

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 752 175**

51 Int. Cl.:

**C12N 15/09** (2006.01)

**C07K 19/00** (2006.01)

**C12N 9/78** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.03.2015 PCT/JP2015/056436**

87 Fecha y número de publicación internacional: **11.09.2015 WO15133554**

96 Fecha de presentación y número de la solicitud europea: **04.03.2015 E 15758734 (6)**

97 Fecha y número de publicación de la concesión europea: **02.10.2019 EP 3115457**

54 Título: **Método de modificación de la secuencia genómica para convertir específicamente bases de ácido nucleico de una secuencia de ADN seleccionada como diana, y complejo molecular para su uso en el mismo**

30 Prioridad:

**05.03.2014 JP 2014043348**  
**30.09.2014 JP 2014201859**

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

**03.04.2020**

73 Titular/es:

**NATIONAL UNIVERSITY CORPORATION KOBE UNIVERSITY (100.0%)**  
**1-1 Rokkodai-cho, Nada-ku**  
**Kobe-shi, Hyogo 657-8501, JP**

72 Inventor/es:

**NISHIDA, KEIJI;**  
**KONDO, AKIHIKO y**  
**KOJIMA, SATOMI**

74 Agente/Representante:

**FÚSTER OLAGUIBEL, Gustavo Nicolás**

**ES 2 752 175 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**

Método de modificación de la secuencia genómica para convertir específicamente bases de ácido nucleico de una secuencia de ADN seleccionada como diana, y complejo molecular para su uso en el mismo

**[Campo técnico]**

La presente invención se refiere a un método de modificación de una secuencia genómica, que permite la modificación de una base de ácido nucleico en una región particular de un genoma, sin escindir el ADN bicatenario (sin escisión o escisión de cadena sencilla), y sin insertar un fragmento de ADN foráneo, y a un complejo de un módulo de reconocimiento de secuencia de ácido nucleico y a una enzima convertidora de bases de ácido nucleico usada para eso.

**[Antecedentes de la técnica]**

En los últimos años, la edición genómica está atrayendo la atención como una técnica para modificar el gen y la región genómica de interés en diversas especies. De manera convencional, como método de edición genómica, se ha propuesto un método que utiliza una nucleasa artificial que comprende una combinación de una molécula que tiene una capacidad de escisión del ADN independiente de la secuencia y una molécula que tiene una capacidad de reconocimiento de secuencia (documento no de patente 1).

Por ejemplo, se han notificado un método de realización de recombinación en un locus génico diana en el ADN en una célula vegetal o célula de insecto como huésped, usando una nucleasa con dedos de cinc (ZFN) en el que se unen un dominio de unión al ADN con dedos de cinc y un dominio de escisión de ADN no específico (documento de patente 1); un método de escisión o modificación de un gen diana en una secuencia de nucleótidos particular o un sitio adyacente al mismo usando TALEN, en el que un efector similar al activador de transcripción (TAL), que es un módulo de unión al ADN que tiene la bacteria patógena vegetal *Xanthomonas*, y una endonucleasa de ADN se unen (documento de patente 2); un método que utiliza un sistema CRISPR-Cas9 en el que la secuencia de ADN CRISPR (repeticiones palindrómicas cortas agrupadas regularmente interespaciadas), que funciona en un sistema inmunitario adquirido poseído por *eubacterium* y *archaebacterium*, y la familia de proteínas nucleasa Cas (asociada a CRISPR) que tiene una función importante junto con CRISPR se combinan (documento de patente 3) y similares. Además, también se ha notificado un método de escisión de un gen diana en la proximidad de una secuencia particular, usando nucleasa artificial en el que una proteína PPR configurada para reconocer una secuencia de nucleótidos particular mediante una serie de motivos PPR que consiste cada uno en 35 aminoácidos y que reconoce una base de ácido nucleico, y nucleasa se unen (documento de patente 4).

El documento WO 2010/132092 A2 se refiere a fusiones de citidina desaminasa y a métodos relacionados. El documento US 2011/0104787 A1 se refiere a péptidos de fusión que se unen a y modifican secuencias de ácido nucleico diana.

**[Lista de documentos]**

[Documentos de patente]

Documento de patente 1: JP-B-4968498

Documento de patente 2: publicación nacional de solicitud de patente internacional n.º 2013-513389

Documento de patente 3: publicación nacional de solicitud de patente internacional n.º 2010-519929

Documento de patente 4: JP-A-2013-128413

[Documento no de patente]

Documento no de patente 1: Kelvin M Esvelt, Harris H Wang (2013) Genome-scale engineering for systems and synthetic biology, *Molecular Systems Biology* 9: 641

**[Sumario de la invención]**

[Problemas que van a resolverse mediante la invención]

Las técnicas de edición genómica que se han propuesto hasta ahora presuponen básicamente roturas de ADN bicatenario (DSB). Sin embargo, ya que implican modificaciones genómicas inesperadas, se producen efectos secundarios tales como fuerte citotoxicidad, reordenamiento cromosómico y similares, y tienen problemas comunes de fiabilidad deteriorada en terapia génica, un número sumamente pequeño de células supervivientes mediante modificación de nucleótidos, y dificultad en la propia modificación genética en un óvulo de primate y

microorganismos unicelulares.

Por tanto, un objeto de la presente invención es proporcionar un nuevo método de edición genómica para modificar una base de ácido nucleico de una secuencia particular de un gen sin DSB o la inserción de un fragmento de ADN foráneo, es decir, sin escisión de un ADN bicatenario ni escisión de cadena sencilla, y un complejo de un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico del mismo.

[Medios para resolver los problemas]

La presente invención se define, entre otros, mediante los siguientes puntos:

1. Un método de modificación de un sitio seleccionado como diana de un ADN bicatenario, que comprende una etapa de poner en contacto un complejo que comprende un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a una secuencia de nucleótidos diana en una ADN bicatenario dado unido a una enzima convertidora de bases de ácido nucleico, con dicho ADN bicatenario, para convertir uno o más nucleótidos en el sitio seleccionado como diana en uno o más de otros nucleótidos o delecionar uno o más nucleótidos, o insertar uno o más nucleótidos en dicho sitio seleccionado como diana, en el que el módulo de reconocimiento de secuencia de ácido nucleico es un sistema CRISPR-Cas, en el que el sistema CRISPR-Cas comprende una proteína Cas que tiene actividad nickasa capaz de escindir sólo una de las cadenas del ADN bicatenario, en el que el método no es un método para el tratamiento del cuerpo humano o animal mediante cirugía o terapia, y en el que el método no es un método para modificar la identidad genética de la línea germinal de seres humanos.
2. El método según el punto 1, en el que la Cas es una Cas9 en la que el 10º residuo de Asp se convierte en un residuo de Ala o el 840º residuo de His se convierte en un residuo de Ala.
3. El método según el punto 1 ó 2, que usa dos o más clases de módulos de reconocimiento de secuencia de ácido nucleico que se une cada uno específicamente a una secuencia de nucleótidos diana diferente.
4. El método según el punto 3, en el que las secuencias de nucleótidos diana diferentes están presentes en genes diferentes.
5. El método según uno cualquiera de los puntos 1 a 4, en el que la enzima convertidora de bases de ácido nucleico es una desaminasa, preferiblemente una citidina desaminasa.
6. El método según uno cualquiera de los puntos 1 a 5, en el que el ADN bicatenario se pone en contacto con el complejo introduciendo un ácido nucleico que codifica para el complejo en una célula que tiene el ADN bicatenario.
7. El método según el punto 6, en el que la célula es una célula procariota.
8. El método según el punto 6, en el que la célula es una célula microbiana.
9. El método según el punto 6, en el que la célula es una célula eucariota, preferiblemente una célula vegetal, una célula de insecto o una célula animal.
10. El método según el punto 9, en el que la célula animal es una célula de vertebrado, preferiblemente una célula de mamífero.
11. El método según uno cualquiera de los puntos 6 a 10, en el que la célula es una célula poliploide, y se modifican todos los sitios seleccionados como diana en alelos en cromosomas homólogos.
12. El método según uno cualquiera de los puntos 6 a 11, que comprende una etapa de introducir un vector de expresión que comprende un ácido nucleico que codifica para el complejo en una forma que permite el control de un periodo de expresión en la célula, y una etapa de inducir la expresión del ácido nucleico durante un periodo necesario para fijar la modificación del sitio seleccionado como diana en el ADN bicatenario.
13. El método según el punto 12, en el que la secuencia de nucleótidos diana en el ADN bicatenario está presente en un gen esencial para la célula.
14. Un complejo enzimático de modificación de ácido nucleico que comprende un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a una secuencia de nucleótidos diana en un ADN bicatenario dado unido a una enzima convertidora de bases de ácido nucleico y el módulo de reconocimiento de secuencia de ácido nucleico es un sistema CRISPR-Cas en el que el sistema CRISPR-Cas comprende una proteína Cas que tiene actividad nickasa capaz de escindir sólo una de las cadenas del ADN bicatenario, complejo que convierte uno o más nucleótidos en el sitio seleccionado como diana en uno o más de otros nucleótidos o deleciona uno o más nucleótidos, o inserta uno o más nucleótidos en dicho sitio seleccionado como diana.

15. Un ácido nucleico que codifica para el complejo enzimático de modificación de ácido nucleico según el punto 14.

Los presentes inventores han realizado intensos estudios en un intento por resolver los problemas mencionados anteriormente y han tomado nota de la adopción de una conversión de bases mediante una reacción de conversión de base de ADN, sin acompañar al DSB. La reacción de conversión de bases mediante una reacción de desaminación de base de ADN ya se conoce; sin embargo, la selección como diana de cualquier sitio reconociendo una secuencia particular de ADN, y específicamente la modificación del ADN seleccionado como diana mediante la conversión de bases de bases de ADN aún no se ha realizado.

Por tanto, se usó desaminasa, que cataliza una reacción de desaminación, como enzima para tal conversión de bases de ácido nucleico, y se unió a una molécula que tenía una capacidad de reconocimiento de secuencia de ADN, de ese modo se modificó una secuencia genómica mediante conversión de bases de ácido nucleico en una región que contiene una secuencia de ADN particular.

Específicamente, se usó el sistema CRISPR-Cas (CRISPR-Cas mutante). Es decir, se produce un ADN que codifica para una molécula de ARN, en el que CRISPR-ARN:ARNcr (ARNg) específico de genoma que contiene una secuencia complementaria a una secuencia diana de un gen que va a modificarse se une a un ARN para reclutar la proteína Cas (ARNcr de transactivación: ARNtracr). Por otro lado, se produjo un ADN en el que un ADN que codifica para una proteína Cas mutante (dCas), en el que la capacidad de escisión de una o ambas cadenas de un ADN bicatenario se inactivan y un gen de desaminasa se unen. Estos ADN se introdujeron en una célula huésped de levadura que comprende un gen que va a modificarse. Como resultado, pudo introducirse una mutación de manera aleatoria dentro del intervalo de varios cientos de nucleótidos del gen de interés que incluye la secuencia diana. En comparación a cuando se usó proteína Cas doble mutante, que no escinde ambas cadenas de ADN en el ADN bicatenario, la eficiencia de la introducción de la mutación aumentó cuando se usó una proteína Cas mutante que escinde cualquiera de las dos cadenas. Además, se clarificó que el área de la región de mutación y la variedad de mutación varía dependiendo de qué cadena doble del ADN se escinde. Además, pudo introducirse una mutación de manera extremadamente eficaz seleccionando como diana una pluralidad de regiones en el gen de interés. Es decir, se sembró una célula huésped introducida con ADN en un medio no selectivo, y se examinó la secuencia del gen de interés en colonias seleccionadas de manera aleatoria. Como resultado, se confirmó la introducción de mutación en casi todas las colonias. Además, se confirmó que la edición genómica puede realizarse simultáneamente en una pluralidad de sitios seleccionando como diana una determinada región en dos o más genes de interés. Se demostró adicionalmente que el método puede introducir simultáneamente una mutación en alelos de genomas diploides o poliploides, que puede introducir una mutación no sólo en células eucariotas sino también células procariotas tales como *Escherichia coli*, y que es ampliamente aplicable independientemente de la especie. Además, se encontró que la edición de un gen esencial, que mostró baja eficiencia hasta ahora, puede realizarse de manera eficaz realizando de manera transitoria una reacción de conversión de bases de ácido nucleico en una fase deseada.

El presente inventor ha realizado estudios adicionales basándose en estos hallazgos y ha completado la presente invención.

Por consiguiente, la presente divulgación es tal como se describe a continuación.

[1] Un método de modificación de un sitio seleccionado como diana de un ADN bicatenario, que comprende una etapa de poner en contacto un complejo en el que un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a una secuencia de nucleótidos diana en un ADN bicatenario seleccionado y una enzima convertidora de bases de ácido nucleico se unen, con dicho ADN bicatenario, para convertir uno o más nucleótidos en el sitio seleccionado como diana en uno o más de otros nucleótidos o delecionar uno o más nucleótidos, o insertar uno o más nucleótidos en dicho sitio seleccionado como diana, sin escindir al menos una cadena de dicho ADN bicatenario en el sitio seleccionado como diana.

[2] El método según [1], en el que el módulo de reconocimiento de secuencia de ácido nucleico se selecciona del grupo que consiste en un sistema CRISPR-Cas en el que al menos se inactiva una capacidad de escisión de ADN de Cas, un motivo de dedos de cinc, un efector TAL y un motivo PPR.

[3] El método según [1], en el que el módulo de reconocimiento de secuencia de ácido nucleico es un sistema CRISPR-Cas en el que al menos se inactiva una capacidad de escisión de ADN de Cas.

[4] El método según cualquiera de [1] - [3], que usa dos o más clases de módulos de reconocimiento de secuencia de ácido nucleico que se une cada uno específicamente a una secuencia de nucleótidos diana diferente.

[5] El método según [4], en el que la secuencia de nucleótidos diana diferente está presente en un gen diferente.

[6] El método de cualquiera de [1] - [5], en el que la enzima convertidora de bases de ácido nucleico es desaminasa.

[7] El método según [6], en el que la desaminasa es AID (AICDA).

[8] El método según cualquiera de [1] - [7], en el que el ADN bicatenario se pone en contacto con el complejo introduciendo un ácido nucleico que codifica para el complejo en una célula que tiene el ADN bicatenario.

5 [9] El método según [8], en el que la célula es una célula procariota.

[10] El método según [8], en el que la célula es una célula eucariota.

[11] El método según [8], en el que la célula es una célula de un microorganismo.

10

[12] El método según [8], en el que la célula es una célula vegetal.

[13] El método según [8], en el que la célula es una célula de insecto.

15 [14] El método según [8], en el que la célula es una célula animal.

[15] El método según [8], en el que la célula es una célula de un vertebrado.

[16] El método según [8], en el que la célula es una célula de mamífero.

20

[17] El método según cualquiera de [9] - [16], en el que la célula es una célula poliploide, y se modifica un sitio en cualquier alelo seleccionado como diana en un cromosoma homólogo.

25 [18] El método según cualquiera de [8] - [17], que comprende una etapa de introducir un vector de expresión que comprende un ácido nucleico que codifica para el complejo en una forma que permite el control de un periodo de expresión en la célula, y una etapa de inducir la expresión del ácido nucleico durante un periodo necesario para estabilizar la modificación del sitio seleccionado como diana en el ADN bicatenario.

30 [19] El método según [18], en el que la secuencia de nucleótidos diana en el ADN bicatenario está presente en un gen esencial para la célula.

[20] Un complejo enzimático de modificación de ácido nucleico en el que un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a una secuencia de nucleótidos diana en un ADN bicatenario seleccionado y una enzima convertidora de bases de ácido nucleico se unen, lo que convierte uno o más nucleótidos en el sitio seleccionado como diana en uno o más de otros nucleótidos o deleciona uno o más nucleótidos, o inserta uno o más nucleótidos en dicho sitio seleccionado como diana, sin escindir al menos una cadena de dicho ADN bicatenario en el sitio seleccionado como diana.

35

[21] Un ácido nucleico que codifica para el complejo enzimático de modificación de ácido nucleico según [20].

40

[Efecto de la invención]

Según la técnica de edición genómica de la presente invención, debido a que no se asocia con la inserción de un ADN foráneo o roturas de ADN bicatenario, la técnica es superior en cuanto a seguridad. La técnica tiene alguna posibilidad de proporcionar una solución en casos en los que los métodos convencionales se consideraron como una recombinación de genes, y por tanto biológica o legalmente controvertidos. Además, es teóricamente posible establecer una amplia gama de introducción de mutación desde una posición identificada de una base hasta varios cientos de bases, y la técnica también puede aplicarse a la inducción de evolución local mediante la introducción de una mutación aleatoria en una región limitada particular, que ha sido casi imposible hasta ahora.

50

**[Breve descripción de los dibujos]**

La figura 1 es una ilustración esquemática que muestra un mecanismo del método de modificación genética de la presente invención usando el sistema CRISPR-Cas.

55

La figura 2 muestra los resultados de verificación, usando una levadura en gemación, del efecto del método de modificación genética de la presente invención que comprende una combinación de un sistema CRISPR-Cas y desaminasa PmCDA1 de *Petromyza marinus*.

La figura 3 muestra los cambios en el número de células supervivientes después de la inducción de la expresión cuando un sistema CRISPR-Cas9 que usa un mutante de D10A de Cas9 que tiene una actividad nickasa y una desaminasa, PmCDA1, se usan en combinación (nCas9 D10A-PmCDA1), y cuando se usa Cas9 convencional que tiene una capacidad de escisión de la doble cadena del ADN.

60

La figura 4 muestra los resultados cuando se construye una pluralidad de constructos de expresión de manera que las desaminasa AID humana y dCas9 se unen a través de un dominio SH3 y un ligando de unión del mismo, en la

65

que los constructos expresos se introducen en una levadura en gemación junto con dos clases de ARNg (secuencias de selección como diana de diana 4 y diana 5).

5 La figura 5 muestra que la eficiencia de la introducción de mutación se aumenta mediante el uso de Cas9 que escinde cualquier cadena sencilla del ADN.

La figura 6 muestra que en el caso en el que un ADN bicatenario no se escinde, el área de la región de introducción de mutación y la frecuencia de la misma cambian dependiendo de qué cadena sencilla se escinda.

10 La figura 7 muestra que puede conseguirse una eficiencia de la introducción de mutación sumamente alta seleccionado como diana dos regiones en la proximidad.

La figura 8 muestra que el método de modificación genética de la presente invención no requiere la selección mediante un marcador. Se encontró que la mutación se introdujo en todas las colonias secuenciadas.

15 La figura 9 muestra que pueden editarse una pluralidad de sitios en un genoma simultáneamente mediante el método de modificación genética de la presente invención. El panel superior muestra la secuencia de nucleótidos y la secuencia de aminoácidos del sitio diana de cada gen, y una flecha en la secuencia de nucleótidos muestra la secuencia de nucleótidos diana. El número en el extremo de la flecha o en la punta de la flecha indica la posición del extremo terminal de la secuencia de nucleótidos diana en ORF. El panel inferior muestra los resultados de secuenciación del sitio diana en cada uno de los 5 clones de las colonias roja (R) y blanca (B). En las secuencias, los nucleótidos indicados con caracteres subrayados muestran la aparición de conversión de bases. En cuanto a la receptividad a canavanina (CanR), R muestra resistencia, y S muestra sensibilidad.

25 La figura 10 muestra que puede introducirse una mutación simultáneamente en ambos alelos en el cromosoma homólogo del genoma diploide mediante el método de modificación genética de la presente invención. La figura 10A muestra la eficiencia de la introducción de mutación homóloga del gen *ade1* (panel superior) y el gen *can1* respectivamente. La figura 10B muestra que la mutación homóloga se introdujo realmente en la colonia roja (panel inferior). Además, se demostró la aparición de mutación heteróloga en la colonia blanca (panel superior).

30 La figura 11 muestra que es posible la edición genómica de *Escherichia coli*, una célula procariota, mediante el método de modificación genética de la presente invención. La figura 11A es una ilustración esquemática que muestra el plásmido usado. La figura 11B muestra que puede introducirse una mutación (CAA→TAA) de manera eficaz seleccionando como diana una región en el gen *galk*. La figura 11C muestra los resultados del análisis de secuencia de cada uno de los dos clones de las respectivas colonias en un medio no selectivo (ninguno), un medio que contiene rifampicina 25 µg/ml (Rif25) o un medio que contiene rifampicina 50 µg/ml (Rif50). Se confirmó la introducción de una mutación que confiere resistencia a la rifampicina (panel superior). La frecuencia de aparición de la cepa resistente a la rifampicina se estimó que era aproximadamente del 10% (panel inferior).

40 La figura 12 muestra el control de los sitios de bases editados por la longitud del ARN guía. La figura 12A es una figura conceptual del sitio de bases de edición cuando la longitud de la secuencia de nucleótidos diana es de 20 bases o de 24 bases. La figura 12B muestra los resultados de la edición seleccionando como diana el gen *gsiA* y cambiando la longitud de la secuencia de nucleótidos diana. Los sitios mutados se muestran con letras en negrita, "T" y "A" muestran la introducción de una mutación completa (C→T o G→A) en el clon, "t" muestra que se introduce no menos del 50% de mutación (C→T) en el clon (clonación incompleta), y "c" muestra que la eficiencia de la introducción de la mutación (C→T) en el clon es de menos del 50%.

La figura 13 es una ilustración esquemática que muestra un plásmido sensible a la temperatura para la introducción de mutación, que se usó en el ejemplo 11.

50 La figura 14 muestra el protocolo de introducción de mutación en el ejemplo 11.

La figura 15 muestra los resultados de introducción de mutación en el gen *rpoB* en el ejemplo 11.

55 La figura 16 muestra los resultados de introducción de mutación en el gen *galk* en el ejemplo 11.

### [Descripción de las realizaciones]

60 La presente descripción proporciona un método de modificación de un sitio seleccionado como diana de un ADN bicatenario convirtiendo la secuencia de nucleótidos diana y nucleótidos en la proximidad del mismo en el ADN bicatenario en otros nucleótidos, sin escindir al menos una cadena del ADN bicatenario que va a modificarse. El método comprende de manera característica una etapa de poner en contacto un complejo en el que se unen un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a la secuencia de nucleótidos diana en el ADN bicatenario y una enzima convertidora de bases de ácido nucleico, con el ADN bicatenario para convertir el sitio seleccionado como diana, es decir, la secuencia de nucleótidos diana y nucleótidos en la proximidad del mismo, en otros nucleótidos.

En la presente invención, la “modificación” de un ADN bicatenario significa que un nucleótido (por ejemplo, dC) en una cadena de ADN se convierte en otro nucleótido (por ejemplo, dT, dA o dG), o se deleta, o se inserta un nucleótido o una secuencia de nucleótidos entre determinados nucleótidos en la cadena de ADN. Mientras que el ADN bicatenario que va a modificarse no está particularmente limitado, es preferiblemente un ADN genómico. El “sitio seleccionado como diana” de un ADN bicatenario significa que la “secuencia de nucleótidos diana” completa o parcial, en la que un módulo de reconocimiento de secuencia de ácido nucleico reconoce y se une específicamente a, o la proximidad de la secuencia de nucleótidos diana (una o ambas en el sentido de 5' y en el sentido de 3'), y la longitud de la misma puede ajustarse de manera apropiada entre 1 base y varios cientos de bases según el objeto.

En la presente invención, el “módulo de reconocimiento de secuencia de ácido nucleico” significa una molécula o complejo de moléculas que tiene una capacidad para reconocer y unirse específicamente a una secuencia de nucleótidos particular (es decir, secuencia de nucleótidos diana) en una cadena de ADN. La unión del módulo de reconocimiento de secuencia de ácido nucleico a una secuencia de nucleótidos diana permite que una enzima convertidora de bases de ácido nucleico unida al módulo actúe específicamente en un sitio seleccionado como diana de un ADN bicatenario.

En la presente invención, la “enzima convertidora de bases de ácido nucleico” significa una enzima capaz de convertir un nucleótido diana en otro nucleótido catalizando una reacción para convertir un sustituyente en un anillo de purina o pirimidina en una base de ADN en otro grupo o átomo, sin escindir la cadena de ADN.

En la presente invención, el “complejo enzimático de modificación de ácido nucleico” significa un complejo molecular que comprende un complejo del módulo de reconocimiento de secuencia de ácido nucleico mencionado anteriormente unido con una enzima convertidora de bases de ácido nucleico, en el que el complejo tiene actividad enzimática convertidora de bases de ácido nucleico y se confiere con una capacidad de reconocimiento de secuencia de nucleótidos particular. El “complejo” usado en el presente documento abarca no sólo uno compuesto de una pluralidad de moléculas, sino también una molécula sencilla que tiene un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico tal como una proteína de fusión.

La enzima convertidora de bases de ácido nucleico usada en la presente invención no está particularmente limitada siempre que pueda catalizar la reacción mencionada anteriormente, y los ejemplos de la misma incluyen desaminasa que pertenece a la superfamilia de desaminasas de ácido nucleico/nucleótido, que cataliza una reacción de desaminación que convierte un grupo amino en un grupo carbonilo. Los ejemplos preferibles de la misma incluyen citidina desaminasa capaz de convertir citosina o 5-metilcitosina en uracilo o timina, respectivamente, adenosina desaminasa capaz de convertir adenina en hipoxantina, guanosina desaminasa capaz de convertir guanina en xantina y similares. Como citidina desaminasa, se prefiere más citidina desaminasa inducida por activación (a continuación en el presente documento también denominada AID), que es una enzima que introduce una mutación en un gen de la inmunoglobulina en la inmunidad adquirida de un vertebrado o similar.

Mientras que el origen de la enzima convertidora de bases de ácido nucleico no está particularmente limitado, por ejemplo, puede usarse PmCDA1 (citocina desaminasa 1 de *Petromyzon marinus*) de *Petromyzon marinus*, o AID (citidina desaminasa inducida por activación; AICDA) de mamífero (por ejemplo, humano, porcino, bovino, de caballo, de mono, etc.). La secuencia de bases y secuencia de aminoácidos de CDS de PmCDA1 se muestran en las SEQ ID NO: 1 y 2, respectivamente, y la secuencia de bases y secuencia de aminoácidos de CDS de AID humana se muestran en las SEQ ID NO: 3 y 4, respectivamente.

Una secuencia de nucleótidos diana en un ADN bicatenario que va a ser reconocida por el módulo de reconocimiento de secuencia de ácido nucleico en el complejo enzimático de modificación de ácido nucleico de la presente invención no está particularmente limitada siempre que el módulo se una específicamente a cualquier secuencia en el ADN bicatenario. La longitud de la secuencia de nucleótidos diana sólo es necesario que sea suficiente para la unión específica del módulo de reconocimiento de secuencia de ácido nucleico. Por ejemplo, cuando se introduce mutación en un sitio particular en el ADN genómico de un mamífero, no es menos de 12 nucleótidos, preferiblemente no menos de 15 nucleótidos, más preferiblemente no menos de 17 nucleótidos, según el tamaño del genoma del mismo. Mientras que el límite superior de la longitud no está particularmente limitado, es preferiblemente no más de 25 nucleótidos, más preferiblemente no más de 22 nucleótidos.

Como el módulo de reconocimiento de secuencia de ácido nucleico en el complejo enzimático de modificación de ácido nucleico de la presente descripción, puede usarse un sistema CRISPR-Cas en el que al menos se inactiva una capacidad de escisión de ADN de Cas (CRISPR-Cas mutante), motivo de dedos de cinc, efector TAL y motivo PPR y similares, así como un fragmento que contiene un dominio de unión al ADN de una proteína que se une específicamente al ADN tal como una enzima de restricción, un factor de transcripción, una ARN polimerasa o similares, y que no tiene una capacidad de escisión de la doble cadena del ADN y similares, pero el módulo no se limita al mismo. Preferiblemente, los módulos incluyen CRISPR-Cas mutante, motivo de dedos de cinc, efector TAL, motivo PPR y similares.

Un motivo de dedos de cinc se construye uniendo 3 - 6 unidades de dedos de cinc de tipo Cys2His2 diferentes (1

dedo reconoce aproximadamente 3 bases), y puede reconocer una secuencia de nucleótidos diana de 9 - 18 bases. Un motivo de dedos de cinc puede producirse mediante un método conocido tal como un método de ensamblaje modular (Nat Biotechnol (2002) 20: 135-141), método OPEN (Mol Cell (2008) 31: 294-301), método CoDA (Nat Methods (2011) 8: 67-69), método de un híbrido de *Escherichia coli* (Nat Biotechnol (2008) 26:695-701) y similares. Puede referirse al documento de patente 1 mencionado anteriormente para el detalle de la producción de motivos de dedos de cinc.

Un efector TAL tiene una estructura de repetición de módulos con aproximadamente 34 aminoácidos como una unidad, y los residuos de aminoácido 12° y 13° (denominados RVD) de un módulo determinan la estabilidad de la unión y la especificidad de bases. Debido a que cada módulo es altamente independiente, el efector TAL específico para una secuencia de nucleótidos diana puede producirse simplemente uniendo los módulos. Para el efector TAL, se han establecido métodos de producción que utilizan un recurso abierto (método REAL (Curr Protoc Mol Biol (2012) capítulo 12: unidad 12.15), método FLASH (Nat Biotechnol (2012) 30: 460-465) y método Golden Gate (Nucleic Acids Res (2011) 39: e82), etc.), y puede diseñarse de manera relativamente fácil un efector TAL para una secuencia de nucleótidos diana. Puede referirse al documento de patente 2 mencionado anteriormente para el detalle de la producción de un efector TAL.

Un motivo PPR se construye de manera que una secuencia de nucleótidos particular se reconoce por una serie de motivos PPR que consiste cada uno en 35 aminoácidos y que reconoce una base de ácido nucleico, y reconoce una base diana sólo por los aminoácidos 1, 4 y ii(-2) de cada motivo. La configuración del motivo no tiene dependencia, y está libre de interferencia de motivos en ambos lados. Por tanto, similar al efector TAL, puede producirse una proteína PPR específica para la secuencia de nucleótidos diana simplemente uniendo motivos PPR. Puede referirse al documento de patente 4 mencionado anteriormente para el detalle de la producción de un motivo PPR.

Cuando se usa un fragmento de enzima de restricción, un factor de transcripción, una ARN polimerasa o similares, debido a que los dominios de unión al ADN de estas proteínas se conocen bien, puede diseñarse y construirse fácilmente un fragmento que contiene dicho dominio y que no tiene una capacidad de escisión de la doble cadena del ADN.

