 34, en el que el primer polinucleótido adyuvante comprende además una primera secuencia de nucleótidos de afinidad.

50. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 49, en el que un polipéptido está unido covalentemente con la primera secuencia de nucleótidos de afinidad.

51. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 49, en el que la primera secuencia de nucleótidos de afinidad comprende un ligando.

52. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 51, en el que el ligando está unido covalentemente con la primera secuencia de nucleótidos de afinidad.
53. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 49, en el que la primera secuencia de nucleótidos de afinidad comprende un resto de unión a ligando.
54. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 53, en el que el resto de unión a ligando está unido covalentemente con la primera secuencia de nucleótidos de afinidad.
55. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 34 o 35, en el que el segundo polinucleótido adyuvante comprende además una segunda secuencia de nucleótidos de afinidad.
56. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 55, en el que un polipéptido está unido covalentemente con la segunda secuencia de nucleótidos de afinidad.
57. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 55, en el que la segunda secuencia de nucleótidos de afinidad comprende un ligando.
58. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 57, en el que el ligando está unido covalentemente con la segunda secuencia de nucleótidos de afinidad.
59. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 55, en el que la segunda secuencia de nucleótidos de afinidad comprende un resto de unión a ligando.
60. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 59, en el que el resto de unión a ligando está unido covalentemente con la segunda secuencia de nucleótidos de afinidad.
61. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 6, 7, 8, 9 o 10, en el que (i) la secuencia de nucleótidos de elemento de tallo de nexos I comprende dos nucleótidos contiguos y la secuencia de nucleótidos de elemento de tallo de nexos II comprende dos nucleótidos contiguos complementarios de los dos nucleótidos contiguos de la secuencia de nucleótidos de elemento de tallo de nexos I y (ii) el elemento de tallo de nexos comprende enlaces de hidrógeno de pares de bases entre los dos nucleótidos contiguos de la secuencia de nucleótidos de elemento de tallo de nexos I y los dos nucleótidos contiguos de la secuencia de nucleótidos de elemento de tallo de nexos II.
62. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 61, en el que los dos nucleótidos contiguos de la secuencia de nucleótidos de elemento de tallo de nexos I se seleccionan del grupo que consiste en un nucleótido con una base de purina y un nucleótido con una base de pirimidina, un nucleótido con una base de pirimidina y un nucleótido con una base de purina, y un nucleótido con una base de purina y un nucleótido con una base de pirimidina.
63. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 6 a 10 o 13 a 60, en el que la secuencia de nucleótidos de elemento de tallo de nexos I comprende dos nucleótidos contiguos, un primer nucleótido (N1) contiguo a un segundo nucleótido (N2) en el que N2 está 3' de N1, y la secuencia de nucleótidos de elemento de tallo de nexos II comprende dos nucleótidos contiguos, un primer nucleótido (Nc1) contiguo a un segundo nucleótido (Nc2) en el que Nc1 está 3' de Nc2, en el que N1-N2 de la secuencia de nucleótidos de elemento de tallo de nexos I son complementarios de Nc1-Nc2 de la secuencia de nucleótidos de elemento de tallo de nexos II.
64. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 11, 12 o 63, en el que N1 tiene una base de purina, N2 tiene una base de pirimidina, Nc1 tiene una base de pirimidina y Nc2 tiene una base de pirimidina.
65. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 64, la base de purina de N1 es adenina, la base de purina de N2 es adenina, la base de pirimidina de Nc1 es timina o uracilo y la base de pirimidina de Nc2 es timina o uracilo, en el que la base de pirimidina de Nc1 y la base de pirimidina de Nc2 son iguales.
66. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 64, la base de purina de N1 es guanosina, la base de purina de N2 es guanosina, la base de pirimidina de Nc1 es citosina y la base de pirimidina de Nc2 es citosina.
67. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 11, 12 o 63, en el que N1 tiene una base de pirimidina, N2 tiene una base de pirimidina, Nc1 tiene una base de purina y Nc2 tiene una base de purina.
68. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 67, la base de pirimidina de N1 es timina o uracilo, la base de pirimidina de N2 es timina o uracilo, la base de purina de Nc1 es adenina y la base de purina de Nc2 es adenina, en el que la base de pirimidina de N1 y la base de pirimidina de N2 son iguales.
69. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 11, 12 o 63, en el que N1 tiene una base de pirimidina, N2

tiene una base de purina, Nc1 tiene una base de purina y Nc2 tiene una base de pirimidina.

70. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 11, 12 o 63, en el que N1 tiene una base de purina, N2 tiene una base de pirimidina, Nc1 tiene una base de pirimidina y Nc2 tiene una base de purina.

71. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8, 11, 13, 14 o 18, en el que el primer polinucleótido y el segundo polinucleótido comprenden bases de ARN, bases de ADN o una combinación de bases de ARN y bases de ADN.

72. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, en el que el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido comprenden bases de ARN, bases de ADN o una combinación de bases de ARN y bases de ADN.

73. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 70, en el que al menos uno de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética comprende bases de ARN, bases de ADN o una combinación de bases de ARN y bases de ADN.

74. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8, 11, 13, 14 o 18, en el que el primer polinucleótido y el segundo polinucleótido comprenden ARN.

75. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, en el que el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido comprenden ARN.

76. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 70, en el que los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética comprenden ARN.

77. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 76, en el que al menos un polinucleótido de CRISPR-Cas9 de tipo II modificado por ingeniería genética comprende una secuencia 5' terminal y/o una secuencia 3' terminal, y la secuencia 5' terminal, la secuencia 3' terminal, o la secuencia 5' terminal y la secuencia 3' terminal comprende un resto de resistencia a exonucleasa.

78. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8, 11, 13, 14 o 18, en el que el primer polinucleótido y el segundo polinucleótido comprenden ácidos nucleicos de fosforotioato, ácidos nucleicos peptídicos, ácido nucleico de treosa o combinaciones de los mismos.

79. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, en el que el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido comprenden ácidos nucleicos de fosforotioato, ácidos nucleicos peptídicos, ácido nucleico de treosa o combinaciones de los mismos.

80. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 70, en el que al menos uno de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética comprende ácidos nucleicos de fosforotioato, ácidos nucleicos peptídicos, ácido nucleico de treosa o combinaciones de los mismos.

81. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8, 11, 13, 14 o 18, en el que al menos uno del primer polinucleótido y el segundo polinucleótido es circular.

82. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, en el que al menos uno del primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido es circular.

83. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 19 o 20, en el que al menos uno del primer polinucleótido, el segundo polinucleótido, el tercer polinucleótido y el polinucleótido espaciador son circulares.

84. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8, 11, 13, 14 o 18 que comprende además una proteína Cas9, en el que el primer polinucleótido y el segundo polinucleótido comprenden ARN, y el primer polinucleótido y el segundo polinucleótido están en complejo con la proteína Cas9.

85. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, que comprende además una proteína Cas9, en el que el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido comprenden ARN, y el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido están en complejo con la proteína Cas9.

86. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 83, que comprende además una proteína Cas9, en el que al menos uno de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética están en complejo con la proteína Cas9.

87. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 83, que comprende además una secuencia codificante

de proteína Cas9.

88. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 83, que comprende además una proteína Cas9.

5 89. El sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 88, que comprende además la proteína Cas9 en complejo con los polinucleótidos del sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética.

10 90. Un procedimiento para fabricar el sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 3, 7, 8, 11, 13, 14 o 18, en el que el procedimiento comprende sintetizar químicamente el primer polinucleótido, el segundo polinucleótido o tanto el primer polinucleótido como el segundo polinucleótido.

15 91. Un procedimiento para fabricar el sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, el procedimiento comprende sintetizar químicamente el primer polinucleótido, el segundo polinucleótido, el tercer polinucleótido o cualquier combinación de los mismos.

92. Un procedimiento para fabricar el sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 89, en el que el procedimiento comprende sintetizar químicamente al menos uno de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética.

20 93. Uno o más casetes de expresión que comprenden promotores de transcripción unidos operativamente con una o más secuencias de ADN que codifican uno o más de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 76.

25 94. El o los casetes de expresión del artículo 93, que comprenden además un casete de expresión que comprende una secuencia de ADN que comprende un promotor de la transcripción y una secuencia reguladora de la traducción unidos operativamente a una secuencia codificante de proteína Cas9.

30 95. Uno o más casetes de expresión que comprenden un promotor de la transcripción unido operativamente a una primera secuencia de ADN que codifica el primer polinucleótido del artículo 75, un promotor de la transcripción unido operativamente a una segunda secuencia de ADN que codifica el segundo polinucleótido del artículo 75 y un promotor de la transcripción unido operativamente a una tercera secuencia de ADN que codifica el tercer polinucleótido del artículo 75.

96. El o los casetes de expresión del artículo 95, que comprenden además un casete de expresión que comprende una cuarta secuencia de ADN que comprende un promotor de la transcripción y una secuencia reguladora de la traducción unidos operativamente a una secuencia codificante de proteína Cas9.

35 97. Una célula recombinante, que comprende el o los casetes de expresión del artículo 93.

98. Una célula recombinante, que comprende el o los casetes de expresión del artículo 95.

99. La célula recombinante del artículo 97, en la que al menos uno de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética se transcribe a partir de una secuencia de ADN.

40 100. La célula recombinante del artículo 98, en la que el primer polinucleótido se transcribe a partir de la primera secuencia de ADN, el segundo polinucleótido se transcribe a partir de la segunda secuencia de ADN y el tercer polinucleótido se transcribe a partir del tercer ADN.

101. Una célula recombinante, que comprende el o los casetes de expresión del artículo 94.

102. Una célula recombinante, que comprende el o los casetes de expresión del artículo 96.

45 103. La célula recombinante del artículo 101, en la que al menos uno de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética se transcribe a partir de una secuencia de ADN y la proteína Cas9 se expresa en la célula recombinante.

104. La célula recombinante del artículo 102, en la que el primer polinucleótido se transcribe a partir de la primera secuencia de ADN, el segundo polinucleótido se transcribe a partir de la segunda secuencia de ADN, el tercer polinucleótido se transcribe a partir del tercer ADN y la proteína Cas9 se expresa en la célula recombinante.

50 105. La célula recombinante de uno cualquiera de los artículos 97 a 104, en el que la célula es una célula vegetal, una célula de levadura, una célula bacteriana, una célula de insecto, una célula de alga o una célula de mamífero.

55 106. La célula recombinante del artículo 105, en el que la célula es una célula vegetal de una planta seleccionada del grupo que consiste en sorgo, soja, arroz, trigo, maíz, *Brassica* oleaginosa, algodón, caña de azúcar, girasol, mijo y alfalfa.

107. La célula recombinante de uno cualquiera de los artículos 97 a 106, en la que al menos un casete de expresión del o los casetes de expresión se integra en un sitio en ADN genómico de la célula recombinante.

60 108. La célula recombinante del artículo 107, en la que un casete de expresión que comprende la segunda secuencia de ADN que codifica el segundo polinucleótido se integra en un sitio en ADN genómico de la célula recombinante y un casete de expresión que comprende la secuencia de ADN que comprende el promotor de transcripción y la secuencia reguladora de la traducción unida operativamente a la secuencia codificante de la proteína Cas9 se integra en un sitio en ADN genómico de la célula recombinante.

109. La célula recombinante del artículo 107, en la que un vector de expresión comprende al menos uno del o los casetes de expresión.

65 110. La célula recombinante de uno cualquiera de los artículos 97 a 106, en la que un vector de expresión comprende al menos uno del o los casetes de expresión.

111. Una célula recombinante, que comprende uno o más ADN que codifican uno o más de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 76.
112. Una célula recombinante, que comprende una primera secuencia de ADN que codifica el primer polinucleótido del artículo 75, una segunda secuencia de ADN que codifica el segundo polinucleótido del artículo 75 y una tercera secuencia de ADN que codifica el tercer polinucleótido del artículo 75.
113. Una célula recombinante, que comprende el sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 76 y una proteína Cas9 en la que uno o más de los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética y la proteína Cas9 están en un complejo.
114. Una célula recombinante, que comprende el sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 75 y una proteína Cas9 en la que el primer polinucleótido, el segundo polinucleótido, el tercer polinucleótido y la proteína Cas9 están en un complejo.
115. Un procedimiento para fabricar los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 99, que comprende aislar los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética de la célula recombinante.
116. Un procedimiento para fabricar el primer polinucleótido y el segundo polinucleótido del artículo 100, que comprende aislar el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido de la célula recombinante.
117. Un kit, que comprende uno o más polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética del sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 83 e instrucciones para el uso del sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética.
118. El kit del artículo 117, que comprende además una proteína Cas9.
119. Un kit, que comprende el primer polinucleótido, el segundo polinucleótido y el tercer polinucleótido del sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20 e instrucciones para el uso del sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética.
120. El kit del artículo 119, que comprende además una proteína Cas9.
121. Un kit, que comprende uno o más casetes de expresión del artículo 93, 94, 95 o 96 e instrucciones para el uso del sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética.
122. El kit del artículo 121, en el que un vector de expresión comprende al menos uno del o los casetes de expresión.
123. Una composición farmacéutica, que comprende:
una proteína Cas9; el sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 76, en el que uno o más de los polinucleótidos del sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética están en complejo con la proteína Cas9; y un excipiente farmacéuticamente aceptable.
124. Una composición farmacéutica, que comprende:
una proteína Cas9; el sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética del artículo 75,
en el que el primer polinucleótido y el segundo polinucleótido están en un complejo con una proteína Cas9; y un excipiente farmacéuticamente aceptable.
125. Una composición farmacéutica, que comprende:
el o los casetes de expresión de uno cualquiera del artículo 93, 94, 95, 96; y un excipiente farmacéuticamente aceptable.
126. Un procedimiento para modificación de ADN que comprende:
poner en contacto una secuencia diana de ADN en el ADN con un sistema de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 83, y una proteína Cas9,
en el que los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética y la proteína Cas9 forman un complejo que se une con y corta la secuencia diana de ADN dando como resultado una modificación del ADN.
127. El procedimiento del artículo 126, en el que el procedimiento es un procedimiento *in vivo* de modificación de ADN genómico en una célula eucariota.
128. El procedimiento del artículo 126, en el que el procedimiento es un procedimiento *ex vivo* de modificación de ADN genómico en una célula eucariota aislada de un organismo.
129. El procedimiento de uno cualquiera de los artículos 126 a 128, que comprende además poner en contacto la secuencia diana de ADN en el ADN genómico con un molde de ADN donante, en el que la modificación comprende que al menos una parte del molde de ADN donante se integre en la secuencia diana de ADN.
130. El procedimiento del artículo 129, en el que un vector de dirección comprende el molde de ADN donante.
131. Un procedimiento para modificación de ADN que comprende:

poner en contacto una secuencia diana de ADN en el ADN con un sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, y una proteína Cas9,

5 en el que el primer polinucleótido, el segundo polinucleótido, el tercer polinucleótido y la proteína Cas9 forman un complejo que se une con y corta la secuencia diana de ADN dando como resultado una modificación del ADN.

132. El procedimiento del artículo 131, en el que el procedimiento es un procedimiento *in vivo* de modificación de ADN genómico en una célula eucariota.

10 133. El procedimiento del artículo 131, en el que el procedimiento es un procedimiento *ex vivo* de modificación de ADN genómico en una célula eucariota aislada de un organismo.

134. El procedimiento de uno cualquiera de los artículos 131 a 133, que comprende además poner en contacto la secuencia diana de ADN en el ADN genómico con un molde de ADN donante en el que la modificación comprende que al menos una parte del molde de ADN donante se integra en la secuencia diana de ADN.

15 135. El procedimiento del artículo 134 en el que un vector de dirección comprende el molde de ADN donante.

136. Un procedimiento para modulación de la expresión de un gen que comprende elementos reguladores de la transcripción, que comprende:

poner en contacto una secuencia diana de ADN en el gen con un sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II modificados por ingeniería genética de uno cualquiera de los artículos 1 a 83, y una proteína Cas9,

20 en el que los polinucleótidos de CRISPR-Cas9 de tipo II modificados por ingeniería genética y la proteína Cas9 forman un complejo que se une con la secuencia diana de ADN dando como resultado modulación de la expresión del gen.

137. El procedimiento del artículo 136, en el que la proteína Cas9 es una Cas9 que es deficiente en nucleasa.

25 138. El procedimiento del artículo 136 o 137, en el que el complejo comprende además una proteína de fusión.

139. Un procedimiento para modulación de la expresión de un gen que comprende elementos reguladores de la transcripción, que comprende:

poner en contacto una secuencia diana de ADN en el gen con un sistema de polinucleótidos de nexos dividido asociados con CRISPR-Cas9 de tipo II de uno cualquiera de los artículos 6, 9, 10, 12, 15, 16, 17, 19 o 20, y una proteína Cas9,

30 en el que el primer polinucleótido, el segundo polinucleótido, el tercer polipéptido y la proteína Cas9 forman un complejo que se une con la secuencia diana de ADN dando como resultado modulación de la expresión del gen.

140. El procedimiento del artículo 139, en el que la proteína Cas9 es una Cas9 que es deficiente en nucleasa.

35 141. El procedimiento del artículo 139 o 140, en el que el complejo comprende además una proteína de fusión de Cas9.

142. Un vector, que comprende uno o más casetes de expresión de uno cualquiera de los artículos 93, 94, 95 o 96.

40 143. El vector del artículo 142, en el que el vector se selecciona del grupo que consiste en vectores bacterianos, vectores de levadura, vectores de algas, vectores de células de insectos, vectores de mamíferos y vectores víricos.

LISTADO DE SECUENCIAS

<110> Caribou Biosciences, Inc.

