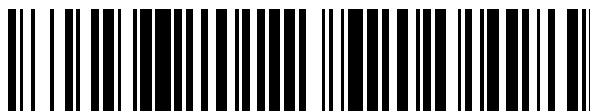


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 553 782**

51 Int. Cl.:

C12N 15/63

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.12.2013** **E 15154539 (9)**

97 Fecha y número de publicación de la concesión europea: **02.09.2015** **EP 2896697**

54 Título: **Ingeniería de sistemas, métodos y composiciones de guía optimizadas para manipulación de secuencias**

30 Prioridad:

12.12.2012 US 201261736527 P
02.01.2013 US 201361748427 P
30.01.2013 US 201361758468 P
25.02.2013 US 201361769046 P
15.03.2013 US 201361791409 P
15.03.2013 US 201361802174 P
28.03.2013 US 201361806375 P
20.04.2013 US 201361814263 P
06.05.2013 US 201361819803 P
28.05.2013 US 201361828130 P
17.06.2013 US 201361835931 P
17.06.2013 US 201361836127 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.12.2015

73 Titular/es:

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

72 Inventor/es:

ZHANG, FENG;
CONG, LE;
HSU, PATRICK y
RAN, FEI

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 553 782 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

Ingeniería de sistemas, métodos y composiciones de guía optimizadas para manipulación de secuencias.

APLICACIONES RELACIONADAS E INCORPORACIÓN POR REFERENCIA

Campo de la invención

- 5 La presente invención se refiere en general a sistemas, métodos y composiciones usados para el control de la expresión de genes que implica fijar como objetivo una secuencia, tal como perturbación del genoma o edición de genes, que pueden usar sistemas vector relacionados con Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR, por sus siglas en inglés) y componentes de los mismos.

Informe en cuanto a investigación patrocinada federalmente

- 10 Esta invención se realizó con soporte del gobierno otorgado por el Pioneer Award del NIH, Institutos Nacionales de Salud, DP1MH100706. El gobierno tiene ciertos derechos en la invención.

Antecedentes de la invención

- 15 Los recientes avances en las técnicas de secuenciación del genoma y métodos de análisis han acelerado significativamente la capacidad para catalogar y cartografiar factores genéticos asociados a un diverso intervalo de funciones biológicas y enfermedades. Se requieren tecnologías precisas para fijar como objetivo genoma para permitir la ingeniería inversa sistemática de variaciones genéticas causales permitiendo la perturbación selectiva de elementos genéticos individuales, así como avanzar las aplicaciones de la biología sintética, biotecnológicas y médicas. Aunque están disponibles técnicas de edición de genoma tales como dedos de cinc diseñadores, efectores del tipo activador de la transcripción (los TALE, por sus siglas en inglés) o meganucleasas de migración para producir perturbaciones del genoma fijado como objetivo, sigue existiendo la necesidad de nuevas tecnologías de ingeniería del genoma que tengan un precio asequible, fáciles de montar, escalables y susceptibles de fijar como objetivo múltiples posiciones dentro del genoma eucariota.
- 20

Sumario de la descripción

- 25 Existe una necesidad apremiante de sistemas y técnicas alternativos y robustos para fijar como objetivo secuencias con una amplia serie de aplicaciones. Esta invención estudia esta necesidad y proporciona ventajas relacionadas. El CRISPR/Cas o el sistema CRISPR-Cas (ambos términos se usan de manera intercambiable por toda esta solicitud) no requiere la generación de proteínas personalizadas para secuencias específicas fijadas como objetivo sino más bien se puede programar una enzima Cas única mediante una molécula de ARN corta para reconocer un objetivo de ADN específico, en otras palabras la enzima Cas se puede reclutar para un objetivo de ADN específico usando dicha molécula de ARN corta. Añadir el sistema CRISPR-Cas al repertorio de técnicas de secuenciación del genoma y métodos de análisis puede simplificar de manera significativa la metodología y acelerar la capacidad para catalogar y cartografiar factores genéticos asociados a un diverso intervalo de funciones biológicas y enfermedades. Para utilizar el sistema CRISPR-Cas de manera eficaz para editar genoma sin efectos perjudiciales, es crítico comprender aspectos de ingeniería y optimización de estas herramientas de ingeniería del genoma, que son aspectos de la presente descripción.
- 30
- 35

De acuerdo con la presente invención, se proporciona un sistema vector de CRISPR-Cas de ingeniería, que no está presente en la naturaleza de acuerdo con la reivindicación 1. En las reivindicaciones dependientes se describen aspectos adicionales de la invención. La invención también proporciona los usos del sistema de acuerdo con las reivindicaciones 19 y 20.

- 40 En un aspecto, la descripción proporciona un sistema vector que comprende uno o más vectores. En algunas realizaciones, el sistema comprende: (a) un primer elemento regulador ligado de manera operable a una secuencia de emparejado tracr y uno o más sitios de inserción para insertar una o más secuencias guía aguas arriba de la secuencia de emparejado tracr, en la que cuando se expresa, la secuencia guía se dirige a unión específica de secuencias de un complejo CRISPR a una secuencia fijada como objetivo en una célula, por ejemplo, célula eucariota, en la que el complejo CRISPR comprende una enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia fijada como objetivo y (2) la secuencia de emparejado tracr que se hibrida a la secuencia tracr y (b) un segundo elemento regulador ligado de manera operable a una secuencia codificadora de enzimas que codifica dicha enzima CRISPR que comprende una secuencia de localización nuclear; en el que los componentes (a) y (b) están situados en los mismos vectores o diferentes del sistema. En algunas realizaciones, el componente (a) comprende además la secuencia tracr aguas abajo de la secuencia de emparejamiento tracr bajo el control del primer elemento regulador. En algunas realizaciones, el componente (a) comprende además dos o más secuencias guía ligadas de manera operable al primer elemento regulador, en el que cuando se expresan, cada una de las dos o más secuencias guía dirige unión específica de la secuencia de un complejo CRISPR a una secuencia objetivo diferente en una célula eucariota. En algunas realizaciones, el sistema comprende la secuencia tracr bajo el control de un tercer elemento regulador, tal como un activador de polimerasa III. En algunas realizaciones, la secuencia tracr presenta al menos 50%, 60%, 70%, 80%, 90%, 95% o 99% de complementariedad de secuencias a lo largo de
- 55

la longitud de la secuencia de emparejado tracr cuando se alinean de manera óptima. En algunas realizaciones, el complejo CRISPR comprende una o más secuencias de localización nuclear de suficiente fortaleza para conducir la acumulación de dicho complejo CRISPR en una cantidad detectable en el núcleo de una célula eucariota. Sin desear estar ligados por la teoría, se cree que una secuencia de localización nuclear no es necesariamente para actividad del complejo CRISPR en eucariotas, sino que incluyendo tales secuencias se mejora la actividad del sistema, especialmente en cuanto a moléculas de ácido nucleico fijadas como objetivo en el núcleo. En algunas realizaciones, la enzima CRISPR es una enzima del sistema CRISPR de tipo II. En algunas realizaciones, la enzima CRISPR es una enzima Cas9. En algunas realizaciones, la enzima Cas9 es Cas9 *S. pneumoniae*, *S. pyogenes* o *S. thermophilus* y puede incluir Cas9 mutada derivada de estos organismos. La enzima puede ser un homólogo u ortólogo de Cas9. En algunas realizaciones, la enzima CRISPR es codón-optimizada para expresión en una célula eucariota. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o dos hebras en la posición de la secuencia objetivo. En algunas realizaciones, la enzima CRISPR carece de actividad de escisión de la cadena de ADN. En algunas realizaciones, el primer elemento regulador es un activador de la polimerasa III. En algunas realizaciones, el segundo elemento regulador es un activador de la polimerasa II. En algunas realizaciones, la secuencia guía tiene al menos 15, 16, 17, 18, 19, 20, 25 nucleótidos o entre 10-30 o entre 15-25 o entre 15-20 nucleótidos de longitud. En general y por toda esta memoria descriptiva, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha unido. Los vectores incluyen, pero no se limitan a, moléculas de ácido nucleico que son monocatenarias, bicatenarias o parcialmente bicatenarias; moléculas de ácido nucleico que comprenden uno o más extremos libres, ningún extremo libre (por ejemplo, circular); moléculas de ácido nucleico que comprenden ADN, ARN o ambos y otras variedades de polinucleótidos conocidos en la técnica. Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN de doble cadena circular en que se pueden insertar segmentos de ADN adicionales, tal como por técnicas de clonación molecular clásicas. Otro tipo de vector es un vector vírico, en el que están presentes secuencias de ADN o ARN derivadas víricamente en el vector para empaquetamiento en un virus (por ejemplo, retrovirus, retrovirus de replicación defectuosa, adenovirus, adenovirus de replicación defectuosa y virus adenoasociados). Los vectores víricos también incluyen polinucleótidos soportados por un virus para transfección a una célula huésped. Algunos vectores son capaces de replicación autónoma en una célula huésped en que se introducen (por ejemplo, vectores bacterianos con un origen bacteriano de replicación y vectores de mamífero episomales). Otros vectores (por ejemplo, vectores de mamíferos no episomales) se integran en el genoma de una célula huésped en la introducción en la célula huésped y se replican de ese modo junto con el genoma huésped. Por otra parte, algunos vectores son capaces de dirigir la expresión de los genes a los que están unidos de manera operativa. Dichos vectores se refieren en la presente memoria como "vectores de expresión". Los vectores de expresión comunes de utilidad en técnicas de ADN recombinante están con frecuencia en la forma de plásmidos.

Los vectores de expresión recombinantes pueden comprender un ácido nucleico de la descripción en una forma adecuada para expresión del ácido nucleico en una célula huésped, que significa que los vectores de expresión recombinantes incluyen uno o más elementos reguladores, que se pueden seleccionar sobre la base de las células huésped que se tienen que usar para expresión, que están unidos de manera operativa a la secuencia de ácidos nucleicos que se tiene que expresar. En un vector de expresión recombinante, "unido de manera operable" significa que la secuencia de nucleótidos de interés está unida al elemento o a los elementos reguladores de una manera que permite la expresión de la secuencia de nucleótidos (por ejemplo, en un sistema de transcripción/traducción in vitro o en una célula huésped cuando el vector se introduce en la célula huésped).

Se pretende que el término "elemento regulador" incluya activadores, potenciadores, sitios de entrada de ribosomas internos (IRES, por sus siglas en inglés) y otros elementos de control de la expresión (por ejemplo, señales de terminación de la transcripción, tales como señales de poliadenilación y secuencias poli-U). Tales elementos reguladores se describen, por ejemplo, en Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1.990). Los elementos reguladores incluyen aquellos que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de célula huésped y aquellos que dirigen la expresión de la secuencia de nucleótidos sólo en ciertas células huésped (por ejemplo, secuencias reguladoras específicas del tejido). Un activador específico del tejido puede dirigir la expresión principalmente en un tejido deseado de interés, tal como músculo, neurona, hueso, piel, sangre, órganos específicos (por ejemplo, hígado, páncreas) o tipos de células particulares (por ej., linfocitos). Los elementos reguladores también pueden dirigir la expresión de una manera dependiente temporal, tal como de una manera dependiente del ciclo celular o dependiente de la fase de desarrollo, que puede ser o no también específica del tejido o del tipo de célula. En algunas realizaciones, un vector comprende uno o más activadores de pol III (por ej., 1, 2, 3, 4, 5 o más activadores de pol III), uno o más activadores de pol II (por ej., 1, 2, 3, 4, 5 o más activadores de pol II), uno o más activadores de pol I (por ej., 1, 2, 3, 4, 5 o más activadores de pol I) o combinaciones de los mismos. Ejemplos de activadores de pol III incluyen, pero no se limitan a, activadores U6 y H1. Ejemplos de activadores de pol II incluyen, pero no se limitan a, el activador LTR del virus del sarcoma de Rous (RSV, por sus siglas en inglés) retrovírico (opcionalmente con el potenciador de RSV), el activador de citomegalovirus (CMV) (opcionalmente con el estimulador de CMV) [véase, por ej., Boshart et al, Cell, 41: 521-530 (1.985)], el activador de SV40, el activador de dihidrofolato reductasa, el activador de β -actina, el activador de fosfoglicerol kinasa (PGK, por sus siglas en inglés) y el activador de EF1 α . También están incluidos en el término "elemento regulador" los elementos estimuladores, tales como WPRE; estimuladores de CMV; el segmento R-U5' en LTR de HTLV-1 (Mol. Cell. Biol., Vol. 8 (1), pág. 466-472, 1.988); estimulador de SV40 y la secuencia del intrón entre los exones 2 y 3 de β -globina de conejo (Proc. Natl. Acad. Sci.

USA., Vol. 78 (3), pág. 1.527-31, 1.981). Se apreciará por los expertos en la materia que el diseño del vector de expresión puede depender de factores tales como la elección de la célula huésped que se tiene que transformar, el nivel de expresión deseado, etc. Se puede introducir un vector en células huésped para producir de ese modo transcritos, proteínas o péptidos, incluyendo proteínas o péptidos de fusión, codificados por ácidos nucleicos como se describe en la presente memoria (por ejemplo, transcritos de repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR), proteínas, enzimas, formas mutantes de los mismos, proteínas de fusión de los mismos, etc.).

Vectores ventajosos incluyen lentivirus y virus adenoasociados y también se pueden seleccionar tipos de tales vectores para fijar como objetivo tipos particulares de células.

En un aspecto, la descripción proporciona un vector que comprende un elemento regulador unido de manera operable a una secuencia codificadora de enzimas que codifica una enzima CRISPR que comprende una o más secuencias de localización nucleares. En algunas realizaciones, dicho elemento regulador conduce la transcripción de la enzima CRISPR en una célula eucariota de manera que dicha enzima CRISPR se acumula en una cantidad detectable en el núcleo de la célula eucariota. En algunas realizaciones, el elemento regulador es un activador de la polimerasa II. En algunas realizaciones, la enzima CRISPR es una enzima del sistema CRISPR de tipo II. En algunas realizaciones, la enzima CRISPR es una enzima Cas9. En algunas realizaciones, la enzima Cas9 es Cas9 de *S. pneumoniae*, *S. pyogenes* o *S. thermophilus* y puede incluir Cas9 mutada procedente de estos organismos. En algunas realizaciones, la enzima CRISPR es codón-optimizada para expresión en una célula eucariota. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o dos hebras en la posición de la secuencia objetivo. En algunas realizaciones, la enzima CRISPR carece de actividad de escisión de la cadena de ADN.

En un aspecto, la descripción proporciona una enzima CRISPR que comprende una o más secuencias de localización nuclear de suficiente fortaleza para conducir la acumulación de dicha enzima CRISPR en una cantidad detectable en el núcleo de una célula eucariota. En algunas realizaciones, la enzima CRISPR es una enzima del sistema CRISPR de tipo II. En algunas realizaciones, la enzima CRISPR es una enzima Cas9. En algunas realizaciones, la enzima Cas9 es Cas9 de *S. pneumoniae*, *S. pyogenes* o *S. thermophilus* y puede incluir Cas9 mutada procedente de estos organismos. La enzima pueden ser un homólogo u ortólogo de Cas9. En algunas realizaciones, la enzima CRISPR carece de capacidad para escindir una o más hebras de una secuencia objetivo a la que se une.

En un aspecto, la descripción proporciona una célula eucariota huésped que comprende: (a) un primer elemento regulador unido de manera operable a una secuencia de emparejamiento tracr y uno o más sitios de inserción para insertar una o más secuencias guía aguas arriba de la secuencia de emparejamiento tracr, en la que cuando se expresa, la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a una secuencia objetivo en una célula eucariota, en la que el complejo CRISPR comprende una enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr y/o (b) un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica a dicha enzima CRISPR que comprende una secuencia de localización nuclear. En algunas realizaciones, la célula huésped comprende los componentes (a) y (b). En algunas realizaciones, el componente (a), el componente (b) o los componentes (a) y (b) están integrados de manera estable en un genoma de la célula eucariota huésped. En algunas realizaciones, el componente (a) comprende además la secuencia tracr aguas abajo de la secuencia de emparejamiento tracr bajo el control del primer elemento regulador. En algunas realizaciones, el componente (a) comprende además dos o más secuencias guía unidas de manera operable al primer elemento regulador, en el que cuando se expresan, cada una de las dos o más secuencias guía dirige la unión específica de la secuencia de un complejo CRISPR a una secuencia objetivo diferente en una célula eucariota. En algunas realizaciones, la célula huésped eucariota comprende además un tercer elemento regulador, tal como un activador de la polimerasa III, unido de manera operable a una dicha secuencia tracr. En algunas realizaciones, la secuencia tracr presenta al menos 50%, 60%, 70%, 80%, 90%, 95% o 99% de complementariedad de secuencias a lo largo de la longitud de la secuencia de emparejamiento tracr cuando se alinean de manera óptima. En algunas realizaciones, la enzima CRISPR comprende una o más secuencias de localización nuclear de suficiente fortaleza para conducir la acumulación de dicha enzima CRISPR en una cantidad detectable en el núcleo de una célula eucariota. En algunas realizaciones, la enzima CRISPR es una enzima del sistema CRISPR de tipo II. En algunas realizaciones, la enzima CRISPR es una enzima Cas9. En algunas realizaciones, la enzima Cas9 es Cas9 de *S. pneumoniae*, *S. pyogenes* o *S. thermophilus* y puede incluir Cas9 mutada procedente de estos organismos. La enzima puede ser un homólogo u ortólogo de Cas9. En algunas realizaciones, la enzima CRISPR es codón-optimizada para expresión en una célula eucariota. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o dos hebras en la posición de la secuencia objetivo. En algunas realizaciones, la enzima CRISPR carece de actividad de escisión de la hebra de ADN. En algunas realizaciones, el primer elemento regulador es un activador de la polimerasa III. En algunas realizaciones, el segundo elemento regulador es un activador de la polimerasa II. En algunas realizaciones, la secuencia guía es al menos 15, 16, 17, 18, 19, 20, 25 nucleótidos o entre 10-30 o entre 15-25 o entre 15-20 nucleótidos de longitud. En un aspecto, la descripción proporciona un organismo eucariota no humano; preferiblemente un organismo eucariota multicelular, que comprende una célula eucariota huésped según cualquiera de las realizaciones descritas. En otros aspectos, la descripción proporciona un organismo eucariota; preferiblemente un organismo eucariota multicelular, que comprende una célula eucariota huésped según cualquiera

de las realizaciones descritas. El organismo en algunas realizaciones de estos aspectos puede ser un animal; por ejemplo un mamífero. También, el organismo puede ser un artrópodo tal como un insecto. El organismo puede también ser una planta. Además, el organismo puede ser un hongo.

En un aspecto, la descripción proporciona un estuche que comprende uno o más de los componentes descritos en la presente memoria. En algunas realizaciones, el estuche comprende un sistema vector e instrucciones para usar el estuche. En algunas realizaciones, el sistema vector comprende: (a) un primer elemento regulador unido de manera operable a una secuencia de emparejamiento tracr y uno o más sitios de inserción para insertar una o más secuencias guía aguas arriba de la secuencia de emparejamiento tracr, en el que cuando se expresan, la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a una secuencia fijada como objetivo en una célula eucariota, en la que el complejo CRISPR comprende una enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr y/o (b) un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica dicha enzima CRISPR que comprende una secuencia de localización nuclear. En algunas realizaciones, el estuche comprende los componentes (a) y (b) situados en los mismos vectores del sistema o diferentes. En algunas realizaciones, el componente (a) comprende además la secuencia tracr aguas abajo de la secuencia de emparejamiento tracr bajo el control del primer elemento regulador. En algunas realizaciones, el componente (a) comprende además dos o más secuencias guía unidas de manera operable al primer elemento regulador, en el que cuando se expresan, cada una de las dos o más secuencias guía dirige la unión específica de la secuencia de un complejo CRISPR a una secuencia objetivo diferente en una célula eucariota. En algunas realizaciones, el sistema comprende además un tercer elemento regulador, tal como un activador de la polimerasa III, unido de manera operable a dicha secuencia tracr. En algunas realizaciones, la secuencia tracr presenta al menos 50%, 60%, 70%, 80%, 90%, 95% o 99% de complementariedad de secuencias a lo largo de la longitud de la secuencia de emparejamiento tracr cuando se alinea de manera óptima. En algunas realizaciones, la enzima CRISPR comprende una o más secuencias de localización nuclear de suficiente fortaleza para conducir la acumulación de dicha enzima CRISPR en una cantidad detectable en el núcleo de una célula eucariota. En algunas realizaciones, la enzima CRISPR es una enzima del sistema CRISPR de tipo II. En algunas realizaciones, la enzima CRISPR es una enzima Cas9. En algunas realizaciones, la enzima Cas9 es Cas9 de *S. pneumoniae*, *S. pyogenes* o *S. thermophilus* y puede incluir Cas9 mutada procedente de estos organismos. La enzima puede ser un homólogo u ortólogo de Cas9. En algunas realizaciones, la enzima CRISPR es codón-optimizada para expresión en una célula eucariota. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o dos hebras en la posición de la secuencia objetivo. En algunas realizaciones, la enzima CRISPR carece de actividad de escisión de la hebra de ADN. En algunas realizaciones, el primer elemento regulador es un activador de la polimerasa III. En algunas realizaciones, el segundo elemento regulador es un activador de la polimerasa II. En algunas realizaciones, la secuencia guía es al menos 15, 16, 17, 18, 19, 20, 25 nucleótidos o entre 10-30 o entre 15-25 o entre 15-20 nucleótidos de longitud.

En un aspecto, la descripción proporciona un método para modificar un polinucleótido objetivo en una célula eucariota. En algunas realizaciones, el método comprende permitir que se una un complejo CRISPR al polinucleótido objetivo para efectuar la escisión de dicho polinucleótido objetivo modificando de ese modo el polinucleótido objetivo, en el que el complejo CRISPR comprende una enzima CRISPR complejada con una secuencia guía se transcribe a una secuencia fijada como objetivo dentro de dicho polinucleótido objetivo, en el que dicha secuencia guía está ligada a una secuencia de emparejamiento tracr que a su vez se transcribe a una secuencia tracr. En algunas realizaciones, dicha escisión comprende escindir una o dos hebras en la posición de la secuencia objetivo mediante dicha enzima CRISPR. En algunas realizaciones, dicha escisión da como resultado la transcripción disminuida de un gen objetivo. En algunas realizaciones, el método comprende además reparar dicho polinucleótido objetivo escindido por recombinación homóloga con un polinucleótido exógeno de plantilla, en el que dicha reparación da como resultado una mutación que comprende una inserción, supresión o sustitución de uno o más nucleótidos de dicho polinucleótido objetivo. En algunas realizaciones, dicha mutación da como resultado uno o más cambios de aminoácido en una proteína expresada de un gen que comprende la secuencia objetivo. En algunas realizaciones, el método comprende además suministrar uno o más vectores a dicha célula eucariota, en el que uno o más vectores conducen la expresión de uno o más de: la enzima CRISPR, la secuencia guía ligada a la secuencia de emparejamiento tracr y la secuencia tracr. En algunas realizaciones, dichos vectores se suministran a la célula eucariota en un individuo. En algunas realizaciones, dicha modificación tiene lugar en dicha célula eucariota en un cultivo celular. En algunas realizaciones, el método comprende además aislar dicha célula eucariota de un individuo antes de dicha modificación. En algunas realizaciones, el método comprende además devolver dicha célula y/o células eucariotas procedentes de ahí a dicho individuo.

En un aspecto, la descripción proporciona un método para modificar la expresión de un polinucleótido en una célula eucariota. En algunas realizaciones, el método comprende permitir que un complejo CRISPR se una al polinucleótido de manera que dicha unión dé como resultado la expresión aumentada o disminuida de dicho polinucleótido; en el que el complejo CRISPR comprende una enzima CRISPR complejada con una secuencia guía se transcribe a una secuencia fijada como objetivo dentro de dicho polinucleótido, en el que dicha secuencia guía está ligada a una secuencia de emparejamiento tracr que a su vez se transcribe a una secuencia tracr. En algunas realizaciones, el método comprende además suministrar uno o más vectores a dichas células eucariotas, en el que uno o más vectores conducen la expresión de una o más de: la enzima CRISPR, la secuencia guía ligada a la secuencia de emparejamiento tracr y la secuencia tracr.

En un aspecto, la descripción proporciona un método para generar una célula eucariota modelo que comprende un gen mutado de la enfermedad. En algunas realizaciones, un gen de la enfermedad es cualquier gen asociado a un aumento del riesgo de tener o desarrollar una enfermedad. En algunas realizaciones, el método comprende (a) introducir uno o más vectores en una célula eucariota, en la que uno o más vectores conducen la expresión de uno o más de: una enzima CRISPR, una secuencia guía ligada a una secuencia de emparejamiento tracr y una secuencia tracr y (b) permitir que un complejo CRISPR se una a un polinucleótido objetivo para efectuar la escisión del polinucleótido objetivo dentro de dicho gen de la enfermedad, en el que el complejo CRISPR comprende la enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo dentro del polinucleótido objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr, generando de ese modo una célula eucariota modelo que comprende un gen mutado de la enfermedad. En algunas realizaciones, dicha escisión comprende escindir una o dos hebras en la posición de la secuencia objetivo mediante dicha enzima CRISPR. En algunas realizaciones, dicha escisión da como resultado la transcripción disminuida de un gen objetivo. En algunas realizaciones, el método comprende además reparar dicho polinucleótido objetivo escindido por recombinación homóloga con un polinucleótido exógeno de plantilla, en el que dicha reparación da como resultado una mutación que comprende una inserción, supresión o sustitución de uno o más nucleótidos de dicho polinucleótido objetivo. En algunas realizaciones, dicha mutación da como resultado uno o más cambios de aminoácido en una expresión de proteína de un gen que comprende la secuencia objetivo.

En un aspecto, la descripción proporciona un método para desarrollar un agente biológicamente activo que module un caso de señalización celular asociado a un gen de la enfermedad. En algunas realizaciones, un gen de la enfermedad es cualquier gen asociado a un aumento del riesgo de tener o desarrollar una enfermedad. En algunas realizaciones, el método comprende (a) poner en contacto un compuesto de ensayo con una célula modelo de una cualquiera de las realizaciones descritas y (b) detectar un cambio en una lectura que sea indicativo de una reducción o un aumento del caso de señalización celular asociado a dicha mutación en dicho gen de la enfermedad, desarrollando de ese modo dicho agente biológicamente activo que module dicho caso de señalización celular asociado a dicho gen de la enfermedad.

En un aspecto, la descripción proporciona un polinucleótido recombinante que comprende una secuencia guía aguas arriba de una secuencia de emparejamiento tracr, en el que la secuencia guía cuando se expresa dirige la unión específica de la secuencia de un complejo CRISPR a una correspondiente secuencia objetivo presente en una célula eucariota. En algunas realizaciones, la secuencia objetivo es una secuencia vírica presente en una célula eucariota. En algunas realizaciones, la secuencia objetivo es un proto-oncogén o un oncogén.

En un aspecto, la descripción proporciona un método para seleccionar una o más células procariotas por introducción de una o más mutaciones en un gen en una o más células procariotas, comprendiendo el método: introducir uno o más vectores en la célula o las células procariotas, en las que uno o más vectores conducen la expresión de uno o más de: una enzima CRISPR, una secuencia guía ligada a una secuencia de emparejamiento tracr, una secuencia tracr y una plantilla de edición; en la que la plantilla de edición comprende una o más mutaciones que anula la escisión de la enzima CRISPR; permitir la recombinación homóloga de la plantilla de edición con el polinucleótido objetivo en la célula o las células que se tienen que seleccionar; permitir que un complejo CRISPR se una a un polinucleótido objetivo para efectuar la escisión del polinucleótido objetivo dentro de dicho gen, en el que el complejo CRISPR comprende la enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo dentro del polinucleótido objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr, en la que la unión del complejo CRISPR al polinucleótido objetivo induce la muerte celular, permitiendo de ese modo que se seleccione una o más células procariotas en que se han introducido una o más mutaciones. En una realización preferida, la enzima CRISPR es Cas9. En otro aspecto de la descripción, la célula que se tiene que seleccionar puede ser una célula eucariota. Los aspectos de la descripción permiten la selección de células específicas sin que se requiera un marcador de selección o un procedimiento de dos etapas que pueda incluir un sistema de contraselección.

En algunos aspectos, la descripción proporciona una composición que no se encuentra en la naturaleza o de ingeniería que comprende una secuencia de polinucleótidos de ARN quimérico (ARNqui) del sistema CRISPR-Cas, en el que la secuencia de polinucleótidos comprende: (a) una secuencia guía capaz de hibridar una secuencia fijada como objetivo en una célula eucariota, (b) una secuencia de emparejamiento tracr y (c) una secuencia tracr en la que (a), (b) y (c) se disponen en una orientación 5' a 3', en la que cuando se transcribe, la secuencia de emparejamiento tracr se hibrida a la secuencia tracr y la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo, en la que el complejo CRISPR comprende una enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr o un sistema de enzima CRISPR, en el que el sistema está codificado por un sistema vector que comprende uno o más vectores que comprenden I. un primer elemento regulador unido de manera operable a una secuencia de polinucleótidos de ARN quimérico (ARNqui) del sistema CRISPR-Cas, en el que la secuencia de polinucleótidos comprende (a) una o más secuencias guía capaces de hibridar una o más secuencias objetivo en una célula eucariota, (b) una secuencia de emparejamiento tracr y (c) una o más secuencias tracr y II. un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica una enzima CRISPR que comprende al menos una o más secuencias de localización nuclear, en la que (a), (b) y (c) se disponen en una orientación 5' a 3', en la que los componentes I y II se sitúan en los mismos vectores del

sistema o diferentes, en la que cuando se transcribe, la secuencia de emparejamiento tracr se hibrida a la secuencia tracr y la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo, en la que el complejo CRISPR comprende la enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr o un sistema de enzima CRISPR multiplexado, en la que el sistema se codifica mediante un sistema vector que comprende uno o más vectores que comprenden I. un primer elemento regulador unido de manera operable a (a) una o más secuencias guía capaces de hibridar una secuencia objetivo en una célula y (b) al menos una o más secuencias de emparejamiento tracr, II. un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica una enzima CRISPR y III. un tercer elemento regulador unido de manera operable a una secuencia tracr, en la que los componentes I, II y III se sitúan en los mismos vectores del sistema o diferentes, en la que cuando se transcribe, la secuencia de emparejamiento tracr se hibrida a la secuencia tracr y la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo, en la que el complejo CRISPR comprende la enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr y en la que en el sistema multiplexado se usan múltiples secuencias guía y una secuencia tracr única y en la que se modifica una o más de las secuencias guía, tracr y de emparejamiento tracr, para mejorar la estabilidad.

En aspectos de la descripción, la modificación comprende una estructura secundaria de ingeniería. Por ejemplo, la modificación puede comprender una reducción en la región de hibridación entre la secuencia de emparejamiento tracr y la secuencia tracr. Por ejemplo, la modificación también puede comprender fusionar la secuencia de emparejamiento tracr y la secuencia tracr por un bucle artificial. La modificación puede comprender la secuencia tracr que tiene una longitud entre 40 y 120 pb. En las realizaciones de la descripción, la secuencia tracr tiene entre 40 pb y la longitud total de la secuencia tracr. En algunas realizaciones, la longitud de ARNtracr incluye al menos 1-67 nucleótidos y en algunas realizaciones al menos 1-85 nucleótidos del ARNtracr de tipo natural. En algunas realizaciones, se pueden usar al menos los nucleótidos que corresponden a los nucleótidos 1-67 o 1-85 de ARNtracr de Cas9 de *S. pyogenes* de tipo natural. En el caso de que el sistema CRISPR use enzimas distintas de Cas9 o distintas de SpCas9, entonces pueden estar presentes los correspondientes nucleótidos en el ARNtracr de tipo natural relevante. En algunas realizaciones, la longitud de ARNtracr incluye no más de los nucleótidos 1-67 o 1-85 de ARNtracr de tipo natural. La modificación puede comprender optimización de la secuencia. En algunos aspectos, la optimización de secuencias puede comprender reducir la incidencia de secuencias poliT en la secuencia tracr y/o de emparejamiento tracr. La optimización de secuencias se puede combinar con la reducción en la región de hibridación entre la secuencia de emparejamiento tracr y la secuencia tracr; por ejemplo, una secuencia tracr de longitud reducida.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o el sistema de enzima CRISPR en el que la modificación comprende la reducción de secuencias poliT en la secuencia tracr y/o de emparejamiento tracr. En algunos aspectos de la descripción, una o más T presentes en una secuencia poli-T de la secuencia de tipo natural relevante (esto es, una extensión de más de 3, 4, 5, 6 o más bases T contiguas; en algunas realizaciones, un tramo de no más de 10, 9, 8, 7, 6 bases T contiguas) puede ser sustituido con un nucleótido no T, por ej. una A, de manera que se rompa la cadena en tramos más pequeños de T teniendo cada tramo 4 o menos de 4 (por ejemplo 3 ó 2) T contiguos. Las bases distintas de A se pueden usar para sustitución, por ejemplo C o G o nucleótidos que no se encuentran en la naturaleza o nucleótidos modificados. Si la cadena de T está implicada en la formación de una horquilla (o tallo-lazo), entonces es ventajoso que la base complementaria para la base no T cambie para complementar el nucleótido no T. Por ejemplo, si la base no T es una A, entonces un complemento se puede cambiar a una T, por ejemplo para conservar o ayudar a la conservación de estructura secundaria. Por ejemplo, se puede modificar 5'-TTTTT para convertirlo en 5'-TTTAT y se puede cambiar el 5'-AAAAA complementario a 5'-ATAAA.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende añadir una secuencia terminadora poliT. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende añadir una secuencia terminadora poliT en secuencias tracr y/o de emparejamiento tracr. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en la que la modificación comprende añadir una secuencia terminadora poliT en la secuencia guía. La secuencia terminadora poliT puede comprender 5 bases T contiguas o más de 5.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende modificar lazos y/u horquillas. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende proporcionar un mínimo de dos horquillas en la secuencia guía. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende proporcionar una horquilla formada por complementación entre la secuencia tracr y de emparejamiento tracr (repetición directa). En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende proporcionar una o más horquillas adicionales en o hacia el extremo 3' de la secuencia ARNtracr. Por ejemplo, se puede formar una horquilla proporcionando secuencias autocomplementarias dentro de la secuencia de ARNtracr unida por un lazo de manera que se forma una horquilla en autoplegado. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o

sistema de enzima CRISPR en el que la modificación comprende proporcionar horquillas adicionales añadidas a la 3' de la secuencia guía. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende extender el extremo 5' de la secuencia guía. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende proporcionar una o más horquillas en el extremo 5' de la secuencia guía. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende agregar la secuencia (5'-AGGACGAAGTCCTAA) al extremo 5' de la secuencia guía. Otras secuencias adecuadas para formar horquillas serán conocidas para el experto y se pueden usar en ciertos aspectos de la descripción. En algunos aspectos de la descripción, se proporcionan al menos 2, 3, 4, 5 o más horquillas adicionales. En algunos aspectos de la descripción se proporcionan no más de 10, 9, 8, 7, 6 horquillas adicionales. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende dos horquillas. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende tres horquillas. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende a lo sumo cinco horquillas.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende proporcionar unión cruzada o proporcionar uno o más nucleótidos modificados en la secuencia de polinucleótidos. Los nucleótidos modificados y/o ligados cruzados se pueden proporcionar en cualquiera o en todas las secuencias tracr, de acoplamiento tracr y/o guía y/o en la secuencia codificadora de enzimas y/o en secuencias de vectores. Las modificaciones pueden incluir la inclusión de al menos un nucleótido que no se encuentra la naturaleza o un nucleótido modificado o análogos de los mismos. Se puede modificar nucleótidos modificados en el resto ribosa, fosfato y/o base. Los nucleótidos modificados pueden incluir 2'-O-metil-análogos, 2'-desoxi-análogos o 2'-fluoro-análogos. La cadena principal de ácidos nucleicos se puede modificar, por ejemplo, se puede usar una cadena principal de fosforotioato. El uso de ácidos nucleicos bloqueados (LNA, por sus siglas en inglés) o ácidos nucleicos de puente (BNA, por sus siglas en inglés) también puede ser posible. Más ejemplos de bases modificadas incluyen, pero no se limitan a, 2-aminopurina, 5-bromo-uridina, pseudouridina, inosina, 7-metilguanosa.

Se entenderá que cualquiera o todas las modificaciones anteriores se pueden proporcionar en aislamiento o en combinación en un sistema CRISPR-Cas o sistema de enzima CRISPR determinado. Dicho sistema puede incluir una, dos, tres, cuatro, cinco o más de dichas modificaciones.

En un aspecto, la invención proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la enzima CRISPR es una enzima del sistema CRISPR de tipo II, por ej., una enzima Cas9. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la enzima CRISPR está constituida por menos de mil aminoácidos, o menos de cuatro mil aminoácidos. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la enzima Cas9 es StCas9 o StlCas9 o la enzima Cas9 es una enzima Cas9 de un organismo seleccionado del grupo que consiste en el género *Streptococcus*, *Campylobacter*, *Nitratifactor*, *Staphylococcus*, *Parvibaculum*, *Roseburia*, *Neisseria*, *Gluconacetobacter*, *Azospirillum*, *Sphaerochaeta*, *Lactobacillus*, *Eubacterium* o *Corynebacter*. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la enzima CRISPR es una nucleasa que dirige la escisión de ambas cadenas en la posición de la secuencia objetivo.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que el primer elemento regulador es un activador de la polimerasa III. En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que el segundo elemento regulador es un activador de la polimerasa II.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la secuencia guía comprende al menos quince nucleótidos.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la modificación comprende ARN de secuencia tracr optimizada y/o de secuencia guía optimizada y/o estructura co-plegada de secuencia tracr y/o secuencia o secuencias de acoplamiento tracr y/o estructuras secundarias estabilizantes de secuencia tracr y/o secuencia tracr con una región reducida de emparejamiento de bases y/o elementos de ARN condensados de secuencia tracr y/o en el sistema multiplexado hay dos ARN que comprenden un trazador y que comprenden una pluralidad de guías o ARN que comprende una pluralidad de quimeras.

En aspectos de la descripción, la arquitectura de ARN quimérico se optimiza además según los resultados de estudios de mutagénesis. En ARN quimérico con dos o más horquillas, las mutaciones en la repetición directa proximal para estabilizar la horquilla pueden dar como resultado ablación de la actividad del complejo CRISPR. Las mutaciones en la repetición directa distal para acortar o estabilizar la horquilla pueden no tener efecto sobre la actividad del complejo CRISPR. La aleatorización de secuencias en la región de la protuberancia entre las repeticiones proximal y distal puede reducir de manera significativa la actividad del complejo CRISPR. Los cambios en los pares de bases únicos o aleatorización de secuencias en la región ligadora entre horquillas pueden dar como resultado la pérdida completa de actividad del complejo CRISPR. La estabilización de horquillas de las horquillas distales que siguen a la primera horquilla después de la secuencia guía puede dar como resultado el mantenimiento

o la mejora de la actividad del complejo CRISPR. De acuerdo con esto, en realizaciones preferidas de la descripción, la arquitectura de ARN quimérico se puede optimizar además por generación de un ARN quimérico más pequeño que puede ser beneficioso para opciones de suministro terapéutico y otros usos y esto se puede conseguir modificando la repetición directa distal de manera que se acorte o estabilice la horquilla. En realizaciones preferidas adicionales de la descripción la arquitectura de ARN quimérico se puede optimizar además por estabilización de una o más horquillas distales. La estabilización de horquillas puede incluir modificar secuencias adecuadas para formar horquillas. En algunos aspectos de la descripción, se proporcionan al menos 2, 3, 4, 5 o más horquillas adicionales. En algunos aspectos de la descripción, se proporcionan no más de 10, 9, 8, 7, 6 horquillas adicionales. En algunos aspectos de la descripción, la estabilización puede ser unión cruzada u otras modificaciones. Las modificaciones pueden incluir la inclusión de al menos un nucleótido que no se encuentre en la naturaleza o un nucleótido modificado o análogos de los mismos. Los nucleótidos modificados se pueden modificar en el resto ribosa, fosfato y/o base. Los nucleótidos modificados pueden incluir 2'-O-metil-análogos, 2'-desoxi-análogos o 2'-fluoro-análogos. Se puede modificar la cadena principal de ácidos nucleicos, por ejemplo, se puede usar una cadena principal de fosforotioato. El uso de ácidos nucleicos bloqueados (LNA) o ácidos nucleicos de puente (BNA) también puede ser posible. Más ejemplos de bases modificadas incluyen, pero no se limitan a, 2-aminopurina, 5-bromo-uridina, pseudouridina, inosina, 7-metilguanina.

En un aspecto, la descripción proporciona el sistema CRISPR-Cas o sistema de enzima CRISPR en el que la enzima CRISPR es codón-optimizada para expresión en una célula eucariota.

De acuerdo con esto, en algunos aspectos de la descripción, no se requiere necesariamente que sea fijada la longitud del ARNtracr requerida en una construcción de la descripción, por ejemplo, una construcción quimérica, y en algunos aspectos puede tener entre 40 y 120 pb y en algunos aspectos de la descripción hasta la longitud total de la tracr, por ejemplo, en algunos aspectos de la descripción hasta el extremo 3' de tracr como se marca por la señal de terminación de la transcripción en el genoma bacteriano. En algunas realizaciones, la longitud del ARNtracr incluye al menos 1-67 nucleótidos y en algunas realizaciones al menos 1-85 nucleótidos del ARNtracr de tipo natural. En algunas realizaciones, se pueden usar al menos los nucleótidos que corresponden a los nucleótidos 1-67 o 1-85 de ARNtracr de Cas9 de *S. pyogenes* de tipo natural. En el caso de que el sistema CRISPR use enzimas distintas de Cas9 u otras distintas de SpCas9, entonces pueden estar presentes los nucleótidos correspondientes en el ARNtracr de tipo natural relevante. En algunas realizaciones, la longitud del ARNtracr incluye no más de 1-67 o 1-85 nucleótidos del ARNtracr de tipo natural. Con respecto a la optimización de secuencias (por ejemplo, reducción en las secuencias poli-T), por ejemplo en cuanto a las cadenas de Ts internas a la secuencia de emparejamiento tracr (repetición directa) o al ARNtracr, en algunos aspectos de la descripción, uno o más Ts presentes en una secuencia poli-T de la secuencia de tipo natural relevante (esto es, un segmento de más de 3, 4, 5, 6 o más bases T contiguas; en algunas realizaciones, un segmento de no más de 10, 9, 8, 7, 6 bases T contiguas) pueden ser sustituidos con un nucleótido no T, por ejemplo, una A, de manera que la cadena se rompa en fragmentos más pequeños de T teniendo cada fragmento 4, o menos de 4 (por ejemplo, 3 o 2) T contiguas. Si la cadena de T está implicada en la formación de una horquilla (o tallo-lazo), entonces es ventajoso que la base complementaria para la base no T se cambie a complemento del nucleótido no T. Por ejemplo, si la base no T es una A, entonces su complemento se puede cambiar a una T, por ejemplo, para conservar o ayudar a la conservación de estructura secundaria. Por ejemplo, se puede modificar 5'-TTTTT para que se convierta en 5'-TTTAT y se puede cambiar la 5'-AAAAA complementaria a 5'-ATAAA. En cuanto a la presencia de secuencias terminadoras poli-T en transcrito de tracr + emparejamiento tracr, por ejemplo, un terminador poli-T (TTTTT o más), en algunos aspectos de la descripción es ventajoso que se añada al extremo del transcrito, si está en forma de dos ARN (tracr y emparejamiento de tracr) o único ARN guía. Con respecto a los lazos y horquillas en los transcritos tracr y de emparejamiento tracr, en algunos aspectos de la descripción es ventajoso que esté presente un mínimo de dos horquillas en el ARN guía quimérico. Una primera horquilla puede ser la horquilla formada por complementación entre la secuencia tracr y la secuencia de emparejamiento tracr (repetición directa). Una segunda horquilla puede estar en el extremo 3' de la secuencia ARNtracr y esto puede proporcionar estructura secundaria para interacción con Cas9. Se pueden añadir horquillas adicionales al 3' del ARN guía, por ejemplo, en algunos aspectos de la descripción para aumentar la estabilidad del ARN guía. Adicionalmente, el extremo 5' del ARN guía, en algunos aspectos de la descripción, se puede extender. En algunos aspectos de la descripción, se puede considerar 20 pb en el extremo 5' como una secuencia guía. Se puede extender la porción 5'. Se puede proporcionar una o más horquillas en la porción 5', por ejemplo, en algunos aspectos de la descripción, esto también puede mejorar la estabilidad del ARN guía. En algunos aspectos de la descripción, se puede proporcionar la horquilla específica adjuntando la secuencia (5'-AGGACGAAGTCCTAA) al extremo 5' de la secuencia guía y, en algunos aspectos de la descripción, esto puede ayudar a mejorar la estabilidad. Otras secuencias adecuadas para formar horquillas serán conocidas para el experto y se pueden usar en ciertos aspectos de la descripción. En algunos aspectos de la descripción, se proporcionan al menos 2, 3, 4, 5 o más horquillas adicionales. En algunos aspectos de la descripción se proporcionan no más de 10, 9, 8, 7, 6 horquillas adicionales. Lo anterior también proporciona aspectos de la descripción que implican estructura secundaria en secuencias guía. En algunos aspectos de la descripción puede haber unión cruzada y otras modificaciones, por ejemplo, para mejorar la estabilidad. Las modificaciones pueden incluir inclusiones de al menos un nucleótido que no se encuentre en la naturaleza o un nucleótido modificado o análogos de los mismos. Los nucleótidos modificados se pueden modificar en el resto ribosa, fosfato y/o base. Los nucleótidos modificados pueden incluir 2'-O-metil-análogos, 2'-desoxi-análogos o 2'-fluoro-análogos. Se puede modificar la cadena principal del ácido nucleico, por ejemplo, se puede usar una cadena principal de fosforotioato. El

uso de ácidos nucleicos bloqueados (LNA) o ácidos nucleicos de puente (BNA) también puede ser posible. Más ejemplos de bases modificadas incluyen, pero no se limitan a, 2-aminopurina, 5-bromo-uridina, pseudouridina, inosina, 7-metilguanosa. Dichas modificaciones o unión cruzada pueden estar presentes en la secuencia guía u otras secuencias adyacentes a la secuencia guía.

5 Breve descripción de los dibujos

Las características novedosas de esta descripción se explican particularmente en las reivindicaciones que la acompañan. Se tendrá un mejor entendimiento de las características y ventajas de la presente descripción por referencia a la siguiente descripción detallada que explica realizaciones ilustrativas, en que se utilizan los principios de la descripción, y los dibujos adjuntos de los que:

- 10 La Figura 1 muestra un modelo esquemático del sistema CRISPR. La nucleasa Cas9 de *Streptococcus pyogenes* (amarillo) se fija como objetivo para ADN genómico por un ARN guía sintético (ARNsg) que consiste en una secuencia guía de 20-nt (azul) y un andamio (rojo). Los pares de bases de la secuencia guía con el ADN objetivo (azul), directamente aguas arriba de una unidad adyacente protoespaciadora (PAM, por sus siglas en inglés; magenta) 5'-NGG requisito y Cas9 media una doble rotura de cadena (DSB, por sus siglas en inglés) ~3 pb aguas
15 arriba de la PAM (triángulo rojo).

La Figura 2A-F ilustra un sistema CRISPR ejemplar, un posible mecanismo de acción, una adaptación de ejemplo para la expresión en células eucariotas y los resultados de los ensayos que evalúan la localización nuclear y la actividad de CRISPR.

- 20 La Figura 3A-C ilustra una casete de expresión ejemplar para expresión de elementos del sistema CRISPR en células eucariotas, estructuras previstas de secuencias guía de ejemplo y actividad del sistema CRISPR cuando se mide en células eucariotas y procariotas.

La Figura 4A-D ilustra los resultados de una evaluación de especificidad de SpCas9 para un objetivo de ejemplo.

La Figura 5A-G ilustra un sistema de vectores ejemplar y los resultados para su uso en dirigir recombinación homóloga en células eucariotas.

- 25 La Figura 6A-C ilustra una comparación de diferentes transcritos de ARNtracr para fijar como objetivo genes mediados por Cas9.

La Figura 7A-D ilustra un sistema CRISPR ejemplar, una adaptación de ejemplo para expresión en células eucariotas y los resultados de los ensayos que evalúan la actividad de CRISPR.

- 30 La Figura 8A-C ilustra manipulaciones ejemplares de un sistema CRISPR para fijar como objetivo sitios genómicos en células de mamíferos.

La Figura 9A-B ilustra los resultados de un análisis por el método Northern de tratamiento de ARNcr en células de mamífero.

La Figura 10A-C ilustra una representación esquemática de ARN quiméricos y los resultados de ensayos SURVEYOR para actividad del sistema CRISPR en células eucariotas.

- 35 La Figura 11A-B ilustra una representación gráfica de los resultados de ensayos SURVEYOR para actividad del sistema CRISPR en células eucariotas.

La Figura 12 ilustra estructuras secundarias previstas para ARN quiméricos ejemplares que comprenden una secuencia guía, secuencia de acoplamiento tracr y secuencia tracr.

La Figura 13 es un árbol filogenético de genes Cas.

- 40 La Figura 14A-F muestra el análisis filogenético que revelan cinco familias de Cas9s, incluyendo tres grupos de Cas9s grandes (~1.400 aminoácidos) y dos de Cas9s pequeños (~1.100 aminoácidos).

La Figura 15 muestra un gráfico que representa la función de diferentes ARN guía optimizados.

La Figura 16 muestra la secuencia y estructura de diferentes ARN quiméricos guía.

La Figura 17 muestra la estructura co-plegada del ARNtracr y repetición dirigida.

- 45 La Figura 18 A y B muestra datos de la optimización de ARN guía quimérico de St1Cas9 *in vitro*.

La Figura 19A-B muestra la escisión de objetivos no metilados o metilados mediante lisado de células de SpCas9.

La Figura 20A-G muestra la optimización de la arquitectura de ARN guía para edición de genomas de mamífero mediada por SpCas9. (a) Esquema de vector de expresión bicistrónico (PX330) para ARN guía único conducido por

activador U6 (ARNsg) y Cas9 de *Streptococcus pyogenes* codón-optimizada humana conducida por activador CBh (*hSpCas9*) usada para todos los experimentos posteriores. El ARNsg consta de una secuencia guía de 20-nt (azul) y andamio (rojo), truncados en diversas posiciones como se indica. (b) Ensayo SURVEYOR para inserciones o supresiones mediadas por SpCas9 en los sitios *EMX1* y *PVALB* humanos. Las flechas indican los fragmentos SURVEYOR esperados ($n = 3$). (c) Análisis por el método Northern para las cuatro arquitecturas de truncado de ARNsg, con U1 como control de carga. (d) Ambos mutantes de tipo natural (wt, por sus siglas en inglés) o nickase (D10A) de inserción activada por SpCas9 de un sitio *HindIII* en el gen *EMX1* humano. Se usaron oligonucleótidos monocatenarios (los ssODN, por sus siglas en inglés) orientados en la dirección sentido o antisentido en relación a la secuencia genómica, como plantillas de recombinación homólogas. (e) Esquema del sitio *SERPINB5* humano. Se indican los ARNsg y las PAM mediante barras coloreadas por encima de las secuencias; se resaltan metilcitosina (Me) (rosa) y se enumeran en relación al sitio de inicio de transcripción (TSS, +1). (f) Estado de metilación de *SERPINB5* ensayado mediante secuenciación de bisulfito de 16 clones. Círculos rellenos, CpG metilado; círculos abiertos, CpG no metilado. (g) Eficacia de modificación por tres ARNsg fijando como objetivo la región metilada de *SERPINB5*, ensayada por secuenciación profunda ($n = 2$). Las barras de error indican intervalos de Wilson (Métodos Online).

La Figura 21A-B muestra la optimización adicional de la arquitectura de ARNsg de CRISPR-Cas. (a) Esquema de cuatro arquitecturas adicionales de ARNsg, I-IV. Cada una consta de una secuencia guía de 20-nt (azul) unida a la repetición directa (RD, gris), que hibrida a la ARNtracr (rojo). El híbrido RD-ARNtracr es truncado en +12 o +22, como se indicaron con un tallo-lazo GAAA artificial. Las posiciones de truncado de ARNtracr se enumeran de acuerdo con el sitio de inicio de la transcripción indicado previamente para ARNtracr. Las arquitecturas de ARNsg II y IV soportan mutaciones dentro de sus tractos poli-U, que podían servir como terminadores de transcripción prematura. (b) Ensayo SURVEYOR para inserciones o supresiones mediadas por SpCas9 en el sitio *EMX1* humano para los sitios 1-3 objetivo. Las flechas indican los fragmentos SURVEYOR expresados ($n = 3$).

La Figura 22 ilustra la visualización de algunos sitios objetivo en el genoma humano.

La Figura 23A-B muestra (A) un esquema del ARNsg y (B) el análisis SURVEYOR para cinco variantes de ARNsg para SaCas9 para una arquitectura truncada óptima con la mayor eficacia de escisión.

Las figuras en la presente memoria son sólo para fines ilustrativos y no están necesariamente dibujadas a escala.

Descripción detallada de la invención

Los términos "polinucleótido", "nucleótido", "secuencia de nucleótidos", "ácido nucleico" y "oligonucleótido" se usan de forma intercambiable. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, desoxirribonucleótidos o ribonucleótidos o análogos de los mismos. Los polinucleótidos pueden tener cualquier estructura tridimensional y pueden realizar cualquier función, conocida o desconocida. Los siguientes no son ejemplos limitantes de los polinucleótidos: regiones codificadoras o no codificadoras de un gen o fragmento de gen, sitios (sitio) definidos a partir de análisis de unión, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ARN de interferencia corta (ARNsi), ARN de horquilla corta (ARNsh), micro-ARN (ARNmi), ribozimas, y ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácidos nucleicos y cebadores. Un polinucleótido puede comprender uno o más nucleótidos modificados, tales como nucleótidos metilados y análogos de nucleótidos. Si hay, se pueden impartir modificaciones a la estructura del nucleótido antes o después del ensamblado del polímero. La secuencia de nucleótidos puede ser interrumpida por componentes no nucleótidos. Un polinucleótido puede ser modificado además después de polimerización, tal como por conjugación con un componente etiquetado.

En aspectos de la invención, los términos "ARN quimérico", "ARN guía quimérico", "ARN guía", "ARN guía único" y "ARN guía sintético" se usan de forma intercambiable y se refieren a la secuencia de polinucleótidos que comprende la secuencia guía, la secuencia tracr y la secuencia de emparejamiento tracr. El término "secuencia guía" se refiere a la secuencia de aproximadamente 20 pb dentro del ARN guía que especifica el sitio objetivo y se puede usar de manera intercambiable con los términos "guía" o "espaciador". El término "secuencia de emparejamiento tracr" también se puede usar de forma intercambiable con el término "repetición o repeticiones directas".

Como se usa en la presente memoria el término "de tipo natural" es un término de la técnica entendido por los expertos y significa la forma típica de un organismo, cepa, gen o característica como se encuentra en la naturaleza como se distingue de formas mutantes o variantes.

Como se usa en la presente memoria el término "variante" se debería considerar que significa la exhibición de cualidades que presentan un patrón que se desvía de lo que ocurre en la naturaleza.

Los términos "que no se encuentra en la naturaleza" o "de ingeniería" se usan de forma intercambiable e indican la implicación de la mano del hombre. Los términos, cuando se refieren a moléculas de ácido nucleico o polipéptidos significan que la molécula de ácido nucleico o el polipéptido está al menos sustancialmente exento de al menos otro componente con el que se asocian de manera natural en la naturaleza y como se encuentran en la naturaleza.

"Complementariedad" se refiere a la capacidad de un ácido nucleico para formar el enlace o enlaces de hidrógeno con otra secuencia de ácidos nucleicos por emparejamiento de bases de Watson-Crick tradicional u otros tipos no tradicionales. Un porcentaje de complementariedad indica el porcentaje de restos en una molécula de ácido nucleico que puede formar enlaces de hidrógeno (por ej., emparejamiento de bases de Watson-Crick) con una segunda secuencia de ácidos nucleicos (por ej., 5, 6, 7, 8, 9, 10 de un total de 10 siendo la complementariedad del 50%, 60%, 70%, 80%, 90% y 100%). "Perfectamente complementario" significa que todos los restos contiguos de una secuencia de ácidos nucleicos formarán un enlace de hidrógeno con el mismo número de restos contiguos en una segunda secuencia de ácidos nucleicos. " Sustancialmente complementario" como se usa en la presente memoria se refiere a un grado de complementariedad que es al menos 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% o 100% por una región de 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50 o más nucleótidos o se refiere a dos ácidos nucleicos que se hibridan en condiciones rigurosas.

Como se usa en la presente memoria, "condiciones rigurosas" para hibridación se refiere a las condiciones en las cuales un ácido nucleico que tiene complementariedad a una secuencia objetivo se hibrida predominantemente con la secuencia objetivo y sustancialmente no se hibrida a las secuencias no objetivo. Las condiciones rigurosas son en general dependientes de la secuencia, y varían dependiendo de una serie de factores. En general, cuanto más larga la secuencia, mayor la temperatura a la que se hibrida de manera específica la secuencia a su secuencia objetivo. Se describen con detalle ejemplos no limitantes de condiciones rigurosas en Tijssen (1.993), *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization With Nucleic Acid Probes Part I*, Segundo Capítulo "Overview of principles of hybridization and the strategy of nucleic acid probe assay", Elsevier, N.Y.

"Hibridación" se refiere a una reacción en que uno o más polinucleótidos reaccionan para formar un complejo que se estabiliza vía enlace de hidrógeno entre las bases de los restos de nucleótido. El enlace de hidrógeno puede tener lugar por apareamiento de bases de Watson Crick, unión de Hoogsteen o de cualquier otra manera específica de la secuencia. El complejo puede comprender dos hebras que forman una estructura dúplex, tres o más hebras que forman un complejo multicatenario, una sola hebra que se autohibrida o cualquier combinación de éstas. Una reacción de hibridación puede constituir una etapa en un procedimiento más extenso, tal como la iniciación de PCR o la escisión de un polinucleótido mediante una enzima. Una secuencia capaz de hibridación con una secuencia determinada se refiere como el "complemento" de la secuencia determinada.

Como se usa en la presente memoria, "estabilización" o "estabilidad aumentada" con respecto a componentes del sistema CRISPR se refieren a asegurar o estabilizar la estructura de la molécula. Esto se puede llevar a cabo por introducción de una o más mutaciones, incluyendo cambios de pares de base únicos o múltiples, aumentando el número de horquillas, unión cruzada, rompiendo tramos particulares de nucleótidos y otras modificaciones. Las modificaciones pueden incluir la inclusión de al menos un nucleótido que no se encuentra en la naturaleza o un nucleótido modificado o análogos de los mismos. Se pueden modificar nucleótidos modificados en el resto ribosa, fosfato y/o base. Los nucleótidos modificados pueden incluir 2'-O-metil-análogos, 2'-desoxi-análogos o 2'-fluoro-análogos. La cadena principal de ácidos nucleicos se puede modificar, por ejemplo, se puede usar una cadena principal de fosforotioato. El uso de ácidos nucleicos bloqueados (LNA, por sus siglas en inglés) o ácidos nucleicos de puente (BNA, por sus siglas en inglés) también puede ser posible. Más ejemplos de bases modificadas incluyen, pero no se limitan a, 2-aminopurina, 5-bromo-uridina, pseudouridina, inosina, 7-metilguanosina. Estas modificaciones se pueden aplicar a cualquier componente del sistema CRSIPR. En una realización preferida estas modificaciones se realizan a los componentes de ARN, por ejemplo, el ARN guía o secuencia de polinucleótidos químicos.

Como se usa en la presente memoria, "expresión" se refiere al procedimiento por el que un polinucleótido se transcribe de una plantilla de ADN (tal como en ARNm u otro transcrito de ARN) y/o el procedimiento por el que se traduce con posterioridad un ARNm transcrito en péptidos, polipéptidos o proteínas. Los transcritos y polipéptidos codificados se pueden referir de manera colectiva como "producto génico". Si el polinucleótido procede de ADN genómico, la expresión puede incluir proceso de corte y empalme del ARNm en una célula eucariota.

Los términos "polipéptido", "péptido" y "proteína" se usan de forma intercambiable en la presente memoria para referirse a polímeros de aminoácidos de cualquier longitud. El polímero puede ser lineal o ramificado y puede comprender aminoácidos modificados y puede ser interrumpido por no aminoácidos. Los términos también incluyen un polímero de aminoácidos que ha sido modificado; por ejemplo, formación de enlace disulfuro, glucosilación, lipidación, acetilación, fosforilación o cualquier otra manipulación, tal como conjugación con un componente etiquetado. Como se usa en la presente memoria el término "aminoácido" incluye aminoácidos naturales y/o no naturales o sintéticos, incluyendo glicina y los dos isómeros ópticos D o L y análogos de aminoácido y peptidomiméticos.

Los términos "individuo," "individual" y "paciente" se usan de forma intercambiable en la presente memoria para referirse a un vertebrado, preferiblemente un mamífero, más preferiblemente un ser humano. Los mamíferos incluyen, pero no están limitados a, murinos, simios, seres humanos, animales de granja, animales deportivos y mascotas. Los tejidos, las células y su progenie de una entidad biológica obtenida in vivo o cultivada in vitro también se incluyen. En algunas realizaciones, un individuo puede ser un animal invertebrado, por ejemplo, un insecto o un nematodo; mientras que en otros, un individuo puede ser una planta o un hongo.

Los términos "agente terapéutico", "agente terapéutico capaz" o "agente de tratamiento" se usan de forma intercambiable y se refieren a una molécula o compuesto que confiere algún efecto beneficioso en la administración a un individuo. El efecto beneficioso incluye permitir determinaciones de diagnóstico; alivio de una enfermedad, síntoma, trastorno o afección patológica; reducir o prevenir el comienzo de una enfermedad, síntoma, trastorno o afección y en general contrarrestar una enfermedad, síntoma, trastorno o afección patológica.

Como se usa en la presente memoria, "tratamiento" o "tratar," o "paliar" o "aliviar" se usan de forma intercambiable. Estos términos se refieren a una propuesta para obtener resultados beneficiosos o deseados incluyendo pero no limitándose a un beneficio terapéutico y/o un beneficio profiláctico. Por beneficio terapéutico se quiere decir cualquier mejora terapéuticamente relevante en, o efecto sobre, una o más enfermedades, afecciones o síntomas bajo tratamiento. Para beneficio profiláctico, las composiciones se pueden administrar a un individuo con riesgo de desarrollar una enfermedad particular, afección o síntoma o a un individuo que indica uno o más de los síntomas fisiológicos de una enfermedad, incluso aunque la enfermedad, afección o síntoma pueda no haberse manifestado aún.

El término "cantidad eficaz" o "cantidad terapéuticamente eficaz" se refiere a la cantidad de un agente que es suficiente para efectuar resultados beneficiosos o deseados. La cantidad terapéuticamente eficaz puede variar dependiendo de uno o más de: el individuo y la enfermedad que se está tratando, el peso y la edad del individuo, la gravedad de la enfermedad, la manera de administración y similares, que se puede determinar fácilmente por un experto en la materia. El término también se aplica a una dosis que proporcionará una imagen para detección por uno cualquiera de los métodos de formación de imagen descritos en la presente memoria. La dosis específica puede variar dependiendo de uno o más de: el agente particular elegido, el régimen de dosificación que se tenga que seguir, si se administra en asociación con otros compuestos, la sincronización de la administración, el tejido que se tiene que someter a formación de imagen y el sistema de suministro físico en el cual esté soportado.