Cualquier módulo de reconocimiento de secuencia de ácido nucleico mencionado anteriormente puede proporcionarse como una proteína de fusión con la enzima convertidora de bases de ácido nucleico mencionada anteriormente, o un dominio de unión a proteína tal como un dominio SH3, dominio PDZ, dominio GK, dominio GB y similares y un componente de unión del mismo pueden fusionarse con un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico, respectivamente, y proporcionarse como un complejo de proteínas a través de una interacción del dominio y un componente de unión del mismo. Alternativamente, un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico pueden fusionarse cada uno con inteína, y pueden unirse mediante ligación después de la síntesis de proteínas.

El complejo enzimático de modificación de ácido nucleico de la presente invención que contiene un complejo (incluyendo una proteína de fusión), en el que se unen un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico, puede ponerse en contacto con un ADN bicatenario como una reacción enzimática en un sistema libre de células. En vista del objeto principal de la presente invención, es deseable realizar el contacto introduciendo un ácido nucleico que codifica para el complejo en una célula que tiene el ADN bicatenario de interés (por ejemplo, ADN genómico).

Por tanto, el módulo de reconocimiento de secuencia de ácido nucleico y la enzima convertidora de bases de ácido nucleico se preparan preferiblemente como un ácido nucleico que codifica para una proteína de fusión del mismo, o como ácidos nucleicos que codifican para cada uno de ellos en una forma capaz de formar un complejo en una célula huésped después de la traducción para dar una proteína utilizando un dominio de unión, inteína o similar. El ácido nucleico puede ser en el presente documento un ADN o un ARN. Cuando es un ADN, es preferiblemente un ADN bicatenario, y se proporciona en forma de un vector de expresión dispuesto bajo la regulación de un promotor funcional en una célula huésped. Cuando es un ARN, es preferiblemente un ARN monocatenario.

Debido a que el complejo de la presente invención en el que se unen un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico, no está asociado con roturas de ADN bicatenario (DSB), es posible la edición genómica con baja toxicidad, y el método de modificación genética de la presente invención puede aplicarse a una amplia gama de materiales biológicos. Por tanto, las células en las que se introduce el ácido nucleico que codifica para el módulo de reconocimiento de secuencia de ácido nucleico y/o la enzima convertidora de bases de ácido nucleico puede abarcar células de cualquier especie, desde células de microorganismos, tales como bacterias, tal como *Escherichia coli* y similares que son procariontes, tales como levadura y similares que son eucariotas inferiores, hasta células de eucariotas superiores tales como insecto, planta y similares, y células de vertebrado que incluyen mamíferos tales como humano y similares.

Un ADN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico tal como un motivo de dedos de cinc, un efector TAL, un motivo PPR y similares puede obtenerse mediante cualquier método mencionado

anteriormente para cada módulo. Un ADN que codifica para un módulo de reconocimiento de secuencia de enzima de restricción, un factor de transcripción, una ARN polimerasa y similares pueden clonarse, por ejemplo, sintetizando un cebador de oligoADN que cubre una región que codifica para una parte deseada de la proteína (una parte que contiene un dominio de unión al ADN) basándose en la información de secuencia de ADNc del mismo, y amplificando mediante el método de RT-PCR usando, el ARN total o la fracción de ARNm preparada a partir de las células productoras de proteínas como molde.

Un ADN que codifica para una enzima convertidora de bases de ácido nucleico también puede clonarse de manera similar sintetizando un cebador de oligoADN basándose en la información de secuencia de ADNc del mismo, y amplificando mediante el método de RT-PCR usando, el ARN total o la fracción de ARNm preparada a partir de las células productoras de enzimas como molde. Por ejemplo puede clonarse, un ADN que codifica para PmCDA1 de *Petromyzon marinus* diseñando cebadores adecuados para el sentido de 5' y el sentido de 3' de CDS basándose en la secuencia de ADNc (n.º de registro EF094822) registrada en la base de datos del NCBI, y clonando a partir de ARNm de *Petromyzon marinus* mediante el método de RT-PCR. Un ADN que codifica para la AID humana puede clonarse diseñando cebadores adecuados para el sentido de 5' y el sentido de 3' de CDS basándose en la secuencia de ADNc (n.º de registro AB040431) registrada en la base de datos del NCBI, y clonando a partir de, por ejemplo, ARNm de ganglio linfático humano mediante el método de RT-PCR.

El ADN clonado puede unirse directamente, o después de la digestión con una enzima de restricción si se desea, o después de la adición de un ligador y/o una señal de localización nuclear adecuados (cada organelo transfiere señal cuando el ADN bicatenario de interés es ADN de mitocondria o cloroplasto), con un ADN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico para preparar un ADN que codifica para una proteína de fusión. Alternativamente, un ADN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico, y un ADN que codifica para una enzima convertidora de bases de ácido nucleico pueden fusionarse cada uno con un ADN que codifica para un dominio de unión o un componente de unión del mismo, o ambos ADN pueden fusionarse con un ADN que codifica para una inteína de separación, mediante lo cual el módulo de conversión de reconocimiento de secuencia de ácido nucleico y la enzima convertidora de bases de ácido nucleico se traducen en una célula huésped para formar un complejo. En estos casos, un ligador y/o una señal de localización nuclear pueden unirse en una posición adecuada de uno o ambos ADN si se desea.

Un ADN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico y un ADN que codifica para una enzima convertidora de bases de ácido nucleico pueden obtenerse sintetizando químicamente la cadena de ADN, o uniendo cadenas cortas de oligoADN sintetizadas parcialmente solapadas utilizando el método de PCR y el método de ensamblaje de Gibson para construir un ADN que codifica para la longitud completa del mismo. La ventaja de la construcción de un ADN de longitud completa mediante síntesis química o una combinación del método de PCR o el método de ensamblaje de Gibson es que puede diseñarse el codón usado en CDS de longitud completa según el huésped en el que se introduce el ADN. En la expresión de un ADN heterólogo, se espera que el nivel de expresión de proteínas aumente convirtiendo la secuencia de ADN del mismo en un codón que se usa de manera muy frecuente en el organismo huésped. Como datos de frecuencia de uso del codón en el huésped usado, por ejemplo, puede usarse la base de datos de frecuencia de uso de códigos genéticos (<http://www.kazusa.o.jp/codon/index.html>) divulgada en página web principal de Kazusa ADN Research Institute, o pueden referirse a documentos que muestran la frecuencia de uso del codón en cada huésped. Con referencia a los datos obtenidos y la secuencia de ADN que va a introducirse, los codones que muestran baja frecuencia de uso en el huésped a partir de los usados para la secuencia de ADN pueden convertirse en un codón que codifica para el mismo aminoácido y que muestra alta frecuencia de uso.

Un vector de expresión que contiene un ADN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico y/o una enzima convertidora de bases de ácido nucleico puede producirse, por ejemplo, uniendo el ADN en el sentido de 3' de un promotor en un vector de expresión adecuado.

Como vector de expresión, se usan plásmidos de *Escherichia coli* (por ejemplo, pBR322, pBR325, pUC12, pUC13); plásmidos de *Bacillus subtilis* (por ejemplo, pUB110, pTP5, pC194); plásmidos de levadura (por ejemplo, pSH19, pSH15); plásmidos de expresión de células de insecto (por ejemplo, pFast-Bac); plásmidos de expresión de células animales (por ejemplo, pA1-11, pXT1, pRc/CMV, pRc/RSV, pcDNA1/Neo); bacteriófagos tales como fago  $\lambda$  y similares; vectores de virus de insecto tales como baculovirus y similares (por ejemplo, BmNPV, AcNPV); vectores de virus animales tales como retrovirus, virus vaccinia, adenovirus y similares.

Como promotor, puede usarse cualquier promotor apropiado para un huésped usado para expresión génica. En un método convencional que implica DSB, debido a que la tasa de supervivencia de la célula huésped algunas veces disminuye de manera pronunciada debido a la toxicidad, es deseable aumentar el número de células al inicio de la inducción usando un promotor inductivo. Sin embargo, ya que también puede alcanzarse una proliferación celular suficiente expresando el complejo enzimático de modificación de ácido nucleico de la presente invención, también puede usarse un promotor constitutivo sin limitación.

Por ejemplo, cuando el huésped es una célula animal, promotor de SR $\alpha$ , promotor de VS40, promotor de LTR, promotor de CMV (citomegalovirus), promotor de VSR (virus del sarcoma de Rous), LTR de MoMuLV (virus de la

## ES 2 752 175 T3

leucemina murina de Moloney), promotor de VHS-TK (timidina cinasa del virus del herpes simple) y similares. De éstos, son preferibles el promotor de CMV, el promotor de SR $\alpha$  y similares.

5 Cuando el huésped es *Escherichia coli*, son preferibles el promotor de trp, el promotor de lac, el promotor de recA, el promotor de  $\lambda$ PL, el promotor de lpp, el promotor de T7 y similares.

Cuando el huésped es del género *Bacillus*, son preferibles el promotor de SPO1, el promotor de SPO2, el promotor de penP y similares.

10 Cuando el huésped es una levadura, son preferibles el promotor de Gal1/10, el promotor de PHO5, el promotor de PGK, el promotor de GAP, el promotor de ADH y similares.

Cuando el huésped es una célula de insecto, son preferibles el promotor de polihedrina, el promotor de P10 y similares.

15 Cuando el huésped es una célula vegetal, son preferibles el promotor de CaMV35S, el promotor de CaMV19S, el promotor de NOS y similares.

20 Como vector de expresión, junto con los mencionados anteriormente, puede usarse uno que contenga un potenciador, una señal de corte y empalme, un terminador, una señal de adición de poliA, un marcador de selección tal como un gen resistente al fármaco, un gen complementario auxotrófico y similares, un origen de replicación y similares según demanda.

25 Un ARN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico y/o una enzima convertidora de bases de ácido nucleico puede prepararse, por ejemplo, mediante transcripción a ARNm en un sistema de transcripción *in vitro* conocido *per se* usando un vector que codifica para ADN que codifica para el módulo de reconocimiento de secuencia de ácido nucleico y/o una enzima convertidora de bases de ácido nucleico mencionados anteriormente como molde.

30 Un complejo de un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico puede expresarse de manera intracelular introduciendo un vector de expresión que contiene un ADN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico y/o una enzima convertidora de bases de ácido nucleico en una célula huésped, y cultivando la célula huésped.

35 Como huésped, se usan género *Escherichia*, género *Bacillus*, levadura, célula de insecto, insecto, célula animal y similares.

40 Como género *Escherichia*, se usan *Escherichia coli* K12-DH1 [Proc. Natl. Acad. Sci. USA, 60, 160 (1968)], *Escherichia coli* JM103 [Nucleic Acids Research, 9, 309 (1981)], *Escherichia coli* JA221 [Journal of Molecular Biology, 120, 517 (1978)], *Escherichia coli* HB101 [Journal of Molecular Biology, 41, 459 (1969)], *Escherichia coli* C600 [Genetics, 39, 440 (1954)] y similares.

45 Como género *Bacillus*, se usan *Bacillus subtilis* M114 [Gene, 24, 255 (1983)], *Bacillus subtilis* 207-21 [Journal of Biochemistry, 95, 87 (1984)] y similares.

Como levadura, se usan *Saccharomyces cerevisiae* AH22, AH22R, NA87-11A, DKD-5D, 20B-12, *Schizosaccharomyces pombe* NCYC1913, NCYC2036, *Pichia pastoris* KM71 y similares.

50 Como célula de insecto cuando el virus es AcNPV, se usan células de una línea establecida a partir de larva del noctuido de la col (célula de *Spodoptera frugiperda*; célula de Sf), células de MG1 del intestino medio de *Trichoplusia ni*, células High Five™ de un huevo de *Trichoplusia ni*, células de *Mamestra brassicae*, células de *Estigmene acrea* y similares. Cuando el virus es BmNPV, se usan células de línea establecida de *Bombyx mori* (célula N de *Bombyx mori*; célula de BmN) y similares como células de insecto. Como célula de Sf, por ejemplo, célula de Sf9 (ATCC CRL1711), célula de Sf21 [todas las anteriores, *In vivo*, 13, 213-217 (1977)] y similares.

55 Como insecto, por ejemplo, se usan larva de *Bombyx mori*, *Drosophila*, grillo y similares [Nature, 315, 592 (1985)].

60 Como célula animal, se usan líneas celulares tales como célula de mono COS-7, célula de mono Vero, célula de ovario de hámster chino (CHO), célula de CHO deficiente en gen dhfr, célula L de ratón, célula de ratón AtT-20, célula de mieloma de ratón, célula de rata GH3, célula humana FL y similares, células madre pluripotentes tales como célula iPS, célula ES y similares de humano y otros mamíferos, y células cultivadas primarias preparadas a partir de diversos tejidos. Además, también pueden usarse embrión de pez cebra, *Xenopus oocyte* y similares.

65 Como célula vegetal, se usan células cultivadas suspendidas, callo, protoplasto, segmento de hojas, segmento de raíces y similares preparados a partir de diversas plantas (por ejemplo, cereales tales como arroz, trigo, maíz y similares, cultivos de productos tales como tomate, pepino, berenjena y similares, plantas de jardín tales como

clavel, *Eustoma russellianum* y similares, plantas de ensayo tales como tabaco, *arabidopsis thaliana* y similares).

Todas las células huésped mencionadas anteriormente pueden ser haploides (monoploides), o poliploides (por ejemplo, diploides, triploides, tetraploides y similares). En los métodos de introducción de mutación convencional, la mutación se introduce, en principio, en un sólo cromosoma homólogo para producir un genotipo heterólogo. Por tanto, la característica deseada no se expresa a menos que sea una mutación dominante, y hacerla homóloga requiere de manera inoportuna trabajo y tiempo. Por el contrario, según la presente invención, debido a que pueden introducirse mutaciones en todos los alelos en el cromosoma homólogo en el genoma, la característica deseada puede expresarse en una única generación incluso en el caso de mutación recesiva (figura 10), que es extremadamente útil ya que puede resolverse el problema del método convencional.

Un vector de expresión puede introducirse mediante un método conocido (por ejemplo, método de lizozima, método competente, método PEG, método de coprecipitación de CaCl<sub>2</sub>, método de electroporación, el método de microinyección, método de pistola de partículas, método de lipofección, método de *Agrobacterium* y similares) según la clase del huésped.

*Escherichia coli* puede transformarse según los métodos descritos en, por ejemplo, Proc. Natl. Acad. Sci. USA, 69, 2110 (1972), Gene, 17, 107 (1982) y similares.

Un vector puede introducirse en el género *Bacillus* según los métodos descritos en, por ejemplo, Molecular & General Genetics, 168, 111 (1979) y similares.

Un vector puede introducirse en una levadura según los métodos descritos en, por ejemplo, Methods in Enzymology, 194, 182-187 (1991), Proc. Natl. Acad. Sci. USA, 75, 1929 (1978) y similares.

Un vector puede introducirse en una célula de insecto y en un insecto según los métodos descritos en, por ejemplo, Bio/Technology, 6, 47-55 (1988) y similares.

Un vector puede introducirse en una célula animal según los métodos descritos en, por ejemplo, Cell Engineering volumen adicional 8, New Cell Engineering Experiment Protocol, 263-267 (1995) (publicado por Shujunsha), y Virology, 52, 456 (1973).

Una célula introducida con un vector puede cultivarse según un método conocido según la clase del huésped.

Por ejemplo, cuando se cultiva *Escherichia coli* o género *Bacillus*, es preferible un medio líquido como medio usado para el cultivo. El medio contiene preferiblemente una fuente de carbono, fuente de nitrógeno, sustancia inorgánica y similares necesarios para el crecimiento del transformante. Los ejemplos de la fuente de carbono incluyen glucosa, dextrina, almidón soluble, sacarosa y similares; los ejemplos de la fuente de nitrógeno incluyen sustancias inorgánicas u orgánicas tales como sales de amonio, sales de nitrato, licor de maíz macerado, peptona, caseína, extracto de carne, torta de soja, extracto de patata y similares; y los ejemplos de la sustancia inorgánica incluyen cloruro de calcio, dihidrogenofosfato de sodio, cloruro de magnesio y similares. El medio puede contener extracto de levadura, vitaminas, factor promotor del crecimiento y similares. El pH del medio es preferiblemente de aproximadamente 5 - aproximadamente 8.

Como medio para cultivar *Escherichia coli*, por ejemplo, es preferible un medio M9 que contiene glucosa, casaminoácido [Journal of Experiments in Molecular Genetics, 431-433, Cold Spring Harbor Laboratory, Nueva York 1972]. En caso necesario, por ejemplo, pueden añadirse agentes tales como ácido 3β-indolilacrílico al medio para asegurar una función eficiente de un promotor. *Escherichia coli* se cultiva generalmente a aproximadamente 15 - aproximadamente 43°C. En caso necesario, puede realizarse aireación y agitación.

El género *Bacillus* se cultiva generalmente a aproximadamente 30 - aproximadamente 40°C. En caso necesario, puede realizarse aireación y agitación.

Los ejemplos del medio para cultivar levadura incluyen medio mínimo de Burkholder [Proc. Natl. Acad. Sci. USA, 77, 4505 (1980)], medio SD que contiene casaminoácido al 0,5% [Proc. Natl. Acad. Sci. USA, 81, 5330 (1984)] y similares. El pH del medio es preferiblemente de aproximadamente 5 - aproximadamente 8. El cultivo se realiza generalmente a aproximadamente 20°C - aproximadamente 35°C. En caso necesario, puede realizarse aireación y agitación.

Como medio para cultivar una célula de insecto o un insecto, por ejemplo, se usan medio para insectos de Grace [Nature, 195, 788 (1962)] que contiene un aditivo tal como suero bovino al 10% inactivado y similares como sea apropiado y similares. El pH del medio es preferiblemente de aproximadamente 6,2 - aproximadamente 6,4. El cultivo se realiza generalmente a aproximadamente 27°C. En caso necesario, puede realizarse aireación y agitación.

Como medio para cultivar una célula animal se usan, por ejemplo, medio esencial mínimo (MEM) que contiene aproximadamente el 5 - aproximadamente el 20% de suero bovino fetal [Science, 122, 501 (1952)], medio de Eagle

modificado por Dulbecco (DMEM) [Virology, 8, 396 (1959)], medio RPMI 1640 [The Journal of the American Medical Association, 199, 519 (1967)], medio 199 [Proceeding of the Society for the Biological Medicine, 73, 1 (1950)] y similares. El pH del medio es preferiblemente de aproximadamente 6 - aproximadamente 8. El cultivo se realiza generalmente a aproximadamente 30°C - aproximadamente 40°C. En caso necesario, puede realizarse aireación y agitación.

Como medio para cultivar una célula vegetal se usan, por ejemplo, medio MS, medio LS, medio B5 y similares. El pH del medio es preferiblemente de aproximadamente 5 - aproximadamente 8. El cultivo se realiza generalmente a aproximadamente 20°C - aproximadamente 30°C. En caso necesario, puede realizarse aireación y agitación.

Tal como se mencionó anteriormente, puede expresarse de manera intracelular un complejo de un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico, es decir, un complejo enzimático de modificación de ácido nucleico.

Un ARN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico y/o una enzima convertidora de bases de ácido nucleico puede introducirse en una célula huésped mediante un método de microinyección, método de lipofección y similares. La introducción de ARN puede realizarse una o múltiples veces (por ejemplo, 2 - 5 veces) en intervalos adecuados.

Cuando un complejo de un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico se expresan mediante un vector de expresión o una molécula de ARN introducidos en la célula, el módulo de reconocimiento de secuencia de ácido nucleico reconoce y se une específicamente a una secuencia de nucleótidos diana en el ADN bicatenario (por ejemplo, ADN genómico) de interés y, debido a la acción de la enzima convertidora de bases de ácido nucleico unida al módulo de reconocimiento de secuencia de ácido nucleico, se produce la conversión de bases en la cadena sentido o cadena antisentido del sitio seleccionado como diana (secuencia de nucleótidos diana completa o parcial o ajustada de manera apropiada con varios cientos de bases incluyendo la proximidad de la misma) y se produce un apareamiento erróneo en el ADN bicatenario (por ejemplo, cuando se usa citidina desaminasa tal como PmCDA1, AID y similares como enzima convertidora de bases de ácido nucleico, la citosina en la cadena sentido o cadena antisentido en el sitio seleccionado como diana se convierte en uracilo para producir apareamiento erróneo U:G o G:U). Cuando el apareamiento erróneo no se repara de manera correcta, y cuando se repara de manera que una base de la cadena opuesta forma un par con una base de la cadena convertida (T-A o A-T en el ejemplo mencionado anteriormente), o cuando otro nucleótido se sustituye adicionalmente (por ejemplo, U→A, G) o cuando se delecionan o se insertan de una a varias docenas de bases durante la reparación, se insertan diversas mutaciones.

En cuanto al motivo de dedos de cinc, la producción de muchos motivos de dedos de cinc realmente funcionales no es fácil, debido a que la eficiencia de producción de un dedo de cinc que se une específicamente a una secuencia de nucleótidos diana no es alta y la selección de un dedo de cinc que tenga alta especificidad de unión no es fácil. Mientras que el efector TAL y el motivo PPR tienen un alto grado de libertad de reconocimiento de secuencias diana de ácido nucleico en comparación con el motivo de dedos de cinc, permanece un problema en la eficiencia debido a que es necesario diseñar y construir una proteína grande cada vez según la secuencia de nucleótidos diana.

Por el contrario, ya que el sistema CRISPR-Cas reconoce la secuencia de ADN bicatenario de interés mediante un ARN guía complementario a la secuencia de nucleótidos diana, cualquier secuencia puede seleccionarse como diana sintetizando simplemente un oligoADN capaz de formar específicamente un híbrido con la secuencia de nucleótidos diana.

Por tanto, en una realización más preferible de la presente invención, un sistema CRISPR-Cas en el que al menos se inactiva una capacidad de escisión de ADN de Cas (CRISPR-Cas mutante) se usa como módulo de reconocimiento de secuencia de ácido nucleico.

La figura 1 es una ilustración esquemática que muestra el método de modificación del ADN bicatenario de la presente invención usando CRISPR-Cas mutante como módulo de reconocimiento de secuencia de ácido nucleico.

El módulo de reconocimiento de secuencia de ácido nucleico de la presente invención usando CRISPR-Cas mutante se proporciona como un complejo de una molécula de ARN que consiste en un ARN guía complementario a la secuencia de nucleótidos diana y ARNtracr necesarios para reclutar una proteína Cas mutante, y una proteína Cas mutante.

La proteína Cas usada en la presente invención no está particularmente limitada siempre que pertenezca al sistema CRISPR, y preferiblemente Cas9. Los ejemplos de Cas9 incluyen, pero no se limitan a, Cas9 (SpCas9) de *Streptococcus pyogenes*, Cas9 (StCas9) de *Streptococcus thermophilus* y similares, preferiblemente SpCas9. Como Cas mutante usada en la presente descripción, puede usarse o bien una Cas que tiene capacidad de escisión de ambas cadenas del ADN bicatenario, o bien una Cas que tiene actividad nickasa en la que sólo se inactiva una de la capacidad de escisión de sólo una de las cadenas. Por ejemplo, en el caso de SpCas9, puede usarse un mutante de D10A en el que el 10º residuo de Asp se convierte en un residuo de Ala y que carece de capacidad de escisión de

una cadena opuesta a la cadena que forma una cadena complementaria con un ARN guía, o un mutante de H840A en el que el 840° residuo de His se convierte en un residuo de Ala y que carece de capacidad de escisión de cadena complementaria al ARN guía, o un mutante doble del mismo, y puede usarse de manera similar otra Cas mutante.

5 Una enzima convertidora de bases de ácido nucleico se proporciona como un complejo con Cas mutante mediante un método similar al esquema de unión con el dedo de cinc mencionado anteriormente y similares. Alternativamente, una enzima convertidora de bases de ácido nucleico y una Cas mutante también pueden unirse utilizando un  
 10 armazón proteínico de ARN con aptómeros de ARN MS2F6, PP7 y similares y proteínas de unión de los mismos. El ARN guía forma una cadena complementaria con la secuencia de nucleótidos diana, se recluta la Cas mutante por el ARNtracr unido y la Cas mutante reconoce la secuencia de reconocimiento del sitio de escisión del ADN PAM (motivo adyacente de protospaciador) (cuando se usa SpCas9, PAM es 3 bases de NGG (N es cualquier base), y, teóricamente, puede seleccionar como diana cualquier posición en el genoma). Uno o ambos ADN no pueden  
 15 escindirse, y, debido a la acción de la enzima convertidora de bases de ácido nucleico unida a la Cas mutante, se produce la conversión de bases en el sitio seleccionado como diana (ajustada de manera apropiada con varios cientos de bases incluyendo la secuencia de nucleótidos diana completa o parcial) y se produce un apareamiento erróneo en el ADN bicatenario. Cuando el apareamiento erróneo no se repara de manera correcta, y cuando se repara de manera que una base de la cadena opuesta forma un par con una base de la cadena convertida, o cuando otro nucleótido se convierte adicionalmente o cuando se delecionan o insertan de una a varias docenas de bases durante la reparación, se introducen diversas mutaciones (véase, por ejemplo, la figura 2).

20 Incluso cuando se usa CRISPR-Cas mutante como módulo de reconocimiento de secuencia de ácido nucleico, se introducen de manera deseable un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico, en forma de un ácido nucleico que codifica para los mismos, en una célula que tiene un ADN bicatenario de interés, similar a cuando se usan dedo de cinc y similares como módulo de  
 25 reconocimiento de secuencia de ácido nucleico.

Un ADN que codifica para Cas puede clonarse mediante un método similar al método mencionado anteriormente para un ADN que codifica para una enzima convertidora de bases de ácido nucleico, a partir de una célula que produce la enzima. Una Cas mutante puede obtenerse introduciendo una mutación para convertir un residuo de  
 30 aminoácido de la parte importante para la actividad de escisión de ADN (por ejemplo, el 10° residuo de Asp y el 840° residuo de His para Cas9, aunque no se limitan a los mismos) en otro aminoácido, en un ADN que codifica para la Cas clonada, mediante un método de inducción de la mutación específica del sitio conocido *per se*.

Alternativamente, también puede construirse un ADN que codifica para Cas mutante como un ADN que tiene un uso  
 35 de codón adecuado para la expresión en una célula huésped que va a usarse, mediante un método similar a los mencionados anteriormente para un ADN que codifica para un módulo de reconocimiento de secuencia de ácido nucleico y un ADN que codifica para una enzima convertidora de bases de ácido nucleico, y en una combinación con síntesis química o método de PCR o método de ensamblaje de Gibson. Por ejemplo, la secuencia de CDS y la secuencia de aminoácidos optimizada para la expresión de SpCas9 en células eucariotas se muestran en las SEQ  
 40 ID NO: 5 y 6. En la secuencia mostrada en la SEQ ID NO: 5, cuando "A" se convierte en "C" en la base n.º 29, puede obtenerse un ADN que codifica para un mutante de D10A, y cuando "CA" se convierte en "GC" en las bases n.ºs 2518-2519, puede obtenerse un ADN que codifica para un mutante de H840A.

Un ADN que codifica para una Cas mutante y un ADN que codifica para una enzima convertidora de bases de ácido  
 45 nucleico pueden unirse para permitir la expresión como una proteína de fusión, o diseñarse para expresarse por separado usando un dominio de unión, inteína o similar, y formar un complejo en una célula huésped a través de interacción proteína-proteína o ligación de proteínas.

El ADN obtenido que codifica para una Cas mutante y/o una enzima convertidora de bases de ácido nucleico  
 50 pueden insertarse en el sentido de 3' de un promotor de un vector de expresión similar al mencionado anteriormente, según el huésped.

Por otro lado, un ADN que codifica para ARN guía y ARNtracr puede obtenerse diseñando una secuencia de  
 55 oligoADN que se une a una secuencia de ARN guía complementaria a la secuencia de nucleótidos diana y la secuencia de ARNtracr conocida (por ejemplo, gttttagagctagaatagcaagttaaaataaggctagtcggtatcaactgaaaaagtgccaccgagtcggtggtgctttt; SEQ ID NO: 7) y sintetizando químicamente usando un sintetizador de ADN/ARN.

Aunque la longitud de la secuencia de ARN guía no está particularmente limitada siempre que pueda unirse  
 60 específicamente a una secuencia de nucleótidos diana, por ejemplo, es de 15 - 30 nucleótidos, preferiblemente de 18 - 24 nucleótidos.

Aunque un ADN que codifica para ARN guía y ARNtracr también puede insertarse en un vector de expresión similar  
 65 al mencionado anteriormente, según el huésped. Como promotor, se usan preferiblemente promotor de pol III (por ejemplo, promotor de SNR6, SNR52, SCR1, RPR1, U6, H1. etc.) y terminador (por ejemplo, secuencia T<sub>6</sub>).

Un ARN que codifica para una Cas mutante y/o una enzima convertidora de bases de ácido nucleico puede prepararse, por ejemplo, mediante transcripción a ARNm en un sistema de transcripción *in vitro* conocido *per se* usando un vector que codifica para la Cas mutante mencionada anteriormente y/o el ADN que codifica para una enzima convertidora de bases de ácido nucleico como molde.

5 Puede obtenerse ARN-ARNtracr guía diseñando una secuencia de oligoADN en la que se unen una secuencia complementaria a la secuencia de nucleótidos diana y la secuencia de ARNtracr conocida, y sintetizando químicamente un sintetizador de ADN/ARN.

10 Un ADN o ARN que codifica para una Cas mutante y/o una enzima convertidora de bases de ácido nucleico, un ARN-ARNtracr guía o un ADN que codifica para el mismo pueden introducirse en una célula huésped mediante un método similar al anterior, según el huésped.

15 Debido a que la nucleasa artificial convencional acompaña a las roturas de ADN bicatenario (DSB), la inhibición del crecimiento y la muerte celular probablemente producida por la escisión alterada (escisión fuera de diana) del cromosoma se producen seleccionando como diana una secuencia en el genoma. El efecto de la misma es particularmente mortal para muchos microorganismos y procariontes, e impide su aplicabilidad. En la presente invención, la mutación no se introduce por escisión de ADN sino por una reacción de conversión del sustituyente en la base de ADN (particularmente reacción de desaminación), y por tanto, puede realizarse una reducción drástica de la toxicidad. De hecho, tal como se muestra en las pruebas de comparación usando una levadura en gemación como huésped en los ejemplos mencionados a continuación, cuando se usa Cas9 que tiene un tipo convencional de actividad de DSB, el número de células supervivientes disminuye por la inducción de expresión, mientras que se confirmó que las células continuaron creciendo y el número de células supervivientes aumentó mediante la técnica de la presente invención usando Cas mutante y una enzima convertidora de bases de ácido nucleico en combinación (figura 3).

