

COMPOSICIONES QUE COMPRENDEN UN POLIPÉPTIDO CAS12I4 VARIANTE Y USOS DE LAS MISMAS

LISTA DE SECUENCIAS

5 La presente solicitud contiene una Lista de Secuencias que se ha enviado electrónicamente en formato ASCII y se incorpora en el presente documento por referencia en su totalidad. Dicha copia ASCII, creada el 11 de febrero de 2022, se denomina 51451-023WO3_Sequence_Listing_2_10_22_ST25, y tiene un tamaño de 607,386 bytes.

SOLICITUDES RELACIONADAS

10 La presente solicitud reivindica la prioridad del N.º de serie de los EE. UU. 63/148,421, presentado el 11 de febrero de 2021, y el N.º de serie de los EE. UU. 63/154,437, presentado el 26 de febrero de 2021. El contenido de cada una de estas solicitudes anteriores se incorpora en el presente documento por referencia en su totalidad.

ANTECEDENTES

15 Las repeticiones palindrómicas cortas agrupadas y regularmente espaciadas (CRISPR, por sus siglas en inglés) y los genes asociados a CRISPR (Cas), conocidos en conjunto como sistemas CRISPR-Cas o CRISPR/Cas, son sistemas inmunitarios adaptativos en arqueas y bacterias que defienden especies particulares frente a elementos genéticos extraños.

SUMARIO DE LA INVENCION

20 La presente invención proporciona determinadas ventajas y avances sobre la técnica anterior frente a los antecedentes anteriores. Aunque la presente invención desvelada en el presente documento no se limita a ventajas o funcionalidades específicas, la invención proporciona un polipéptido Cas12i4 variante que comprende una secuencia que tiene al menos un 95 % de identidad con una secuencia expuesta en una cualquiera de las SEQ ID NO: 3-59.

25 En un aspecto de la variante, el polipéptido Cas12i4 variante es una variante de un polipéptido precursor de la SEQ ID NO: 2.

En otro aspecto de la variante, el polipéptido Cas12i4 variante comprende una sustitución de la Tabla 2.

30 En otro aspecto de la variante, la variante comprende la secuencia expuesta en una cualquiera de las SEQ ID NO: 3-59.

En otro aspecto de la variante, la variante comprende la secuencia expuesta en la SEQ ID NO: 3.

En otro aspecto de la variante, la variante comprende la secuencia expuesta en la SEQ ID NO: 4.

En otro aspecto de la variante, el polipéptido Cas12i4 variante presenta una formación de complejos binarios aumentada con una guía de ARN, con respecto a un polipéptido precursor.

35 En otro aspecto de la variante, un complejo binario que comprende el polipéptido Cas12i4 variante presenta una estabilidad aumentada, con respecto a un complejo binario precursor.

En otro aspecto de la variante, el polipéptido Cas12i4 variante presenta una actividad nucleasa aumentada, con respecto a un polipéptido precursor.

En otro aspecto de la variante, el polipéptido Cas12i4 variante comprende además una sustitución de la Tabla 4.

5 En otro aspecto de la variante, la sustitución de la Tabla 4 aumenta la formación de complejos binarios con una guía de ARN, con respecto a un polipéptido precursor.

En otro aspecto de la variante, la sustitución de la Tabla 4 aumenta la estabilidad de un complejo binario que comprende el polipéptido Cas12i4 variante, con respecto a un complejo binario precursor.

10 En otro aspecto de la variante, el polipéptido Cas12i4 variante comprende además una sustitución que aumenta la formación de complejos ternarios con una guía de ARN y un ácido nucleico diana, con respecto a un polipéptido precursor.

En otro aspecto de la variante, el polipéptido Cas12i4 variante comprende además una sustitución que aumenta la estabilidad de los complejos ternarios, con respecto a un polipéptido precursor.

15 En otro aspecto de la variante, la sustitución es una sustitución de la Tabla 5, la Tabla 6, la Tabla 7, la Tabla 8, la Tabla 9 y/o la Tabla 10.

En otro aspecto de la variante, el polipéptido Cas12i4 variante comprende además una sustitución que aumenta la unión en la diana a un ácido nucleico diana, con respecto a un polipéptido precursor.

En otro aspecto de la variante, la sustitución es una sustitución de la Tabla 11.

20 Además, la invención proporciona además una composición que comprende un polipéptido Cas12i4 variante como se describe en el presente documento, en donde la composición comprende además una guía de ARN o un ácido nucleico que codifica la guía de ARN, en donde la guía de ARN comprende una secuencia repetitiva directa y una secuencia espaciadora.

En un aspecto de la composición, la secuencia repetitiva directa comprende:

25 a. del nucleótido 1 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

b. del nucleótido 2 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

30 c. del nucleótido 3 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

d. del nucleótido 4 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 35 119, 120, 121, 122, 123 o 124;

e. del nucleótido 5 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

f. del nucleótido 6 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

g. del nucleótido 7 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

h. del nucleótido 8 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

i. del nucleótido 9 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

j. del nucleótido 10 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

k. del nucleótido 11 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

l. del nucleótido 12 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

m. del nucleótido 13 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

n. del nucleótido 14 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o

o. una secuencia que es al menos un 90 % idéntica a una secuencia de la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la composición, la secuencia repetitiva directa comprende:

a. del nucleótido 1 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

b. del nucleótido 2 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

c. del nucleótido 3 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

d. del nucleótido 4 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

e. del nucleótido 5 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

f. del nucleótido 6 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

g. del nucleótido 7 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

h. del nucleótido 8 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

i. del nucleótido 9 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

j. del nucleótido 10 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

k. del nucleótido 11 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

l. del nucleótido 12 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

m. del nucleótido 13 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

n. del nucleótido 14 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o

o. una secuencia que es al menos un 95% idéntica a una secuencia de la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la composición, la repetición directa comprende:

- a. del nucleótido 1 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- b. del nucleótido 2 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- c. del nucleótido 3 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- d. del nucleótido 4 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- e. del nucleótido 5 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- f. del nucleótido 6 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- g. del nucleótido 7 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- h. del nucleótido 8 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- i. del nucleótido 9 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- j. del nucleótido 10 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- k. del nucleótido 11 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- l. del nucleótido 12 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- m. del nucleótido 13 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- n. del nucleótido 14 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o
- o. la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la composición, la secuencia repetitiva directa comprende $AGN_1N_2N_3N_4GUGUN_5N_6N_7CAGN_8GACN_9C$ (SEQ ID NO: 125), en donde N_1 es A o G, N_2 es C o U, N_3 es A o G, N_4 es U o C, N_5 es C o U, N_6 es C o U, N_7 es U, A, C o G, N_8 es U o C, y N_9 es A o C.

En otro aspecto de la composición, la secuencia espaciadora comprende de aproximadamente 15 nucleótidos a aproximadamente 35 nucleótidos de longitud.

En otro aspecto de la composición, la secuencia espaciadora se une a una secuencia de cadena diana de un ácido nucleico diana, y una secuencia de cadena no diana de la secuencia de ácido nucleico

diana es adyacente a una secuencia de motivo adyacente al protoespaciador (PAM, por sus siglas en inglés).

En otro aspecto de la composición, la secuencia de PAM es 5'-TTN-3', 5'-NTTN-3', 5'-NTN'-3', 5'-NNTN-3', 5'-VTN-3' o 5'-NVTN-3', en donde N es cualquier nucleótido y V es A, G o C.

5 En otro aspecto de la variante o la composición, el polipéptido Cas12i4 variante comprende además una señal de localización nuclear (NLS, por sus siglas en inglés).

En otro aspecto de la variante o la composición, el polipéptido Cas12i4 variante comprende además un marcador de péptido, una proteína fluorescente, un dominio de edición de base, un dominio de metilación de ADN, un dominio de modificación de residuo de histona, un factor de localización, un factor de modificación de la transcripción, un factor de control dependiente de la luz, un factor químicamente inducible o un factor de visualización de cromatina.

La invención proporciona además un ácido nucleico que codifica un polipéptido Cas12i4 o una composición como se describe en el presente documento.

15 En un aspecto de la composición, el ácido nucleico tiene codones optimizados para la expresión en una célula.

En otro aspecto de la composición, el ácido nucleico se une operativamente a un promotor.

En otro aspecto de la composición, el ácido nucleico está en un vector.

En otro aspecto de la composición, el vector comprende un vector retrovírico, un vector lentivírico, un vector fágico, un vector adenovírico, un vector adenoasociado o un vector de herpes simple.

20 En otro aspecto de la variante o la composición, el polipéptido Cas12i4 variante está presente en un sistema de suministro que comprende una nanopartícula (por ejemplo, una nanopartícula lipídica), un liposoma, un exosoma, una microvesícula o un cañón génico.

La invención proporciona además una célula que comprende un polipéptido Cas12i4 variante o una composición como se describe en el presente documento.

25 En un aspecto de la célula, la célula es una célula eucariota.

En otro aspecto de la célula, la célula es una célula de mamífero o una célula vegetal.

En otro aspecto de la célula, la célula es una célula humana.

La invención proporciona además una composición que comprende un polipéptido Cas12i4 variante o un complejo que comprende el polipéptido Cas12i4 variante, en donde el polipéptido Cas12i4 variante comprende una secuencia que tiene al menos un 95 % de identidad con una secuencia expuesta en una cualquiera de las SEQ ID NO: 3-59, y en donde el polipéptido Cas12i4 variante o el complejo presenta una actividad enzimática potenciada, una actividad de unión potenciada, una especificidad de unión potenciada y/o una estabilidad potenciada, con respecto a un polipéptido precursor o un complejo que comprende el polipéptido precursor.

35 En un aspecto de la composición, el polipéptido Cas12i4 variante comprende una sustitución de la Tabla 2, la Tabla 4, la Tabla 5, la Tabla 6, la Tabla 7, la Tabla 8, la Tabla 9, la Tabla 10 y/o la Tabla 11.

En otro aspecto de la composición, el polipéptido Cas12i4 variante comprende la secuencia expuesta en una cualquiera de las SEQ ID NO: 3-59.

En otro aspecto de la composición, el polipéptido Cas12i4 variante comprende la secuencia expuesta en la SEQ ID NO: 3.

En otro aspecto de la composición, el polipéptido Cas12i4 variante comprende la secuencia expuesta en la SEQ ID NO: 4.

5 En otro aspecto de la composición, la actividad enzimática potenciada es la actividad nucleasa potenciada.

En otro aspecto de la composición, el polipéptido Cas12i4 variante presenta una actividad de unión potenciada a una guía de ARN, con respecto al polipéptido precursor.

10 En otro aspecto de la composición, el polipéptido Cas12i4 variante presenta una especificidad de unión potenciada a una guía de ARN, con respecto al polipéptido precursor.

En otro aspecto de la composición, el complejo que comprende el polipéptido Cas12i4 variante es un complejo binario variante que comprende además una guía de ARN, y el complejo binario variante presenta una actividad de unión potenciada a un ácido nucleico diana (por ejemplo, actividad de unión en la diana), con respecto a un complejo binario precursor.

15 En otro aspecto de la composición, el complejo que comprende el polipéptido Cas12i4 variante es un complejo binario variante que comprende además una guía de ARN, y el complejo binario variante presenta una especificidad de unión potenciada a un ácido nucleico diana (por ejemplo, especificidad de unión en la diana), con respecto a un complejo binario precursor.

20 En otro aspecto de la composición, el complejo que comprende el polipéptido Cas12i4 variante es un complejo binario variante que comprende además una guía de ARN, y el complejo binario variante presenta una estabilidad potenciada, con respecto a un complejo binario precursor.

En otro aspecto de la composición, el complejo binario variante y un ácido nucleico diana forman un complejo ternario variante, y el complejo ternario variante presenta una estabilidad aumentada, con respecto a un complejo ternario precursor.

25 En otro aspecto de la composición, el polipéptido Cas12i4 variante presenta además una formación de complejos binarios potenciada, interacciones proteína-ARN potenciadas y/o una disociación disminuida de una guía de ARN, con respecto al polipéptido precursor.

30 En otro aspecto de la composición, el complejo binario variante presenta además una disociación disminuida de un ácido nucleico diana y/o una unión fuera de la diana disminuida a un ácido nucleico no diana, con respecto al complejo binario precursor.

En otro aspecto de la composición, la actividad enzimática potenciada, la actividad de unión potenciada, la especificidad de unión potenciada y/o la estabilidad potenciada se producen en un intervalo de temperaturas, por ejemplo, de 20 °C a 65 °C.

35 En otro aspecto de la composición, la actividad enzimática potenciada, la actividad de unión potenciada, la especificidad de unión potenciada y/o la estabilidad potenciada se producen en un intervalo de tiempos de incubación.

En otro aspecto de la composición, la actividad enzimática potenciada, la actividad de unión potenciada, la especificidad de unión potenciada y/o la estabilidad potenciada se producen en un tampón que tiene un pH en el intervalo de aproximadamente 7.3 a aproximadamente 8.6.

5 En otro aspecto de la composición, la actividad enzimática potenciada, la actividad de unión potenciada, la especificidad de unión potenciada y/o la estabilidad potenciada se producen cuando un valor de T_m del polipéptido Cas12i4 variante, el complejo binario variante o el complejo ternario variante es al menos 8 °C superior al valor de T_m del polipéptido precursor, el complejo binario precursor o el complejo ternario precursor.

10 En otro aspecto de la composición, el polipéptido Cas12i4 variante comprende un dominio RuvC o un dominio RuvC dividido.

En otro aspecto de la composición, el polipéptido precursor comprende la secuencia de la SEQ ID NO: 2.

En otro aspecto de la composición, la guía de ARN comprende una secuencia repetitiva directa y una secuencia espaciadora.

15 En otro aspecto de la composición, la repetición directa comprende:

- a. del nucleótido 1 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- b. del nucleótido 2 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una
20 secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- c. del nucleótido 3 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- 25 d. del nucleótido 4 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- e. del nucleótido 5 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118,
30 119, 120, 121, 122, 123 o 124;
- f. del nucleótido 6 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- g. del nucleótido 7 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una
35 secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

h. del nucleótido 8 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

i. del nucleótido 9 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

j. del nucleótido 10 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

k. del nucleótido 11 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

l. del nucleótido 12 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

m. del nucleótido 13 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

n. del nucleótido 14 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o

o. una secuencia que es al menos un 90 % idéntica a una secuencia de la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la composición, la repetición directa comprende:

a. del nucleótido 1 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

b. del nucleótido 2 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

c. del nucleótido 3 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

d. del nucleótido 4 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

e. del nucleótido 5 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

f. del nucleótido 6 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

g. del nucleótido 7 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

h. del nucleótido 8 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

i. del nucleótido 9 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

j. del nucleótido 10 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

k. del nucleótido 11 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

l. del nucleótido 12 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

m. del nucleótido 13 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

n. del nucleótido 14 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o

o. una secuencia que es al menos un 95% idéntica a una secuencia de la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la composición, la repetición directa comprende:

a. del nucleótido 1 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

b. del nucleótido 2 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

c. del nucleótido 3 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

d. del nucleótido 4 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

5 e. del nucleótido 5 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

f. del nucleótido 6 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

10 g. del nucleótido 7 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

h. del nucleótido 8 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

i. del nucleótido 9 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

15 j. del nucleótido 10 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

k. del nucleótido 11 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

20 l. del nucleótido 12 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

m. del nucleótido 13 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

n. del nucleótido 14 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o

25 o. la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la composición, la secuencia repetitiva directa comprende $AGN_1N_2N_3N_4GUGUN_5N_6N_7CAGN_8GACN_9C$ (SEQ ID NO: 125), en donde N_1 es A o G, N_2 es C o U, N_3 es A o G, N_4 es U o C, N_5 es C o U, N_6 es C o U, N_7 es U, A, C o G, N_8 es U o C, y N_9 es A o C.

30 En otro aspecto de la composición, la secuencia espaciadora comprende entre 15 y 35 nucleótidos de longitud.

En otro aspecto de la composición, la secuencia espaciadora comprende complementariedad con una secuencia de cadena diana de un ácido nucleico diana.

En otro aspecto de la composición, el ácido nucleico diana comprende una secuencia de cadena no diana adyacente a una secuencia de motivo adyacente al protoespaciador (PAM).

35 En otro aspecto de la composición, la secuencia de PAM es una 5'-TTN-3', 5'-NTTN-3', 5'-NTN'-3', 5'-NNTN-3', 5'-VTN-3' o 5'-NVTN-3', en donde N es cualquier nucleótido (por ejemplo, A, G, T o C) y V es A, G o C.

En otro aspecto de la composición, el polipéptido Cas12i4 variante comprende además un marcador de péptido, una proteína fluorescente, un dominio de edición de base, un dominio de metilación de ADN, un dominio de modificación de residuo de histona, un factor de localización, un factor de modificación de la transcripción, un factor de control dependiente de la luz, un factor químicamente inducible o un factor de visualización de cromatina.

La invención proporciona además una composición que comprende un ácido nucleico que codifica un polipéptido Cas12i4 como se describe en el presente documento, en donde, opcionalmente, el ácido nucleico tiene codones optimizados para la expresión en una célula.

En un aspecto de la composición, la célula es una célula eucariota.

En otro aspecto de la composición, la célula es una célula de mamífero o una célula vegetal.

En otro aspecto de la composición, la célula es una célula humana.

En otro aspecto de la composición, el ácido nucleico que codifica el polipéptido Cas12i4 variante se une operativamente a un promotor.

En otro aspecto de la composición, el ácido nucleico que codifica el polipéptido Cas12i4 variante está en un vector.

En otro aspecto de la composición, el vector comprende un vector retrovílico, un vector lentivílico, un vector fágico, un vector adenovílico, un vector adenoasociado o un vector de herpes simple.

En otro aspecto de la composición, la composición está presente en un sistema de suministro que comprende una nanopartícula, un liposoma, un exosoma, una microvesícula, o un cañón génico.

La invención proporciona además una guía de ARN o un ácido nucleico que codifica la guía de ARN, en donde la guía de ARN comprende una secuencia repetitiva directa que comprende:

a. del nucleótido 1 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

b. del nucleótido 2 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

c. del nucleótido 3 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

d. del nucleótido 4 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

e. del nucleótido 5 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

f. del nucleótido 6 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

g. del nucleótido 7 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

h. del nucleótido 8 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

i. del nucleótido 9 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

j. del nucleótido 10 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

k. del nucleótido 11 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

l. del nucleótido 12 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

m. del nucleótido 13 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

n. del nucleótido 14 al nucleótido 36 de una secuencia que es al menos un 90 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o

o. una secuencia que es al menos un 90 % idéntica a una secuencia de la SEQ ID NO: 61 o una porción de la misma.

En un aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, la repetición directa comprende:

a. del nucleótido 1 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

b. del nucleótido 2 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;

- c. del nucleótido 3 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- d. del nucleótido 4 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una
5 secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- e. del nucleótido 5 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- f. del nucleótido 6 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una
10 secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- g. del nucleótido 7 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118,
15 119, 120, 121, 122, 123 o 124;
- h. del nucleótido 8 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- i. del nucleótido 9 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una
20 secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- j. del nucleótido 10 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- k. del nucleótido 11 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- l. del nucleótido 12 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117,
30 118, 119, 120, 121, 122, 123 o 124;
- m. del nucleótido 13 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- n. del nucleótido 14 al nucleótido 36 de una secuencia que es al menos un 95 % idéntica a
35 una secuencia de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o
- o. una secuencia que es al menos un 95% idéntica a una secuencia de la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, la repetición directa comprende:

- a. del nucleótido 1 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- 5 b. del nucleótido 2 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- c. del nucleótido 3 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- d. del nucleótido 4 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 10 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- e. del nucleótido 5 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- f. del nucleótido 6 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- 15 g. del nucleótido 7 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- h. del nucleótido 8 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- i. del nucleótido 9 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 20 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- j. del nucleótido 10 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- k. del nucleótido 11 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- 25 l. del nucleótido 12 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- m. del nucleótido 13 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124;
- n. del nucleótido 14 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 30 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, o 124; o
- o. la SEQ ID NO: 61 o una porción de la misma.

En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, la secuencia repetitiva directa comprende $AGN_1N_2N_3N_4GUGUN_5N_6N_7CAGN_8GACN_9C$ (SEQ ID NO: 125), N_1 es A o G, N_2 es C o U, N_3 es A o G, N_4 es U o C, N_5 es C o U, N_6 es C o U, N_7 es U, A, C o G, N_8 es U o C, y N_9 es A 35 o C.

En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, la guía de ARN comprende además una secuencia espaciadora.

En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, la secuencia espaciadora comprende de aproximadamente 15 a aproximadamente 35 nucleótidos de longitud.

En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, la secuencia espaciadora reconoce un ácido nucleico diana.

5 En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, el ácido nucleico diana comprende una secuencia diana adyacente a una secuencia de motivo adyacente al protoespaciador (PAM), en donde la secuencia de PAM comprende una secuencia de nucleótidos expuesta como 5'-TTN-3', 5'-NTTN-3', 5'-NTN-3', 5'-NNTN-3', 5'-VTN-3' o 5'-NVTN-3', en donde N es cualquier nucleótido (por ejemplo, A, G, T o C) y V es A, G o C.

10 La invención proporciona además una composición que comprende una guía de ARN o un ácido nucleico que codifica la guía de ARN como se describe en el presente documento.

En un aspecto de la composición, la composición es una composición de suministro que comprende una nanopartícula (por ejemplo, una nanopartícula lipídica), un liposoma, un exosoma, una microvesícula o un cañón génico.

15 En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN descrita en el presente documento, el ácido nucleico que codifica la guía de ARN está unido operativamente a un promotor.

En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, el ácido nucleico que codifica la guía de ARN está en un vector.

20 En otro aspecto de la guía de ARN o el ácido nucleico que codifica la guía de ARN, el vector comprende un vector retroviral, un vector lentiviral, un vector fágico, un vector adenoviral, un vector adenoasociado o un vector de herpes simple.

La invención proporciona además una célula que comprende la guía de ARN o el ácido nucleico que codifica la guía de ARN descrita en el presente documento.

25 En un aspecto de la célula, la célula es una célula eucariota.

En otro aspecto de la célula, la célula es una célula de mamífero o una célula vegetal.

En otro aspecto de la célula, la célula es una célula humana.

30 La invención proporciona además un método para editar un gen en una célula, comprendiendo el método poner en contacto la célula con una variante, una composición, una guía de ARN o una molécula de ácido nucleico como se describe en el presente documento.

La invención proporciona además una molécula de ácido nucleico que codifica una variante de Cas12i4 de la SEQ ID NO: 4, en donde la secuencia de la molécula de ácido nucleico es un 95 % idéntica a la seleccionada del grupo que consiste en las SEQ ID NO: 222-228.

35 En una realización, la secuencia de la molécula de ácido nucleico comprende una secuencia seleccionada del grupo que consiste en las SEQ ID NO: 222-228

Definiciones

La presente invención se describirá con respecto a realizaciones particulares y con referencia a determinadas figuras, pero la invención no se limita a las mismas sino únicamente a las reivindicaciones.

En general, los términos que se exponen a continuación en el presente documento deben entenderse en su sentido común, a menos que se indique otra cosa.

Como se usa en el presente documento, el término "actividad" se refiere a una actividad biológica. En algunas realizaciones, la actividad nucleasa incluye la actividad enzimática, por ejemplo, la capacidad catalítica de una nucleasa. Por ejemplo, una actividad nucleasa puede incluir la actividad nucleasa. En algunas realizaciones, la actividad nucleasa incluye la actividad de unión, por ejemplo, la actividad de unión de una nucleasa a una guía de ARN y/o ácido nucleico diana.

Como se usa en el presente documento, el término "complejo" se refiere a una agrupación de dos o más moléculas. En algunas realizaciones, el complejo comprende un polipéptido y una molécula de ácido nucleico que interactúan (por ejemplo, se unen, entran en contacto, se adhieren) entre sí.

Como se usa en el presente documento, la expresión "complejo binario" se refiere a una agrupación de dos moléculas (por ejemplo, un polipéptido y una molécula de ácido nucleico). En algunas realizaciones, un complejo binario se refiere a una agrupación de un polipéptido y un resto de direccionamiento (por ejemplo, una guía de ARN). En algunas realizaciones, un complejo binario se refiere a una ribonucleoproteína (RNP). Como se usa en el presente documento, la expresión "complejo binario variante" se refiere a la agrupación de un polipéptido Cas12i4 variante y una guía de ARN. Como se usa en el presente documento, la expresión "complejo binario precursor" se refiere a la agrupación de un polipéptido precursor y una guía de ARN o un polipéptido de referencia y una guía de ARN.

Como se usa en el presente documento, la expresión "complejo ternario" se refiere a una agrupación de tres moléculas (por ejemplo, un polipéptido y dos moléculas de ácido nucleico). En algunas realizaciones, un "complejo ternario" se refiere a una agrupación de un polipéptido, una molécula de ARN y una molécula de ADN. En algunas realizaciones, un complejo ternario se refiere a una agrupación de un polipéptido, un resto de direccionamiento (por ejemplo, una guía de ARN) y un ácido nucleico diana (por ejemplo, una molécula de ADN diana). En algunas realizaciones, un "complejo ternario" se refiere a una agrupación de un complejo binario (por ejemplo, una ribonucleoproteína) y una tercera molécula (por ejemplo, un ácido nucleico diana).

Como se usa en el presente documento, el término "dominio" se refiere a una unidad estructural y/o funcional distinta de un polipéptido. En algunas realizaciones, un dominio puede comprender una secuencia de aminoácidos conservada.

Como se usa en el presente documento, el término "interfaz" se refiere a uno o más residuos de un polipéptido Cas12i4 variante (por ejemplo, un dominio/motivo o una porción de un dominio/motivo) en contacto con (por ejemplo, que interactúan con o son adyacentes a) una molécula de ácido nucleico o un dominio/motivo distinto o una porción de un dominio/motivo distinto del polipéptido Cas12i4 variante. En algunos aspectos, una interfaz es un área de superficie enterrada entre dominios o motivos adyacentes. En algunos aspectos, una interfaz es un área superficial entre un polipéptido y un ligando (por ejemplo, ADN o ARN) donde el polipéptido y el ligando hacen contacto. Como se usa en el presente documento, la expresión "interfaz de ácido nucleico" se refiere a residuos del polipéptido Cas12i4 variante que están muy cerca (por ejemplo, son adyacentes a) o interactúan con una secuencia de ácido nucleico (por ejemplo,

una secuencia de ADN o una secuencia de ARN). Como se usa en el presente documento, la expresión "interfaz de unión a ARN" se refiere a los residuos del polipéptido Cas12i4 variante que están muy cerca (por ejemplo, son adyacentes a) o interactúan con una guía de ARN (por ejemplo, la repetición directa de la guía de ARN). Como se usa en el presente documento, la expresión "interfaz de unión a ADN bicatenario" se refiere a los residuos del polipéptido Cas12i4 variante que están en estrecha proximidad (por ejemplo, son adyacentes a) y/o interactúan con ADN bicatenario. Como se usa en el presente documento, la expresión "interfaz de unión a ADN monocatenario" se refiere a los residuos del polipéptido Cas12i4 variante que están en estrecha proximidad (por ejemplo, son adyacentes a) y/o interactúan con ADN monocatenario. Como se usa en el presente documento, la expresión "interfaz dominio-dominio" se refiere a un dominio en estrecha proximidad a (por ejemplo, adyacente a) un dominio separado. En algunas realizaciones, se forma una interfaz dominio-dominio (por ejemplo, una interfaz dominio Helicoidal II-dominio Nuc) tras la formación del complejo (por ejemplo, la formación del complejo ternario).

Como se usan en el presente documento, los términos "precursor", "polipéptido precursor" y "secuencia precursora" se refieren a un polipéptido precursor (por ejemplo, polipéptido de partida) al que se le hace una alteración para producir un polipéptido Cas12i4 variante de la presente invención. En algunas realizaciones, el precursor es un polipéptido que tiene una secuencia de aminoácidos idéntica de la variante a una o más de las posiciones especificadas. El precursor puede ser un polipéptido de origen natural (tipo silvestre). En una realización particular, el precursor es un polipéptido con al menos un 60 %, al menos un 61 %, al menos un 62 %, al menos un 63 %, al menos un 64 %, al menos un 65 %, al menos un 70 %, al menos un 72 %, al menos un 73 %, al menos un 74 %, al menos un 75 %, al menos un 80 %, al menos un 81 %, al menos un 82 %, al menos un 83 %, al menos un 84 %, al menos un 85 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % o un 100 % de identidad con un polipéptido de la SEQ ID NO: 2.

Como se usa en el presente documento, la expresión "motivo adyacente al protoespaciador" o "PAM" se refiere a una secuencia de ADN adyacente a una "secuencia diana" a la que se une un complejo que comprende un polipéptido Cas12i4 y una guía de ARN. El "ácido nucleico diana" es una molécula bicatenaria: una cadena comprende la secuencia diana adyacente al PAM y se denomina "cadena PAM" (por ejemplo, la cadena no diana o la cadena no complementaria al espaciador) y la otra cadena complementaria se denomina la "cadena no PAM" (por ejemplo, la cadena diana o la cadena complementaria al espaciador). Como se usa en el presente documento, el término "adyacente" incluye casos en los que una guía de ARN del complejo se une, interactúa con o se asocia específicamente a una secuencia diana que está inmediatamente adyacente a un PAM. En dichos casos, no hay nucleótidos entre la secuencia diana y el PAM. El término "adyacente" también incluye casos en los que hay un número pequeño (por ejemplo, 1, 2, 3, 4 o 5) de nucleótidos entre la secuencia diana, a la que se une el resto de direccionamiento, y el PAM.

Como se usan en el presente documento, las expresiones "composición de referencia", "molécula de referencia", "secuencia de referencia" y "referencia" se refieren a un control, tal como un control negativo

o un precursor (por ejemplo, una secuencia precursora, una proteína precursora o una proteína de tipo silvestre). Por ejemplo, una molécula de referencia se refiere a un polipéptido con el que se compara un polipéptido Cas12i4 variante. Análogamente, una guía de ARN de referencia se refiere a un resto de direccionamiento con el que se compara una guía de ARN modificada. La molécula variante o modificada puede compararse con la molécula de referencia basándose en la secuencia (por ejemplo, la molécula variante o modificada puede tener el X % de identidad u homología de secuencia con la molécula de referencia), la termoestabilidad o la actividad (por ejemplo, la molécula variante o modificada puede tener el X % de la actividad de la molécula de referencia). Por ejemplo, una molécula variante o modificada puede caracterizarse por tener no más del 10 % de una actividad del polipéptido de referencia o puede caracterizarse por tener al menos el 10 % más de una actividad del polipéptido de referencia. Los ejemplos de polipéptidos de referencia incluyen polipéptidos sin modificar de origen natural, por ejemplo, polipéptidos de origen natural de arqueas u otras especies bacterianas. En determinadas realizaciones, el polipéptido de referencia es un polipéptido de origen natural que tiene la identidad de secuencia u homología más cercana con el polipéptido Cas12i4 variante con el que se está comparando. En determinadas realizaciones, el polipéptido de referencia es una molécula precursora que tiene una secuencia de origen natural o conocida en la que se ha realizado una mutación para llegar al polipéptido Cas12i4 variante.

Como se usan en el presente documento, las expresiones "guía de ARN" o "secuencia guía de ARN" se refieren a cualquier molécula de ARN que facilita el direccionamiento de un polipéptido Cas12i4 descrito en el presente documento a un ácido nucleico diana. Por ejemplo, una guía de ARN puede ser una molécula que reconoce (por ejemplo, se une a) un ácido nucleico diana. Una guía de ARN puede diseñarse para que sea complementaria a una cadena diana (por ejemplo, la cadena no PAM) de una secuencia de ácido nucleico diana. Una guía de ARN comprende una secuencia de direccionamiento de ADN y una secuencia repetitiva directa (DR, por sus siglas en inglés). Las expresiones ARN de CRISPR (ARNcr), pre-ARNcr, ARNcr maduro y ARNg también se usan en el presente documento para hacer referencia a una guía de ARN. Como se usa en el presente documento, la expresión "pre-ARNcr" se refiere a una molécula de ARN sin procesar que comprende una secuencia DR-espaciador-DR. Como se usa en el presente documento, la expresión "ARNcr maduro" se refiere a una forma procesada de un pre-ARNcr; un ARNcr maduro puede comprender una secuencia DR-espaciador, en donde la DR es una forma truncada de la DR de un pre-ARNcr y/o el espaciador es una forma truncada del espaciador de un pre-ARNcr.

Como se usa en el presente documento, la expresión "sustancialmente idéntico/a" se refiere a una secuencia, polinucleótido o polipéptido, que tiene un determinado grado de identidad con una secuencia de referencia.

Como se usan en el presente documento, las expresiones "ácido nucleico diana", "secuencia diana" y "sustrato diana" se refieren a un ácido nucleico al que se le une específicamente una guía de ARN. En algunas realizaciones, la secuencia de direccionamiento de ADN de una guía de ARN se une a un ácido nucleico diana.

Como se usan en el presente documento, las expresiones "polipéptido Cas12i4 variante" y "polipéptido nucleasa variante" se refieren a un polipéptido que comprende una alteración, por ejemplo,

una sustitución, inserción, supresión y/o fusión, a uno o más posiciones de residuos, en comparación con un polipéptido precursor. Como se usan en el presente documento, las expresiones "polipéptido Cas12i4 variante" y "polipéptido nucleasa variante" se refieren a un polipéptido que comprende una alteración en comparación con el polipéptido de la SEQ ID NO: 2.

5

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **FIG. 1A** es un gel de EMSA de ADN que muestra la capacidad de las RNP preparadas con a) Cas12i4 de tipo silvestre (SEQ ID NO: 2) o Cas12i4 variante de la SEQ ID NO: 4 y b) una guía de ARN de la SEQ ID NO: 62 para unirse a una diana de ADNbc de AAVS1 (SEQ ID NO: 65). Se indican las bandas de ADNbc no unido.

10

La **FIG. 1B** es un gel de EMSA de ADN que muestra la capacidad de las RNP preparadas con a) Cas12i4 de tipo silvestre (SEQ ID NO: 2) o Cas12i4 variante de la SEQ ID NO: 4 y b) una guía de ARN de la SEQ ID NO: 63 para unirse a una diana de ADNbc de AAVS1 (SEQ ID NO: 66). Se indican las bandas de ADNbc unido y ADNbc no unido.

15

La **FIG. 1C** es un gel de EMSA de ADN que muestra la capacidad de las RNP preparadas con a) Cas12i4 de tipo silvestre (SEQ ID NO: 2) o Cas12i4 variante de la SEQ ID NO: 4 y b) una guía de ARN de la SEQ ID NO: 64 para unirse a una diana de ADNbc de EMX1 (SEQ ID NO: 67). Se indican las bandas de ADNbc unido y ADNbc no unido.

20

La **FIG. 1D** es un gel de EMSA de ADN de control que muestra la capacidad de las RNP preparadas con a) Cas12i4 de tipo silvestre (SEQ ID NO: 2) o Cas12i4 variante de la SEQ ID NO: 4 y b) una guía de ARN de la SEQ ID NO: 62 para unirse a una diana de ADNbc de EMX1 (SEQ ID NO: 67). Se indican las bandas de ADNbc no unido.

25

La **FIG. 2A** es un gel que muestra la escisión de una diana de ADNbc de AAVS1 (SEQ ID NO: 65) por RNP preparadas con a) Cas12i4 de tipo silvestre (SEQ ID NO: 2) o Cas12i4 variante de la SEQ ID NO: 4 y b) una guía de ARN de la SEQ ID NO: 62. Se indican las bandas de ADN de longitud completa y escindidas.

30

La **FIG. 2B** es un gel que muestra la escisión de una diana de ADNbc de AAVS1 (SEQ ID NO: 66) por RNP preparadas con a) Cas12i4 de tipo silvestre (SEQ ID NO: 2) o Cas12i4 variante de la SEQ ID NO: 4 y b) una guía de ARN de la SEQ ID NO: 63. Se indican las bandas de ADN de longitud completa y escindidas.

35

La **FIG. 2C** es un gel que muestra la escisión de una diana de ADNbc de EMX1 (SEQ ID NO: 67) por RNP preparadas con a) Cas12i4 de tipo silvestre (SEQ ID NO: 2) o Cas12i4 variante de la SEQ ID NO: 4 y b) una guía de ARN de la SEQ ID NO: 64. Se indican las bandas de ADN de longitud completa y escindidas.

La **FIG. 3** es un gráfico que muestra indels inducidos en dianas de AAVS1, EMX1 y VEGFA (SEQ ID NO: 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105 y 107) por Cas12i4 de tipo silvestre (SEQ ID NO: 2) y las variantes de Cas12i4 de la SEQ ID NO: 3 y la SEQ ID NO: 4 en células de mamíferos.

La **FIG. 4** es un gráfico que muestra indels inducidos en dianas de AAVS1, EMX1 y VEGFA adyacentes a secuencias PAM 5'-NTTN-3' o 5'-NVTN-3' por Cas12i4 de tipo silvestre (SEQ ID NO: 2) y la variante de Cas12i4 de la SEQ ID NO: 4 en células de mamíferos.

La **FIG. 5** es un esquema que muestra la estructura del dominio del polipéptido Cas12i4.

La **FIG. 6A** representa la ubicación de la sustitución V592R en la estructura de Cas12i4. La sustitución V592R puede interactuar con la cadena no diana monocatenaria.

La **FIG. 6B** representa las ubicaciones de las sustituciones E480R y G564R en la estructura de Cas12i4, que están cerca de la secuencia de PAM del ADN bicatenario. Las sustituciones E480R y G564R pueden estabilizar interacciones con ADN bicatenario.

DESCRIPCIÓN DETALLADA

La presente divulgación se refiere a variantes novedosas del polipéptido de la SEQ ID NO: 2 y a métodos de producción y uso de las mismas. La presente divulgación se refiere además a complejos que comprenden una variante del efector de la SEQ ID NO: 2 y a métodos de producción y uso de las mismas. En algunos aspectos, en el presente documento se describe una composición que comprende un complejo que tiene una o más características. En algunos aspectos, se describe un método de suministro de una composición que comprende el complejo.

COMPOSICIONES

En algunas realizaciones, una composición de la invención incluye un polipéptido Cas12i4 variante que presenta una actividad enzimática potenciada, una actividad de unión potenciada, una especificidad de unión potenciada y/o una estabilidad potenciada, con respecto a un polipéptido precursor. En algunas realizaciones, una composición de la invención incluye un complejo que comprende un polipéptido Cas12i4 variante que presenta una actividad enzimática potenciada, una actividad de unión potenciada, una especificidad de unión potenciada y/o una estabilidad potenciada con respecto a un complejo precursor.

En algunas realizaciones, una composición de la invención incluye un polipéptido Cas12i4 variante y una guía de ARN. En algunas realizaciones, una composición de la invención incluye un complejo binario variante que comprende un polipéptido Cas12i4 variante y una guía de ARN.

En algunos aspectos de la composición, el polipéptido Cas12i4 variante tiene una formación de complejos aumentada (por ejemplo, una formación de complejos binarios aumentada) con la guía de ARN en comparación con un polipéptido precursor. En algunos aspectos de la composición, el polipéptido Cas12i4 variante y la guía de ARN tienen una mayor afinidad de unión, en comparación con un polipéptido precursor y la guía de ARN. En algunos aspectos de la composición, el polipéptido Cas12i4 variante y la guía de ARN tienen interacciones proteína-ARN más fuertes (por ejemplo, interacciones iónicas), en comparación con un polipéptido precursor y la guía de ARN. En algunos aspectos de la composición, el complejo binario variante es más estable que un complejo binario precursor.

En algunas realizaciones, una composición de la invención incluye un polipéptido Cas12i4 variante, una guía de ARN y un ácido nucleico diana. En algunas realizaciones, una composición de la invención incluye un complejo ternario variante que comprende un polipéptido Cas12i4 variante, una guía de ARN y un ácido nucleico diana.

En algunos aspectos de la composición, el polipéptido Cas12i4 variante tiene una formación de complejos aumentada (por ejemplo, una formación de complejos ternarios aumentada) con la guía de ARN y el ácido nucleico diana en comparación con un polipéptido precursor. En algunos aspectos de la composición, el polipéptido Cas12i4 variante y la guía de ARN (por ejemplo, el complejo binario variante) tienen una mayor afinidad de unión a un ácido nucleico diana, en comparación con un polipéptido precursor y la guía de ARN (por ejemplo, un complejo binario precursor). En algunos aspectos de la composición, el complejo ternario variante es más estable que un complejo ternario precursor.

Polipéptido Cas12i4 Variante

En algunas realizaciones, la composición de la presente invención incluye un polipéptido Cas12i4 variante descrito en el presente documento.

En algunas realizaciones, el polipéptido de la presente invención es una variante de un polipéptido precursor, en donde el precursor es codificado por un polinucleótido que comprende una secuencia de nucleótidos tal como la SEQ ID NO: 1 o comprende una secuencia de aminoácidos tal como la SEQ ID NO: 2.

Tabla 1. Secuencias precursoras.