La práctica de la presente invención emplea, a menos que se indique de otro modo, técnicas convencionales de inmunología, bioquímica, química, biología molecular, microbiología, biología celular, ADN genómico y recombinante, que están dentro de la destreza de la técnica. Véase Sambrook, Fritsch y Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2ª edición (1.989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, et al. eds., (1.987)); las series METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M. J. MacPherson, B. D. Hames y G. R. Taylor eds. (1.995)), Harlow and Lane, eds. (1.988) ANTIBODIES, A LABORATORY MANUAL AND ANIMAL CELL CULTURE (R. I. Freshney, ed. (1.987)).

Varios aspectos de la invención se refieren a sistemas vector que comprenden uno o más vectores o vectores como tales. Se pueden diseñar vectores para expresión de transcritos CRISPR (por ejemplo, transcritos de ácidos nucleicos, proteínas o enzimas) en células procariotas o eucariotas. Por ejemplo, se pueden expresar transcritos de CRISPR en células bacterianas tales como *Escherichia coli*, células de insecto (usando vectores de expresión de baculovirus), células de levaduras o células de mamíferos. Las células huésped adecuadas se discuten además en Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego; Calif. (1.990). Como alternativa, el vector de expresión recombinante se puede transcribir y traducir in vitro, por ejemplo usando secuencias reguladoras de activador T7 y polimerasa T7.

Se pueden introducir vectores y propagar en una procariota. En algunas realizaciones, se usó una procariota para multiplicar copias de un vector que se tiene que introducir en una célula eucariota o como un vector intermedio en la producción de un vector que se tiene que introducir en una célula eucariota (por ej., multiplicando un plásmido como parte de un sistema de empaquetamiento de vector vírico). En algunas realizaciones, se usa una procariota para multiplicar copias de un vector y expresar uno o más ácidos nucleicos, tal como para proporcionar una fuente de una o más proteínas para suministro a una célula huésped u organismo huésped. La expresión de proteínas en procariotas se lleva a cabo lo más frecuentemente en *Escherichia coli* con vectores que contienen activadores constitutivos o inducibles que dirigen la expresión de proteínas de fusión o de no fusión. Los vectores de fusión añaden una serie de aminoácidos a una proteína codificada allí, tal como al término amino de la proteína recombinante. Dichos vectores de fusión pueden servir para uno o más fines, tales como: (i) para aumentar la expresión de proteína recombinante; (ii) para aumentar la solubilidad de la proteína recombinante y (iii) para ayudar en la purificación de la proteína recombinante actuando como un ligando de purificación de afinidad. Con frecuencia, en vectores de expresión de fusión, se introduce un sitio de escisión proteolítico en la unión del resto de fusión y la proteína recombinante para permitir la separación de la proteína recombinante del resto de fusión posterior a la purificación de la proteína de fusión. Tales enzimas, y sus secuencias de reconocimiento o de origen similar, incluyen Factor Xa, trombina y enterocinasa. Vectores de expresión de fusión de ejemplo incluyen pGEX (Pharmacia Biotech Inc; Smith and Johnson, 1.988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) y pRIT5 (Pharmacia, Piscataway, N. J.) que condensan glutatona S-transferasa (GST), proteína de unión maltosa E o proteína A, respectivamente, a la proteína recombinante objetivo.

Ejemplos de vectores de expresión *E. coli* no de fusión, inducibles, adecuados, incluyen pTrc (Amrann et al., (1.988) *Gene* 69: 301-315) y pET 11d (Studier et al., GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1.990) 60-89).

En algunas realizaciones, un vector es un vector de expresión de levadura. Ejemplos de vectores para expresión en

levadura *Saccharomyces cerevisiae* incluyen pYepSec1 (Baldari, et al., 1.987. EMBO J. 6: 229-234), pMFa (Kuijan and Herskowitz, 1.982. Cell 30: 933-943), pJRY88 (Schultz et al., 1.987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.) y picZ (Invitrogen Corp, San Diego, Calif.).

En algunas realizaciones, un vector conduce la expresión de proteínas en células de insecto usando vectores de expresión de baculovirus. Los vectores baculovirus disponibles para expresión de proteínas en células de insecto cultivadas (por ej., células SF9) incluyen la serie pAc (Smith, et al., 1.983. Mol. Cell. Biol. 3: 2.156-2.165) y la serie pVL (Lucklow and Summers, 1.989. Virology 170: 31-39).

En algunas realizaciones, un vector es capaz de conducir la expresión de una o más secuencias en células de mamífero usando un vector de expresión de mamífero. Ejemplos de vectores de expresión de mamífero incluyen pCDM8 (Seed, 1.987. Nature 329: 840) y pMT2PC (Kaufman, et al., 1.987. EMBO J. 6: 187-195). Cuando se usa en células de mamífero, las funciones de control del vector de expresión se proporcionan típicamente por uno o más elementos reguladores. Por ejemplo, los activadores usados comúnmente proceden de polioma, adenovirus 2, citomegalovirus, virus 40 de simio y otros descritos en la presente memoria y conocidos en la técnica. Para otros sistemas de expresión adecuados para ambas células procariotas y eucariotas véanse, por ejemplo, los Capítulos 16 y 17 de Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL. 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1.989.

En algunas realizaciones, el vector de expresión de mamíferos recombinante es capaz de dirigir la expresión del ácido nucleico preferentemente en un tipo de célula particular (por ejemplo, se usan elementos reguladores específicos del tejido para expresar el ácido nucleico). Los elementos reguladores específicos del tejido son conocidos en la técnica. Ejemplos no limitantes de activadores específicos del tejido adecuados incluyen el activador de albúmina (específico del hígado; Pinkert, et al., 1.987. Genes Dev. 1: 268-277), activadores específicos linfoides (Calame and Eaton, 1.988. Adv. Immunol. 43: 235-275), en particular activadores de receptores de las células T (Winoto and Baltimore, 1.989. EMBO J. 8: 729-733) e inmunoglobulinas (Baneiji, et al., 1.983. Cell 33: 729-740; Queen and Baltimore, 1.983. Cell 33: 741-748), activadores específicos de las neuronas (por ej., el activador de neurofilamentos; Byrne and Ruddle, 1.989. Proc. Natl. Acad. Sci. USA 86: 5.473-5.477), activadores específicos del páncreas (Edlund, et al., 1.985. Science 230: 912-916) y activadores específicos de las glándulas mamarias (por ej., activador de suero de leche; Patente de EE.UU. N° 4. 873. 316 y Publicación de la Solicitud de Patente Europea N° 264.166). También se incluyen activadores regulados por el desarrollo, por ej., los activadores de hox de murina (Kessel and Gruss, 1.990. Science 249: 374-379) y el activador de α -fetoproteína (Campes and Tilghman, 1.989. Genes Dev. 3: 537-546).

En algunas realizaciones, un elemento regulador se une de manera operable a uno o más elementos de un sistema CRISPR a fin de conducir la expresión de uno o más elementos del sistema CRISPR. En general, las CRISPR (Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas), también conocidas como las SPIDR (Repeticiones Directas Intercaladas Separadas, por sus siglas en inglés), constituyen una familia de sitios de ADN que son normalmente específicos para una especie bacteriana particular. El sitio de CRISPR comprende una clase distinta de repeticiones de secuencias cortas intercaladas (las SSR) que se reconocieron en *E. coli* (Ishino et al., J. Bacteriol., 169: 5.429-5.433 [1.987] y Nakata et al., J. Bacteriol., 171: 3.553-3.556 [1.989]) y genes asociados. Las SSR intercaladas similares han sido identificadas en *Haloferax mediterranei*, *Streptococcus pyogenes*, *Anabaena* y *Mycobacterium tuberculosis* (Véase, Groenen et al., Mol. Microbiol., 10: 1.057-1.065 [1.993]; Hoe et al., Emerg. Infect. Dis., 5: 254-263 [1.999]; Masepohl et al., Biochim. Biophys. Acta 1.307: 26-30 [1.996] y Mojica et al., Mol. Microbiol., 17: 85-93 [1.995]). El sitio de CRISPR difiere típicamente de las otras SSR por la estructura de las repeticiones, que se han denominado repeticiones espaciadas regularmente cortas (las SRSR) (Janssen et al. OMICS J. Integ. Biol., 6: 23-33 [2.002] y Mojica et al., Mol. Microbiol., 36: 244-246 [2.000]). En general, las repeticiones son elementos cortos que tienen lugar en agrupaciones que están regularmente espaciadas por secuencias de intervención únicas con una longitud sustancialmente constante (Mojica et al., [2.000], supra). Aunque las secuencias de repeticiones se conservan mucho entre cepas, el número de repeticiones intercaladas y las secuencias de las regiones espaciadoras difiere típicamente de cepa a cepa (van Embden et al., J. Bacteriol., 182: 2.393-2.401 [2.000]). Se ha identificado el sitio de CRISPR en más de 40 procariotas (Véase por ej., Jansen et al., Mol. Microbiol., 43: 1.565-1.575 [2.002] y Mojica et al., [2.005]) incluyendo, pero no limitado a *Aeropyrum*, *Pyrobaculum*, *Sulfolobus*, *Archaeoglobus*, *Halocarcularia*, *Methanobacterium*, *Methanococcus*, *Methanosarcina*, *Methanopyrus*, *Pyrococcus*, *Picrophilus*, *Thermoplasma*, *Corynebacterium*, *Mycobacterium*, *Streptomyces*, *Aquifex*, *Porphyromonas*, *Chlorobium*, *Thermus*, *Bacillus*, *Listeria*, *Staphylococcus*, *Clostridium*, *Thermoanaerobacter*, *Mycoplasma*, *Fusobacterium*, *Azarcus*, *Chromobacterium*, *Neisseria*, *Nitrosomonas*, *Desulfovibrio*, *Geobacter*, *Myxococcus*, *Campylobacter*, *Wolinella*, *Acinetobacter*, *Erwinia*, *Escherichia*, *Legionella*, *Methylococcus*, *Pasturella*, *Photobacterium*, *Salmonella*, *Xanthomonas*, *Yersinia*, *Treponema* y *Thermotoga*.

En general, "sistema CRISPR" se refiere colectivamente a transcritos y otros elementos implicados en la expresión de o dirigidos a la actividad de genes asociados a CRISPR ("Cas"), incluyendo secuencias que codifican un gen Cas, una secuencia tracr (CRISPR trans-activante) (por ej., ARNtracr o un ARNtracr parcial activo), una secuencia de emparejamiento tracr (incluyendo una "repeticón dirigida" y una repeticón directa parcial tratada por ARNtracr en el contexto de un sistema CRISPR endógeno), una secuencia guía (también referida como un "espaciador" en el contexto de un sistema CRISPR endógeno) u otras secuencias y transcritos de un sitio CRISPR. En algunas realizaciones, uno o más elementos de un sistema CRISPR procede de un sistema CRISPR tipo I, tipo II o tipo III.

En algunas realizaciones, uno o más elementos de un sistema CRISPR procede de un organismo particular que comprende un sistema CRISPR endógeno, tal como *Streptococcus pyogenes*. En general, un sistema CRISPR se caracteriza por elementos que activan la formación de un complejo CRISPR en el sitio de una secuencia objetivo (también referido como un protoespaciador en el contexto de un sistema CRISPR endógeno). En el contexto de formación de un complejo CRISPR, "secuencia objetivo" se refiere a una secuencia para la que se diseña que una secuencia guía presente complementariedad, donde la hibridación entre una secuencia objetivo y una secuencia guía active la formación de un complejo CRISPR. No se requiere necesariamente complementariedad completa, siempre que haya suficiente complementariedad para producir hibridación y activar la formación de un complejo CRISPR. Una secuencia objetivo puede comprender cualquier polinucleótido, tal como polinucleótido de ADN o ARN. En algunas realizaciones, se sitúa una secuencia objetivo en el núcleo o citoplasma de una célula. En algunas realizaciones, la secuencia objetivo puede estar dentro de un organelo de una célula eucariota, por ejemplo, mitocondria o cloroplasto. Una secuencia o plantilla que se puede usar para recombinación en el sitio fijado como objetivo que comprende las secuencias objetivo se refiere como "plantilla de edición" o "polinucleótido de edición" o "secuencia de edición". En aspectos de la invención, un polinucleótido de plantilla exógeno se puede referir como una plantilla de edición. En un aspecto de la invención la recombinación es recombinación homóloga.

Típicamente, en el contexto de un sistema CRISPR endógeno, la formación de un complejo CRISPR (que comprende una secuencia guía hibridada a una secuencia objetivo y complejada con una o más proteínas Cas) da como resultado la escisión de una o más hebras en o cerca de (por ej., dentro de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 o más pares de bases de) la secuencia objetivo. Sin desear estar limitados por la teoría, la secuencia tracr, que puede comprender o constar de toda o una porción de una secuencia tracr de tipo natural (por ej., aproximadamente o más de aproximadamente 20, 26, 32, 45, 48, 54, 63, 67, 85 o más nucleótidos de una secuencia tracr de tipo natural), también puede formar parte de un complejo CRISPR, tal como por hibridación a lo largo de al menos una porción de la secuencia tracr a toda o una porción de una secuencia de emparejamiento tracr que esté ligada de manera operable a la secuencia guía. En algunas realizaciones, la secuencia tracr presenta suficiente complementariedad a una secuencia de emparejamiento tracr para hibridar y participar en la formación de un complejo CRISPR. Como con la secuencia objetivo, se cree que no es necesaria la complementariedad completa, siempre que haya suficiente para que sea funcional. En algunas realizaciones, la secuencia tracr presenta al menos 50%, 60%, 70%, 80%, 90%, 95% o 99% de complementariedad de secuencias a lo largo de la longitud de la secuencia de emparejamiento tracr cuando se alinea de manera óptima. En algunas realizaciones, uno o más vectores que conducen la expresión de uno o más elementos de un sistema CRISPR se introducen en una célula huésped de manera que la expresión de los elementos del sistema CRISPR dirija la formación de un complejo CRISPR en uno o más sitios objetivo. Por ejemplo, una enzima Cas, una secuencia guía ligada a una secuencia de emparejamiento tracr y una secuencia tracr podían estar unidas cada una de manera operable a elementos reguladores separados en vectores separados. Como alternativa, dos o más de los elementos expresados de los mismos o diferentes elementos reguladores, se pueden combinar en un vector único, proporcionando uno o más vectores adicionales cualquier componente del sistema CRISPR no incluido en el primer vector. Los elementos del sistema CRISPR que se combinan en un único vector se pueden disponer en cualquier orientación adecuada, tal como un elemento situado en 5' con respecto a ("aguas arriba" de) o 3' con respecto a ("aguas abajo" de) un segundo elemento. La secuencia codificadora de un elemento puede estar situada en la misma hebra u opuesta de la secuencia codificadora de un segundo elemento y orientada en la misma dirección u opuesta. En algunas realizaciones, un activador único conduce la expresión de una transcripción que codifica una enzima CRISPR y una o más de la secuencia guía, la secuencia de emparejamiento tracr (opcionalmente unida de manera operable a la secuencia guía) y una secuencia tracr embebida dentro de una o más secuencias intrón (por ej., cada una en un intrón diferente, dos o más en al menos un intrón o todas en un solo intrón). En algunas realizaciones, la enzima CRISPR, secuencia guía, secuencia de emparejamiento tracr y secuencia tracr se unen de manera operable a y se expresan a partir del mismo activador.

En algunas realizaciones, un vector comprende uno o más sitios de inserción, tal como una secuencia de reconocimiento de endonucleasa de restricción (también referido como un "sitio de clonación"). En algunas realizaciones, uno o más sitios de inserción (por ej., aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más sitios de inserción) están situados aguas arriba y/o aguas abajo de uno o más elementos de secuencia de uno o más vectores. En algunas realizaciones, un vector comprende un sitio de inserción aguas arriba de una secuencia de emparejamiento tracr y opcionalmente aguas abajo de un elemento regulador unido de manera operable a la secuencia de emparejamiento tracr, de manera que después de la inserción de una secuencia guía en el sitio de inserción y en la expresión de la secuencia guía dirija la unión específica de la secuencia de un complejo CRISPR a una secuencia objetivo en una célula eucariota. En algunas realizaciones, un vector comprende dos o más sitios de inserción, estando situado cada sitio de inserción entre dos secuencias de emparejamiento tracr de manera que se permita la inserción de una secuencia guía en cada sitio. En dicha disposición, dos o más secuencias guía pueden comprender dos o más copias de una secuencia guía única, dos o más secuencias guía diferentes o combinaciones de éstas. Cuando se usan múltiples secuencias guía diferentes, se puede usar una construcción de expresión única para fijar como objetivo la actividad de CRISPR a múltiples secuencias objetivo correspondientes, diferentes, dentro de una célula. Por ejemplo, un solo vector puede comprender aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 o más secuencias guía. En algunas realizaciones, se pueden proporcionar aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más de tales vectores que contienen secuencia guía y se suministran opcionalmente a una célula.

En algunas realizaciones, un vector comprende un elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica una enzima CRISPR, tal como una proteína Cas. Ejemplos no limitantes de proteínas Cas incluyen Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocidas como Csn1 y Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homólogos de las mismas o versiones modificadas de las mismas. Estas enzimas son conocidas; por ejemplo, la secuencia de aminoácidos de proteína Cas9 de *S. pyogenes* se puede encontrar en la base de datos SwissProt con el número de acceso Q99ZW2. En algunas realizaciones, la enzima CRISPR no modificada presenta actividad de escisión de ADN, tal como Cas9. En algunas realizaciones la enzima CRISPR es Cas9 y puede ser Cas9 de *S. pyogenes* o *S. pneumoniae*. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o ambas hebras en la posición de una secuencia objetivo, tal como dentro de la secuencia objetivo y/o dentro del complemento de la secuencia objetivo. En algunas realizaciones, la enzima CRISPR dirige la escisión de una o ambas hebras dentro de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500 o más pares de bases del primer o el último nucleótido de una secuencia objetivo. En algunas realizaciones, un vector codifica una enzima CRISPR que muta con respecto a una enzima de tipo natural correspondiente de manera que la enzima CRISPR mutada carece de la capacidad para escindir una o ambas hebras de un polinucleótido objetivo que contiene una secuencia objetivo. Por ejemplo, una sustitución de aspartato a alanina (D10A) en el dominio catalítico RuvC I de Cas9 de *S. pyogenes* convierte Cas9 de una nucleasa que escinde ambas hebras a una nickasa (escinde una única hebra). Otros ejemplos de mutaciones que hacen Cas9 una nickasa incluyen, sin limitación, H840A, N854A y N863A. En algunas realizaciones, se puede usar una nickasa de Cas9 junto con secuencia o secuencias guía, por ejemplo, dos secuencias guía, que fijan como objetivo hebras transcrita y complementaria respectivamente del objetivo de ADN. Esta combinación permite que ambas hebras sean nickeadas y usadas para inducir NHEJ. Los solicitantes han demostrado (datos no mostrados) la eficacia de dos objetivos de nickasa (es decir, los ARNsg fijados como objetivo en la misma posición pero para diferentes hebras de ADN) en la inducción de NHEJ mutagénico. Una única nickasa (Cas9-D10A con un único ARNsg) es incapaz de inducir NHEJ y crear inserciones o supresiones pero los Solicitantes han demostrado en la nickasa doble (Cas9-D10A y dos ARNsg fijados como objetivo a diferentes hebras en la misma posición) lo puede hacer en células madre embrionarias humanas (las hESC). La eficacia es aproximadamente 50% de nucleasa (es decir, Cas9 regular sin mutación D10) en las hESC.

Como un ejemplo más, dos o más dominios catalíticos de Cas9 (RuvC I, RuvC II y RuvC III) pueden mutar para producir un Cas9 mutado que carece sustancialmente de toda actividad de escisión de ADN. En algunas realizaciones, una mutación de D10A se combina con una o más de las mutaciones H840A, N854A o N863A para producir una enzima Cas9 que carece sustancialmente de toda actividad de escisión de ADN. En algunas realizaciones, se considera que una enzima CRISPR carece sustancialmente de toda actividad de escisión de ADN cuando la actividad de escisión de ADN de la enzima mutada es menor que aproximadamente 25%, 10%, 5%, 1%, 0,1%, 0,01% o menor con respecto a su forma no mutada. Pueden ser útiles otras mutaciones; en el caso de que la Cas9 u otra enzima CRISPR sea de una especie distinta de *S. pyogenes*, se pueden realizar mutaciones en los correspondientes aminoácidos para conseguir efectos similares.

En algunas realizaciones, una secuencia codificadora de enzimas que codifica una enzima CRISPR es codón optimizada para expresión en células particulares, tales como células eucariotas. Las células eucariotas pueden ser aquellas de, o procedentes de, un organismo particular, tal como un mamífero, incluyendo pero no limitado a, ser humano, ratón, rata, conejo, perro o primate no ser humano. En general, optimización de codón se refiere a un procedimiento para modificar una secuencia de ácidos nucleicos para mejorar la expresión en las células huésped de interés reemplazando al menos un codón (por ej., aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 o más codones) de la secuencia natural con codones que se usan más frecuentemente o lo más frecuentemente en los genes de esa célula huésped al tiempo que se mantiene la secuencia de aminoácidos natural. Diversas especies presentan sesgo particular para ciertos codones de un aminoácido particular. El sesgo de codones (diferencias en el uso de codones entre organismos) con frecuencia se correlaciona con la eficacia de traducción de ARN mensajero (ARNm), que se cree a su vez que depende de, entre otras cosas, las propiedades de los codones que se están traduciendo y la disponibilidad de moléculas de ARN de transferencia (ARNt) particulares. La predominancia de los ARNt seleccionados en una célula es en general una reflexión de los codones usados lo más frecuentemente en síntesis de péptidos. De acuerdo con esto, se pueden adaptar los genes para expresión óptima de los genes en un organismo determinado basándose en optimización de codones. Las tablas de utilización de codones están fácilmente disponibles, por ejemplo, en la "Base de Datos de Uso de Codones", y estas tablas se pueden adaptar de una serie de maneras. Véase Nakamura Y., et al. "Codon usage tabulated from the international DNA sequence databases: status for the year 2000" Nucl. Acids Res. 28: 292 (2.000). También están disponibles algoritmos de ordenador para codones optimizando una secuencia particular para expresión en una célula huésped particular, tales como Gene Forge (Aptagen; Jacobus, PA). En algunas realizaciones, uno o más codones (por ej., 1, 2, 3, 4, 5, 10, 15, 20, 25, 50 o más o todos los codones) en una secuencia codificando una enzima CRISPR corresponden al codón usado lo más frecuentemente para un aminoácido particular.

En algunas realizaciones, un vector codifica una enzima CRISPR que comprende una o más secuencias de localización nuclear (las NLS, por sus siglas en inglés), tales como aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más NLS. En algunas realizaciones, la enzima CRISPR comprende aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más NLS en o cerca del amino terminal, aproximadamente o

más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más NLS en o cerca del carboxi terminal o una combinación de estos (por ejemplo, una o más NLS en el amino terminal y una o más NLS en el carboxi terminal). Cuando está presente más de una NLS, cada una se puede seleccionar independientemente de las otras, de manera que puede estar presente una sola NLS en más de una copia y/o junto con otra u otras NLS más presentes en una o más copias. En una realización preferida de la invención, la enzima CRISPR comprende a lo sumo 6 NLS. En algunas realizaciones, se considera una NLS cerca del N- o C-terminal cuando el aminoácido más cercano de la NLS está dentro de aproximadamente 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50 o más aminoácidos a lo largo de la cadena polipeptídica del N o C-terminal. Típicamente, una NLS consta de una o más secuencias cortas de lisinas o argininas cargadas positivamente expuestas en la superficie de la proteína, pero se conocen otros tipos de NLS. Ejemplos no limitantes de las NLS incluyen una secuencia de NLS procedente de: la NLS del antígeno T grande del virus SV40, que tiene la secuencia de aminoácidos PKKKRKV; la NLS de nucleoplasmina (por ej., la NLS bipartita de nucleoplasmina con la secuencia KRPAATKKAGQAKKKK); la NLS de c-myc que tiene la secuencia de aminoácidos PAAKRVKLD o RQRRNELKRSP; la NLS M9 del hRNP1 que tiene la secuencia NQSSNFGPMKGGNFGGRSSGPGYGGGGYFAKPRNQGGY; la secuencia RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV del dominio IBB de importina-alfa; las secuencias VSRKRPRP y PPKKARED de la proteína T de mioma; la secuencia POPKKKPL de p53 humano; la secuencia SALIKKKKKMAP de c-abl IV de ratón; las secuencias DRLRR y PKQKKRK del virus de la influenza NS1; la secuencia RKLKKIKKL del antígeno delta del virus de la Hepatitis; la secuencia REKKKFLKRR de la proteína Mx1 de ratón; la secuencia KRKGDEVGDGVDEVAKKKSKK de la poli(ADP-ribosa) polimerasa humana y la secuencia RKCLQAGMNLARKTKK del glucocorticoide de receptores de hormonas esteroideas (humano).

En general, una o más NLS son de suficiente fortaleza para conducir la acumulación de la enzima CRISPR en una cantidad detectable en el núcleo de una célula eucariota. En general, la fortaleza de actividad de localización nuclear puede proceder del número de las NLS en la enzima CRISPR, la o las NLS particulares usadas o una combinación de estos factores. La detección de acumulación en el núcleo se puede realizar por cualquier técnica adecuada. Por ejemplo, se puede condensar un marcador detectable a la enzima CRISPR, de manera que se pueda visualizar la localización dentro de una célula, tal como junto con un medio para detectar la localización del núcleo (por ejemplo, una mancha específica para el núcleo tal como DAPI). Ejemplos de marcadores detectables incluyen proteínas fluorescentes (tales como proteínas fluorescentes Verdes o GFP; RFP; CFP) y etiquetas de epítomos (etiqueta HA, etiqueta bandera, etiqueta SNAP). También se pueden aislar los núcleos celulares de las células, los contenidos de los cuales se pueden analizar después por cualquier procedimiento adecuado para detectar proteína, tal como inmunohistoquímica, ensayo de Western o ensayo de actividad enzimática. La acumulación en el núcleo también se puede determinar de una manera indirecta, tal como por un ensayo para el efecto de la formación de complejo CRISPR (por ejemplo, ensayo para escisión de ADN o mutación en la secuencia objetivo o ensayo para actividad de expresión de genes modificada afectada por la formación de complejo CRISPR y/o actividad de la enzima CRISPR), cuando se compara con un control no expuesto a la enzima o complejo CRISPR o expuesto a una enzima CRISPR que carece de una o más NLS.

En general, una secuencia guía es cualquier secuencia de polinucleótidos que tiene suficiente complementariedad con una secuencia de polinucleótidos objetivo para hibridación con la secuencia objetivo y dirigir la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo. En algunas realizaciones, el grado de complementariedad entre una secuencia guía y su correspondiente secuencia objetivo, cuando se alinean de manera óptima usando un algoritmo de alineación adecuado, es aproximadamente o más de aproximadamente 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97,5%, 99% o más. La alineación óptima se puede determinar con el uso de cualquier algoritmo adecuado para alinear secuencias, ejemplo no limitante de lo cual incluye el algoritmo Smith-Waterman, el algoritmo Needleman-Wunsch, los algoritmos basados en la transformación de Burrows-Wheeler (por ej., el Alineador de Burrows Wheeler), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, CA), SOAP (disponible en soap.genomics.org.cn) y Maq (disponible en maq.sourceforge.net). En algunas realizaciones, una secuencia guía es aproximadamente o más de aproximadamente 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75 o más nucleótidos de longitud. En algunas realizaciones, una secuencia guía es menor que aproximadamente 75, 50, 45, 40, 35, 30, 25, 20, 15, 12 o menos nucleótidos de longitud. La capacidad de una secuencia guía para dirigir la unión específica de la secuencia de un complejo CRISPR a una secuencia objetivo se puede evaluar por cualquier ensayo adecuado. Por ejemplo, los componentes de un sistema CRISPR suficientes para formar un complejo CRISPR, incluyendo la secuencia guía que se tiene que ensayar, se puede proporcionar a una célula huésped que tenga la correspondiente secuencia objetivo, tal como por transfección con vectores que codifiquen los componentes de la secuencia CRISPR, seguido por una evaluación de la escisión preferente dentro de la secuencia objetivo, tal como mediante ensayo Surveyor como se describe en la presente memoria. De manera similar, la escisión de una secuencia de polinucleótido objetivo se puede evaluar en un tubo de ensayo proporcionando la secuencia objetivo, los componentes de un complejo CRISPR, incluyendo la secuencia guía que se tiene que ensayar y una secuencia guía de control diferente de la secuencia guía de ensayo y comparando la unión o velocidad de escisión en la secuencia objetivo entre las reacciones de la secuencia guía de ensayo y de control. Son posibles otros ensayos, y tendrán lugar para los expertos en la materia.

Se puede seleccionar una secuencia guía para fijar como objetivo cualquier secuencia objetivo. En algunas realizaciones, la secuencia objetivo es una secuencia dentro de un genoma de una célula. Las secuencias objetivo

ejemplos incluyen aquellas que son únicas en el genoma objetivo. Por ejemplo, para la Cas9 de *S. pyogenes*, una secuencia objetivo única en un genoma puede incluir un sitio objetivo de Cas9 de la forma MMMMMMMMMNNNNNNNNNNNNXGG donde NNNNNNNNNNNNNXGG (N es A, G, T o C y X puede ser cualquiera) presenta un solo caso en el genoma. Una única secuencia objetivo en un genoma puede incluir un sitio objetivo de Cas9 de *S. pyogenes* de la forma MMMMMMMMMNNNNNNNNNNNNXGG donde NNNNNNNNNNNXGG (N es A, G, T o C y X puede ser cualquiera) presenta un solo caso en el genoma. Para la Cas9 de CRISPR1 de *S. thermophilus*, una única secuencia objetivo en un genoma puede incluir un sitio objetivo de Cas9 de la forma MMMMMMMMMNNNNNNNNNNNNXXAGAA W donde NNNNNNNNNNNXXAGAAW (N es A, G, T o C; X puede ser cualquiera y W es A o T) presenta un solo caso en el genoma. Una única secuencia objetivo en un genoma puede incluir un sitio objetivo de Cas9 de CRISPR1 de *S. thermophilus* de la forma MMMMMMMMMNNNNNNNNNNNNXXAGAAW donde NNNNNNNNNNNXXAGAAW (N es A, G, T o C; X puede ser cualquiera y W es A o T) presenta un solo caso en el genoma. Para la Cas9 de *S. pyogenes*, una sola secuencia objetivo en un genoma puede incluir un sitio objetivo de Cas9 de la forma MMMMMMMMMNNNNNNNNNNNNXGGXG donde NNNNNNNNNNNXGGXG (N es A, G, T o C y X puede ser cualquiera) presenta un solo caso en el genoma. Una única secuencia objetivo en un genoma puede incluir un sitio objetivo de Cas9 de *S. pyogenes* de la forma MMMMMMMMMNNNNNNNNNNNNXGGXG donde NNNNNNNNNNNXGGXG (N es A, G, T o C y X puede ser cualquiera) presenta un único caso en el genoma. En cada una de estas secuencias "M" puede ser A, G, T o C y requiere que no se considere en la identificación de una secuencia como única.

En algunas realizaciones, una secuencia guía se selecciona para reducir el grado de estructura secundaria dentro de la secuencia guía. La estructura secundaria se puede determinar por cualquier algoritmo de plegamiento de polinucleótidos adecuado. Algunos programas se basan en calcular la energía libre mínima de Gibbs. Un ejemplo de uno de tales algoritmos es mFold, como se describe por Zuker y Stiegler (Nucleic Acid Res. 9 (1.981), 133-148). Otro algoritmo de plegamiento de ejemplo es el RNAfold webserver online, desarrollado en el Instituto para Química Teórica de la Universidad de Viena, usando el algoritmo de predicción de estructura centroide (véase por ej., A. R. Gruber et al., 2.008, Cell 106 (1): 23-24 y PA Carr y GM Church, 2.009, Nature Biotechnology 27 (12): 1.151-62).

En general, una secuencia de emparejamiento tracr incluye cualquier secuencia que presente suficiente complementariedad con una secuencia tracr para activar uno o más de: (1) escisión de una secuencia guía flanqueada por secuencias de emparejamiento tracr en una célula que contiene la correspondiente secuencia tracr y (2) formación de un complejo CRISPR en una secuencia objetivo, en la que el complejo CRISPR comprende la secuencia de emparejamiento tracr hibridada a la secuencia tracr. En general, el grado de complementariedad es con referencia al alineamiento óptimo de la secuencia de emparejamiento tracr y la secuencia tracr, junto con la longitud del acortador de las dos secuencias. El alineamiento óptimo se puede determinar por cualquier algoritmo de alineamiento adecuado y puede justificar además estructuras secundarias, tal como autocomplementariedad dentro de cualquiera la secuencia tracr o la secuencia de emparejamiento tracr. En algunas realizaciones, el grado de complementariedad entre la secuencia tracr y la secuencia de emparejamiento tracr junto con la longitud del acortador de las dos cuando se alinean de manera óptima es aproximadamente o más de aproximadamente 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97,5%, 99% o mayor. Las ilustraciones de ejemplo de alineación óptima entre una secuencia tracr y una secuencia de emparejamiento tracr se proporcionan en las Figuras 12B y 13B. En algunas realizaciones, la secuencia tracr es aproximadamente o más de aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50 o más nucleótidos de longitud. En algunas realizaciones, la secuencia tracr y secuencia de emparejamiento tracr están contenidas dentro de una transcripción única, de manera que la hibridación entre las dos produce un transcripto con una estructura secundaria, tal como una horquilla. Las secuencias formadoras de lazo preferidas para uso en estructuras de horquilla tienen cuatro nucleótidos de longitud y lo más preferiblemente tienen la secuencia GAAA. Sin embargo, se pueden usar secuencias de lazo más largas o más cortas, como se pueden usar secuencias alternativas. Las secuencias incluyen preferiblemente un triplete nucleótido (por ejemplo, AAA) y un nucleótido adicional (por ejemplo C o G). Ejemplos de secuencias formadoras de lazo incluyen CAAA y AAAG. En una realización de la descripción, el transcripto o secuencia de polinucleótidos transcrita presenta al menos dos o más horquillas. En realizaciones preferidas, el transcripto presenta dos, tres, cuatro o cinco horquillas. En una realización más de la descripción, el transcripto presenta a lo sumo cinco horquillas. En algunas realizaciones, el transcripto único incluye además una secuencia de terminación de la transcripción; preferiblemente esta es una secuencia poliT, por ejemplo seis nucleótidos T. Una ilustración de ejemplo de dicha estructura de horquilla se proporciona en la porción inferior de la Figura 13B, donde la porción de la secuencia 5' del "N" final y aguas arriba del lazo corresponde a la secuencia de emparejamiento tracr y la porción de la secuencia 3' del lazo corresponde a la secuencia tracr. Ejemplos no limitantes adicionales de polinucleótidos únicos que comprenden una secuencia guía, una secuencia de emparejamiento tracr y una secuencia tracr son como sigue (enumerados 5' a 3'), donde "N" representa una base de una secuencia guía, el primer bloque de letras minúsculas representan la secuencia de emparejamiento tracr y el segundo bloque de letras minúsculas secuencia representa la secuencia tracr y la secuencia poli-T final representa el terminador de la transcripción: (1) NNNNNNNNNNNNNNNNNNNNNNgttttgtactctcaagatttaGAAAtaaactctgcagaagctacaaagataaggcttcatgccgaatacaacacctgtcatttatggcagggtgttttcgtatttaaTTTTTT; (2) NNNNNNNNNNNNNNNNNNNNNNgttttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaatacaacacctgtcatttatggcagggtgttttcgtatttaaTTTTTT; (3) NNNNNNNNNNNNNNNNNNNNNNgttttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaatacaacacctgtcatttatggcagggtgttttcgtatttaaTTTTTT; (4)

NNNNNNNNNNNNNNNNNNNNNgttttagagctaGAAAtagcaagttaaaataaggctaqtccgttatcaacttgaaaa

agtggcaccgagtcggtgcTTTTT;

NNNNNNNNNNNNNNNNNNNNNNNNNNNNgtttagagctaGAAATAGcaagttaaataaaggctagtccgttatcaactgaa aaagtgTTTTTTT; y (6)

NNNNNNNNNNNNNNNNNNNNgttttagagctaqAAATAGcaagttaaataaggcctaqtccgttatcaTTTT TTT. En algunas

realizaciones, las secuencias (1) a (3) se usan junto con Cas9 de *S. thermophilus*. En algunas realizaciones, las secuencias (4) a (6) se usan junto con Cas9 de *S. pyogenes*. En algunas realizaciones, la secuencia tracr es un transcrito separado de una transcripción que comprende la secuencia de emparejamiento tracr (tal como se ilustra en la porción superior de la Figura 13B).

En algunas realizaciones, también se proporciona una plantilla de recombinación. Una plantilla de recombinación puede ser un componente de otro vector como se describe en la presente memoria, contenido en un vector separado o proporcionado como un polinucleótido separado. En algunas realizaciones, se diseña una plantilla de recombinación para servir como una plantilla en recombinación homóloga, tal como dentro de o cerca de una secuencia objetivo nickeada o escindida por una enzima CRISPR como parte de un complejo CRISPR. Un polinucleótido de plantilla puede ser de cualquier longitud adecuada, tal como aproximadamente o más de aproximadamente 10, 15, 20, 25, 50, 75, 100, 150, 200, 500, 1.000 o más nucleótidos de longitud. En algunas realizaciones, el polinucleótido de plantilla es complementario a una porción de un polinucleótido que comprende la secuencia objetivo. Cuando se alinea de manera óptima, un polinucleótido de plantilla puede solaparse con uno o más nucleótidos de una secuencia objetivo (por ej., aproximadamente o más de aproximadamente 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 o más nucleótidos). En algunas realizaciones, cuando una secuencia de plantilla y un polinucleótido que comprende una secuencia objetivo se alinean de manera óptima, el nucleótido más cercano del polinucleótido de plantilla está dentro de aproximadamente 1, 5, 10, 15, 20, 25, 50, 75, 100, 200, 300, 400, 500, 1.000, 5.000, 10.000 o más nucleótidos de la secuencia objetivo.

En algunas realizaciones, la enzima CRISPR es parte de una proteína de fusión que comprende uno o más dominios de proteína heteróloga (por ej., aproximadamente o más de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más dominios además de la enzima CRISPR). Una proteína de fusión de enzima CRISPR puede comprender cualquier secuencia de proteínas adicional y opcionalmente una secuencia ligadora entre dos dominios cualesquiera. Ejemplos de dominios de proteínas que se pueden condensar a una enzima CRISPR incluyen, sin limitación, etiquetas de epítomos, secuencias de genes informadores y dominios de proteínas con una o más de las siguientes actividades: actividad de metilasa, actividad de desmetilasa, actividad de activación de la transcripción, actividad de represión de la transcripción, actividad del factor de liberación de la transcripción, actividad de modificación de histona, actividad de escisión de ARN y actividad de unión de ácidos nucleicos. Ejemplos no limitantes de etiquetas de epítomos incluyen etiquetas de histidina (His), etiquetas V5, etiquetas FLAG, etiquetas de hemaglutinina de la influenza (HA), etiquetas Myc, etiquetas VSV-G y etiquetas de tioredoxina (Trx). Los ejemplos de genes informadores incluyen, pero no se limitan a, glutatión-S-transferasa (GST), peroxidasa de rábano (HRP, por sus siglas en inglés), cloramfenicol acetiltransferasa (CAT, por sus siglas en inglés) beta-galactosidasa, beta-glucuronidasa, luciferasa, proteína fluorescente verde (PFV), HcRed, DsRed, proteína fluorescente cian (pFC), proteína fluorescente amarilla (PFAM) y proteínas autofluorescentes incluyendo proteína fluorescente azul (PFA). Una enzima CRISPR se puede condensar a una secuencia de genes que codifique una proteína o un fragmento de una proteína que una moléculas de ADN u otras moléculas celulares, incluyendo pero no limitándose a (por sus siglas en inglés), proteína de unión de maltosa (MBP), etiqueta S, fusiones de dominios de unión de ADN Lex A (DBD), fusiones de dominios de unión de ADN GAL4 y fusiones de proteínas BP16 de virus del herpes simple (HSV). Los dominios adicionales que pueden formar parte de una proteína de fusión que comprende una enzima CRISPR se describen en la patente de EE.UU. 20110059502 incorporada a la presente por referencia. En algunas realizaciones, se usa una enzima CRISPR etiquetada para identificar la posición de una secuencia objetivo.

En algunos aspectos, la descripción proporciona métodos que comprenden suministrar uno o más polinucleótidos, tales como o uno o más vectores como se describe en la presente memoria, uno o más transcritos de los mismos y/o una o proteínas transcritas de allí, a una célula huésped. En algunos aspectos, la descripción proporciona además células producidas por tales métodos y organismos (tales como animales, plantas u hongos) que comprenden o son producidos de tales células. En algunas realizaciones, se suministra una enzima CRISPR junto con (y opcionalmente complejada con) una secuencia guía a una célula. Se pueden usar métodos de transferencia de genes con base vírica y no vírica convencionales para introducir ácidos nucleicos en células de mamífero o tejidos objetivo. Se pueden usar dichos métodos para administrar componentes que codifican ácidos nucleicos de un sistema CRISPR a células en cultivo o en un organismo huésped. Los sistemas de suministro de vectores no víricos incluyen plásmidos de ADN, ARN (por ej., un transcrito de un vector descrito en la presente memoria), ácido nucleico desnudo y ácido nucleico complejado con un vehículo de suministro, tal como un liposoma. Los sistemas de suministro de vectores víricos incluyen virus de ADN y ARN, que tienen genomas episomales o integrados después de suministro a la célula. Para una revisión de procedimientos de tratamiento de genes, véase Anderson, *Science* 256: 808-813 (1.992); Nabel & Felgner, *TIBTECH* 11: 211-217 (1.993); Mitani & Caskey, *TIBTECH* 11: 162-166 (1.993); Dillon, *TIBTECH* 11: 167-175 (1.993); Miller, *Nature* 357: 455-460 (1.992); Van Brunt, *Biotechnology* 6 (10): 1.149-1.154 (1.988); Vigne, *Restorative Neurology and Neuroscience* 8: 35-36 (1.995); Kremer & Perricaudet, *British Medical Bulletin* 51 (1): 31-44 (1.995); Haddada et al., en *Current Topics in Microbiology and Immunology* Doerfler and Böhm (eds) (1.995) y Yu et al., *Gene Therapy* 1: 13-26 (1.994).

Los métodos de suministro no vírico de ácidos nucleicos incluyen lipofección, nucleofección, microinyección, biolística, virosomas, liposomas, inmunoliposomas, conjugados de polimerización o lípido:ácido nucleico, ADN desnudo, viriones artificiales y absorción de ADN aumentada por agente. La lipofección se describe en por ej., Las Patentes de EE.UU. N° 5.049.386, 4.946.787 y 4.897.355) y se venden comercialmente los reactivos de lipofección (por ej., Transfectam™ y Lipofectin™). Los lípidos catiónicos y neutros que son adecuados para lipofección de reconocimiento de receptores eficaz de polinucleótidos incluyen aquéllos de Felgner, Patente Internacional WO 91/17424; Patente Internacional WO 91/16024. El suministro puede ser a células (por ej., administración in vitro o ex vivo) o tejidos objetivo (por ej., administración in vivo).

La preparación de complejos de lípido:ácido nucleico, incluyendo liposomas fijados como objetivo tales como complejos de inmunolípidos, es conocida para un experto en la materia (véase, por ej., Crystal, Science 270: 404-410 (1.995); Blaese et al., Cancer Gene Ther. 2: 291-297 (1.995); Behr et al., Bioconjugate Chem. 5: 382-389 (1.994); Remy et al., Bioconjugate Chem. 5: 647-654 (1.994); Gao et al., Gene Therapy 2: 710-722 (1.995); Ahmad et al., Cancer Res. 52: 4.817-4.820 (1.992); Patentes de EE.UU. N° 4.186.183; 4.217.344; 4.235.871; 4.261.975; 4.485.054; 4.501.728; 4.774.085; 4.837.028 y 4.946.787).

El uso de sistemas de base vírica de ARN o ADN para el suministro de ácidos nucleicos toma ventaja de los procedimientos altamente desarrollados para fijar como objetivo un virus a células específicas en el cuerpo y traficar la carga útil vírica al núcleo. Se pueden administrar vectores víricos directamente a pacientes (in vivo) o se pueden usar para tratar células in vitro y las células modificadas se pueden administrar opcionalmente a los pacientes (ex vivo). Los sistemas con base vírica convencionales podían incluir vectores retrovirales, de lentivirus, adenovirales, adenoasociados y del virus del herpes simple para transferencia de genes. La integración en el genoma huésped es posible con los métodos de transferencia de genes de retrovirus, lentivirus y virus adenoasociados, dando como resultado con frecuencia expresión a largo plazo del transgen insertado. Adicionalmente, se han observado altas eficacias de transducción en muchos tipos celulares diferentes y tejidos fijados como objetivo.

El tropismo de un retrovirus se puede modificar por incorporación de proteínas de sobre extrañas, expandiendo la población objetivo potencial de células objetivo. Los vectores lentivirales son vectores retrovirales que son capaces de transducir o infectar células no de división y típicamente producen títulos víricos altos. La selección de un sistema de transferencia de genes retrovirales dependería por lo tanto del tejido objetivo. Los vectores retrovirales están constituidos por repeticiones terminales largas que actúan en cis con capacidad de empaquetamiento para hasta 6-10 kb de secuencia extraña. Las mínimas LTR que actúan en cis son suficientes para la replicación y empaquetamiento de los vectores, que se usan después para integrar el gen terapéutico en la célula objetivo para proporcionar expresión de transgenes permanente. Los vectores retrovirales ampliamente usados incluyen los basados en virus de la leucemia murina (MuLV), virus de la leucemia del mono gibón (GaLV), virus de la inmunodeficiencia simia (VIS), virus de la inmunodeficiencia humana (VIH) y combinaciones de los mismos (véase, por ej., Buchscher et al., J. Virol. 66: 2.731-2.739 (1.992); Johann et al., J. Virol. 66: 1.635-1.640 (1.992); Sommnerfelt et al., Virol. 176: 58-59 (1.990); Wilson et al., J. Virol. 63: 2.374-2.378 (1.989); Miller et al., J. Virol. 65: 2.220-2.224 (1991); Patente de EE.UU. PCT/US94/05700).

En aplicaciones donde se prefiere la expresión transitoria, se pueden usar sistemas a base de adenovirus. Los vectores a base de adenovirus son capaces de una eficacia de transducción muy alta en muchos tipos de células y no requieren división celular. Con tales vectores, se han obtenido altos títulos y niveles de expresión. Este vector se puede producir en grandes cantidades en un sistema relativamente simple. También se pueden usar vectores de virus adenoasociados ("AAV", por sus siglas en inglés) para transducir células con ácidos nucleicos objetivo, por ejemplo, en la producción in vitro de ácidos nucleicos y péptidos y para procedimientos de tratamiento con genes in vivo y ex vivo (véase, por ej., West et al., Virology 160: 38-47 (1.987); la Patente de EE.UU. N° 4.797.368; la patente internacional WO 93/24641; Kotin, Human Gene Therapy 5: 793-801 (1.994); Muzyczka, J. Clin. Invest. 94: 1.351 (1.994). Se describe la construcción de vectores AAV recombinantes en una serie de publicaciones, incluyendo la Patente de EE.UU. N° 5.173.414; Tratschin et al., Mol. Cell. Biol. 5: 3.251-3.260 (1.985); Tratschin, et al., Mol. Cell Biol. 4: 2.072-2.081 (1.984); Hermonat & Muzyczka, PNAS 81: 6.466-6.470 (1.984) y Samulski et al., J. Virol. 63: 03822-3828 (1.989).

Típicamente se usan células de empaquetamiento para formar partículas de virus que sean capaces de infectar una célula huésped. Tales células incluyen células 293, que empaquetan adenovirus, y células ψ 2 o células PA317, que empaquetan retrovirus. Los vectores víricos usados en terapia génica se generan habitualmente produciendo una estirpe celular que empaqueta un vector de ácidos nucleicos en una partícula vírica. Los vectores contienen típicamente las mínimas secuencias víricas requeridas para empaquetamiento y posterior integración en un huésped, siendo reemplazadas otras secuencias víricas por una casete de expresión para el polinucleótido o los polinucleótidos que se tienen que expresar. Las funciones víricas que faltan se suministran típicamente en trans por la estirpe celular de empaquetamiento. Por ejemplo, los vectores AAV usados en terapia génica típicamente poseen sólo secuencias ITR del genoma AAV que se requieren para empaquetamiento e integración en el genoma huésped. El ADN vírico se empaqueta en una estirpe celular, que contiene un plásmido auxiliar que codifica los otros genes AAV, es decir rep y cap, pero carecen de secuencias ITR. La estirpe celular también puede infectarse con adenovirus como auxiliar. El virus auxiliar activa la replicación del vector AAV y la expresión de genes AAV del plásmido auxiliar. El plásmido auxiliar no se empaqueta en cantidades significativas debido a ausencia de secuencias ITR. La contaminación con adenovirus se puede reducir por, por ejemplo, tratamiento térmico al cual es

más sensible el adenovirus que AAV. Los métodos adicionales para el suministro de ácidos nucleicos a células son conocidos para los expertos en la materia. Véase, por ejemplo, la Patente de EE.UU. 20030087817.

En algunas realizaciones, una célula huésped se transfecta de manera transitoria o de manera no transitoria con uno o más vectores descritos en la presente memoria. En algunas realizaciones, una célula se transfecta como tiene lugar de manera natural en un individuo. En algunas realizaciones, una célula que se transfecta es tomada de un individuo. En algunas realizaciones, la célula procede de células tomadas de un individuo, tales como una estirpe celular. Una amplia variedad de estirpes celulares de cultivos de tejido se conocen en la técnica. Ejemplos de estirpes celulares incluyen, pero no se limitan a, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panc1, PC-3, TF1, CTL2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, epitelial de riñón de mono BS-C-1, fibroblasto de embrión de ratón BALB/ 3T3, 3T3 Swiss, 3T3-L1, fibroblastos fetales humanos 132-d5; fibroblastos de ratón 10.1, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, células BCP-1, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr -/-, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepa1c1c7, HL-60, HMEC, HT-29, Jurkat, células JY, células K562, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCK II, MDCK II, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145 estirpes celulares OPCN / OPCT, Peer, PNT-1A / PNT 2, RenCa, RIN-5F, RMA/RMAS, células Saos-2, Sf-9, SkBr3, T2, T-47D, T84, estirpe celular THP1, U373, U87, U937, VCaP, células Vero, WM39, WT-49, X63 YAC-1, YAR y variedades transgénicas de los mismos. Las estirpes celulares están disponibles de una variedad de fuentes conocidas por los expertos en la materia (véase, por ej., La Colección de Cultivos de Tipo Americano (ATCC, por sus siglas en inglés) (Manassus, Va.)). En algunas realizaciones, una célula transfectada con uno o más vectores descritos en la presente memoria se usa para establecer una nueva estirpe celular que comprende una o más secuencias procedentes de vector. En algunas realizaciones, una célula transfectada de manera transitoria con los componentes de un sistema CRISPR como se describe en la presente memoria (tal como por transfección transitoria de uno o más vectores o transfección con ARN) y modificada por la actividad de un complejo CRISPR, se usa para establecer una nueva estirpe celular que comprende células que contienen la modificación pero carecen de otra secuencia exógena. En algunas realizaciones, las células transfectadas de manera transitoria o de manera no transitoria con uno o más vectores descritos en la presente memoria o estirpes celulares procedentes de dichas células se usan en la valoración de uno o más compuestos de ensayo.

En algunas realizaciones, uno o más vectores descritos en la presente memoria se usan para producir un animal transgénico no humano o planta transgénica. En algunas realizaciones, el animal transgénico es un mamífero, tal como un ratón, rata o conejo. En algunas realizaciones, el organismo o individuo es una planta. En algunas realizaciones, el organismo o individuo o planta es un alga. Los métodos para producir plantas y animales transgénicos son conocidos en la técnica y generalmente empiezan con un método de transfección de células, tal como se describe en la presente memoria. También se proporcionan animales transgénicos como son plantas transgénicas, especialmente cultivos y algas. El animal o la planta transgénicos pueden ser útiles en aplicaciones fuera de proporcionar un modelo de enfermedad. Éstas pueden incluir producción de alimentos o de alimentación por expresión de, por ejemplo, mayores niveles de proteína, carbohidrato, nutriente o vitaminas que lo que se observaría normalmente en el tipo natural. Con respecto a esto, se prefieren las plantas transgénicas, especialmente legumbres y tubérculos y animales, especialmente los mamíferos tales como ganado (vacas, ovejas, cabras y cerdos), pero también aves de corral e insectos comestibles.

Las algas transgénicas u otras plantas tales como colza pueden ser útiles en particular en la producción de aceites vegetales o biocombustibles tales como alcoholes (especialmente metanol y etanol), por ejemplo. Estos se pueden conseguir por ingeniería para expresar o sobreexpresar altos niveles de aceite o alcoholes para uso en las industrias de aceites o biocombustibles.

En un aspecto, la descripción proporciona métodos para modificar un polinucleótido objetivo en una célula eucariota. En algunas realizaciones, el método comprende permitir que un complejo CRISPR se una al polinucleótido objetivo para efectuar la escisión de dicho polinucleótido objetivo modificando de ese modo el polinucleótido objetivo, en la que el complejo CRISPR comprende una enzima CRISPR complejada con una secuencia guía hibridada a una secuencia fijada como objetivo dentro de dicho polinucleótido objetivo, en el que dicha secuencia guía está ligada a una secuencia de emparejamiento *tracr* que a su vez híbrida a una secuencia *tracr*.

En un aspecto, la descripción proporciona un método para modificar la expresión de un polinucleótido en una célula eucariota. En algunas realizaciones, el método comprende permitir que un complejo CRISPR se una al polinucleótido de manera que dicha unión dé como resultado expresión aumentada o disminuida de dicho polinucleótido; en el que el complejo CRISPR comprende una enzima CRISPR complejada con una secuencia guía híbrida a una secuencia fijada como objetivo dentro de dicho polinucleótido, en el que dicha secuencia guía está ligada a una secuencia de emparejamiento tracr que a su vez híbrida a una secuencia tracr.

Con los recientes avances en la genómica de cultivos, la capacidad para usar sistemas CRISPR-Cas para realizar edición de genes y manipulación eficaz y de coste eficaz permitirá la rápida selección y la comparación de manipulaciones genéticas únicas y multiplexadas para transformar dichos genomas para producción mejorada y cebos mejorados. Con respecto a esto se hace referencia a las Patentes y publicaciones de EE.UU.: Patente de EE.UU. Nº 6.603.061 - Agrobacterium-Mediated Plant Transformation Method; la Patente de EE.UU. Nº 7.868.149 - Plant Genome Sequences and Uses Thereof y la Patente de EE.UU. 2009/0100536 - Transgenic Plants with Enhanced Agronomic Traits. En la práctica de la descripción, también se hace referencia a los contenidos y la descripción de Morrell et al "Crop genomics: advances and applications" Nat Rev Genet. 29 de diciembre de 2.011; 13 (2): 85-96. En una realización ventajosa de la descripción, el sistema CRISPR/Cas9 se usa para lograr por ingeniería micro algas. De acuerdo con esto, la referencia de la presente memoria a células de animales puede aplicarse también, mutatis mutandis, a células de plantas a menos que sea evidente de otro modo.

En un aspecto, la descripción proporciona métodos para modificar un polinucleótido objetivo en una célula eucariota, que puede ser in vivo, ex vivo o in vitro. En algunas realizaciones, el método comprende muestrear una célula o población de células de un animal humano o no humano o planta (incluyendo microalgas) y modificar la célula o células. Puede tener lugar cultivo en cualquier fase ex vivo. La célula o las células pueden ser introducidas de nuevo incluso en el animal no humano o planta (incluyendo microalgas).

En un aspecto, la descripción proporciona estuches que contienen uno o más cualesquiera de los elementos descritos en los métodos y composiciones anteriores. En algunas realizaciones, el estuche comprende un sistema vector e instrucciones para usar el estuche. En algunas realizaciones, el sistema vector comprende (a) un primer elemento regulador unido de manera operable a una secuencia de emparejamiento tracr y uno o más sitios de inserción para insertar una secuencia guía aguas arriba de la secuencia de emparejamiento tracr, en el que cuando se expresa, la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a una secuencia fijada como objetivo en una célula eucariota, en la que el complejo CRISPR comprende una enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr y/o (b) un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica dicha enzima CRISPR que comprende una secuencia de localización nuclear. Los elementos se pueden proporcionar de manera individual o en combinaciones y se pueden proporcionar en cualquier envase adecuado, tal como un vial, un frasco o un tubo. En algunas realizaciones, el estuche incluye instrucciones en uno o más idiomas, por ejemplo en más de un idioma.

En algunas realizaciones, un estuche comprende uno o más reactivos para uso en un procedimiento utilizando uno o más de los elementos descritos en la presente memoria. Los reactivos se pueden proporcionar en cualquier envase adecuado. Por ejemplo, un estuche puede proporcionar uno o más tampones de reacción o de almacenamiento. Los reactivos se pueden proporcionar en una forma que sea utilizable en un ensayo particular o en una forma que requiera la adición de otro u otros componentes más antes de uso (por ejemplo, en forma concentrada o liofilizada). Un tampón puede ser cualquier tampón, incluyendo pero no limitándose a tampón de carbonato de sodio, un tampón de bicarbonato de sodio, un tampón de borato, un tampón de Tris, un tampón de MOPS, un tampón de HEPES y combinaciones de los mismos. En algunas realizaciones, el tampón es alcalino. En algunas realizaciones, el tampón presenta un pH de aproximadamente 7 a aproximadamente 10. En algunas realizaciones, el estuche comprende uno o más oligonucleótidos correspondientes a una secuencia guía para inserción en un vector de manera que se una de manera operable a la secuencia guía y un elemento regulador. En algunas realizaciones, el estuche comprende un polinucleótido de plantilla de recombinación homólogo.

En un aspecto, la descripción proporciona métodos para usar uno o más elementos de un sistema CRISPR. El complejo CRISPR de la descripción proporciona un medio eficaz para modificar un polinucleótido objetivo. El complejo CRISPR de la descripción presenta una amplia variedad de utilidad incluyendo modificación (por ej., supresión, inserción, traslocación, inactivación, activación) de un polinucleótido objetivo en una multiplicidad de tipos de células. Como tal el complejo CRISPR de la descripción presenta un amplio espectro de aplicaciones en, por ej., tratamiento génico, investigación de fármacos, diagnóstico de enfermedades y prognosis. Un complejo CRISPR ejemplar comprende una enzima CRISPR complejada con una secuencia guía hibridada a una secuencia fijada como objetivo dentro del polinucleótido objetivo. La secuencia guía está ligada a una secuencia de emparejamiento tracr, que a su vez hibrida a una secuencia tracr.

El polinucleótido objetivo de un complejo CRISPR puede ser cualquier polinucleótido endógeno o exógeno a la célula eucariota. Por ejemplo, el polinucleótido objetivo puede ser un polinucleótido que resida en el núcleo de la célula eucariota. El polinucleótido objetivo puede ser una secuencia que codifique un producto génico (por ej., una proteína) o una secuencia no codificadora (por ej., un polinucleótido regulador o un ADN basura). Sin desear estar limitados por la teoría, se cree que la secuencia objetivo debería estar asociada a un PAM (unidad adyacente protoespaciadora); esto es, una secuencia corta reconocida por el complejo CRISPR. Los requerimientos de secuencia y longitud precisos para el PAM difieren dependiendo de la enzima CRISPR usada, pero los PAM son típicamente secuencias de 2-5 pares de bases adyacentes al protoespaciador (esto es, la secuencia objetivo). Ejemplos de secuencias PAM se proporcionan en la sección de ejemplos a continuación y el experto podrá identificar secuencias PAM adicionales para uso con una enzima CRISPR determinada.

El polinucleótido objetivo de un complejo CRISPR puede incluir una serie de genes y polinucleótidos asociados a

enfermedades así como genes y polinucleótidos asociados a ruta bioquímica de señalización.

Ejemplos de polinucleótidos objetivo incluyen una secuencia asociada a una ruta bioquímica de señalización, por ej., un gen o polinucleótido asociado a una ruta bioquímica de señalización. Ejemplos de polinucleótidos objetivo incluyen un gen o polinucleótido asociado a una enfermedad. Un gen o polinucleótido "asociado a una enfermedad" se refiere a cualquier gen o polinucleótido que esté proporcionando productos de transcripción o traducción en un nivel anormal o en una forma anormal en las células procedentes de tejidos afectados por una enfermedad comparado con tejidos o células de un control no de enfermedad. Puede ser un gen que llegue a expresarse a un nivel anormalmente alto; puede ser un gen que llegue a expresarse a un nivel anormalmente bajo, donde la expresión modificada se correlaciona con la aparición y/o progresión de la enfermedad. Un gen asociado a la enfermedad también se refiere a un gen que posee mutación o mutaciones o variación genética que es directamente responsable o está ligada a desequilibrio con un gen, o genes, que es responsable de la etiología de una enfermedad. Los productos transcritos o traducidos pueden ser conocidos o desconocidos y puede ser a un nivel normal o anormal.

Ejemplos de genes y polinucleótidos asociados a una enfermedad están disponibles en Instituto McKusick-Nathans de Medicina Genética, Universidad de Johns Hopkins (Baltimore, Md.) y Centro Nacional para Información de Biotecnología, Biblioteca Nacional de Medicina (Bethesda, Md.), disponible en la Red Informática Mundial.

Ejemplos de genes y polinucleótidos asociados a una enfermedad se enumeran en las Tablas A y B. La información específica de la enfermedad está disponible en el Instituto McKusick-Nathans de Medicina Genética, Universidad de Johns Hopkins (Baltimore, Md.) y Centro Nacional para Información de Biotecnología, Biblioteca Nacional de Medicina (Bethesda, Md.), disponible en la Red Informática Mundial. Ejemplos de genes y polinucleótidos asociados a una ruta bioquímica de señalización se enumeran en la Tabla C.

Las mutaciones en estos genes y rutas pueden dar como resultado la producción de proteínas inapropiadas o proteínas en cantidades inapropiadas que afecten a la función. Tales genes, proteínas y rutas pueden ser el polinucleótido objetivo de un complejo CRISPR.