25 La modificación del ADN bicatenario en la presente invención no evita la aparición de escisión del ADN bicatenario en un sitio distinto del sitio seleccionado como diana (ajustado de manera apropiada dentro de varios cientos de bases incluyendo la secuencia de nucleótidos diana completa o parcial). Sin embargo, una de las mayores ventajas de la presente invención es evasión de la toxicidad por escisión fuera de diana, que es generalmente aplicable a cualquier especie. En una realización preferible, por tanto, la modificación del ADN bicatenario en la presente invención no se asocia con la escisión de cadena de ADN no sólo en un sitio seleccionado como diana de un ADN bicatenario seleccionado sino en otros sitios.

30 Tal como se muestra en los ejemplos mencionados a continuación, cuando se usa una Cas que tiene una actividad nickasa capaz de escindir sólo una de las cadenas del ADN bicatenario como Cas mutante (figura 5), la eficiencia de la introducción de mutación aumenta en comparación a cuando se usa Cas mutante que no es capaz de escindir ambas cadenas. Por tanto, por ejemplo, junto a un módulo de reconocimiento de secuencia de ácido nucleico y una enzima convertidora de bases de ácido nucleico, que se unen a una proteína que tiene una actividad nickasa, escindiendo de ese modo sólo una cadena sencilla de ADN en la proximidad de la secuencia de nucleótidos diana, la eficiencia de la introducción de mutación puede mejorarse mientras que se evita la fuerte toxicidad de DSB.

35 Además, una comparación de los efectos de Cas mutante que tiene dos clases de actividad nickasa de escisión de cadena diferente revela que usando una de la Cas mutante da como resultado sitios mutados que se acumulan cerca del centro de la secuencia de nucleótidos diana, y usando otra Cas mutante da como resultado diversas mutaciones que se introducen de manera aleatoria en una región de varios cientos de bases de la secuencia de nucleótidos diana (figura 6). Por tanto, seleccionando una cadena que va a escindirse por la nickasa, puede introducirse una mutación en un nucleótido o una región de nucleótidos particular en una posición identificada, o pueden introducirse de manera aleatoria diversas mutaciones en un intervalo comparativamente amplio, que puede adoptarse de manera apropiada según el objeto. Por ejemplo, cuando la primera técnica se aplica a una célula iPS genéticamente enferma, puede producirse agente terapéutico de trasplante celular con un menor riesgo de rechazo reparando la mutación del gen patógeno en una célula iPS producida a partir de la propia célula de los pacientes, y diferenciando la célula en la célula somática de interés.

40 El ejemplo 7 y los ejemplos posteriores mencionados a continuación muestran que puede introducirse una mutación en un nucleótido particular casi en una posición identificada. Para la introducción de una posición identificada de una mutación en un nucleótido deseado, la secuencia de nucleótidos diana debe establecerse para mostrar una determinada regularidad en la relación posicional entre un nucleótido deseado que va a introducirse con una mutación y la secuencia de nucleótidos diana. Se usa un sistema CRISPR-Cas como módulo de reconocimiento de secuencia de ácido nucleico y se usa AID como enzima convertidora de bases de ácido nucleico, la secuencia de nucleótidos diana puede diseñarse de manera que C (o G en la cadena opuesta) en la que se desea una mutación se introducirá en 2 - 5 nucleótidos desde el extremo 5' de la secuencia de nucleótidos diana. Tal como se mencionó anteriormente, la longitud de la secuencia de ARN guía puede determinarse de manera apropiada para que esté dentro de entre 15 - 30 nucleótidos, preferiblemente 18 - 24 nucleótidos. Debido a que la secuencia de ARN guía es una secuencia complementaria a la secuencia de nucleótidos diana, la longitud de la secuencia de nucleótidos diana cambia cuando se cambia la longitud de la secuencia de ARN guía; sin embargo, se mantiene la regularidad de que

es probable que se introduzca una mutación en C o G en 2 - 5 nucleótidos desde el extremo 5' independientemente de la longitud del nucleótido (figura 12). Por tanto, determinando de manera apropiada la longitud de la secuencia de nucleótidos diana (ARN guía como una cadena complementaria de la misma), el sitio de una base en el que puede introducirse una mutación puede desplazarse. Como resultado, la restricción por secuencia de reconocimiento del sitio de escisión del ADN PAM (NGG) también puede eliminarse, y el grado de libertad de la introducción de mutación se vuelve mayor.

Tal como se muestra en los ejemplos mencionados a continuación, cuando se producen módulos de reconocimiento de secuencias correspondientes a una pluralidad de secuencias de nucleótidos diana en la proximidad, y se usan simultáneamente, la eficiencia de la introducción de mutación aumenta de manera drástica con relación a cuando se usa una secuencia de nucleótidos sencilla como diana (figura 7). Como efecto del mismo, se consigue una inducción de la mutación similar incluso cuando ambas secuencias de nucleótidos diana se solapan parcialmente cuando ambas están separadas por aproximadamente 600 pb. Puede producirse cuando ambas secuencias de nucleótidos diana están en la misma dirección (las secuencias de nucleótidos diana están presentes en la misma cadena) (figura 7), y cuando son opuestas (las secuencias de nucleótidos diana están presentes en cada cadena del ADN bicatenario) (figura 4).

Tal como se muestran en los ejemplos mencionados a continuación, el método de modificación de secuencia genómica de la presente invención puede introducir mutación en casi todas las células en las que se ha expresado el complejo enzimático de modificación de ácido nucleico de la presente invención, seleccionando una secuencia de nucleótidos diana adecuada (figura 8). Por tanto, la inserción y selección de un gen marcador de selección, que son esenciales en la edición genómica convencional, no son necesarias. Esto facilita y simplifica de manera drástica la manipulación génica y extiende la aplicabilidad a la reproducción de cultivos y similares debido a que no se produce un organismo recombinante con ADN foráneo.

Debido a que el método de modificación de secuencia genómica de la presente invención muestra una eficiencia de introducción de mutación muy alta, y no requiere selección por marcadores, pueden modificarse una pluralidad de regiones de ADN en posiciones completamente diferentes como dianas (figura 9). Por tanto, en una realización preferible de la presente invención, pueden usarse dos o más clases de módulos de reconocimiento de secuencia de ácido nucleico que se unen específicamente a diferentes secuencias de nucleótidos diana (que pueden estar presentes en un gen de interés, o dos o más genes de interés diferentes, que pueden estar presentes en el mismo cromosoma o diferentes cromosomas). En este caso, cada uno de estos módulos de reconocimiento de secuencia de ácido nucleico y enzima convertidora de bases de ácido nucleico forman un complejo enzimático de modificación de ácido nucleico. En el presente documento, puede usarse una enzima convertidora de bases de ácido nucleico habitual. Por ejemplo, cuando se usa un sistema CRISPR-Cas como módulo de reconocimiento de secuencia de ácido nucleico, se usa un complejo habitual de una proteína Cas y una enzima convertidora de bases de ácido nucleico (incluyendo una proteína de fusión), y se producen dos o más clases de ARN quimérico de ARNtracr y cada uno de los dos o más ARN guía que forman respectivamente una cadena complementaria con diferentes secuencias de nucleótidos diana y se usan como ARN-ARNtracr guías. Por otro lado, cuando se usan motivo de dedos de cinc, efector TAL y similares como módulos de reconocimiento de secuencia de ácido nucleico, por ejemplo, puede fusionarse una enzima convertidora de bases de ácido nucleico con un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a un nucleótido diana diferente.

Para expresar el complejo enzimático de modificación de ácido nucleico de la presente invención en una célula huésped, tal como se mencionó anteriormente, se introduce un vector de expresión que contiene un ADN que codifica para el complejo enzimático de modificación de ácido nucleico, o un ARN que codifica para el complejo enzimático de modificación de ácido nucleico en una célula huésped. Para la introducción de mutación eficiente, es deseable mantener la expresión del complejo enzimático de modificación de ácido nucleico en un nivel dado o superior durante no menos que un periodo dado. A partir de tal aspecto, la introducción de un vector de expresión que puede replicarse de manera autónoma en una célula huésped (plásmido, etc.) es fiable. Sin embargo, debido a que el plásmido, etc. son ADN foráneos, se retiran preferiblemente de manera rápida después de la introducción exitosa de mutación. Por tanto, aunque varía dependiendo de la clase de célula huésped y similares, por ejemplo, el plásmido introducido se retira de manera deseable de la célula huésped después de un lapso de 6 h - 2 días desde la introducción de un vector de expresión usando diversos métodos de retirada de plásmidos que se conocen bien en la técnica.

Alternativamente, siempre que se logre suficiente expresión de un complejo enzimático de modificación de ácido nucleico para la introducción de mutación, es también preferible introducir una mutación en el ADN bicatenario de interés mediante expresión transitoria usando un vector de expresión sin capacidad de replicación autónoma en una célula huésped (por ejemplo, un vector que carece de origen de replicación que funciona en una célula huésped y/o un gen que codifica para la proteína necesaria para la replicación, etc.) o ARN.

La expresión del gen diana se suprime mientras que el complejo enzimático de modificación de ácido nucleico de la presente invención se expresa en una célula huésped para realizar una reacción de conversión de bases de ácido nucleico. Por tanto, fue difícil editar directamente un gen esencial para la supervivencia de la célula huésped como gen diana (da como resultado efectos secundarios tales como inhibición del crecimiento del huésped, eficiencia de la

introducción de mutación inestable, mutación de un sitio diferente de la diana y similares). En la presente invención, la edición directa de un gen esencial se ha conseguido de manera exitosa y eficiente produciendo una reacción de conversión de bases de ácido nucleico en una fase deseada, y expresando de manera transitoria el complejo enzimático de modificación de ácido nucleico de la presente invención en una célula huésped durante un periodo necesario para estabilizar la modificación del sitio seleccionado como diana. Aunque el periodo necesario para una reacción de conversión de bases de ácido nucleico y la estabilización de la modificación del sitio seleccionado como diana varía dependiendo de la clase de célula huésped, las condiciones de cultivo y similares, se considera que son necesarias generalmente 2 - 20 generaciones. Por ejemplo, cuando la célula huésped es una levadura o bacteria (por ejemplo, *Escherichia coli*), la expresión de un complejo enzimático de modificación de ácido nucleico necesita inducirse durante 5 - 10 generaciones. Los expertos habituales en la técnica pueden determinar de manera apropiada un periodo de inducción de la expresión preferible basándose en el tiempo de duplicación de la célula huésped en las condiciones de cultivo usadas. Por ejemplo, cuando una levadura en gemación se somete a cultivo líquido en un medio inductor de galactosa al 0,02%, el periodo de inducción de la expresión es de, por ejemplo, 20 - 40 h. El periodo de inducción de la expresión del ácido nucleico que codifica para el complejo enzimático de modificación de ácido nucleico de la presente invención puede extenderse más allá del "periodo necesario para establecer la modificación del sitio seleccionado como diana" mencionado anteriormente en la medida que no cause efectos secundarios a la célula huésped.

Como medio para expresar de manera transitoria el complejo enzimático de modificación de ácido nucleico de la presente invención en una fase deseada durante un periodo deseado, puede usarse un método que comprende producir un constructo (vector de expresión) que contiene un ácido nucleico que codifica para el complejo enzimático de modificación de ácido nucleico (un ADN que codifica para un ARN-ARNtracr guía y un ADN que codifica para una Cas mutante y enzima de sustitución de bases de ácido nucleico en el caso del sistema CRISPR-Cas), de manera que el periodo de expresión pueda controlarse, e introducir el constructo en una célula huésped. La "manera en la que el periodo de expresión puede controlarse" es específicamente, por ejemplo, un ácido nucleico que codifica para el complejo enzimático de modificación de ácido nucleico de la presente invención colocado bajo regulación de una región reguladora inducible. Aunque la "región reguladora inducible" no está particularmente limitada, es, por ejemplo, un operón de un represor de mutación sensible a la temperatura (ts) y un operador regulado de ese modo en células de microorganismos de bacteria (por ejemplo, *Escherichia coli*), levadura y similares. Los ejemplos del represor de mutación ts incluyen, pero no se limitan a, mutación ts del represor de cl del fago  $\lambda$ . En el caso del represor de cl (ts) del fago  $\lambda$ , se une a un operador para suprimir la expresión del gen en el sentido de 3' a no más de 30°C (por ejemplo, 28°C). A una temperatura alta de no menos de 37°C (por ejemplo, 42°C), se disocia del operador para permitir la inducción de la expresión génica (figuras 13 y 14). Por tanto, el periodo cuando se suprime la expresión del gen diana puede minimizarse cultivando una célula huésped introducida con un ácido nucleico que codifica para el complejo enzimático de modificación de ácido nucleico generalmente a no más de 30°C, elevando la temperatura a no menos de 37°C en una fase apropiada, realizando un cultivo durante un periodo dado para llevar a cabo una reacción de conversión de bases de ácido nucleico y, después la introducción de la mutación en el gen diana, disminuyendo rápidamente la temperatura a no más de 30°C. Por tanto, incluso cuando se selecciona como diana un gen esencial para la célula huésped, puede editarse de manera eficaz mientras que se suprimen los efectos secundarios (figura 15).

Cuando se utiliza una mutación sensible a la temperatura, por ejemplo, se incluye un mutante sensible a la temperatura de una proteína necesaria para la replicación autónoma de un vector en un vector que contiene un ADN que codifica para el complejo enzimático de modificación de ácido nucleico de la presente invención. Como resultado, la replicación autónoma se vuelve rápidamente imposible después de la expresión del complejo enzimático de modificación de ácido nucleico, y el vector decae de manera natural durante la división celular. Los ejemplos de proteína mutante sensible a la temperatura incluyen, pero no se limitan a, un mutante sensible a la temperatura de Rep101 ori necesario para la replicación de pSC101 ori. Rep101 ori (ts) actúa sobre pSC101 ori para permitir la replicación autónoma del plásmido a no más de 30°C (por ejemplo, 28°C), pero pierde la función a no menos de 37°C (por ejemplo, 42°C), y el plásmido no puede replicarse de manera autónoma. Por tanto, un uso combinado con el represor de cl (ts) del fago  $\lambda$  mencionado anteriormente de manera simultánea permite la expresión transitoria del complejo enzimático de modificación de ácido nucleico de la presente invención, y la retirada del plásmido.

Por otro lado, cuando se usa una célula eucariota superior tal como una célula animal, célula de insecto, célula vegetal y similares como célula huésped, un ADN que codifica para el complejo enzimático de modificación de ácido nucleico de la presente invención que se introduce en una célula huésped bajo regulación de un promotor inducible (por ejemplo, un promotor de metalotioneína (inducido por un ion de metal pesado), promotor de proteína de choque térmico (inducido por choque térmico), promotor del sistema Tet-ON/Tet-OFF (inducido por la adición o retirada de tetraciclina o un derivado del mismo), promotor sensible a compuestos esteroideos (inducido por la hormona esteroidea o un derivado de la misma), etc.), la sustancia de inducción se añade al medio (o se retira del medio) en una fase apropiada para inducir la expresión del complejo enzimático de modificación de ácido nucleico, el cultivo se realiza durante un periodo dado para llevar a cabo una reacción de conversión de bases de ácido nucleico y, la introducción de la mutación en el gen diana, puede realizarse una expresión transitoria del complejo enzimático de modificación de ácido nucleico.

En células procariotas tales como *Escherichia coli* y similares, también pueden usarse promotores inducibles. Los ejemplos de tales promotores inducibles incluyen, pero no se limitan a, promotor de lac (inducido por IPTG), promotor de cspA (inducido por choque frío), promotor de araBAD (inducido por arabinosa) y similares.

5 Alternativamente, los promotores inducibles mencionados anteriormente también pueden utilizarse como un mecanismo de retirada de vectores cuando se usan células eucariotas superiores tales como una célula animal, célula de insecto, célula vegetal y similares como célula huésped. Es decir, se carga un vector con un origen de replicación que puede funcionar en una célula huésped, y un ácido nucleico que codifica para una proteína necesaria para la replicación del mismo (por ejemplo, VS40 ori y antígeno T grande, oriP y EBNA-1, etc., para células animales), y la expresión del ácido nucleico que codifica para la proteína se regula por el promotor inducible mencionado anteriormente. Como resultado, aunque el vector puede replicarse de manera autónoma en presencia de una sustancia de inducción, cuando se retira la sustancia de inducción, no se produce la replicación autónoma, y el vector de manera natural decae durante la división celular (por otro lado, la replicación autónoma se vuelve imposible por la adición de tetraciclina y doxiciclina en el caso del vector del sistema Tet-OFF).

15 La presente invención se explica a continuación mediante referencia a los siguientes ejemplos, que no se interpretan como limitativos.

### [Ejemplos]

20 En los ejemplos mencionados a continuación 1 - 6, se realizaron ensayos de la siguiente manera.

<Línea celular, cultivo, transformación e inducción de la expresión>

25 Se cultivó levadura en gemación de *Saccharomyces cerevisiae* cepa BY4741 (que requiere leucina y uracilo) en un medio YPDA o medio SD convencional con una composición de abandono que satisface la auxotroficidad. El cultivo realizado fue cultivo fijo en una placa de agar o cultivo en agitación en un medio líquido entre 25°C y 30°C. Se realizó la transformación mediante un método de acetato de litio, y se realizó la selección en medio SD que corresponde a auxotroficidad apropiada. Para determinar la inducción de la expresión por galactosa, después del cultivo previo durante la noche en un medio SD apropiado, se realizaron el cultivo en medio SR durante la noche con fuente de carbono cambiada desde el 2% de glucosa hasta el 2% de rafinosa, y un cultivo adicional en medio SGal durante de 3 h a aproximadamente dos noches con fuente de carbono cambiada del 0,2 - 2% de galactosa para determinar la inducción de la expresión.

35 Para la medición del número de células supervivientes y la tasa de mutación de Can1, se diluyó de manera apropiada una suspensión celular, y se aplicó sobre un medio de placa con SD y medio de placa con SD-Arg+canavanina 60 mg/l o medio de placa con SD+canavanina 300 mg/l, y se contó el número de colonias que surgieron 3 días más tarde como el número células supervivientes. Usando el número de colonias supervivientes en placa con SD como el número total de células, y el número de colonias supervivientes en placa con canavanina como el número de cepa mutante resistente, se calculó y se evaluó la tasa de mutación. Se identificó el sitio de la introducción de mutación amplificando fragmentos de ADN que contenían la región del gen diana de cada cepa mediante un método de PCR de colonias, realizando secuenciación de ADN, y realizando un análisis de alineación basándose en la secuencia de la base de datos de genomas de *Saccharomyces* (<http://www.yeastgenome.org/>).

45 <Manejo del ácido nucleico>

Se procesó o construyó ADN mediante cualquiera de método de PCR, tratamiento de enzimas de restricción, ligación, método de ensamblaje de Gibson y síntesis química artificial. Para el plásmido, se usaron como levadura vector lanzadera de *Escherichia coli*, pRS315 para la selección de leucina y pRS426 para la selección de uracilo como la estructura principal. Se amplificó el plásmido por *Escherichia coli* línea XL-10 gold o DH5 $\alpha$ , y se introdujo en una levadura mediante el método de acetato de litio.

<Constructo>

55 Para determinar la expresión inducible, se usó levadura en gemación pGal1/10 (SEQ ID NO: 8), que es un promotor inducible bidireccional por galactosa. En el sentido de 3' de un promotor, se añadió una señal de localización nuclear (ccc aag agg aag gtg; SEQ ID NO: 9 (PKKKRV; que codifica SEQ ID NO: 10)) al ORF del gen Cas9 de *Streptococcus pyogenes* que tiene un codón optimizado para expresión en eucariotes (SEQ ID NO: 5) y se unió el ORF (SEQ ID NO: 1 ó 3) del gen de desaminasa (PmCDA1 de *Petromyzon marinus* o hAID de humano) a través de una secuencia de ligador, y se expresó como una proteína de fusión. Como secuencia de ligador, se seleccionaron y usaron en combinación ligador GS (repetición de ggt gga ggt tct; SEQ ID NO: 11 (que codifica para GGGGS; SEQ ID NO: 12)), etiqueta Flag (gac tat aag gac cacgac gga gac tac aag gat cat gat att gat tac aaa gac gat gac gat aag; SEQ ID NO: 13 (que codifica para DYKDHDGDYKDHDIDYKDDDDK; SEQ ID NO: 14)), etiqueta Strep (tgg agc cac ccg cag ttc gaa aaa; SEQ ID NO: 15 (que codifica para WSHQPFEK; SEQ ID NO: 16)), y otros dominios. En el presente documento, particularmente, se unieron y se usaron 2xGS, dominio SH3 (SEQ ID NO: 17 y 18) y etiqueta Flag. Como terminador, se unieron terminador ADH1 de levadura en gemación (SEQ ID NO: 19) y terminador Top2 (SEQ

ID NO: 20). En el método de integración de dominios, se unió el ORF del gen Cas9 al dominio SH3 a través de ligador 2xGS para proporcionar una proteína, se añadió desaminasa con secuencia de ligando de SH3 (las SEQ ID NO: 21 y 22) como otra proteína, y se unieron al promotor de Gal1/10 en ambos lados. Y se expresaron simultáneamente. Se incorporaron en el plásmido pRS315.

En Cas9, se introdujeron la mutación para convertir el 10° ácido aspártico en alanina (D10A, que corresponde a la mutación de secuencia de ADN a29c) y la mutación para convertir la 840° histidina en alanina (H840A, que corresponde a la mutación de secuencia de ADN ca2518gc) para eliminar la capacidad de escisión de cualquier lado de la cadena de ADN.

Se dispuso el ARNg como una estructura quimérica con ARNtracr (de *Streptococcus pyogenes*; SEQ ID NO: 7) entre el promotor de SNR52 (SEQ ID NO: 23) y el terminador Sup4 (SEQ ID NO: 24), y se incorporó en el plásmido pRS426. Como secuencia de bases diana de ARNg, se usaron el ORF del gen CAN1, 187-206 (gatacgttctctatggagga; SEQ ID NO: 25) (diana 1), 786-805 (ttggagaaccaggtgcct; SEQ ID NO: 26) (diana 3), 793-812 (aacccaggtgcctgggtcc; SEQ ID NO: 27) (diana 4), 563-582 (ttggccaagtcattcaatt; SEQ ID NO: 28) (diana 2) y secuencia de cadena complementaria de 767-786 (ataacggaatccaactgggc; SEQ ID NO: 29) (diana 5r). Para la determinación de la expresión simultánea de una pluralidad de dianas, usando una secuencia desde un promotor a un terminador como un conjunto, y una pluralidad de los conjuntos se incorporaron en el mismo plásmido. Se introdujeron en células junto con un plásmido de expresión de Cas9-desaminasa, se expresaron de manera intracelular, y se formó un complejo de ARNg-ARNtracr y Cas9-desaminasa.

Ejemplo 1: modificación de secuencia genómica uniendo la capacidad de reconocimiento de secuencia de ADN de CRISPR-Cas a desaminasa PmCDA1

Para someter a prueba el efecto de la técnica de modificación de secuencia genómica de la presente invención utilizando desaminasa y la capacidad de reconocimiento de secuencia de ácido nucleico de CRISPR-Cas, se intentó la introducción de mutación en el gen CAN1 que codifica para el transportador de canavanina, cuyo déficit génico da como resultado una resistencia a la canavanina. Como ARNg, se usó una secuencia complementaria a 187-206 (diana 1) del ORF del gen CAN1, se construyeron un vector de expresión de ARN quimérico obtenido uniendo al mismo ARNtracr de *Streptococcus pyogenes*, y un vector que expresa una proteína obtenida fusionando dCas9 con actividad nucleasa alterada introduciendo mutaciones (D10A y H840A) en Cas9 (SpCas9) de *Streptococcus pyogenes* con PmCDA1 de *Petromyzon marinus* como desaminasa, se introdujeron en la levadura en gemación mediante el método de acetato de litio, y se expresaron de manera conjunta. Los resultados se muestran en la figura 2. Cuando se cultivaron sobre una placa con SD que contenía canavanina, sólo las células sometidas a introducción y expresión de ARNg-ARNtracr y dCas9-PmCDA1 formaron colonia. Se recogió la colonia resistente y se determinó la secuencia de región del gen CAN1. Como resultado, se confirmó que una mutación se introdujo en la secuencia de nucleótidos diana (diana 1) y la proximidad de la misma.

Ejemplo 2: reducción drástica de efectos secundarios-toxicidad

En la Cas9 convencional y otras nucleasas artificiales (ZFN, TALEN), la inhibición del crecimiento y la muerte celular producida probablemente por la escisión alterada del cromosoma se produce seleccionando como diana una secuencia en el genoma. El efecto de la misma es particularmente mortal para muchos microorganismos y procariontes, e impide su aplicabilidad.

Por tanto, para verificar la seguridad y la toxicidad celular de la técnica de modificación de secuencia genómica de la presente invención, se realizó una prueba comparativa con CRISPR-Cas9 convencional. Usando secuencias (dianas 3, 4) en el gen CAN1 como diana de ARNg, las células supervivientes se contaron de manera inmediata después del inicio de la inducción de la expresión por galactosa y a las 6 h después de la inducción basándose en la capacidad de formación de colonias sobre una placa con SD. Los resultados se muestran en la figura 3. En Cas9 convencional, se inhibió el crecimiento y se indujo la muerte celular, lo que disminuyó el número de células supervivientes. Por el contrario, mediante la presente técnica (nCas9 D10A-PmCDA1), las células podían continuar creciendo, y el número de células supervivientes aumentó de manera drástica.

Ejemplo 3: uso de un esquema de unión diferente

Se examinó si la mutación puede introducirse en un gen seleccionado como diana incluso cuando no se usan Cas9 y desaminasa como proteína de fusión pero cuando se forma un complejo enzimático de modificación de ácido nucleico a través de un dominio de unión y un ligando del mismo. Como Cas9, se usó dCas9 usada en el ejemplo 1 y se usó AID humana en vez de PmCDA1 como desaminasa. Se fusionó el dominio SH3 con la primera, y un ligando de unión del mismo se fusionó con la última para producir diversos constructos conocidos en la figura 4. Además, se usaron las secuencias (diana 4, 5r) en el gen CAN1 como dianas de ARNg. Se introdujeron estos constructos en una levadura en gemación. Como resultado, incluso cuando se unieron dCas9 y desaminasa a través del dominio de unión, la mutación se introdujo de manera eficaz en el sitio seleccionado como diana del gen CAN1 (figura 4). La eficiencia de la introducción de mutación se mejoró de manera notable introduciendo una pluralidad de dominios de unión en dCas9. El sitio principal de introducción de mutación fue el 782° (g782c) de ORF.

Ejemplo 4: alta eficiencia y cambios en el patrón de mutación mediante el uso de nickasa

De la misma manera que en el ejemplo 1 excepto que se usó un mutante de D10A nCas9 (D10A) que escinde sólo una cadena complementaria a ARNg, o un mutante de H840A nCas9 (H840A) que escinde sólo una cadena opuesta a una cadena complementaria a ARNg en vez de dCas9, se introdujo una mutación en el gen CAN1, y se examinó la secuencia en la región del gen CAN1 de la colonia generada sobre una placa con SD que contenía canavanina. Se encontró que la eficiencia aumenta en el primero (nCas9 (D10A)) en comparación con dCas9 (figura 5), y la mutación se reúne en el centro de la secuencia diana (figura 6). Por tanto, este método permite la introducción de una posición identificada de mutación. Por otro lado, se encontró en el último (nCas9 (H840A)) que se introdujeron una pluralidad de mutaciones aleatorias en una región de varios cientos de bases del nucleótido seleccionado como diana (figura 6) junto con un aumento en la eficiencia en comparación con dCas9 (figura 5).

Una introducción similar notable de mutación pudo confirmarse incluso cuando se cambió la secuencia de nucleótidos diana. En este sistema de edición genómica usando un sistema CRISPR-Cas9 y citidina desaminasa, se confirmó tal como se muestra en la tabla 1 que la citosina presente dentro del intervalo de aproximadamente 2 - 5 pb desde el lado de 5' de la secuencia de nucleótidos diana (20 pb) se desaminó de manera preferente. Por tanto, ajustando la secuencia de nucleótidos diana basándose en esta regularidad y combinando además con nCas9 (D10A), es posible la edición genómica precisa de 1 unidad de nucleótido. Por otro lado, puede insertarse una pluralidad de mutaciones simultáneamente dentro del intervalo de aproximadamente varios cientos de pb en la proximidad de la secuencia de nucleótidos diana usando nCas9 (H840A). Además, la especificidad de sitio posiblemente puede variarse adicionalmente cambiando el esquema de unión de desaminasa.

Estos resultados muestran que la clase de proteína Cas9 puede cambiarse de manera apropiada según el objeto.

Tabla 1

Posición del ORF del gen CAN1	Secuencia (SEQ ID NO:)	Sitio de introducción de la mutación principal
187 - 206 (diana 1)	gatacgttctctatggagga (25)	c191a, g226t
563 - 582 (diana 2)	ttggccaagtcattcaattt (28)	cc567at, c567del
786 - 805 (diana 3) y 793 - 812 (diana 4)	ttggagaaacccaggtgcct (26) aaccaggtgcctgggtcc (27)	cc795tt, cc796tt
767 - 786 (cadena complementaria) (diana 5r)	ataacggaatccaactgggc (29)	g782c

Ejemplo 5: la eficiencia aumenta de manera sinérgica seleccionando como diana una pluralidad de secuencias de ADN en la proximidad

La eficiencia aumentó de manera drástica usando simultáneamente una pluralidad de dianas en la proximidad en vez de una única diana (figura 7). De hecho, el 10 - 20% de células tenían una mutación resistente a la canavanina (dianas 3, 4). En la figura, ARNg1 y ARNg2 seleccionan como diana la diana 3 y diana 4, respectivamente. Como desaminasa, se usó PmCDA1. El efecto de la misma se confirmó que se producía no sólo cuando las secuencias se solaparon de manera parcial (dianas 3, 4), sino también cuando estaban separadas por aproximadamente 600 pb (dianas 1, 3). El efecto se encontró tanto cuando las secuencias de ADN estaban en la misma dirección (dianas 3, 4) como opuestas (dianas 4, 5) (figura 4).