Identificador de secuencia	Secuencia	Descripción
SEQ ID NO: 1	ATGGCTTCCATCTCTAGGCCATACGGCACCAAGCTGCGACCG GACGCACGGAAGAAGGAGATGCTCGATAAGTTCTTTAATACA CTGACTAAGGGTCAGCGCGTGTTGCGAGACCTGGCCCTGTG CATCTATGGCTCCCTGACCCTGGAGATGGCCAAGTCTCTGGA GCCAGAAAGTGATTGAGAACTGGTGTGCGCTATTGGGTGGTT TCGGCTGGTGGACAAGACCATCTGGTCCAAGGATGGCATCA AGCAGGAGAATCTGGTGAAACAGTACGAAGCCTATTCCGGAA AGGAGGCTTCTGAAGTGGTCAAAACATACCTGAACAGCCCCA GCTCCGACAAGTACGTGTGGATCGATTGCAGGCAGAAATTCC TGAGGTTTCAGCGCGAGCTCGGCACTCGCAACCTGTCCGAG GACTTCGAATGTATGCTCTTTGAACAGTACATTAGACTGACCA AGGGCGAGATCGAAGGGTATGCCGCTATTTCAAATATGTTTCG GAAACGGCGAGAAGGAAGACCGGAGCAAGAAAAGAATGTAC GCTACACGGATGAAAGATTGGCTGGAGGCAAACGAAAATATC ACTTGGGAGCAGTATAGAGAGGCCCTGAAGAACCAGCTGAAT GCTAAAAACCTGGAGCAGGTTGTGGCCAATTACAAGGGGAAC GCTGGCGGGGCGAGACCCCTTCTTTAAGTATAGCTTCTCCAAA GAGGGAATGGTGAGCAAGAAAGAACATGCACAGCAGCTCGA CAAGTTCAAACCGTCCTGAAGAACAAGCCCGGGACCTGAA TTTTCCAAACAAGGAGAAGCTGAAGCAGTACCTGGAGGCCGA	Secuencia de nucleótidos que codifica el polipéptido precursor

5	AATCGGCATTCCGGTCGACGCTAACGTGTACTCCCAGATGTT CTCTAACGGGGTGAGTGAGGTCCAGCCTAAGACCACACGGA ATATGTCTTTTAGTAACGAGAACTGGATCTGCTCACTGAACT GAAGGACCTGAACAAGGGCGATGGGTTTCGAGTACGCCAGAG AAGTGCTGAACGGGTTCTTTGACTCCGAGCTCCACACTACCG AGGATAAGTTTAATATCACCTCTAGGTACCTGGGAGGCGACA AATCAAACCGCCTGAGCAAACCTCTATAAGATCTGGAAGAAAG AGGGTGTGGACTGCGAGGAAGGCATTACAGCAGTTCTGTGAA GCCGTCAAAGATAAGATGGGCCAGATCCCCATTGAAATGTG 10 CTGAAGTACCTGTGGCAGTTCCGGGAGACAGTCAGTGCCGA GGATTTTGAAGCAGCCGCTAAGGCTAACCATCTGGAGGAAAA GATCAGCCGGGTGAAAGCCCACCCAATCGTGATTAGCAATAG GTACTGGGCTTTTGGGACTTCCGCACTGGTGGGAAACATTAT GCCCCGAGACAAGAGGCATCAGGGAGAGTATGCCGGTCAGA 15 ATTTCAAATGTGGCTGGAGGCTGAACTGCACTACGATGGCA AGAAAGCAAAGCACCATCTGCCTTTTTATAACGCCCGCTTCTT TGAGGAAGTGTACTGCTATCACCCCTCTGTGCCGAGATCAC TCCTTTCAAAACCAAGCAGTTTGGCTGTGAAATCGGGAAGGA CATTCCAGATTACGTGAGCGTCGCTCTGAAGGACAATCCGTA 20 TAAGAAAGCAACCAAACGAATCCTGCGTGCAATCTACAATCC CGTCGCCAACACAACCTGGCGTTGATAAGACCACAAACTGCAG CTTCATGATCAAACGCGAGAATGACGAATATAAGCTGGTCAT CAACCGAAAAATTTCCGTGGATCGGCCTAAGAGAATCGAAGT GGGCAGGACAATTATGGGGTACGACCGCAATCAGACAGCTA 25 GCGATACTTATTGGATTGGCCGGCTGGTGCCACCTGGAACC CGGGGCGCATACCGCATCGGAGAGTGGAGCGTCCAGTATAT TAAGTCCGGGCCTGTCTGTCTAGTACTCAGGGAGTTAACA TTCCACTACCGACCAGCTGGTGTACAACGGCATGCCATCAAG CTCCGAGCGGTTCAAGGCCTGGAAGAAAGCCAGAATGGCTT 30 TTATCCGAAAACCTATTCGTACAGCTGAATGACGAGGGACTGG AATCTAAGGGTCAGGATTATATCCCCGAGAACCCTTCTAGTTT CGATGTGCGGGGCGAAACCCTGTACGTCTTTAACAGTAATTA TCTGAAGGCCCTGGTGAGCAAACACAGAAAGGCCAAGAAAC CTGTTGAGGGGATCCTGGACGAGATTGAAGCCTGGACATCTA 35 AAGACAAGGATTCATGCAGCCTGATGCGGCTGAGCAGCCTG AGCGATGCTTCCATGCAGGGAATCGCCAGCCTGAAGAGTCT GATTAACAGCTACTTCAACAAGAATGGCTGTAAAACCATCGA	
---	--	--

5	GGACAAAGAAAAGTTTAATCCCGTGCTGTATGCCAAGCTGGT TGAGGTGGAACAGCGGAGAACAAACAAGCGGTCTGAGAAAG TGGGAAGAATCGCAGGTAGTCTGGAGCAGCTGGCCCTGCTG AACGGGGTTGAGGTGGTCATCGGCGAAGCTGACCTGGGGGA GGTCGAAAAAGGAAAGAGTAAGAAACAGAATTCACGGAACAT GGATTGGTGCGCAAAGCAGGTGGCACAGCGGCTGGAGTACA AACTGGCCTTCCATGGAATCGGTTACTTTGGAGTGAACCCCA TGTATACCAGCCACCAGGACCCTTTCGAACATAGGCGCGTGG 10 CTGATCACATCGTCATGCGAGCACGTTTTGAGGAAGTCAACG TGGAGAACATTGCCGAATGGCACGTGCGAAATTTCTCAAAC ACCTGCGTGCAGACAGCGGCACTGGGCTGTACTATAAGCAG GCCACCATGGACTTCCTGAAACATTACGGTCTGGAGGAACAC GCTGAGGGCCTGGAAAATAAGAAAATCAAGTTCTATGACTTTA 15 GAAAGATCCTGGAGGATAAAAACCTGACAAGCGTGATCATTC CAAAGAGGGGCGGGCGCATCTACATGGCCACCAACCCAGTG ACATCCGACTCTACCCCGATTACATACGCCGGCAAGACTTAT AATAGGTGTAACGCTGATGAGGTGGCAGCCGCTAATATCGTT ATTCTGTGCTGGCTCCCCGCAGTAAGAAAAACGAGGAACAG 20 GACGATATCCCTCTGATTACCAAGAAAGCCGAGAGTAAGTCA CCACCGAAAGACCGGAAGAGATCAAAAACAAGCCAGCTGCC TCAGAAA	
25	SEQ ID NO: 2 MASISRPYGTKLRPDARKKEMLDKFFNTLTKGQRVFADLALCIY GSLTLEMAKSLEPESDSELVCAIGWFRLVDKTIWSKDGIKQENLV KQYEAYSGKEASEVVKTYLNSPSSDKYVWIDCRQKFLRFQREL GTRNLSEDFECMLFEQYIRLTKGEIEGYAAISNMFGNGEKEDRS KKRMYATRMKDWLEANENITWEQYREALKNQLNAKNLEQVVAN YKGNAGGADPFFKYFSKEGMVSKKEHAQQLDKFKTVLKNKAR DLNFPNKEKLKQYLEAEIGIPVDANVYSQMFSNGVSEVQPKTTR 30 NMSFSNEKLDLLTELKDLNKGDFEYAREVLNGFFDSELHTTED KFNITSRYLGGDKSNRLSKLYKIWKKEGVDC EEGIQQFCEAVKD KMGQIPIRNVLKYLWQFRETVSAEDFEAAAKANHLEEKISRVKA HPIVISNRYWAFGTSALVGNIMPADKRHQGEYAGQNFKMWLEA ELHYDGKKAKHHLFPFYNARFFEEVYCYHPSVAEITPFKTKQFGC EIGKDIPDYVSVALKDNPYKKATKRILRAIYNPVANTTGVDKTTNC 35 SFMIKRENDEYKLVINRKISVDRPKRIEVGRTIMGYDRNQTASDT YWIGRLVPPGTRGAYRIGESVQYIKSGPVLSSSTQGVNNSTTD QLVYNGMPSSSERFKAWKKARMAFIRKLIRQLNDEGLESKGQD	Polipéptido precursor

5	YIPENPSSFDVRGETLYVFNSNYLKALVSKHRKAKKPVEGILDEI EAWTSKDKDSCSLMRLSSLSDASMQGIASLKSLINSYFNKNGCK TIEDKEKFNPVLYAKLVEVEQRRTNKRSEKVGRIAGSLEQLALLN GVEVVIGEADLGEVEKGKSKKQNSRNMDWCAKQVAQRLEYKL AFHGIGYFGVNPMYTSHQDPFEHRRVADHIVMRARFEEVNVENI AEWHVRNFSNYLRADSGTGLYYKQATMDFLKHYLEEHAEGLE NKKIKFYDFRKILEDKNLTSVIIPKRGGRIYMATNPVTSDSTPITYA GKTYNRCNADEVAAANIVISVLAPRSKKNEEQDDIPLITKKAESK SPPKDRKRSKTSQLPQK	
---	---	--

Una secuencia de ácido nucleico que codifica el polipéptido precursor descrito en el presente documento puede ser sustancialmente idéntica a una secuencia de ácido nucleico de referencia, por ejemplo, la SEQ ID NO: 1. En algunas realizaciones, el polipéptido Cas12i4 variante está codificado por un ácido nucleico que comprende una secuencia que tiene al menos aproximadamente un 60 %, al menos aproximadamente un 65 %, al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 91 %, al menos aproximadamente un 92 %, al menos aproximadamente un 93 %, al menos aproximadamente un 94 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 % o al menos aproximadamente un 99.5 % de identidad de secuencia con la secuencia de ácido nucleico de referencia, por ejemplo, la secuencia de ácido nucleico que codifica el polipéptido precursor, por ejemplo, la SEQ ID NO: 1. El porcentaje de identidad entre dos de dichos ácidos nucleicos puede determinarse manualmente mediante inspección de las dos secuencias de ácido nucleico alineadas de manera óptima o usando programas de software o algoritmos (por ejemplo, BLAST, ALIGN, CLUSTAL) usando parámetros convencionales. Una indicación de que dos secuencias de ácido nucleico son sustancialmente idénticas es que las moléculas de ácido nucleico se hibridan con la secuencia complementaria de la otra en condiciones rigurosas (por ejemplo, dentro de un intervalo de rigurosidad media a alta).

En algunas realizaciones, el polipéptido Cas12i4 variante está codificado por una secuencia de ácido nucleico que tiene al menos aproximadamente un 60 %, al menos aproximadamente un 65 %, al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 91 %, al menos aproximadamente un 92 %, al menos aproximadamente un 93 %, al menos aproximadamente un 94 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 % o más de identidad de secuencia, pero no un 100 % de identidad de secuencia, con una secuencia de ácido nucleico de referencia, por ejemplo, la secuencia de ácido nucleico que codifica el polipéptido precursor, por ejemplo, la SEQ ID NO: 1.

En algunas realizaciones, el polipéptido Cas12i4 variante de la presente invención comprende una secuencia polipeptídica que tiene un 50 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 %, o un 99 %, pero no un 100 %, de identidad con la SEQ ID NO: 2. En algunas realizaciones, el polipéptido Cas12i4 variante de la presente invención comprende una secuencia polipeptídica que tiene más de un 50 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 %, un 91 %, un 92 %, un 93 %, un 94 %, un 95 %, un 96 %, un 97 %, un 98 %, o un 99 %, pero no un 100 %, de identidad con la SEQ ID NO: 2. En algunas realizaciones, el polipéptido Cas12i4 variante mantiene los cambios de aminoácidos (o al menos 1, 2, 3, 4, 5, etc. de estos cambios) que diferencian el polipéptido de su respectiva secuencia precursora/de referencia.

En algunas realizaciones, la presente invención describe un polipéptido Cas12i4 variante que tiene un grado especificado de identidad de secuencia de aminoácidos con uno o más polipéptidos de referencia, por ejemplo, un polipéptido precursor, por ejemplo, al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 % o incluso al menos un 99 %, pero no un 100 %, de identidad de secuencia con la secuencia de aminoácidos de la SEQ ID NO: 2. La homología o identidad puede determinarse mediante alineación de la secuencia de aminoácidos, por ejemplo, usando un programa tal como BLAST, ALIGN o CLUSTAL, como se describe en el presente documento. En algunas realizaciones, el polipéptido Cas12i4 variante mantiene los cambios de aminoácidos (o al menos 1, 2, 3, 4, 5, etc. de estos cambios) que diferencian el polipéptido de su respectiva secuencia precursora/de referencia.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración a uno o más (por ejemplo, varios) aminoácidos de un polipéptido precursor, en donde al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 162, 164, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 193, 194, 195, 196, 197, 198, 199, 200 o más están alterados. En algunas realizaciones, el polipéptido Cas12i4 variante mantiene los cambios de aminoácidos (o al menos 1, 2, 3, 4, 5, etc. de estos cambios) que diferencian el polipéptido de su respectiva secuencia precursora/de referencia.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una o más de las sustituciones de aminoácidos enumeradas en la Tabla 2.

Tabla 2. Sustituciones de un solo aminoácido en el polipéptido Cas12i4 variante.

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	1	M	R, G, A, K, Q, N, H
	2	A	R, G, K, Q, N, H
	3	S	R, G, A, K, Q, N, H
	4	I	R, G, A, K, Q, N, H
	5	S	R, G, A, K, Q, N, H
10	6	R	G, A, K, Q, N, H
	7	P	R, G, A, K, Q, N, H
	8	Y	R, G, A, K, Q, N, H
	9	G	R, A, K, Q, N, H
	10	T	R, G, A, K, Q, N, H
15	11	K	R, G, A, Q, N, H
	12	L	R, G, A, K, Q, N, H
	13	R	G, A, K, Q, N, H
	14	P	R, G, A, K, Q, N, H
	15	D	R, G, A, K, Q, N, H
20	16	A	R, G, K, Q, N, H
	17	R	G, A, K, Q, N, H
	18	K	R, G, A, Q, N, H
	19	K	R, G, A, Q, N, H
	20	E	R, G, A, K, Q, N, H
25	21	M	R, G, A, K, Q, N, H
	22	L	R, G, A, K, Q, N, H
	23	D	R, G, A, K, Q, N, H
	24	K	R, G, A, Q, N, H
	25	F	R, G, A, K, Q, N, H
30	26	F	R, G, A, K, Q, N, H
	27	N	R, G, A, K, Q, H
	28	T	R, G, A, K, Q, N, H
	29	L	R, G, A, K, Q, N, H
	30	T	R, G, A, K, Q, N, H
35	31	K	R, G, A, Q, N, H
	32	G	R, A, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
33	Q	R, G, A, K, N, H
34	R	G, A, K, Q, N, H
35	V	R, G, A, K, Q, N, H
36	F	R, G, A, K, Q, N, H
37	A	R, G, K, Q, N, H
38	D	R, G, A, K, Q, N, H
39	L	R, G, A, K, Q, N, H
40	A	R, G, K, Q, N, H
41	L	R, G, A, K, Q, N, H
42	C	R, G, A, K, Q, N, H
43	I	R, G, A, K, Q, N, H
44	Y	R, G, A, K, Q, N, H
45	G	R, A, K, Q, N, H
46	S	R, G, A, K, Q, N, H
47	L	R, G, A, K, Q, N, H
48	T	R, G, A, K, Q, N, H
49	L	R, G, A, K, Q, N, H
50	E	R, G, A, K, Q, N, H
51	M	R, G, A, K, Q, N, H
52	A	R, G, K, Q, N, H
53	K	R, G, A, Q, N, H
54	S	R, G, A, K, Q, N, H
55	L	R, G, A, K, Q, N, H
56	E	R, G, A, K, Q, N, H
57	P	R, G, A, K, Q, N, H
58	E	R, G, A, K, Q, N, H
59	S	R, G, A, K, Q, N, H
60	D	R, G, A, K, Q, N, H
61	S	R, G, A, K, Q, N, H
62	E	R, G, A, K, Q, N, H
63	L	R, G, A, K, Q, N, H
64	V	R, G, A, K, Q, N, H

Tabla 2		
	Posición	Residuo de tipo silvestre
5	65	C
	66	A
	67	I
	68	G
10	69	W
	70	F
	71	R
	72	L
15	73	V
	74	D
	75	K
	76	T
20	77	I
	78	W
	79	S
	80	K
25	81	D
	82	G
	83	I
	84	K
30	85	Q
	86	E
	87	N
	88	L
35	89	V
	90	K
	91	Q
	92	Y
	93	E
	94	A
	95	Y
	96	S

Tabla 2		
	Posición	Residuo de tipo silvestre
	97	G
	98	K
	99	E
	100	A
	101	S
	102	E
	103	V
	104	V
	105	K
	106	T
	107	Y
	108	L
	109	N
	110	S
	111	P
	112	S
	113	S
	114	D
	115	K
	116	Y
	117	V
	118	W
	119	I
	120	D
	121	C
	122	R
	123	Q
	124	K
	125	F
	126	L
	127	R
	128	F

Tabla 2		
	Posición	Residuo de tipo silvestre
5	129	Q
	130	R
	131	E
	132	L
10	133	G
	134	T
	135	R
	136	N
15	137	L
	138	S
	139	E
	140	D
20	141	F
	142	E
	143	C
	144	M
25	145	L
	146	F
	147	E
	148	Q
30	149	Y
	150	I
	151	R
	152	L
35	153	T
	154	K
	155	G
	156	E
	157	I
	158	E
	159	G
	160	Y

Tabla 2		
	Posición	Residuo de tipo silvestre
	161	A
	162	A
	163	I
	164	S
	165	N
	166	M
	167	F
	168	G
	169	N
	170	G
	171	E
	172	K
	173	E
	174	D
	175	R
	176	S
	177	K
	178	K
	179	R
	180	M
	181	Y
	182	A
	183	T
	184	R
	185	M
	186	K
	187	D
	188	W
	189	L
	190	E
	191	A
	192	N

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	193	E	R, G, A, K, Q, N, H
	194	N	R, G, A, K, Q, H
	195	I	R, G, A, K, Q, N, H
	196	T	R, G, A, K, Q, N, H
	197	W	R, G, A, K, Q, N, H
10	198	E	R, G, A, K, Q, N, H
	199	Q	R, G, A, K, N, H
	200	Y	R, G, A, K, Q, N, H
	201	R	G, A, K, Q, N, H
	202	E	R, G, A, K, Q, N, H
15	203	A	R, G, K, Q, N, H
	204	L	R, G, A, K, Q, N, H
	205	K	R, G, A, Q, N, H
	206	N	R, G, A, K, Q, H
	207	Q	R, G, A, K, N, H
20	208	L	R, G, A, K, Q, N, H
	209	N	R, G, A, K, Q, H
	210	A	R, G, K, Q, N, H
	211	K	R, G, A, Q, N, H
	212	N	R, G, A, K, Q, H
25	213	L	R, G, A, K, Q, N, H
	214	E	R, G, A, K, Q, N, H
	215	Q	R, G, A, K, N, H
	216	V	R, G, A, K, Q, N, H
	217	V	R, G, A, K, Q, N, H
30	218	A	R, G, K, Q, N, H
	219	N	R, G, A, K, Q, H
	220	Y	R, G, A, K, Q, N, H
	221	K	R, G, A, Q, N, H
	222	G	R, A, K, Q, N, H
35	223	N	R, G, A, K, Q, H
	224	A	R, G, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
225	G	R, A, K, Q, N, H
226	G	R, A, K, Q, N, H
227	A	R, G, K, Q, N, H
228	D	R, G, A, K, Q, N, H
229	P	R, G, A, K, Q, N, H
230	F	R, G, A, K, Q, N, H
231	F	R, G, A, K, Q, N, H
232	K	R, G, A, Q, N, H
233	Y	R, G, A, K, Q, N, H
234	S	R, G, A, K, Q, N, H
235	F	R, G, A, K, Q, N, H
236	S	R, G, A, K, Q, N, H
237	K	R, G, A, Q, N, H
238	E	R, G, A, K, Q, N, H
239	G	R, A, K, Q, N, H
240	M	R, G, A, K, Q, N, H
241	V	R, G, A, K, Q, N, H
242	S	R, G, A, K, Q, N, H
243	K	R, G, A, Q, N, H
244	K	R, G, A, Q, N, H
245	E	R, G, A, K, Q, N, H
246	H	R, G, A, K, Q, N
247	A	R, G, K, Q, N, H
248	Q	R, G, A, K, N, H
249	Q	R, G, A, K, N, H
250	L	R, G, A, K, Q, N, H
251	D	R, G, A, K, Q, N, H
252	K	R, G, A, Q, N, H
253	F	R, G, A, K, Q, N, H
254	K	R, G, A, Q, N, H
255	T	R, G, A, K, Q, N, H
256	V	R, G, A, K, Q, N, H

Tabla 2		
	Posición	Residuo de tipo silvestre
5	257	L
	258	K
	259	N
	260	K
10	261	A
	262	R
	263	D
	264	L
15	265	N
	266	F
	267	P
	268	N
20	269	K
	270	E
	271	K
	272	L
25	273	K
	274	Q
	275	W
	276	L
30	277	E
	278	A
	279	E
	280	I
35	281	G
	282	I
	283	P
	284	V
	285	D
	286	A
	287	N
	288	V

Tabla 2		
	Posición	Residuo de tipo silvestre
	289	Y
	290	S
	291	Q
	292	M
	293	F
	294	S
	295	N
	296	G
	297	V
	298	S
	299	E
	300	V
	301	Q
	302	P
	303	K
	304	T
	305	T
	306	R
	307	N
	308	M
	309	S
	310	F
	311	S
	312	N
	313	E
	314	K
	315	L
	316	D
	317	L
	318	L
	319	T
	320	E

Tabla 2		
	Posición	Residuo de tipo silvestre
5	321	L
	322	K
	323	D
	324	L
10	325	N
	326	K
	327	G
	328	D
	329	G
15	330	F
	331	E
	332	Y
	333	A
	334	R
20	335	E
	336	V
	337	L
	338	N
	339	G
25	340	F
	341	F
	342	D
	343	S
	344	E
30	345	L
	346	H
	347	T
	348	T
35	349	E
	350	D
	351	K
	352	F

Tabla 2		
	Posición	Residuo de tipo silvestre
	353	N
	354	I
	355	T
	356	S
	357	R
	358	Y
	359	L
	360	G
	361	G
	362	D
	363	K
	364	S
	365	N
	366	R
	367	L
	368	S
	369	K
	370	L
	371	Y
	372	K
	373	I
	374	W
	375	K
	376	K
	377	E
	378	G
	379	V
	380	D
	381	C
	382	E
	383	E
	384	G

Tabla 2		
	Posición	Residuo de tipo silvestre
5	385	I
	386	Q
	387	Q
	388	F
10	389	C
	390	E
	391	A
	392	V
15	393	K
	394	D
	395	K
	396	M
20	397	G
	398	Q
	399	I
	400	P
25	401	I
	402	R
	403	N
	404	V
30	405	L
	406	K
	407	Y
	408	L
35	409	W
	410	Q
	411	F
	412	R
	413	E
	414	T
	415	V
	416	S

Tabla 2		
	Posición	Residuo de tipo silvestre
	417	A
	418	E
	419	D
	420	F
	421	E
	422	A
	423	A
	424	A
	425	K
	426	A
	427	N
	428	H
	429	L
	430	E
	431	E
	432	K
	433	I
	434	S
	435	R
	436	V
	437	K
	438	A
	439	H
	440	P
	441	I
	442	V
	443	I
	444	S
	445	N
	446	R
	447	Y
	448	W

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	449	A	R, G, K, Q, N, H
	450	F	R, G, A, K, Q, N, H
	451	G	R, A, K, Q, N, H
	452	T	R, G, A, K, Q, N, H
10	453	S	R, G, A, K, Q, N, H
	454	A	R, G, K, Q, N, H
	455	L	R, G, A, K, Q, N, H
	456	V	R, G, A, K, Q, N, H
15	457	G	R, A, K, Q, N, H
	458	N	R, G, A, K, Q, H
	459	I	R, G, A, K, Q, N, H
	460	M	R, G, A, K, Q, N, H
20	461	P	R, G, A, K, Q, N, H
	462	A	R, G, K, Q, N, H
	463	D	R, G, A, K, Q, N, H
	464	K	R, G, A, Q, N, H
25	465	R	G, A, K, Q, N, H
	466	H	R, G, A, K, Q, N
	467	Q	R, G, A, K, N, H
	468	G	R, A, K, Q, N, H
30	469	E	R, G, A, K, Q, N, H
	470	Y	R, G, A, K, Q, N, H
	471	A	R, G, K, Q, N, H
	472	G	R, A, K, Q, N, H
35	473	Q	R, G, A, K, N, H
	474	N	R, G, A, K, Q, H
	475	F	R, G, A, K, Q, N, H
	476	K	R, G, A, Q, N, H
	477	M	R, G, A, K, Q, N, H
	478	W	R, G, A, K, Q, N, H
	479	L	R, G, A, K, Q, N, H
	480	E	R, G, A, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
481	A	R, G, K, Q, N, H
482	E	R, G, A, K, Q, N, H
483	L	R, G, A, K, Q, N, H
484	H	R, G, A, K, Q, N
485	Y	R, G, A, K, Q, N, H
486	D	R, G, A, K, Q, N, H
487	G	R, A, K, Q, N, H
488	K	R, G, A, Q, N, H
489	K	R, G, A, Q, N, H
490	A	R, G, K, Q, N, H
491	K	R, G, A, Q, N, H
492	H	R, G, A, K, Q, N
493	H	R, G, A, K, Q, N
494	L	R, G, A, K, Q, N, H
495	P	R, G, A, K, Q, N, H
496	F	R, G, A, K, Q, N, H
497	Y	R, G, A, K, Q, N, H
498	N	R, G, A, K, Q, H
499	A	R, G, K, Q, N, H
500	R	G, A, K, Q, N, H
501	F	R, G, A, K, Q, N, H
502	F	R, G, A, K, Q, N, H
503	E	R, G, A, K, Q, N, H
504	E	R, G, A, K, Q, N, H
505	V	R, G, A, K, Q, N, H
506	Y	R, G, A, K, Q, N, H
507	C	R, G, A, K, Q, N, H
508	Y	R, G, A, K, Q, N, H
509	H	R, G, A, K, Q, N
510	P	R, G, A, K, Q, N, H
511	S	R, G, A, K, Q, N, H
512	V	R, G, A, K, Q, N, H

Tabla 2		
	Posición	Residuo de tipo silvestre
5	513	A
	514	E
	515	I
	516	T
10	517	P
	518	F
	519	K
	520	T
15	521	K
	522	Q
	523	F
	524	G
20	525	C
	526	E
	527	I
	528	G
25	529	K
	530	D
	531	I
	532	P
30	533	D
	534	Y
	535	V
	536	S
35	537	V
	538	A
	539	L
	540	K
	541	D
	542	N
	543	P
	544	Y

Tabla 2		
	Posición	Residuo de tipo silvestre
	545	K
	546	K
	547	A
	548	T
	549	K
	550	R
	551	I
	552	L
	553	R
	554	A
	555	I
	556	Y
	557	N
	558	P
	559	V
	560	A
	561	N
	562	T
	563	T
	564	G
	565	V
	566	D
	567	K
	568	T
	569	T
	570	N
	571	C
	572	S
	573	F
	574	M
	575	I
	576	K

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	577	R	G, A, K, Q, N, H
	578	E	R, G, A, K, Q, N, H
	579	N	R, G, A, K, Q, H
	580	D	R, G, A, K, Q, N, H
10	581	E	R, G, A, K, Q, N, H
	582	Y	R, G, A, K, Q, N, H
	583	K	R, G, A, Q, N, H
	584	L	R, G, A, K, Q, N, H
15	585	V	R, G, A, K, Q, N, H
	586	I	R, G, A, K, Q, N, H
	587	N	R, G, A, K, Q, H
	588	R	G, A, K, Q, N, H
20	589	K	R, G, A, Q, N, H
	590	I	R, G, A, K, Q, N, H
	591	S	R, G, A, K, Q, N, H
	592	V	R, G, A, K, Q, N, H
25	593	D	R, G, A, K, Q, N, H
	594	R	G, A, K, Q, N, H
	595	P	R, G, A, K, Q, N, H
	596	K	R, G, A, Q, N, H
30	597	R	G, A, K, Q, N, H
	598	I	R, G, A, K, Q, N, H
	599	E	R, G, A, K, Q, N, H
	600	V	R, G, A, K, Q, N, H
35	601	G	R, A, K, Q, N, H
	602	R	G, A, K, Q, N, H
	603	T	R, G, A, K, Q, N, H
	604	I	R, G, A, K, Q, N, H
	605	M	R, G, A, K, Q, N, H
	606	G	R, A, K, Q, N, H
	607	Y	R, G, A, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
608	D	A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
609	R	G, A, K, Q, N, H
610	N	R, G, A, K, Q, H
611	Q	R, G, A, K, N, H
612	V	R, G, A, K, Q, N, H
613	A	R, G, K, Q, N, H
614	S	R, G, A, K, Q, N, H
615	D	R, G, A, K, Q, N, H
616	T	R, G, A, K, Q, N, H
617	Y	R, G, A, K, Q, N, H
618	W	R, G, A, K, Q, N, H
619	I	R, G, A, K, Q, N, H
620	G	R, A, K, Q, N, H
621	R	G, A, K, Q, N, H
622	L	R, G, A, K, Q, N, H
623	V	R, G, A, K, Q, N, H
624	P	R, G, A, K, Q, N, H
625	P	R, G, A, K, Q, N, H
626	G	R, A, K, Q, N, H
627	T	R, G, A, K, Q, N, H
628	R	G, A, K, Q, N, H
629	G	R, A, K, Q, N, H
630	A	R, G, K, Q, N, H
631	Y	R, G, A, K, Q, N, H
632	R	G, A, K, Q, N, H
633	I	R, G, A, K, Q, N, H
634	G	R, A, K, Q, N, H
635	E	R, G, A, K, Q, N, H
636	W	R, G, A, K, Q, N, H
637	S	R, G, A, K, Q, N, H

Tabla 2		
	Posición	Residuo de tipo silvestre
5	638	V
	639	Q
	640	Y
	641	I
10	642	K
	643	S
	644	G
	645	P
15	646	V
	647	L
	648	S
	649	S
20	650	T
	651	Q
	652	G
	653	V
25	654	N
	655	N
	656	S
	657	T
30	658	T
	659	D
	660	Q
	661	L
35	662	V
	663	Y
	664	N
	665	G
	666	M
	667	P
	668	S
	669	S

Tabla 2		
	Posición	Residuo de tipo silvestre
	670	S
	671	E
	672	R
	673	F
	674	K
	675	A
	676	W
	677	K
	678	K
	679	A
	680	R
	681	M
	682	A
	683	F
	684	I
	685	R
	686	K
	687	L
	688	I
	689	R
	690	Q
	691	L
	692	N
	693	D
	694	E
	695	G
	696	L
	697	E
	698	S
	699	K
	700	G
	701	Q

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	702	D	R, G, A, K, Q, N, H
	703	Y	R, G, A, K, Q, N, H
	704	I	R, G, A, K, Q, N, H
	705	P	R, G, A, K, Q, N, H
10	706	E	R, G, A, K, Q, N, H
	707	N	R, G, A, K, Q, H
	708	P	R, G, A, K, Q, N, H
	709	S	R, G, A, K, Q, N, H
15	710	S	R, G, A, K, Q, N, H
	711	F	R, G, A, K, Q, N, H
	712	D	R, G, A, K, Q, N, H
	713	V	R, G, A, K, Q, N, H
20	714	D	R, G, A, K, Q, N, H
	715	R	G, A, K, Q, N, H
	716	G	R, A, K, Q, N, H
	717	T	R, G, A, K, Q, N, H
25	718	L	R, G, A, K, Q, N, H
	719	Y	R, G, A, K, Q, N, H
	720	V	R, G, A, K, Q, N, H
	721	F	R, G, A, K, Q, N, H
30	722	N	R, G, A, K, Q, H
	723	S	R, G, A, K, Q, N, H
	724	N	R, G, A, K, Q, H
	725	Y	R, G, A, K, Q, N, H
35	726	L	R, G, A, K, Q, N, H
	727	K	R, G, A, Q, N, H
	728	A	R, G, K, Q, N, H
	729	L	R, G, A, K, Q, N, H
	730	V	R, G, A, K, Q, N, H
	731	S	R, G, A, K, Q, N, H
	732	K	R, G, A, Q, N, H
	733	H	R, G, A, K, Q, N

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
734	R	G, A, K, Q, N, H
735	K	R, G, A, Q, N, H
736	A	R, G, K, Q, N, H
737	K	R, G, A, Q, N, H
738	K	R, G, A, Q, N, H
739	P	R, G, A, K, Q, N, H
740	V	R, G, A, K, Q, N, H
741	E	R, G, A, K, Q, N, H
742	G	R, A, K, Q, N, H
743	I	R, G, A, K, Q, N, H
744	L	R, G, A, K, Q, N, H
745	D	R, G, A, K, Q, N, H
746	E	R, G, A, K, Q, N, H
747	I	R, G, A, K, Q, N, H
748	E	R, G, A, K, Q, N, H
749	A	R, G, K, Q, N, H
750	W	R, G, A, K, Q, N, H
751	T	R, G, A, K, Q, N, H
752	S	R, G, A, K, Q, N, H
753	K	R, G, A, Q, N, H
754	D	R, G, A, K, Q, N, H
755	K	R, G, A, Q, N, H
756	D	R, G, A, K, Q, N, H
757	S	R, G, A, K, Q, N, H
758	C	R, G, A, K, Q, N, H
759	S	R, G, A, K, Q, N, H
760	L	R, G, A, K, Q, N, H
761	M	R, G, A, K, Q, N, H
762	R	G, A, K, Q, N, H
763	L	R, G, A, K, Q, N, H
764	S	R, G, A, K, Q, N, H
765	S	R, G, A, K, Q, N, H

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	766	L	R, G, A, K, Q, N, H
	767	S	R, G, A, K, Q, N, H
	768	D	R, G, A, K, Q, N, H
	769	A	R, G, K, Q, N, H
	770	S	R, G, A, K, Q, N, H
10	771	M	R, G, A, K, Q, N, H
	772	Q	R, G, A, K, Q, N, H
	773	G	R, A, K, Q, N, H
	774	I	R, G, A, K, Q, N, H
	775	A	R, G, K, Q, N, H
15	776	S	R, G, A, K, Q, N, H
	777	L	R, G, A, K, Q, N, H
	778	K	R, G, A, Q, N, H
	779	S	R, G, A, K, Q, N, H
	780	L	R, G, A, K, Q, N, H
20	781	I	R, G, A, K, Q, N, H
	782	N	R, G, A, K, Q, H
	783	S	R, G, A, K, Q, N, H
	784	Y	R, G, A, K, Q, N, H
	785	F	R, G, A, K, Q, N, H
25	786	N	R, G, A, K, Q, H
	787	K	R, G, A, Q, N, H
	788	N	R, G, A, K, Q, H
	789	G	R, A, K, Q, N, H
	790	C	R, G, A, K, Q, N, H
30	791	K	R, G, A, Q, N, H
	792	T	R, G, A, K, Q, N, H
	793	I	R, G, A, K, Q, N, H
	794	E	R, G, A, K, Q, N, H
	795	D	R, G, A, K, Q, N, H
35	796	K	R, G, A, Q, N, H
	797	E	R, G, A, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
798	K	R, G, A, Q, N, H
799	F	R, G, A, K, Q, N, H
800	N	R, G, A, K, Q, H
801	P	R, G, A, K, Q, N, H
802	V	R, G, A, K, Q, N, H
803	L	R, G, A, K, Q, N, H
804	Y	R, G, A, K, Q, N, H
805	A	R, G, K, Q, N, H
806	K	R, G, A, Q, N, H
807	L	R, G, A, K, Q, N, H
808	V	R, G, A, K, Q, N, H
809	E	R, G, A, K, Q, N, H
810	V	R, G, A, K, Q, N, H
811	E	R, G, A, K, Q, N, H
812	Q	R, G, A, K, N, H
813	R	G, A, K, Q, N, H
814	R	G, A, K, Q, N, H
815	T	R, G, A, K, Q, N, H
816	N	R, G, A, K, Q, H
817	K	R, G, A, Q, N, H
818	R	G, A, K, Q, N, H
819	S	R, G, A, K, Q, N, H
820	E	R, G, A, K, Q, N, H
821	K	R, G, A, Q, N, H
822	V	R, G, A, K, Q, N, H
823	G	R, A, K, Q, N, H
824	R	G, A, K, Q, N, H
825	I	R, G, A, K, Q, N, H
826	A	R, G, K, Q, N, H
827	G	R, A, K, Q, N, H
828	S	R, G, A, K, Q, N, H
829	L	R, G, A, K, Q, N, H

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	830	E	R, G, A, K, Q, N, H
	831	Q	R, G, A, K, N, H
	832	L	R, G, A, K, Q, N, H
	833	A	R, G, K, Q, N, H
10	834	L	R, G, A, K, Q, N, H
	835	L	R, G, A, K, Q, N, H
	836	N	R, G, A, K, Q, H
	837	G	R, A, K, Q, N, H
15	838	V	R, G, A, K, Q, N, H
	839	E	R, G, A, K, Q, N, H
	840	V	R, G, A, K, Q, N, H
	841	V	R, G, A, K, Q, N, H
20	842	I	R, G, A, K, Q, N, H
	843	G	R, A, K, Q, N, H
	844	E	A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
	845	A	R, G, K, Q, N, H
25	846	D	R, G, A, K, Q, N, H
	847	L	R, G, A, K, Q, N, H
	848	G	R, A, K, Q, N, H
	849	E	R, G, A, K, Q, N, H
30	850	V	R, G, A, K, Q, N, H
	851	E	R, G, A, K, Q, N, H
	852	K	R, G, A, Q, N, H
	853	G	R, A, K, Q, N, H
35	854	K	R, G, A, Q, N, H
	855	S	R, G, A, K, Q, N, H
	856	K	R, G, A, Q, N, H
	857	K	R, G, A, Q, N, H
	858	Q	R, G, A, K, N, H
	859	N	R, G, A, K, Q, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
860	S	R, G, A, K, Q, N, H
861	R	G, A, K, Q, N, H
862	N	R, G, A, K, Q, H
863	M	R, G, A, K, Q, N, H
864	D	R, G, A, K, Q, N, H
865	W	R, G, A, K, Q, N, H
866	C	R, G, A, K, Q, N, H
867	A	R, G, K, Q, N, H
868	K	R, G, A, Q, N, H
869	Q	R, G, A, K, N, H
870	V	R, G, A, K, Q, N, H
871	A	R, G, K, Q, N, H
872	Q	R, G, A, K, N, H
873	R	G, A, K, Q, N, H
874	L	R, G, A, K, Q, N, H
875	E	R, G, A, K, Q, N, H
876	Y	R, G, A, K, Q, N, H
877	K	R, G, A, Q, N, H
878	L	R, G, A, K, Q, N, H
879	A	R, G, K, Q, N, H
880	F	R, G, A, K, Q, N, H
881	H	R, G, A, K, Q, N
882	G	R, A, K, Q, N, H
883	I	R, G, A, K, Q, N, H
884	G	R, A, K, Q, N, H
885	Y	R, G, A, K, Q, N, H
886	F	R, G, A, K, Q, N, H
887	G	R, A, K, Q, N, H
888	V	R, G, A, K, Q, N, H
889	N	R, G, A, K, Q, H
890	P	R, G, A, K, Q, N, H
891	M	R, G, A, K, Q, N, H

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	892	Y	R, G, A, K, Q, N, H
	893	T	R, G, A, K, Q, N, H
	894	S	R, G, A, K, Q, N, H
	895	H	R, G, A, K, Q, N
	896	Q	R, G, A, K, N, H
10	897	D	R, G, A, K, Q, N, H
	898	P	R, G, A, K, Q, N, H
	899	F	R, G, A, K, Q, N, H
	900	E	R, G, A, K, Q, N, H
	901	H	R, G, A, K, Q, N
15	902	R	G, A, K, Q, N, H
	903	R	G, A, K, Q, N, H
	904	V	R, G, A, K, Q, N, H
	905	A	R, G, K, Q, N, H
	906	D	R, G, A, K, Q, N, H
20	907	H	R, G, A, K, Q, N
	908	I	R, G, A, K, Q, N, H
	909	V	R, G, A, K, Q, N, H
	910	M	R, G, A, K, Q, N, H
	911	R	G, A, K, Q, N, H
25	912	A	R, G, K, Q, N, H
	913	R	G, A, K, Q, N, H
	914	F	R, G, A, K, Q, N, H
	915	E	R, G, A, K, Q, N, H
	916	E	R, G, A, K, Q, N, H
30	917	V	R, G, A, K, Q, N, H
	918	N	R, G, A, K, Q, H
	919	V	R, G, A, K, Q, N, H
	920	E	R, G, A, K, Q, N, H
	921	N	R, G, A, K, Q, H
35	922	I	R, G, A, K, Q, N, H
	923	A	R, G, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
924	E	R, G, A, K, Q, N, H
925	W	R, G, A, K, Q, N, H
926	H	R, G, A, K, Q, N
927	V	R, G, A, K, Q, N, H
928	R	G, A, K, Q, N, H
929	N	R, G, A, K, Q, H
930	F	R, G, A, K, Q, N, H
931	S	R, G, A, K, Q, N, H
932	N	R, G, A, K, Q, H
933	Y	R, G, A, K, Q, N, H
934	L	R, G, A, K, Q, N, H
935	R	G, A, K, Q, N, H
936	A	R, G, K, Q, N, H
937	D	R, G, A, K, Q, N, H
938	S	R, G, A, K, Q, N, H
939	G	R, A, K, Q, N, H
940	T	R, G, A, K, Q, N, H
941	G	R, A, K, Q, N, H
942	L	R, G, A, K, Q, N, H
943	Y	R, G, A, K, Q, N, H
944	Y	R, G, A, K, Q, N, H
945	K	R, G, A, Q, N, H
946	Q	R, G, A, K, N, H
947	A	R, G, K, Q, N, H
948	T	R, G, A, K, Q, N, H
949	M	R, G, A, K, Q, N, H
950	D	R, G, A, K, Q, N, H
951	F	R, G, A, K, Q, N, H
952	L	R, G, A, K, Q, N, H
953	K	R, G, A, Q, N, H
954	H	R, G, A, K, Q, N
955	Y	R, G, A, K, Q, N, H

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	956	G	R, A, K, Q, N, H
	957	L	R, G, A, K, Q, N, H
	958	E	R, G, A, K, Q, N, H
	959	E	R, G, A, K, Q, N, H
	960	H	R, G, A, K, Q, N
10	961	A	R, G, K, Q, N, H
	962	E	R, G, A, K, Q, N, H
	963	G	R, A, K, Q, N, H
	964	L	R, G, A, K, Q, N, H
	965	E	R, G, A, K, Q, N, H
15	966	N	R, G, A, K, Q, H
	967	K	R, G, A, Q, N, H
	968	K	R, G, A, Q, N, H
	969	I	R, G, A, K, Q, N, H
	970	K	R, G, A, Q, N, H
20	971	F	R, G, A, K, Q, N, H
	972	Y	R, G, A, K, Q, N, H
	973	D	R, G, A, K, Q, N, H
	974	F	R, G, A, K, Q, N, H
	975	R	G, A, K, Q, N, H
25	976	K	R, G, A, Q, N, H
	977	I	R, G, A, K, Q, N, H
	978	L	R, G, A, K, Q, N, H
	979	E	R, G, A, K, Q, N, H
	980	D	R, G, A, K, Q, N, H
30	981	K	R, G, A, K, Q, N, H
	982	N	R, G, A, K, Q, H
	983	L	R, G, A, K, Q, N, H
	984	T	R, G, A, K, Q, N, H
	985	S	R, G, A, K, Q, N, H
35	986	V	R, G, A, K, Q, N, H
	987	I	R, G, A, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
988	I	R, G, A, K, Q, N, H
989	P	R, G, A, K, Q, N, H
990	K	R, G, A, Q, N, H
991	R	G, A, K, Q, N, H
992	G	R, A, K, Q, N, H
993	G	R, A, K, Q, N, H
994	R	G, A, K, Q, N, H
995	I	R, G, A, K, Q, N, H
996	Y	R, G, A, K, Q, N, H
997	M	R, G, A, K, Q, N, H
998	A	R, G, K, Q, N, H
999	T	R, G, A, K, Q, N, H
1000	N	R, G, A, K, Q, H
1001	P	R, G, A, K, Q, N, H
1002	V	R, G, A, K, Q, N, H
1003	T	R, G, A, K, Q, N, H
1004	S	R, G, A, K, Q, N, H
1005	D	R, G, A, K, Q, N, H
1006	S	R, G, A, K, Q, N, H
1007	T	R, G, A, K, Q, N, H
1008	P	R, G, A, K, Q, N, H
1009	I	R, G, A, K, Q, N, H
1010	T	R, G, A, K, Q, N, H
1011	Y	R, G, A, K, Q, N, H
1012	A	R, G, K, Q, N, H
1013	G	R, A, K, Q, N, H
1014	K	R, G, A, Q, N, H
1015	T	R, G, A, K, Q, N, H
1016	Y	R, G, A, K, Q, N, H
1017	N	R, G, A, K, Q, H
1018	R	G, A, K, Q, N, H
1019	C	R, G, A, K, Q, N, H

	Tabla 2		
	Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
5	1020	N	R, G, A, K, Q, H
	1021	A	R, G, K, Q, N, H
	1022	D	A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y
10	1023	E	R, G, A, K, Q, N, H
	1024	V	R, G, A, K, Q, N, H
	1025	A	R, G, K, Q, N, H
	1026	A	R, G, K, Q, N, H
15	1027	A	R, G, K, Q, N, H
	1028	N	R, G, A, K, Q, H
	1029	I	R, G, A, K, Q, N, H
	1030	V	R, G, A, K, Q, N, H
20	1031	I	R, G, A, K, Q, N, H
	1032	S	R, G, A, K, Q, N, H
	1033	V	R, G, A, K, Q, N, H
	1034	L	R, G, A, K, Q, N, H
25	1035	A	R, G, K, Q, N, H
	1036	P	R, G, A, K, Q, N, H
	1037	R	G, A, K, Q, N, H
	1038	S	R, G, A, K, Q, N, H
30	1039	K	R, G, A, Q, N, H
	1040	K	R, G, A, Q, N, H
	1041	N	R, G, A, K, Q, H
	1042	E	R, G, A, K, Q, N, H
35	1043	E	R, G, A, K, Q, N, H
	1044	Q	R, G, A, K, N, H
	1045	D	R, G, A, K, Q, N, H
	1046	D	R, G, A, K, Q, N, H
	1047	I	R, G, A, K, Q, N, H
	1048	P	R, G, A, K, Q, N, H
	1049	L	R, G, A, K, Q, N, H

Tabla 2		
Posición	Residuo de tipo silvestre	Sustitución o Sustituciones
1050	I	R, G, A, K, Q, N, H
1051	T	R, G, A, K, Q, N, H
1052	K	R, G, A, Q, N, H
1053	K	R, G, A, Q, N, H
1054	A	R, G, K, Q, N, H
1055	E	R, G, A, K, Q, N, H
1056	S	R, G, A, K, Q, N, H
1057	K	R, G, A, Q, N, H
1058	S	R, G, A, K, Q, N, H
1059	P	R, G, A, K, Q, N, H
1060	P	R, G, A, K, Q, N, H
1061	K	R, G, A, Q, N, H
1062	D	R, G, A, K, Q, N, H
1063	R	G, A, K, Q, N, H
1064	K	R, G, A, Q, N, H
1065	R	G, A, K, Q, N, H
1066	S	R, G, A, K, Q, N, H
1067	K	R, G, A, Q, N, H
1068	T	R, G, A, K, Q, N, H
1069	S	R, G, A, K, Q, N, H
1070	Q	R, G, A, K, N, H
1071	L	R, G, A, K, Q, N, H
1072	P	R, G, A, K, Q, N, H
1073	Q	R, G, A, K, N, H
1074	K	R, G, A, Q, N, H

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que aumenta las interacciones del polipéptido Cas12i4 variante con la guía de ARN. En algunas realizaciones, la alteración que aumenta las interacciones con la guía de ARN es una sustitución de arginina, lisina, glutamina, asparagina o histidina. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que aumenta las interacciones del polipéptido Cas12i4 variante con el ácido nucleico diana. En algunas realizaciones, la alteración que aumenta las interacciones con el ácido nucleico diana es una sustitución de arginina, lisina, glutamina, asparagina o histidina. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de alanina.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de arginina con respecto al polipéptido precursor de la SEQ ID NO: 2. Por ejemplo, en algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de arginina en el residuo 480, 482, 484, 486, 487, 490, 503, 545, 564, 566, 568, 569, 570, 587, 591, 592, 595, 598, 599, 612, 625, 629, 633, 635, 641, 668, 679, 713, 727, 735, 753, 754, 812, 825, 826, 831, 845, 846, 863, 865, 867, 870, 875, 886, 906, 945, 1028, 1032, 1042, 1049, 1055, 1058, 1059, 1071 de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de glicina con respecto al polipéptido precursor de la SEQ ID NO: 2. Por ejemplo, en algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de glicina en el residuo 480, 482, 484, 486, 490, 503, 545, 566, 568, 569, 570, 587, 591, 592, 595, 598, 599, 612, 621, 625, 633, 635, 641, 668, 679, 689, 713, 727, 735, 753, 754, 812, 818, 825, 826, 831, 845, 846, 863, 865, 867, 870, 875, 886, 906, 945, 1028, 1032, 1042, 1049, 1055, 1058, 1059, 1071 de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende dos o más sustituciones con respecto al polipéptido precursor de la SEQ ID NO: 2. Por ejemplo, el polipéptido variante puede comprender dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más sustituciones en comparación con la SEQ ID NO: 2. En la Tabla 3 se muestran ejemplos no limitantes de las dos o más sustituciones. En algunas realizaciones, un polipéptido Cas12i4 variante comprende las dos o más sustituciones enumeradas en la Tabla 3 y comprende además una sustitución enumerada en la Tabla 2.