<120> COMPOSICIONES DE CRISPR-CAS9 MODIFICADAS POR INGENIERÍA GENÉTICA Y PROCEDIMIENTOS DE USO

45 <130> CBI017.30

<150> 62/202.715

<151> 07/08/2015

<150> 62/209.334

<151> 24/08/2015

50 <160> 128

<170> PatentIn versión 3.5

<210> 1

<211>31

<212> ARN

55 <213> Secuencia artificial

ES 2 705 540 T3

	<220>		
	<223> Procedente de <i>Streptococcus pyogenes</i>		
	<400> 1		
	caaaacagca uagcaaguua aaauaaggcu a	31	
5	<210>2		
	<211> 46		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
10	<223> Procedente de <i>Streptococcus pyogenes</i>		
	<400> 2		
	guccguuauc aacuugaaaa aguggcaccg agucggugcu uuuuuu	46	
	<210>3		
15	<211> 33		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> Procedente de <i>Streptococcus thermophilus</i> CRISPR-I		
20	<400> 3		
	uaaaucuugc agaagcuaca acgauaaggc uuc	33	
	<210>4		
	<211> 54		
	<212> ARN		
	<213> Secuencia artificial		
25	<220>		
	<223> Procedente de <i>Streptococcus thermophilus</i> CRISPR-I		
	<400> 4		
	augccgaaau caacaccug ucauuuuuug gcaggguguu uucguuuuu uuuu	54	
30	<210>5		
	<211> 32		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> Procedente de <i>Listeria innocua</i>		
35	<400> 5		
	caaaauaaca uagcaaguua aaauaaggcu uu	32	
	<210>6		
	<211> 48		
40	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> Procedente de <i>Listeria innocua</i>		
	<400> 6		
	guccguuauc aacuuuuuau uaaguagcgc uguuucggcg cuuuuuuu	48	
45	<210>7		
	<211> 40		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
50	<223> Procedente de <i>Neisseria meningitidis</i>		

ES 2 705 540 T3

	<400> 7 cugcgaaaug agaaccguug cuacaauaag gccgucugaa	40
5	<210>8 <211> 56 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Neisseria meningitidis</i>	
10	<400> 8 aagaugugcc gcaacgcucu gccccuuaaa gcuucugcuu uaaggggcau cguuuu	56
15	<210>9 <211> 42 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Streptococcus gallolyticus</i>	
20	<400> 9 uugggacua ucgaaacaac acagcgaguu aaaauaaggc uu	42
25	<210> 10 <211> 50 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Streptococcus gallolyticus</i>	
30	<400> 10 uguccguaca caacuuguaa aaguggcacc cgauucgggu gcguuuuuuu	50
35	<210> 11 <211> 46 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Staphylococcus aureus</i>	
40	<400> 11 auuguacuua uaccuuaaa uacagaau cuuaaaaca ggcaaa	46
45	<210> 12 <211> 39 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Staphylococcus aureus</i>	
50	<400> 12 augccguguu uaucucguca acuuguuggc gagauuuuu	39
	<210> 13 <211> 48 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Corynebacterium diphtheriae</i>	
	<400> 13 agucacuaac uaaauuaau agaacugaac cucaguaagc auuggcuc	48

ES 2 705 540 T3

	<210> 14	
	<211> 58	
	<212> ARN	
	<213> Secuencia artificial	
5	<220>	
	<223> Procedente de <i>Corynebacterium diphtheriae</i>	
	<400> 14	
	guuuccaaug uugauugcuc cgccggugcu ccuuuuuuu aagggcgccg gcuuucuu	58
10	<210> 15	
	<211> 46	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Procedente de <i>Parvibaculum lavamentivorans</i>	
15	<400> 15	
	uagcaaaucg agagggcguc gcuuuucgca agcaaauga cccuu	46
20	<210> 16	
	<211> 77	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Procedente de <i>Parvibaculum lavamentivorans</i>	
	<400> 16	
	gugcggguc ggcaucccaa ggcagcugc cgguuuuau cgaaaaggcc caccgcaagc	60
	agcgcguggg ccuuuuu	77
25	<210> 17	
	<211>51	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
30	<223> Procedente de <i>Campylobacter lari</i>	
	<400> 17	
	aaucucugcu aaagaaauu aaaaagagac uaaaauaagu gguuuuuggu c	51
35	<210> 18	
	<211> 73	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> Procedente de <i>Campylobacter lari</i>	
	<400> 18	
	auccacgcag gguuacaauc ccuuuaaaac cauuaaaau caaauaaacu agguuguauc	60
40	aacuuaguuu uuu	73
	<210> 19	
	<211> 50	
	<212> ARN	
	<213> Secuencia artificial	
45	<220>	
	<223> Procedente de <i>Neisseria cinerea</i>	

ES 2 705 540 T3

	<400> 19 auugucgcac ugcgaaauga gaaccguugc uacaauaagg ccgucugaaa	50
5	<210> 20 <211> 89 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Neisseria cinerea</i> <400> 20 agaugugccg caacgcucug ccccuaaaag cuucugcuuu aaggggcauc guuuuuuucg	60
10	guuaaaaaug ccgucugaaa ccgguuuuu	89
	<210>21 <211> 47 <212> ARN <213> Secuencia artificial	
15	<220> <223> Procedente de <i>Streptococcus pasteurianus</i> <400> 21 cuugcaggu uacuuaaauc uugcugagcc uacaagaua aggcuuu	47
20	<210> 22 <211> 43 <212> ARN <213> Secuencia artificial <220> <223> Procedente de <i>Streptococcus pasteurianus</i>	
25	<400> 22 augccgaauu caagcacccc auguuuugac augaggugcu uuu	43
30	<210> 23 <211> 23 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico <400> 23 agtaataata cgactcacta tag	23
35	<210> 24 <211> 24 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	
40	<400> 24 aagcaccgac tcggtgccac tttt	24
45	<210> 25 <211> 58 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	

ES 2 705 540 T3

	<400> 25 taatacgact cactatagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctt	58
5	<210> 26 <211> 49 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	
10	<400> 26 taatacgact cactatagca ggacagcata gcaagtgag ataaggcta	49
15	<210> 27 <211> 49 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	
20	<400> 27 tagccttacc tcaacttgct atgctgtcct gctatagtga gtcgtatta	49
25	<210> 28 <211> 55 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	
30	<400> 28 taatacgact cactataggg gccactaggg acaggatgac tcagagctat gctgt	55
35	<210> 29 <211> 55 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	
40	<400> 29 acagcatagc tctgagacat cctgtcccta gtggccccta tagtgagtcg tatta	55
45	<210> 30 <211> 20 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	
50	<400> 30 ccccgttctc ctgtggattc	20
	<210>31 <211> 20 <212> ADN <213> Secuencia artificial <220> <223> Cebador oligonucleotídico	
	<400> 31 atcctctctg gctccatcgt	20

ES 2 705 540 T3

	<210> 32	
	<211> 52	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Cebador oligonucleotídico	
	<400> 32	
	cactctttcc ctacacgacg ctctccgat ctctggcaa ggagagagat gg	52
10	<210> 33	
	<211> 48	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
15	<400> 33	
	ggagttcaga cgtgtgctct tccgatctta tattcccagg gccggtta	48
20	<210> 34	
	<211> 57	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 34	
	caagcagaag acggcatacg agattacgtg atgtgactgg agttcagacg tgtgctc	57
25	<210> 35	
	<211> 58	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Cebador oligonucleotídico	
	<400> 35	
	aatgatacgg cgaccaccga gatctacacc gtctaataca ctcttcct acacgacg	58
35	<210> 36	
	<211> 58	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 36	
40	aatgatacgg cgaccaccga gatctacact ctctccgaca ctcttcct acacgacg	58
45	<210> 37	
	<211> 58	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 37	
	aatgatacgg cgaccaccga gatctacact cgactagaca ctcttcct acacgacg	58
50	<210> 38	
	<211> 58	
	<212> ADN	

ES 2 705 540 T3

	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
5	<400> 38 aatgatacgg cgaccaccga gatctacact tctagctaca ctctttccct acacgacg	58
	<210> 39	
	<211> 58	
	<212> ADN	
	<213> Secuencia artificial	
10	<220>	
	<223> Cebador oligonucleotídico	
	<400> 39 aatgatacgg cgaccaccga gatctacacc ctagagtaca ctctttccct acacgacg	58
	<210> 40	
15	<211> 58	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
20	<400> 40 aatgatacgg cgaccaccga gatctacacc tattaagaca ctctttccct acacgacg	58
	<210> 41	
25	<211> 62	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 41 ggcagtagcc ttatctcaac ttgctatgct gtcctgtttc caggacagca tagctctgag	60
	ac	62
30	<210> 42	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
35	<223> Cebador oligonucleotídico	
	<400> 42 ggcagtagcc ttatctcaac	20
	<210> 43	
40	<211> 63	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 43 taatacgact cactataggg aggtccgtta tcaacttgaa aaagtggcac cgagtcggtg	60
45	ctt	63

ES 2 705 540 T3

	<210> 44		
	<211> 67		
	<212> ADN		
	<213> Secuencia artificial		
5	<220>		
	<223> Cebador oligonucleotídico		
	<400> 44		
	ggcagtgaac tagccttatac tcaacttgct atgctgtcct gttccagga cagcatagct		60
	ctgagac		67
10	<210> 45		
	<211> 20		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Cebador oligonucleotídico		
15	<400> 45		
	ggcagtgaac tagccttatac	20	
	<210> 46		
	<211> 68		
20	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Cebador oligonucleotídico		
	<400> 46		
	taatacgact cactataggg agctaaggctc cgttatcaac ttgaaaaagt ggcaccgagt		60
	cggtgctt		68
25	<210> 47		
	<211> 56		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
30	<223> Cebador oligonucleotídico		
	<400> 47		
	taatacgact cactataggg gccactaggg acaggatgctc tcagagctat gctgctc	56	
	<210> 48		
	<211> 56		
35	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Cebador oligonucleotídico		
	<400> 48		
40	taatacgact cactataggt tgtgttcca taaactggctc tcagagctat gctgctc	56	
	<210> 49		
	<211> 56		
	<212> ADN		
	<213> Secuencia artificial		
45	<220>		

	<223> Cebador oligonucleotídico	
	<400> 49	
	taatacgact cactatagcc cgccaccacc aggatgtgctc tcagagctat gctgctc	56
5	<210> 50 <211> 56 <212> ADN <213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
10	<400> 50	
	taatacgact cactatagcc agccagcatg atgagacgctc tcagagctat gctgctc	56
	<210>51 <211> 20 <212> ADN <213> Secuencia artificial	
15	<220>	
	<223> Cebador oligonucleotídico	
	<400> 51	
	aagcaccgac tcggtgccac	20
20	<210> 52 <211> 52 <212> ADN <213> Secuencia artificial	
	<220>	
25	<223> Cebador oligonucleotídico	
	<400> 52	
	cactctttcc ctacacgacg ctcttccgat ctacatgcac acccatgttt tg	52
	<210> 53 <211> 48 <212> ADN <213> Secuencia artificial	
30	<220>	
	<223> Cebador oligonucleotídico	
	<400> 53	
	ggagttcaga cgtgtgctct tccgatctaa catttccagg tgacaggc	48
	<210> 54 <211> 50 <212> ADN <213> Secuencia artificial	
40	<220>	
	<223> Cebador oligonucleotídico	
	<400> 54	
	cactctttcc ctacacgacg ctcttccgat ctgttccgac gctcctttaa	50
	<210> 55 <211> 48 <212> ADN <213> Secuencia artificial	
45	<220>	
	<223> Cebador oligonucleotídico	
50	<400> 55	

ES 2 705 540 T3

	ggagttcaga cgtgtgctct tccgatctca gatgcatga cctttgtg	48
	<210> 56	
	<211> 52	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 56	
	cactctttcc ctacacgacg ctcttccgat ctaagaaagg caagaagcct gg	52
10	<210> 57	
	<211> 48	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> Cebador oligonucleotídico	
	<400> 57	
	ggagttcaga cgtgtgctct tccgatctgc tggcctgaga cattccta	48
	<210> 58	
20	<211> 57	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 58	
25	tagccttatc tcaacttgc atgctgtcct gttccagga cagcatagct ctgagac	57
	<210> 59	
	<211> 57	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Cebador oligonucleotídico	
	<400> 59	
	taatacgact cactataggg gccactaggg acaggatgtc tcagagctat gcagtcc	57
35	<210> 60	
	<211> 60	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
40	<400> 60	
	cagtagcctt atctcaact gctatgcagt cctgtttcca ggactgcata gctctgagac	60
	<210>61	
45	<211> 40	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 61	
	ctgcctatac ggcagtagcc ttatctcaac ttgctatgca	40
50	<210> 62	

ES 2 705 540 T3

	<211> 57	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador oligonucleotídico	
	<400> 62	
	taatacgact cactatagct gccgtatagg caggtccgtt atcaactga aaaagtg	57
	<210> 63	
10	<211> 43	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 63	
15	aagcaccgac tcggtgccac ttttcaagt tgataacgga cct	43
	<210> 64	
	<211> 60	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> Cebador oligonucleotídico	
	<400> 64	
	gtctagcctt atctcaactt gctatgcagt cctgtttcca ggactgcata gctctgagac	60
	<210> 65	
25	<211> 40	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 65	
30	ctgcctatac ggcagtgctt agccttatct caacttgcta	40
	<210> 66	
35	<211> 57	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 66	
	taatacgact cactatagct gccgtatagg cagagacagt ccggtatcaa ctgaaa	57
40	<210> 67	
	<211> 48	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
45	<223> Cebador oligonucleotídico	
	<400> 67	
	aagcaccgac tcggtgccac ttttcaagt tgataacgga ctgtctct	48
	<210> 68	
50	<211>41	
	<212> ARN	
	<213> Secuencia artificial	

ES 2 705 540 T3

	<220>		
	<223> ARN asociado con Cas9 de nexos dividido		
	<400> 68		
	guccguuauc aacuugaaaa aguggcaccg agucggugcu u	41	
5	<210> 69		
	<211> 32		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
10	<223> ARN asociado con Cas9 de nexos dividido		
	<400> 69		
	gcaggacagc auagcaaguu gagauaaggc ua	32	
	<210> 70		
15	<211> 38		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexos dividido		
	<400> 70		
20	ggggccacua gggacaggau gucucagagc uaugcugu	38	
	<210> 71		
	<211> 82		
	<212> ARN		
	<213> Secuencia artificial		
25	<220>		
	<223> ARN asociado con Cas9 de nexos dividido		
	<400> 71		
	ggggccacua gggacaggau gucucagagc uaugcugucc uggaaacagg acagcauagc	60	
	aaguugagau aaggcuacug cc	82	
30	<210> 72		
	<211> 82		
	<212> ARN		
	<213> Secuencia artificial		
35	<220>		
	<223> ARN asociado con Cas9 de nexos dividido		
	<400> 72		
	guuuguguuu ccauaaacug gucucagagc uaugcugucc uggaaacagg acagcauagc	60	
	aaguugagau aaggcuacug cc	82	
	<210> 73		
40	<211> 82		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexos dividido		
	<400> 73		
	gcccgcacc accaggaugu gucucagagc uaugcugucc uggaaacagg acagcauagc	60	
45	aaguugagau aaggcuacug cc	82	

ES 2 705 540 T3

	<210> 74		
	<211> 82		
	<212> ARN		
	<213> Secuencia artificial		
5	<220>		
	<223> ARN asociado con Cas9 de nexa dividido		
	<400> 74		
	ggcagccagc augaugagac gucucagagc uaugcugucc uggaaacagg acagcauagc		60
	aaguugagau aaggcuacug cc		82
10	<210> 75		
	<211> 46		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexa dividido		
15	<400> 75		
	ggcagguccg uuaucaacu gaaaaagugg caccgagucg gugcuu	46	
20	<210> 76		
	<211> 87		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexa dividido		
	<400> 76		
	ggggccacua gggacaggau gucucagagc uaugcugucc uggaaacagg acagcauagc		60
	aaguugagau aaggcuaguu cacugcc		87
25	<210> 77		
	<211> 87		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
30	<223> ARN asociado con Cas9 de nexa dividido		
	<400> 77		
	guuuguguuu ccauaaacug gucucagagc uaugcugucc uggaaacagg acagcauagc		60
	aaguugagau aaggcuaguu cacugcc		87
35	<210> 78		
	<211> 87		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexa dividido		
	<400> 78		
	gcccgccacc accaggaugu gucucagagc uaugcugucc uggaaacagg acagcauagc		60
40	aaguugagau aaggcuaguu cacugcc		87

ES 2 705 540 T3

	<210> 79	
	<211> 87	
	<212> ARN	
	<213> Secuencia artificial	
5	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 79	
	ggcagccagc augaugagac gucucagagc uaugcugucc uggaaacagg acagcauagc	60
	aaguugagau aaggcuaguu cacugcc	87
10	<210> 80	
	<211> 68	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
15	<400> 80	
	gggcagugaa cuagccuuau cucaacuugc uaugcugucc uguuuccagg acagcauagc	60
	ucugagac	68
20	<210>81	
	<211> 77	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 81	
	ggggccacua gggacaggau gucucagagc uaugcugucc uggaaacagg acagcauagc	60
	aaguugagau aaggcua	77
25	<210> 82	
	<211> 77	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
30	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 82	
	gcccgccacc accaggaugu gucucagagc uaugcugucc uggaaacagg acagcauagc	60
	aaguugagau aaggcua	77
35	<210> 83	
	<211> 77	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 83	

ES 2 705 540 T3

	ggcagccagc augaugagac gucucagagc uaugcugucc uggaaacagg acagcauagc	60
	aaguugagau aaggcua	77
5	<210> 84 <211>41 <212> ARN <213> Secuencia artificial	
	<220> <223> ARN asociado con Cas9 de nexa dividido	
	<400> 84 gucgguuau aacuugaaaa aguggcaccg agucggugcu u	41
10	<210> 85 <211> 92 <212> ARN <213> Secuencia artificial	
15	<220> <223> ARN asociado con Cas9 de nexa dividido	
	<400> 85	
	ggggccacua gggacaggau gucucagagc uaugcagucc uggaaacagg acugcauagc	60
	aaguugagau aaggcuacug ccguauaggc ag	92
20	<210> 86 <211> 56 <212> ARN <213> Secuencia artificial	
	<220> <223> ARN asociado con Cas9 de nexa dividido	
	<400> 86	
25	cugccguaua ggcagguccg uuaucaacu gaaaaagugg caccgagucg gugcuu	56
	<210> 87 <211> 96 <212> ARN <213> Secuencia artificial	
30	<220> <223> ARN asociado con Cas9 de nexa dividido	
	<400> 87	
	ggggccacua gggacaggau gucucagagc uaugcagucc uggaaacagg acugcauagc	60
	aaguugagau aaggcuagac acugccguau aggag	96
35	<210> 88 <211>61 <212> ARN <213> Secuencia artificial	
	<220> <223> ARN asociado con Cas9 de nexa dividido	
40	<400> 88	