Ejemplo 6: modificación genética que no requiere un marcador de selección

En cuanto a las células (dianas 3, 4) en las que se seleccionaron como diana la diana 3 y la diana 4 en el ejemplo 5, se seleccionaron de manera aleatoria 10 colonias de las que se hicieron crecer sobre una placa no seleccionada (libre de canavanina) (placa con SD que no contiene ni Leu ni Ura) y se determinaron las secuencias de la región del gen CAN1. Como resultado, se introdujo una mutación en el sitio seleccionado como diana del gen CAN1 en todas las colonias examinadas (figura 8). Es decir, puede esperarse la edición en casi todas las células expresadas seleccionando una secuencia diana adecuada según la presente invención. Por tanto, la inserción de un gen marcador y la selección, que son esenciales para la manipulación génica convencional, no son necesarias. Esto facilita y simplifica de manera drástica la manipulación génica y amplía la aplicabilidad a reproducción de cultivos y similares debido a que no se produce un organismo recombinante con ADN foráneo.

En los siguientes ejemplos, se realizaron las técnicas de ensayo compartidas en los ejemplos 1 - 6 de la misma manera que anteriormente.

Ejemplo 7: edición simultánea de una pluralidad de sitios (gen diferente)

En un método de manipulación génica general, la mutación de sólo un sitio se logra generalmente mediante un

manejo debido a diversas restricciones. Por tanto, se sometió a prueba si un manejo de mutación simultánea de una pluralidad de sitios es posible usando el método de la presente invención.

5 Usando el ORF de las posiciones 3 a 22 del gen Ade1 de la cepa BY4741 de levadura en gemación como la primera secuencia de nucleótidos diana (Ade1, diana 5: GTCAATTACGAAGACTGAAC; SEQ ID NO: 30), y el ORF de las posiciones 767-786 (cadena complementaria) del gen Can1 como la segunda secuencia de nucleótidos diana (Can1, diana 8 (786-767; ATAACGGAATCCAACCTGGGC; SEQ ID NO: 29), codificando ambos ADN para ARN quiméricos de dos clases de ARNg que contienen cada uno una secuencia de nucleótidos complementaria al mismo y ARNtracr (SEQ ID NO: 7) se colocaron en el mismo plásmido (pRS426), y se introdujeron en la cepa BY4741 junto con el plásmido nCas9 D10A-PmCDA1 que contiene un ácido nucleico que codifica para una proteína de fusión de Cas9 mutante y PmCDA1, y se expresaron, y se verificó la introducción de mutación en ambos genes. Se cultivaron las células en un medio SD de abandono (deficiencia de uracilo y leucina; SD-UL) como base, que mantiene el plásmido. Se diluyeron de manera apropiada las células, se aplicaron sobre medio de adición con SD-UL y canavanina y se dejó que formaran una colonia. Después de 2 días de cultivo a 28°C, se observaron colonias, y se contaron la incidencia de la colonia roja debido a la mutación ade1, y la tasa de supervivencia en un medio con canavanina respectivamente. Los resultados se muestran en la tabla 2.

Tabla 2

Medio	Incidencia de la colonia roja	Tasa de supervivencia en medio con canavanina (Can)	Tasa de supervivencia de la colonia roja y Can
SD-UL	0,54±0,04		
+ canavanina	0,64±0,14	0,51±0,15	0,31±0,04

20 Como fenotipo, la proporción de introducción de mutación tanto en el gen Ade1 como en el gen Can1 fue alta y aproximadamente del 31%.

25 Luego, se sometió una colonia sobre un medio SD-UL a amplificación por PCR seguido por secuenciación. Se amplificaron las regiones que contenían el ORF de cada uno de Ade1 y Can1, y se obtuvo una información de secuencia de aproximadamente secuencias de 500 b que rodean la secuencia diana. Para ser específico, se analizaron 5 colonias rojas y 5 colonias blancas para encontrar la conversión de la 5° C del ORF del gen Ade1 en G en todas las colonias rojas y la 5° C en T en todas las colonias blancas (figura 9). Aunque la tasa de mutación de la diana es del 100%, como la tasa de mutación según el objeto de la destrucción génica, la tasa de mutación deseada es del 50% ya que es necesario que la 5° C se cambie por G para ser un codón de parada. De manera similar, en cuanto al gen Can1, se confirmó la mutación en la 782° G del ORF en todos los clones (figura 9); sin embargo, ya que sólo la mutación en C proporciona resistencia a la canavanina, la tasa de mutación deseada es del 70%. Las mutaciones deseadas en ambos genes se obtuvieron simultáneamente en el 40% de los clones (4 clones de cada 10 clones) mediante investigación, y se obtuvo prácticamente una alta eficiencia.

### 35 Ejemplo 8: edición de genoma poliploide

40 Muchos organismos tienen genoma diploide o poliploide. En los métodos de introducción de mutación convencionales, se introduce la mutación, en principio, en un único cromosoma homólogo para producir un genotipo heterólogo. Por tanto, no se obtiene una característica deseada a menos que no sea una mutación dominante, y hacerla homóloga requiere trabajo y tiempo. Por tanto, se sometió a prueba si la técnica de la presente invención puede introducir mutación en todos los alelos diana en el cromosoma homólogo en el genoma.

45 Es decir, se realizó la edición simultánea de los genes Ade1 y Can1 en una cepa YPH501 de levadura en gemación como cepa diploide. El fenotipo de estas mutaciones génicas (colonia roja y resistente a canavanina) es un fenotipo recesivo, y por tanto, estos fenotipos no aparecen a menos que se introduzcan ambas mutaciones del gen homólogo (mutación homóloga).

50 Usando el ORF de las posiciones 1173-1154 (cadena complementaria) del gen Ade1 (Ade1, diana 1: GTCAATAGGATCCCCTTTT; SEQ ID NO: 31) o de las posiciones 3-22 (Ade1, diana 5: GTCAATTACGAAGACTGAAC; SEQ ID NO: 30) como la primera secuencia de nucleótidos diana, y el ORF de las posiciones 767-786 (cadena complementaria) del gen Can1 como la segunda secuencia de nucleótidos diana (Can1, diana 8: ATAACGGAATCCAACCTGGGC; SEQ ID NO: 29), se colocaron ambos ADN que codifican para ARN quiméricos de dos clases de ARNg que contienen cada uno una secuencia de nucleótidos complementaria al mismo y ARNtracr (SEQ ID NO: 7) en el mismo plásmido (pRS426), y se introdujeron en la cepa BY4741 junto con el plásmido nCas9 D10A-PmCDA1 que contiene un ácido nucleico que codifica para una proteína de fusión de Cas9 mutante y PmCDA1, y se expresaron, y se verificó la introducción de mutación en cada gen.

60 Como resultado del recuento de colonias, se encontró que cada característica de fenotipo podía obtenerse a una alta probabilidad (40% - 70%) (figura 10A).

Para confirmar la mutación, se secuenciaron la región diana de Ade1 de cada una de la colonia blanca y la colonia

roja para confirmar el solapamiento de las señales de secuencia que indican una mutación heteróloga en el sitio diana de la colonia blanca (figura 10B, panel superior, las señales de G y T solapan en ↓). Se confirmó que estaba ausente el fenotipo en la colonia con mutación heteróloga. Por otro lado, se confirmó la mutación homóloga libre de señal de solapamiento en la colonia roja (figura 10B, panel inferior, señal de T en ↓).

#### Ejemplo 9: edición genómica en *Escherichia coli*

En este ejemplo, se demuestra que esta técnica funciona de manera eficaz en *Escherichia coli*, que es un organismo modelo de bacteria representativo. Particularmente, la técnica de edición genómica de tipo nucleasa convencional es mortal para las bacterias, y la aplicación es difícil. Por tanto, se destaca la superioridad de esta técnica. En combinación con levadura, que es una célula modelo de eucariota, se demuestra que esta técnica es ampliamente aplicable a cualquier especie independientemente de procarionte y eucarionte.

Se introdujo mutación de aminoácido de D10A y H840A (dCas9) en el gen Cas9 de *Streptococcus pyogenes* que contiene una región promotora bidireccional, y se construyó un constructo que va a expresarse como una proteína de fusión con PmCDA1 a través de una secuencia de ligador, se incluyeron simultáneamente ARN quiméricos que codifican para una secuencia complementaria a cada una de las secuencias de nucleótidos diana en un plásmido (la secuencia de nucleótidos de longitud completa se muestra en SEQ ID NO: 32, en cuya secuencia, se introduce una secuencia complementaria a cada una de las secuencias diana en el sitio de  $n_{20}$ ) (figura 11A).

En primer lugar, el ORF de las posiciones 426-445 (T CAA TGG GCT AAC TAC GTT C; SEQ ID NO: 33) del gen galK de *Escherichia coli* se introdujo como secuencia de nucleótidos diana en un plásmido, se transformaron diversas cepas de *Escherichia coli* (XL10-gold, DH5a, MG1655, BW25113) con el plásmido mediante el método de calcio o el método de electroporación, se añadió medio SOC, se realizó un cultivo de recuperación durante la noche, se seleccionaron células que portan plásmido del medio LB que contiene ampicilina, y se formó la colonia. Se verificó la introducción de mutación mediante secuencia directa a partir de la PCR de la colonia. Los resultados se muestran en la figura 11B.

Se seleccionó de manera aleatoria una colonia independiente (1 - 3), y se analizó la secuencia. Como resultado, la posición 427 de C del ORF se convirtió en T (clones 2, 3) a una probabilidad de no menos del 60%, y se confirmó la aparición de destrucción génica que genera un codón de parada (TAA).

Luego, con una secuencia complementaria (5'-GGTCCATAAACTGAGACAGC-3'; SEQ ID NO: 34) de la región de bases 1530-1549 del ORF del gen rpoB, que es un gen esencial, como diana, se introdujo una mutación puntual particular mediante un método similar al método mencionado anteriormente para tratar de conferir una función resistente a la rifampicina para *Escherichia coli*. Las secuencias de colonias seleccionadas se analizaron en un medio no selectivo (ninguno), un medio que contenía rifampicina 25 µg/ml (Rif25) y rifampicina 50 µg/ml (Rif50). Como resultado, se confirmó que la conversión de la posición 1546 de G del ORF en A introdujo una mutación de aminoácidos desde Asp(GAC) hasta Asn(AAC), y se confirió la resistencia a la rifampicina (figura 11C, panel superior). Se añadió gota a gota una serie de diluciones de 10 veces de la suspensión celular después del tratamiento de transformación en un medio no selectivo (ninguno), un medio que contenía rifampicina 25 µg/ml (Rif25) y rifampicina 50 µg/ml (Rif50) y se cultivó. Como resultado, se estimó que la cepa resistente a la rifampicina se obtuvo a aproximadamente el 10% de frecuencia (figura 11C, panel inferior).

Tal como se muestra a continuación, mediante esta técnica, puede añadirse una nueva función mediante una mutación puntual particular, en vez de una destrucción génica simple. Esta técnica es superior ya que se edita directamente un gen esencial.

#### Ejemplo 10: ajuste del de sitio de bases de edición por la longitud del ARNg

De manera convencional, la longitud del ARNg en relación con una secuencia de nucleótidos diana era de 20 b como básica, y se usó citosina (o guanina en la cadena opuesta) en un sitio de 2 – 5 b desde el extremo 5' terminal de la misma (15 – 19 b en el sentido 5' de la secuencia PAM) como diana de mutación. Se examinó si la expresión de una longitud de ARNg diferente puede desplazar el sitio de la base que va a ser la diana (figura 12A).

El ejemplo experimental realizado en *Escherichia coli* se muestra en la figura 12B. Se buscó un sitio que contiene muchas citosinas en el genoma de *Escherichia coli*, y se realizó el ensayo usando el gen *gsiA*, que es un supuesto transportador de ABC. Se examinó la citosina sustituida mientras que se cambiaba la longitud de la diana a 24 pb, 22 pb, 20 pb, 18 pb para encontrar que la citosina 898<sup>a</sup>, 899<sup>a</sup> se sustituyó por timina en el caso de 20 pb (longitud habitual). Cuando el sitio diana es más largo que 20 pb, las citosinas 896<sup>a</sup> y 897<sup>a</sup> también se sustituyeron, y cuando el sitio diana era más corto, las citosinas 900<sup>a</sup> y 901<sup>a</sup> también se sustituyeron. De hecho, el sitio diana pudo desplazarse cambiando la longitud del ARNg.

#### Ejemplo 11: desarrollo de un plásmido de edición genómica dependiente de la temperatura

Se diseñó un plásmido que induce la expresión del complejo enzimático de modificación de ácido nucleico de la

presente invención en condiciones de alta temperatura. Mientras se optimizaba la eficiencia controlando de manera limitativa el estado de la expresión, se aspiró a la reducción de los efectos secundarios (inhibición del crecimiento del huésped, eficiencia de la introducción de mutación inestable, mutación de un sitio diferente de la diana y similares). Simultáneamente, se pretendía una retirada simultánea y fácil del plásmido después de la edición combinando un mecanismo para cesar la replicación de plásmido a una alta temperatura. El detalle del ensayo se muestra a continuación.

Con un sistema de plásmido sensible a la temperatura pSC101-Rep101 (la secuencia de pSC101 ori se muestra en SEQ ID NO: 35, y la secuencia de Rep 101 sensible a la temperatura Rep101 se muestra en SEQ ID NO: 36) como estructura principal, se usó un sistema de represor de  $\lambda$  sensible a la temperatura (cl857) para determinar la inducción de la expresión. Para determinar la edición genómica, se introdujo una mutación G113E que confiere resistencia a RecA en el represor de  $\lambda$ , para asegurar la función normal incluso bajo la respuesta SOS (SEQ ID NO: 37). Se unió dCas9-PmCDA1 (SEQ ID NO: 38) al operador derecho (SEQ ID NO: 39), y se unió ARNg (SEQ ID NO: 40) en el sentido de 3' en el operador izquierdo (SEQ ID NO: 41) para regular la expresión (la secuencia de nucleótidos de longitud completa del vector de expresión construido se muestra en SEQ ID NO: 42). Durante el cultivo a no más de 30°C, se suprimió la transcripción de ARNg y la expresión de dCas9-PmCDA1, y las células pueden crecer normalmente. Cuando se cultivó a no menos de 37°C, se indujo la transcripción de ARNg y la expresión de dCas9-PmCDA1, y se suprimió la replicación del plásmido simultáneamente. Por tanto, se suministró de manera transitoria un complejo enzimático de modificación de ácido nucleico necesario para la edición genómica, y el plásmido puede retirarse fácilmente después de la edición (figura 13).

El protocolo específico de la sustitución de bases se muestra en la figura 14.

La temperatura del cultivo para la construcción del plásmido se establece a aproximadamente 28°C, y se establece en primer lugar una colonia de *Escherichia coli* que retiene el plásmido deseado. Luego, se usa directamente la colonia, o después de la extracción del plásmido cuando se cambia la cepa, se realiza la transformación con la cepa diana de nuevo, y se usa la colonia. Se realiza el cultivo líquido a 28°C durante la noche. Después de eso, se diluye la colonia con el medio, se realiza el cultivo de inducción a 42°C durante de 1 h a durante la noche, se diluye de manera apropiada la suspensión celular y se siembra en superficie o se añade gota a gota sobre una placa para conseguir una única colonia.

Como ensayo de verificación, se realizó la introducción de mutación puntual en un gen rpoB esencial. Cuando rpoB, que es uno de los factores que constituyen la ARN polimerasa, se deleta o se pierde su función, la *Escherichia coli* no sobrevivirá. Por otro lado, se sabe que la resistencia al antibiótico rifampicina (Rif) se consigue cuando la mutación puntual se introduce en un sitio particular. Por tanto, con el objetivo de dicha introducción de la mutación puntual, se selecciona un sitio diana y se realiza el ensayo.

Los resultados se muestran en la figura 15. En el panel izquierdo superior, la parte izquierda muestra una placa con LB (adición de cloramfenicol), y la parte derecha muestra una placa con LB con rifampicina añadida (adición de cloramfenicol), y se prepararon muestras con o sin cloramfenicol y se cultivaron a 28°C o 42°C. Cuando se cultivan a 28°C, la tasa de resistencia a Rif es menor; sin embargo, cuando se cultivan a 42°C, se obtuvo resistencia a la rifampicina con una eficiencia muy alta. Cuando se secuenciaron las colonias (sin selección) obtenidas sobre LB mediante 8 colonias, se sustituyó la 1546<sup>o</sup> guanina (G) por adenina (A) en no menos del 60% de la cepa cultivada a 42°C (paneles izquierdos inferior y superior). Está claro que la base también se sustituye de manera completa en el espectro de secuencias real (panel derecho inferior).

De manera similar, se realizó la sustitución de bases de galK, que es uno de los factores implicados en el metabolismo de la galactosa. Debido a que el metabolismo de la 2-desoxi-galactosa (2DOG), que es un análogo de la galactosa, por galK es mortal para *Escherichia coli*, ésta se usó como método de selección. Se estableció el sitio diana de manera que se induce una mutación de cambio de sentido en la diana 8, y se introduce el codón de parada en la diana 12 (figura 16, parte inferior derecha).

Los resultados se muestran en la figura 16. En los paneles superior izquierdo e inferior izquierdo, la parte izquierda muestra una placa con LB (adición de cloramfenicol), y la parte derecha muestra una placa con LB con 2-DOG añadida (adición de cloramfenicol), y se prepararon muestras con o sin cloramfenicol y se cultivaron a 28°C o 42°C. En la diana 8, se produjo ligeramente sólo una colonia sobre una placa con adición de 2-DOG (panel superior izquierdo), se secuenciaron 3 colonias sobre LB (marco rojo) para determinar que la 61<sup>a</sup> citosina (C) se sustituyó por timina (T) en todas las colonias (parte superior derecha). Esta mutación se supone que es insuficiente para perder la función de galK. Por otro lado, en la diana 12, se obtuvo la colonia sobre una placa con adición de 2-DOG mediante cultivo a 28°C y 42°C (panel inferior izquierdo). Se secuenciaron 3 colonias sobre LB para determinar que la 271<sup>a</sup> citosina se sustituyó por timina en todas las colonias (parte inferior derecha). Se demostró que la mutación también puede introducirse de manera estable y muy eficaz en dichas diferentes dianas.

#### [Aplicabilidad industrial]

La presente invención hace posible introducir de manera segura una mutación específica de sitio en cualquier

especie sin la inserción de un ADN foráneo o roturas de ADN bicatenario. Además, es posible establecer una amplia gama de introducción de mutación desde una posición identificada de una base a varios cientos bases, y la técnica también puede aplicarse a la inducción de la evolución oportuna mediante la introducción de mutación aleatoria en una región restringida particular, que ha sido casi imposible hasta ahora, y es sumamente útil.

- 5 **Lista de secuencias**
- <110> NATIONAL UNIVERSITY CORPORATION KOBE UNIVERSITY
- 10 <120> Método de modificación de secuencias genómicas que comprende específicamente la conversión de una nucleobase en una secuencia de ADN seleccionada como diana y complejo molecular usado para ello
- <130> 092301
- 15 <150> Documento JP 2014-043348  
<151> 05-03-2014
- <150> Documento JP 2014-201859  
<151> 30-09-2014
- 20 <160> 42
- <170> PatentIn versión 3.5
- 25 <210> 1  
<211> 624  
<212> ADN  
<213> *Petromyzon marinus*
- 30 <220>  
<221> CDS  
<222> (1)..(624)
- 35 <400> 1

ES 2 752 175 T3

atg acc gac gct gag tac gtg aga atc cat gag aag ttg gac atc tac 48  
 Met Thr Asp Ala Glu Tyr Val Arg Ile His Glu Lys Leu Asp Ile Tyr  
 1 5 10 15

acg ttt aag aaa cag ttt ttc aac aac aaa aaa tcc gtg tcg cat aga 96  
 Thr Phe Lys Lys Gln Phe Phe Asn Asn Lys Lys Ser Val Ser His Arg  
 20 25 30

tgc tac gtt ctc ttt gaa tta aaa cga cgg ggt gaa cgt aga gcg tgt 144  
 Cys Tyr Val Leu Phe Glu Leu Lys Arg Arg Gly Glu Arg Arg Ala Cys  
 35 40 45

ttt tgg ggc tat gct gtg aat aaa cca cag agc ggg aca gaa cgt ggc 192  
 Phe Trp Gly Tyr Ala Val Asn Lys Pro Gln Ser Gly Thr Glu Arg Gly  
 50 55 60

att cac gcc gaa atc ttt agc att aga aaa gtc gaa gaa tac ctg cgc 240  
 Ile His Ala Glu Ile Phe Ser Ile Arg Lys Val Glu Glu Tyr Leu Arg  
 65 70 75 80

gac aac ccc gga caa ttc acg ata aat tgg tac tca tcc tgg agt cct 288  
 Asp Asn Pro Gly Gln Phe Thr Ile Asn Trp Tyr Ser Ser Trp Ser Pro  
 85 90 95

tgt gca gat tgc gct gaa aag atc tta gaa tgg tat aac cag gag ctg 336  
 Cys Ala Asp Cys Ala Glu Lys Ile Leu Glu Trp Tyr Asn Gln Glu Leu  
 100 105 110

cgg ggg aac ggc cac act ttg aaa atc tgg gct tgc aaa ctc tat tac 384  
 Arg Gly Asn Gly His Thr Leu Lys Ile Trp Ala Cys Lys Leu Tyr Tyr  
 115 120 125

gag aaa aat gcg agg aat caa att ggg ctg tgg aac ctc aga gat aac 432  
 Glu Lys Asn Ala Arg Asn Gln Ile Gly Leu Trp Asn Leu Arg Asp Asn  
 130 135 140

ggg gtt ggg ttg aat gta atg gta agt gaa cac tac caa tgt tgc agg 480  
 Gly Val Gly Leu Asn Val Met Val Ser Glu His Tyr Gln Cys Cys Arg  
 145 150 155 160

aaa ata ttc atc caa tcg tcg cac aat caa ttg aat gag aat aga tgg 528  
 Lys Ile Phe Ile Gln Ser Ser His Asn Gln Leu Asn Glu Asn Arg Trp  
 165 170 175

ctt gag aag act ttg aag cga gct gaa aaa cga cgg agc gag ttg tcc 576  
 Leu Glu Lys Thr Leu Lys Arg Ala Glu Lys Arg Arg Ser Glu Leu Ser  
 180 185 190

att atg att cag gta aaa ata ctc cac acc act aag agt cct gct gtt 624  
 Ile Met Ile Gln Val Lys Ile Leu His Thr Thr Lys Ser Pro Ala Val  
 195 200 205

<210> 2

5 <211> 208

<212> PRT

<213> *Petromyzon marinus*

<400> 2

10

ES 2 752 175 T3

Met Thr Asp Ala Glu Tyr Val Arg Ile His Glu Lys Leu Asp Ile Tyr  
 1 5 10 15

Thr Phe Lys Lys Gln Phe Phe Asn Asn Lys Lys Ser Val Ser His Arg  
 20 25 30

Cys Tyr Val Leu Phe Glu Leu Lys Arg Arg Gly Glu Arg Arg Ala Cys  
 35 40 45

Phe Trp Gly Tyr Ala Val Asn Lys Pro Gln Ser Gly Thr Glu Arg Gly  
 50 55 60

Ile His Ala Glu Ile Phe Ser Ile Arg Lys Val Glu Glu Tyr Leu Arg  
 65 70 75 80

Asp Asn Pro Gly Gln Phe Thr Ile Asn Trp Tyr Ser Ser Trp Ser Pro  
 85 90 95

Cys Ala Asp Cys Ala Glu Lys Ile Leu Glu Trp Tyr Asn Gln Glu Leu  
 100 105 110

Arg Gly Asn Gly His Thr Leu Lys Ile Trp Ala Cys Lys Leu Tyr Tyr  
 115 120 125

Glu Lys Asn Ala Arg Asn Gln Ile Gly Leu Trp Asn Leu Arg Asp Asn  
 130 135 140  
 Gly Val Gly Leu Asn Val Met Val Ser Glu His Tyr Gln Cys Cys Arg  
 145 150 155 160

Lys Ile Phe Ile Gln Ser Ser His Asn Gln Leu Asn Glu Asn Arg Trp  
 165 170 175

Leu Glu Lys Thr Leu Lys Arg Ala Glu Lys Arg Arg Ser Glu Leu Ser  
 180 185 190

Ile Met Ile Gln Val Lys Ile Leu His Thr Thr Lys Ser Pro Ala Val  
 195 200 205

5 <210> 3  
 <211> 600  
 <212> ADN  
 <213> *Homo sapiens*

10 <220>  
 <221> CDS  
 <222> (1)..(600)

<400> 3

ES 2 752 175 T3

atg gac agc ctc ttg atg aac cgg agg aag ttt ctt tac caa ttc aaa	48
Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys	
1 5 10 15	
aat gtc cgc tgg gct aag ggt cgg cgt gag acc tac ctg tgc tac gta	96
Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val	
20 25 30	
gtg aag agg cgt gac agt gct aca tcc ttt tca ctg gac ttt ggt tat	144
Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr	
35 40 45	
ctt cgc aat aag aac ggc tgc cac gtg gaa ttg ctc ttc ctc cgc tac	192
Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr	
50 55 60	
atc tcg gac tgg gac cta gac cct ggc cgc tgc tac cgc gtc acc tgg	240
Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp	
65 70 75 80	
ttc acc tcc tgg agc ccc tgc tac gac tgt gcc cga cat gtg gcc gac	288
Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp	
85 90 95	
ttt ctg cga ggg aac ccc tac ctc agt ctg agg atc ttc acc gcg cgc	336
Phe Leu Arg Gly Asn Pro Tyr Leu Ser Leu Arg Ile Phe Thr Ala Arg	
100 105 110	
ctc tac ttc tgt gag gac cgc aag gct gag ccc gag ggg ctg cgg cgg	384
Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg	
115 120 125	
ctg cac cgc gcc ggg gtg caa ata gcc atc atg acc ttc aaa gat tat	432
Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr	
130 135 140	
ttt tac tgc tgg aat act ttt gta gaa aac cat gaa aga act ttc aaa	480
Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys	
145 150 155 160	
gcc tgg gaa ggg ctg cat gaa aat tca gtt cgt ctc tcc aga cag ctt	528
Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu	
165 170 175	
cgg cgc atc ctt ttg ccc ctg tat gag gtt gat gac tta cga gac gca	576
Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala	
180 185 190	
ttt cgt act ttg gga ctt ctc gac	600
Phe Arg Thr Leu Gly Leu Leu Asp	
195 200	

<210> 4  
 5 <211> 200  
 <212> PRT  
 <213> *Homo sapiens*

<400> 4

10

ES 2 752 175 T3

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys  
 1 5 10 15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val  
 20 25 30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr  
 35 40 45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr  
 50 55 60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp  
 65 70 75 80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp  
 85 90 95

Phe Leu Arg Gly Asn Pro Tyr Leu Ser Leu Arg Ile Phe Thr Ala Arg  
 100 105 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg  
 115 120 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr  
 130 135 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys  
 145 150 155 160  
 Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu  
 165 170 175

Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala  
 180 185 190

Phe Arg Thr Leu Gly Leu Leu Asp  
 195 200

- <210> 5
- 5 <211> 4116
- <212> ADN
- <213> Secuencia artificial
- <220>
- 10 <223> CDS de Cas9 derivada de *Streptococcus pyogenes* optimizada para expresión eucariota.
- <220>
- <221> CDS
- 15 <222> (1)..(4116)
- <400> 5

ES 2 752 175 T3

atg gac aag aag tac tcc att ggg ctc gat atc ggc aca aac agc gtc	48
Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val	
1 5 10 15	
ggt tgg gcc gtc att acg gac gag tac aag gtg ccg agc aaa aaa ttc	96
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe	
20 25 30	
aaa gtt ctg ggc aat acc gat cgc cac agc ata aag aag aac ctc att	144
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile	
35 40 45	
ggc gcc ctc ctg ttc gac tcc ggg gag acg gcc gaa gcc acg cgg ctc	192
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu	
50 55 60	
aaa aga aca gca cgg cgc aga tat acc cgc aga aag aat cgg atc tgc	240
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys	
65 70 75 80	
tac ctg cag gag atc ttt agt aat gag atg gct aag gtg gat gac tct	288
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser	
85 90 95	
ttc ttc cat agg ctg gag gag tcc ttt ttg gtg gag gag gat aaa aag	336
Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys	
100 105 110	
cac gag cgc cac cca atc ttt ggc aat atc gtg gac gag gtg gcg tac	384
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr	
115 120 125	
cat gaa aag tac cca acc ata tat cat ctg agg aag aag ctt gta gac	432
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp	

ES 2 752 175 T3

130	135	140	
agt act gat aag gct gac ttg cgg ttg atc tat ctc gcg ctg gcg cat Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His 145 150 155 160			480
atg atc aaa ttt cgg gga cac ttc ctc atc gag ggg gac ctg aac cca Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro 165 170 175			528
gac aac agc gat gtc gac aaa ctc ttt atc caa ctg gtt cag act tac Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr 180 185 190			576
aat cag ctt ttc gaa gag aac ccg atc aac gca tcc gga gtt gac gcc Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala 195 200 205			624
aaa gca atc ctg agc gct agg ctg tcc aaa tcc cgg cgg ctc gaa aac Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn 210 215 220			672
ctc atc gca cag ctc cct ggg gag aag aag aac ggc ctg ttt ggt aat Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn 225 230 235 240			720
ctt atc gcc ctg tca ctc ggg ctg acc ccc aac ttt aaa tct aac ttc Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe 245 250 255			768
gac ctg gcc gaa gat gcc aag ctt caa ctg agc aaa gac acc tac gat Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp 260 265 270			816
gat gat ctc gac aat ctg ctg gcc cag atc ggc gac cag tac gca gac Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp 275 280 285			864
ctt ttt ttg gcg gca aag aac ctg tca gac gcc att ctg ctg agt gat Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp 290 295 300			912
att ctg cga gtg aac acg gag atc acc aaa gct ccg ctg agc gct agt Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser 305 310 315 320			960
atg atc aag cgc tat gat gag cac cac caa gac ttg act ttg ctg aag Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys 325 330 335			1008
gcc ctt gtc aga cag caa ctg cct gag aag tac aag gaa att ttc ttc Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe 340 345 350			1056
gat cag tct aaa aat ggc tac gcc gga tac att gac ggc gga gca agc Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser 355 360 365			1104
cag gag gaa ttt tac aaa ttt att aag ccc atc ttg gaa aaa atg gac Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp 370 375 380			1152
ggc acc gag gag ctg ctg gta aag ctt aac aga gaa gat ctg ttg cgc			1200