Tabla A

ENFERMEDAD/TRASTORNOS	GEN (ES)
Neoplasia	PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4;
	Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF;
	HIF1a; HIF3a; Met; HRG; Bcl2; PPAR alfa; PPAR
	gamma; WT1 (Tumor Wilms); Familia Receptor FGF
	miembros (5 miembros: 1, 2, 3, 4, 5); CDKN2a; APC; RB
	(retinoblastoma); MEN1; VHL; BRCA1; BRCA2; AR
	(Receptor Andrógenos); TSG101; IGF; Receptor IGF; Igf1 (4
	variantes); Igf2 (3 variantes); Receptor Igf 1; Receptor Igf 2;
	Bax; Bcl2; familia caspases (9 miembros:
	1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Apc
Degeneración	Abcr; Cc12; Cc2; cp (ceruloplasmina); Timp3; catepsina D;
Macular relacionada con la edad	Vldlr; Ccr2
Esquizofrenia	Neuregulina1 (Nrg1); Erb4 (receptor para Neuregulina);
	Complexina1 (Cplx1); Tph1 Triptófano hidroxilasa; Tph2
	Triptófano hidroxilasa 2; Neurexina 1; GSK3; GSK3a;
	GSK3b
Trastornos	5-HTT (Slc6a4); COMT; DRD (Drd1a); SLC6A3; DAOA;
	DTNBP1; Dao (Dao1)
Repetición del trinucleótido	HTT (Huntington's Dx); SBMA/SMAX1/AR (Kennedy's
Trastornos	Dx); FXN/X25 (Friedrich's Ataxia); ATX3 (Machado-
	Joseph's Dx); ATXN1 y ATXN2 (ataxia

ENFERMEDAD/TRASTORNOS	GEN (ES)
	espinocerebelar); DMPK (distrofia miotónica); Atrophin-1 y Atn1
	(DRPLA Dx); CBP (Creb-BP - inestabilidad global); VLDLR
	(Alzheimer's); Atxn7; Atxn10
Síndrome frágil X	FMR2; FXR1; FXR2; mGLUR5
Trastornos	APH-1 (alfa y beta); Presenilina (Psen1); nicastrina
Relacionados con la secretasa	(Ncstn); PEN-2
Otros	Nos1; Parp1; Nat1; Nat2
Trastornos relacionados con priones	Prp
ALS	SOD1; ALS2; STEX; FUS; TARDBP; VEGF (VEGF-a;
	VEGF-b; VEGF-c)
Drogadicción	Prkce (alcohol); Drd2; Drd4; ABAT (alcohol); GRIA2;
	Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Gria1 (alcohol)
Autismo	Mecp2; BZRAP1; MDGA2; Sema5A; Neurexina 1; Frágil X
	(FMR2 (AFF2); FXR1; FXR2; Mglur5)
enfermedad de Alzheimer	E1; CHIP; UCH; UBB; Tau; LRP; PICALM.; Clusterina; PS1;
	SORL1; CR1; Vldlr; Uba1; Uba3; CHIP28 (Aqp1,
	Aquaporina 1); Uchl1; Uchl3; APP
Inflamación	IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL-
	17b; IL-17c; IL-17d; IL-17f); IL-23; Cx3cr1; ptpn22; TNFa;
	NOD2/CARD15 para IBD; IL-6; IL-12 (IL-12a; IL-12b);
	CTLA4; Cx3cl1
enfermedad de Parkinson	x-Synuclein; DJ-1; LRRK2; Parkin; PINK1

Tabla B

Enfermedades y trastornos de la sangre y la coagulación	Anemia (CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, BUMPH1, PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB, ABCB7, ABC7, ASAT); síndrome de linfocito desnudo (TAPBP, TPSN, TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP, RFX5); Trastornos de sangrado (TBXA2R, P2RX1, P2X1); Factor H y factor H tipo - 1 (HF1, CFH, HUS); Factor V y factor VIII (MCFD2); deficiencia de Factor VII (F7); deficiencia de Factor X (F10); deficiencia de Factor XI (F11); deficiencia de Factor XII (F12, HAF); deficiencia de Factor XIIIa (F13A1, F13A); deficiencia de Factor XIIIb (F13B); anemia Fanconi (FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064, FANCB, FANCC, FACC, BRCA2, FANCD1, FANCD2, FANCD, FACD, FAD, FANCE, FACE, FANCF, XRCC9, FANCG, BRIP1, BACH1, FANCI, PHF9, FANCL, FANCM, KIAA1596); trastornos Linfocitosis Hemofagocítica (PRF1, HPLH2, UNC13D, MUNC13-4, HPLH3, HLH3, FHL3); Hemofilia A (F8, F8C, HEMA); Hemofilia B (F9, HEMB); trastornos hemorrágicos (PI, ATT, F5); deficiencias y trastornos de leucocia (ITGB2, CD18, LCAMB, LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LWMM, CACH, CLE, EIF2B4); anemia depreanocítica (HBB); Talasemia (HBA2, HBB, HBD, LCRB, HBA1).
Desregulación celular y enfermedades y trastornos oncológicos	Linfoma no de Hodgkin de células B (BCL7A, BCL7); Leucemia (TAL1, TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1, HOXD4, HOX4B, BCR, CML, PHL, ALL, ARNT, KRAS2, RASK2, GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH, CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214, D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1L1, NSD3, FLT3, AF1Q, NPM1, NUMA1, ZNF 145, PLZF, PML, MYL, STAT5B, AF 10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML, PHL, ALL, GRAF, NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2, NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1, NFE1,

	ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN).
Inflamación y enfermedades y trastornos inmunorelacionados	AIDS (KIR3DL1, NKAT3, NKB1, AMB11, KIR3DS1, IFNG, CXCL12, SDF1); síndrome linfoproliferativo autoinmunitario (TNFRSF6, APT1, FAS, CD95, ALPS1A); inmunodeficiencia combinada (IL2RG, SCIDX1, SCIDX, IMD4); HIV-1 (CCL5, SCYA5, D17S136E, TCP228), susceptibilidad o infección por VIH (IL10, CSIF, CMKBR2, CCR2, CMKBR5, CCCR5 (CCR5)); Inmunodeficiencias (CD3E, CD3G, AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4, TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX, TNFRSF14B, TACI); Inflamación (IL-10, IL-1 (IL-1a, IL-1b), IL-13, IL-17 (IL-17a (CTLA8), IL-17b, IL-17c, IL-17d, IL-17f), IL-23, Cx3cr1, ptpn22, TNFa, NOD2/CARD15 for IBD, IL-6, IL-12 (IL-12a, IL-12b), CTLA4, Cx3cl1); inmunodeficiencias combinadas severas (SCIDs)(JAK3, JAKL, DCLRE1C, ARTEMIS, SCIDA, RAG1, RAG2, ADA, PTPRC, CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCIDX1, SCIDX, IMD4).
Enfermedades y trastornos metabólicos, hepáticos, renales y proteínicos	Neuropatía amiloide (TTR, PALB); Amiloidosis (APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); Cirrosis (KRT18, KRT8, C1RH1A, NAIC, TEX292, K1AA1988); fibrosis quística (CFTR, ABCC7, CF, MRP7); enfermedades de almacenamiento de glucógeno (SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM); adenoma hepático, 142330 (TCF1, HNF1A, MODY3), fallo hepático, comienzo temprano y trastorno neurológico (SCOD1, SCO1), deficiencia de lipasa hepática(LIPC), Hepatoblastoma, cáncer y carcinomas (CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5; enfermedad del riñón quístico medular (UMOD, HNFJ, FJHN, MCKD2, ADMCKD2); Fenilcetonuria (PAH, PKU1, QDPR, DHP, PTS); enfermedad renal y hepática poliquística (FCYT, PKHD1, ARPKD, PKD 1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63).
enfermedades y trastornos musculares / del esqueleto	Distrofia muscular de Becker (DMD, BMD, MYF6), Distrofia muscular de Duchenne (DMD, BMD); Distrofia muscular de Emery-Dreifuss (LMNA, LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1, EMD2, FPLD, CMD1A); Distrofia muscular de Facioscapulohumeral (FSHMD1A, FSHD1A); Distrofia muscular (FKRP, MDC1C, LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD, TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C, DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB, LGMD2E, SGCD, SGD, LGMD2F, CMD1L, TCAP, LGMD2G, CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD21, TTN, CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPNI, SELN, RSMD1, PLEC1, PLTN, EBS1); Osteopetrosis (LRP5, BMND1, LRP7, LR3, OPPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1, TIRC7, OC116, OPTB1); Atrofia muscular (VAPB, VAPC, ALS8, SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1, CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1).
enfermedades y trastornos neurológicos y neuronales	ALS (SOD1, ALS2, STEX, FUS, TARDBP, VEGF (VEGF-a, VEGF-b, VEGF-c); enfermedad de Alzheimer (APP, AAA, CVAP, AD1, APOE, AD2, PSEN2, AD4, STM2, APBB2, FE65L1, NOS3, PLAUI, URK, ACE, DCP1, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH, PSEN1, AD3); Autismo (MECP2, BZRAP1, MDGA2, Sema5A, Neurexina 1, GLO1, MECP2, RTT, PPMX, MRX16, MRX79, NLGN3, NLGN4, KIAA1260, AUTSX2); Síndrome frágil X (FMR2, FXR1, FXR2, mGLUR5); enfermedad de Huntington y trastornos tipo enfermedad (HD, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17); enfermedad de Parkinson (NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, SNCA, NACP, PARK1, PARK4, DJ1, PARK7, LRRK2, PARK8, PINK1, PARK6, UCHL1, PARK5, SNCA, NACP, PARK1, PARK4, PRKN, PARK2, PDJ, DBH, NDUFV2); síndrome de Rett (MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, x-Synuclein, DJ-1); Esquizofrenia (Neuregulin1 (Nrg1), Erb4 (receptor de Neuregulin), Complexin1 (Cplx1), Tph1 Triptófano hidroxilasa, Tph2, Triptófano hidroxilasa 2, Neurexina 1, GSK3, GSK3a, GSK3b, 5-HTT (Slc6a4), COMT, DRD (Drd1a), SLC6A3, DAOA, DTNBP1, Dao (Dao1)); Trastornos Relacionados con la secretasa (APH-1 (alfa y beta), Presenilin (Psen1), nicastrin, (Ncstn), PEN-2, Nos1, Parp1, Nat1, Nat2); Trastornos Repetición del trinucleótido (HTT (Huntington's Dx), SBMA/SMAX1/AR (Kennedy's Dx), FXN/X25 (Ataxia de Friedrich), ATX3 (Machado- Joseph's Dx), ATXN1 y ATXN2 (ataxias espinocerebelares), DMPK (distrofia miotónica), Atrofina-1 y Atn1 (DRPLA Dx), CBP (Creb-BP - inestabilidad global), VLDLR (Alzheimer's), Atnx7, Atnx10).

enfermedades y trastornos oculares	degeneración macular relacionada con la edad (Abcr, Ccl2, Cc2, cp (ceruloplasmina), Timp3, cathepsinaD, Vldlr, Ccr2); Catarata (CRYAA, CRYA1, CRYBB2, CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6, AN2, MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19, CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM, MIP, AQP0, CRYAB, CRYA2, CTPP2, CRYBB1, CRYGD, CRYG4, CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1, GJA8, CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1); distrofia y turbidez de la córnea (APOA1, TGFB1, CSD2, CDGG1, CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD, PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD); Cornea plana congénita (KERA, CNA2); Glaucoma (MYOC, TIGR, GLC1A, JOAG, GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1, NTG, NPG, CYP1B1, GLC3A); amaurosis congénita de Leber (CRB1, RP12, CRX, CORD2, CRD, RPGRIP1, LCA6, CORD9, RPE65, RP20, AIPL1, LCA4, GUCY2D, GUC2D, LCA1, CORD6, RDH12, LCA3); distrofia macular (ELOVL4, ADMD, STGD2, STGD3, RDS, RP7, PRPH2, PRPH, AVMD, AOFMD, VMD2).
------------------------------------	---

Tabla C

FUNCIÓN CELULAR	GENES
Señalización PI3K/AKT	PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2;
	PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1;
	AKT2; IKBKB; PIK3CA; CDK8; CDKN1B; NFKB2; BCL2;
	PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2;
	ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3;
	PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7;
	YWHAZ; ILK; TP53; RAF1; IKBKG; RELB; DYRK1A;
	CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1;
	CHUK; PDPK1; PPP2R5C; CTNNB1; MAP2K1; NFKB1;
	PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2;
	TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK;
	HSP90AA1; RPS6KB1
Señalización ERK/MAPK	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2;
	EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6;
	MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1;
	PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A;
	PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN;
	EIF4EBP1; PPARG; PRKCD; PRKCA1; MAPK9; SRC;
	CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ;
	PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1;
	PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1;
	CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK
Señalización	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1;
Receptor Glucocorticoide	MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I;
	PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2;
	MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPKB; BCL2L1;
	MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13;

FUNCIÓN CELULAR	GENES
	RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1;
	PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3;
	MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP;
	CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2;
	PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1;
	ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1; STAT1; IL6; HSP90AA1
Señalización Guía Axonal	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12;
	IGF1; RAC1; RAP1A; EIF4E; PRKCZ; NRP1; NTRK2;
	ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2;
	PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2;
	CFL1; GNAQ; PIK3CB; CXCL 12; PIK3C3; WNT11;
	PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1;
	FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1;
	GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3;
	CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B;
	AKT3; PRKCA
Señalización Receptor Ephrin	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1;
	PRKAA2; EIF2AK2; RAC 1; RAP1A; GRK6; ROCK2;
	MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2;
	DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14;
	CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1;
	KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2;
	PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4; AKT1; JAK2; STAT3; ADAM10;
	MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2;
	EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4;
	AKT3; SGK
Señalización	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1;
cicloesqueleto de actina	PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6;
	ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8;
	PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8;
	F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD;
	PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7;
	PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3;
	ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL;
	BRAF; VAV3; SGK
Señalización	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2;
enfermedad de Huntington	MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2;

FUNCIÓN CELULAR	GENES
	PIK3CA; HDAC5; CREB1; PRKCI; HSPA5; REST;
	GNAQ; PIK3CB; PIK3C3; MAPKB; IGF1R; PRKD1;
	GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2;
	HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A;
	HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1;
	PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX;
	ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3
Señalización de muerte celular programada	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1;
	BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB;
	CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8;
	BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF;
	RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2;
	CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2;
	BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK;
	CASP3; BIRC3; PARP1
Señalización de Receptor de células B	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11;
	AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1;
	MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9;
	EGRI; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB;
	MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1;
	NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN;
	GSK3B; ATF4; AKT3; VAV3; RPS6KB1
Señalización de diapédesis	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA;
	RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11;
	MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12;
	PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB;
	MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK;
	MAPK14; NOX1; PXN; VIL2; VASP; TTGB1; MAP2K2;
	CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK;
	CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9
Señalización de Integrina	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A;
	TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2;
	CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8;
	CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP;
	RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1;
	TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2;
	CRKL; BRAF; GSK3B; AKT3

FUNCIÓN CELULAR	GENES
Señalización	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11;
de Respuesta de fase aguda	AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14;
	PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS;
	MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1;
	TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1;
	IKBKG; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1;
	CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN;
	AKT3; IL1R1; IL6
Señalización de PTEN	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L 11;
	MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA;
	CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1;
	MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR;
	RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2;
	AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1;
	NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2;
	GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
Señalización de p53	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A;
	BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2;
	PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1;
	PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9;
	CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A;
	HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1;
	SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN;
	SNAI2; GSK3B; BAX; AKT3
Señalización	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1;
de Receptor Arilhidrocarbonado	NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1;
	SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1;
	MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1;
	SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF;
	CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1;
	CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1;
	HSP90AA1
Señalización de	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1;
Metabolismo Xenobiótico	NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A;
	PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1;
	ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD;
	GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL;
	NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1;
	CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1;
	NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1;

FUNCIÓN CELULAR	GENES
	HSP90AA1
Señalización de SAPK/JNK	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC I ; ELK1;
	GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA;
	FADD; CDK8; PIK3CB; PIK3C3; MAPK8; R1PK1;
	GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A;
	TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2;
	PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1;
	CRKL; BRAF; SGK
Señalización de PPAR/RXR	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN;
	RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2;
	ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8;
	IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A;
	NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7;
	CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB 1;
	TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1; ADIPOQ
Señalización de NF-KB	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ; TRAF6;
	TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2;
	KRAS; RELA; PIK3C2A; TRAF2; TLR4; PDGFRB; TNF;
	INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1;
	PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10;
	GSK3B; AKT3; TNFAIP3; IL1R1
Señalización de Neuregulina	ERBB4; PRKCE; ITGAM; ITGA5; PTEN; PRKCZ; ELK1;
	MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI;
	CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS;
	PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2;
	ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3;
	EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL;
	AKT3; PRKCA; HSP90AA1; RPS6KB1
Señalización de	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO;
catenina Wnt y Beta	AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A;
	WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2; ILK;
	LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1;
	PPP2R5C; WNT5A; LRP5; CTNNA1; TGFBR1; CCND1;
	GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B;
	AKT3; SOX2
Señalización de Receptor de Insulina	PTEN; INS; EIF4E; PTPN1; PRKCZ; MAPK1; TSC1;
	PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3;
	MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1;

FUNCIÓN CELULAR	GENES
	SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN;
	MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1;
	GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK;
	RPS6KB1
Señalización de IL-6	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK3;
	MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1;
	MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG;
	RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3;
	MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6
Colestasis Hepática	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA;
	RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8;
	PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1;
	TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8;
	CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4;
	JUN; IL1R1; PRKCA; IL6
Señalización de IGF-1	IGF 1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2;
	PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8;
	IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A;
	YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1;
	PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3;
	FOXO1; SRF; CTGF; RPS6KB1
Respuesta de Estrés	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1;
Oxidativo mediado por NRF2	NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL;
	NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP;
	MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1;
	GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
Fibrosis Hepática /	EDN1; 1GF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF;
Activación de células estelares hepáticas	SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9;
	IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8;
	PDGFRA; NFKB1; TGFB1; SMAD4; VEGFA; BAX;
	IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
Señalización de PPAR	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB;
	NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3;
	NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2;
	PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG;
	RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA;
	MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1
Señalización de Fc Epsilon RI	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11;

FUNCIÓN CELULAR	GENES
	AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD;
	MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN;
	MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3;
	VAV3; PRKCA
Señalización de Receptor	PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB;
Acoplado a proteína G	PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB;
	PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1;
	IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK;
	PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3;
	PRKCA
Metabolismo de	PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6;
Inositol fosfato	MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2;
	PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A; PIK3R1;
	MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK
Señalización de PDGF	EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB;
	PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC;
	PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2;
	PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC;
	JUN; CRKL; PRKCA; SRF; STAT1; SPHK2
Señalización de VEGF	ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF;
	AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3;
	BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN;
	RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN;
	VEGFA; AKT3; FOXO1; PRKCA
Señalización de células	PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11;
aniquilantes naturales	KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB;
	PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6;
	PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1;
	PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA
Ciclo celular: G1/S	HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC;
Regulación de control	ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11;
	HDAC9; CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1;
	E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1;
	GSK3B; RBL1; HDAC6
Señalización de Receptor de células T	RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS;
	NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	RELA; PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB; FYN;
	MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10;

FUNCIÓN CELULAR	GENES
	JUN; VAV3
Señalización de Receptor de la muerte	CRADD; HSPB1; BID; BIRC4; TBK1; IKKBK; FADD;
	FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8;
	DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBK; RELB;
	CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3;
	BIRC3
Señalización de FGF	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11;
	AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8;
	MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1;
	AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4;
	AKT3; PRKCA; HGF
Señalización de GM-CSF	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A;
	STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3;
	ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3;
	STAT1
Señalización de esclerosis lateral amiotrófica	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2;
	PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1;
	PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1;
	APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
Señalización de JAK/Stat	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B;
	PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A;
	PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3;
	STAT1
Metabolismo de Nicotinato y Nicotinamida	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1;
	PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1;
	PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2;
	MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK
Señalización de quimiocinas	CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ;
	CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13;
	RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1;
	MAP2K2; MAP2K1; FLAN; CCL2; PRKCA
Señalización de IL-2	ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS;
	STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2;
	JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3
Depresión sináptica a largo lazo	PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS;
	PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3;

FUNCIÓN CELULAR	GENES
	KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA;
	YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA
Señalización de	TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2;
Receptor de Estrógenos	SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1;
	HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP;
	MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2
Ruta de	TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4;
Ubiquitinación de proteínas	CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7;
	USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8;
	USP1; VHL; HSP90AA1; BIRC3
Señalización de IL-10	TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2;
	MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF;
	IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1;
	JUN; IL1R1; IL6
Activación de VDR/RXR	PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1;
	NCOR2; SP1; PRKCI; CDKN1B; PRKD1; PRKCD;
	RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1;
	LRP5; CEBPB; FOXO1; PRKCA
Señalización de TGF-beta	EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1;
	FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2;
	SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2;
	MAP2K1; TGFB1; SMAD4; JUN; SMAD5
Señalización de Receptor tipo Toll	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13;
	RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK;
	NFKB1; TLR2; JUN
Señalización de p38 MAPK	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS;
	CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2;
	MAPK14; TNF; MAP3K7; TGFB1; MYC; ATF4; IL1R1;
	SRF; STAT1
Señalización de Neurotrofina/TRK	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS;
	PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A;
	RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1;
	CDC42; JUN; ATF4
Activación de FXR/RXR	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8;
	APOB; MAPK10; PPARG; MTTP; MAPK9; PPARGC1A;
	TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
Potenciación sináptica	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1;
a largo plazo	PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS;
	PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1;

FUNCIÓN CELULAR	GENES
	ATF4; PRKCA
Señalización de calcio	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1;
	CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11;
	HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4;
	HDAC6
Señalización de EGF	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1;
	STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
Señalización de Hipoxia en el	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT;
Sistema Cardiovascular	HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM;
	VEGFA; JUN; ATF4; VHL; HSP90AA1
Inhibición mediada por LPS/IL-1	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1;
de función RXR	MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2;
	TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1
Activación de LXR/RXR	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA;
	NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1;
	SREBF1; IL1R1; CCL2; IL6; MMP9
Proceso amiloide	PRKCE; CSNK1E; MAPK1; CAPNS 1; AKT2; CAPN2;
	CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1;
	PSEN1; CSNK1A1; GSK3B; AKT3; APP
Señalización de IL-4	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCKS1;
	PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1;
	FRAP1; AKT3; RPS6KB1
Regulación	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC;
Control de daño	CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A;
Ciclo celular: ADN G2/M	PRKDC; ATM; SFN; CDKN2A
Señalización de óxido nítrico en el	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3;
Sistema Cardiovascular	CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1;
	VEGFA; AKT3; HSP90AA1
Metabolismo de Purina	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4;
	PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C;
	NT5E; POLD1; NME1
Señalización mediada por cAMP	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3 ;
	SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
Disfunción mitocondrial	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9;
	PARK7; PSEN1; PARK2; APP; CASP3
Señalización de Notch	HES1; JAG1; NUMB; NOTCH4; ADAM 17; NOTCH2;
	PSEN1; NOTCH3; NOTCH1; DLL4
Ruta del estrés del	HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4;
Retículo endoplasmático	EIF2AK3; CASP3

FUNCIÓN CELULAR	GENES
Metabolismo de pirimidina	NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B;
	NT5E; POLD1; NME1
Señalización de Parkinson	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7;
	PARK2; CASP3
Señalización	GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC;
Cardíaca y beta adrenérgica	PPP2R5C
Glucolisis/Gluconeogénesis	HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1
Señalización de interferón	IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3
Señalización Sonic Hedgehog	ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRK1B
Metabolismo	PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2
Glicerofosfolípidos	
Degradación de fosfolípidos	PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2
Metabolismo de triptofano	SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1
Degradación de lisina	SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C
Ruta	ERCC5; ERCC4; XPA; XPC; ERCC1
Reparación escisión de nucleótidos	
Metabolismo de	UCHL1; HK2; GCK; GPI; HK1
Almidón y sacarosa	
Metabolismo de aminoazúcares	NQO1; HK2; GCK; HK1
Metabolismo de	PRDX6; GRN; YWHAZ; CYP1B1
Ácido araquidónico	
Señalización de ritmo circadiano	CSNK1E; CREB1; ATF4; NR1D1
Sistema de coagulación	BDKRB1; F2R; SERPINE1; F3
Receptor de Dopamina	PPP2R1A; PPP2CA; PPP1CC; PPP2R5C
Señalización	
Metabolismo de glutatona	IDH2; GSTP1; ANPEP; IDH1
Metabolismo de glicerolípidos	ALDH1A1; GPAM; SPHK1; SPHK2
Metabolismo de ácido linoleico	PRDX6; GRN; YWHAZ; CYP1B1
Metabolismo de Metionina	DNMT1; DNMT3B; AHCY; DNMT3A
Metabolismo de piruvato	GLO1; ALDH1A1; PKM2; LDHA
Metabolismo de	ALDH1A1; NOS3; NOS2A
Arginina y Prolina	
Señalización Eicosanoide	PRDX6; GRN; YWHAZ
metabolismo de	HK2; GCK; HK I
Fructosa y Manosa	
Metabolismo de Galactosa	HK2; GCK; HK1
Biosíntesis de estilbeno,	PRDX6; PRDX1; TYR
cumarina y lignina	
Ruta de	CALR; B2M
Presentación de antígenos	

FUNCIÓN CELULAR	GENES
Biosíntesis de esteroides	NQO1; DHCR7
Metabolismo de Butanoato	ALDH1A1; NLGN1
Ciclo del Citrato	IDH2; IDH1
Metabolismo de ácidos grasos	ALDH1A1; CYP1B1
Metabolismo de	PRDX6; CHKA
Glicerofosfolípidos	
Metabolismo de Histidina	PRMT5; ALDH1A1
Metabolismo de Inositol	ERO1L; APEX1
Metabolismo de xenobióticos	GSTP1; CYP1B1
por citocromo p450	
Metabolismo de metano	PRDX6; PRDX1
Metabolismo de fenilalanina	PRDX6; PRDX1
Metabolismo de Propanoato	ALDH1A1; LDHA
Metabolismo de	PRMT5; AHCY
Selenoaminoácido	
Metabolismo de esfingolípidos	SPHK1; SPHK2
Aminofosfonato	PRMT5
Metabolismo	
Andrógenos y estrógenos	PRMT5
Metabolismo de	
Ascorbato y Aldarato	ALDH1A1
Metabolismo de	
Biosíntesis de ácidos biliares	ALDH1A1
Metabolismo de cisteína	LDHA
Biosíntesis de ácidos grasos	FASN
Receptor de Glutamato	GNB2L1
Señalización de	
Respuesta de estrés	PRDX1
Oxidativa mediada por NRF2	
Ruta de	GPI
Pentosa fosfato	
Interconversiones de	UCHL1
Pentosa y Glucuronato	
Metabolismo de Retinol	ALDH1A1
Metabolismo de Riboflavina	TYR
Metabolismo de tirosina	PRMT5, TYR
Biosíntesis de Ubiquinona	PRMT5
Degradación de Valina, Leucina e	ALDH1A1
Isoleucina	
Metabolismo de Glicina, Serina y	CHKA

FUNCIÓN CELULAR	GENES
Treonina	
Degradación de lisina	ALDH1A1
dolor/sabor	TRPM5; TRPA1
dolor	TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2;
	Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca;
	Prkacb; Prkar1a; Prkar2a
Función mitocondrial	AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2
Neurología de desarrollo	BMP-4; Chordin (Chrd); Noggin (Nog); WNT (Wnt2;
	Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b;
	Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); beta-catenin;
	Dkk-1; proteínas relacionadas con Frizzled; Otx-2; Gbx2; FGF-8;
	Reelin; Dab1; unc-86 (Pou4f1 or Brn3a); Numb; Reln

Las realizaciones de la descripción también se refieren a métodos y composiciones relacionadas con eliminación de genes, la multiplicación de genes y reparación de mutaciones particulares asociadas a la inestabilidad de repeticiones de ADN y trastornos neurológicos (Robert D. Wells, Tetsuo Ashizawa, Genetic Instabilities and Neurological Diseases, Segunda Edición, Academic Press 13 de octubre de 2.011-Medical). Aspectos específicos de secuencias de repetición en tándem se han encontrado responsable de más de veinte enfermedades humanas (Nuevos conocimientos en la inestabilidad de la repetición: función de híbridos de ARN·ADN. Mclvor EI, Polak U, Napierala M. RNA Biol. Sep-Oct 2.010; 7 (5): 551-8). El sistema CRISPR-Cas puede ser aprovechado para corregir estos defectos de inestabilidad genómica.

- 5 Un aspecto más de la descripción se refiere a la utilización del sistema CRISPR-Cas para corregir defectos en los genes EMP2A y EMP2B que se ha identificado que están asociados a la enfermedad Lafora. La enfermedad de Lafora es un trastorno autosómico recesivo que se caracteriza por epilepsia mioclónica progresiva que puede empezar como ataques epilépticos en la adolescencia. Algunos casos de la enfermedad pueden estar causados por mutaciones en genes que se tienen que identificar sin embargo. La enfermedad produce ataques, espasmos musculares, dificultad para caminar, demencia y eventualmente la muerte. En la actualidad no hay tratamiento que se haya demostrado eficaz contra el progreso de la enfermedad. Otras anomalías genéticas asociadas a la epilepsia también pueden ser fijadas como objetivo mediante el sistema CRISPR-Cas y la genética subyacente se describe además en Genetics of Epilepsy and Genetic Epilepsies, editado por Giuliano Avanzini, Jeffrey L. Noebels, Mariani Foundation Paediatric Neurology: 20; 2.009).
- 10 En otro aspecto más de la descripción, el sistema CRISPR-Cas puede ser usado para corregir defectos oculares que surgen de mutaciones genéticas diversas descritas además en Genetic Diseases of the Eye, Segunda Edición , editado por Elias I. Traboulsi, Oxford University Press, 2.012.

25 Diversos aspectos adicionales de la descripción se refiere a corregir defectos asociados a un amplio intervalo de enfermedades genéticas que se describen además en la página web del Instituto Nacional para la Salud bajo la subsección tónica Trastornos Genéticos. Las enfermedades cerebrales genéticas pueden incluir pero no se limitan a, Adrenoleucodistrofia, Angenesia del Cuerpo Calloso, Síndrome de Aicardi, Enfermedad de Alper, Enfermedad de Alzheimer, Síndrome de Barth, Enfermedad de Batten, CADASIL, Degeneración Cerebelar, Enfermedad de Fabry, Enfermedad de Gerstmann-Straussler-Scheinker, Enfermedad de Huntington y otros Trastornos de Triple Repetición, Enfermedad de Leigh, Síndrome de Lesch-Nyhan, Enfermedad de Menkes, Miopatías Mitocondriales y Colpocefalia de NINDS (por sus siglas en inglés). Estas enfermedades se describen además en la página web del Instituto Nacional de la Salud bajo la subsección Trastornos Cerebrales Genéticos.

30 En algunas realizaciones, la afección puede ser neoplasia. En algunas realizaciones, donde la afección es neoplasia, los genes que se tienen que fijar como objetivo son cualquiera de los enumerados en la Tabla A (en este caso PTEN etc). En algunas realizaciones, la afección puede ser Degeneración Macular Relacionada con la Edad.

35 En algunas realizaciones, la afección puede ser un Trastorno Esquizofrénico. En algunas realizaciones, la afección puede ser un trastorno de Repetición del Trinucleótido. En algunas realizaciones, la afección puede ser Síndrome de Frágil X. En algunas realizaciones, la afección puede ser un Trastorno Relacionado con la Secretasa. En algunas realizaciones, la afección puede ser un trastorno relacionado con el Prión. En algunas realizaciones, la afección puede ser ALS (por sus siglas en inglés). En algunas realizaciones, la afección puede ser una drogadicción.

40 En algunas realizaciones, la afección puede ser Autismo. En algunas realizaciones, la afección puede ser Enfermedad de Alzheimer. En algunas realizaciones, la afección puede ser inflamación. En algunas realizaciones, la afección

puede ser Enfermedad de Parkinson.

Ejemplos de proteínas asociadas a la enfermedad de Parkinson incluyen pero no se limitan a α -sinucleína, DJ-1, LRRK2, PINK1, Parkin, UCHL1, Synphilin-1 y NURR1.

Ejemplos de proteínas relacionadas con la adicción pueden incluir ABAT por ejemplo.

- 5 Ejemplos de proteínas relacionadas con la inflamación pueden incluir la proteína quimioatrayente de monocitos -1 (MCP1) codificada por el gen Ccr2, el receptor de quimiocina C-C tipo 5 (CCR5) codificada por el gen Ccr5, el receptor IIB de IgG (FCGR2b, también denominado CD32) codificada por el gen Fcgr2b o la proteína de R1g Fc épsilon (FCER1g) codificada por el gen Fcer1g, por ejemplo.

- 10 Ejemplos de proteínas asociadas a enfermedades cardiovasculares pueden incluir IL1B (interleucina 1, beta), XDH (xantina deshidrogenasa), TP53 (proteína tumoral p53), PTGIS (prostaglandina I2 (prostaciclina) sintasa), MB (mioglobina), IL4 (interleucina 4), ANGPT1 (angiopoyetina 1), ABCG8 (casete de unión a ATP, sub-familia G (WHITE), miembro 8) o CTSK (cathepsina K), por ejemplo.

- 15 Ejemplos de proteínas asociadas a la enfermedad de Alzheimer pueden incluir la proteína receptora de lipoproteína de densidad muy baja (VLDLR, por sus siglas en inglés) codificada por el gen VLDLR, la enzima activadora de modificador de tipo ubiquitina 1 (UBA1) codificada por el gen UBA1 o la proteína de subunidad catalítica de enzima E1 activadora de NEDD8 (UBE1C) codificada por el gen UBA3, por ejemplo.

- 20 Ejemplos de proteínas asociadas a Trastorno por Espectro Autista pueden incluir la proteína 1 asociada a receptor de benzodiazapina (periférica) (BZRAP1) codificada por el gen BZRAP1, la proteína del miembro 2 de la familia AF4/FMR2 (AFF2) codificada por el gen AFF2 (también denominado MFR2), la proteína de homólogo 1 autosómico de retraso mental frágil X (FXR1) codificada por el gen FXR1 o la proteína de homólogo 2 autosómico del retraso mental frágil X (FXR2) codificada por el gen FXR2, por ejemplo.

- 25 Los ejemplos de proteínas asociadas a Degeneración Macular pueden incluir la casete de unión de ATP, proteína del miembro 4 (ABCA4) de la sub-familia A (ABC 1) codificada por el gen ABCR, la proteína apolipoproteína E (APOE) codificada por el gen APOE o la proteína de Ligando 2 (CCL2) de quimiocina (unidad C-C) codificada por el gen CCL2, por ejemplo.

Ejemplos de proteínas asociadas a Esquizofrenia pueden incluir NRG1, ErbB4, CPLX1, TPH1, TPH2, NRXN1, GSK3A, BDNF, DISC1, GSK3B y combinaciones de las mismas.

- 30 Ejemplos de proteínas implicadas en la supresión de tumores pueden incluir ATM (ataxia telangiectasia mutada), ATR (ataxia telangiectasia y relacionada con Rad3), EGFR (receptor de factor de crecimiento epidérmico), ERBB2 (homólogo 2 de oncogén vírico de leucemia eritroblástica v-erb-b2), ERBB3 (homólogo 3 de oncogén vírico de leucemia eritroblástica v-erb-b2), ERBB4 (homólogo 4 de oncogén vírico de leucemia eritroblástica v-erb-b2), Notch 1, Notch2, Notch 3 o Notch 4, por ejemplo.

- 35 Ejemplos de proteínas asociadas a trastorno de la secretas pueden incluir PSENEN (homólogo de estimulador 2 de presenilina (C. elegans)), CTSB (cathepsina B), PSEN1 (presenilina 1), APP (proteína precursora de amiloide beta (A4)), APH1B (homólogo B defectuoso 1 anterior de la faringe (C. elegans)), PSEN2 (presenilina 2 (enfermedad de Alzheimer 4)) o BACE1 (enzima 1 de escisión de APP de sitio beta), por ejemplo.

- 40 Los ejemplos de proteínas asociadas a Esclerosis Lateral Amiotrófica pueden incluir SOD1 (superóxido dismutasa 1), ALS2 (esclerosis lateral amiotrófica 2), FUS (fusionado en sarcoma), TARDBP (proteína de unión de ADN TAR), VAGFA (factor A de crecimiento endotelial vascular), VAGFB (factor B de crecimiento endotelial vascular) y VAGFC (factor C de crecimiento endotelial vascular) y cualquier combinación de los mismos.

Ejemplos de proteínas asociadas a enfermedades por priones pueden incluir SOD1 (superóxido dismutasa 1), ALS2 (esclerosis lateral amiotrófica 2), FUS (fusionado en sarcoma), TARDBP (proteína de unión de ADN TAR), VAGFA (factor A de crecimiento endotelial vascular), VAGFB (factor B de crecimiento endotelial vascular) y VAGFC (factor C de crecimiento endotelial vascular) y cualquier combinación de los mismos.

- 45 Los ejemplos de proteínas relacionadas con enfermedades neurodegenerativas en trastornos por priones pueden incluir A2M (Alfa-2-Macroglobulina), AATF (Factor de transcripción de antagonización de muerte celular programada), ACPP (fosfatasa ácida prostática), ACTA2 (Actina alfa 2 de músculo liso aorta), ADAM22 (dominio de la metalopeptidasa ADAM), ADORA3 (Receptor de adenosina A3) o ADRA1D (Receptor adrenérgico Alfa-1D para adrenoreceptor Alfa-1D), por ejemplo.

- 50 Los ejemplos de proteínas asociadas a Inmunodeficiencia pueden incluir A2M [alfa-2-macroglobulina]; AANAT [arilalquilamina N-acetiltransferasa]; ABCA1 [casete de unión a ATP, sub-familia A (ABC1), miembro 1]; ABCA2 [casete de unión a ATP, sub-familia A (ABC1), miembro 2] o ABCA3 [casete de unión a ATP, sub-familia A (ABC1), miembro 3]; por ejemplo.

Ejemplos de proteínas asociadas a Trastornos de Repetición del Trinucleótido incluyen AR (receptor de andrógeno),

FMR1 (retraso mental 1 frágil X), HTT (huntingtina) o DMPK (distrofia miotónica-proteína cinasa), FXN (frataxina), ATXN2 (ataxina 2), por ejemplo.

Ejemplos de proteínas asociadas a Trastornos de Neurotransmisión incluyen SST (somatostatina), NOS1 (óxido nítrico sintasa 1 (neuronal)), ADRA2A (receptor alfa-2A- adrenérgico), ADRA2C (receptor alfa-2C- adrenérgico), TACR1 (receptor de taquicína 1) o HTR2c (receptor 2C de 5-hidroxitriptamina (serotonina)), por ejemplo.

Ejemplos de secuencias asociadas al neurodesarrollo incluyen A2BP1 [proteína 1 de unión de ataxina 2], AADAT [aminoadipato aminotransferasa], AANAT [arilalquilamina N-acetiltransferasa], ABAT [4-aminobutirato aminotransferasa], ABCA1 [casete de unión a ATP, sub-familia A (ABC1), miembro 1] o ABCA13 [casete de unión a ATP, sub-familia A (ABC1), miembro 13], por ejemplo.

Más ejemplos de trastornos preferidos tratables con el presente sistema se pueden seleccionar de: Síndrome de Aicardi-Goutières; Enfermedad de Alexander; Síndrome de Allan-Herndon-Dudley; Trastornos Relacionados con POLG; Alfa-Manosidosis (Tipo II y III); Síndrome de Alström; Angelman; Síndrome; Ataxia-Telangiectasia; Lipofuscinosis Ceroide Neuronal; Beta-Talasemia; Atrofia Óptica Bilateral y Atrofia Óptica (Infantil) Tipo 1; Retinoblastoma (bilateral); Enfermedad de Canavan; Síndrome 1 Cerebroculofacioesquelético [COFS1]; Xantomatosis Cerebrotendinosa; Síndrome de Cornelia de Lange; Trastornos Relacionados con MAPT; Enfermedades Genéticas por Priones; Síndrome de Dravet; Enfermedad de Alzheimer Familiar de Comienzo Temprano; Ataxia de Friedreich [FRDA]; Síndrome de Fyryns; Fucosidosis; Distrofia Muscular Congénita de Fukuyama; Galactosialidosis; Enfermedad de Gaucher; Acidemias Orgánicas; Linfocitosis Hemofagocítica; Síndrome de Progeria de Hutchinson-Gilford; Mucopolidosis II; Enfermedad de Almacenamiento de Ácido Siálico Libre Infantil; Neurodegeneración Asociada a PLA2G6; Síndrome de Jervell y Lange-Nielsen; Epidermolisis Bullosa sobre las Articulaciones; Enfermedad de Huntington; Enfermedad de Krabbe (Infantil); Síndrome de Leigh Asociado al ADN Mitocondrial y NARP; Síndrome de Lesch-Nyhan; Lisencefalia Asociada a LIS1; Síndrome de Lowe; Enfermedad Urinaria de Jarabe de Maple; Síndrome de Duplicación de MECP2; Trastornos de Transporte de Cobre Relacionado con ATP7A; Distrofia Muscular Relacionada con LAMA2; Deficiencia de Arilsulfatasa A; Mucopolisacaridosis Tipos I, II o III; Trastornos de la Biogénesis del Peroxisoma, Espectro del Síndrome de Zellweger; Neurodegeneración con Trastornos de Acumulación de Hierro en el Cerebro; Deficiencia de Esfingomielinasa Ácida; Enfermedad de Niemann-Pick Tipo C; Encefalopatía por Glicina; Trastornos Relacionados con ARX; Trastornos del Ciclo de la Urea; Osteogénesis Imperfecta Relacionada con COL1A1/2; Síndrome de Supresión de ADN Mitocondrial; Trastornos Relacionados con PLP1; Síndrome de Perry; Síndrome de Phelan-McDermid; Enfermedad de Almacenamiento de Glucógeno Tipo II (Enfermedad de Pompe) (Infantil); Trastornos Relacionados con MAPT; Trastornos Relacionados con MECP2; Condrodisplasia Punctata Rizomélica Tipo 1; Síndrome de Roberts; Enfermedad de Sandhoff; Enfermedad de Schindler - Tipo 1; Deficiencia de Adenosina Deaminasa; Síndrome de Smith-Lemli-Opitz; Atrofia Muscular Espinal; Ataxia Espinocerebelar de Comienzo Infantil; Deficiencia de Hexosaminidasa A; Displasia Tanatófórica de Tipo 1; Trastornos Relacionados con el Colágeno Tipo VI; Síndrome de Usher Tipo I; Distrofia Muscular Congénita; Síndrome de Wolf-Hirschhorn; Deficiencia de Lipasa Ácida Lisosomal y Xeroderma Pigmentosa.

La administración crónica de terapéutica proteínica puede provocar respuestas inmunitarias inaceptables a la proteína específica. La inmunogenicidad de fármacos de proteínas puede ser atribuida a algunos epítomos de linfocitos de auxiliar T inmunodominante (HTL, por sus siglas en inglés). Reducir la afinidad de unión de MHC de estos epítomos HTL contenidos dentro de estas proteínas puede generar fármacos con inmunogenicidad menor (Tangri S, et al. ("Rationally engineered therapeutic proteins with reduced immunogenicity" J Immunol. 15 de marzo de 2.005; 174(6): 3.187-96.) En la presente descripción, la inmunogenicidad de la enzima CRISPR en particular se puede reducir siguiendo la propuesta explicada primero en Tangri et al con respecto a eritropoyetina y desarrollada con posterioridad. De acuerdo con esto, se puede usar la evolución directa o el diseño racional para reducir la inmunogenicidad de la enzima CRISPR (por ejemplo una Cas9) en la especie huésped (ser humano u otras especies).

En las plantas, los patógenos son con frecuencia específicos del huésped. Por ejemplo, *Fusarium oxysporum* f. sp. *lycopersici* produce marchitamiento del tomate pero ataca sólo al tomate y *F. oxysporum* f. *dianthii* *Puccinia graminis* f sp. *tritici* ataca sólo al trigo. Las plantas tienen defensas existentes e inducidas para resistir a la mayoría de los patógenos. Las mutaciones y casos de recombinación a lo largo de las generaciones de plantas conducen a variabilidad genética que da lugar a susceptibilidad, especialmente cuando se reproducen los patógenos con más frecuencia que las plantas. En las plantas puede haber resistencia al no huésped, por ejemplo, el huésped y el patógeno son incompatibles. También puede ser Resistencia Horizontal, por ejemplo, resistencia parcial contra todas las razas de un patógeno, controlado típicamente por muchos genes y Resistencia Vertical, por ejemplo, resistencia completa a algunas razas de un patógeno pero no a las demás razas, típicamente controlado por algunos genes. En un nivel gen por gen, las plantas y los patógenos evolucionan juntos y los cambios genéticos en uno equilibran los cambios en otro. De acuerdo con esto, usando Variabilidad Natural, los criadores combinan la mayoría de los genes útiles para Rendimiento, Calidad, Uniformidad, Dureza, Resistencia. Las fuentes de genes de la resistencia incluyen Variedades naturales o extrañas, Variedades Heirloom, Mutaciones Relativas a las Plantas Silvestres e Inducidas, por ej., tratando material vegetal con agentes mutagénicos. Usando la presente invención, se proporciona a los criadores de plantas una nueva herramienta para inducir mutaciones. De acuerdo con esto, un experto en la materia puede analizar el genoma de fuentes de genes de resistencia y en Variedades con

características o rasgos deseados emplea la presente invención para inducir el aumento de genes de resistencia, con más precisión que los agentes mutagénicos previos y por lo tanto acelera y mejora los programas de cultivo.

Como será evidente, se prevé que el presente sistema se pueda usar para fijar como objetivo cualquier secuencia de polinucleótidos de interés. Algunos ejemplos de afecciones o enfermedades que se podían tratar positivamente usando el presente sistema se incluyen en las Tablas anteriores y también se proporcionan allí ejemplos de genes asociados en la actualidad a esas enfermedades. Sin embargo, los genes ejemplificados no son exhaustivos.

Ejemplos

Los siguientes ejemplos se proporcionan con fines ilustrativos de diversas realizaciones de la invención y no están destinados a limitar la presente invención de ningún modo. Los presentes ejemplos, junto con los métodos descritos en la presente memoria son representativos en el momento presente de realizaciones preferidas, son ejemplares, y no se destinan a limitar el alcance de la invención.

Ejemplo 1: Actividad del complejo CRISPR en el núcleo de una célula eucariota.

Un sistema CRISPR de ejemplo de tipo II es el sitio CRISPR de tipo II de *Streptococcus pyogenes* SF370, que contiene una agrupación de cuatro genes Cas9, Cas1, Cas2 y Csn1, así como dos elementos de ARN no codificante, ARNtracr y una serie característica de secuencias repetitivas (repeticiones directas) inter espaciadas por tramos cortos de secuencias no repetitivas (espaciadores, aproximadamente 30 pb cada uno). En este sistema, se genera rotura de doble cadena (RDC) de ADN fijado como objetivo en cuatro etapas secuenciales (Figura 2A). Primero, dos ARN no codificantes, la serie pre- ARNcr y ARNtracr, se transcriben del sitio CRISPR. Segundo, ARNtracr se hibrida a las repeticiones directas de pre- ARNcr, que se tratan después en ARNcr maduros que contienen secuencias espaciadoras individuales. Tercero, el complejo ARNcr:ARNtracr maduro dirige Cas9 al ADN objetivo que consiste en el protoespaciador y la correspondiente PAM vía la formación de heterodúplex entre la región espaciadora del ARNcr y el ADN protoespaciador. Finalmente, Cas9 media la escisión de ADN objetivo aguas arriba de PAM para crear una RDC dentro del protoespaciador (Figura 2A). Este ejemplo describe un procedimiento de ejemplo para adaptar este sistema de nucleasa programable de ARN para dirigir la actividad del complejo CRISPR en el núcleo de las células eucariotas.

Para mejorar la expresión de los componentes CRISPR en células de mamífero, se codón-optimizaron dos genes del sitio 1 de SF370 de *Streptococcus pyogenes* (*S. pyogenes*), *Cas9* (*SpCas9*) y *RNasa III* (*SpRNasa III*). Para facilitar la localización nuclear, se incluyó una señal de localización nuclear en el término (N) amino o (C) carboxilo de ambas *SpCas9* y *SpRNasa III* (Figura 2B). Para facilitar la visualización de la expresión de proteínas, también se incluyó un marcador de proteína fluorescente en el N- o C-terminal de ambas proteínas (Figura 2B). También se generó una versión de *SpCas9* con un NLS unido a ambos N- y C-terminal (2xNLS-*SpCas9*). Se transfectaron construcciones que contenían *SpCas9* NLS-fusionada y *SpRNasa III* en células de riñón embrionario humano (HEK, por sus siglas en inglés) 293FT y se encontró el posicionamiento relativo de la NLS a *SpCas9* y *SpRNasa III* para afectar a su eficacia de localización nuclear. Mientras la NLS C-terminal fue suficiente para fijar como objetivo *SpRNasa III* para los núcleos, la unión de una sola copia de estas NLS particulares en cualquiera de los N- o C-terminal de *SpCas9* no pudo conseguir la localización nuclear adecuada en este sistema. En este ejemplo, la NLS C-terminal fue la de nucleoplasmina (KRPAATKKAGQAKKKK) y la NLS C-terminal fue la del antígeno T grande SV40 (PKKKRKV). De las versiones de *SpCas9* ensayadas, sólo 2xNLS-*SpCas9* presentó localización nuclear (Figura 2B).

La secuencia ARNtracr del sitio CRISPR de *S. pyogenes* SF370 presenta dos sitios de comienzo transcripcional, dando lugar a dos transcritos de 89 nucleótidos (nt) y 171 nt que se tratan con posterioridad en ARNtracr maduros de 75 nt idénticos. Se seleccionó el ARNtracr de 89 nt más corto para expresión en células de mamífero (construcciones de expresión ilustradas en la Figura 6, con neutralidad cuando se determina por los resultados del ensayo de Surveryor mostrado en la Figura 6B). Los sitios de comienzo de la transcripción se señalan como +1 y el terminador de la transcripción y la secuencia probada por el ensayo de northern se indican también. La expresión de ARNtracr tratado también se confirmó por ensayo de Northern. La Figura 7C muestra los resultados de un ensayo de Northern de ARN total extraído de células 293FT transfectadas con construcciones de expresión U6 que soportan ARNtracr largo o corto, así como *SpCas9* y DR-EMX1(1)-DR. Los paneles izquierdo y derecho son de células 293FT transfectadas sin o con *SpRNasa III*, respectivamente. U6 indica control de carga representado gráficamente con una sonda que fija como objetivo ARNsn U6 humano. La transfección de la construcción de expresión de ARNtracr corto conduce a niveles abundantes de la forma tratada de ARNtracr (~75 pb). Se detectan cantidades muy bajas de ARNtracr largo en el análisis de Northern.

Para activar la iniciación precisa transcripcional, se seleccionó el activador de U6 basado en ARN polimerasa III para conducir la expresión de ARNtracr (Figura 2C). De manera similar, se desarrolló una construcción a base de activador U6 para expresar una serie pre-ARNcr que consiste en un único espaciador flanqueado por dos repeticiones directas (las RD, también incluida por el término "secuencias de emparejamiento tracr"; Figura 2C). Se diseñó el espaciador inicial para fijar como objetivo un sitio objetivo de 33 pares de bases (pb) (protoespaciador de 30 pb más una secuencia de unidad (PAM) CRISPR de 3 pb que satisface la unidad de reconocimiento NGG de Cas9) en el sitio *EMX1* humano (Figura 2C), un gen clave en el desarrollo de la corteza cerebral.

Para ensayar si la expresión heteróloga del sistema CRISPR (SpCas9, SpRNasa III, ARNtracr y pre-ARNcr) en células de mamífero puede conseguir la escisión fijada como objetivo de cromosomas de mamífero, se transfectaron células HEK 293FT con combinaciones de componentes CRISPR. Puesto que las RDC en núcleos de mamífero se reparan parcialmente por la ruta de unión de extremos no homólogos (NHEJ, por sus siglas en inglés), que conduce a la formación de inserciones o supresiones, se usó el ensayo Surveyor para detectar potencial actividad de escisión en el sitio *EMX1* objetivo (véase por ej., Guschin *et al.*, 2.010, *Methods Mol Biol* 649: 247). La transfección conjunta de los cuatro componentes CRISPR pudo inducir una escisión hasta del 5,0% en el protoespaciador (véase la Figura 2D). La transfección conjunta de todos los componentes CRISPR menos SpRNasa III también indujo hasta 4,7% de inserción o supresión en el protoespaciador, que sugiere que puede haber RNasas de mamífero endógenas que son capaces de ayudar a la maduración de ARNcr, tal como por ejemplo las enzimas Dicer y Drosha relacionadas. Retirar cualquiera de los tres componentes restantes anula la actividad de escisión del genoma del sistema CRISPR (Figura 2D). La secuenciación Sanger de amplicones que contiene el sitio fijado como objetivo verificó la actividad de escisión: en 43 clones secuenciados, se encontraron 5 alelos mutados (11,6%). Experimentos similares usando una variedad de secuencias guía produjeron porcentajes de inserción o supresión tan altos como 29% (véanse las Figuras 4-8, 10 y 11). Estos resultados definen un sistema de tres componentes para modificación de genomas mediada por CRISPR eficaz en células de mamífero.

Para optimizar la eficacia de escisión, los solicitantes también ensayaron si diferentes isoformas de ARNtracr afectaban a la eficacia de la escisión y encontraron que, en este sistema de ejemplo, sólo la forma transcrita corta (89 pb) era capaz de mediar la escisión del sitio genómico *EMX1* humano. La Figura 9 proporciona un análisis de Northern adicional de tratamiento de ARNcr en células de mamífero. La Figura 9A ilustra un esquema que muestra el vector de expresión para un único espaciador flanqueado por dos repeticiones directas (RD-*EMX1*(1)- RD). El espaciador de 30 pb que fija como objetivo el protoespaciador 1 del sitio *EMX1* humano y las secuencias de repetición directa se muestran en la secuencia debajo de la Figura 9A. La línea indica la región cuya secuencia de complemento inverso se usó para generar sondas de Northern para detección de ARNcr de *EMX1*(1). La Figura 9B muestra un análisis Northern de ARN total extraído de células 293FT transfectadas con construcciones de expresión U6 que soportan RD-*EMX1*(1)- RD. Los paneles izquierdo y derecho son de células 293FT transfectadas sin o con SpRNasa III respectivamente. Se trató RD-*EMX1*(1)- RD en ARNcr maduros sólo en presencia de SpCas9 y ARNtracr corto y no fue dependiente de la presencia de SpRNasa III. El ARNcr maduro detectado de ARN total de 293FT transfectado es ~33 pb y es más corto que el ARNcr maduro de 39-42 pb de *S. pyogenes*. Estos resultados demuestran que un sistema CRISPR se puede trasplantar en células eucariotas y reprogramar para facilitar la escisión de polinucleótidos objetivo de mamíferos endógenos.

La Figura 2 ilustra el sistema CRISPR bacteriano descrito en este ejemplo. La Figura 2A ilustra un esquema que muestra el sitio 1 de CRISPR de *Streptococcus pyogenes* SF370 y un mecanismo propuesto de escisión de ADN mediada por CRISPR por este sistema. El ARNcr maduro tratado de la serie repetición directa – espaciador dirige Cas9 a objetivos genómicos que constan de protoespaciadores complementario y una unidad adyacente al protoespaciador (PAM). En el emparejamiento de base objetivo-espaciador, Cas9 media una rotura de doble cadena en el ADN objetivo. La Figura 2B ilustra el logro de Cas9 de *S. pyogenes* (SpCas9) y RNasa III (SpRNasa III) con señales de localización nuclear (las NLS) para permitir la importación al núcleo del mamífero. La Figura 2C ilustra la expresión de mamífero de SpCas9 y SpRNasa III conducida por el activador EF1a constitutivo y ARNtracr y la serie pre-ARNcr (RD-Espaciador-RD) conducidos por el activador U6 de ARN Pol3 para activar iniciación y terminación de transcripción precisa. Un protoespaciador del sitio *EMX1* humano con una secuencia PAM satisfactoria se usa como el espaciador en la serie pre-ARNcr. La Figura 2D ilustra ensayo de la nucleasa surveyor para inserciones y supresiones minoritarias mediadas por SpCas9. SpCas9 se expresó con y sin SpRNasa III, ARNtracr y una serie pre-ARNcr que soporta el espaciador objetivo *EMX1*. La Figura 2E ilustra una representación esquemática de emparejamiento de bases entre el sitio objetivo y ARNcr que fija como objetivo *EMX1*, así como un cromatograma de ejemplo que muestra una microsupresión adyacente al sitio de escisión de SpCas9. La Figura 2F ilustra alelos mutados identificados de análisis de secuenciación de 43 amplicones clonales que muestran una variedad de microinserciones y supresiones. Las líneas discontinuas indican bases suprimidas y bases no alineadas o desajustadas indican inserciones o mutaciones. Barra de escala = 10 µm.

Para simplificar además el sistema de tres componentes, se adaptó un diseño híbrido de ARNcr-ARNtracr quimérico, donde un ARNcr maduro (que comprende una secuencia guía) se fusiona a un ARNtracr parcial vía un tallo-lazo para imitar el dúplex ARNcr:ARNtracr natural (Figura 3A).

Las secuencias guía se pueden insertar entre sitios BbsI usando oligonucleótidos hibridados. Los protoespaciadores en las hebras transcrita y complementaria se indican por encima y por debajo de las secuencias de ADN, respectivamente. Se consiguió una tasa de modificación de 6,3% y 0,75% para los sitios *PVALB* humano y *Th* de ratón, respectivamente, que demuestra la amplia aplicabilidad del sistema CRISPR en la modificación de diferentes sitios a través de múltiples organismos. Aunque la escisión sólo se detectó con uno de tres espaciadores para cada sitio usando las construcciones quiméricas, todas las secuencias objetivo se escindieron con eficacia de producción de inserción o supresión que alcanzó el 27% cuando se usó la disposición de pre-ARNcr expresada conjuntamente (Figuras 4 y 5).

La Figura 5 proporciona una ilustración más de que SpCas9 puede ser reprogramado para fijar como objetivo múltiples sitios genómicos en células de mamífero. La Figura 5A proporciona un esquema del sitio *EMX1* humano

que muestra la posición de cinco protoespaciadores, indicado por las secuencias subrayadas. La Figura 5B proporciona un esquema del complejo pre-ARNcr/ARNtracr que muestra la hibridación entre la región de repetición directa del pre-ARNcr y ARNtracr (parte superior) y un esquema de un diseño de ARN quimérico que comprende una secuencia guía de 20 pb y secuencias de emparejamiento tracr y tracr que consisten en secuencias de repetición directa parcial y ARNtracr hibridadas según la estructura de horquilla (fondo). Los resultados de un ensayo Surveyor comparando la eficacia de escisión mediada por Cas9 en cinco protoespaciadores en el sitio *EMX1* humano se ilustra en la Figura 5C. Cada protoespaciador se fija como objetivo usando complejo pre-ARNcr/ARNtracr tratado (ARNcr) o ARN quimérico (ARNqui).

Puesto que la estructura secundaria de ARN puede ser crucial para interacciones intermoleculares, se usó un conjunto de algoritmo de predicción de estructuras basado en energía libre mínima y estructura pesada de Boltzmann para comparar la estructura secundaria putativa de todas las secuencias guía usadas en nuestro experimento de fijación como objetivo de genoma (Figura 3B) (véase por ej., Gruber et al., 2.008, Nucleic Acids Research, 36: W70). El análisis reveló que en la mayoría de los casos, las secuencias guía eficaces en el contexto de ARNcr quimérico estaban sustancialmente exentas de unidades de estructura secundaria, mientras las secuencias guía ineficaces era más probable que formaran estructuras secundarias internas que podían evitar el emparejamiento de bases con el ADN de protoespaciador objetivo. Así es posible que la variabilidad en la estructura secundaria del espaciador puede impactar en la eficacia de la interferencia mediada por CRISPR cuando se usa un ARNcr quimérico.

La Figura 3 ilustra vectores de expresión de ejemplo. La Figura 3A proporciona un esquema de un vector bicistónico para conducir la expresión de una quimera ARNcr-ARNtracr sintética (ARN quimérico) así como SpCas9. El ARN guía quimérico contiene una secuencia guía de 20 pb correspondiendo al protoespaciador en el sitio objetivo genómico. La Figura 3B proporciona un esquema que muestra secuencias guía que fijan como objetivo los sitios *EMX1*, *PVALB* humano y *Th* de ratón, así como sus estructuras secundarias previstas. La eficacia de la modificación en cada sitio objetivo se indica debajo del dibujo de la estructura secundaria del ARN (*EMX1*, $n = 216$ lecturas de secuenciación de amplicones; *PVALB*, $n = 224$ lecturas; *Th*, $n = 265$ lecturas). El algoritmo de plegamiento produjo una salida con cada base coloreada de acuerdo con su probabilidad de asumir la estructura secundaria prevista, como se indica por una escala en arco iris que se reproduce en la Figura 3B en escala de grises. Más diseños de vectores para SpCas9 se presentan en la Figura 3A, incluyendo vectores únicos de expresión que incorporan un activador U6 ligado a un sitio de inserción para un oligo guía y un activador Cbh ligado a secuencia de codificación SpCas9.

Para ensayar si los espaciadores que contienen estructuras secundarias son capaces de funcionar en células procariotas en las cuales los CRISPR operan de manera natural, se ensayó interferencia de transformación de plásmidos que soportan protoespaciadores en una cepa *E. coli* que expresa de manera heteróloga el sitio 1 CRISPR SF370 de *S. pyogenes* (Figura 3C). Se clonó el sitio CRISPR en un vector de expresión *E. coli* de copia baja y se reemplazó la serie ARNcr con un espaciador único flanqueado por un par de las RD (pCRISPR). Se transformaron cepas *E. coli* que albergaban diferentes plásmidos pCRISPR con plásmidos cuestionados conteniendo el correspondiente protoespaciador y secuencias PAM (Figura 3C). En el ensayo de bacterias, todos los espaciadores facilitaron interferencia de CRISPR eficaz (Figura 3C). Estos resultados sugieren que puede haber factores adicionales que afecten a la eficacia de la actividad de CRISPR en células de mamífero.

Para investigar la especificidad de la escisión mediada por CRISPR, se analizó el efecto de mutaciones de único nucleótido en la secuencia guía sobre la escisión del protoespaciador en el genoma de mamífero usando una serie de ARNcr quiméricos que fijan como objetivo *EMX1* con mutación de puntos únicos (Figura 4A). La Figura 4B ilustra los resultados de un ensayo de nucleasa Surveyor comparando la eficacia de la escisión de Cas9 cuando se empareja con diferentes ARN quiméricos mutantes. El desajuste de bases únicas hasta 12 pb 5' de la PAM sustancialmente anuló la escisión genómica por SpCas9, mientras los espaciadores con mutaciones en posiciones aguas arriba más lejos retuvieron la actividad frente al objetivo del protoespaciador original (Figura 4B). Además de la PAM, SpCas9 presenta especificidad de base única dentro de los últimos 12 pb del espaciador. Además, CRISPR es capaz de mediar la escisión genómica tan eficazmente como un par de nucleasas TALE (TALEN) fijando como objetivo el mismo protoespaciador *EMX1*. La Figura 4C proporciona un esquema que muestra el diseño de las TALEN que fijan como objetivo *EMX1* y la Figura 4D muestra un gel Surveyor comparando la eficacia de TALEN y Cas9 ($n=3$).

Teniendo establecida una serie de componentes para conseguir edición de genes mediada por CRISPR en células de mamífero a través del mecanismo NHEJ con predisposición a errores, se ensayó la capacidad de CRISPR para estimular la recombinación homóloga (RH), una ruta de reparación de genes de alta fidelidad para preparar mediciones precisas en el genoma. SpCas9 de tipo natural es capaz de mediar las RDC específicas del sitio, que se pueden reparar a través de tanto NHEJ como RH. Además, se logró una sustitución de aspartato a alanina (D10A) en el dominio catalítico RuvC I de SpCas9 para convertir la nucleasa en una nickasa (SpCas9n; ilustrado en la Figura 5A) (véase por ej., Sapranasaks et al., 2.011, Nucleic Acids Research, 39: 9.275; Gasiunas et al., 2.012, Proc. Natl. Acad. Sci. USA, 109: E2579), de manera que el ADN genómico nickeado experimenta la reparación dirigida por homología de alta fidelidad (HDR, por sus siglas en inglés). El ensayo Surveyor confirmó que SpCas9n no generaba inserciones o supresiones en el objetivo del protoespaciador *EMX1*. Como se ilustra en la Figura 5B, la expresión conjunta de ARNcr quimérico que fija como objetivo *EMX1* con SpCas9 produjo inserciones o supresiones en el sitio

fijado como objetivo, mientras la expresión conjunta con SpCas9n no lo hizo (n=3). Por otra parte, la secuenciación de 327 amplicones no detectó ninguna inserción o supresión inducida por SpCas9n. Se seleccionó el mismo sitio para ensayar RH mediada por CRISPR por transfección conjunta de células HEK 293FT con el ARN quimérico que fija como objetivo *EMX1*, hSpCas9 o hSpCas9n, así como una plantilla de RH para introducir un par de sitios de restricción (*HindIII* y *NheI*) cerca del protoespaciador. La Figura 5C proporciona una ilustración en esquema de la estrategia de la RH, con posiciones relativas de puntos de recombinación y secuencias de hibridación de cebadores (flechas). SpCas9 y SpCas9n por supuesto catalizaron la integración de la plantilla de RH en el sitio *EMX1*. La multiplicación por PCR de la región fijada como objetivo seguida por digestión de restricción con *HindIII* reveló productos de escisión correspondiendo a tamaños de fragmentos esperados (flechas en análisis de gel de polimorfismo de longitud del fragmento de restricción mostrado en la Figura 5D), mediando SpCas9 y SpCas9n niveles similares de eficacias de la RH. Los solicitantes verificaron además RH usando secuenciación de Sanger de amplicones genómicos (La Figura 5E). Estos resultados demuestran la utilidad de CRISPR para facilitar la inserción de genes fijados como objetivo en el genoma del mamífero. Dada la especificidad objetivo de 14 pb (12 pb del espaciador y 2 pb de la PAM) de SpCas9 de tipo natural, la disponibilidad de una nickasa puede reducir significativamente la probabilidad de modificaciones fuera de lo proyectado, puesto que las roturas de hebra única no son sustratos para la ruta NHEJ con predisposición a errores.