ES 2 705 540 T3

	cugccguaua ggcagagaca guccguuauc aacuugaaaa aguggcaccg agucggugcu	60
	u	61
5	<210> 89 <211> 20 <212> ADN <213> Secuencia artificial	
	<220> <223> Espaciador Tet1	
	<400> 89 ggctgctgtc agggagctca	20
10	<210> 90 <211> 20 <212> ADN <213> Secuencia artificial	
	<220> <223> Espaciador Tet2	
15	<400> 90 gaaagtgccacacagatatcc	20
20	<210>91 <211> 27 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 91 tgtaagtcca ggatgaacta ggagaaa	27
25	<210> 92 <211> 27 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 92 tgttcaact tagaacgaat acagaaa	27
30	<210> 93 <211> 27 <212> ADN <213> <i>Homo sapiens</i>	
35	<400> 93 agggtggggc aggaaggaag ccagaat	27
	<210> 94 <211> 27 <212> ADN <213> <i>Homo sapiens</i>	
40	<400> 94 ctccctcagg gtctgtaaga aaagaaa	27
45	<210> 95 <211> 27 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 95 atacacccaa agtatcagga cgagaat	27
	<210> 96 <211> 27	

ES 2 705 540 T3

	<212> ADN		
	<213> <i>Homo sapiens</i>		
	<400> 96		
	gcggtcgggg aataaggtat aaagaaa	27	
5	<210> 97		
	<211> 99		
	<212> ARN		
	<213> Secuencia artificial		
10	<220>		
	<223> ARNgi asociado con Cas9		
	<400> 97		
	uguaagucca ggaugaacua guuuuuguac ucgaaagaag cuacaaagau aaggcuucau	60	
	gccgaaauca acaccuguc auuuuauaggc aggguguuu	99	
15	<210> 98		
	<211> 99		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARNgi asociado con Cas9		
	<400> 98		
	uguucaacuu uagaacgaau guuuuuguac ucgaaagaag cuacaaagau aaggcuucau	60	
20	gccgaaauca acaccuguc auuuuauaggc aggguguuu	99	
	<210> 99		
	<211> 99		
	<212> ARN		
	<213> Secuencia artificial		
25	<220>		
	<223> ARNgi asociado con Cas9		
	<400> 99		
	aggguggggc aggaaggaag guuuuuguac ucgaaagaag cuacaaagau aaggcuucau	60	
	gccgaaauca acaccuguc auuuuauaggc aggguguuu	99	
30	<210> 100		
	<211> 99		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARNgi asociado con Cas9		
35	<400> 100		
	cuccucagg gucuguaaga guuuuuguac ucgaaagaag cuacaaagau aaggcuucau	60	
	gccgaaauca acaccuguc auuuuauaggc aggguguuu	99	
40	<210> 101		
	<211> 99		
	<212> ARN		
	<213> Secuencia artificial		

ES 2 705 540 T3

	<220>		
	<223> ARNgi asociado con Cas9		
	<400> 101		
	auacacccaa aguaucagga guuuuuguac ucgaaagaag cuacaaagau aaggcuucau		60
	gccgaaauca acaccuguc auuuuauggc aggguguuu		99
5	<210> 102		
	<211> 99		
	<212> ARN		
	<213> Secuencia artificial		
10	<220>		
	<223> ARNgi asociado con Cas9		
	<400> 102		
	gcggucgggg aaauagguau guuuuuguac ucgaaagaag cuacaaagau aaggcuucau		60
	gccgaaauca acaccuguc auuuuauggc aggguguuu		99
15	<210> 103		
	<211> 58		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexo dividido		
20	<400> 103		
	uguaagucca ggauaacua guuuuuguac ucgaaagaag cuacaaagau aaggcuuc		58
	<210> 104		
	<211> 54		
	<212> ARN		
	<213> Secuencia artificial		
25	<220>		
	<223> ARN asociado con Cas9 de nexo dividido		
	<400> 104		
	uguucaacu uagaacgau guuuuuguac ucgaaagaag cuacaaagau aagg		54
30	<210> 105		
	<211> 58		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexo dividido		
35	<400> 105		
	aggguggggc aggaaggaag guuuuuguac ucgaaagaag cuacaaagau aaggcuuc		58
	<210> 106		
	<211> 58		
	<212> ARN		
40	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexo dividido		
	<400> 106		
	cuccucagg gucuguaaga guuuuuguac ucgaaagaag cuacaaagau aaggcuuc		58

ES 2 705 540 T3

	<210> 107	
	<211> 58	
	<212> ARN	
	<213> Secuencia artificial	
5	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 107	
	auacacccaa aguaucagga guuuuuguac ucgaaagaag cuacaaagau aaggcuuc	58
10	<210> 108	
	<211> 54	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
15	<400> 108	
	gcggucgggg auaagguau guuuuuguac ucgaaagaag cuacaaagau aagg	54
20	<210> 109	
	<211>41	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 109	
	augccgaaau caacaccug ucauuuuug gcagguguu u	41
25	<210> 110	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
30	<223> Cebador oligonucleotídico	
	<400> 110	
	tgtaaggaag ctgcagcacc agg	23
35	<210> 111	
	<211> 22	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador oligonucleotídico	
	<400> 111	
40	tcttgctg gacacccgt tc	22
	<210> 112	
	<211> 118	
	<212> ARN	
	<213> Secuencia artificial	
45	<220>	
	<223> ARNgi asociado con Cas9	
	<400> 112	
	agacccaaua ucaggagacu gucucagagc uaugcugucc uggaaacagg acagcauagc	60
	aaguugagau gaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugcuu	118

ES 2 705 540 T3

	<210> 113		
	<211> 118		
	<212> ARN		
	<213> Secuencia artificial		
5	<220>		
	<223> ARNgi asociado con Cas9		
	<400> 113		
	ugucccuagu ggccccacug gucucagagc uaugcugucc uggaaacagg acagcauagc		60
	aaguugagau gaggcuaugc cguuaucaac uugaaaaagu ggcaccgagu cggugcuu		118
10	<210> 114		
	<211> 118		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARNgi asociado con Cas9		
15	<400> 114		
	cggggacaca ggaucuccugg gucucagagc uaugcugucc uggaaacagg acagcauagc		60
	aaguugagau gaggcuaugc cguuaucaac uugaaaaagu ggcaccgagu cggugcuu		118
20	<210> 115		
	<211> 73		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexos divididos		
	<400> 115		
	agacccaaua ucaggagacu gucucagagc uaugcugucc uggaaacagg acagcauagc		60
	aaguugagau aag		73
25	<210> 116		
	<211> 73		
	<212> ARN		
	<213> Secuencia artificial		
30	<220>		
	<223> ARN asociado con Cas9 de nexos divididos		
	<400> 116		
	ugucccuagu ggccccacug gucucagagc uaugcugucc uggaaacagg acagcauagc		60
	aaguugagau aag		73
35	<210> 117		
	<211> 73		
	<212> ARN		
	<213> Secuencia artificial		
	<220>		
	<223> ARN asociado con Cas9 de nexos divididos		
	<400> 117		

ES 2 705 540 T3

	cggggacaca ggauc<u>cc</u>ugg gu<u>cuc</u>agagc uau<u>gc</u>ugucc uggaaacagg acagcauagc	60
	aaguugagau aag	73
5	<210> 118 <211> 45 <212> ARN <213> Secuencia artificial	
	<220> <223> ARN asociado con Cas9 de nexo dividido	
	<400> 118 gcuaguccgu uaucaacuug aaaaaguggc accgagucgg ugcuu	45
10	<210> 119 <211> 76 <212> ARN <213> Secuencia artificial	
15	<220> <223> ARN asociado con Cas9 de nexo dividido	
	<400> 119	
	agacc<u>ca</u>aua ucag<u>g</u>agacu gu<u>cuc</u>agagc uau<u>gc</u>ugucc uggaaacagg acagcauagc	60
	aaguugagau aag<u>g</u>cu	76
20	<210> 120 <211> 76 <212> ARN <213> Secuencia artificial	
	<220> <223> ARN asociado con Cas9 de nexo dividido	
	<400> 120	
	ug<u>cc</u>cuagu gg<u>cc</u>ccacug gu<u>cuc</u>agagc uau<u>gc</u>ugucc uggaaacagg acagcauagc	60
25	aaguugagau aag<u>g</u>cu	76
30	<210> 121 <211> 76 <212> ARN <213> Secuencia artificial	
	<220> <223> ARN asociado con Cas9 de nexo dividido	
	<400> 121	
	cggggacaca ggauc<u>cc</u>ugg gu<u>cuc</u>agagc uau<u>gc</u>ugucc uggaaacagg acagcauagc	60
	aaguugagau aag<u>g</u>cu	76
35	<210> 122 <211> 43 <212> ARN <213> Secuencia artificial	
	<220> <223> ARN asociado con Cas9 de nexo dividido	
40	<400> 122	

ES 2 705 540 T3

	gaguccguua ucaacuugaa aaaguggcac cgagucggug cuu	43
	<210> 123	
	<211> 78	
5	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 123	
	agacc ca uaa ucaggagacu gucucagagc uaugcugucc uggaaacagg acagcauagc	60
	aaguugagau aaggcuag	78
10	<210> 124	
	<211> 78	
	<212> ARN	
	<213> Secuencia artificial	
	<220>	
15	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 124	
	ugucccuagu ggcccacug gucucagagc uaugcugucc uggaaacagg acagcauagc	60
	aaguugagau aaggcuag	78
	<210> 125	
	<211> 78	
20	<212> ARN	
	<213> Secuencia artificial	
	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 125	
	ggggacaca ggauc cc ugg gucucagagc uaugcugucc uggaaacagg acagcauagc	60
25	aaguugagau aaggcuag	78
	<210> 126	
	<211> 41	
	<212> ARN	
	<213> Secuencia artificial	
30	<220>	
	<223> ARN asociado con Cas9 de nexa dividido	
	<400> 126	
	guccguuauc aacuug aaa aguggcaccg agucggugcu u	41
35	<210> 127	
	<211> 1387	
	<212> PRT	
	<213> Streptococcus pyogenes	
	<400> 127	