ES 2 752 175 T3

Gly 385	Thr	Glu	Glu	Leu	Leu	Val	Lys	Leu	Asn	Arg	Glu	Asp	Leu	Leu	Arg	
					390					395					400	
aaa	cag	cgc	act	ttc	gac	aat	gga	agc	atc	ccc	cac	cag	att	cac	ctg	1248
Lys	Gln	Arg	Thr	Phe	Asp	Asn	Gly	Ser	Ile	Pro	His	Gln	Ile	His	Leu	
				405					410					415		
ggc	gaa	ctg	cac	gct	atc	ctc	agg	cgg	caa	gag	gat	ttc	tac	ccc	ttt	1296
Gly	Glu	Leu	His	Ala	Ile	Leu	Arg	Arg	Gln	Glu	Asp	Phe	Tyr	Pro	Phe	
			420					425					430			
ttg	aaa	gat	aac	agg	gaa	aag	att	gag	aaa	atc	ctc	aca	ttt	cgg	ata	1344
Leu	Lys	Asp	Asn	Arg	Glu	Lys	Ile	Glu	Lys	Ile	Leu	Thr	Phe	Arg	Ile	
		435				440						445				
ccc	tac	tat	gta	ggc	ccc	ctc	gcc	cgg	gga	aat	tcc	aga	ttc	gcg	tgg	1392
Pro	Tyr	Tyr	Val	Gly	Pro	Leu	Ala	Arg	Gly	Asn	Ser	Arg	Phe	Ala	Trp	
	450					455					460					
atg	act	cgc	aaa	tca	gaa	gag	acc	atc	act	ccc	tgg	aac	ttc	gag	gaa	1440
Met	Thr	Arg	Lys	Ser	Glu	Glu	Thr	Ile	Thr	Pro	Trp	Asn	Phe	Glu	Glu	
465					470					475					480	
gtc	gtg	gat	aag	ggg	gcc	tct	gcc	cag	tcc	ttc	atc	gaa	agg	atg	act	1488
Val	Val	Asp	Lys	Gly	Ala	Ser	Ala	Gln	Ser	Phe	Ile	Glu	Arg	Met	Thr	
				485					490					495		
aac	ttt	gat	aaa	aat	ctg	cct	aac	gaa	aag	gtg	ctt	cct	aaa	cac	tct	1536
Asn	Phe	Asp	Lys	Asn	Leu	Pro	Asn	Glu	Lys	Val	Leu	Pro	Lys	His	Ser	
			500					505					510			
ctg	ctg	tac	gag	tac	ttc	aca	gtt	tat	aac	gag	ctc	acc	aag	gtc	aaa	1584
Leu	Leu	Tyr	Glu	Tyr	Phe	Thr	Val	Tyr	Asn	Glu	Leu	Thr	Lys	Val	Lys	
		515				520						525				
tac	gtc	aca	gaa	ggg	atg	aga	aag	cca	gca	ttc	ctg	tct	gga	gag	cag	1632
Tyr	Val	Thr	Glu	Gly	Met	Arg	Lys	Pro	Ala	Phe	Leu	Ser	Gly	Glu	Gln	
	530					535					540					
aag	aaa	gct	atc	gtg	gac	ctc	ctc	ttc	aag	acg	aac	cgg	aaa	gtt	acc	1680
Lys	Lys	Ala	Ile	Val	Asp	Leu	Leu	Phe	Lys	Thr	Asn	Arg	Lys	Val	Thr	
545					550					555				560		
gtg	aaa	cag	ctc	aaa	gaa	gac	tat	ttc	aaa	aag	att	gaa	tgt	ttc	gac	1728
Val	Lys	Gln	Leu	Lys	Glu	Asp	Tyr	Phe	Lys	Lys	Ile	Glu	Cys	Phe	Asp	
				565					570					575		
tct	gtt	gaa	atc	agc	gga	gtg	gag	gat	cgc	ttc	aac	gca	tcc	ctg	gga	1776
Ser	Val	Glu	Ile	Ser	Gly	Val	Glu	Asp	Arg	Phe	Asn	Ala	Ser	Leu	Gly	
			580					585					590			
acg	tat	cac	gat	ctc	ctg	aaa	atc	att	aaa	gac	aag	gac	ttc	ctg	gac	1824
Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp	
		595				600						605				
aat	gag	gag	aac	gag	gac	att	ctt	gag	gac	att	gtc	ctc	acc	ctt	acg	1872
Asn	Glu	Glu	Asn	Glu	Asp	Ile	Leu	Glu	Asp	Ile	Val	Leu	Thr	Leu	Thr	
	610					615					620					
ttg	ttt	gaa	gat	agg	gag	atg	att	gaa	gaa	cgc	ttg	aaa	act	tac	gct	1920
Leu	Phe	Glu	Asp	Arg	Glu	Met	Ile	Glu	Glu	Arg	Leu	Lys	Thr	Tyr	Ala	
625					630					635					640	

ES 2 752 175 T3

cat ctc ttc gac gac aaa gtc atg aaa cag ctc aag agg cgc cga tat	1968
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr	
645 650 655	
aca gga tgg ggg cgg ctg tca aga aaa ctg atc aat ggg atc cga gac	2016
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp	
660 665 670	
aag cag agt gga aag aca atc ctg gat ttt ctt aag tcc gat gga ttt	2064
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe	
675 680 685	
gcc aac cgg aac ttc atg cag ttg atc cat gat gac tct ctc acc ttt	2112
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe	
690 695 700	
aag gag gac atc cag aaa gca caa gtt tct ggc cag ggg gac agt ctt	2160
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu	
705 710 715 720	
cac gag cac atc gct aat ctt gca ggt agc cca gct atc aaa aag gga	2208
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly	
725 730 735	
ata ctg cag acc gtt aag gtc gtg gat gaa ctc gtc aaa gta atg gga	2256
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly	
740 745 750	
agg cat aag ccc gag aat atc gtt atc gag atg gcc cga gag aac caa	2304
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln	
755 760 765	
act acc cag aag gga cag aag aac agt agg gaa agg atg aag agg att	2352
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile	
770 775 780	
gaa gag ggt ata aaa gaa ctg ggg tcc caa atc ctt aag gaa cac cca	2400
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro	
785 790 795 800	
gtt gaa aac acc cag ctt cag aat gag aag ctc tac ctg tac tac ctg	2448
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu	
805 810 815	
cag aac ggc agg gac atg tac gtg gat cag gaa ctg gac atc aat cgg	2496
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg	
820 825 830	
ctc tcc gac tac gac gtg gat cat atc gtg ccc cag tct ttt ctc aaa	2544
Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys	
835 840 845	
gat gat tct att gat aat aaa gtg ttg aca aga tcc gat aaa aat aga	2592
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg	
850 855 860	
ggg aag agt gat aac gtc ccc tca gaa gaa gtt gtc aag aaa atg aaa	2640
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys	
865 870 875 880	
aat tat tgg cgg cag ctg ctg aac gcc aaa ctg atc aca caa cgg aag	2688
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys	
885 890 895	

ES 2 752 175 T3

ttc gat aat ctg act aag gct gaa cga ggt ggc ctg tct gag ttg gat	2736
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp	
900 905 910	
aaa gcc ggc ttc atc aaa agg cag ctt gtt gag aca cgc cag atc acc	2784
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr	
915 920 925	
aag cac gtg gcc caa att ctc gat tca cgc atg aac acc aag tac gat	2832
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp	
930 935 940	
gaa aat gac aaa ctg att cga gag gtg aaa gtt att act ctg aag tct	2880
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser	
945 950 955 960	
aag ctg gtc tca gat ttc aga aag gac ttt cag ttt tat aag gtg aga	2928
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg	
965 970 975	
gag atc aac aat tac cac cat gcg cat gat gcc tac ctg aat gca gtg	2976
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val	
980 985 990	
gta ggc act gca ctt atc aaa aaa tat ccc aag ctt gaa tct gaa ttt	3024
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe	
995 1000 1005	
gtt tac gga gac tat aaa gtg tac gat gtt agg aaa atg atc gca	3069
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala	
1010 1015 1020	
aag tct gag cag gaa ata ggc aag gcc acc gct aag tac ttc ttt	3114
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe	
1025 1030 1035	
tac agc aat att atg aat ttt ttc aag acc gag att aca ctg gcc	3159
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala	
1040 1045 1050	
aat gga gag att cgg aag cga cca ctt atc gaa aca aac gga gaa	3204
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu	
1055 1060 1065	
aca gga gaa atc gtg tgg gac aag ggt agg gat ttc gcg aca gtc	3249
Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val	
1070 1075 1080	
cgg aag gtc ctg tcc atg ccg cag gtg aac atc gtt aaa aag acc	3294
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr	
1085 1090 1095	
gaa gta cag acc gga ggc ttc tcc aag gaa agt atc ctc ccg aaa	3339
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys	
1100 1105 1110	
agg aac agc gac aag ctg atc gca cgc aaa aaa gat tgg gac ccc	3384
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro	
1115 1120 1125	
aag aaa tac ggc gga ttc gat tct cct aca gtc gct tac agt gta	3429
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val	

ES 2 752 175 T3

1130		1135		1140		
ctg gtt	gtg gcc	aaa gtg	gag	aaa ggg	aag tct	aaa aaa ctc aaa
Leu Val	Val Ala	Lys Val	Glu	Lys Gly	Lys Ser	Lys Lys Leu Lys
1145			1150			1155
agc gtc	aag gaa	ctg ctg	ggc	atc aca	atc atg	gag cga tca agc
Ser Val	Lys Glu	Leu Leu	Gly	Ile Thr	Ile Met	Glu Arg Ser Ser
1160			1165			1170
ttc gaa	aaa aac	ccc atc	gac	ttt ctc	gag gcg	aaa gga tat aaa
Phe Glu	Lys Asn	Pro Ile	Asp	Phe Leu	Glu Ala	Lys Gly Tyr Lys
1175			1180			1185
gag gtc	aaa aaa	gac ctc	atc	att aag	ctt ccc	aag tac tct ctc
Glu Val	Lys Lys	Asp Leu	Ile	Ile Lys	Leu Pro	Lys Tyr Ser Leu
1190			1195			1200
ttt gag	ctt gaa	aac ggc	cg	aaa cga	atg ctc	gct agt gcg ggc
Phe Glu	Leu Glu	Asn Gly	Arg	Lys Arg	Met Leu	Ala Ser Ala Gly
1205			1210			1215
gag ctg	cag aaa	ggt aac	gag	ctg gca	ctg ccc	tct aaa tac gtt
Glu Leu	Gln Lys	Gly Asn	Glu	Leu Ala	Leu Pro	Ser Lys Tyr Val
1220			1225			1230
aat ttc	ttg tat	ctg gcc	agc	cac tat	gaa aag	ctc aaa ggg tct
Asn Phe	Leu Tyr	Leu Ala	Ser	His Tyr	Glu Lys	Leu Lys Gly Ser
1235			1240			1245
ccc gaa	gat aat	gag cag	aag	cag ctg	ttc gtg	gaa caa cac aaa
Pro Glu	Asp Asn	Glu Gln	Lys	Gln Leu	Phe Val	Glu Gln His Lys
1250			1255			1260
cac tac	ctt gat	gag atc	atc	gag caa	ata agc	gaa ttc tcc aaa
His Tyr	Leu Asp	Glu Ile	Ile	Glu Gln	Ile Ser	Glu Phe Ser Lys
1265			1270			1275
aga gtg	atc ctc	gcc gac	gct	aac ctc	gat aag	gtg ctt tct gct
Arg Val	Ile Leu	Ala Asp	Ala	Asn Leu	Asp Lys	Val Leu Ser Ala
1280			1285			1290
tac aat	aag cac	agg gat	aag	ccc atc	agg gag	cag gca gaa aac
Tyr Asn	Lys His	Arg Asp	Lys	Pro Ile	Arg Glu	Gln Ala Glu Asn
1295			1300			1305
att atc	cac ttg	ttt act	ctg	acc aac	ttg ggc	gcg cct gca gcc
Ile Ile	His Leu	Phe Thr	Leu	Thr Asn	Leu Gly	Ala Pro Ala Ala
1310			1315			1320
ttc aag	tac ttc	gac acc	acc	ata gac	aga aag	cg
Phe Lys	Tyr Phe	Asp Thr	Thr	Ile Asp	Arg Lys	Arg Tyr Thr Ser
1325			1330			1335
aca aag	gag gtc	ctg gac	gcc	aca ctg	att cat	cag tca att acg
Thr Lys	Glu Val	Leu Asp	Ala	Thr Leu	Ile His	Gln Ser Ile Thr
1340			1345			1350
ggg ctc	tat gaa	aca aga	atc	gac ctc	tct cag	ctc ggt gga gac
Gly Leu	Tyr Glu	Thr Arg	Ile	Asp Leu	Ser Gln	Leu Gly Gly Asp
1355			1360			1365
agc agg	gct gac					
Ser Arg	Ala Asp					
1370						4116

ES 2 752 175 T3

<210> 6  
 <211> 1372  
 <212> PRT  
 5 <213> Secuencia artificial  
  
 <220>  
 <223> Constructo sintético  
  
 10 <400> 6  
  
 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

ES 2 752 175 T3

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				

ES 2 752 175 T3

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

ES 2 752 175 T3

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

ES 2 752 175 T3

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg  
 965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val  
 980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe  
 995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala  
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe  
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala  
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu  
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val  
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr  
 1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys  
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro  
 1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val  
 1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys  
 1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser  
 1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys  
 1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu



ES 2 752 175 T3

<211> 665  
 <212> ADN  
 <213> *Saccharomyces cerevisiae*

5 <400> 8

tttcaaaaat tcttactttt tttttggatg gacgcaaaga agtttaataa tcatattaca	60
tggcattacc accatataca tatccatata catatccata tctaacttta cttatatggt	120
gtggaaatgt aaagagcccc attatccttag cctaaaaaaa ccttctcttt ggaactttca	180
gtaatacgct taactgctca ttgctatatt gaagtacgga ttagaagccg ccgagcgggt	240
gacagccctc cgaaggaaga ctctcctccg tgcgtcctcg tcttcaccgg tcgcgttcct	300
gaaacgcaga tgtgcctcgc gccgcactgc tccgaacaat aaagattcta caatactagc	360
ttttatgggt atgaagagga aaaattggca gtaacctggc cccacaaacc ttcaaatgaa	420
cgaatcaaat taacaacccat aggatgataa tgcgattagt tttttagcct tatttctggg	480
gtaattaatc agcgaagcga tgatttttga tctattaaca gatatataaa tgcaaaaact	540
gcataaccac tttaactaat actttcaaca ttttcggttt gtattacttc ttattcaaat	600
gtaataaaag tatcaacaaa aaattgtaa tatacctcta tactttaacg tcaaggagaa	660
aaaac	665

10 <210> 9  
 <211> 21  
 <212> ADN  
 <213> Secuencia artificial

15 <220>  
 <223> Señal de transición nuclear.

<220>  
 <221> CDS  
 <222> (1)..(21)

20 <400> 9

ccc aag aag aag agg aag gtg	21
Pro Lys Lys Lys Arg Lys Val	
1 5	

25 <210> 10  
 <211> 7  
 <212> PRT  
 <213> Secuencia artificial

30 <220>  
 <223> Constructo sintético

<400> 10

Pro Lys Lys Lys Arg Lys Val
1 5

35 <210> 11  
 <211> 15  
 <212> ADN

ES 2 752 175 T3

<213> Secuencia artificial

<220>

<223> ligador GS

5 <220>

<221> CDS

<222> (1)..(15)

10 <400> 11

ggt gga gga ggt tct  
 Gly Gly Gly Gly Ser  
 1 5

15 <210> 12

<211> 5

<212> PRT

<213> Secuencia artificial

<220>

20 <223> Constructo sintético

<400> 12

Gly Gly Gly Gly Ser  
 1 5

25 <210> 13

<211> 66

<212> ADN

<213> Secuencia artificial

30 <220>

<223> etiqueta Flag

<220>

35 <221> CDS

<222> (1)..(66)

<400> 13

gac tat aag gac cac gac gga gac tac aag gat cat gat att gat tac  
 Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr  
 1 5 10 15

40 aaa gac gat gac gat aag  
 Lys Asp Asp Asp Asp Lys  
 20

<210> 14

<211> 22

<212> PRT

45 <213> Secuencia artificial

<220>

<223> Constructo sintético

50 <400> 14

ES 2 752 175 T3

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

Lys Asp Asp Asp Asp Lys  
 20

5 <210> 15  
 <211> 24  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> etiqueta Strep

<220>  
 <221> CDS  
 <222> (1)..(24)

15 <400> 15

tgg agc cac ccg cag ttc gaa aaa  
 Trp Ser His Pro Gln Phe Glu Lys  
 1 5

24

20 <210> 16  
 <211> 8  
 <212> PRT  
 <213> Secuencia artificial

25 <220>  
 <223> Constructo sintético

<400> 16

Trp Ser His Pro Gln Phe Glu Lys  
 1 5

30 <210> 17  
 <211> 171  
 <212> ADN  
 <213> Secuencia artificial

35 <220>  
 <223> dominio SH3

40 <220>  
 <221> CDS  
 <222> (1)..(171)

<400> 17

gca gag tat gtg cgg gcc ctc ttt gac ttt aat ggg aat gat gaa gaa  
 Ala Glu Tyr Val Arg Ala Leu Phe Asp Phe Asn Gly Asn Asp Glu Glu  
 1 5 10 15

48

45 gat ctt ccc ttt aag aaa gga gac atc ctg aga atc cgg gat aag cct

96

ES 2 752 175 T3

	Asp	Leu	Pro	Phe	Lys	Lys	Gly	Asp	Ile	Leu	Arg	Ile	Arg	Asp	Lys	Pro	
				20					25					30			
	gaa	gag	cag	tgg	tgg	aat	gca	gag	gac	agc	gaa	gga	aag	agg	ggg	atg	144
	Glu	Glu	Gln	Trp	Trp	Asn	Ala	Glu	Asp	Ser	Glu	Gly	Lys	Arg	Gly	Met	
			35					40					45				
	att	cct	gtc	cct	tac	gtg	gag	aag	tat								171
	Ile	Pro	Val	Pro	Tyr	Val	Glu	Lys	Tyr								
		50					55										
	<210>	18															
	<211>	57															
5	<212>	PRT															
	<213>	Secuencia artificial															
	<220>																
	<223>	Constructo sintético															
10	<400>	18															
	Ala	Glu	Tyr	Val	Arg	Ala	Leu	Phe	Asp	Phe	Asn	Gly	Asn	Asp	Glu	Glu	
	1				5					10					15		
	Asp	Leu	Pro	Phe	Lys	Lys	Gly	Asp	Ile	Leu	Arg	Ile	Arg	Asp	Lys	Pro	
				20					25					30			
	Glu	Glu	Gln	Trp	Trp	Asn	Ala	Glu	Asp	Ser	Glu	Gly	Lys	Arg	Gly	Met	
			35					40					45				
	Ile	Pro	Val	Pro	Tyr	Val	Glu	Lys	Tyr								
		50					55										
15	<210>	19															
	<211>	188															
	<212>	ADN															
	<213>	<i>Saccharomyces cerevisiae</i>															
20	<400>	19															
	gcgaatttct	tatgatttat	gatttttatt	attaaataag	ttataaaaaa	aataagtgta	60										
	tacaaatttt	aaagtgactc	ttaggtttta	aaacgaaaat	tcttattctt	gagtaactct	120										
	ttcctgtagg	tcaggttgct	ttctcaggta	tagcatgagg	tcgctcttat	tgaccacacc	180										
	tctaccgg						188										
	<210>	20															
	<211>	417															
25	<212>	ADN															
	<213>	<i>Saccharomyces cerevisiae</i>															
	<400>	20															
30	ataccaggca	tggagcttat	ctgggccggt	cgagttttcg	acgagtttgg	agacattctt	60										
	tatagatgtc	cttttttttt	aatgatattc	gttaaagaac	aaaaagtcaa	agcagtttaa	120										

ES 2 752 175 T3

cctaacacct gttgttgatg ctacttgaaa caaggcttct aggcgaatac ttaaaaaggt 180  
aatttcaata gcggtttata tatctgtttg cttttcaaga tattatgtaa acgcacgatg 240  
tttttcgccc aggctttatt ttttttgttg ttgttgcctt ctccaagaat tttctcgggc 300  
agatctttgt cggaatgtaa aaaagcgcgt aattaaactt tctattatgc tgactaaaat 360  
ggaagtgatc accaaaggct atttctgatt atataatcta gtcattactc gctcgag 417

5 <210> 21  
<211> 33  
<212> ADN  
<213> Secuencia artificial

10 <220>  
<223> ligando de unión a SH3

<220>  
<221> CDS  
<222> (1)..(33)

15 <400> 21

cct cca cct gct ctg cca cct aag aga agg aga 33  
Pro Pro Pro Ala Leu Pro Pro Lys Arg Arg Arg  
1 5 10

20 <210> 22  
<211> 11  
<212> PRT  
<213> Secuencia artificial

25 <220>  
<223> Constructo sintético

<400> 22

Pro Pro Pro Ala Leu Pro Pro Lys Arg Arg Arg  
1 5 10

30 <210> 23  
<211> 269  
<212> ADN  
<213> *Saccharomyces cerevisiae*

35 <400> 23

tctttgaaaa gataatgtat gattatgctt tcaactcatat ttatacagaa acttgatggt 60  
ttctttcgag tatatacaag gtgattacat gtacgtttga agtacaactc tagattttgt 120  
agtgcctct tgggctagcg gtaaagggtgc gcattttttc acaccctaca atgttctggt 180  
caaaagattt tgggtcaaacg ctgtagaagt gaaagttggt gcgcatgttt cggcgttcga 240  
aacttctccg cagtgaaaga taaatgatc 269

40 <210> 24  
<211> 14  
<212> ADN  
<213> *Saccharomyces cerevisiae*

45 <400> 24

	tgtttttat gtct	14
5	<210> 25 <211> 20 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
10	<400> 25 gatacgttct ctatggagga	20
15	<210> 26 <211> 20 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
20	<400> 26 ttggagaaac ccaggtgcct	20
25	<210> 27 <211> 20 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
30	<400> 27 aaccaggtg cctgggtcc	20
35	<210> 28 <211> 20 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
40	<400> 28 ttggccaagt cattcaatt	20
45	<210> 29 <211> 20 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
50	<400> 29 ataacggaat ccaactgggc	20
55	<210> 30 <211> 20 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
60	<400> 30 gtcaattacg aagactgaac	20
65	<210> 31 <211> 19 <212> ADN <213> <i>Saccharomyces cerevisiae</i>	
	<400> 31 gtcaatagga tccccttt	19

ES 2 752 175 T3

<210> 32  
 <211> 10126  
 <212> ADN  
 <213> Secuencia artificial

5 <220>  
 <223> Plásmido que porta proteína de fusión dCas9-PmCDA1 y ARN quimérico que selecciona como diana el gen galK de *E.coli*

10 <220>  
 <221> misc\_feature  
 <222> (5561)..(5580)  
 <223> n e s a , c , g , o t

15 <400> 32

atcgccattc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg cgggcctctt	60
cgctattacg ccagctggcg aaaggggat gtgctgcaag gcgattaagt tgggtaacgc	120
cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag ctccggtacct	180
ggccgcaaac aacagataaa acgaaaggcc cagtctttcg actgagcctt tcgttttatt	240
tgatgcctgt caagtaacag caggactctt agtgggtgtg agtattttta cctgaatcat	300
aatggacaac tcgctccgtc gtttttcagc tcgcttcaaa gtcttctcaa gccatctatt	360
ctcattcaat tgattgtgcg acgattggat gaatattttc ctgcaacatt ggtagtgttc	420
acttaccatt acattcaacc caaccccgtt atctctgagg ttccacagcc caatttgatt	480
cctcgcattt ttctcgtaat agagtttgca agcccagatt ttcaaagtgt ggccgttccc	540
ccgcagctcc tggttatacc attctaagat cttttcagcg caatctgcac aaggactcca	600
ggatgagtac caatttatcg tgaattgtcc ggggttgtcg cgcaggtatt cttcgacttt	660
tctaattgta aagatttcg cgtgaatgcc acgttctgtc ccgctctgtg gtttattcac	720
agcatagccc caaaaacacg ctctacgttc accccgtcgt ttaattcaa agagaacgta	780
gcatctatgc gacacggatt ttttgttgtt gaaaaactgt ttcttaaacg ttagatgtc	840
caacttctca tggattctca cgtactcagc gtcggtcatc ctgacttat cgtcatcgtc	900
tttgtaatca atatcatgat ccttgtagtc tccgtcgtgg tccttatagt ctccggactc	960
gagcctagac ttatcgtcat cgtctttgta atcaatatca tgatccttgt agtctccgtc	1020
gtggtcctta tagtctccg aatacttctc cacgtaagg acaggaatca tccccctctt	1080
tccttcgctg tcctctgcat tccaccactg ctccctcaggc ttatcccgga ttctcaggat	1140
gtctcctttc ttaaaggga gatcctcttc atcattccca ttaaagtcaa agagggctcg	1200

ES 2 752 175 T3

cacatactca gcagaacctc cacctccaga acctcctcca ccgtcacctc ctagctgact 1260  
caaatcaatg cgtgtttcat aaagaccagt gatggattga tggataagag tggcatctaa 1320  
aacttctttt gtagacgtat atcgtttacg atcaattggt gtatcaaaaat atttaaaagc 1380  
agcgggagct ccaagattcg tcaacgtaaa taaatgaata atattttctg cttgttcacg 1440  
tattggtttg tctctatggt tgttatatgc actaagaact ttatctaaat tggcatctgc 1500  
taaaataaca cgcttagaaa attcactgat ttgctcaata atctcatcta aataatgctt 1560  
atgctgctcc acaacaatt gtttttgttc gttatcttct ggactaccct tcaacttttc 1620  
ataatgacta gctaaatata aaaaattcac atatttgctt ggagagcca gctcatttcc 1680  
tttttgtaat tctccggcac tagccagcat ccgtttacga ccgttttcta actcaaaaag 1740  
actatattta ggtagtttaa tgattaagtc ttttttaact tccttatatc ctttagcttc 1800  
taaaaagtca atcggatttt tttcaaagga acttctttcc ataattgtga tccctagtaa 1860  
ctctttaacg gattttaact tcttcgattt ccctttttcc accttagcaa cactaggac 1920  
tgaataagct accgttggac tatcaaaacc accatatttt tttggatccc agtctttttt 1980  
acgagcaata agcttgcctg aatttctttt tggtaaaatt gactccttgg agaatccgcc 2040  
tgtctgtact tctgttttct tgacaatatt gacttggggc atggacaata ctttgcgcac 2100  
tgtggcaaaa tctcgcctt tatcccagac aatttctcca gtttcccat tagtttcgat 2160  
tagagggcgt ttgcgaatct ctccatttgc aagtgttaatt tctgttttga agaagttcat 2220  
gatattagag taaaagaaat attttgcggt tgctttgcct atttcttgcct cagacttagc 2280  
aatcatttta cgaacatcat aaactttata atcacatag acaaactccg attcaagttt 2340  
tggatatttc ttaatcaaag cagttccaac gacggcattt agatacgcac catgggcatg 2400  
atggtaattg ttaatctcac gtactttata gaattggaaa tcttttcgga agtcagaaac 2460  
taatttagat tttaaggtaa tcaacttaac ctctcgaata agtttatcat tttcatcgta 2520  
tttagtattc atgcgactat ccaaaatttg tgccacatgc ttagtgattt ggcgagtttc 2580  
aaccaattgg cgtttgataa aaccagcttt atcaagttca ctcaaacctc cacgttcagc 2640  
tttcgttaaa ttatcaaact tacgttgagt gattaacttg gcgttttagaa gttgtctcca 2700  
atagtttttc atctttttga ctacttcttc acttggaacg ttatccgatt taccacgatt 2760  
tttatcagaa cgcgttaaga ccttattgtc tattgaatcg tctttaagga aactttgtgg 2820  
aacaatggca tcgacatcat aatcacttaa acgattaata tctaattctt ggtccacata 2880  
catgtctctt ccattttgga gataatagag atagagcttt tcattttgca attgagtatt 2940  
ttcaacagga tgctctttaa gaatctgact tcctaattct ttgatacctt cttogattcg 3000  
tttcatacgc tctcgcgaat ttttctggcc cttttgagtt gtctgatttt cacgtgccat 3060  
ttcaataacg atattttctg gcttatgccc cccattact ttgaccaatt catcaacaac 3120