Las construcciones de expresión que imitan la arquitectura natural de sitios CRISPR con espaciadores formados (Figura 2A) se construyeron para ensayar la posibilidad de fijar como objetivo secuencias multiplexadas. Usando una única serie CRISPR codificando un par de espaciadores que fijan como objetivo *EMX1* y *PVALB*, se detectó la escisión eficaz en ambos sitios (Figura 4F, que muestra tanto un diseño esquemático de la serie ARNcr como un análisis Surveyor mostrando la mediación eficaz de la escisión). La supresión fijada como objetivo de regiones genómicas mayores a través de las RDC usando espaciadores frente a dos objetivos dentro de *EMX1* espaciado por 119 pb también se ensayó y se detectó una eficacia de la supresión del 1,6% (3 de un total de 182 amplicones; Figura 5G). Esto demuestra que el sistema CRISPR puede mediar la edición multiplexada dentro de un único genoma.

Ejemplo 2: Modificaciones y alternativas del sistema CRISPR.

La capacidad para usar ARN para programar escisión de ADN específica de la secuencia define una nueva clase de herramientas de ingeniería del genoma para una variedad de aplicaciones de investigación e industriales. Diversos aspectos del sistema CRISPR se pueden mejorar además para aumentar la eficacia y la versatilidad de la fijación como objetivo de CRISPR. La actividad óptima de Cas9 puede depender de la disponibilidad de Mg^{2+} libre en niveles mayores que los presentes en el núcleo de los mamíferos (véase por ej., Jinek et al., 2012, Science, 337: 816) y la preferencia para una unidad de NGG inmediatamente aguas abajo del protoespaciador restringe la capacidad para fijar como objetivo de promedio cada 12 pb en el genoma humano. Algunas de estas restricciones se pueden superar explorando la diversidad de los sitios CRISPR a través del metagenoma microbiano (véase por ej., Makarova et al., 2011, Nat Rev Microbiol, 9: 467). Otros sitios CRISPR pueden ser trasplantados en el medio celular de los mamíferos mediante un procedimiento similar al descrito en el Ejemplo 1. La eficacia de la modificación en cada sitio fijado como objetivo se indica debajo de las estructuras secundarias de ARN. El algoritmo que genera las estructuras colorea cada base de acuerdo con su probabilidad de asumir la estructura secundaria prevista. Los espaciadores 1 y 2 guía de ARN indujeron 14% y 6,4%, respectivamente. El análisis estadístico de la actividad de escisión a través de réplicas biológicas en estos dos sitios del protoespaciador se proporciona también en la Figura 7.

Ejemplo 3: Algoritmo de selección de secuencias objetivo de muestra.

Se diseña un programa informático para identificar secuencias objetivo de CRISPR candidato en ambas cadenas de una secuencia de ADN de entrada basada en longitud deseada de secuencias guía y una secuencia de unidad CRISPR (PAM) para una enzima CRISPR especificada. Por ejemplo, los sitios objetivo para Cas9 de *S. pyogenes*, con secuencias NGG de PAM, se pueden identificar por investigación para 5'-N_x-NGG-3' ambas en la secuencia de entrada y en el complemento inverso de la entrada. Asimismo los sitios objetivo para Cas9 de CRISPR1 de *S. thermophilus*, con secuencia NNAGAAW de PAM, se pueden identificar por investigación para 5'-N_x-NNAGAAW-3' ambos en la secuencia de entrada y en el complemento inverso de la entrada. Asimismo, los sitios objetivo para Cas9 de CRISPR3 de *S. thermophilus*, con secuencia NGGNG de PAM, se pueden identificar por investigación para 5'-N_x-NGGNG-3' ambos en la secuencia de entrada y en el complemento inverso de la entrada. El valor "x" en N_x puede ser fijado por el programa o especificado por el usuario, tal como 20.

Puesto que múltiples apariciones en el genoma del sitio objetivo de ADN pueden conducir a edición de genoma no específica, después de identificar todos los sitios potenciales, el programa filtra secuencias basadas en el número de veces que aparecen en el genoma de referencia relevante. Para esas enzimas CRISPR para las cuales se determina la especificidad de secuencia mediante una secuencia "simiente", tal como la 5' de 11-12 pb de la secuencia PAM, incluyendo la propia secuencia PAM, la etapa de filtración puede estar basada en la secuencia simiente. Así, para evitar la edición en sitios genómicos adicionales, los resultados se filtran basándose en el número de apariciones de la secuencia simiente:PAM en el genoma relevante. Se puede permitir que el usuario elija la longitud de la secuencia simiente. También se puede permitir que el usuario especifique el número de apariciones de la secuencia simiente:PAM en un genoma para fines de pase del filtro. El defecto es investigar secuencias

únicas. El nivel de filtración se modifica cambiando tanto la longitud de la secuencia simiente como el número de apariciones de la secuencia en el genoma. El programa puede proporcionar además o como alternativa la secuencia de una secuencia guía complementaria a la secuencia o secuencias objetivo indicadas proporcionando el complemento inverso de la secuencia o secuencias objetivo identificadas.

5 Ejemplo 4: Evaluación de híbridos ARNcr-ARNtracr quiméricos múltiples.

Este ejemplo describe los resultados obtenidos para los ARN quiméricos (los ARNqui; que comprenden una secuencia guía, una secuencia de emparejamiento tracr y una secuencia tracr en un transcrito único) que tienen secuencias tracr que incorporan diferentes longitudes de secuencia ARNtracr de tipo natural. La Figura 18a ilustra un esquema de un vector bicistrónico de expresión para ARN quimérico y Cas9. Cas9 es conducido por el activador CBh y el ARN quimérico es conducido por un activador U6. El ARN guía quimérico consta de una secuencia guía de 20 pb (Ns) unida a la secuencia tracr (barriendo desde la primera "U" de la hebra menor al extremo del transcrito), que se trunca en diversas posiciones como se indica. Las secuencias guía y tracr se separan mediante la secuencia de emparejamiento tracr GUUUUAGAGCUA seguido por la secuencia lazo GAAA. Los resultados de los ensayos SURVEYOR para inserciones o supresiones mediadas por Cas9 en los sitios *EMX1* y *PVALB* humanos se ilustran en la Figura 18b y 18c, respectivamente. Las flechas indican los fragmentos SURVEYOR esperados. Los ARNqui se indican por su designación "+n" y ARNcr se refiere a un ARN híbrido donde las secuencias guía y tracr se expresan como transcritos separados. La cuantificación de estos resultados, realizados por triplicado, se ilustran por histograma en las Figuras 11a y 11b, que corresponden a las Figuras 10b y 10c, respectivamente ("N. D." indica no inserciones o supresiones detectadas). Los ID del protoespaciador y su correspondiente objetivo genómico, secuencia del protoespaciador, secuencia PAM y posición de la cadena se proporcionan en la Tabla D. Las secuencias guía se diseñan para que sean complementarias a la secuencia completa del protoespaciador en el caso de transcritos separados en el sistema híbrido o sólo a la porción subrayada en el caso de ARN quiméricos.

Tabla D:

ID del protoespaciador	objetivo genómico	secuencia del protoespaciador (5' a 3')	PAM	Hebra
1	<i>EMX1</i>	GGACATCGATGTCACCTCCAATGACTAG GG	TGG	+
2	<i>EMX1</i>	CATTGGAGGTGACATCGATGTCCTCCCC AT	TGG	-
3	<i>EMX1</i>	GGAAGGGCCTGAGTCCGAGCAGAAGAA GAA	GGG	+
4	<i>PVALB</i>	GGTGGCGAGAGGGGCGGAGATTGGGTGT TC	AGG	+
5	<i>PVALB</i>	ATGCAGGAGGGTGGCGAGAGGGGCCGA GAT	TGG	+

25 Cultivo celular y transfección

Se mantuvo la estirpe celular 293FT (Life Technologies) de riñón embrionario humano (HEK, por sus siglas en inglés) en Medio Eagle modificado de Dulbecco (DMEM) enriquecido con suero fetal bovino al 10% (HyClone), GlutaMAX 2 mM (Life Technologies), 100 U/ml de penicilina y 100 µg/ml de estreptomycin a 37°C con incubación en CO₂ al 5%. Se sembraron células 293FT en placas de 24 pozos (Corning) 24 horas previamente a transfección a una densidad de 150.000 células por pozo. Se transfectaron las células usando Lipofectamine 2000 (Life Technologies) siguiendo el protocolo recomendado por el fabricante. Para cada pozo de una placa de 24 pozos, se usó un total de 500 ng de plásmido.

Ensayo SURVEYOR para modificación de genomas

Se transfectaron células 293FT con ADN de plásmido como se describió anteriormente. Se incubaron las células a 37°C durante 72 horas post-transfección previamente a extracción de ADN genómico. Se extrajo ADN genómico usando la solución de extracción de ADN QuickExtract (Epicentre) siguiendo el protocolo del fabricante. En resumen, se volvieron a suspender las células del botón en solución QuickExtract y se incubaron a 65 °C durante 15 minutos y 98 °C durante 10 minutos. La región genómica que flanquea el sitio objetivo de CRISPR para cada gen se multiplicó por PCR (cebadores enumerados en la Tabla E) y se purificaron los productos usando una columna de QiaQuick Spin (Qiagen) siguiendo el protocolo del fabricante. Se mezclaron 400 ng totales de los productos PCR purificados con 2 µl 10X de tampón PCR ADN Taq Polimerasa (Enzymatics) y agua ultra pura a un volumen final de 20 µl y se sometieron a un procedimiento de rehibridación para permitir la formación de heteroduplex: 95°C durante 10 min, 95°C a 85°C descendiendo a -2°C/s, 85°C a 25°C a -0,25°C/s y 25°C mantenidos durante 1 minuto. Después de rehibridación, se trataron los productos con nucleasa SURVEYOR y estimulador S SURVEYOR (Transgenomics)

siguiendo el protocolo recomendado por el fabricante y se analizaron sobre geles de poliacrilamida Novex TBE al 4-20% (Life Technologies). Se tiñeron los geles con tinción de ADN SYBR Gold (Life Technologies) durante 30 minutos y se formaron imágenes con un sistema de formación de imágenes de gel Gel Doc (Bio-rad). La cuantificación estuvo basada en intensidades de banda relativas.

5 Tabla E:

nombre del cebador	objetivo genómico	secuencia del cebador (5' a 3')
Sp-EMX1-F	<i>EMX1</i>	AAAACCACCCTTCTCTCTGGC
Sp-EMX1-R	<i>EMX1</i>	GGAGATTGGAGACACGGAGA G
Sp-PVALB-F	<i>PVALB</i>	CTGGAAAGCCAATGCCTGAC
Sp-PVALB-R	<i>PVALB</i>	GGCAGCAAACCTCTTGTCTCT

Identificación computacional de sitios fijados como objetivo de CRISPR únicos

Para identificar sitios fijados como objetivo únicos para la enzima SF370 Cas9 (SpCas9) de *S. pyogenes* en genoma de ser humano, ratón, rata, pez cebra, mosca de la fruta y *C. elegans*, se desarrolló un paquete informático para escanear ambas hebras de una secuencia de ADN e identificar todos los posibles sitios objetivo de SpCas9. Para este ejemplo, cada sitio objetivo SpCas9 se definió operacionalmente como una secuencia de 20 pb seguido por una secuencia de unidad adyacente del protoespaciador (PAM) de NGG y se identificaron todas las secuencias que satisficieran esta definición de 5'-N₂₀-NGG-3' en todos los cromosomas. Para evitar la edición de genoma no específica, después de identificar todos los sitios potenciales, se filtraron todos los sitios objetivo basándose en el número de veces que aparecían en el genoma de referencia relevante. Para sacar provecho de la especificidad de la secuencia de la actividad de Cas9 conferida por una secuencia "simiente", que puede ser, por ejemplo, secuencia 5' de aproximadamente 11-12 pb de la secuencia PAM, se seleccionaron secuencias 5'-NNNNNNNNNN-NGG-3' para que fueran únicas en el genoma relevante. Todas las secuencias genómicas se descargaron de UCSC Genome Browser (genoma humano hg19, genoma de ratón mm9, genoma de rata m5, genoma de pez cebra danRer7, genoma dm4 de *D. melanogaster* y genoma ce10 de *C. elegans*). Los resultados de la investigación completa están disponibles para navegar usando la información de UCSC Genome Browser. Una visualización de ejemplo de algunos sitios objetivo en el genoma humano se proporciona en la Figura 22.

Inicialmente, se fijaron como objetivo tres sitios dentro del sitio de *EMX1* en células HEK 293FT humanas. Se evaluó la eficacia de la modificación del genoma de cada ARNqui usando el ensayo de la nucleasa SURVEYOR, que detecta mutaciones que resultan de las roturas de la doble cadena de ADN (las RDC) y su posterior reparación por la ruta de reparación de daños de ADN de unión del extremo no homólogo (NHEJ). Las construcciones designadas ARNqui(+n) indican que hasta el nucleótido +n de ARNtracr de tipo natural se incluye en la construcción de ARN quimérico, con valores de 48, 54, 67 y 85 usados para n. Los ARN quiméricos que contienen fragmentos mayores de ARNtracr de tipo natural (ARNqui (+67) y ARNqui (+85)) mediaron la escisión de ADN en los tres sitios objetivo de *EMX1*, demostrando ARNqui (+85) en particular niveles significativamente mayores de escisión de ADN que los correspondientes híbridos ARNcr/ARNtracr que expresaron secuencias guía y tracr en transcritos separados (Figuras 10b y 10a). También se fijaron como objetivo dos sitios en el sitio *PVALB* que no proporcionaron escisión detectable usando el sistema híbrido (secuencia guía y secuencia tracr expresadas como transcritos separados) usando ARNquis. ARNqui(+67) y ARNqui(+85) fueron capaces de mediar escisión significativa en los dos protoespaciadores *PVALB* (Figuras 10c y 10b).

Para los cinco objetivos en los *EMX1* y *PVALB*, se observó un incremento consistente en la eficacia de modificación del genoma con longitud creciente de la secuencia tracr. Sin desear estar limitados por ninguna teoría, la estructura secundaria formada por el extremo 3' de la ARNtracr puede desempeñar una función en la activación de la tasa de formación de complejo CRISPR. Una ilustración de las estructuras secundarias previstas para cada uno de los ARN quiméricos usados en este ejemplo se proporciona en la Figura 21. Se predijo la estructura secundaria usando ARNfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) usando energía libre mínima y algoritmo de función de partición. El pseudocolor para cada base (reproducido en escala de grises) indica la probabilidad de emparejamiento. Debido a que los ARNquis con secuencias tracr más largas eran capaces de escindir objetivos que no eran escindidos por híbridos ARNcr/ARNtracr de CRISPR naturales, es posible que se pueda cargar ARN quimérico en Cas9 más eficazmente que su contrapartida híbrida natural. Para facilitar la aplicación de Cas9 para edición de genoma específica del sitio en células y organismos eucariotas, todos los sitios objetivo únicos previstos para la Cas9 de *S. pyogenes* se identificaron de manera computacional en los genomas del ser humano, ratón, rata,

pez cebra, *C. elegans* y *D. melanogaster*. Se pueden diseñar ARN quiméricos para enzimas Cas9 de otros microbios para expandir el espacio fijado como objetivo de las nucleasas programables de ARN de CRISPR.

- 5 Las Figuras 11 y 21 ilustran vectores de expresión bicistronicos ejemplares para expresión de ARN quimérico incluyendo hasta el nucleótido +85 de secuencia de ARNtracr de tipo natural y SpCas9 con secuencias de localización nuclear. SpCas9 se expresa a partir de un activador CBh y termina con la señal poliA bGH (bGH pA). La
- 10 secuencia expandida ilustrada inmediatamente debajo del esquema corresponde a la región que rodea al sitio de inserción de la secuencia guía e incluye, desde 5' a 3', la porción 3' del activador U6 (primera región sombreada), sitios de escisión BbsI (flechas), repetición directa parcial (secuencia de emparejamiento tracr GTTTTAGAGCTA, subrayada), secuencia lazo GAAA y secuencia tracr +85 (secuencia subrayada después de la secuencia lazo). Un inserto de la secuencia guía ejemplar se ilustra debajo del sitio de la inserción de la secuencia guía, con nucleótidos de la secuencia guía para un objetivo seleccionado representado por una "N".

Las secuencias descritas en los ejemplos anteriores son como sigue (las secuencias de polinucleótidos son 5' a 3'):

ARNtracr corto U6 (*Streptococcus pyogenes* SF370):

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGA
GAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTG
ACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAAT
GGACTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTGGCTTTATATATC
TTGTGGAAAGGACGAAACACCGGAACCATTCAAAACAGCATAGCAAGTTAAAAT
AAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

- 15 (negrita = secuencia ARNtracr; subrayado = secuencia terminadora)

ARNtracr largo U6 (*Streptococcus pyogenes* SF370):
GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGA
GAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTG
ACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAAT
GGACTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTGGCTTTATATATC
TTGTGGAAAGGACGAAACACCGGTAGTATTAAGTATTGTTTTATGGCTGATAAATTT
CTTTGAATTTCTCCTTGATTATTTGTTATAAAAGTTATAAAATAATCTTGTTGGAACC
ATTCAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG
TGGCACCGAGTCGGTGCTTTTTTTT

- 20 U6-DR-BbsI cadena principal-RD (*Streptococcus pyogenes* SF370):
GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGA
GAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTG
ACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAAT
GGACTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTGGCTTTATATATC
TTGTGGAAAGGACGAAACACCGGGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAA
ACGGGTCTTCGAGAAGACGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAAC

U6-ARN quimérico-Cadena principal BbsI (*Streptococcus pyogenes* SF370)

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGA
GAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTG
ACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAT
GGACTATCATATGCTTACCGTAACTTGAAAAGTATTTGATTTCTTGGCTTTATATATC
TTGTGGAAAGGACGAAACACCGGGTCTTCGAGAAGACCTGTTTTAGAGCTAGAAAT
AGCAAGTTAAAATAAGGCTAGTCCG

NLS-SpCas9-EGFP:

MDYKDHGDYKDHIDYKDDDDKMAPKKRKVGIHGVPAAADKKYSIGLDIGTNSVG
WAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFHRLSESLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYH
LRKKLVSTDKADLRILIYLAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFE
ENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDL
AEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS
ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFDQSKNGYAGYIDGGASQEEFYKFIK
PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNR
EKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTN
FDKNLPNEKVLPHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILE
DIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG
ILQTVKVDELVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIEELGKELGSQI
LKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSID
NKVLTRSDKNRGSNDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSLLVSDFRKDF
QFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQE
IGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLFELE
NGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVLDTLHQSTGLYETRIDLSQLGGDAAAVSKGEELFTGVVPILV
ELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWPTLVTTLTYGVCFSRYP
DHMKQHDFFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKED
GNILGHKLEYNYNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGP
VLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYK

5

SpCas9-EGFP-NLS:

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
 EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIF
 GNIVDEVAYHEKYPTIYHLRKKLVDDSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNS
 DVDKLFQQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFG
 NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSD
 AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY
 AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL
 HAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEE
 VVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPA
 FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL
 KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVSGQG
 DSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQTTQKGQKN
 SRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
 YDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLIT
 QRKFDNLTKAERGGLSELDAKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE
 VKVITLKSGLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYG
 DYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
 VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPPK
 YGGFDSPTVAYSVLVVAKEKVGKSKKLKSVKELLGITIMERSSEKKNPIDFLEAKGYKE
 VKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKS
 PEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVL SAYNKHHRDKPIREQAENI
 IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGDAA
 AVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWP
 TLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG
 DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSV
 QLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMD
 ELYKKRPAATKKAGQAKKKK

NLS-SpCas9-EGFP-NLS:

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGIHGVPAADKKYSIGLDJGTNSVG
WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYH
LRKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFE
ENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDL
AEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS
ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK
PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNR
EKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTN
FDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILE
DIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQDLSHEHIANLAGSPAIKKG
ILQTVKVVDLVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGKELGSQI
LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHIVPQSFLKDDSID
NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDF
QFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQE
IGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIILPKYSLFELE
NGRKRMLASAGELQKGNEALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDAAAVSKGEELFTGVVPILV
ELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWPTLVTTLT YGVQCFSRYP
DHMKQHDFFSAMPEGYVQERTIFFKDDGNYKTRA EVKFEGDTLVNRIELKGIDFKED
GNILGHKLEYNYNSHN VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGP
VLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYKKRPAATKKAGQA
KKKK

NLS-SpCas9-NLS:

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGIHGVPAADKKYSIGLDIGTNSVG
WAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYH
LRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLF
ENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDL
AEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLS
ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK
PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNR
EKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTN
FDKNLPNEKVLPHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILE
DIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDLSHEHIANLAGSPAIKKG
ILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKNSRERMKRIEELGELGSKI
LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSID
NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDF
QFYKVRINNYHHAHDAYLNAVVGTLIKKYPKLESEFVYGDYKVYDVRKMIKSEQE
IGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIILPKYSLFELE
NGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKHRRDKPIREQAENIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVLDTLHQSIQGLYETRIDLSQLGGDKRPAATKKAGQAKKKK

NLS-mCherry-SpRNasa3:

MFLFLSLTSFLSSSRTLVSKEEDNMAIIEFMRFKVHMEGSVNGHEFEIEGEGEGRPYE
GTQTAKLKVTGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYKLKSFPEGFKWERVM
NFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPDGPVMQKKTMGWEASSERMYPED
GALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVE
QYERAEGRHSTGGMDELYKGSQLEELLSTSFDIQFNDLTLLTAFTHTSYANEHRLN
VSHNERLEFLGDAVLQLIIEYLFAYPKKTEGDMSKLRSMIVREESLAGFSRFSFDAYI
KLKGEEKSGGRRRDTILGDLFEAFLGALLLDKGIDAVRRFLKQVMIPQVEKGNFERVK
DYKTCLQEFLQTKGDVAIDYQVISEKGPAHAKQFEVSIVVNGAVLSKGLGKSKKLAEQD
AAKNALAEQSEV

5

SpRNasa3-mCherry-NLS:

MKQLEELLSTSFDIQFNDLTLLETAFTHTSYANEHRLNLSHNERLEFLGDAVLQLIIEY
LFAKYPKKTEGDMSKLRSMIVREESLAGFSRFSFDAYIKLGKGEEKSGGRRRDITLGD
FEAFLGALLLDKGIDAVRRFLKQVMIPQVEKGNFERVKDYKTCLQEFLQTKGDVAIDYQ
VISEKGPAHAKQFEVSIVVNGAVLSKGLGKSKKLAEQDAAKNALAQLSEVGSVSKGEE
DNMAIIEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTGGPLPFAWDIL
SPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFI
YKVKLRGTNFPDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLDGGHYDAE
VKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKKRP
AATKKAGQAKKKK

NLS-SpCas9n-NLS (la mutación de nickasa D10A es minúscula):

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGIIHGVPAAADKKYSIGLaIGTNSVGW
AVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKN
RICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
RKKLV DSTDKADRLRLIALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEE
NPINASGVDAKILSARLSKSRLENLIAQLPGEKKNLFGNLIASLGLTPNFKSNFDLA
EDAKLQLSKD TYDDDLNLLAQIGDQYADFLAAKNLSDAILSDILRVNTEITKAPLSA
SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI
LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREK
IEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFD
KNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE
DIVLTLTLFEDREMIERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKSDGFANRNFQMILHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKG
ILQTVKVVDELVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIEIGIKELGSQI
LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSID
NKVLTRSDKNRGKSDNPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDF
QFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQE
IGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLEFE
NGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVLDTLIHQSI TGLYETRIDLSQLGGDKRPAATKKAGQAKKKK

hEMX1-Plantilla

GAATGCTGCCCTCAGACCCGCTTCCTCCCTGTCCTTGTCTGTCCAAGGAGAATGAGG
TCTCACTGGTGGATTTCGGACTACCCTGAGGAGCTGGCACCTGAGGGACAAGGCCC
CCCACCTGCCCAGCTCCAGCCTCTGATGAGGGGTGGGAGAGAGCTACATGAGGTTG
CTAAGAAAGCCTCCCCTGAAGGAGACCACACAGTGTGTGAGGTTGGAGTCTCTAGC
AGCGGGTTCTGTGCCCCCAGGGATAGTCTGGCTGTCCAGGCACTGCTCTTGATATAA
ACACCACCTCCTAGTTATGAAACCATGCCATTCTGCCTCTCTGTATGGAAAAGAGC

RH-HindII-NheI:

ATGGGGCTGGCCCGTGGGGTGGTGTCCACTTTAGGCCCTGTGGGAGATCATGGGAA
 CCCACGCAGTGGGTCATAGGCTCTCTCATTTACTACTCACATCCACTCTGTGAAGAA
 GCGATTATGATCTCTCCTCTAGAAACTCGTAGAGTCCCATGTCTGCCGGCTTCCAGA
 GCCTGCACTCCTCCACCTTGGCTTGGCTTTGCTGGGGCTAGAGGAGCTAGGATGCAC
 AGCAGCTCTGTGACCCTTTGTTTGAGAGGAACAGGAAAACCACCCTTCTCTCTGGCC
 CACTGTGTCCTCTTCCTGCCCTGCCATCCCCTTCTGTGAATGTTAGACCCATGGGAGC
 AGCTGGTCAGAGGGGACCCCGGCCTGGGGCCCCCTAACCCCTATGTAGCCTCAGTCTTC
 CCATCAGGCTCTCAGCTCAGCCTGAGTGTTGAGGCCCCAGTGGCTGCTCTGGGGGCC
 TCCTGAGTTTCTCATCTGTGCCCCCTCCCTCCCTGGCCCAGGTGAAGGTGTGGTTCCAG
 AACC GGAGGACAAAGTACAAACGGCAGAAGCTGGAGGAGGAAGGGCCTGAGTCCG
 AGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCACGAAGCAG
 GCCAATGGGGAGGACATCGATGTACCTCCAATGACaagcttgctagcGGTGGGCAACCAC
 AAACCCACGAGGGCAGAGTGCTGCTTGCTGCTGGCCAGGCCCTGCGTGGGGCCAA
 GCTGGACTCTGGCCACTCCCTGGCCAGGCTTTGGGGAGGCCTGGAGTCATGGCCCCA
 CAGGGCTTGAAGCCCGGGGCCGCGCATTTGACAGAGGGACAAGCAATGGGCTGGCTGA
 GGCCTGGGACCACTTGGCCTTCTCCTCGGAGAGCCTGCCTGCCTGGGCGGGCCCCGCC
 CGCCACCGCAGCCTCCAGCTGCTCTCCGTGTCTCCAATCTCCCTTTTGTGTTTGATGC
 ATTTCTGTTTTAATTTATTTTCCAGGCACCACTGTAGTTTAGTGATCCCCAGTGTCCC
 CCTTCCCTATGGGAATAATAAAAGTCTCTCTCTTAATGACACGGGCATCCAGCTCCA
 GCCCCAGAGCCTGGGGTGGTAGATTCCGGCTCTGAGGGCCAGTGGGGGCTGGTAGA
 GCAAACGCGTTCAGGGCCTGGGAGCCTGGGGTGGGGTACTGGTGGAGGGGGTCAAG
 GGTAATTCATTAACCTCTCTTTTGTGGGGGACCCTGGTCTCTACCTCCAGCTCCA
 CAGCAGGAGAAACAGGCTAGACATAGGGAAGGGCCATCCTGTATCTTGAGGGAGGA
 CAGGCCCAGGTCTTTCTTAACGTATTGAGAGGTGGGAATCAGGCCCAGGTAGTTCAA
 TGGGAGAGGGAGAGTGCTTCCCTCTGCCTAGAGACTCTGGTGGCTTCTCCAGTTGAG
 GAGAAACCAGAGGAAAGGGGAGGATTGGGGTCTGGGGGAGGGAACACCATTACACA
 AAGGCTGACGGTTCCAGTCCGAAGTCGTGGGCCCCACCAGGATGCTCACCTGTCCTTG
 GAGAACCCTGGGCAGGTTGAGACTGCAGAGACAGGGCTTAAGGCTGAGCCTGCAA
 CCAGTCCCCAGTGACTCAGGGCCTCCTCAGCCCAAGAAAGAGCAACGTGCCAGGGC
 CCGCTGAGCTCTTGTGTTACCTG

NLS-StCsn1-NLS:

MKRPAAATKKAGQAKKKKSDLVLGLDIGIGSVGVGILNKVTGEIHKNSRJPAAQAENN
LVRRTNRQGRRLARRKKHRRVRLNRLFEESSLITDFTKISINLNPYQLRVKGLTDELSNE
ELFIALKNMVKHRGISYLLDDASDDGNSSVGDYAQIVKENSQKLETKTPGQIQLERYQTY
GQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKR
KYYHGPNGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLND
LNNLTVPPTETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKS
GKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLTNTEREGIQEALEHEFADGSFS
QKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSS
SNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKK
AIQIKQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLY
TGKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQALDSMD
DAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNA
LQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHVAVDALIAASSQLNLWKK
QKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSILFSYQVDSK
FNRKISDATIYATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMY
RHDPQTFEKVIEPILENYPNKQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNQPEIKSLK
YYDSKLGNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEILGLKYADLQFEKGT
GTYKISQEKYNDIKKKEGVSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPKQK
HYVELKPYDKQKFEGGEALIKVLGNVANSQGCKKGLGKSNISIYKVRTDVLGNQHIKN
EGDKPKLDFKRPAATKKAGQAKKKK

U6-St ARNtracr(7-97):

GAGGGCCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGA
GAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTG
ACGTAGAAAAGTAATAATTTCTTGCGTAGTTTGCAGTTTTAAAATTATGTTTTAAAAT
GGACTATCATATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGCTTTATATATC
TTGTGGAAAGGACGAAACACCGTTACTTAAATCTTGCAGAAGCTACAAAGATAAGG
CTTCATGCCGAAATCAACACCTGTCTATTTTATGGCAGGGTGTTCCTCGTTATTTAA

U6-RD-espaciador-RD(S.

pyogenes

SF370)

gaggggcctattcccatgattccctcatatttgcatafacgatacaaggctgffagagagataattggaatttaattgactgtaaacacaaagatat
tagtacaaaatacgtgacgtagaaaftaataattcttgggtagtttgcagttttaaattatgtttaaaatggactatcatatgcttaccgtactt

5 gaaagtatttcgatttcttggccttatatacttctgtggaaggacgaaacaccgggttttagagctatgctgtttgaatggccccaaaacNNN

[illegible]

(subrayado minúsculas = repetición directa; N = secuencia guía; negrita = terminador)

ARN quimérico que contiene +48 ARNtracr (S. *pyogenes* SF370)

[illegible]

15 ARN quimérico que contiene +54 ARNtracr (*S. pyogenes* SF370)

[illegible]

ARN quimérico que contiene +67 ARNtracr (*S. pyogenes* SF370)
 gagggcctatttcccatgattccttcataattgcatatacgatacaaggctgtagagagataattggaattaatttgactgtaaacacaaagatat
 tagtacaaaatacgtgacgtagaaagtaataatttctgggtagttgcagttttaaattatgttttaaatggactatcatatgcttaccgtaact
 gaaagtattcgatttcttgctttatatactgtggaaaggacgaaacaccNNNNNNNNNNNNNNNNNNNNgttttaga
 5 gctagaaatagcaagttaaataaggctagtcggttatcaactgaaaaagtTTTTTT (N = secuencia guía; primer subrayado =
 secuencia de emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)

ARN quimérico que contiene +85 ARNtracr (*S. pyogenes* SF370)
 gagggcctatttcccatgattccttcataattgcatatacgatacaaggctgtagagagataattggaattaatttgactgtaaacacaaagatat
 tagtacaaaatacgtgacgtagaaagtaataatttctgggtagttgcagttttaaattatgttttaaatggactatcatatgcttaccgtaact
 10 gaaagtattcgatttcttgctttatatactgtggaaaggacgaaacaccNNNNNNNNNNNNNNNNNNNNgttttaga
gctagaaatagcaagttaaataaggctagtcggttatcaactgaaaaagtggcaccgagtcggtgcTTTTTT (N = secuencia guía; primer
 subrayado = secuencia de emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)

CBh-NLS-SpCas9-NLS
 CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCC
 ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG
 ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTA

TCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCTGGCA
TTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTCGAGGTGAGCCCCACGTTCCTGCTTCACTCTCCCCATCT
CCCCCCCCCTCCCCACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCG
ATGGGGGCGGGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAG
GGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGC
TCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGA
AGCGCGCGGCGGGCGGGAGTCGCTGCGACGCTGCCTTCGCCCCGTGCCCCGCTCCG
CCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCACAGGTGA
GCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCTGAGCAAGAGGTAAGGG
TTTAAGGGATGGTTGGTTGGTGGGGTATTAATGTTTAATTACCTGGAGCACCTGCCT
GAAATCACTTTTTTTCAGGTTGGaccggtgccaccATGGACTATAAGGACCACGACGGAGA
CTACAAGGATCATGATATTGATTACAAAGACGATGACGATAAGATGGCCCCAAAGA
AGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATC
GGCCTGGACATCGGCACCAACTCTGTGGGGCTGGGCCGTGATCACCGACGAGTACAA
GGTGGCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGA
AGAACCTGATCGGAGCCCTGCTGTTCGACAGCGCGGAAACAGCCGAGGCCACCCGG
CTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCT
GCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGAC
TGGAAGAGTCCTTCCTGGTGGAAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTC
GGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCACCATCTACCACCT
GAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGG
CCCTGGCCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACC
CCGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCAG
CTGTTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCT
TGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCCGCG
AGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCC
AACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGA
CACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAAGTACGCGG
ACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGA
GAGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCCTCTATGATCAAGAGATAC

GACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCC
TGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACA
TTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAA
AAGATGGACGGCACCGAGGAAGTCTCGTGAAGCTGAACAGAGAGGACCTGCTGCG
GAAGCAGCGGACCTTCGACAACGGCAGCATCCCCACCAGATCCACCTGGGAGAGC
TGCACGCCATTCTGCGGCGGCAGGAAGATTTTACCCATTCTGAAGGACAACCGGG
AAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCCTCTGGCCA
GGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCC
CTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGC
GGATGACCAACTTCGATAAAGAACCTGCCAACGAGAAGGTGCTGCCCAAGCACAGC
CTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACC
GAGGGAATGAGAAAAGCCCGCCTTCTGAGCGGCGAGCAGAAAAAGGCCATCGTGG
ACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTAC
TTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTT
AACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTT
CCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACAC
TGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCACCTGTTC
GACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGC
TGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTG
GATTTCTGAAGTCCGACGGCTTCGCCAACAGAACTTCATGCAGCTGATCCACGAC
GACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGA
TAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCA
TCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCGGCACAAG
CCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGAC
AGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGG
CAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAG
CTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGA
CATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAA
GGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAG
AGCGACAACGTGCCCTCCGAAGAGGTCTGAAGAAGATGAAGAACTACTGGCGGCA
GCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTTCGACAATCTGACCAAGGCCG

AGAGAGGCGGCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTG
GAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACAC
TAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCTGAAGT
CCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCA
ACAACCTACCACCACGCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTG
ATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA
CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCC
AAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCTGGCC
AACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGA
TCGTGTGGGATAAGGGCCGGGATTTTGGCACCGTGCGGAAAGTGCTGAGCATGCCC
CAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTC
TATCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACC
CTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTG
CCAAAGTGGAAGAGGGCAAGTCCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGG
GATCACCATCATGGAAAAGAAGCAGCTTCGAGAAAGAAATCCCATCGACTTTCTGGAAG
CCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCC
CTGTTGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCA
GAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACCTTCTGTACCTGGCCA
GCCACTATGAGAAGCTGAAGGGCTCCCCGAGGATAATGAGCAGAAACAGCTGTTT
GTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTC
CAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACA
AGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACC
CTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGG
AAGAGGTACACCAGCACCAAGAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCAT
CACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACTTTCTTTT
TCTTAGCTTGACCAGCTTTCTTAGTAGCAGCAGGACGCTTTAA (subrayado = NLS-
hSpCas9-NLS)

- ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
NNNNNNNNNNNNNNNNNNNNNNgTTTTgtactctcaagatttaGAAAtaaatctgcagaagctacaaagataaggctt
catgccgaaatcaacacctgtcattttatggcagggtgttttcgttatttaaTTTTTT (N = secuencia guía; primer subrayado = secuencia
5 de emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)
- ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
NNNNNNNNNNNNNNNNNNNNNNgTTTTgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatca
acacctgtcattttatggcagggtgttttcgttatttaaTTTTTT (N = secuencia guía; primer subrayado = secuencia de
emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)
- 10 ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
NNNNNNNNNNNNNNNNNNNNNNgTTTTgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccgaaatca
acacctgtcattttatggcagggtgtTTTTTT (N = secuencia guía; primer subrayado = secuencia de emparejamiento tracr;
segundo subrayado = secuencia tracr; negrita = terminador)
- 15 ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
NNNNNNNNNNNNNNNNNNNNNNgTTTTgtactctcaagatttaGAAAtaaatctgcagaagctacaaagataaggctt
catgccgaaatcaacacctgtcattttatggcagggtgttttcgttatttaaTTTTTT (N = secuencia guía; primer subrayado = secuencia
de emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)

ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
 NNNNNNNNNNNNNNNNNNNNNNNgttattgtactctcaGAAAtgcagaaectacaaagataaggcttcagccgaaatc
aacaccctgtcattttatggcagggtgttttcgttatttaaTTTTTT (N = secuencia guía; primer subrayado = secuencia de
 emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)

- 5 ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
 NNNNNNNNNNNNNNNNNNNNNNNgttattgtactctcaGAAAtgcagaagctacaaagataaggcttcagccgaaatc
aacaccctgtcattttatggcagggtgtTTTTTT (N = secuencia guía; primer subrayado = secuencia de emparejamiento tracr;
 segundo subrayado = secuencia tracr; negrita = terminador)

- 10 ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
 NNNNNNNNNNNNNNNNNNNNNNNgttattgtactctcaagatttaGAAAtaaatcttcagaagctacaatgataaggctt
catgccgaaatcaacaccctgtcattttatggcagggtgttttcgttatttaaTTTTTT (N = secuencia guía; primer subrayado = secuencia
 de emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)

- 15 ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
 NNNNNNNNNNNNNNNNNNNNNNNgttattgtactctcaGAAAtgcagaagctacaatgataaggcttcagccgaaatc
acaccctgtcattttatggcagggtgttttcgttatttaaTTTTTT (N = secuencia guía; primer subrayado = secuencia de
 emparejamiento tracr; segundo subrayado = secuencia tracr; negrita = terminador)

- 20 ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR1 Cas9 (con PAM de NNAGAAW)
 NNNNNNNNNNNNNNNNNNNNNNNgttattgtactctcaGAAAtgcagaagctacaatgataaasscttcagccgaaatc
acaccctgtcattttatggcagggtgtTTTTTT (N = secuencia guía; primer subrayado = secuencia de emparejamiento tracr;
 segundo subrayado = secuencia tracr; negrita = terminador)

ARN quimérico de ejemplo para *S. thermophilus* LMD-9 CRISPR3 Cas9 (con PAM de NGGNG)
 NNNNNNNNNNNNNNNNNNNNNNNtttttagagctgtgGAAAcacagcaggttaaaataaggcttagtccgtactcaactt
gaaaagggtggcaccgattcgggtTTTTTT (N = secuencia guía; primer subrayado = secuencia de emparejamiento tracr;
 segundo subrayado = secuencia tracr; negrita = terminador)

- 25 Versión codón-optimizada de Cas9 de sitio CRISPR3 LMD-9 de *S. thermophilus* (con una NLS en ambos extremos
 5' y 3')
 ATGAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGACCA
 AGCCCTACAGCATCGGCCTGGACATCGGCACCAATAGCGTGGGCTGGGCCGTGACC
 ACCGACAACCTACAAGGTGCCAGCAAGAAAATGAAGGTGCTGGGCAACACCTCCAA
 GAAGTACATCAAGAAAAACCTGCTGGGCGTGCTGCTGTTTCGACAGCGGCATTACAG
 CCGAGGGCAGACGGCTGAAGAGAACCGCCAGACGGCGGTACACCCGGCGGAGAAA
 CAGAATCCTGTATCTGCAAGAGATCTTCAGCACCGAGATGGCTACCCTGGACGACG
 CCTTCTCCAGCGGCTGGACGACAGCTTCCTGGTGCCCGACGACAAGCGGGACAGC
 AAGTACCCCATCTTCGGCAACCTGGTGGAAGAGAAGGCCTACCACGACGAGTTCCC
 CACCATCTACCACCTGAGAAAGTACCTGGCCGACAGCACCAAGAAGGCCGACCTGA
 GACTGGTGTATCTGGCCCTGGCCACATGATCAAGTACCGGGGCCACTTCCTGATCG
 AGGGCGAGTTCAACAGCAAGAACAACGACATCCAGAAGAACTTCCAGGACTTCCTG

GACACCTACAACGCCATCTTCGAGAGCGACCTGTCCCTGGAAAAACAGCAAGCAGCT
 GGAAGAGATCGTGAAGGACAAGATCAGCAAGCTGGAAAAGAAGGACCGCATCCTG
 AAGCTGTTCCCCGGCGAGAAGAACAGCGGAATCTTCAGCGAGTTTCTGAAGCTGAT
 CGTGGGCAACCAGGCCGACTTCAGAAAGTGCTTCAACCTGGACGAGAAAGCCAGCC
 TGCACTTCAGCAAAGAGAGCTACGACGAGGACCTGGAAACCCTGCTGGGATATATC
 GGCGACGACTACAGCGACGTGTTCCCTGAAGGCCAAGAAGCTGTACGACGCTATCCT
 GCTGAGCGGCTTCCTGACCGTGACCGACAACGAGACAGAGGCCCCACTGAGCAGCG
 CCATGATTAAGCGGTACAACGAGCACAAAGAGGATCTGGCTCTGCTGAAAGAGTAC
 ATCCGGAACATCAGCCTGAAAACCTACAATGAGGTGTTCAAGGACGACACCAAGAA
 CGGCTACGCCGGCTACATCGACGGCAAGACCAACCAGGAAGATTTCTATGTGTACC
 TGAAGAAAGCTGCTGGCCGAGTTCGAGGGGGCCGACTACTTTCTGGAAAAAATCGAC
 CGCGAGGATTTCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCTACCA
 GATCCATCTGCAGGAAATGCGGGCCATCCTGGACAAGCAGGCCAAGTTCTACCCAT
 TCCTGGCCAAGAACAAGAGCGGATCGAGAAGATCCTGACCTCCGCATCCCTTACT
 ACGTGGGCCCCCTGGCCAGAGGCAACAGCGATTTTGCCTGGTCCATCCGGAAGCGC
 AATGAGAAGATACCCCCCTGGAACCTCGAGGACGTGATCGACAAAGAGTCCAGCGC
 CGAGGCCCTTCATCAACCGGATGACCAGCTTCGACCTGTACCTGCCCAGGAAAAGG
 TGCTGCCCAAGCACAGCCTGCTGTACGAGACATTCAATGTGTATAACGAGCTGACCA
 AAGTGCGGTTTATCGCCGAGTCTATGCGGGACTACCAAGTTCTGGACTCCAAGCAGA
 AAAAGGACATCGTGCGCTGTACTTCAAGGACAAGCGGAAAGTGACCGATAAGGAC
 ATCATCGAGTACCTGCACGCCATCTACGGCTACGATGGCATCGAGCTGAAGGGCAT
 CGAGAAGCAGTTCAACTCCAGCCTGAGCACATAACCACGACCTGCTGAACATTATCA
 ACGACAAAGAATTTCTGGACGACTCCAGCAACGAGGCCATCATCGAAGAGATCATC
 CACACCCTGACCATCTTTGAGGACCGCGAGATGATCAAGCAGCGGCTGAGCAAGTT
 CGAGAACATCTTCGACAAGAGCGTGCTGAAAAAGCTGAGCAGACGGCACTACACCG
 GCTGGGGCAAGCTGAGCGCCAAGCTGATCAACGGCATCCGGGACGAGAAGTCCGGC
 AACACAATCCTGGACTACCTGATCGACGACGGCATCAGCAACCGGAACCTCATGCA
 GCTGATCCACGACGACGCCCTGAGCTTCAAGAAGAAGATCCAGAAGGCCCAGATCA
 TCGGGGACGAGGACAAGGGCAACATCAAAGAAGTCGTGAAGTCCCTGCCCGGCAGC
 CCCGCCATCAAGAAGGGAATCCTGCAGAGCATCAAGATCGTGGACGAGCTCGTGAA
 AGTGATGGGCGGCAGAAAGCCCGAGAGCATCGTGGTGGAAATGGCTAGAGAGAAC

CAGTACACCAATCAGGGCAAGAGCAACAGCCAGCAGAGACTGAAGAGACTGGAAA
 AGTCCCTGAAAAGAGCTGGGCAGCAAGATTCTGAAAGAGAATATCCCTGCCAAGCTG
 TCCAAGATCGACAACAACGCCCTGCAGAACGACCGGCTGTACCTGTACTACCTGCA
 GAATGGCAAGGACATGTATACAGGCGACGACCTGGATATCGACCGCCTGAGCAACT
 ACGACATCGACCATATTATCCCCCAGGCCTTCCTGAAAGACAACAGCATTGACAAC
 AAAGTGCTGGTGTCTCCGCCAGCAACCGCGGCAAGTCCGATGATGTGCCCAGCCT
 GGAAGTCGTGAAAAAGAGAAAGACCTTCTGGTATCAGCTGCTGAAAAGCAAGCTGA
 TTAGCCAGAGGAAGTTCGACAACCTGACCAAGGCCGAGAGAGGCGGCCTGAGCCCT
 GAAGATAAGGCCGGCTTCATCCAGAGACAGCTGGTGGAAACCCGGCAGATCACCAA
 GCACGTGGCCAGACTGCTGGATGAGAAGTTTAAACAACAAGAAGGACGAGAACAACC
 GGGCCGTGCGGACCGTGAAGATCATCACCTGAAGTCCACCCTGGTGTCCCAGTTCC
 GGAAGGACTTCGAGCTGTATAAAGTGCGCGAGATCAATGACTTTCACCACGCCAC
 GACGCCTACCTGAATGCCGTGGTGGCTTCGCCCTGCTGAAGAAGTACCCTAAGCTG
 GAACCCGAGTTCGTGTACGGCGACTACCCCAAGTACAACCTCCTTCAGAGAGCGGAA
 GTCCGCCACCGAGAAGGTGTACTTCTACTCCAACATCATGAATATCTTTAAGAAGTC
 CATCTCCCTGGCCGATGGCAGAGTGATCGAGCGGCCCTGATCGAAGTGAACGAAG
 AGACAGGCGAGAGCGTGTGGAACAAAGAAAGCGACCTGGCCACCGTGCGGCGGGT
 GCTGAGTTATCCTCAAGTGAATGTCGTGAAGAAGGTGGAAGAACAGAACCGGCC
 TGGATCGGGGCAAGCCCAAGGGCCTGTTCAACGCCAACCTGTCCAGCAAGCCTAAG
 CCCAACTCCAACGAGAATCTCGTGGGGGCCAAAGAGTACCTGGACCCTAAGAAGTA
 CGGCGGATACGCCGGCATCTCCAATAGCTTCACCGTGCTCGTGAAGGGCACAATCG
 AGAAGGGCGCTAAGAAAAAGATCACAAACGTGCTGGAATTTAGGGGATCTCTATC
 CTGGACCGGATCAACTACCGGAAGGATAAGCTGAACCTTCTGCTGGAAAAAGGCTA
 CAAGGACATTGAGCTGATTATCGAGCTGCCTAAGTACTCCCTGTTTGAAGTGAAGCGA
 CGGCTCCAGACGGATGCTGGCCTCCATCCTGTCCACCAACAACAAGCGGGGCGAGA
 TCCACAAGGGAAACCAGATCTTCTGAGCCAGAAATTTGTGAAACTGCTGTACCACG
 CCAAGCGGATCTCCAACACCATCAATGAGAACCACCGGAAATACGTGGAAAACCAC
 AAGAAAGAGTTTGAGGAACTGTTCTACTACATCCTGGAGTTCAACGAGAAGTATGTG
 GGAGCCAAGAAGAACGGCAAACCTGCTGAACCTCCGCCTTCAGAGCTGGCAGAACCA
 CAGCATCGACGAGCTGTGCAGCTCCTTCATCGGCCCTACCGGCAGCGAGCGGAAGG
 GACTGTTTGAGCTGACCTCCAGAGGCTCTGCCGCCGACTTTGAGTTCCTGGGAGTGA
 AGATCCCCCGGTACAGAGACTACACCCCTCTAGTCTGCTGAAGGACGCCACCCTGA
 TCCACCAGACCGTGACCGGCTGTACGAAACCCGGATCGACCTGGCTAAGCTGGGC
 GAGGGAAAGCGTCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAATA
 A

Ejemplo 5: Optimización del ARN guía para Cas9 de *Streptococcus pyogenes* (referido como SpCas9).

- 5 Los solicitantes mutaron las secuencias ARNtracr y de repetición directa o mutaron el ARN guía quimérico para activar los ARN en las células.

La optimización se basa en la observación de que hay tramos de timinas (Ts) en la ARNtracr y ARN guía, que podía conducir a terminación de la transcripción temprana por el activador Pol 3. Por lo tanto los Solicitantes generaron las siguientes secuencias optimizadas. ARNtracr optimizada y la repetición directa optimizada correspondiente se presentan en parejas.

- 10

ARNtracr optimizado 1 (mutación subrayada):

GGAACCATTCaAACAGCATAGCAAGTTaATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGCACCGAGTCGGTGCTTTTT

Repetición directa optimizada 1 (mutación subrayada):

GTTaTAGAGCTATGCTGTTaTGAATGGTCCCAAAAC

- 5 ARNtracr optimizado 2 (mutación subrayada):

GGAACCATTCAAaACAGCATAGCAAGTTAAaATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGCACCGAGTCGGTGCTTTTT

Repetición directa optimizada 2 (mutación subrayada):

GTaTTAGAGCTATGCTGTaTTGAATGGTCCCAAAAC

- 10 Los solicitantes también optimizaron el ARN guía quimérico para actividad óptima en células eucariotas.

ARN guía original :

NNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

Secuencia de ARN guía quimérico optimizado 1:

NNNNNNNNNNNNNNNNNNNNNGTATTAGAGCTAGAAATAGCAAGTTAATATAAGGC
TAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT

- 15 Secuencia de ARN guía quimérico optimizado 2:

NNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTATGCTGTTTTGGAAACAAAACAGCA
TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG
GTGCTTTTTTTT

Secuencia de ARN guía quimérico optimizado 3:

NNNNNNNNNNNNNNNNNNNNNGTATTAGAGCTATGCTGTATTGGAACAATACAGC
ATAGCAAGTTAATATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTC
GGTGCTTTTTTTT

- 20 Los solicitantes mostraron que el ARN guía quimérico optimizado funciona mejor como se indica en la Figura 9. El experimento se realizó por co-transfección de células 293FT con Cas9 y una casete de ADN ARN guía U6 para expresar una de las cuatro formas de ARN mostradas anteriormente. El objetivo del ARN guía es el mismo sitio objetivo en el lugar Emx1 humano: "GTCACCTCCAATGACTAGGG"

Ejemplo 6: Optimización de Cas9 CRISPR1 LMD-9 de *Streptococcus thermophilus* (referido como StlCas9).

Los solicitantes diseñaron los ARN quiméricos guía como se muestra en la Figura 12.

- 25 Los ARN guía de StlCas9 pueden experimentar el mismo tipo de optimización que para los ARN guía SpCas9, por rotura de los tramos de poli timinas (Ts).

Ejemplo 7: Mejora del sistema Cas9 para aplicación in vivo

Los solicitantes realizaron una búsqueda Metagenómica para un Cas9 con un peso molecular pequeño. La mayoría de los homólogos de Cas9 fue bastante grande. Por ejemplo el SpCas9 tiene alrededor de 1368aa de largo, que es

demasiado largo para ser empaquetado fácilmente en vectores víricos para suministro. Algunas de las secuencias puede que se hayan desnotado y por lo tanto la frecuencia exacta para cada longitud puede no ser necesariamente precisa. Sin embargo, proporciona un atisbo de la distribución de proteínas Cas9 y sugiere que hay homólogos de Cas9 más cortos.

- 5 Por análisis computacional, los solicitantes encontraron que en la cepa bacteriana *Campylobacter*, hay dos proteínas Cas9 con menos de 1.000 aminoácidos. La secuencia para una Cas9 de *Campylobacter jejuni* se presenta a continuación. A esta longitud, CjCas9 se puede empaquetar fácilmente en AAV, lentivirus, Adenovirus, y otros vectores víricos para suministro robusto en células primarias e in vivo en modelos de animales.

Cas9 de *Campylobacter jejuni* (CjCas9)

MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRLARSARKRLARR
KARLNHLKHLIANEFKLNEDYQSFDESLAKAYKGSLISPYELRFRALNELLSKQDFAR
VILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENSK
EFTNVRNKKESYERCIAQSFLKDELKLIFFKKQREFGFSFSKKFEEVLSVAFYKRALKDFS
HLVGNCSSFTDEKRAPKNSPLAFMFVALTRIINLLNNLNKNTGILYTKDDLNALLNEVLK
NGTLTYKQTKLLGLSDDYEFKGEKGTIFYEFKKYKEFIKALGEHNLSQDDLNEIAKDIT
LIKDEIKLKKALAKYDLNQNDISLSKLEFKDHLNISFKALKLVTPMLLEGKKYDEACNE
LNLKVAINEDKKDFLPAFNETYYKDEVTPNPVVLRAIKEYRKVLNALLKKYGVHKNIE
LAREVGKNHSQRAKIEKEQNENYKAKKDAELECEKLGLKINSKNILKRLKFKEQKEFCA
YSGEKIKISDLQDEKMLEIDHIYPYSRSDSYMNKVLVFTKQNEKLNQTPFEAFGNDS
AKWQKIEVLAKNLPKKQKRILDKNYKDKEQKNFKDRNLNDTRYIARLVNLYTKDYL
DFLPLSDDENTKLNLTQKGSKVHVEAKSGMLTSALRHTWGFSAKDRNNHLHHDAVI
IAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGFRQKVLDKIDEIF
VSKPERKKPSGALHEETFRKEEFYQSYGGKEGVKALELGKIRKVNGKIVKNGDMFR
VDIFKHKKTNKFYAVPIYTMDFAKVLPNKAVARSKKGEIKDWILMDENYEFCSLYK
DSLILIQTKDMQEPEFVYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKS
10 IGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK.

El elemento de ARNtracr putativo para esta CjCas9 es:

TATAATCTCATAAGAAATTTAAAAAGGGACTAAAATAAAGAGTTTGCGGGACTCTG
CGGGGTTACAATCCCCTAAAACCGCTTTTAAAT

La secuencia de repetición directa es:

ATTTTACCATAAAGAAATTTAAAAAGGGACTAAAAC

- 15 La estructura co-plegada del ARNtracr y repetición directa se proporciona en la Figura 6.

Un ejemplo de un ARN guía quimérico para CjCas9 es:

NNNNNNNNNNNNNNNNNNNNNGUUUUAGUCCCGAAAGGGACUAAAAUAAAGAGUU
UGCGGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU

Los solicitantes también optimizaron ARN guía de Cas9 usando métodos *in vitro*. La Figura 18 muestra datos de la optimización de ARN guía quimérico St1Cas9 *in vitro*.

- 20 Aunque se han presentado realizaciones preferidas de la presente invención y se han descrito en la presente memoria, será evidente para los expertos en la materia que tales realizaciones se proporcionan como ejemplo solamente. Ahora tendrán lugar numerosas variaciones, cambios y sustituciones para los expertos en la materia sin apartarse de la invención. Se debería entender que se pueden emplear en la práctica de la invención diversas alternativas a las realizaciones de la invención descritas en la presente memoria. Se desea que las siguientes reivindicaciones definan el alcance de la invención y que los métodos y las estructuras dentro del alcance de las
- 25 reivindicaciones y sus equivalentes estén cubiertos de ese modo.

Ejemplo 8: Optimización de ARNsg Sa

5 Los solicitantes diseñaron cinco variantes de ARNsg para SaCas9 para una arquitectura truncada óptima con la mayor eficacia de escisión. Además, se ensayó el sistema dúplex tracr:repetición directa natural junto a los ARNsg. Se transfectaron conjuntamente guías con longitudes indicadas con SaCas9 y se ensayaron en células HEK 293FT para la actividad. Se transfectaron conjuntamente un total de 100 ng de amplicon U6-PCR de ARNsg (o 50 ng de repetición directa y 50 ng de ARNtracr) y 400 ng de plásmido de SaCas9 en 200.000 hepatocitos de ratón Hepa1-6 y se recogió el ADN en 72 horas post -transfección para análisis SURVEYOR. Los resultados se presentan en la Fig. 23.

REFERENCIAS

1. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636-646 (2010).
2. Bogdanove, A.J. & Voytas, D.F. TAL effectors: customizable proteins for DNA targeting. *Science* **333**, 1843-1846 (2011).
3. Stoddard, B.L. Homing endonuclease structure and function. *Q. Rev. Biophys.* **38**, 49-95 (2005).
4. Bae, T. & Schneewind, O. Allelic replacement in *Staphylococcus aureus* with inducible counter-selection. *Plasmid* **55**, 58-63 (2006).
5. Sung, C.K., Li, H., Claverys, J.P. & Morrison, D.A. An *rpsL* cassette, janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. *Appl. Environ. Microbiol.* **67**, 5190-5196 (2001).
6. Sharan, S.K., Thomason, L.C., Kuznetsov, S.G. & Court, D.L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.* **4**, 206-223 (2009).
7. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
8. Deveau, H., Garneau, J.E. & Moineau, S. CRISPR-Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* **64**, 475-493 (2010).
9. Horvath, P. & Barrangou, R. CRISPR-Cas, the immune system of bacteria and archaea. *Science* **327**, 167-170 (2010).

10. Terns, M.P. & Terns, R.M. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* **14**, 321-327 (2011).
11. van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M. & Brouns, S.J. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends. Biochem. Sci.* **34**, 401-407 (2009).
12. Brouns, S.J. et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960-964 (2008).
13. Carte, J., Wang, R., Li, H., Terns, R.M. & Terns, M.P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22**, 3489-3496 (2008).
14. Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602-607 (2011).
15. Hatoum-Aslan, A., Maniv, I. & Marraffini, L.A. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 21218-21222 (2011).
16. Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J.A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355-1358 (2010).
17. Deveau, H. et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190**, 1390-1400 (2008).
18. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* (2012).
19. Makarova, K.S., Aravind, L., Wolf, Y.I. & Koonin, E.V. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct.* **6**, 38 (2011).
20. Barrangou, R. RNA-mediated programmable DNA cleavage. *Nat. Biotechnol.* **30**, 836-838 (2012).
21. Brouns, S.J. Molecular biology. A Swiss army knife of immunity. *Science* **337**, 808-809 (2012).
22. Carroll, D. A CRISPR Approach to Gene Targeting. *Mol. Ther.* **20**, 1658-1660 (2012).

23. Bikard, D., Hatoum-Aslan, A., Mucida, D. & Marraffini, L.A. CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* **12**, 177-186 (2012).
24. Sapranaukas, R. et al. The *Streptococcus thermophilus* CRISPR-Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* (2011).
25. Semenova, E. et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. U.S.A.* (2011).
26. Wiedenheft, B. et al. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. U.S.A.* (2011).
27. Zahner, D. & Hakenbeck, R. The *Streptococcus pneumoniae* beta-galactosidase is a surface protein. *J. Bacteriol.* **182**, 5919-5921 (2000).
28. Marraffini, L.A., Dedent, A.C. & Schneewind, O. Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 192-221 (2006).
29. Motamedi, M.R., Szigety, S.K. & Rosenberg, S.M. Double-strand-break repair recombination in *Escherichia coli*: physical evidence for a DNA replication mechanism in vivo. *Genes Dev.* **13**, 2889-2903 (1999).
30. Hosaka, T. et al. The novel mutation K87E in ribosomal protein S12 enhances protein synthesis activity during the late growth phase in *Escherichia coli*. *Mol. Genet. Genomics* **271**, 317-324 (2004).
31. Costantino, N. & Court, D.L. Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 15748-15753 (2003).
32. Edgar, R. & Qimron, U. The *Escherichia coli* CRISPR system protects from lambda lysogenization, lysogens, and prophage induction. *J. Bacteriol.* **192**, 6291-6294 (2010).
33. Marraffini, L.A. & Sontheimer, E.J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**, 568-571 (2010).
34. Fischer, S. et al. An archaeal immune system can detect multiple Protospacer Adjacent Motifs (PAMs) to target invader DNA. *J. Biol. Chem.* **287**, 33351-33363 (2012).
35. Gudbergssdottir, S. et al. Dynamic properties of the *Sulfolobus* CRISPR-Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol. Microbiol.* **79**, 35-49 (2011).

36. Wang, H.H. et al. Genome-scale promoter engineering by coselection MAGE. *Nat Methods* **9**, 591-593 (2012).
37. Cong, L. et al. Multiplex Genome Engineering Using CRISPR-Cas Systems. *Science* **In press** (2013).
38. Mali, P. et al. RNA-Guided Human Genome Engineering via Cas9. *Science* **In press** (2013).
39. Hoskins, J. et al. Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.* **183**, 5709-5717 (2001).
40. Havarstein, L.S., Coomaraswamy, G. & Morrison, D.A. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11140-11144 (1995).
41. Horinouchi, S. & Weisblum, B. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. *J. Bacteriol.* **150**, 815-825 (1982).
42. Horton, R.M. In Vitro Recombination and Mutagenesis of DNA : SOEing Together Tailor-Made Genes. *Methods Mol. Biol.* **15**, 251-261 (1993).
43. Podbielski, A., Spellerberg, B., Woischnik, M., Pohl, B. & Lutticken, R. Novel series of plasmid vectors for gene inactivation and expression analysis in group A streptococci (GAS). *Gene* **177**, 137-147 (1996).
44. Husmann, L.K., Scott, J.R., Lindahl, G. & Stenberg, L. Expression of the Arp protein, a member of the M protein family, is not sufficient to inhibit phagocytosis of *Streptococcus pyogenes*. *Infection and immunity* **63**, 345-348 (1995).
45. Gibson, D.G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345 (2009).
46. Tangri S, et al. ("Rationally engineered therapeutic proteins with reduced immunogenicity" *J Immunol.* 2005 Mar 15;174(6):3187-96.

Aunque se han presentado realizaciones preferidas de la presente invención y se han descrito en la presente memoria, será evidente para los expertos en la materia que tales realizaciones se proporcionan como ejemplo solamente. Ahora tendrán lugar numerosas variaciones, cambios y sustituciones para los expertos en la materia sin apartarse de la invención. Se debería entender que se pueden emplear diversas alternativas a las realizaciones de la invención descritas en la presente memoria en la práctica de la invención.

Se definen aspectos adicionales de la descripción en las siguientes cláusulas numeradas.