Tabla 3. Sustituciones de múltiples aminoácidos en el polipéptido Cas12i4 variante.

Identificador de secuencia	Secuencia	Sustituciones
SEQ ID NO: 3	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM	V592R E1042R

5		VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLPFYNNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV	
10		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSTT QGVNNSTTDQ LRVNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSLAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
15		VIGEADLGEV EKGKSKKQNS RNMDWCAQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
20			
25	SEQ ID NO: 4	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSILTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV	E480R G564R V592R E1042R
30			
35			

5		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSUFNKNKC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
10		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15	SEQ ID NO: 5	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV	E480R G564R
20		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSUFNKNKC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
25			
30			
35			

		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
5	6	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLS DAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRNTKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R V592R
30	7	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM	E480R E1042R
35			

5		VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKTN CSFMIKREND EYKLVINRKI SVDRPKRIEV	
10		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LRVNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLS DAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
15		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
20			
25	SEQ ID NO: 8	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGS LTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKD WLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AGLHYDGKKA KHHLPFYNNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKTN CSFMIKREND EYKLVINRKI SVDRPKRIEV	E480R E482G
30			
35			

5		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
10		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15	SEQ ID NO: 9	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV	G564R V592R
20		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
25			
30			
35			

		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
5	SEQ ID NO: 10	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSTT QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLS DAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGSKKQNS RNMDWCAQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	G564R E1042R
30	SEQ ID NO: 11	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM	G564R E482G
35			

5		VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AGLHYDGKKA KHHLPFYNNR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV	
10		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSTT QGVNNSTTDQ LRVNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSLAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
15		VIGEADLGEV EKGKSKKQNS RNMDWCAQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
20			
25	SEQ ID NO: 12	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSILTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AGLHYDGKKA KHHLPFYNNR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV	V592R E482G
30			
35			

5		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST	
		QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG	
		QDYIPENPSS FDVRGETLYV	
		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS	
		MQGIASLKSL INSUFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE	
		KVGRIAGSLE QLALLNGVEV	
		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP	
		MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT	
		GLYYKQATMD FLKHYGLEEH	
10		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT	
		YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP	
		KDRKRSKTSQ LPQK	
15	SEQ ID NO: 13	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE	E1042R
		MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA	E482G
		SEVVKTYLNS PSSDKYVWID	
		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG	
		EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY	
		KGNAGGADPF FKYSFSKEGM	
20		VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KKLQYLEAEI GIPVDANVYS	
		QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF	
		FDSELHTTED KFNITSRYLG	
		GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ	
		FRETSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM	
		PADKRHQGEY AGQNFKMWLE	
25		AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD	
		IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND	
		EYKLVINRKI SVDRPKRIEV	
		GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST	
		QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG	
30		QDYIPENPSS FDVRGETLYV	
		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS	
		MQGIASLKSL INSUFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE	
		KVGRIAGSLE QLALLNGVEV	
		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP	
35		MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT	
		GLYYKQATMD FLKHYGLEEH	

		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
5	SEQ ID NO: 14	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPPFYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSTT QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R G564R V592R
30	SEQ ID NO: 15	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG	E480R G564R E1042R

5		GD KSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRET VSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKR HQGEY AGQNFKMWLR AELHYDGKKA KHHL PFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
10		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
15		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
20	SEQ ID NO: 16	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GD KSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRET VSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKR HQGEY AGQNFKMWLR AGLHYDGKKA KHHL PFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	E480R E482G G564R
25			
30			
35			

5		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYL RADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 17	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNV LKYLWQ FRET VSAEDF EAAAKANHLE EKISR VKAHP IVISNRYWAF GTSALVGNIM PADKR HQGEY AGQNFKMWLR AELHYDGKKA KHHLPFY NAR FFEEVYCYHP SVAEITP FKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKT TN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVL SST QGVNNSTTDQ LVIYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYL RADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R V592R E1042R
15			
20			
25			
30			
35			

5	18	SEQ ID NO:	MASISRPYGT	KLRPDARKKE	MLDKFFNTLT	KGQRVFADLA	LCIYGSLTLE	E480R
			MAKSLEPESD	SELVCAIGWF	RLVDKTIWSK	DGIKQENLVK	QYEAYSGKEA	E482G
			SEVVKTYLNS	PSSDKYVWID				V592R
			CRQKFLRFQR	ELGTRNLSED	FECMLFEQYI	RLTKGEIEGY	AAISNMFGNG	
			EKEDRSKKRM	YATRMKDWLE	ANENITWEQY	REALKNQLNA	KNLEQVVANY	
10			KGNAAGADPF	FKYSFSKEGM				
			VSKKEHAQQL	DKFKTVLKNK	ARDLNFPNKE	KLKQYLEAEI	GIPVDANVYS	
			QMFSNGVSEV	QPKTTRNMSF	SNEKLDLLTE	LKDLNKGDDF	EYAREVLNGF	
			FDSELHTTED	KFNITSRYLG				
			GDKSNRLSKL	YKIWKKEGVD	CEEGIQQFCE	AVKDKMGQIP	IRNVLKYLWQ	
15			FRETVSAEDF	EAAAKANHLE	EKISRKAHP	IVISNRYWAF	GTSALVGNIM	
			PADKRHQGEY	AGQNFKMWLR				
			AGLHYDGKKA	KHHLPPFYAR	FFEEVYCYHP	SVAEITPFKT	KQFGCEIGKD	
			IPDYVSVALK	DNPYKKATKR	ILRAIYNPVA	NTTGVDKTTN	CSFMIKREND	
			EYKLVINRKI	SRDRPKRIEV				
20			GRTIMGYDRN	QTASDTYWIG	RLVPPGTRGA	YRIGESVQY	IKSGPVLSTT	
			QGVNNSTTDQ	LVYNGMPSSS	ERFKAWKKAR	MAFIRKLIRQ	LNDEGLESKG	
			QDYIPENPSS	FDVRGETLYV				
			FNSNYLKALV	SKHRKAKKPV	EGILDEIEAW	TSKDKDSCSL	MRLSSLSLAS	
			MQGIASLKSL	INSYFNKNGC	KTIEDKEKFN	PVLYAKLVEV	EQRRTNKRSE	
25			KVGRIAGSLE	QLALLNGVEV				
			VIGEADLGEV	EKGKSKKQNS	RNMDWCAKQV	AQRLEYKLAF	HGIGYFGVNP	
			MYTSHQDPFE	HRRVADHIVM	RARFEEVNVE	NIAEWHVRNF	SNYLRADSGT	
			GLYYKQATMD	FLKHYLEEH				
			AEGLENKKIK	FYDFRKILED	KNLTSVIIPK	RGGRIYMATN	PVTSSTPIT	
30			YAGKTYNRCN	ADEVAAANIV	ISVLAPRSKK	NEEQDDIPLI	TKKAESKSP	
			KDRKRSKTSQ	LPQK				
		SEQ ID NO:	MASISRPYGT	KLRPDARKKE	MLDKFFNTLT	KGQRVFADLA	LCIYGSLTLE	E480R
		19	MAKSLEPESD	SELVCAIGWF	RLVDKTIWSK	DGIKQENLVK	QYEAYSGKEA	E482G
			SEVVKTYLNS	PSSDKYVWID				E1042R
35			CRQKFLRFQR	ELGTRNLSED	FECMLFEQYI	RLTKGEIEGY	AAISNMFGNG	
			EKEDRSKKRM	YATRMKDWLE	ANENITWEQY	REALKNQLNA	KNLEQVVANY	
			KGNAAGADPF	FKYSFSKEGM				
			VSKKEHAQQL	DKFKTVLKNK	ARDLNFPNKE	KLKQYLEAEI	GIPVDANVYS	
			QMFSNGVSEV	QPKTTRNMSF	SNEKLDLLTE	LKDLNKGDDF	EYAREVLNGF	
			FDSELHTTED	KFNITSRYLG				

5		GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SVDPRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
10		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
15		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
20	SEQ ID NO: 20	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	G564R V592R E1042R
25			
30			
35			

5		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLyakLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGeADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTsvIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 21	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETvSAEDF EAAAKANHLE EKISRvKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLsST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLyakLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGeADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTsvIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E482G G564R V592R
35	SEQ ID NO: 22	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E482G G564R E1042R

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE	
10		AGLHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYL RADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 23	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE	E482G V592R E1042R
30			
35			

5		AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
20	SEQ ID NO: 24	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDFG EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	E480R E482G G564R V592R
25			
30			
35			

5		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 25	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R E482G G564R E1042R
35	SEQ ID NO: 26	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E480R E482G V592R E1042R

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	
10		AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVL SST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLS DAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVL YAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 27	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGS LTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE	E482G G564R V592R E1042R
30			
35			

5		AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSIFYNKNKC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
20	SEQ ID NO: 28	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLLTE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AGLHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSIFYNKNKC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	E480R E482G G564R V592R E1042R
25			
30			
35			

5		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 29	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R G564R V592R E1042R T568R
35	SEQ ID NO: 30	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E480R G564R V592R E1042R S591R

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	
10		AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LVIYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 31	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	E480R G564R V592R E1042R D846G
30			
35			

5		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
	SEQ ID NO: 32	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	E480R G564R V592R E1042R F886R
20			
25			
30			
35			

5		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 33	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R G564R V592R E1042R T568R S591R
35	SEQ ID NO: 34	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E480R G564R V592R E1042R

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	T568R D846G
10		AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LVIYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSLAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 35	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	E480R G564R V592R E1042R T568R F886R
30			
35			

5		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYLEEH AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
20	SEQ ID NO: 36	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	E480R G564R V592R E1042R S591R D846G
25			
30			
35			

5		VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 37	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R G564R V592R E1042R S591R F886R
35	SEQ ID NO: 38	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E480R G564R V592R E1042R

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	D846G F886R
25	SEQ ID NO: 39	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	E480R G564R V592R E1042R T568R S591R D846G
30			
35			

5		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSTT QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
20	SEQ ID NO: 40	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSTT QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	E480R G564R V592R E1042R T568R S591R F886R
25			
30			
35			

5		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 41	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R G564R V592R E1042R T568R D846G F886R
35	SEQ ID NO: 42	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E480R G564R V592R E1042R S591R

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	D846G F886R
10		AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKTTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
20		AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 43	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	E480R G564R V592R E1042R T568R S591R D846G F886R
35			

5		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTRVDKRTN CSFMIKREND EYKLVINRKI RRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSIFYNKNKC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NREQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
	SEQ ID NO: 44	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSIFYNKNKC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	E480R T568R V592R
20			
25			
30			
35			

5		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 45	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKRTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R T568R D846G
35	SEQ ID NO: 46	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E480R T568R F886R

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	
10		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKRTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVL SST QGVNNSTTDQ LVIYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLS DAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVL YAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 47	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGS LTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	E480R V592R D846G
30			
35			

5		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
	SEQ ID NO: 48	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDFG EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKLWQ FRETVSAEDF EAAAKANHLE EKISRKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	E480R V592R F886R
20			
25			
30			
35			

5		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 49	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFMMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R D846G F886R
35	SEQ ID NO: 50	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	T568R V592R D846G

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE	
10		AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVL SST QGVNNSTTDQ LVIYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLS DAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVL YAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 51	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGS LTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTLYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE	T568R V592R F886R
30			
35			

5		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	
10		VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
15			
	SEQ ID NO: 52	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFVADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKGDF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYLWQ FRETVSAEDF EAAAKANHLE EKISRKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKRTN CSFMIKREND EYKLVINRKI SVDPRKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV	T568R D846G F886R
20			
25			
30			
35			

5		VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
10	SEQ ID NO: 53	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKTTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	V592R D846G F886R
35	SEQ ID NO: 54	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID	E480R T568R V592R D846G

5		CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR	
10		AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVL SST QGVNNSTTDQ LVIYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
15		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYFGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
25	SEQ ID NO: 55	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGV DKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN	E480R T568R V592R F886R
30			
35			

5		QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEADLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP	
10		KDRKRSKTSQ LPQK	
15	SEQ ID NO: 56	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSLTLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KKLQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLYKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLFPYNAR FFEEVYCYHP SVAEITPFFT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKRTN CSFMIKREND EYKLVINRKI SVDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LVYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	E480R T568R D846G F886R
20			
25			
30			
35			

5	57	SEQ ID NO:	MASISRPYGT	KLRPDARKKE	MLDKFFNTLT	KGQRVFADLA	LCIYGSLTLE	E480R
			MAKSLEPESD	SELVCAIGWF	RLVDKTIWSK	DGIKQENLVK	QYEAYSGKEA	V592R
			SEVVKTYLNS	PSSDKYVWID				D846G
			CRQKFLRFQR	ELGTRNLSED	FECMLFEQYI	RLTKGEIEGY	AAISNMFGNG	F886R
			EKEDRSKKRM	YATRMKDWLE	ANENITWEQY	REALKNQLNA	KNLEQVVANY	
10			KGNAGGADPF	FKYSFSKEGM				
			VSKKEHAQQL	DKFKTVLKNK	ARDLNFPNKE	KLKQYLEAEI	GIPVDANVYS	
			QMFSNGVSEV	QPKTTRNMSF	SNEKLDLLTE	LKDLNKG DGF	EYAREVLNGF	
			FDSELHTTED	KFNITSRYLG				
			GDKSNRLSKL	YKIWKKEGVD	CEEGIQQFCE	AVKDKMGQIP	IRNVLKYLWQ	
15			FRETVSAEDF	EAAAKANHLE	EKISRVKAHP	IVISNRYWAF	GTSALVGNIM	
			PADKRHQGEY	AGQNFKMWLR				
			AELHYDGKKA	KHHLPPFYNA	FFEEVYCYHP	SVAEITPFFT	KQFGCEIGKD	
			IPDYVSVALK	DNPYKKATKR	ILRAIYNPVA	NTTGVDKTTN	CSFMIKREND	
			EYKLVINRKI	SRDRPKRIEV				
20			GRTIMGYDRN	QTASDTYWIG	RLVPPGTRGA	YRIGESVQY	IKSGPVLSTT	
			QGVNNSTTDQ	LVYNGMPSSS	ERFKAWKKAR	MAFIRKLIRQ	LNDEGLESKG	
			QDYIPENPSS	FDVRGETLYV				
			FNSNYLKALV	SKHRKAKKPV	EGILDEIEAW	TSKDKDSCSL	MRLSSLS DAS	
			MQGIASLKSL	INSYFNKNGC	KTIEDKEKFN	PVLYAKLVEV	EQRRTNKRSE	
25			KVGRIAGSLE	QLALLNGVEV				
			VIGEAGLGEV	EKGKSKKQNS	RNMDWCAKQV	AQRLEYKLAF	HGIGYRGVNP	
			MYTSHQDPFE	HRRVADHIVM	RARFEEVNVE	NIAEWHVRNF	SNYLRADSGT	
			GLYYKQATMD	FLKHYGLEEH				
			AEGLENKKIK	FYDFRKILED	KNLTSVIIPK	RGGRIYMATN	PVTS DSTPIT	
30			YAGKTYNRCN	ADEVAAANIV	ISVLAPRSKK	NEEQDDIPLI	TKKAESKSP	
			KDRKRSKTSQ	LPQK				
		SEQ ID NO:	MASISRPYGT	KLRPDARKKE	MLDKFFNTLT	KGQRVFADLA	LCIYGSLTLE	T568R
		58	MAKSLEPESD	SELVCAIGWF	RLVDKTIWSK	DGIKQENLVK	QYEAYSGKEA	V592R
			SEVVKTYLNS	PSSDKYVWID				D846G
35			CRQKFLRFQR	ELGTRNLSED	FECMLFEQYI	RLTKGEIEGY	AAISNMFGNG	F886R
			EKEDRSKKRM	YATRMKDWLE	ANENITWEQY	REALKNQLNA	KNLEQVVANY	
			KGNAGGADPF	FKYSFSKEGM				
			VSKKEHAQQL	DKFKTVLKNK	ARDLNFPNKE	KLKQYLEAEI	GIPVDANVYS	
			QMFSNGVSEV	QPKTTRNMSF	SNEKLDLLTE	LKDLNKG DGF	EYAREVLNGF	
			FDSELHTTED	KFNITSRYLG				

5		GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLE AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	
10		FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLYAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH	
15		AEGLNKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSPP KDRKRSKTSQ LPQK	
20	SEQ ID NO: 59	MASISRPYGT KLRPDARKKE MLDKFFNTLT KGQRVFADLA LCIYGSITLE MAKSLEPESD SELVCAIGWF RLVDKTIWSK DGIKQENLVK QYEAYSGKEA SEVVKTYLNS PSSDKYVWID CRQKFLRFQR ELGTRNLSED FECMLFEQYI RLTKGEIEGY AAISNMFGNG EKEDRSKKRM YATRMKDWLE ANENITWEQY REALKNQLNA KNLEQVVANY KGNAGGADPF FKYSFSKEGM VSKKEHAQQL DKFKTVLKNK ARDLNFPNKE KLKQYLEAEI GIPVDANVYS QMFSNGVSEV QPKTTRNMSF SNEKLDLLTE LKDLNKG DGF EYAREVLNGF FDSELHTTED KFNITSRYLG GDKSNRLSKL YKIWKKEGVD CEEGIQQFCE AVKDKMGQIP IRNVLKYLWQ FRETVSAEDF EAAAKANHLE EKISRVKAHP IVISNRYWAF GTSALVGNIM PADKRHQGEY AGQNFKMWLR AELHYDGKKA KHHLPFYNAR FFEEVYCYHP SVAEITPFKT KQFGCEIGKD IPDYVSVALK DNPYKKATKR ILRAIYNPVA NTTGVDKRTN CSFMIKREND EYKLVINRKI SRDRPKRIEV GRTIMGYDRN QTASDTYWIG RLVPPGTRGA YRIGESVQY IKSGPVLSSST QGVNNSTTDQ LUYNGMPSSS ERFKAWKKAR MAFIRKLIRQ LNDEGLESKG QDYIPENPSS FDVRGETLYV	E480R T568R V592R D846G F886R
25			
30			
35			

5	FNSNYLKALV SKHRKAKKPV EGILDEIEAW TSKDKDSCSL MRLSSLSDAS MQGIASLKSL INSYFNKNGC KTIEDKEKFN PVLAKLVEV EQRRTNKRSE KVGRIAGSLE QLALLNGVEV VIGEAGLGEV EKGKSKKQNS RNMDWCAKQV AQRLEYKLAF HGIGYRGVNP MYTSHQDPFE HRRVADHIVM RARFEEVNVE NIAEWHVRNF SNYLRADSGT GLYYKQATMD FLKHYGLEEH AEGLENKKIK FYDFRKILED KNLTSVIIPK RGGRIYMATN PVTSDSTPIT YAGKTYNRCN ADEVAAANIV ISVLAPRSKK NEEQDDIPLI TKKAESKSP KDRKRSKTSQ LPQK	
---	---	--

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una o más sustituciones adicionales además de la secuencia de la SEQ ID NO: 3 (por ejemplo, el polipéptido Cas12i4 variante comprende una sustitución V592R y una sustitución E1042R y comprende además una o más sustituciones que se muestran en la Tabla 2 o la Tabla 3). En algunas realizaciones, el polipéptido Cas12i4 variante comprende una o más sustituciones adicionales además de la secuencia de la SEQ ID NO: 4 (por ejemplo, el polipéptido Cas12i4 variante comprende una sustitución E480R, una sustitución G564R, una sustitución V592R, una sustitución E1042R y comprende además una o más sustituciones que se muestran en la Tabla 2 o la Tabla 3). En algunas realizaciones, el polipéptido Cas12i4 variante comprende una o más sustituciones adicionales además de una cualquiera de las secuencias de las SEQ ID NO: 5-59 (por ejemplo, el polipéptido Cas12i4 variante comprende además una o más sustituciones que se muestran en la Tabla 2 o la Tabla 3). Como se ha indicado anteriormente, en algunas realizaciones, el polipéptido Cas12i4 variante mantiene los cambios de aminoácidos (o al menos 1, 2, 3, 4, 5, etc. de estos cambios) que diferencian el polipéptido de su respectiva secuencia precursora/de referencia.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos un motivo RuvC o un dominio RuvC.

Los dominios de los polipéptidos Cas12i4 desvelados en el presente documento se representan en la **FIG. 5**. El dominio Wedge comprende los residuos 1-14 y 447-593 del polipéptido Cas12i4. El dominio Rec1 comprende los residuos 15-171 y 266-446 del polipéptido Cas12i4. El dominio PI comprende los residuos 172-265 del polipéptido Cas12i4. El dominio Rec2 comprende los residuos 647-839 del polipéptido Cas12i4. El dominio Nuc comprende los residuos 891-11018 del polipéptido Cas12i4. El dominio RuvC comprende los residuos 594-646 (motivo RuvC1), los residuos 840-890 (motivo RuvC2) y los residuos 1019-1074 (motivo RuvC3) del polipéptido Cas12i4.

Aunque los cambios descritos en el presente documento pueden ser uno o más cambios de aminoácidos, los cambios en el polipéptido Cas12i4 variante también pueden ser de naturaleza sustancial, tales como la fusión de polipéptidos como extensiones amino y/o carboxilo terminales. Por ejemplo, el polipéptido Cas12i4 variante puede contener péptidos adicionales, por ejemplo, uno o más péptidos. Los ejemplos de péptidos adicionales pueden incluir péptidos epítomos para marcaje, tales como un marcador de polihistidina (His-tag), Myc y FLAG. En algunas realizaciones, el polipéptido Cas12i4 variante descrito

en el presente documento puede fusionarse con un resto detectable tal como una proteína fluorescente (por ejemplo, proteína fluorescente verde (GFP) o proteína fluorescente amarilla (YFP)).

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una (por ejemplo, dos, tres, cuatro, cinco, seis o más) señal de localización nuclear (NLS). En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una (por ejemplo, dos, tres, cuatro, cinco, seis o más) señal de exportación nuclear (NES). En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una (por ejemplo, dos, tres, cuatro, cinco, seis o más) NLS y al menos una (por ejemplo, dos, tres, cuatro, cinco, seis o más) NES.

En algunas realizaciones, el polipéptido Cas12i4 variante descrito en el presente documento puede ser auto-inactivante. Véase, Epstein *et al.*, "Engineering a Self-Inactivating CRISPR System for AAV Vectors", *Mol. Ther.*, 24 (2016): S50, que se incorpora por referencia en su totalidad.

En algunas realizaciones, la secuencia de nucleótidos que codifica el polipéptido Cas12i4 variante descrito en el presente documento puede tener codones optimizados para su uso en un organismo o célula hospedadora particular. Por ejemplo, el ácido nucleico puede tener codones optimizados para cualquier eucariota no humano, incluyendo ratones, ratas, conejos, perros, ganado o primates no humanos. Las tablas de uso de codones están disponibles, por ejemplo, en la "Base de datos de uso de codones" disponible en www.kazusa.or.jp/codon/ y estas tablas pueden adaptarse de varias maneras. Véase Nakamura *et al. Nucl. Acids Res.* 28:292 (2000), que se incorpora en el presente documento por referencia en su totalidad. También hay disponibles algoritmos informáticos para la optimización de codones de una secuencia particular para la expresión en una célula hospedadora particular, tal como Gene Forge (Aptagen; Jacobus, PA).

Funcionalidad de los polipéptidos variantes

Como se usa en el presente documento, una "porción biológicamente activa" es una porción que conserva al menos una función (por ejemplo, completamente, parcialmente, mínimamente) del polipéptido precursor (por ejemplo, un dominio "mínimo" o "central"). En algunas realizaciones, el polipéptido Cas12i4 variante conserva la actividad enzimática al menos tan activa como la del polipéptido precursor. En consecuencia, en algunas realizaciones, un polipéptido Cas12i4 variante tiene una actividad enzimática mayor que la del polipéptido precursor.

También se proporciona un polipéptido Cas12i4 variante de la presente invención que tiene actividad enzimática, por ejemplo, actividad nucleasa o endonucleasa, y que comprende una secuencia de aminoácidos que difiere de las secuencias de aminoácidos de uno cualquiera de un polipéptido precursor y la SEQ ID NO: 2 en 50, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, o 1 residuos de aminoácidos, cuando se alinea usando cualquiera de los métodos de alineación descritos anteriormente.

En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución V592R presenta una actividad enzimática potenciada. En algunas realizaciones, el residuo V592R interactúa con la NTS. En algunas realizaciones, el residuo V592R contacta con la NTS cerca de la secuencia de PAM. Véase la **FIG. 6A**.

En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución E480R presenta una actividad enzimática potenciada. En algunas realizaciones, la sustitución E480R interactúa con ADN bicatenario. En algunas realizaciones, la sustitución E480R interactúa con ADN bicatenario en dirección 5' de la secuencia de PAM. En algunas realizaciones, la sustitución E480R estabiliza las interacciones del polipéptido Cas12i4 variante con un ácido nucleico diana. Véase la **FIG. 6B**.

En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución G564R presenta una actividad enzimática potenciada. En algunas realizaciones, la sustitución G564R interactúa con ADN bicatenario. En algunas realizaciones, la sustitución G564R interactúa con ADN bicatenario en dirección 5' de la secuencia de PAM. En algunas realizaciones, la sustitución G564R estabiliza las interacciones del polipéptido Cas12i4 variante con un ácido nucleico diana. Véase la **FIG. 6B**.

En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución E1042R presenta una actividad enzimática potenciada.

En algunas realizaciones, el polipéptido Cas12i4 variante tiene actividad nucleasa reducida o es un polipéptido de nucleasa muerta. Como se usa en el presente documento, los residuos catalíticos de un polipéptido desvelado en el presente documento son D608, E844 y D1022. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución en uno o más de D608, E844 y D1022 (por ejemplo, D608A, E844A y D1022A) presenta una actividad nucleasa reducida o ninguna actividad nucleasa con respecto a un polipéptido precursor.

En algunas realizaciones, el polipéptido Cas12i4 variante de la presente invención tiene una actividad enzimática equivalente o mayor que la del polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante de la presente invención tiene actividad enzimática a un intervalo de temperatura de aproximadamente 20 °C hasta aproximadamente 90 °C. En algunas realizaciones, el polipéptido Cas12i4 variante de la presente invención tiene actividad enzimática a una temperatura de aproximadamente 20 °C a aproximadamente 25 °C o a una temperatura de aproximadamente 37 °C.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que potencia la afinidad por ARN (por ejemplo, afinidad por el ARN), en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una afinidad por el ARN potenciada, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una afinidad por el ARN potenciada, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una afinidad por el ARN potenciada, en comparación con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta una afinidad por el ARN potenciada cuando el

valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática disminuida y (b) una afinidad por el ARN potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática aumentada y (b) una afinidad por el ARN potenciada, con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática conservada y (b) una afinidad por el ARN potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que potencia la formación de complejos con una guía de ARN (por ejemplo, formación de complejos binarios), en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una formación de complejos binarios potenciada, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una formación de complejos binarios potenciada, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una formación de complejos binarios potenciada, en comparación con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta una formación de complejos binarios potenciada cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática disminuida y (b) una formación de complejos binarios potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática aumentada y (b) una formación de complejos binarios potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática conservada y (b) una formación de complejos binarios potenciada, con respecto al polipéptido precursor de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que potencia la actividad de unión a una guía de ARN, en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad de unión a una guía de ARN potenciada, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad de unión a una guía de ARN potenciada, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad de unión a una guía de ARN potenciada, en comparación con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta una actividad de unión a una guía de ARN potenciada cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática disminuida y (b) una actividad de unión a una guía de ARN potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática aumentada y (b) una actividad de unión a una guía de ARN potenciada, con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática conservada y (b) una actividad de unión a una guía de ARN potenciada, con respecto al polipéptido precursor de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que potencia la especificidad de unión a una guía de ARN, en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una especificidad de unión a una guía de ARN potenciada, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una especificidad de unión a una guía de ARN potenciada, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una especificidad de unión a una guía de ARN potenciada, en comparación con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C,

6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta una especificidad de unión a una guía de ARN potenciada cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

5 En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una especificidad enzimática disminuida y (b) una especificidad de unión a una guía de ARN potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a)
10 una especificidad enzimática aumentada y (b) una especificidad de unión a una guía de ARN potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una especificidad enzimática conservada y (b) una especificidad de unión a una guía de ARN potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2.

15 En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que potencia las interacciones proteína-ARN, en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta interacciones proteína-ARN potenciadas, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C,
20 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta interacciones proteína-ARN potenciadas, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta interacciones proteína-ARN potenciadas, en
25 comparación con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta interacciones proteína-ARN potenciadas cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

30 En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática disminuida y (b) interacciones proteína-ARN potenciadas con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática
35 aumentada y (b) interacciones proteína-ARN potenciadas con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática

conservada y (b) interacciones proteína-ARN potenciadas con respecto al polipéptido precursor de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que potencia la estabilidad de la proteína, en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una estabilidad de la proteína potenciada, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una estabilidad de la proteína potenciada, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una estabilidad de la proteína potenciada, en comparación con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta una estabilidad de la proteína potenciada cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática disminuida y (b) una estabilidad de la proteína potenciada, con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática aumentada y (b) una estabilidad de la proteína potenciada, con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática conservada y (b) una estabilidad de la proteína potenciada, con respecto al polipéptido precursor de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que disminuye la disociación de una guía de ARN (por ejemplo, disociación de complejo binario), en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una disociación disminuida de una guía de ARN, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una disociación disminuida de una guía de ARN, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una disociación disminuida de una guía de ARN, en comparación

con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta una disociación disminuida de una guía de ARN cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una disociación disminuida de una guía de ARN, en comparación con un polipéptido precursor, a lo largo de un periodo de incubación de al menos aproximadamente uno cualquiera de 10 minutos, 15 minutos, 20 minutos, 25 minutos, 30 minutos, 35 minutos, 40 minutos, 45 minutos, 50 minutos, 55 minutos, 1 h, 2 h, 3 h, 4 h o más horas. En algunas realizaciones, un complejo de ribonucleoproteína (RNP) variante no intercambia la guía de ARN con un ARN diferente.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática disminuida y (b) una disociación disminuida de una guía de ARN con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática aumentada y (b) una disociación disminuida de una guía de ARN, con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática conservada y (b) una disociación disminuida de una guía de ARN, con respecto al polipéptido precursor de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración que potencia la formación de complejos ternarios con una guía de ARN y un ácido nucleico diana, en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una formación de complejos ternarios potenciada, en comparación con un polipéptido precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una formación de complejos ternarios potenciada, en comparación con un polipéptido precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una formación de complejos ternarios potenciada, en comparación con un polipéptido precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el polipéptido Cas12i4 variante presenta una formación de complejos ternarios potenciada cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática disminuida y (b) una formación de complejos ternarios potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática aumentada y (b) una formación de complejos ternarios potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que presenta (a) una actividad enzimática conservada y (b) una formación de complejos ternarios potenciada con respecto al polipéptido precursor de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración de modo que un complejo binario que comprende el polipéptido Cas12i4 variante (por ejemplo, un complejo binario variante) presenta una afinidad de unión potenciada con un ácido nucleico diana, en comparación con un complejo binario precursor. En algunas realizaciones, el complejo binario variante presenta una afinidad de unión potenciada a un ácido nucleico diana, en comparación con un complejo binario precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el complejo binario variante presenta una afinidad de unión potenciada a un ácido nucleico diana, en comparación con un complejo binario precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el complejo binario variante presenta una afinidad de unión potenciada con un ácido nucleico diana, en comparación con un complejo binario precursor, cuando el valor de T_m del complejo binario variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un complejo binario precursor. En una realización, el complejo binario variante presenta una afinidad de unión potenciada a un ácido nucleico diana cuando el valor de T_m del complejo binario variante es al menos 8 °C superior al valor de T_m del complejo binario precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática disminuida y (b) una afinidad de unión potenciada a un ácido nucleico diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática aumentada y (b) una afinidad de unión potenciada a un ácido nucleico diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una

actividad enzimática conservada y (b) una afinidad de unión potenciada a un ácido nucleico diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración de modo que un complejo binario que comprende el polipéptido Cas12i4 variante (por ejemplo, un complejo binario variante) presenta una actividad de unión potenciada en la diana, en comparación con un complejo binario precursor. En algunas realizaciones, el complejo binario variante presenta una actividad de unión en la diana potenciada, en comparación con un complejo binario precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el complejo binario variante presenta una actividad de unión en la diana potenciada, en comparación con un complejo binario precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el complejo binario variante presenta una actividad de unión en la diana potenciada, en comparación con un complejo binario precursor, cuando el valor de T_m del complejo binario variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un complejo binario precursor. En una realización, el complejo binario variante presenta una actividad de unión en la diana potenciada cuando el valor de T_m del complejo binario variante es al menos 8 °C superior al valor de T_m del complejo binario precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática disminuida y (b) una actividad de unión en la diana potenciada, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática aumentada y (b) una actividad de unión en la diana potenciada, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática conservada y (b) una actividad de unión en la diana potenciada, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración de modo que un complejo binario que comprende el polipéptido Cas12i4 variante (por ejemplo, un complejo binario variante) presenta una especificidad de unión en la diana potenciada, en comparación con un complejo binario precursor. En algunas realizaciones, el complejo binario variante presenta una especificidad de unión en la diana potenciada, en comparación con un complejo binario precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C,

42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el complejo binario variante presenta una especificidad de unión en la diana potenciada, en comparación con un complejo binario precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el complejo binario variante presenta una especificidad de unión en la diana potenciada, en comparación con un complejo binario precursor, cuando el valor de T_m del complejo binario variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un complejo binario precursor. En una realización, el complejo binario variante presenta una especificidad de unión en la diana potenciada cuando el valor de T_m del complejo binario variante es al menos 8 °C superior al valor de T_m del complejo binario precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática disminuida y (b) una especificidad de unión en la diana potenciada, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática aumentada y (b) una especificidad de unión en la diana potenciada, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática conservada y (b) una especificidad de unión en la diana potenciada, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración de modo que un complejo binario que comprende el polipéptido Cas12i4 variante (por ejemplo, un complejo binario variante) presenta una unión fuera de la diana disminuida a un ácido nucleico no diana, en comparación con un complejo binario precursor. En algunas realizaciones, el complejo binario variante presenta una unión fuera de la diana reducida a un ácido nucleico no diana, en comparación con un complejo binario precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el complejo binario variante presenta una unión fuera de la diana reducida a un ácido nucleico no diana, en comparación con un complejo binario precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el complejo binario variante presenta una unión fuera de la diana disminuida a un ácido nucleico no diana, en comparación con un complejo binario precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el complejo binario variante presenta una unión fuera de la diana reducida a

un ácido nucleico no diana cuando el valor de T_m del complejo binario variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática disminuida y (b) una unión fuera de la diana disminuida a un ácido nucleico no diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática aumentada y (b) una unión fuera de la diana disminuida a un ácido nucleico no diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática conservada y (b) una unión fuera de la diana disminuida a un ácido nucleico no diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración de modo que un complejo binario que comprende el polipéptido Cas12i4 variante (por ejemplo, un complejo binario variante) presenta una disociación disminuida del ácido nucleico diana, en comparación con un complejo binario precursor. En algunas realizaciones, el complejo binario variante presenta una disociación disminuida del ácido nucleico diana, en comparación con un complejo binario precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el complejo binario variante presenta una disociación disminuida del ácido nucleico diana, en comparación con un complejo binario precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el complejo binario variante presenta una disociación disminuida del ácido nucleico diana, en comparación con un complejo binario precursor, cuando el valor de T_m del polipéptido Cas12i4 variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un polipéptido precursor. En una realización, el complejo binario variante presenta una disociación disminuida del ácido nucleico diana cuando el valor de T_m del complejo binario variante es al menos 8 °C superior al valor de T_m del polipéptido precursor.

En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática disminuida y (b) una disociación disminuida del ácido nucleico diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una

actividad enzimática aumentada y (b) una disociación disminuida del ácido nucleico diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2. En algunas realizaciones, se introduce al menos una alteración en el polipéptido precursor de la SEQ ID NO: 2 para producir un polipéptido Cas12i4 variante que forma un complejo binario variante que presenta (a) una actividad enzimática conservada y (b) una disociación disminuida del ácido nucleico diana, con respecto a un complejo binario precursor que comprende el polipéptido de la SEQ ID NO: 2.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende al menos una alteración de modo que un complejo ternario que comprende el polipéptido Cas12i4 variante (por ejemplo, un complejo ternario variante) presenta una estabilidad potenciada, en comparación con un complejo ternario precursor.

En algunas realizaciones, el complejo ternario variante presenta una estabilidad potenciada, en comparación con un complejo ternario precursor, a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C, 51 °C, 52 °C, 53 °C, 54 °C, 55 °C, 56 °C, 57 °C, 58 °C, 59 °C, 60 °C o 65 °C. En algunas realizaciones, el complejo ternario variante presenta una estabilidad potenciada, en comparación con un complejo ternario precursor, en un tampón que tiene un pH en un intervalo de aproximadamente 7.3 a aproximadamente 8.6. En algunas realizaciones, el complejo ternario variante presenta una estabilidad potenciada, en comparación con un complejo ternario precursor, cuando el valor de T_m del complejo ternario variante es al menos 1 °C, 2 °C, 3 °C, 4 °C, 5 °C, 6 °C, 7 °C, 8 °C, 9 °C, 10 °C, 11 °C, 12 °C, 13 °C, 14 °C, 15 °C, 16 °C, 17 °C, 18 °C, 19 °C o 20 °C superior al valor de T_m de un complejo ternario precursor. En una realización, el complejo ternario variante presenta una estabilidad potenciada cuando el valor de T_m del complejo ternario variante es al menos 8 °C superior al valor de T_m del complejo ternario precursor.