ES 2 752 175 T3

ttttacagtc tgtaaaatac cttttttaat agcagggcta ccagctaaat ttgcaatatg 3180  
 ttcattgtaa ctatogcctt gtccagacac ttgtgctttt tgaatgtctt ctttaaagt 3240  
 caaactatca tcatggatca gctgcataaa attgcgattg gcaaaacat ctgatttcaa 3300  
 aaaatcta attggtttgc cagattgctt atccctaata ccattaatca attttcgaga 3360  
 caaacgtccc caaccagtat aacggcgacg ttttaagctgt ttcattcacct tatcatcaaa 3420  
 gaggtgagca tatgttttaa gtctttcctc aatcatctcc ctatcttcaa ataagggtcaa 3480  
 tgttaaaaca atatcctcta agatatcttc attttcttca ttatccaaaa aatctttatc 3540  
 ttttaataatt tttagcaaat catggtaggt acctaataaa gcattaaatc tatcttcaac 3600  
 tcttgaaatt tcaaacactat caaaacattc tatttttttg aaataatctt cttttaattg 3660  
 cttaacgggt acttttcgat ttgttttgaa gagtaaatca acaatggctt tcttctgttc 3720  
 acctgaaaga aatgctgggt ttgcattcc ttcagtaaca tatttgacct ttgtcaattc 3780  
 gttataaacc gtaaaatact cataaagcaa actatgtttt ggtagtactt tttcatttgg 3840  
 aagattttta tcaaagtttg tcatgcgctc aataaatgat tgagctgaag cacctttatc 3900  
 gacaacttct tcaaaattcc atggggtaat tgtttcttca gacttccgag tcatccatgc 3960  
 aaaacgacta ttgccacgcg ccaatggacc aacataataa ggaattcgaa aagtcaagat 4020  
 tttttcaatc ttctcacgat tgtcttttaa aaatggataa aagtcttctt gtcttctcaa 4080  
 aatagcatgc agctcaccca agtgaatttg atggggaata gagccgttgt caaagggtccg 4140  
 ttgcttgccg agcaaatctt cacgatttag tttcaccaat aattcctcag taccatccat 4200  
 tttttctaaa attggtttga taaatttata aaattcttct tggctagctc ccccatcaat 4260  
 ataacctgca tatcogtttt ttgattgatc aaaaaagatt tctttatact tttctggaag 4320  
 ttgttgtoga actaaagctt ttaaaagagt caagtcttga tgatgttcat cgtagcgttt 4380  
 aatcattgaa gctgataggg gagccttagt tatttcagta tttactotta ggatatctga 4440  
 aagtaaaata gcatctgata aattcttagc tgccaaaaac aaatcagcat attgatctcc 4500  
 aatttgogcc aataaattat ctaaactatc atcgtaagta tcttttgaaa gctgtaattt 4560  
 agcatcttct gccaaatcaa aatttgattt aaaattaggg gtcaaacca atgacaaagc 4620  
 aatgagattc ccaaataagc ctttttctt ctcaccgggg agctgagcaa tgagattttc 4680  
 taatogtctt gatttactca atcgtgcaga aagaatcgtt ttagcatcta ctccacttgc 4740  
 gttaataggg ttttcttcaa ataattgatt gtaggtttgt accaactgga taaatagttt 4800  
 gtccacatca ctattatcag gatttaaatc tccctcaatc aaaaaatgac cacgaaactt 4860  
 aatcatatgc gctaaggcca aatagattaa ggcgaaatcc gctttatcag tagaatctac 4920  
 caattttttt cgcagatgat agatagttgg atatttctca tgataagcaa ctcatctac 4980

ES 2 752 175 T3

tatatttcca aaaataggat gacgttcatg cttcttgtct tcttccacca aaaaagactc 5040  
 ttcaagtcga tgaagaaac tatcatctac tttcgccatc tcatttgaaa aaatctcctg 5100  
 tagataacaa atacgattct tccgacgtgt ataccttcta cgagctgtcc gtttgagacg 5160  
 agtcgcttcc gctgtctctc cactgtcaaa taaaagagcc cctataagat tttttttgat 5220  
 actgtggcgg tctgtatttc ccagaacctt gaacttttta gacggaacct tatattcatc 5280  
 agtgatcacc gcccatccga cgctatttgt gccgatagct aagcctattg agtatttctt 5340  
 atccattttt gcctcctaaa atgggccctt taaattaaat ccataatgag tttgatgatt 5400  
 tcaataatag ttttaatgac ctccgaaatt agtttaatat gctttaattt ttctttttca 5460  
 aaatatctct tcaaaaaata ttaccaata cttaataata aatagattat aacacaaaat 5520  
 tcttttgaca agtagtttat tttgttataa ttctatagta nnnnnnnnnn nnnnnnnnnn 5580  
 gttttagagc tagaaatagc aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt 5640  
 ggcaccgagt cgggtgcttt tttgatactt ctattctact ctgactgcaa accaaaaaaa 5700  
 caagcgcttt caaacgctt gttttatcat ttttagggaa attaactctt taatcctttt 5760  
 atcattctac attaggcgc tgccatcttg ctaaacctac taagctccac aggatgattt 5820  
 cgtaatcccg caagaggccc ggcagtaccg gcataaccaa gcctatgcct acagcatcca 5880  
 gggtgacggg gccgaggatg acgatgagcg cattgttaga tttcatcac ggtgcctgac 5940  
 tgcgttagca atttaactgt gataaaactac cgcattaaag cttatcgatg ataagctgtc 6000  
 aaacatgaga attacaactt atatcgtatg gggctgactt caggtgctac atttgaagag 6060  
 ataaattgca ctgaaatcta gtcggatcct cgctcactga ctcgctgcgc tcggctgctt 6120  
 ggctgcggcg agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag 6180  
 gggataacgc aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa 6240  
 aggccgcggt gctggcgttt ttccatagge tccgcccccc tgacgagcat cacaaaaatc 6300  
 gacgctcaag tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc 6360  
 ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc gcttaccgga tacctgtccg 6420  
 cctttctccc ttcgggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt 6480  
 cgggtgtaggt cgttcgctcc aagctgggct gtgtgcacga acccccgtt cagcccgacc 6540  
 gctgcgcctt atccgtaac tatcgtcttg agtccaacct ggtaagacac gacttatcgc 6600  
 cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 6660  
 agttcttgaa gtgggtggcct aactacggct aactagaag gacagtattt ggtatctgcg 6720  
 ctctgctgaa gccagttacc ttcggaaaaa gagttgtag ctcttgatcc ggcaaaaaa 6780  
 ccaccgctgg tagcgggtgt ttttttgtt gcaagcagca gattacgcgc agaaaaaag 6840  
 gatctcaaga agatcctttg atcttttcta cggggctctga cgctcagtgg aacgaaaact 6900

ES 2 752 175 T3

cacgттаagg gatttttggtc atgagattat caaaaaggat cttcacctag atccttttaa 6960  
attaanaatg aagtttttaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 7020  
accaatgctt aatcagtgag gcacctatct cagegatctg tctatttcgt tcatccatag 7080  
ttgcctgact ccccgctgtg tagataacta cgatacggga gggcttacca tctggcccca 7140  
gtgctgcaat gataccgcga gaaccacgct caccggctcc agatttatca gcaataaacc 7200  
agccagccgg aagggccgag cgcagaagtg gtctgcaac tttatccgcc tccatccagt 7260  
ctattaattg ttgcccggaa gctagagtaa gtagttcgcc agttaatagt ttgcccgaacg 7320  
ttgttgccat tgctgcaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca 7380  
gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 7440  
ttagctcctt cggctcctcg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca 7500  
tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 7560  
tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcccgga ccgagttgct 7620  
cttgcccggc gtcaacacgg gataataccg cgcacatag cagaacttta aaagtgctca 7680  
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 7740  
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg 7800  
tttctgggtg agcaaaaaaca ggaaggcaaa atgccgcaaa aaaggaata agggcgacac 7860  
ggaaatgttg aataactcata ctcttccttt ttcaatatta ttgaagcatt tatcaggggtt 7920  
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaaca ataggggttc 7980  
cgcgcacatt tccccgaaa gtgccacctg acgtcaatgc cgagcgaaag cgagccgaag 8040  
ggtagcattt acgttagata acccctgat atgctccgac gctttatata gaaaagaaga 8100  
ttcaactagg taaaatctta atataggttg agatgataag gtttataagg aatttgtttg 8160  
ttctaatttt tcaactcatt tgttctaatt tcttttaaca aatgttcttt tttttttaga 8220  
acagttatga tatagttaga atagtttaa ataaggagtg agaaaaagat gaaagaaaga 8280  
tatggaacag tctataaagg ctctcagagg ctcatagacg aagaaagtgg agaagtcata 8340  
gaggtagaca agttataccg taaacaaacg tctggtaact tcgtaaaggc atatatagtg 8400  
caattaataa gtatgttaga tatgattggc ggaaaaaac ttaaatcgt taactatctc 8460  
ctagataatg tccacttaag taacaataca atgatagcta caacaagaga aatagcaaaa 8520  
gctacaggaa caagtctaca aacagtaata acaacactta aatctttaga agaaggaaat 8580  
attataaaaa gaaaaactgg agtattaatg ttaaaccctg aactactaat gagaggcgac 8640  
gacaaaaaac aaaaatacct ctactcgaa tttgggaact ttgagcaaga ggcaaatgaa 8700  
atagattgac ctccaataa caccacgtag ttattgggag gtcaatctat gaaatgcgat 8760

ES 2 752 175 T3

taagcttttt ctaattcaca taagcgtgca ggtttaaagt acataaaaa tataatgaaa 8820  
 aaaagcatca ttatactaac gttataccaa cattatactc tcattatact aattgcttat 8880  
 tccaatttcc tattggttgg aaccaacagg cgttagtgtg ttgttgagtt ggtactttca 8940  
 tgggattaat cccatgaaac cccaaccaa ctgcgcaaag ctttggctaa cacacacgcc 9000  
 attccaacca atagttttct cggcataaag ccatgctctg acgcttaaat gcactaatgc 9060  
 cttaaaaaaa cattaagtc taacacacta gacttattta cttcgttaatt aagtcgtaa 9120  
 accgtgtgct ctacgaccaa aagtataaaa cctttaagaa ctttcttttt tcttgtaaaa 9180  
 aaagaaacta gataaatctc tcatatcttt tattcaataa tcgcatcaga ttgcagtata 9240  
 aatttaacga tcaactcatca tgttcatatt tatcagagct cgtgctataa ttatactaat 9300  
 tttataagga ggaaaaaata aagagggtta taatgaacga gaaaaatata aaacacagtc 9360  
 aaaactttat tacttcaaaa cataatatag ataaaataat gacaaatata agattaaatg 9420  
 aacatgataa tatctttgaa atcggctcag gaaaaggga tttaccctt gaattagtag 9480  
 agaggtgtaa tttcgttaact gccattgaaa tagaccataa attatgcaa actacagaaa 9540  
 ataaacttgt tgatcacgat aatttccaag ttttaacaa ggatatattg cagtttaaat 9600  
 ttcttaaaaa ccaatcctat aaaatatttg gtaatatacc ttataacata agtacggata 9660  
 taatacgsaa aattgttttt gatagtatag ctgatgagat ttatttaatc gtggaatacg 9720  
 ggtttgctaa aagattatta aatacaaaac gctcattggc attattttta atggcagaag 9780  
 ttgatatttc tatattaagt atggttcaa gagaatattt tcatcctaaa cctaaagtga 9840  
 atagctcact tatcagatta aatagaaaaa aatcaagaat atcacacaaa gataaacaga 9900  
 agtataatta tttcgttatg aaatgggtta acaagaata caagaaaata tttacaaaaa 9960  
 atcaatttaa caattcctta aaacatgcag gaattgacga tttaaacaat attagctttg 10020  
 aacaattctt atctcttttc aatagctata aattatttaa taagtaagtt aagggatgca 10080  
 taaactgcat cccttaactt gtttttcgtg tacctatttt ttgtga 10126

<210> 33  
 <211> 20  
 <212> ADN  
 <213> *Escherichia coli*

5

<400> 33

10 tcaatgggct aactacgtc 20

<210> 34  
 <211> 20  
 <212> ADN  
 <213> *Escherichia coli*

15

<400> 34

ggtccataaa ctgagacagc 20

ES 2 752 175 T3

<210> 35  
 <211> 223  
 <212> ADN  
 5 <213> *Escherichia coli*  
 <400> 35  
 gagttataca cagggctggg atctattcct tttatctttt tttattcttt ctttattcta 60  
 taaattataa ccacttgaat ataaacaaaa aaaacacaca aaggtctagc ggaatttaca 120  
 gagggctctag cagaatttac aagttttcca gcaaaggtct agcagaattt acagatacc 180  
 acaactcaaa ggaaaaggac tagtaattat cattgactag ccc 223  
 10  
 <210> 36  
 <211> 951  
 <212> ADN  
 <213> *Escherichia coli*  
 15 <400> 36  
 atgtctgaat tagttgtttt caaagcaaat gaactagcga ttagtcgcta tgacttaacg 60  
 gagcatgaaa ccaagcta atttatgctgt gtggcactac tcaacccac gattgaaaac 120  
 cctacaatga aagaacggac ggtatcgctc acttataacc aatacgttca gatgatgaac 180  
 atcagtaggg aaaatgctta tgggtgatta gctaaagcaa ccagagagct gatgacgaga 240  
 actgtggaaa tcaggaatcc tttggttaaa ggctttgaga ttttccagtg gacaaactat 300  
 gccaaagtct caagcgaaaa attagaatta gtttttagtg aagagatatt gccttatctt 360  
 ttccagttaa aaaaattcat aaaatataat ctggaacatg ttaagtcttt tgaaaacaaa 420  
 tactctatga ggatttatga gtggttatta aaagaactaa cacaaaagaa aactcacaag 480  
 gcaaatatag agattagcct tgatgaattt aagttcatgt taatgcttga aaataactac 540  
 catgagttta aaaggcttaa ccaatgggtt ttgaaaccaa taagtaaaga tttaaactact 600  
 tacagcaata tgaaattggt ggttgataag cgaggccgcc cgactgatac gttgattttc 660  
 caagttgaac tagatagaca aatggatctc gtaaccgaac ttgagaacaa ccagataaaa 720  
 atgaatggtg acaaaaatacc aacaaccatt acatcagatt cctacctaca taacggacta 780  
 agaaaaacac tacacgatgc ttttaactgca aaaattcagc tcaccagttt tgaggcaaaa 840  
 tttttgagtg acatgcaaag taagcatgat ctcaatggtt cgttctcatg gctcacgcaa 900  
 aaacaacgaa ccacactaga gaacatactg gctaaatacg gaaggatctg a 951  
 20  
 <210> 37  
 <211> 714  
 <212> ADN  
 <213> Bacteriófago lambda  
 25 <400> 37  
 tcagccaaac gtctcttcag gccactgact agcgataact ttccccacaa cggaacaact 60

ES 2 752 175 T3

ctcattgcat gggatcattg ggtactgtgg gtttagtggt tgtaaaaaca cctgaccgct 120  
atccctgatac agtttcttga aggtaaactc atcccccca agtctggcta tgcagaaatc 180  
acctggctca acagcctgct cagggtcaac gagaattaac attccgctcag gaaagcttgg 240  
cttgagacct gttggtgctg tcatggaatt accttcaacc tcaagccaga atgcagaatc 300  
actggctttt ttggttgtgc ttacccatct ctccgcatca ctttggtaa aggttctaag 360  
cttaggtgag aacatccctg cctgaacatg agaaaaaaca ggtactcat actcacttct 420  
aagtgacggc tgcatactaa ccgcttcata catctcgtag atttctctgg cgattgaagg 480  
gctaaattct tcaacgctaa ctttgagaat ttttgtaagc aatgcggcgt tataagcatt 540  
taatgcattg atgccattaa ataaagcacc aacgcctgac tgccccatcc ccatcttctc 600  
tgcgacagat tcctgggata agccaagttc attttcttt ttttcataaa ttgctttaag 660  
gcgacgtgcg tcctcaagct gctcttgtgt taatggtttc tttttgtgc tcat 714

<210> 38  
<211> 5097  
5 <212> ADN  
<213> Secuencia artificial

<220>  
10 <223> dCas9-PmCDA1

<400> 38

atggataaga aataactcaat aggcttagct atcggcacia atagcgtcgg atgggagggtg 60  
atcactgatg aatataaggc tccgtctaaa aagttcaagg ttctgggaaa tacagaccgc 120  
cacagtatca aaaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa 180  
gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt 240  
tatctacagg agattttttc aaatgagatg gcgaaagtag atgatagttt ctttcatcga 300  
cttgaagagt cttttttggt ggaagaagac aagaagcatg aacgtcatcc tatttttgga 360  
aatatagtag atgaagttgc ttatcatgag aatatccaa ctatctatca tctgcaaaaa 420  
aaattggtag attctactga taaagcggat ttgctgctaa tctatttggc cttagcgcac 480  
atgattaagt ttcgtggtca ttttttgatt gagggagatt taaatcctga taatagtgat 540  
gtggacaaac tatttatcca gttggtacaa acctacaatc aattatttga agaaaaccct 600  
attaacgcaa gtggagtaga tgctaaagcg attctttctg cacgattgag taaatcaaga 660  
cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggctt atttgggaat 720  
ctcattgctt tgtcattggg tttgaccctt aattttaaat caaattttga tttggcagaa 780  
gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg 840  
caaattggag atcaatatgc tgatttgttt ttggcagcta agaatttatc agatgctatt 900

ES 2 752 175 T3

ttactttcag	atatacctaag	agtaaataact	gaaataacta	aggctcccct	atcagcttca	960
atgattaaac	gctacgatga	acatcatcaa	gacttgactc	ttttaaagc	tttagttcga	1020
caacaacttc	cagaaaagta	taaagaaatc	ttttttgatc	aatcaaaaaa	cggatatgca	1080
ggttatattg	atgggggagc	tagccaagaa	gaattttata	aatttatcaa	accaatttta	1140
gaaaaaatgg	atgggtactga	ggaattattg	gtgaaactaa	atcgtgaaga	tttgctgcgc	1200
aagcaacgga	cctttgacaa	cggctctatt	ccccatcaa	ttcacttggg	tgagctgcat	1260
gctattttga	gaagacaaga	agacttttat	ccatttttaa	aagacaatcg	tgagaagatt	1320
gaaaaaatct	tgacttttcg	aattccttat	tatgttggtc	cattggcgcg	tggaatagt	1380
cgttttgcat	ggatgactcg	gaagtctgaa	gaaacaatta	ccccatggaa	ttttgaagaa	1440
gttgctgata	aaggtgcttc	agctcaatca	tttattgaac	gcatgacaaa	ctttgataaa	1500
aatcttccaa	atgaaaaagt	actacaaaa	catagtttgc	tttatgagta	ttttacggtt	1560
tataacgaat	tgacaaaggt	caaatatggt	actgaaggaa	tgcgaaaacc	agcatttctt	1620
tcaggtgaac	agaagaaagc	cattgttgat	ttactcttca	aaacaaatcg	aaaagtaacc	1680
gtaagcaat	taaaagaaga	ttatttcaaa	aaaatagaat	gttttgatag	tggtgaaatt	1740
tcaggagttg	aagatagatt	taatgcttca	ttaggtacct	accatgattt	gctaaaaatt	1800
attaagata	aagatttttt	ggataatgaa	gaaaatgaag	atatacttaga	ggatattggt	1860
ttaacattga	ccttatttga	agatagggag	atgattgagg	aaagacttaa	aacatagct	1920
cacctctttg	atgataaggt	gatgaaacag	cttaaacgtc	gccgttatac	tggttgggga	1980
cgtttgtctc	gaaaattgat	taatggtatt	agggataagc	aatctggcaa	aacaatatta	2040
gattttttga	aatcagatgg	ttttgccaat	cgcaatttta	tgacagctgat	ccatgatgat	2100
agtttgacat	ttaaagaaga	cattcaaaaa	gcacaagtgt	ctggacaagg	cgatagttta	2160
catgaacata	ttgcaaattt	agctggttagc	cctgctatta	aaaaaggtat	tttacagact	2220
gtaaaagttg	ttgatgaatt	ggtcaaagta	atggggcggc	ataagccaga	aaatatcggt	2280
attgaaatgg	cacgtgaaaa	tcagacaact	caaaagggcc	agaaaaattc	gcgagagcgt	2340
atgaaacgaa	tcgaagaag	tatcaaagaa	ttaggaagtc	agattcttaa	agagcatcct	2400
gttgaaaata	ctcaattgca	aaatgaaaag	ctctatctct	attatctcca	aaatggaaga	2460
gacatgtatg	tggaccaaga	attagatatt	aatcgtttta	gtgattatga	tgtcgatgcc	2520
attgttccac	aaagtttctt	taaagaogat	tcaatagaca	ataaggtcct	aacgcgttct	2580
gataaaaatc	gtggtaaadc	ggataacggt	ccaagtgaag	aagtagtcaa	aaagatgaaa	2640
aactattgga	gacaacttct	aaacgccaa	ttaatcactc	aacgtaagtt	tgataattta	2700
acgaaagctg	aacgtggagg	tttgagtgaa	cttgataaag	ctggttttat	caaacgccaa	2760
ttggttgaaa	ctcgccaaat	cactaagcat	gtggcacaaa	ttttggatag	tcgcatgaat	2820

ES 2 752 175 T3

actaaatacg atgaaaatga taaacttatt cgagaggtta aagtgattac cttaaaatct 2880  
aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat 2940  
taccatcatg cccatgatgc gtatctaaat gccgtcgttg gaactgcttt gattaagaaa 3000  
tatccaaaac ttgaatcgga gtttgtctat ggtgattata aagtttatga tgttcgtaaa 3060  
atgattgcta agtctgagca agaaataggc aaagcaaccg caaaatattt cttttactct 3120  
aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaaccg 3180  
cctctaatac gaaactaatg ggaaactgga gaaattgtct gggataaagg gcgagatttt 3240  
gccacagtgc gcaaagtatt gtccatgcc caagtcaata ttgtcaagaa aacagaagta 3300  
cagacaggcg gattctccaa ggagtcaatt ttacaaaaa gaaattcgga caagcttatt 3360  
gctcgtaaaa aagactggga tccaaaaaaa tatggtggtt ttgatagtcc aacggtagct 3420  
tattcagttc tagtggttgc taaggtggaa aaagggaaat cgaagaagtt aaaatccggt 3480  
aaagagttac tagggatcac aattatggaa agaagttcct ttgaaaaaaaa tccgattgac 3540  
tttttagaag ctaaaggata taaggaagtt aaaaaagact taatcattaa actacctaaa 3600  
tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaatta 3660  
caaaaaggaa atgagctggc tctgccaaagc aaatatgtga attttttata tttagctagt 3720  
cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780  
cagcataagc attatttaga tgagattatt gagcaaatca gtgaattttc taagcgtggt 3840  
attttagcag atgccaatth agataaagtt cttagtgcac ataacaaaca tagagacaaa 3900  
ccaatacgtg aacaagcaga aaatattatt catttattta cgttgacgaa tcttgagct 3960  
cccgctgctt ttaaataatt tgatacaaca attgatcgt aacgatatac gtctacaaaa 4020  
gaagttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080  
gatttgagtc agctaggagg tgacggtgga ggaggttctg gaggtggagg ttctgctgag 4140  
tatgtgcgag ccctctttga ctttaatggg aatgatgaag aggatcttcc ctttaagaaa 4200  
ggagacatcc tgagaatccg ggataagcct gaggagcagt ggtggaatgc agaggacagc 4260  
gaaggaaaga ggggatgat tccgtgcoct tacgtggaga agtattccgg agactataag 4320  
gaccacgacg gagactacaa ggatcatgat attgattaca aagacgatga cgataagtct 4380  
aggctcgagt ccggagacta taaggaccac gacggagact acaaggatca tgatattgat 4440  
tacaagacg atgacgataa gtctaggatg accgacgctg agtacgtgag aatccatgag 4500  
aagttggaca tctacacggt taagaaacag tttttcaaca acaaaaaatc cgtgtcgcac 4560  
agatgctacg ttctctttga attaaaacga cggggtgaac gtagagcgtg tttttggggc 4620  
tatgctgtga ataaaccaca gagcgggaca gaacgtggca ttcacgccga aatctttagc 4680

ES 2 752 175 T3

	attagaaaag tcgaagaata cctgcgcgac aacccccggac aattcacgat aaattggtac	4740
	tcatacctgga gtccttgtgc agattgcgct gaaaagatct tagaatggta taaccaggag	4800
	ctgccccggga acggccacac tttgaaaatc tgggcttgca aactctatta cgagaaaaat	4860
	gcgaggaatc aaattgggct gtggaacctc agagataacg gggttgggtt gaatgtaatg	4920
	gtaagtgaac actaccaatg ttgcaggaaa atattcatcc aatcgtcgca caatcaattg	4980
	aatgagaata gatggcttga gaagactttg aagcgagctg aaaaacgacg gagcgagttg	5040
	tccattatga ttcaggtaaa aatactccac accactaaga gtcctgctgt tacttga	5097
	<210> 39	
	<211> 105	
5	<212> ADN	
	<213> <i>Escherichia coli</i>	
	<400> 39	
	acgttaaatc tatcaccgca agggataaat atctaacacc gtgctgttg actattttac	60
10	ctctggcggg gataatgggt gcagggccca ttttaggagg caaaa	105
	<210> 40	
	<211> 247	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> ARNg	
20	<400> 40	
	ggtttagcaa gatggcagcg cctaaatgta gaatgataaa aggattaaga gattaatttc	60
	cctaaaaatg ataaaaacaag cgttttgaaa gcgcttgttt ttttggttg cagtcagagt	120
	agaatagaag tatcaaaaaa agcaccgact cggtgccact ttttcaagtt gataacggac	180
	tagccttatt ttaacttgct atttctagct ctaaaactga gaccatccc ggtctctact	240
	gcagaat	247
25	<210> 41	
	<211> 64	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
	<400> 41	
30	tatcaccgcc agtgggtattt atgtcaacac cgccagagat aatttatcac cgcagatggt	60
	tatc	64
35	<210> 42	
	<211> 10867	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<223> Plásmido	

ES 2 752 175 T3

<400> 42

gtcggaaactg actaaagtag tgagttatac acagggctgg gatctattct ttttatcttt	60
ttttattctt tctttattct ataaattata accacttgaa tataaacaaa aaaaacacac	120
aaaggtctag cggaatttac agaggttcta gcagaattta caagttttcc agcaaaggtc	180
tagcagaatt tacagatacc cacaactcaa aggaaaagga ctagtaatta tcattgacta	240
gccccatctca attggtatag tgattaaaat cacctagacc aattgagatg tatgtctgaa	300
ttagttgttt tcaaagcaaa tgaactagcg attagtcgct atgacttaac ggagcatgaa	360
accaagctaa ttttatgctg tgtggcacta ctcaaccca cgattgaaaa ccctacaatg	420
aaagaacgga cggtatcgtt cacttataac caatacgttc agatgatgaa catcagtagg	480
gaaaatgctt atggtgtatt agctaaagca accagagagc tgatgacgag aactgtggaa	540
atcaggaatc ctttggttaa aggctttgag attttccagt ggacaaacta tgccaagttc	600
tcaagcgaaa aattagaatt agtttttagt gaagagatat tgccttatct tttccagtta	660
aaaaaattca taaaatataa tctggaacat gttaagtctt ttgaaaaca atactctatg	720
aggatttatg agtggttatt aaaagaacta acacaaaaga aaactcaca ggcaaatata	780
gagattagcc ttgatgaatt taagttcatg ttaatgcttg aaaataacta ccatgagttt	840
aaaaggctta accaatgggt tttgaaacca ataagtaaag atttaaacac ttacagcaat	900
atgaaattgg tggttgataa gcgaggccgc ccgactgata cgttgatttt ccaagttgaa	960
ctagatagac aaatggatct cgtaaccgaa cttgagaaca accagataaa aatgaatggt	1020
gacaaaatac caacaacctat tacatcagat tcctacctac ataacggact aagaaaaaca	1080
ctacacgatg ctttaactgc aaaaattcag ctcaccagtt ttgaggcaaa atttttgagt	1140
gacatgcaaa gtaagcatga tctcaatggt tcgttctcat ggctcacgca aaaacaacga	1200
accacactag agaacatact ggctaaatac ggaaggatct gaggttctta tggctcttgt	1260
atctatcagt gaagcatcaa gactaacaaa caaaagtaga acaactgttc accgttacat	1320
atcaaagggg aaactgtcca tatgcacaga gataatctca tgaccaaaac cggtagctag	1380
aggggccgca ttaggcaccc caggctttac actttatgct tccggctcgt ataatgtgtg	1440
gattttgagt taggatccgg cgagattttc aggagctaag gaagctaaaa tggagaaaaa	1500
aatcactgga tataccaccg ttgatatac ccaatggcat cgtaaagaac attttgaggc	1560
atctcagtca gttgctcaat gtacctataa ccagaccgtt cagctggata ttacggcctt	1620
tttaaagacc gtaaagaaaa ataagcaca gttttatccg gcctttattc acattcttgc	1680
ccgcctgatg aatgctcatc cggaattccg tatggcaatg aaagacggtg agctggtgat	1740
atgggatagt gttcaccctt gttacaccgt tttccatgag caaactgaaa cgttttcatc	1800

ES 2 752 175 T3

gctctggagt gaataccacg acgatttccg gcagtttcta cacatatatt cgcaagatgt 1860  
ggcgtgttac ggtgaaaacc tggcctatth ccctaaaggg tttattgaga atatgttttt 1920  
cgtctcagcc aatccctggg tgagtttcac cagttttgat ttaaactgtg ccaatatgga 1980  
caacttcttc gccccgtht tcaccatggg caaatattat acgcaaggcg acaagggtgct 2040  
gatgccgctg gcgattcagg ttcacatgc cgtctgtgat ggcttccatg tcggcagaat 2100  
gcttaatgaa ttacaacagt actgcgatga gtggcagggc gggcgtaaa cgcgtggatc 2160  
cggcttacta aaagccagat aacagtatgc gtatttgcgc gctgattttt gcggctctaga 2220  
ggtttagcaa gatggcagcg cctaaatgta gaatgataaa aggattaaga gattaatttc 2280  
cctaaaaatg ataaaaaag cgttttgaaa gcgcttgtht ttttgtht gctcagagt 2340  
agaatagaag tatcaaaaa agcaccgact cgggtgccact tttcaagth gataacggac 2400  
tagccttatt ttaacttgct atttctagct ctaaaactga gaccatcccg ggtctctact 2460  
gcagaattat caccgccagt ggtatthtg tcaacaccgc cagagataat ttatcaccgc 2520  
agatggtht cgatgaagat tcttgctcaa ttgthtctag ctatgcgccg accagaacac 2580  
cttgccgatc agccaaactg ctcttcaggg cactgactag cgataactth cccacaacg 2640  
gaacaactct cattgcatgg gatcattggg tactgtgggt ttagtgtht taaaaacacc 2700  
tgaccgctat ccctgatcag tttcttgaag gtaaactcat caccccaag tctggctatg 2760  
cagaaatcac ctggctcaac agcctgctca gggcaacga gaattaacat tccgtcagga 2820  
aagcttggct tggagcctg tggcgggc atggaattac cttcaacctc aagccagaat 2880  
gcagaatcac tggcttht gthtggct acccatctct ccgcatcacc tttgtht 2940  
gthtctaagct taggtgagaa catccctgcc tgaacatgag aaaaaacagg gtactcatac 3000  
tcaacttctaa gtgacggctg catactaacc gthtcataca tctcgtagat thtctggcg 3060  
attgaagggc taaattctc aacgctaact ttgagaatth ttgtaagcaa tggggcgtta 3120  
taagcattta atgcattgat gccattaaat aaagcaccaa cgcctgactg cccatcccc 3180  
atcttgtctg cgacagattc ctgggataag ccaagthcat tthtcttht thcataaatt 3240  
gctthttaggc gacgtgctc ctcaagctgc tcttgtgtht atgthtctt thtthtgcctc 3300  
atacgthttaa tctatcaccg caagggataa atatctaaca ccgtgcgtg thgactattht 3360  
acctctggcg gtgataatgg ttgcagggcc cthtthttagga ggcaaaaatg gataagaaat 3420  
actcaatagg cttagctatc ggcacaaaata gcgtcggatg ggcggtgatc actgatgaat 3480  
ataaggthtcc gtctaaaaag thcaagtht tgggaaatac agaccgccac agtatcaaaa 3540  
aaaatcttat agggctctt thatttgaca gtggagagac agcggagcg actcgtctca 3600  
aacggacagc tcgtagaag thatacagtc ggaagaatcg thtthttht ctacaggaga 3660  
thtthtcaaa tgagatggcg aaagtagatg atagthtctt tcatcgact thagagtht 3720