1. Una composición que no está presente en la naturaleza o de ingeniería que comprende:

A) una secuencia de polinucleótidos de ARN quimérico (ARNqui) del sistema CRISPR-Cas, en el que la secuencia de polinucleótidos comprende

(a) una secuencia guía capaz de hibridizar una secuencia fijada como objetivo en una célula eucariota,

(b) una secuencia de emparejamiento tracr y

(c) una secuencia tracr

en la que (a), (b) y (c) se disponen en una orientación 5' a 3',

en la que cuando se transcribe, la secuencia de emparejamiento tracr se hibrida a la secuencia tracr y la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo,

en la que el complejo CRISPR comprende una enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr,

5 o

B) un sistema de enzima CRISPR, en el que el sistema está codificado por un sistema vector que comprende uno o más vectores que comprenden

I. un primer elemento regulador unido de manera operable a una secuencia de polinucleótidos de ARN quimérico (ARNqui) del sistema CRISPR-Cas, en el que la secuencia de polinucleótidos comprende

10 (a) una o más secuencias guía capaces de hibridar una o más secuencias objetivo en una célula eucariota,

(b) una secuencia de emparejamiento tracr y

(c) una o más secuencias tracr y

II. un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica una enzima CRISPR que comprende al menos una o más secuencias de localización nuclear,

15 en la que (a), (b) y (c) se disponen en una orientación 5' a 3',

en la que los componentes I y II se sitúan en los mismos vectores del sistema o diferentes,

en la que cuando se transcribe, la secuencia de emparejamiento tracr se hibrida a la secuencia tracr y la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo,

20 en la que el complejo CRISPR comprende la enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr,

o

C) un sistema de enzima CRISPR multiplexado, en la que el sistema se codifica mediante un sistema vector que comprende uno o más vectores que comprenden

25 I. un primer elemento regulador unido de manera operable a una secuencia de polinucleótidos de ARN quimérico (ARNqui) del sistema CRISPR-Cas, en el que la secuencia de polinucleótidos comprende

(a) una o más secuencias guía capaces de hibridar una o más secuencias objetivo en una célula eucariota,

(b) una secuencia de emparejamiento tracr, y

(c) una o más secuencias tracr, y

30 II. un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica una enzima CRISPR que comprende al menos una o más secuencias de localización nuclear,

en la que (a), (b) y (c) se disponen en una orientación 5' a 3',

en la que los componentes I y II se sitúan en los mismos vectores del sistema o diferentes,

en la que cuando se transcribe, la secuencia de emparejamiento tracr se hibrida a la secuencia tracr y la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo,

35 en la que el complejo CRISPR comprende la enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr,

en la que se utilizan secuencias de polinucleótidos de ARNqui en el sistema multiplexado

o

40 D) un sistema de enzima CRISPR multiplexado, en la que el sistema se codifica mediante un sistema vector que comprende uno o más vectores que comprenden

I. un primer elemento regulador unido de manera operable a

(a) una o más secuencias guía capaces de hibridar una secuencia objetivo en una célula y

(b) al menos una o más secuencias de emparejamiento tracr,

II. un segundo elemento regulador unido de manera operable a una secuencia codificadora de enzima que codifica una enzima CRISPR y

III. un tercer elemento regulador unido de manera operable a una secuencia tracr,

5 en la que los componentes I, II y III se sitúan en los mismos vectores del sistema o diferentes,

en la que cuando se transcribe, la secuencia de emparejamiento tracr se hibrida a la secuencia tracr y la secuencia guía dirige la unión específica de la secuencia de un complejo CRISPR a la secuencia objetivo,

en la que el complejo CRISPR comprende la enzima CRISPR complejada con (1) la secuencia guía que se hibrida a la secuencia objetivo y (2) la secuencia de emparejamiento tracr que se hibrida a la secuencia tracr y

10 en la que en el sistema multiplexado se usan múltiples secuencias guía y una secuencia tracr única;

y

en la que, en la secuencia del polinucleótido de A), o en el sistema de B), C) o D), se modifican una o más de las secuencias guía, tracr y de emparejamiento tracr, para mejorar la estabilidad.

15 2. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 1, en el que la modificación comprende una estructura secundaria de ingeniería.

3. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 1 o la cláusula 2, en el que la modificación comprende una reducción en una región de hibridación entre la secuencia de emparejamiento tracr y la secuencia tracr.

20 4. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente en el que la modificación comprende condensar la secuencia de emparejamiento tracr y la secuencia tracr mediante un lazo artificial.

5. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente en el que la modificación comprende que la secuencia tracr tenga una longitud comprendida entre 40 y 120 pb.

25 6. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente en el que la secuencia tracr está comprendida entre 40 pb y la longitud completa de tracr.

7. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la secuencia tracr incluye al menos los nucleótidos 1-67 del ARNtracr de tipo natural correspondiente.

8. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la secuencia tracr incluye al menos los nucleótidos 1-85 del ARNtracr de tipo natural correspondiente.

30 9. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la secuencia tracr comprende los nucleótidos correspondientes a los nucleótidos 1-67 del ARNtracr de Cas9 de *S. pyogenes* de tipo natural.

35 10. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la secuencia tracr comprende los nucleótidos correspondientes a los nucleótidos 1-85 del ARNtracr de Cas9 de *S. pyogenes* de tipo natural.

11. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 9, en el que la secuencia tracr consiste esencialmente en los nucleótidos correspondientes a los nucleótidos 1-67 del ARNtracr de Cas9 de *S. pyogenes* de tipo natural.

40 12. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 10, en el que la secuencia tracr consiste esencialmente en los nucleótidos correspondientes a los nucleótidos 1-85 del ARNtracr de Cas9 de *S. pyogenes* de tipo natural.

13. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende la optimización de la secuencia.
14. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 13, en el que la modificación comprende la reducción en las secuencias de poliT en la secuencia tracr y/o de emparejamiento tracr.
- 5 15. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 14, en el que una o más T presentes en una secuencia de poliT de la secuencia de tipo natural relevante se han sustituido con un nucleótido que no es T.
16. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 13,14 o 15, en el que la secuencia modificada no comprende ninguna secuencia de poliT que tenga más de cuatro T contiguas.
- 10 17. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende añadir una secuencia terminadora de poliT.
18. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 17, en el que la modificación comprende añadir una secuencia terminadora de poliT en las secuencias tracr y/o de emparejamiento tracr.
- 15 19. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 17 o 18, en el que la modificación comprende añadir una secuencia terminadora de poliT en la secuencia guía.
20. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende alterar lazos y/u horquillas.
21. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 20, en el que la modificación comprende proporcionar como mínimo dos horquillas en la secuencia guía.
- 20 22. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 20 o 21, en el que la modificación comprende proporcionar una horquilla formada mediante la complementación entre la secuencia tracr y de emparejamiento de tracr.
23. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 20 o 22, donde la modificación comprende proporcionar una o más horquillas adicionales en el extremo 3' de la secuencia de ARNtracr.
- 25 24. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 20-23, en el que la modificación comprende proporcionar una o más horquillas adicionales añadidas a la parte 3' de la secuencia guía.
25. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende extender el extremo 5' de la secuencia guía.
- 30 26. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 25, en el que la modificación comprende proporcionar una o más horquillas en el extremo 5' de la secuencia guía.
27. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 25 o 26, donde la modificación comprende unir la secuencia (5'-AGGACGAAGTCCTAA) al extremo 5' de la secuencia guía.
28. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende proporcionar unión cruzada o proporcionar uno o más nucleótidos modificados en la secuencia de polinucleótidos.
- 35 29. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la reivindicación 28, en el que los nucleótidos modificados se proporcionan en cualquiera o en todas las secuencias tracr, de emparejamiento tracr y/o guía y/o en la secuencia codificadora de enzimas y/o en secuencias de vectores.
- 40 30. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 28 o 29, en el que proporcionar nucleótidos modificados comprende la inclusión de al menos un nucleótido que no se encuentra en la naturaleza o un nucleótido modificado o análogos de los mismos.
31. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 30, en el que los nucleótidos

modificados se modifican en el resto ribosa, fosfato y/o base.

32. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 30, en el que el nucleótido modificado se selecciona del grupo constituido por 2'-O-metil-análogos, 2'-desoxi-análogos o 2'-fluoro-análogos.

5 33. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de la cláusula 30, en el que el nucleótido modificado se selecciona del grupo constituido por 2-aminopurina, 5-bromo-uridina, pseudouridina, inosina, 7-metilguanosa.

34. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende dos horquillas.

10 35. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende tres horquillas.

36. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende como mucho cinco horquillas.

37. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la enzima CRISPR es una enzima del sistema CRISPR de tipo II.

15 38. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la enzima CRISPR es una enzima Cas9.

39. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la enzima CRISPR comprende menos de mil aminoácidos.

20 40. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquiera de las cláusulas 1-38, en el que la enzima CRISPR comprende menos de cuatro mil aminoácidos.

41. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente en el que la enzima Cas9 es StCas9 o St1Cas9.

25 42. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la enzima Cas9 es una enzima Cas9 procedente de un organismo seleccionado del grupo que comprende los géneros *Streptococcus*, *Campylobacter*, *Nitratifactor*, *Staphylococcus*, *Parvibaculum*, *Roseburia*, *Neisseria*, *Gluconacetobacter*, *Azospirillum*, *Sphaerochaeta*, *Lactobacillus*, *Eubacterium* o *Corynebacter*.

43. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la enzima CRISPR es una nucleasa que dirige la escisión de ambas hebras en la posición de la secuencia objetivo.

30 44. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que el primer elemento regulador es un activador de la polimerasa III.

45. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que el segundo elemento regulador es un activador de la polimerasa II.

46. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la secuencia guía comprende al menos quince nucleótidos.

35 47. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquier cláusula precedente, en el que la modificación comprende ARN de secuencia tracr optimizada y/o de secuencia guía optimizada y/o estructura co-plegada de secuencia tracr y/o secuencia o secuencias de emparejamiento tracr y/o estructuras secundarias estabilizantes de secuencia tracr y/o secuencia tracr con una región reducida de emparejamiento de bases y/o elementos de ARN condensados de secuencia tracr; y/o en el sistema multiplexado hay dos ARN que comprenden un trazador y que comprenden una pluralidad de guías o ARN que comprende una pluralidad de quimeras.

40

48. El ARNqui del sistema CRISPR-Cas o sistema de enzima CRISPR de cualquiera de las cláusulas 1-47, en el que la enzima CRISPR es codón-optimizada para la expresión en una célula eucariota.

49. La composición de cualquier cláusula precedente, en la que la composición comprende una secuencia de polinucleótidos del ARN quimérico (ARNqui) del sistema CRISPR-Cas.

50. La composición de la cláusula 49, en la que la composición comprende además una secuencia de polinucleótidos que codifica una enzima CRISPR que comprende al menos una o más secuencias de localización nuclear.

5

51. El sistema de enzima CRISPR de cualquier cláusula precedente.

52. El sistema de enzima CRISPR multiplexado de cualquier cláusula precedente.

53. El producto de la transcripción o traducción de la composición de la cláusula 49 o 50, el sistema de enzima CRISPR de la cláusula 51 o el sistema de enzima CRISPR multiplexado de la cláusula 52.

LISTADO DE SECUENCIAS

<110> THE BROAD INSTITUTE, INC.

INSTITUTO DE TECNOLOGÍA DE MASSACHUSETTS

<120> INGENIERÍA DE SISTEMAS, MÉTODOS Y COMPOSICIONES GUÍA OPTIMIZADAS PARA MANIPULACIÓN DE SECUENCIAS

5 <130> PC927381EPB

<140> Patente de EE.UU. PCT/US2013/074819

< 141> 2013-12-12

<150> 61/836.127

< 151> 2013-06-17

10 <150> 61/835.931
< 151> 2013-06-17

<150> 61/828.130

< 151> 2013-05-28

15 <150> 61/819.803
< 151> 2013-05-06

<150> 61/814.263

< 151> 2013-04-20

<150> 61/806.375

< 151> 2013-03-28

20 <150> 61/802.174
< 151> 2013-03-15

<150> 61/791.409

< 151> 2013-03-15

25 <150> 61/769.046
< 151> 2013-02-25

<150> 61/758.468

< 151> 2013-01-30

<150> 61/748.427

< 151> 2013-01-02

30 <150> 61/736.527
< 151> 2012-12-12

<160> 264

<170> PatentIn versión 3.5

35 <210> 1
< 211> 15

< 212> ADN

< 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

5 <400> 1
 aggacgaagt cctaa 15

<210> 2
 < 211> 7
 < 212> PRT
 < 213> Virus 40 del simio

10 <400> 2
 Pro Lys Lys Lys Arg Lys Val
 1 5

<210> 3
 < 211> 16
 < 212> PRT
 15 < 213> Desconocido

<220>
 < 221> fuente
 < 223> /nota="Descripción de Desconocido: Secuencia de NLS bipartita de nucleoplasmina"

20 <400> 3
 Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
 1 5 10 15

<210> 4
 < 211> 9
 < 212> PRT
 < 213> Desconocido

25 <220>
 < 221> fuente
 < 223> /nota="Descripción de Desconocido: Secuencia NLS C-myc"

<400> 4
 Pro Ala Ala Lys Arg Val Lys Leu Asp
 1 5

30 <210> 5
 < 211> 11
 < 212> PRT
 < 213> Desconocido

<220>
 35 < 221> fuente
 < 223> /nota="Descripción de Desconocido: Secuencia NLS C-myc"

<400> 5
 Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Pro
 1 5 10

40 <210> 6
 < 211> 38
 < 212> PRT
 < 213> Homo sapiens

<400> 6

Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly
1 5 10 15

Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro
20 25 30

Arg Asn Gln Gly Gly Tyr
35

<210> 7

<211> 42

5 <212> PRT

<213> Desconocido

<220>

<221> fuente

<223> /nota="Descripción de Desconocido: dominio IBB de secuencia de importina-alfa"

10 <400> 7

Arg Met Arg Ile Glx Phe Lys Asn Lys Gly Lys Asp Thr Ala Glu Leu
1 5 10 15

Arg Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys
20 25 30

Asp Glu Gln Ile Leu Lys Arg Arg Asn Val
35 40

<210> 8

<211> 8

<212> PRT

15 <213> Desconocido

<220>

<221> fuente

<223> /nota="Descripción de Desconocido: Secuencia de proteína de mioma T"

<400> 8

Val Ser Arg Lys Arg Pro Arg Pro
1 5

20

<210> 9

<211> 8

<212> PRT

<213> Desconocido

<220>

<221> fuente

<223> /nota="Descripción de Desconocido: Secuencia de proteína de mioma T"

<400> 9

Pro Pro Lys Lys Ala Arg Glu Asp
1 5

30

<210> 10

<211> 8

<212> PRT

<213> Homo sapiens

<400> 10
Pro Gln Pro Lys Lys Lys Pro Leu
 1 5

5 <210> 11
 <211> 12
 <212> PRT
 <213> Mus musculus

<400> 11
Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro
 1 5 10

10 <210> 12
 <211> 5
 <212> PRT
 <213> Virus de la influenza

<400> 12
Asp Arg Leu Arg Arg
 1 5

15 <210> 13
 <211> 7
 <212> PRT
 <213> Virus de la influenza

20 <400> 13
Pro Lys Gln Lys Lys Arg Lys
 1 5

<210> 14
 <211> 10
 <212> PRT
 <213> Virus de Hepatitis delta

25 <400> 14
Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu
 1 5 10

30 <210> 15
 <211> 10
 <212> PRT
 <213> Mus musculus

<400> 15
Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg
 1 5 10

35 <210> 16
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 16
Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys
 1 5 10 15

Lys Ser Lys Lys
 20

<210> 17
 < 211> 17
 < 212> PRT
 < 213> Homo sapiens

5 <400> 17
Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys
 1 5 10 15

Lys

<210> 18
 < 211> 27
 < 212> ADN
 10 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

15 <220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t o g

20 <220>
 < 221> base modificada
 < 222> (21)..(22)
 < 223> a, c, t, g, Desconocido u otro

<400> 18
 nnnnnnnnnn nnnnnnnnnn nnagaaw 27

25 <210> 19
 < 211> 19
 < 212> ADN
 < 213> Secuencia Artificial

30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<220>
 < 221> base modificada
 < 222> (1)..(12)
 < 223> a, c, t o g

35 <220>
 < 221> base modificada
 < 222> (13)..(14)
 < 223> a, c, t, g, Desconocido u otro

40 <400> 19
 nnnnnnnnnn nnnnagaaw 19

<210> 20
 < 211> 27
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

5 <220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t o g

10 <220>
 < 221> base modificada
 < 222> (21)..(22)
 < 223> a, c, t, g, Desconocido u otro

<400> 20
 nnnnnnnnnnn nnnnnnnnnnn nnagaaw 27

15 <210> 21
 < 211> 18
 < 212> ADN
 < 213> Secuencia Artificial

20 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<220>
 < 221> base modificada
 < 222> (1)..(11)
 < 223> a, c, t o g

25 <220>
 < 221> base modificada
 < 222> (12)..(13)
 < 223> a, c, t, g, Desconocido u otro

30 <400> 21
 nnnnnnnnnnn nnnagaaw 18

<210> 22
 < 211> 137
 < 212> ADN
 < 213> Secuencia Artificial

35 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

40 <220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t, g, Desconocido u otro

<400> 22
 nnnnnnnnnnn nnnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa 60
 gctacaaaga taaggcttca tgccgaaatc aacaccctgt cattttatgg cagggtgttt 120
 tcgttatttta atttttt 137

<210> 23
 < 211> 123
 < 212> ADN
 < 213> Secuencia Artificial

5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

10 <220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t, g, Desconocido u otro

<400> 23
 nnnnnnnnnnn nnnnnnnnnn gtttttgtac tctcagaaat gcagaagcta caaagataag 60
 gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttcgt tatttaattt 120
 ttt 123

15 <210> 24
 < 211> 110
 < 212> ADN
 < 213> Secuencia Artificial

20 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t, g, Desconocido u otro

25 <400> 24
 nnnnnnnnnnn nnnnnnnnnn gtttttgtac tctcagaaat gcagaagcta caaagataag 60
 gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttttt 110

30 <210> 25
 < 211> 102
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

35 <220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t, g, Desconocido u otro

<400> 25
 nnnnnnnnnnn nnnnnnnnnn gtttttagagc tagaaatagc aagttaaaat aaggctagtc 60
 cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tt 102

40 <210> 26
 < 211> 88

	< 212> ADN		
	< 213> Secuencia Artificial		
	<220>		
	< 221> fuente		
5	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<220>		
	< 221> base modificada		
	< 222> (1)..(20)		
	< 223> a, c, t, g, Desconocido u otro		
10	<400> 26		
	nnnnnnnnnn nnnnnnnnnn gtttttagagc tagaaatagc aagttaaaat aaggctagtc	60	
	cgttatcaac ttgaaaaagt gttttttt	88	
	<210> 27		
	< 211> 76		
	< 212> ADN		
15	< 213> Secuencia Artificial		
	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<220>		
20	< 221> base modificada		
	< 222> (1)..(20)		
	< 223> a, c, t, g, Desconocido u otro		
	<400> 27		
	nnnnnnnnnn nnnnnnnnnn gtttttagagc tagaaatagc aagttaaaat aaggctagtc	60	
	cgttatcatt tttttt	76	
25	<210> 28		
	< 211> 12		
	< 212> ARN		
	< 213> Secuencia Artificial		
	<220>		
30	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<400> 28		
	guuuuagagc ua	12	
	<210> 29		
35	< 211> 33		
	< 212> ADN		
	< 213> Homo sapiens		
	<400> 29		
	ggacatcgat gtcacctcca atgactaggg tgg	33	
40	<210> 30		
	< 211> 33		
	< 212> ADN		
	< 213> Homo sapiens		

<400> 30
 cattggaggt gacatcgatg tcctcccat tgg 33

5 <210> 31
 < 211> 33
 < 212> ADN
 < 213> Homo sapiens

<400> 31
 ggaagggcct ggtccgagc agaagaaga ggg 33

10 <210> 32
 < 211> 33
 < 212> ADN
 < 213> Homo sapiens

<400> 32
 ggtggcgaga ggggccgaga ttgggtgtc agg 33

15 <210> 33
 < 211> 33
 < 212> ADN
 < 213> Homo sapiens

20 <400> 33
 atgcaggagg gtggcgagag gggccgagat tgg 33

<210> 34
 < 211> 21
 < 212> ADN
 < 213> Secuencia Artificial

25 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

<400> 34
 aaaaccaccc ttctctctgg c 21

30 <210> 35
 < 211> 21
 < 212> ADN
 < 213> Secuencia Artificial

35 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

<400> 35
 ggagattgga gacacggaga g 21

40 <210> 36
 < 211> 20
 < 212> ADN
 < 213> Secuencia Artificial

45 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

<400> 36
ctggaaagcc aatgcctgac 20

5 <210> 37
< 211> 20
< 212> ADN
< 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

10 <400> 37
ggcagcaaac tcctgtcct 20

15 <210> 38
< 211> 12
< 212> ADN
< 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

20 <400> 38
gttttagagc ta 12

<210> 39
< 211> 335
< 212> ADN
< 213> Secuencia Artificial

25 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 39
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag 60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180
atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240
cgaaacaccg gaaccattca aaacagcata gcaagttaaa ataaggctag tccgttatca 300
acttgaaaaa gtggcaccga gtcggtgctt ttttt 335

30 <210> 40
< 211> 423
< 212> ADN
< 213> Secuencia Artificial

35 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 40

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag	60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga	120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat	180
atgcttaccg taacttgaaa gtatttcgat ttcttggcct tatatatctt gtggaaagga	240
cgaacacccg gtagtattaa gtattgtttt atggctgata aatttctttg aatttctcct	300
tgattatttg ttataaaagt tataaaataa tcttgttgga accattcaaa acagcatagc	360
 aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtgctttt	420
ttt	423

5 <210> 41
< 211> 339
< 212> ADN
< 213> Secuencia Artificial

10 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 41

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag	60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga	120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat	180
atgcttaccg taacttgaaa gtatttcgat ttcttggcct tatatatctt gtggaaagga	240
cgaacacccg ggttttagag ctatgctgtt ttgaatggc ccaaacggg tcttcgagaa	300
gacgttttag agctatgctg ttttgaatgg tcccaaac	339

15 <210> 42
< 211> 309
< 212> ADN
< 213> Secuencia Artificial

20 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

```

<400> 42
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag      60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga      120
aagtaataat ttcttgggta gtttgcagtt ttaaattat gttttaaaat ggactatcat      180
atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga      240
cgaaacaccg ggtcttcgag aagacctgtt ttagagctag aaatagcaag taaaataag      300
gctagtccg                                     309

```

<210> 43

< 211> 1648

5 < 212> PRT

< 213> Secuencia Artificial

<220>

< 221> fuente

< 223> /nota="Descripción de secuencia artificial: polipéptido sintético"

<400> 43

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
 1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
 20 25 30

Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly Leu
 35 40 45

Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr
 50 55 60

Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His
 65 70 75 80

Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu
 85 90 95

Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr
 100 105 110

Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu
 115 120 125

Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe
 130 135 140

Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn
 145 150 155 160

Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His
 165 170 175

Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu
 180 185 190

Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu
 195 200 205

Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe
 210 215 220

Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile
 225 230 235 240

Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser
 245 250 255

Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys
 260 265 270

Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr
 275 280 285

Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln
 290 295 300

Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln
 305 310 315 320

Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser
 325 330 335

Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr
 340 345 350

Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His
 355 360 365

Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu
 370 375 380

Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly
 385 390 395 400

Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys
 405 410 415

Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu
 420 425 430

Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser
 435 440 445

Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg
 450 455 460

Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu
 465 470 475 480

Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg
 485 490 495

Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile

ES 2 553 782 T3

500					505					510					
Thr	Pro	Trp	Asn	Phe	Glu	Glu	Val	Val	Asp	Lys	Gly	Ala	Ser	Ala	Gln
		515					520					525			
Ser	Phe	Ile	Glu	Arg	Met	Thr	Asn	Phe	Asp	Lys	Asn	Leu	Pro	Asn	Glu
	530					535					540				
Lys	Val	Leu	Pro	Lys	His	Ser	Leu	Leu	Tyr	Glu	Tyr	Phe	Thr	Val	Tyr
545					550					555					560
Asn	Glu	Leu	Thr	Lys	Val	Lys	Tyr	Val	Thr	Glu	Gly	Met	Arg	Lys	Pro
				565					570					575	
Ala	Phe	Leu	Ser	Gly	Glu	Gln	Lys	Lys	Ala	Ile	Val	Asp	Leu	Leu	Phe
			580					585					590		
Lys	Thr	Asn	Arg	Lys	Val	Thr	Val	Lys	Gln	Leu	Lys	Glu	Asp	Tyr	Phe
		595					600					605			
Lys	Lys	Ile	Glu	Cys	Phe	Asp	Ser	Val	Glu	Ile	Ser	Gly	Val	Glu	Asp
	610					615					620				
Arg	Phe	Asn	Ala	Ser	Leu	Gly	Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile
625					630					635					640
Lys	Asp	Lys	Asp	Phe	Leu	Asp	Asn	Glu	Glu	Asn	Glu	Asp	Ile	Leu	Glu
				645					650					655	
Asp	Ile	Val	Leu	Thr	Leu	Thr	Leu	Phe	Glu	Asp	Arg	Glu	Met	Ile	Glu
			660					665					670		
Glu	Arg	Leu	Lys	Thr	Tyr	Ala	His	Leu	Phe	Asp	Asp	Lys	Val	Met	Lys
		675					680					685			
Gln	Leu	Lys	Arg	Arg	Arg	Tyr	Thr	Gly	Trp	Gly	Arg	Leu	Ser	Arg	Lys
	690					695					700				
Leu	Ile	Asn	Gly	Ile	Arg	Asp	Lys	Gln	Ser	Gly	Lys	Thr	Ile	Leu	Asp
705					710					715					720
Phe	Leu	Lys	Ser	Asp	Gly	Phe	Ala	Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile
				725					730					735	
His	Asp	Asp	Ser	Leu	Thr	Phe	Lys	Glu	Asp	Ile	Gln	Lys	Ala	Gln	Val
			740					745					750		

Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly
 755 760 765

Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp
 770 775 780

Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile
 785 790 795 800

Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser
 805 810 815

Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser
 820 825 830

Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu
 835 840 845

Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp
 850 855 860

Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile
 865 870 875 880

Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu
 885 890 895

Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu
 900 905 910

Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala
 915 920 925

Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg
 930 935 940

Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu
 945 950 955 960

Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser
 965 970 975

Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val
 980 985 990

Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp
 995 1000 1005

Phe	Gln	Phe	Tyr	Lys	Val	Arg	Glu	Ile	Asn	Asn	Tyr	His	His	Ala
1010						1015					1020			
His	Asp	Ala	Tyr	Leu	Asn	Ala	Val	Val	Gly	Thr	Ala	Leu	Ile	Lys
1025						1030					1035			
Lys	Tyr	Pro	Lys	Leu	Glu	Ser	Glu	Phe	Val	Tyr	Gly	Asp	Tyr	Lys
1040						1045					1050			
Val	Tyr	Asp	Val	Arg	Lys	Met	Ile	Ala	Lys	Ser	Glu	Gln	Glu	Ile
1055						1060					1065			
Gly	Lys	Ala	Thr	Ala	Lys	Tyr	Phe	Phe	Tyr	Ser	Asn	Ile	Met	Asn
1070						1075					1080			
Phe	Phe	Lys	Thr	Glu	Ile	Thr	Leu	Ala	Asn	Gly	Glu	Ile	Arg	Lys
1085						1090					1095			
Arg	Pro	Leu	Ile	Glu	Thr	Asn	Gly	Glu	Thr	Gly	Glu	Ile	Val	Trp
1100						1105					1110			
Asp	Lys	Gly	Arg	Asp	Phe	Ala	Thr	Val	Arg	Lys	Val	Leu	Ser	Met
1115						1120					1125			
Pro	Gln	Val	Asn	Ile	Val	Lys	Lys	Thr	Glu	Val	Gln	Thr	Gly	Gly
1130						1135					1140			
Phe	Ser	Lys	Glu	Ser	Ile	Leu	Pro	Lys	Arg	Asn	Ser	Asp	Lys	Leu
1145						1150					1155			
Ile	Ala	Arg	Lys	Lys	Asp	Trp	Asp	Pro	Lys	Lys	Tyr	Gly	Gly	Phe
1160						1165					1170			
Asp	Ser	Pro	Thr	Val	Ala	Tyr	Ser	Val	Leu	Val	Val	Ala	Lys	Val
1175						1180					1185			
Glu	Lys	Gly	Lys	Ser	Lys	Lys	Leu	Lys	Ser	Val	Lys	Glu	Leu	Leu
1190						1195					1200			
Gly	Ile	Thr	Ile	Met	Glu	Arg	Ser	Ser	Phe	Glu	Lys	Asn	Pro	Ile
1205						1210					1215			
Asp	Phe	Leu	Glu	Ala	Lys	Gly	Tyr	Lys	Glu	Val	Lys	Lys	Asp	Leu
1220						1225					1230			
Ile	Ile	Lys	Leu	Pro	Lys	Tyr	Ser	Leu	Phe	Glu	Leu	Glu	Asn	Gly
1235						1240					1245			

Arg Lys	Arg Met	Leu Ala	Ser	Ala Gly	Glu Leu	Gln	Lys Gly	Asn
1250			1255			1260		
Glu Leu	Ala Leu	Pro Ser	Lys	Tyr Val	Asn Phe	Leu	Tyr Leu	Ala
1265			1270			1275		
Ser His	Tyr Glu	Lys Leu	Lys	Gly Ser	Pro Glu	Asp	Asn Glu	Gln
1280			1285			1290		
Lys Gln	Leu Phe	Val Glu	Gln	His Lys	His Tyr	Leu	Asp Glu	Ile
1295			1300			1305		
Ile Glu	Gln Ile	Ser Glu	Phe	Ser Lys	Arg Val	Ile	Leu Ala	Asp
1310			1315			1320		
Ala Asn	Leu Asp	Lys Val	Leu	Ser Ala	Tyr Asn	Lys	His Arg	Asp
1325			1330			1335		
Lys Pro	Ile Arg	Glu Gln	Ala	Glu Asn	Ile Ile	His	Leu Phe	Thr
1340			1345			1350		
Leu Thr	Asn Leu	Gly Ala	Pro	Ala Ala	Phe Lys	Tyr	Phe Asp	Thr
1355			1360			1365		
Thr Ile	Asp Arg	Lys Arg	Tyr	Thr Ser	Thr Lys	Glu	Val Leu	Asp
1370			1375			1380		
Ala Thr	Leu Ile	His Gln	Ser	Ile Thr	Gly Leu	Tyr	Glu Thr	Arg
1385			1390			1395		
Ile Asp	Leu Ser	Gln Leu	Gly	Gly Asp	Ala Ala	Ala	Val Ser	Lys
1400			1405			1410		
Gly Glu	Glu Leu	Phe Thr	Gly	Val Val	Pro Ile	Leu	Val Glu	Leu
1415			1420			1425		
Asp Gly	Asp Val	Asn Gly	His	Lys Phe	Ser Val	Ser	Gly Glu	Gly
1430			1435			1440		
Glu Gly	Asp Ala	Thr Tyr	Gly	Lys Leu	Thr Leu	Lys	Phe Ile	Cys
1445			1450			1455		
Thr Thr	Gly Lys	Leu Pro	Val	Pro Trp	Pro Thr	Leu	Val Thr	Thr
1460			1465			1470		
Leu Thr	Tyr Gly	Val Gln	Cys	Phe Ser	Arg Tyr	Pro	Asp His	Met

1475		1480		1485
Lys Gln His Asp Phe Phe	Lys Ser Ala Met Pro Glu Gly Tyr Val			
1490	1495	1500		
Gln Glu Arg Thr Ile Phe	Phe Lys Asp Asp Gly Asn Tyr Lys Thr			
1505	1510	1515		
Arg Ala Glu Val Lys Phe	Glu Gly Asp Thr Leu Val Asn Arg Ile			
1520	1525	1530		
Glu Leu Lys Gly Ile Asp	Phe Lys Glu Asp Gly Asn Ile Leu Gly			
1535	1540	1545		
His Lys Leu Glu Tyr Asn	Tyr Asn Ser His Asn Val Tyr Ile Met			
1550	1555	1560		
Ala Asp Lys Gln Lys Asn	Gly Ile Lys Val Asn Phe Lys Ile Arg			
1565	1570	1575		
His Asn Ile Glu Asp Gly	Ser Val Gln Leu Ala Asp His Tyr Gln			
1580	1585	1590		
Gln Asn Thr Pro Ile Gly	Asp Gly Pro Val Leu Leu Pro Asp Asn			
1595	1600	1605		
His Tyr Leu Ser Thr Gln	Ser Ala Leu Ser Lys Asp Pro Asn Glu			
1610	1615	1620		
Lys Arg Asp His Met Val	Leu Leu Glu Phe Val Thr Ala Ala Gly			
1625	1630	1635		
Ile Thr Leu Gly Met Asp	Glu Leu Tyr Lys			
1640	1645			

<210> 44

5 <211> 1625

<212> PRT

<213> Secuencia Artificial

<220>

<221> fuente

10 <223> /nota="Descripción de Secuencia Artificial: Polipéptido sintético"

<400> 44

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

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

20					25					30					
Lys	Val	Leu	Gly	Asn	Thr	Asp	Arg	His	Ser	Ile	Lys	Lys	Asn	Leu	Ile
		35					40					45			
Gly	Ala	Leu	Leu	Phe	Asp	Ser	Gly	Glu	Thr	Ala	Glu	Ala	Thr	Arg	Leu
	50					55					60				
Lys	Arg	Thr	Ala	Arg	Arg	Arg	Tyr	Thr	Arg	Arg	Lys	Asn	Arg	Ile	Cys
65					70					75					80
Tyr	Leu	Gln	Glu	Ile	Phe	Ser	Asn	Glu	Met	Ala	Lys	Val	Asp	Asp	Ser
				85					90					95	
Phe	Phe	His	Arg	Leu	Glu	Glu	Ser	Phe	Leu	Val	Glu	Glu	Asp	Lys	Lys
			100					105					110		
His	Glu	Arg	His	Pro	Ile	Phe	Gly	Asn	Ile	Val	Asp	Glu	Val	Ala	Tyr
		115					120					125			
His	Glu	Lys	Tyr	Pro	Thr	Ile	Tyr	His	Leu	Arg	Lys	Lys	Leu	Val	Asp
	130					135					140				
Ser	Thr	Asp	Lys	Ala	Asp	Leu	Arg	Leu	Ile	Tyr	Leu	Ala	Leu	Ala	His
145					150					155					160
Met	Ile	Lys	Phe	Arg	Gly	His	Phe	Leu	Ile	Glu	Gly	Asp	Leu	Asn	Pro
				165					170					175	
Asp	Asn	Ser	Asp	Val	Asp	Lys	Leu	Phe	Ile	Gln	Leu	Val	Gln	Thr	Tyr
			180					185					190		
Asn	Gln	Leu	Phe	Glu	Glu	Asn	Pro	Ile	Asn	Ala	Ser	Gly	Val	Asp	Ala
		195					200					205			
Lys	Ala	Ile	Leu	Ser	Ala	Arg	Leu	Ser	Lys	Ser	Arg	Arg	Leu	Glu	Asn
	210					215					220				
Leu	Ile	Ala	Gln	Leu	Pro	Gly	Glu	Lys	Lys	Asn	Gly	Leu	Phe	Gly	Asn
225					230					235					240
Leu	Ile	Ala	Leu	Ser	Leu	Gly	Leu	Thr	Pro	Asn	Phe	Lys	Ser	Asn	Phe
				245					250					255	
Asp	Leu	Ala	Glu	Asp	Ala	Lys	Leu	Gln	Leu	Ser	Lys	Asp	Thr	Tyr	Asp
			260					265					270		

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285
 Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300
 Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320
 Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335
 Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 340 345 350
 Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365
 Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415
 Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430
 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445
 Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460
 Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480
 Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495
 Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510
 Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525

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

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

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

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

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

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

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

 His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655

 Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670

 Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685

 Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700

 Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720

 His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
 725 730 735

 Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
 740 745 750

 Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765

 Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
 785 790 795 800
 Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
 805 810 815
 Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
 820 825 830
 Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
 835 840 845
 Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
 850 855 860
 Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
 865 870 875 880
 Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
 885 890 895
 Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
 900 905 910
 Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925
 Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940
 Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960
 Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975
 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 1040	Phe Lys Thr Glu 1045	Ile Thr Leu Ala 1050
Asn Gly Glu Ile Arg Lys Arg 1055	Pro Leu Ile Glu 1060	Thr Asn Gly Glu 1065
Thr Gly Glu Ile Val Trp Asp 1070	Lys Gly Arg Asp 1075	Phe Ala Thr Val 1080
Arg Lys Val Leu Ser Met Pro 1085	Gln Val Asn Ile 1090	Val Lys Lys Thr 1095
Glu Val Gln Thr Gly Gly Phe 1100	Ser Lys Glu Ser 1105	Ile Leu Pro Lys 1110
Arg Asn Ser Asp Lys Leu Ile 1115	Ala Arg Lys Lys 1120	Asp Trp Asp Pro 1125
Lys Lys Tyr Gly Gly Phe Asp 1130	Ser Pro Thr Val 1135	Ala Tyr Ser Val 1140
Leu Val Val Ala Lys Val Glu 1145	Lys Gly Lys Ser 1150	Lys Lys Leu Lys 1155
Ser Val Lys Glu Leu Leu Gly 1160	Ile Thr Ile Met 1165	Glu Arg Ser Ser 1170
Phe Glu Lys Asn Pro Ile Asp 1175	Phe Leu Glu Ala 1180	Lys Gly Tyr Lys 1185
Glu Val Lys Lys Asp Leu Ile 1190	Ile Lys Leu Pro 1195	Lys Tyr Ser Leu 1200
Phe Glu Leu Glu Asn Gly Arg 1205	Lys Arg Met Leu 1210	Ala Ser Ala Gly 1215
Glu Leu Gln Lys Gly Asn Glu 1220	Leu Ala Leu Pro 1225	Ser Lys Tyr Val 1230
Asn Phe Leu Tyr Leu Ala Ser 1235	His Tyr Glu Lys 1240	Leu Lys Gly Ser 1245
Pro Glu Asp Asn Glu Gln Lys 1250	Gln Leu Phe Val 1255	Glu Gln His Lys 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			
Ala	Ala	Ala	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val
1370						1375					1380			
Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe
1385						1390					1395			
Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu
1400						1405					1410			
Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp
1415						1420					1425			
Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser
1430						1435					1440			
Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala
1445						1450					1455			
Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp
1460						1465					1470			
Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp
1475						1480					1485			
Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu
1490						1495					1500			

Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser
1505 1510 1515

His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys
1520 1525 1530

Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln
1535 1540 1545

Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
1550 1555 1560

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
1565 1570 1575

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu
1580 1585 1590

Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr
1595 1600 1605

Lys Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys
1610 1615 1620

Lys Lys
1625

<210> 45

< 211> 1664

5 < 212> PRT

< 213> Secuencia Artificial

<220>

< 221> fuente

< 223> /nota="Descripción de Secuencia Artificial: Polipéptido sintético"

10 <400> 45

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
20 25 30

Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly Leu
35 40 45

Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr
50 55 60

Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His
 65 70 75 80
 Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu
 85 90 95
 Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr
 100 105 110
 Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu
 115 120 125
 Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe
 130 135 140
 Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn
 145 150 155 160
 Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His
 165 170 175
 Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu
 180 185 190
 Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu
 195 200 205
 Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe
 210 215 220
 Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile
 225 230 235 240
 Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser
 245 250 255
 Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys
 260 265 270
 Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr
 275 280 285
 Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln
 290 295 300
 Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln
 305 310 315 320

Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser
325 330 335

Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr
340 345 350

Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His
355 360 365

Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu
370 375 380

Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly
385 390 395 400

Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys
405 410 415

Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu
420 425 430

Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser
435 440 445

Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg
450 455 460

Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu
465 470 475 480

Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg
485 490 495

Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile
500 505 510

Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln
515 520 525

Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu
530 535 540

Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr
545 550 555 560

Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro

ES 2 553 782 T3

565								570				575			
Ala	Phe	Leu	Ser	Gly	Glu	Gln	Lys	Lys	Ala	Ile	Val	Asp	Leu	Leu	Phe
		580						585					590		
Lys	Thr	Asn	Arg	Lys	Val	Thr	Val	Lys	Gln	Leu	Lys	Glu	Asp	Tyr	Phe
		595					600					605			
Lys	Lys	Ile	Glu	Cys	Phe	Asp	Ser	Val	Glu	Ile	Ser	Gly	Val	Glu	Asp
	610					615					620				
Arg	Phe	Asn	Ala	Ser	Leu	Gly	Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile
625					630				635						640
Lys	Asp	Lys	Asp	Phe	Leu	Asp	Asn	Glu	Glu	Asn	Glu	Asp	Ile	Leu	Glu
				645					650					655	
Asp	Ile	Val	Leu	Thr	Leu	Thr	Leu	Phe	Glu	Asp	Arg	Glu	Met	Ile	Glu
		660						665					670		
Glu	Arg	Leu	Lys	Thr	Tyr	Ala	His	Leu	Phe	Asp	Asp	Lys	Val	Met	Lys
	675						680					685			
Gln	Leu	Lys	Arg	Arg	Arg	Tyr	Thr	Gly	Trp	Gly	Arg	Leu	Ser	Arg	Lys
	690					695					700				
Leu	Ile	Asn	Gly	Ile	Arg	Asp	Lys	Gln	Ser	Gly	Lys	Thr	Ile	Leu	Asp
705					710					715					720
Phe	Leu	Lys	Ser	Asp	Gly	Phe	Ala	Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile
				725					730					735	
His	Asp	Asp	Ser	Leu	Thr	Phe	Lys	Glu	Asp	Ile	Gln	Lys	Ala	Gln	Val
			740					745					750		
Ser	Gly	Gln	Gly	Asp	Ser	Leu	His	Glu	His	Ile	Ala	Asn	Leu	Ala	Gly
	755						760					765			
Ser	Pro	Ala	Ile	Lys	Lys	Gly	Ile	Leu	Gln	Thr	Val	Lys	Val	Val	Asp
	770					775					780				
Glu	Leu	Val	Lys	Val	Met	Gly	Arg	His	Lys	Pro	Glu	Asn	Ile	Val	Ile
785					790				795						800
Glu	Met	Ala	Arg	Glu	Asn	Gln	Thr	Thr	Gln	Lys	Gly	Gln	Lys	Asn	Ser
				805					810					815	

Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser
820 825 830

Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu
835 840 845

Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp
850 855 860

Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile
865 870 875 880

Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu
885 890 895

Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu
900 905 910

Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala
915 920 925

Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg
930 935 940

Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu
945 950 955 960

Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser
965 970 975

Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val
980 985 990

Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys A
995 1000 1005

Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala
1010 1015 1020

His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys
1025 1030 1035

Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys
1040 1045 1050

Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile
1055 1060 1065

Gly	Lys	Ala	Thr	Ala	Lys	Tyr	Phe	Phe	Tyr	Ser	Asn	Ile	Met	Asn
1070						1075					1080			
Phe	Phe	Lys	Thr	Glu	Ile	Thr	Leu	Ala	Asn	Gly	Glu	Ile	Arg	Lys
1085						1090					1095			
Arg	Pro	Leu	Ile	Glu	Thr	Asn	Gly	Glu	Thr	Gly	Glu	Ile	Val	Trp
1100						1105					1110			
Asp	Lys	Gly	Arg	Asp	Phe	Ala	Thr	Val	Arg	Lys	Val	Leu	Ser	Met
1115						1120					1125			
Pro	Gln	Val	Asn	Ile	Val	Lys	Lys	Thr	Glu	Val	Gln	Thr	Gly	Gly
1130						1135					1140			
Phe	Ser	Lys	Glu	Ser	Ile	Leu	Pro	Lys	Arg	Asn	Ser	Asp	Lys	Leu
1145						1150					1155			
Ile	Ala	Arg	Lys	Lys	Asp	Trp	Asp	Pro	Lys	Lys	Tyr	Gly	Gly	Phe
1160						1165					1170			
Asp	Ser	Pro	Thr	Val	Ala	Tyr	Ser	Val	Leu	Val	Val	Ala	Lys	Val
1175						1180					1185			
Glu	Lys	Gly	Lys	Ser	Lys	Lys	Leu	Lys	Ser	Val	Lys	Glu	Leu	Leu
1190						1195					1200			
Gly	Ile	Thr	Ile	Met	Glu	Arg	Ser	Ser	Phe	Glu	Lys	Asn	Pro	Ile
1205						1210					1215			
Asp	Phe	Leu	Glu	Ala	Lys	Gly	Tyr	Lys	Glu	Val	Lys	Lys	Asp	Leu
1220						1225					1230			
Ile	Ile	Lys	Leu	Pro	Lys	Tyr	Ser	Leu	Phe	Glu	Leu	Glu	Asn	Gly
1235						1240					1245			
Arg	Lys	Arg	Met	Leu	Ala	Ser	Ala	Gly	Glu	Leu	Gln	Lys	Gly	Asn
1250						1255					1260			
Glu	Leu	Ala	Leu	Pro	Ser	Lys	Tyr	Val	Asn	Phe	Leu	Tyr	Leu	Ala
1265						1270					1275			
Ser	His	Tyr	Glu	Lys	Leu	Lys	Gly	Ser	Pro	Glu	Asp	Asn	Glu	Gln
1280						1285					1290			
Lys	Gln	Leu	Phe	Val	Glu	Gln	His	Lys	His	Tyr	Leu	Asp	Glu	Ile
1295						1300					1305			

Ile Glu	Gln Ile Ser Glu Phe	Ser Lys Arg Val	Ile Leu Ala Asp
1310	1315		1320
Ala Asn	Leu Asp Lys Val Leu	Ser Ala Tyr Asn Lys	His Arg Asp
1325	1330		1335
Lys Pro	Ile Arg Glu Gln Ala	Glu Asn Ile Ile His	Leu Phe Thr
1340	1345		1350
Leu Thr	Asn Leu Gly Ala Pro	Ala Ala Phe Lys Tyr	Phe Asp Thr
1355	1360		1365
Thr Ile	Asp Arg Lys Arg Tyr	Thr Ser Thr Lys Glu	Val Leu Asp
1370	1375		1380
Ala Thr	Leu Ile His Gln Ser	Ile Thr Gly Leu Tyr	Glu Thr Arg
1385	1390		1395
Ile Asp	Leu Ser Gln Leu Gly	Gly Asp Ala Ala Ala	Val Ser Lys
1400	1405		1410
Gly Glu	Glu Leu Phe Thr Gly	Val Val Pro Ile Leu	Val Glu Leu
1415	1420		1425
Asp Gly	Asp Val Asn Gly His	Lys Phe Ser Val Ser	Gly Glu Gly
1430	1435		1440
Glu Gly	Asp Ala Thr Tyr Gly	Lys Leu Thr Leu Lys	Phe Ile Cys
1445	1450		1455
Thr Thr	Gly Lys Leu Pro Val	Pro Trp Pro Thr Leu	Val Thr Thr
1460	1465		1470
Leu Thr	Tyr Gly Val Gln Cys	Phe Ser Arg Tyr Pro	Asp His Met
1475	1480		1485
Lys Gln	His Asp Phe Phe Lys	Ser Ala Met Pro Glu	Gly Tyr Val
1490	1495		1500
Gln Glu	Arg Thr Ile Phe Phe	Lys Asp Asp Gly Asn	Tyr Lys Thr
1505	1510		1515
Arg Ala	Glu Val Lys Phe Glu	Gly Asp Thr Leu Val	Asn Arg Ile
1520	1525		1530
Glu Leu	Lys Gly Ile Asp Phe	Lys Glu Asp Gly Asn	Ile Leu Gly

1535		1540		1545
His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met				
1550		1555		1560
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg				
1565		1570		1575
His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln				
1580		1585		1590
Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn				
1595		1600		1605
His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu				
1610		1615		1620
Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly				
1625		1630		1635
Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Lys Arg Pro Ala Ala				
1640		1645		1650
Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys				
1655		1660		

- 5 <210> 46
 <211> 1423
 <212> PRT
 <213> Secuencia Artificial
- 10 <220>
 <221> fuente
 <223> /nota="Descripción de Secuencia Artificial: Polipéptido sintético"
- <400> 46

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
20 25 30

Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly Leu
35 40 45

Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr
50 55 60

Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His

ES 2 553 782 T3

65		70		75		80									
Ser	Ile	Lys	Lys	Asn	Leu	Ile	Gly	Ala	Leu	Leu	Phe	Asp	Ser	Gly	Glu
				85					90					95	
Thr	Ala	Glu	Ala	Thr	Arg	Leu	Lys	Arg	Thr	Ala	Arg	Arg	Arg	Tyr	Thr
			100					105						110	
Arg	Arg	Lys	Asn	Arg	Ile	Cys	Tyr	Leu	Gln	Glu	Ile	Phe	Ser	Asn	Glu
		115					120					125			
Met	Ala	Lys	Val	Asp	Asp	Ser	Phe	Phe	His	Arg	Leu	Glu	Glu	Ser	Phe
	130					135					140				
Leu	Val	Glu	Glu	Asp	Lys	Lys	His	Glu	Arg	His	Pro	Ile	Phe	Gly	Asn
145					150					155					160
Ile	Val	Asp	Glu	Val	Ala	Tyr	His	Glu	Lys	Tyr	Pro	Thr	Ile	Tyr	His
				165					170					175	
Leu	Arg	Lys	Lys	Leu	Val	Asp	Ser	Thr	Asp	Lys	Ala	Asp	Leu	Arg	Leu
			180					185					190		
Ile	Tyr	Leu	Ala	Leu	Ala	His	Met	Ile	Lys	Phe	Arg	Gly	His	Phe	Leu
	195						200					205			
Ile	Glu	Gly	Asp	Leu	Asn	Pro	Asp	Asn	Ser	Asp	Val	Asp	Lys	Leu	Phe
	210					215					220				
Ile	Gln	Leu	Val	Gln	Thr	Tyr	Asn	Gln	Leu	Phe	Glu	Glu	Asn	Pro	Ile
225					230					235					240
Asn	Ala	Ser	Gly	Val	Asp	Ala	Lys	Ala	Ile	Leu	Ser	Ala	Arg	Leu	Ser
				245					250					255	
Lys	Ser	Arg	Arg	Leu	Glu	Asn	Leu	Ile	Ala	Gln	Leu	Pro	Gly	Glu	Lys
			260					265					270		
Lys	Asn	Gly	Leu	Phe	Gly	Asn	Leu	Ile	Ala	Leu	Ser	Leu	Gly	Leu	Thr
	275						280					285			
Pro	Asn	Phe	Lys	Ser	Asn	Phe	Asp	Leu	Ala	Glu	Asp	Ala	Lys	Leu	Gln
	290					295					300				
Leu	Ser	Lys	Asp	Thr	Tyr	Asp	Asp	Asp	Leu	Asp	Asn	Leu	Leu	Ala	Gln
305					310					315					320

Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser
325 330 335

Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr
340 345 350

Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His
355 360 365

Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu
370 375 380

Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly
385 390 395 400

Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys
405 410 415

Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu
420 425 430

Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser
435 440 445

Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg
450 455 460

Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu
465 470 475 480

Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg
485 490 495

Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile
500 505 510

Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln
515 520 525

Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu
530 535 540

Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr
545 550 555 560

Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro
565 570 575

Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe
 580 585 590
 Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe
 595 600 605
 Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp
 610 615 620
 Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile
 625 630 635 640
 Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu
 645 650 655
 Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu
 660 665 670
 Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys
 675 680 685
 Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys
 690 695 700
 Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp
 705 710 715 720
 Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile
 725 730 735
 His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val
 740 745 750
 Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly
 755 760 765
 Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp
 770 775 780
 Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile
 785 790 795 800
 Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser
 805 810 815
 Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser
 820 825 830

Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu
 835 840 845

Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp
 850 855 860

Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile
 865 870 875 880

Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu
 885 890 895

Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu
 900 905 910

Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala
 915 920 925

Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg
 930 935 940

Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu
 945 950 955 960

Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser
 965 970 975

Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val
 980 985 990

Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp
 995 1000 1005

Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala
 1010 1015 1020

His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys
 1025 1030 1035

Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys
 1040 1045 1050

Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile
 1055 1060 1065

Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn

1070		1075		1080
Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys 1085 1090 1095				
Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp 1100 1105 1110				
Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met 1115 1120 1125				
Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly 1130 1135 1140				
Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu 1145 1150 1155				
Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe 1160 1165 1170				
Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val 1175 1180 1185				
Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu 1190 1195 1200				
Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile 1205 1210 1215				
Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu 1220 1225 1230				
Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly 1235 1240 1245				
Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn 1250 1255 1260				
Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala 1265 1270 1275				
Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln 1280 1285 1290				
Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp Glu Ile 1295 1300 1305				

Ile	Glu	Gln	Ile	Ser	Glu	Phe	Ser	Lys	Arg	Val	Ile	Leu	Ala	Asp
1310						1315					1320			
Ala	Asn	Leu	Asp	Lys	Val	Leu	Ser	Ala	Tyr	Asn	Lys	His	Arg	Asp
1325						1330					1335			
Lys	Pro	Ile	Arg	Glu	Gln	Ala	Glu	Asn	Ile	Ile	His	Leu	Phe	Thr
1340						1345					1350			
Leu	Thr	Asn	Leu	Gly	Ala	Pro	Ala	Ala	Phe	Lys	Tyr	Phe	Asp	Thr
1355						1360					1365			
Thr	Ile	Asp	Arg	Lys	Arg	Tyr	Thr	Ser	Thr	Lys	Glu	Val	Leu	Asp
1370						1375					1380			
Ala	Thr	Leu	Ile	His	Gln	Ser	Ile	Thr	Gly	Leu	Tyr	Glu	Thr	Arg
1385						1390					1395			
Ile	Asp	Leu	Ser	Gln	Leu	Gly	Gly	Asp	Lys	Arg	Pro	Ala	Ala	Thr
1400						1405					1410			
Lys	Lys	Ala	Gly	Gln	Ala	Lys	Lys	Lys	Lys					
1415						1420								

- 5 <210> 47
 <211> 483
 <212> PRT
 <213> Secuencia Artificial
- 10 <220>
 <221> fuente
 <223> /nota="Descripción de Secuencia Artificial: Polipéptido sintético"

ES 2 553 782 T3

<400> 47

Met Phe Leu Phe Leu Ser Leu Thr Ser Phe Leu Ser Ser Ser Arg Thr
1 5 10 15

Leu Val Ser Lys Gly Glu Glu Asp Asn Met Ala Ile Ile Lys Glu Phe
20 25 30

Met Arg Phe Lys Val His Met Glu Gly Ser Val Asn Gly His Glu Phe
35 40 45

Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr
50 55 60

Ala Lys Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp
65 70 75 80

Ile	Leu	Ser	Pro	Gln	Phe	Met	Tyr	Gly	Ser	Lys	Ala	Tyr	Val	Lys	His	85	90	95
Pro	Ala	Asp	Ile	Pro	Asp	Tyr	Leu	Lys	Leu	Ser	Phe	Pro	Glu	Gly	Phe	100	105	110
Lys	Trp	Glu	Arg	Val	Met	Asn	Phe	Glu	Asp	Gly	Gly	Val	Val	Thr	Val	115	120	125
Thr	Gln	Asp	Ser	Ser	Leu	Gln	Asp	Gly	Glu	Phe	Ile	Tyr	Lys	Val	Lys	130	135	140
Leu	Arg	Gly	Thr	Asn	Phe	Pro	Ser	Asp	Gly	Pro	Val	Met	Gln	Lys	Lys	145	150	155
Thr	Met	Gly	Trp	Glu	Ala	Ser	Ser	Glu	Arg	Met	Tyr	Pro	Glu	Asp	Gly	165	170	175
Ala	Leu	Lys	Gly	Glu	Ile	Lys	Gln	Arg	Leu	Lys	Leu	Lys	Asp	Gly	Gly	180	185	190
His	Tyr	Asp	Ala	Glu	Val	Lys	Thr	Thr	Tyr	Lys	Ala	Lys	Lys	Pro	Val	195	200	205
Gln	Leu	Pro	Gly	Ala	Tyr	Asn	Val	Asn	Ile	Lys	Leu	Asp	Ile	Thr	Ser	210	215	220
His	Asn	Glu	Asp	Tyr	Thr	Ile	Val	Glu	Gln	Tyr	Glu	Arg	Ala	Glu	Gly	225	230	235
Arg	His	Ser	Thr	Gly	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Gly	Ser	Lys	Gln	245	250	255
Leu	Glu	Glu	Leu	Leu	Ser	Thr	Ser	Phe	Asp	Ile	Gln	Phe	Asn	Asp	Leu	260	265	270
Thr	Leu	Leu	Glu	Thr	Ala	Phe	Thr	His	Thr	Ser	Tyr	Ala	Asn	Glu	His	275	280	285
Arg	Leu	Leu	Asn	Val	Ser	His	Asn	Glu	Arg	Leu	Glu	Phe	Leu	Gly	Asp	290	295	300
Ala	Val	Leu	Gln	Leu	Ile	Ile	Ser	Glu	Tyr	Leu	Phe	Ala	Lys	Tyr	Pro	305	310	315
Lys	Lys	Thr	Glu	Gly	Asp	Met	Ser	Lys	Leu	Arg	Ser	Met	Ile	Val	Arg	325	330	335

Glu Glu Ser Leu Ala Gly Phe Ser Arg Phe Cys Ser Phe Asp Ala Tyr
 340 345 350

Ile Lys Leu Gly Lys Gly Glu Glu Lys Ser Gly Gly Arg Arg Arg Asp
 355 360 365

Thr Ile Leu Gly Asp Leu Phe Glu Ala Phe Leu Gly Ala Leu Leu Leu
 370 375 380

Asp Lys Gly Ile Asp Ala Val Arg Arg Phe Leu Lys Gln Val Met Ile
 385 390 395 400

Pro Gln Val Glu Lys Gly Asn Phe Glu Arg Val Lys Asp Tyr Lys Thr
 405 410 415

Cys Leu Gln Glu Phe Leu Gln Thr Lys Gly Asp Val Ala Ile Asp Tyr
 420 425 430

Gln Val Ile Ser Glu Lys Gly Pro Ala His Ala Lys Gln Phe Glu Val
 435 440 445

Ser Ile Val Val Asn Gly Ala Val Leu Ser Lys Gly Leu Gly Lys Ser
 450 455 460

Lys Lys Leu Ala Glu Gln Asp Ala Ala Lys Asn Ala Leu Ala Gln Leu
 465 470 475 480

Ser Glu Val

<210> 48

< 211> 483

< 212> PRT

5 < 213> Secuencia Artificial

<220>

< 221> fuente

< 223> /nota="Descripción de Secuencia Artificial: Polipéptido sintético"

<400> 48

Met Lys Gln Leu Glu Glu Leu Leu Ser Thr Ser Phe Asp Ile Gln Phe
 1 5 10 15

Asn Asp Leu Thr Leu Leu Glu Thr Ala Phe Thr His Thr Ser Tyr Ala
 20 25 30

Asn Glu His Arg Leu Leu Asn Val Ser His Asn Glu Arg Leu Glu Phe
 35 40 45

ES 2 553 782 T3

Leu Gly Asp Ala Val Leu Gln Leu Ile Ile Ser Glu Tyr Leu Phe Ala
 50 55 60
 Lys Tyr Pro Lys Lys Thr Glu Gly Asp Met Ser Lys Leu Arg Ser Met
 65 70 75 80
 Ile Val Arg Glu Glu Ser Leu Ala Gly Phe Ser Arg Phe Cys Ser Phe
 85 90 95
 Asp Ala Tyr Ile Lys Leu Gly Lys Gly Glu Glu Lys Ser Gly Gly Arg
 100 105 110
 Arg Arg Asp Thr Ile Leu Gly Asp Leu Phe Glu Ala Phe Leu Gly Ala
 115 120 125
 Leu Leu Leu Asp Lys Gly Ile Asp Ala Val Arg Arg Phe Leu Lys Gln
 130 135 140
 Val Met Ile Pro Gln Val Glu Lys Gly Asn Phe Glu Arg Val Lys Asp
 145 150 155 160
 Tyr Lys Thr Cys Leu Gln Glu Phe Leu Gln Thr Lys Gly Asp Val Ala
 165 170 175
 Ile Asp Tyr Gln Val Ile Ser Glu Lys Gly Pro Ala His Ala Lys Gln
 180 185 190
 Phe Glu Val Ser Ile Val Val Asn Gly Ala Val Leu Ser Lys Gly Leu
 195 200 205
 Gly Lys Ser Lys Lys Leu Ala Glu Gln Asp Ala Ala Lys Asn Ala Leu
 210 215 220
 Ala Gln Leu Ser Glu Val Gly Ser Val Ser Lys Gly Glu Glu Asp Asn
 225 230 235 240
 Met Ala Ile Ile Lys Glu Phe Met Arg Phe Lys Val His Met Glu Gly
 245 250 255
 Ser Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu Gly Glu Gly Arg
 260 265 270
 Pro Tyr Glu Gly Thr Gln Thr Ala Lys Leu Lys Val Thr Lys Gly Gly
 275 280 285
 Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Phe Met Tyr Gly
 290 295 300

Ser Lys Ala Tyr Val Lys His Pro Ala Asp Ile Pro Asp Tyr Leu Lys
305 310 315 320

Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val Met Asn Phe Glu
325 330 335

Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser Leu Gln Asp Gly
340 345 350

Glu Phe Ile Tyr Lys Val Lys Leu Arg Gly Thr Asn Phe Pro Ser Asp
355 360 365

Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu Ala Ser Ser Glu
370 375 380

Arg Met Tyr Pro Glu Asp Gly Ala Leu Lys Gly Glu Ile Lys Gln Arg
385 390 395 400

Leu Lys Leu Lys Asp Gly Gly His Tyr Asp Ala Glu Val Lys Thr Thr
405 410 415

Tyr Lys Ala Lys Lys Pro Val Gln Leu Pro Gly Ala Tyr Asn Val Asn
420 425 430

Ile Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr Thr Ile Val Glu
435 440 445

Gln Tyr Glu Arg Ala Glu Gly Arg His Ser Thr Gly Gly Met Asp Glu
450 455 460

Leu Tyr Lys Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys
465 470 475 480

Lys Lys Lys

<210> 49

<211> 1423

<212> PRT

5 <213> Secuencia Artificial

<220>

<221> fuente

<223> /nota="Descripción de Secuencia Artificial: Polipéptido sintético"

<400> 49

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
 20 25 30
 Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly Leu
 35 40 45
 Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr
 50 55 60
 Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His
 65 70 75 80
 Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu
 85 90 95
 Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr
 100 105 110
 Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu
 115 120 125
 Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe
 130 135 140
 Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn
 145 150 155 160
 Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His
 165 170 175
 Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu
 180 185 190
 Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu
 195 200 205
 Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe
 210 215 220
 Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile
 225 230 235 240
 Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser
 245 250 255
 Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys

260	265	270
Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr 275 280 285		
Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln 290 295 300		
Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln 305 310 315 320		
Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser 325 330 335		
Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr 340 345 350		
Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His 355 360 365		
Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu 370 375 380		
Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly 385 390 395 400		
Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys 405 410 415		
Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu 420 425 430		
Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser 435 440 445		
Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg 450 455 460		
Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu 465 470 475 480		
Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg 485 490 495		
Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile 500 505 510		

Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln
515 520 525

Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu
530 535 540

Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr
545 550 555 560

Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro
565 570 575

Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe
580 585 590

Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe
595 600 605

Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp
610 615 620

Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile
625 630 635 640

Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu
645 650 655

Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu
660 665 670

Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys
675 680 685

Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys
690 695 700

Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp
705 710 715 720

Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile
725 730 735

His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val
740 745 750

Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly
755 760 765

Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp
770 775 780

Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile
785 790 795 800

Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser
805 810 815

Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser
820 825 830

Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu
835 840 845

Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp
850 855 860

Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile
865 870 875 880

Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu
885 890 895

Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu
900 905 910

Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala
915 920 925

Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg
930 935 940

Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu
945 950 955 960

Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser
965 970 975

Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val
980 985 990

Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp
995 1000 1005

Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala
1010 1015 1020

His	Asp	Ala	Tyr	Leu	Asn	Ala	Val	Val	Gly	Thr	Ala	Leu	Ile	Lys
1025						1030					1035			
Lys	Tyr	Pro	Lys	Leu	Glu	Ser	Glu	Phe	Val	Tyr	Gly	Asp	Tyr	Lys
1040						1045					1050			
Val	Tyr	Asp	Val	Arg	Lys	Met	Ile	Ala	Lys	Ser	Glu	Gln	Glu	Ile
1055						1060					1065			
Gly	Lys	Ala	Thr	Ala	Lys	Tyr	Phe	Phe	Tyr	Ser	Asn	Ile	Met	Asn
1070						1075					1080			
Phe	Phe	Lys	Thr	Glu	Ile	Thr	Leu	Ala	Asn	Gly	Glu	Ile	Arg	Lys
1085						1090					1095			
Arg	Pro	Leu	Ile	Glu	Thr	Asn	Gly	Glu	Thr	Gly	Glu	Ile	Val	Trp
1100						1105					1110			
Asp	Lys	Gly	Arg	Asp	Phe	Ala	Thr	Val	Arg	Lys	Val	Leu	Ser	Met
1115						1120					1125			
Pro	Gln	Val	Asn	Ile	Val	Lys	Lys	Thr	Glu	Val	Gln	Thr	Gly	Gly
1130						1135					1140			
Phe	Ser	Lys	Glu	Ser	Ile	Leu	Pro	Lys	Arg	Asn	Ser	Asp	Lys	Leu
1145						1150					1155			
Ile	Ala	Arg	Lys	Lys	Asp	Trp	Asp	Pro	Lys	Lys	Tyr	Gly	Gly	Phe
1160						1165					1170			
Asp	Ser	Pro	Thr	Val	Ala	Tyr	Ser	Val	Leu	Val	Val	Ala	Lys	Val
1175						1180					1185			
Glu	Lys	Gly	Lys	Ser	Lys	Lys	Leu	Lys	Ser	Val	Lys	Glu	Leu	Leu
1190						1195					1200			
Gly	Ile	Thr	Ile	Met	Glu	Arg	Ser	Ser	Phe	Glu	Lys	Asn	Pro	Ile
1205						1210					1215			
Asp	Phe	Leu	Glu	Ala	Lys	Gly	Tyr	Lys	Glu	Val	Lys	Lys	Asp	Leu
1220						1225					1230			
Ile	Ile	Lys	Leu	Pro	Lys	Tyr	Ser	Leu	Phe	Glu	Leu	Glu	Asn	Gly
1235						1240					1245			
Arg	Lys	Arg	Met	Leu	Ala	Ser	Ala	Gly	Glu	Leu	Gln	Lys	Gly	Asn

1250		1255		1260
Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala				
1265		1270		1275
Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln				
1280		1285		1290
Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp Glu Ile				
1295		1300		1305
Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp				
1310		1315		1320
Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp				
1325		1330		1335
Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr				
1340		1345		1350
Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr				
1355		1360		1365
Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp				
1370		1375		1380
Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg				
1385		1390		1395
Ile Asp Leu Ser Gln Leu Gly Gly Asp Lys Arg Pro Ala Ala Thr				
1400		1405		1410
Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys				
1415		1420		

5

<210> 50
 <211> 2012
 <212> ADN
 <213> Secuencia Artificial

10

<220>
 <221> fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

ES 2 553 782 T3

<400> 50

gaatgctgcc ctcagacccg cttectccct gtccctgtct gtccaaggag aatgaggtct	60
cactggtgga ttccggacta ccctgaggag ctggcaactg agggacaagg cccccaact	120
gccagctcc agcctctgat gaggggtggg agagagctac atgaggttgc taagaaagcc	180

5

tcccctgaag gagaccacac agtgtgtgag gttggagtct ctagcagcgg gttctgtgcc	240
cccagggata gtctggctgt ccaggcactg ctcttgatat aaacaccacc tcctagttat	300
gaaaccatgc ccattctgcc tctctgtatg gaaaagagca tggggctggc ccgtggggtg	360
gtgtccactt taggccctgt gggagatcat gggaaaccac gcagtgggtc ataggctctc	420
tcatttacta ctcacatcca ctctgtgaag aagcgattat gatctctcct ctagaaactc	480
gtagagtccc atgtctgccg gcttccagag cctgcactcc tccaccttgg cttggctttg	540
ctggggctag aggagctagg atgcacagca gctctgtgac cctttgtttg agaggaacag	600
gaaaaccacc cttctctctg gccactgtg tcctcttcct gccctgccat ccccttctgt	660
gaatgttaga cccatgggag cagctggtca gaggggaccc cggcctgggg cccctaaccc	720
tatgtagcct cagtcttccc atcaggctct cagctcagcc tgagtgttga ggccccagt	780
gctgctctgg gggcctcctg agtttctcat ctgtgcccct ccctccctgg ccaggtgaa	840
ggtgtggttc cagaaccgga ggacaaagta caaacggcag aagctggagg aggaagggcc	900
tgagtccgag cagaagaaga agggctccca tcacatcaac cggtggcgca ttgccacgaa	960
gcaggccaat ggggaggaca tcgatgtcac ctccaatgac aagcttgcta gcggtgggca	1020
accacaaacc cacgagggca gagtgtctgt tgctgtctggc caggccccctg cgtgggcccc	1080
agctggactc tggccactcc ctggccaggc tttggggagg cctggagtca tggccccaca	1140
gggcttgaag cccggggccg ccattgacag agggacaagc aatgggctgg ctgaggcctg	1200
ggaccacttg gccttctcct cggagagcct gcctgcctgg gcgggcccgc ccgccaccgc	1260
agcctcccag ctgctctccg tgtctccaat ctcccttttg ttttgatgca tttctgtttt	1320
aatttatattt ccaggcacca ctgtagttta gtgatcccca gtgtccccct tccctatggg	1380
aataataaaa gtctctctct taatgacacg ggcattccagc tccagcccc	1440
tggtagattc cggctctgag ggccagtggg ggctggtaga gcaaacgcgt tcagggcctg	1500
ggagcctggg gtgggggtact ggtggagggg gtcaagggta attcattaac tcctctcttt	1560
tgttggggga ccctggtctc tacctccagc tccacagcag gagaaacagg ctagacatag	1620
ggaagggcca tcctgtatct tgaggaggga caggcccagg tctttcttaa cgtattgaga	1680
ggtgggaatc aggccaggt agttcaatgg gagagggaga gtgcttcct ctgcctagag	1740
actctggtgg cttctccagt tgaggagaaa ccagaggaaa ggggaggatt ggggtctggg	1800
ggagggaaca ccattcacia aggctgacgg ttccagtccg aagtcgtggg cccaccagga	1860
tgtcacctg tccttggaaga accgtgggc aggttgagac tgcagagaca gggcttaagg	1920
ctgagcctgc aaccagtccc cagtgactca gggcctcctc agcccaagaa agagcaacgt	1980
gccagggccc gctgagctct tgtgttcacc tg	2012

<210> 51
< 211> 1153
< 212> PRT
5 < 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de secuencia artificial: polipéptido sintético"

<400> 51

Met Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
 1 5 10 15

Lys Ser Asp Leu Val Leu Gly Leu Asp Ile Gly Ile Gly Ser Val Gly
 20 25 30

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

Arg Ile Phe Pro Ala Ala Gln Ala Glu Asn Asn Leu Val Arg Arg Thr
 50 55 60

Asn Arg Gln Gly Arg Arg Leu Ala Arg Arg Lys Lys His Arg Arg Val
 65 70 75 80

Arg Leu Asn Arg Leu Phe Glu Glu Ser Gly Leu Ile Thr Asp Phe Thr
 85 90 95

Lys Ile Ser Ile Asn Leu Asn Pro Tyr Gln Leu Arg Val Lys Gly Leu
 100 105 110

Thr Asp Glu Leu Ser Asn Glu Glu Leu Phe Ile Ala Leu Lys Asn Met
 115 120 125

Val Lys His Arg Gly Ile Ser Tyr Leu Asp Asp Ala Ser Asp Asp Gly
 130 135 140

Asn Ser Ser Val Gly Asp Tyr Ala Gln Ile Val Lys Glu Asn Ser Lys
 145 150 155 160

Gln Leu Glu Thr Lys Thr Pro Gly Gln Ile Gln Leu Glu Arg Tyr Gln
 165 170 175

Thr Tyr Gly Gln Leu Arg Gly Asp Phe Thr Val Glu Lys Asp Gly Lys
 180 185 190

Lys His Arg Leu Ile Asn Val Phe Pro Thr Ser Ala Tyr Arg Ser Glu
 195 200 205

Ile Leu Thr Arg Leu Gly Lys Gln Lys Thr Thr Ser Ser Ser Asn Lys
465 470 475 480

Thr Lys Tyr Ile Asp Glu Lys Leu Leu Thr Glu Glu Ile Tyr Asn Pro
485 490 495

Val Val Ala Lys Ser Val Arg Gln Ala Ile Lys Ile Val Asn Ala Ala
500 505 510

Ile Lys Glu Tyr Gly Asp Phe Asp Asn Ile Val Ile Glu Met Ala Arg
515 520 525

Glu Thr Asn Glu Asp Asp Glu Lys Lys Ala Ile Gln Lys Ile Gln Lys
530 535 540

Ala Asn Lys Asp Glu Lys Asp Ala Ala Met Leu Lys Ala Ala Asn Gln
545 550 555 560

Tyr Asn Gly Lys Ala Glu Leu Pro His Ser Val Phe His Gly His Lys
565 570 575

Gln Leu Ala Thr Lys Ile Arg Leu Trp His Gln Gln Gly Glu Arg Cys
580 585 590

Leu Tyr Thr Gly Lys Thr Ile Ser Ile His Asp Leu Ile Asn Asn Ser
595 600 605

Asn Gln Phe Glu Val Asp His Ile Leu Pro Leu Ser Ile Thr Phe Asp
610 615 620

Asp Ser Leu Ala Asn Lys Val Leu Val Tyr Ala Thr Ala Asn Gln Glu
625 630 635 640

Lys Gly Gln Arg Thr Pro Tyr Gln Ala Leu Asp Ser Met Asp Asp Ala
645 650 655

Trp Ser Phe Arg Glu Leu Lys Ala Phe Val Arg Glu Ser Lys Thr Leu
660 665 670

Ser Asn Lys Lys Lys Glu Tyr Leu Leu Thr Glu Glu Asp Ile Ser Lys
675 680 685

Phe Asp Val Arg Lys Lys Phe Ile Glu Arg Asn Leu Val Asp Thr Arg
690 695 700

Tyr Ala Ser Arg Val Val Leu Asn Ala Leu Gln Glu His Phe Arg Ala

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

His Ile Asp Ile Thr Pro Lys Asp Ser Asn Asn Lys Val Val Leu Gln
965 970 975

Ser Val Ser Pro Trp Arg Ala Asp Val Tyr Phe Asn Lys Thr Thr Gly
980 985 990

Lys Tyr Glu Ile Leu Gly Leu Lys Tyr Ala Asp Leu Gln Phe Glu Lys
995 1000 1005

Gly Thr Gly Thr Tyr Lys Ile Ser Gln Glu Lys Tyr Asn Asp Ile
1010 1015 1020

Lys Lys Lys Glu Gly Val Asp Ser Asp Ser Glu Phe Lys Phe Thr
1025 1030 1035

Leu Tyr Lys Asn Asp Leu Leu Leu Val Lys Asp Thr Glu Thr Lys
1040 1045 1050

Glu Gln Gln Leu Phe Arg Phe Leu Ser Arg Thr Met Pro Lys Gln
1055 1060 1065

Lys His Tyr Val Glu Leu Lys Pro Tyr Asp Lys Gln Lys Phe Glu
1070 1075 1080

Gly Gly Glu Ala Leu Ile Lys Val Leu Gly Asn Val Ala Asn Ser
1085 1090 1095

Gly Gln Cys Lys Lys Gly Leu Gly Lys Ser Asn Ile Ser Ile Tyr
1100 1105 1110

Lys Val Arg Thr Asp Val Leu Gly Asn Gln His Ile Ile Lys Asn
1115 1120 1125

Glu Gly Asp Lys Pro Lys Leu Asp Phe Lys Arg Pro Ala Ala Thr
1130 1135 1140

Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1145 1150

<210> 52

<211> 340

<212> ADN

<213> Secuencia Artificial

<220>

<221> fuente

<223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 52
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag 60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180
atgcttaccg taacttgaaa gtatttcgat ttcttggcctt tatatatctt gtggaaagga 240
cgaacaccg ttacttaaat cttgcagaag ctacaaagat aaggcttcat gccgaaatca 300
acaccctgtc attttatggc aggggtgttt cgttatttaa 340

<210> 53
 < 211> 360
 5 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

10 <220>
 < 221> base modificada
 < 222> (288)..(317)
 < 223> a, c, t, g, Desconocido u otro

<400> 53
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag 60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180
atgcttaccg taacttgaaa gtatttcgat ttcttggcctt tatatatctt gtggaaagga 240
cgaacaccg ggttttagag ctatgctgtt ttgaatggc ccaaaacnnn nnnnnnnnnn 300
nnnnnnnnnn nnnnnnngtt ttagagctat gctgttttga atggtcccaa aacttttttt 360

15 <210> 54
 < 211> 318
 < 212> ADN
 < 213> Secuencia Artificial
 20 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

25 <220>
 < 221> base modificada
 < 222> (250).. (269)
 < 223> a, c, t, g, Desconocido u otro

<400> 54
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag 60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120

aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180
atgcttaccg taacttgaaa gtatttcgat ttcttggcctt tatatatctt gtggaaagga 240
cgaaacaccn nnnnnnnnnn nnnnnnnnng ttttagagct agaaatagca agttaaata 300
aggctagtcc gttttttt 318

5 <210> 55
 <211> 325
 <212> ADN
 <213> Secuencia Artificial

<220>
 <221> fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

10 <220>
 <221> base modificada
 <222> (250)..(269)
 <223> a, c, t, g, Desconocido u otro

<400> 55
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag 60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180
atgcttaccg taacttgaaa gtatttcgat ttcttggcctt tatatatctt gtggaaagga 240
cgaaacaccn nnnnnnnnnn nnnnnnnnng ttttagagct agaaatagca agttaaata 300
aggctagtcc gttatcattt ttttt 325

15 <210> 56
 <211> 337
 <212> ADN
 <213> Secuencia Artificial

20 <220>
 <221> fuente
 <223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

25 <220>
 <221> base modificada
 <222> (250)..(269)
 <223> a, c, t, g, Desconocido u otro

<400> 56

```

gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag      60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga      120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat      180
atgcttaccg taacttgaaa gtatttcgat ttcttggcct tatatatctt gtggaaagga      240
cgaaacaccn nnnnnnnnnn nnnnnnnnng ttttagagct agaaatagca agttaaata      300
aggctagtcc gttatcaact tgaaaaagtg tttttttt      337

```

5 <210> 57
 < 211> 352
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

10 <220>
 < 221> base modificada
 < 222> (250) .. (269)
 < 223> a, c, t, g, Desconocido u otro

```

<400> 57
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag      60
ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga      120
aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat      180
atgcttaccg taacttgaaa gtatttcgat ttcttggcct tatatatctt gtggaaagga      240
cgaaacaccn nnnnnnnnnn nnnnnnnnng ttttagagct agaaatagca agttaaata      300
aggctagtcc gttatcaact tgaaaaagtg gcaccgagtc ggtgcttttt tt      352

```

15 <210> 58
 < 211> 5101
 < 212> ADN
 < 213> Secuencia Artificial

20 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 58

cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc ccgcccatt	60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	120
atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc	180
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	300
catggtcgag gtgagcccca cgttctgctt cactctcccc atctcccccc cctccccacc	360
cccaattttg tattttattta ttttttaatt attttgtgca gcgatggggg cggggggggg	420
ggggggggcgc gcgccaggcg gggcggggcg gggcgagggg cggggcgggg cgaggcggag	480

aggtgcggcg	gcagccaatc	agagcggcgc	gctccgaaag	tttcctttta	tggcgaggcg	540
gcggcggcgg	cggccctata	aaaagcgaag	cgcgcggcgg	gcgggagtcg	ctgcgacgct	600
gccttcgccc	cgtgccccgc	tccgcgcgcg	cctcgcgcgcg	cccgcgcccg	ctctgactga	660
ccgcgttact	cccacagggtg	agcggggcggg	acggcccttc	tcctccgggc	tgtaattagc	720
tgagcaagag	gtaagggttt	aagggatggt	tggttggtgg	ggtattaatg	tttaattacc	780
tggagcacct	gcctgaaatc	actttttttc	aggttggaac	ggtgccacca	tggactataa	840
ggaccacgac	ggagactaca	aggatcatga	tattgattac	aaagacgatg	acgataagat	900
ggccccaag	aagaagcga	aggtcgggat	ccacggagtc	ccagcagccg	acaagaagta	960
cagcatcggc	ctggacatcg	gcaccaactc	tgtgggctgg	gccgtgatca	ccgacgagta	1020
caaggtgcc	agcaagaaat	tcaaggtgct	gggcaacacc	gaccggcaca	gcatcaagaa	1080
gaacctgatc	ggagccctgc	tgttcgacag	cggcgaaaca	gccgaggcca	cccggctgaa	1140
gagaaccgcc	agaagaagat	acaccagacg	gaagaaccgg	atctgctatc	tgcaagagat	1200
cttcagcaac	gagatggcca	aggtggacga	cagcttcttc	cacagactgg	aagagtcctt	1260
cctggtggaa	gaggataaga	agcacgagcg	gcaccccatc	ttcggcaaca	tcgtggacga	1320
ggtggcctac	cacgagaagt	acccaccat	ctaccacctg	agaaagaaac	tggtggacag	1380
caccgacaag	gccgacctgc	ggctgatcta	tctggccctg	gccacatga	tcaagttccg	1440
gggccacttc	ctgatcgagg	gcgacctgaa	ccccgacaac	agcgacgtgg	acaagctggt	1500
catccagctg	gtgcagacct	acaaccagct	gttcgaggaa	aaccccatca	acgccagcgg	1560
cgtggacgcc	aaggccatcc	tgtctgccag	actgagcaag	agcagacggc	tggaaaatct	1620
gatcgcccag	ctgcccggcg	agaagaagaa	tggcctgttc	ggcaacctga	ttgccctgag	1680
cctgggcctg	acccccaact	tcaagagcaa	cttcgacctg	gccgaggatg	ccaaactgca	1740
gctgagcaag	gacacctacg	acgacgacct	ggacaacctg	ctggcccaga	tcggcgacca	1800
gtacgccgac	ctgtttctgg	ccgccaagaa	cctgtccgac	gccatcctgc	tgagcgacat	1860
cctgagagtg	aacaccgaga	tcaccaaggc	cccctgagc	gcctctatga	tcaagagata	1920
cgacgagcac	caccaggacc	tgacctgct	gaaagctctc	gtgcggcagc	agctgcctga	1980
gaagtacaaa	gagattttct	tcgaccagag	caagaacggc	tacgccggct	acattgacgg	2040
cggagccagc	caggaagagt	tctacaagtt	catcaagccc	atcctggaaa	agatggacgg	2100
caccgaggaa	ctgctcgtga	agctgaacag	agaggacctg	ctgcggaagc	agcggacctt	2160
cgacaacggc	agcatcccc	accagatcca	cctgggagag	ctgcacgcca	ttctgcggcg	2220
gcaggaagat	ttttacccat	tcctgaagga	caaccgggaa	aagatcgaga	agatcctgac	2280
cttccgcatac	ccctactacg	tggggccctct	ggccagggga	aacagcagat	tcgcctggat	2340

gaccagaaag agcgaggaaa ccatcacccc ctggaacttc gaggaagtgg tggacaaggg	2400
cgcttccgcc cagagcttca tcgagcggat gaccaacttc gataagaacc tgcccaacga	2460
gaaggtgctg cccaagcaca gcctgctgta cgagtacttc accgtgtata acgagctgac	2520
caaagtgaag tacgtgaccg agggaatgag aaagcccgcc ttcttgagcg gcgagcagaa	2580
aaaggccatc gtggacctgc tgttcaagac caaccgaaa gtgacctga agcagctgaa	2640
agaggactac ttcaagaaaa tcgagtgcct cgactccgtg gaaatctccg gcgtggaaga	2700
tcggttcaac gcctccctgg gcacatacca cgatctgctg aaaattatca aggacaagga	2760
cttctggac aatgaggaaa acgaggacat tctggaagat atcgtgctga ccctgacact	2820
gtttgaggac agagagatga tcgaggaacg gctgaaaacc tatgccacc tgttcgacga	2880
caaagtgatg aagcagctga agcggcggag atacaccggc tggggcaggc tgagccggaa	2940
gctgatcaac ggcacccggg acaagcagtc cggcaagaca atcctggatt tcctgaagtc	3000
cgacggcttc gccaacagaa acttcatgca gctgatccac gacgacagcc tgaccttaaa	3060
agaggacatc cagaaagccc aggtgtccgg ccaggcggat agcctgcacg agcacattgc	3120
caatctggcc ggcagccccg ccattaagaa gggcatcctg cagacagtga aggtggtgga	3180
cgagctcgtg aaagtgatgg gccggcacia gcccgagaac atcgtgatcg aaatggccag	3240
agagaaccag accaccaga agggacagaa gaacagccgc gagagaatga agcggatcga	3300
agagggcatc aaagagctgg gcagccagat cctgaaagaa cccccgtgg aaaacaccca	3360
gctgcagaac gagaagctgt acctgtacta cctgcagaat gggcgggata tgtacgtgga	3420
ccaggaactg gacatcaacc ggctgtccga ctacgatgtg gaccatatcg tgcctcagag	3480
ctttctgaag gacgactcca tcgacaacaa ggtgctgacc agaagcgaca agaaccgggg	3540
caagagcgac aacgtgccct ccgaagaggt cgtgaagaag atgaagaact actggcggca	3600
gctgctgaac gccaaactga ttaccagag aaagttcgac aatctgacca aggccgagag	3660
aggcggcctg agcgaactgg ataaggccgg ctatcatcaag agacagctgg tggaaacccg	3720
gcagatcaca aagcacgtgg cacagatcct ggactcccgg atgaacacta agtacgacga	3780
gaatgacaag ctgatccggg aagtgaaggt gatcaccttg aagtccaagc tgggtgtccga	3840
tttccggaag gatttccagt tttaaaaagt gcgcgagatc aacaactacc accacgcccc	3900
cgacgcctac ctgaacgccg tcgtgggaac cgccctgatc aaaaagtacc ctaagctgga	3960
aagcgagttc gtgtacggcg actacaaggt gtacgacgtg cggaagatga tcgccaagag	4020
cgagcaggaa atcggcaagg ctaccgcccc gtacttcttc tacagcaaca tcatgaactt	4080
tttcaagacc gagattaccc tggccaacgg cgagatccgg aagcggcctc tgatcgagac	4140
aaacggcgaa accggggaga tcgtgtggga taagggccgg gatctttgcca ccgtgcggaa	4200
agtgtgtgagc atgccccaaag tgaatatcgt gaaaaagacc gaggtgcaga caggcggctt	4260

cagcaaagag tctatcctgc ccaagaggaa cagcgataag ctgatcgcca gaaagaagga 4320
ctgggaccct aagaagtacg gcggcttcga cagccccacc gtggcctatt ctgtgctggt 4380
ggtggccaaa gtggaaaag gcaagtccaa gaaactgaag agtgtgaaag agctgctggg 4440
gatcaccatc atggaaagaa gcagcttcga gaagaatccc atcgactttc tggaagccaa 4500
gggctacaaa gaagtgaaaa aggacctgat catcaagctg cctaagtact ccctgttcga 4560
gctggaaaac ggccggaaga gaatgctggc ctctgccggc gaactgcaga agggaaacga 4620
actggccctg ccctccaaat atgtgaactt cctgtacctg gccagccact atgagaagct 4680
gaagggtcc cccgaggata atgagcagaa acagctgttt gtggaacagc acaagcacta 4740
cctggacgag atcatcgagc agatcagcga gttctccaag agagtgatcc tggccgacgc 4800
taatctggac aaagtgctgt ccgcctacaa caagcaccgg gataagccca tcagagagca 4860
ggccgagaat atcatccacc tgtttaccct gaccaatctg ggagcccctg ccgccttcaa 4920
gtactttgac accaccatcg accggaagag gtacaccagc accaaagagg tgctggacgc 4980
caccctgatc caccagagca tcaccggcct gtacgagaca cggatcgacc tgtctcagct 5040
gggaggcgac tttctttttc ttagcttgac cagctttctt agtagcagca ggacgcttta 5100
a 5101

<210> 59
< 211> 137
5 < 212> ADN
< 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

10 <220>
< 221> base modificada
< 222> (1)..(20)
< 223> a, c, t, g, Desconocido u otro

<400> 59
nnnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa 60
gctacaaaga taaggcttca tgccgaaatc aacaccctgt cattttatgg cagggtgttt 120
tcgttattta atttttt 137

15

<210> 60
< 211> 123
< 212> ADN
< 213> Secuencia Artificial

20 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

	<220> < 221> base modificada < 222> (1)..(20) < 223> a, c, t, g, Desconocido u otro	
5	<400> 60 nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcagaaat gcagaagcta caaagataag gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttcgt tatttaattt ttt	60 120 123
10	<210> 61 < 211> 110 < 212> ADN < 213> Secuencia Artificial <220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
15	<220> < 221> base modificada < 222> (1)..(20) < 223> a, c, t, g, Desconocido u otro <400> 61 nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcagaaat gcagaagcta caaagataag gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttttt	60 110
20	<210> 62 < 211> 137 < 212> ADN < 213> Secuencia Artificial	
25	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
30	<220> < 221> base modificada < 222> (1) .. (20) < 223> a, c, t, g, Desconocido u otro <400> 62 nnnnnnnnnn nnnnnnnnnn gttattgtac tctcaagatt tagaaataaa tcttgcagaa gctacaaaga taaggcttca tgccgaaatc aacaccctgt cattttatgg cagggtgttt tcgttattta atttttt	60 120 137
35	<210> 63 < 211> 123 < 212> ADN < 213> Secuencia Artificial	

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

5 <220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t, g, Desconocido u otro

<400> 63
 nnnnnnnnnn nnnnnnnnnn gttattgtac tctcagaaat gcagaagcta caaagataag 60
 gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttcgt tattaattt 120
 ttt 123

10 <210> 64
 < 211> 110
 < 212> ADN
 < 213> Secuencia Artificial

15 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

20 <220>
 < 221> base modificada
 < 222> (1)..(20)
 < 223> a, c, t, g, Desconocido u otro

<400> 64
 nnnnnnnnnn nnnnnnnnnn gttattgtac tctcagaaat gcagaagcta caaagataag 60
 gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttttt 110

25 <210> 65
 < 211> 137
 < 212> ADN
 < 213> Secuencia Artificial

30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

30 <220>
 < 221> base modificada
 < 222> (1) .. (20)
 < 223> a, c, t, g, Desconocido u otro

<400> 65
 nnnnnnnnnn nnnnnnnnnn gttattgtac tctcaagatt tagaaataaa tcttgcagaa 60
 gctacaatga taaggcttca tgccgaaatc aacaccctgt ctttttatgg cagggtgttt 120
 tcgttattta atttttt 137

35 <210> 66
 < 211> 123

< 212> ADN
< 213> Secuencia Artificial

<220>
< 221> fuente
5 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
< 221> base modificada
< 222> (1)..(20)
< 223> a, c, t, g, Desconocido u otro

10 <400> 66
nnnnnnnnnnn nnnnnnnnnn gttattgtac tctcagaaat gcagaagcta caatgataag 60
gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttcgt tatttaattt 120
ttt 123

<210> 67
< 211> 110
< 212> ADN
15 < 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
20 < 221> base modificada
< 222> (1)..(20)
< 223> a, c, t, g, Desconocido u otro

<400> 67
nnnnnnnnnnn nnnnnnnnnn gttattgtac tctcagaaat gcagaagcta caatgataag 60
gcttcatgcc gaaatcaaca ccctgtcatt ttatggcagg gtgttttttt 110

25 <210> 68
< 211> 107
< 212> ADN
< 213> Secuencia Artificial

<220>
30 < 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
< 221> base modificada
< 222> (1) .. (20)
35 < 223> a, c, t, g, Desconocido u otro

<400> 68
nnnnnnnnnnn nnnnnnnnnn gtttttagagc tgtggaaaca cagcgagtta aaataaggct 60
tagtcogtac tcaacttgaa aaggtggcac cgattcgggtg tttttttt 107

<210> 69
 < 211> 4263
 < 212> ADN
 < 213> Secuencia Artificial

5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<400> 69
 atgaaaaggc cggcggccac gaaaaaggcc ggccaggcaa aaaagaaaaa gaccaagccc 60
 tacagcatcg gcctggacat cggcaccaat agcgtgggct gggccgtgac caccgacaac 120
 tacaagggtgc ccagcaagaa aatgaagggtg ctgggcaaca cctccaagaa gtacatcaag 180
 aaaaacctgc tgggcgtgct gctgttcgac agcggcatta cagccgaggg cagacggctg 240
 aagagaaccg ccagacggcg gtacacccgg cggagaaaca gaatcctgta tctgcaagag 300
 atcttcagca ccgagatggc taccctggac gacgccttct tccagcggct ggacgacagc 360
 ttctgtgtgc ccgacgacaa gggggacagc aagtacccca tcttcggcaa cctggtggaa 420
 gagaaggcct accacgacga gttccccacc atctaccacc tgagaaagta cctggccgac 480
 agcaccaaga aggccgacct gagactggtg tatctggccc tggccacat gatcaagtac 540
 cggggccact tcctgatcga gggcgagttc aacagcaaga acaacgacat ccagaagaac 600
 ttccaggact tcctggacac ctacaacgcc atcttcgaga gcgacctgtc cctggaaaac 660
 agcaagcagc tggaagagat cgtgaaggac aagatcagca agctggaaaa gaaggaccgc 720
 atcctgaagc tgttccccgg cgagaagaac agcggaatct tcagcgagtt tctgaagctg 780
 atcgtgggca accaggccga cttcagaaag tgcttcaacc tggacgagaa agccagcctg 840
 cacttcagca aagagagcta cgacgaggac ctggaaaccc tgctgggata tatcggcgac 900
 gactacagcg acgtgttcct gaaggccaag aagctgtacg acgctatcct gctgagcggc 960
 ttcttgaccg tgaccgacaa cgagacagag gccccactga gcagcgccat gattaagcgg 1020
 tacaacgagc acaaagagga tctggctctg ctgaaagagt acatccggaa catcagcctg 1080
 aaaacctaca atgaggtgtt caaggacgac accaagaacg gctacgccgg ctacatcgac 1140
 ggcaagacca accaggaaga tttctatgtg tacctgaaga agctgctggc cgagttcgag 1200

ggggccgact	acttttctgga	aaaaatcgac	cgcgaggatt	tcctgcggaa	gcagcggacc	1260
ttcgacaacg	gcagcatccc	ctaccagatc	catctgcagg	aaatgcgggc	catcctggac	1320
aagcaggcca	agttctaccc	attcctggcc	aagaacaaag	agcggatcga	gaagatcctg	1380
accttccgca	tcccttacta	cgtgggcccc	ctggccagag	gcaacagcga	ttttgcctgg	1440
tccatccgga	agcgcaatga	gaagatcacc	ccctggaact	tcgaggacgt	gatcgacaaa	1500
gagtccagcg	ccgaggcctt	catcaaccgg	atgaccagct	tcgacctgta	cctgcccag	1560
gaaaagggtgc	tgcccaagca	cagcctgctg	tacgagacat	tcaatgtgta	taacgagctg	1620
accaaagtgc	ggtttatcgc	cgagtctatg	cgggactacc	agttcctgga	ctccaagcag	1680
aaaaaggaca	tcgtgcggt	gtacttcaag	gacaagcggga	aagtgaccga	taaggacatc	1740
atcgagtacc	tgcacgccat	ctacggctac	gatggcatcg	agctgaagg	catcgagaag	1800
cagttcaact	ccagcctgag	cacataccac	gacctgctga	acattatcaa	cgacaaagaa	1860
tttctggacg	actccagcaa	cgaggccatc	atcgaagaga	tcatccacac	cctgaccatc	1920
tttgaggacc	gcgagatgat	caagcagcgg	ctgagcaagt	tcgagaacat	cttcgacaag	1980
agcgtgctga	aaaagctgag	cagacggcac	tacaccggct	ggggcaagct	gagcgccaag	2040
ctgatcaacg	gcatccggga	cgagaagtcc	ggcaacacaa	tcctggacta	cctgatcgac	2100
gacggcatca	gcaaccggaa	cttcatgcag	ctgatccacg	acgacgccct	gagcttcaag	2160
aagaagatcc	agaaggccca	gatcatcggg	gacgaggaca	agggcaacat	caaagaagtc	2220
gtgaagtccc	tgcccggcag	ccccgccatc	aagaagggaa	tcctgcagag	catcaagatc	2280
gtggacgagc	tcgtgaaagt	gatgggcggc	agaaagccc	agagcatcgt	ggtggaaatg	2340
gctagagaga	accagtacac	caatcagggc	aagagcaaca	gccagcagag	actgaagaga	2400
ctggaaaagt	ccctgaaaga	gctgggcagc	aagattctga	aagagaatat	ccctgccaa	2460
ctgtccaaga	tcgacaacaa	cgcctgcag	aacgaccggc	tgtacctgta	ctacctgcag	2520
aatggcaagg	acatgtatac	aggcgacgac	ctggatatcg	accgcctgag	caactacgac	2580
atcgaccata	ttatccccca	ggccttcctg	aaagacaaca	gcattgacaa	caaagtgtg	2640
gtgtcctccg	ccagcaaccg	cggcaagtcc	gatgatgtgc	ccagcctgga	agtcgtgaaa	2700
aagagaaaga	ccttctggta	tcagctgctg	aaaagcaagc	tgattagcca	gaggaaagttc	2760
gacaacctga	ccaaggccga	gagaggcggc	ctgagccctg	aagataaggc	cggcttcac	2820
cagagacagc	tggtggaaac	ccggcagatc	accaagcacg	tgccagact	gctggatgag	2880
aagtttaaca	acaagaagga	cgagaacaac	cgggcctg	ggaccgtgaa	gatcatcacc	2940
ctgaagtcca	ccctggtgtc	ccagttccgg	aaggacttcg	agctgtataa	agtgcgcgag	3000
atcaatgact	ttcaccacgc	ccacgacgcc	tacctgaatg	ccgtggtggc	ttccgccctg	3060
ctgaagaagt	accctaagct	ggaacccgag	ttcgtgtacg	gcgactaccc	caagtacaac	3120

tccttcagag agcggaagtc cgccaccgag aaggtgtact tctactccaa catcatgaat 3180
atctttaaga agtccatctc cctggccgat ggcagagtga tcgagcggcc cctgatcgaa 3240
gtgaacgaag agacaggcga gagcgtgtgg aacaaagaaa gcgacctggc caccgtgagg 3300
cgggtgctga gttatcctca agtgaatgtc gtgaagaagg tggaagaaca gaaccacggc 3360
ctggatcggg gcaagcccaa gggcctgttc aacgccaacc tgtccagcaa gcctaagccc 3420
aactccaacg agaatctcgt gggggccaaa gaggacctgg accctaagaa gtacggcgga 3480
tacgcgggca tctccaatag cttcaccgtg ctctgtgaagg gcacaatcga gaagggcgct 3540
aagaaaaaga tcacaaacgt gctggaattt caggggatct ctatcctgga ccggatcaac 3600
taccggaagg ataagctgaa ctttctgctg gaaaaaggct acaaggacat tgagctgatt 3660
atcgagctgc ctaagtactc cctgttcgaa ctgagcgacg gctccagacg gatgctggcc 3720
tccatcctgt ccaccaacaa caagcggggc gagatccaca agggaaacca gatcttcctg 3780
agccagaaat ttgtgaaact gctgtaccac gccaaagcga tctccaacac catcaatgag 3840
aaccaccgga aatacgtgga aaaccacaag aaagagtttg aggaactgtt ctactacatc 3900
ctggagttca acgagaacta tgtgggagcc aagaagaacg gcaaactgct gaactccgcc 3960
ttccagagct ggcagaacca cagcatcgac gagctgtgca gctccttcac cggccctacc 4020
ggcagcgagc ggaagggact gtttgagctg acctccagag gctctgccgc cgactttgag 4080
ttcctggggag tgaagatccc ccggtacaga gactacaccc cctctagtct gctgaaggac 4140
gccaccctga tccaccagag cgtgaccggc ctgtacgaaa cccggatcga cctggctaag 4200
ctgggcgagg gaaagcgtcc tgctgctact aagaaagctg gtcaagctaa gaaaaagaaa 4260
taa 4263

- 5 <210> 70
<211> 84
<212> ADN
<213> Secuencia Artificial
- 10 <220>
<221> fuente
<223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"
- <400> 70
ggaaccattc ataacagcat agcaagttat aataaggcta gtccgttata aacttgaaaa 60
agtggcaccg agtcgggtgct tttt 84

<210> 71
<211> 36

	< 212> ADN		
	< 213> Secuencia Artificial		
	<220>		
	< 221> fuente		
5	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<400> 71		
	ggtatagagc tatgctgta tgaatggtcc caaaac	36	
	<210> 72		
	< 211> 84		
10	< 212> ADN		
	< 213> Secuencia Artificial		
	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
15	<400> 72		
	ggaaccattc aatacagcat agcaagttaa tataaggcta gtccgttatc aacttgaaaa	60	
	agtggcaccg agtcggtgct tttt	84	
	<210> 73		
	< 211> 36		
	< 212> ADN		
20	< 213> Secuencia Artificial		
	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
25	<400> 73		
	gtattagagc tatgctgtat tgaatggtcc caaaac	36	
	<210> 74		
	< 211> 103		
	< 212> ADN		
	< 213> Secuencia Artificial		
30	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"		
	<220>		
	< 221> base modificada		
35	< 222> (1) .. (20)		
	< 223> a, c, t, g, Desconocido u otro		
	<400> 74		
	nnnnnnnnnnn nnnnnnnnnn gtttttagagc tagaaatagc aagttaaaat aaggctagtc	60	
	cgttatcaac ttgaaaaagt ggcaccgagt cggcgctttt ttt	103	
	<210> 75		
40	< 211> 103		
	< 212> ADN		
	< 213> Secuencia Artificial		

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
 5 < 221> base modificada
 < 222> (1) .. (20)
 < 223> a, c, t, g, Desconocido u otro

<400> 75
 nnnnnnnnnn nnnnnnnnnn gtattagagc tagaaatagc aagttaatat aaggctagtc 60
 cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt 103

10 <210> 76
 < 211> 123
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 15 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
 < 221> base modificada
 < 222> (1) .. (20)
 20 < 223> a, c, t, g, Desconocido u otro

<400> 76
 nnnnnnnnnn nnnnnnnnnn gtttttagagc tatgctgttt ttgaaacaaa acagcatagc 60
 aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 120
 ttt 123

<210> 77
 < 211> 123
 25 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
 30 < 221> base modificada
 < 222> (1) .. (20)
 < 223> a, c, t, g, Desconocido u otro

<400> 77
 35 nnnnnnnnnn nnnnnnnnnn gtattagagc tatgctgtat ttgaaacaat acagcatagc 60
 aagttaatat aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 120
 ttt 123

<210> 78
 40 < 211> 20

< 212> ADN
< 213> Homo sapiens

<400> 78
gtcacctcca atgactaggg 20

5 <210> 79
< 211> 984
< 212> PRT
< 213> Campylobacter jejuni

<400> 79

10

Met Ala Arg Ile Leu Ala Phe Asp Ile Gly Ile Ser Ser Ile Gly Trp
1 5 10 15

Ala Phe Ser Glu Asn Asp Glu Leu Lys Asp Cys Gly Val Arg Ile Phe
20 25 30

Thr Lys Val Glu Asn Pro Lys Thr Gly Glu Ser Leu Ala Leu Pro Arg
35 40 45

Arg Leu Ala Arg Ser Ala Arg Lys Arg Leu Ala Arg Arg Lys Ala Arg
50 55 60

Leu Asn His Leu Lys His Leu Ile Ala Asn Glu Phe Lys Leu Asn Tyr
65 70 75 80

Glu Asp Tyr Gln Ser Phe Asp Glu Ser Leu Ala Lys Ala Tyr Lys Gly
85 90 95

Ser Leu Ile Ser Pro Tyr Glu Leu Arg Phe Arg Ala Leu Asn Glu Leu
100 105 110

Leu Ser Lys Gln Asp Phe Ala Arg Val Ile Leu His Ile Ala Lys Arg
115 120 125

Arg Gly Tyr Asp Asp Ile Lys Asn Ser Asp Asp Lys Glu Lys Gly Ala
130 135 140

Ile Leu Lys Ala Ile Lys Gln Asn Glu Glu Lys Leu Ala Asn Tyr Gln
145 150 155 160

Ser Val Gly Glu Tyr Leu Tyr Lys Glu Tyr Phe Gln Lys Phe Lys Glu
165 170 175

Asn Ser Lys Glu Phe Thr Asn Val Arg Asn Lys Lys Glu Ser Tyr Glu
 180 185 190
 Arg Cys Ile Ala Gln Ser Phe Leu Lys Asp Glu Leu Lys Leu Ile Phe
 195 200 205
 Lys Lys Gln Arg Glu Phe Gly Phe Ser Phe Ser Lys Lys Phe Glu Glu
 210 215 220
 Glu Val Leu Ser Val Ala Phe Tyr Lys Arg Ala Leu Lys Asp Phe Ser
 225 230 235 240
 His Leu Val Gly Asn Cys Ser Phe Phe Thr Asp Glu Lys Arg Ala Pro
 245 250 255
 Lys Asn Ser Pro Leu Ala Phe Met Phe Val Ala Leu Thr Arg Ile Ile
 260 265 270
 Asn Leu Leu Asn Asn Leu Lys Asn Thr Glu Gly Ile Leu Tyr Thr Lys
 275 280 285
 Asp Asp Leu Asn Ala Leu Leu Asn Glu Val Leu Lys Asn Gly Thr Leu
 290 295 300
 Thr Tyr Lys Gln Thr Lys Lys Leu Leu Gly Leu Ser Asp Asp Tyr Glu
 305 310 315 320
 Phe Lys Gly Glu Lys Gly Thr Tyr Phe Ile Glu Phe Lys Lys Tyr Lys
 325 330 335
 Glu Phe Ile Lys Ala Leu Gly Glu His Asn Leu Ser Gln Asp Asp Leu
 340 345 350
 Asn Glu Ile Ala Lys Asp Ile Thr Leu Ile Lys Asp Glu Ile Lys Leu
 355 360 365
 Lys Lys Ala Leu Ala Lys Tyr Asp Leu Asn Gln Asn Gln Ile Asp Ser
 370 375 380
 Leu Ser Lys Leu Glu Phe Lys Asp His Leu Asn Ile Ser Phe Lys Ala
 385 390 395 400
 Leu Lys Leu Val Thr Pro Leu Met Leu Glu Gly Lys Lys Tyr Asp Glu
 405 410 415
 Ala Cys Asn Glu Leu Asn Leu Lys Val Ala Ile Asn Glu Asp Lys Lys
 420 425 430

Asp Phe Leu Pro Ala Phe Asn Glu Thr Tyr Tyr Lys Asp Glu Val Thr
 435 440 445
 Asn Pro Val Val Leu Arg Ala Ile Lys Glu Tyr Arg Lys Val Leu Asn
 450 455 460
 Ala Leu Leu Lys Lys Tyr Gly Lys Val His Lys Ile Asn Ile Glu Leu
 465 470 475 480
 Ala Arg Glu Val Gly Lys Asn His Ser Gln Arg Ala Lys Ile Glu Lys
 485 490 495
 Glu Gln Asn Glu Asn Tyr Lys Ala Lys Lys Asp Ala Glu Leu Glu Cys
 500 505 510
 Glu Lys Leu Gly Leu Lys Ile Asn Ser Lys Asn Ile Leu Lys Leu Arg
 515 520 525
 Leu Phe Lys Glu Gln Lys Glu Phe Cys Ala Tyr Ser Gly Glu Lys Ile
 530 535 540
 Lys Ile Ser Asp Leu Gln Asp Glu Lys Met Leu Glu Ile Asp His Ile
 545 550 555 560
 Tyr Pro Tyr Ser Arg Ser Phe Asp Asp Ser Tyr Met Asn Lys Val Leu
 565 570 575
 Val Phe Thr Lys Gln Asn Gln Glu Lys Leu Asn Gln Thr Pro Phe Glu
 580 585 590
 Ala Phe Gly Asn Asp Ser Ala Lys Trp Gln Lys Ile Glu Val Leu Ala
 595 600 605
 Lys Asn Leu Pro Thr Lys Lys Gln Lys Arg Ile Leu Asp Lys Asn Tyr
 610 615 620
 Lys Asp Lys Glu Gln Lys Asn Phe Lys Asp Arg Asn Leu Asn Asp Thr
 625 630 635 640
 Arg Tyr Ile Ala Arg Leu Val Leu Asn Tyr Thr Lys Asp Tyr Leu Asp
 645 650 655
 Phe Leu Pro Leu Ser Asp Asp Glu Asn Thr Lys Leu Asn Asp Thr Gln
 660 665 670
 Lys Gly Ser Lys Val His Val Glu Ala Lys Ser Gly Met Leu Thr Ser

ES 2 553 782 T3

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

Glu Thr Leu Ser Lys Asn Gln Lys Ile Leu Phe Lys Asn Ala Asn Glu
930 935 940

Lys Glu Val Ile Ala Lys Ser Ile Gly Ile Gln Asn Leu Lys Val Phe
945 950 955 960

Glu Lys Tyr Ile Val Ser Ala Leu Gly Glu Val Thr Lys Ala Glu Phe
965 970 975

Arg Gln Arg Glu Asp Phe Lys Lys
980

5 <210> 80
< 211> 91
< 212> ADN
< 213> Secuencia Artificial

10 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 80
tataatctca taagaaattt aaaaaggac taaaataaag agtttgcggg actctgcggg 60
gttacaatcc cctaaaaccg cttttaaaat t 91

15 <210> 81
< 211> 36
< 212> ADN
< 213> Secuencia Artificial

20 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 81
atttaccat aaagaaattt aaaaaggac taaaac 36

25 <210> 82
< 211> 95
< 212> ARN
< 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

30 <220>
< 221> base modificada
< 222> (1)..(20)
< 223> a, c, u, g, Desconocido u otro

<400> 82		
	nnnnnnnnnn nnnnnnnnnn guuuuagucc cgaaagggac uaaaauaaag aguuugcggg	60
	acucugcggg guuacaaucc ccuaaaaccg cuuuu	95
<210> 83		
< 211> 69		
5	< 212> ARN	
	< 213> Secuencia Artificial	
<220>		
< 221> fuente		
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
<400> 83		
10	gucaccucca augacuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuuuuuu	69
<210> 84		
< 211> 69		
< 212> ARN		
15	< 213> Secuencia Artificial	
<220>		
< 221> fuente		
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
<400> 84		
	ggggccgaga uuggguguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuuuuuu	69
20		
<210> 85		
< 211> 69		
< 212> ARN		
< 213> Secuencia Artificial		
<220>		
25	< 221> fuente	
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
<400> 85		
	guggcgagag gggccgagau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuuuuuu	69
30		
<210> 86		
< 211> 69		
< 212> ARN		
< 213> Secuencia Artificial		
<220>		
35	< 221> fuente	
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		

	<400> 86		
	ggggccgaga uuggguguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60	
	cguuuuuuu	69	
5	<210> 87 < 211> 69 < 212> ARN < 213> Secuencia Artificial		
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
10	<400> 87		
	guggcgagag gggccgagau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60	
	cguuuuuuu	69	
	<210> 88 < 211> 76 < 212> ARN < 213> Secuencia Artificial		
15	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<400> 88		
	gucaccucca augacuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60	
20	cguuaucauu uuuuuu	76	
	<210> 89 < 211> 76 < 212> ARN < 213> Secuencia Artificial		
25	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<400> 89		
	gacaucgaug uccuccccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60	
	cguuaucauu uuuuuu	76	
30	<210> 90 < 211> 76 < 212> ARN < 213> Secuencia Artificial		
35	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		

	<400> 90	
	gaguccgagc agaagaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucauu uuuuuu	76
5	<210> 91 < 211> 76 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
10	<400> 91	
	gaguccgagc agaagaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucauu uuuuuu	76
15	<210> 92 < 211> 76 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
20	<400> 92	
	ggggccgaga uuggguguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucauu uuuuuu	76
25	<210> 93 < 211> 88 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
30	<400> 93	
	gucaccucca augacuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu guuuuuuu	88
35	<210> 94 < 211> 88 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	

	<400> 94	
	gacauccgaug uccuccccau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguaaucaac uugaaaaagu guuuuuuu	88
5	<210> 95 < 211> 88 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
10	<400> 95	
	gaguccgagc agaagaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguaaucaac uugaaaaagu guuuuuuu	88
15	<210> 96 < 211> 88 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
20	<400> 96	
	ggggccgaga uuggguguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguaaucaac uugaaaaagu guuuuuuu	88
25	<210> 97 < 211> 88 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
30	<400> 97	
	guggcgagag gggccgagau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguaaucaac uugaaaaagu guuuuuuu	88
35	<210> 98 < 211> 103 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	

	<400> 98	
	gucaccucca augacuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	103
	<210> 99	
	< 211> 103	
5	< 212> ARN	
	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
10	<400> 99	
	gucaccucca augacuaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	103
	<210> 100	
	< 211> 103	
	< 212> ARN	
15	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 100	
	gaguccgagc agaagaagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	103
20		
	<210> 101	
	< 211> 103	
	< 212> ARN	
	< 213> Secuencia Artificial	
25	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 101	
	ggggccgaga ugggguguuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu	103
30	<210> 102	
	< 211> 103	
	< 212> ARN	
	< 213> Secuencia Artificial	
	<220>	
35	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	

<400> 102
guggcgagag gggccgagau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60
cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu uuu 103

<210> 103
 < 211> 102
 5 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

10 <400> 103
gttttagagc tatgctgttt tgaatggtcc caaaacggaa ggcctgagt cgcagcagaa 60
gaagaagttt tagagctatg ctgttttgaa tggccccaaa ac 102

<210> 104
 < 211> 100
 < 212> ADN
 15 < 213> Homo sapiens

<400> 104
cggaggacaa agtacaaacg gcagaagctg gaggaggaag ggcctgagtc cgagcagaag 60
aagaagggct cccatcacat caaccggtgg cgcattgcca 100

<210> 105
 < 211> 50
 20 < 212> ADN
 < 213> Homo sapiens

<400> 105
 agctggagga ggaagggcct ggtccgagc agaagaagaa gggctccac 50

<210> 106
 < 211> 30
 25 < 212> ARN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 30 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 106
 gaguccgagc agaagaagaa guuuuagagc 30

<210> 107
 < 211> 49
 35 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

	<400> 107 agctggagga ggaagggcct gagtccgagc agaagagaag ggctcccat	49
5	<210> 108 < 211> 53 < 212> ADN < 213> Homo sapiens	
	<400> 108 ctggaggagg aagggcctga gtccgagcag aagaagaagg gctcccatca cat	53
10	<210> 109 < 211> 52 < 212> ADN < 213> Homo sapiens	
	<400> 109 ctggaggagg aagggcctga gtccgagcag aagagaaggg ctcccatcac at	52
15	<210> 110 < 211> 54 < 212> ADN < 213> Homo sapiens	
20	<400> 110 ctggaggagg aagggcctga gtccgagcag aagaagaagg ggctcccatc acat	54
	<210> 111 < 211> 50 < 212> ADN < 213> Homo sapiens	
25	<400> 111 ctggaggagg aagggcctga gtccgagcag aagaagggt cccatcacat	50
	<210> 112 < 211> 47 < 212> ADN < 213> Homo sapiens	
30		
	<400> 112 ctggaggagg aagggcctga gcccgagcag aagggtccc atcacat	47
	<210> 113 < 211> 66 < 212> ADN < 213> Secuencia Artificial	
35		
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
40	<220> < 221> fuente < 223> /nota="Descripción de Molécula de ADN/ARN Combinada: Oligonucleótido sintético"	
	<220> < 221> base modificada < 222> (1) .. (20) < 223> a, c, t, g, Desconocido u otro	
45		

	<400> 113		
	nnnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggctagtc		60
	cguuuu		66
5	<210> 114 < 211> 20 < 212> ARN < 213> Secuencia Artificial		
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
10	<400> 114 gaguccgagc agaagaagaa	20	
15	<210> 115 < 211> 20 < 212> ARN < 213> Secuencia Artificial		
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
20	<400> 115 gacaucgaug uccuccccau	20	
	<210> 116 < 211> 20 < 212> ARN < 213> Secuencia Artificial		
25	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<400> 116 gucaccucca augacuaggg	20	
30	<210> 117 < 211> 20 < 212> ARN < 213> Secuencia Artificial		
35	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
	<400> 117 auuggguguu cagggcagag	20	
40	<210> 118 < 211> 20 < 212> ARN < 213> Secuencia Artificial		

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

5 <400> 118
 guggcgagag gggccgagau 20

<210> 119
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

10 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

 <400> 119
 ggggccgaga uuggguguuc 20

15 <210> 120
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

20 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

 <400> 120
 gugccauuag cuaaugcau 20

25 <210> 121
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

 <400> 121
 guaccacca caggugccag 20

35 <210> 122
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

40 <400> 122
 gaaagccucu gggccaggaa 20

 <210> 123
 < 211> 48
 < 212> ADN

45 < 213> Homo sapiens

<400> 123
ctggaggagg aagggcctga gtccgagcag aagaagaagg gctcccat 48

<210> 124
< 211> 20
5 < 212> ARN
< 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

10 <400> 124
gaguccgagc agaagaagau 20

<210> 125
< 211> 20
< 212> ARN
15 < 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

20 <400> 125
gaguccgagc agaagaagua 20

<210> 126
< 211> 20
< 212> ARN
< 213> Secuencia Artificial

25 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 126
gaguccgagc agaagaaca 20

30 <210> 127
< 211> 20
< 212> ARN
< 213> Secuencia Artificial

<220>
35 < 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 127
gaguccgagc agaagaugaa 20

<210> 128
40 < 211> 20
< 212> ARN
< 213> Secuencia Artificial

<220>
45 < 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 128
 gaguccgagc agaaguagaa 20

5 <210> 129
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

10 <400> 129
 gaguccgagc agaugaagaa 20

15 <210> 130
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

20 <400> 130
 gaguccgagc acaagaagaa 20

<210> 131
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

25 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 131
 gaguccgagg agaagaagaa 20

30 <210> 132
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

35 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 132
 gaguccgugc agaagaagaa 20

40 <210> 133
 < 211> 20
 < 212> ARN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 45 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

<400> 133
 gagucggagc agaagaagaa 20

 <210> 134
 < 211> 20
 5 < 212> ARN
 < 213> Secuencia Artificial

 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

 10 <400> 134
 gagaccgagc agaagaagaa 20

 <210> 135
 < 211> 24
 < 212> ADN
 15 < 213> Secuencia Artificial

 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

 20 <400> 135
 aatgacaagc ttgctagcgg tggg 24

 <210> 136
 < 211> 39
 < 212> ADN
 < 213> Secuencia Artificial

 25 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

 <400> 136
 aaaacggaag ggcctgagtc cgagcagaag aagaagttt 39

 30 <210> 137
 < 211> 39
 < 212> ADN
 < 213> Secuencia Artificial

 <220>
 35 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

 <400> 137
 aaacaggggc cgagattggg tggtcagggc agaggtttt 39

 <210> 138
 40 < 211> 38
 < 212> ADN
 < 213> Secuencia Artificial

 <220>
 < 221> fuente
 45 < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

	<400> 138 aaaacggaag ggctgagtc cgagcagaag aagaagtt	38
5	<210> 139 < 211> 40 < 212> ADN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
10	<400> 139 aacggaggga ggggcacaga tgagaaactc agggtttttag	40
15	<210> 140 < 211> 38 < 212> ADN < 213> Homo sapiens	
	<400> 140 agcccttctt cttctgctcg gactcaggcc cttcctcc	38
20	<210> 141 < 211> 40 < 212> ADN < 213> Homo sapiens	
	<400> 141 cagggaggga ggggcacaga tgagaaactc aggaggcccc	40
25	<210> 142 < 211> 80 < 212> ADN < 213> Secuencia Artificial	
30	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 142 ggcaatgcg caccggttga tgtgatggga gcccttctag gaggccccca gagcagccac	60
	tggggcctca acactcaggc	80
35	<210> 143 < 211> 98 < 212> ADN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
40	<400> 143 ggacgaaaca ccggaaccat tcaaaacagc atagcaagtt aaaataaggc tagtccgtta	60
	tcaacttgaa aaagtggcac cgagtcggtg cttttttt	98

	<210> 144	
	< 211> 186	
	< 212> ADN	
	< 213> Secuencia Artificial	
5	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<400> 144	
	ggacgaaaca ccggtagtat taagtattgt tttatggctg ataaatttct ttgaatttct	60
	ccttgattat ttgttataaa agttataaaa taatcttggt ggaaccattc aaaacagcat	120
	agcaagttaa aataaggcta gtccgttatc aacttgaaaa agtggcaccc agtcggtgct	180
	tttttt	186
10	<210> 145	
	< 211> 46	
	< 212> ARN	
	< 213> Secuencia Artificial	
	<220>	
15	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<220>	
	< 221> base modificada	
	< 222> (1) .. (19)	
20	< 223> a, c, u, g, Desconocido u otro	
	<400> 145	
	nnnnnnnnnn nnnnnnnnng uuauuguacu cucaagauuu auuuuu	46
	<210> 146	
	< 211> 91	
25	< 212> ARN	
	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
30	<400> 146	
	guuacuuaaa ucuugcagaa gcuacaaaga uaaggcuuca ugccgaaauc aacacccugu	60
	cauuuuauagg caggguguuu ucguuuuuu a	91
	<210> 147	
	< 211> 70	
	< 212> ADN	
35	< 213> Homo sapiens	
	<400> 147	
	ttttctagtg ctgagtttct gtgactcctc tacattctac ttctctgtgt ttctgtatac	60
	tacctcctcc	70

	<210> 148		
	< 211> 122		
	< 212> ADN		
	< 213> Homo sapiens		
5	<400> 148		
	ggaggaaggg cctgagtccg agcagaagaa gaagggctcc catcacatca accggtggcg	60	
	cattgccacg aagcaggcca atggggagga catcgatgtc acctccaatg actaggggtgg	120	
	gc	122	
10	<210> 149		
	< 211> 48		
	< 212> ARN		
	< 213> Secuencia Artificial		
15	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
20	<220>		
	< 221> base modificada		
	< 222> (3) .. (32)		
	< 223> a, c, u, g, Desconocido u otro		
	<400> 149		
	acnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnguuuuaga gcuaugcu	48	
25	<210> 150		
	< 211> 67		
	< 212> ADN		
	< 213> Secuencia Artificial		
	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		
30	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Molécula de ADN/ARN Combinada: Oligonucleótido sintético"		
	<400> 150		
	agcauagcaa guuaaaaauaa ggctaguccg uuaucaacuu gaaaaagugg caccgagucg	60	
	gugcuuu	67	
35	<210> 151		
	< 211> 62		
	< 212> ARN		
	< 213> Secuencia Artificial		
40	<220>		
	< 221> fuente		
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"		

	<220> < 221> base modificada < 222> (1) .. (20) < 223> a, c, u, g, Desconocido u otro	
5	<400> 151 nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cg	60 62
10	<210> 152 < 211> 73 < 212> ADN < 213> Secuencia Artificial <220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
15	<400> 152 tgaatggtcc caaaacggaa gggcctgagt ccgagcagaa gaagaagttt tagagctatg ctgttttgaa tgg	60 73
20	<210> 153 < 211> 99 < 212> ARN < 213> Secuencia Artificial <220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
25	<220> < 221> base modificada < 222> (1) .. (20) < 223> a, c, u, g, Desconocido u otro	
30	<400> 153 nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugcuuu	60 99
35	<210> 154 < 211> 127 < 212> ARN < 213> Secuencia Artificial <220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	

	<400> 154	
	guuuuuuguac ucucaagauu uaaguaacug uacaacguua cuuaaaucuu gcagaagcua	60
	caaagauaag gcuucaugcc gaaaucaaca cccugucuu uuauggcagg guguuuucgu	120
	uauuuuaa	127
	<210> 155	
	< 211> 56	
5	< 212> ADN	
	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
10	<220>	
	< 221> base modificada	
	< 222> (1) .. (20)	
	< 223> a, c, t, g, Desconocido u otro	
	<400> 155	
15	nnnnnnnnnn nnnnnnnnnn gttttgtac tctcaagatt taagtaactg tacaac	56
	<210> 156	
	< 211> 91	
	< 212> ADN	
	< 213> Secuencia Artificial	
20	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 156	
	gttacttaaa tcttgcagaa gctacaaaga taaggcttca tgccgaaatc aacaccotgt	60
	cattttatgg caggggtgtt tcgttattta a	91
25	<210> 157	
	< 211> 134	
	< 212> ADN	
	< 213> Secuencia Artificial	
	<220>	
30	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<220>	
	< 221> base modificada	
	< 222> (1) .. (20)	
35	< 223> a, c, t, g, Desconocido u otro	
	<400> 157	
	nnnnnnnnnn nnnnnnnnnn gttttgtac tctcaagatt taaggaaact aaatcttgca	60
	gaagctacaa agataaggct tcatgccgaa atcaacaccc tgtcatttta tggcaggggtg	120
	ttttcgttat ttaa	134

<210> 158
 < 211> 131
 < 212> ADN
 < 213> Secuencia Artificial

5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
 < 221> base modificada
 10 < 222> (1) .. (20)
 < 223> a, c, t, g, Desconocido u otro

<400> 158
 nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa 60
 gctacaaaga taaggcttca tgccgaaatc aacaccctgt cattttatgg cagggtgttt 120
 tcgttattta a 131

15 <210> 159
 < 211> 125
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 20 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
 < 221> base modificada
 < 222> (1) .. (20)
 < 223> a, c, t, g, Desconocido u otro

25 <400> 159
 nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatg aaaatcttgc agaagctaca 60
 aagataaggc ttcattgccga aatcaacacc ctgtcatttt atggcagggt gttttcgtta 120
 tttaa 125

30 <210> 160
 < 211> 112
 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

35 <220>
 < 221> base modificada
 < 222> (1) .. (20)
 < 223> a, c, t, g, Desconocido u otro

<400> 160
 nnnnnnnnnn nnnnnnnnnn gtttttgtac tctgaaaaga agctacaaag ataaggcttc 60

atgccgaaat caacaccctg tcattttatg gcagggtgtt ttcgttattt aa

112

5 <210> 161
< 211> 107
< 212> ADN
< 213> Secuencia Artificial

<220>
< 221> fuente
10 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
< 221> base modificada
< 222> (1)..(20)
< 223> a, c, t, g, Desconocido u otro

15 <400> 161
nnnnnnnnnn nnnnnnnnnn gtttttgtac tgaaaagcta caaagataag gcttcatgcc 60
gaaatcaaca ccctgtcatt ttatggcagg gtgttttcgt tatttaa 107

20 <210> 162
< 211> 108
< 212> ADN
< 213> Secuencia Artificial

<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

25 <220>
< 221> base modificada
< 222> (1) .. (20)
< 223> a, c, t, g, Desconocido u otro

<400> 162
nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa 60
gctacaaaga taaggcttca tgccgaaatc aacaccctgt cattttat 108

30 <210> 163
< 211> 86
< 212> ADN
< 213> Secuencia Artificial

35 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"

40 <220>
< 221> base modificada
< 222> (1) .. (20)
< 223> a, c, t, g, Desconocido u otro

<400> 163
nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa 60
gctacaaaga taaggcttca tgccga 86