Interacciones aumentadas de la guía de ARN

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y la guía de ARN, en comparación con un polipéptido precursor. En algunas realizaciones, la alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y la guía de ARN es sustituir uno o más aminoácidos por un residuo de arginina, lisina, glutamina, asparagina, histidina, serina o tirosina. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de uno o más aminoácidos en la interfaz de unión a ARN a un residuo de arginina, lisina, glutamina, asparagina, histidina, serina, tirosina, fenilalanina, ácido glutámico o metionina. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos en al menos un dominio (por ejemplo, el dominio Wedge, el motivo RuvC1, el motivo RuvC2 o el dominio Rec2). En algunas realizaciones, la sustitución o sustituciones de la interfaz de unión a ARN aumenta la unión de la guía de ARN o la afinidad de unión de la guía de ARN en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%

74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor.

En algunas realizaciones, la sustitución aumenta la formación de complejo de guía de ARN (complejo binario) con respecto a un polipéptido precursor. En la Tabla 4 se muestran ejemplos no limitantes de sustituciones que pueden alterar la capacidad de un polipéptido Cas12i4 variante para interactuar con la secuencia repetitiva directa de una guía de ARN. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 presenta una formación potenciada de complejo de guía de ARN (complejo binario) con respecto a un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 forma un complejo binario más estable con una guía de ARN, en comparación con un complejo binario que comprende un polipéptido precursor.

Tabla 4. Sustituciones que aumentan el contacto de la secuencia repetitiva directa.

Residuo	Sustitución o Sustituciones
K11	R
D15	H, Q, Y, F, M, E
N474	Q, E
E469	R, W
F475	Y
F476	M
Y497	K
T520	K, R
Q522	K
Y544	R, K
K545	R,
K546	R
T612	K, R
Q651	R, K
N654	R, K
T657	K, R
T658	K
Y633	R, K
N654	R
Y719	K
C757	T

V808	K, R
E809	R, K, Q
Q812	H, R, K
N816	K
E830	N
Q831	K

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 4. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 4.

En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una formación potenciada de complejo de guía de ARN (complejo binario) comprende dos o más sustituciones. En algunas realizaciones, un polipéptido Cas12i4 variante comprende además K545R y K546R. En algunas realizaciones, un polipéptido Cas12i4 variante comprende además K545R y K546R y N654R.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende además una o más sustituciones enumeradas en la Tabla 4 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende además una o más sustituciones enumeradas en la Tabla 4 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad enzimática aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

Aumento de las interacciones de ADN bicatenario

En algunos aspectos, un polipéptido Cas12i4 variante comprende una alteración que aumenta las interacciones con el ADN bicatenario con respecto a un polipéptido precursor. En algunas realizaciones, las interacciones aumentadas con el ADN bicatenario son interacciones electrostáticas aumentadas. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que aumenta la afinidad entre el polipéptido Cas12i4 variante y el ADN bicatenario con respecto a un polipéptido precursor. En algunas realizaciones, la alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y el ADN bicatenario aumenta la unión del polipéptido Cas12i4 variante a una secuencia de PAM.

En algunas realizaciones, la alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y el ADN bicatenario es la sustitución de uno o más aminoácidos. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de uno o más aminoácidos en la interfaz de unión a ADN bicatenario. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos en al menos un dominio (por ejemplo, el dominio Rec1, el dominio PI o el dominio Wedge) a un residuo de arginina, lisina, glutamina, asparagina, histidina, triptófano, glicina, leucina, alanina o serina. En algunas realizaciones, la sustitución o sustituciones de la interfaz de unión a ADN bicatenario aumentan las interacciones y/o la afinidad del ADN bicatenario en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor. En algunas realizaciones, la sustitución o sustituciones de la interfaz de unión a ADN bicatenario aumentan la unión del polipéptido Cas12i4 variante a una secuencia de PAM en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor.

En algunas realizaciones, la sustitución que aumenta las interacciones de ADN bicatenario aumenta la formación de complejos ternarios con respecto a un polipéptido precursor. En la Tabla 5 se muestran ejemplos no limitantes de sustituciones que pueden alterar la capacidad de un polipéptido Cas12i4 variante para interactuar con el ADN bicatenario. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 5 presenta interacciones de ADN bicatenario aumentadas (formación de complejo ternario) con respecto a un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 5 forma un complejo ternario más estable, en comparación con un polipéptido precursor.

Tabla 5. Sustituciones que alteran las interacciones bicatenarias.

Residuo	Sustitución o Sustituciones
S3	R, K

5	T568	K, H
	N165	R, K
	E173	S
	Y233	R, K, Q, W
	D285	K, R, N
10	T568	R
	V456	K, A, R
	V458	R, K
	A227	R, K
	D228	K, N, A
15	Q248	R, K
	K252	R
	T255	N, K, R
	N259	Q, S, H, G
	K260	R
20	K264	R, K
	P461	R, K
	S5	K, R
	A161	R, K
	A449	N
25	R175	K
	A218	R, K
	Y220	K, R
	K221	R
	N297	K
30	N570	K, R
	V217	R, K
	K232	R
	K178	R
	N287	K
	A286	R
	Y160	L

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 5. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 5.

En algunas realizaciones, un polipéptido Cas12i4 variante que presenta interacciones de ADN bicatenario aumentadas comprende dos o más sustituciones enumeradas en la Tabla 5. En algunas

realizaciones, un polipéptido Cas12i4 variante que presenta interacciones de ADN bicatenario aumentadas comprende K232R y D228A. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta interacciones de ADN bicatenario aumentadas comprende A286R y Y160L. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta interacciones de ADN bicatenario aumentadas comprende N287K y V456A. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta interacciones de ADN bicatenario aumentadas comprende K178R y E173S.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una cualquiera o más sustituciones de la Tabla 4 y/o la Tabla 5. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 presenta interacciones y/o una afinidad por el ADN bicatenario aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad enzimática aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%,

160%, 170%, 180%, 190%, 200%, or más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

Interacciones de ADN monocatenario aumentadas

En algunas realizaciones, un polipéptido Cas12i4 variante comprende una alteración que aumenta las interacciones con el ADN monocatenario con respecto a un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que aumenta la afinidad entre el polipéptido Cas12i4 variante y el ADN bicatenario con respecto a un polipéptido precursor. En algunas realizaciones, el ADN monocatenario comprende la cadena no diana (NTS). En algunas realizaciones, las interacciones aumentadas con el ADN monocatenario (por ejemplo, la NTS) son interacciones entre la secuencia de PAM y el sitio activo del polipéptido Cas12i4 variante. En algunas realizaciones, el ADN monocatenario comprende ADN monocatenario que interactúa con el polipéptido Cas12i4 variante en o cerca del sitio activo del polipéptido Cas12i4 variante. En algunas realizaciones, una alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y el ADN monocatenario estabiliza el bucle R. Como se usa en el presente documento, el "bucle R" se refiere a un ácido nucleico que comprende una guía de ARN emparejada con la cadena diana (TS) y la cadena no diana (NTS) monocatenaria.

En algunas realizaciones, la alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y el ADN monocatenario es la sustitución de uno o más aminoácidos. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de uno o más aminoácidos en la interfaz de unión a ADN monocatenario. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos en al menos un dominio/motivo (por ejemplo, el dominio PI, el dominio Rec1, el motivo RuvC1, el dominio Rec2, el motivo RuvC2, el dominio Nuc o el motivo RuvC3) a una arginina, lisina o alanina. En algunas realizaciones, la sustitución o sustituciones de la interfaz de unión a ADN monocatenario aumentan las interacciones y/o la afinidad del ADN monocatenario en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor.

En algunas realizaciones, la sustitución que aumenta las interacciones de ADN monocatenario aumenta la formación de complejos ternarios con respecto a un polipéptido precursor. En la Tabla 6 se muestran ejemplos no limitantes de sustituciones que pueden alterar la capacidad de un polipéptido Cas12i4 variante para interactuar con el ADN monocatenario. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 6 presenta interacciones de ADN monocatenario aumentadas (formación de complejo ternario) con respecto a un polipéptido

precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 6 forma un complejo ternario más estable, en comparación con un polipéptido precursor.

Tabla 6. Sustituciones que alteran las interacciones monocatenarias.

5	Residuo	Sustitución o Sustituciones	
	Y233	R	en la NTS del ADNmc
	P625	R, K	
	E635	R, K	
10	W636	R, K	
	P1036	R, K	
	T612	R, K	Cerca del sitio activo
	D362	R, K	
	N724	K	
15	A728	R, K	
	D768	N, A	
	Q772	R, K	
	S938	R, K	
20	L942	R, K	
	V720	R, K	
	Q858	K	
	Y892	R, K	
	D937	R, K	
25	G939	K	
	E962	R, K	
	E965	R, K	

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 6. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 6.

En algunas realizaciones, un polipéptido Cas12i4 variante que muestra interacciones de ADN monocatenario aumentadas comprende dos o más sustituciones enumeradas en la Tabla 6. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una formación/estabilidad de complejos ternarios aumentada comprende dos o más sustituciones enumeradas en la Tabla 6. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una cualquiera o más sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 presenta interacciones y/o una afinidad por el ADN monocatenario aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%,

11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, un polipéptido Cas12i4 variante comprende una sustitución que aumenta la estabilidad del ADN monocatenario (por ejemplo, la sustitución aumenta las interacciones electrostáticas entre el ADN monocatenario y el sitio activo del polipéptido Cas12i4 variante). En algunas realizaciones, el polipéptido Cas12i4 variante aumenta la estabilidad del ADN monocatenario en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor. En la Tabla 6 se muestran ejemplos no limitantes de sustituciones que pueden alterar la capacidad de un polipéptido Cas12i4 variante para estabilizar el ADN monocatenario. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 6 presenta una estabilidad del ADN monocatenario aumentada con respecto a un polipéptido precursor.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y presenta una actividad enzimática aumentada. En algunas realizaciones, el

polipéptido Cas12i4 variante presenta una actividad enzimática aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

Interacciones heterodúplex aumentadas

En algunas realizaciones, un polipéptido Cas12i4 variante comprende una sustitución que aumenta las interacciones con una molécula híbrida de ADN/ARN con respecto a un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que aumenta la afinidad entre el polipéptido Cas12i4 variante y un híbrido de ADN/ARN con respecto a un polipéptido precursor. En algunas realizaciones, la molécula híbrida de ADN/ARN es un heterodúplex. Como se usa en el presente documento, el "heterodúplex" se refiere a una doble hélice formada por el espaciador de una guía de ARN y la cadena diana (TS). Como se usa en el presente documento, la expresión "región semilla" se refiere a la parte de la TS del heterodúplex que está inmediatamente en dirección 3' de la secuencia de PAM. La región semilla comprende las primeras bases que se emparejan con la guía de ARN en el heterodúplex y son necesarias para la unión de ARN-ADN y el desplazamiento de la TS. En algunas realizaciones, una alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y el heterodúplex aumenta los contactos de ácidos nucleicos no específicos. En algunas realizaciones, una alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y el heterodúplex aumenta la formación/estabilidad de los complejos ternarios con respecto a un polipéptido precursor.

En algunas realizaciones, la alteración que aumenta las interacciones y/o la afinidad entre el polipéptido Cas12i4 variante y el heterodúplex es la sustitución de uno o más aminoácidos. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de uno o más aminoácidos que contactan con el heterodúplex. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos en al menos un dominio/motivo (por ejemplo, el dominio Wedge, el dominio Rec1, el dominio Rec2 o el motivo RuvC2) a un lisina, arginina, histidina, serina, glutamina o asparagina. En algunas realizaciones, la sustitución o sustituciones de la interfaz de ácido nucleico aumentan las interacciones y/o la afinidad heterodúplex en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%,

100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor.

En algunas realizaciones, la sustitución que aumenta las interacciones heterodúplex aumenta la formación/estabilidad de complejos ternarios con respecto a un polipéptido precursor. En la Tabla 7 se muestran ejemplos no limitantes de sustituciones que pueden alterar la capacidad de un polipéptido Cas12i4 variante para interactuar con el heterodúplex. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 7 presenta interacciones heterodúplex aumentadas (formación de complejos ternarios) con respecto a un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 7 forma un complejo ternario más estable, en comparación con un polipéptido precursor.

Tabla 7. Sustituciones que alteran las interacciones heterodúplex.

Residuo	Sustitución o Sustituciones
S294*	K, Q, R
T355	R, K
T792	R, K
I793	R, K, Q
E811	K, R, Q
T815	R, K, H, N
S819	R, K
Q872	R, K
G9*	R, K
E156*	R, K, Q
T347	R, K, H
E349	R, K, Q
Y447*	R, K, S
V585*	R, K
S731	R, K
A775	R, K
S779	R, K
E794	R, K, Q
S855	R, K, Q
D74*	R, K, N
V442*	A, K
S113	R, K
N353	R, K, H

5	I401	R, K
	E578*	R, K
	D846*	R, K, N
	S860	K, Q
	Q869	R, K
10	Y116*	R, K
	E313*	R, K
	H428	R, K
	S776	R, K
	P7*	N
15	D114	R
	Q301*	H
	V730	K
	S309*	K
	V436	K
20	N445*	R, K
	C866	S
	G789	R, K

* Sustitución en región semilla

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 7. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 7.

En algunas realizaciones, un polipéptido Cas12i4 variante que presenta interacciones heterodúplex aumentadas comprende dos o más sustituciones enumeradas en la Tabla 7. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una formación/estabilidad de complejos ternarios aumentada comprende dos o más sustituciones enumeradas en la Tabla 7. En algunas realizaciones, un polipéptido Cas12i4 variante comprende V585R e Y447S. En algunas realizaciones, un polipéptido Cas12i4 variante comprende V585K e Y447S. En algunas realizaciones, un polipéptido Cas12i4 variante comprende V585R e Y447K. En algunas realizaciones, un polipéptido Cas12i4 variante comprende V585K e Y447K. En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 7. En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además V585R y Y447S, V585K y Y447S, V585R y Y447K o V585K y Y447K. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una cualquiera o más sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 presenta interacciones y/o una afinidad heterodúplex aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%,

15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad enzimática aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

30 *Estabilidad aumentada de dúplex y heterodúplex de ADN bicatenario*

Durante la formación del complejo ternario, el ADN bicatenario en dirección 3' de la secuencia de PAM se funde (por ejemplo, se desenrolla) en una cadena diana (TS) y una cadena no diana (NTS). El espaciador de una guía de ARN se une a la TS, formando una doble hélice que se denomina heterodúplex. La secuencia de PAM no se funde y permanece intacta como ADN bicatenario. Esto da como resultado una exposición parcial de este par de bases de ADNbc de PAM terminal al medio ambiente y la proteína, lo que puede ser desfavorable energéticamente. De manera similar, el par de bases terminal del heterodúplex está expuesto y puede ser desfavorable energéticamente. En algunas realizaciones, una alteración que aumenta las interacciones aromáticas, hidrófobas, de Van der Waals y/o catión-pi entre el

polipéptido Cas12i4 variante y las bases de PAM terminales expuestas del dúplex de ADN bicatenario o las bases terminales del heterodúplex aumenta la estabilidad de la fusión del ADN durante la formación de complejos ternarios. En algunas realizaciones, una alteración que aumenta las interacciones aromáticas, hidrófobas, de Van der Waals y/o catión-pi entre el polipéptido Cas12i4 variante y las bases expuestas del dúplex o heterodúplex de ADN bicatenario aumenta la estabilidad del bucle R durante la formación de complejos ternarios. En algunas realizaciones, una alteración que aumenta las interacciones aromáticas, hidrófobas, de Van der Waals y/o catión-pi entre el polipéptido Cas12i4 variante y las bases expuestas del dúplex o heterodúplex de ADN bicatenario aumenta la formación de complejos ternarios. En algunas realizaciones, una alteración que aumenta las interacciones aromáticas, hidrófobas, de Van der Waals y/o catión-pi entre el polipéptido Cas12i4 variante y las bases expuestas del dúplex o heterodúplex de ADN bicatenario aumenta la estabilidad de complejos ternarios.

En algunas realizaciones, la alteración que aumenta las interacciones aromáticas, hidrófobas, de Van der Waals y/o catión-pi es la sustitución de uno o más residuos. En algunas realizaciones, la alteración que aumenta las interacciones aromáticas, hidrófobas, de Van der Waals y/o catión-pi es la sustitución de uno o más residuos en contacto con el dúplex y/o heterodúplex de ADN bicatenario. En algunas realizaciones, la alteración que aumenta las interacciones aromáticas, hidrófobas, de Van der Waals y/o catión-pi es una sustitución enumerada en la Tabla 8. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución enumerada en la Tabla 8 presenta interacciones aromáticas, hidrófobas, Van der Waals y/o catión-pi aumentadas entre el polipéptido Cas12i4 variante y las bases expuestas del dúplex o heterodúplex de ADN bicatenario en comparación con un polipéptido precursor. En algunas realizaciones, la alteración incluye sustituir aminoácidos adyacentes a los pares de bases del dúplex terminal con aminoácidos de cadena ramificada, hidrófoba, aromática o con carga positiva para crear condiciones energéticamente más favorables para el ADN bicatenario y el heterodúplex.

Tabla 8. Sustituciones que estabilizan el bucle R.

Residuos	Sustitución o Sustituciones
I4	A
S5	Q, I, M
Y876	W, H
E156	R
E158	Q, K, R
A161	M, R, Y

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 8. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 8.

En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, estabilizando la

fusión del ADN y/o el bucle R) comprende dos o más sustituciones enumeradas en la Tabla 8. En algunas realizaciones, un polipéptido Cas12i4 variante comprende I4A y Y876W. En algunas realizaciones, un polipéptido Cas12i4 variante comprende E156R y E158Q. En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 8. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una cualquiera o más sustituciones de la Tabla 4, la Tabla 5, la Tabla 6, la Tabla 7 y/o la Tabla 8. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4, la Tabla 5, la Tabla 6, la Tabla 7 y/o la Tabla 8 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4, la Tabla 5, la Tabla 6, la Tabla 7 y/o la Tabla 8 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4, la Tabla 5, la Tabla 6, la Tabla 7 y/o la Tabla 8 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad enzimática aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

Cambios conformacionales aumentados

Los cambios conformacionales, por ejemplo, tras la unión de la guía de ARN o el ADN diana, afectan a la función de un polipéptido Cas12i4 variante, por ejemplo, los cambios conformacionales pueden alterar la cinética del polipéptido Cas12i4 variante. El dominio Rec1 (Helicoidal II) de Cas12i4 se mueve y gira para adaptarse a la unión del ADN durante la formación de complejos ternarios. En algunas realizaciones, una alteración que aumenta el movimiento (por ejemplo, la flexibilidad o los cambios conformacionales) del dominio Helicoidal II aumenta la unión a ADN/afinidad de unión a ADN. En algunas realizaciones, una sustitución para aumentar la flexibilidad, por ejemplo, una sustitución de un aminoácido voluminoso por un aminoácido con una cadena lateral pequeña o más pequeña (residuo de alanina, valina,

glicina o serina), en el dominio Helicoidal II aumenta la formación de complejos ternarios. En algunas realizaciones, una alteración que aumenta el movimiento (por ejemplo, la flexibilidad o los cambios conformacionales) del dominio Helicoidal II aumenta la estabilidad de los complejos ternarios. En algunas realizaciones, la alteración que aumenta los cambios conformacionales del dominio Helicoidal II es la sustitución de uno o más residuos con un residuo de alanina, valina, glicina o serina. En algunas realizaciones, la alteración que aumenta la flexibilidad del dominio Helicoidal II es la sustitución de uno o más residuos. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos cerca del dominio Helicoidal II. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos cerca del dominio Helicoidal II. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una sustitución expuesta en la Tabla 9.

15

20

25

30

35

Tabla 9. Sustituciones que alteran la flexibilidad del dominio Helicoidal II.

5	Sustituciones de aminoácidos
	D328G+F330V
	D328G+F330N
	D328A+F330V
	D328A+F330N
10	D328G
	D328A
	F330V
	F330N
	P440A
15	P440G
	L324A
	A437G
	K326G
	K326S

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 9. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 9.

En algunas realizaciones, la alteración que aumenta la flexibilidad del dominio Helicoidal II es una sustitución enumerada en la Tabla 9. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones enumeradas en la Tabla 9 presenta una mayor flexibilidad del dominio Helicoidal II en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor. En algunas realizaciones, la alteración que aumenta la unión a ADN/afinidad por el ADN es una sustitución enumerada en la Tabla 9. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones enumeradas en la Tabla 9 presenta una unión a ADN/afinidad por el ADN aumentada en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%

74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor.

5 En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución enumerada en la Tabla 9 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 10 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

15 En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una mayor flexibilidad del dominio Helicoidal II comprende dos o más sustituciones enumeradas en la Tabla 9. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una unión la ADN/afinidad por el ADN aumentada comprende dos o más sustituciones enumeradas en la Tabla 9. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una formación/estabilidad de complejos ternarios aumentada comprende dos o más sustituciones enumeradas en la Tabla 9. En algunas realizaciones, un polipéptido 20 Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 9. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una o más sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 presenta una unión/afinidad por el ADN 25 aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 30 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por 35 ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%,

72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

5 En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 presenta una actividad

10 enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad enzimática aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%,

15 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, una alteración que aumenta las conexiones entre la interfaz Nuc y

20 Helicoidal II, que se forma cuando el ADN monocatenario objetivo está en el sitio activo de un polipéptido Cas12i4, aumenta la transición de complejo binario a complejo ternario. En algunas realizaciones, una alteración que aumenta las conexiones entre la interfaz Nuc y Helicoidal II aumenta la formación de complejos ternarios. En algunas realizaciones, una alteración que aumenta las conexiones entre la interfaz Nuc y Helicoidal II aumenta la estabilidad de los complejos ternarios. En algunas realizaciones, la alteración

25 que aumenta las conexiones entre la interfaz Nuc y Helicoidal II es sustituir uno o más residuos con un residuo de ácido aspártico, ácido glutámico, arginina o lisina. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una sustitución expuesta en la Tabla 10.

Tabla 10. Sustituciones que aumentan las conexiones en la interfaz Nuc y Helicoidal II.

30	Sustituciones de aminoácidos
	Q386E+Q387D+N966R
	Q386E+Q387E+N966R
35	Q386E+N966R
	Q386E+Q387D+A936K+N966R
	Q386E+Q387E+A936K+N966R
	Q387D
	A936K
	Q387E

S931K
N932K

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 10. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 10.

En algunas realizaciones, una sustitución de la Tabla 10 aumenta las conexiones entre la interfaz Nuc y Helicoidal II. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 10 aumenta las conexiones entre la interfaz Nuc y Helicoidal II en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una sustitución enumerada en la Tabla 10 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, un polipéptido Cas12i4 variante que presenta conexiones aumentadas entre la interfaz Nuc y Helicoidal II comprende dos o más sustituciones enumeradas en la Tabla 10. En algunas realizaciones, un polipéptido Cas12i4 variante que presenta una formación/estabilidad de complejos ternarios aumentada comprende dos o más sustituciones enumeradas en la Tabla 10. En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 10. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una o más sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 presenta conexiones aumentadas entre la interfaz Nuc y Helicoidal II (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%,

35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante con una o más de las sustituciones de la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 presenta una formación de complejos ternarios y/o una estabilidad de complejos ternarios aumentadas (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 presenta una actividad enzimática aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad enzimática aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, una alteración reduce las conexiones entre la interfaz Nuc y Helicoidal II. En algunas realizaciones, una alteración que disminuye las conexiones entre la interfaz Nuc y Helicoidal II aumenta la formación de complejos ternarios. En algunas realizaciones, una alteración que disminuye las conexiones entre la interfaz Nuc y Helicoidal II es sustituir uno o más residuos. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 10.

Fidelidad aumentada

En algunos aspectos, un polipéptido Cas12i4 variante comprende una alteración que aumenta la especificidad en la diana con respecto a un polipéptido precursor. En algunos aspectos, un polipéptido Cas12i4 variante comprende una alteración que aumenta la unión en la diana con respecto a un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que aumenta las interacciones (por ejemplo, la afinidad) entre el polipéptido Cas12i4 variante y el ADN en la diana con respecto a un polipéptido precursor.

En algunas realizaciones, la alteración que aumenta la especificidad en la diana es sustituir uno o más aminoácidos. En algunos aspectos, la alteración que aumenta la especificidad en la diana es truncar un residuo que contacta con la secuencia espaciadora de una guía de ARN (por ejemplo, sustituir un residuo que contacta con la secuencia espaciadora con un residuo que tiene una cadena lateral más pequeña). En algunos aspectos, la alteración que aumenta la especificidad en la diana es truncar un residuo que contacta con la secuencia espaciadora de una guía de ARN.

En algunas realizaciones, el polipéptido Cas12i4 variante comprende una sustitución de uno o más aminoácidos que contactan con la secuencia espaciadora de una guía de ARN. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos en al menos un dominio/motivo (por ejemplo, el dominio Wedge, el dominio Rec1, el dominio Rec2 o el motivo RuvC2). En algunas realizaciones, una sustitución truncada en el dominio Helicoidal II da como resultado un polipéptido Cas12i4 variante que presenta una especificidad de unión potenciada en la diana.

En algunas realizaciones, la sustitución o sustituciones aumentan la especificidad en la diana con el polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor. En algunas realizaciones, la sustitución o sustituciones aumentan la unión en la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor.

En algunas realizaciones, la sustitución o sustituciones aumentan la afinidad de unión en la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%,

14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos, en comparación con un polipéptido precursor.

Los ejemplos no limitantes de alteraciones que pueden alterar la capacidad de un polipéptido Cas12i4 variante para unirse selectivamente al ADN en la diana son las sustituciones enumeradas en la Tabla 11. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 11 presenta una especificidad en la diana aumentada con respecto a un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 11 presenta una unión en la diana aumentada con respecto a un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 11 presenta una afinidad de unión en la diana aumentada con respecto a un polipéptido precursor.

Tabla 11. Sustituciones que aumentan la especificidad en la diana.

Sustitución	o
Sustituciones	
E349A	
D350A	
E349A+E350A	
R446A	
R306A	
T348A	
H439A	
N786A	
M863A	
K303A	
T305A	
K437A	
M574A	
E811A	
V850A	
K852A	
S855A	
K856A	

5	K857A
	N859A
	S860A
	F352A
	R357A
10	K393A
	Q398A
	P302A
	S309A
	K432A
15	R435A
	Y447A
	R734A
	S779A
	T815A
20	R818A
	K868A
	R873A
	K115A
	K154A
25	T347A
	Y358A
	V442A
	K576A
	S783A
30	K787A
	K791A
	I793A
	Q872A
	Y116A
35	P400A
	N445A
	N782A
	C866A
	Q869A

En algunas realizaciones, la alteración que aumenta la especificidad en la diana (por ejemplo, una sustitución enumerada en la Tabla 11) aumenta adicionalmente la formación de complejos ternarios en la diana y/o la estabilidad de los complejos ternarios en la diana (por ejemplo, la formación/estabilidad de complejos ternarios en la diana). En algunas realizaciones, la alteración que aumenta la especificidad en la diana aumenta la formación de complejo ternarios en la diana y/o la estabilidad de los complejos ternarios en la diana en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos en comparación con un polipéptido precursor.

En algunos aspectos, un polipéptido Cas12i4 variante comprende una alteración que disminuye la especificidad fuera de la diana con respecto a un polipéptido precursor. En algunos aspectos, un polipéptido Cas12i4 variante comprende una alteración que disminuye la unión fuera de la diana con respecto a un polipéptido precursor. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración que reduce las interacciones (por ejemplo, la afinidad) entre el polipéptido Cas12i4 variante y el ADN fuera de la diana con respecto a un polipéptido precursor.

Se conocen en la técnica métodos para detectar actividad fuera de la diana. En algunas realizaciones, la actividad fuera de la diana se detecta mediante la secuenciación del sitio de integración de marcadores basada en la tagmentación (TTISS) o la identificación imparcial de DSB en todo el genoma habilitada mediante la secuenciación (GUIDE-Seq). Por ejemplo, en algunas realizaciones, se realiza la TTISS usando un polipéptido Cas12i4 o un polipéptido Cas12i4 variante usando el método de TTISS descrito en el documento PCT/US2021/025257, que se incorpora por referencia en su totalidad.

En algunas realizaciones, la alteración que disminuye la especificidad fuera de la diana es sustituir uno o más aminoácidos por un residuo de alanina, serina, valina, glutamina o asparagina. En algunos aspectos, la alteración que disminuye la especificidad fuera de la diana es truncar un residuo que contacta con la secuencia espaciadora de una guía de ARN (por ejemplo, sustituir un residuo que contacta con la secuencia espaciadora con un residuo que tiene una cadena lateral más pequeña). En algunos aspectos, la alteración que disminuye la especificidad fuera de la diana es truncar un residuo, por ejemplo, la sustitución por alanina, serina o valina, que contacta con la secuencia espaciadora de una guía de ARN. En algunas realizaciones, el polipéptido Cas12i4 variante comprende una alteración de uno o más aminoácidos en al menos un dominio/motivo (por ejemplo, el dominio Wedge, el dominio Rec1, el dominio Rec2 o el motivo RuvC2) a una alanina. En algunas realizaciones, una sustitución truncada en el dominio Helicoidal II da como resultado un polipéptido Cas12i4 variante que presenta una especificidad de unión fuera de la diana disminuida. En algunas realizaciones, la sustitución o sustituciones disminuyen la especificidad fuera de la diana con el polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%,

7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99 % o el 100 %, en comparación con un polipéptido precursor. En algunas realizaciones, la sustitución o sustituciones disminuyen la unión fuera de la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o el 100 %, en comparación con un polipéptido precursor. En algunas realizaciones, la sustitución o sustituciones disminuyen la afinidad de unión fuera de la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o el 100 %, en comparación con un polipéptido precursor.

Los ejemplos no limitantes de alteraciones que pueden alterar la capacidad de un polipéptido Cas12i4 variante para unirse a ADN fuera de la diana son las sustituciones enumeradas en la Tabla 11. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 11 presenta una especificidad fuera de la diana disminuida con respecto a un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 11 presenta una unión fuera de la diana disminuida con respecto a un polipéptido precursor. En algunas realizaciones, un polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 11 presenta una afinidad de unión fuera de la diana reducida con respecto a un polipéptido precursor.

En algunas realizaciones, la sustitución o sustituciones que aumentan la especificidad en la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más, o cualquier porcentaje entre ellos, disminuye

adicionalmente la especificidad fuera de la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o el 100%, en comparación con un polipéptido precursor. En algunas realizaciones, la sustitución o sustituciones que aumentan la unión en la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más, o cualquier porcentaje entre ellos, disminuye adicionalmente la unión fuera de la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o el 100%, en comparación con un polipéptido precursor. En algunas realizaciones, la sustitución o sustituciones que aumentan la afinidad de unión en la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más, o cualquier porcentaje entre ellos, disminuye adicionalmente la afinidad de unión fuera de la diana del polipéptido Cas12i4 variante en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o el 100%, en comparación con un polipéptido precursor.

En algunas realizaciones, un polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 comprende además una o más sustituciones enumeradas en la Tabla 11. En algunas realizaciones, un polipéptido Cas12i4 variante comprende una o más sustituciones enumeradas en la Tabla 2 y la Tabla 11.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11 presenta una actividad enzimática en la diana aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11 presenta una actividad enzimática en la diana aumentada. En algunas realizaciones, el polipéptido Cas12i4 variante presenta una actividad enzimática en la diana aumentada (por ejemplo, en aproximadamente el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos) en comparación con un polipéptido precursor.

En algunas realizaciones, el polipéptido Cas12i4 variante de una cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11 presenta una relación aumentada de actividad enzimática en la diana con respecto a actividad enzimática fuera de la diana. En algunas realizaciones, el polipéptido Cas12i4 variante que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11 presenta una relación aumentada de actividad enzimática en la diana con respecto a actividad enzimática fuera de la diana. En algunas realizaciones, la actividad enzimática en la diana del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) es al menos el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos superior a la actividad enzimática fuera de la diana del polipéptido Cas12i4 variante, en comparación con un polipéptido precursor. En algunas realizaciones, la actividad enzimática en la diana del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante que comprende una o más sustituciones

enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) es al menos el 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, o más o cualquier porcentaje entre ellos superior a la actividad enzimática fuera de la diana del polipéptido Cas12i4 variante, en comparación con un polipéptido precursor.

En algunas realizaciones, la actividad enzimática del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 en un locus fuera de la diana no es más del 10 % (por ejemplo, el 10 %, el 9 %, el 8 %, el 7 %, el 6 %, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la actividad enzimática en el locus en la diana. En algunas realizaciones, la actividad enzimática del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) en un locus fuera de la diana no es más del 10 % (por ejemplo, el 10 %, el 9 %, el 8 %, el 7 %, el 6 %, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la actividad enzimática en el locus en la diana. En algunas realizaciones, la actividad enzimática del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 en un locus fuera de la diana no es más del 5 % (por ejemplo, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la actividad enzimática en el locus en la diana. En algunas realizaciones, la actividad enzimática del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) en un locus fuera de la diana no es más del 5 % (por ejemplo, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la actividad enzimática en el locus en la diana. En comparación, la actividad enzimática de SpCas9 en un locus fuera de la diana es de hasta el 95 % (por ejemplo, el 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42%, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, o el 0 %) de la actividad enzimática en el locus en la diana.

En algunas realizaciones, la eficiencia de edición del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 en un locus fuera de la diana no es más del 10 % (por ejemplo, el 10 %, el 9 %, el 8 %, el 7 %, el 6 %, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la eficiencia de edición en el locus en la diana. En algunas realizaciones, la eficiencia de edición del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID

NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) en un locus fuera de la diana no es más del 10 % (por ejemplo, el 10 %, el 9 %, el 8 %, el 7 %, el 6 %, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la eficiencia de edición en el locus en la diana. En algunas realizaciones, la eficiencia de edición del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 en un locus fuera de la diana no es más del 5 % (por ejemplo, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la eficiencia de edición en el locus en la diana. En algunas realizaciones, la eficiencia de edición del polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) en un locus fuera de la diana no es más del 5 % (por ejemplo, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la eficiencia de edición en el locus en la diana. En comparación, la eficiencia de edición de SpCas9 en un locus fuera de la diana es de hasta el 95 % (por ejemplo, el 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42%, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, o 0%) de la eficiencia de edición en el locus en la diana.

En algunas realizaciones, la edición mediante el polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 en un locus fuera de la diana no es más del 10 % (por ejemplo, el 10 %, el 9 %, el 8 %, el 7 %, el 6 %, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la edición en el locus en la diana. En algunas realizaciones, la edición mediante el polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) en un locus fuera de la diana no es más del 10 % (por ejemplo, el 10 %, el 9 %, el 8 %, el 7 %, el 6 %, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la edición en el locus en la diana. En algunas realizaciones, la edición mediante el polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 en un locus fuera de la diana no es más del 5 % (por ejemplo, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la edición en el locus en la diana. En algunas realizaciones, la edición mediante el polipéptido Cas12i4 variante (por ejemplo, el polipéptido Cas12i4 variante de cualquiera de las SEQ ID NO: 2-59 que comprende una o más sustituciones enumeradas en la Tabla 4 y/o la Tabla 5 y/o la Tabla 6 y/o la Tabla 7 y/o la Tabla 8 y/o la Tabla 9 y/o la Tabla 10 y/o la Tabla 11) en un locus fuera de la diana no es más del 5 % (por ejemplo, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % o el 0 %) de la edición en el locus en la diana. En comparación, la edición de SpCas9 en un locus fuera de la diana es de hasta el 95 % (por ejemplo, el 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%,

56%, 55%, 54%, 53%, 52%, 51%, 50%, 49%, 48%, 47%, 46%, 45%, 44%, 43%, 42%, 41%, 40%, 39%, 38%, 37%, 36%, 35%, 34%, 33%, 32%, 31%, 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, o 0%) de la edición en el locus en la diana.

5 **Guía de ARN**

En algunas realizaciones, una composición o complejo como se describe en el presente documento comprende un resto de direccionamiento (por ejemplo, una guía de ARN, antisentido, oligonucleótidos, conjugados de péptido oligonucleótido) que se une al ácido nucleico diana e interactúa con el polipéptido Cas12i4 (por ejemplo, el polipéptido precursor o el polipéptido Cas12i4 variante). El resto de direccionamiento puede unirse a un ácido nucleico diana (por ejemplo, con una afinidad de fuera de la diana al ácido nucleico diana).

En algunas realizaciones, el resto de direccionamiento comprende, o es, una guía de ARN. En algunas realizaciones, la guía de ARN se dirige el polipéptido Cas12i4 (por ejemplo, el polipéptido precursor o el polipéptido Cas12i4 variante) a una secuencia de ácido nucleico particular. Los expertos en la materia que lean los siguientes ejemplos de tipos particulares de guías de ARN comprenderán que, en algunas realizaciones, una guía de ARN es específica de sitio. Es decir, en algunas realizaciones, una guía de ARN se asocia específicamente a una o más secuencias de ácido nucleico diana (por ejemplo, ADN específico o secuencias de ADN genómico) y no a secuencias de ácido nucleico no dirigidas (por ejemplo, ADN no específico o secuencias aleatorias).

En algunas realizaciones, la composición como se describe en el presente documento comprende una guía de ARN que se asocia al polipéptido Cas12i4 (por ejemplo, el polipéptido precursor o el polipéptido Cas12i4 variante) y dirige el polipéptido Cas12i4 a una secuencia de ácido nucleico diana (por ejemplo, ADN).

La guía de ARN puede dirigir (por ejemplo, se asocia a, se dirige a, entra en contacto, o se une a) uno o más nucleótidos de una secuencia diana, por ejemplo, una secuencia específica de sitio o una diana específica de sitio. En algunas realizaciones, la nucleoproteína (por ejemplo, el polipéptido precursor o el polipéptido Cas12i4 variante más una guía de ARN) se activa tras la unión a un ácido nucleico diana que es complementario a una secuencia de direccionamiento de ADN en la guía de ARN (por ejemplo, un sustrato específico de la secuencia o ácido nucleico diana).

En algunas realizaciones, una guía de ARN comprende un espaciador que tiene una longitud de aproximadamente 11 nucleótidos a aproximadamente 100 nucleótidos. Por ejemplo, el segmento de direccionamiento al ADN puede tener una longitud de aproximadamente 11 nucleótidos a aproximadamente 80 nucleótidos, de aproximadamente 11 nucleótidos a aproximadamente 50 nucleótidos, de aproximadamente 11 nucleótidos a aproximadamente 40 nucleótidos, de aproximadamente 11 nucleótidos a aproximadamente 30 nucleótidos, de aproximadamente 11 nucleótidos a aproximadamente 25 nucleótidos, de aproximadamente 11 nucleótidos a aproximadamente 20 nucleótidos o de aproximadamente 11 nucleótidos a aproximadamente 19 nucleótidos. Por ejemplo, el espaciador puede tener una longitud de aproximadamente 19 nucleótidos a aproximadamente 20 nucleótidos, de

aproximadamente 19 nucleótidos a aproximadamente 25 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 30 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 35 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 40 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 45 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 50 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 60 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 70 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 80 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 90 nucleótidos, de aproximadamente 19 nucleótidos a aproximadamente 100 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 25 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 30 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 35 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 40 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 45 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 50 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 60 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 70 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 80 nucleótidos, de aproximadamente 20 nucleótidos a aproximadamente 90 nucleótidos o de aproximadamente 20 nucleótidos a aproximadamente 100 nucleótidos.

En algunas realizaciones, en general, el espaciador de la guía de ARN puede diseñarse para que tenga una longitud de entre 11 y 50 nucleótidos (por ejemplo, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 o 35 nucleótidos) y sea complementario a una secuencia de ácido nucleico diana específica. En algunas realizaciones particulares, la guía de ARN puede diseñarse para que sea complementaria a una cadena de ADN específica, por ejemplo, de un locus genómico. En algunas realizaciones, la secuencia de direccionamiento de ADN se diseña para que sea complementaria a una cadena de ADN específica, por ejemplo, de un locus genómico.

La guía de ARN puede ser sustancialmente idéntica a una cadena complementaria de una secuencia de ácido nucleico de referencia. En algunas realizaciones, la guía de ARN comprende una secuencia que tiene al menos aproximadamente un 60 %, al menos aproximadamente un 65 %, al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 91 %, al menos aproximadamente un 92 %, al menos aproximadamente un 93 %, al menos aproximadamente un 94 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 % o al menos aproximadamente un 99.5 % de identidad de secuencia con una cadena complementaria de una secuencia de ácido nucleico de referencia, por ejemplo, ácido nucleico diana. El porcentaje de identidad entre dos de dichos ácidos nucleicos puede determinarse manualmente mediante inspección de las dos secuencias de ácido nucleico alineadas de manera óptima o usando programas de software o algoritmos (por ejemplo, BLAST, ALIGN, CLUSTAL) usando parámetros convencionales.

En algunas realizaciones, la guía de ARN tiene al menos aproximadamente un 60 %, al menos aproximadamente un 65 %, al menos aproximadamente un 70 %, al menos aproximadamente un 75 %, al

menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 91 %, al menos aproximadamente un 92 %, al menos aproximadamente un 93 %, al menos aproximadamente un 94 %, al menos aproximadamente un 95 %, al menos aproximadamente un 96 %, al menos aproximadamente un 97 %, al menos aproximadamente un 98 %, al menos aproximadamente un 99 %, o al menos aproximadamente un 99.5 % de identidad de secuencia con una cadena complementaria de un ácido nucleico diana.

En algunas realizaciones, la guía de ARN comprende un espaciador que tiene una longitud de entre 11 y 50 nucleótidos (por ejemplo, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 o 35 nucleótidos) y al menos un 80 %, al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % complementario a un ácido nucleico diana. En algunas realizaciones, la guía de ARN comprende una secuencia al menos un 80 %, al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % complementaria a una secuencia de ADN diana. En algunas realizaciones, la guía de ARN comprende una secuencia al menos un 80 %, al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % complementaria a una secuencia genómica diana. En algunas realizaciones, la guía de ARN comprende una secuencia, por ejemplo, secuencia de ARN, que tiene una longitud de hasta el 50 y al menos un 80 %, al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % complementaria a un ácido nucleico diana. En algunas realizaciones, la guía de ARN comprende una secuencia al menos un 80 %, al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % complementaria a una secuencia de ADN diana. En algunas realizaciones, la guía de ARN comprende una secuencia al menos un 80 %, al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % complementaria a una secuencia genómica diana.

En determinadas realizaciones, la guía de ARN incluye, consiste esencialmente en, o comprende una secuencia repetitiva de directa unida a una secuencia de direccionamiento de ADN. En algunas realizaciones, la guía de ARN incluye una secuencia repetitiva directa y una secuencia de direccionamiento de ADN o una repetición directa- secuencia de direccionamiento de ADN -secuencia repetitiva directa. En algunas realizaciones, la guía de ARN incluye una secuencia repetitiva directa truncada y una secuencia de direccionamiento de ADN, que es típica de ARNcr procesado o maduro. En algunas realizaciones, el polipéptido Cas12i4 (por ejemplo, el polipéptido precursor o el polipéptido Cas12i4 variante) forma un complejo con la guía de ARN, y la guía de ARN dirige el complejo para que se asocie al ácido nucleico diana específico de sitio que es complementario a al menos una porción de la guía de ARN.

En algunas realizaciones, la secuencia repetitiva directa de la guía de ARN tiene una longitud de entre 12-100, 13-75, 14-50 o 15-40 nucleótidos (por ejemplo, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40 nucleótidos).