ES 2 705 540 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

```

ES 2 705 540 T3

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

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

ES 2 705 540 T3

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

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

ES 2 705 540 T3

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

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
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
 1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
 1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
 1145 1150 1155

ES 2 705 540 T3

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
 1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
 1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
 1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
 1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
 1220 1225 1230

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
 1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
 1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
 1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

Gly Ser Pro Lys Lys Lys Arg Lys Val Glu Asp Pro Lys Lys Lys
 1370 1375 1380

Arg Lys Val Asp
 1385

ES 2 705 540 T3

<211> 1168
 <212> PRT
 <213> *Streptococcus thermophilus*

5 <400> 128

Met Pro Lys Lys Lys Arg Lys Val Asp Gly Ser Asp Leu Val Leu Gly
 1 5 10 15

Leu Asp Ile Gly Ile Gly Ser Val Gly Val Gly Ile Leu Asn Lys Val
 20 25 30

Thr Gly Glu Ile Ile His Lys Asn Ser Arg Ile Phe Pro Ala Ala Gln
 35 40 45

Ala Glu Asn Asn Leu Val Arg Arg Thr Asn Arg Gln Gly Arg Arg Leu
 50 55 60

Thr Arg Arg Lys Lys His Arg Arg Val Arg Leu Asn Arg Leu Phe Glu
 65 70 75 80

Glu Ser Gly Leu Ile Thr Asp Phe Thr Lys Ile Ser Ile Asn Leu Asn
 85 90 95

Pro Tyr Gln Leu Arg Val Lys Gly Leu Thr Asp Glu Leu Ser Asn Glu
 100 105 110

Glu Leu Phe Ile Ala Leu Lys Asn Met Val Lys His Arg Gly Ile Ser
 115 120 125

Tyr Leu Asp Asp Ala Ser Asp Asp Gly Asn Ser Ser Val Gly Asp Tyr
 130 135 140

Ala Gln Ile Val Lys Glu Asn Ser Lys Gln Leu Glu Thr Lys Thr Pro
 145 150 155 160

Gly Gln Ile Gln Leu Glu Arg Tyr Gln Thr Tyr Gly Gln Leu Arg Gly
 165 170 175

Asp Phe Thr Val Glu Lys Asp Gly Lys Lys His Arg Leu Ile Asn Val
 180 185 190

Phe Pro Thr Ser Ala Tyr Arg Ser Glu Ala Leu Arg Ile Leu Gln Thr
 195 200 205

Gln Gln Glu Phe Asn Pro Gln Ile Thr Asp Glu Phe Ile Asn Arg Tyr
 210 215 220

ES 2 705 540 T3

Leu Glu Ile Leu Thr Gly Lys Arg Lys Tyr Tyr His Gly Pro Gly Asn
 225 230 235 240

Glu Lys Ser Arg Thr Asp Tyr Gly Arg Tyr Arg Thr Ser Gly Glu Thr
 245 250 255

Leu Asp Asn Ile Phe Gly Ile Leu Ile Gly Lys Cys Thr Phe Tyr Pro
 260 265 270

Asp Glu Phe Arg Ala Ala Lys Ala Ser Tyr Thr Ala Gln Glu Phe Asn
 275 280 285

Leu Leu Asn Asp Leu Asn Asn Leu Thr Val Pro Thr Glu Thr Lys Lys
 290 295 300

Leu Ser Lys Glu Gln Lys Asn Gln Ile Ile Asn Tyr Val Lys Asn Glu
 305 310 315 320

Lys Ala Met Gly Pro Ala Lys Leu Phe Lys Tyr Ile Ala Lys Leu Leu
 325 330 335

Ser Cys Asp Val Ala Asp Ile Lys Gly Tyr Arg Ile Asp Lys Ser Gly
 340 345 350

Lys Ala Glu Ile His Thr Phe Glu Ala Tyr Arg Lys Met Lys Thr Leu
 355 360 365

Glu Thr Leu Asp Ile Glu Gln Met Asp Arg Glu Thr Leu Asp Lys Leu
 370 375 380

Ala Tyr Val Leu Thr Leu Asn Thr Glu Arg Glu Gly Ile Gln Glu Ala
 385 390 395 400

Leu Glu His Glu Phe Ala Asp Gly Ser Phe Ser Gln Lys Gln Val Asp
 405 410 415

Glu Leu Val Gln Phe Arg Lys Ala Asn Ser Ser Ile Phe Gly Lys Gly
 420 425 430

Trp His Asn Phe Ser Val Lys Leu Met Met Glu Leu Ile Pro Glu Leu
 435 440 445

Tyr Glu Thr Ser Glu Glu Gln Met Thr Ile Leu Thr Arg Leu Gly Lys
 450 455 460

Gln Lys Thr Thr Ser Ser Ser Asn Lys Thr Lys Tyr Ile Asp Glu Lys

ES 2 705 540 T3

Ser Val Val Arg Gly Gln Phe Thr Ser Gln Leu Arg Arg His Trp Gly
725 730 735

Ile Glu Lys Thr Arg Asp Thr Tyr His His His Ala Val Asp Ala Leu
740 745 750

Ile Ile Ala Ala Ser Ser Gln Leu Asn Leu Trp Lys Lys Gln Lys Asn
755 760 765

Thr Leu Val Ser Tyr Ser Glu Asp Gln Leu Leu Asp Ile Glu Thr Gly
770 775 780

Glu Leu Ile Ser Asp Asp Glu Tyr Lys Glu Ser Val Phe Lys Ala Pro
785 790 795 800

Tyr Gln His Phe Val Asp Thr Leu Lys Ser Lys Glu Phe Glu Asp Ser
805 810 815

Ile Leu Phe Ser Tyr Gln Val Asp Ser Lys Phe Asn Arg Lys Ile Ser
820 825 830

Asp Ala Thr Ile Tyr Ala Thr Arg Gln Ala Lys Val Gly Lys Asp Lys
835 840 845

Ala Asp Glu Thr Tyr Val Leu Gly Lys Ile Lys Asp Ile Tyr Thr Gln
850 855 860

Asp Gly Tyr Asp Ala Phe Met Lys Ile Tyr Lys Lys Asp Lys Ser Lys
865 870 875 880

Phe Leu Met Tyr Arg His Asp Pro Gln Thr Phe Glu Lys Val Ile Glu
885 890 895

Pro Ile Leu Glu Asn Tyr Pro Asn Lys Gln Ile Asn Glu Lys Gly Lys
900 905 910

Glu Val Pro Cys Asn Pro Phe Leu Lys Tyr Lys Glu Glu His Gly Tyr
915 920 925

Ile Arg Lys Tyr Ser Lys Lys Gly Asn Gly Pro Glu Ile Lys Ser Leu
930 935 940

Lys Tyr Tyr Asp Ser Lys Leu Gly Asn His Ile Asp Ile Thr Pro Lys
945 950 955 960

Asp Ser Asn Asn Lys Val Val Leu Gln Ser Val Ser Pro Trp Arg Ala