	<210> 164	
	< 211> 79	
	< 212> ADN	
	< 213> Secuencia Artificial	
5	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<220>	
	< 221> base modificada	
10	< 222> (1) .. (20)	
	< 223> a, c, t, g, Desconocido u otro	
	<400> 164	
	nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa	60
	gctacaaaga taaggcttc	79
	<210> 165	
15	< 211> 73	
	< 212> ADN	
	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
20	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<220>	
	< 221> base modificada	
	< 222> (1)..(20)	
	< 223> a, c, t, g, Desconocido u otro	
25	<400> 165	
	nnnnnnnnnn nnnnnnnnnn gtttttgtac tctcaagatt tagaaataaa tcttgcagaa	60
	gctacaaaga taa	73
	<210> 166	
	< 211> 125	
	< 212> ARN	
30	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<400> 166	
35	guuuuagucc cuuuuuuuuu uucuuuauugg uaaaauuuaa aucucauaag aaauuuuuuu	60
	agggacuuaa auuuuaguu ugcgggacuc ugcggggguu caaucccuua aaaccgcuuu	120
	uaaaa	125
	<210> 167	
	< 211> 91	
40	< 212> ARN	
	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	

	<400> 167	
	guuacuuaaa ucuugcagaa gcuacaaaga uaaggcuuca ugccgaauc aacaccugu	60
	cauuuuauagg caggguguuu ucguuuuuu a	91
5	<210> 168 < 211> 56 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
10	<400> 168 gggacucaac caagucuuu guuuuuuguac ucucaagauu uaaguaacug uacaac 56	
	<210> 169 < 211> 147 < 212> ARN	
15	< 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<400> 169	
	gggacucaac caagucuuu guuuuuuguac ucucaagauu uaaguaacug uacaacguua	60
	cuuaaaucuu gcagaagcua caaagauaag gcuucaugcc gaaaucaaca cccugucuuu	120
20	uuauaggcagg guguuuucgu uauuuua	147
	<210> 170 < 211> 70 < 212> ARN < 213> Secuencia Artificial	
25	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 170	
	cuugcagaag cuacaaagau aaggcuucau gccgaaauca acaccuguc auuuuauggc	60
	aggguguuuu	70
30	<210> 171 < 211> 42 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
35	<400> 171 gggacucaac caagucuuu guuuuuuguac ucucaagauu ua	42
40	<210> 172 < 211> 112	

	< 212> ARN	
	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
5	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<400> 172	
	gggacucaac caagucauuc guuuuuguac ucucaagauu uacuugcaga agcuacaaag	60
	auaaggcuuc augccgaaau caacacccug ucauuuuuug gcaggguugu uu	112
	<210> 173	
	< 211> 116	
10	< 212> ARN	
	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
15	<400> 173	
	gggacucaac caagucauuc guuuuuguac ucucaagauu uagaaacuug cagaagcuac	60
	aaagauaagg cuucaugccg aaaucaacac ccugucauuu uauggcaggg uguuuu	116
	<210> 174	
	< 211> 116	
	< 212> ARN	
20	< 213> Secuencia Artificial	
	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<400> 174	
	gggacucaac caagucauuc guuuuuguac ucucaagauu uagaaacuug cagaagcuac	60
25	aaagauaagg cuucaugccg aaaucaacac ccugucauuu uauggcaggg uguuuu	116
	<210> 175	
	< 211> 102	
	< 212> ARN	
	< 213> Secuencia Artificial	
30	<220>	
	< 221> fuente	
	< 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<400> 175	
	gggacucaac caagucauuc guuuuuguag aaauacaaag auaaggcuuc augccgaaau	60
	caacacccug ucauuuuuug gcaggguuguu uucguuuuu aa	102
35	<210> 176	
	< 211> 102	
	< 212> ARN	
	< 213> Secuencia Artificial	

	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<400> 176 gggacucaac caaguc <u>auuc guuuuuguag aaauacaaag auaaggcuuc augccgaaau</u>	60
5	caacacccug ucauuuu <u>aug gcaggguguu uucguuauuu aa</u>	102
	<210> 177 < 211> 57 < 212> ARN < 213> Secuencia Artificial	
10	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 177 gggacucaac caaguc <u>auuc guuuuuguag aaauacaaag auaaggcuuc augccga</u>	57
15	<210> 178 < 211> 57 < 212> ARN < 213> Secuencia Artificial	
20	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 178 gggacucaac caaguc <u>auuc guuuuuguag aaauacaaag auaaggcuuc augccga</u>	57
25	<210> 179 < 211> 23 < 212> ADN < 213> Secuencia Artificial	
30	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 179 gtggtgtcac gctcgtcgtt tgg	23
35	<210> 180 < 211> 23 < 212> ADN < 213> Secuencia Artificial	
40	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
	<400> 180 tccagtctat taattgtgc cgg	23
45	<210> 181 < 211> 64 < 212> ADN < 213> Homo sapiens	

	<400> 181	
	caagaggctt gagtaggaga ggagtgccgc cgagggcgggg cggggcgggg cgtggagctg	60
	ggct	64
5	<210> 182 < 211> 99 < 212> ARN < 213> Secuencia Artificial	
	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Oligonucleótido sintético"	
10	<220> < 221> base modificada < 222> (1)..(20) < 223> a, c, u, g, Desconocido u otro	
	<400> 182	
	nnnnnnnnnnn nnnnnnnnnnn guauuagagc uagaaauagc aaguuaauau aaggcuaguc	60
15	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuu	99
	<210> 183 < 211> 119 < 212> ARN < 213> Secuencia Artificial	
20	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
25	<220> < 221> base modificada < 222> (1)..(20) < 223> a, c, u, g, Desconocido u otro	
	<400> 183	
	nnnnnnnnnnn nnnnnnnnnnn guuuuagagc uaugcuguuu uggaaacaaa acagcauagc	60
	aaguuaaaau aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugcuuu	119
30	<210> 184 < 211> 119 < 212> ARN < 213> Secuencia Artificial	
35	<220> < 221> fuente < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"	
	<220> < 221> base modificada < 222> (1) .. (20) < 223> a, c, u, g, Desconocido u otro	
40	<400> 184	
	nnnnnnnnnnn nnnnnnnnnnn guauuagagc uaugcuguuu uggaaacaa uacagcauagc	60
	aaguuaauau aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugcuuu	119

<210> 185
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

5 <400> 185
 tagcgggtaa gc 12

<210> 186
 < 211> 12
 < 212> ADN
 10 < 213> Homo sapiens

<400> 186
 tcggtgacat gt 12

<210> 187
 < 211> 12
 15 < 212> ADN
 < 213> Homo sapiens

<400> 187
 actccccgta gg 12

<210> 188
 < 211> 12
 20 < 212> ADN
 < 213> Homo sapiens

<400> 188
 actgcgtgtt aa 12

25 <210> 189
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

30 <400> 189
 acgtcgctg at 12

<210> 190
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

35 <400> 190
 taggtcgacc ag 12

<210> 191
 < 211> 12
 < 212> ADN
 40 < 213> Homo sapiens

<400> 191
 ggcgttaatg at 12

<210> 192
 < 211> 12
 45 < 212> ADN
 < 213> Homo sapiens

	<400> 192	
	tgtcgcattg ta	12
5	<210> 193	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 193	
	atggaaacgc at	12
10	<210> 194	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 194	
	gccgaattcc tc	12
15	<210> 195	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
20	<400> 195	
	gcatgttacg ga	12
	<210> 196	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
25	<400> 196	
	cgttactctt ac	12
	<210> 197	
	< 211> 12	
	< 212> ADN	
30	< 213> Homo sapiens	
	<400> 197	
	gcctgtgccg ta	12
	<210> 198	
	< 211> 12	
35	< 212> ADN	
	< 213> Homo sapiens	
	<400> 198	
	tacggttaagt cg	12
	<210> 199	
40	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 199	
	cacgaaatta cc	12

	<210> 200	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
5	<400> 200	
	aaccaagata cg	12
	<210> 201	
	< 211> 12	
	< 212> ADN	
10	< 213> Homo sapiens	
	<400> 201	
	gagtcgatac gc	12
	<210> 202	
	< 211> 12	
15	< 212> ADN	
	< 213> Homo sapiens	
	<400> 202	
	gtctcacgat cg	12
	<210> 203	
20	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 203	
	tcgtcgggtg ca	12
25	<210> 204	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 204	
30	actccgtagt ga	12
	<210> 205	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
35	<400> 205	
	caggacgtcc gt	12
	<210> 206	
	< 211> 12	
	< 212> ADN	
40	< 213> Homo sapiens	
	<400> 206	
	tcgtatccct ac	12
	<210> 207	
	< 211> 12	
45	< 212> ADN	
	< 213> Homo sapiens	

	<400> 207 tttcaaggcc gg	12
5	<210> 208 < 211> 12 < 212> ADN < 213> Homo sapiens	
	<400> 208 cgccggtgga at	12
10	<210> 209 < 211> 12 < 212> ADN < 213> Homo sapiens	
	<400> 209 gaaccgtcc ta	12
15	<210> 210 < 211> 12 < 212> ADN < 213> Homo sapiens	
20	<400> 210 gattcatcag cg	12
	<210> 211 < 211> 12 < 212> ADN < 213> Homo sapiens	
25	<400> 211 acaccgtct tc	12
	<210> 212 < 211> 12 < 212> ADN < 213> Homo sapiens	
30		
	<400> 212 atcgtgccct aa	12
	<210> 213 < 211> 12 < 212> ADN < 213> Homo sapiens	
35		
	<400> 213 gcgtaatgt tc	12
	<210> 214 < 211> 12 < 212> ADN < 213> Homo sapiens	
40		
	<400> 214 ctccgtatct cg	12

<210> 215
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

5 <400> 215
 ccgattcctt cg 12

<210> 216
 < 211> 12
 < 212> ADN
 10 < 213> Homo sapiens

<400> 216
 tgcgcctcca gt 12

<210> 217
 < 211> 12
 15 < 212> ADN
 < 213> Homo sapiens

<400> 217
 taacgtcgga gc 12

<210> 218
 < 211> 12
 20 < 212> ADN
 < 213> Homo sapiens

<400> 218
 aaggtcgccc at 12

25 <210> 219
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

30 <400> 219
 gtcggggact at 12

<210> 220
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

35 <400> 220
 ttcgagcgat tt 12

<210> 221
 < 211> 12
 < 212> ADN
 40 < 213> Homo sapiens

<400> 221
 tgagtcgtcg ag 12

<210> 222
 < 211> 12
 45 < 212> ADN
 < 213> Homo sapiens

	<400> 222	
	tttacgcaga gg	12
	<210> 223	
	< 211> 12	
5	< 212> ADN	
	< 213> Homo sapiens	
	<400> 223	
	aggaagtatc gc	12
	<210> 224	
10	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 224	
	actcgatacc at	12
15	<210> 225	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 225	
20	cgctacatag ca	12
	<210> 226	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
25	<400> 226	
	ttcataaccg gc	12
	<210> 227	
	< 211> 12	
	< 212> ADN	
30	< 213> Homo sapiens	
	<400> 227	
	ccaaacggtt aa	12
	<210> 228	
	< 211> 12	
35	< 212> ADN	
	< 213> Homo sapiens	
	<400> 228	
	cgattccttc gt	12
	<210> 229	
40	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 229	
	cgtcatgaat aa	12

<210> 230
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

5 <400> 230
 agtggcgatg ac 12

<210> 231
 < 211> 12
 < 212> ADN
 10 < 213> Homo sapiens

<400> 231
 cccctacggc ac 12

<210> 232
 < 211> 12
 15 < 212> ADN
 < 213> Homo sapiens

<400> 232
 gccaacccgc ac 12

<210> 233
 < 211> 12
 20 < 212> ADN
 < 213> Homo sapiens

<400> 233
 tgggacaccg gt 12

25 <210> 234
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

30 <400> 234
 ttgactgcgg cg 12

<210> 235
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

35 <400> 235
 actatgcgta gg 12

<210> 236
 < 211> 12
 < 212> ADN
 40 < 213> Homo sapiens

<400> 236
 tcacccaaag cg 12

<210> 237
 < 211> 12
 45 < 212> ADN
 < 213> Homo sapiens

	<400> 237 gcaggacgtc cg	12
5	<210> 238 < 211> 12 < 212> ADN < 213> Homo sapiens	
	<400> 238 acaccgaaaa cg	12
10	<210> 239 < 211> 12 < 212> ADN < 213> Homo sapiens	
	<400> 239 cgggtgtattg ag	12
15	<210> 240 < 211> 12 < 212> ADN < 213> Homo sapiens	
20	<400> 240 cacgaggtat gc	12
	<210> 241 < 211> 12 < 212> ADN < 213> Homo sapiens	
25	<400> 241 taaagcgacc cg	12
	<210> 242 < 211> 12 < 212> ADN < 213> Homo sapiens	
30		
	<400> 242 cttagtcggc ca	12
	<210> 243 < 211> 12 < 212> ADN < 213> Homo sapiens	
35		
	<400> 243 cgaaaacgtg gc	12
	<210> 244 < 211> 12 < 212> ADN < 213> Homo sapiens	
40		
	<400> 244 cgtgccctga ac	12

<210> 245
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

5 <400> 245
 ttaccatcg aa 12

<210> 246
 < 211> 12
 < 212> ADN
 10 < 213> Homo sapiens

<400> 246
 cgtagccatg tt 12

<210> 247
 < 211> 12
 15 < 212> ADN
 < 213> Homo sapiens

<400> 247
 cccaaacggt ta 12

<210> 248
 < 211> 12
 20 < 212> ADN
 < 213> Homo sapiens

<400> 248
 gcgttatcag aa 12

25 <210> 249
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

30 <400> 249
 tcgatggtaa ac 12

<210> 250
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

35 <400> 250
 cgactttttg ca 12

<210> 251
 < 211> 12
 < 212> ADN
 40 < 213> Homo sapiens

<400> 251
 tcgacgactc ac 12

<210> 252
 < 211> 12
 45 < 212> ADN
 < 213> Homo sapiens

	<400> 252	
	acgcgtcaga ta	12
	<210> 253	
	< 211> 12	
5	< 212> ADN	
	< 213> Homo sapiens	
	<400> 253	
	cgtacggcac ag	12
	<210> 254	
10	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 254	
	ctatgccgtg ca	12
	<210> 255	
15	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 255	
20	cgcgtcagat at	12
	<210> 256	
	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 256	12
25	aagatcggtg gc	12
	<210> 257	
	< 211> 12	
	< 212> ADN	
30	< 213> Homo sapiens	
	<400> 257	
	cttcgcaagg ag	12
	<210> 258	
	< 211> 12	
35	< 212> ADN	
	< 213> Homo sapiens	
	<400> 258	
	gtcgtggact ac	12
	<210> 259	
40	< 211> 12	
	< 212> ADN	
	< 213> Homo sapiens	
	<400> 259	
	ggcgtcatc aa	12

<210> 260
 < 211> 12
 < 212> ADN
 < 213> Homo sapiens

5 <400> 260
 gttacacg cg 12

<210> 261
 < 211> 12
 < 212> ADN
 10 < 213> Homo sapiens

<400> 261
 tagctaaccg tt 12

<210> 262
 < 211> 12
 15 < 212> ADN
 < 213> Homo sapiens

<400> 262
 agtaaaggcg ct 12

<210> 263
 < 211> 12
 20 < 212> ADN
 < 213> Homo sapiens

<400> 263
 ggtaatttcg tg 12

25 <210> 264
 < 211> 147
 < 212> ARN
 < 213> Secuencia Artificial

30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Polinucleótido sintético"

<220>
 < 221> base modificada
 < 222> (1)..(20)
 35 < 223> a, c, u, g, Desconocido u otro

<400> 264
 nnnnnnnnnnnn nnnnnnnnnnnn guuuuaguac ucuguaauuu uagguaugag guagacgaaa 60
 auuguacuua uaccuaaaaau uacagaaucu acuaaaacaa ggcaaaaugc cguguuuauuc 120
 ucgucaacuu guuggcgaga uuuuuuu 147

40

REIVINDICACIONES

1. Un sistema vector relacionado con Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR)-CRISPR asociado a (Cas) (CRISPR-Cas) de ingeniería, que no está presente en la naturaleza, que comprende uno o más vectores que comprenden:

- 5 a) un primer elemento regulador unido de manera operable a una secuencia de nucleótidos que codifica una secuencia de polinucleótidos del sistema CRISPR-Cas que comprende una secuencia guía, un ARNtracr y una secuencia de emparejamiento tracr, en la que la secuencia guía se hibrida con una o más secuencias objetivo en los sitios de los polinucleótidos en una célula eucariota,
- 10 b) un segundo elemento regulador unido de manera operable a la secuencia de nucleótidos que codifica una proteína Cas9 de Tipo II,
- c) una plantilla de recombinación

en el que los componentes (a), (b) y (c) están situados en los mismos vectores o diferentes vectores del sistema,

en el que el sistema comprende además una o más señales de localización nuclear (las NLS) expresadas con la secuencia de nucleótidos que codifica la proteína Cas9,

- 15 según lo cual la secuencia guía fija como objetivo uno o más sitios de polinucleótidos en una célula eucariota y la proteína Cas9 escinde uno o más sitios de polinucleótidos, según lo cual se modifica la secuencia de uno o más sitios de polinucleótidos.

2. Un sistema vector CRISPR-Cas de tipo II de ingeniería, que no está presente en la naturaleza de acuerdo con la reivindicación 1,

- 20 en el que en el que la proteína Cas9 está mutada con respecto a la proteína Cas9 de tipo natural correspondiente de manera que la proteína mutada es una nickasa que carece de capacidad para escindir una hebra de un polinucleótido objetivo,

según lo cual la secuencia guía fija como objetivo uno o más sitios de polinucleótidos en una célula eucariota y la proteína Cas9 solo escinde una hebra de los sitios de polinucleótidos, según lo cual se modifica la secuencia de uno o más sitios de polinucleótidos.

- 25

3. El sistema de la reivindicación 2, en el que la proteína Cas9 comprende una o más mutaciones en los dominios catalíticos RuvC I, RuvC II o RuvC III.

4. El sistema de la reivindicación 2, en el que la proteína Cas9 comprende una mutación seleccionada del grupo que consiste en D10A, H840A, N854A y N863A con referencia a la numeración de la posición de una proteína Cas9 de *Streptococcus pyogenes* (SpCas9).

- 30

5. El sistema de cualquiera de las reivindicaciones 1-4, en el que la plantilla de recombinación (c) está contenida en un vector separado de los componentes (a) y (b).

6. El sistema de cualquiera de las reivindicaciones 1-4, en el que la plantilla de recombinación (c) es un componente de un vector que comprende el componente (a) y/o (b).

- 35 7. El sistema de cualquiera de las reivindicaciones 1-6, en el que los componentes (a) y (b) son componentes del mismo vector.

8. El sistema de cualquiera de las reivindicaciones precedentes, en el que la plantilla de recombinación es complementaria a una porción de un polinucleótido que comprende la secuencia objetivo.

9. El sistema de cualquiera de las reivindicaciones precedentes, en el que el sistema CRISPR-Cas comprende dos o más señales de localización nuclear expresadas con la secuencia de nucleótidos que codifica la proteína Cas9.

- 40

10. El sistema de la reivindicación 9, en el que al menos una NLS está en o cerca del término amino de la proteína Cas9 y/o al menos una NLS está en o cerca del término carboxi de la proteína Cas9.

11. El sistema de la reivindicación 10, en el que al menos una NLS está en o cerca del término amino de Cas9 y al menos una NLS está en o cerca del término carboxi de la proteína Cas9.
12. El sistema de cualquiera de las reivindicaciones precedentes, en el que los vectores son vectores víricos.
13. El sistema de la reivindicación 12, en el que los vectores víricos son vectores víricos retrovíricos, lentivíricos, adenovíricos, adenoasociados o de herpes simple.
14. El sistema de cualquiera de las reivindicaciones precedentes, en el que la secuencia de polinucleótidos del sistema CRISPR-Cas comprende una secuencia guía condensada con una secuencia cr de activación trans (tracr).
15. El sistema de cualquiera de las reivindicaciones precedentes, en el que la secuencia de polinucleótidos del sistema CRISPR-Cas es un ARN quimérico que comprende la secuencia guía, la secuencia tracr y una secuencia de emparejamiento tracr.
16. El sistema de cualquiera de las reivindicaciones precedentes, en el que la célula eucariota es una célula de mamífero o una célula humana.
17. El sistema de cualquiera de las reivindicaciones precedentes, en el que la proteína Cas9 es codón-optimizada para la expresión en una célula eucariota.
18. El sistema de cualquiera de las reivindicaciones precedentes, en el que la guía incluye una secuencia tracr con una longitud de 50 o más nucleótidos.
19. El uso del sistema de cualquiera de las reivindicaciones 1-18 para la ingeniería de genomas, siempre que el uso no comprenda un procedimiento para modificar la identidad genética de la línea germinal de seres humanos y siempre que dicho uso no sea un método para tratamiento del cuerpo humano o de un animal por cirugía o terapia.
20. El uso del sistema de cualquiera de las reivindicaciones 1-18 en la producción de un animal transgénico no humano o una planta transgénica.
21. El sistema de cualquiera de las reivindicaciones 1-18 para su uso en la terapia génica o edición genómica.

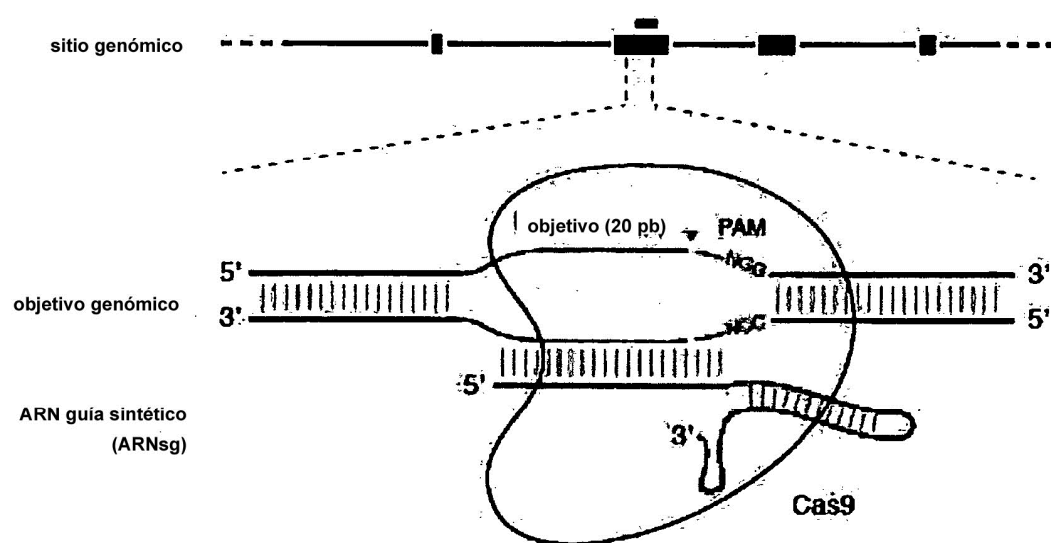


FIG. 1

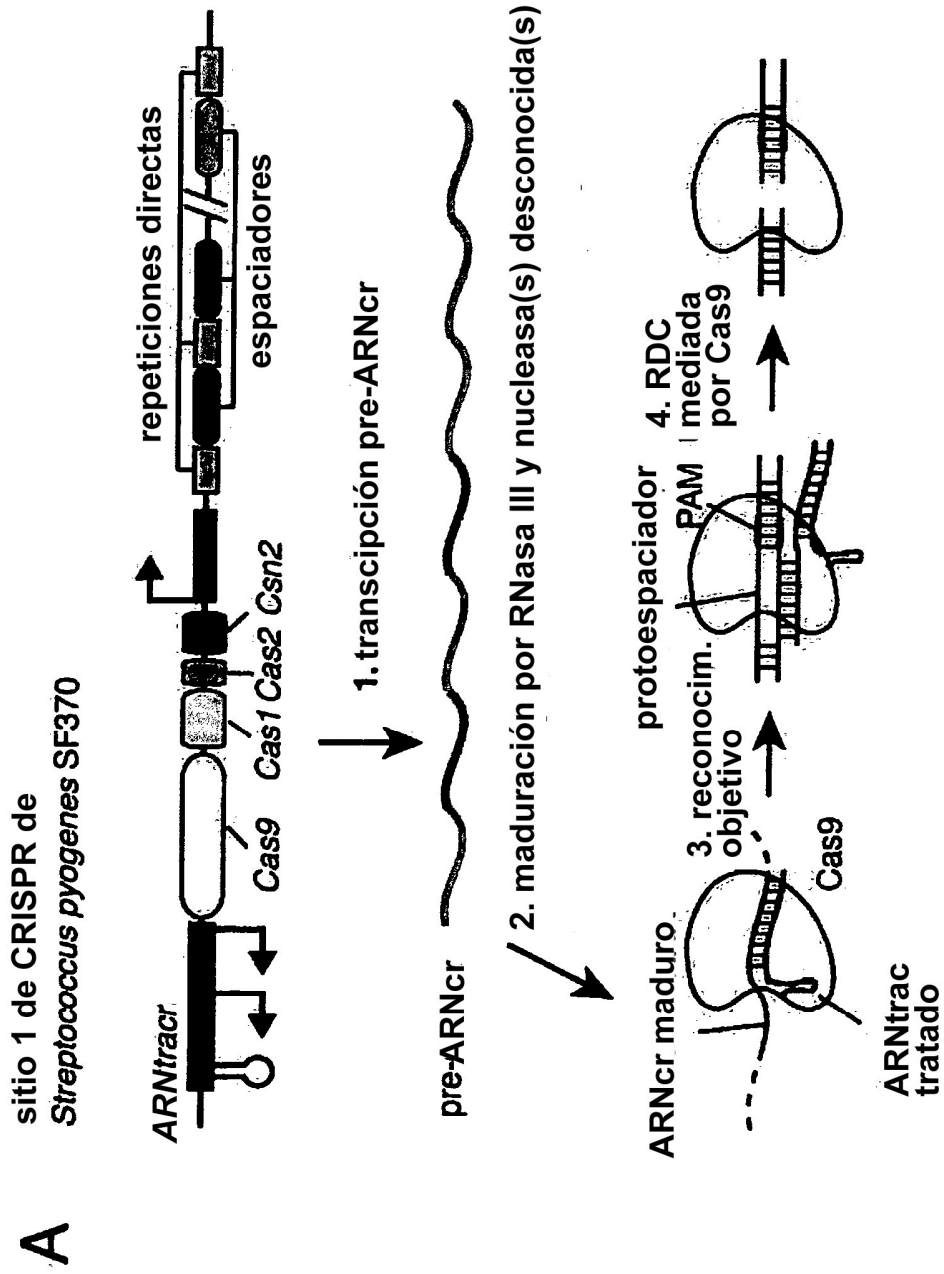


FIG. 2A

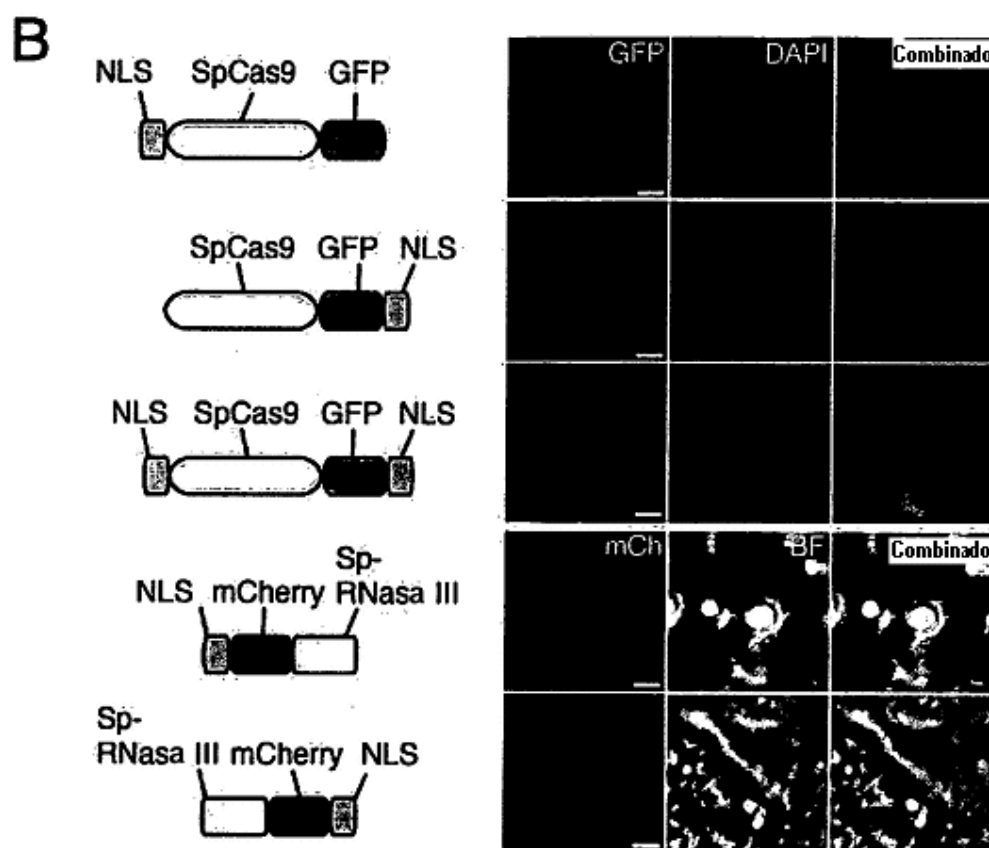


FIG. 2B

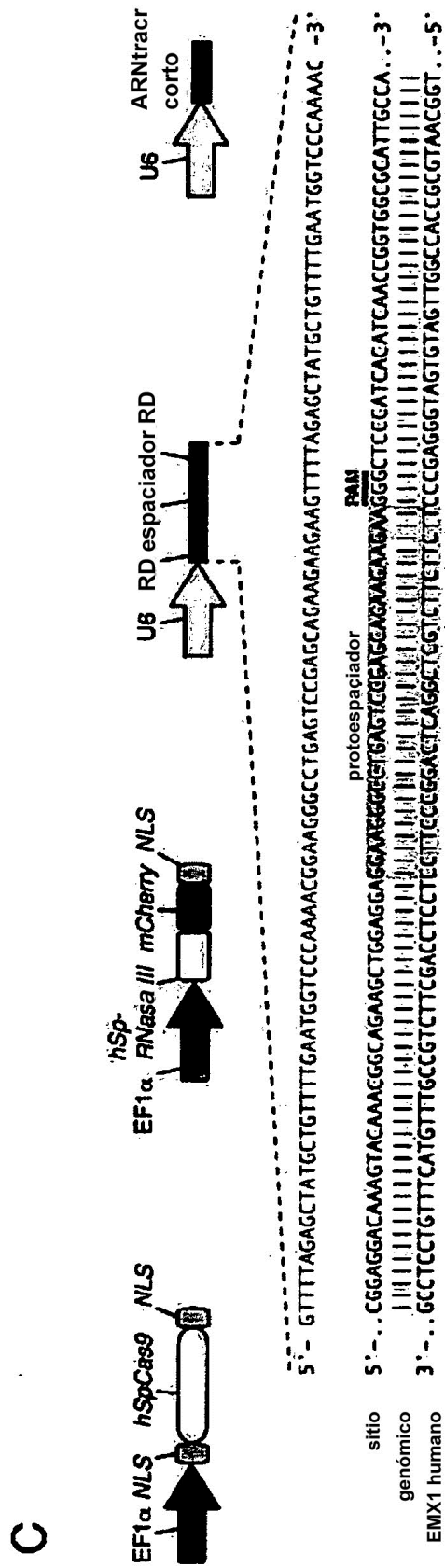
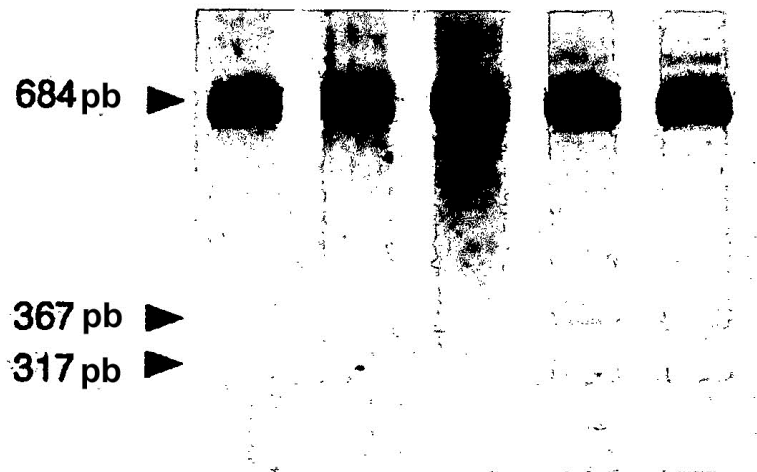


FIG. 2C

D

2xNLS-SpCas9	+	+	+	+	+
SpRNAsa III	-	+	+	-	+
ARNtracr corto	-	+	-	+	+
RD-EMX1-RD	+	-	+	+	+



inserción o supresión (%):

4,7 5,0

FIG. 2D

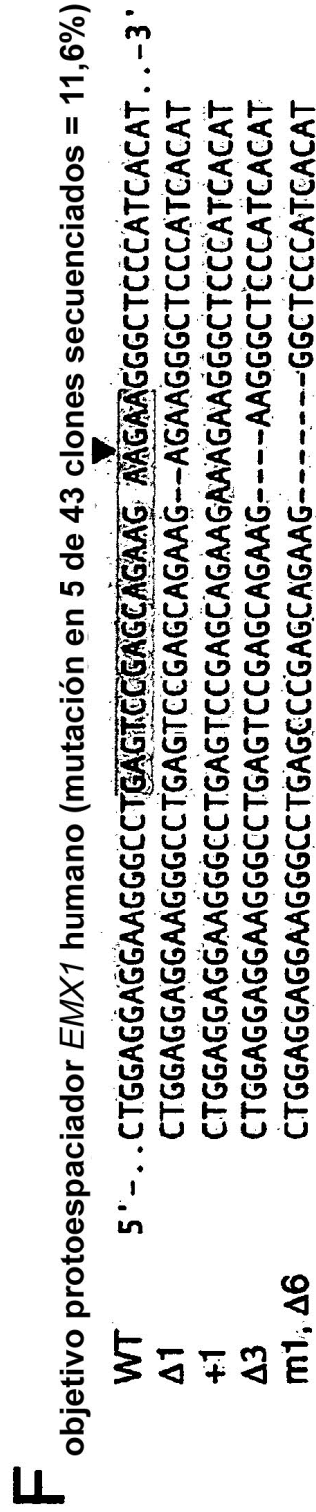
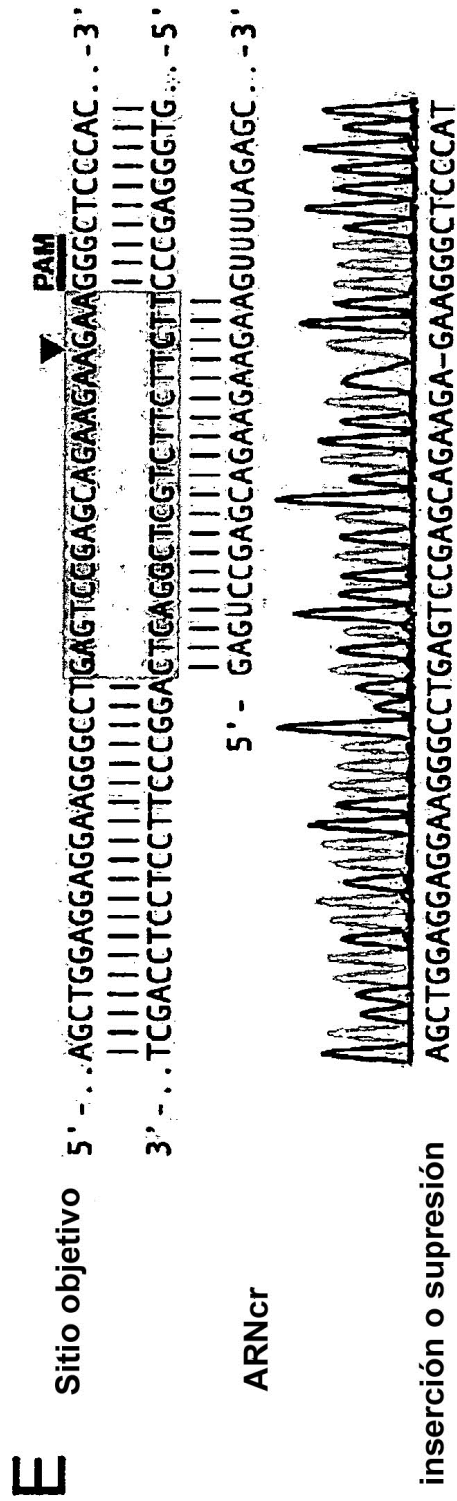
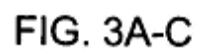


FIG. 2E-F



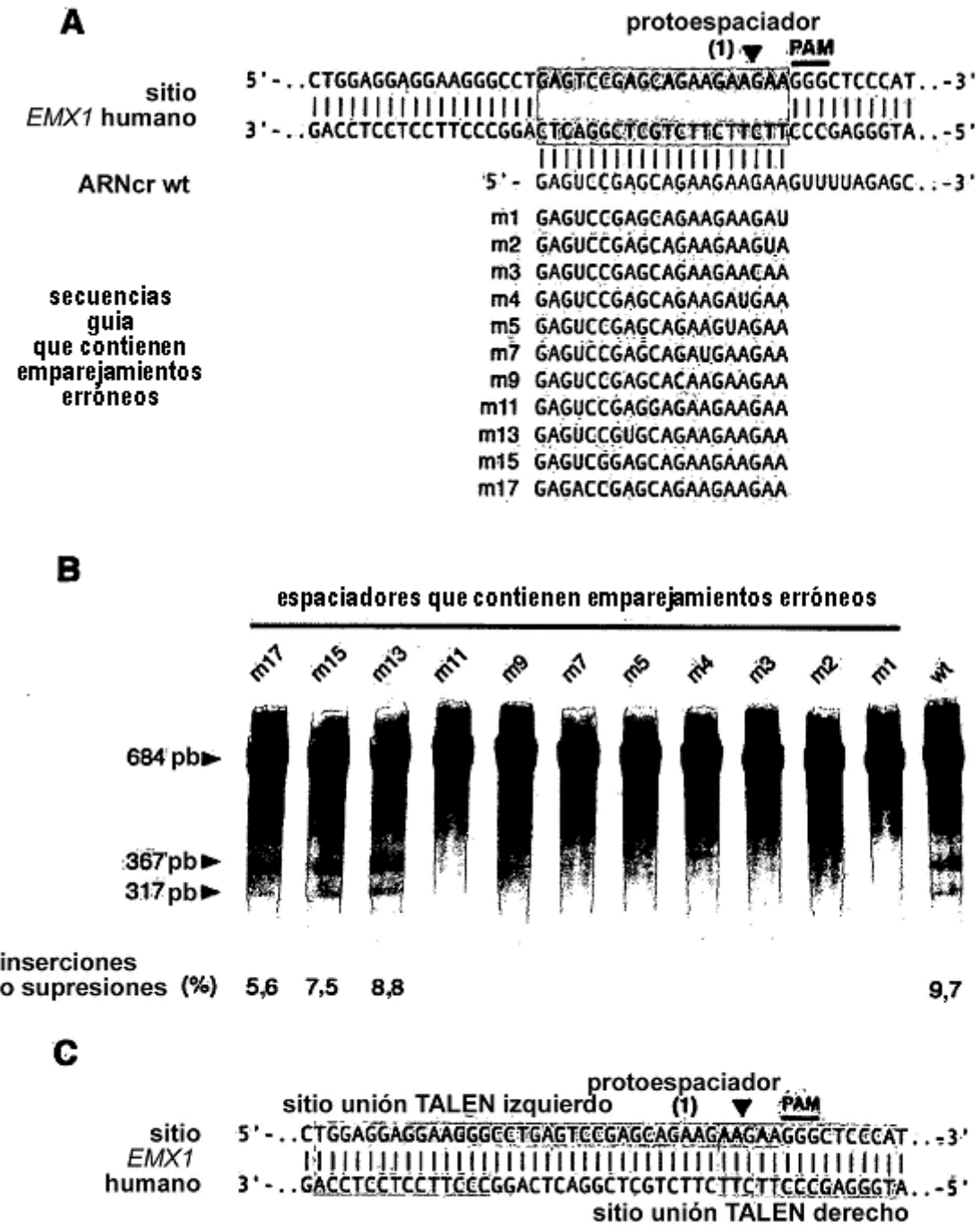


FIG. 4A-C

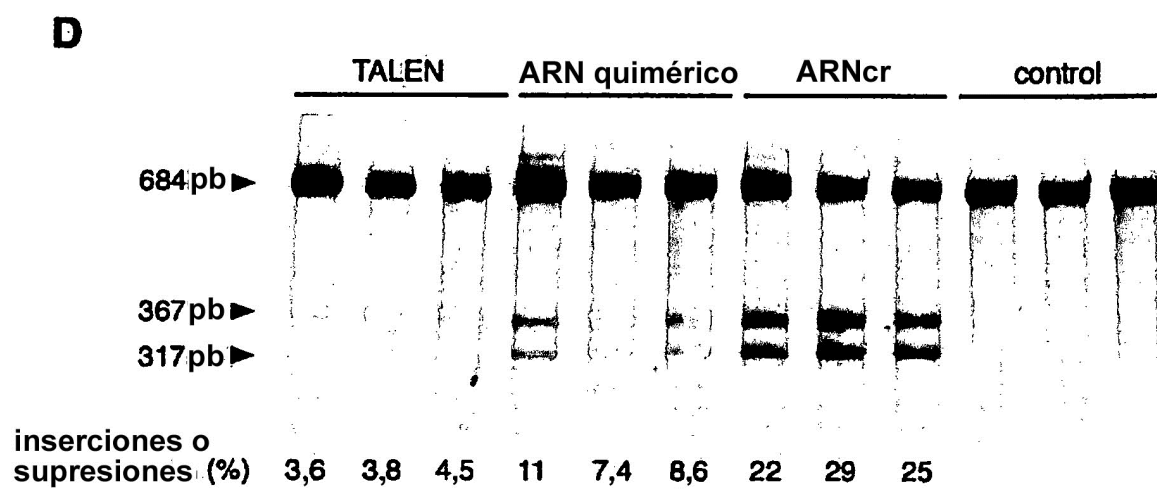


FIG. 4D

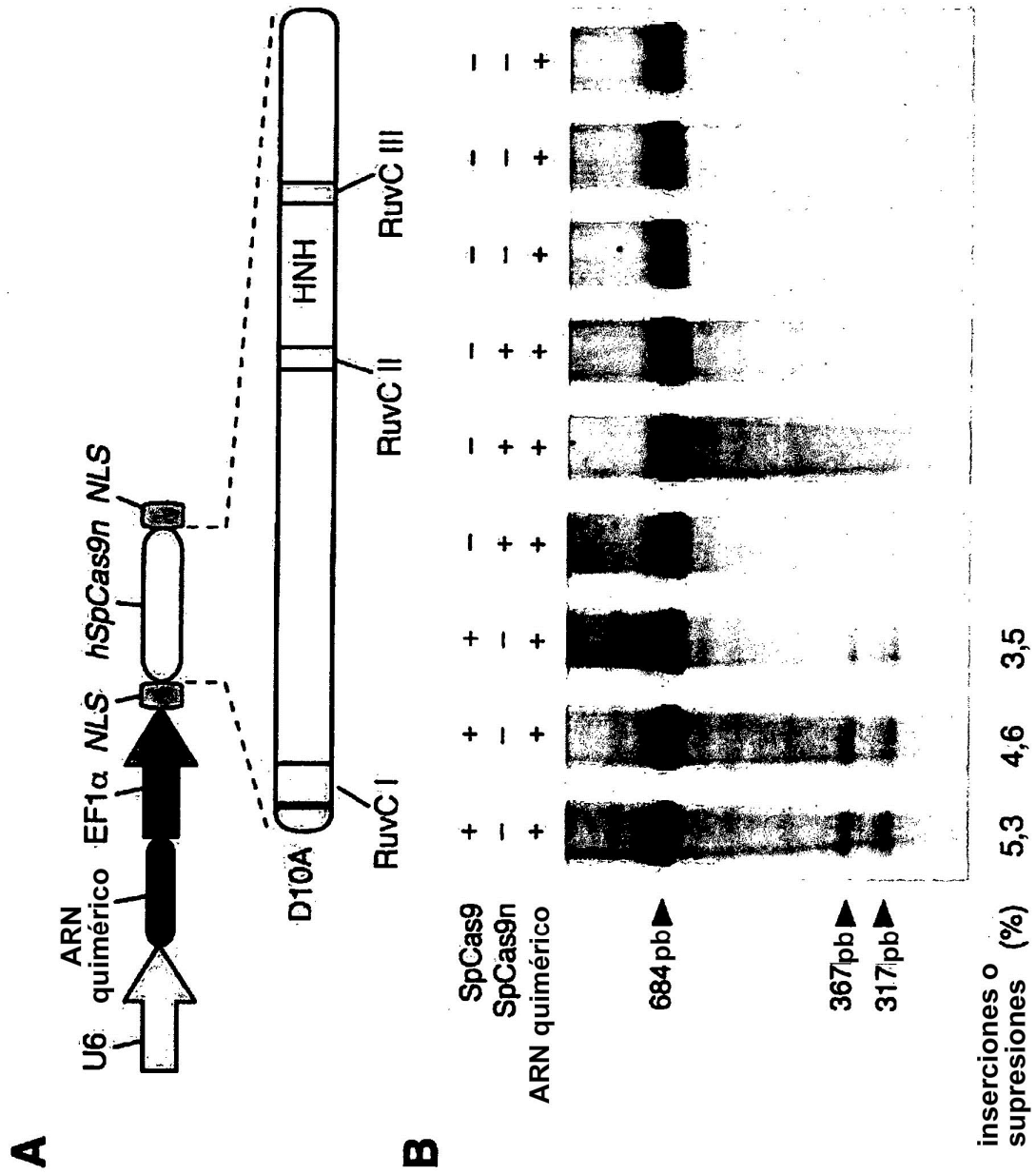
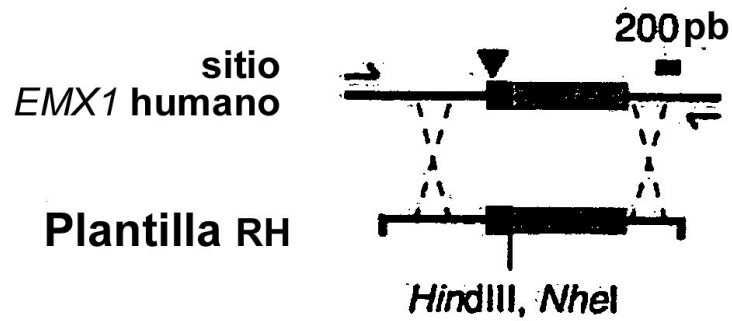
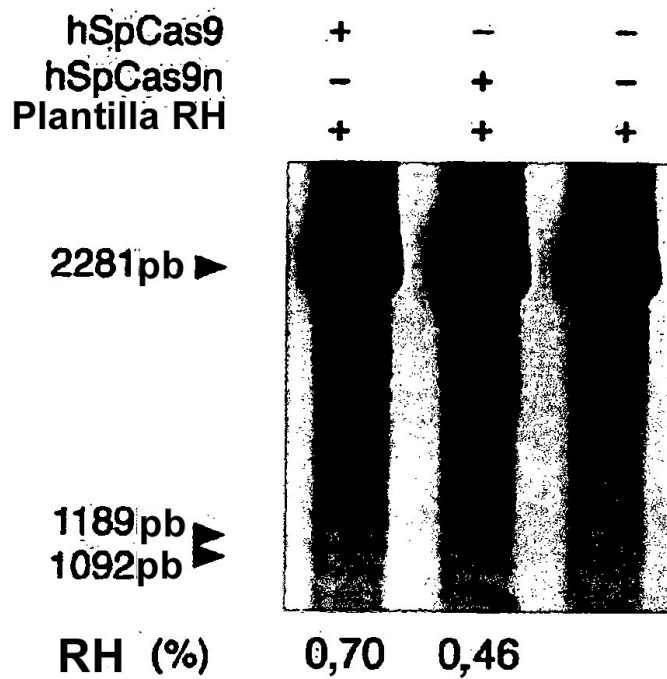


FIG. 5A-B

C



D



E

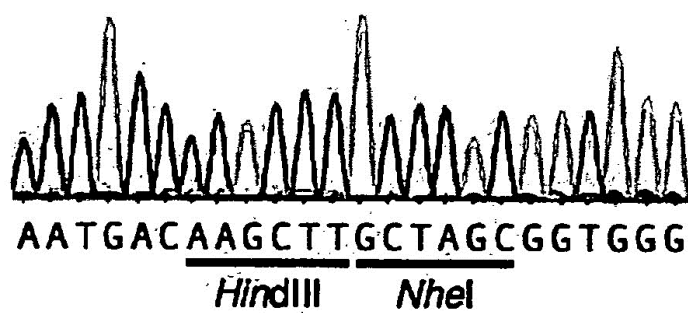


FIG. 5C-E

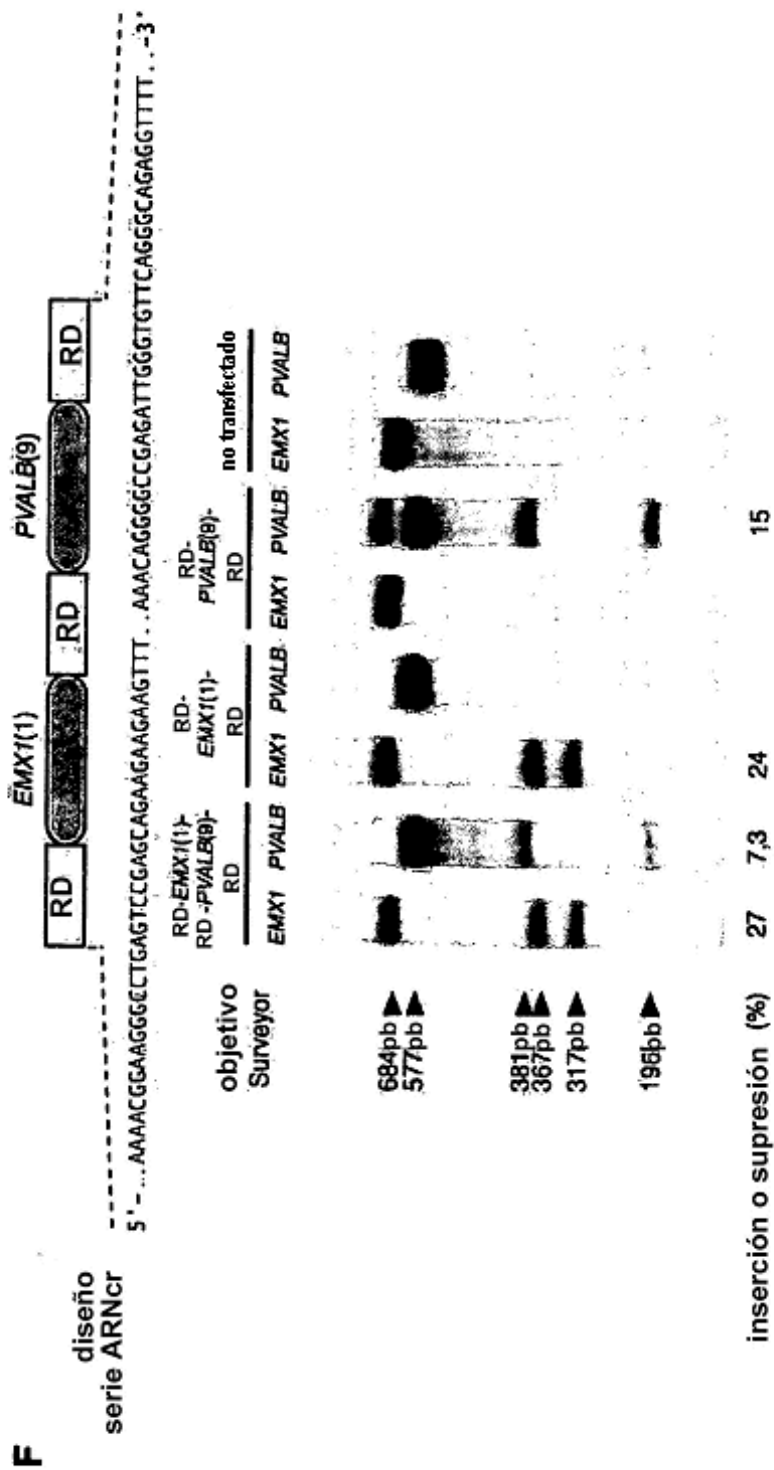


FIG. 5F

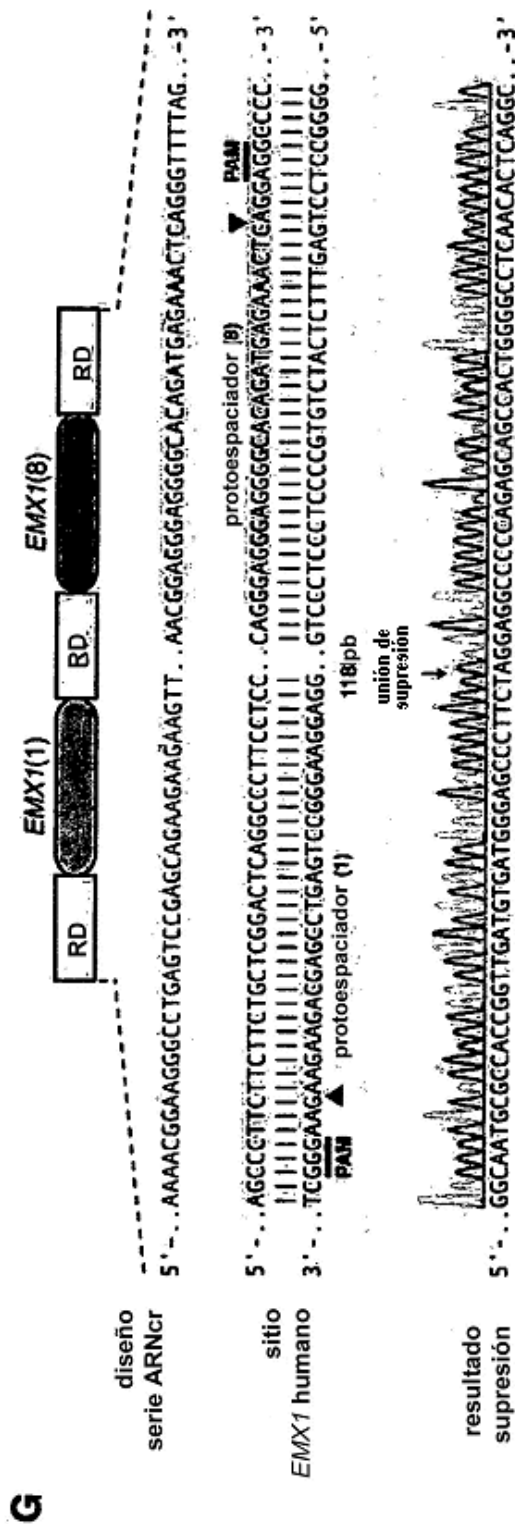
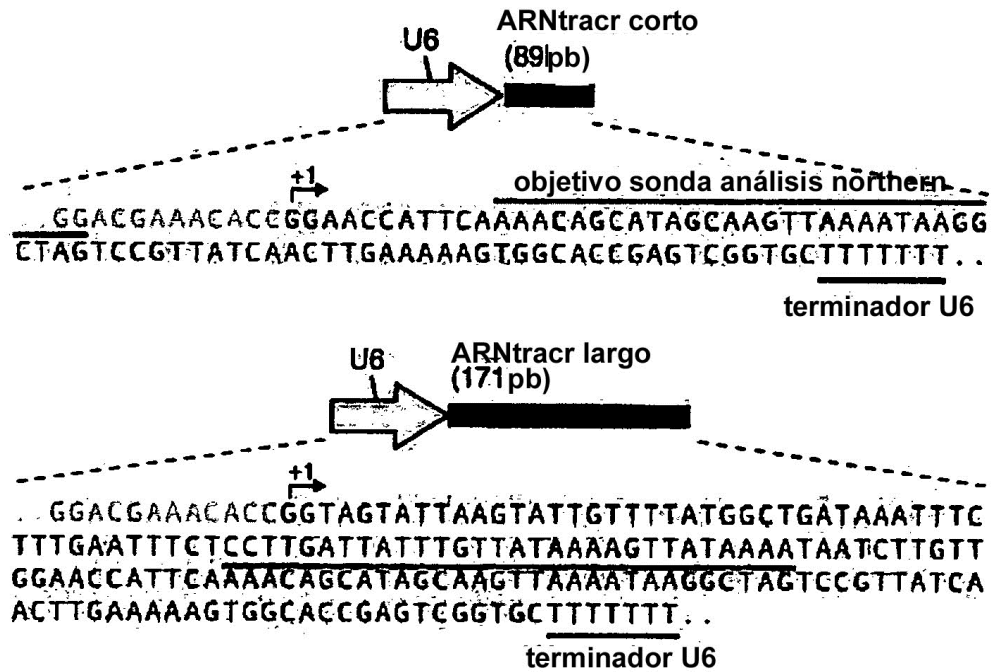


FIG. 5G

A



B

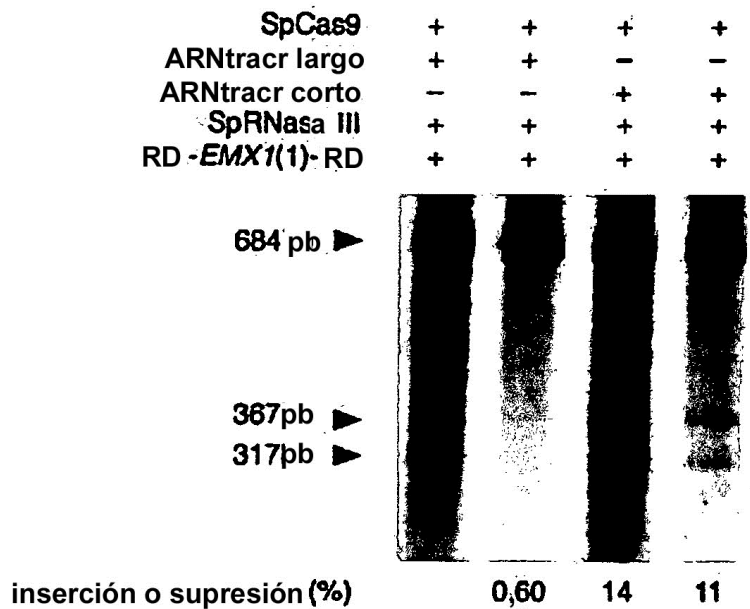


FIG. 6A-B

C

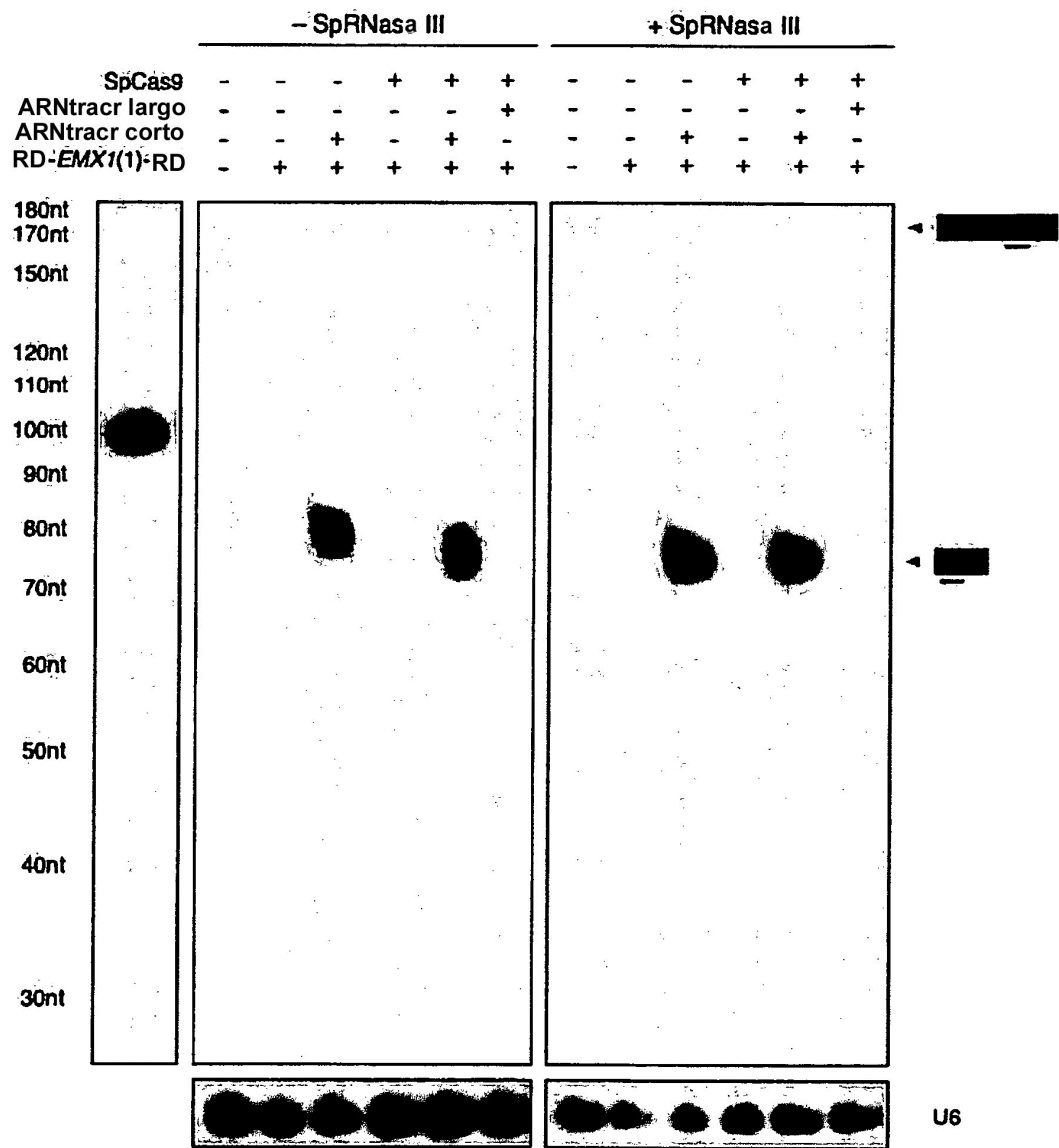


FIG. 6C

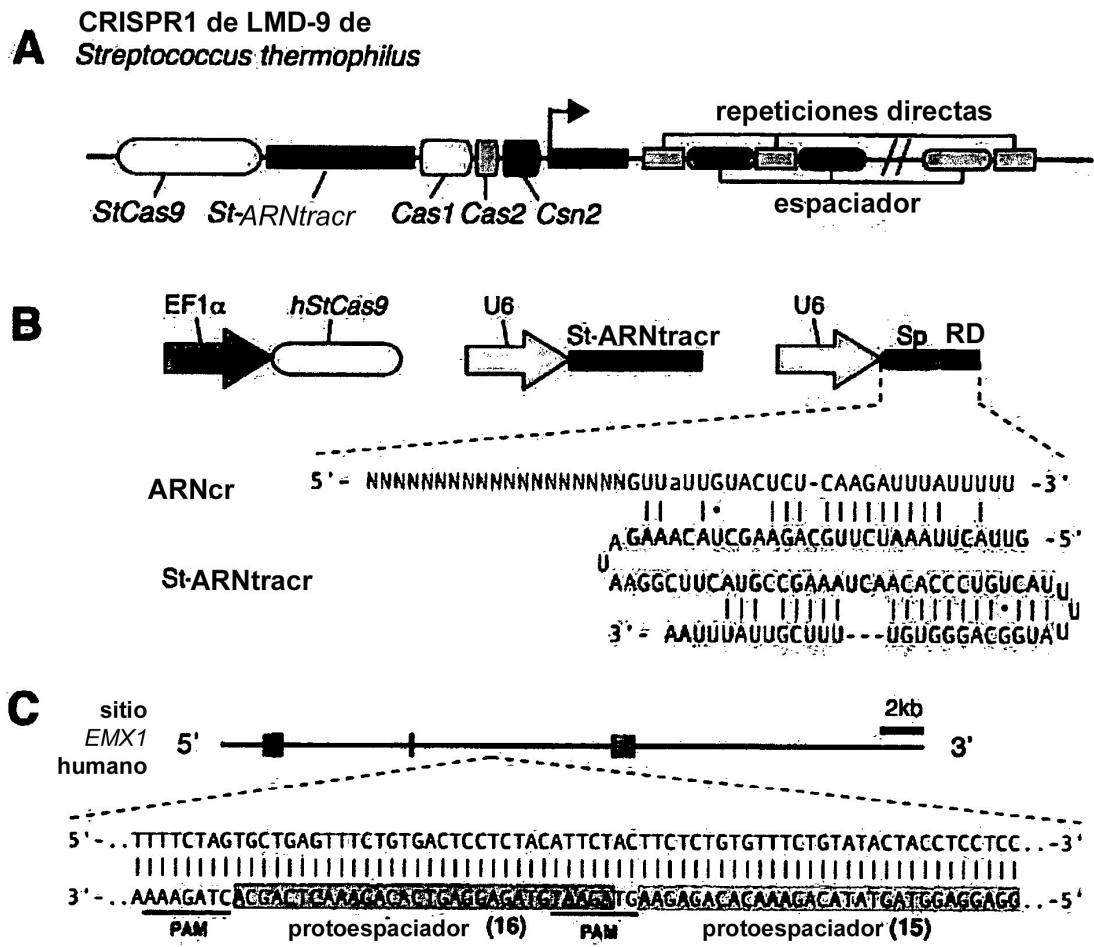


FIG. 7A-C

D

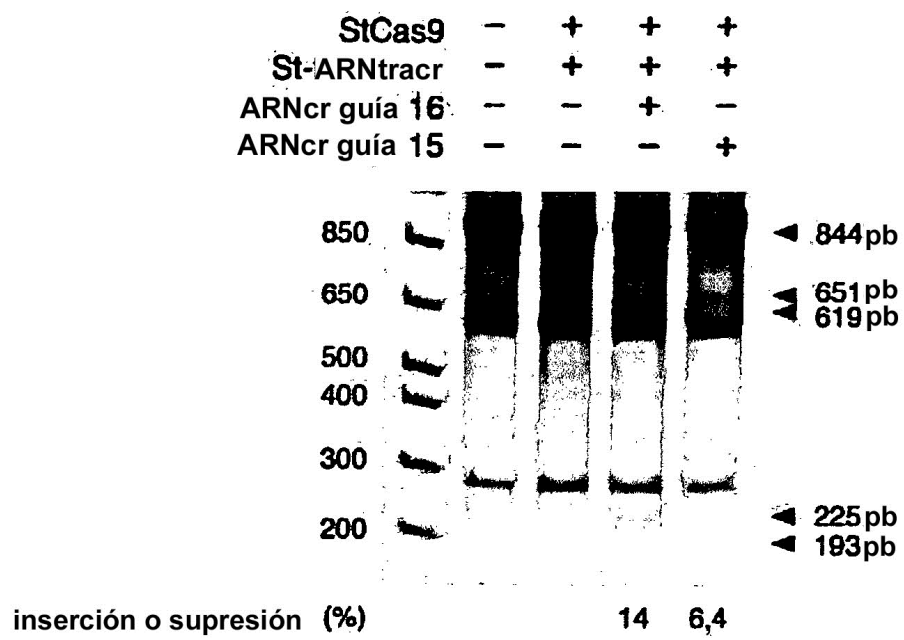
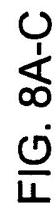
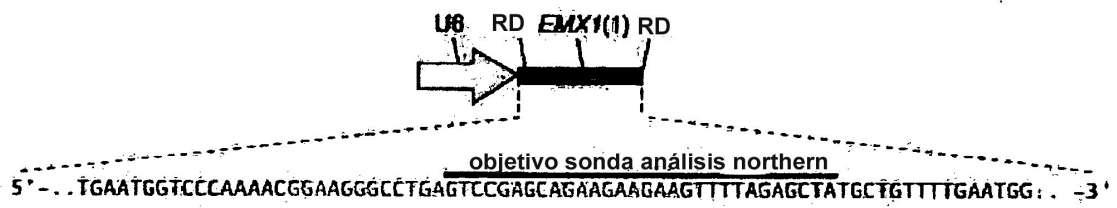