En algunas realizaciones, la secuencia repetitiva directa es una secuencia de la Tabla 12 o una porción de una secuencia de la Tabla 12. La secuencia repetitiva directa puede comprender del nucleótido 1 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116,

117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 2 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 3 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 4 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 5 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 6 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 7 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 8 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 9 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 10 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 11 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 12 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 13 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede comprender del nucleótido 14 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124.

En algunas realizaciones, la secuencia repetitiva directa tiene al menos un 95 % de identidad (por ejemplo, al menos un 95 %, 96 %, 97 %, 98 % o 99 % de identidad) con una secuencia de la Tabla 12 o una porción de una secuencia de la Tabla 12. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del nucleótido 1 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 2 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 3 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que

comprende del 4 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 5 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 6 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 7 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 8 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 9 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 10 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 11 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 12 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 95 % de identidad con una secuencia que comprende del 13 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124.

En algunas realizaciones, la secuencia repetitiva directa tiene al menos un 90 % de identidad (por ejemplo, al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad) con una secuencia de la Tabla 12 o a una porción de una secuencia de la Tabla 12. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del nucleótido 1 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 2 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 3 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 4 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia

repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 5 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 6 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 7 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 8 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 9 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 10 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 11 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 12 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. La secuencia repetitiva directa puede tener al menos un 90 % de identidad con una secuencia que comprende del 13 al nucleótido 36 de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124.

En algunas realizaciones, la secuencia repetitiva directa es al menos un 90 % idéntica al complemento inverso de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. En algunas realizaciones, la secuencia repetitiva directa es al menos un 95 % idéntica al complemento inverso de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124. En algunas realizaciones, la secuencia repetitiva directa es el complemento inverso de una cualquiera de las SEQ ID NO: 60, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123 o 124.

Tabla 12. Secuencias de repetición directa.

Identificador de secuencia	Secuencia de repetición directa
SEQ ID NO: 60	UCUCAACGAUAGUCAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 108	UUUUAAACAACACUCAGGCAUGUGUCCACAGUGACAC
SEQ ID NO: 109	UUGAACGGAUACUCAGACAUGUGUUUCCAGUGACAC
SEQ ID NO: 110	UGCCCUCAAUAGUCAGAUGUGUGUCCACAGUGACAC
SEQ ID NO: 111	UCUCAAUAGAUACUUAAGAUACGUGUCCUCAGUGACAC
SEQ ID NO: 112	UCUCAAUAGAUACUCAGACAUGUGUCCCCAGUGACAC
SEQ ID NO: 113	UCUCAAUAGAUACUAAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 114	UCUCAACUAUACUCAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 115	UCUCAACGAUACUCAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 116	UCUCAACGAUACUAAGAUUAUGUGUCCUCAGCGACAC
SEQ ID NO: 117	UCUCAACGAUACUAAGAUUAUGUGUCCCCAGUGACAC
SEQ ID NO: 118	UCUCAACGAUACUAAGAUUAUGUGUCCACAGUGACAC
SEQ ID NO: 119	UCUCAACAAUACUCAGACAUGUGUCCCCAGUGACAC
SEQ ID NO: 120	UCUCAACAAUACUAAGGCAUGUGUCCCCAGUGACCC
SEQ ID NO: 121	UCUCAAGAUACUCAGACACGUGUCCCCAGUGACAC
SEQ ID NO: 122	UCUCAAAAAUACUCAGACAUGUGUCCUCAGUGACAC
SEQ ID NO: 123	GCGAAACAACAGUCAGACAUGUGUCCCCAGUGACAC
SEQ ID NO: 124	CCUCAACGAUAUUAAGACAUGUGUCCGCAGUGACAC
SEQ ID NO: 61	AGACAUGUGUCCUCAGUGACAC

En algunas realizaciones, la secuencia repetitiva directa es AGN₁N₂N₃N₄GUGUN₅N₆N₇CAGN₈GACN₉C (SEQ ID NO: 125), en donde N₁ es A o G, N₂ es C o U, N₃ es A o G, N₄ es U o C, N₅ es C o U, N₆ es C o U, N₇ es U, A, C o G, N₈ es U o C, y N₉ es A o C. En algunas realizaciones, la secuencia repetitiva directa de la SEQ ID NO: 125 se conoce como la DR madura de Cas12i4.

En algunas realizaciones, la secuencia repetitiva directa es al menos un 90 % idéntica a la SEQ ID NO: 61 o una porción de la SEQ ID NO: 61. En algunas realizaciones, la secuencia repetitiva directa es al menos un 95% idéntica a la SEQ ID NO: 61 o una porción de la SEQ ID NO: 61. En algunas realizaciones, la secuencia repetitiva directa es un 100 % idéntica a la SEQ ID NO: 61 o una porción de la SEQ ID NO: 61. En algunas realizaciones, la secuencia repetitiva directa de la SEQ ID NO: 61 se conoce como la DR madura de Cas12i4.

En algunas realizaciones, la composición o el complejo descritos en el presente documento incluye uno o más (por ejemplo, dos, tres, cuatro, cinco, seis, siete, ocho o más) guías de ARN, por ejemplo, una pluralidad de guías de ARN.

En algunas realizaciones, la guía de ARN tiene una arquitectura similar a, por ejemplo las Publicaciones Internacionales N.º WO 2014/093622 y WO 2015/070083, cuyo contenido completo se incorporan en el presente documento por referencia.

A menos que se indique otra cosa, todas las composiciones y complejos y polipéptidos proporcionados en el presente documento se preparan en referencia al nivel activo de esta composición o complejo o polipéptido, y excluyen impurezas, por ejemplo, disolventes residuales o subproductos, que pueden estar presentes en fuentes disponibles en el mercado. Los pesos de los componentes enzimáticos se basan en la proteína activa total. Todos los porcentajes y relaciones se calculan en peso a menos que se indique otra cosa. Todos los porcentajes y relaciones se calculan basándose en la composición total a menos que se indique otra cosa. En la composición ejemplificada, los niveles enzimáticos se expresan por enzima pura en peso de la composición total y, a menos que se especifique otra cosa, los componentes se expresan en peso de las composiciones totales.

Modificaciones

La guía de ARN o cualquiera de las secuencias de ácido nucleico que codifican los polipéptidos Cas12i4 puede incluir una o más modificaciones covalentes con respecto a una secuencia de referencia, en particular el polirribonucleótido precursor, que se incluyen dentro del alcance de la presente invención.

Las modificaciones a modo de ejemplo pueden incluir cualquier modificación con respecto al azúcar, la nucleobase, la unión internucleosídica (por ejemplo, a un fosfato de unión/a una unión de fosfodiéster/a la cadena principal del fosfodiéster) y cualquier combinación de los mismos. Algunas de las modificaciones a modo de ejemplo proporcionadas en el presente documento se describen en detalle a continuación.

La guía de ARN o cualquiera de las secuencias de ácido nucleico que codifican los componentes de los polipéptidos Cas12i4 descritos en el presente documento pueden incluir cualquier modificación útil, tal como al azúcar, la nucleobase o la unión internucleosídica (por ejemplo, a un fosfato de unión/a una unión fosfodiéster/a la cadena principal fosfodiéster). Uno o más átomos de una nucleobase de pirimidina pueden reemplazarse o sustituirse con amino opcionalmente sustituido, tiol opcionalmente sustituido, alquilo opcionalmente sustituido (por ejemplo, metilo o etilo) o halo (por ejemplo, cloro o flúor). En determinadas realizaciones, las modificaciones (por ejemplo, una o más modificaciones) están presentes en cada uno de entre el azúcar y la unión internucleosídica. Las modificaciones pueden ser modificaciones de ácidos ribonucleicos (ARN) a ácidos desoxirribonucleicos (ADN), ácidos nucleicos de treosa (TNA), ácidos nucleicos de glicol (GNA), ácidos nucleicos peptídicos (PNA), ácidos nucleicos bloqueados (LNA) o híbridos de los mismos). En el presente documento se describe modificaciones adicionales.

En algunas realizaciones, la modificación puede incluir una modificación inducida química o celularmente. Por ejemplo, algunos ejemplos no limitantes de modificaciones de ARN intracelular se describen por Lewis y Pan en "RNA modifications and structures cooperate to guide RNA-protein interactions" de Nat Reviews Mol Cell Biol, 2017, 18:202-210.

Pueden existir diferentes modificaciones de azúcar, modificaciones de nucleótidos y/o uniones internucleosídicas (por ejemplo, estructuras de cadena principal) en diversas posiciones de la secuencia.

Un experto en la materia apreciará que los análogos de nucleótidos u otras modificaciones pueden ubicarse en cualquier posición o posiciones de la secuencia, de modo que la función de la secuencia no se reduce sustancialmente. La secuencia puede incluir de aproximadamente el 1 % a aproximadamente el 100 % de nucleótidos modificados (o bien con respecto a todo el contenido de nucleótidos, o con respecto a uno o más tipos de nucleótidos, es decir, uno cualquiera o más de A, G, U o C) o cualquier porcentaje intermedio (por ejemplo, del 1 % al 20 %, del 1 % al 25 %, del 1 % al 50 %, del 1 % al 60 %, del 1 % al 70 %, del 1 % al 80 %, del 1 % al 90 %, del 1 % al 95 %, del 10 % al 20 %, del 10 % al 25 %, del 10 % al 50 %, del 10 % al 60 %, del 10 % al 70 %, del 10 % al 80 %, del 10 % al 90 %, del 10 % al 95 %, del 10 % al 100 %, del 20 % al 25 %, del 20 % al 50 %, del 20 % al 60 %, del 20 % al 70 %, del 20 % al 80 %, del 20 % al 90 %, del 20 % al 95 %, del 20 % al 100 %, del 50 % al 60 %, del 50 % al 70 %, del 50 % al 80 %, del 50 % al 90 %, del 50 % al 95 %, del 50 % al 100 %, del 70 % al 80 %, del 70 % al 90 %, del 70 % al 95 %, del 70 % al 100 %, del 80 % al 90 %, del 80 % al 95 %, del 80 % al 100 %, del 90 % al 95 %, del 90 % al 100 % y del 95 % al 100 %).

En algunas realizaciones, las modificaciones del azúcar (por ejemplo, en la posición 2' o la posición 4') o el reemplazo del azúcar en uno o más ribonucleótidos de la secuencia pueden, así como modificaciones de la cadena principal, incluir la modificación o el reemplazo de las uniones fosfodiéster. Los ejemplos específicos de una secuencia incluyen, pero no se limitan a, secuencias que incluyen cadenas principales modificadas o uniones internucleosídicas no naturales tales como modificaciones internucleosídicas, incluyendo la modificación o el reemplazo de las uniones fosfodiéster. Las secuencias que tienen cadenas principales modificadas incluyen, entre otras, aquellas que no tienen un átomo de fósforo en la cadena principal. Para los fines de la presente solicitud, y como a veces se hace referencia en la técnica, los ARN modificados que no tienen un átomo de fósforo en su cadena principal internucleosídica también pueden considerarse oligonucleósidos. En realizaciones particulares, una secuencia incluirá ribonucleótidos con un átomo de fósforo en su cadena principal internucleosídica.

Las cadenas principales de secuencia modificada pueden incluir, por ejemplo, fosforotioatos, fosforotioatos quirales, fosforoditioatos, fosfotriésteres, aminoalquilfosfotriésteres, fosfonatos de metilo y otros alquilos tales como 3'-alquilen fosfonatos y fosfonatos quirales, fosfinatos, aminofosforamidatos-tioramidato-fosfonatos tales como 3'-alquilen fosfonatos, tionoalquilfosfonatos, tionoalquilfosfotriésteres y boranofosfatos que tienen uniones 3'-5' normales, análogos unidos en 2'-5' de estos y aquellos que tienen polaridad invertida en donde los pares adyacentes de unidades de nucleósidos se unen 3'-5' a 5'-3' o 2'-5' a 5'-2'. También se incluyen diversas sales, sales mixtas y formas de ácido libre. En algunas realizaciones, la secuencia puede tener carga negativa o positiva.

Los nucleótidos modificados, que pueden incorporarse en la secuencia, pueden modificarse en la unión internucleosídica (por ejemplo, cadena principal de fosfato). En el presente documento, en el contexto de la cadena principal del polinucleótido, los términos "fosfato" y "fosfodiéster" se usan indistintamente. Los grupos fosfato de la cadena principal pueden modificarse reemplazando uno o más átomos de oxígeno por un sustituyente diferente. Además, los nucleósidos y nucleótidos modificados pueden incluir el reemplazo total de un resto fosfato no modificado con otra unión internucleosídica como se describe en el presente

documento. Los ejemplos de grupos fosfato modificados incluyen, pero no se limitan a, fosforotioato, fosforoselenatos, boranofosfatos, ésteres de boranofosfato, hidrogenofosfonatos, fosforamidatos, fosforodiamidatos, fosfonatos de alquil o arilo y fosfotriésteres. Los fosforoditioatos tienen ambos oxígenos no de unión reemplazados por azufre. El enlazador de fosfato también se puede modificar mediante el reemplazo de un oxígeno de unión con nitrógeno (fosforamidatos unidos por puente), azufre (fosforotioatos unidos por puente) y carbono (fosfonatos de metileno unidos por puente).

El resto fosfato α -tio sustituido se proporciona para conferir estabilidad a los polímeros de ARN y ADN a través de la uniones de la cadena principal de fosforotioato no natural. El ADN y el ARN de fosforotioato tienen una mayor resistencia a las nucleasas y, posteriormente, una semivida más prolongada en un entorno celular.

En realizaciones específicas, un nucleósido modificado incluye un alfa-tio-nucleósido (por ejemplo, 5'-O-(1-tiofosfato)-adenosina, 5'-O-(1-tiofosfato)-citidina (α -tio-citidina), 5'-O-(1-tiofosfato)-guanosina, 5'-O-(1-tiofosfato)-uridina o 5'-O-(1-tiofosfato)-pseudouridina).

Se describen en el presente documento otras uniones internucleosídicas que pueden emplearse de acuerdo con la presente invención, incluyendo las uniones internucleosídicas que no contienen un átomo fosforoso.

En algunas realizaciones, la secuencia puede incluir uno o más nucleósidos citotóxicos. Por ejemplo, los nucleósidos citotóxicos pueden incorporarse en la secuencia, tal como una modificación bifuncional. Los nucleósidos citotóxicos pueden incluir, pero no se limita a, arabinósido de adenosina, 5-azacitidina, 4'-tio-aracitidina, ciclopentenilcitosina, cladribina, clofarabina, citarabina, arabinósido de citosina, 1-(2-C-ciano-2-desoxi-beta -D-arabino-pentofuranosil)-citosina, decitabina, 5-fluorouracilo, fludarabina, floxuridina, gemcitabina, una combinación de tegafur y uracilo, tegafur ((RS)-5-fluoro-1-(tetrahydrofuran-2-il)pirimidina-2,4(1H,3H)-diona), troxacitabina, tezacitabina, 2'-desoxi-2'-metilidencitidina (DMDC) y 6-mercaptopurina. Los ejemplos adicionales incluyen fosfato de fludarabina, N4-behenoil-1-beta-D-arabinofuranosilcitosina, N4-octadecil-1-beta-D-arabinofuranosilcitosina, N4-palmitoil-1-(2-C-ciano-2-desoxi-beta-D-arabino- pentofuranosil) citosina y P-4055 (éster de ácido 5'-elaídico de citarabina).

En algunas realizaciones, la secuencia incluye una o más modificaciones postranscripcionales (por ejemplo, protección con caperuza, escisión, poliadenilación, corte y empalme, secuencia poli-A, metilación, acilación, fosforilación, metilación de residuos de lisina y arginina, acetilación, y nitrosilación de grupos tiol y residuos de tirosina, etc.). La una o más modificaciones postranscripcionales pueden ser cualquier modificación postranscripcional, tal como cualquiera de las más de cien modificaciones de nucleósidos diferentes que se han identificado en el ARN (Rozenski, J, Crain, P, y McCloskey, J. (1999). The RNA Modification Database: 1999 update. Nucl Acids Res 27: 196-197) En algunas realizaciones, el primer ácido nucleico aislado comprende ARN mensajero (ARNm). En algunas realizaciones, el ARNm comprende al menos un nucleósido seleccionado del grupo que consiste en ribonucleósido de piridin-4-ona, 5-aza-uridina, 2-tio-5-aza-uridina, 2-tiouridina, 4-tio-pseudouridina, 2-tio-pseudouridina, 5-hidroxiuridina, 3-metiluridina, 5-carboximetil-uridina, 1-carboximetil-pseudouridina, 5-propinil-uridina, 1-propinil-pseudouridina, 5-taurinometiluridina, 1-taurinometil-pseudouridina, 5-taurinometil-2-tio-uridina, 1-

taurinometil-4-tiouridina, 5-metil-uridina, 1-metil-pseudouridina, 4-tio-1-metil-pseudouridina, 2-tio-1-metil-pseudouridina, 1-metil-1-desaza-pseudouridina, 2-tio-1-metil-1-desaza-pseudouridina, dihidrouridina, dihidropseudouridina, 2-tio-dihidrouridina, 2-tio-dihidropseudouridina, 2-metoxiuridina, 2-metoxi-4-tio-uridina, 4-metoxi-pseudouridina y 4-metoxi-2-tio-pseudouridina. En algunas realizaciones, el ARNm

5 comprende al menos un nucleósido seleccionado del grupo que consiste en 5-aza-citidina, pseudoisocitidina, 3-metil-citidina, N4-acetilcitidina, 5-formilcitidina, N4-metilcitidina, 5-hidroximetilcitidina, 1-metil-pseudoisocitidina, pirrolo-citidina, pirrolo-pseudoisocitidina, 2-tio-citidina, 2-tio-5-metil-citidina, 4-tio-pseudoisocitidina, 4-tio-1-metil-pseudoisocitidina, 4-tio-1-metil-1-desaza-pseudoisocitidina, 1-metil-1-desaza-pseudoisocitidina, zebularina, 5-aza-zebularina, 5-metil-zebularina, 5-aza-2-tio-zebularina, 2-tio-

10 zebularina, 2-metoxi-citidina, 2-metoxi-5-metil-citidina, 4-metoxi-pseudoisocitidina y 4-metoxi-1-metil-pseudoisocitidina. En algunas realizaciones, el ARNm comprende al menos un nucleósido seleccionado del grupo que consiste en 2-aminopurina, 2, 6-diaminopurina, 7-desaza-adenina, 7-desaza-8-aza-adenina, 7-desaza-2-aminopurina, 7-desaza-8-aza-2-aminopurina, 7-desaza-2,6-diaminopurina, 7-desaza-8-aza-2,6-diaminopurina, 1-metiladenosina, N6-metiladenosina, N6-isopenteniladenosina, N6-(cis-

15 hidroxisopentenil)adenosina, 2-metiltio-N6-(cis-hidroxisopentenil)adenosina, N6-glicinilcarbamoiladenosina, N6-treonilcarbamoiladenosina, 2-metiltio-N6-treonil carbamoiladenosina, N6,N6-dimetiladenosina, 7-metiladenina, 2-metiltio-adenina y 2-metoxi-adenina. En algunas realizaciones, el ARNm comprende al menos un nucleósido seleccionado del grupo que consiste en inosina, 1-metil-inosina, wiosina, wibutosina, 7-desaza-guanosina, 7-desaza-8-aza-guanosina, 6-tio-guanosina, 6-tio-7-

20 desaza-guanosina, 6-tio-7-desaza-8-aza-guanosina, 7-metil-guanosina, 6-tio-7-metil-guanosina, 7-metilinosina, 6-metoxi-guanosina, 1-metilguanosina, N2-metilguanosina, N2, N2-dimetilguanosina, 8-oxo-guanosina, 7-metil-8-oxo-guanosina, 1-metil-6-tio-guanosina, N2-metil-6-tio-guanosina y N2, N2-dimetil-6-tio-guanosina.

La secuencia puede no modificarse uniformemente a lo largo de toda la longitud de la molécula.

25 Por ejemplo, uno o más o todos los tipos de nucleótido (por ejemplo, nucleótidos de origen natural, purina o pirimidina, o uno cualquiera o más o todos de A, G, U, C, I, pU) pueden modificarse uniformemente o no en la secuencia, o en una región de la secuencia predeterminada dada de la misma. En algunas realizaciones, la secuencia incluye una pseudouridina. En algunas realizaciones, la secuencia incluye una inosina, que puede ayudar al sistema inmunitario a caracterizar la secuencia como ARN endógeno frente

30 a vírico. La incorporación de inosina también puede mediar una estabilidad potenciada del ARN/degradación reducida. Véase, por ejemplo, Yu, Z. *et al.* (2015) RNA editing by ADAR1 marks dsRNA as "self". Cell Res. 25, 1283-1284, que se incorpora por referencia en su totalidad.

Ácido Nucleico Diana

Los métodos desvelados en el presente documento son aplicables para una diversidad de ácidos

35 nucleicos diana. En algunas realizaciones, el ácido nucleico diana es un ADN, tal como un locus de ADN. En algunas realizaciones, el ácido nucleico diana es un ARN, tal como un locus de ARN o ARNm. En algunas realizaciones, el ácido nucleico diana es monocatenario (por ejemplo, ADN monocatenario). En algunas realizaciones, el ácido nucleico diana es bicatenario (por ejemplo, ADN bicatenario). En algunas

realizaciones, el ácido nucleico diana comprende regiones tanto monocatenarias como bicatenarias. En algunas realizaciones, el ácido nucleico diana es lineal. En algunas realizaciones, el ácido nucleico diana es circular. En algunas realizaciones, el ácido nucleico diana comprende uno o más nucleótidos modificados, tales como nucleótidos metilados, nucleótidos dañados o análogos de nucleótidos. En algunas realizaciones, no se modifica el ácido nucleico diana.

El ácido nucleico diana puede tener cualquier longitud, tal como aproximadamente al menos una cualquiera de 100 pb, 200 pb, 500 pb, 1000 pb, 2000 pb, 5000 pb, 10 kb, 20 kb, 50 kb, 100 kb, 200 kb, 500 kb, 1 Mb o más. El ácido nucleico diana también puede comprender cualquier secuencia. En algunas realizaciones, el ácido nucleico diana es rico en GC, tal como que tiene al menos aproximadamente uno cualquiera del 40 %, el 45 %, el 50 %, el 55 %, el 60 %, el 65 % o mayor contenido en GC. En algunas realizaciones, el ácido nucleico diana tiene un contenido en GC de al menos aproximadamente el 70 %, el 80 % o más. En algunas realizaciones, el ácido nucleico diana es un fragmento rico en GC en un ácido nucleico diana no rico en GC. En algunas realizaciones, el ácido nucleico diana no es rico en GC. En algunas realizaciones, el ácido nucleico diana tiene una o más estructuras secundarias o estructuras de orden superior. En algunas realizaciones, el ácido nucleico diana no está en un estado condensado, tal como en una cromatina, para hacer que el ácido nucleico diana sea inaccesible por el complejo de polipéptido Cas12i4/guía de ARN.

En algunas realizaciones, el ácido nucleico diana está presente en una célula. En algunas realizaciones, el ácido nucleico diana está presente en el núcleo de la célula. En algunas realizaciones, el ácido nucleico diana es endógeno a la célula. En algunas realizaciones, el ácido nucleico diana es un ADN genómico. En algunas realizaciones, el ácido nucleico diana es un ADN cromosómico. En algunas realizaciones, el ácido nucleico diana es un gen codificante de la proteína o una región funcional del mismo, tal como una región codificante, o un elemento regulador, tal como un promotor, potenciador, una región no traducida en 5' o 3', etc. En algunas realizaciones, el ácido nucleico diana es un gen no codificante, tal como transposón, miARN, ARNt, ARN ribosómico, ribozima o ARNncl. En algunas realizaciones, el ácido nucleico diana es un plásmido.

En algunas realizaciones, el ácido nucleico diana es exógeno a una célula. En algunas realizaciones, el ácido nucleico diana es un ácido nucleico vírico, tal como ADN vírico o ARN vírico. En algunas realizaciones, el ácido nucleico diana es un plásmido transferido horizontalmente. En algunas realizaciones, el ácido nucleico diana se integra en el genoma de la célula. En algunas realizaciones, el ácido nucleico diana no se integra en el genoma de la célula. En algunas realizaciones, el ácido nucleico diana es un plásmido en la célula. En algunas realizaciones, el ácido nucleico diana está presente en una matriz extracromosómica.

En algunas realizaciones, el ácido nucleico diana es un ácido nucleico aislado, tal como un ADN aislado o un ARN aislado. En algunas realizaciones, el ácido nucleico diana está presente en un entorno sin células. En algunas realizaciones, el ácido nucleico diana es un vector aislado, tal como un plásmido. En algunas realizaciones, el ácido nucleico diana es un plásmido ultrapuro.

El ácido nucleico diana es un segmento del ácido nucleico diana que se hibrida con la guía de ARN. En algunas realizaciones, el ácido nucleico diana tiene solamente una copia del ácido nucleico diana. En algunas realizaciones, el ácido nucleico diana tiene más de una copia, tal como al menos aproximadamente una cualquiera de 2, 3, 4, 5, 10, 100 o más copias del ácido nucleico diana. Por ejemplo, un ácido nucleico
 5 diana que comprende una secuencia repetida en un genoma de un ácido nucleico vírico o una bacteria puede ser la diana de la nucleoproteína.

La secuencia diana es adyacente a un motivo adyacente al protoespaciador o PAM de la divulgación como se describe en el presente documento. El PAM puede estar inmediatamente adyacente a la secuencia diana o, por ejemplo, dentro de un pequeño número (por ejemplo, 1, 2, 3, 4 o 5) de
 10 nucleótidos de la secuencia diana. En el caso de una diana bicatenaria, el resto de direccionamiento (por ejemplo, una guía de ARN) se une a una primera cadena de la diana, y una secuencia de PAM como se describe en el presente documento está presente en la segunda cadena complementaria. En un caso de este tipo, la secuencia de PAM es inmediatamente adyacente a (o dentro de un número pequeño, por ejemplo, 1, 2, 3, 4 o 5 nucleótidos de) una secuencia en la segunda cadena que es complementaria a la
 15 secuencia en la primera cadena a la que se une el resto de unión. En algunas realizaciones, la especificidad de secuencia requiere una coincidencia completa de la secuencia espaciadora en la guía de ARN con la cadena no PAM de un ácido nucleico diana. En otras realizaciones, la especificidad de secuencia requiere una coincidencia parcial (contigua o no contigua) de la secuencia espaciadora en la guía de ARN con la cadena no PAM de un ácido nucleico diana.

En algunas realizaciones, la guía de ARN o un complejo que comprende la guía de ARN y un polipéptido Cas12i4 descrito en el presente documento se une a un ácido nucleico diana en una secuencia definida por la región de complementariedad entre la guía de ARN y el ácido nucleico diana. En algunas realizaciones, la secuencia de PAM descrita en el presente documento se ubica directamente en dirección
 20 5' de la secuencia diana del ácido nucleico diana (por ejemplo, directamente 5' de la secuencia diana). En algunas realizaciones, la secuencia de PAM descrita en el presente documento se ubica directamente en 5' de la secuencia diana en la cadena complementaria no espaciadora (por ejemplo, cadena no diana) del ácido nucleico diana.

En algunas realizaciones, las secuencias de PAM correspondientes a Cas12i4 (por ejemplo, el polipéptido Cas12i4 precursor o polipéptidos Cas12i4 variantes) incluyen 5'-TTN-3' y 5'-NTTN-3', en donde
 30 N es cualquier nucleótido (por ejemplo, A, G, T o C). En algunas realizaciones, la secuencia de PAM comprende 5'-TTH-3', 5'-TTY-3', 5'-TTC-3', 5'-NTTH-3', 5'-NTTY-3' o 5'-NTTC-3', en donde N es cualquier nucleótido, H es A, C o T e Y es C o T. En algunas realizaciones, la secuencia de PAM comprende 5'-TTA-3', 5'-TTC-3', 5'-TTG-3', 5'-TTT-3', 5'-NTTA-3', 5'-NTTC-3', 5'-NTTG-3' o 5'-NTTT-3'. Por ejemplo, en algunas realizaciones, el PAM comprende 5'-ATTA-3', 5'-CTTA-3', 5'-TTTC-3', 5'-TTA-3', 5'-GTTA-3', 5'-
 35 CTTC-3', 5'-CTTG-3', 5'-TTC-3', 5'-TTTA-3', 5'-GTTC-3', 5'-GTTG-3', 5'-TTG-3', 5'-ATTC-3', 5'-TTTT-3', 5'-GTTT-3', 5'-ATTT-3', 5'-TTT-3', 5'-CTTT-3', 5'-TTTG-3' o 5'-CTT-3'.

En algunas realizaciones, las secuencias de PAM correspondientes a Cas12i4 (por ejemplo, el polipéptido Cas12i4 precursor o polipéptidos Cas12i4 variantes) incluyen 5'-NTN-3', 5'-NNTN-3', 5'-VTN-3'

y 5'-NVTN-3', en donde N es cualquier nucleótido (por ejemplo, A, G, T o C) y V es A, G o C. En algunas realizaciones, la secuencia de PAM comprende 5'-NTC-3', 5'-NTA-3', 5'-NTG-3', 5'-NTT-3', 5'-NNTC-3', 5'-NNTA-3', 5'-NNTG-3' o 5'-NNTT-3'. Por ejemplo, en algunas realizaciones, la secuencia de PAM comprende 5'-AATC-3', 5'-CCTG-3', 5'-CTA-3', 5'-TCTC-3', 5'-CTG-3', 5'-GCTG-3', 5'-CTC-3', 5'-GCTC-3', 5'-TCTG-3', 5'-ACTG-3', 5'-GATA-3', 5'-TATC-3', 5'-ATC-3', 5'-ATA-3', 5'-GATC-3', 5'-ACTA-3', 5'-GATG-3', 5'-TGTG-3', 5'-TCTT-3', 5'-CCTT-3', GCTT-3' o 5'-ACTT-3'.

En algunas realizaciones, un polipéptido Cas12i4 (por ejemplo, el polipéptido Cas12i4 precursor o un polipéptido Cas12i4 variante) reconoce una secuencia de PAM expuesta como 5'-ATAA-3', 5'-CAAT-3', 5'-CGAT-3', 5'-GAGA-3', 5'-CAAG-3', 5'-ACGT-3', 5'-GGCC-3', 5'-GGAC-3', 5'-GGCA-3', 5'-GTAC-3', 5'-GACC-3' o 5'-TTAC-3'.

En algunas realizaciones, el ácido nucleico diana está presente en una región de fácil acceso del ácido nucleico diana. En algunas realizaciones, el ácido nucleico diana está en un exón de un gen diana. En algunas realizaciones, el ácido nucleico diana está a través de una unión exón-intrón de un gen diana. En algunas realizaciones, el ácido nucleico diana está presente en una región no codificante, tal como una región reguladora de un gen. En algunas realizaciones, en donde el ácido nucleico diana es exógeno a una célula, el ácido nucleico diana comprende una secuencia que no se encuentra en el genoma de la célula.

Las condiciones de unión de ADN/ARN adecuadas incluyen condiciones fisiológicas normalmente presentes en una célula. Se conocen en la técnica otras condiciones de unión de ADN/ARN adecuadas (por ejemplo, condiciones en un sistema sin células); véase, por ejemplo, Sambrook, citado anteriormente. La cadena del ácido nucleico diana que es complementaria a y se hibrida con la guía de ARN se denomina la "cadena complementaria" y la cadena del ácido nucleico diana que es complementaria a la "cadena complementaria" (y, por lo tanto, no es complementaria a la guía de ARN) se denomina la "cadena no complementaria" o "cadena no-complementaria".

PREPARACIÓN

En algunas realizaciones, el polipéptido Cas12i4 variante de la presente invención puede prepararse (a) cultivando bacterias que producen el polipéptido Cas12i4 variante de la presente invención, aislando el polipéptido Cas12i4 variante, opcionalmente, purificando el polipéptido Cas12i4 variante, y formando complejos del polipéptido Cas12i4 variante con una guía de ARN. El polipéptido Cas12i4 variante también puede prepararse mediante (b) una técnica de ingeniería genética conocida, específicamente, aislando un gen que codifica el polipéptido Cas12i4 variante de la presente invención de bacterias, construyendo un vector de expresión recombinante, y después transfiriendo el vector en una célula hospedadora apropiada que expresa la guía de ARN para la expresión de una proteína recombinante que forma complejo con la guía de ARN en la célula hospedadora. Como alternativa, el polipéptido Cas12i4 variante puede prepararse mediante (c) un sistema de transcripción-traducción acoplado *in vitro* y después forma complejo con la guía de ARN. Las bacterias que pueden usarse para la preparación del polipéptido Cas12i4 variante de la presente invención no se limitan particularmente siempre que puedan producir el polipéptido Cas12i4 variante de la presente invención. Algunos ejemplos no limitantes de las bacterias incluyen células de *E. coli* descritas en el presente documento.

Vectores

La presente invención proporciona un vector para expresar el polipéptido Cas12i4 variante descrito en el presente documento o pueden incorporarse en un vector ácidos nucleicos que codifican la variante descrita en el presente documento. En algunas realizaciones, un vector de la invención incluye una
 5 secuencia de nucleótidos que codifica el polipéptido Cas12i4 variante. En algunas realizaciones, un vector de la invención incluye una secuencia de nucleótidos que codifica el polipéptido Cas12i4 variante.

La presente invención también proporciona un vector que puede usarse para la preparación del polipéptido Cas12i4 variante o composiciones que comprenden el polipéptido Cas12i4 variante como se describe en el presente documento. En algunas realizaciones, la invención incluye la composición o vector
 10 descrito en el presente documento en una célula. En algunas realizaciones, la invención incluye un método para expresar la composición que comprende el polipéptido Cas12i4 variante, o vector o ácido nucleico que codifica el polipéptido Cas12i4 variante, en una célula. El método puede comprender las etapas de proporcionar la composición, por ejemplo, vector o ácido nucleico, y administrar la composición a la célula.

La expresión de polinucleótidos naturales o sintéticos se logra normalmente uniendo
 15 operativamente un polinucleótido que codifica el gen de interés, por ejemplo, una secuencia de nucleótidos que codifica el polipéptido Cas12i4 variante, a un promotor e incorporando la construcción en un vector de expresión. El vector de expresión no se limita particularmente siempre que incluya un polinucleótido que codifique el polipéptido Cas12i4 variante de la presente invención y pueda ser adecuado para la replicación e integración en células eucariotas.

Los vectores de expresión típicos incluyen terminadores de transcripción y traducción, secuencias de inicio y promotores útiles para la expresión del polinucleótido deseado. Por ejemplo, pueden usarse vectores plasmídicos que portan una secuencia de reconocimiento para la ARN polimerasa (pSP64, pBluescript, etc.). Los vectores que incluyen los derivados de retrovirus tales como los lentivirus son
 20 herramientas adecuadas para lograr la transferencia génica a largo plazo, ya que permiten la integración estable a largo plazo de un transgén y su propagación en células hijas. Los ejemplos de vectores incluyen vectores de expresión, vectores de replicación, vectores de generación de sonda y vectores de secuenciación. El vector de expresión puede proporcionarse a una célula en la forma de un vector vírico.

La tecnología de vectores víricos se conoce bien en la técnica descrita en una diversidad de manuales de virología y biología molecular. Los virus que son útiles como vectores incluyen, pero no se
 30 limitan a, virus de fagos, retrovirus, adenovirus, virus adenoasociados, virus del herpes y lentivirus. En general, un vector adecuado contiene un origen de replicación funcional en al menos un organismo, una secuencia promotora, sitios de endonucleasa de restricción convenientes y uno o más marcadores seleccionables.

El tipo de vector no se limita particularmente, y puede seleccionarse de manera apropiada un vector
 35 que se puede expresar en células hospedadoras. Para ser más específicos, dependiendo del tipo de célula hospedadora, se selecciona apropiadamente una secuencia promotora para garantizar la expresión del polipéptido Cas12i4 variante del polinucleótido, y esta secuencia promotora y el polinucleótido se insertan en cualquiera de los diversos plásmidos, etc. para la preparación del vector de expresión.

Los elementos promotores adicionales, por ejemplo, secuencias potenciadoras, regulan la frecuencia de inicio de la transcripción. Normalmente, se ubican en la región 30-110 pb en dirección 5' del sitio de inicio, aunque recientemente se ha demostrado que varios promotores también contienen elementos funcionales en dirección 3' del sitio de inicio. Dependiendo del promotor, parece que los elementos individuales pueden actuar de forma cooperativa o independientemente para activar la transcripción.

Además, la divulgación no debe limitarse al uso de promotores constitutivos. Los promotores inducibles también se contemplan como parte de la divulgación. El uso de un promotor inducible proporciona un interruptor molecular que puede activar la expresión de la secuencia polinucleotídica a la que está unida operativamente cuando se desea dicha expresión o desactivar la expresión cuando no se desea la expresión. Los ejemplos de promotores inducibles incluyen, pero no se limitan a, un promotor de metalotionina, un promotor de glucocorticoides, un promotor de progesterona y un promotor de tetraciclina.

El vector de expresión que ha de introducirse también puede contener un gen marcador seleccionable o un gen indicador o ambos para facilitar la identificación y selección de células de expresión de la población de células que se busca transfectar o infectar a través de vectores víricos. En otros aspectos, el marcador seleccionable puede portarse en un trozo independiente de ADN y usarse en un procedimiento de cotransfección. Tanto los genes indicadores como los marcadores seleccionables pueden estar flanqueados por secuencias de control transcripcionales apropiadas para permitir la expresión en las células hospedadoras. Los ejemplos de un marcador de este tipo incluyen un gen de dihidrofolato reductasa y un gen de resistencia a neomicina para cultivo de células eucariotas; y un gen de resistencia a la tetraciclina y un gen de resistencia a la ampicilina para el cultivo de *E. coli* y otras bacterias. Mediante el uso de un marcador de selección de este tipo, puede confirmarse si el polinucleótido que codifica el polipéptido Cas12i4 variante de la presente invención se ha transferido a las células hospedadoras y después se ha expresado sin fallos.

El método de preparación de vectores de expresión recombinantes no se limita particularmente y los ejemplos del mismo incluyen métodos que usan un plásmido, un fago o un cósmido.

Métodos de expresión

La presente invención incluye un método para la expresión de proteínas, que comprende traducir el polipéptido Cas12i4 variante descrito en el presente documento.

En algunas realizaciones, una célula hospedadora descrita en el presente documento se usa para expresar el polipéptido Cas12i4 variante. La célula hospedadora no se limita particularmente, y pueden usarse preferentemente diversas células conocidas. Los ejemplos específicos de la célula hospedadora incluyen bacterias tales como de *E. coli*, levaduras (levadura de gemación, *Saccharomyces cerevisiae*, y levadura de fisión, *Schizosaccharomyces pombe*), nematodos (*Caenorhabditis elegans*), ovocitos de *Xenopus laevis*, y células animales (por ejemplo, células CHO, células COS y células HEK293). El método para transferir el vector de expresión descrito anteriormente a las células hospedadoras, es decir, el método de transformación, no se limita particularmente, y pueden usarse métodos conocidos tales como la electroporación, el método del fosfato de calcio, el método del liposoma y el método de dextrano DEAE.

Después de que un hospedador se transforma con el vector de expresión, las células hospedadoras pueden cultivarse o reproducirse para la producción del polipéptido Cas12i4 variante. Después de la expresión del polipéptido Cas12i4 variante, las células hospedadoras pueden recogerse y el polipéptido Cas12i4 variante puede purificarse de los cultivos etc. de acuerdo con métodos convencionales (por ejemplo, filtración, centrifugación, disrupción celular, cromatografía de filtración en gel, cromatografía de intercambio iónico, etc.).

En algunas realizaciones, los métodos para la expresión del polipéptido Cas12i4 variante comprenden la traducción de al menos 5 aminoácidos, al menos 10 aminoácidos, al menos 15 aminoácidos, al menos 20 aminoácidos, al menos 50 aminoácidos, al menos 100 aminoácidos, al menos 150 aminoácidos, al menos 200 aminoácidos, al menos 250 aminoácidos, al menos 300 aminoácidos, al menos 400 aminoácidos, al menos 500 aminoácidos, al menos 600 aminoácidos, al menos 700 aminoácidos, al menos 800 aminoácidos, al menos 900 aminoácidos o al menos 1000 aminoácidos del polipéptido Cas12i4 variante. En algunas realizaciones, los métodos para la expresión de la proteína comprenden la traducción de aproximadamente 5 aminoácidos, aproximadamente 10 aminoácidos, aproximadamente 15 aminoácidos, aproximadamente 20 aminoácidos, aproximadamente 50 aminoácidos, aproximadamente 100 aminoácidos, aproximadamente 150 aminoácidos, aproximadamente 200 aminoácidos, aproximadamente 250 aminoácidos, aproximadamente 300 aminoácidos, aproximadamente 400 aminoácidos, aproximadamente 500 aminoácidos, aproximadamente 600 aminoácidos, aproximadamente 700 aminoácidos, aproximadamente 800 aminoácidos, aproximadamente 900 aminoácidos, aproximadamente 1000 aminoácidos o más del polipéptido Cas12i4 variante.

Puede usarse una diversidad de métodos para determinar el nivel de producción de un polipéptido Cas12i4 variante maduro en una célula hospedadora. Dichos métodos incluyen, pero no se limitan a, por ejemplo, métodos que utilizan anticuerpos o bien policlonales o monoclonales específicos para el polipéptido Cas12i4 variante o un marcador de marcaje como se describe en otra parte en el presente documento. Los métodos a modo de ejemplo incluyen, pero no se limitan a, ensayos de inmunoabsorción ligados a enzimas (ELISA), radioinmunoensayos (RIA), inmunoensayos fluorescentes (FIA), y clasificación de células activadas por fluorescencia (FACS). Estos y otros ensayos son bien conocidos en la técnica (véase, por ejemplo, Maddox *et al.*, J. Exp. Med. 158:1211 [1983]).

La presente divulgación proporciona métodos de expresión *in vivo* del polipéptido Cas12i4 variante en una célula, que comprenden proporcionar un polirribonucleótido que codifica el polipéptido Cas12i4 variante a una célula hospedadora en donde el polirribonucleótido codifica el polipéptido Cas12i4 variante, expresar el polipéptido Cas12i4 variante en la célula y obtener el polipéptido Cas12i4 variante de la célula.

Introducción de alteración o mutación

Las secuencias de ácido nucleico que codifican polipéptidos variantes o polipéptidos variantes pueden generarse mediante métodos de síntesis conocidos en la técnica. Usando la secuencia de ácido nucleico que codifica el propio polipéptido precursor como marco, pueden insertarse alternancias o mutaciones una o más a la vez para alterar la secuencia de ácido nucleico que codifica el polipéptido precursor. Del mismo modo, el polipéptido precursor puede alterarse o mutarse introduciendo los cambios

en la secuencia del polipéptido a medida que se sintetiza sintéticamente. Esto puede lograrse mediante métodos bien conocidos en la técnica.