965 970 975

ES 2 705 540 T3

Asp Val Tyr Phe Asn Lys Thr Thr Gly Lys Tyr Glu Ile Leu Gly Leu
 980 985 990

Lys Tyr Ala Asp Leu Gln Phe Glu Lys Gly Thr Gly Thr Tyr Lys Ile
 995 1000 1005

Ser Gln Glu Lys Tyr Asn Asp Ile Lys Lys Lys Glu Gly Val Asp
 1010 1015 1020

Ser Asp Ser Glu Phe Lys Phe Thr Leu Tyr Lys Asn Asp Leu Leu
 1025 1030 1035

Leu Val Lys Asp Thr Glu Thr Lys Glu Gln Gln Leu Phe Arg Phe
 1040 1045 1050

Leu Ser Arg Thr Met Pro Lys Gln Lys His Tyr Val Glu Leu Lys
 1055 1060 1065

Pro Tyr Asp Lys Gln Lys Phe Glu Gly Gly Glu Ala Leu Ile Lys
 1070 1075 1080

Val Leu Gly Asn Val Ala Asn Ser Gly Gln Cys Lys Lys Gly Leu
 1085 1090 1095

Gly Lys Ser Asn Ile Ser Ile Tyr Lys Val Arg Thr Asp Val Leu
 1100 1105 1110

Gly Asn Gln His Ile Ile Lys Asn Glu Gly Asp Lys Pro Lys Leu
 1115 1120 1125

Asp Phe Gly Ser Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser
 1130 1135 1140

Leu Gly Ser Gly Ser Pro Lys Lys Lys Arg Lys Val Asp Gly Ser
 1145 1150 1155

Pro Lys Lys Lys Arg Lys Val Asp Ser Gly
 1160 1165

REIVINDICACIONES

1. Una composición de polinucleótidos de nexos divididos asociados con CRISPR-Cas9 de tipo II que comprende:

un primer polinucleótido de nexos divididos asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn1-casPN) que comprende, en la dirección de 5' a 3', una primera secuencia de nucleótidos de elemento de tallo I y una secuencia de nucleótidos de elemento de tallo de nexos I;

un segundo polinucleótido de nexos divididos asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn2-casPN) que comprende, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo de nexos II, en el que la secuencia de nucleótidos de elemento de tallo de nexos I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo de nexos II del sn2-casPN son capaces de formar un elemento de tallo de nexos mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexos I y la secuencia de nucleótidos de elemento de tallo de nexos II; y

un tercer polinucleótido de nexos divididos asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn3-casPN) que comprende, en la dirección de 5' a 3', una secuencia de unión a diana de ácido nucleico y una primera secuencia de nucleótidos de elemento de tallo II, en el que la primera secuencia de nucleótidos de elemento de tallo I del sn1-casPN y la primera secuencia de nucleótidos de elemento de tallo II del sn3-casPN son capaces de formar un primer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la primera secuencia de nucleótidos de elemento de tallo I y la primera secuencia de nucleótidos de elemento de tallo II.

2. La composición de la reivindicación 1, en la que

la primera secuencia de nucleótidos de elemento de tallo I del sn1-casPN comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I y una secuencia de nucleótidos de elemento de tallo inferior I, y la primera secuencia de nucleótidos de elemento de tallo II del sn3-casPN comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, en la que la secuencia de nucleótidos de elemento de tallo superior I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo superior II del sn3-casPN son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo inferior II del sn3-casPN son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

3. La composición de la reivindicación 1 o 2, que comprende además:

un primer polinucleótido adjunto que comprende un extremo 5' y un extremo 3' que comprende una segunda secuencia de nucleótidos de elemento de tallo II; en el que el sn2-casPN comprende, en la dirección de 5' a 3', la secuencia de nucleótidos de elemento de tallo de nexos II y una segunda secuencia de nucleótidos de elemento de tallo I, y la segunda secuencia de nucleótidos de elemento de tallo I del sn2-casPN y la segunda secuencia de nucleótidos de elemento de tallo II del primer polinucleótido adjunto son capaces de formar un segundo elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la segunda secuencia de nucleótidos de elemento de tallo I y la segunda secuencia de nucleótidos de elemento de tallo II.

4. La composición de la reivindicación 3, en la que el primer polinucleótido adjunto comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la segunda secuencia de nucleótidos de elemento de tallo II, y en la que el extremo 5' del primer polinucleótido adjunto se une covalentemente con el extremo 3' del sn2-casPN.

5. La composición de una cualquiera de las reivindicaciones 1 a 4, que comprende además:

un segundo polinucleótido adjunto que comprende un extremo 5' y un extremo 3' que comprende una tercera secuencia de nucleótidos de elemento de tallo II; en el que el primer polinucleótido adjunto comprende, en la dirección de 5' a 3', la secuencia de nucleótidos de elemento de bucle, la segunda secuencia de nucleótidos de elemento de tallo II y una tercera secuencia de nucleótidos de elemento de tallo I, y la tercera secuencia de nucleótidos de elemento de tallo I del primer polinucleótido adjunto y la tercera secuencia de nucleótidos de elemento de tallo II del segundo polinucleótido adjunto son capaces de formar un tercer elemento de tallo mediante enlaces de hidrógeno de pares de bases entre la tercera secuencia de nucleótidos de elemento de tallo I y la tercera secuencia de nucleótidos de elemento de tallo II.