ES 2 752 175 T3

ttttggtgga agaagacaag aagcatgaac gtcacocctat ttttggaaat atagtagatg 3780  
 aagttgctta tcatgagaaa tatccaacta tctatcatct gcgaaaaaaaa ttggtagatt 3840  
 ctactgataa agcggatttg cgcttaatct atttggcctt agcgcataatg attaaagtttc 3900  
 gtggtcattt tttgattgag ggagatttaa atcctgataa tagtgatgtg gacaaaactat 3960  
 ttatccagtt ggtacaaaacc tacaatcaat tatttgaaga aaaccctatt aacgcaagtg 4020  
 gagtagatgc taaagcgatt ctttctgcac gattgagtaa atcaagacga ttagaaaatc 4080  
 tcattgctca gctccccggt gagaagaaaa atggccttatt tgggaatctc attgctttgt 4140  
 cattgggttt gaccocctaat tttaaatcaa attttgattt ggcagaagat gctaaattac 4200  
 agctttcaaa agatacttac gatgatgatt tagataatth attggcgcaa attggagatc 4260  
 aatatgctga tttgtttttg gcagctaaga atttatcaga tgctatttta ctttcagata 4320  
 tccaaagagt aaatactgaa ataactaagg ctcccctatc agcttcaatg attaaacgct 4380  
 acgatgaaca tcatcaagac ttgactcttt taaaagcttt agttcgacaa caacttccag 4440  
 aaaagtataa agaaatcttt tttgatcaat caaaaaacgg atatgcaggt tatattgatg 4500  
 ggggagctag ccaagaagaa ttttataaat ttatcaaacc aattttagaa aaaatggatg 4560  
 gtactgagga attattggtg aaactaaatc gtgaagattt gctgcgcaag caacggacct 4620  
 ttgacaacgg ctctattccc catcaaattc acttgggtga gctgcatgct attttgagaa 4680  
 gacaagaaga cttttatcca tttttaaaag acaatcgtga gaagattgaa aaaatcttga 4740  
 cttttcgaat tcttattat gttggtccat tggcgcgtgg caatagtcgt tttgcatgga 4800  
 tgactcggaa gtctgaagaa acaattacc catggaattt tgaagaagtt gtcgataaag 4860  
 gtgcttcagc tcaatcattt attgaacgca tgacaaactt tgataaaaat cttccaaatg 4920  
 aaaaagtact accaaaacat agtttgcttt atgagtattt tacggtttat aacgaattga 4980  
 caaaggctca atatgttact gaaggaatgc gaaaaccagc atttctttca ggtgaacaga 5040  
 agaaagccat tgttgattta ctcttcaaaa caaatcgaaa agtaaccggt aagcaattaa 5100  
 aagaagatta tttcaaaaaa atagaatggt ttgatagtggt tgaaatttca ggagttgaag 5160  
 atagatttaa tgcttcatta ggtacctacc atgatttgct aaaaattatt aaagataaag 5220  
 attttttgga taatgaagaa aatgaagata tcttagagga tattgtttta acattgacct 5280  
 tatttgaaga tagggagatg attgaggaaa gacttaaaac atatgctcac ctctttgatg 5340  
 ataaggtgat gaaacagctt aaacgtcgcc gttatactgg ttggggacgt ttgtctcgaa 5400  
 aattgattaa tgggtattagg gataagcaat ctggcaaaac aatattagat tttttgaaat 5460  
 cagatgggtt tgccaatcgc aattttatgc agctgatcca tgatgatagt ttgacattta 5520  
 aagaagacat tcaaaaagca caagtgtctg gacaaggcga tagtttacat gaacatattg 5580

ES 2 752 175 T3

caaatttagc tggtagccct gctattaanaa aaggatattt acagactgta aaagttggtg 5640  
 atgaattggt caaagtaatg gggcggcata agccagaaaa tatcgttatt gaaatggcac 5700  
 gtgaaaatca gacaactcaa aagggccaga aaaattcgcg agagcgtatg aaacgaatcg 5760  
 aagaaggat caaagaatta ggaagtcaga ttcttaaaga gcatcctggt gaaaatactc 5820  
 aattgcaaaa tgaaaagctc tatctctatt atctcaaaa tggaagagac atgtatgtgg 5880  
 accaagaatt agatattaat cgtttaagtg attatgatgt cgatgccatt gttccacaaa 5940  
 gtttccttaa agacgattca atagacaata aggtcttaac gcgttctgat aaaaatcgtg 6000  
 gtaaactcga taacgttcca agtgaagaag tagtcaaaaa gatgaaaaac tattggagac 6060  
 aacttctaaa cgccaagtta atcactcaac gtaagtttga taatttaacg aaagctgaac 6120  
 gtggagggtt gagtgaactt gataaagctg gttttatcaa acgccaattg gttgaaactc 6180  
 gccaaatcac taagcatgtg gcacaaatth tgatagtcg catgaatact aaatacgtg 6240  
 aaaatgataa acttattcga gaggttaaag tgattacctt aaaatctaaa ttagtttctg 6300  
 acttccgaaa agatttccaa ttctataaag tacgtgagat taacaattac catcatgccc 6360  
 atgatgcgta tctaaatgcc gtcgttgga ctgctttgat taagaaatat ccaaaacttg 6420  
 aatcggagtt tgtctatggt gattataaag tttatgatgt tcgtaaaatg attgctaagt 6480  
 ctgagcaaga aatagcaca gcaaccgcaa aatatttctt ttactctaata atcatgaact 6540  
 tcttcaaaa acgaaattaca cttgcaaatg gagagattcg caaacgccct ctaatcgaaa 6600  
 ctaatgggga aactggagaa attgtctggg ataaagggcg agattttgcc acagtgcgca 6660  
 aagtattgtc catgccccaa gtcaatattg tcaagaaaa agaatgacag acaggcggat 6720  
 tctccaagga gtcaatttta ccaaaaagaa attcggacaa gcttattgct cgtaaaaaag 6780  
 actgggatcc aaaaaatat ggtgggtttg atagtccaac ggtagcttat tcagtcctag 6840  
 tggttgctaa ggtggaaaaa gggaaatcga agaagttaaa atccgttaa gagttactag 6900  
 ggatcacaat tatggaaaga agttcctttg aaaaaatcc gattgacttt ttagaagcta 6960  
 aaggatataa ggaagttaa aaagacttaa tcattaaact acctaaatat agtctttttg 7020  
 agttagaaaa cggtcgtaaa cggatgctgg ctagtgccgg agaattacaa aaaggaaatg 7080  
 agctggctct gccaaagcaa tatgtgaatt ttttatattt agctagtcata tatgaaaagt 7140  
 tgaagggtag tccagaagat aacgaacaaa aacaattggt tgtggagcag cataagcatt 7200  
 atttagatga gattattgag caaatcagtg aattttctaa gcgtgttatt ttagcagatg 7260  
 ccaatttaga taaagttctt agtgcataa acaaacatag agacaaacca atacgtgaac 7320  
 aagcagaaaa tattattcat ttatttacgt tgacgaatct tggagctccc gctgctttta 7380  
 aatattttga tacaacaatt gatcgtaaac gatatacgtc taaaaagaa gtttttagatg 7440  
 ccactcttat ccatcaatcc atcactggtc tttatgaaac acgcattgat ttgagtcagc 7500

ES 2 752 175 T3

taggaggtga cgggtggagga ggttctggag gtggaggttc tgctgagtat gtgcgagccc 7560  
 tctttgactt taatgggaat gatgaagagg atcttccctt taagaaagga gacatcctga 7620  
 gaatccggga taagcctgag gagcagtggg ggaatgcaga ggacagcgaa ggaaagaggg 7680  
 ggatgattcc tgtcccttac gtggagaagt attccggaga ctataaggac cacgacggag 7740  
 actacaagga tcatgatatt gattacaaag acgatgacga taagtctagg ctcgagtccg 7800  
 gagactataa ggaccacgac ggagactaca aggatcatga tattgattac aaagacgatg 7860  
 acgataagtc taggatgacc gacgctgagt acgtgagaat ccatgagaag ttggacatct 7920  
 acacgtttta gaaacagttt ttcaacaaca aaaaatccgt gtcgcataga tgctacgttc 7980  
 tctttgaatt aaaacgacgg ggtgaacgta gagcgtgttt ttggggctat gctgtgaata 8040  
 aaccacagag cgggacagaa cgtggcattc acgccgaaat ctttagcatt agaaaagtcg 8100  
 aagaatacct gcgcgacaac cccggacaat tcacgataaa ttggtactca tcctggagtc 8160  
 cttgtgcaga ttgcgctgaa aagatcttag aatggtataa ccaggagctg cgggggaacg 8220  
 gccacacttt gaaaatctgg gcttgcaaac tctattacga gaaaaatgcg aggaatcaaa 8280  
 ttgggctgtg gaacctcaga gataacgggg ttgggttgaa tgtaatggta agtgaacact 8340  
 accaatgttg caggaaaata ttcatccaat cgtcgcacaa tcaattgaat gagaatagat 8400  
 ggcttgagaa gactttgaag cgagctgaaa aacgacggag cgagttgtcc attatgattc 8460  
 aggtaaaaat actccacacc actaagagtc ctgctgttac ttgacaggca tcaataaaaa 8520  
 cgaaaggctc agtcgaaaga ctgggccttt cgttttatct gttgtttgcg gccgggtacc 8580  
 gagctcgaat tcactggccg tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac 8640  
 ccaacttaat cgcttgcag cacatcccc tttcgccagc tggcgtaata gcgaagaggc 8700  
 ccgcaccgat cgcccttccc aacagttgcg cagcctgaat ggcgaatggc gattcacaaa 8760  
 aataggtac acgaaaaaca agttaaggga tgcagtttat gcatccotta acttacttat 8820  
 taaataatth atagctattg aaaagagata agaattgttc aaagctaata ttgtttaaat 8880  
 cgtcaattcc tgcagtgttt aaggaattgt taaattgatt ttttgtaaat attttcttgt 8940  
 attctttgtt aaccatttc ataacgaaat aattatactt ctgtttatct ttgtgtgata 9000  
 ttcttgattt ttttctatth aatctgataa gtgagctatt cactttaggt ttaggatgaa 9060  
 aatattctct tggaaaccata cttaatatag aaatatcaac ttctgccatt aaaaataatg 9120  
 ccaatgagcg ttttgtatth aataatctth tagcaaaccg gtattccacg attaaataaa 9180  
 tctcatcagc tatactatca aaaacaatth tgcgtattat atccgtactt atgttataag 9240  
 gtatattacc aatattthta taggattggg ttttaggaaa tttaaactgc aatatacct 9300  
 tgtthaaaac ttggaaatta tcgtgatcaa caagtttatt ttctgtagtt ttgcataatt 9360

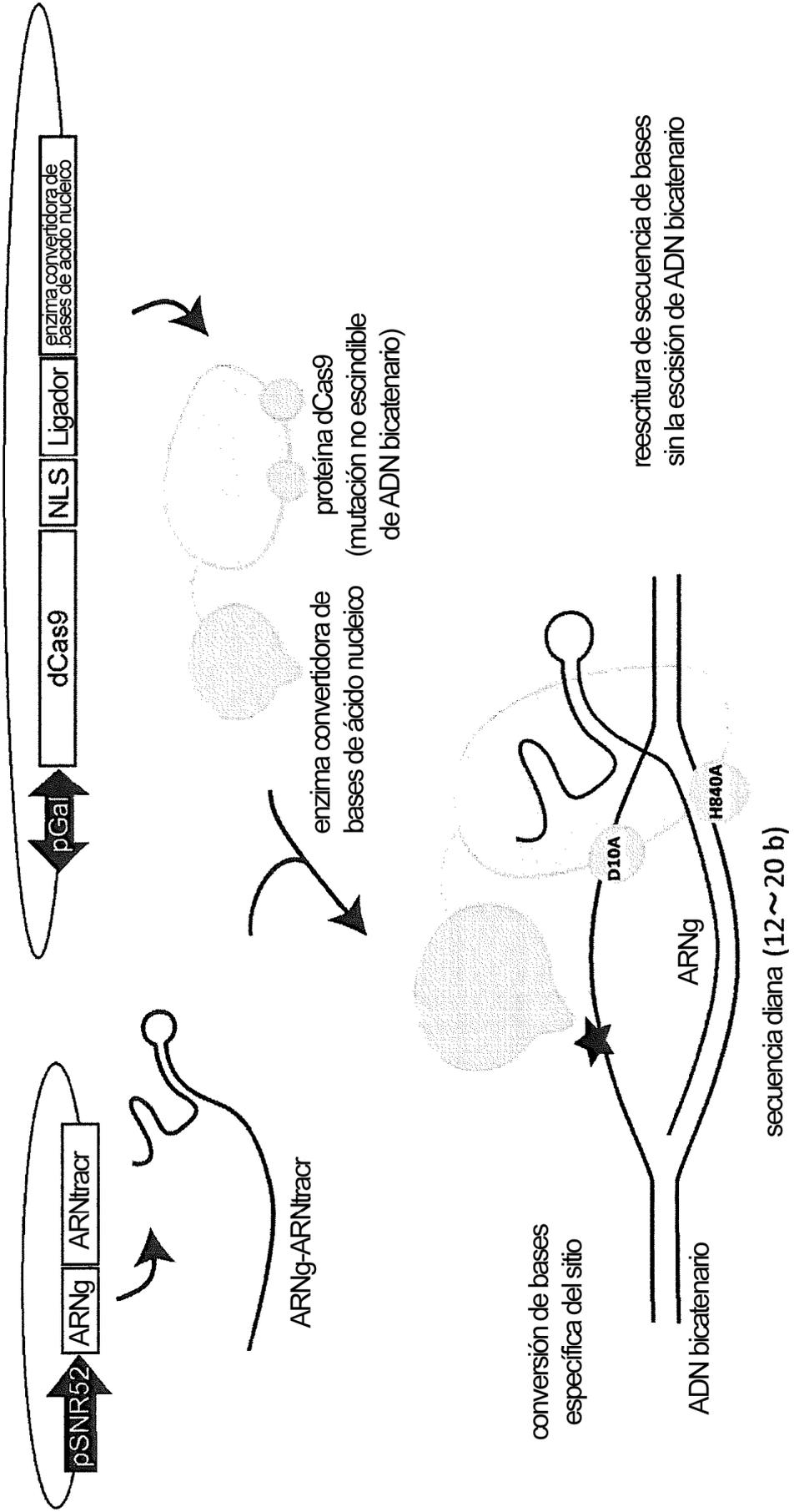
ES 2 752 175 T3

tatggtctat	ttcaatggca	gttacgaaat	tacacctctg	tactaattca	agggtaaaat	9420
gcccttttcc	tgagccgatt	tcaaagatat	tatcatgttc	atthaatctt	atatttgtca	9480
ttattttatc	tatattatgt	tttgaagtaa	taaagttttg	actgtgtttt	atatttttct	9540
cgttcattat	aaccctcttt	atthtttctt	ccttataaaa	ttagtataat	tatagcacga	9600
gctctgataa	atatgaacat	gatgagtgat	cgtaaattt	atactgcaat	ctgatgcat	9660
tattgaataa	aagatatgag	agatttatct	agtttctttt	ttacaagaa	aaaagaaagt	9720
tcttaaaggt	tttatacttt	tggtcgtaga	gcacacgggt	taacgactta	attacgaagt	9780
aaataagtct	agtgtgttag	actttaatgt	ttttttaagg	cattagtgca	tttaagcgtc	9840
agagcatggc	tttatgccga	gaaaactatt	ggttggaaatg	gcgtgtgtgt	tagccaaagc	9900
tttggcgagt	tggttggggg	tttcatggga	ttaatcccat	gaaagtacca	actcaacaac	9960
acactaacgc	ctgttggttc	caaccaatag	gaaattggaa	taagcaatta	gtataatgag	10020
agtataatgt	tggtataacg	ttagtataat	gatgcttttt	ttcattatat	tttttatgta	10080
ctttaaacct	gcacgcttat	gtgaattaga	aaaagcttaa	tcgcatttca	tagattgacc	10140
tccaataaac	tacgtggtgt	tattgggagg	tcaatctatt	tcatttgcct	cttgctcaaa	10200
gttcccaaat	tcgagtaaga	ggtatthttg	tttttggctg	tcgcctctca	ttagtagttc	10260
agggtttaac	attaatactc	cagthtttct	ttttataata	ttccttctt	ctaagattht	10320
aagtgttggt	attactgttt	gtagacttgt	tcctgtagct	tttgctatth	ctcttgttgt	10380
agctatcatt	gtattgttac	ttaagtggac	attatctagg	atatagttaa	cgattthtaag	10440
tttttttccg	ccaatcatat	ctaacatact	tattaattgc	actatatatg	cctttacgaa	10500
gttaccagac	gthttgtttac	ggtataactt	gtctacctct	atgacttctc	cactthtctc	10560
gtctatgagc	ctctgagagc	ctttatagac	tgthccatat	ctthctthca	tctthttctc	10620
actccttatt	ttaaacctatt	ctaactatat	cataactgth	ctaaaaaaaa	aagaacatth	10680
gttaaaagaa	attagaacaa	aatgagttaa	aaattagaac	aaacaaattc	cttataaacc	10740
ttatcatctc	aacctatatt	aagatthttac	ctagthtgaat	ctthctthtct	atataaagcg	10800
tcggagcata	tcaggggggtt	atctaacgta	aatgctacc	ttcggctcgc	tttcgctcgg	10860
cattgac						10867

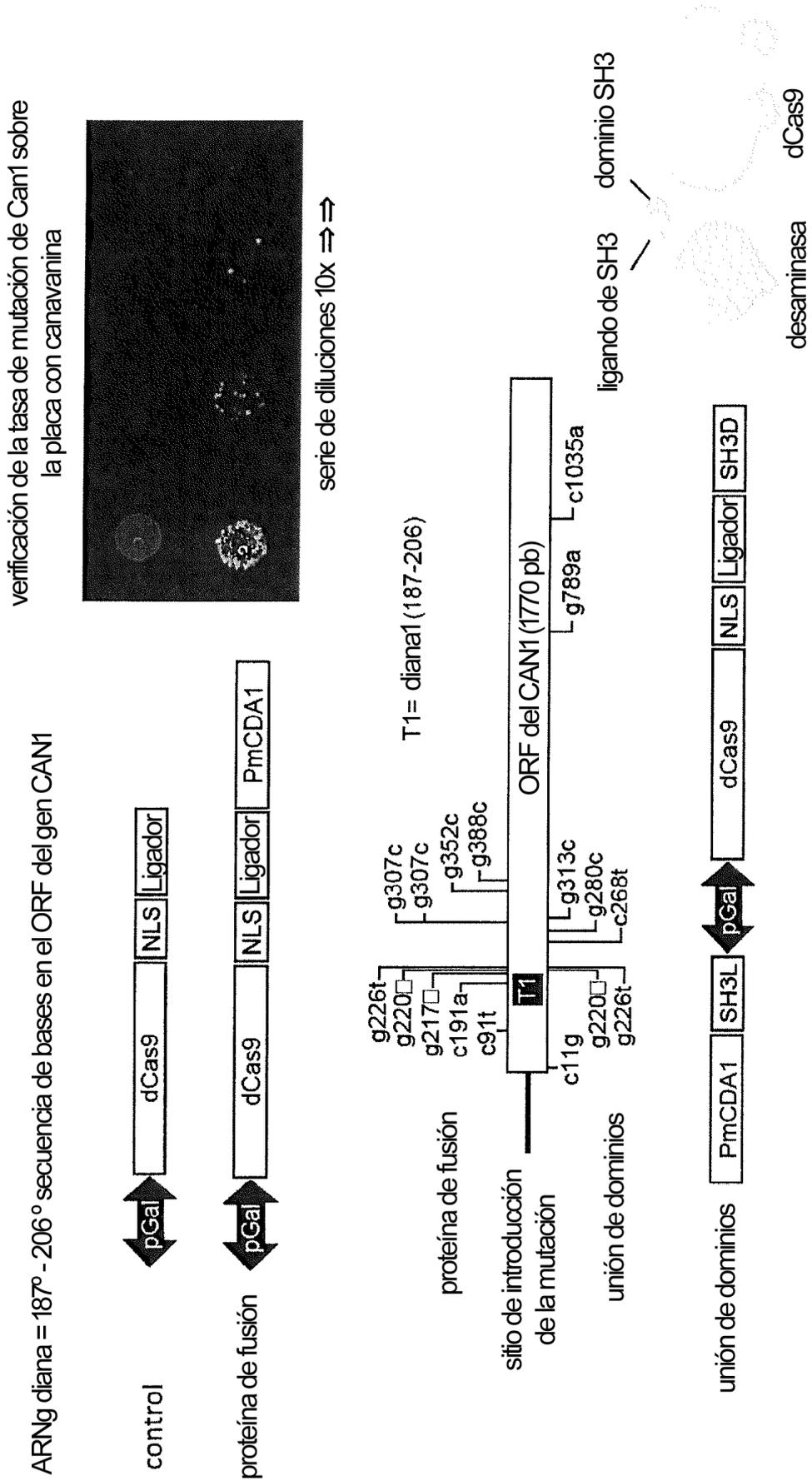
**REIVINDICACIONES**

1. Método de modificación de un sitio seleccionado como diana de un ADN bicatenario, que comprende una etapa de poner en contacto un complejo que comprende un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a una secuencia de nucleótidos diana en un ADN bicatenario dado unido a una enzima convertidora de bases de ácido nucleico, con dicho ADN bicatenario, para convertir uno o más nucleótidos en el sitio seleccionado como diana en uno o más de otros nucleótidos o delecionar uno o más nucleótidos, o insertar uno o más nucleótidos en dicho sitio seleccionado como diana, en el que el módulo de reconocimiento de secuencia de ácido nucleico es un sistema CRISPR-Cas, en el que el sistema CRISPR-Cas comprende una proteína Cas que tiene actividad nickasa capaz de escindir sólo una de las cadenas del ADN bicatenario, en el que el método no es un método para el tratamiento del cuerpo humano o animal mediante cirugía o terapia, y en el que el método no es un método para modificar la identidad genética de línea germinal de seres humanos.
2. Método según la reivindicación 1, en el que la Cas es una Cas9 en el que el 10º residuo de Asp se convierte en un residuo de Ala o el 840º residuo de His se convierte en un residuo de Ala.
3. Método según la reivindicación 1 ó 2, que usa dos o más clases de módulos de reconocimiento de secuencia de ácido nucleico que se une cada uno específicamente a una secuencia de nucleótidos diana diferente.
4. Método según la reivindicación 3, en el que las secuencias de nucleótidos diana diferentes están presentes en genes diferentes.
5. Método según una cualquiera de las reivindicaciones 1 a 4, en el que la enzima convertidora de bases de ácido nucleico es una desaminasa, preferiblemente una citidina desaminasa.
6. Método según una cualquiera de las reivindicaciones 1 a 5, en el que el ADN bicatenario se pone en contacto con el complejo introduciendo un ácido nucleico que codifica para el complejo en una célula que tiene el ADN bicatenario.
7. Método según la reivindicación 6, en el que la célula es una célula procariota.
8. Método según la reivindicación 6, en el que la célula es una célula microbiana.
9. Método según la reivindicación 6, en el que la célula es una célula eucariota, preferiblemente una célula vegetal, una célula de insecto o una célula animal.
10. Método según la reivindicación 9, en el que la célula animal es una célula de vertebrado, preferiblemente una célula de mamífero.
11. Método según una cualquiera de las reivindicaciones 6 a 10, en el que la célula es una célula poliploide, y se modifican todos los sitios seleccionados como diana en alelos en cromosomas homólogos.
12. Método según una cualquiera de las reivindicaciones 6 a 11, que comprende una etapa de introducir un vector de expresión que comprende un ácido nucleico que codifica para el complejo en una forma que permite el control de un periodo de expresión en la célula, y una etapa de inducir la expresión del ácido nucleico durante un periodo necesario para fijar la modificación del sitio seleccionado como diana en el ADN bicatenario.
13. Método según la reivindicación 12, en el que la secuencia de nucleótidos diana en el ADN bicatenario está presente en un gen esencial para la célula.
14. Complejo enzimático de modificación de ácido nucleico que comprende un módulo de reconocimiento de secuencia de ácido nucleico que se une específicamente a una secuencia de nucleótidos diana en un ADN bicatenario dado unido a una enzima convertidora de bases de ácido nucleico y el módulo de reconocimiento de secuencia de ácido nucleico es un sistema CRISPR-Cas, en el que el sistema CRISPR-Cas comprende una proteína Cas que tiene actividad nickasa capaz de escindir sólo una de las cadenas del ADN bicatenario, complejo que convierte uno o más nucleótidos en el sitio seleccionado como diana en uno o más de otros nucleótidos o deleciona uno o más nucleótidos, o inserta uno o más nucleótidos en dicho sitio seleccionado como diana.
15. Ácido nucleico que codifica para el complejo enzimático de modificación de ácido nucleico según la reivindicación 14.