La producción e introducción de una alteración o mutación en una secuencia polipeptídica precursora puede lograrse usando cualquier método conocido por los expertos en la materia. En particular, en algunas realizaciones, pueden usarse cebadores de oligonucleótidos para PCR para la síntesis rápida de un molde de ADN que incluye una o más alteraciones o mutaciones en la secuencia de ácido nucleico que codifica el polipéptido variante. También puede usarse mutagénesis específica de sitio como una técnica útil en la preparación de péptidos individuales, o proteínas o péptidos equivalentes biológicamente funcionales, mediante mutagénesis específica del ADN subyacente. La técnica proporciona además una capacidad inmediata para preparar y someter a ensayo variantes, incorporando una o más de las consideraciones anteriores, mediante la introducción de uno o más cambios en la secuencia de nucleótidos en el ADN. La mutagénesis específica de sitio permite la producción de variantes mediante el uso de secuencias de oligonucleótidos específicas que codifican la secuencia de ADN de la mutación deseada, así como un número suficiente de nucleótidos adyacentes, para proporcionar una secuencia de cebador de tamaño y complejidad de secuencia suficientes para formar un dúplex estable a ambos lados de la unión de supresión que se está atravesando. Normalmente, se prefiere un cebador de aproximadamente 17 a 25 nucleótidos de longitud, con aproximadamente 5 a 10 residuos en ambos lados de la unión de la secuencia que se está alterando.

La introducción de variaciones estructurales, tales como la fusión de polipéptidos como extensiones amino y/o carboxilo terminales, puede lograrse de manera similar a la introducción de alteraciones o mutaciones en el polipéptido precursor. Los péptidos adicionales pueden añadirse al polipéptido precursor o al polipéptido variante incluyendo la secuencia de ácido nucleico apropiada que codifica los péptidos adicionales a la secuencia de ácido nucleico que codifica el polipéptido precursor o el polipéptido variante. Opcionalmente, los péptidos adicionales pueden añadirse directamente al polipéptido variante a través de la producción de polipéptidos sintéticos.

En un aspecto, la invención también proporciona métodos para introducir una alteración o mutación en la secuencia polipeptídica precursora para producir un polipéptido Cas12i4 variante que tiene una unión en la diana aumentada con dos o más loci (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o más) de un ácido nucleico diana, en comparación con un polipéptido precursor.

En un aspecto, la invención también proporciona métodos para introducir una alteración o mutación en la secuencia polipeptídica precursora para producir una pluralidad de polipéptidos Cas12i4 variantes (por ejemplo, polipéptidos Cas12i4 variantes independientes que tienen la misma secuencia de aminoácidos), que cuando forman complejos individualmente con una pluralidad de guías de ARN distintas, tienen una unión en la diana aumentada con dos o más loci de un ácido nucleico diana, en comparación con una pluralidad de polipéptidos precursores y guías de ARN.

En un aspecto, la invención también proporciona métodos para introducir una alteración o mutación en la secuencia polipeptídica precursora para producir un polipéptido Cas12i4 variante que tiene una

formación de complejos ternarios aumentada en la diana con dos o más loci diana de un ácido nucleico diana, en comparación con un polipéptido precursor.

En un aspecto, la invención también proporciona métodos para introducir una alteración o mutación en la secuencia polipeptídica precursora para producir una pluralidad de polipéptidos Cas12i4 variantes (por ejemplo, polipéptidos Cas12i4 variantes independientes que tienen la misma secuencia de aminoácidos), que cuando forman complejos individualmente con una pluralidad de guías de ARN distintas, tienen una formación de complejos ternarios aumentada con dos o más loci de un ácido nucleico diana, en comparación con una pluralidad de polipéptidos precursores y guías de ARN.

En un aspecto, la invención también proporciona métodos para introducir una alteración o mutación en la secuencia polipeptídica precursora para producir polipéptidos Cas12i4 que presentan el direccionamiento de un mayor número de ácidos nucleicos diana o loci diana, en comparación con un polipéptido precursor.

En un aspecto, la invención también proporciona métodos para introducir una alteración o mutación en la secuencia polipeptídica precursora para producir una pluralidad de polipéptidos Cas12i4 variantes (por ejemplo, polipéptidos Cas12i4 variantes independientes que tienen la misma secuencia de aminoácidos), que cuando forman complejos individualmente con una pluralidad de guías de ARN distintas, presentan el direccionamiento de un mayor número de ácidos nucleicos diana o loci diana, en comparación con una pluralidad de polipéptidos precursores y guías de ARN.

En un aspecto, la invención también proporciona métodos para introducir una alteración o mutación en la secuencia polipeptídica precursora para potenciar la estabilidad del polipéptido Cas12i4. La estabilidad del polipéptido Cas12i4 puede determinarse por o puede incluir una técnica no limitada a ensayos de desnaturalización térmica, ensayos de desplazamiento térmico, calorimetría de barrido diferencial (DSC), fluorimetría de barrido diferencial (DSF), calorimetría de titulación isotérmica (ITC), métodos de pulso y caza, métodos de persecución de blanqueamiento, métodos para la persecución de cicloheximida, espectroscopía de dicroísmo circular (CD), ensayos de actividad basados en fluorescencia y cristalización.

Complejo binario variante

En general, el polipéptido Cas12i4 variante y la guía de ARN se unen entre sí en una relación molar de aproximadamente 1:1 para formar el complejo binario variante. El polipéptido Cas12i4 variante y la guía de ARN, o bien solos o bien juntos, no son de origen natural.

En algunas realizaciones, el polipéptido Cas12i4 variante puede sobreexpresarse en una célula hospedadora y purificarse como se describe en el presente documento, después formar complejo con la guía de ARN (por ejemplo, en un tubo de ensayo) para formar una ribonucleoproteína (RNP) variante (por ejemplo, complejo binario variante).

En algunas realizaciones, el complejo binario variante presenta una afinidad de unión aumentada a un ácido nucleico diana, una actividad de unión aumentada en la diana, una especificidad de unión aumentada en la diana, una formación de complejos ternarios aumentada con un ácido nucleico diana y/o una estabilidad aumentada en un intervalo de tiempos de incubación. En algunas realizaciones, el

complejo binario variante presenta una unión fuera de la diana disminuida a un ácido nucleico no diana y/o una disociación disminuida de un ácido nucleico diana en un intervalo de tiempos de incubación. En algunas realizaciones, el complejo binario variante presenta una formación de complejos de ácido nucleico diana, una actividad del ácido nucleico diana y/o una especificidad del ácido nucleico diana aumentadas en un intervalo de tiempos de incubación.

En algunas realizaciones, la formación de complejo de un complejo binario se produce a una temperatura inferior a aproximadamente una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C, 44 °C, 45 °C, 50 °C o 55 °C. En algunas realizaciones, el polipéptido Cas12i4 variante no se disocia de la guía de ARN ni se une a un ARN libre a aproximadamente 37 °C a lo largo de un periodo de incubación de al menos aproximadamente uno cualquiera de 10 minutos, 15 minutos, 20 minutos, 25 minutos, 30 minutos, 35 minutos, 40 minutos, 45 minutos, 50 minutos, 55 minutos, 1 h, 2 h, 3 h, 4 h o más horas. En algunas realizaciones, después de la formación de complejos binarios, el complejo de ribonucleoproteína variante no intercambia la guía de ARN con un ARN diferente.

En algunas realizaciones, el polipéptido Cas12i4 variante y una guía de ARN forman complejo en un tampón de formación de complejo binario. En algunas realizaciones, el polipéptido Cas12i4 variante se almacena en un tampón que se reemplaza con un tampón de formación de complejo binario para formar un complejo con la guía de ARN. En algunas realizaciones, el polipéptido Cas12i4 variante se almacena en un tampón de formación de complejo binario.

En algunas realizaciones, el tampón de formación de complejo binario tiene un pH en un intervalo de aproximadamente 7.3 a 8.6. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 7.3. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 7.4. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 7.5. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 7.6. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 7.7. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 7.8. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 7.9. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 8.0. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 8.1. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 8.2. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 8.3. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 8.4. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 8.5. En una realización, el pH del tampón de formación de complejo binario es de aproximadamente 8.6.

La termoestabilidad del polipéptido Cas12i4 variante puede aumentar en condiciones favorables tales como la adición de una guía de ARN, por ejemplo, la unión de una guía de ARN.

En algunas realizaciones, el polipéptido Cas12i4 variante puede sobreexpresarse y formar complejo con la guía de ARN en una célula hospedadora antes de la purificación como se describe en el presente documento. En algunas realizaciones, el ARNm o el ADN que codifica el polipéptido Cas12i4 variante se introduce en una célula de modo que el polipéptido Cas12i4 variante se expresa en la célula.

5 La guía de ARN, que guía el polipéptido Cas12i4 variante hacia el ácido nucleico diana deseado, también se introduce en la célula, ya sea de forma simultánea, separada o secuencial a partir de una única construcción de ARNm o ADN, de modo que se forme el complejo de ribonucleoproteína necesario en la célula.

Evaluación de la estabilidad y la funcionalidad del complejo binario variante

10 En determinadas realizaciones, en el presente documento se proporcionan métodos para identificar un complejo de polipéptido Cas12i4 variante óptimo/guía de ARN (denominado en el presente documento el complejo binario variante) que incluyen (a) combinar un polipéptido Cas12i4 variante y una guía de ARN en una muestra para formar el complejo binario variante; (b) medir un valor del complejo binario variante; y (c) determinar si el complejo binario variante es óptimo con respecto a la molécula de referencia, si el
15 valor del complejo binario variante es superior a un valor de una molécula de referencia. En algunas realizaciones, el valor puede incluir, pero no se limita a, una medición de estabilidad (por ejemplo, valor de T_m , termoestabilidad), una tasa de formación de complejos binarios, especificidad de unión a una guía de ARN y/o actividad del complejo.

20 En algunas realizaciones, un complejo de polipéptido Cas12i4 variante óptimo/guía de ARN (es decir, un complejo binario variante) se identifica por las etapas de: (a) combinar un polipéptido Cas12i4 variante y una guía de ARN en una muestra para formar el complejo binario variante; (b) detectar un valor de T_m del complejo binario variante; y (c) determinar que el complejo binario variante es estable si el valor de T_m del complejo binario variante es superior a un valor de T_m de una molécula de referencia o un valor de referencia de T_m en al menos 8 °C.

25 Los métodos que implican una etapa de medir la termoestabilidad de un complejo polipéptido Cas12i4 variante/guía de ARN (es decir, un complejo binario variante) pueden incluir, sin limitación, métodos para determinar la estabilidad de un complejo binario variante, métodos para determinar una condición que promueve un complejo binario variante estable, métodos para seleccionar un complejo binario variante estable y métodos para identificar un ARNg óptimo para formar un complejo binario variante
30 estable. En determinadas realizaciones, puede medirse un valor de termoestabilidad de un complejo binario variante.

Adicionalmente, en determinadas realizaciones, también puede medirse un valor de termoestabilidad de una molécula de referencia. En determinadas realizaciones, puede determinarse que
35 un complejo binario variante es estable si el valor de termoestabilidad medido del complejo binario variante es superior al valor de termoestabilidad medido de la molécula de referencia o un valor de referencia de termoestabilidad, medido en las mismas condiciones experimentales, como se describe en el presente documento. En determinadas realizaciones, la molécula de referencia puede ser el polipéptido Cas12i4 variante sin una guía de ARN.

En determinadas realizaciones, el valor de termoestabilidad que se mide puede ser un valor de temperatura de desnaturalización. En estas realizaciones, el valor de referencia de termoestabilidad es un valor de referencia de temperatura de desnaturalización. En determinadas realizaciones, el valor de termoestabilidad que se mide puede ser un valor de T_m . En estas realizaciones, el valor de referencia de termoestabilidad puede ser un valor de referencia de T_m . En determinadas realizaciones, el valor de termoestabilidad puede medirse usando un ensayo de desplazamiento térmico. En determinadas realizaciones, un ensayo utilizado para medir la termoestabilidad puede implicar una técnica descrita en el presente documento incluyendo, pero no limitada a, ensayos de desnaturalización térmica, ensayos de desplazamiento térmico, calorimetría de barrido diferencial (DSC), fluorimetría de barrido diferencial (DSF), calorimetría de titulación isotérmica (ITC), métodos de persecución de pulso, métodos de persecución de blanqueamiento, métodos de persecución de cicloheximida, espectroscopía de dicroísmo circular (CD), ensayos de actividad basados en fluorescencia y cristalización.

En determinadas realizaciones, puede identificarse un complejo binario variante si la tasa de formación de complejos polipéptido Cas12i4 variante/guía de ARN, la especificidad de unión a una guía de ARN y/o la actividad de complejo del complejo binario variante es superior a un valor de la molécula de referencia o el valor de referencia (por ejemplo, un valor de un complejo polipéptido precursor/guía de ARN, denominado en el presente documento un complejo binario precursor). Por ejemplo, en determinadas realizaciones, puede identificarse el complejo binario variante si el valor de una tasa de formación de complejos polipéptido Cas12i4 variante/guía de ARN, una especificidad de unión a una guía de ARN y/o una actividad de complejo del complejo binario variante es al menos un X % superior a un valor de la molécula de referencia o el valor de referencia (por ejemplo, un valor de un complejo binario precursor). En determinadas realizaciones, los métodos descritos en el presente documento pueden comprender además etapas que incluyen medir la actividad del complejo binario variante como se describe en el presente documento.

Complejación ternaria variante

En algunas realizaciones, el polipéptido Cas12i4 variante, la guía de ARN y el ácido nucleico diana, como se describe en el presente documento, forman un complejo ternario variante (por ejemplo, en un tubo de ensayo o célula). En general, el polipéptido Cas12i4 variante, la guía de ARN, y el ácido nucleico diana se asocian entre sí en una relación molar de aproximadamente 1:1:1 para formar el complejo ternario variante. El polipéptido Cas12i4 variante, la guía de ARN, y el ácido nucleico diana, o bien solos o bien juntos, no son de origen natural.

En algunas realizaciones, el complejo binario variante (por ejemplo, complejo de polipéptido Cas12i4 variante y una guía de ARN) como se describe en el presente documento, forma complejo adicionalmente con el ácido nucleico diana (por ejemplo, en un tubo de ensayo o célula) para formar un complejo ternario variante.

En algunas realizaciones, la formación de complejo del complejo ternario se produce a una temperatura inferior a una cualquiera de 20 °C, 21 °C, 22 °C, 23 °C, 24 °C, 25 °C, 26 °C, 27 °C, 28 °C, 29 °C, 30 °C, 31 °C, 32 °C, 33 °C, 34 °C, 35 °C, 36 °C, 37 °C, 38 °C, 39 °C, 40 °C, 41 °C, 42 °C, 43 °C,

44 °C, 45 °C, 50 °C o 55 °C. En algunas realizaciones, el complejo binario variante no se disocia del ácido nucleico diana ni se une a un ácido nucleico libre (por ejemplo, ADN libre) a aproximadamente 37 °C a lo largo de un período de incubación de al menos aproximadamente uno cualquiera de 10 minutos, 15 minutos, 20 minutos, 25 minutos, 30 minutos, 35 minutos, 40 minutos, 45 minutos, 50 minutos, 55 minutos, 1 h, 2 h, 3 h, 4 h o más horas. En algunas realizaciones, tras la formación del complejo ternario, un complejo binario variante no intercambia el ácido nucleico diana con un ácido nucleico diferente.

En algunas realizaciones, el polipéptido Cas12i4 variante, una guía de ARN, y ácido nucleico diana forman complejo en un tampón de formación de complejo ternario. En algunas realizaciones, el polipéptido Cas12i4 variante se almacena en un tampón que se reemplaza con un tampón de formación de complejo ternario para formar un complejo con la guía de ARN y ácido nucleico diana. En algunas realizaciones, el polipéptido Cas12i4 variante se almacena en un tampón de formación de complejo ternario.

En algunas realizaciones, el complejo binario variante y ácido nucleico diana forman complejo en un tampón de formación de complejo ternario. En algunas realizaciones, el complejo binario variante se almacena en un tampón que se reemplaza con un tampón de formación de complejo ternario para formar un complejo con el ácido nucleico diana. En algunas realizaciones, el complejo binario variante se almacena en un tampón de formación de complejo ternario.

En algunas realizaciones, el tampón de formación de complejo ternario tiene un pH en un intervalo de aproximadamente 7.3 a 8.6. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 7.3. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 7.4. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 7.5. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 7.6. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 7.7. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 7.8. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 7.9. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 8.0. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 8.1. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 8.2. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 8.3. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 8.4. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 8.5. En una realización, el pH del tampón de formación de complejo ternario es de aproximadamente 8.6.

La termoestabilidad de un polipéptido Cas12i4 variante puede aumentar en condiciones favorables tales como la adición de una guía de ARN y ácido nucleico diana.

Evaluación de la estabilidad y funcionalidad del complejo ternario variante

En determinadas realizaciones, en el presente documento se proporcionan métodos para identificar un complejo ternario variante óptimo incluyendo (a) combinar un polipéptido Cas12i4 variante, una guía de ARN y un ácido nucleico diana en una muestra para formar el complejo ternario variante; (b) medir un valor

del complejo ternario variante; y (c) determinar que el complejo ternario variante es óptimo sobre la molécula de referencia, si el valor del complejo ternario variante es superior al valor de una molécula de referencia. En algunas realizaciones, el valor puede incluir, pero no se limita a, una medición de estabilidad (por ejemplo, valor de T_m , termoestabilidad), una tasa de formación de complejos ternarios, una medición de la afinidad de unión a ADN, una medición de la especificidad de unión a ADN y/o una medición de actividad del complejo (por ejemplo, medición de actividad nucleasa).

En algunas realizaciones, un complejo ternario variante óptimo se identifica por las etapas de: (a) combinar un polipéptido Cas12i4 variante, una guía de ARN y un ácido nucleico diana en una muestra para formar el complejo ternario variante; (b) detectar un valor de T_m del complejo ternario variante; y (c) determinar que el complejo ternario variante es estable si el valor de T_m del complejo ternario variante es superior a un valor de T_m de una molécula de referencia o un valor de referencia de T_m en al menos 8 °C.

Los métodos que implican una etapa de medir la termoestabilidad de un complejo ternario variante puede incluir, sin limitación, métodos para determinar la estabilidad de un complejo ternario variante, métodos para determinar una condición que promueve un complejo ternario variante estable, métodos para seleccionar un complejo ternario variante estable, y métodos para identificar un complejo binario óptimo para formar un complejo ternario variante estable. En determinadas realizaciones, puede medirse un valor de termoestabilidad de un complejo ternario variante.

Adicionalmente, en determinadas realizaciones, también puede medirse un valor de termoestabilidad de una molécula de referencia. En determinadas realizaciones, puede determinarse que un complejo ternario variante es estable si el valor de termoestabilidad medido del complejo ternario variante es superior al valor de termoestabilidad medido de la molécula de referencia o un valor de referencia de termoestabilidad, medido en las mismas condiciones experimentales, como se describe en el presente documento. En determinadas realizaciones, la molécula de referencia puede ser el polipéptido Cas12i4 variante sin una guía de ARN y/o ácido nucleico diana.

En determinadas realizaciones, el valor de termoestabilidad que se mide puede ser un valor de temperatura de desnaturalización. En estas realizaciones, el valor de referencia de termoestabilidad es un valor de referencia de temperatura de desnaturalización. En determinadas realizaciones, el valor de termoestabilidad que se mide puede ser un valor de T_m . En estas realizaciones, el valor de referencia de termoestabilidad puede ser un valor de referencia de T_m . En determinadas realizaciones, el valor de termoestabilidad puede medirse usando un ensayo de desplazamiento térmico. En determinadas realizaciones, un ensayo utilizado para medir la termoestabilidad puede implicar una técnica descrita en el presente documento que incluye, pero no se limita a, fluorimetría diferencial de barrido (DSF), calorimetría diferencial de barrido (DSC) o calorimetría de titulación isotérmica (ITC).

En determinadas realizaciones, un complejo ternario variante puede identificarse si la tasa de formación de complejos ternarios, la afinidad de unión a ADN, la especificidad de unión a ADN y/o la actividad del complejo (por ejemplo, actividad nucleasa) del complejo ternario variante es superior a un valor de la molécula de referencia o el valor de referencia (por ejemplo, un valor de un complejo ternario precursor). Por ejemplo, en determinadas realizaciones, el complejo ternario variante puede identificarse si

el valor de una tasa de formación de complejos ternarios, una afinidad de unión a ADN, una especificidad de unión a ADN y/o una actividad de complejo del complejo ternario variante es al menos un X % superior a un valor de la molécula de referencia o el valor de referencia (por ejemplo, un valor de un complejo ternario precursor). En determinadas realizaciones, los métodos descritos en el presente documento pueden comprender además etapas que incluyen medir la actividad del complejo ternario variante como se describe en el presente documento.

SUMINISTRO

Las composiciones o complejos descritos en el presente documento pueden formularse, por ejemplo, incluyendo un portador, tal como un portador y/o un portador polimérico, por ejemplo, un liposoma, y administrarse mediante métodos conocidos a una célula (por ejemplo, una planta procariota, eucariota, de mamífero, etc.). Dichos métodos incluyen, pero no se limitan a, transfección (por ejemplo, polímeros catiónicos mediados por lípidos, fosfato de calcio, dendrímeros); electroporación u otros métodos de disrupción de la membrana (por ejemplo, nucleofección), administración vírica (por ejemplo, lentivirus, retrovirus, adenovirus, AAV), microinyección, bombardeo con microproyectiles ("cañón génico"), Eugene, carga sónica directa, compresión celular, transfección óptica, fusión de protoplastos, impalefección, magnetofección, transferencia mediada por exosomas, transferencia mediada por nanopartículas lipídicas y cualquier combinación de los mismos.

En algunas realizaciones, el método comprende administrar uno o más ácidos nucleicos (por ejemplo, ácidos nucleicos que codifican el polipéptido Cas12i4 variante, una guía de ARN, ADN donante, etc.), uno o más transcritos de los mismos y/o un complejo polipéptido Cas12i4 variante/guía de ARN preformado (es decir, complejo binario variante) a una célula. Los métodos de suministro intracelular a modo de ejemplo incluyen, pero no se limitan a: virus o agentes similares a virus; métodos de transfección de base química, tales como los que usan fosfato cálcico, dendrímeros, liposomas o polímeros catiónicos (por ejemplo, DEAE-dextrano o polietilenimina); métodos no químicos, tales como microinyección, electroporación, compresión celular, sonoporación, transfección óptica, impalefección, fusión de protoplastos, conjugación bacteriana, suministro de plásmidos o transposones; métodos basados en partículas, tales como el uso de un cañón génico, magnetofección o transfección asistida por imán, bombardeo de partículas; y métodos híbridos, tales como nucleofección. En algunas realizaciones, la presente solicitud proporciona además células producidas mediante dichos métodos, y organismos (tales como animales, plantas, u hongos) que comprenden o se producen a partir de dichas células.

Células

Las composiciones o complejos descritos en el presente documento pueden suministrarse a una diversidad de células. En algunas realizaciones, la célula es una célula aislada. En algunas realizaciones la célula está en cultivo celular. En algunas realizaciones, la célula es *ex vivo*. En algunas realizaciones, la célula se obtiene de un organismo vivo, y se mantiene en un cultivo celular. En algunas realizaciones, la célula es un organismo unicelular.

En algunas realizaciones, la célula es una célula procariota. En algunas realizaciones, la célula es una célula bacteriana o derivada de una célula bacteriana. En algunas realizaciones, la célula bacteriana

no está relacionada con la especie bacteriana de la que deriva el polipéptido precursor. En algunas realizaciones, la célula es una célula de arquea o derivada de una de arquea. En algunas realizaciones, la célula es una célula eucariota. En algunas realizaciones, la célula es una célula vegetal o derivada de una célula vegetal. En algunas realizaciones, la célula es una célula fúngica o derivada de una célula fúngica.

5 En algunas realizaciones, la célula es una célula animal o derivada de una célula animal. En algunas realizaciones, la célula es una célula de invertebrado o derivada de una célula de invertebrado. En algunas realizaciones, la célula es una célula de vertebrado o derivada de una célula de vertebrado. En algunas realizaciones, la célula es una célula de mamífero o derivada de una célula de mamífero. En algunas realizaciones, la célula es una célula humana. En algunas realizaciones, la célula es una célula de pez cebra. En algunas realizaciones, la célula es una célula de roedor. En algunas realizaciones, la célula se prepara sintéticamente, a veces denominada célula artificial.

En algunas realizaciones, la célula se obtiene de una estirpe celular. Se conocen en la técnica una amplia diversidad de estirpes celulares para cultivo de tejidos. Los ejemplos de estirpes celulares incluyen, pero sin limitación, 293T, MF7, K562, HeLa y variedades transgénicas de las mismas. Las estirpes celulares
15 están disponibles de una diversidad de fuentes conocidas por los expertos en la materia (véase, por ejemplo, la Colección Americana de Cultivos Tipo (ATCC) (Manassas, Va.)). En algunas realizaciones, una célula transfectada con uno o más ácidos nucleicos (tales como vector codificante Ago y ADNg) o complejo de Ago-ADNg descrito en el presente documento se usa para establecer una nueva estirpe celular que comprende una o más secuencias derivadas de vectores para establecer una nueva estirpe celular que
20 comprende modificación al ácido nucleico diana. En algunas realizaciones, se usan células transfectadas transitoria o no transitoriamente con uno o más ácidos nucleicos (tales como el vector que codifica el polipéptido Cas12i4 variante y la guía de ARN) o el complejo polipéptido Cas12i4 variante/guía de ARN (es decir, complejo binario variante) descrito en el presente documento, o estirpes celulares derivadas de dichas células, para evaluar uno o más compuestos de ensayo.

25 En algunas realizaciones, la célula es una célula primaria. Por ejemplo, pueden hacerse pases de los cultivos de células primarias 0 veces, 1 vez, 2 veces, 4 veces, 5 veces, 10 veces, 15 veces o más. En algunas realizaciones, las células primarias se recogen de un individuo mediante cualquier método conocido. Por ejemplo, los leucocitos pueden recogerse por aféresis, leucocitaféresis, separación por gradiente de densidad, etc. Las células de tejidos tales como piel, músculo, médula ósea, bazo, hígado, páncreas, pulmón, intestino, estómago, etc. pueden recogerse mediante biopsia. Puede usarse una
30 disolución apropiada para la dispersión o suspensión de las células recogidas. En general, dicha disolución puede ser una solución salina equilibrada (por ejemplo, solución salina normal, solución salina tamponada con fosfato (PBS), solución salina equilibrada de Hank, etc.), convenientemente complementada con suero de ternero fetal u otros factores de origen natural, junto con un tampón aceptable a baja concentración. Los
35 tampones pueden incluir HEPES, tampones de fosfato, tampones de lactato, etc. Las células pueden usarse inmediatamente o pueden almacenarse (por ejemplo, mediante congelación). Las células congeladas pueden descongelarse y pueden reutilizarse. Las células pueden congelarse en un DMSO,

suero, tampón medio (por ejemplo, DMSO al 10 %, suero al 50 %, medio tamponado al 40 %) y/o alguna otra dicha disolución común utilizada para conservar las células a temperaturas de congelación.

En algunas realizaciones, el polipéptido Cas12i4 variante tiene actividad nucleasa que induce roturas bicatenarias o monocatenarias en un ácido nucleico diana, (por ejemplo ADN genómico). La rotura bicatenaria puede estimular vías de reparación del ADN endógeno celular, incluyendo Recombinación Dirigida por Homología (HDR), Unión de Extremos no Homólogos (NHEJ) o Unión de Extremos no Homólogos Alternativos (A-NHEJ). La NHEJ puede reparar el ácido nucleico diana escindido sin necesidad de un molde homólogo. Esto puede dar como resultado la supresión o inserción de uno o más nucleótidos en el ácido nucleico diana. La HDR puede producirse con un molde homólogo, tal como el ADN donante. El molde homólogo puede comprender secuencias que son homólogas de las secuencias que flanquean el sitio de escisión del ácido nucleico diana. En algunos casos, HDR puede insertar una secuencia polinucleotídica exógena en el ácido nucleico diana escindido. Las modificaciones del ADN diana debido a NHEJ y/o HDR pueden conducir a, por ejemplo, mutaciones, supresiones, alteraciones, integraciones, corrección de genes, reemplazo de genes, marcaje de genes, activación de transgenes, alteración de genes y/o inactivación de genes.

En algunas realizaciones, el cultivo celular se sincroniza para potenciar la eficiencia de los métodos. En algunas realizaciones, las células en las fases S y G2 se usan para la edición de genes mediada por HDR. En algunas realizaciones, la célula puede someterse al método a cualquier ciclo celular. En algunas realizaciones, el recubrimiento de células reduce significativamente la eficacia del método. En algunas realizaciones, el método se aplica a un cultivo celular a no más de aproximadamente uno cualquiera del 40 %, el 45 %, el 50 %, el 55 %, el 60 %, el 65 %, o el 70 % de confluencia.

En algunas realizaciones, la unión del complejo polipéptido Cas12i4 variante/guía de ARN (es decir, complejo binario variante) al ácido nucleico diana en la célula recluta una o más moléculas celulares endógenas o vías distintas de las vías de reparación del ADN para modificar el ácido nucleico diana. En algunas realizaciones, la unión del complejo binario variante bloquea el acceso de una o más moléculas celulares endógenas o vías al ácido nucleico diana, modificando de ese modo el ácido nucleico diana. Por ejemplo, la unión del complejo binario variante puede bloquear la transcripción endógena o la maquinaria de traducción para disminuir la expresión del ácido nucleico diana.

Kits

La invención también proporciona kits que pueden usarse, por ejemplo, para realizar un método descrito en el presente documento. En algunas realizaciones, los kits incluyen un polipéptido Cas12i4 variante de la invención, por ejemplo, una variante que comprende una sustitución de la Tabla 2 o un polipéptido variante de la Tabla 3. En algunas realizaciones, los kits incluyen un polinucleótido que codifica un polipéptido Cas12i4 variante de este tipo, y opcionalmente el polinucleótido está comprendido dentro de un vector, por ejemplo, como se describe en el presente documento. Los kits también pueden incluir opcionalmente una guía de ARN, por ejemplo, como se describe en el presente documento. La guía de ARN de los kits de la invención puede diseñarse para que se dirija a una secuencia de interés, como se conoce en la técnica. La variante de nucleasa y la guía de ARN pueden envasarse dentro del mismo vial o

en otro recipiente dentro de un kit o puede envasarse en viales independientes o en otros recipientes, cuyo contenido puede mezclarse antes de su uso. Los kits pueden incluir adicionalmente, opcionalmente, un tampón y/o instrucciones para el uso de la variante de nucleasa y/o guía de ARN.

Todas las referencias y publicaciones citadas en el presente documento se incorporan en el presente documento por referencia.

Secuencias que codifican la variante de Cas12i4 de la SEQ ID NO: 4

Identificador de secuencia	Secuencia
222	ATGGCCAGCATCTCACGCCCCTACGGGACCAAGCTGCGGCCTGATGCCC GGAAAAAGGAAATGCTGGACAAATTTTTCAACACTCTCACCAAGGGCCAA AGGGTATTTGCCGATCTGGCGCTGTGTATATACGGGAGCCTGACACTCGA GATGGCCAAGAGCCTAGAGCCGGAATCTGACAGCGAGCTCGTTTGTGCC ATCGGGTGGTTTAGACTCGTAGACAAAACCATTTGGAGTAAGGATGGGAT AAAGCAGGAGAATCTCGTGAAGCAATATGAGGCCTATTCGGAAAAGAGG CGTCAGAGGTGGTGAAGACTTACCTGAATAGCCCCTCATCCGACAAATAC GTATGGATTGATTGCCGACAAAAGTTTTTACGGTTCCAGCGGGAACCTGG AACGAGGAACCTGAGCGAAGATTTTGAATGTATGCTGTTTGAGCAGTATA TCCGGCTTACCAAAGGCGAGATTGAAGGATACGCCGCCATTTCTAATATG TTTGGTAATGGCGAGAAGGAGGACAGATCAAAGAAGAGGATGTACGCTA CTCGTATGAAGGACTGGTTGGAGGCAAATGAAAATATTACCTGGGAGCAG TACCGAGAAGCGCTCAAAAACCAAGTTGAACGCAAAGAATCTGGAGCAAGT GGTGGCCAACATAAGGGCAATGCCGGCGGCGCCGATCCATTTTTCAAG TATAGTTTCTCGAAGGAAGGTATGGTGTCCAAGAAAGAGCACGCGCAGCA GCTGGACAAGTTCAAAACAGTCCTGAAGAATAAAGCCCGCGATTAAATTT CCCCAACAAGGAGAAGCTCAAGCAGTACTTAGAAGCTGAGATTGGTATCC CAGTTGATGCAAACGTATACTCACAGATGTTCTCTAACGGGGTGTCCGAG GTCCAACCAAAAACAACACGAAACATGTCCTTTAGCAATGAGAAGCTAGA TCTGTTGACTGAACTGAAGGACTTAAACAAGGGCGATGGATTGGAATACG CCAGAGAAGTGCTTAATGGCTTCTTCGATAGTGAAGTCCACACTACAGAA GATAAATTCAACATTACTAGTCGGTACCTTGGTGGGGACAAATCCAACAG GCTCAGCAAGTTGTATAAGATTTGGAAGAAGGAGGGGGTGTATTGTGAAG AAGGAATTCAGCAGTTCTGCGAGGCTGTGAAGGATAAAATGGGCCAGATC CCCATCCGGAATGTCCTCAAATATTTATGGCAGTTTAGGGAGACCGTCAG TGCCGAGGACTTCGAAGCTGCAGCAAAGGCAAACCACTAGAGGAGAAA ATAAGCCGAGTGAAAGCTACCCGATTGTGATTTCCAACAGGTATTGGGC TTTCGGCACAAGCGCTCTGGTTGGCAACATCATGCCAGCTGACAAGCGTC ACCAGGGGGAGTATGCCGGACAGAACTTCAAAATGTGGCTGCGCGCAGA

5	GCTGCATTATGATGGCAAGAAAGCCAAACATCACCTCCCGTTCTATAATG CCAGGTTTTTCGAAGAAGTCTATTGTTACCATCCATCTGTGCTGAAATCA CTCCTTTTAAAACCAAACAATTCTGGCTGCGAGATCGGGAAGGATATCCG GATTATGTCTCTGTGGCTCTGAAGGACAATCCCTACAAGAAGGCGACTAA AAGGATTCTACGGGCCATCTACAACCCCGTTGCTAACACTACACGAGTGG ATAAAACAACCAATTGCTCCTTCATGATCAAAAGAGAGAACGACGAGTATA AACTGGTCATAAATAGGAAGATCTCGCGAGACCGCCCTAAGAGGATAGAA GTCGGACGCACCATCATGGGCTATGACCGAAACCAGACCGCGTCTGACA CCTACTGGATCGGTCGGCTTGTGCCTCCTGGGACCAGAGGAGCTTACAG 10AATTGGGGAGTGGAGTGTGCAGTATATCAAATCCGGACCAGTGCTGTCTT CCACACAGGGTGTTAATAACTCCACAACCGATCAGCTCGTCTACAACGGT ATGCCTTCAAGTAGCGAGCGCTTTAAGGCGTGGAAGAAGGCCAGAATGG CATTTATCCGCAAACTCATCAGACAACCTGAATGATGAGGGGTTAGAATCAA AAGGGCAGGACTATATTCCTGAAAAATCCAAGTTCCTTCGACGTGAGGGGG 15GAAACGTTGTATGTGTTCAACTCCAATTACCTTAAGGCCCTGGTATCAAAA CACAGGAAGGCTAAGAAGCCTGTGGAAGGCATCCTTGACGAGATCGAAG CCTGGACCTCCAAAGACAAAGATTCTGTTCACTGATGCGGCTCTCTAGC CTGAGTGATGCCTCCATGCAAGGTATAGCCTCACTAAAGAGCCTGATTAA CTCTTACTTTAATAAAAAATGGTTGCAAGACAATAGAGGATAAAGAAAAATTT 20AACCCAGTCTTGTATGCAAACTGGTGGAGGTGGAACAGAGACGTACAAA CAAACGGAGCGAGAAAAGTGGGAAGAATCGCTGGATCTCTAGAGCAGCTG GCGCTGCTTAACGGCGTCGAAGTGGTTATTGGAGAGGCAGATCTGGGAG AAGTTGAGAAAGGGAAGTCTAAGAAACAGAATAGCCGTAACATGGACTGG TGCGCCAAGCAGGTGGCACAGAGATTGGAGTACAAGCTGGCTTTTTCACG 25GCATCGGTTACTTTGGCGTTAATCCCATGTACACGAGTCACCAGGACCCC TTCGAGCATCGCCGTGTAGCCGACCATATCGTGATGCGTGCAAGATTTGA GGAAGTTAACGTAGAGAACATCGCTGAATGGCATGTGAGAACTTTAGCA ATTACCTCCGCGCCGACAGCGGCACCGGCCTTTACTACAAGCAGGCCAC GATGGACTTTTTGAAGCATTATGGACTGGAGGAGCACGCCGAGGGCTTG 30GAAAACAAAAAATTAAGTTCTATGACTTCAGGAAGATTCTTGAAGACAAA AACCTGACGTCTGTGATCATACCTAAACGCGGAGGGCGCATTACATGGC TACAAACCCTGTTACTTCCGACAGCACACCCATCACTTACGCCGGAAAAA CCTATAATCGGTGCAATGCAGACGAGGTGGCAGCTGCCAATATAGTGATC TCCGTCCTGGCACCAAGAAGTAAAAAGAATAGGGAACAAGACGATATCCC 35CCTCATAACTAAAAAGGCAGAGTCGAAGTCTCCCCCAAAGGATCGAAAC GGTCTAAGACCTCACAGTTGCCCCAAAAG
---	---

223	ATGGCTAGCATCAGCAGACCCTACGGCACCAAGCTGAGACCCGACGCTA GAAAGAAGGAGATGCTGGACAAGTTTTTCAATACCCTGACCAAGGGGCAG CGTGTGTTTCGCCGACCTGGCCCTGTGCATCTACGGCAGCCTGACCCTGG AGATGGCCAAGAGCCTGGAGCCCGAGAGCGACAGCGAACTGGTGTGTG CCATCGGCTGGTTCAGACTGGTGGATAAAACCATCTGGAGCAAGGACGG CATCAAGCAAGAGAACCTGGTGAAGCAGTACGAGGCCTACAGCGGCAAG GAGGCTAGCGAGGTGGTGAAGACCTACCTGAACAGCCCTAGCAGCGACA AGTACGTGTGGATCGACTGCAGACAGAAGTTCCTGAGATTTTCAGAGAGAG CTGGGCACAAGAAACCTGAGCGAGGATTTTCGAGTGCATGCTGTTTCGAGC AGTACATCAGACTGACCAAGGGAGAGATCGAGGGCTACGCCGCCATCAG CAACATGTTTCGGCAACGGCGAGAAGGAGGACCGGAGCAAGAAGAGAATG TACGCCACAAGAATGAAGGACTGGCTGGAGGCCAACGAGAACATCACCT GGGAGCAGTACAGAGAGGCCCTGAAGAATCAGCTGAACGCCAAGAACCT GGAGCAAGTGGTGGCCAACCTACAAGGGCAACGCCGGCGCGCCGACCC CTTCTTCAAGTACAGCTTCAGCAAGGAGGGCATGGTGAGCAAGAAGGAG CACGCTCAGCAGCTCGATAAGTTCAAACCGTGCTGAAGAACAAGGCTAG AGACCTGAACTTCCCCAACAAGGAGAAGCTGAAGCAGTACCTGGAGGCC GAGATCGGCATCCCCGTGGACGCCAACGTGTACTCTCAGATGTTTCAGCA ACGGCGTGAGCGAGGTGCAGCCCAAGACCACAAGAAACATGAGCTTCAG CAACGAGAAGCTGGACCTGCTGACCGAGCTGAAGGACCTGAACAAGGGC GACGGCTTCGAGTACGCTAGAGAGGTGCTGAACGGCTTCTTCGATAGCG AACTGCATACAACCGAGGACAAGTTCAACATTACAAGCAGATACCTGGGC GGCGACAAGAGCAACAGACTGAGCAAGCTGTACAAGATCTGGAAGAAGG AGGGCGTGACTGCGAGGAGGGCATTTCAGCAGTTCTGCGAGGCCGTGA AGGACAAGATGGGGCAGATCCCCATCAGAAACGTGCTGAAGTACCTGTG GCAGTTCAGAGAGACCGTGAGCGCCGAAGACTTCGAGGCAGCCGCCAAA GCCAACCACCTGGAGGAGAAAATTAGCAGAGTGAAAGCCCACCCCATCG TGATTAGCAATAGATACTGGGCCTTCGGCACAAGCGCCCTGGTGGGCAA CATCATGCCCCGCCGACAAGAGACACCAAGGCGAGTACGCCGGGCAGAAC TTCAAGATGTGGCTGAGAGCCGAGCTGCACTACGACGGCAAGAAGGCCA AGCACCACCTGCCCTTCTACAACGCTAGATTCTTCGAAGAGGTGTACTGC TACCACCCTAGCGTGGCCGAGATCACCCCTTCAAGACCAAGCAGTTTCG GCTGCGAGATCGGCAAGGACATCCCCGACTACGTGAGCGTGGCCCTGAA GGACAACCCCTACAAGAAGGCCACCAAGAGAATCCTGAGAGCCATCTAC AACCCCGTGGCCAACACCACAAGAGTCGATAAGACCACCAACTGCAGCTT CATGATCAAGAGAGAGAACGACGAGTATAAGCTGGTAATCAACAGAAAAA TTTCCCGAGACAGACCCAAGAGAATCGAGGTTCGGCAGAACCATAATGGG
5	
10	
15	
20	
25	
30	
35	

5		CTACGACAGAAATCAGACCGCTAGCGACACCTACTGGATCGGCAGACTG GTGCCCCCGGCACAAGAGGCGCCTACAGAATCGGCGAGTGGAGCGTG CAGTACATCAAGAGCGGCCCCGTGCTGAGCAGCACCCAAGGCGTGAACA ACAGCACCACCGATCAGCTGGTGTACAACGGCATGCCTAGCAGCAGCGA GAGATTCAAGGCCTGGAAGAAGGCTAGAATGGCCTTCATCAGAAAGCTGA TCAGACAGCTGAACGACGAGGGTCTGGAGAGCAAGGGCCAAGACTACAT CCCCGAGAACCCTAGCAGCTTCGACGTGAGAGGCGAGACCCTGTACGTG TTCAACTCCAATATCTGAAAGCTCTGGTGAGCAAGCACAGAAAGGCCAA GAAGCCCGTGGAGGGCATCCTGGACGAGATCGAGGCCTGGACAAGCAA 10 GGACAAGGACAGCTGCAGCCTGATGAGACTGAGCAGCCTGAGCGACGCT AGCATGCAAGGCATCGCTAGCCTGAAGAGCCTGATCAACAGCTACTTCAA CAAGAACGGCTGCAAGACCATCGAGGACAAGGAGAAGTTCAACCCCGTG CTGTACGCCAAGCTGGTGGAGGTGGAGCAGAGAAGAACCAACAAGAGAA GCGAGAAGGTAGGAAGAATCGCCGGCAGCCTGGAGCAGCTGGCCCTGC 15 TGAACGGCGTGGAGGTGGTGATCGGCGAGGCCGACCTGGGCGAGGTGG AGAAGGGCAAGAGCAAGAAGCAGAACAGCAGAAACATGGACTGGTGCGC CAAGCAAGTGGCTCAGAGACTGGAGTACAAGCTGGCCTTCCACGGCATC GGCTACTTCGGCGTGAACCCCATGTACACAAGCCACCAAGACCCCTTCG AGCACAGAAGAGTGGCCGACCACATCGTGATGAGAGCTAGATTCGAGGA 20 AGTAAACGTGGAGAACATCGCCGAGTGGCACGTGAGAACTTCAGCAAC TACCTGCGCGCGGACAGCGGCACCGGCCTGTACTACAAGCAAGCCACCA TGGACTTCCTGAAGCACTACGGCCTGGAGGAGCACGCCGAGGGCCTGG AGAACAAGAAGATCAAGTTCTACGACTTCAGAAAGATCCTGGAGGACAAG AACCTGACAAGCGTGATCATCCCCAAGAGAGGCGGCAGAATCTACATGG 25 CCACCAACCCCGTGACAAGCGACAGCACCCCATCACCTACGCCGGCAA GACCTACAACAGATGCAACGCCGACGAGGTGGCAGCCGCGAATATAGTG ATCAGCGTGCTAGCCCCCGAAGCAAGAAGAACAGAGAGCAAGACGACA TCCCCCTGATACCAAGAAGGCCGAGAGCAAGAGCCCCCCCCAAGGACAG AAAGAGAAGCAAGACATCTCAGCTGCCTCAGAAG
30	224	ATGGCCAGCATCAGCCGGCCCTACGGCACCAAGCTGCGGCCCGACGCC CGGAAGAAGGAGATGCTGGACAAGTTCTTCAACACCCTGACCAAGGGCC AGCGGGTGTTGCGCCGACCTGGCCCTGTGCATCTACGGCAGCCTGACCCT GGAGATGGCCAAGAGCCTGGAGCCCAGAGCGACAGCGAGCTGGTGTG CGCCATCGGCTGGTTCCGGCTGGTGGACAAGACCATCTGGAGCAAGGAC 35 GGCATCAAGCAGGAGAACCTGGTGAAGCAGTACGAGGCCTACAGCGGCA AGGAGGCCAGCGAGGTGGTGAAGACCTACCTGAACAGCCCCAGCAGCG ACAAGTACGTGTGGATCGACTGCCGGCAGAAGTTCCTGCGGTTCCAGCG