6. La composición de la reivindicación 5, en la que el segundo polinucleótido adjunto comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de bucle y la tercera secuencia de nucleótidos de elemento de tallo II, en la que el extremo 5' del segundo polinucleótido adjunto se une covalentemente con el extremo 3' del primer polinucleótido adjunto.

7. La composición de una cualquiera de las reivindicaciones 1 a 6, en la que el sn1-casPN comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I.
8. La composición de una cualquiera de las reivindicaciones 1 a 7, en la que el sn2-casPN comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II.
- 5 9. La composición de una cualquiera de las reivindicaciones 1 a 6, en la que el sn1-casPN comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I y el primer polinucleótido adyuvante comprende además una secuencia de nucleótidos de elemento de unión efector I, el sn2-casPN comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II y el segundo polinucleótido adyuvante comprende
- 10 además una secuencia de nucleótidos de elemento de unión efector II, y la secuencia de nucleótidos de elemento de unión efector I del primer polinucleótido adyuvante y la secuencia de nucleótidos de elemento de unión efector II del segundo polinucleótido adyuvante son capaces de formar un elemento de unión efector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de unión efector I y la secuencia de nucleótidos de elemento de unión efector II.
- 15 10. La composición de la reivindicación 9, en la que el primer polinucleótido adyuvante comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento de unión efector I, el segundo polinucleótido adyuvante comprende, en la dirección de 5' a 3', la secuencia de nucleótidos de elemento de unión efector II y una secuencia de nucleótidos de elemento conector II, y la secuencia de nucleótidos de elemento conector I del primer polinucleótido adyuvante y la secuencia de nucleótidos de elemento conector II del segundo polinucleótido adyuvante son capaces de formar un elemento conector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento conector II.
- 20 11. La composición de la reivindicación 9 o 10, en la que el elemento de unión efector es un ARN bicatenario, y el elemento de unión efector comprende un sitio de unión a proteína Csy4.
- 25 12. La composición de la reivindicación 7, en la que el primer polinucleótido adyuvante comprende una horquilla.
13. La composición de la reivindicación 8, en la que el segundo polinucleótido adyuvante comprende una horquilla.
14. La composición de una cualquiera de las reivindicaciones 1-6, en la que el sn1-casPN comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I y el primer polinucleótido adyuvante comprende además, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento conector I y una horquilla,
- 30 el sn2-casPN comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II y el segundo polinucleótido adyuvante comprende además, en la dirección de 5' a 3', una horquilla y una secuencia de nucleótidos de elemento conector II, y
- 35 la secuencia de nucleótidos de elemento conector I del primer polinucleótido adyuvante y la secuencia de nucleótidos de elemento conector II del segundo polinucleótido adyuvante son capaces de formar un elemento conector mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento conector I y la secuencia de nucleótidos de elemento conector II.
15. Una composición de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II que comprende:
- 40 un primer polinucleótido de nexa dividido asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn1-casPN) que comprende, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo superior I, una secuencia de nucleótidos de elemento de protuberancia I, una secuencia de nucleótidos de elemento de tallo inferior I y una secuencia de nucleótidos de elemento de tallo de nexa I;
- 45 un segundo polinucleótido de nexa dividido asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn2-casPN) que comprende, en la dirección de 5' a 3', una secuencia de nucleótidos de elemento de tallo de nexa II, un segundo elemento de tallo que comprende una horquilla, y un tercer elemento de tallo que comprende una horquilla, en el que la secuencia de nucleótidos de elemento de tallo de nexa I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo de nexa II del sn2-casPN son capaces de formar un elemento de tallo de nexa mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo de nexa I y la secuencia de nucleótidos de elemento de tallo de nexa II; y