FIG. 7D



A



B

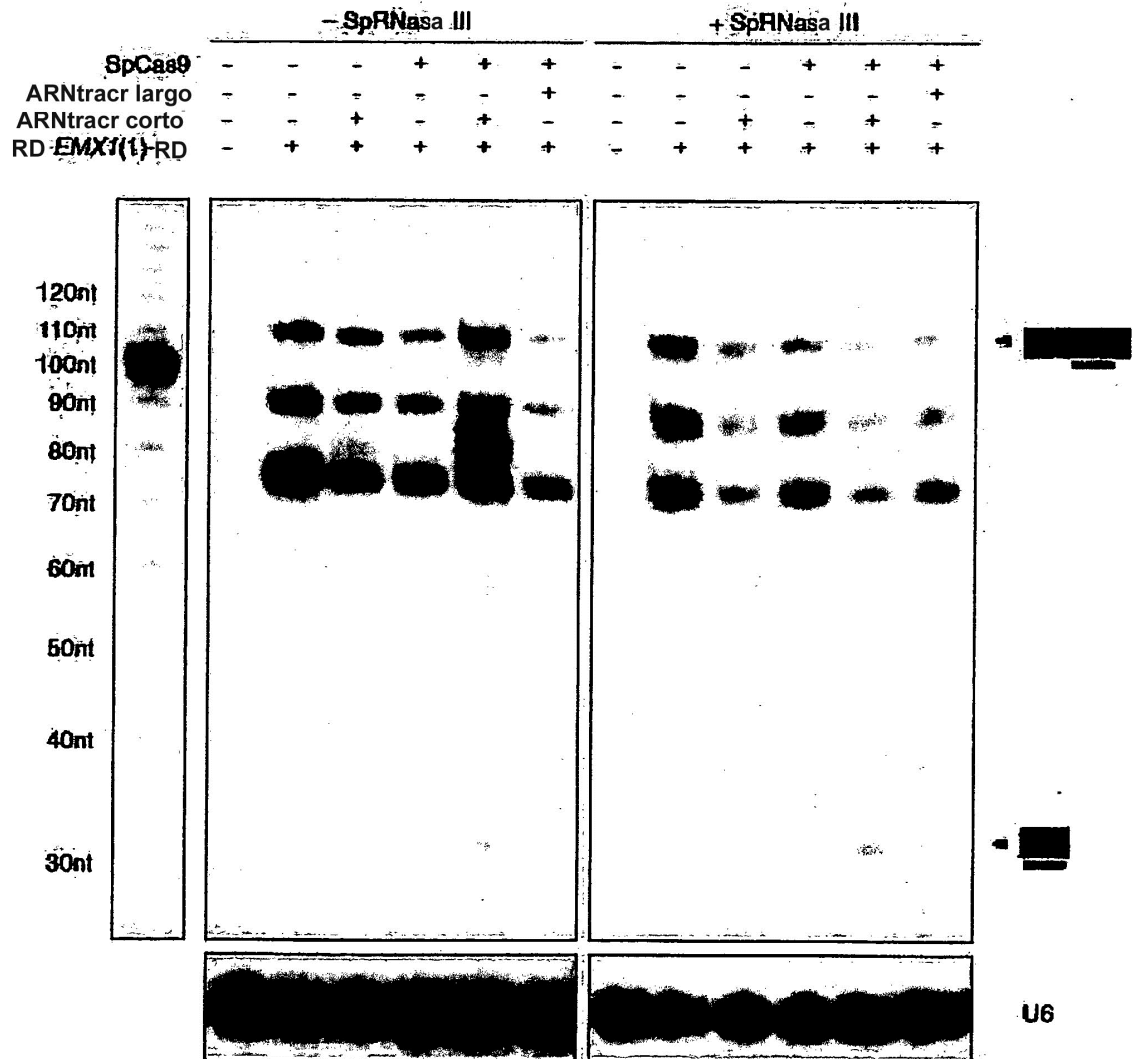


FIG. 9A-B

a

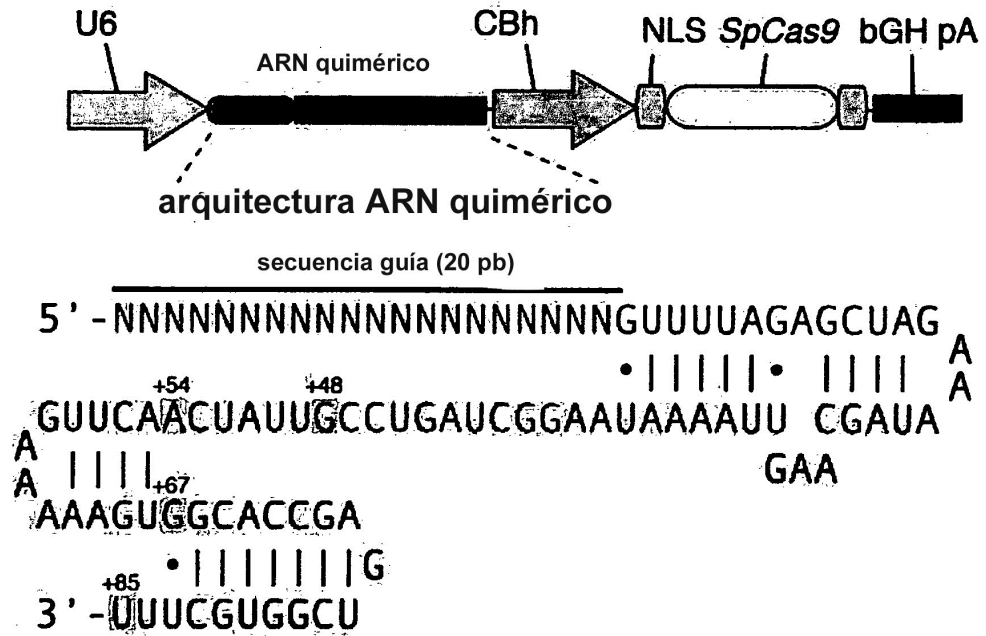
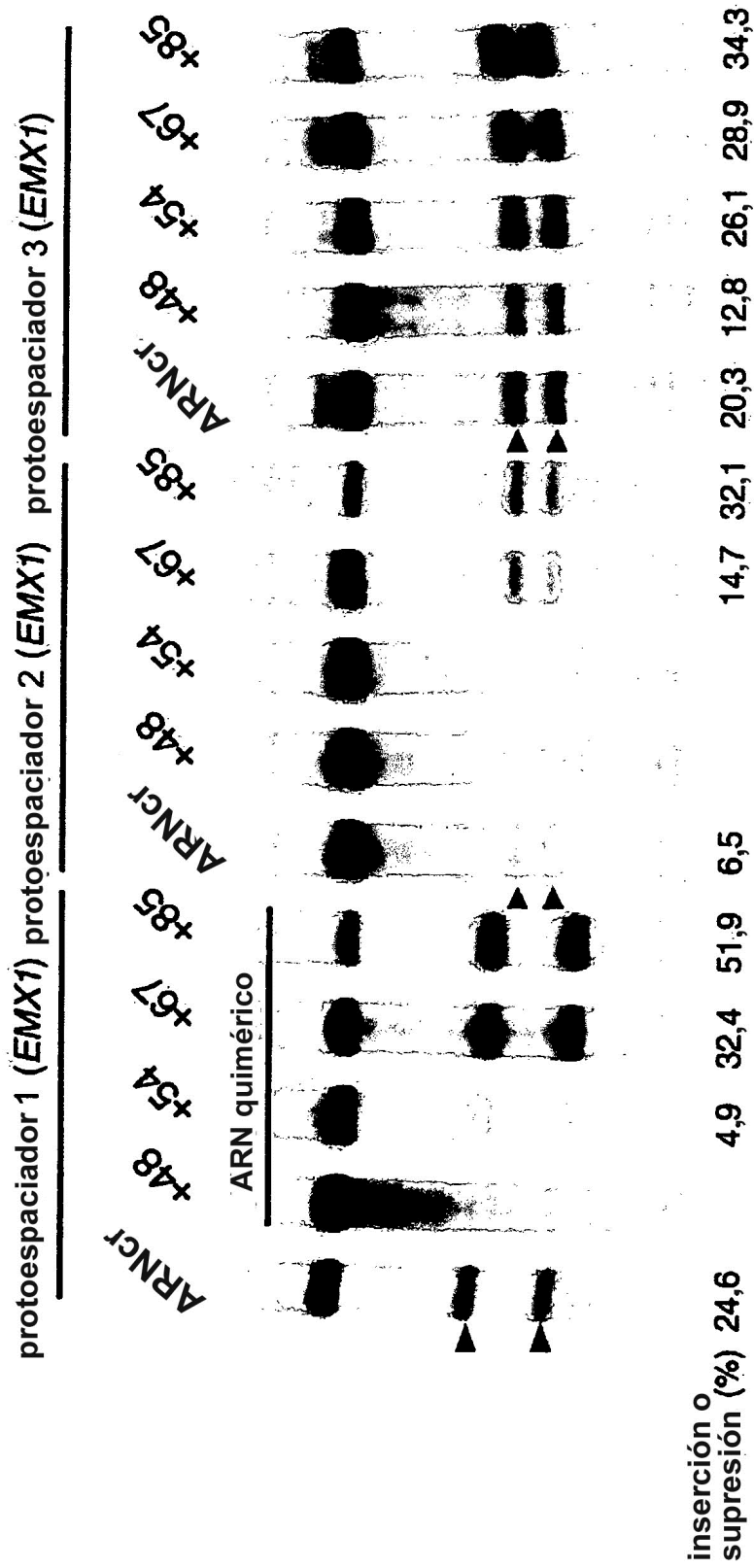


FIG. 10A



C

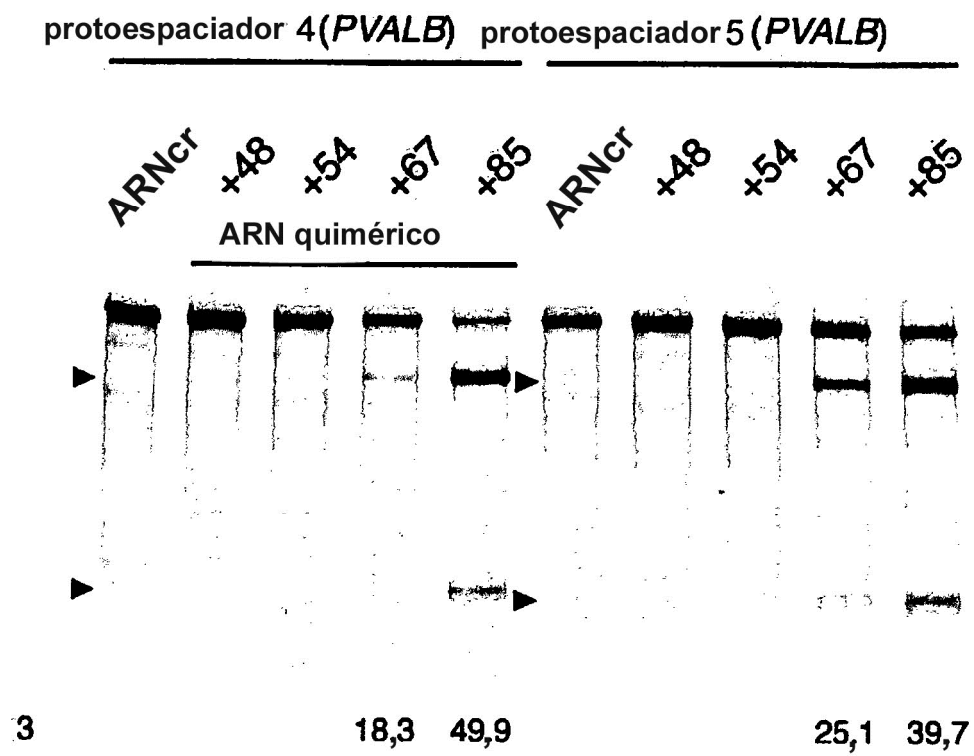


FIG. 10C

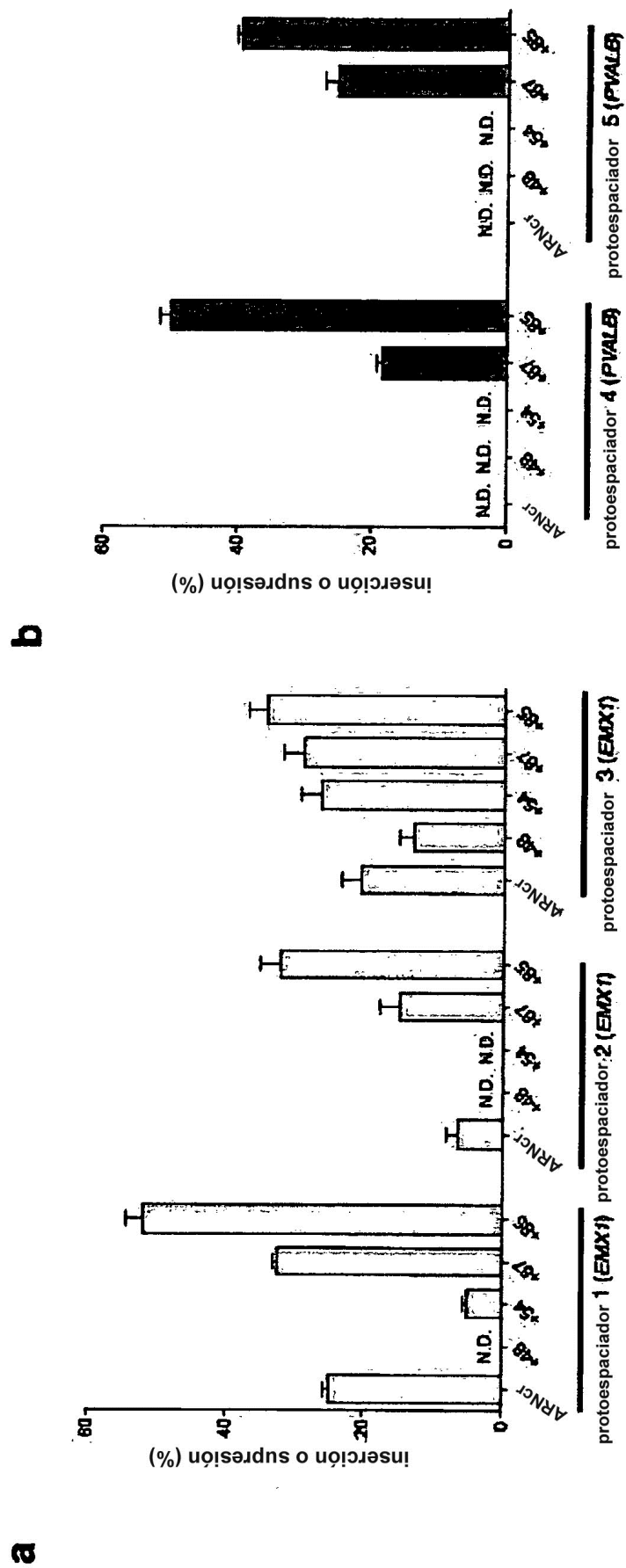


FIG. 11A-B

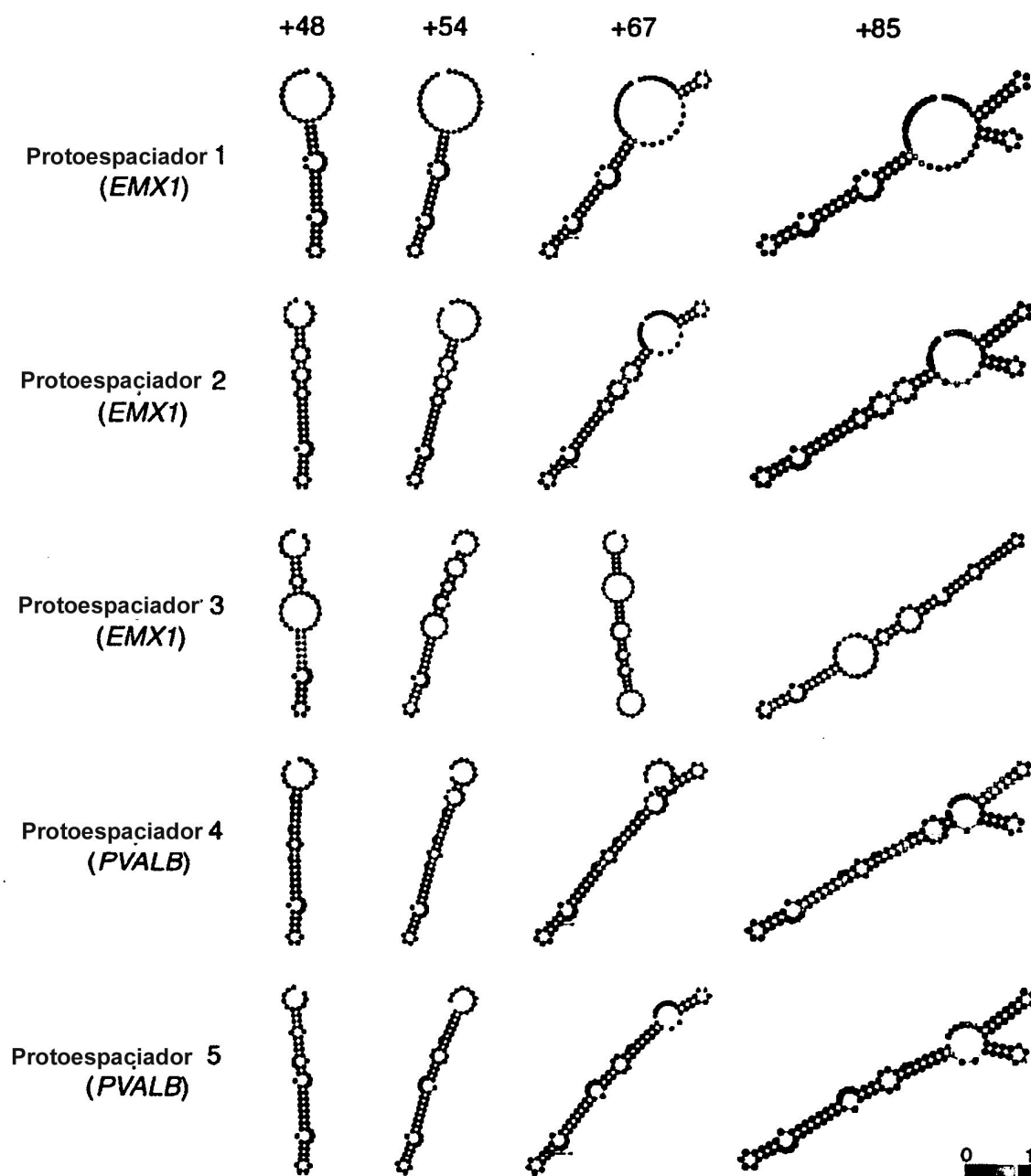


FIG. 12

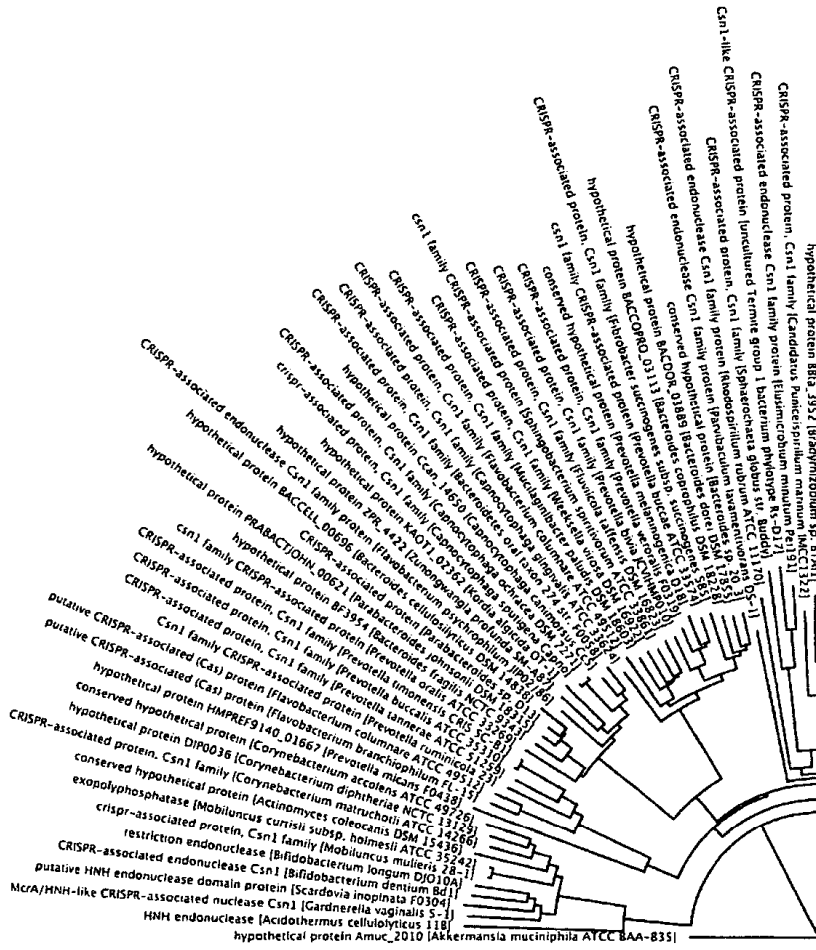


FIG. 13A

hypothetical protein = proteína hipotética; HNH endonuclease = endonucleasa HNH; Mcra/HNH-like CRISPR-associated nuclease Csn1 = Csn1 de endonucleasa asociada a CRISPR de tipo Mcra/HNH; putative HNH endonuclease domain protein = proteína de dominio de la endonucleasa HNH putativa; CRISPR-associated endonuclease Csn1 = Csn1 de endonucleasa asociada a CRISPR; Restriction endonuclease = endonucleasa de restricción; exopolphosphatase = exopolifosfatasa; conserved hypothetical protein = proteína hipotética conservada; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; putative CRISPR-associated (Cas) Protein = Proteína (Cas) asociada a CRISPR putativa;

FIG. 13B

hypothetical protein = proteína hipotética; CRISPR-associated protein, Csn1 family = proteína asociada a CRISPR, familia Csn1=; CRISPR-associated endonuclease Csn1 family protein, = proteína de familia Csn1 de endonucleasa asociada a CRISPR; CRISPR-associated protein Cas9/Csn1, subtype II= proteína Cas9/Csn1 asociada a CRISPR, subtipo II; CRISPR-associated Cas5e family protein= proteína de la familia Cas5e asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada

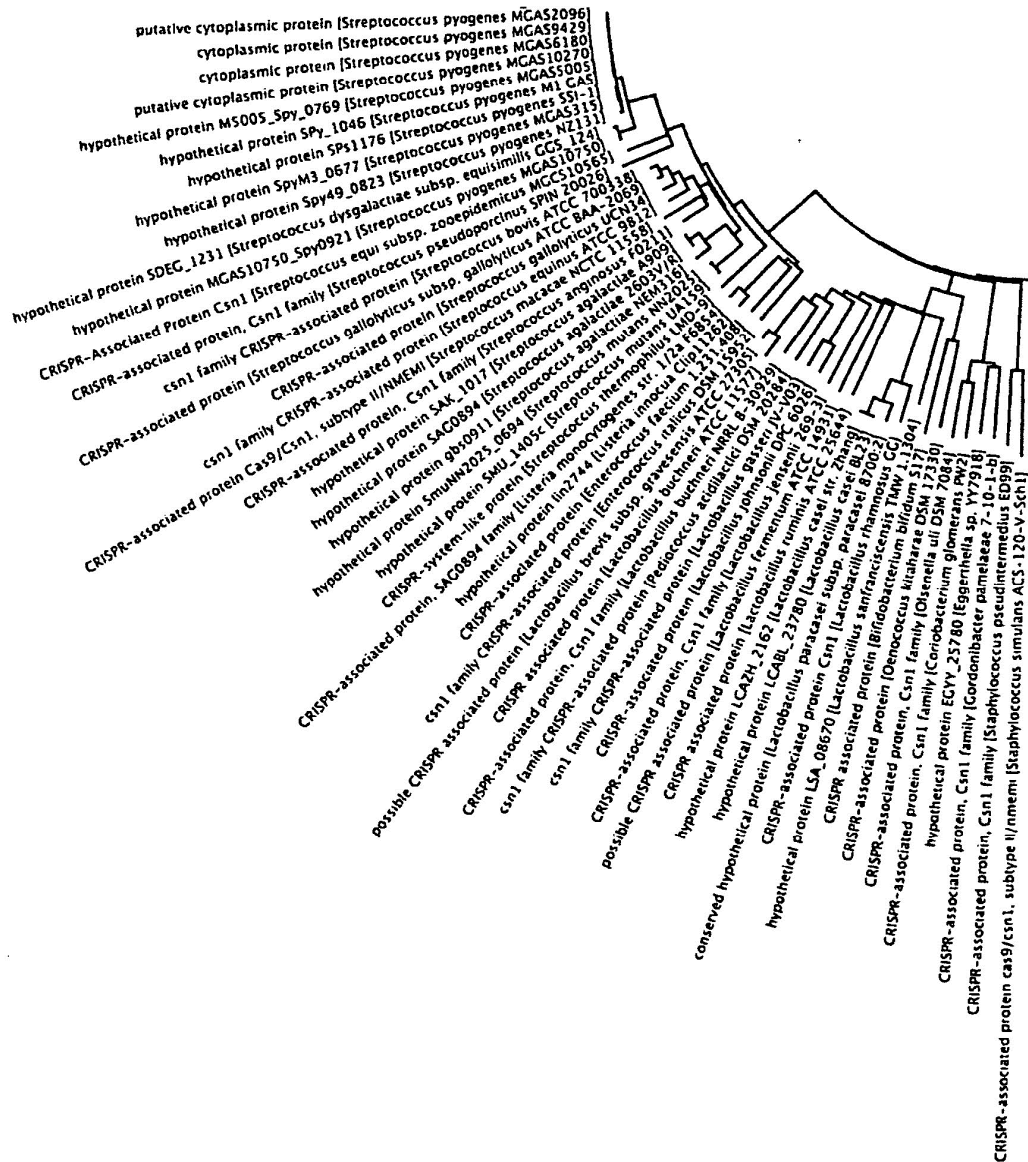
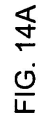


FIG. 13C

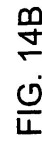
Putative cytoplasmic protein = proteína citoplasmática putativa; cytoplasmic protein = proteína citoplasmática; hypothetical protein = proteína hipotética; CRISPR-associated Protein Csn1 = Csn1 Proteína asociada a CRISPR; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated Protein Cas9/Csn1, subtype II = Proteína Cas9/Csn1 asociada a CRISPR, subtipo II; possible CRISPR-associated protein = Proteína asociada a CRISPR posible; conserved hypothetical protein = proteína hipotética conservada; CRISPR-system-like protein = protein de tipo sistema CRISPR

FIG. 13D

Putative CRISPR-associated Protein Csx12 family = familia Csx12 de proteína asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated Protein Cas9/Csn1 = Proteína Cas9/Csn1 asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada; CRISPR-system related protein = proteína relacionada con el sistema CRISPR; CRISPR-system like protein = proteína de tipo sistema CRISPR; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; membrane protein = proteína de membrana; CRISPR-associated large protein (provisional), putative = proteína grande asociada a CRISPR (provisional) putativa; CRISPR-associated CasSe = CasSe asociada a CRISPR.



Putative CRISPR-associated Protein = proteína asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated Protein Cas9/Csn1, subtype II = Proteína Cas9/Csn1 asociada a CRISPR, subtipo II; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; CRISPR-associated Cas5e family protein = proteína de la familia Cas5e asociada a CRISPR



Putative CRISPR-associated (Cas) Protein = proteína (Cas) asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada; Csn1-like CRISPR-associated protein [uncultured Termite Group 1 bacterium phylotype R_s-D17] = proteína asociada a CRISPR de tipo Csn1 [bacteria del grupo 1 termita no cultivada filotipo R_s-D17];

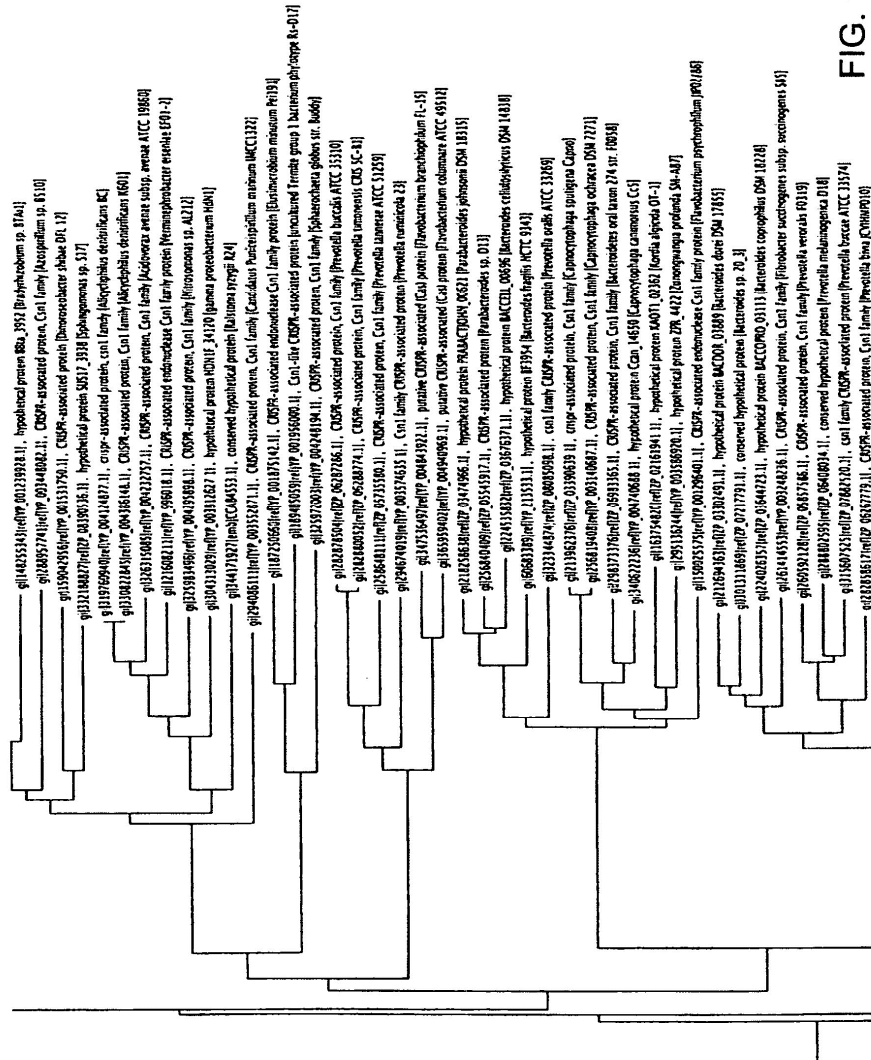
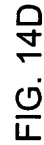


FIG. 14C

Putative CRISPR-associated (Cas) Protein = proteína (Cas) asociada a CRISPR putativa; hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada; Csn1-like CRISPR-associated protein [uncultured Termite Group 1 bacterium phylotype R_s-D17] = proteína asociada a CRISPR de tipo Csn1 [bacteria del grupo 1 termita no cultivada filotipo R_sD17];



Membrane protein = proteína de membrana; Hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated endonuclease Csn1 family protein = proteína de la familia Csn1 de endonucleasa asociada a CRISPR; conserved hypothetical protein = proteína hipotética conservada; Csn1-like CRISPR-associated protein = proteína asociada a CRISPR de tipo Csn1; CRISPR-system-like protein = Proteína de tipo sistema CRISPR; CRISPR-associated protein Cas9/Csn1= proteína Cas9/Csn1 asociada a CRISPR; CRISPR-system related protein = proteína relacionada con el sistema CRISPR

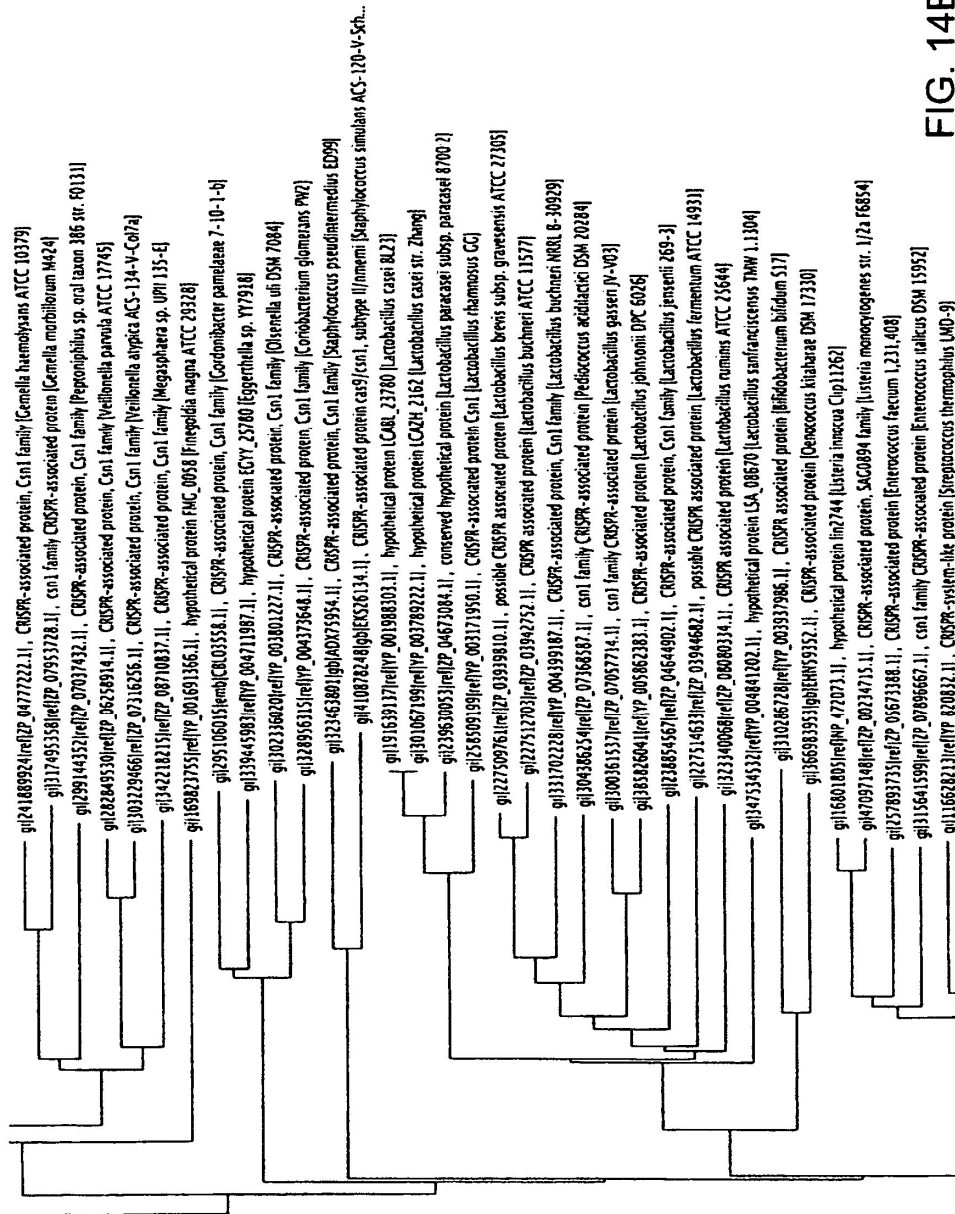


FIG. 14E

Hypothetical protein = proteína hipotética; CRISPR-associated Protein Csn1 family = familia Csn1 Proteína asociada a CRISPR; possible CRISPR-associated protein = Proteína asociada a CRISPR posible; conserved hypothetical protein = proteína hipotética conservada; CRISPR-associated protein cas9/csn1, subtype II/nmemi = proteína Cas9/Csn1 asociada a CRISPR, subtipo II nmemi; CRISPR-system-like protein = Proteína de tipo sistema CRISPR.

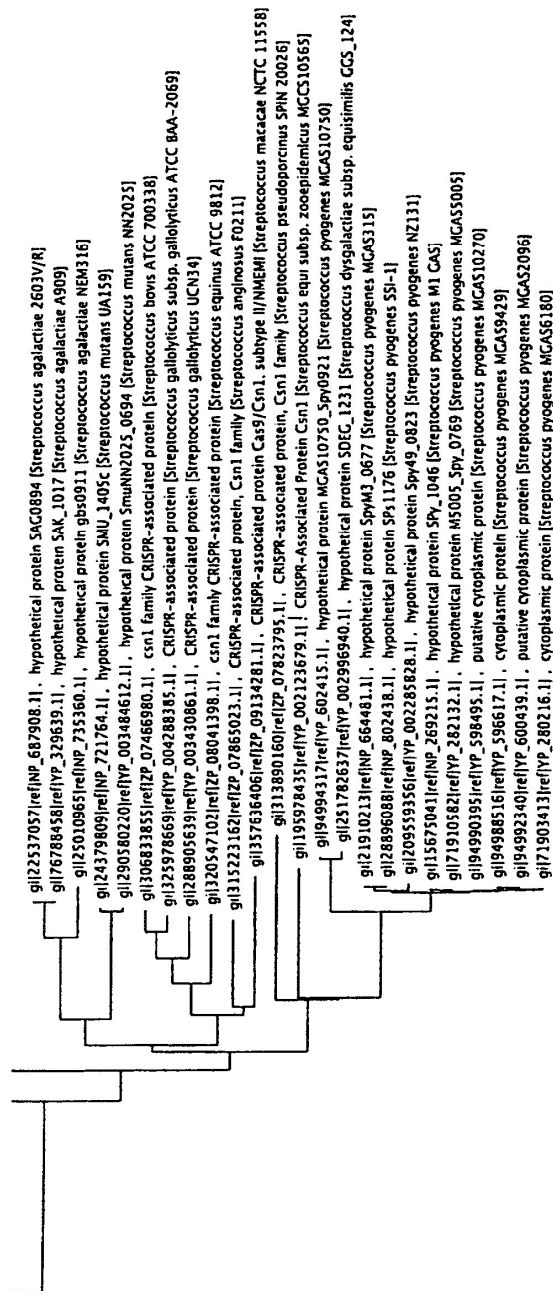


FIG. 14F

Hypothetical protein = proteína hipotética; CRISPR-associated Protein, Csn1 family = Proteína asociada a CRISPR, familia Csn1; CRISPR-associated protein Cas9/Csn1, subtype II= proteína Cas9/Csn1 asociada a CRISPR, subtipo II; Putative cytoplasmic protein = proteína citoplasmática putativa; cytoplasmic protein = proteína citoplasmática

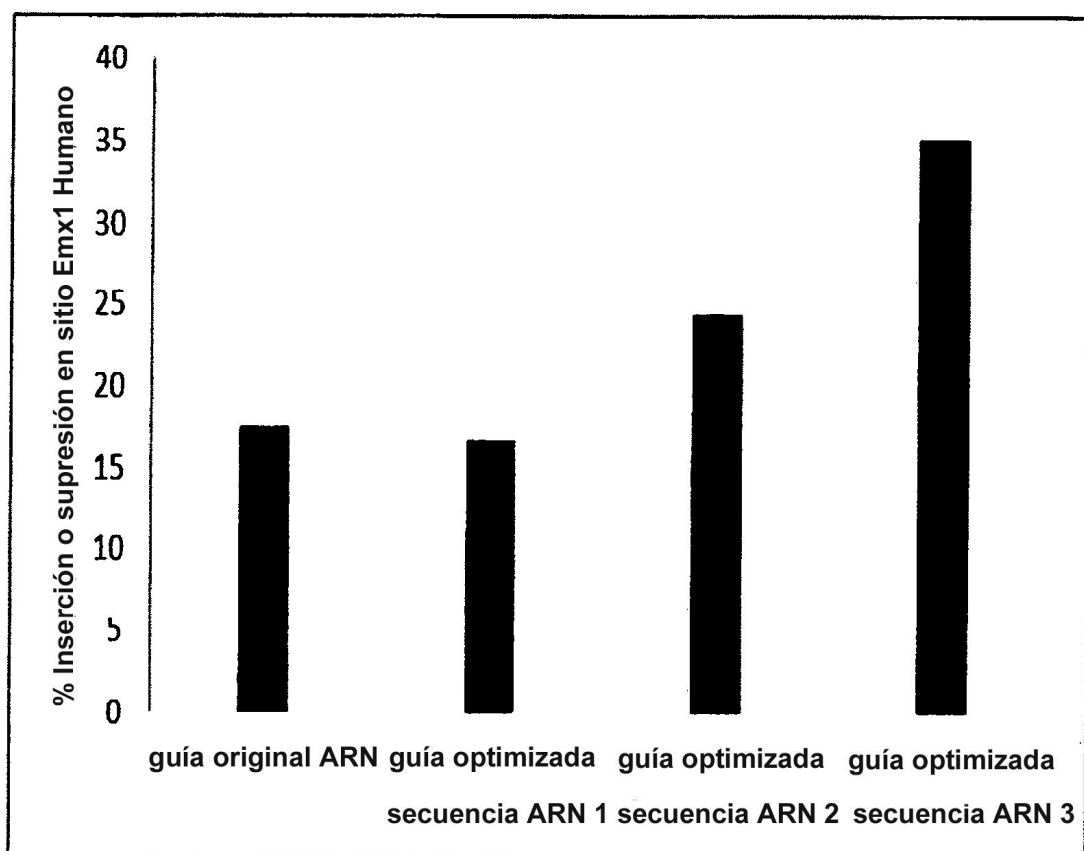
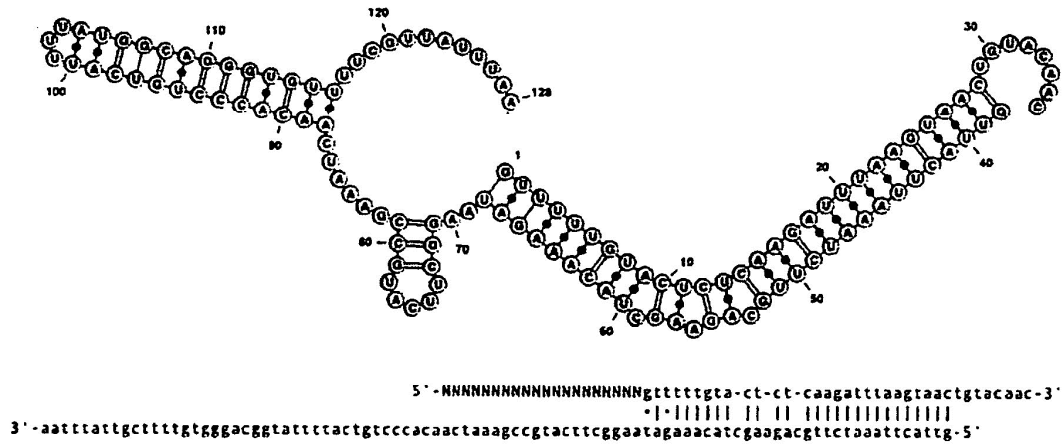


FIG. 15

ARN CRISPR1 de *S. thermophilus* co-plegado de ARNtracr con repetición directa



diseño ARN guía quimérico



FIG. 16

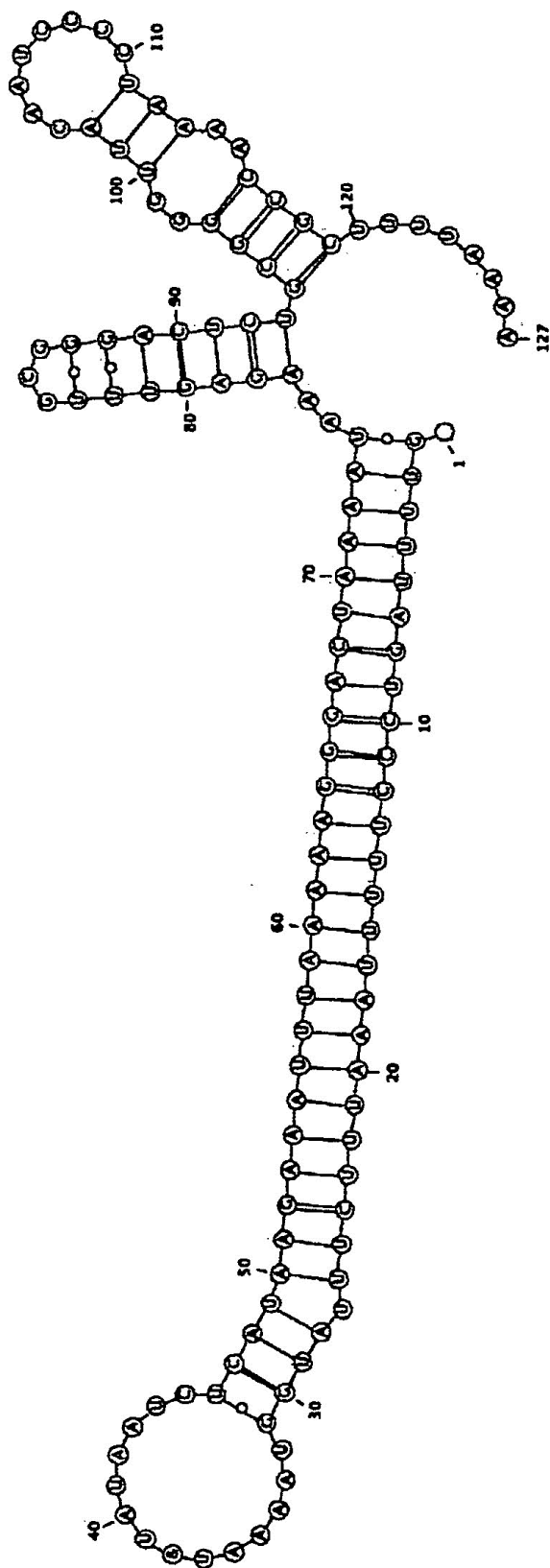
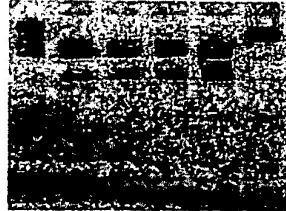


FIG. 17

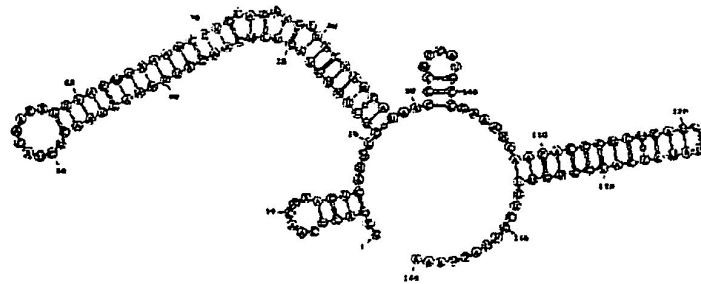
Ensayo S1Cas9

Escalera 1 2 3 4 5



Ruta 1: ARNtracr completo + espaciador completo_RD

full tracrRNA: GUUACUUAUUUUCUUGCAGAAGCUACAAGAUAGGCUUCAUGCCGAAUACAACACCCUGUCAUUUUUAGGCAGGGUGUUUUCGUUAL
full spacer_DR: GGGACUCAACCAAGUCAUUCGUUUUUGUACUCUCAAGAUUUAAGUAAACUGUACAAC



Ruta 2: ARNtracr maduro + espaciador maduro_RD

mature tracrRNA: CUUGCAGAAGCUACAAGAUAGGCUUCAUGCCGAAUACAACACCCUGUCAUUUUUAGGCAGGGUGUUUU
mature spacer_DR: GGGACUCAACCAAGUCAUUCGUUUUUGUACUCUCAAGAUUUA

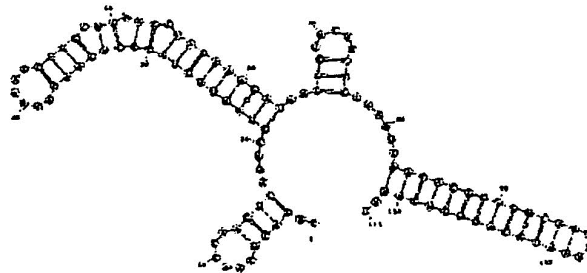
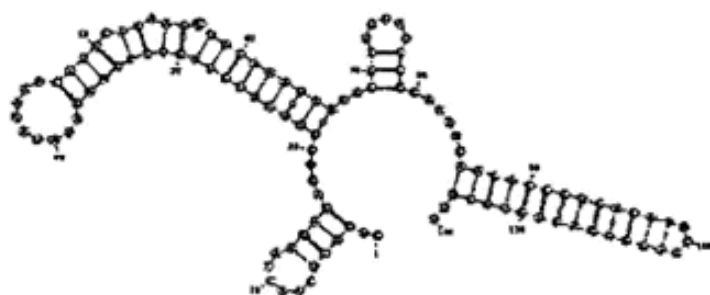


FIG. 18A

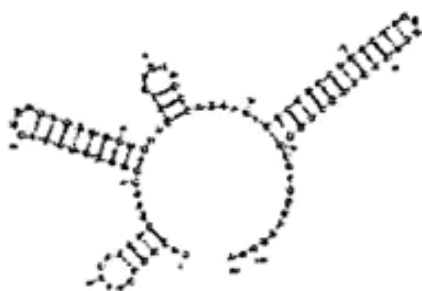
Ruta 3: diseño ARN guía quimérico 1 (1-22 RD, 12-81 ARNtracr)

secuencia : GGGACGACACCAAGGTCAGCCGCTGTTGGGCGCCGCGAGGATGAGGACCTGCGAGGCTGCGAGGCGAGGAGCCGACGCGGATGACACACCCGCGCTGTTGGGATGGGCGAGGCGGCGG



Ruta 4: diseño ARN guía quimérico 2 (1-9 RD, 20-91 ARNtracr)

secuencia : GGGACGACACCAAGGTCAGCCGCTGTTGGGCGCCGCGAGGATGAGGACCTGCGAGGCTGCGAGGCGAGGAGCCGACGCGGATGACACACCCGCGCTGTTGGGATGGGCGAGGCGGCGG



Ruta 5: diseño ARN guía quimérico 3 (1-9 RD, 20-46 ARNtracr)

secuencia : GGGACGACACCAAGGTCAGCCGCTGTTGGGCGCCGCGAGGATGAGGACCTGCGAGGCTGCGAGGCGAGGAGCCGACGCGGATGACACACCCGCGCTGTTGGGATGGGCGAGGCGGCGG

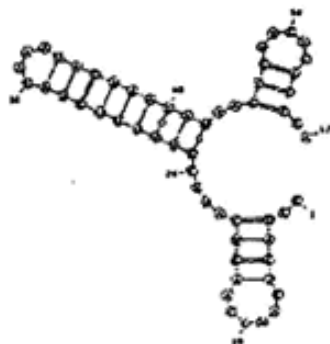
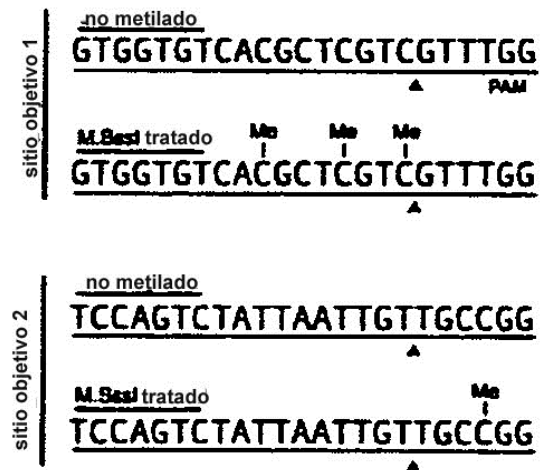


FIG. 18B

a



b

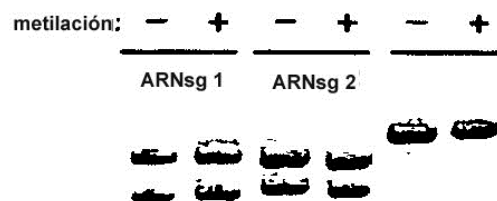


FIG. 19A-B

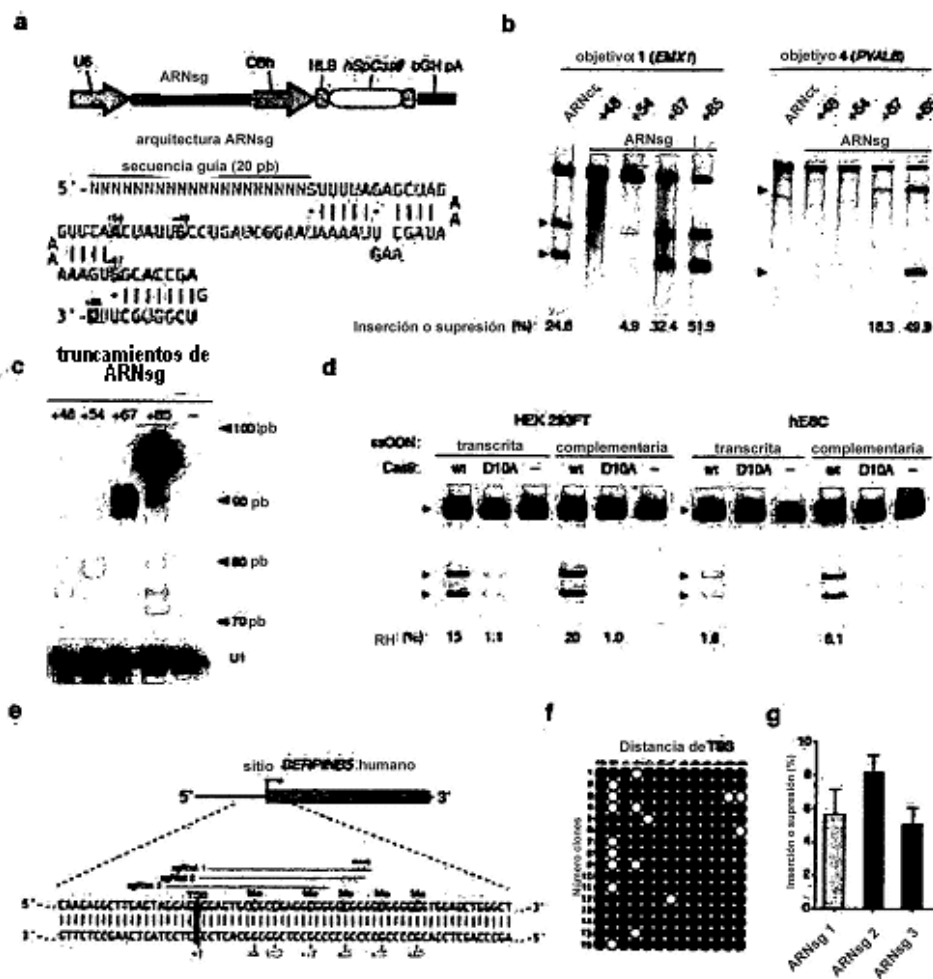


FIG. 20A-G

Navegador Genomas UCSC en Ser Humano Feb. 2.009 (GRCh37/hg19) Montaje

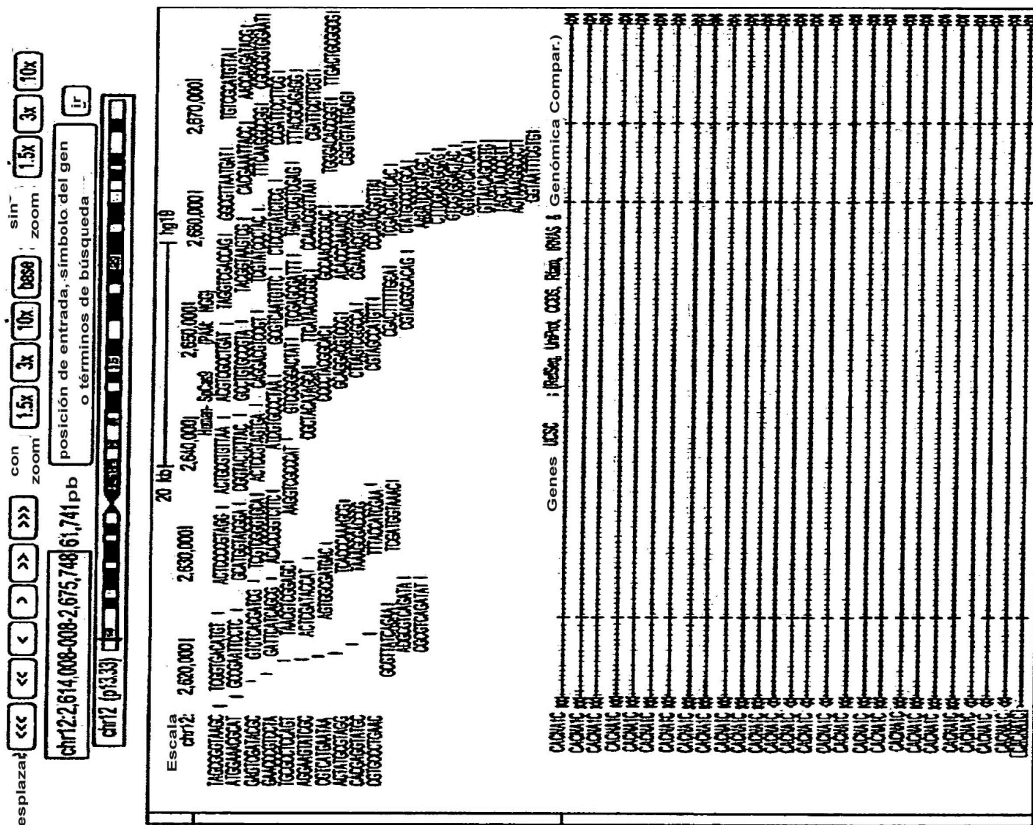
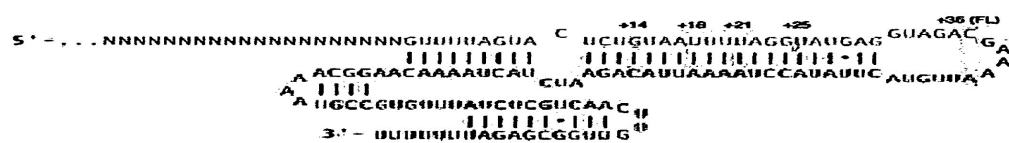
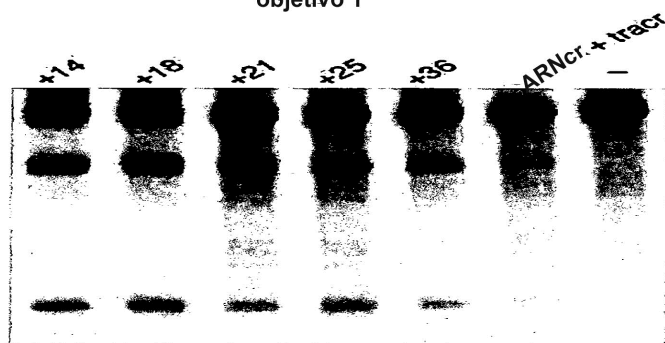


FIG. 22

A

**B**

objetivo 1



Insertión o supresión (%):

20 20 12 14 11

FIG. 23A-B