Fig. 1



**Fig. 2**



**Fig. 3**

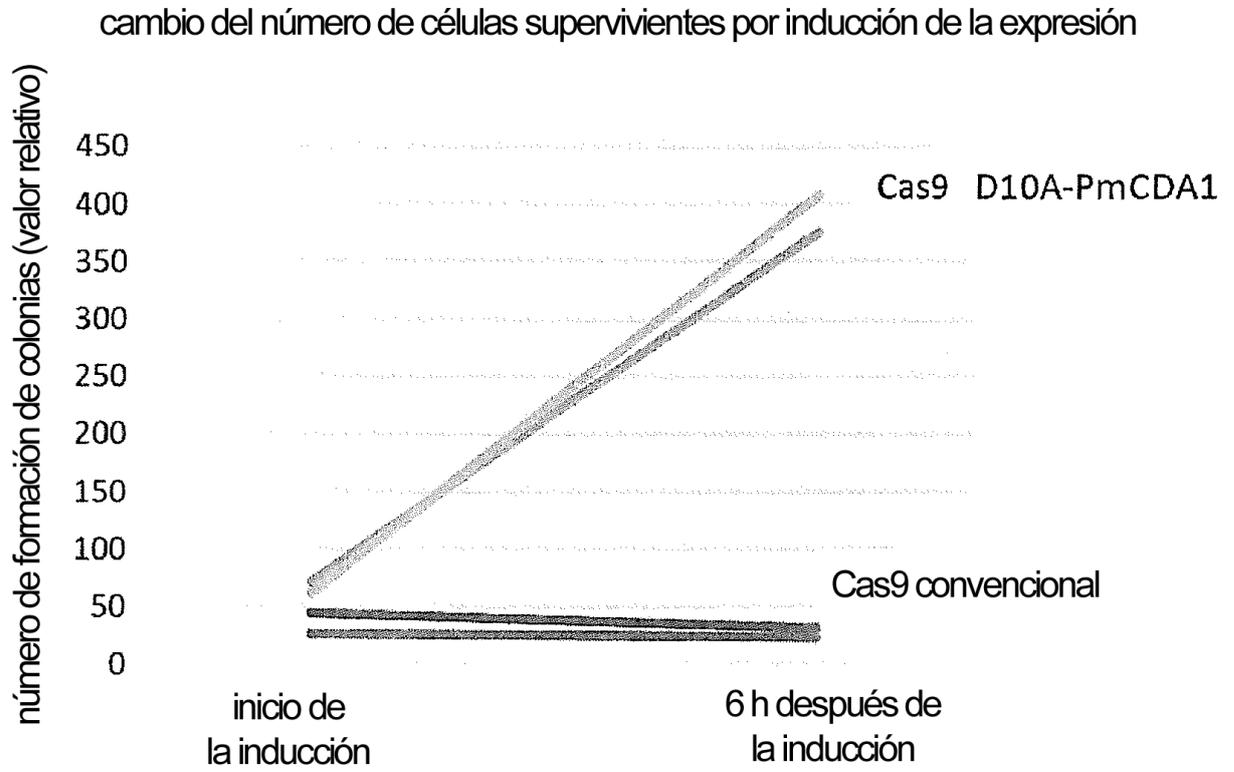




Fig. 5

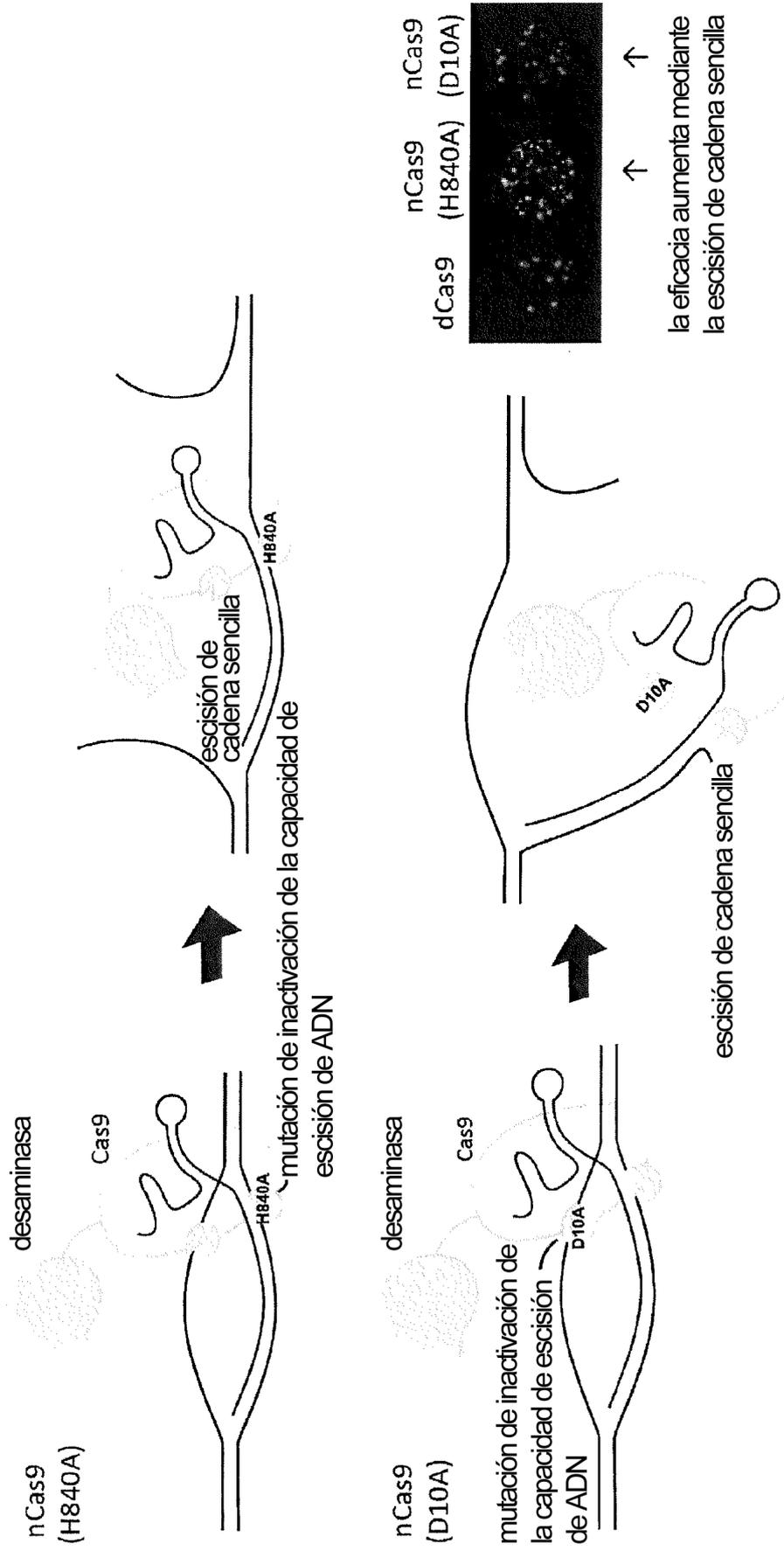




Fig. 7

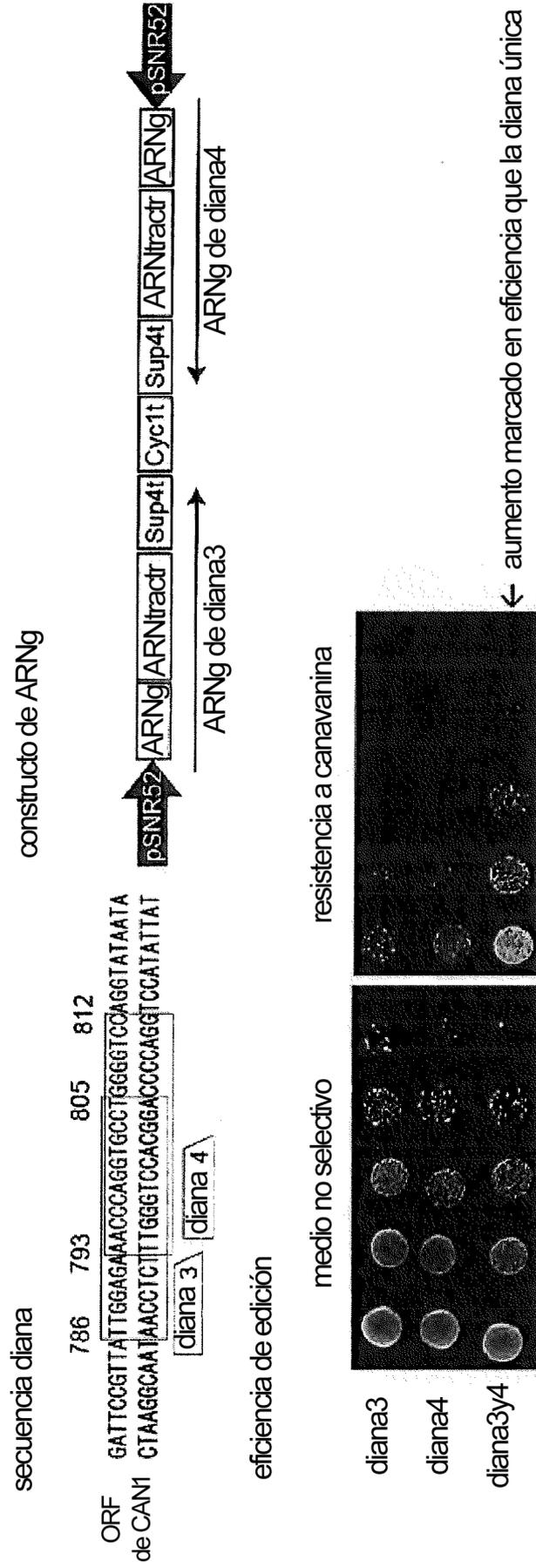
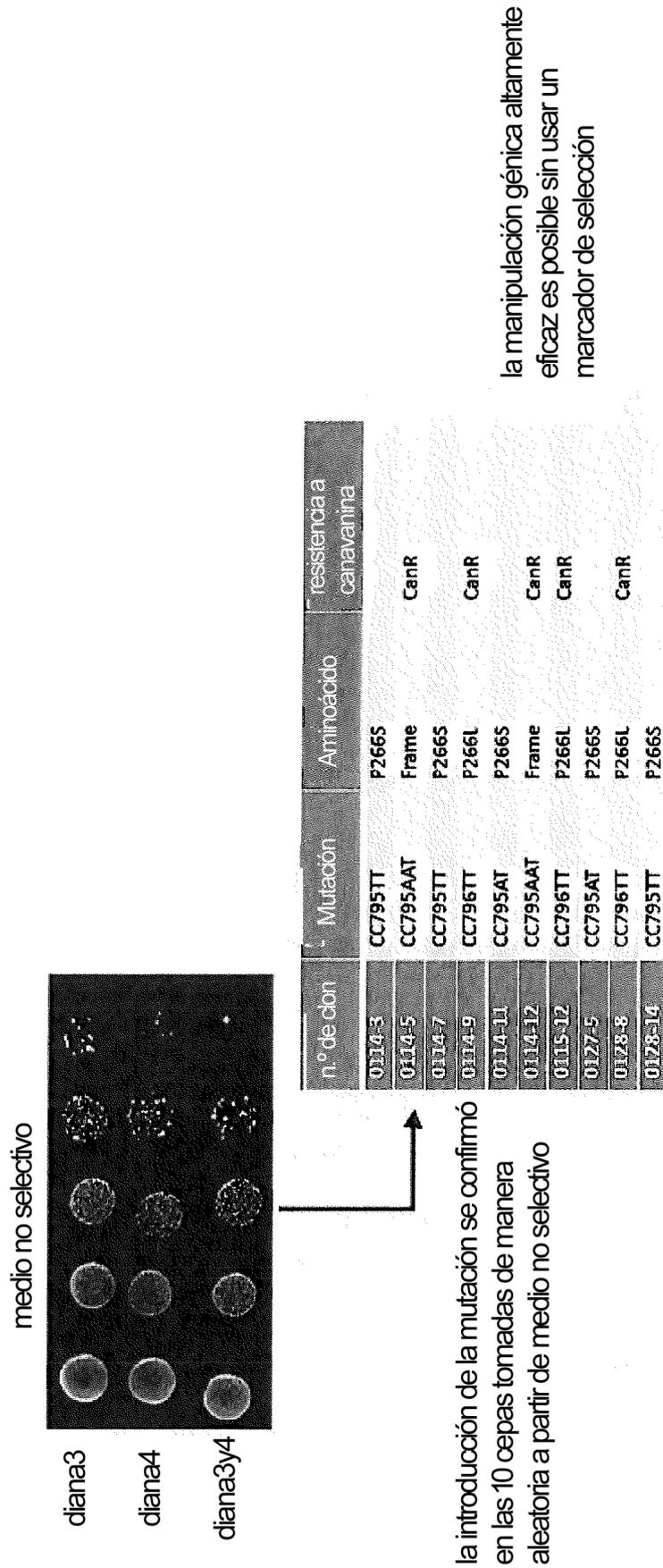
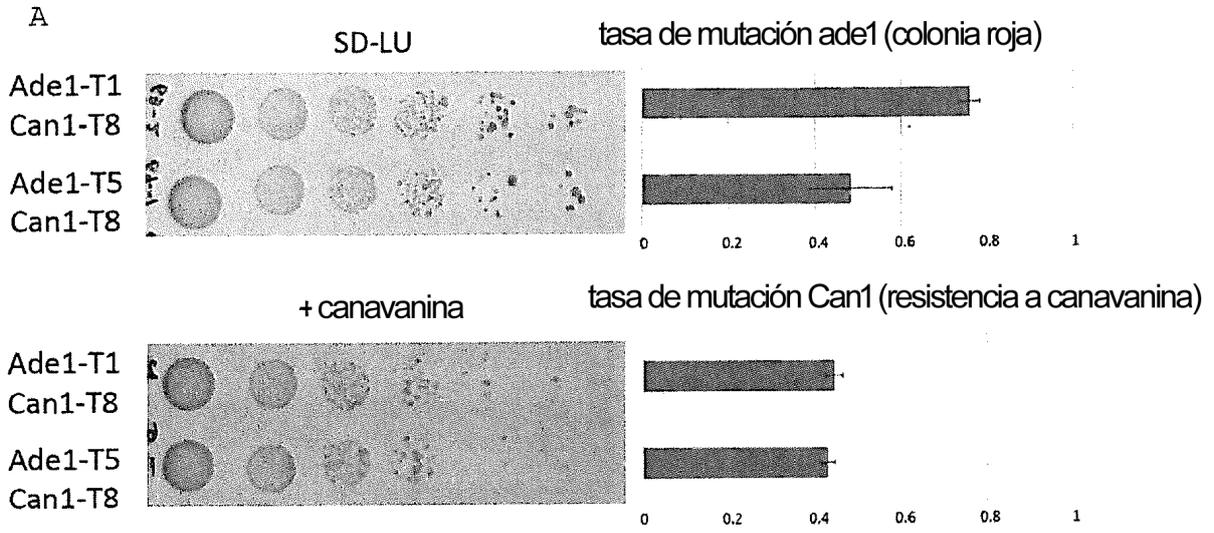


Fig. 8





**Fig. 10**



**B**

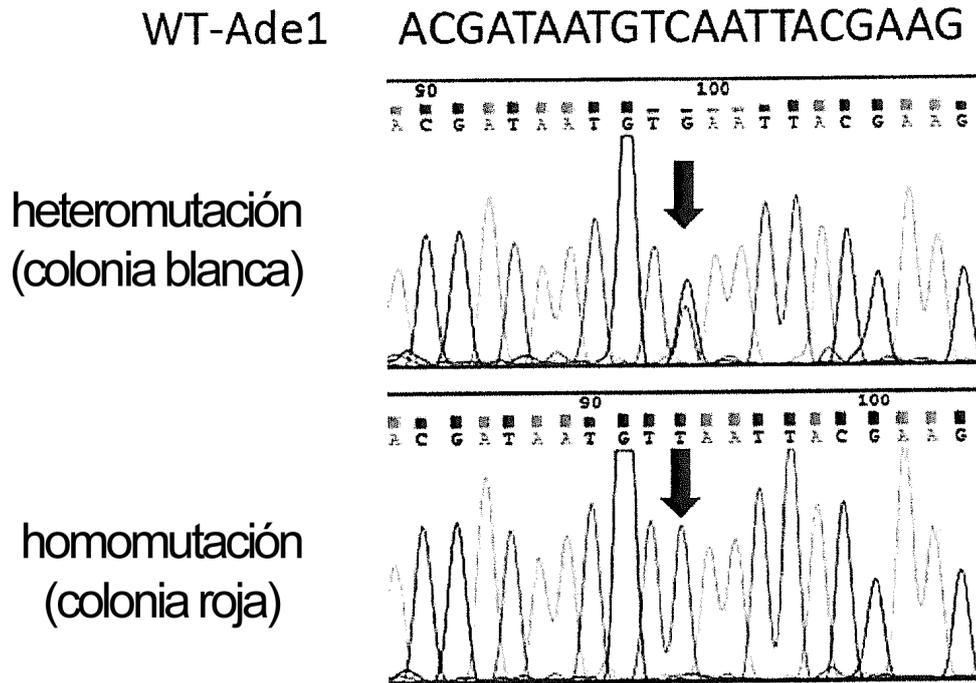
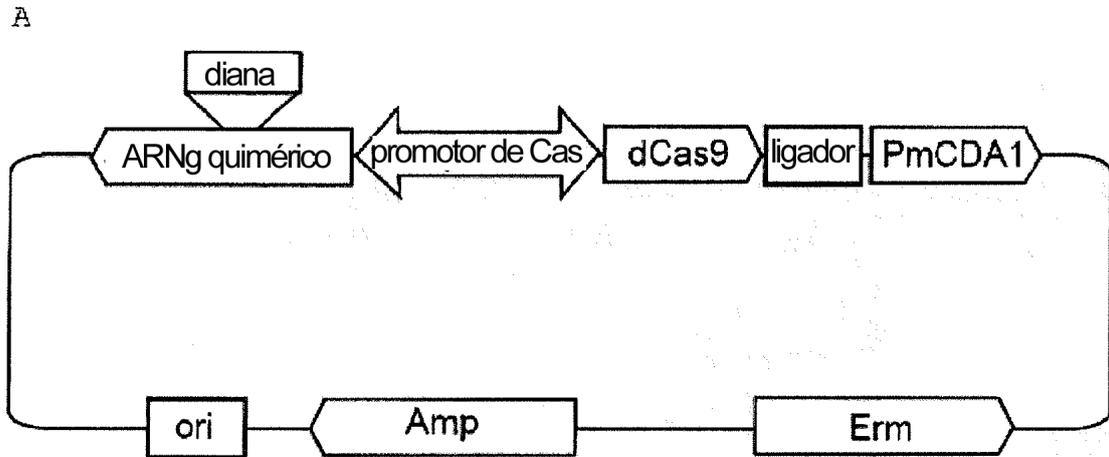


Fig. 11



B

***galK*** TCAATGGGCTAACTACGTTTC-3'  
 1 TCAATGGGCTAACTACGTTTC  
 2 TTAATGGGCTAACTACGTTTC  
 3 TTAATGGGCTAACTACGTTTC

C

***rpoB*** GCTGTCTCAGTTTATGGACC-3'  
 [CGACAGAGTCAAATACCTGG-5']

ninguno [ 1 GCTGTCTCAGTTTATGGACC  
 2 GCTGTCTCAGTTTATGGACC  
 Rif25 [ 1 GCTGTCTCAGTTTATGAACC  
 2 GCTGTCTCAGTTTATGAACC  
 Rif50 [ 1 GCTGTCTCAGTTTATAAACC  
 2 GCTGTCTCAGTTTATGAACC  
 [CGACAGAGTCAAATACTTGG-5']

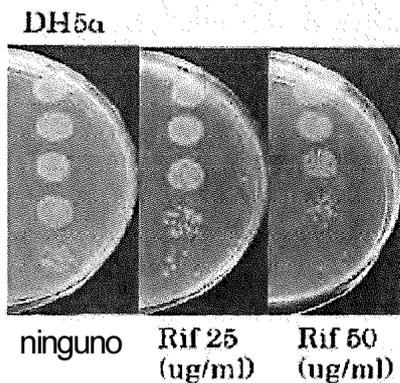
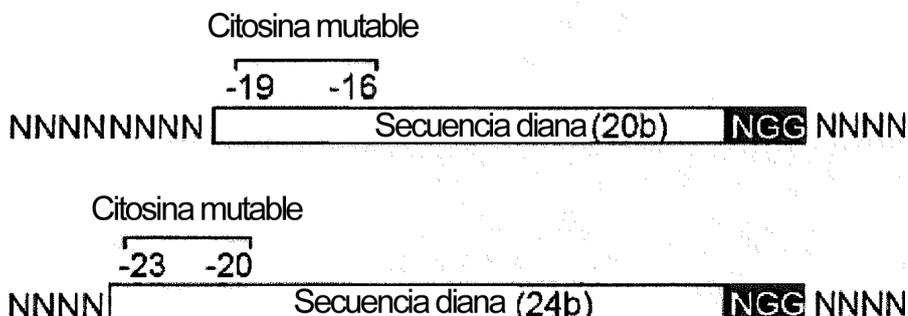


Fig. 12

A



B

TCGCTTGAACATCCAGCGAAACAGGG**CCCCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
diana 24 pb

TCGCTTGAACATCCAGCGAAACAGGG**CCCCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCCCATCGAGCAGAAA**CGGTGGTGGATGGC

diana 22 pb

TCGCTTGAACATCCAGCGAAACAGGG**CCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCCATCGAGCAGAAA**CGGTGGTGGATGGC

diana 20 pb

TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC

diana 18 pb

TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC  
TCGCTTGAACATCCAGCGAAACAGGG**CCCATCGAGCAGAAA**CGGTGGTGGATGGC

**Fig. 13**

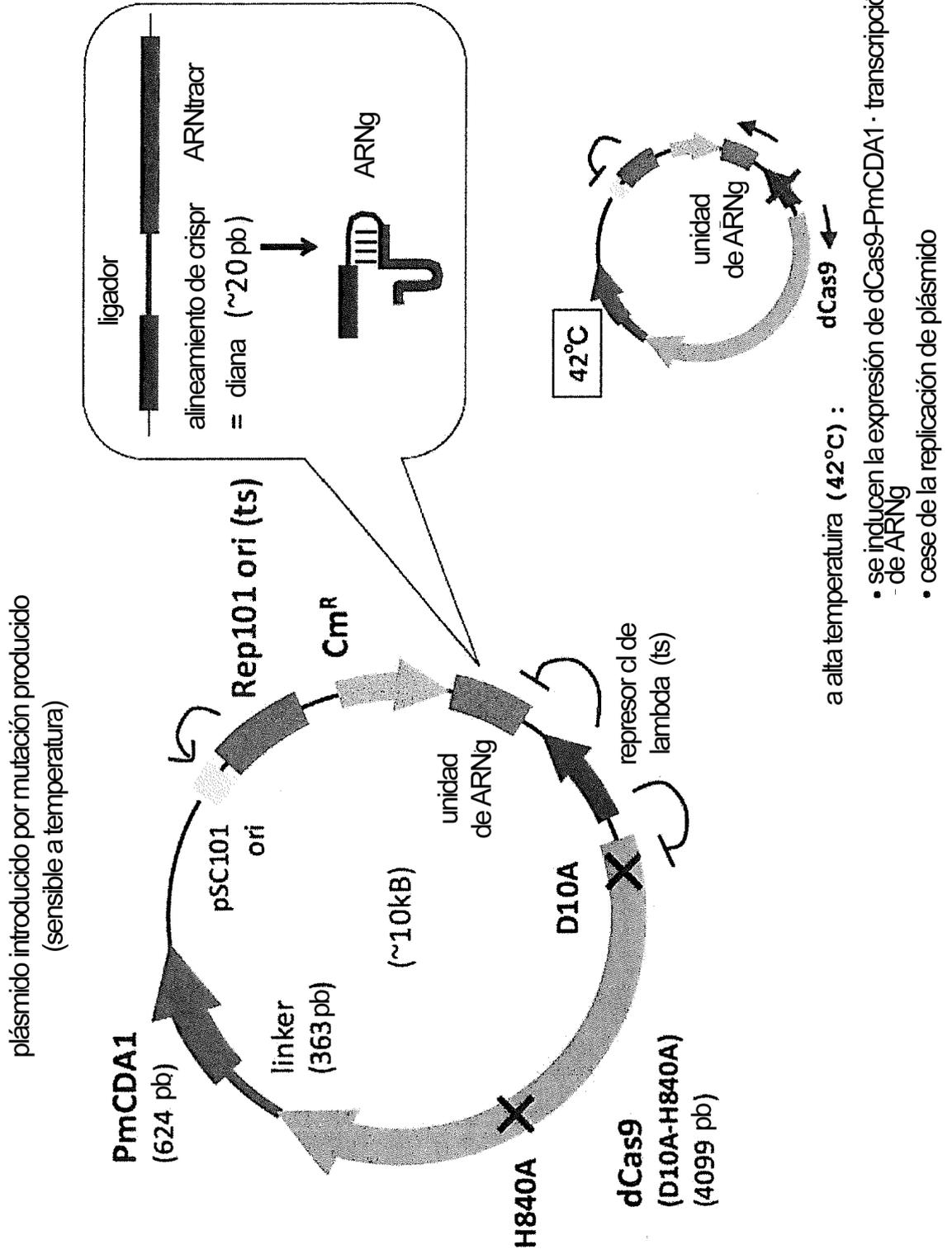
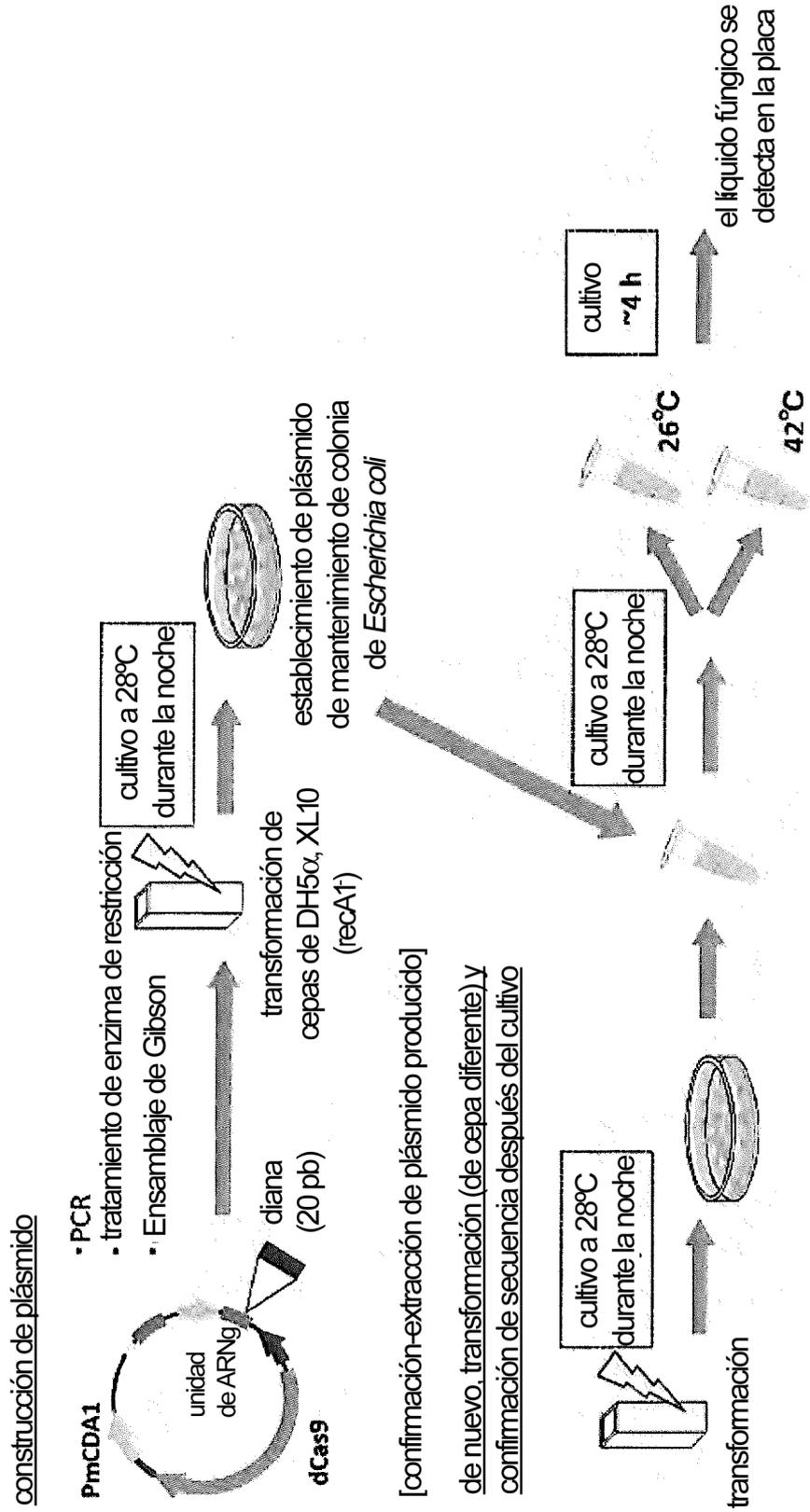


Fig. 14

método de introducción de mutación (plásmido sensible a la temperatura)



inicio de la inducción de la expresión de dCas9-PmCDA1- parada de la transcripción de ARNg de la replicación de plásmido

Fig. 15

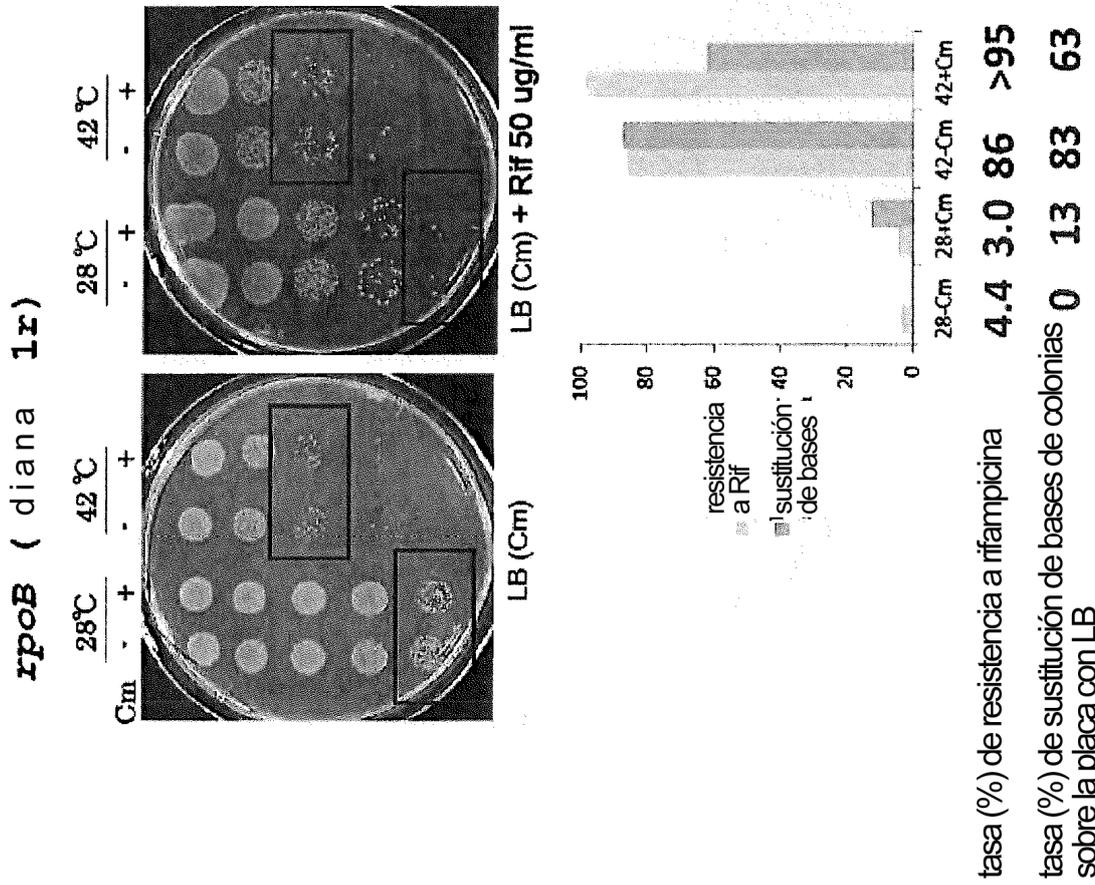
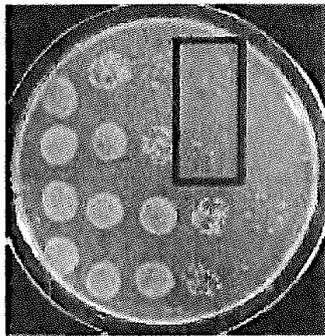


Fig. 16

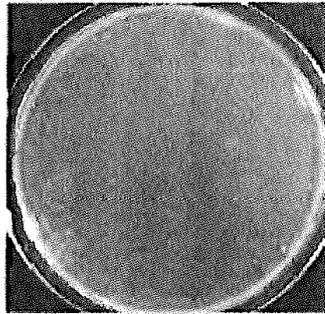
*galK* (diana 8, diana 12)

	28 °C	42 °C	
Cm	-	+	+
	-	+	-



LB (Cm)

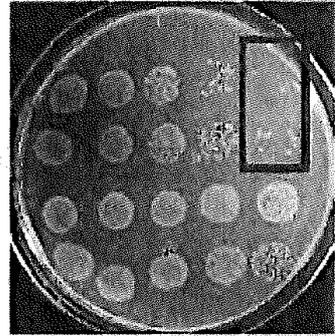
	28 °C	42 °C	
	-	+	+
	-	+	-



LB (Cm) + 2-DOG at 0.2%

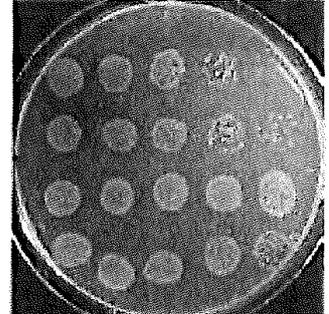
<i>galK</i> 8	ACTCACACCATTCAGGGGCC -3'
colonia1	ACTTACACCATTCAGGGGCC
colonia2	ACTTACACCATTCAGGGGCC (c→t)

	28 °C	42 °C	
Cm	-	+	+
	-	+	-



LB (Cm)

	28 °C	42 °C	
	-	+	+
	-	+	-



LB (Cm) + 2-DOG at 0.2%

<i>galK</i> 12	TCAATGGGCTAACTACGTTGG -3'
colonia1	TTAATGGGCTAACTACGTTGG
colonia2	TTAATGGGCTAACTACGTTGG (c→t)