- 50 un tercer polinucleótido de nexa dividido asociado con CRISPR-Cas9 de tipo II que comprende un extremo 5' y un extremo 3' (sn3-casPN) que comprende, en la dirección de 5' a 3', una secuencia de unión a diana de ADN, una secuencia de nucleótidos de elemento de tallo inferior II, una secuencia de nucleótidos de elemento de protuberancia II y una secuencia de nucleótidos de elemento de tallo superior II, en el que la secuencia de nucleótidos de elemento de tallo superior I del sn1-casPN y la secuencia de nucleótidos de elemento de tallo superior II del sn3-casPN son capaces de formar un elemento de tallo superior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo superior I y la secuencia de nucleótidos de elemento de tallo superior II, y la secuencia de nucleótidos de elemento de tallo inferior I del sn1-
- 55

casPN y la secuencia de nucleótidos de elemento de tallo inferior II del sn3-casPN son capaces de formar un elemento de tallo inferior mediante enlaces de hidrógeno de pares de bases entre la secuencia de nucleótidos de elemento de tallo inferior I y la secuencia de nucleótidos de elemento de tallo inferior II.

- 5 16. La composición de la reivindicación 15, en la que el sn1-casPN comprende además un primer polinucleótido adyuvante adyacente 3' a la secuencia de nucleótidos de elemento de tallo de nexa I y el sn2-casPN comprende además un segundo polinucleótido adyuvante adyacente 5' a la secuencia de nucleótidos de elemento de tallo de nexa II.
17. Un sistema de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9, que comprende:
- 10 la composición de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II de una cualquiera de las reivindicaciones 1 a 14; y una proteína Cas9 o un polinucleótido que codifica una proteína Cas9.
18. Una composición farmacéutica, que comprende:
- la composición de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II de una cualquiera de las reivindicaciones 1 a 14; y un excipiente farmacéuticamente aceptable.
- 15 19. Un complejo nucleoproteico de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9, que comprende:
- la composición de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II de una cualquiera de las reivindicaciones 1 a 14; y una proteína Cas9;
- 20 en la que el complejo nucleoproteico de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 tiene capacidad de unión dirigida a sitio con un ácido nucleico diana complementario de la unión a diana de ácido nucleico.
20. Un procedimiento de corte de un ácido nucleico diana, que comprende:
- poner en contacto un ácido nucleico que comprende el ácido nucleico diana con el complejo nucleoproteico de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 de la reivindicación 19,
- 25 facilitando de este modo la unión del complejo nucleoproteico de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 con el ácido nucleico diana, en el que el complejo nucleoproteico de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 unido corta el ácido nucleico diana,
- 30 en el que dicho procedimiento no comprende un procedimiento para el tratamiento del cuerpo humano o animal por cirugía o terapia.
21. Un procedimiento de unión de un ácido nucleico diana, que comprende:
- poner en contacto un ácido nucleico que comprende el ácido nucleico diana con el complejo nucleoproteico de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 de la reivindicación 19,
- 35 facilitando de este modo la unión del complejo nucleoproteico de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II/proteína Cas9 con el ácido nucleico diana, en el que dicho procedimiento no comprende un procedimiento para el tratamiento del cuerpo humano o animal por cirugía o terapia.
22. Un kit que comprende:
- la composición de polinucleótidos de nexa dividido asociados con CRISPR-Cas9 de tipo II de una cualquiera de las reivindicaciones 1 a 14 y un tampón.
- 40 23. El kit de la reivindicación 22, que comprende además una proteína Cas9 o un polinucleótido que codifica una proteína Cas9.

FIG. 1A

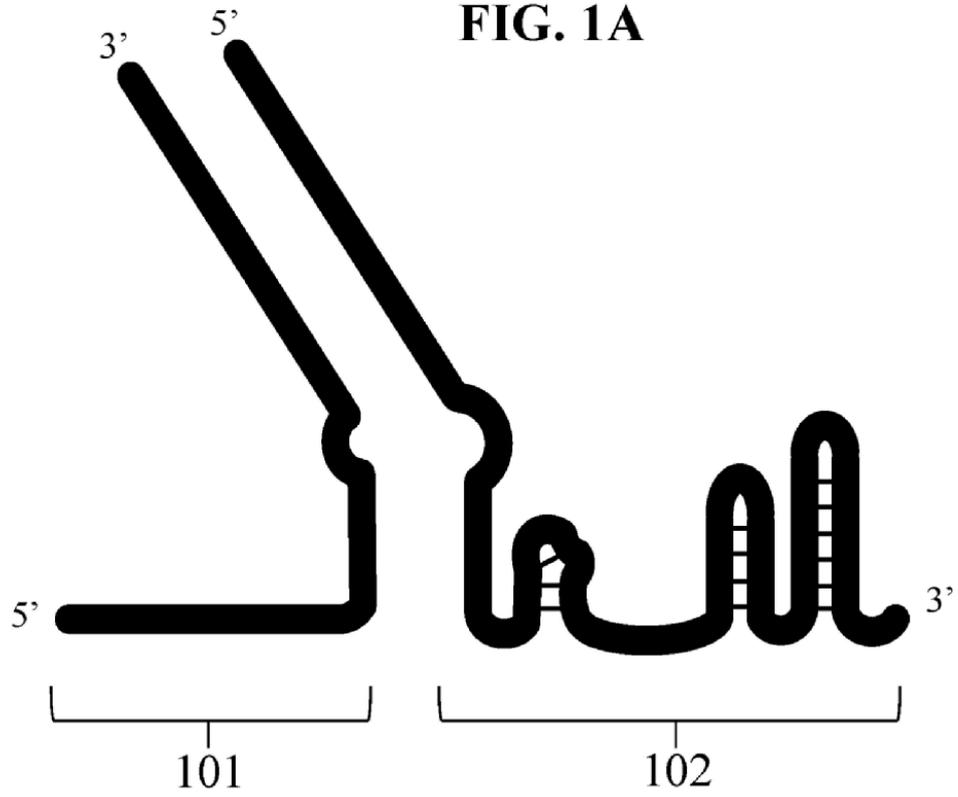
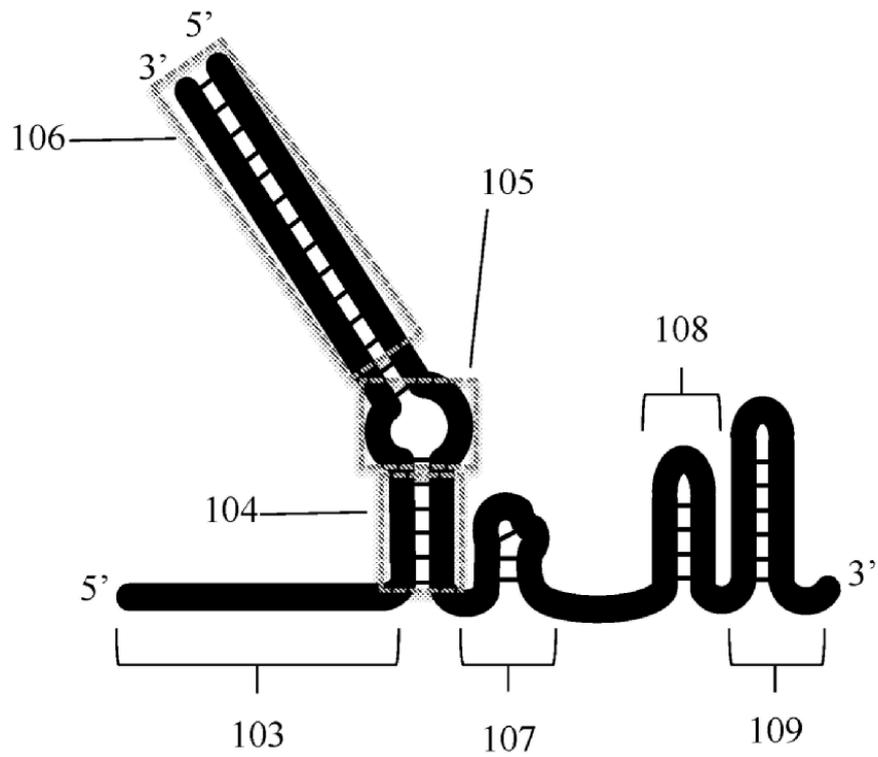


FIG. 1B



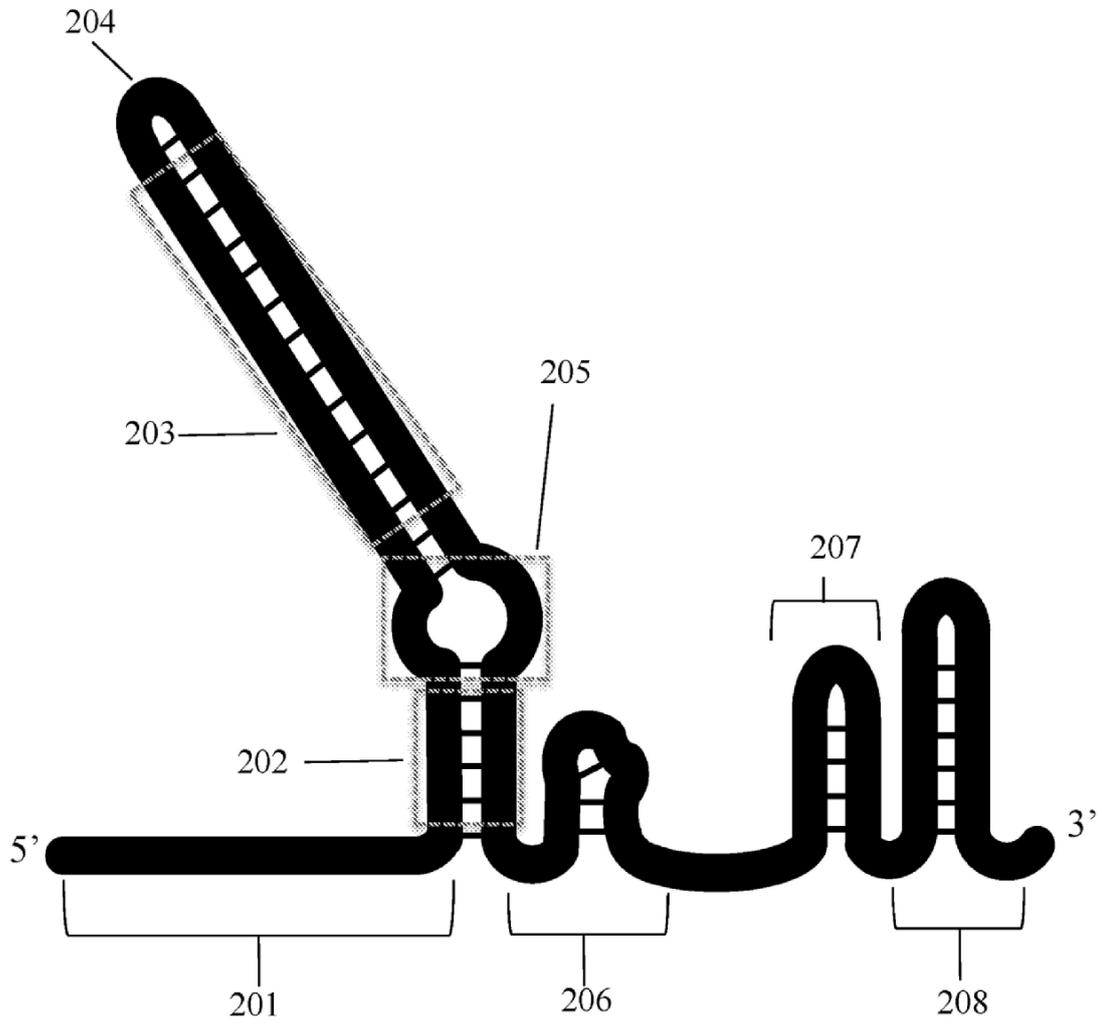


FIG. 2

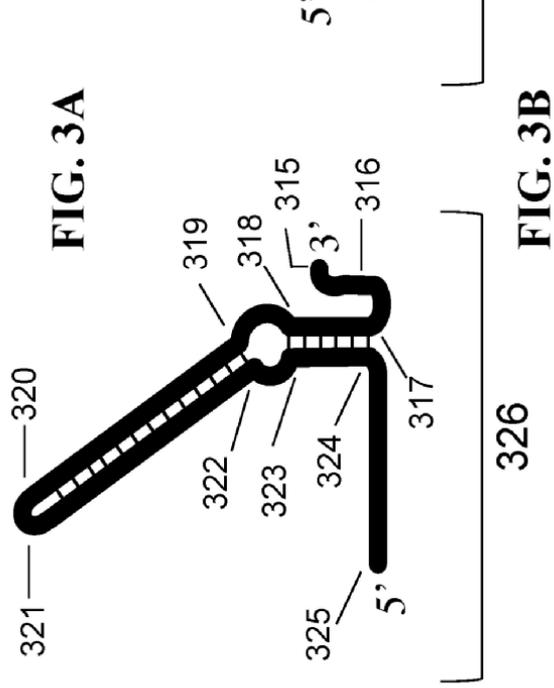
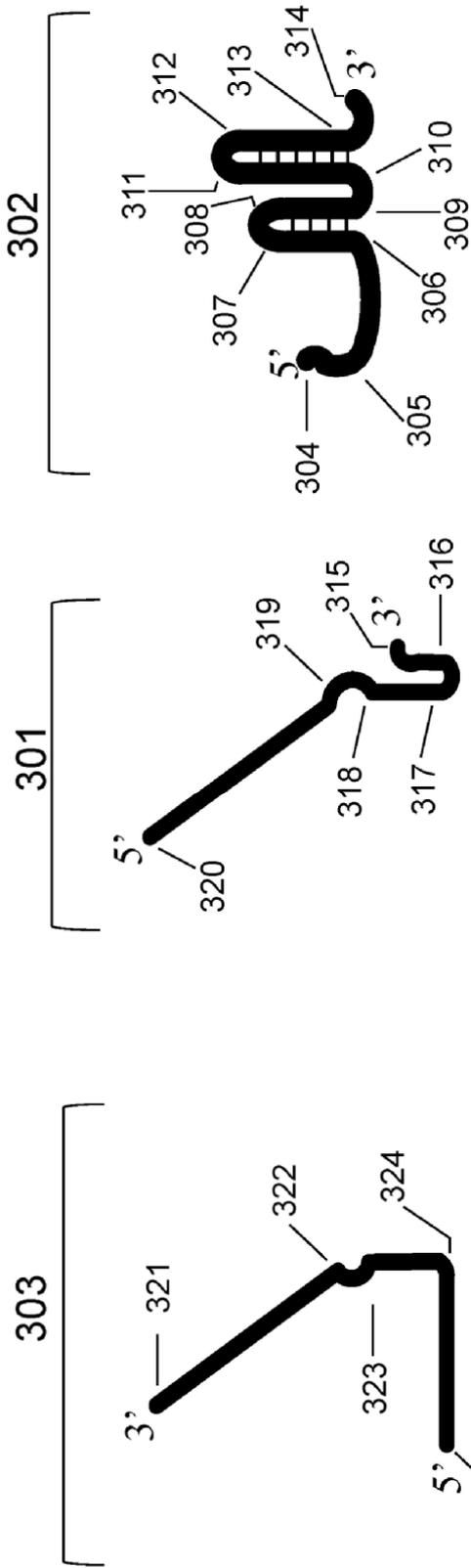


FIG. 3A

FIG. 3B

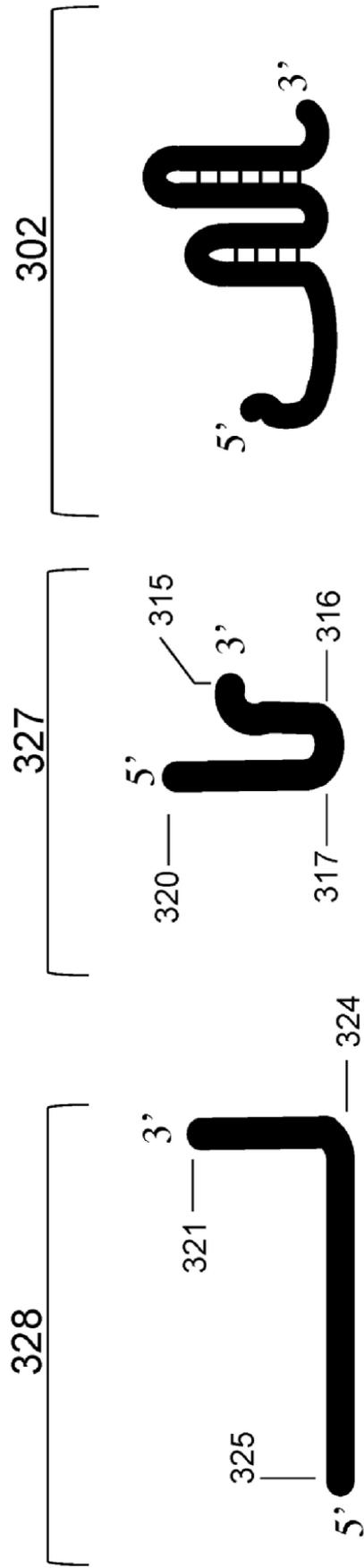


FIG. 3C

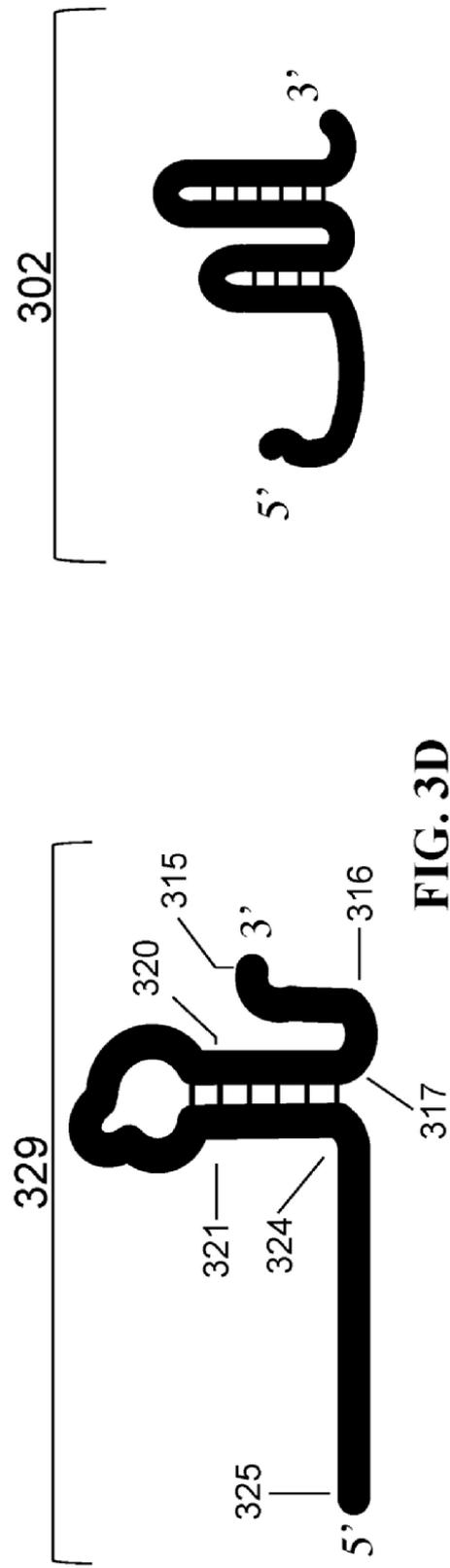


FIG. 3D

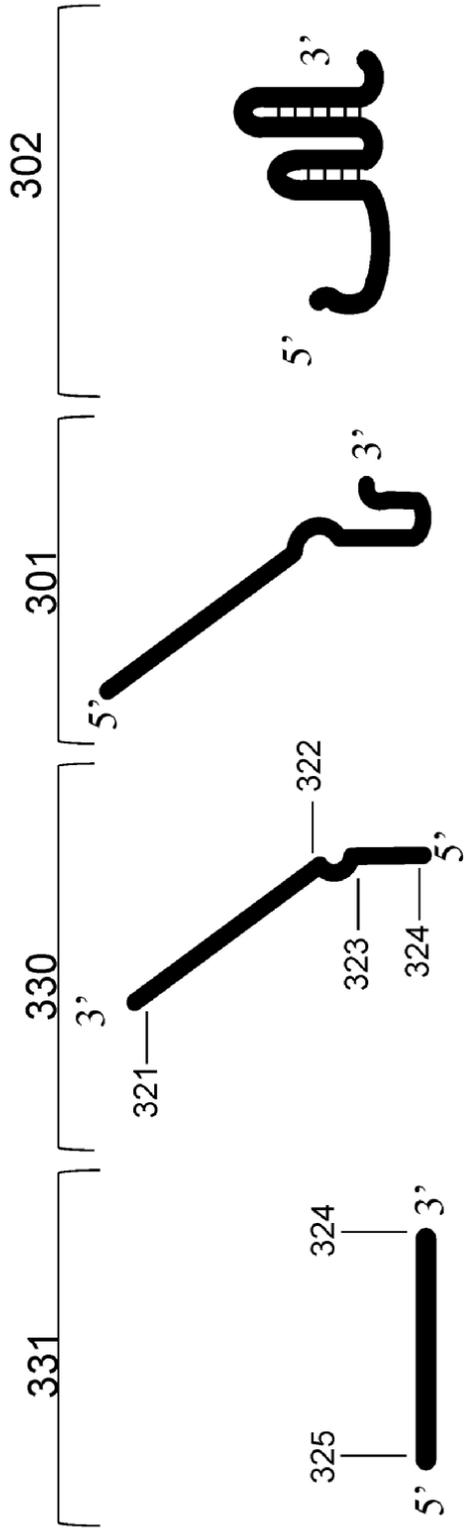


FIG. 3E

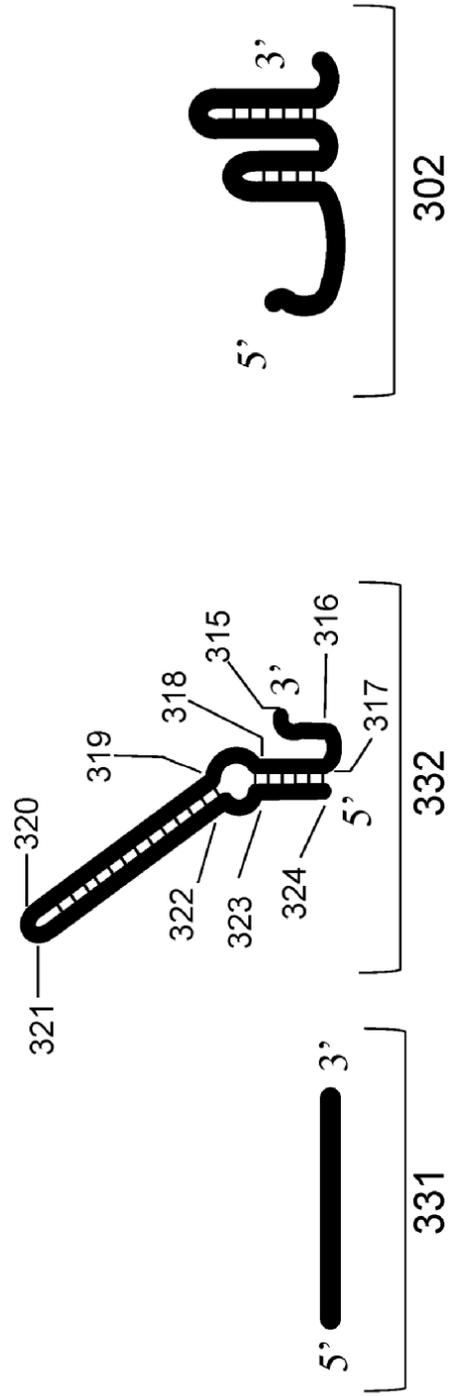


FIG. 3F

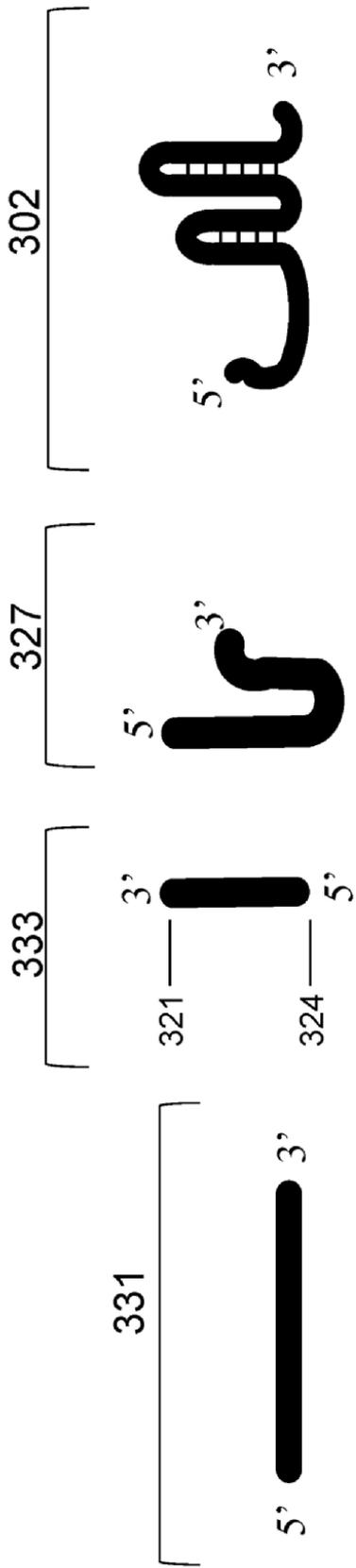


FIG. 3G

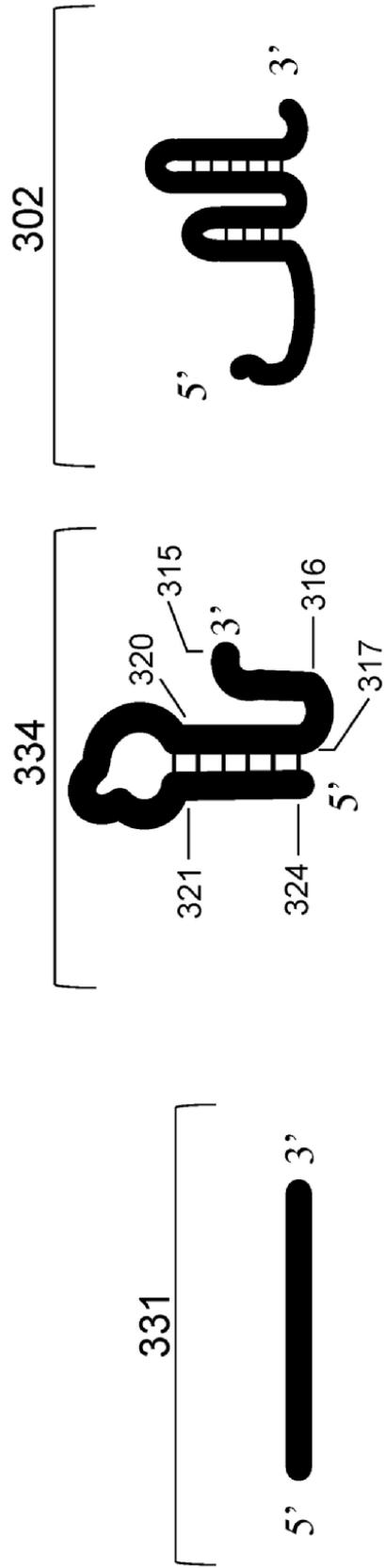


FIG. 3H

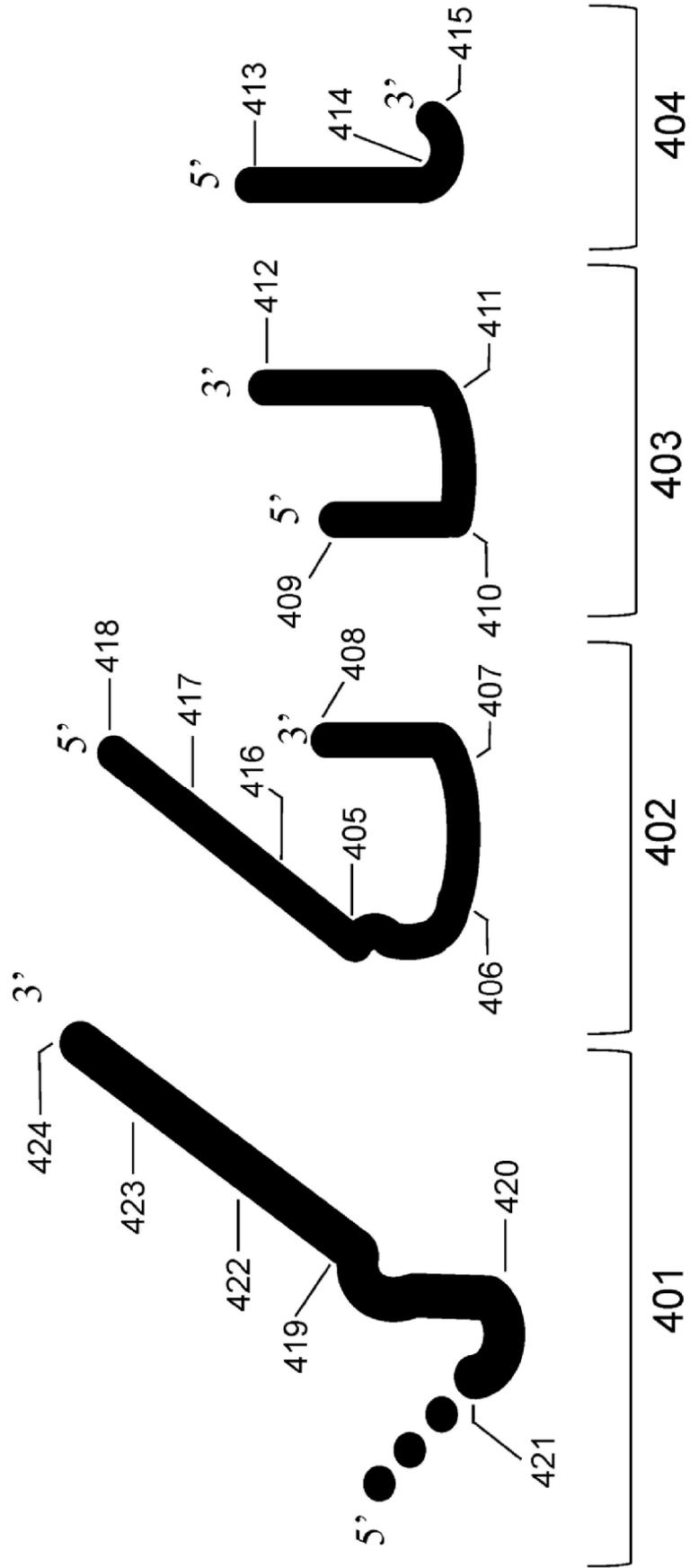


FIG. 4A

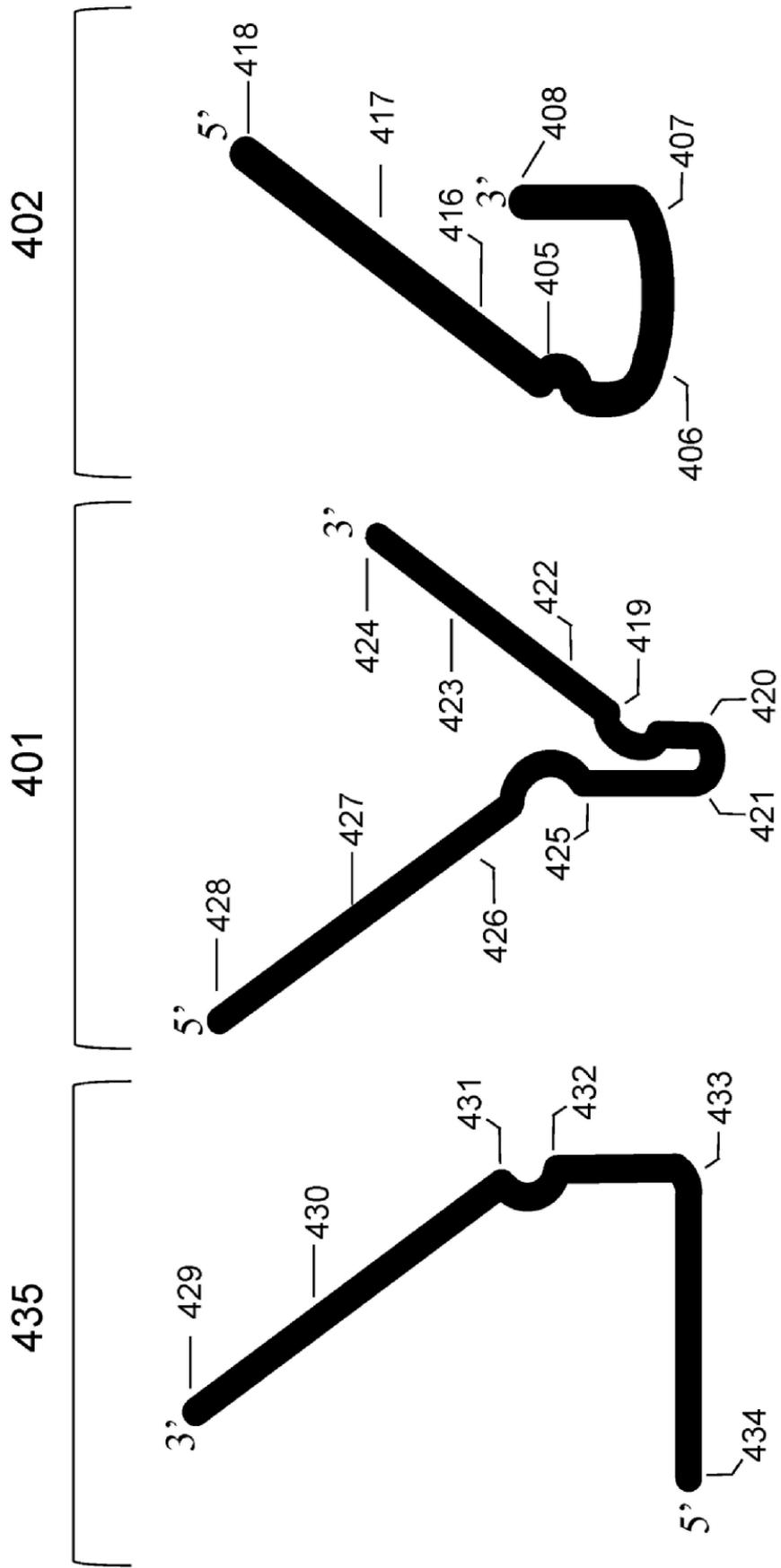


FIG. 4B

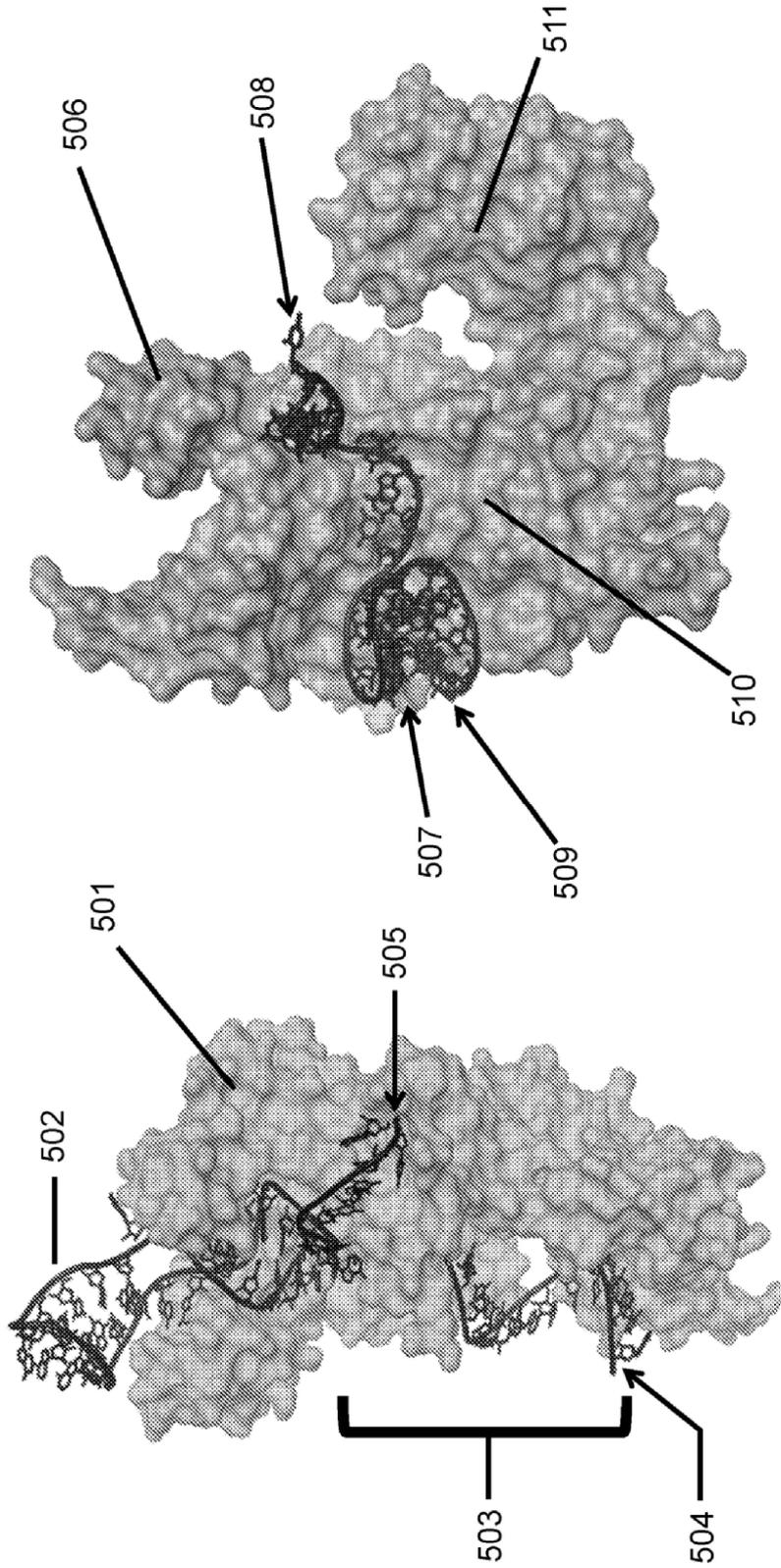


FIG. 5B

FIG. 5A

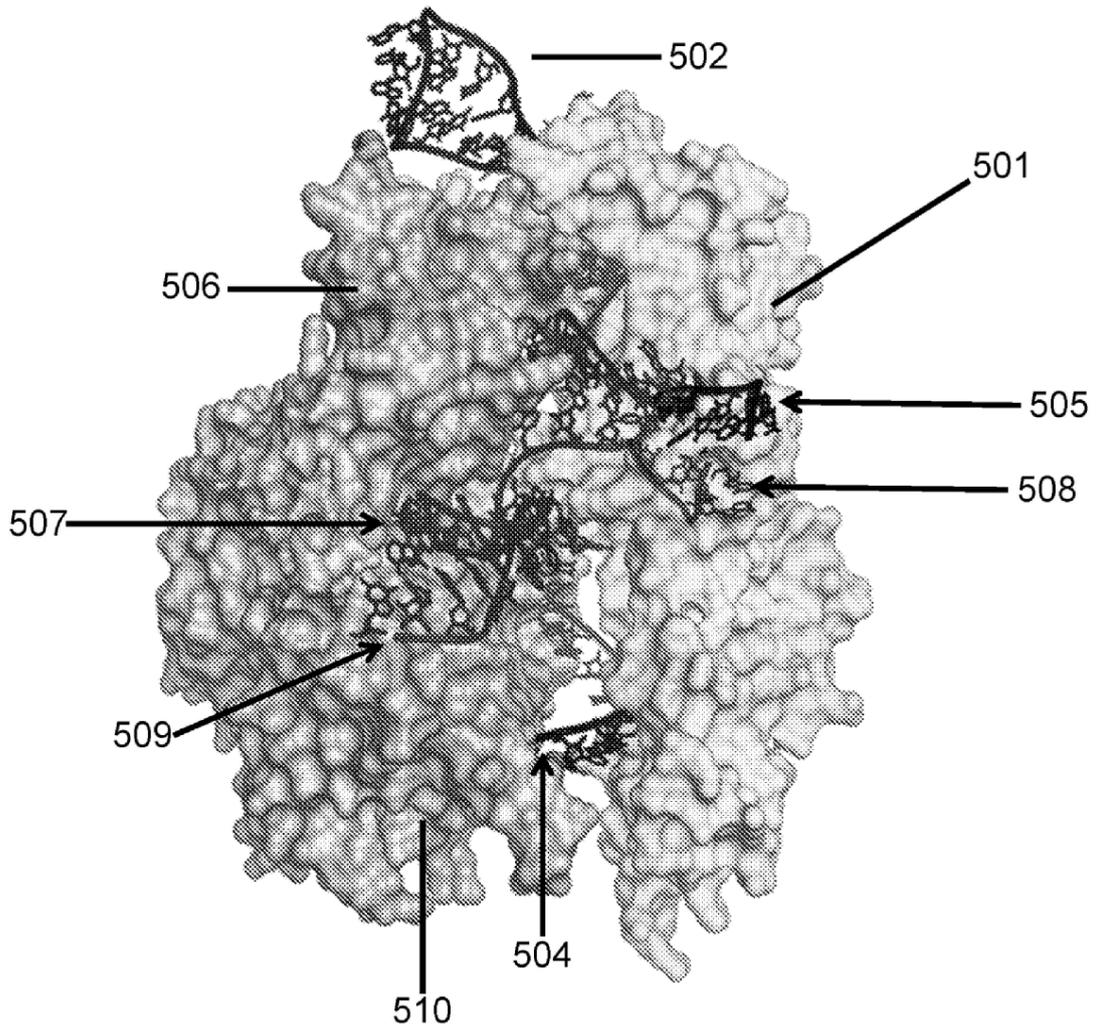


FIG. 5C

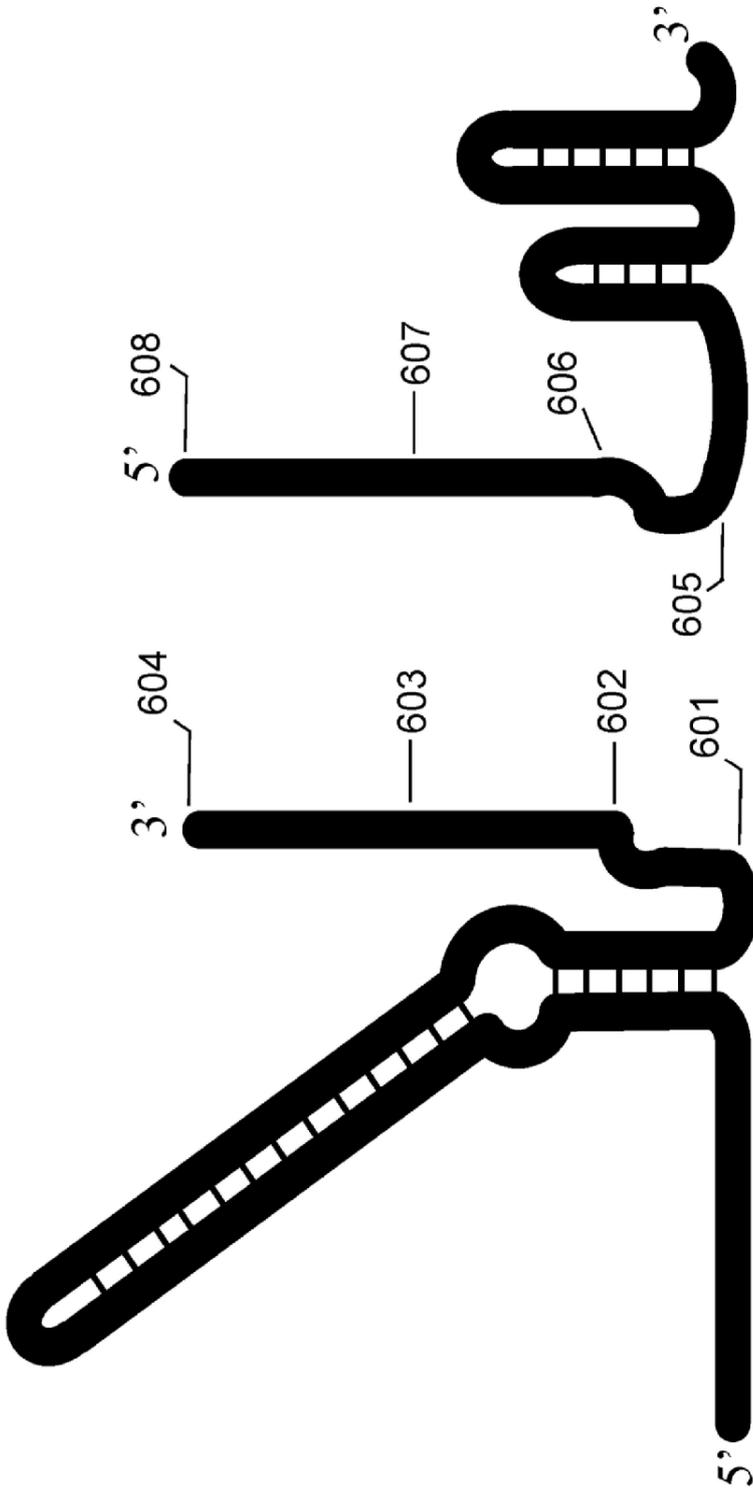


FIG. 6A

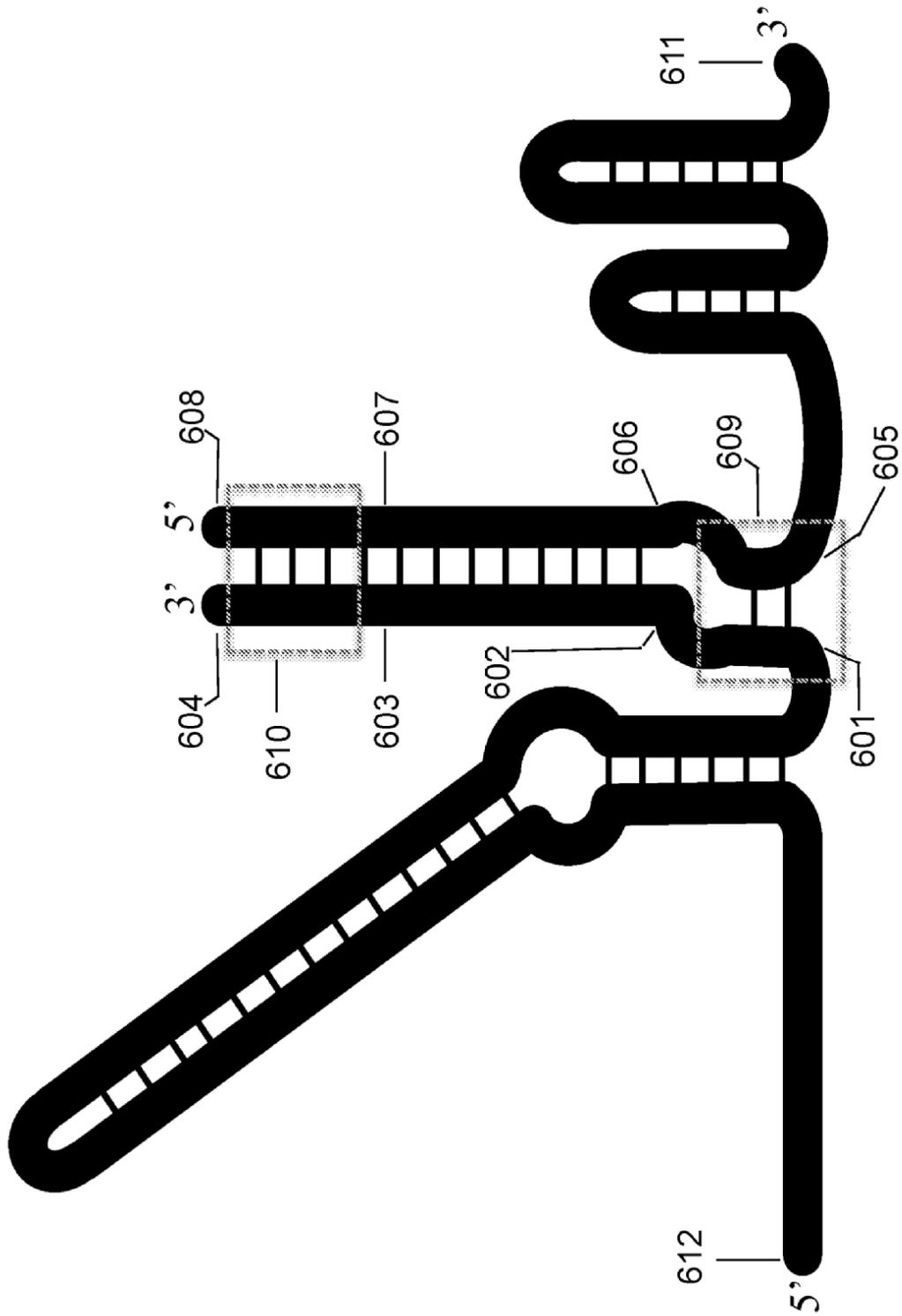