5	GGAGCTGGGCACCCGGAACCTGAGCGAGGACTTCGAGTGCATGCTGTT GAGCAGTACATCCGGCTGACCAAGGGCGAGATCGAGGGCTACGCCGCC ATCAGCAACATGTTTCGGCAACGGCGAGAAGGAGGACCGGAGCAAGAAGC GGATGTACGCCACCCGGATGAAGGACTGGCTGGAGGCCAACGAGAACAT CACCTGGGAGCAGTACCGGGAGGCCCTGAAGAACCAGCTGAACGCCAA GAACCTGGAGCAGGTGGTGGCCAACTACAAGGGCAACGCCGGCGGCGC CGACCCCTTCTTCAAGTACAGCTTCAGCAAGGAGGGCATGGTGAGCAAG AAGGAGCACGCCCAGCAGCTGGACAAGTTCAAGACCGTGCTGAAGAACA AGGCCCCGGGACCTGAACTTCCCCAACAAGGAGAAGCTGAAGCAGTACCT 10 GGAGGCCGAGATCGGCATCCCCGTGGACGCCAACGTGTACAGCCAGAT GTTTACAGCAACGGCGTGAGCGAGGTGCAGCCCAAGACCACCCGGAACATG AGCTTCAGCAACGAGAAGCTGGACCTGCTGACCGAGCTGAAGGACCTGA ACAAGGGCGACGGCTTCGAGTACGCCCGGGAGGTGCTGAACGGCTTCTT CGACAGCGAGCTGCACACCACCGAGGACAAGTTCAACATCACCAGCCGG 15 TACCTGGGCGGCGACAAGAGCAACCGGCTGAGCAAGCTGTACAAGATCT GGAAGAAGGAGGGCGTGACTGCGAGGAGGGCATCCAGCAGTTCTGCG AGGCCGTGAAGGACAAGATGGGCCAGATCCCCATCCGGAACGTGCTGAA GTACCTGTGGCAGTTCCGGGAGACCGTGAGCGCCGAGGACTTCGAGGC CGCCGCCAAGGCCAACACCTGGAGGAGAAGATCAGCCGGGTGAAGGC 20 CCACCCCATCGTGATCAGCAACCGGTACTGGGCCTTCGGCACCAAGCGCC CTGGTGGGCAACATCATGCCCCGCCGACAAGCGGCACCAGGGCGAGTAC GCCGGCCAGAACTTCAAGATGTGGCTGCGGGCCGAGCTGCACTACGACG GCAAGAAGGCCAAGCACCACTGCCCTTCTACAACGCCCGGTTCTTCGA GGAGGTGTACTGCTACCACCCAGCGTGGCCGAGATACCCCTTCAAG 25 ACCAAGCAGTTCGGCTGCGAGATCGGCAAGGACATCCCCGACTACGTGA GCGTGGCCCTGAAGGACAACCCCTACAAGAAGGCCACCAAGCGGATCCT GCGGGCCATCTACAACCCCGTGGCCAACACCACCCGGGTGACAAGACC ACCAACTGCAGCTTCATGATCAAGCGGGAGAACGACGAGTACAAGCTGG TGATCAACCGGAAGATCAGCCGGGACCGGCCCAAGCGGATCGAGGTGG 30 GCCGGACCATCATGGGCTACGACCGGAACCAGACCGCCAGCGACACCTA CTGGATCGGCCGGCTGGTGCCCCCGGCACCCGGGGCGCCTACCGGAT CGGCGAGTGGAGCGTGACGTACATCAAGAGCGGCCCCGTGCTGAGCAG CACCCAGGGCGTGAACAACAGCACCAACCGACCGAGCTGGTGTACAACGGC ATGCCAGCAGCAGCGAGCGGTTCAAGGCCTGGAAGAAGGCCCGGATG 35 GCCTTCATCCGGAAGCTGATCCGGCAGCTGAACGACGAGGGCCTGGAGA GCAAGGGCCAGGACTACATCCCCGAGAACCCAGCAGCTTCGACGTGCG GGGCGAGACCCTGTACGTGTTCAACAGCAACTACCTGAAGGCCCTGGTG
---	--

5	AGCAAGCACCGGAAGGCCAAGAAGCCCGTGGAGGGCATCCTGGACGAG ATCGAGGCCTGGACCAGCAAGGACAAGGACAGCTGCAGCCTGATGCGG CTGAGCAGCCTGAGCGACGCCAGCATGCAGGGCATCGCCAGCCTGAAG AGCCTGATCAACAGCTACTTCAACAAGAACGGCTGCAAGACCATCGAGGA CAAGGAGAAGTTCAACCCCGTGCTGTACGCCAAGCTGGTGGAGGTGGAG CAGCGGCGGACCAACAAGCGGAGCGAGAAGGTGGGCCGGATCGCCGGC AGCCTGGAGCAGCTGGCCCTGCTGAACGGCGTGGAGGTGGTGATCGGC GAGGCCGACCTGGGCGAGGTGGAGAAGGGCAAGAGCAAGAAGCAGAAC AGCCGGAACATGGACTGGTGCGCCAAGCAGGTGGCCCAGCGGCTGGAG 10 TACAAGCTGGCCTTCCACGGCATCGGCTACTTCGGCGTGAACCCCATGTA CACCAGCCACCAGGACCCCTTCGAGCACCGGCGGGTGGCCGACCACAT CGTGATGCGGGCCCGGTTTCGAGGAGGTGAACGTGGAGAACATCGCCGA GTGGCACGTGCGGAACCTTCAGCAACTACCTGCGGGCCGACAGCGGCAC CGGCCTGTACTACAAGCAGGCCACCATGGACTTCCTGAAGCACTACGGC 15 CTGGAGGAGCACGCCGAGGGCCTGGAGAACAAGAAGATCAAGTTCTACG ACTTCCGGAAGATCCTGGAGGACAAGAACCTGACCAGCGTGATCATCCC CAAGCGGGGCGGCCGGATCTACATGGCCACCAACCCCGTGACCAGCGA CAGCACCCCATCACCTACGCCGGCAAGACCTACAACCGGTGCAACGCC GACGAGGTGGCCGCCGCCAACATCGTGATCAGCGTGCTGGCCCCCGG 20 AGCAAGAAGAACC GGGAGCAGGACGACATCCCCCTGATCACCAAGAAGG CCGAGAGCAAGAGCCCCCCCCAAGGACCGGAAGCGGAGCAAGACCAGCC AGCTGCCCCAGAAG
25	225 ATGGCCTCAATAAGTCGGCCGTACGGAACAAAACCTCAGACCAGATGCCA GGAAAAAGGAAATGCTCGATAAATTCTTCAATACCCTGACAAAAGGACAG CGAGTCTTTGCGGATCTTGCGCTCTGTATTTATGGTTCACTGACACTGGA GATGGCGAAGTCACTCGAGCCAGAATCAGATAGTGAACCTGTATGTGCCA TCGGCTGGTTTAGATTGGTGGACAAGACTATATGGAGCAAGGATGGCATC AAGCAAGAAAACCTTGGTCAAGCAGTACGAGGCGTATAGTGGTAAAGAGG CGTCAGAGGTCGTGAAAACGTATCTTAACAGTCCTAGTTCAGACAAGTAT 30 GTCTGGATAGACTGTGCCAAAAGTTTCTTCGCTTCCAGCGGGAACCTCGG GACCCGAAATCTTAGTGAGGACTTTGAGTGCATGTTGTTCAACAATATAT CCGGCTGACTAAAGGTGAGATCGAGGGATACGCCGCAATTAGTAACATGT TCGGAAACGGAGAAAAAGAGGATAGGTCTAAGAAGCGGATGTACGCGAC ACGAATGAAGGATTGGCTGGAAGCAAATGAGAACATCACCTGGGAGCAG 35 TATAGGGAGGCTTTGAAAAATCAACTGAATGCTAAAAACTTGAGCAAGT CGTCGCAAATTATAAGGGAAACGCAGGTGGCGCCGACCCATTCTTTAAGT ATAGCTTCAGTAAGGAAGGAATGGTTTCAAAGAAAGAGCACGCCACGAG

5	CTTGATAAGTTCAAGACCGTACTGAAAAATAAAGCGCGGGACCTCAATTT CCCTAATAAGGAAAAATTGAAGCAATACTTGGAGGCTGAGATTGGTATAC CGGTAGATGCAAATGTCTATAGCCAAATGTTTAGTAACGGTGTGAGTGAG GTACAACCAAAGACAACGCGAAATATGAGTTTTTCAAATGAGAAGTTGGAT CTTTTGACGGAATTGAAGGATCTTAACAAGGGTGACGGCTTCGAGTACGC TCGGGAAGTCTTGAACGGTTTTTTTTGATTCCGAGTTGCACACCACTGAGG ACAAGTTTAACATCACCAGTCGATACCTGGGGGGCGATAAATCTAACAGG CTCAGTAAACTCTACAAGATATGGAAGAAAGAAGGAGTCGATTGCGAGGA AGGTATCCAACAGTTCTGCGAAGCTGTGAAGGACAAAATGGGACAAATCC
10	CCATAAGGAATGTGCTTAAATATCTTTGGCAGTTCCGCGAAACAGTCAGT GCAGAAGACTTCGAAGCTGCAGCCAAAGCCAACCACCTCGAAGAGAAAA TCAGCAGAGTAAAAGCGCATCCTATCGTCATAAGTAATCGCTACTGGGCG TTTGGTACTTCTGCGCTCGTTGGGAATATCATGCCGGCAGACAAAAGACA CCAAGGGGAGTACGCTGGGCAAATTTCAAATGTGGCTCAGGGCGGAG
15	CTCCATTATGATGGAAAGAAAGCAAAGCATCATCTGCCTTTTTATAACGCG CGGTTCTTTGAAGAAGTCTACTGTTATCATCCAAGCGTAGCTGAAATAACG CCCTTTAAACTAAACAGTTTGGGTGCGAAATAGGGAAAGATATTCCCGAT TATGTGTCCGTGGCGCTGAAAGATAATCCATACAAAAAGGCTACGAAGCG GATCCTGCGCGCCATTTATAATCCCGTCGCGAACACCACCCGCGTGATA
20	AGACAACATAATTGTTCTTTATGATAAAGCGCGAAAACGATGAGTATAAAC TGGTCATTAACCGCAAGATCTCTCGAGACAGGCCAAAACGCATAGAGGTA GGCCGAACCATTATGGGTTATGACAGGAATCAGACCGCCTCTGATACATA TTGGATTGGGAGGCTCGTGCCTCCTGGTACGAGGGGCGCTTACCGCATT GGAGAATGGTCAGTGCAGTACATCAAGTCCGGGCCCGTGCTTAGTTCTAC
25	CCAAGGGGTAAATAACTCAACTACGGACCAACTGGTGTATAACGGAATGC CAAGTAGTTCCGAACGGTTTAAAGCATGGAAGAAGGCTAGAATGGCGTTT ATACGGAACTCATACGACAATTGAATGATGAGGGACTTGAGAGCAAGGG TCAAGATTACATCCCAGAGAATCCAAGCTCTTTTGACGTCAGGGGTGAGA CACTGTATGTTTTCAATAGCAACTATTTGAAAGCACTCGTTTCTAAACACC
30	GGAAGGCCAAAAAACCTGTGGAAGGGATACTCGACGAGATTGAAGCCTG GACTTCTAAAGATAAAGATAGTTGTTCCCTTATGCGGCTCTCTAGCTTGAG CGATGCGTCAATGCAAGGGATTGCCTCTTTGAAAAGTCTCATCAACAGCT ACTTCAATAAGAACGGTTGCAAGACGATCGAGGATAAGGAGAAGTTCAAT
35	CCTGTTTTGTATGCCAAATTGGTAGAAGTGGAGCAGAGAAGAACTAACAA GAGATCTGAGAAGGTAGGCAGGATTGCCGGATCCCTTGAACAGCTGGCA CTCCTTAATGGGGTCGAAGTGGTCATTGGTGAAGCCGACCTTGGCGAAG TCGAAAAGGGCAAGTCCAAGAAGCAGAACAGTCGCAACATGGATTGGTG

5	CGCAAAACAGGTAGCACAAAGGCTCGAATATAAGCTCGCCTTCCACGGCA TTGGGTACTTCGGCGTTAACCCAATGTACACCAGTCACCAAGACCCCTTT GAGCATAGAAGAGTAGCAGATCATATAGTGATGAGGGCCAGATTCTGAAGA AGTGAACGTCGAGAATATCGCAGAATGGCACGTAAGGAATTTCTCCAATT ATCTGCGCGCTGATTCTGGTACAGGCCTCTACTACAAGCAGGCCACCATG GATTTTCTGAAACATTACGGGCTCGAGGAGCACGCCGAAGGTCTGGAGA ATAAGAAGATTAAGTTTTATGACTTCCGAAAGATTCTGGAGGACAAGAATC TTACCTCCGTGATCATCCCAAAGCGAGGGGGACGCATCTATATGGCTACC AATCCCGTGACTAGCGACAGCACTCCAATAACGTATGCCGGCAAAACCTA 10 CAATCGCTGTAACGCTGACGAGGTGGCTGCCGCCAATATAGTCATATCCG TGCTTGCTCCCCGAAGTAAAAAGAATCGGGAGCAAGACGATATTCCTTTG ATAACGAAAAAAGCCGAGAGTAAATCTCCACCCAAAGATCGGAAGAGATC AAAGACCTCACAACTCCCGCAAAAG
15	226 ATGGCATCTATCAGCAGACCATACGGAACCAAACCTGAGACCAGATGCTCG GAAAAAGGAGATGCTGGACAAGTTCTTCAACACCCTGACCAAGGGACAGA GGGTGTTGCGCGATCTGGCCCTGTGCATCTACGGCTCTCTGACCCTGGA AATGGCTAAGTCGCTCGAACCTGAGAGCGACTCCGAGCTGGTTTGTGCC ATTGGATGGTTCAGACTGGTCGATAAGACCATCTGGAGCAAGGACGGCAT 20 CAAGCAGGAGAACCTGGTGAAACAGTACGAGGCCTACAGCGGCAAGGAG GCGTCTGAAGTCGTGAAGACCTACCTGAACAGCCCTTCTAGTGATAAGTA CGTGTGGATCGACTGTAGACAGAAGTTCCTGAGATTTACGCGGGAACCTG GGCACCAGAAACCTGAGCGAGGACTTTGAATGCATGCTGTTTCGAGCAGT ACATCAGACTGACCAAGGGCGAAATCGAGGGATATGCCGCCATTAGCAA CATGTTGCGCAACGGCGAGAAAGAGGATAGAAGCAAGAAGAGAATGTAC 25 GCTACACGGATGAAGGACTGGCTGGAGGCCAACGAGAACATCACCTGGG AGCAGTATAGAGAAGCCCTGAAGAACCAGCTGAACGCCAAGAACCTCGA GCAGGTGGTGGCTAACTACAAGGGCAACGCCGGCGGCGCCGATCCTTTC TTCAAGTACTCCTTCAGCAAGGAGGGCATGGTGTCCAAGAAGGAGCATG CCCAGCAACTGGACAAATTCAAGACAGTGCTGAAGAACAAGGCCCGGGA 30 TCTGAACTTCCCCAACAAGGAGAAGCTCAAACAGTACCTGGAAGCCGAGA TCGGCATCCCCGTCGACGCCAATGTGTACTCTCAGATGTTCTCCAACGGC GTGTCTGAAGTGCAACCTAAGACAACAAGAAATATGAGCTTTAGCAATGA GAAGCTGGACCTGCTGACAGAACTGAAAGATCTGAACAAAGGCGATGGG TTCGAATACGCCC GCGAAGTGCTGAACGGGTTCTTTGATTCTGAGCTGCA 35 CACGACAGAAGATAAGTTCAATATCACCTCGCGGTACCTGGGAGGCGAC AAGAGCAATAGACTGAGCAAGCTGTATAAGATCTGGAAGAAGGAGGGCG TGGACTGCGAGGAGGGCATCCAACAGTTCTGCGAGGCTGTGAAGGATAA

5	GATGGGCCAAATCCCTATCAGGAACGTTCTCAAGTACCTGTGGCAGTTCA GAGAAACCGTGAGCGCCGAGGATTTTCGAGGCCGCCGCTAAGGCCAACCA CCTGGAGGAGAAGATCAGCAGAGTGAAGGCCACCCTATCGTGATCAGC AACAGATACTGGGCCTTCGGCACCTCTGCTCTGGTCGGAAATATCATGCC CGCCGATAAGCGGCACCAGGGCGAGTACGCCGGCCAGAACTTCAAGATG TGGCTGCGGGCCGAACCTTATTACGACGGCAAAAAGGCTAAACACCACC TGCCTTTCTACAACGCCAGATTCTTCGAGGAGGTGTACTGCTACCACCCC AGCGTGGCCGAAATCACACCTTTCAAGACTAAGCAGTTTGGATGTGAAAT CGGTAAGGATATCCCCGACTACGTCAGCGTGGCACTGAAAGACAACCCTT 10 ACAAAAAAGCTACCAAACGATTCTGAGAGCCATCTACAACCCCGTTGCC AATACCACAAGAGTGGACAAAACAACCAACTGCTCTTTCATGATCAAAAGA GAGAATGACGAATACAAGCTGGTAATAAACAGAAAGATCAGCAGAGACCG GCCTAAGCGCATCGAGGTGGGAAGAACCATTATGGGCTACGATAGAAAC CAGACCGCCAGCGATACCTACTGGATCGGCAGACTGGTGCCCCCTGGCA 15 CAAGAGGCGCCTACAGAATCGGCGAATGGTCCGTGCAGTACATCAAGAG CGGCCCTGTGCTGAGCTCTACCCAGGGAGTGAACAACAGCACCACCGAT CAGCTGGTGTACAACGGTATGCCTAGCAGCAGCGAGCGGTTCAAGGCAT GGAAGAAGGCCCGGATGGCCTTCATCCGGAAGCTGATCAGACAGCTGAA TGACGAGGGCCTGGAAAGCAAGGGACAGGACTACATCCCAGAGAACCCT 20 AGCAGCTTCGACGTGCGGGGCGAGACGCTGTACGTGTTCAACAGCAACT ATCTGAAAGCCCTGGTCAGCAAGCACAGAAAGGCCAAGAAGCCCGTGGA AGGTATCCTGGATGAGATCGAGGCCTGGACCAGCAAGGACAAGGACAGC TGCAGCCTGATGCGGCTGTCTTCTCTGAGCGACGCCTCCATGCAGGGCA TCGCCAGCCTGAAAAGCCTAATCAACAGCTACTTTAACAAGAACGGCTGC 25 AAGACAATCGAGGACAAGGAAAAGTTTAACCCTGTGCTGTATGCCAACT GGTGGAGGTGGAACAGCGGCGGACCAACAAGCGGAGCGAAAAAGTGGG CAGAATCGCCGGAAGCCTGGAGCAGCTTGCCCTGCTGAATGGCGTGGA GTGGTGATAGGCGAGGCCGACCTGGGCGAAGTGAGAAAGGGCAAGAGC AAGAAGCAGAACTCCAGAAACATGGACTGGTGCGCCAAACAGGTGGCCC 30 AGAGACTGGAATATAAGCTGGCTTTTCACGGCATCGGCTACTTCGGCGTT AATCCTATGTACACCAGCCACCAGGACCCCTTCGAGCACCGGAGAGTGG CCGACCACATAGTGATGAGAGCCCGGTTTCGAGGAAGTGAACGTGGAGAA CATCGCCGAGTGGCACGTGCGGAATTTTTCTAATTACCTGAGAGCCGACA GCGGAACAGGCCTGTACTACAAGCAGGCCACAATGGACTTCCTGAAGCA 35 CTACGGCCTGGAAGAGCACGCCGAGGGCCTGGAAAACAAGAAGATCAAG TTCTACGACTTCCGGAATCCTGGAGGATAAGAACCTCACCTCTGTCAT CATCCCTAAGCGAGGCGGAAGAATCTACATGGCCACAAACCCAGTGACC
---	---

5		AGCGACTCCACCCCTATCACCTACGCCGGCAAGACATACAACAGGTGTAA CGCCGACGAAGTGGCCGCTGCCAACATCGTGATCTCTGTGCTGGCTCCT AGATCAAAGAAGAATAGAGAACAAGACGACATTCCCCTGATCACAAAGAA AGCAGAGAGCAAGTCCCCACCTAAGGACAGAAAGAGAAGCAAAACCTCC CAGTTGCCTCAAAAA
10	227	ATGGCCTCAATCTCTAGGCCATATGGGACCAAATTGAGACCTGATGCTCG AAAAAAGGAGATGCTGGATAAGTTTTTCAACACACTTACCAAAGGCCAGA GAGTATTCGCTGACCTGGCTCTGTGTATCTATGGCTCTCTGACCCTGGAG ATGGCCAAATCTCTGGAGCCTGAGAGCGATTCCGAACCTTGTGTGCGCTAT TGGTTGGTTCAGGCTGGTTGACAAAACAATCTGGTCTAAAGATGGAATTA AGCAGGAAAACCTGGTGAAGCAATATGAGGCATATTCAGGAAAAGAGGCT TCCGAGGTGGTTAAGACTTACCTTAACTCACCATCAAGTGATAAGTACGTC TGGATCGACTGTAGGCAGAAATTTCTGCGCTTTCAGAGGGAACTCGGCAC TCGCAATCTGTCCGAAGATTTTGAGTGCAATGCTGTTTGAACAGTATATCCG
15		CCTCACAAAGGGCGAAATTGAGGGTTACGCTGCAATCTCCAACATGTTTCG GTAATGGCGAGAAGGAAGATAGGTCCAAGAAGCGCATGTACGCAACACG AATGAAAGACTGGCTCGAAGCCAACGAGAATATTACATGGGAGCAGTACC GGGAAGCTCTGAAGAATCAACTCAATGCGAAAAACCTGGAACAGGTGGTT GCGAATTACAAAGGGAATGCTGGTGGTGCTGACCCCTTCTTTAAATACTC
20		CTTCTCAAAGGAGGGTATGGTTTCAAAGAAAGAGCATGCTCAGCAGCTCG ACAAGTTCAAGACAGTGTTGAAGAATAAGGCCAGGGATTGAACTTCCCA AACAAAGAAAAGCTGAAGCAATACCTGGAAGCTGAGATTGGCATTCCCGT TGATGCTAACGTGTACAGCCAAATGTTCTCCAATGGCGTCAGTGAGGTCC
25		AACCGAAAACAACAAGAAACATGTCCTTCTCTAACGAGAAGCTCGATTGT TGACTGAATTGAAGGATCTGAACAAAGGAGACGGCTTCGAATATGCTCGG GAAGTGTTGAACGGCTTTTTCGACAGCGAGTTGCACACTACTGAAGATAA ATTCAACATCACCTCTAGGTATCTCGGCGGGGATAAGAGCAATAGACTCT
30		CTAAGTTGTACAAGATATGGAAAAAGGAGGGCGTGATTGTGAGGAGGG AATCCAGCAGTTCTGTGAGGCCGTGAAGGACAAGATGGGTCAAATCCCTA TCCGGAACGTGCTGAAGTACCTGTGGCAATTCCGAGAGACGGTGTCCGC TGAAGATTTTGAGGCCGCTGCCAAAGCAAATCACCTGGAGGAGAAGATAA
35		GTAGGGTGAAGGCACACCCCATCGTGATTAGTAACAGATATTGGGCATTT GGAACCTCAGCGTTGGTTGGAAACATTATGCCCGCTGATAAAAGACATCA AGGAGAGTATGCCGGGCAGAAATTTCAAATGTGGCTCCGCGCAGAACTC CACTATGACGGGAAAAAGGCCAAGCATCACTTGCCATTTTACAACGCCCG CTTCTTCGAGGAGGTCTATTGCTACCACCCCTCCGTGCGAGAGATCACAC CATTTAAACCAAACAGTTTGGTTGCGAGATCGGGAAGGACATTCCAGAT

5	TACGTAAGCGTCGCACTTAAAGACAATCCTTACAAGAAGGCGACAAAAAG GATCCTCAGAGCCATTTATAACCCCGTGGCCAACACCACAAGGGTGGACA AGACTACCAACTGTTTCCTTCATGATTAAGCGGGAGAACGACGAGTACAAA TTGGTGATTAACCGCAAGATTAGCAGAGACAGACCAAAAAGGATTGAAGT AGGACGGACCATCATGGGGTATGATCGGAATCAGACTGCCAGCGATACA TACTGGATCGGAAGATTGGTGCCACCTGGTACCAGGGGAGCATACCGGA TCGGAGAGTGGTCTGTACAGTACATTAAATCTGGCCCCGTGCTTTCCTCT ACCCAGGGCGTTAACAACCTCTACTACAGACCAGCTCGTTTACAACGGAAT GCCAAGTTCTTCCGAAAGATTTAAGGCCTGGAAAAAGGCCCGGATGGCCT 10 TCATCCGAAAGCTGATCCGCCAGCTGAATGACGAAGGGTTGGAATCTAAG GGCCAGGACTACATTCCTGAGAATCCTAGCAGTTTTGATGTTGCGGAGA GACGCTGTACGTGTTTAATTCTAACTATCTTAAAGCCCTCGTGAGTAAGCA TAGGAAGGCTAAAAAACCAGTCGAAGGTATATTGGACGAAATCGAAGCAT GGACCAGCAAGGACAAAGACTCTTGTTCTCTGATGCGACTGTCCAGCTTG 15 AGCGATGCTTCCATGCAGGGCATTGCAAGCCTGAAAAGTCTTATTAACAG CTACTTCAACAAAAATGGGTGCAAACTATCGAGGACAAAGAGAAGTTCA ACCCCGTGCTCTATGCAAAGTTGGTTGAAGTGGAGCAGCGACGGACAAA TAAACGGAGTGAGAAGGTCGGACGGATTGCTGGGAGCCTCGAACAATTG GCCCTGTTGAATGGGGTGGAGGTGGTGATCGGGGAAGCAGACCTTGGA 20 GAAGTAGAGAAGGGCAAAAGTAAAAAGCAGAATCCCGAAATATGGATTG GTGTGCCAAACAGGTGGCTCAGAGGCTGGAGTATAAACTCGCCTTTCATG GTATCGGGTATTTCCGCGTGAATCCTATGTACACCAGTCATCAGGACCCG TTTGAACACAGGAGGGTCGCTGACCATATTGTGATGAGAGCCAGGTTTGA AGAAGTCAATGTAGAGAACATCGCCGAATGGCACGTGCGAAATTTCTCAA 25 ACTATCTCCGGGCCGACTCCGGAACGGGTCTTTATTACAAACAAGCTACC ATGGATTTCTGAAGCATTACGGCCTGGAAGAGCATGCCGAGGGTCTGG AAAACAAGAAGATAAAATTCTACGATTTCCGGAAGATCCTCGAGGACAAG AACCTGACCTCCGTCAATCCCAAACGGGGTGGACGAATCTACATGGC CACAAATCCCGTTACGTCCGACAGCACCCCTATTACATACGCCGGCAAGA 30 CCTATAACCGGTGCAACGCAGATGAAGTCGCCGCTGCAAAATAGTTATC TCCGTTCTGGCCCCGAGGTCCAAGAAAAACAGAGAACAGGACGACATCC CCCTGATTACCAAAAAAGCTGAGTCAAAATCTCCGCCCAAAGACAGGAAG CGGAGCAAGACCTCCCAGCTGCCCCAGAAG
228	ATGGCTTCAATTTCCCGCCCCTATGGCACTAAGCTGCGCCCTGACGCCC GGAAAAAGGAGATGCTGGACAAGTTTTTTAATACTGACCAAGGGACAG 35 CGCGTGTTGCGCGACCTGGCCCTGTGTATCTACGGCTCTCTGACGCTGG AGATGGCTAAGTCCCTGGAGCCCGAGTCTGACTCTGAGCTGGTGTGCGC

5	TATCGGGTGGTTCAGACTGGTGGATAAGACCATCTGGTCTAAAGATGGCA TTAAGCAGGAGAACCTGGTGAAGCAATACGAGGCCTACTCAGGGAAGGA GGCCAGCGAAGTGGTGA AACCTACCTCAATAGCCCAAGCAGCGACAAG TACGTGTGGATTGATTGCCGCCAGAAGTTTCTCCGCTTCCAGCGGGAGCT GGGGACTAGGAATCTGAGCGAAGATTTTGAGTGCATGCTGTTTGAACAGT ACATCCGGCTGACTAAAGGGGAGATCGAGGGCTATGCCGCCATCAGCAA CATGTTTGGCAACGGGGAGAAAAGAGGACAGAAGTAAAAACGGATGTAT GCAACCCGCATGAAGGACTGGCTGGAAGCCAATGAGAACATCACCTGGG AACAGTATCGCGAAGCTCTGAAGAACCAGCTGAATGCCAAGAATCTGGAA 10 CAGGTGGTGGCCAATTACAAAGGGAACGCCGGCGGGGCCGATCCCTTCT TCAAATACTCTTTCAGTAAGGAAGGCATGGTGAGTAAGAAGGAGCACGCC CAGCAGCTGGATAAGTTTAAACGGTGCTCAAGAACAAGGCCAGGGACC TGAAC TTTCCCAATAAGGAGAAGCTGAAGCAGTACCTGGAGGCCGAGATC GGCATCCCCGTGGACGCGAACGTGTACTCCCAGATGTT CAGCAATGGAG 15 TGAGCGAGGTGCAGCCCAAGACCACCCGGAACATGAGCTTTTCTAACGA AAA ACTGGACCTGCTGACCGAGCTGAAGGACCTGAATAAGGGCGACGGA TTTGAGTACGCACGGGAAGTGCTGAATGGCTTCTTTGATAGCGAGCTGCA CACCACAGAGGATAAGTTCAATATCACCTCCAGGTACCTGGGAGGCGATA AGAGCAACAGACTCTCTAAGCTGTATAAGATTTGGAAGAAGGAAGGGGTG 20 GACTGCGAGGAGGGCATCCAGCAGTTCTGCGAGGCCGTGAAGGACAAG ATGGGCCAGATCCCTATCAGAAACGTGCTGAAGTATCTGTGGCAGTTCCG CGAGACCGTGAGCGCCGAGGACTTTGAGGCCGCCGCTAAGGCTAACCAC CTGGAAGAAAAGATCTCCCGGGTGAAAGCCCACCCTATTGTGATCTCCAA TAGATACTGGGCCTTCGGAAC TTTCTGCCCTGGTGGGAAATATCATGCCCCG 25 CCGACAAAAGACACCAGGGGGAGTATGCTGGCCAGAACTTCAAGATGTG GCTTAGGGCCGAGCTGCACTATGATGGCAAGAAGGCCAAGCATCACCTG CCTTTCTACAATGCTAGATTCTTTGAAGAGGTGTACTGTTACCACCCTAGC GTGGCCGAGATCACCCCTTTAAGACTAAACAGTTTGGCTGTGAGATTGG CAAGGACATCCCCGATTACGTGAGCGTGGCTCTGAAGGACAACCCATATA 30 AGAAAGCCACCAAACGCATCCTCCGGGCTATCTATAACCCCGTGGCCAAT ACTACCCGGGTGGACAAGACAACCAACTGTAGCTTCATGATCAAAAGAGA GAACGACGAGTATAAGCTGGTGATCAACAGAAAAATCTCCCGGGACCGC CCCAAAAGGATTGAGGTGGGACGCACCATTATGGGATACGATAGGAACC AGACCGCCTCAGACACCTACTGGATCGGCCGGCTGGTGCCTCCTGGCAC 35 TAGGGGGGCCTACCGCATCGGCGAATGGTCCGTGCAGTACATTAAATCC GGCCCCGTGCTGAGCTCCACACAGGGAGTGAATAATTCCACCACCGACC AGCTGGTGTACAACGGCATGCCCAGCAGCAGCGAGCGGTTCAAGGCCTG
---	--

5	GAAGAAGGCCCGGATGGCTTTTATACGGAAGCTGATCCGCCAGCTGAAC GATGAGGGCCTGGAATCCAAGGGCCAGGACTACATTCCCGAAAACCCTT CATCCTTCGACGTGAGAGGCGAAACTCTGTACGTGTTCAATTCCAACACTAC CTCAAGGCCCTGGTGTCTAAGCACAGGAAGGCCAAGAAGCCCGTGGAAG GCATCCTGGACGAGATTGAGGCATGGACCAGCAAGGACAAGGATAGCTG TTCTCTCATGAGACTGAGCAGCCTGTCCGATGCAAGCATGCAGGGGATC GCCTCCCTGAAGAGCCTGATTAACCTTTACTTTAACAAAAATGGCTGCAAG ACCATCGAGGATAAAGAGAAGTTTAATCCCGTGCTGTACGCAAACTCGT GGAGGTGGAGCAGAGGCGCACCAACAAGAGGAGCGAGAAAGTGGGGCG 10 GATCGCTGGAAGTCTGGAACAGCTGGCCCTGCTGAACGGCGTGAGGTC GTGATTGGCGAAGCGGACCTGGGCGAGGTGGAGAAGGGGAAGTCTAAG AAGCAGAACTCTAGGAATATGGACTGGTGCGCCAAGCAGGTGGCCCAGA GACTGGAATACAACTGGCCTTTTCATGGCATTGGATACTTCGGCGTGAAT CCTATGTACACATCACACCAGGATCCATTGAGCACAGGAGAGTGGCCG 15 ACCACATCGTGATGAGAGCCAGATTGAGGAGGTGAACGTGGAGAACAT CGCAGAGTGGCACGTGAGGAACTTTTCCAACATCTGCGGGCCGACTCT GGGACTGGACTGTATTACAAGCAGGCCACCATGGACTTCCTGAAGCACTA TGGCCTGGAGGAGCACGCTGAAGGGCTGGAAAACAAGAAAATAAAGTTTT ACGACTTCCGGAAGATTCTGGAGGATAAGAACCTGACCTCTGTTATCATC 20 CCAAAGCGGGGCGGCAGAACTACATGGCCACCAACCCCGTGACCTCCG ACAGCACCCCATACCTACGCCGAAAGACATACAACAGATGCAATGCT GACGAGGTGGCCGCCGCCAACATAGTGATTTCCTGTGCTGGCCCCAAGGA GTAAGAAGAACCGAGAGCAGGACGACATTCCACTGATTACCAAGAAGGCT GAATCCAAATCCCCACCAAGGACAGGAAGAGGAGCAAGACCTCTCAGC 25 TGCCTCAGAAG
---	--

EJEMPLOS

Los siguientes ejemplos se proporcionan para ilustrar adicionalmente algunas realizaciones de la presente invención pero no pretenden limitar el alcance de la invención; se entenderá por su naturaleza de ejemplo que se pueden usar como alternativa otros procedimientos, metodologías o técnicas conocidas por los expertos en la materia.

Ejemplo 1 - Preparación de construcciones variantes

En este Ejemplo, se generaron construcciones variantes.

Se construyeron moldes de ADN que comprendían mutaciones únicas a través de dos etapas de PCR usando cebadores mutagénicos directos y mutagénicos inversos pedidos a IDT™ (Integrated DNA Technologies, Inc.). En la primera etapa, se realizaron dos conjuntos de reacciones de PCR en 384 placas para generar dos fragmentos. Las regiones superpuestas de dos fragmentos de PCR contenían las mutaciones únicas deseadas y permitieron el ensamblaje de todo el molde de ADN a través de una segunda

PCR. En la segunda etapa, los fragmentos purificados de la primera etapa se usaron como molde para la PCR superpuesta (OL PCR) y los oligos Dir e Inv que se hibridan con la cadena principal del vector como los cebadores de OL PCR. Los moldes de ADN lineales resultantes contenían un promotor T7, un terminador T7 y el marco de lectura abierto para el polipéptido.

Estos moldes de ADN lineales se usaron directamente en un sistema de transcripción y traducción sin células para expresar las variantes de polipéptidos que contienen las mutaciones únicas. Se transfirieron además individualmente las construcciones variantes a vectores de transfección transitorios. Adicionalmente, se prepararon moldes de ADN que comprendían mutaciones combinatorias mediante PCR y posteriormente se transfirieron a vectores de transfección transitoria.

Ejemplo 2 – Ensayo de polarización de fluorescencia para la detección de complejos binarios variantes

En este Ejemplo, la capacidad de un polipéptido de nucleasa de tipo silvestre o variante y una guía de ARN para formar un complejo binario se evalúa a través de un ensayo de polarización de fluorescencia.

Se sintetizan mediante IDT™ fragmentos de ADNmc lineales que comprenden el complemento inverso de la secuencia promotora de la ARN polimerasa T7 en dirección 5' de la secuencia repetitiva directa y la guía de ARN diana de 20 pb deseada. Después, se generan moldes de transcripción *in vitro* (IVT) de ADCbc lineal hibridando un oligo directo de T7 universal (95-4 °C a 5 °C/minuto) con el ADNmc del complemento inverso y se rellenan con el fragmento de Klenow (New England Biolabs®) durante 15 minutos a 25 °C. Después se transcribe la plantilla de IVT resultante en una guía de ARN usando el kit de síntesis de ARN de alto rendimiento HiScribe T7 (New England Biolabs) a 37 °C durante 4 horas. Después de la transcripción, cada guía de ARN se purifica usando un kit concentrador y de limpieza de ARN (Zymo) y se almacena a -20 °C hasta su uso.

Después, la guía de ARN se marca con 6-carboxifluoresceína (6-FAM) (IDT™). El polipéptido de nucleasa 25 nM (polipéptido Cas12i4 de tipo silvestre o variante) en tampón de ensayo 1X (Tris-HCl 20 mM (pH 7.5), KCl 150 mM, MgCl₂ 5 mM, DTT 1 mM) se titula con concentraciones crecientes de una guía de ARN marcada (7.5 –250 nM). Se incuban los complejos a 37 °C durante 30 minutos antes de tomar medidas de polarización de fluorescencia usando un lector de microplacas (Infinite 200® Pro, Tecan).

También se investiga la formación de complejos binarios a diferentes temperaturas. Se realizan experimentos de unión adicionales descritos anteriormente de forma isotérmica a 25, 50, 60 y 70 °C.

La formación de un complejo binario tras la titulación de un polipéptido de nucleasa (polipéptido Cas12i4 de tipo silvestre o variante) con concentraciones crecientes de una guía de ARN (o la formación de un complejo binario tras la titulación de una guía de ARN con concentraciones crecientes de un polipéptido de nucleasa) da como resultado cambios en la señal de polarización de fluorescencia, en unidades de milipolarización (mP). Se genera una curva de unión trazando los cambios en la señal de polarización de fluorescencia en un intervalo de concentraciones de una guía de ARN.

Este Ejemplo indica cómo pueden determinarse y compararse las afinidades de unión de los polipéptidos de nucleasa (polipéptido Cas12i4 de tipo silvestre o variante) a guías de ARN.

Ejemplo 3 – Ensayo de desplazamiento de movilidad electroforética de ARN para la detección del complejo binario variante

Este Ejemplo describe el uso de un EMSA de ARN para determinar la capacidad de un polipéptido de nucleasa (de tipo silvestre o variante) para unirse a una guía de ARN.

Las guías de ARN sintético de IDT™ se marcan con un IRDye® 800CW 5' (también conocido como colorante IR800 o IR800) usando el kit de marcaje EndTag 5' (Vector® Laboratories) y maleimida IRDye® 800CW (LI-COR® Biosciences), como se detalló anteriormente en Yan *et al.*, 2018. Después del marcaje, se limpian y se concentran las guías de ARN mediante extracción con fenol cloroformo. Las concentraciones se cuantifican mediante Nanodrop™.

Para los ensayos de unión de ARN, se diluyen polipéptidos de nucleasa (polipéptidos Cas12i4 de tipo silvestre o variante) a 2.5 µM en tampón de unión 1X (NaCl 50 mM, Tris-HCl 10 mM, MgCl₂ 10 mM, DTT 1 mM, pH 7.9). Después, los polipéptidos se diluyen en serie de 2.5 µM a 37.5 µM en tampón de unión 1X. Los polipéptidos se diluyen de nuevo 1:10 en tampón de unión 1X más una guía de ARN marcada con IR800 50 nM y se mezclan minuciosamente. Estas reacciones pueden incluir además 0.5-5 µg de ARNt, que sirve como inhibidor competitivo para disminuir la unión no específica del polipéptido al ARN y, por lo tanto, facilitar determinaciones precisas de unión específica. Se incuban las reacciones a 37 °C durante 1 hora. Se añade 1 µl de azul de bromofenol 100X a las reacciones para la visualización del frente del colorante, después toda la reacción se carga en un gel de retardo de ADN al 6 % (ThermoFisher Scientific™), que funciona durante 90 minutos a 80 V. Se forman imágenes del gel en Licor® Odyssey® CLx.

Este ensayo se basa en el principio de que la velocidad a la que el ARN migra a través del gel está determinada por su tamaño. Una muestra de ARN solamente puede migrar una distancia particular. Sin embargo, si el ARN se une a un polipéptido, aparece una banda que representa un complejo de ARN más grande y menos móvil, que se "desplaza hacia arriba" en el gel.

Se miden las intensidades de dos tipos de bandas: 1) una banda de ARN solamente y 2) una banda de ARN "desplazada hacia arriba" unida a polipéptido. Si todo el ARN está unido a un polipéptido, solamente se observa una banda desplazada hacia arriba. A medida que disminuye la concentración de polipéptido, la intensidad de la banda desplazada hacia arriba disminuye, mientras que la intensidad de la banda de ARN solo aumenta. En comparación con las afinidades de unión del ARN por los polipéptidos de nucleasa (polipéptidos Cas12i4 de tipo silvestre o variante), una mayor afinidad polipéptido/ARN se caracteriza por una unión más específica a concentraciones más bajas de polipéptido.