FIG. 6B

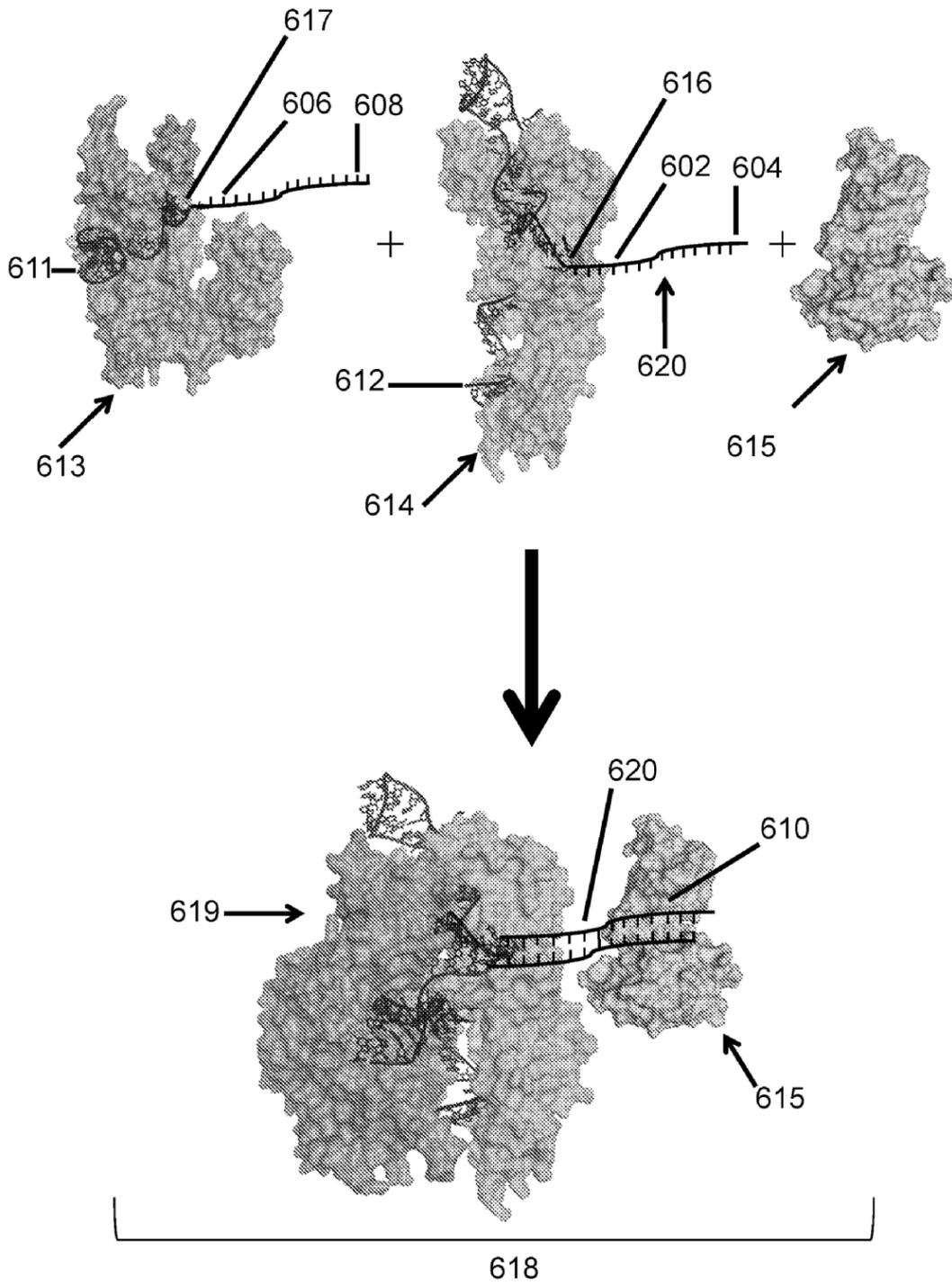


FIG. 6C

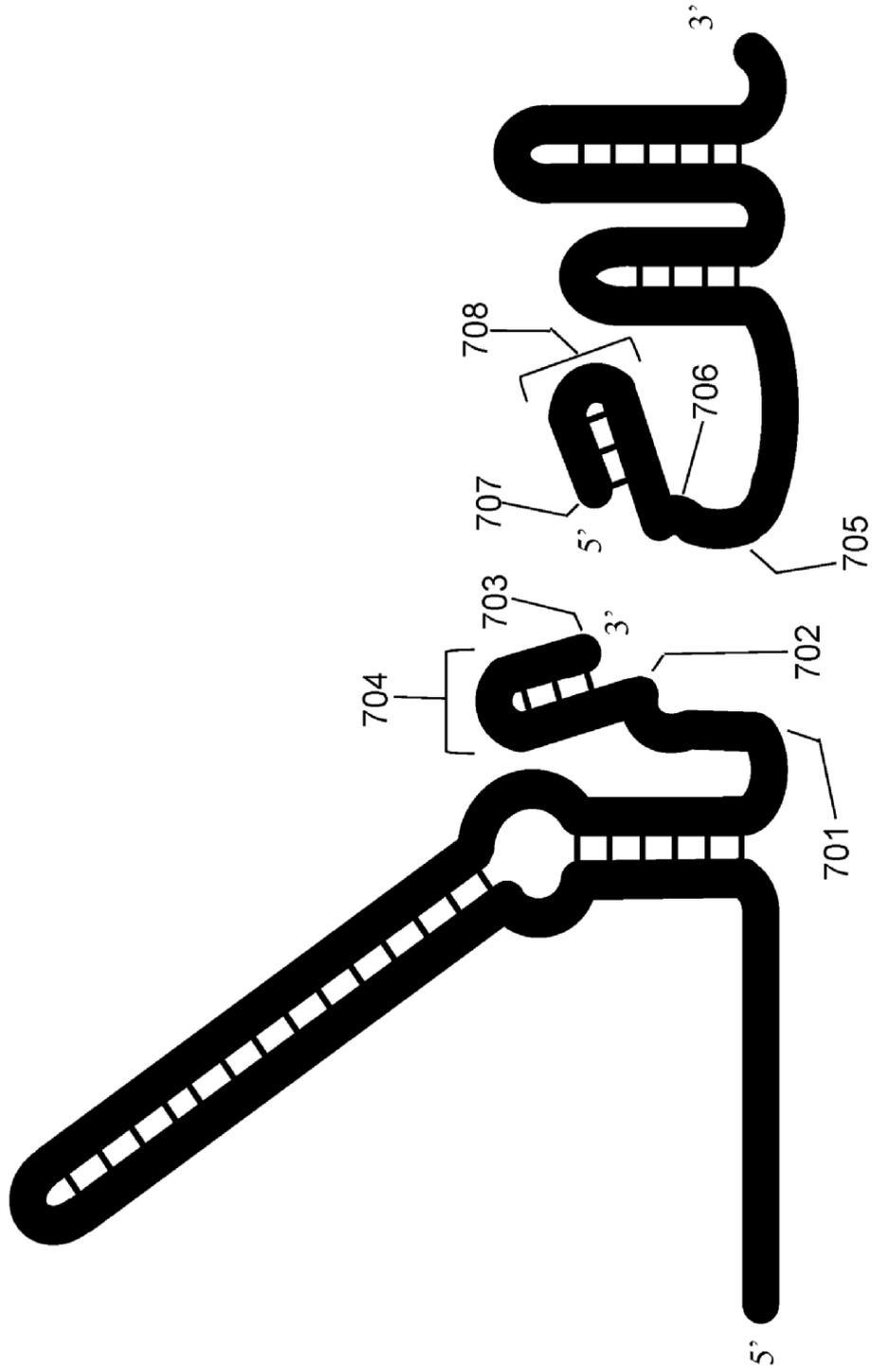


FIG. 7A

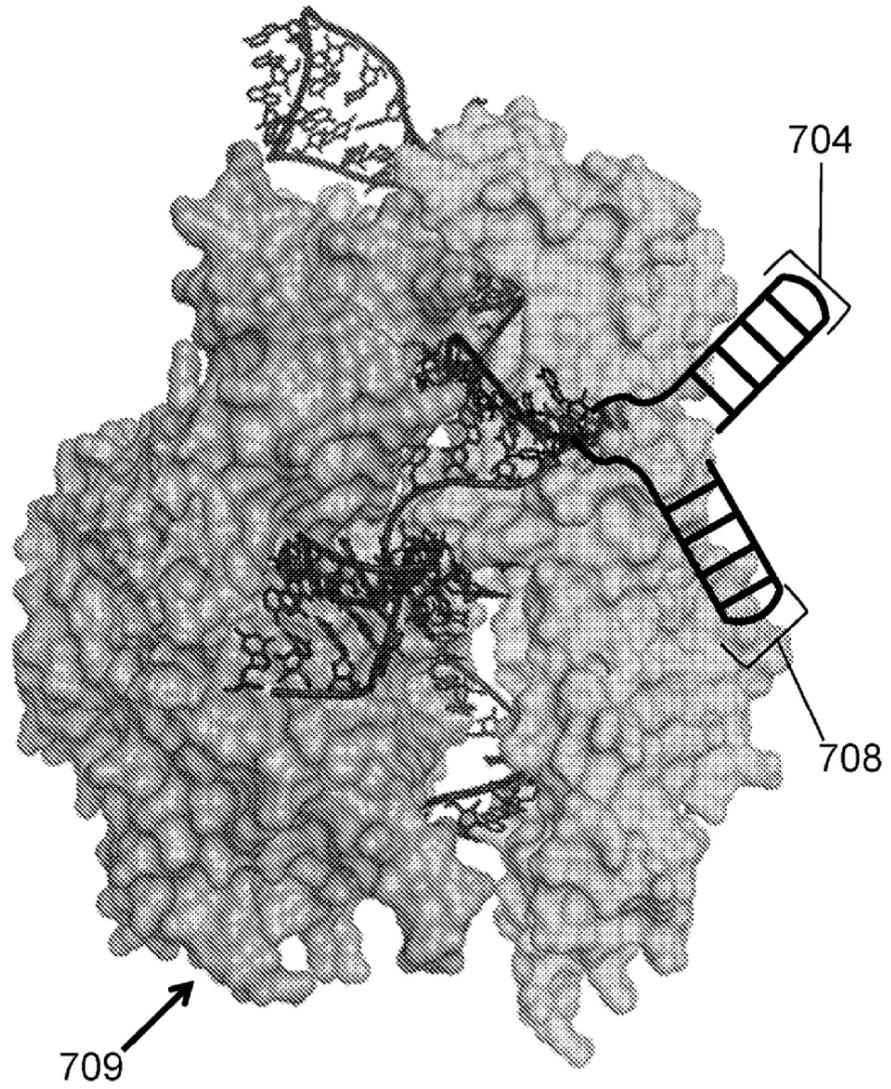


FIG. 7B

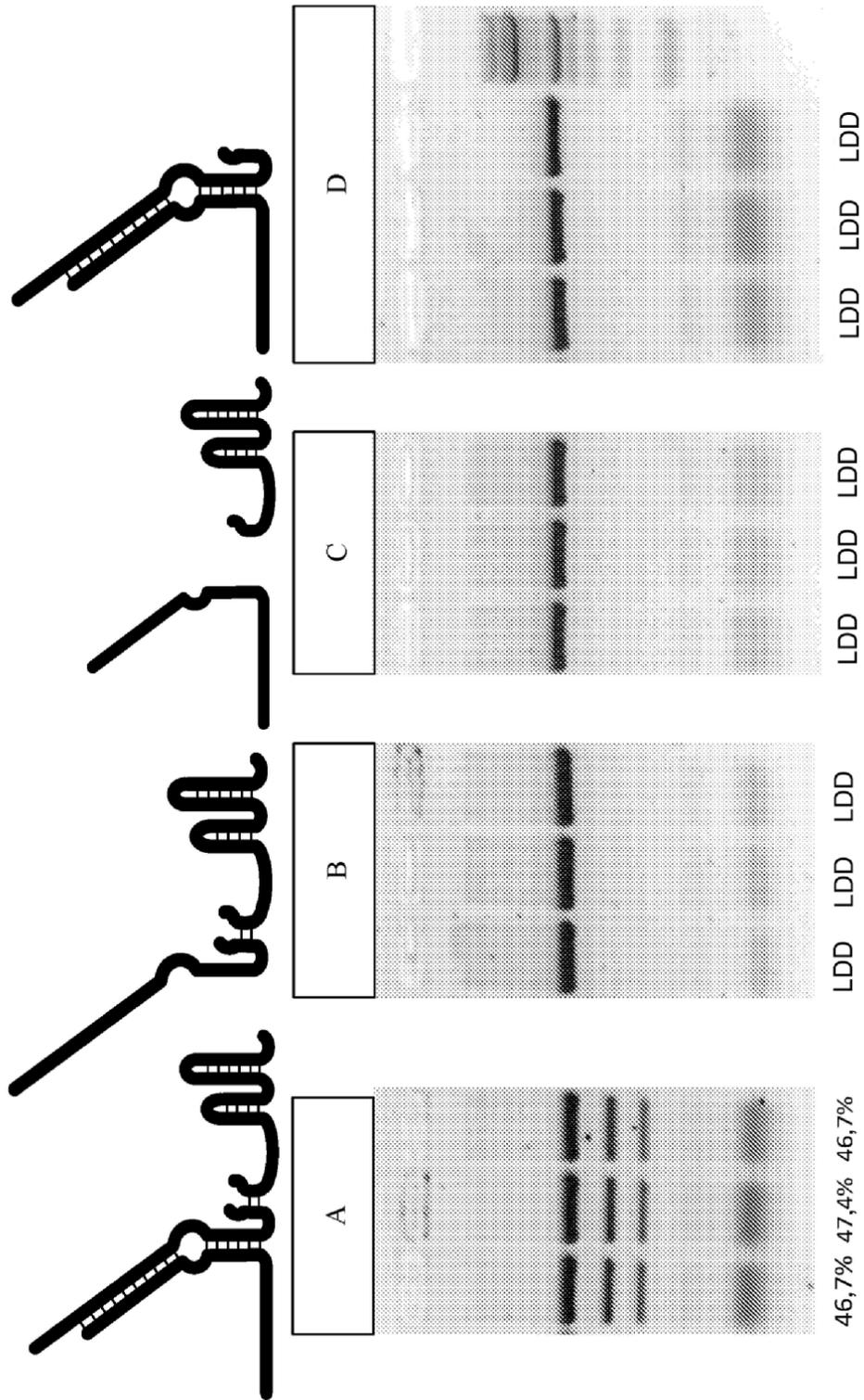


FIG. 8

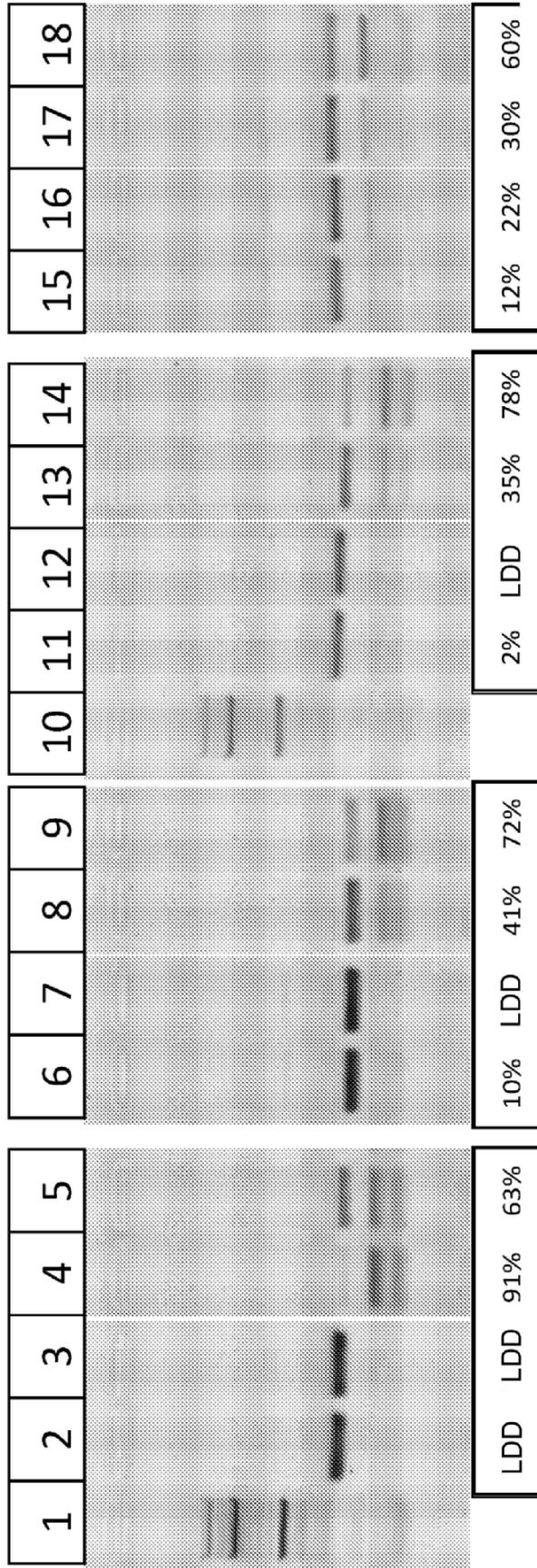


FIG. 9

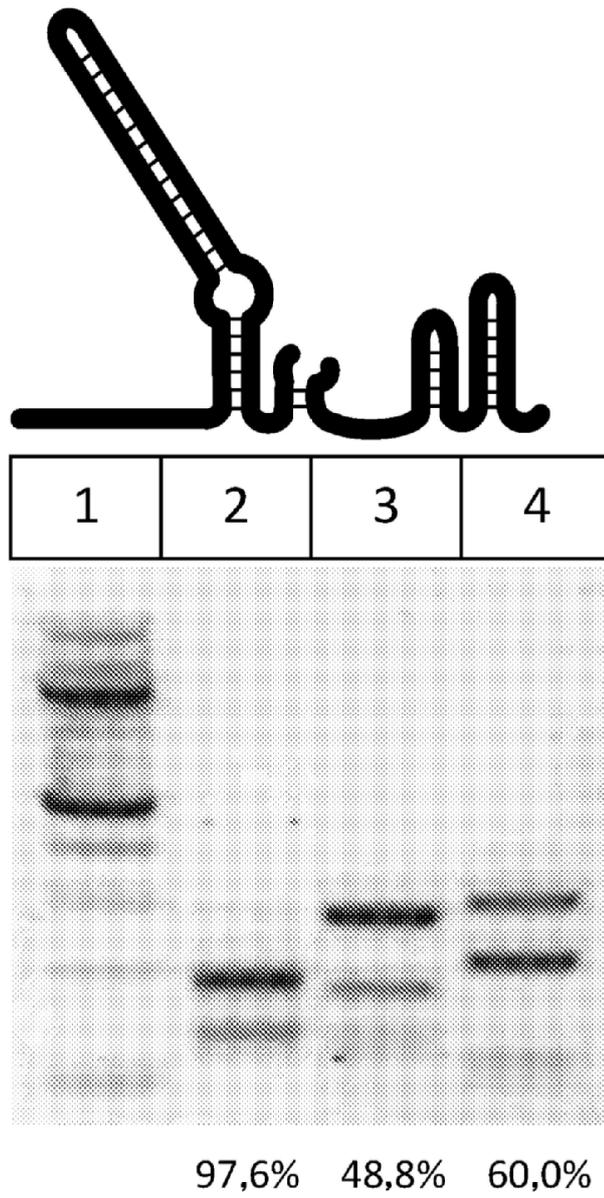


FIG. 10

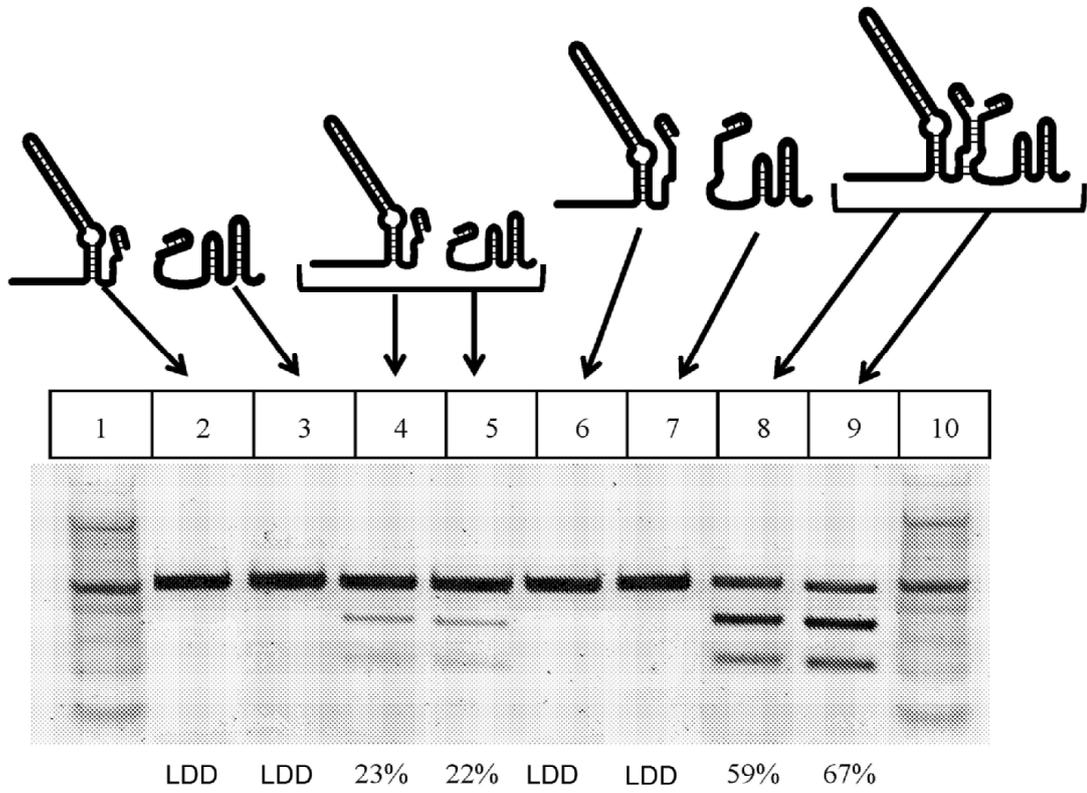


FIG. 11

1	CAAAACAGCAUAGCAAGUAAAAUAAGGCUA/GUCC GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU UUUUUU	SEQ ID NO:1/ SEQ ID NO:2
2	UAAAUCUUGCAGAAGCUACAACGAUAAGGCUUC/AU GCCGAAAUCAACACCCUGUCAUUUUAUGGCAGGGUG UUUUCGUUAUUUUUUU	SEQ ID NO:3/ SEQ ID NO:4
3	CAAAUAACAUAAGCAAGUAAAAUAAGGCUUU/GUC CGUUAUCAACUUUUAUUUAAGUAGCGCUGUUUCGGC GCUUUUUUU	SEQ ID NO:5/ SEQ ID NO:6
4	CUGCGAAUUGAGAACCGUUGCUACAAUAAGGCCGUC UGAA/AAGAUGUGCCGCAACGCUCUGCCCUUAAAGC UUCUGCUUUAAGGGGCAUCGUUUA	SEQ ID NO:7/ SEQ ID NO:8
5	UUGGAGCUAUUCGAAACAACACAGCGAGUAAAAUA AGGCUU/UGUCCGUACACAACUUGUAAAAGUGGCAC CCGAUUCGGGUGCGUUUUUU	SEQ ID NO:9/ SEQ ID NO:10
6	AUUGUACUUUACCUAAAAUUACAGAAUCUACUAAA ACAAGGCAAA/AUGCCGUGUUUAUCUCGUCAACUUG UUGGGGAGAUUUUU	SEQ ID NO:11/ SEQ ID NO:12
7	AGUCACUAAACUUAAUUAAAUAAGAACUGAACCCUCAGU AAGCAUUGGCUC/GUUUCCAUGUUGAUUGCUCGCC GGUGCUCCUUAUUUUUAAGGGCGCCGGCUUUUCU	SEQ ID NO:13/ SEQ ID NO:14
8	UAGCAAUCGAGAGGCGGUCGUUUUCGCAAGCAAA UUGACCCCUU/GUGCGGGCUCGGCAUCCCAAGGUCAG CUGCCGGUUAUUUCGAAAAGGCCACCGCAAGCAG CGCGUGGGCCUUUUU	SEQ ID NO:15/ SEQ ID NO:16
9	AAUUCUUGC UAAAAGAAAUUUAAAAAGAGACUAAAA UAAGUGGUUUUUGGUC/AUCCACGCAGGGUUACAAU CCUUUAAAACCAUUAAAUAUCAAAUAAACUAGGUU GUAUCAACUUAGUUUUUU	SEQ ID NO:17/ SEQ ID NO:18
10	AUUGUCGCACUGCGAAAUGAGAACCGUUGCUACAAU AAGGCCGUCUGAAA/AGAUGUGCCGCAACGCUCUGCC CCUUAAAAGCUUCUGCUUUAAGGGGCAUCGUUUUUUU CGGUUAAAAAUGCCGUCUGAAACCGGUUUUU	SEQ ID NO:19/ SEQ ID NO:20
11	CUUGCACGGUACUAAAUCUUGCUGAGCCUACAAA GAUAAGGCUUU/AUGCCGAAUUCAAGCACCCCAUGU UUUGACAUGAGGUGCUUUU	SEQ ID NO:21/ SEQ ID NO:22

FIG. 12

ES 2 705 540 T3

A	AGTAATAATACGACTCACTATAG	SEQ ID NO:23
B	AAGCACCGACTCGGTGCCACTTTT	SEQ ID NO:24
C	TAATACGACTCACTATAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCTT	SEQ ID NO:25
D	TAATACGACTCACTATAGCAGGACAGCATAGCAAGTTGA GATAAGGCTA	SEQ ID NO:26
E	TAGCCTTATCTCAACTTGCTATGCTGTCCTGCTATAGTGA GTCGTATTA	SEQ ID NO:27
F	TAATACGACTCACTATAGGGGCCACTAGGGACAGGATGT CTCAGAGCTATGCTGT	SEQ ID NO:28
G	ACAGCATAGCTCTGAGACATCCTGTCCCTAGTGGCCCCTA TAGTGAGTCGTATTA	SEQ ID NO:29
H	CCCCGTTCTCCTGTGGATTC	SEQ ID NO:30
I	ATCCTCTCTGGCTCCATCGT	SEQ ID NO:31
J	CACTCTTCCCTACACGACGCTCTTCCGATCTTCTGGCAA GGAGAGAGATGG	SEQ ID NO:32
K	GGAGTTCAGACGTGTGCTCTTCCGATCTTATATTTCCAGG GCCGGTTA	SEQ ID NO:33
L	CAAGCAGAAGACGGCATAACGAGATTACGTGATGTGACTG GAGTTCAGACGTGTGCTC	SEQ ID NO:34
M	AATGATACGGCGACCACCGAGATCTACACCGTCTAATAC ACTCTTTCCTACACGACG	SEQ ID NO:35
N	AATGATACGGCGACCACCGAGATCTACACTCTCTCCGAC ACTCTTTCCTACACGACG	SEQ ID NO:36
O	AATGATACGGCGACCACCGAGATCTACACTCGACTAGAC ACTCTTTCCTACACGACG	SEQ ID NO:37
P	AATGATACGGCGACCACCGAGATCTACACTTCTAGCTAC ACTCTTTCCTACACGACG	SEQ ID NO:38
Q	AATGATACGGCGACCACCGAGATCTACACCCTAGAGTAC ACTCTTTCCTACACGACG	SEQ ID NO:39
R	AATGATACGGCGACCACCGAGATCTACACCTATTAAGAC ACTCTTTCCTACACGACG	SEQ ID NO:40
S	GGCAGTAGCCTTATCTCAACTTGCTATGCTGTCCTGTTTCC AGGACAGCATAGCTCTGAGAC	SEQ ID NO:41
T	GGCAGTAGCCTTATCTCAAC	SEQ ID NO:42
U	TAATACGACTCACTATAGGCAGGTCCGTTATCAACTTGAA AAAGTGGCACCGAGTCGGTGCTT	SEQ ID NO:43
V	GGCAGTGAAGTAGCCTTATCTCAACTTGCTATGCTGTCCT GTTTCCAGGACAGCATAGCTCTGAGAC	SEQ ID NO:44
W	GGCAGTGAAGTAGCCTTATC	SEQ ID NO:45
X	TAATACGACTCACTATAGGCAGCTAAGGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGGTGCTT	SEQ ID NO:46
Y	TAATACGACTCACTATAGGGGCCACTAGGGACAGGATGT CTCAGAGCTATGCTGTC	SEQ ID NO:47
Z	TAATACGACTCACTATAGTTTGTGTTTCCATAAACTGGTC TCAGAGCTATGCTGTC	SEQ ID NO:48

FIG. 13

AA	TAATACGACTCACTATAGCCCGCCACCACCAGGATGTGTC TCAGAGCTATGCTGTC	SEQ ID NO:49
AB	TAATACGACTCACTATAGGCAGCCAGCATGATGAGACGT CTCAGAGCTATGCTGTC	SEQ ID NO:50
AC	AAGCACCGACTCGGTGCCAC	SEQ ID NO:51
AD	CACTCTTTCCCTACACGACGCTCTTCCGATCTACATGCAC ACCCATGTTTTG	SEQ ID NO:52
AE	GGAGTTCAGACGTGTGCTCTTCCGATCTAACATTTCCAGG TGACAGGC	SEQ ID NO:53
AF	CACTCTTTCCCTACACGACGCTCTTCCGATCTGTTCCGAC GCTCCTTGAA	SEQ ID NO:54
AG	GGAGTTCAGACGTGTGCTCTTCCGATCTCAGATGCGATGA CCTTTGTG	SEQ ID NO:55
AH	CACTCTTTCCCTACACGACGCTCTTCCGATCTAAGAAAGG CAAGAAGCCTGG	SEQ ID NO:56
AI	GGAGTTCAGACGTGTGCTCTTCCGATCTGCTGGCCTGAGA CATTCTA	SEQ ID NO:57
AJ	TAGCCTTATCTCAACTTGCTATGCTGTCCTGTTTCCAGGAC AGCATAGCTCTGAGAC	SEQ ID NO:58
AK	TAATACGACTCACTATAGGGGCCACTAGGGACAGGATGT CTCAGAGCTATGCAGTCC	SEQ ID NO:59
AL	CAGTAGCCTTATCTCAACTTGCTATGCAGTCCTGTTTCCA GGACTGCATAGCTCTGAGAC	SEQ ID NO:60
AM	CTGCCTATACGGCAGTAGCCTTATCTCAACTTGCTATGCA	SEQ ID NO:61
AN	TAATACGACTCACTATAGCTGCCGTATAGGCAGGTCCGTT ATCAACTTGAAAAAGTG	SEQ ID NO:62
AO	AAGCACCGACTCGGTGCCACTTTTCAAGTTGATAACGGA CCT	SEQ ID NO:63
AP	GTCTAGCCTTATCTCAACTTGCTATGCAGTCCTGTTTCCAG GACTGCATAGCTCTGAGAC	SEQ ID NO:64
AQ	CTGCCTATACGGCAGTGICTAGCCTTATCTCAACTTGCTA	SEQ ID NO:65
AR	TAATACGACTCACTATAGCTGCCGTATAGGCAGAGACAG TCCGTTATCAACTTGAAA	SEQ ID NO:66
AS	AAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGA CTGTCTCT	SEQ ID NO:67

FIG. 13 (Cont.)