Este Ejemplo indica cómo pueden determinarse y compararse las afinidades de unión de polipéptidos de nucleasa de tipo silvestre a las guías de ARN y las afinidades de unión de los polipéptidos Cas12i4 variantes a las guías de ARN.

Ejemplo 4 – Ensayo de desplazamiento de movilidad electroforética de ADN para la detección del complejo ternario de Cas12i4 Variante

Este ejemplo describe el uso de un Ensayo de desplazamiento de movilidad electroforética de ADN (EMSA) para determinar la capacidad de una guía de ARN, un polipéptido Cas12i4 (Cas12i4 de tipo silvestre o variante) y un sustrato de ADN diana para formar un complejo ternario.

Se transformaron Cas12i4 de tipo silvestre de la SEQ ID NO: 2 y variante de Cas12i4 de la SEQ ID NO: 4 en *E. coli* BL21 (DE3) (New England BioLabs®) y BL21(DE3)pLySS (Novagen®), respectivamente, y se expresaron bajo un promotor T7. Se cultivaron inicialmente las células transformadas durante la noche en 5 ml de caldo Luria (TEKNOVA™) + kanamicina 50 µg/ml, seguido de la inoculación en 1 l de medio de caldo Terrific (TEKNOVA™) + kanamicina 50 µg/ml. Se cultivaron células de Cas12i4 de tipo silvestre y variantes a 37 °C hasta una DO₆₀₀ de 0.6-0.8 y 3, respectivamente, después se indujo la expresión de la proteína con IPTG 0.5 mM. Después, se cultivaron cultivos a 18 °C durante 14-18 horas adicionales. Se recogieron los cultivos y se sedimentaron a través de centrifugación, después se resuspendieron en 1 ml de tampón de extracción por 5 g de sedimento celular (HEPES 50 mM, pH 7.5, NaCl 500 mM, glicerol al 5 %, TCEP 0.5 mM). Se lisaron las células mediante un disruptor celular (Constant System Limited), después se centrifugaron a 20,000 x g durante 20 minutos a 4 °C para aclarar el lisado. Se añadió polietilenimina (PEI) al 0.2 % al lisado aclarado y la mezcla se incubó a 4 °C con rotación constante de un extremo a otro durante 20 minutos. Después, el lisado se centrifugó de nuevo a 20,000 x g durante 10 minutos. El Cas12i4 de tipo silvestre se purificó mediante cromatografía de intercambio iónico e hidrófoba, y Cas12i4 variante se purificó mediante cromatografía de intercambio iónico y afinidad con metales inmovilizados. Después de la purificación, se procesaron las fracciones en geles SDS-PAGE y las fracciones que contenían proteína del tamaño apropiado se combinaron y se concentraron usando unidades centrífugas Amicon® Ultra15 de 30 kD. Se intercambiaron los tampones de las proteínas en HEPES 12.5 mM pH 7.0, NaCl 120 mM, TCEP 0.5 mM y glicerol al 50 %. Después se midieron las concentraciones usando Nanodrop™ (ThermoFisher Scientific™) y las proteínas se almacenaron a -20 °C.

Se prepararon RNP usando una relación 2:1 de una guía de ARN sintético (Integrated DNA Technologies, IDT™) con respecto al polipéptido. Las secuencias de la guía de ARN se muestran en la Tabla 13. ARNcr 1 (SEQ ID NO: 62) correspondió a la Diana 1 (SEQ ID NO: 65), ARNcr 2 (SEQ ID NO: 63) correspondió a la Diana 2 (SEQ ID NO: 66) y ARNcr 3 (SEQ ID NO: 64) correspondió a la Diana 3 (SEQ ID NO: 67). Las RNP formaron complejo durante 30 minutos a 37 °C en NEBuffer™ 2 1X (NEB2; New England Biolabs®; NaCl 50 mM, Tris-HCl 10 mM, MgCl₂ 10 mM, DTT 1 mM, pH 7.9). Después de la formación de complejo, se realizó una dilución en serie 1:2 de 5 puntos de 5 µM a 37.5 µM, usando NEB2 1X como tampón de dilución. Las reacciones Apo (polipéptido sin guía de ARN) se prepararon de la misma manera, completando el volumen de la guía de ARN con H₂O.

Tabla 13. Secuencias de una guía de ARN de EMSA de ADN.

Guía de ARN	Secuencia
-------------	-----------

crRNA 1 (AAVS1_T6)	AGACAUGUGUCCUCAGUGACACGUAGCCUCUCCCGCUCUG (SEQ ID NO: 62)
crRNA 2 (AAVS1_T7)	AGACAUGUGUCCUCAGUGACACGGAAGUGGUUGGUCAGC (SEQ ID NO: 63)
crRNA 3 (EMX1_T4)	AGACAUGUGUCCUCAGUGACACGGGGAGGCCUGGAGUCAU (SEQ ID NO: 64)

Se generaron los sustratos diana de ADNbc de las secuencias de la Tabla 14 mediante PCR a partir de un oligo (Integrated DNA Technologies, Inc.) usando los cebadores de la Tabla 15. Antes de la PCR, el extremo 5' del cebador directo se marcó con un colorante IR800, como se describe en Yan *et al.*, 2018. Usando Amplitaq Gold® (ThermoFisher Scientific™), el sustrato de ADNbc se amplificó después con el cebador directo marcado IR800 y el cebador inverso sin marcar. El ADNbc resultante se purificó con un kit concentrador y de limpieza de ARN (Zymo) y se cuantificó mediante Nanodrop™ (ThermoFisher Scientific™).

Tabla 14. Sustratos diana de EMSA de ADN.

Identificador de la diana	Secuencia
Diana 1 (AAVS1_T6)	CGCAAAGTGTGGGATTACAGGCGTGAGCCCGGCCATTCTGAGACTGTGT TGCAGGCATTGTACATCTTCGCCTGATGCACAGCAGGTATCTCCTGCCAC AAGGAAAACCTCCTGCAGAACCACAGTAGGGATGCAACACGCTACCCCT GTGTTGACCTTGATGCTACACTCTCACCACCGCACCAACCTTGATGCTAC ACTCTCACCACCGCACCAACCTTGATGCTACACTCTCACCACCGCACCC AACCTTGATGCTACACTCTCACCACCGCACCAACCTTGATGCTACACTCT CACCCACCGCACCAACCTTGATGCTACACTCTCACCACCTGCACCAACCTT GATGCTACACTGTTGCCTGCGTTTCTCCTTGACATTCTTTGTAGCCTCTCC CGCTCTGGTTCAGGGCCCAGCTAGGGATCCAGATCTGGGTGATTTAGGCT CCCTCTGTCTGGATCAGTCCTCTTTTCCCTTGACCCCAGGGAGGCCGG GAATGGCCTCCAGGGGGTCTGTGAACTTTCTGACGTTGTATTTCTGCAG AAATTGCTCATAACTTGCATCAGCTTCTCAGAGGGGG (SEQ ID NO: 65)
Diana 2 (AAVS1_T7)	CCTGAGCCCATCACTGTTGCAAAGGTGACAGGAAGGCCTGGTGATGTGCG CACCCTGGAGCCAGGCTATGGGCCCGGTCACATTGAAACCATATGGGGC AAAGTGTGGGTGAGGAAAGTCAAGATGAGGTCACAGGGGAAGGGAGAAT GGATTTTCGTAGGCCCAAGCAGCAGCTGTGCTGCAGGGACACGCAGCAG CACCATGTCCTGTGCAGAAGGGACCCTCCCTGGCCACTTTGCACAGGGG CATGGAAGTGGCAGGAAGAAGACATGATGTGTTTTTGAACATTTGAAGCC AGCTCACTTGGAATTCCAGCATCCAAGTCAGCTGGAAGAGGGGGAGTTAC CCTTGGAGGCAGGCGGAATCGACCATTGGATAGCTCCAAGTGCTGACAAG GGCGGACACGGGAGCTGATTTCTGCCTGGTGGGAAAGGTGATGATTCCA

		GCTACTTTGGGAAGTGGTTGGTCAGCATGGATTATAGCCGAAGGCCCCAG CTTTGCCTTGTCTAGCAGTTCCACTCCTGGGCAGCCCGAGAGAGGCCTT CCCAACCATGGGCAGATGTTTCATCATAGTATTGTTTGCAGTAGTAAGAGGT CGGAGCCCACACCAAAG (SEQ ID NO: 66)
5	Diana 3 (EMX1_T4)	AGGACAAAGTACAAACGGCAGAAGCTGGAGGAGGAAGGGCCTGAGTCCG AGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCAC GAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATGACTAGGGT GGGCAACCACAAACCCACGAGGGCAGAGTGCTGCTTGCTGCTGGCCAGG CCCCTGCGTGGGCCCCAAGCTGGACTCTGGCCACTCCCTGGCCAGGCTTT 10 GGGGAGGCCTGGAGTCATGGCCCCACAGGGCTTGAAGCCCGGGGCCGC CATTGACAGAGGGACAAGCAATGGGCTGGCTGAGGCCTGGGACCACTTG GCCTTCTCCTCGGAGAGCCTGCCTGCCTGGGCGGGCCCCGCCACCG CAGCCTCCCAGCTGCTCTCCGTGTCTCCAATCTCCCTTTTGTGTTTGATGCA 15 TTTCTGTTTTAATTTATTTCCAGGCACCACTGTAGTTTAGTGATCCC (SEQ ID NO: 67)

Tabla 15. Cebadores para la generación de sustrato diana de EMSA de ADN.

	Identificador de la diana	Secuencia de Cebador Directo	Secuencia de Cebador Inverso
20	Diana 1 (AAVS1_T6)	CGCAAAGTGTGGGATTACAGGCGT (SEQ ID NO: 68)	CCCCCTCTGAGAAGCTGATGCAA GT (SEQ ID NO: 69)
	Diana 2 (AAVS1_T7)	CCCCCTCTGAGAAGCTGATGCAAGT (SEQ ID NO: 70)	CTTTGGTGTGGGCTCCGACCTCT TA (SEQ ID NO: 71)
25	Diana 3 (EMX1_T4)	AGGACAAAGTACAAACGGCAGAAGC TGG (SEQ ID NO: 72)	GGGATCACTAACTACAGTGGTG CCTGG (SEQ ID NO: 73)

Las muestras de RNP y las muestras de Apo (control) se diluyeron 1:10 en tampón de unión 1X (NaCl 50 mM, Tris-HCl 10 mM, TCEP 1 mM, glicerol al 10 %, EDTA 2 mM, pH 8.0) más sustrato de ADN diana marcado con IR800 20 nM y se mezclaron bien. Se incubaron las reacciones a 37 °C durante 1 hora. Se añadió azul de bromofenol a las reacciones para la visualización del frente del colorante, después toda la reacción se cargó en un gel de retardo de ADN al 6 % (ThermoFisher Scientific™), que funcionó durante 90 minutos a 80 V. Se formaron imágenes del gel en Licor® Odyssey® CLx.

Las **FIG. 1A**, **FIG. 1B** y **FIG. 1C** muestran geles de EMSA para la Diana 1 (AAVS1_T6), la Diana 2 (AAVS1_T7) y la Diana 3 (EMX1_T4), respectivamente. En cada gel, las calles "Apo" (calles 1 y 8) incluían ADN diana más Cas12i4 de tipo silvestre (calle 1) o variante de Cas12i4 de la SEQ ID NO: 4 (calle 8). Las calles "Ref" incluían ADN diana solo. Las calles 2-6 de las **FIG. 1A**, **FIG. 1B** y **FIG. 1C** correspondían a

concentraciones decrecientes de las RNP que comprendían Cas12i4 de tipo silvestre de la SEQ ID NO: 2), de 500 nM a 37 nM. Las calles 9-13 en las **FIG. 1A, FIG. 1B y FIG. 1C** correspondían a concentraciones decrecientes de las RNP que comprendían la variante de Cas12i4 de la SEQ ID NO: 4, de 500 nM a 37 nM.

Los geles de la **FIG. 1A, FIG. 1B, y FIG. 1C** muestran bandas de ADN que migraron a diferentes distancias. En este ensayo, la velocidad a la que el ADN migra a través del gel se determina por su tamaño. Una muestra de ADN solamente escapaz de migrar una distancia particular. Sin embargo, si un RNP se une al ADN, aparece una banda que representa un complejo de ADN más grande y menos móvil, que se "desplaza hacia arriba" en el gel. Por lo tanto, las flechas de las **FIG. 1A, FIG. 1B y FIG. 1C** apuntan a "ADNbc no unido" y al "ADNbc unido", en donde el "ADNbc unido" migró menos que el "ADNbc no unido".

La **FIG. 1A** muestra que para la concentración más alta de RNP de Cas12i4 de tipo silvestre (calle 2) y para la concentración más alta de RNP de Cas12i4 variante (calle 9) sólo había presentes bandas de ADNbc no unido, lo que indica que los RNP de Cas12i4 de tipo silvestre y variantes no formaron un complejo ternario con ADN diana AAVS1_T6.

La **FIG. 1B** muestra que incluso a las concentraciones más altas de RNP de Cas12i4 de tipo silvestre (calle 2), sólo había presentes bandas de ADNbc no unido, lo que indica que las RNP de Cas12i4 de tipo silvestre no formaron un complejo ternario con el ADN diana AAVS1_T7. Sin embargo, se observaron bandas de ADNbc unido con RNP preparadas con Cas12i4 variante (calles 9-10). Por lo tanto, las RNP preparadas con Cas12i4 variante tuvieron una mayor afinidad por el ADN diana AAVS1_T7 que el Cas12i4 de tipo silvestre.

Análogamente, la **FIG. 1C** muestra que incluso a las concentraciones más altas de RNP de Cas12i4 de tipo silvestre (calle 2), sólo había presentes bandas de ADNbc no unido, lo que indica que las RNP de Cas12i4 de tipo silvestre no formaron un complejo ternario con el ADN diana EMX1. Sin embargo, se observaron bandas de ADNbc unido con RNP preparadas con Cas12i4 variante (calle 10). Por lo tanto, las RNP preparadas con variante tuvieron una mayor afinidad por el ADN diana EMX1 que el Cas12i4 de tipo silvestre.

Basándose en los datos de las **FIG. 1A, FIG. 1B y FIG. 1C**, las RNP preparadas con Cas12i4 variante tenían una mayor afinidad por múltiples dianas de ADNbc, en comparación con la afinidad de las RNP de Cas12i4 de tipo silvestre por las dianas de ADNbc.

Con el fin de demostrar que el desplazamiento hacia arriba del ADN del sustrato dependía de la secuencia, se incubaron las RNP con sustratos diana que no coincidían. Se realizaron estas reacciones de la misma manera, formando cualquier volumen de polipéptido con tampón NEB2 1X. Las reacciones que comprenden el polipéptido Cas12i4 (de tipo silvestre o variante), ARNcr 1 (SEQ ID NO: 62) y Diana 3 de ADN (SEQ ID NO: 67) se muestran en la **FIG. 1D**.

En el gel de la **FIG. 1D**, las calles "Apo" (calles 1 y 8) incluían el ADN de la Diana 3 (SEQ ID NO: 67) más Cas12i4 de tipo silvestre (calle 1) y Cas12i4 variante (calle 8). Las calles "Ref" incluían ADN de Diana 3 solo. Las calles 2-6 de la **FIG. 1D** correspondían a concentraciones decrecientes de RNP de Cas12i4 de tipo silvestre preparadas con ARNcr 1 (SEQ ID NO: 62), de 500 nM a 37 nM. Las calles 9-13

en la **FIG. 1D** correspondían a concentraciones decrecientes de RNP preparadas con Cas12i4 variante de la SEQ ID NO: 4 y ARNcr 1 (SEQ ID NO: 62), de 500 nM a 37 nM.

Como se muestra en la **FIG. 1D**, los ADNbc permanecieron sin unirse a RNP en todas las concentraciones, lo que indica que las RNP para Cas12i4 de tipo silvestre y variante no pudieron formar un complejo ternario. Por lo tanto, la capacidad de una RNP para unirse a un sustrato de ADN diana, como se muestra en la **FIG. 1B** y la **FIG. 1C**, dependía de las secuencias de la guía de ARN y del sustrato de ADN diana.

En general, este Ejemplo muestra que RNP (complejos binarios) preparados con polipéptidos Cas12i4 variantes tenía mayor afinidad por múltiples dianas de ADN (para producir un complejo ternario) que la afinidad de las RNP de Cas12i4 de tipo silvestre por las dianas de ADN.

Ejemplo 5 – Ensayo de escisión *in vitro* para la determinación de la formación de complejos ternarios de Cas12i4 Variante

Este Ejemplo describe métodos para evaluar la actividad bioquímica *in vitro* de RNP de Cas12i4 (Cas12i4 de tipo silvestre o variante) en un sustrato de ADN diana como medio para determinar la formación de complejos ternarios.

Las guías de ARN y sustratos de ADNbc en este Ejemplo son idénticos a aquellos en la Tabla 13 y Tabla 14, respectivamente. Los sustratos de ADNbc en este ensayo permanecen sin marcar. Se generaron muestras de RNP y apo y se incubaron de la misma manera que se describió en el Ejemplo 4, después se diluyen en serie de 1 μ M a 15.7 nM en NEB2 1X. Después, las muestras de RNP y apo se diluyen adicionalmente 1:10 en NEB2 1X y se añadió un sustrato de ADNbc diana a 20 nM. Las reacciones se mezclaron minuciosamente, después se incubaron a 37 °C durante 1 hora y después se interrumpieron con 1 μ l de Proteinasa K 20 mg/ml (ThermoFisher Scientific™). Las reacciones se incubaron durante otros 15 a 50 °C, después la reacción entera se ejecutó en un E-gel de agarosa al 2 % (ThermoFisher Scientific™). Los geles se visualizaron con bromuro de etidio en un Gel Doc™ EZ Gel Imager (BioRad®).

Las **FIG. 2A**, **FIG. 2B** y **FIG. 2C** muestran geles de escisión para la Diana 1 (AAVS1_T6), la Diana 2 (AAVS1_T7) y la Diana 3 (EMX1_T4), respectivamente. En cada gel, la calles "Apo" (calles 1 y 11) incluían ADN diana más Cas12i4 de tipo silvestre (calle 1) o variante de Cas12i4 de la SEQ ID NO: 4 (calle 11). Las calles "Ref" incluían ADN diana solo. Las calles 2-9 de las **FIG. 2A**, **FIG. 2B** y **FIG. 2C** corresponden a concentraciones decrecientes de las RNP que comprendían Cas12i4 de tipo silvestre de la SEQ ID NO: 2), de 1 μ M a 15.7 nM. Las calles 12-19 de las **FIG. 2A**, **FIG. 2B** y **FIG. 2C** corresponden a concentraciones decrecientes de las RNP que comprendían la variante de Cas12i4 de la SEQ ID NO: 4, de 1 μ M a 15.7 nM.

En las **FIG. 2A**, **FIG. 2B** y **FIG. 2C**, se midieron las intensidades de dos tipos de bandas: 1) una banda de ADN de longitud completa (sin escindir) y 2) una o más bandas de ADN escindidas desplazadas hacia abajo. Una RNP inactiva se caracterizó por una banda de ADN de longitud completa (por ejemplo, la RNP no pudo formar un complejo ternario con el sustrato de ADN). Una RNP activa produjo una o más bandas de ADN escindidas con desplazamiento hacia abajo (por ejemplo, la RNP pudo formar un complejo ternario con el sustrato de ADN). A medida que disminuyó la concentración de una RNP activa, aumentó la intensidad de la banda de longitud completa y disminuyó la intensidad de la banda o bandas escindidas.

En comparación con la actividad de múltiples RNP, una RNP que tenía mayor actividad que otra se caracterizó por bandas escindidas más intensas a concentraciones de RNP más bajas.

Las **FIG. 2A**, **FIG. 2B** y **FIG. 2C** muestran que Cas12i4 de tipo silvestre y Cas12i4 variante escindieron cada una de las dianas *in vitro*. Sin embargo, Cas12i4 variante pudo escindir cada una de las dianas a concentraciones de RNP más bajas. Por lo tanto, el Cas12i4 variante de la SEQ ID NO: 4 presentó una mayor actividad de escisión que Cas12i4 de tipo silvestre.

Ejemplo 6 – Ensayos de estabilidad *in vitro* de polipéptidos Cas12i4 variantes y complejos binarios variantes

En este Ejemplo, se evalúa la estabilidad de una RNP variante.

Para el estudio de estabilidad acelerada se generan RNP (5 μ M) de la misma forma que se describe en el Ejemplo 4, y posteriormente se almacenan las muestras a 25 °C durante 48 horas.

Los ensayos de escisión *in vitro* (como se describe en el Ejemplo 5) se realizan en las muestras de RNP. Se comparan estos resultados con los del Ejemplo 5 para determinar el grado en el que las RNP variantes almacenadas a 25 °C durante 48 horas conservan la actividad bioquímica.

También se incubó el polipéptido Apo (sin guía de ARN) a 25 °C durante 48 horas. Se realizan ensayos de EMSA de ARN en las muestras de apo usando el método descrito en el Ejemplo 3. Se comparan estos resultados con los del Ejemplo 3 para determinar el grado en el que una nucleasa variante es capaz de formar un complejo binario con una guía de ARN.

Las muestras de Apo incubadas a 25 °C durante 48 horas también forman complejo con guías de ARN para formar RNP, usando el método descrito en el Ejemplo 4. Después se realizan ensayos de escisión *in vitro* de acuerdo con los métodos del Ejemplo 5. Se comparan los resultados del ensayo con los del Ejemplo 5 para evaluar los niveles de actividad de RNP variantes formadas con la proteína incubada a 25 °C.

Los métodos de este Ejemplo permiten comparar la estabilidad de los polipéptidos Cas12i4 de tipo silvestre y variantes y las RNP (complejos binarios) de tipo silvestre y variantes. Un polipéptido de nucleasa que demuestra una mayor unión específica a una guía de ARN que otro polipéptido de nucleasa a la guía de ARN es indicativo de un polipéptido más estable. Una RNP que demuestra una escisión *in vitro* más robusta de un ADN diana que la escisión por otra RNP con un polipéptido de nucleasa diferente es indicativa de un complejo binario más estable.

Ejemplo 7 – Direccionamiento de genes de mamíferos por nucleasas variantes

Este Ejemplo describe la evaluación de indels en múltiples dianas usando Cas12i4 de tipo silvestre y variantes de Cas12i4 introducidos en células de mamífero por transfección transitoria.

Las nucleasas de la SEQ ID NO: 2, la SEQ ID NO: 3 y la SEQ ID NO: 4 se clonaron en una cadena principal pcda3.1 (Invitrogen®). Las guías de ARN se clonaron en una cadena principal pUC19 (New England Biolabs). Después, se realizaron maxipreps de los plásmidos y se diluyeron a 1 μ g/ μ l. La guía de ARN y las secuencias diana se muestran en la Tabla 16.

Tabla 16. Dianas de mamíferos y ARNcr correspondientes.

	Identificador de la diana	Secuencia de ARNcr	Secuencia de la diana
5	AAVS1_T1	AGACAUGUGUCCUCAGUGACACUGUCCCCC CAAGUUUUGGACCCCU (SEQ ID NO: 74)	TGTCCCCCAAGTTTTGGAC CCCT (SEQ ID NO: 75)
	AAVS1_T3	AGACAUGUGUCCUCAGUGACACGUGAGAAU GGUGCGUCCUAGGUGU (SEQ ID NO: 76)	GTGAGAATGGTGCGTCCTAG GTGT (SEQ ID NO: 77)
	AAVS1_T5	AGACAUGUGUCCUCAGUGACACAACUGGCC CUGGCUUUGGCAGCCU (SEQ ID NO: 78)	AACTGGCCCTGGCTTTGGCA GCCT (SEQ ID NO: 79)
10	AAVS1_T6	AGACAUGUGUCCUCAGUGACACGUAGCCUC UCCCGCUCUGGUUCAG (SEQ ID NO: 80)	GTAGCCTCTCCCGCTCTGGT TCAG (SEQ ID NO: 81)
	AAVS1_T7	AGACAUGUGUCCUCAGUGACACGGGAAGUG GUUGGUCAGCAUGGAU (SEQ ID NO: 82)	GGGAAGTGTTGGTCAGCAT GGAT (SEQ ID NO: 83)
15	EMX1_T1	AGACAUGUGUCCUCAGUGACACGGGAAGUG GUUGGUCAGCAUGGAU (SEQ ID NO: 84)	GGGCGCAGGGCCACCTGGA CCCTG (SEQ ID NO: 85)
	EMX1_T2	AGACAUGUGUCCUCAGUGACACGGAUGGCG ACUUCAGGCACAGGAU (SEQ ID NO: 86)	GGATGGCGACTTCAGGCACA GGAT (SEQ ID NO: 87)
20	EMX1_T4	AGACAUGUGUCCUCAGUGACACGGGGAGGC CUGGAGUCAUGGCCCC (SEQ ID NO: 88)	GGGGAGGCCTGGAGTCATG GCCCC (SEQ ID NO: 89)
	EMX1_T6	AGACAUGUGUCCUCAGUGACACGAGCCAGU GUUGCUAGUCAAGGGC (SEQ ID NO: 90)	GAGCCAGTGTTGCTAGTCAA GGGC (SEQ ID NO: 91)
	EMX1_T7	AGACAUGUGUCCUCAGUGACACAGCAAGGG ACUAUUCAGGGAUGAA (SEQ ID NO: 92)	AGCAAGGGACTATTCAGGGA TGAA (SEQ ID NO: 93)
25	EMX1_T8	AGACAUGUGUCCUCAGUGACACAAAUUGA GCAAUCUACCCUGGUC (SEQ ID NO: 94)	AAAATTGAGCAATCTACCCTG GTC (SEQ ID NO: 95)
	VEGFA_T1	AGACAUGUGUCCUCAGUGACACUGGGGGUG ACCGCCGGAGCGCGGC (SEQ ID NO: 96)	TGGGGGTGACCGCCGGAGC GCGGC (SEQ ID NO: 97)
30	VEGFA_T2	AGACAUGUGUCCUCAGUGACACAAUCCUCC ACCAGUCAUGGUGACA (SEQ ID NO: 98)	AATCCTCCACCAGTCATGGT GACA (SEQ ID NO: 99)
	VEGFA_T3	AGACAUGUGUCCUCAGUGACACGUUGACAU UGUCCACACCUGGAU (SEQ ID NO: 100)	GTTGACATTGTCCACACCTG GAAT (SEQ ID NO: 101)
	VEGFA_T5	AGACAUGUGUCCUCAGUGACACUUAACUC UCCAUGGACCAGGCUC (SEQ ID NO: 102)	TTAAACTCTCCATGGACCAGG CTC (SEQ ID NO: 103)
35	VEGFA_T6	AGACAUGUGUCCUCAGUGACACGCCCAUAC UGGGGACCAAGGAAGU (SEQ ID NO: 104)	GCCCATACTGGGGACCAAGG AAGT (SEQ ID NO: 105)
	VEGFA_T7	AGACAUGUGUCCUCAGUGACACGCCGUAAC CCUUCGUGGGUAGAGA (SEQ ID NO: 106)	GCCGTAACCTTCGTGGGTA GAGA (SEQ ID NO: 107)

Aproximadamente 16 horas antes de la transfección, se sembraron en placa 100 µl de 25,000 células HEK293T en DMEM/FBS al 10 % + Pen/Estrep en cada pocillo de una placa de 96 pocillos. El día de la transfección, las células eran confluyentes al 70-90 %. Para cada pocillo que va a transfectarse, se preparó una mezcla de 0.5 µl de Lipofectamine™ 2000 y 9.5 µl de Opti-MEM y después se incubó a temperatura ambiente durante 5-20 minutos (Solución 1). Después de la incubación, la mezcla de lipofectamine™:OptiMEM™ se añadió a una mezcla separada que contenía 126 ng de plásmido de nucleasa y 174 ng de plásmido guía y agua hasta 10 µl (Solución 2). En el caso de los controles negativos, el ARNcr no se incluyó en la Solución 2. Se mezclaron las mezclas de la solución 1 y la solución 2 pipeteando hacia arriba y hacia abajo y después se incubaron a temperatura ambiente durante 25 minutos. Después de la incubación, se añadieron gota a gota 20 µl de la mezcla de la solución 1 y la solución 2 a cada pocillo de una placa de 96 pocillos que contenía las células. 72 horas después de la transfección, las células se tripsinizan añadiendo 10 µl de TrypLE™ al centro de cada pocillo y se incuban durante aproximadamente 5 minutos. Después, se añadieron 100 µl de medios D10 a cada pocillo y se mezclaron para resuspender las células. Después, las células se centrifugaron a 500 g durante 10 minutos y se desechó el sobrenadante. Se añadió tampón QuickExtract™ (Lucigen®) a 1/5 de la cantidad del volumen de suspensión celular original. Las células se incubaron a 65 °C durante 15 minutos, 68 °C durante 15 minutos y 98 °C durante 10 minutos.

Se prepararon muestras para la secuenciación de última generación mediante dos ciclos de PCR. El primer ciclo (PCR1) se usó para amplificar regiones genómicas específicas en función de la diana. Los productos de PCR1 se purificaron mediante purificación en columna. Se realizó el ciclo 2 de PCR (PCR2) para añadir adaptadores e índices de Illumina®. Después, las reacciones se agruparon y se purificaron mediante purificación en columna. Se realizaron ejecuciones de secuenciación con un equipo de rendimiento medio o alto NextSeq™ v2.5 de 150 ciclos.

La **FIG. 3** muestra la actividad indel para Cas12i4 de tipo silvestre de la SEQ ID NO: 2, Cas12i4 variante de la SEQ ID NO: 3, o Cas12i4 variante de la SEQ ID NO: 4. Cas12i4 variante de la SEQ ID NO: 3 o Cas12i4 variante de la SEQ ID NO: 4 demostraron una mayor actividad de indel en cada una de las dianas en comparación con Cas12i4 de tipo silvestre de la SEQ ID NO: 2. Por lo tanto, las variantes de Cas12i4 modificadas por ingeniería genética demostraron una actividad nucleasa aumentada en células de mamífero.

Ejemplo 8 – Direccionamiento de genes de mamíferos por nucleasa variante usando secuencias de PAM 5'-NTTN-3' y 5'-NVTN-3'

Este Ejemplo describe la evaluación de indel en múltiples dianas adyacentes a un PAM 5'-NTTN-3' o 5'-NVTN-3' usando Cas12i4 de tipo silvestre y variantes de Cas12i4 introducidos en células de mamífero mediante transfección transitoria.

Las construcciones de guía de ARN y nucleasa se prepararon y se transfectaron en células HEK293T como se describe en el Ejemplo 7. La guía de ARN y las secuencias diana se muestran en la Tabla 17.

Tabla 17. Dianas de mamíferos, PAM y ARNcr correspondientes.

	Gen diana	Secuencia de la diana	Secuencia de ARNcr	PAM
5	AAVS1	CTCATTCTTCCCTTAGGGGT (SEQ ID NO: 126)	AGACAUGUGUCCUCAGUGACACCUCA UUCUUCCCUUAGGGGU (SEQ ID NO: 127)	NTTN
	AAVS1	CCCCCAAGTCCCTCACCTC (SEQ ID NO: 128)	AGACAUGUGUCCUCAGUGACACCCC CCCAAGUCCCUCACCUC (SEQ ID NO: 129)	NTTN
10	AAVS1	ACCAGGTCGTGGCCGCCTCT (SEQ ID NO: 130)	AGACAUGUGUCCUCAGUGACACACCA GGUCGUGGCCGCCUCU (SEQ ID NO: 131)	NTTN
	AAVS1	TAGGCCTGCATCATCACCGT (SEQ ID NO: 132)	AGACAUGUGUCCUCAGUGACACUAG GCCUGCAUCAUCACCGU (SEQ ID NO: 133)	NTTN
15	AAVS1	ACTGGCCCTGGCTTTGGCAG (SEQ ID NO: 134)	AGACAUGUGUCCUCAGUGACACACU GGCCCUGGCUUUGGCAG (SEQ ID NO: 135)	NTTN
	AAVS1	TAGCCTCTCCCGCTCTGGTT (SEQ ID NO: 136)	AGACAUGUGUCCUCAGUGACACUAG CCUCUCCCGCUCUGGUU (SEQ ID NO: 137)	NTTN
20	AAVS1	TAGCCGAAGGCCCCAGCTTT (SEQ ID NO: 138)	AGACAUGUGUCCUCAGUGACACUAG CCGAAGGCCCCAGCUUU (SEQ ID NO: 139)	NTTN
	AAVS1	GCGGGTATGGGAAGGGCTTT (SEQ ID NO: 140)	AGACAUGUGUCCUCAGUGACACGCG GGUAUGGGAAGGGCUUU (SEQ ID NO: 141)	NTTN
25	AAVS1	ACACGGGCCACCGTTTCTCA (SEQ ID NO: 142)	AGACAUGUGUCCUCAGUGACACACAC GGGCCACCGUUUCUCA (SEQ ID NO: 143)	NVTN
30	AAVS1	ACCCCCAAGTCCCTCACCT (SEQ ID NO: 144)	AGACAUGUGUCCUCAGUGACACACCC CCCAAGUCCCUCACCU (SEQ ID NO: 145)	NVTN
35	AAVS1	GGTGTTCAACAGGTCGTGGC (SEQ ID NO: 146)	AGACAUGUGUCCUCAGUGACACGGU GUUACACAGGUCGUGGC (SEQ ID NO: 147)	NVTN

5	AAVS1	GGCCTGCATCATCACCGTTT (SEQ ID NO: 148)	AGACAUGUGUCCUCAGUGACACGGC CUGCAUCAUCACCGUUU (SEQ ID NO: 149)	NVTN
	AAVS1	GCCCTGGCTTTGGCAGCCTG (SEQ ID NO: 150)	AGACAUGUGUCCUCAGUGACACGCC CUGGCUUUGGCAGCCUG (SEQ ID NO: 151)	NVTN
	AAVS1	GCCTCTCCCGCTCTGGTTCA (SEQ ID NO: 152)	AGACAUGUGUCCUCAGUGACACGCC UCUCCCGCUCUGGUUCA (SEQ ID NO: 153)	NVTN
10	AAVS1	ATTATAGCCGAAGGCCCCAG (SEQ ID NO: 154)	AGACAUGUGUCCUCAGUGACACAUUA UAGCCGAAGGCCCCAG (SEQ ID NO: 155)	NVTN
15	AAVS1	GTGCAGAGGGTGGGCCGGGG (SEQ ID NO: 156)	AGACAUGUGUCCUCAGUGACACGUG CAGAGGGUGGGCCGGGG (SEQ ID NO: 157)	NVTN
	EMX1	TGCTGAGAACCACCCAGGGT (SEQ ID NO: 158)	AGACAUGUGUCCUCAGUGACACUGC UGAGAACCACCCAGGGU (SEQ ID NO: 159)	NTTN
20	EMX1	GGTGCCCTAGGAAGCTGCCT (SEQ ID NO: 160)	AGACAUGUGUCCUCAGUGACACGGU GCCCUAGGAAGCUGCCU (SEQ ID NO: 161)	NTTN
	EMX1	ATGCCCAAAGGTCAGATGAT (SEQ ID NO: 162)	AGACAUGUGUCCUCAGUGACACAUG CCCAAAGGUCAGAUGAU (SEQ ID NO: 163)	NTTN
25	EMX1	GGGGAGGCCTGGAGTCATGG (SEQ ID NO: 164)	AGACAUGUGUCCUCAGUGACACGGG GAGGCCUGGAGUCAUGG (SEQ ID NO: 165)	NTTN
30	EMX1	GCACCACTGTAGTTTAGTGA (SEQ ID NO: 166)	AGACAUGUGUCCUCAGUGACACGCA CCACUGUAGUUUAGUGA (SEQ ID NO: 167)	NTTN
	EMX1	TTTGAGCCAGTGTTGCTAGT (SEQ ID NO: 168)	AGACAUGUGUCCUCAGUGACACUUU GAGCCAGUGUUGCUAGU (SEQ ID NO: 169)	NTTN
35	EMX1	CTTTAGCAAGGGACTATTCA (SEQ ID NO: 170)	AGACAUGUGUCCUCAGUGACACCUU UAGCAAGGGACUAUUCA (SEQ ID NO: 171)	NTTN

5	EMX1	AGCAATCTACCCTGGTCCTC (SEQ ID NO: 172)	AGACAUGUGUCCUCAGUGACACAGCA AUCUACCCUGGUCCUC (SEQ ID NO: 173)	NTTN
	EMX1	AGAACCACCCAGGGTCCAGG (SEQ ID NO: 174)	AGACAUGUGUCCUCAGUGACACAGAA CCACCCAGGGUCCAGG (SEQ ID NO: 175)	NVTN
	EMX1	GGGTGCCCTAGGAAGCTGCC (SEQ ID NO: 176)	AGACAUGUGUCCUCAGUGACACGGG UGCCCUAGGAAGCUGCC (SEQ ID NO: 177)	NVTN
10	EMX1	AGATGATAGCATAGGTACAC (SEQ ID NO: 178)	AGACAUGUGUCCUCAGUGACACAGAU GAUAGCAUAGGUACAC (SEQ ID NO: 179)	NVTN
15	EMX1	ACTCCAGGCCTCCCCAAAGC (SEQ ID NO: 180)	AGACAUGUGUCCUCAGUGACACACUC CAGGCCUCCCCAAAGC (SEQ ID NO: 181)	NVTN
	EMX1	ACTAAACTACAGTGGTGCCT (SEQ ID NO: 182)	AGACAUGUGUCCUCAGUGACACACUA AACUACAGUGGUGCCU (SEQ ID NO: 183)	NVTN
20	EMX1	CTTTGAGCCAGTGTTGCTAG (SEQ ID NO: 184)	AGACAUGUGUCCUCAGUGACACCUU UGAGCCAGUGUUGCUAG (SEQ ID NO: 185)	NVTN
25	EMX1	CCTTGCTAAAGAAACATGTG (SEQ ID NO: 186)	AGACAUGUGUCCUCAGUGACACCCU UGCUAAGAAACAUGUG (SEQ ID NO: 187)	NVTN
	EMX1	GAGCAATCTACCCTGGTCCT (SEQ ID NO: 188)	AGACAUGUGUCCUCAGUGACACGAG CAAUCUACCCUGGUCCU (SEQ ID NO: 189)	NVTN
	VEGFA	TGGGGGTGACCGCCGGAGCG (SEQ ID NO: 190)	AGACAUGUGUCCUCAGUGACACUGG GGGUGACCGCCGGAGCG (SEQ ID NO: 191)	NTTN
30	VEGFA	TGGGCTGCTTGGGGTTGTCA (SEQ ID NO: 192)	AGACAUGUGUCCUCAGUGACACUGG GCUGCUUGGGGUUGUCA (SEQ ID NO: 193)	NTTN
35	VEGFA	TCCACACCTGGAATCGGCTT (SEQ ID NO: 194)	AGACAUGUGUCCUCAGUGACACUCCA CACCUGGAAUCGGCUU (SEQ ID NO: 195)	NTTN

5	VEGFA	GTGTAGAGAGGAAAATGTGG (SEQ ID NO: 196)	AGACAUGUGUCCUCAGUGACACGUG UAGAGAGGAAAAUGUGG (SEQ ID NO: 197)	NTTN
	VEGFA	GGGGCTTTGTTTGGGAAGCT (SEQ ID NO: 198)	AGACAUGUGUCCUCAGUGACACGGG GCUUUGUUUGGGAAGCU (SEQ ID NO: 199)	NTTN
	VEGFA	ACACTTCCTTGGTCCCCAGT (SEQ ID NO: 200)	AGACAUGUGUCCUCAGUGACACACAC UCCUUGGUCCCCAGU (SEQ ID NO: 201)	NTTN
10	VEGFA	GTGGGTAGAGAAGGATTCTG (SEQ ID NO: 202)	AGACAUGUGUCCUCAGUGACACGUG GGUAGAGAAGGAUUCUG (SEQ ID NO: 203)	NTTN
15	VEGFA	AAATCCTCCCTTGACCCACC (SEQ ID NO: 204)	AGACAUGUGUCCUCAGUGACACAAAU CCUCCCUUGACCCACC (SEQ ID NO: 205)	NTTN
	VEGFA	CCAAGGGGGAGGGCTCACGC (SEQ ID NO: 206)	AGACAUGUGUCCUCAGUGACACCCAA GGGGGAGGGCUCACGC (SEQ ID NO: 207)	NVTN
20	VEGFA	ACAACCCCAAGCAGCCCACA (SEQ ID NO: 208)	AGACAUGUGUCCUCAGUGACACACAA CCCCAAGCAGCCCACA (SEQ ID NO: 209)	NVTN
25	VEGFA	AAAGCCGATTCCAGGTGTGG (SEQ ID NO: 210)	AGACAUGUGUCCUCAGUGACACAAAG CCGAUUCCAGGUGUGG (SEQ ID NO: 211)	NVTN
	VEGFA	TGTGTAGAGAGGAAAATGTG (SEQ ID NO: 212)	AGACAUGUGUCCUCAGUGACACUGU GUAGAGAGGAAAAUGUG (SEQ ID NO: 213)	NVTN
	VEGFA	ATCCAGCTTCCCAAACAAAG (SEQ ID NO: 214)	AGACAUGUGUCCUCAGUGACACAUCC AGCUUCCCAAACAAAG (SEQ ID NO: 215)	NVTN
30	VEGFA	CTGGGGACCAAGGAAGTGTC (SEQ ID NO: 216)	AGACAUGUGUCCUCAGUGACACCUG GGGACCAAGGAAGUGUC (SEQ ID NO: 217)	NVTN
35	VEGFA	GGTAGAGAAGGATTCTGTGC (SEQ ID NO: 218)	AGACAUGUGUCCUCAGUGACACGGU AGAGAAGGAUUCUGUGC (SEQ ID NO: 219)	NVTN

VEGFA	TTCAAATCCTCCCTTGACCC (SEQ ID NO: 220)	AGACAUGUGUCCUCAGUGACACUUCA AAUCCUCCCUUGACCC (SEQ ID NO: 221)	NVTN
-------	--	--	------

Los indels se analizaron como se describe en el Ejemplo 7 y los resultados se muestran en la **FIG.**

- 5 4. Las formas abiertas representan dianas con PAM 5'-NVTN-3' y las formas cerradas representan dianas con PAM 5'-NTTN-3'. Los círculos representan Cas12i4 de tipo silvestre (SEQ ID NO: 2) y los cuadrados representan la variante de Cas12i4 de la SEQ ID NO: 4. Las barras representan indels medios en todas las dianas. Cas12i4 variante de la SEQ ID NO: 4 mostró una mayor actividad indel que Cas12i4 de tipo silvestre de la SEQ ID NO: 2, y el uso de un PAM 5'-NTTN-3' dio como resultado niveles de indel más elevados que
- 10 el uso de un PAM 5'-NVTN-3'.

Este ejemplo muestra que Cas12i4 (Cas12i4 de tipo silvestre o variante) puede inducir indels usando un PAM 5'-NTTN-3' o PAM 5'-NVTN-3'.

15

20

25

30

35